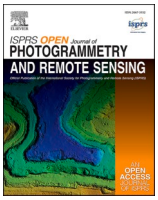
ISPRS Open Journal of Photogrammetry and Remote Sensing 9 (2023) 100040

Contents lists available at ScienceDirect

ISPRS Open Journal of Photogrammetry and Remote Sensing

journal homepage: www.journals.elsevier.com/isprs-open-journal-of-photogrammetry-and-remote-sensing



Branch information extraction from Norway spruce using handheld laser scanning point clouds in Nordic forests

Olli Winberg [a], Jiri Pyörälä [a,c,d,*], Xiaowei Yu [a], Harri Kaartinen [a], Antero Kukko [a,b], Markus Holopainen [c], Johan Holmgren [d], Matti Lehtomäki [a], Juha Hyyppä [a,b]

[a] Department of Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute FGI, The National Land Survey of Finland, Vuorimiehentie 5, FI-02150, Espoo, Finland
[b] Department of Built Environment, Aalto University, School of Engineering, P.O. Box 11000, FI-00076, Aalto, Finland
[c] Department of Forest Sciences, University of Helsinki, FI-00014, Helsinki, Finland
[d] Department of Forest Resource Management, Swedish University of Agricultural Sciences, SE-901 83, Umeå, Sweden

A B S T R A C T

We showed that a mobile handheld laser scanner (HHLS) provides useful features concerning the wood quality-influencing external structures of trees. When linked with wood properties measured at a sawmill utilizing state-of-the-art X-ray scanners, these data enable the training of various wood quality models for use in targeting and planning future wood procurement. A total of 457 Norway spruce sample trees (*Picea abies* (L.) H. Karst.) from 13 spruce-dominated stands in southeastern Finland were used in the study. All test sites were recorded with a ZEB Horizon HHLS, and the sample trees were tracked to a sawmill and subjected to X-rays. Two branch extraction techniques were applied to the HHLS point clouds: 1) a method developed in this study that was based on the density-based spatial clustering of applications with noise (DBSCAN) and 2) segmentation-based quantitative structure model (treeQSM). On average, the treeQSM method detected 46% more branches per tree than the DBSCAN did. However, compared with the X-rayed references, some of the branches detected by the treeQSM may either be false positives or so small in size that the X-rays are unable to detect them as knots, as the method overestimated the whorl count by 19% when compared with the X-rays. On the other hand, the DBSCAN method only detected larger branches and showed a −11% bias in whorl count. Overall, the DBSCAN underestimated knot volumes within trees by 6%, while the treeQSM overestimated them by 25%. When we input the HHLS features into a Random Forest model, the knottiness variables measured at the sawmill were predicted with R^2s of 0.47–0.64. The results were comparable with previous results derived with the static terrestrial laser scanners. The obtained stem branching data are relevant for predicting wood quality attributes but do not provide data that are directly comparable with the X-ray features. Future work should combine terrestrial point clouds with dense above-canopy point clouds to overcome the limitations related to vertical coverage.

1. Introduction

The global goals of the UN Agenda 2030 aim to balance the three dimensions of sustainable development: economic viability, environmental conservation, and social equity. In addition, solutions are needed for the urgent challenge of shifting away from fossil-based production industries to bio-based, renewable ones. Stakeholders involved in the bio-based economy require precise and up-to-date information to control the use of renewable natural resources and related industrial processes. Nordic countries are the world's most extensively forested industrialized and temperate countries. Especially in Finland and Sweden, the forest sector accounts for a significant share of export revenues and acts as a major employer. The extent of Nordic forest resources is estimated at 5600 million cubic meters, and the annual increment is approximately 220 million cubic meters, of which 150 million cubic meters is harvested annually. We have over 300 billion (300×10^9) trees in our forests. In Finland and Sweden alone, we have ca. 1 million forest owners with 4 B€ gross stumpage earnings, several thousand harvesting companies, 1000+ transportation companies with long-distance transportation costs close to 1B€, and more than 1000 industrial companies with approximately 50 B€ total direct turnover. Nordic forest companies are among the top forest and paper product companies in Europe.

Next-generation forest inventories record not only basic tree attributes, but also the forest state and product yields in high detail. These data act as the basis for decision-making regarding human interventions in forest ecosystems, ranging from landscape-scale management decisions to harvest planning in forest stands and log-breakdown decisions at the individual-tree level. Great benefits would be obtained if the wood properties and timber quality of individual trees, i.e., their suitability for specific end uses, could be characterized prior to harvest. Such characterization would enable the optimization of forest operations, wood procurement down to details of log breakdown, and raw material

* Corresponding author. Department of Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute, 02150, Espoo, Finland.
E-mail address: jiri.pyorala@nls.fi (J. Pyörälä).

https://doi.org/10.1016/j.ophoto.2023.100040
Received 24 February 2023; Received in revised form 26 June 2023; Accepted 26 June 2023
Available online 28 June 2023

distribution between appropriate processors.

Wood quality is the result of past growth allocation between tree compartments, i.e., the stem and crown wood, foliage, and roots. Depending on the species-specific life strategy, environmental and climatic conditions, and competition, trees optimize their capacities to photosynthesize and conduct water, along with providing structural support to the whole organism (Lachenbruch et al., 2011). The characterization of wood qualities within tree stems can most effectively be achieved with X-rays, which allow the direct interpretation of wood density within logs. This is affected, e.g., by the moisture content, knot presence, internal defects (such as compression wood and decay), and ring width (Oja et al., 2003). The techniques are based on the absorption of X-rays: the denser the wood, the more X-rays it will absorb. The interpretation of the X-ray images is based on artificial intelligence making distinctions between lighter voxels versus denser voxels, to identify wood properties that are influential to log sorting, planning of timber production, and log value determination, e.g., log geometry, knots, and defects (Beaulieu and Dutilleul, 2019). State-of-the-art sawmills in Finland and Sweden deploy multi-directional digital radiographers that provide sparse data on the log interiors at high speed, while some specialized sawmills have begun using rotational X-rays that enable the computed tomography scanning of log interiors in high detail but on the expense of the measurement speed (e.g., (Wei et al., 2011; Lindgren, 1991)). However, wood properties in standing timber can only be assessed indirectly based on, e.g., increment cores, probe penetration techniques, or acoustic velocity measurements (Schimleck et al., 2019), but these are laborious, expensive for operational settings (e.g., Gao et al., 2017), and hard to link with sawmill measurements or to generalize over large areas.

Previous studies have investigated numerous ways to utilize the airborne point cloud data for estimating wood quality and wood properties over larger areas (Pyörälä et al., 2019a; Blanchette et al., 2015; Luther et al., 2014; Hilker et al., 2013; Côté et al., 2021). One of the key bottlenecks of such approaches is the lack of ground-based observations of wood quality indicators, to support the prediction framework. Therefore, as a potential operational concept, we vision that a mobile laser scanner mounted on a harvester could facilitate the detailed point cloud-based reconstructions of the external structures of the harvested trees that were linkable with the wood properties measured in the sawmill. Accumulation of the database with fused laser scans and X-ray data would enable training various wood quality models to be used in targeting and planning future wood procurement or for real-time, quality-based bucking optimization in harvesters.

Past work towards realizing this novel concept includes the following studies, which have mainly considered static terrestrial laser scanning (TLS): Pyörälä et al. (2018a) compared TLS with sawmill X-ray scanning data to find a link between these datasets for whorl locations and each whorl's maximum branch and knot diameters. Pyörälä et al. (2019b) further compared measurements of log geometry (stem dimensions, volume, taper, and sweep) between TLS and sawmill data. Moderate correlations were found between the stem geometry variables and internal stem properties, i.e., whorl-to-whorl distances and mean sapwood density. Several previous studies also exist, showing that the stem curve, stem diameters, and 3D locations can be extracted from terrestrial and mobile point clouds. Liang et al. (2014) extracted the stem curves of a few dozen trees from TLS data at a cm-level compared to harvested measurements. Nguyen et al. (2021) developed TLS-based methods to detect tree surface defects, such as branch scars, branches, epicormics, and bark defects. Bauwens et al. (2017) used terrestrial photogrammetry for measuring and modeling stems and their irregularities. Hyyppä et al. (2020a) compared several mobile laser scanning (MLS) techniques, such as backpack laser scanners, a handheld Zeb-Horizon laser scanner, and an under-canopy unmanned aircraft vehicle (UAV) laser scanning system. They found that the stem curve measurements had root-mean-squared errors (RMSEs) of 2–15% depending on the system and the measurement height of the standing tree. Additionally, several studies confirm the capability of TLS and MLS to extract stem diameters and stem locations in 3D (Hyyppä et al., 2020b, 2021; Balenović et al., 2021; Chen et al., 2019; Cabo et al., 2018; Bauwens et al., 2016; Del Perugia et al., 2019; Bienert et al., 2018; Čerňava et al., 2019). Moreover, the branching structures of trees, such as whorl-to-whorl distances, branch angles, and diameters are also extractable from the static terrestrial point clouds (e.g., (Dassot et al., 2012; Lau et al., 2018; Pyörälä et al., 2018b; Bournez et al., 2017; Raumonen et al., 2013)).

The static TLS thus provides highly detailed tree-specific information at roughly the same rate as manual field work provides data on sparser features, i.e., more information is gathered, but the time consumed remains the same. However, the major promise made by laser-scanned data is the development of rapid and autonomous, mobile data acquisition systems, such as those based on drones flying within the canopy, or harvesters (Hyyppä et al., 2020c, 2021; Wang et al., 2021; Bienert et al., 2021; Vandendaele et al., 2022). For example, a rotating mobile laser scanner is a feasible sensor to be implemented, e.g., in a handheld system, in an under-canopy drone, or in a harvester. The introduction of autonomous sensor systems to practical forestry will expand the amount of point cloud data available for the industry and for use in large-scale mapping. Prior work has presented various strategies for extracting and estimating crown dimensions and structure from mobile point clouds, along with whorl characteristics (Puliti et al., 2023; Hartley et al., 2022), branch volume (Qi et al., 2022), and tree defects (Morgan et al., 2022). Mobile point clouds typically exhibit more stochastic noise, lower point density, and poorer spatial precision than static point clouds, especially when working with global navigation satellite system (GNSS) shadows and when the length of the mobile device trajectory increases (Kaartinen et al., 2015). These are key challenges in the utilization of mobile point clouds to measure the finer details of tree structure. Therefore, further method development studies are still required to process mobile point cloud data.

In this article, we describe the results obtained from using a mobile, handheld laser scanner (HHLS) for automatically extracting branching structures of standing trees. We have correlated our results with industrial sawmill data as a reference. The study's main objective was to identify the feasibility of the HHLS point clouds to provide information of the branching properties of Norway spruce (*Picea abies* (L.) H. Karst.) that are relevant to industrial use. We inspected two algorithms to test the accuracy of the HHLS point clouds in representing the branching structures of Norway spruces, to proxy the industrial knottiness features at the individual-sawlog level, and to serve as input variables in empirical models that predict knottiness.

In summary, the novel contribution of this paper includes:

- Mobile terrestrial point clouds are used for extracting branching structure data of standing trees to derive wood quality features.
- The study compares the automatically derived features and further modeled outputs with the sawmill-measured X-ray measurements towards the abovementioned goal.

