A multi-sensor multi-resolution dataset to support forest inventory methods

Lauris Bocaux[1,2], Narges Takhtkeshha[1,3], Zhenyu Ma[1], Fabio Remondino [1]

[1] 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy
Email: (lbocaux, ntakhtkeshha, zma, remondino)@fbk.eu
[2] Department of Information Engineering and Computer Science, University of Trento, Trento, Italy
Email: lauris.bocaux@unitn.it
[3] Department of Geodesy and Geoinformation, TU Wien, Vienna, Austria
Email: Narges.Takhtkeshha@geo.tuwien.ac.at

Keywords: 3D forestry, Tree inventory, LiDAR, Semantic segmentation, Individual Tree Detection, Benchmark

Abstract
Accurate estimation of forest structural and taxonomic parameters is vital for biodiversity monitoring, carbon accounting and sustainable management. Most of the current methods for estimating these parameters are still developed and tested on site-specific case studies, limiting reproducibility and cross-site generalization. This paper introduces 3D3, a multi-sensor and multi-resolution benchmark dataset designed to evaluate 3D forestry algorithms across diverse European forest types. 3D3 includes data collected by airborne, helicopter, UAV and terrestrial (static and mobile) laser scanning systems along with RGB and hyperspectral imagery, covering a variety of forest types (Boreal, Alpine and Mediterranean). By encompassing both mono- and multi-wavelength laser data, 3D3 represents a unique resource for developing new algorithms and evaluating them on distinct datasets. Each site provides ground truth for at least one task among Individual Tree Segmentation (ITS), Forest Semantic Segmentation (FSS) or tree parameter estimation and species classification.

1. Introduction

Light Detection and Ranging (LiDAR) technology has opened new opportunities for forest monitoring at the individual-tree scale, enabling the estimation of key ecological parameters such as tree height, crown area and species. Forestry decision-making is increasingly relying on detailed stand-level data rather than coarse inventories. Such information is critical for biodiversity assessment, carbon accounting, and sustainable forest management. The emergence of airborne- and Unmanned Aerial Vehicle (UAV)-based LiDAR, together with multispectral and hyperspectral imaging and data fusion methods, has boosted forest analyses in three dimensions and across spectral domains (Balestra et al., 2024; Ma et al., 2025). In this regard, different rule-based and data-driven algorithms for 2D and 3D Individual Tree Segmentation (ITS) and Forest Semantic Segmentation (FSS) have been developed. Most of 2D ITS methods are based on canopy height model (CHM). In this context, the reversed watershed algorithm has been actively used (Dalponte and Coomes, 2016; Douss & Farah, 2022), including its integration with multispectral information (Deschene et al., 2017). 3D-based ITS has gained attention in most recent studies and is moving from traditional methods like Li et al. (2012) and Treeiso (Xi & Hopkinson, 2022) toward advanced deep learning architectures such as ForAINet (Xiang et al., 2024), TreeisoNet (Xi & Degenhardt, 2025) or unsupervised methods (Ruoppa et al., 2025). Although studies driven by tree object semantic segmentation with Quantitative Structure Modelling (QSM) present satisfactory results (Shen et al., 2022), FSS is recently leveraging deep learning (DL) models used inside methods like ForAINet, which seems promising. But despite recent advances, a major challenge remains: most methods are evaluated on site-specific datasets, which limits reproducibility and reduces cross-comparisons. Therefore, evaluating generalization capabilities and limitations of developed methods across different forest types and LiDAR datasets remains a prominent research question. The problem is compounded by differences in LiDAR pulse density and penetration (Jakubowski et al., 2013; Ferraz et al., 2015), acquisition geometry and the availability of complementary data such as hyperspectral imagery (Dalponte et al., 2012; Heinzel & Koch, 2012). As a result, even well-established methods often show inconsistent performance across different forest types or acquisition conditions. Recent initiatives have started to address this issue by providing multi-sensor benchmark datasets for forestry applications (Taher et al., 2025). For example, the NeonTreeEvaluation dataset (Weinstein et al., 2021) integrates LiDAR, hyperspectral and RGB data across 22 NEON sites for consistent evaluation of tree detection methods, while the FOR-instance benchmark dataset (Puliti et al., 2023) provides annotated UAV-based LiDAR point clouds for both semantic and instance segmentation. More general efforts such as FoMo-Bench (Bountos et al., 2023) extend the benchmarking concept to forest monitoring tasks across multiple data sources, including LiDAR, SAR and multispectral imagery. However, there is still no widely adopted benchmark that unifies 2D and 3D ITS, FSS and parameter estimation across heterogeneous forest environments, datasets, and acquisition parameters.
To contextualize these limitations, Table 1 summarizes the main publicly available benchmark datasets for 3D high-resolution forest remote sensing, outlining their modalities, acquisition platforms, supported tasks and remaining gaps. This comparison highlights the growing diversity of multimodal forest datasets, while emphasizing the lack of a unified resource that simultaneously addresses both structural and spectral dimensions across ecosystems. Each benchmark contributes with valuable assets - ranging from multimodal integration to dense LiDAR sampling - but none provides a unified, multi-resolution and multi-modal framework that encompasses 2D and 3D ITS, FSS, and tree-attribute estimation across contrasting forest types. This persistent gap motivates the creation of the 3D3 dataset.

1.1. Paper aims

This paper introduces a unique benchmark dataset - named 3D3 - that directly addresses the abovementioned gap by combining multi-scale, multi-wavelength and multi-modal acquisitions of

https://doi.org/10.5194/isprs-archives-XLVIII-1-W6-2025-33-2025

Dataset	Modalities	Platform	Main tasks	Limitations
NeonTreeEvaluation (Weinstein et al., 2021)	LiDAR, Hyperspectral and RGB imaging	Airborne	ITS	Fixed region; limited LiDAR density variation
Treeiso (Xi & Hopkinson, 2022)	LiDAR	TLS	ITS	Localized; no multimodal data
FOR-instance (Puliti et al., 2023)	LiDAR, RGB imaging	UAV	ITS, FSS	UAV-only; lacks spectral fusion
ForestSemantic (Liang et al., 2024)	LiDAR	TLS	ITS, FSS	Localized; no multimodal data
ETH/SwissTree (Xiang et al., 2024)	LiDAR, RGB imaging	Airborne	ITS	No hyperspectral; limited ecosystem diversity
FoMo-Bench (Bountos et al., 2025)	LiDAR, SAR, Multi- and Hyperspectral imaging	Mixed (satellite & airborne)	Multimodal forest monitoring	Broad tasks; not focused on ITS/FSS
TreeisoNet Benchmark (Xi & Degenhardt, 2025)	LiDAR	Airborne, UAV, TLS	ITS	No spectral data; limited task diversity
Ours (3D3)	Mono- & multispectral LiDAR, Hyperspectral and RGB imaging	Airborne, Helicopter, UAV, MLS, TLS	ITS, FSS, tree species classification	Not all modalities in all datasets

Table 1: Summary of major benchmark datasets for 3D high-resolution forestry.

