

Development of a Precise Tree Structure from LiDAR Point Clouds

Abdul Nurunnabi^{1,2}, Felicia Teferle¹, Debra F Laefer³, Meida Chen⁴, Mir Masoom Ali⁵

¹ Geodesy and Geospatial Engineering, Faculty of Science, Technology and Medicine, University of Luxembourg

² Institute for Advanced Studies, University of Luxembourg, Luxembourg

6, rue Richard Coudenhove-Kalergi, L-1359, Luxembourg – (abdul.nurunnabi, rebecca.teferle)@uni.lu

³ Center for Urban Science + Progress; Department of Civil and Urban Engineering, Tandon School of Engineering,
New York University, NY, USA – debra.laefer@nyu.edu

⁴ Institute for Creative Technologies, University of Southern California, Los Angeles, CA 90094, USA – mechen@ict.usc.edu

⁵ Department of Mathematical Sciences, Ball State University, Muncie, IN 47306, USA – mali@bsu.edu

Keywords: Biomass, Forest, Geometric Feature, Leaf-Wood Separation, Segmentation, Tree Information Modeling

Abstract

A precise tree structure that represents the distribution of tree stem, branches, and leaves is crucial for accurately capturing the full representation of a tree. Light Detection and Ranging (LiDAR)-based three-dimensional (3D) point clouds (PCs) capture the geometry of scanned objects including forests stands and individual trees. PCs are irregular, unstructured, often noisy, and contaminated by outliers. Researchers have struggled to develop methods to separate leaves and wood without losing the tree geometry. This paper proposes a solution that employs only the spatial coordinates (x, y, z) of the PC. The new algorithm works as a filtering approach, utilizing multi-scale neighborhood-based geometric features (GFs) e.g., linearity, planarity, and verticality to classify linear (wood) and non-linear (leaf) points. This involves finding potential wood points and coupling them with an octree-based segmentation to develop a tree architecture. The main contributions of this paper are (i) investigating the potential of different GFs to split linear and non-linear points, (ii) introducing a novel method that pointwise classifies leaf and wood points, and (iii) developing a precise 3D tree structure. The performance of the new algorithm has been demonstrated through terrestrial laser scanning PCs. For a Scots pine tree, the new method classifies leaf and wood points with an overall accuracy of 97.9%.

1. Introduction

Understanding tree topology, including the separation of leaf and woody materials (e.g., stem, branches, and twigs) has important implications for forest-urban management, ecosystem functions, and tree information modeling (Wang et al., 2020; Hui et al., 2021; Shu et al., 2022; Tian and Li, 2022). This involves estimation of precise tree parameters such as diameter at breast height (DBH) [Nurunnabi et al., 2024], canopy profile (Zhao et al., 2015), leaf angle (Zhao et al., 2015), stem volume, and leaf area index (LAI) [Beland et al., 2014; Li et al., 2017]. Additionally, it encompasses the quantification of above-ground biomass (AGB), and carbon sequestration capacity (Claders et al., 2015). Image-based methods are not well-suited for accurately assessing the 3D geometry of objects because objects may appear mixed or overlapped within two-dimensional (2D) image pixels, thereby limiting the ability to accurately assess geometry. Conversely, Light Detection and Ranging (LiDAR)-based scanning systems—Terrestrial, Mobile, and Aerial Laser Scanning (TLS, MLS, and ALS)—provide 3D georeferenced points called point clouds (PCs) that capture detailed geometry of the scanned objects. TLS in forestry and ecology allow individual tree parameters estimation and forest stands structural measurements (Raumonen et al., 2013; Zhao et al., 2015; Tian and Li, 2022). PC processing is primarily challenging due to its unstructured, unlabelled, and inhomogeneous nature, and the presence of outliers (Nurunnabi et al., 2012). Additional difficulties emerge from the mixed and occlusion effects caused by finer branches, tiny twigs, and leaves (Vicari et al., 2019; Hui et al., 2021).

Many existing tree PC processing methods have included manual components that are laborious, time-consuming, and non-scalable (Vicari et al., 2019). A significant ongoing challenge among these remaining obstacles is the separation of leaf points

from wood points (Disney et al., 2018). Existing methods commonly utilize radiometric (Béland et al., 2011, 2014) and/or geometric features (GFs) [Ma et al., 2016; Zhu et al., 2018].

Radiometric features pertain to the intensity and reflectivity of laser returns at specific wavelengths. The availability and reliability of radiometric information, which encompass color and intensity, may be impacted by ambient lighting. The intensity of laser points is contingent upon variations in their optical properties at the LiDAR sensor's wavelength (Béland et al., 2014). Furthermore, the intensity of laser returns is sensitive to both distance and the incidence angle to the scanned objects, making radiometric calibration of intensity challenging (Côté, et al., 2011). In fact, Tao et al. (2015) asserted that the intensity-based approach is unsuitable for dense canopies due to the potential generation of misleading intensity values from partial hits. The presence of a dense canopy can also increase the occurrence of mixed returns. In contrast, GF-based methods mainly consider the coordinates (x, y, z) of the points (including partial returns) and utilize the underlying geometric patterns (features) of points based on their local neighbors. When compared with radiometric feature-based approaches, geometric feature-based approaches often yield better classification accuracy (Vicary et al., 2019; Wang et al., 2020). Some methods use a combination of both radiometric and geometric features to leverage the respective benefits of each feature (Zhu et al., 2018). Grau et al. (2017) pointed out that when utilizing point clouds in a volumetric sense, considering the mixture of materials can be critically important.

With complex tree structures, many methods fail to expose finer branching structures and misclassify tiny twigs as leaf points (Hui et al., 2021). The mislabelling of twigs as leaves may inaccurately inflate AGB and LAI values (Claders et al., 2015). To date, researchers have struggled to develop methods that

efficiently distinguish leaf from wood points without sacrificing geometric accuracy (Wang et al., 2020; Hui et al., 2021). Hence, the primary hurdles in this endeavor include ensuring sufficient accuracy, robustness, and reliability.

