International Journal of Geo-Information

MDPI

Article

Concepts Towards Nation-Wide Individual Tree Data and Virtual Forests

Matti Hyyppä [1], Tuomas Turppa [1], Heikki Hyyti [1], Xiaowei Yu [1], Hannu Handolin [1,2], Antero Kukko [1], Juha Hyyppä [1,2] and Juho-Pekka Virtanen [1,3,*]

[1] Department of Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute, Vuorimiehentie 5, 02150 Espoo, Finland; matti.hyyppa@nls.fi (M.H.); tuomas.turppa@nls.fi (T.T.); heikki.hyyti@nls.fi (H.H.); xiaowei.yu@nls.fi (X.Y.); antero.kukko@nls.fi (A.K.); juha.hyyppa@nls.fi (J.H.)
[2] Department of Built Environment, School of Engineering, Aalto University, P.O. Box 11000, 00076 Aalto, Finland; hannu.handolin@aalto.fi
[3] Forum Virium Helsinki, 00130 Helsinki, Finland
* Correspondence: juho.virtanen@nls.fi

Abstract: Individual tree data could offer potential uses for both forestry and landscape visualization but has not yet been realized on a large scale. Relying on 5 points/m^2 Finnish national laser scanning, we present the design and implementation of a system for producing, storing, distributing, querying, and viewing individual tree data, both in a web browser and in a game engine-mediated interactive 3D visualization, "virtual forest". In our experiment, 3896 km^2 of airborne laser scanning point clouds were processed for individual tree detection, resulting in over 100 million trees detected, but the developed technical infrastructure allows for containing 10+ billion trees (a rough number of log-sized trees in Finland) to be visualized in the same system. About 92% of trees wider than 20 cm in diameter at breast height (corresponding to industrial log-size trees) were detected using national laser scanning data. Obtained relative RMSE for height, diameter, volume, and biomass (stored above-ground carbon) at individual tree levels were 4.5%, 16.9%, 30.2%, and 29.0%, respectively. The obtained RMSE and bias are low enough for operational forestry and add value over current area-based inventories. By combining the single-tree data with open GIS datasets, a 3D virtual forest was produced automatically. A comparison against georeferenced panoramic images was performed to assess the verisimilitude of the virtual scenes, with the best results obtained from sparse grown forests on sites with clear landmarks. Both the online viewer and 3D virtual forest can be used for improved decision-making in multifunctional forestry. Based on the work, individual tree inventory is expected to become operational in Finland in 2026 as part of the third national laser scanning program.

Keywords: virtual forest; individual tree detection; airborne laser scanning; national tree database; precision forestry

check for updates

Citation: Hyyppä, M.; Turppa, T.; Hyyti, H.; Yu, X.; Handolin, H.; Kukko, A.; Hyyppä, J.; Virtanen, J.-P. Concepts Towards Nation-Wide Individual Tree Data and Virtual Forests. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 424. https://doi.org/10.3390/ijgi13120424

Academic Editors: Godwin Yeboah and Wolfgang Kainz

Received: 6 September 2024
Revised: 14 November 2024
Accepted: 21 November 2024
Published: 26 November 2024

Copyright: © 2024 by the authors. Published by MDPI on behalf of the International Society for Photogrammetry and Remote Sensing. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Forests are a significant part of the planetary "life support system". They are not just a source of wood and other resources, but also provide a plethora of ecosystem services ranging from climate regulation, water supply and regulation, wild foraging, air purification, erosion control, and biodiversity preservation [1,2] Managing the economical and ecological benefits is a complicated balancing act. Sometimes, even the ecological services can conflict. For example, the abundance of dead wood is a biodiversity indicator, as many species use decaying wood as habitat. At the same time, dead wood is also a slow carbon source. Should we prioritize biodiversity preservation by leaving dead wood to remain in forests, or focus on faster carbon capture with live trees, thus maximizing carbon sinks with economic exploitation of forests? To balance different interests and respond to topical ecological challenges, there is an urgent need for tools allowing the optimization

of the forest value chain on a societal level [3,4]. Optimization, however, requires detailed knowledge of forests, preferably at the country level.