Dataset	A	B	C	D	E	F
Instrument	Riegl VQ780ii Mono-wavelength LiDAR (1064 nm)	Teledyne Optech GALAXY T1000 - Mono-wavelength LiDAR (1064 nm); SPECIM AISAFENIX - Hyperspectral camera (381.35 nm to 2502.38 nm); PhaseOne iXU-RS 1000 RGB camera	Riegl VUX-120 Mono-wavelength LiDAR (1550 nm); PhaseOne iXM100 RSM35 RGB camera	HeliALS multispectral LiDAR: VQ-840-G (532 nm), miniVUX-1DL (905 nm), VUX-1HA (1550 nm)	DJI-L1 Mono-wavelength LiDAR (905 nm)	Semi-Synthetic
Platform	Aircraft	Aircraft	Helicopter	Helicopter	UAV, MLS and TLS	Aircraft
Approximated density/resolution	10 pts/m²	LiDAR: 75 pts/m² Hyperspectral: 60 cm RGB ortho: 10 cm	375 pts/m²	1200 pts/m²	2000 pts/m²	From 0 to 75 pts/m²
Area of Interest	16 km x 11 km	1680 m x 1550 m	2000 m x 2000 m	100 m x 100 m	275 m x 150 m	Flexible
Type of data	LiDAR: X, Y, Z, I	LiDAR: X, Y, Z, I Hyperspectral: 364 bands OrthoRGB: R, G, B	LiDAR: X, Y, Z, I	LiDAR: X, Y, Z, SWIR, NIR, Green	LiDAR: X, Y, Z, I, RGB	LiDAR: X, Y, Z, I
Biome	Urban, Continental	Dense, Continental	Dense, Mediterranean	Dense, Boreal	Dense, Alpine	Variable
Ground Truth	ITS, species (203 classes)	Species (5 classes)	ITS	FSS (6 classes)	ITS, trunk sizes, tree species	ITS, FSS, species

Table 2: Summary of 3D3 data for testing and validating forest inventory methods.

forestry areas across diverse European areas for evaluating individual tree and performing forest-level analyses. The 3D3 benchmark dataset aims to (i) advance comparability, reproducibility and transferability in forest remote sensing and (ii) provide a guideline for the next generation of data-driven ecological monitoring tools. Unlike existing datasets (Table 1), 3D3 enables the evaluation of various state-of-the-art methods under dense or sparse LiDAR data and multi-modal imaging. Moreover, the paper proposes a summary guideline that could support forest managers in choosing sensor technologies and data types based on their needs.

2. The 3D3 benchmark dataset

The proposed 3D3 benchmark dataset (Table 2) is assembled to include the variability of forest structure, acquisition conditions and sensing technologies that challenge tree-level analyses. 3D3 covers five sites across Europe, representing urban, Mediterranean, temperate and boreal ecosystems, spanning from airborne acquisitions at operational scales (15-75 pts/m²) to UAV-based ultra-dense LiDAR (1000-2000 pts/m²). 3D3 is also completed by a flexible semi-synthetic set of data which allows us to reconstruct plenty of forest scenarios, changing LiDAR density, tree overlap and slope, based on a random distribution of manually selected trees.

The various sites integrate complementary modalities, including RGB orthophotos and hyperspectral imagery with 364 bands, enabling multimodal analysis at the tree and forest level. A distinguishing feature of our dataset is its heterogeneity in wavelength and sensor configuration: from mono-wavelength LiDAR (905 nm, 1064 nm and 1550 nm) to multispectral LiDAR (Green - 532 nm, NIR - 905 nm, SWIR - 1550 nm). This diversity allows users to evaluate not only the potential of conventional mono-wavelength LiDAR but also advanced multi-wavelength LiDAR systems (Takhtkeshha et al., 2024) to strengthen FSS methods. 3D3 features ground truth (GT) data in each available site, either coming from manual measurements or from synthetic data.

GT varies according to canopy complexity and point density. Creating a reliable GT requires manual annotation of images or 3D point clouds at the pixel/point level. This process is both labor-intensive and subject to interpretation biases, particularly in dense or multi-layered canopies where class boundaries are ambiguous. Dataset F consists of a Semi-Synthetic Ground Truth (SSGT) (Figure 1): isolated trees from the available scenes are first extracted based on height and spacing criteria (Figure 1a) and reassembled into artificial forest configurations following a geometric adjacency rule algorithm (Figure 1b). From these semi-synthetic scenes, corresponding CHM and multi-format annotations are derived (Figure 1c–f). On the other hand, for the scenes with clear separability, traditional manual annotation is employed, aided by preliminary segmentation from ForAINet (Xiang et al., 2024) to accelerate label definition.
3D3 forestry data form a resource for testing both geometric methods and state-of-the-art deep learning architectures, guaranteeing that evaluation is not tied to a single acquisition scenario but instead reflects the breadth of existing forest monitoring challenges. The 3D3 benchmark dataset is presented to specifically support, among others, the following tasks in forest remote sensing:

- 2D & 3D Individual Tree Segmentation (ITS);
- Forest Semantic Segmentation (FSS);
- tree parameters estimate (tree height, crown area, DBH, etc.);
- tree species classification.