The approach proposed herein avoids radiometric information, employs only GFs at a multi-scale neighborhood level. This involves finding potential wood points and coupling them with an octree-based segmentation to develop the tree architecture. The main scientific contributions of this paper are (i) benchmarking potential GFs to separate linear and non-linear points, and (ii) developing a pointwise method to classify tree leaf and wood points for a detailed tree architecture.

The remaining paper has 4 sections. Section 2 briefly presents related literature. Section 3 introduces the algorithmic methodology. Section 4 applies the algorithm with two real-world TLS data sets, and Section 5 concludes the paper.

2. Related Literature

As previously mentioned, features for leaf-wood separation and 3D tree structure development using PCs have been radiometric feature-based, geometric feature-based, or a combination. One of the first efforts was done by Côté et al. (2009), who aimed to reconstruct a 3D tree architecture by segmenting the PC between wood and leaf components based on intensity information. Béland et al. (2011, 2014) used the intensity returns for wood-leaf separation.

An early notable GF-based contribution is the work by Raunonen et al. (2013) who focused on tree branch segmentation via flexible cylindrical model to represent both trunk and branches elements. The reconstruction utilizes neighbor relations, as well as geometrical properties. Contemporaneously, Belton et al. (2013) applied GFs for developing a Gaussian mixture model (GMM) that classifies tree points into leaves, trunks and unknown. Subsequently, Tao et al. (2015) developed a geometric-based wood-leaf separator and favorably benchmarked it against an intensity-based approach. Continuing in that spirit, Ma et al. (2016) calculated a set of GFs for another GMM for the classification of photosynthetic (i.e., leaf, flowers, etc.) and non-photosynthetic (wood; stem, branches, etc.) components. Similarly, Wang et al. (2017) assessed 26 GFs via 4 machine learning methods: support vector machine, naive Bayes, random forest (RF), and GMM. The authors concluded that RF was the most accurate classifier. Similar efforts were undertaken by Ferrara et al. (2018) used GFs and the density-based clustering algorithm, DBSCAN. Zhu et al. (2018) employed an adaptive neighborhood search algorithm as an alternative approach using multiple radiometric and geometric features for classification using a RF classifier. They concluded that GFs are more important than radiometric features.

More recently, a generic and fully automatic leaf-wood separation (LeWoS) procedure was introduced by Wang et al. (2020) utilizing GFs. The work employed a set of recursive point cloud segmentation and regularization procedures.

In another novel approach, Tian and Li (2022) developed a graph-based leaf-wood separation (GBS) method. GBS utilized the shortest path-based features in tree point cloud segmentation, cluster recognition and region growing (1) using a weighted undirected connected graph that is constructed based on a single

tree point cloud; (2) then segmenting that into homogeneous clusters at multiple scales; (3) next finding cylindrical and/or linear characteristics; and (4) using points on those elements as initial seed points to find irregular clusters collected using a region-growing-based approach. To demonstrate their method, multiple tree species and sizes were tested.

Identifying the shape and topological structure of a tree has also been a longstanding topic of interest. Early on, Bucksh and Lindenberg (2008) proposed an octree-graph based method for tree skeleton extraction, while Fu et al. (2020) employed level sets and cylindrical shapes. An approach developed by Li et al. (2022) to generate refined tree skeleton include separation, over-segmentation, and shrinkage of main branches. Notably, Hui et al. (2021) developed a hybrid method that combines graph-based and GF-based methods.

Recently, deep learning (DL) methods (Garcia-Garcia et al., 2017) have emerged for various PC processing tasks, such as segmentation and classification, which hold significant implications in forest analysis. Zou et al. (2017) introduced a DL-based algorithm for tree classification, leveraging low-level feature representation through voxel-based rasterization. Xi et al. (2018) employed a 3D fully convolutional neural network for filtering tree stems and branches.

3. Methodology

Herein a 3-step algorithm for wood and leaf separation is proposed. The first two steps aim to overcome problems related to ground points, noise and outliers. The flowchart of the new algorithm is shown in Fig. 1.

3.1 Step 1. Ground points filtering

For this, the ground filtering algorithm proposed by Nurunnabi et al. (2016) is employed, which couples robust statistical approaches with locally weighted regression. For this method, the data set is divided into manageable strips. The algorithm is executed on each strip, and the results are subsequently merged. For each strip, the algorithm operates iteratively on two orthogonal profiles (x - z and y - z). The iterative process consists of two main steps. In the first step, robust locally weighted regression (RLWR) is employed locally to obtain a nonlinear fit for the entire strip. The second step involves iterative fitting for each strip, where the z (elevation) values are down-weighted based on the residuals of the previous fit. The process of fitting and down-weighting continues until it reaches the ground level, and the ground points are identified as those that are close to the ground level within a specified threshold. This method is highly robust in the presence of outliers. The reader is referred to Nurunnabi et al. (2016) for additional details.

3.2 Step 2. Denoising and outliers removal

Denoising and outlier removal are achieved by applying a statistically robust algorithm developed in Nurunnabi et al. (2015). In that, a maximum consistency with minimum distance (MCMD)-based Z-score is employed as a well-known distance-based statistical measure that often used to identify outliers. It locates outliers locally that are significantly distant from the maximum consistent set (MCS) within a local neighborhood. The MCS was developed by utilizing a PCA-based plane fitting. For more detailed information about the algorithm, the reader can refer to the original paper (Nurunnabi et al., 2015).