Forest inventory forms the basis for decision-making, both in the forest industry and broadly in society. Improved forest inventories are needed to meet the needs of multifunctional precision forestry. High-quality forest inventory data are beneficial for multiple purposes: (1) by providing an improved knowledge of the quality of wood; (2) by providing accurate information on the quantity and availability of wood; (3) by enabling advanced electronic wood trade; (4) by improvement to the bucking of trees; (5) improvement to certification of wood origin; (6) improved forest value estimation; (7) improved logistics chain; (8) removal of damaged trees (broken tops, forked trees) from forests in thinning operations; and (9) conservation of the most vital and biodiversity-preserving trees in forests. According to Kangas et al. [5], even the current area-based inventory with a 16 m by 16 m pixel size, when updated every 5 years, could provide added value of approx. 56 €/ha by well-informed decisions by owners and parties involved in the wood trade, in the context of Finland. Additionally, Kangas et al. [5] reported that improving the relative RMSE of stem volume estimates by 1 percent-unit would benefit the Finnish forest owner by 4 €/ha in 10 years. Improved bucking would intensify the revenue.

In society, the decision-making related to forest management is also connected with landscape visualization. The importance of visualizations for communicating the potential impacts of silvicultural operations for different experts and the general public is well acknowledged (e.g., [6]). The general concept of applying interactive 3D visualization for forest environments to support planning tasks in forest management has been known since the 1990s [7]. This also likely marks the advent of the term "virtual forest". With individual tree information not being commonly available, many virtual forests have been realized relying on statistical models and randomized tree positions [8,9]. While this allows visualization on a plot level, it restricts the utility of virtual forests for decisions concerning individual trees. The methods to produce these virtual forests from large quantities of real individual tree data are also largely missing. This is a hindrance for applications like precision forestry [10] and detailed landscape planning, which would benefit from the ability to choose individual trees from the forest for operations planning.

The two main approaches for providing forest inventory data from laser scanning are the area-based approach (ABA, based on e.g., 16 m by 16 m grid) (e.g., [11]) and individual tree detection (ITD) [12], the latter also being called "single-tree inventory". In ABA, a number of statistical descriptors are estimated for a given area (e.g., a grid cell) from the laser scanning data, whereas ITD relies on actually detecting the individual trees [12,13]. Area-based variables can of course be calculated from individual tree characteristics later on, if they are needed.

The ITD requires a dense enough point cloud to reach sufficient accuracies. In addition, the computational capabilities to store and utilize forest inventory data consisting of large collections of individual trees are needed. This has limited ITD's practical application to mostly research settings and potentially as a source of calibration data for wider area ABA [13]. Several methods have been developed for ITD, with varying accuracies found in different forest conditions and point cloud densities [14–18]. The use of full waveform laser scanning has also been found beneficial for ITD processing [19]. Despite these findings, large-scale forest inventories are still performed with ABA. At the same time, there is an ever-increasing need for obtaining further tree-specific information, such as wood quality parameters [20]. Large-scale ITD also remains interesting for "precision forestry" [10]. Currently, ABA is unable to meet use cases that require data concerning individual trees. Neither can ABA differentiate growth, damage of tree tops and fallen/harvested trees, nor the quality of trees within each grid cell. Methods used to store and utilize forest inventories will have to be revamped to operate with ITD data to realize ITD-based forest inventories. This development is required even if future inventories are based on the synergetic use of ABA and ITD, as ABA accurately provides the average value of stand characteristics and ITD provides the variance of such characteristics with a focus on log-size trees.

Ongoing national laser scanning programs produce ALS point clouds approaching the densities needed for ITD, on a national scale. To support decision-making related to forestry, the diameter, volume, and biomass of each tree should also be estimated in ITD. Such large-area ITD are largely missing. Krzystek et al. [21] present the segmentation and classification of over 25 million trees from a specially-acquired combination of ALS (55 p/m^2) and multispectral airborne imagery, but do not mention the estimation of tree parameters (such as stem diameter, height and volume). Ma et al. [22] reported one of the largest known areas modeled at individual tree level, 6 million trees in New York, but the error analysis was done at plot level and no information was given on the detection rate of each diameter class needed for understanding the missing biomass contribution. In the Finnish conditions, an operational country-level ITD based inventory requires an information and visualization system capable of handling 10+ billion log-size trees. It would also be beneficial to be able to derive recognizable virtual forests ("digital twin") from the same data to assist in forest decision making by joining large scale ITD with a suitable immersive visualization system.