2. Materials

2.1. Study site

Our study was conducted in fall 2020 on thirteen test sites located in southeastern Finland in the South-Karelian region, see Fig. 1a (61°03′ N 28°11′ E). All test sites were situated in even-aged managed boreal forests. Norway spruce was the dominant tree species on all test sites. The test sites were selected from the harvest program of the industrial partner (Stora Enso Oyj (SE), Helsinki, Finland), to represent the typical variation of conditions in mature Norway spruce stands available for clear-cutting in the studied area. Stand mean tree heights (H) were ca. 20–25 m, and stocking densities ranged from 300 to 950 stems per hectare, with basal areas of 23–34 m^2/ha. Forest types encompassed mesic (MT) and sub-mesic (OMT) mineral soils and corresponding peat

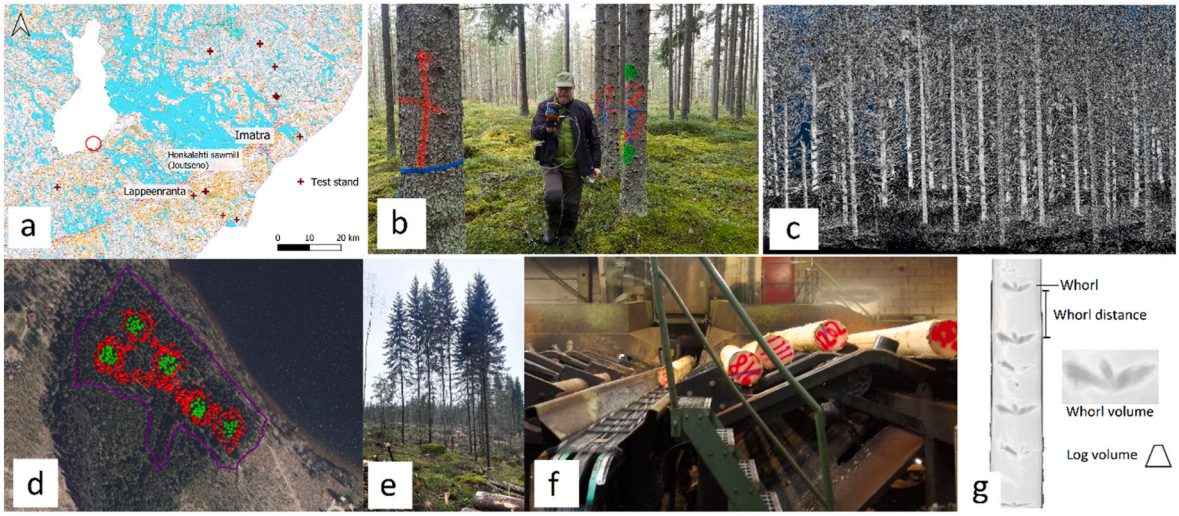


Fig. 1. Overview of the materials used in this study. a) Locations of the studied stands in southeastern Finland. b) Acquisition of the handheld laser scanning (HHLS) data. c) Preview of an HHLS point cloud. d) Examples of tree maps at the stand level from the HHLS (red) and at the plot level (green). e) Sample tree group prior to felling. f) Logs from the sample trees at the sawmill, with identification numbers painted on log ends. g) Schematic illustration of the X-ray images used to detect knot whorls and to obtain log volume (as the volume of a truncated cone). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

lands. Varying levels of species mixtures were present on the stands, mainly consisting of Scots pine (*Pinus sylvestris* L.), silver birch (*Betula pendula* Roth), and Eurasian aspen (*Populus tremula* L.).

We placed two to five circular sample plots with 12-m radius each evenly around each stand, with the number of plots depending on the stand size (i.e., roughly one plot per hectare, but with a minimum of two and a maximum of five plots per stand). From each plot, 6–10 Norway spruces were selected around the plot center that represented the distribution of the Norway spruce diameters at breast height (DBH), i.e., at a height of 1.3 m from the ground), H, and the presence of crown classes or canopy layers (dominant, intermediate, and suppressed) within the stand. In total, 457 Norway spruce sample trees from 49 sample plots were selected. See Table 1 for a summary of the study site and sample tree attributes.

Fig. 1b shows an example of a tree map. The sample trees were annotated with laminated paper sheets containing Tree-ID numbers for identifying the individual sample trees from point clouds and during harvests.

2.2. The acquisition and pre-processing of the handheld laser-scanned data

All test sites were recorded with a handheld ZEB Horizon (GeoSLAM, UK) scanner in October 2020. Each of the test sites was covered by walking in a four-leaf-clover-shaped pattern containing several loops, and the start and end points were the same to improve the accuracy of data collection to provide multiple possibilities for loop closure detection and subsequent drift elimination: these precautions can effectively eliminate the drift and off-set errors in the point clouds. Similar walking patterns were also used in Hyyppä et al. (2020a), see Fig. 4c in the reference. The walking speed was 2 km/h, the system was held upright in hand at the fixed height of 1.2 m above ground. Each handheld acquisition took approximately 20 min.

After data collection, the raw data were pre-processed using the GeoSLAM Hub (version 6.0.0.) software. To this end, we utilized the default values of the processing parameters: Convergence threshold: 0, Window size: 0, Voxel density: 1, Rigidity: 0, Maximum range: 100 m, Closed Loop. After the pre-processing steps, the resulting point cloud data were exported into ".las"-format for further processing. The quality of the scans was checked, as reported by Hyyppä et al. (2022).

Table 1
Standwise forest inventory information of the study sites on the left: Stand age, mean diameter at breast height (DBH), and mean height (H) are reported only for Norway spruces. The characteristics of the sample trees in each stand on the right.

Standwise forest inventory									Sample tree inventory			
Stand	Area (ha)	Forest type	Basal area (m^2/ha)	Stems/ ha	Spruce (%/volume)	Stand age, a	Mean DBH (cm)	Mean H (m)	Plots	Sample trees	Mean DBH (cm)	Mean H (m)
1	3.3	OMT	25.2	952	93.5	71	25.3	21.0	5	42	29.9	23.8
2	1.8	MT	31.1	465	58.7	66	29.8	22.4	4	31	32.7	26.7
3	0.6	MT	31	544	95.2	64	29.1	20.3	2	17	32.5	23.8
4	2.8	MT	29.6	600	66.5	64	26.3	19.5	4	34	25.6	22.2
5	3.4	MT	25.3	524	85.7	71	27.0	21.0	5	44	29.2	23.4
6	1.8	MT	23.3	407	86.1	68	29.3	21.5	4	33	31.0	24.6
7	2.5	OMT, Peat	25.9	482	59.5	71	27.4	20.2	5	39	29.6	23.2
8	2.3	MT	27.3	315	88.9	90	30.6	26.0	4	30	29.6	24.9
9	5.3	MT	33.2	479	75.6	91	33.2	25.4	5	46	33.4	27.4
10	0.9	OMT	32	506	87.1	85	29.0	22.6	2	18	27.2	23.0
11	4.7	OMT	25.5	696	70.1	76	23.7	19.4	5	44	29.8	24.4
12	1.6	MT,Peat	30.6	481	76.2	82	31.2	24.8	4	33	31.7	24.1
13	3.8	MT	33.7	632	77.8	84	28.0	23.1	5	42	29.9	25.3

2.3. Sawmill data acquisition

All test sites were harvested during the study and taken to SE's Honkalahti sawmill in Joutseno, Finland, where the X-ray-based wood quality measurement was undertaken. The selected sample trees were tracked throughout the harvests and transportation to the sawmill. The harvests took place over the course of Oct.–Nov. 2020, and the sawmilling was carried out in Nov.–Dec. 2020.

The harvesting companies first clear-cut the stands around the sample tree groups. At the time of sample tree felling, the logs bucked from the sample trees were annotated with tree identification numbers (Tree-ID) using marking paint at both ends of the logs. The bucking followed the production matrix of the sawmill, and logs were cut to lengths ranging from 4.1 m to 5.7 m at 30-cm intervals. The minimum allowed top diameter was 12 cm. The minimum quality requirement was that each sawlog should stay in one piece after bucking, but the harvester operator ignored factors such as excessive sweep, crook, decay, ramicorn branches, and other flaws that would otherwise have justified downgrading logs to pulp wood.

Paint-marked logs were kept separate from the rest of the logs, hauled to the sawmill, and stored in a designated pile. Due to delays in the harvesting operations from one stand, and some sample logs lost in the process, a total of 1244 logs from 450 (identified) trees eventually found their way to the sawmill in time for the scheduled sawmill measurements. The total fresh volume of the batch was 2945 m^3.

At the sawmill, the logs were debarked and then X-rayed using a two-directional FUSION G3X digital radiographer (Finnos Oy, Lappeenranta, Finland) with 0.7-mm effective resolution. The feeding of the logs into the X-ray machine was videoed, and by reading the Tree-ID from each log on the video, the X-raying data were linked to respective sample trees. From the X-ray measurements, over 100 features were calculated for each log using the in-house algorithms from Finnos. We used five of these features in our analyses: knot volume (m^3), knot index (knot volume relative to log volume), mean whorl volume (knot volume relative to the number of whorls in each log, m^3), maximum whorl volume (m^3), number of whorls, and mean whorl-to-whorl distance (m).

3. Methods

3.1. Overview of the methods

We compared two approaches for detecting branches from the HHLS data. A method based on the density-based spatial clustering of applications with noise (DBSCAN) (Ester et al., 1996) was developed and compared to the segmentation-based quantitative structure model (treeQSM) (Raumonen and Åkerblom, 2019) that is widely used in the literature (e.g., Lau et al., 2018; Raumonen et al., 2013; Hackenberg et al., 2015). Both approaches assume point clouds of individual trees as an input. Individual-tree extraction from the point clouds of a larger forest area with possibly multiple trees can be achieved by implementing different segmentation methods prior to tree modeling, such as those described in Hyyppä et al. (2020c) that were also used here. In short, the point clouds of individual trees were extracted as follows: a canopy height model was first generated and smoothed with a Gaussian filter and then individual trees were segmented by applying the watershed algorithm. Branch growing locations, diameters, and insertion angles were generated from the HHLS point clouds of the segmented, denoised trees with the two algorithms studied, and compared to log X-ray factory data that provided reference values for the internal wood quality.

3.2. Denoising the point cloud data

The pre-processed HHLS data had point densities of 1000–10000 points/m^3. These point cloud data still contain noise due to physical limitations and measurement errors of the measurement device, which complicate the analyses of detailed geometric features. Denoising was therefore the first step in processing the point clouds of individual trees. To reduce the noise and enhance the extraction of relevant features, a heuristic denoising based on the local point density was applied to the point cloud as part of the analysis. Fig. 2 illustrates the results of denoising a point cloud. The assumption is that noisy points are uniformly distributed in space and have smaller local point densities than points arising from actual hits of the measurement laser in various parts of the tree. The local point density for a point is defined as the number of points in a sphere of a fixed radius centered at the point.