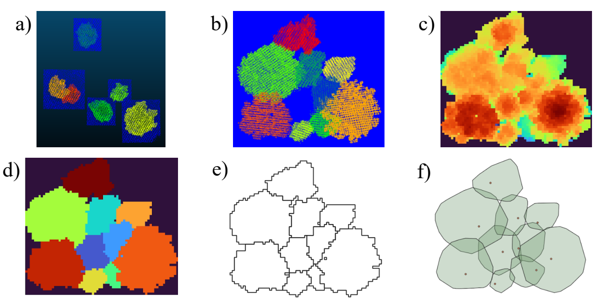

Figure 1: Dataset F with Semi-Synthetic Ground Truth. Single trees (a), clumped trees (b), CHM (c), 2D raster ITS (d), 2D ITS canopy polygons (e), 3D ITS convex hull (f).

3. Data processing

Although LiDAR acquisitions have all their own specificities, particularly density, we propose a shared pre-processing methodology to create a common base for all future processes. Then, methods for ITS and FSS are applied to the different LiDAR datasets of 3D3. Finally, tree parameter estimation methods are employed.

3.1 Pre-processing

We created a replicable pre-processing process applicable to all 3D3 datasets. Nonetheless, the main processing part differs according to the point cloud resolution.

- **Outlier removal and ground filtering:** Spurious points are first removed using the Statistical Outlier Removal (SOR) filter (Rusu & Cousins, 2011), followed by ground extraction with the Cloth Simulation Filter (CSF - Zhang et al., 2016) (Figure 2b). This step establishes a stable reference surface for height normalization.
- **Height normalization:** For 2D analyses, Digital Terrain Model (DTM) and Digital Surface Models (DSM) are derived from the classified ground and top points, respectively. Gaps in both models are filled by spherical kriging, and canopy maxima are refined through an in-house canopy-clothing algorithm. The subtraction of the DTM from the DSM results in the Digital Elevation Model (DEM) (Figure 3a). In 3D, height normalization is performed with OPALS (Pfeifer et al., 2014), by subtracting the ground model elevation from each LiDAR point (Figure 2d).
- **Tree point extraction:** Non-tree points are filtered using a pretrained Kernel Point Convolution (KPConv) network (Thomas et al., 2019), fine-tuned on Dataset A, to isolate tree structures (Figure 2c). For 2D-based analyses, a binary tree mask is generated, then applied to the DEM (Figure 3a) to obtain the Canopy Height Model (CHM) (Figure 3b) in order to constrain canopy regions spatially.

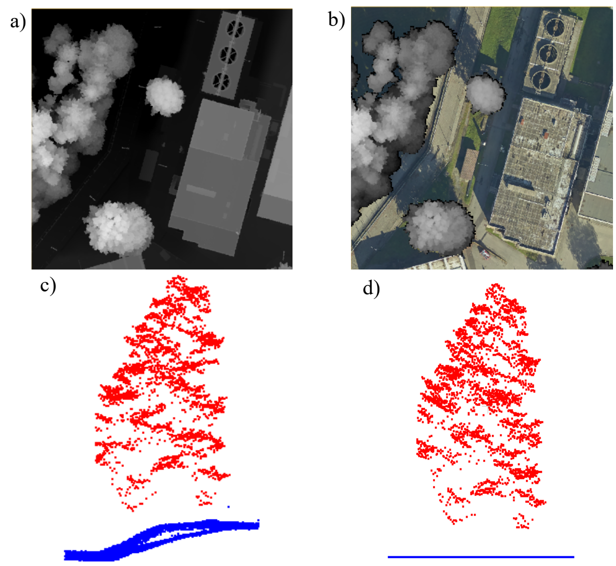

Figure 3: 2D & 3D height normalization; DEM (a), CHM (b), raw PC (c), normalized PC (d).

3.2 Main tasks

3.2.1 Individual Tree Segmentation (ITS)

- **2D approach:** The 2D ITS approach relies on the Canopy Height Model (CHM) of each area of interest. The CHM is inverted to produce a valley-like raster and smoothed with a Gaussian filter to reduce local anomalies. A variable-radius watershed algorithm is then applied, using tree height as a guiding parameter and enforcing constraints on minimum height (e.g. 2 m) and crown expansion. This method ("reversed watershed") can separates nearby canopies while reducing over-segmentation on varied ground.

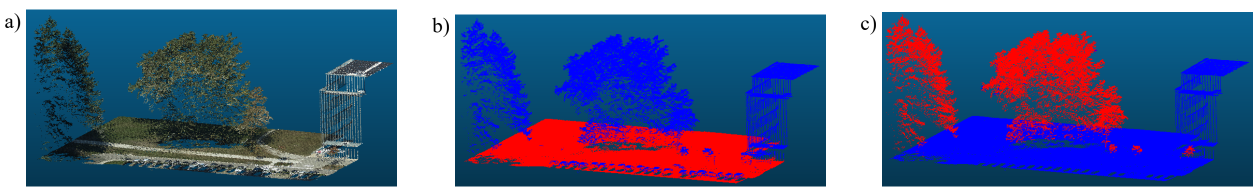

Figure 2: Part of the original point cloud (a), ground (b), tree points (c), in red.

- **3D approach:** For 3D ITS, segmentation is directly performed on the height-normalized point cloud. The algorithm (Li et al., 2021) iteratively identifies the highest unassigned point and grows crowns top-down, governed by 2D Euclidean distance and a minimum-spacing rule to prevent tree merging near the ground. Shape descriptors, including convex hull indices and angular sector checks, are used to refine crown boundaries and separate intertwined branches.

3.2.2 Forest Semantic Segmentation (FSS)

FSS provides valuable insights into ecosystem structure by enabling the classification of vegetation layers, tree components and terrain, which supports accurate estimation of forest attributes such as canopy cover, biomass, and species distribution. FSS on high-resolution 3D datasets is performed using Point Transformer v3 (PTv3), following the attention-based framework proposed by Wu et al. (2024) and adopted in Takhtkeshha et al. (2025). PTv3 leverages vector attention and hierarchical patch-based feature extraction to model long-range spatial dependencies within dense forest point clouds. Unlike convolutional kernels or MLP encoders, PTv3 dynamically learns contextual relationships between points, allowing precise differentiation between foliage, woody components and terrain. Its hierarchical design enhances geometric awareness, enabling fine delineation of canopy layers, trunks, and ground surfaces. In our study, PTv3 is fine-tuned on Dataset D for FSS using six classes: ground, low vegetation, trunk, branches, foliage and woody debris (Takhtkeshha et al., 2025).