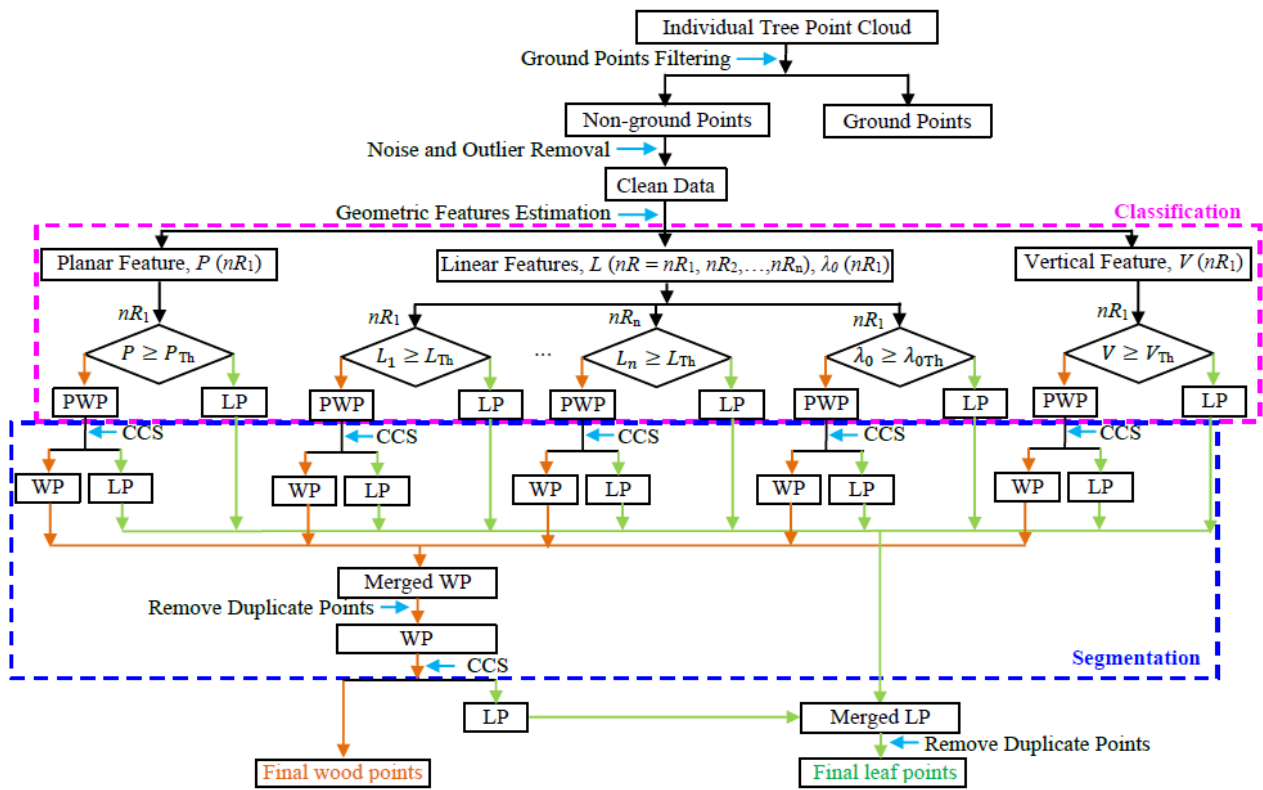


Figure 1. The flowchart of the proposed tree structure development algorithm. L = linearity, P = planarity, V = verticality, λ_0 = 3rd eigenvalue, nR = neighborhood Radius, $*_{th}$ = threshold, PWP = potential wood points, WP = wood points, LP = leaf points, CCS = connected components segmentation.

3.3 Step 3. Identification, classification and reconstruction of linear (wood) and non-linear (leaf) structures

This is the main component of the proposed algorithm and the predominant contribution of this paper. This step includes separating wood (linear) and leaf (non-linear) points to uncover the distribution of tree stem and branches and, ultimately, the tree skeleton. The tree skeleton can be articulated as a non-linear structure, which is mainly a combination of the tree stem and branches that are actually linear structures within a local region. To establish this, an iterative process of (i) classification and (ii) segmentation is employed.

3.3.1 Task (i), Geometric feature estimation and classification: Principal Component Analysis (PCA) is a well-known (Brodu and Logue, 2012; Böhm, et al., 2016) simple statistical method that can identify the related directions of points (in a cloud) in a local neighborhood. Here PCA is performed on the covariance matrix (C) of the local neighbors for each point of a tree PC (P). The matrix C of k points in a neighborhood of a point of interest $p_i (x, y, z) \in P$ can be defined as:

$$C_{3 \times 3} = \frac{1}{k} \sum_{i=1}^k (p_i - \bar{p})(p_i - \bar{p})^T, \quad (1)$$

where $\bar{p} = \frac{1}{k} \sum_{i=1}^k p_i$ is the mean of the points in the local neighborhood. To perform PCA, the following eigenvalue equation (Eq. 2) is solved using the well-known singular value decomposition (SVD) technique,

$$\lambda V = CV, \quad (2)$$

where V is a matrix containing the eigenvectors as its columns, and λ is a diagonal matrix that contains eigenvalues in its

diagonal. The results yield three eigenvectors v_2, v_1 and v_0 , along with three respective eigenvalues λ_2, λ_1 and λ_0 ($\lambda_2 \geq \lambda_1 \geq \lambda_0 \geq 0$). The eigenvectors are known as the principal components. They represent the directions of maximum variance in the data (neighbors). These directions with corresponding eigenvalues can correspond to the main GFs. The 1st principal component (v_2) explains the direction of the most data variability, and the 2nd and 3rd principal components progressively less. Thus, PCA highlights the dominant patterns and structures within a PC. Fig. 2 explains how, a PC geometrically appears at a specific position within a local neighborhood (with a given scale), indicating whether a point resembles a volumetric shape, a line, or a plane. If all three eigenvalues are relatively equal, the points are spread out almost evenly in all directions, suggesting a volumetric or spherical shape. If one eigenvalue (λ_2) is significantly larger than the others (λ_1 and λ_0), it indicates that points are more along one direction, thus suggesting a linear shape. If two eigenvalues (λ_2 and λ_1) are significantly larger than the third (λ_0), it suggests that points are concentrated along a plane, indicating a planar structure.

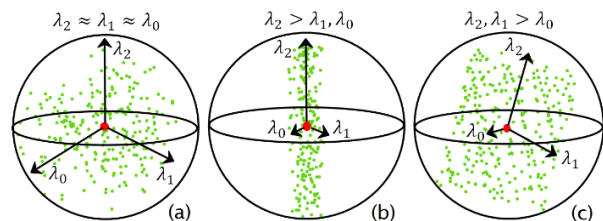


Figure 2. (a) a spherical neighbourhood, points are scattered, (b) points represent a linear surface, and (c) points represent a planar surface.