Having computed a local point density for each point with some fixed neighborhood radius, noisy points were filtered out by setting a threshold on the local point density values. By trial and error, a 7-cm radius was chosen for the spherical neighborhood, and the threshold value used for the local point density was the 20th percentile over all spherical neighborhoods for each point. Such denoising is computationally time efficient and takes only seconds to perform for a coarse point cloud of an individual tree with over one million points. In addition to removing noisy points, the denoising also removed foliage from the point clouds of trees, leaving mostly the woody parts for further analysis, as illustrated in Fig. 2. This foliage removal effect was desired, as it simplified analyzing the branching structure of the trees.

3.3. Branch extraction with DBSCAN

The DBSCAN-based branch detection method can be divided into the following five stages:

1) Denoising (Figs. 2 and 3a)
2) Segmentation and extraction of points belonging to the tree trunk (Fig. 3b)
3) Removal of trunk points (Fig. 3c)
4) Determining disconnected components of the remaining point cloud (Fig. 3d)
5) Post-processing disconnected components to remove clusters not corresponding to individual branches (Fig. 3e)
6) Estimating geometric features (growing location, diameter, and insertion angle) of each branch (Fig. 3f)

After denoising the point cloud of an individual tree, the branch inventory of the tree begins by identifying and extracting its trunk. Points that were part of the trunk were detected with arc detection, random sample consensus (RANSAC) - based circle fitting, and local descriptors. First, a rough identification of trunk points was achieved by detecting circular arcs with RANSAC-based arc detection (Fig. 3b), as the trunk is assumed to produce such arcs in the laser scanning measurement process (Hyyppä et al., 2020c). Secondly, a circle fitting process was applied to the arc-shaped points to remove any arc-shaped clusters of points not belonging to the trunk and to extract the trunk in finer detail (Fig. 3 c). A parallel strategy for trunk detection included computing local descriptors for the denoised point cloud and then selecting points with nearly vertical growth as trunk point candidates. Local descriptors provide an invariant geometric characterization, in which the trunk geometry was typically characterized with high local flatness and a nearly vertical principal axis. As trunk detection is carried out as a combination of arc detection and local descriptors, it also works for curved trees.

After the trunk points were identified from the denoised point cloud, we removed points that were not within a 1-m radius of the trunk (Figs. 2 and 3c). Then the trunk points were removed, and the analysis continued with the remaining point cloud. According to (Fig. 3c). By a heuristic assumption, different branches were connected to each other only by the trunk, meaning disconnected components in the remaining point cloud should correspond to individual branches, as described in Equations (Lachenbruch et al., 2011) and (Oja et al., 2003). The search for disconnected components from the remaining point cloud was performed with the DBSCAN (Fig. 3d).

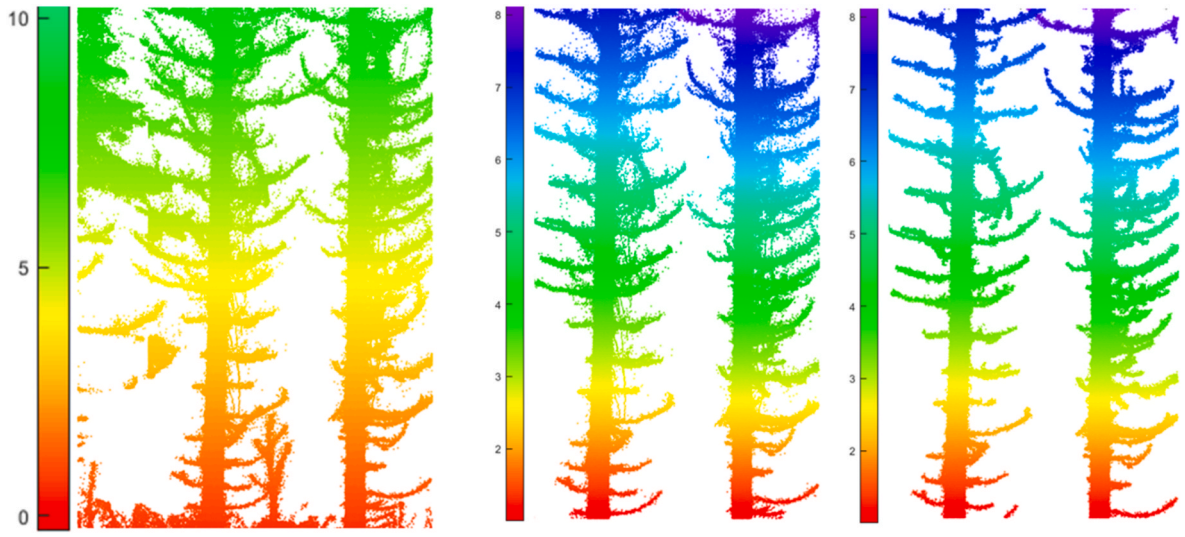


Fig. 2. The preview of the denoising effects. Left: the original point cloud. Middle: the point cloud with points further than 1 m from the trunk removed. Right: the final denoised point cloud.

The estimates for branch insertion angle were thus achieved by selecting a 15-cm section of the branch near the trunk that was subjected to principal component analysis (PCA) (Fig. 3f). The eigenvector corresponding to the largest eigenvalue v_1 was then taken as the principal growth direction of the branch (Equation 3). The dot product of the eigenvector and unit vector of the vertical axis (e_z) gave the cosine of the branch insertion angle, i.e., we assumed the trunk normal as perfectly vertical (Equation 4). The branch diameter estimate was achieved by projecting the 15-cm branch section onto a plane orthogonal to the principal eigenvector. Then, a circle was fitted to the boundary of this two-dimensional set of points, the diameter of which was reported as the branch diameter (Equation 5, Fig. 3f).

$$branches = disconnectedComponents(Tree \backslash Trunk)_{DBSCAN} \quad [1]$$

$$branches = \{B_i\}_{i=1}^{n} \quad [2]$$

$$[eigenvalues, eigenvectors] = PCA(B_i) \quad [3]$$

$$vertical\ branch\ angle = acos\ (v_1 \cdot e_z) \quad [4]$$

$$branch\ diameter = circlefit(P_v(B_i))_{RANSAC} \quad [5]$$

B_i is the set of points belonging to an individual branch i, PCA is the principal component analysis, v_1 is the first principal direction, e_z is the vertical axis [0,0,1], and $P_v(B_i)$ are the points from branch i projected orthogonally towards the principal direction v_1.

In practice, the problem of this approach is that neighboring branches may touch each other outside the trunk, meaning more than one branch may form a joint single cluster. Clusters showing signs of including more than one branch were post-processed or completely removed from the following branch angle and diameter estimation. The criteria used for filtering away poor clusters not containing individual branches included small linearity, a too small point count, and a too small diameter for the set of points, the points being too far away from the trunk, as well as an unreasonable estimate for branch growth angle and diameter.

3.4. Branch extraction with treeQSM

A cylinder-based reconstruction technique for TLS point clouds of individual trees was previously proposed by Raumonen et al. (Raumonen et al., 2013; Raumonen and Åkerblom, 2019). The requirements for the point cloud data that treeQSM can take as inputs are listed in the official documentation of treeQSM (Raumonen and Åkerblom, 2019). HHLS point clouds do not satisfy all the assumptions needed to create the cover-based model for a tree. However, the segmentation and cover set creation methods of treeQSM can be used for determining the branch inventory of a tree. Here, the following functions from the treeQSM toolbox were used to extract the branching inventory from the denoised HHLS point clouds of individual trees: “sets.m”, “segments.m”, “correct \segments.m”, and "tree\sets.m". Parameters used for these methods are listed in Table A1 (in the Appendix) and discussed further in the treeQSM manual (Raumonen and Åkerblom, 2019).

After extracting the individual branches with treeQSM segmentation methods, the branch diameter and insertion angle estimates were derived similarly, as described with the DBSCAN-based method. The segmentation of treeQSM is illustrated in Fig. 4.

3.5. Statistical analysis

First, we analyzed the differences between the DBSCAN and treeQSM methods in finding the branching locations along the tree stems and inspected the properties of the found branches using tree-specific descriptive statistics (mean, standard deviation, minimum, and maximum values) of branch heights, diameters, and insertion angles. The purpose of the comparison was to provide understanding on how the methods performed in reconstructing the branching of the trees and to enable analyzing the implications of these baseline differences to the calculations downstream.

Second, we evaluated the performance of the applied algorithms in reproducing proxies for the knottiness variables measured in the sawmill with X-rays. Log-to-log direct comparisons of the point cloud-based branching data to the sawmill data were conducted by calculating proxies of the X-rayed features directly from the detected branches. Total knot volume within a log was calculated as the sum of the knot volumes estimated for each branch based on the branch diameter and stem radius: the knot was expected to resemble a cone extruding from stem pith to stem surface, linearly expanding, and revolving symmetrically around the branch pith. Knot index was calculated as the ratio of total knot volume to the log volume. Mean whorl volume was calculated as the ratio of knot volume to whorl number. The number of whorls was defined based on hierarchical clustering of the branch heights, using a 10-cm vertical window (i.e., groups of branches within a 10-cm vertical distance are considered to belong to a single whorl). Whorl distance was defined as the mean distance between each whorl and their two nearest neighbors. The evaluation considered the following statistics: bias (or