3.2.3 Panoptic segmentation

ForAInet (Xiang et al., 2024) is a panoptic deep learning framework jointly performing semantic and instance segmentation of 3D point clouds. Instance labels directly provide ITS masks while semantic predictions describe forest structures across classes such as ground, low vegetation, branches, trunk and foliage. The network combines center-offset voting with metric embeddings to generate complementary hypotheses, refined through non-maximum suppression and hierarchical merging, resulting in contiguous, per-tree point clusters suitable for computing structural and taxonomic parameters.

3.2.4 Tree parameter estimation

- **Tree species classification:** Accurate tree species mapping requires, ideally, both geometric and spectral information. In our study, hyperspectral imaging data are combined with 2D ITS outputs to classify species. Non-tree pixels are first filtered using the same tree mask used for DEM creation. The hyperspectral raster is then transformed through Minimum Noise Fraction (MNF) (Green et al., 1988), a dimensionality reduction method similar to a principal component analysis (PCA), that enhances signal-to-noise separation. A Random Forest (RF) classifier is then trained on the MNF components and punctual tree species ground truth, producing species labels per segmented crown.
- **Tree height and tree-top position:** They are computed directly from the ITS results. In 2D, height corresponds to the maximum CHM value within each crown mask whereas the tree-top position is assigned to the centroid pixel. In 3D, the tree-top is defined as the highest point in the normalized point cloud for each tree instance.
- **Crown area:** It is estimated according to the data dimensionality and the ITS method. For 2D reversed watershed results, the crown area represents the exposed canopy footprint derived from the CHM and is computed as the number of crown pixels multiplied by the squared spatial resolution. In 3D, the crown area is obtained from the top-view projection of each individual tree point cloud. Depending on the desired level of accuracy and computational efficiency, the area can be estimated via the convex hull or using alpha-shape polygons.

4. Results

To evaluate the performance of the proposed methods on the 3D3 datasets, we adopt commonly used metrics like precision, recall, F1-score, accuracy, intersection over union (IoU) and their means (Nemmaoui et al., 2024; Liu et al., 2025). Nonetheless, while their formulations remain consistent, their interpretation varies depending on the application: for instance, in ITS, they reflect the correctness of detected tree instances, whereas, in FSS, they quantify per-class classification performance at the point level.

4.1 Individual Tree Segmentation (ITS)

The evaluation of ITS is divided into two complementary components. Individual tree detection assesses whether the number and spatial position of detected trees match the ground truth, considering a detection correct when the IoU between the predicted and reference crown areas exceeds 50% (Lucas et al., 2024; Li et al., 2025). Individual tree delineation, on the other hand, measures how accurately each detected tree is spatially outlined, quantified using the mIoU between predicted and reference crowns. Together, these metrics provide a ample view of both detection accuracy and boundary precision.

4.1.1 2D ITS

Datasets A (10 pts/m²), F (at 75 pts/m²) and E (2000 pts/m²) are considered as they share similar forest structures. 2D ITS is performed with the reverse watershed algorithm. Accuracies are calculated using the available manual 3D ground truth (A, E) or the Semi-Synthetic Ground Truth on 100 random trees (F). According to Table 3, the results of 2D ITS are homogeneous, although the density of LiDAR is at a totally different scale. Since trees are complex objects and the reversed watershed algorithm does not imply any form of adaptive intelligence, but remains a simple applicative algorithm, the results are promising. Delineation results are shown in Figure 4.

2D ITS	Detection			Delineation
Dataset	Precision	Recall	F1-score	mIoU
A	74.68%	61.83%	67.65%	No data
F	65.3%	64.3%	64.8%	58.3%
E	70.5%	66.3%	68.3%	No data

Table 3: Results of 2D ITS with reversed watershed algorithm.

In all 3D3 scenes, the highest part of the canopy is properly described regardless of the LiDAR density. Thus, 2D ITS accuracy is influenced by forest density, much more than LiDAR density. 2D ITS performs well also in case of low-density LiDAR acquired over a large urban area (e.g. Dataset A). Very often, trees from the same street share the same species and the same date of plantation, so quite the same height, and are usually spaced apart from each other, leading to less misunderstanding from the simple reversed algorithm.

4.1.2 3D ITS

All 3D3 are considered and analyzed with ForAINet (Xiang et al., 2024) and Li2012 (Li et al., 2012), although Datasets B, C and D have no ground truth for 3D ITS.