We surveyed the relevant literature (e.g., Hackel et al., 2016; Ma et al., 2016; Nurunnabi et al., 2021) to identify the most informative GFs for detecting points that predominantly exhibit linear patterns, such as those found for wood (stem and branches) points. By employing formulas for GFs, namely linearity (L), planarity (P), and verticality (V) as outlined in Table 1, we are able to effectively extract wood points representing linear, planar, or vertical structures.

Geometric Feature (GF)	Definition
Linearity, L	$(\lambda_2 - \lambda_1)/\lambda_2$
Planarity, P	$(\lambda_1 - \lambda_0)/\lambda_2$
Scatteredness, S	λ_0/λ_2
Verticality, V	$1 - \langle [0 \ 0 \ 1], v_0 \rangle $

Table 1. Definitions of geometric features (GFs).

Brodu and Lague (2012) noted that by integrating information from various scales, creates signatures capable of identifying categories of objects. Wang et al. (2015) demonstrated that multiscale features effectively extract shape features from complex PCs. They asserted that multiscale and hierarchical features are discriminative and robust, particularly when dealing with PCs exhibiting varying point density and missing data.

Due to the variation in tree species, branch/stem sizes, and point density, we estimate GFs based on multiple sizes of the neighborhood Radius (nR). In Fig. 3 we have a motivating finding: points on a stem (which is typically linear) can be represented as linear/planar and vice versa, depending on their position and the size of the local neighborhood. Although we consider L , P , and V , our primary focus is on L . We estimate multi-scale neighborhood based L values to extract the majority of wood points from stems and branches. Fig 4. illustrates that how L values from multi-scale neighborhoods show linear structure from different parts of a tree.

We consider multiple (n) sizes of neighborhood radii, and experienced 5 to 7 different sizes of radii is good enough to collect the wood points from different parts (stump/stem/branch) of the tree. By the definitions of L , P and S in Table 1, we have $L+P+S = 1$. Points with significantly higher values of L (or P) compared to others will be categorized as linear (or planar) accordingly. We classify a point as linear, planar, or vertical if their corresponding values of L , P , or V exceed their respective thresholds ($*_{Th}$), indicating $L \geq L_{Th}$, $P \geq P_{Th}$ or $V \geq V_{Th}$, and consider them potential wood points (PWP). In this paper, we fix $L_{Th} = P_{Th} = V_{Th} = 0.60$.

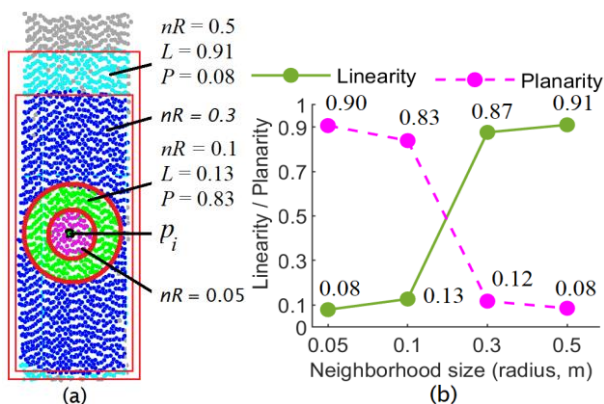


Figure 3. (a) A point p_i on a part of a stem (linear structure) is identified as a planar point with $nR = 0.1m$ (green area), and as a linear point when $nR = 0.5m$ (cyan area), (b) line diagram shows impact of neighborhood size on GFs.

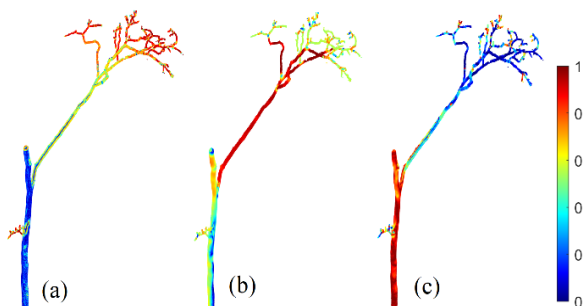


Figure 4. Search for wood (linear/planar/vertical surface) points, (a) linear (red) points from 3rd level branch; $nR=0.05m$, (b) linear (red) points from 1st level branch; $nR=0.3m$, (c) vertical (red) points from stem; $nR=0.1m$. Points on a finer branch shows as linear with a neighborhood of a smaller radius, whereas a point on a heavier branch is revealed as linear when a larger neighborhood is counted.

To investigate any potential relationship between nR and DBH, we selected 15 trees (Pine, Oak and Far) with varying DBH sizes, ranging from 15.1cm to 86.7cm. To ensure consistency in point density, we subsampled the data with a 0.05m point spacing. Setting a planarity threshold of greater than 0.70, we identified the sizes of nR for which a point around the DBH of a tree can be classified as a planar point. We generated a line diagram based on the estimated nRs versus respective DBHs. Fig. 5 illustrates that the size of nR has a positive correlation with the DBH of a tree when identifying planarity. Therefore, we posit that DBH can serve as a suitable reference for determining an appropriate size of nR . Users can fix the sizes of nR based on their specific dataset requirements.

Concurring with existing literature, we note that wood points are predominantly identified by L , P , and/or V . Furthermore, we discover that the 3rd eigenvalue, λ_0 (which quantifies variation along a point normal, v_0) can aid in identifying wood points, particularly for finer branches. We integrate λ_0 with a very small threshold ($\lambda_{0Th} \leq 0.0005$) to identify wood points that might originate from the 3rd and 4th levels of branches.

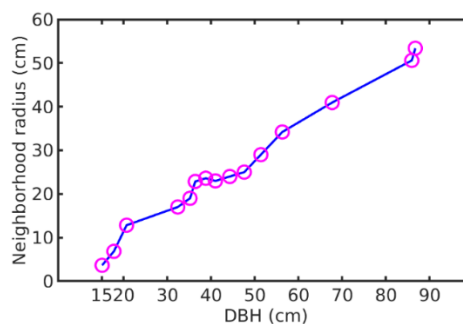


Figure 5. Line diagram of neighborhood radius, nR versus DBH for 15 trees.