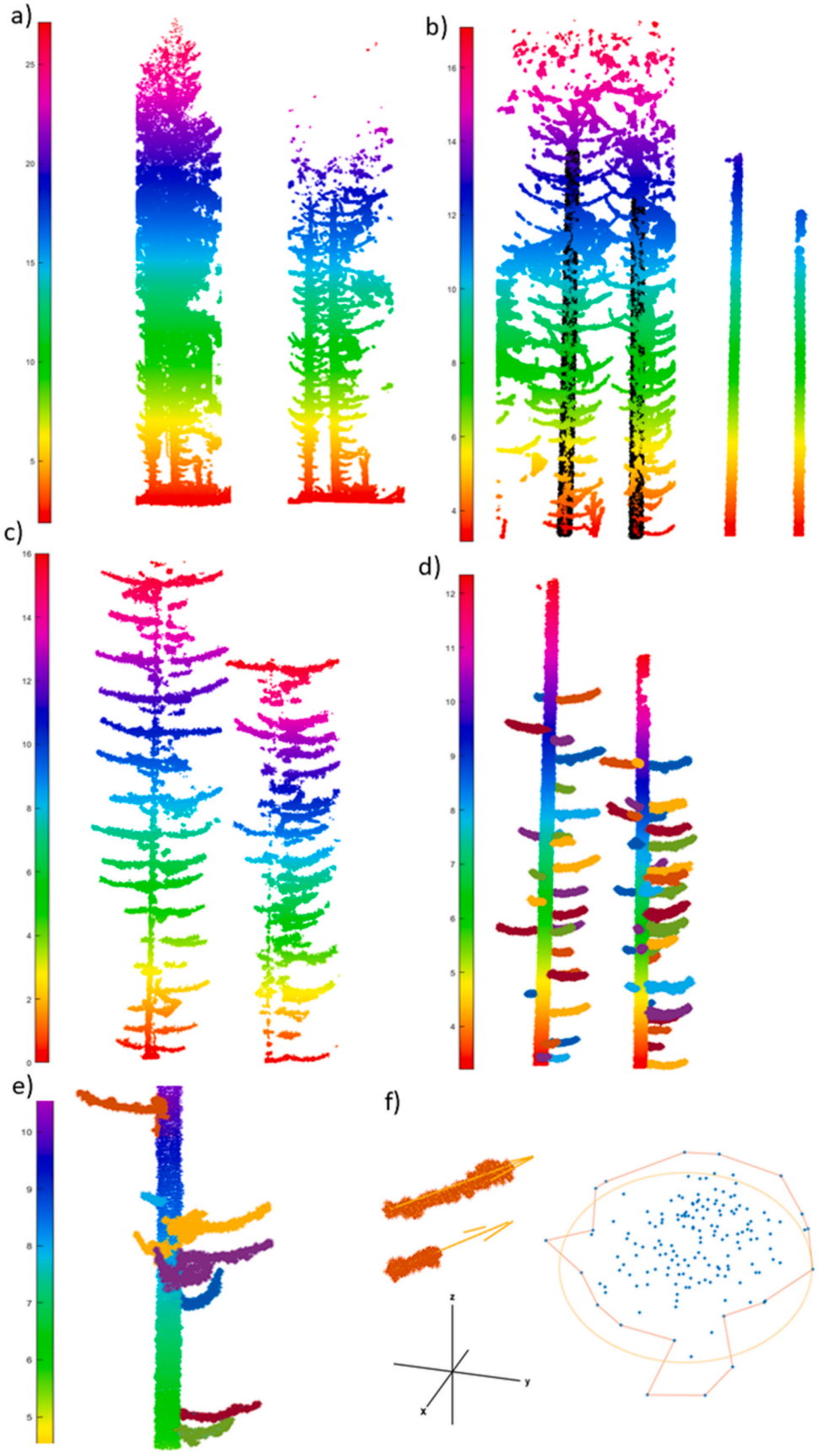


Fig. 3. The stages of the DBSCAN-based branch detection method. a) Point cloud denoising. b) Trunk detection. c) Trunk removal. d) Searching for disconnected components from the remaining point cloud. e) Post-processing and removal of erroneous clusters. f) Estimating geometric features of detected branches.

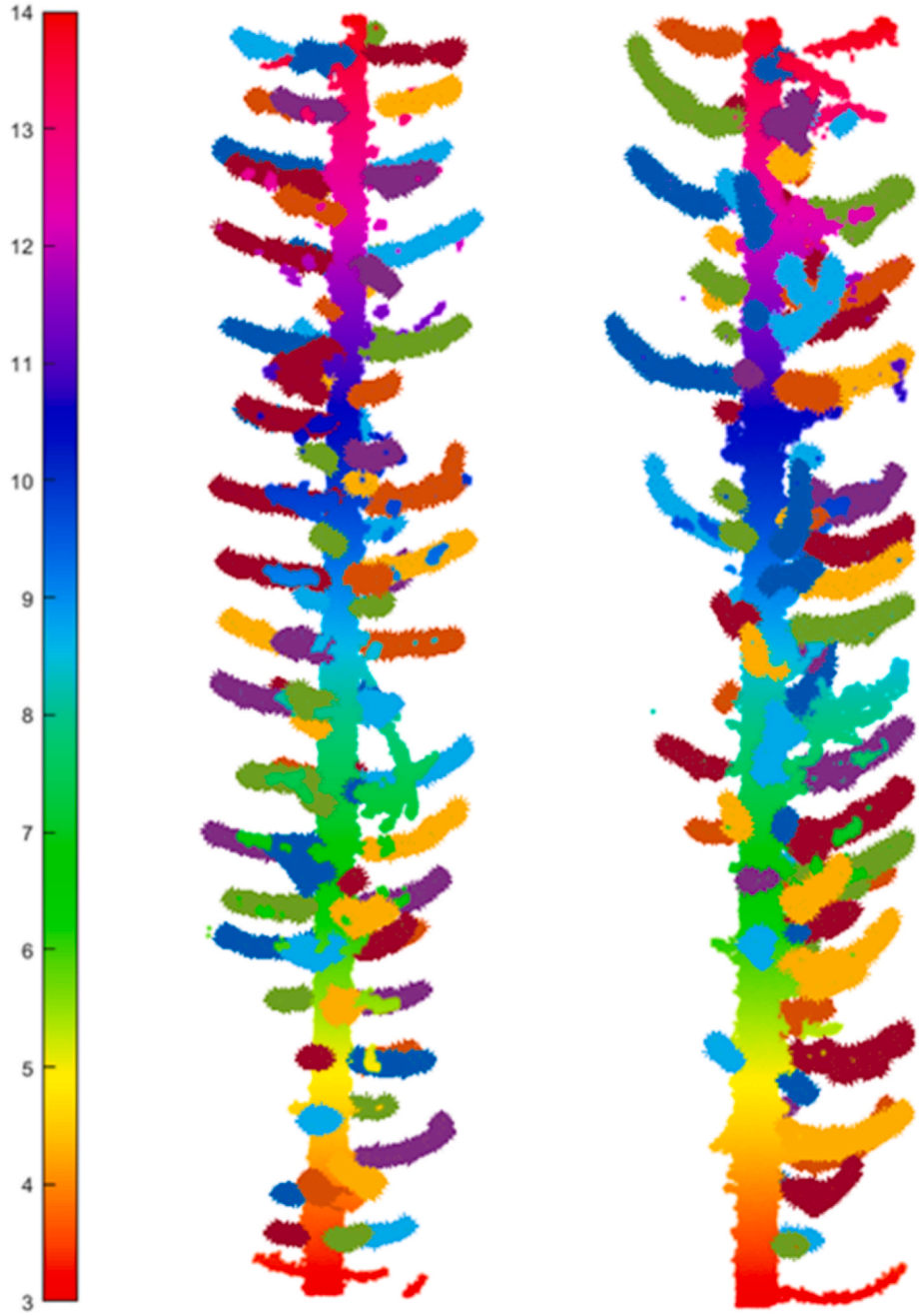


Fig. 4. Visual illustration of the treeQSM segmentation outcome in two example stems. Boldfaced points in different colors correspond to individual branches. Trunk height colored in meters. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

the mean difference), RMSE, both in absolute and relative terms, the coefficient of determination (R^2), and sigma (or the standard error).

Thirdly, for each log, we calculated the minimum, mean, maximum, and standard deviation of the branch diameters and insertion angles. These variables together with the above-listed proxies were then subjected to an ensemble feature selection procedure consisting of regression trees and the Random Forest algorithm (RF) (Breiman, 2001), with 2000 regression trees generated and two variables randomly permuted at each node, similarly to Pyörälä et al. (2019a). RF evaluated the prediction accuracy against one-third of the sample tree data (out-of-the-bag sample) in each regression tree. We used the analysis to identify the most important predictors of the X-rayed features based on their prevalence and averaged the effect across all regression trees. We assessed the final modeling accuracy within our data using the RF, based on the R^2, p-value, and the residual standard error (RSE).

4. Results

4.1. Comparison of the branch detection methods

Of the two tested approaches for determining the branch inventory of individual trees, treeQSM identified a larger number of branches than the DBSCAN. An example case illustrating individual branch detection is given in Fig. 5. On average, treeQSM found 37 branches more (+46%) per tree than the DBSCAN method. Fig. 6 illustrates that the number of detected branches was lower with DBSCAN both among the lowest branches and the highest branches, i.e., treeQSM performed better in branch identification throughout the length of the stem: The minimum

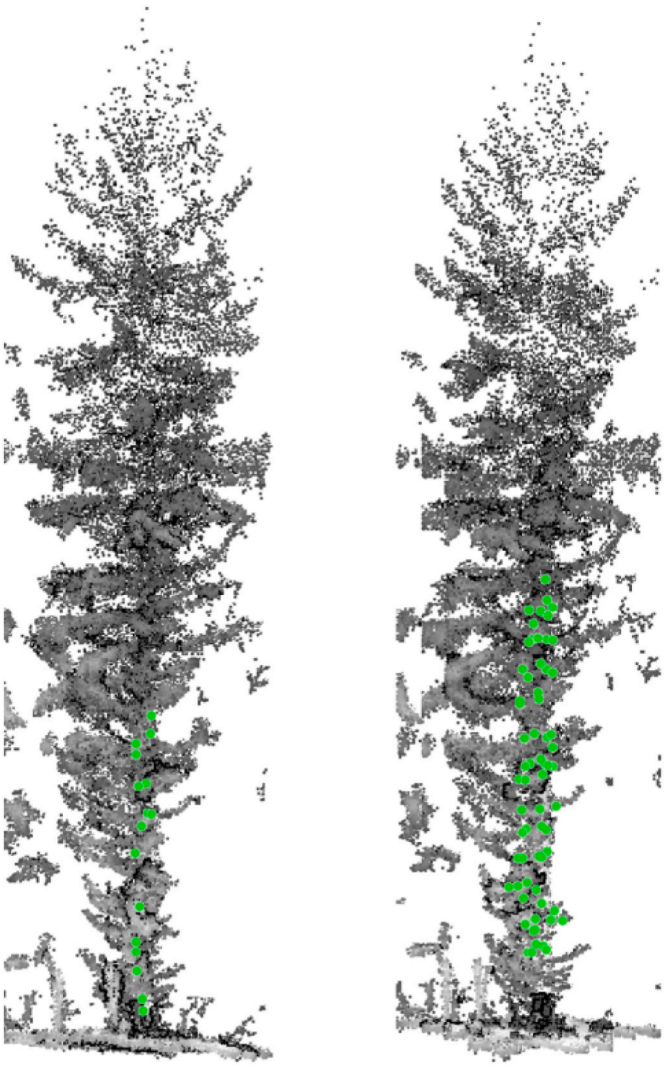


Fig. 5. The illustration of branch detection with both methods, the DBSCAN clustering method (left) and treeQSM segmentation (right).

and maximum heights of the detected branches averaged 0.40 m (17.6%) and 1.44 m (17.9%) lower and higher, respectively, with treeQSM compared with DBSCAN.

On average, DBSCAN found larger branches with steeper insertion angles than treeQSM (Fig. 6): the tree-specific minimum and mean branch diameters found by treeQSM were, on average, 28% and 1.6% smaller, respectively, than with DBSCAN. treeQSM, thus, also detected smaller branches that were overlooked by DBSCAN. However, the tree-specific maximum branch diameters were 4.8% larger with treeQSM than with DBSCAN, while the variance around the mean of the tree-specific maximum branch diameters was slightly greater with DBSCAN than with treeQSM; 3.1 cm and 2.1 cm, respectively. The methods also differed between the estimated branch insertion angles, especially between the minimum values. On average, DBSCAN gave 8.8% higher tree-specific mean branch insertion angles.

4.2. Results of the proxy-based approach

Log volumes from the HHLS point clouds were overestimated with a bias of 0.024 m^3 (10.2%) compared with the X-rayed values (Table 2). Most of the discrepancy was likely caused by bark removal and butt swelling at the sawmill prior to the X-raying, despite the obvious underestimation of the relative stem diameters in the HHLS point clouds (Fig. 7). Stem occlusion in the top logs also affects the representation of the volume. DBH was estimated with high accuracy; a bias of −0.37 cm (−1.2%) compared with the field-measured values.

The comparison of the target knot variable proxies from the point clouds to the values measured with X-rays revealed that both branch extraction methods mostly overestimated small values and underestimated large values (Fig. 7). The comparisons gave relatively high RMSEs above 50% and up to almost 90% for all the variables (Table 2), which highlights the large variation in the detected and modeled branches (compare Figs. 6 and 7).