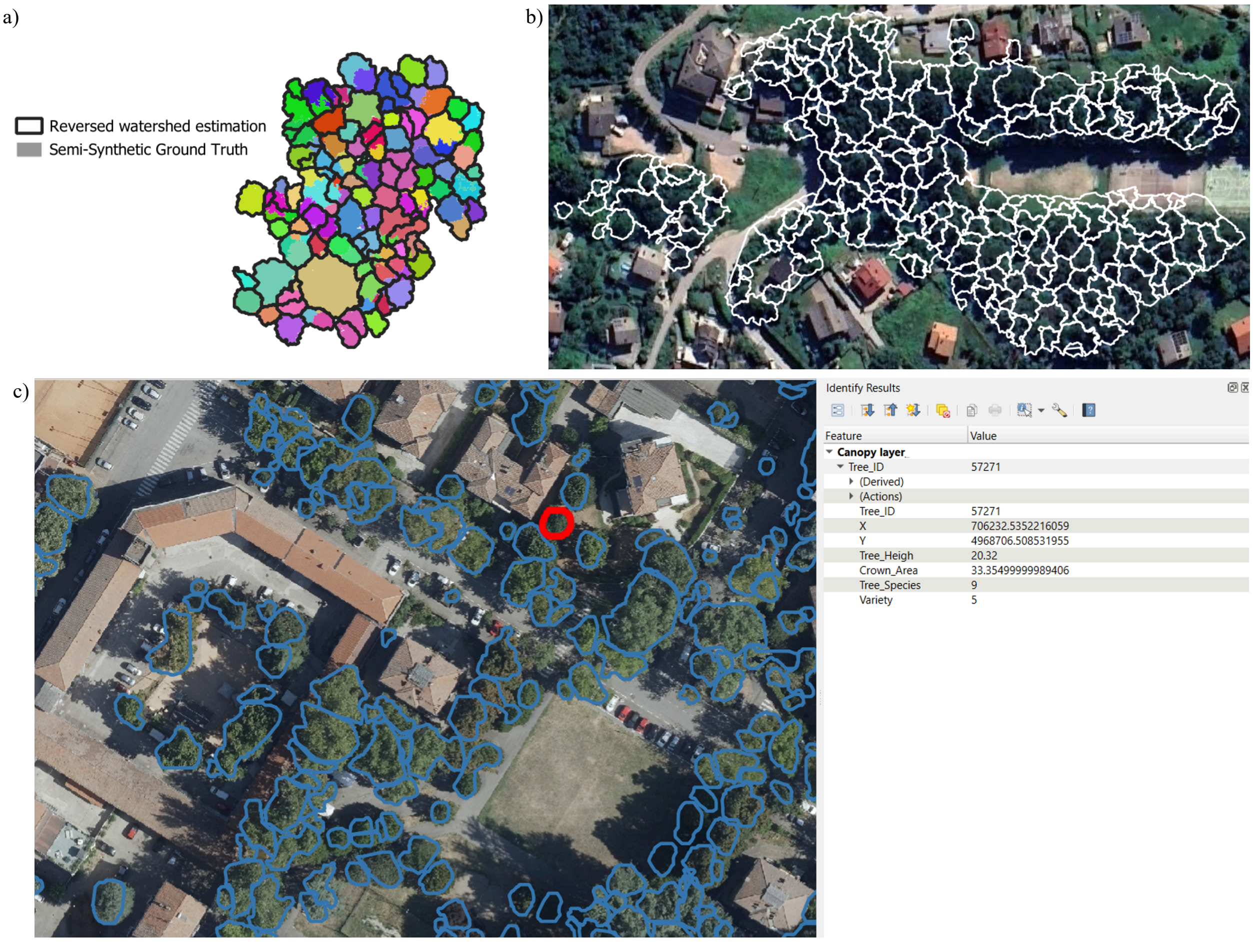

Figure 4: 2D ITS with reversed watershed algorithm for Dataset B - with GT overlaid (a) and Dataset E (b). ITS and tree parameters for dataset A (c).

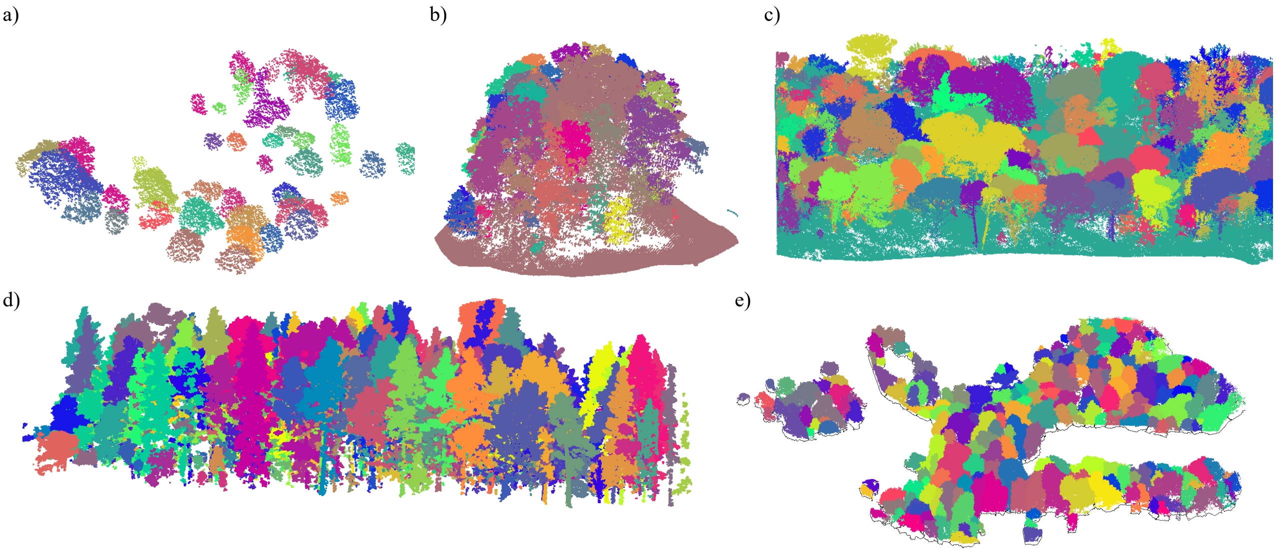

Figure 5: 3D ITS results displayed for parts of Dataset A (a), B (b), C (c) and D (d) and the entire area of E (e).

Figure 5 shows the achieved 3D ITS results on subsets or entire areas of the 3D3 datasets. In general, standalone trees are almost perfectly detected and delineated, whereas clumped trees give more difficulties to both deep-learning models. For instance, with Li2012, we obtain for Dataset A, a detection accuracy of 73.0% whereas for Dataset E, a precision of 81.4%, a recall of 69.3% for an F1-score of 74.9%. However, the more the point clouds are dense and complete (i.e., describing the entire structure of the tree), the more 3D ITS is accurate. For Dataset B (Figure 5b) Li2012 classified ground (pink-brown color) also part of trees due to low penetration of LiDAR points through the dense canopy. On the other hand, for Dataset E, which has the same forest structures as Dataset B but a much consistent penetration of the LiDAR signal, very decent and coherent results are achieved (Figure 5e).

IoU (for all classes)					
Ground	Low veg.	Trunk	Branches	Foliage	Woody debris
93.2%	66.5%	71.5%	26.1%	95.1%	62.2%
mIoU = 69.1%; **mAcc** = 80.0%; **OAcc** = 94.9%					

Table 4: Fine-grained FSS using PTv3 on Dataset D.