For national individual tree based forest inventories, a number of development steps are therefore urgently needed: (1) automation of large scale ITD, (2) development of storage and distribution systems compatible with large scale ITD data and (3) development of systems that allow the interaction with individual tree forest inventories (queries, aggregation, value estimates). With these, the ITD could potentially advance from the research domain towards operational forestry and help attain the benefits expected from "precision forestry".

Thus, this study has three main goals:

(1) Development of a large-scale ITD system: We present a system for producing ITD data using Finnish national laser scanning data (5 pts/m^2), and evaluate its computational performance. We assess the quality of individual tree estimates (detection rate, height, diameter, volume, and biomass, as a function of diameter class). Additionally, we calculate timber assortment estimates including the amount of log wood, pulp and energy wood, and stored carbon.
(2) Development of ITD storage and viewer tools: We propose a fast, multi-user software infrastructure for providing a cloud-based web service for storing, distributing and viewing ITD-based forest inventories at the national scale. The service is designed to also provide APIs for potential third-party viewers, promoting wider usage.
(3) Automation of immersive forest visualization: We prototype the automatic generation of an interactive 3D virtual forest based on ITD data and evaluate its verisimilitude against on-site images from the same location. Notably, the production of the virtual 3D scenes is fully automated, relying on ITD data (accessed via the system's API) and publicly available national geospatial datasets.

2. Related Studies

2.1. Individual Tree Detection from Point Clouds

The concept of individual tree detection using laser scanning data, and its use on ITD or ABA style inventories was presented in [12,23,24]. In [12], ITD was assumed to have four steps: (1) recording of the forest so that individual trees or tree groups are visible, (2) calculation of normalized height, (3) use of segmentation/clustering to detect individual trees or group of trees, and (4) estimation of stand characteristics, such as volume or basal area from the features that can be directly measured (position, height, crown area and species). For improving the separation capacity of overlapping trees, trunk hits [25], full waveform laser scanning [19], canopy height maps created using low returns [26], and clustering points inside the watersheds into several trees [27] have been proposed. Various ALS ITD methods have been benchmarked or reviewed in [14–18]. In typical implementations, the tree positions, heights, canopy areas and crown-related features are directly derived whereas stem diameter and stem volume are predicted by non-parametric estimation (e.g., Random Forest) and directly measured features. For species determination, additional data sources such as RGB images are also applied [28,29].

The segmentation of single trees is complicated by their complex spatial configuration in forests: tree crown adjacency and intersection, differences and changes of tree shape with age, canopy structure, and understory vegetation occur both within a single tree species and between different species [18]. Segmentation methods operating on point clouds can be based on (1) spacing between trees and their spatial distribution, (2) point cloud characteristics describing the tree shape, and (3) intensity, or the combinations of the aforementioned. Combination enables few commission and omission errors. Many established clustering techniques, such as K-means, DBSCAN, hierarchical clustering, and mean-shift, have also been applied in segmentation. Initial solution is typically based on preliminary detection of tree tops. Combination of segmentation and clustering are also applied e.g., in cases where segmentation produces a potential group of trees and then clustering is applied to dense point clouds to separate trees inside the segment/group [27].

Most of the individual tree detection algorithms belong to the following categories:

- Raster-based processing of canopy height [12]
- Voxel-based processing of canopy height [30]
- Vector or point-cloud based delineation of canopy height [31]
- Mixed representations [32]
- Detection of stems [25]
- Direct measurement of stem curve from point cloud data [33]
- Use of deep learning for stem detection [34]

Determination of key parameters for segmentation requires field reference data to adapt to different crown geometries and for different strata. Limited number of key parameters and small computational complexity are typically required. Terrestrial, mobile and UAV laser scanning, all operating at a much closer range than ALS, can be applied for direct estimation of DBH, e.g., [35] and stem curve (e.g., [36]) and in general for deriving the field reference data for ALS-based inventories. While dense close-range observations would potentially increase the detail level of forest inventory, these techniques are not yet feasible for surveying forests on a national scale. Instead, they are aimed at automating field reference collection, which is otherwise costly and laborious. State-of-art of field reference collection using traditional techniques is depicted in [37]. Deep learning is currently improving segmentation accuracies with high-density data obtained from ALS sensors developed for drone or helicopter use (e.g., [34,38]). Here, it can be expected that deep learning methods will improve segmentation results if manually labeled data can be used for teaching. However, manual annotations from ALS point clouds were tested in [13] and in [14] for 10 and 8 points per m^2 (both higher point densities than our source data), respectively. In both studies, automatic methods were found better than manual annotations. More specifically, in [13] automatically derived segmentation based on 509 reference plots were manually improved with the result of reduced accuracy. In [14] it was concluded that several tested segmentation methods were superior to manual annotations in the dominant, co-dominant, and suppressed tree storeys. Hakula et al. [39] demonstrated that results are comparable for the detection of dominant and co-dominant trees between simple (2D) and sophisticated (3D) machine-learning methods. In forests that are typically considered difficult for ITD, such as very dense forests, raster-based methods worked even better than 3D methods. For detecting suppressed or smaller trees, 3D methods performed better than 2D methods. However, the detection capabilities are primarily limited by the decrease in point density due to the block of top layer [15].