The total knot volume within logs and the ratio of knot volume to log volume (knot index) were estimated with −6.7% and 25.3%, and −27% and 4% relative biases with DBSCAN and treeQSM, respectively (Table 2). Fig. 8 illustrates how the knot volume and knot index are overestimated by treeQSM in the butt and middle logs, while both methods underestimate the features in the top logs. Both methods show slightly larger variation in the values compared with the X-rays, except for the knot indexes in the top logs. The mean whorl volumes were estimated with low biases, while the maximum whorl volumes were underestimated by both methods, with biases of −45.9% for DBSCAN and −36.6% for treeQSM (Table 2). The mean whorl volume was overestimated by both methods, especially in the butt logs, and underestimated in the top logs (Fig. 8).

The treeQSM method overestimated the number of whorls compared with the X-rayed data (bias of 19.3%), most notably in the middle logs with 54% overestimation (Table 2). The DBSCAN method underestimated the whorl numbers with an overall bias of −11.4% (Table 2). Both methods underestimated the number of whorls in the upper stem parts. The number of detected whorls was directly reflected on the estimated mean whorl distances, yielding an underestimation of 0.08 m (−29.3% bias) with treeQSM and an overestimation of 0.02 m (6.1% bias) with the DBSCAN method (Table 2). The treeQSM method also gave smaller variation of the whorl distances compared with the X-rays and the DBSCAN (Fig. 8).

Overall, the DBSCAN method obtained the most accurate proxies for the knottiness of the middle logs, i.e., gave nearly unbiased estimates for the number of whorls and the mean whorl volume, and as a result, the knot volume and knot index (Table 2, Fig. 8).

4.3. Results of the model-based approach

When we subjected the HHLS features to the RF models, they predicted the X-ray features with R^2s of 0.47–0.54 using the treeQSM features and 0.56–0.64 using the DBSCAN features, with relative RSEs of 9.6–18.7% and 8.8–16.5%, respectively (Table 3).

Following the regression tree-based selection procedure, the models used 4–11 features in the predictions. Of the treeQSM features, the direct proxy of the knot volume was the most important predictor for the X-rayed knot volume and the whorl features (Table 3). In addition, the treeQSM-derived knot index, mean branch diameter, and mean whorl volume were among the most important predictors of the other X-ray knot variables (Table 3). With the DBSCAN features, the mean whorl volume was the most important factor predicting knot volume and maximum whorl volume (Table 3). The maximum branch insertion angle was the most important predictor of the X-ray knot index, while the maximum branch diameter and mean whorl volume explained most of the variance in the mean whorl volumes (Table 3). The DBSCAN-derived knot index and knot volume best explained the X-ray whorl features (Table 3).

Based on Fig. 9, modeling with either set of the HHLS features yielded quite similar predictions. The models underestimated the largest values and overestimated the smallest values, but the DBSCAN features gave slightly less errors at both ends, as can also be deduced from the slightly smaller RSEs in Table 3.

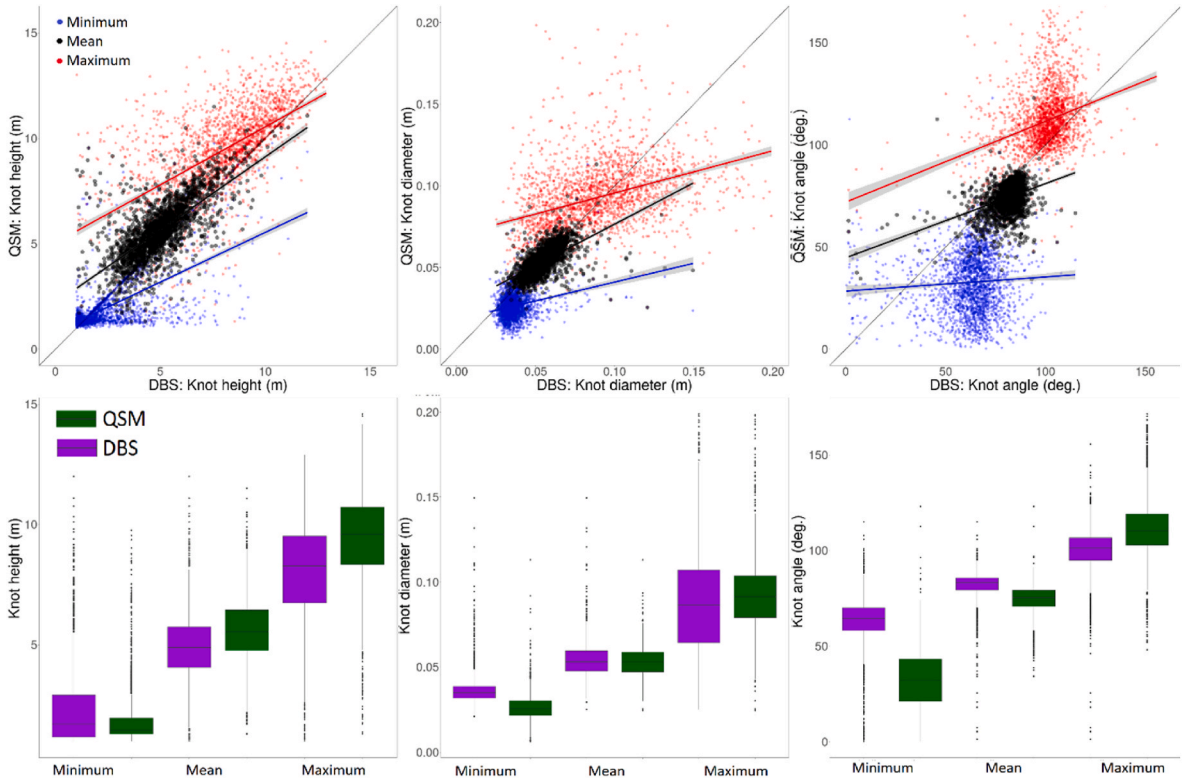


Fig. 6. Comparisons of the two algorithms used for branch extraction: tree-specific statistics of the heights, diameters, and insertion angles of the extracted branches. The comparison entails data from 2185 trees that were modeled with both methods.

5. Discussion

5.1. Overview

In this study, we investigated the performance of two point cloud processing algorithms in the task of extracting individual branches from terrestrial point clouds of Norway spruces, obtained with a handheld laser scanner. We compared the two methods with each other, as well as with knot features measured in an industrial sawmill with X-rays to assess their feasibility in reproducing industrially relevant data on knottiness. The results showed that while the treeQSM method captured more branches and a wider range of branch properties, the sparser features recorded by the developed DBSCAN method more closely reflected the measurements made in the sawmill. However, neither method was capable of directly estimating the important knottiness features in the standing trees. Instead, the use of either of the feature spaces yielded promising results when used as predictors in the Random Forest -based modeling approach, supporting the idea that the tree structures recorded with under-canopy mobile point clouds produce useful data to build models for predicting wood quality in standing timber.

The main motivation for this investigation was that the mobile point clouds have roughly 10 times the measurement speed in the field when compared to static point clouds, and they are also applicable to practical solutions such as harvester integration. These solutions will enable collecting databases that consist of point cloud-based information of tree characteristics, harvester data, and wood quality data from the sawmill. Based on our results, the collection of branching data with mobile laser scanners is justified for the purpose of building novel wood quality modeling frameworks.

5.2. Data and method considerations

The methods previously developed for static terrestrial point cloud data expect are highly accurate (ranging accuracy of a few millimeters) and dense point clouds, while the mobile point clouds are generally less accurate (ranging accuracy 2–3 cm) and sparser. In addition, mobile laser scanners usually have larger footprints, and the width of laser beams can be larger than the branch diameters. Despite these obvious differences, our results were comparable to our previous results with the static terrestrial point clouds and indicated the presence of similar limitations as those found, e.g., by Pyörälä et al., 2018b, 2018c: the accuracy and coverage of the point clouds drastically decrease above the live crown base height, due to the effects of occlusion, increasing point spacing, and footprint. However, the lower tree parts that are more thoroughly recorded, yield sufficient branching data to be used for empirical modeling of wood quality features. Our results indicated that the choice of point cloud processing algorithms (and their parameters) mostly affected the sensitivity and robustness of the branch detection criterion.

The treeQSM toolkit makes several assumptions concerning the input point cloud data (Raumonen and Åkerblom, 2019). One issue is the poor visibility to higher parts of the trees above the live crown base. Also, the needles covering the branches affect the reconstruction of the branches. The methods extracting individual trees from the point clouds of larger forest areas with multiple trees do make errors. Regardless of these shortcomings in our data, the cover-based segmentation method in treeQSM still produced useful branch inventories.

The DBSCAN method produced the segmentation of the point cloud into branches, but as the trunk extraction was not always complete, the search for disconnected components failed to produce individual branches in these parts of the stems. The treeQSM method detected more branches, but in comparison to the X-rayed references, many of the detected branches may either be false positives or so small in size that the X-rays do not detect them as knots. On the other hand, the DBSCAN method only detected larger branches but still underestimated the maximum whorl volumes. In the X-ray analysis, some of the smaller knots (e.g., the inter-nodal branches) may be bundled with the adjacent

Table 2
The results of the direct comparisons of the stem and knot variables proxied from the HHLS point clouds with the X-rayed values.