3.3.2. Task (ii), Segmentation and refinement: We identify potential wood points (PWPs): linear with multi-scale neighborhood, nR_1, \dots, nR_n , and using λ_0 values with nR_1 , also planar and vertical points with nR_1 in Task (i). For each nR , after collecting potential wood points (linear, planar and vertical), we perform an octree (voxel)-based connected component segmentation (CCS) [e.g., Vo et al., 2015] to find connected components and filter out leaf points that are wrongly identified as wood points, logically leaf points are scattered and mostly disconnected as a linear component. Identifying PWP before implementing CCS facilitates breaking the connection between wood and leaf points. Otherwise, the strong nexus between wood

and leaf points impedes effective segmentation. We disregard segments that have few (e.g., 200) points. In the final step, we consolidate wood points identified from all pairs of classification and CCS, eliminated duplicates, and performed a final round of CCS. The outcomes were then categorized into wood and leaf.

4. Experiments, Evaluation, and Discussion

4.1 Methods and metrics used for comparison

In this section, the effectiveness of the proposed method is benchmarked (quantitatively and qualitatively) with two TLS data sets against two recently developed approaches: LeWoS (Wang et al., 2020) employing a geometric approach and GBS, a graph-based method developed by Tian and Li (2022). Those authors asserted that their methods offer automatic functionality and high adaptability across various tree species and sizes.

Assessment involved comparing the predicted labels against manual (ground truth) labels and computing performance metrics based on both the manual and predicted labels obtained from various methods. The methods' performance was assessed utilizing four well-known metrics: precision, recall, F₁-score, and overall accuracy (OA). These metrics are defined as follows:

$$\text{Precision} = \frac{TP}{TP+FP}, \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN}, \quad (4)$$

$$F_1 - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (5)$$

$$\text{Overall accuracy (OA)} = \frac{TP+TN}{TP+FP+TN+FN}, \quad (6)$$

where True Positive, False Positive, True Negative, and False Negative are abbreviated as TP, FP, TN and FN, respectively.

4.2 Experiments and evaluation

4.2.1 Experiment 1, Scots pine tree data: A Scots pine (*Pinus sylvestris*) tree was selected from Weiser et al. (2022), where the data were collected with a RIGEL VZ-400 TLS system in a German Forest in 2019. Excluding the ground points, the tree in Fig. 6(a) consists of 570,337 points. It is of 28.28m height and DBH 46.2cm. Wang et al. (2020) manually classified the tree data into leaf and wood points, and used the classified data for validation of their method. Later, the tree is used in evaluation for the method in Tian and Li (2022). Similar to Tian and Li (2022), in order to mitigate accuracy bias stemming from irregularly distributed points, minimize computational load, and maintain a consistent distance between adjacent points, we subsampled the data with a point spacing of 0.02m. Also we did some manual classification to be precised. Our method was applied to estimate the GFs (L , P , V and λ_0). We find potential linear points using the L values obtained from six different scales of local neighborhood, $nR = 0.05m, 0.1m, 0.2m, 0.3m, 0.4m$ and $0.5m$. Additionally, potential planar and vertical points were identified using P and V values of $nR = 0.1m$. We calculated λ_0 of $nR = 0.1m$ with $\lambda_{0Th} = 0.0005$ to find wood points that originating from thin (3rd and 4th levels) branches. Furthermore, we conducted the GBS method (Tian and Li, 2022), and the LeWoS method (Wang et al., 2020). The outcomes are detailed in Table 2 and visually depicted in Fig. 6. We observe that our method consistently produces superior results across various metrics compared to others, boasting an overall accuracy (OA) and F₁-score of 97.9% and 94% (wood), respectively. LeWoS,

achieves higher OA and F₁-scores for both leaves and wood compared to GBS. However, GBS demonstrates significantly better recall for wood points (77.3%) than LeWoS (68.7%). Fig. 6 (d) reveals that LeWoS misclassifies numerous wood points (depicted in red) as leaves, while in Fig. 6 (g), GBS incorrectly identifies many leaf points (green) as wood. Additionally, as illustrated in Fig. 6 (i), GBS fails to recognize numerous stem points (in the magenta ellipse) as wood points.

Method	Precision (%)		Recall (%)		F ₁ -score (%)		OA (%)
	Leaf	Wood	Leaf	Wood	Leaf	Wood	
LeWoS	93.9	99.1	99.9	68.7	96.8	81.8	94.5
GBS	95.3	76.7	95.1	77.3	95.2	77.0	92.1
Ours	99.1	92.2	98.3	95.9	98.7	94.0	97.9

Table 2. Classification results of Scots pine tree data using LeWoS, GBS, and our methods. OA (overall accuracy).

4.2.2 Experiment 2, Okan tree data: For the second experiment, we selected another open-access data set featuring a tropical Okan (*Cylicodiscus gabunensis Harms*) tree acquired in East Cameroon using the Leica ScanStation C10 (Takoudjou et al., 2018). The tree comprises 462,932 points with a height of 43.68m and a DBH of 116.6cm. Wang et al. (2020) employed this data set to assess the LeWoS algorithm. In their study, Wang et al. (2020) manually categorized the data into leaf and wood points. Subsequently, following Tian and Li (2022), we down-sampled the data with a point spacing of 0.05m. We conducted additional manual corrections to rectify some misclassifications in the data set. The tree data were classified as leaf and wood points of 37,3985 and 88,947, respectively. Our method was performed on the Okan tree data along with the GBS (Tian and Li, 2022), and LeWoS (Wang et al., 2020). Likewise the first experiment, we find potential linear points using the L values obtained from 5 different scales of local neighborhood, $nR = 0.2m, 0.4m, 0.6m, 0.8m$ and $1.0m$. Potential planar and vertical points were identified using P and V values of $nR = 0.2m$. We calculated λ_0 of $nR = 0.2m$ with $\lambda_{0Th} = 0.0005$ to find wood points that originating from thinner branches. The results are presented in Table 3, and visually explored in Fig. 7. We see LeWoS and GBS achieved almost similar OA. Fig 7(d) and 7(g) show that both of them missed to identify many wood (red) points. Whereas, the proposed method in Fig. 7(j) missclassified very few points (red and green), it also achieved significantly higher OA (97.9%) than LeWoS (93.7%) and GBS (93.4%).