4.2 Forest Semantic Segmentation (FSS)

Dataset D is considered, being the only one furnished with a meticulously self-made ground truth for semantic segmentation. FSS is performed with the fine-tuned PTv3 model (Takhtkeshha et al., 2025). Visual are reported in Figure 6a while metrics in Table 4. The three different wavelengths used to survey the forest help to better separate foliage, branches and woody parts, which improves the overall segmentation quality.
On the other hand, ForAINet is applied to Dataset E (Figure 6b). For all other datasets, if LiDAR coverage is not complete, meaning that brunches or trunks are not properly sensed and the density is low, FSS results are poor.

4.3 Tree parameter estimation

After ITS and FSS processes, tree and forest parameters can be derived to support further analyses. These include tree species classification and determination of tree height, tree-top position and crown area. Each parameter is assessed according to the method used (2D or 3D) and the type of subsidiary data available, enabling a broader evaluation of the approaches.

4.3.1 Tree species classification

Dataset B comes with hyperspectral images that are essential for a reliable spices' identification. The Random Forest was repeated 6 times to lower the variance of the results due to its randomness, with a train ratio of 70% and 1000 random trees. Visual results are shown in Figure 7 whereas metrics of tree species classification are presented in Table 5.

Mean metrics	**Other**	**Pinus nigra**	**Pinus strobus**	**Quercus rubra**
Precision	87.5%	84.8%	81.7%	91.2%
Recall	95.0%	72.5%	81.7%	82.3%
F1-score	91.0%	77.3%	81.3%	86.3%

Table 5: Mean metrics of Random Forest on Dataset B for species classification.

4.3.2 Tree height, treetop position, and crown area

The quality of these parameters is mainly based on the accuracy of the results from the previous steps (ITS and FSS). For instance, in the case of an estimated tree being truly detected and having a perfect delineation, the top is almost always the highest pixel of the CHM in 2D or the highest point of the normalized point cloud. The only tricky aspect - that is too rare to be significant - is when the height normalization process (Section 3.1) alters the structure of the tree because of sloping relief below the trees. Figure 8 shows the results of the parameter extraction for Dataset A in 3D and Dataset B in 2D.

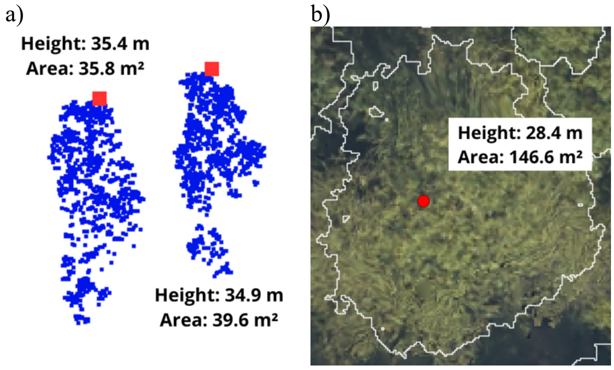

Figure 8: Tree parameter extracted in 3D for Dataset A (a) and in 2D for Dataset B (b). Tree-top positions are shown in red.

5. Conclusions

This work introduced 3D3 (https://github.com/3DOM-FBK/3D3), a multi-sensor multi-resolution benchmark dataset for 3D forestry applications. It is designed to support the development and evaluation of ITS methods in 2D or 3D, FSS algorithms or the extraction of structural and taxonomic parameters. 3D3 is unique given the diverse geographic locations and European forest types, geometric and radiometric resolutions, ground truths and possible tasks. Using state-of-the-art processes, it was reported how segmentation and parameter estimation performance vary with both LiDAR and forest density - as summarized in Table 6.
Moreover, a key innovation of 3D3 is the Semi-Synthetic Ground Truth (SSGT), which enables the creation of controlled forest scenarios based on real tree objects. This allows testing algorithms under varying tree arrangements and densities, improving reproducibility and scalability.

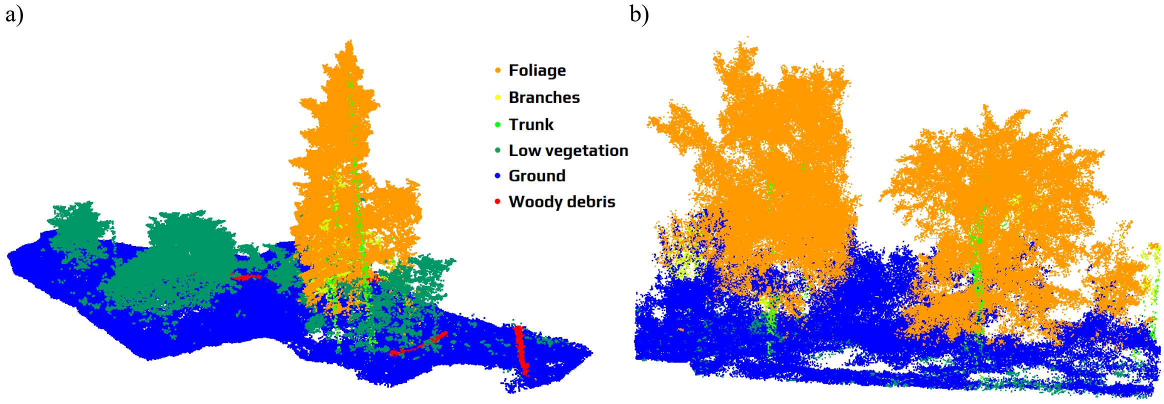

Figure 6: FSS results on a part of Dataset D with fine-tuned PTv3 (a - metrics in Table 4) and Dataset E with ForAINet (b).