Feature	Algorithm	Log	Bias	Bias (%)	RMSE	RMSE (%)	R^2	sigma
DBH (cm)	RANSAC	all	−3.650	−1.20 %	1.765	5.78 %	0.89	0.0170
Log volume (m^3)	RANSAC	all	24.363	10.17 %	41.257	17.22 %	0.93	0.0280
		butt	32.900	11.10 %	41.690	14.06 %	0.97	0.0210
		middle	29.458	13.16 %	44.489	19.88 %	0.87	0.0290
		top	5.205	2.99 %	36.063	20.69 %	0.70	0.0300
Knot volume (m^3)	DBSCAN	all	−0.673	−6.72 %	7.087	70.69 %	0.11	0.0060
		butt	0.746	8.04 %	6.265	67.55 %	0.22	0.0050
		middle	−0.341	−3.34 %	6.123	60.02 %	0.24	0.0050
		top	−2.805	−25.67 %	8.857	81.06 %	0.00	0.0060
	QSM	all	2.539	25.32 %	8.618	85.96 %	0.10	0.0050
		butt	3.898	42.02 %	7.535	81.24 %	0.30	0.0040
		middle	5.071	49.71 %	8.806	86.31 %	0.21	0.0050
		top	−2.637	−24.14 %	9.792	89.62 %	0.00	0.0060
Max whorl volume (m^3)	DBSCAN	all	−1.560	−47.56 %	2.345	71.49 %	0.09	0.0020
		butt	−1.323	−42.05 %	2.052	65.23 %	0.18	0.0020
		middle	−1.553	−47.23 %	2.331	70.89 %	0.11	0.0020
		top	−1.862	−53.59 %	2.677	77.07 %	0.03	0.0020
	QSM	all	−1.264	−38.53 %	2.178	66.39 %	0.07	0.0020
		butt	−0.938	−29.81 %	1.835	58.35 %	0.18	0.0020
		middle	−1.176	−35.77 %	2.065	62.79 %	0.10	0.0020
		top	−1.854	−53.36 %	2.715	78.16 %	0.00	0.0020
Mean whorl volume (m^3)	DBSCAN	all	0.012	1.49 %	0.434	52.77 %	0.09	0.0000
		butt	0.195	27.11 %	0.419	58.05 %	0.11	0.0000
		middle	−0.064	−7.45 %	0.411	47.69 %	0.18	0.0000
		top	−0.125	−13.55 %	0.477	51.51 %	0.09	0.0000
	QSM	all	0.001	0.17 %	0.453	55.04 %	0.05	0.0000
		butt	0.195	27.09 %	0.426	59.02 %	0.10	0.0000
		middle	−0.033	−3.84 %	0.421	48.91 %	0.12	0.0000
		top	−0.240	−25.96 %	0.524	56.60 %	0.05	0.0000
Knot index	DBSCAN	all	−0.012	−26.95 %	0.031	70.41 %	0.00	0.0230
		butt	−0.001	−1.94 %	0.018	56.63 %	0.06	0.0110
		middle	−0.006	−13.24 %	0.026	57.85 %	0.02	0.0190
		top	−0.033	−52.54 %	0.046	73.72 %	0.00	0.0260
	QSM	all	0.002	4.00 %	0.034	78.00 %	0.00	0.0220
		butt	0.009	29.78 %	0.021	67.75 %	0.09	0.0120
		middle	0.017	36.83 %	0.035	78.46 %	0.00	0.0180
		top	−0.028	−44.94 %	0.047	75.29 %	0.00	0.0250
Whorl distance (m)	DBSCAN	all	0.017	6.10 %	0.142	52.17 %	0.04	0.0880
		butt	0.008	3.30 %	0.116	45.06 %	0.07	0.0840
		middle	0.007	2.39 %	0.132	46.46 %	0.07	0.0870
		top	0.038	13.75 %	0.177	63.94 %	0.00	0.0900
	QSM	all	−0.080	−29.30 %	0.132	48.71 %	0.04	0.0930
		butt	−0.067	−26.08 %	0.125	48.61 %	0.06	0.0930
		middle	−0.099	−34.87 %	0.141	49.39 %	0.04	0.0910
		top	−0.073	−26.35 %	0.132	47.63 %	0.02	0.0910
Whorl count	DBSCAN	all	−1.384	−11.37 %	5.235	42.99 %	0.08	3.7650
		butt	−2.347	−18.39 %	5.325	41.73 %	0.13	3.8760
		middle	0.153	1.30 %	4.396	37.32 %	0.16	3.5300
		top	−1.970	−16.70 %	5.961	50.53 %	0.01	3.4540
	QSM	all	2.349	19.30 %	6.660	54.70 %	0.07	3.8450
		butt	1.111	8.71 %	4.992	39.12 %	0.16	3.9540
		middle	6.358	53.97 %	7.796	66.17 %	0.23	3.3400
		top	−0.862	−7.31 %	7.233	61.32 %	0.00	3.6090

larger whorls, thus decreasing the number of counted whorls and increasing their perceived volume.

Both methods used the same denoised point clouds of individual trees. The denoising was essential to remove erroneous points from the point cloud and it also had the effect of removing varying amounts of foliage from the trees, which helps the geometric analysis of the woody parts. Although necessary, the denoising process may have a non-negligible result on the branch detection capability of both studied methods, as some parts of the branches may be lost because of the denoising. This is especially the case when the point density varies between trees and within a tree at different heights, which complicates the accurate classification of noisy points. In addition, the varying point densities of different point clouds of individual trees likely impacts the effectiveness of the denoising method, which was unnormalized with respect to the point density. However, the effect on comparability of the two methods with respect to each other should remain minimal, as the input point clouds went through identical denoising processes.

5.3. Comparisons with previous research

The accuracies of the HHLS point clouds were found to be a few percent in the estimation of the stem diameters and volume by Hyyppä et al. (2020a). However, the processing method underestimates the stem diameters at the lowest part of the first log and overestimates them close to the treetop (Hyyppä et al., 2020b). In our current study, the stem extraction was slightly more inaccurate, which was probably due to the more challenging conditions in the Norway spruce-dominated stands and the discrepancy between the over-bark and under-bark stem diameters.

Previous studies on the extraction of the branching patterns from mobile point clouds are few. Puliti et al. (2023) estimated tree height-growth trajectories using UAV laser scanning data and deep learning techniques to detect individual branch whorls with an accuracy better than 10 cm for the estimation of the mean annual height increment. In our study, mean whorl-to-whorl distances within trees were

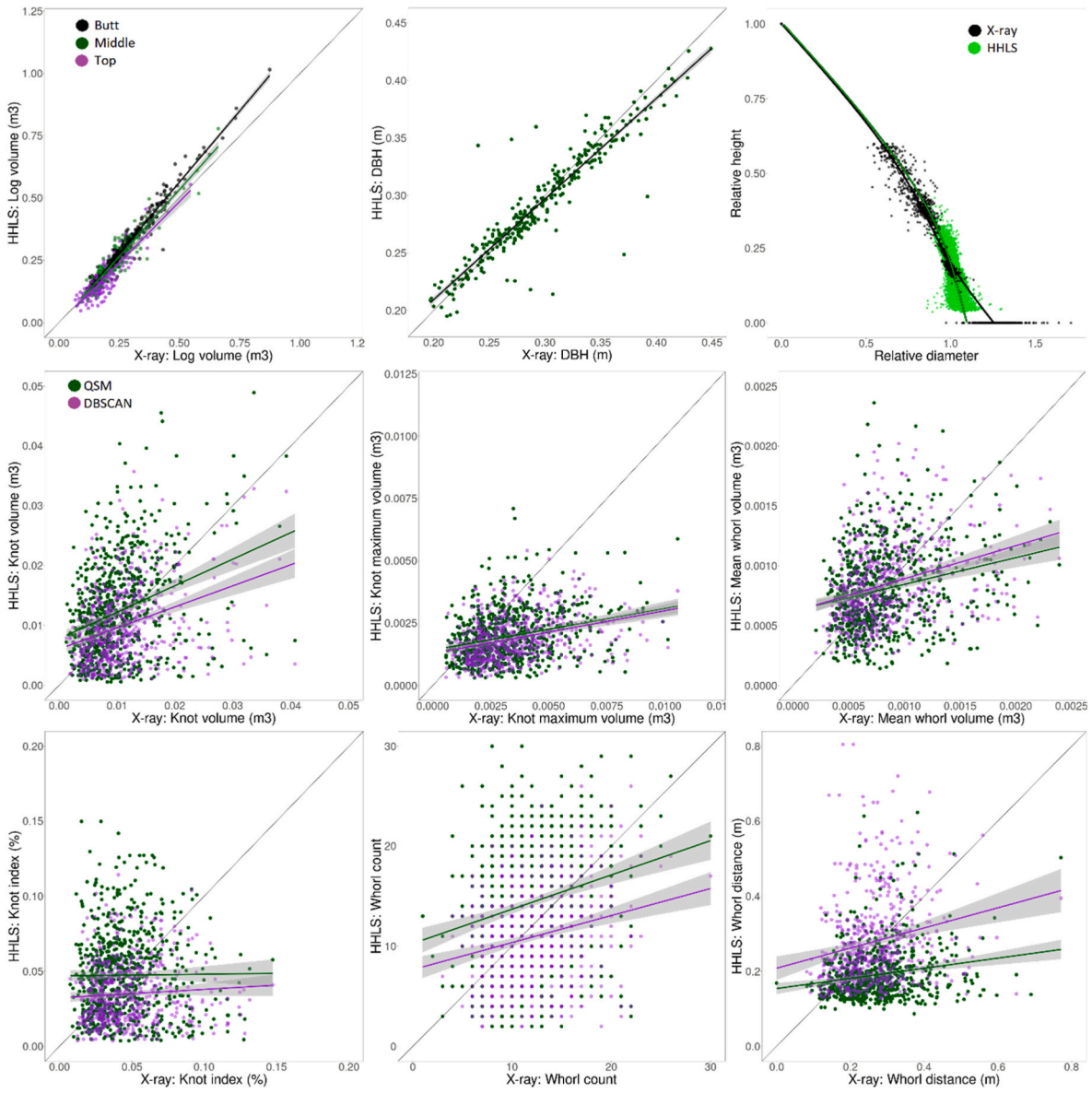


Fig. 7. The results of the direct comparison of the stem and knottiness features proxied from the point clouds against the X-rayed values.

estimated with RMSEs of approximately 13–14 cm (49–52%). Hartley et al. (2022) inferred whorl locations based on the swelling pattern in the stems of Monterey pines (*Pinus radiata* D. Don) in New Zealand: they reported that the locations of the whorls were accurately estimated with this approach compared with the manual measurements, but the detection accuracy remained ca. 40%, slightly lower than what was found, e.g., with a TLS-based whorl extraction when compared with X-ray references by Pyörälä et al. (2018a).

In our current work, we found a 43% RMSE in whorl counts within logs when using the DBSCAN method, which could translate into 50–60% detection accuracy assuming a low number of false positives due to the negative bias. On the other hand, with the treeQSM method we found an overestimate of the whorl numbers in comparison to the X-ray. This higher inaccuracy (55% RMSE) may be due to the false positives, or detection of branches smaller than those measured by the X-rays (e.g., the small internodal branches characteristic for Norway spruce).

The extraction of the branching information has been more thoroughly investigated with TLS point clouds (Dassot et al., 2012; Lau et al., 2018; Pyörälä et al., 2018b; Bournez et al., 2017; Raumonen et al., 2013). Several studies report underestimations of branching volumes, especially with smaller branches (Demol et al., 2022) and with the branches above the live crown base height (Pyörälä et al., 2018b). However, based on our current and previous results, the effects of such underestimations for the prediction of wood quality do not necessarily play any crucial role. Moreover, it is important to remember that the X-ray measurements used here do not represent the absolute truth concerning the explicit knottiness variables (e.g., the number and sizes of individual knots), rather than industrially relevant proxies of their impact on wood quality.