Method	Precision (%)		Recall (%)		F ₁ -score (%)		OA (%)
	Leaf	Wood	Leaf	Wood	Leaf	Wood	
LeWoS	92.9	99.2	99.8	68.4	96.2	80.9	93.7
GBS	92.8	97.2	99.5	67.4	96.0	79.6	93.4
Ours	98.3	96.5	99.2	92.8	98.7	94.5	97.9

Table 3. Classification results of Okan tree data using LeWoS, GBS, and our methods.

5. Conclusions

In this paper, we introduce a method for individual tree leaf-wood classification based on GFs that relies solely on spatial coordinates (x, y, z) of each point, without considering any radiometric feature. The algorithm employs classification and segmentation, utilizing only four GFs (L, P, V , and λ_0).

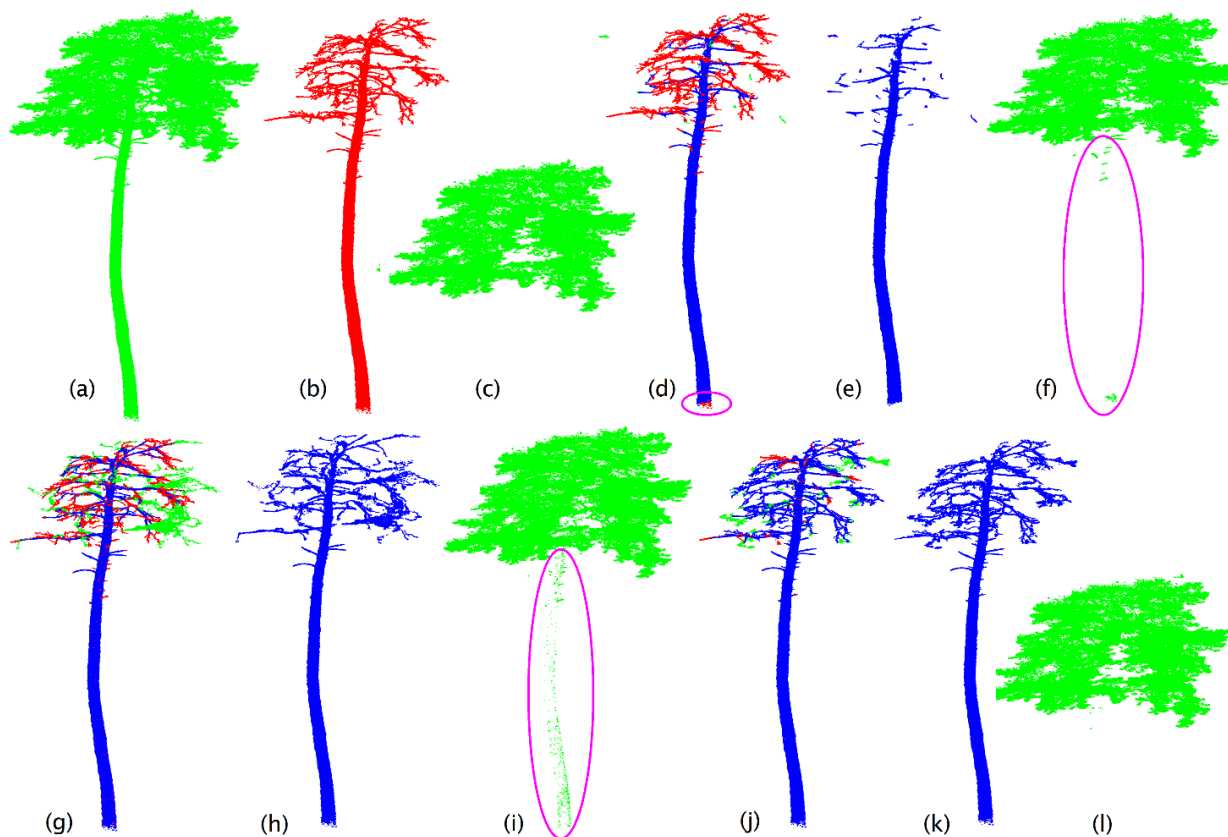


Figure 6. (a) Scots pine tree point cloud. Manually classified points: (b) wood, (c) leaf. Classification results of LeWoS are in (d), (e) and (f): (d) blue points are correctly classified wood points, red and green points are wrongly identified leaf (FN) and wood (FP) points, respectively, identified points are (e) wood, and (f) leaf. Likewise, LeWoS with similar order, classification results of GBS are in (g), (h) and (i), and our results are in (j), (k), and (l).

While leaf points may be incorrectly identified as wood points during the classification step, the incorporation of multi-scale features provides sufficient discriminative power to distinguish between wood and leaf points. Additionally, connected component segmentation (CCS) is employed to reassign misclassified points to their appropriate groups.

Precise wood classification enables the development of an accurate tree structure (skeleton), holding promise for 3D tree modeling. This new method shows potential for comprehending the biophysical processes of trees and estimating various tree parameters such as wood volume, height, and leaf area index. Notably, the method is independent to tree species and size. However, a limitation lies in the necessity for users to determine appropriate sizes for the multiple local neighborhoods to estimate geometric features (GFs). However, neighborhood sizes can be fixed based on the DBH information of a specific tree. Future research will focus on further separation of tree twigs and leaves, as well as reconstructing the irregular geometric shapes of tree stems and branches to enhance the precision of tree modeling. We aim to incorporate data from more intricate environments and mixed forests to broaden the applicability and robustness of our method.

Acknowledgements

Abdul Nurunnabi is funded through the IAS-AUDACITY-PIONEER-2022 project at the University of Luxembourg.