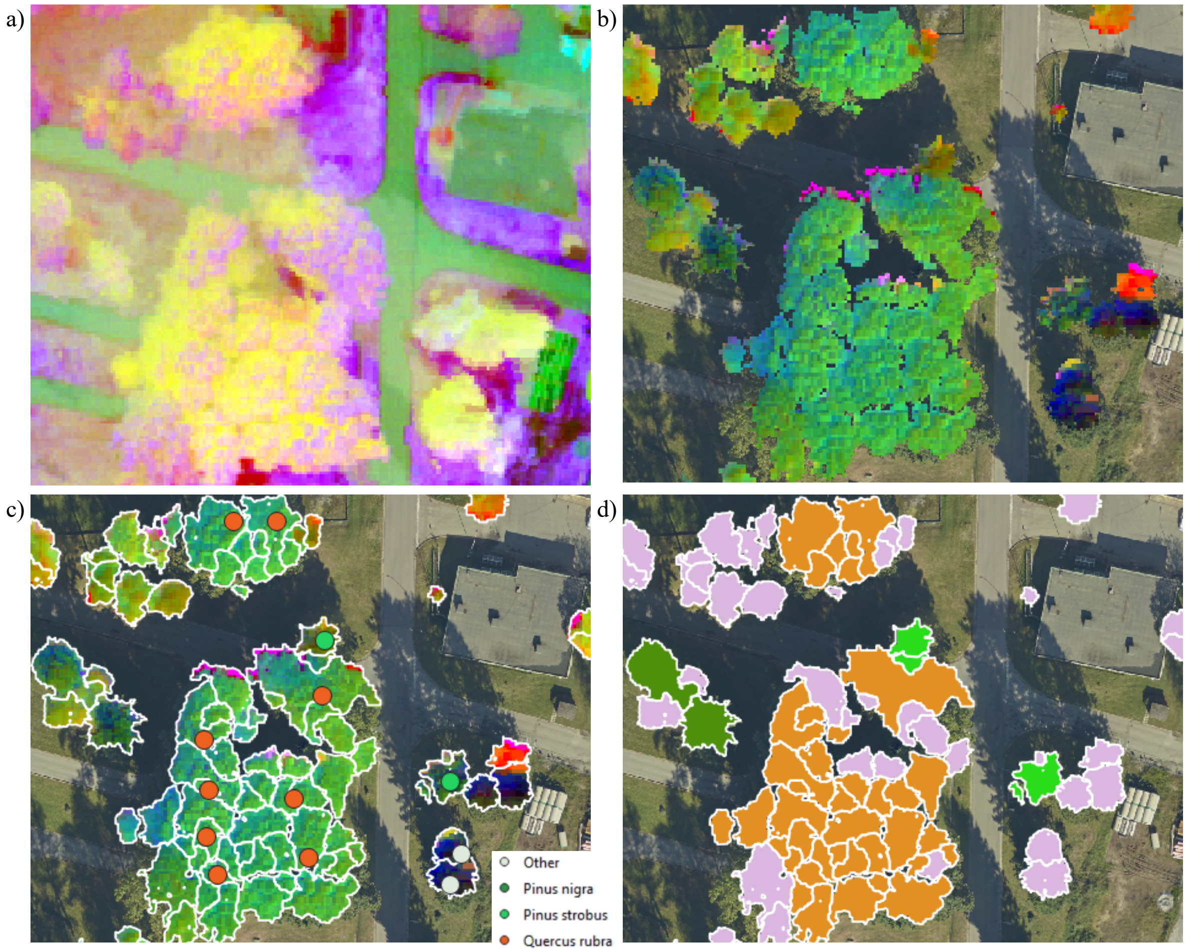

Figure 7: Results on a part of Dataset B: MNF (a); vegetation MNF (b); 2D ITS & tree species ground truth (c); tree species map based on Random Forest (d).

LiDAR density	Individual Tree Segmentation		Forest Semantic Segmentation		Tree Parameter Extraction	
	2D	3D	Ground / Vegetation	Complete structure	Height, Top, Area	Species
Below 100 pts/m²	Yes	Non-touching trees	Yes	No	Yes	Hypersp & GT
100 to 1000 pts/m²	Yes	Non-clumped trees	Yes	Yes	Yes	Hypersp & GT
Over 1000 pts/m²	Yes	Yes	Yes	Yes	Yes	Hypersp & GT

Table 6: Achievability of high-resolution forestry tasks according to LiDAR density.

While the current FSS focus is on general forest attributes, future work could tighten semantic segmentation classes to leaf/wood differentiation, offering simpler structural information. Additional tasks could include trunk instance segmentation combined with Quantitative Structure Models (QSM), as the trunk represents the primary contributor to CO_2 sequestration estimation. Moreover, more ITS, FSS, tree parameters and tree species will be added to further enrich 3D3. These extensions would further leverage 3D3's multi-modal, multi-resolution data, providing a flexible platform for a wide range of ecological and forest monitoring applications.

Acknowledgments

We acknowledge the support of the project *InnovAction: Network Italiano dei Centri per l'Innovazione Tecnologica* (CUP B47H2200437000), funded by MIMIT with NPRR - NextGenerationEU funds. Authors are thankful to AVT Airborne Sensing Italia, Alto Drone and the Finnish Geospatial Research Institute (FGI) - with its funding sources Research Council of Finland through UNITE flagship (grant numbers 337127, 357906, 359175), Scan4est research infrastructure: Measuring Spatiotemporal Changes in Forest Ecosystem (grant number 337811) and MULTIRISK (grant number 353264) - for providing some of the data included in 3D3.

References

Balestra, M., Marselis, S., Sankey, T.T., Cabo, C., Liang, X., Mokroš, M., Peng, X., Singh, A., Stereńczak, K., Vega, C., 2024. LiDAR data fusion to improve forest attribute estimates: A review. *Current Forestry Reports* 10, 281-297.

Bountos, N. I., Ouaknine, A., Papoutsis, I., & Rolnick, D., 2025. FoMo: Multi-modal, multi-scale and multi-task remote sensing foundation models for forest monitoring. Proc. *AAAI Conference on Artificial Intelligence*, Vol. 39, No. 27.

Dalponte, M., Bruzzone, L., & Gianelle, D., 2012. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sensing of Environment*, 123, 258–270.

Dalponte, M., & Coomes, D. A., 2016. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. *Methods in Ecology and Evolution*, 7(10), 1236–1245.

Dechesne, C., Mallet, C., Le Bris, A., & Gouet-Brunet, V., 2017. Semantic segmentation of forest stands of pure species combining airborne LiDAR data and very high resolution multispectral imagery. *ISPRS J. Photogrammetry and Remote Sensing*, 126, 129-145.