Therefore, the model-based approach was included based on these previous findings, which show that direct comparisons between point clouds and X-raying data are not meaningful. In our study, the Random Forest approach yielded results of moderate accuracy. The branching observations made from terrestrial point clouds need to be considered as samples of wood quality indicators or as predictive features that are used to devise locally calibrated models that indirectly predict the desired branching features. Studies approaching a similar task have been carried out, e.g., by Cote et al. (Côté et al., 2012) and Lau et al. (2019). There are also several other modeling techniques developed to solve the task that relates the wood properties with external (e.g., forest management,

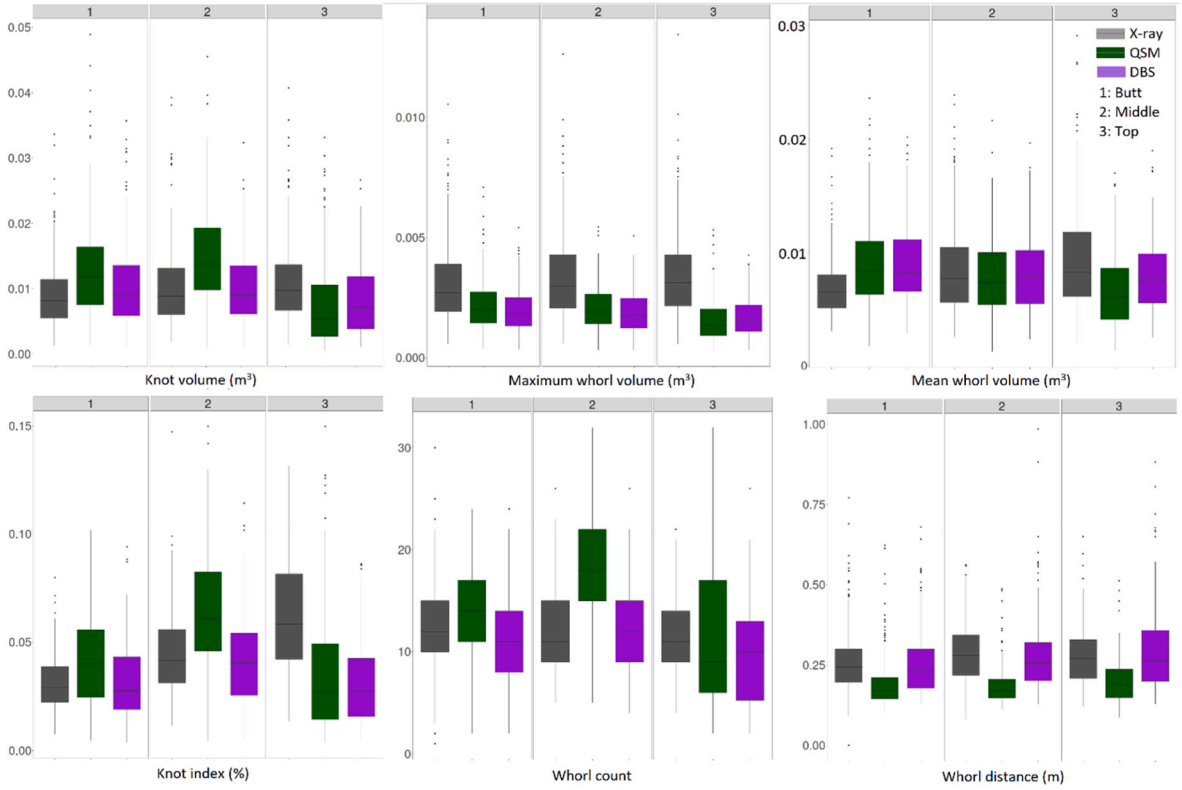


Fig. 8. The variabilities of the X-rayed knot variables and those proxied from the point clouds with QSM and DBSCAN in the logs at different heights.

Table 3
The results of the Random Forest modeling approach for each studied variable: the coefficient of determination (R^2), statistical significance (p) of the model based on the F-test, residual standard error in absolute values and relative to the observed mean (RSE), most important factors in the model based on their effect on mean squared error and root squared error (MIF1 and MIF2, respectively), and number of features in the final model (NoF).

	Response variable	R^2	p	RSE	RSE (%)	MIF1	MIF2	NoF
QSM	Knot volume (m^3)	0.54	<0.01	0.002	18.7	Knot volume	Knot volume	6
	Max whorl volume (m^3)	0.51	<0.01	0.001	16.7	Knot index	Knot volume	7
	Knot index (%)	0.48	<0.01	0.005	11.8	Mean branch diameter	Mean branch diameter	8
	Mean whorl volume (m^3)	0.51	<0.01	0.088	10.6	Knot index	Mean whorl volume	11
	Whorl count	0.52	<0.01	1.172	9.6	Knot volume	Knot volume	6
	Whorl distance (m)	0.47	<0.01	0.03	11.2	Knot volume	Knot volume	4
DBSCAN	Knot volume (m^3)	0.64	<0.01	0.002	16.5	Mean whorl volume	Mean whorl volume	7
	Max whorl volume (m^3)	0.61	<0.01	0.001	15.6	Mean whorl volume	Mean whorl volume	6
	Knot index (%)	0.61	<0.01	0.005	10	Max branch angle	Max branch angle	11
	Mean whorl volume (m^3)	0.59	<0.01	0.109	12.4	Max branch diameter	Mean whorl volume	10
	Whorl count	0.56	<0.01	1.199	9.6	Knot index	Knot volume	7
	Whorl distance (m)	0.61	<0.01	0.023	8.8	Knot index	Knot volume	9

forest type, soil, climate, altitude, wind, and precipitation) and internal (e.g., stem taper, tree height, crown shape and dimensions, the height of the lowest dead branch, maximum branch diameter, and whorl-to-whorl distances) factors. These were reviewed by Drew et al. (2022). There is, however, a lack of sufficient transferability, accuracy, and spatial coverage of the wood quality models, limiting their usability in practical forestry. A sufficient framework for next-generation wood quality estimation should entail unbiased models for the within-tree variability of wood quality, which was transferable to individual trees over large spatial extents. Novel remote and close-range sensing methods may offer a crucial steppingstone for this purpose (Pyörälä et al., 2019a; Van Leeuwen et al., 2011).

5.4. Outlook to practical solutions

In terms of cost efficiency and practical solutions to forestry, our results offer promise that mobile point clouds gathered from, e.g., scanners mounted on harvesters could offer ways to connect meaningful branching features measured in standing timber to industrial data sources. There are still problems to solve before the presented method for wood quality estimations of individual trees would be useful in operational forestry. Automated methods are needed for the traceability of individual tree stems to establish large X-ray datasets to be used as references. Fingerprint matching using images of log ends has been proposed for tracing individual logs from the harvest to the industry (Schraml et al., 2015), which potentially could be used to link X-ray reference data and terrestrial point clouds. Data acquisition with mobile systems performed prior to harvesting operations would provide information useful for forest management planning but would further extend the traceability problem to also include matching tree positions measured by laser scanning with tree positions measured during the harvest, which could be solved using high-precision positioning systems on modern harvesters. Moreover, the development of faster algorithms for branch detection (e.g., based on deep learning) would enable

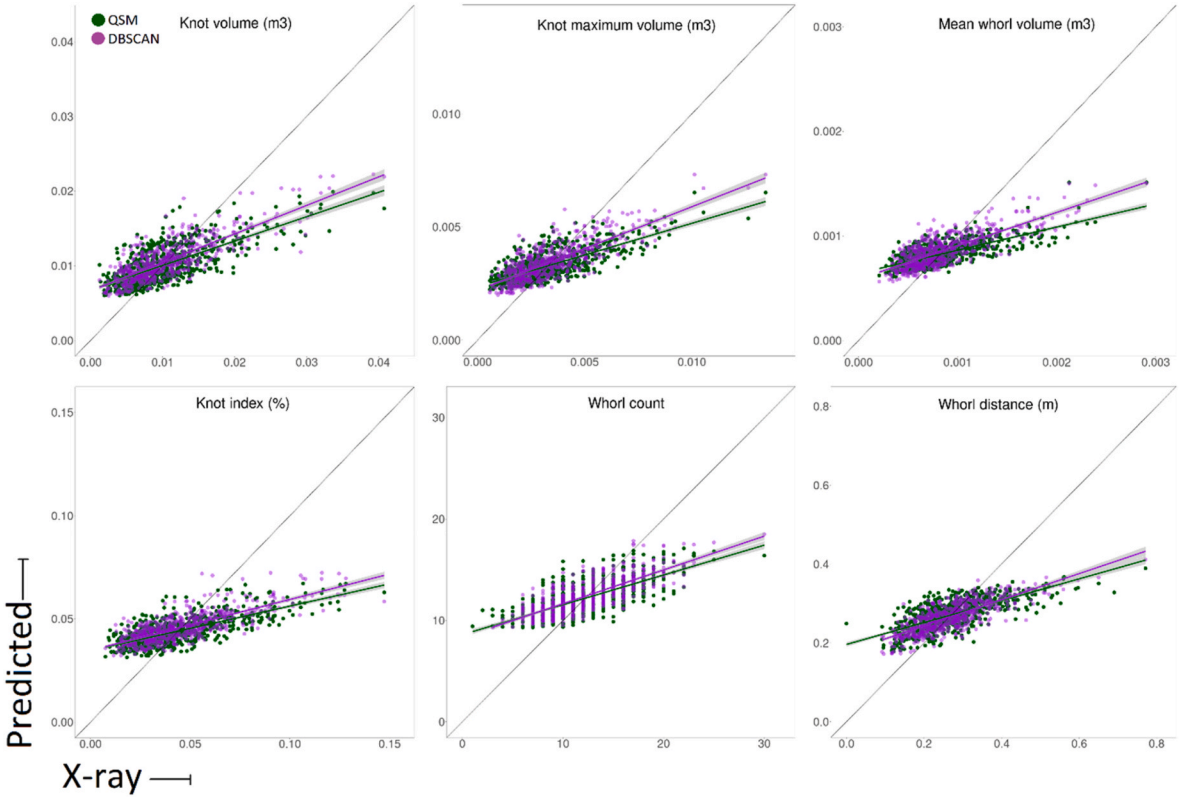


Fig. 9. The results of the Random Forest model approach. Predicted (y-axis) and X-rayed values (x-axis), colors indicate the feature space used as predictors, i.e., the branching features extracted from the point clouds with the quantitative structure model (QSM) and density-based clustering (DBSCAN). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

real-time estimations of optimal quality-bucking decisions during the harvests. An example of a potential real-time solution for point cloud classification was given by Kaijaluoto et al. (2022). The methods should be able to work also during poor weather conditions (snow/rain) and despite the disturbing debris generated by the harvesting process.

6. Conclusions

In conclusion, we demonstrated that mobile laser scanning data, accompanied by geometrical point cloud processing, and linked with within-stem wood properties measured at sawmills with X-ray scanners, can facilitate more extensive collection of wood-quality related data along the forest value chain. The main advantage of the technology used is that such laser scanners can be mounted on a harvester or an under-canopy flying drone, to enable the fast collection of point clouds of the standing timber. The under-canopy point clouds capture stem branching data that are relevant to predicting wood quality attributes used at sawmills, but do not provide data that would be directly comparable to the X-ray features. Possible outlooks for the future include accompanying the terrestrial point clouds with dense within- or above-canopy point clouds to overcome the limitations that are due to the perspective. Moreover, the traceability issue, i.e., the problems related to matching the logs at the sawmill to the corresponding data from the standing timber require further investigations. Once we can trace logs with sufficient precision, we will be able to collect vast amounts of reference data, which will pave the way for deep learning algorithms.

Author contributions

Olli Winberg (OW), Jiri Pyörälä (JP), Xiaowei Yu (XY), Harri Kaartinen (HK), Antero Kukko (AK), Markus Holopainen (MH), Johan Holmgren (JH), Matti Lehtomäki (ML), Juha Hyyppä (JHy):

OW, JP, XY, HK and JHy wrote the first version of the manuscript. OW developed the DBSCAN-based method and modified the treeQSM method for the study at hand. ML helped with and supervised the method development. XY supervised the method development and processed the data. JP conducted the data fusion with the sawmill data and carried out the statistical analyses. HK and AK obtained and pre-processed the handheld laser scanning data. JP selected the study stands and sample trees, collected the field references, and coordinated the harvests and transportation, all in close cooperation with Stora Enso. AK, MH, JH, and JHy created the original project plans and managed the projects. All authors participated in finalizing and revising the manuscript.