References

- Béland, M., Widłowski, J-L., Fournier, A. R., et al., 2011. Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements. *Agric For Meteorol*, 151(9), 1252–1266.
- Béland, M., Baldocchi, D.D., Widłowski, J-L., Fournier, R.A., Verstraete, M.M., 2014. On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial LiDAR. *Agric For Meteorol*, 184, 82–97.
- Belton, D., Moncrieff, S., Chapman, J., 2013. Processing tree point clouds using Gaussian mixture models. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, 2, 43–48.
- Böhm, J., Bredif, M., Gierlinger, T., Krämer, M., et al., 2016. The IQmulus urban showcase: Automatic tree classification and identification in huge mobile mapping point clouds. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 41, 301–307.
- Brodu, N., Lague, D., 2012. 3D Terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: applications in geomorphology. *ISPRS J Photogramm and Remote Sens.*, 68, 121–134.
- Bucksch, A., Lindenbergh, R., 2008. CAMPINO— A skeletonization method for point cloud processing. *ISPRS J Photogramm and Remote Sens.*, 63(1), 115–127.
- Calders, K., Newnham, G., Burt, A., Murphy, S., Raunonen, P., et al., 2015. Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods Ecol. Evol.*, 6, 198–208.

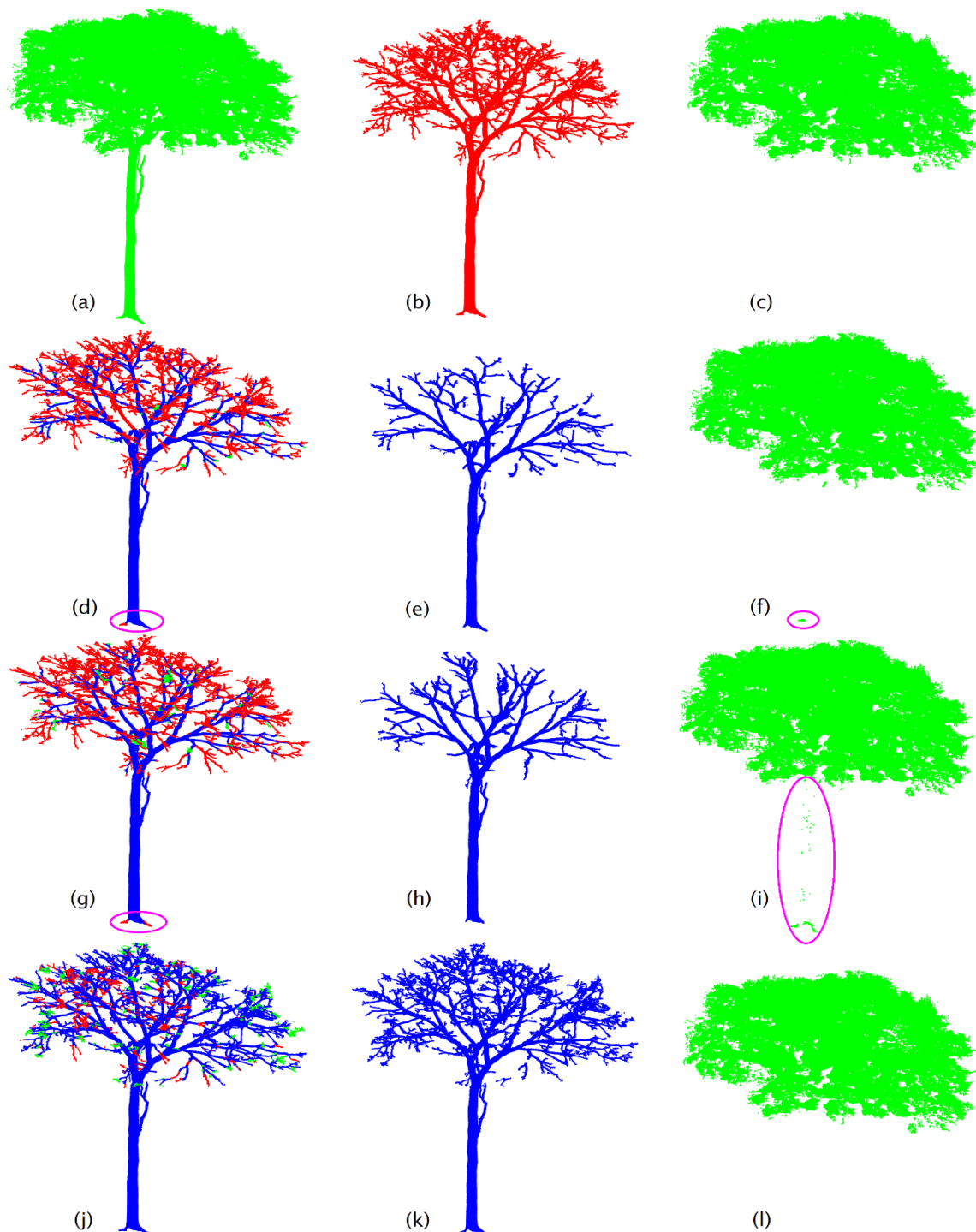


Figure 7. (a) Okan tree point cloud. Manually classified points: (b) wood, (c) leaf. Classification results of LeWoS (d), (e) and (f): (d) blue points are correctly classified wood points, red and green points are wrongly identified leaf (FN) and wood (FP) points, respectively, identified points are of (e) wood, and (f) leaf. Likewise, LeWoS with similar order, classification results of GBS are in (g), (h) and (i), and our results are in (j), (k), and (l).

Côté, J-F., Widlowski, J-L., Fournier, R. A., Verstraete, M. M., 2009. The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. *Remote Sens. Environ.*, 113, 1067–1081.

Côté, J-F., Fournier, R. A., Egli, R., 2011. An architectural model of trees to estimate forest structural attributes using terrestrial LiDAR. *Environ Model Softw*, 26(6), 761–777.

Disney, M. I., Vicari, M. B., Burt, A., Calders, K., et al., 2018. Weighing trees with lasers: advances, challenges, and opportunities. *Interface Focus*, 8(2), 20170048.

Ferrara, R., Salvatore, G.P.V., Ventura, A., Ghisu, T., et al., 2018. An automated approach for wood-leaf separation from terrestrial LIDAR point clouds using the density-based clustering algorithm DBSCAN. *Agric For Meteorol*, 262(15), 434–444.