Douss, R., & Farah, I. R., 2022. Extraction of individual trees based on Canopy Height Model to monitor the state of the forest. *Trees, Forests and People*, 8, 100257.

Ferraz, A., Mallet, C., Jacquemoud, S., Gonçalves, G. R., Tomé, M., Soares, P., Pereira, L. G., & Bretar, F., 2015. Canopy density model: A new ALS-derived product to generate multilayer crown cover maps. *IEEE Trans. on Geoscience and Remote Sensing*, 53(12), 6776–6790.

Heinzel, J., & Koch, B., 2012. Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. *Int. J. Applied Earth Observation and Geoinformation*.

Jakubowski, M. K., Guo, Q., & Kelly, M., 2013. Tradeoffs between lidar pulse density and forest measurement accuracy. *Remote Sensing of Environment*, 130, 245–253.

Li, Q., Hu, B., Shang, J., & Remmel, T.K., 2025. Two-stage deep learning framework for individual tree crown detection and delineation in mixed-wood forests using high-resolution Light Detection and Ranging data. *Remote Sensing*, 17(9), 1578.

Li, W., Guo, Q., Jakubowski, M. K., & Kelly, M., 2012. A new method for segmenting individual trees from the LiDAR point cloud. *Photogrammetric Engineering & Remote Sensing*, 78(1), 75–84.

Liang, X., Qi, H., Deng, X., Chen, J., Cai, S., Zhang, Q., Wang, Y., Kukko, A., Hyyppä, J., 2024: ForestSemantic: A dataset for forest semantic learning from close-range sensing. *Geo-spatial Information Science*, 28(1), 185–211.

Liu, Y., Chen, D., Na, J., Peethambaran, J., Pfeifer, N., & Zhang, L., 2025. Segmentation of individual trees in TLS point clouds via graph optimization. *IEEE Trans. on Geoscience and Remote Sensing*, 63, 1–21.

Lucas, M., Pukrop, M., Beckschäfer, P., & Waske, B., 2024. Individual tree detection and crown delineation in the Harz National Park from 2009 to 2022 using Mask R–CNN and aerial imagery. *ISPRS Open J. Photogrammetry and Remote Sensing*, 13, 100071.

Ma, Z., Bocaux, L., Takhtkeshha, N., Remondino, F., 2025. Towards detailed and accurate forest inventory with multi-source LiDAR data. *Photogrammetric Engineering & Remote Sensing*, in press.

Nemmaoui, A., Aguilar, F. J., & Aguilar, M. A., 2024. Benchmarking of individual tree segmentation methods in mediterranean forest based on point clouds from unmanned aerial vehicle imagery and low-density airborne laser scanning. *Remote Sensing*, 16(21), 3974.

Puliti, S., Pearse, G., Surový, P., Wallace, L., Hollaus, M., Wielgosz, M., & Astrup, R., 2023. FOR-instance: A UAV laser scanning benchmark dataset for semantic and instance segmentation of individual trees (Version 1) [Data set]. *Zenodo*.

Ruoppa, L., Oinonen, O., Taher, J., Lehtomäki, M., Takhtkeshha, N., Kukko, A., Kaartinen, H., Hyyppä, J., 2025. Unsupervised deep learning for semantic segmentation of multispectral LiDAR forest point clouds. *ISPRS J. Photogrammetry and Remote Sensing*, Vol. 228, pp. 694-722

Rusu, R. B., & Cousins, S., 2011. 3D is here: Point cloud library (PCL). IEEE *International Conference on Robotics and Automation (ICRA), Shanghai, China*, 1–4.

Shen, X., Huang, Q., Wang, X., Li, J., & Xi, B., 2022. A Deep Learning-Based Method for Extracting Standing Wood Feature Parameters from Terrestrial Laser Scanning Point Clouds of Artificially Planted Forest. *Remote Sensing*, 14(15), 3842.

Taher, J., Hyyppä, E., Hyyppä, M., Salolahti, K., Yu, X., Matikainen, L., et al., 2025. Multispectral airborne laser scanning for tree species classification: a benchmark of machine learning and deep learning algorithms. *arXiv:2504.14337*.

Takhtkeshha, N., Mandlburger, G., Remondino, F., Hyyppä, J., 2024: Multispectral Light Detection and Ranging Technology and Applications: A Review. *Sensors*; 24(5):1669.

Takhtkeshha, N., Bocaux, L., Ruoppa, L., Remondino, F., Mandlburger, G., Kukko, A., & Hyyppä, J., 2025. 3D Forest Semantic Segmentation using multispectral LiDAR and 3D deep learning. *arXiv preprint* arXiv:2507.08025.

Thomas, H., Qi, C. R., Deschaud, J.-E., Marcotegui, B., Goulette, F., & Guibas, L. J., 2019. KPConv: Flexible and deformable convolution for point clouds. Proc. *ICCV*, pp. 6410–6419.

Weinstein, B. G., Marconi, S., Bohlman, S. A., Zare, A., Singh, A., Graves, S. J., & White, E. P., 2021. A remote sensing derived data set of 100 million individual tree crowns for the National Ecological Observatory Network. *eLife*, 10, e62922.

Wu, X., Jiang, L., Wang, P. S., Liu, Z., Liu, X., Qiao, Y., Ouyang, W., Tong, H., Zhao, H., 2024. Point Transformer V3: Simpler, faster, stronger. *CVPR*, 4840–4851.

Xi, Z., & Degenhardt, D., 2025. A new unified framework for supervised 3D crown segmentation (TreeisoNet) using deep neural networks across airborne, UAV-borne, and terrestrial laser scans. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, 15, 100083.

Xi, Z., Hopkinson, C., 2022. 3D Graph-Based Individual-Tree Isolation (Treeiso) from Terrestrial Laser Scanning Point Clouds. *Remote Sensing*.

Xiang, B., Wielgosz, M., Kontogianni, T., Peters, T., Puliti, S., Astrup, R., & Schindler, K., 2024. Automated forest inventory: Analysis of high-density airborne LiDAR point clouds with 3D deep learning. *Remote Sensing of Environment*, 305, 114078.

Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., & Yan, G., 2016. An easy-to-use airborne LiDAR data filtering method based on cloth simulation. *Remote Sensing*, 8(6), 501.