Declaration of competing interest

The authors declare no conflicting interests.

Acknowledgements

This work was part of the Tandem forest value research project “Estimating Forest Resources and Quality-related Attributes Using Automated Methods and Technologies (ForestQuality)”, with financing from the Academy of Finland, the Kempe foundations, and Stora Enso. The work was also part of the Academy of Finland projects “Feasibility of Inside-Canopy UAV Laser Scanning for Automated Tree Quality Surveying” (Quality4Trees) and “Capturing structural and functional diversity of trees and tree communities for supporting sustainable use of forests” (Diversity4Forests) and the Ministry of Agriculture and Forestry project “Future forest information system at individual tree level”. We gratefully acknowledge the professionals from Stora Enso who contributed to this study: Pekka Alajärvi, Mika Korvenranta (currently with CollectiveCrunch oy), Kalle Kärhä (currently with the University of

Eastern Finland), and all the personnel involved in the harvesting, transportation, and sawmilling that made this work possible. We also thank Mika Pehkonen (currently with the Finnish Geospatial Institute and the University of Helsinki) for help in the field work and in the pre-processing of the sawmill data.

Appendix

Table A1
Parameters, and their values as set in the application of the treeQSM algorithm in this paper, with short explanations. Refer to the treeQSM manual (Raumonen and Åkerblom, 2019) for further details.

Parameter name	Value used	Short explanation of the parameter
PatchDiam	0.03–0.05	Determines the size of open sets covering the point cloud (m)
BallRad	0.07	The radius of spherical patch used in cover generation (m)
nmin	10	Minimum number of points in BallRad-balls
OnlyTree	0	Is 0 if point cloud contains points not only from a single tree
Tria	0	Is 0 if no triangulation is produced
Dist	1	Is 1 if point-model distances are computed

References

Balenović, I., et al., 2021. Hand-held personal laser scanning–current status and perspectives for forest inventory application. Croat. J. For. Eng.: Journal for Theory and Application of Forestry Engineering 42 (1), 165–183.
Bauwens, S., et al., 2016. Forest inventory with terrestrial LiDAR: a comparison of static and hand-held mobile laser scanning. Forests 7 (6), 127.
Bauwens, S., et al., 2017. Terrestrial photogrammetry: a non-destructive method for modelling irregularly shaped tropical tree trunks. Methods Ecol. Evol. 8 (4), 460–471.
Beaulieu, J., Dutilleul, P., 2019. Applications of computed tomography (CT) scanning technology in forest research: a timely update and review. Can. J. For. Res. 49 (10), 1173–1188.
Bienert, A., et al., 2018. Comparison and combination of mobile and terrestrial laser scanning for natural forest inventories. Forests 9 (7), 395.
Bienert, A., et al., 2021. Automatic extraction and measurement of individual trees from mobile laser scanning point clouds of forests. Ann. Bot. 128 (6), 787–804.
Blanchette, D., et al., 2015. Predicting wood fiber attributes using local-scale metrics from terrestrial LiDAR data: a case study of Newfoundland conifer species. For. Ecol. Manag. 347, 116–129.
Bournez, E., et al., 2017. From TLS point clouds to 3D models of trees: a comparison of existing algorithms for 3D tree reconstruction. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 42 (2), 113–120.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Cabo, C., et al., 2018. Comparing terrestrial laser scanning (TLS) and wearable laser scanning (WLS) for individual tree modeling at plot level. Rem. Sens. 10 (4), 540.
Čerňava, J., et al., 2019. Processing chain for estimation of tree diameter from GNSS-IMU-based mobile laser scanning Data. Rem. Sens. 11 (6), 615.
Chen, S., et al., 2019. Applicability of personal laser scanning in forestry inventory. PLoS One 14 (2), e0211392.
Côté, J.-F., et al., 2012. A fine-scale architectural model of trees to enhance LiDAR-derived measurements of forest canopy structure. Agric. For. Meteorol. 166 (1), 72–85.
Côté, J.-F., et al., 2021. Assessing the impact of fine-scale structure on predicting wood fibre attributes of boreal conifer trees and forest plots. For. Ecol. Manag. 479, 118624.
Dassot, M., et al., 2012. Terrestrial laser scanning for measuring the solid wood volume, including branches, of adult standing trees in the forest environment. Comput. Electron. Agric. 89, 86–93.
Del Perugia, B., et al., 2019. Influence of scan density on the estimation of single-tree attributes by hand-held mobile laser scanning. Forests 10 (3), 277.
Demol, M., et al., 2022. Volumetric overestimation of small branches in 3D reconstructions of Fraxinus excelsior. Silva Fenn. 56 (1).
Drew, D., et al., 2022. A review of progress and applications in wood quality modelling. Current Forestry Reports 1–16.
Ester, M., DBSCAN for discovering clusters, 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Ester, M., Kriegel, H.-P., Sander, J., Xu, X. (Eds.), 1996, Proceedings of the Second International C K M (K'96).
Gao, S., et al., 2017. A critical analysis of methods for rapid and nondestructive determination of wood density in standing trees. Ann. For. Sci. 74 (2), 1–13.
Hackenberg, J., et al., 2015. SimpleTree-an efficient open source tool to build tree models from TLS clouds. Forests 6 (11), 4245–4294.
Hartley, R.J., et al., 2022. Assessing the potential of backpack-mounted mobile laser scanning systems for tree phenotyping. Rem. Sens. 14 (14), 3344.
Hilker, T., et al., 2013. Prediction of wood fiber attributes from LiDAR-derived forest canopy indicators. For. Sci. 59 (2), 231–242.
Hyyppä, E., et al., 2020a. Comparison of backpack, handheld, under-canopy UAV, and above-canopy UAV laser scanning for field reference data collection in boreal forests. Rem. Sens. 12 (20), 3327.
Hyyppä, E., et al., 2020b. Accurate derivation of stem curve and volume using backpack mobile laser scanning. ISPRS J. Photogrammetry Remote Sens. 161, 246–262.
Hyyppä, E., et al., 2020c. Under-canopy UAV laser scanning for accurate forest field measurements. ISPRS J. Photogrammetry Remote Sens. 164, 41–60.
Hyyppä, J., et al., 2021. Under-canopy UAV laser scanning providing canopy height and stem volume accurately. Forests 12 (7), 856.
Hyyppä, E., et al., 2022. Direct and automatic measurements of stem curve and volume using a high-resolution airborne laser scanning system. Science of Remote Sensing 5, 100050.
Kaartinen, H., et al., 2015. Accuracy of kinematic positioning using global satellite navigation systems under forest canopies. Forests 6 (9), 3218–3236.
Kaijaluoto, R., et al., 2022. Semantic segmentation of point cloud data using raw laser scanner measurements and deep neural networks. ISPRS Open Journal of Photogrammetry and Remote Sensing 3, 100011.
Lachenbruch, B., Moore, J.R., Evans, R., 2011. Radial variation in wood structure and function in woody plants, and hypotheses for its occurrence. In: Size-and Age-Related Changes in Tree Structure and Function. Springer, Dordrect, Germany, pp. 121–164.
Lau, A., et al., 2018. Quantifying branch architecture of tropical trees using terrestrial LiDAR and 3D modelling. Trees (Berl.) 32 (5), 1219–1231.
Lau, A., et al., 2019. Estimating architecture-based metabolic scaling exponents of tropical trees using terrestrial LiDAR and 3D modelling. For. Ecol. Manag. 439, 132–145.
Liang, X.L., et al., 2014. Automated stem curve measurement using terrestrial laser scanning. IEEE Trans. Geosci. Rem. Sens. 52 (3), 1739–1748.
Lindgren, L., 1991. Medical CAT-scanning: X-ray absorption coefficients, CT-numbers and their relation to wood density. Wood Sci. Technol. 25 (5), 341–349.
Luther, J.E., et al., 2014. Predicting wood quantity and quality attributes of balsam fir and black spruce using airborne laser scanner data. Forestry 87 (2), 313–326.
Morgan, C.J., Powers, M., Strimbu, B.M., 2022. Estimating tree defects with point clouds developed from active and passive sensors. Rem. Sens. 14 (8), 1938.
Nguyen, V.-T., Constant, T., Colin, F., 2021. An innovative and automated method for characterizing wood defects on trunk surfaces using high-density 3D terrestrial LiDAR data. Ann. For. Sci. 78 (2), 1–18.
Oja, J., et al., 2003. Automatic grading of Scots pine (Pinus sylvestris L.) sawlogs using an industrial X-ray log scanner. Comput. Electron. Agric. 41 (1), 63–75.
Puliti, S., et al., 2023. Tree height-growth trajectory estimation using uni-temporal UAV laser scanning data and deep learning. Forestry 96 (1), 37–48.
Pyörälä, J., et al., 2018a. Comparison of terrestrial laser scanning and X-ray scanning in measuring Scots pine (Pinus sylvestris L.) branch structure. Scand. J. For. Res. 33 (3), 291–298.
Pyörälä, J., et al., 2018b. Assessing branching structure for biomass and wood quality estimation using terrestrial laser scanning point clouds. Can. J. Rem. Sens. 44 (5), 462–475.
Pyörälä, J., et al., 2018c. Quantitative assessment of Scots pine (Pinus sylvestris L.) whorl structure in a forest environment using terrestrial laser scanning. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 11 (10), 3598–3607.
Pyörälä, J., et al., 2019a. Variability of Wood Properties Using Airborne and Terrestrial Laser Scanning. Remote Sensing of Environment, p. 235.
Pyörälä, J., et al., 2019b. Assessing log geometry and wood quality in standing timber using terrestrial laser-scanning point clouds. Forestry 92 (2), 177–187.
Qi, Y., et al., 2022. Comparing tree attributes derived from quantitative structure models based on drone and mobile laser scanning point clouds across varying canopy cover conditions. ISPRS J. Photogrammetry Remote Sens. 192, 49–65.
Raumonen, P., Åkerblom, M., 2019. InverseTampere/TreeQSM:, Version 2.3. 1.
Raumonen, P., et al., 2013. Fast automatic precision tree models from terrestrial laser scanner data. Rem. Sens. 5 (2), 491–520.
Schimleck, L., et al., 2019. Non-destructive evaluation techniques and what they tell us about wood property variation. Forests 10 (9), 728.

Schraml, R., et al., 2015. Towards the applicability of biometric wood log traceability using digital log end images. Comput. Electron. Agric. 119, 112–122.
Van Leeuwen, M., et al., 2011. Assessment of standing wood and fiber quality using ground and airborne laser scanning: a review. For. Ecol. Manag. 261 (9), 1467–1478.
Vandendaele, B., et al., 2022. Mobile laser scanning for estimating tree structural attributes in a temperate hardwood forest. Rem. Sens. 14 (18), 4522.
Wang, Y., et al., 2021. Seamless integration of above-and under-canopy unmanned aerial vehicle laser scanning for forest investigation. Forest Ecosystems 8 (1), 1–15.
Wei, Q., Leblon, B., La Rocque, A., 2011. On the use of X-ray computed tomography for determining wood properties: a review. Can. J. For. Res. 41 (11), 2120–2140.