- Fu, L., Liu, J., Zhou, J., Zhang, M., Lin, Y., 2020. Tree skeletonization for raw point cloud exploiting cylindrical shape prior. *IEEE Access*, 8, 27327–27341.
- Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Garcia-Rodriguez, J., 2017. A review on deep learning techniques applied to semantic segmentation. arXiv preprint arXiv:1704.06857.
- Grau, E., Durrieu, S., Fournier, R., Gastellu-Etchegorry, J. P., Yin, T., 2017. Estimation of 3D vegetation density with terrestrial laser scanning data using voxels. A sensitivity analysis of influencing parameters. *Remote Sens. Environ.*, 191, 373–388.
- Hackel, T., Wegner, J.D., Schindler, K., 2016. Contour detection in unstructured 3D point clouds. *IEEE CVPR*, 1610–1618.
- Hui, Z., Jin, S., Xia, Y., Wang, L., Z., Y-Y., Cheng, P., 2021. Wood and leaf separation from terrestrial LiDAR point clouds based on mode points evolution. *ISPRS J Photogramm and Remote Sens.*, 178, 219–239.
- Li, S., Dai, L., Wang, H., Wang, Y., He, Z. E., Lin, S., 2017. Estimating leaf area density of individual trees using the point cloud segmentation of terrestrial lidar data and a voxel-based model. *Remote Sensing*, 9(11), 1202.
- Li, J., Wu, H., Xiao, Z., Lu, H., 2022. 3D modeling of laser-scanned trees based on skeleton refined extraction. *Int J Appl Earth Obs Geoinformation*, 112, 102943.
- Ma, L., Zheng, G., Eitel, U. H. J., Moskal, L. M., He, W., Huang, H., 2016. Improved salient feature-based approach for automatically separating photosynthetic and nonphotosynthetic components within terrestrial lidar point cloud data of forest canopies. *IEEE Trans Geosci Remote Sens*, 54(2), 679–696.
- Nurunnabi, A., Belton, D., West, G., 2012. Diagnostic-robust statistical analysis for local surface fitting in 3D point cloud data. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, I-3, 269–274.
- Nurunnabi, A., West, G., Belton, D., 2015. Outlier detection and robust normal-curvature estimation in mobile laser scanning 3D point cloud data. *Pattern Recognit.*, 48(4), 1404–1419.
- Nurunnabi, A., West, G., Belton, D., 2016. Robust locally weighted regression techniques for ground surface points filtering in mobile laser scanning 3D point cloud data. *IEEE Trans. Geosci. Remote Sens.*, 54(4), 2181–2193.
- Nurunnabi, A., Teferle, F., Li, J., Lindenbergh, R., Hunegnaw, A., 2021. An efficient deep learning approach for ground point filtering in aerial laser scanning point clouds. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 43(B1–2021), 31–38.
- Nurunnabi, A., Teferle, F., Al-Sharadqah, A., Xia, Y., Hunegnaw, A., 2024. Robust estimation of tree diameter at breast height from point clouds. *IEEE IGARSS*, 7-12 July, Athens, Greece.
- Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., et al., 2013. Fast automatic precision tree models from terrestrial laser scanner data. *Remote Sensing*, 5(2), 491–520.
- Shu, Q., Rötzer, T., Detter, A., Ludwig, F., 2022. Tree information modeling: a data exchange platform for tree design and management. *Forests*, 13(11), 1955.
- Tao, S., Guo, Q., Su, Y., Xu, S., et al., 2015. A geometric method for wood-leaf separation using terrestrial and simulated lidar data. *Photogramm Eng Remote Sensing*, 81(10), 767–776.
- Takoudjou, S. M., Ploton, P., Sonké, B., Hackenberg, J., Griffon, S., De Coligny, et al., 2018. Using terrestrial laser scanning data to estimate large tropical trees biomass and calibrate allometric models: A comparison with traditional destructive approach. *Methods Ecol. Evol.*, 9(4), 905–916.
- Tian, Z., Li, S., 2022. Graph-based leaf-wood separation method for individual trees using terrestrial LiDAR point clouds. *IEEE Trans. Geosci. Remote Sens.*, 60, 1–11.
- Vicari, B. M., Disney, M., Wilkes, P., Burt, A., Calders, K., Woodgate, W., 2019. Leaf and wood classification framework for terrestrial LiDAR point clouds. *Methods in Ecol Evol*. 10, 680–694.
- Vo., A-V., Truong-Hong, L., Laefer, D.F., Bertolotto, M., 2015. Octree-based region growing for point cloud segmentation. *ISPRS J Photogramm and Remote Sens.*, 104, 88–100.
- Wang, Z., Zhang, L., Fang, T., Mathiopoulos, P. T., Tong, X., Qu, H., et al., 2015. A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification. *IEEE Trans. Geosci. Remote Sens.*, 53(5), 2409–2425.
- Wang, D., Hollaus, M., Pfeifer, N., 2017. Feasibility of machine learning methods for separating wood and leaf points from terrestrial laser scanning data. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV–2/W4, 157–164.
- Wang, D., Takoudjou, S. M., 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.
- Weiser, H., Schäfer, J., Winiwarter, L., Krašovec, N., et al., 2022. Individual tree point clouds and tree measurements from multi-platform laser scanning in German forests. *Earth Syst. Sci. Data*, 14(7), 2989–3012.
- Xi, Z., Hopkinson, C., Chasmer, L., 2018. Filtering stems and branches from terrestrial laser scanning point clouds using deep 3-D fully convolutional networks. *Remote Sensing*, 10(8), 1215.
- Zhao, K., García, M., Liu, S., Guo, Q., et al., 2015. Terrestrial lidar remote sensing of forests: Maximum likelihood estimates of canopy profile, leaf area index, and leaf angle distribution. *Agric For Meteorol*, 209–210, 100–113.
- Zhu, X., Skidmore, A.K., Darvishzadeh, R., Niemann, K.O., Liu, J., Shi, Y., et al., 2018. Foliar and woody materials discriminated using terrestrial LiDAR in a mixed natural forest. *Int J Appl Earth Obs Geoinformation*, 64, 43–50.
- Zou, X., Cheng, M., Wang, C., Xia, Y., Li, J., 2017. Tree classification in complex forest point clouds based on deep learning. *IEEE Geosci. Remote Sens. Lett.*, 14(12), 2360–2364.