Typically, ALS data are used for ABA, as recently by Hauglin et al. [40]. ALS and deep learning can be used to derive canopy height estimates from airborne imagery, as done by Wagner et al. [41] for California and by Tolan et al. [42] at jurisdictional scale. As stated before, when performing country-level ITD with current national ALS datasets, the diameter, volume, species, and biomass of each tree should be estimated to support decision-making. While country-wide ITD has been demonstrated [43], there are only a few previous studies analyzing ITD by producing diameter, volume, species, and biomass estimates of each tree, while using airborne laser scanning data as the remote sensing data

source. Additionally, most of the studies carried out at individual tree level report errors at the plot (ABA) level, such as the recent paper on the trees of New York [22] which reports overestimation of tree count, mean tree height and maximum tree height by 11.1%, 8.6%, and 5.3%, respectively. In [44], volumes were estimated with errors ranging from 30% to 37% at individual tree level, when using a point density of 10 transmitted pulses per m^2. As a comparison, allometric models and manual field inventories estimated individual tree volumes with 17% random error [44]. Ma et al. [22] are also one of the largest known areas modeled at individual tree levels including height, diameter, and volume/biomass.

2.2. Virtual Forests

Already in the early examples of interactive 3D forest visualization, several important aspects of virtual forests were brought up: Orland et al. [7] implemented the visualization consisting of a DTM and individual trees that carried information of their 3D position, height, DBH, condition, and species. The visualization system was connected to a database storing this data. At the time, computer performance and the availability of individual tree data limited these to some thousands of trees at maximum. Later on, two different views of the same data were introduced, one offering a simplified visual representation for management in larger areas and another one intended for landscape visualization, aiming to maximize the amount of visual realism via texturing [45]. The forest visualizations were linked with growth estimation systems to illustrate future scenarios [6]. Fabrika [46] also mentions visualization of different times of the day in the virtual forest and transferring the decisions made in the virtual forest to other systems. This way, the virtual forest became not the full toolbox for silviculture, but rather a visualization system. Uusitalo and Orland [6] listed encountered obstacles to applying virtual forests: (a) the lack of single-tree inventories, (b) the difficulty of geometric tree modeling or alternatively the production of realistic but varying tree objects for visualization, and (c) issues in managing different datasets, such as elevation models and other geographic data.

Later on, the virtual forests were implemented on various platforms: dedicated landscape visualization tools [47], browser-based 3D for online use [8,48], and 3D game engines [9,49]. Out of these, 3D game engines have become the most applied platform for virtual forests when aiming for a near-photorealistic visualization. In game engines, tree model libraries have simplified the production of virtual forests [50]. They also offer tools for optimizing the rendering, e.g., via the use of different levels of detail, and culling the trees to be rendered according to camera position [50]. An additional motivator for using commercial game engines for producing virtual forests has been their support for virtual reality head-mounted displays, offering immersive visualization. Huang et al. [9] present the use of virtual reality head-mounted displays (VR HMDs) in conjunction with a virtual forest illustrating the impacts of climate change. In their work, the forest modeling for visualization also includes the understory vegetation—not just the trees [9]. VR HMD virtual forest has also been tested as a restorative environment [49].

To summarize, the past two decades have simplified the realization of increasingly photorealistic, immersive virtual forests. In addition, the global open data movement has made GIS data, such as elevation models and various vegetation maps openly available for several countries. Methods to apply these in game engines have been developed as a part of geoinformation research. Combined, this has solved many of the issues mentioned by Uusitalo and Orland [6] concerning the realization of virtual forests, apart from the availability of large-scale single-tree inventories.

2.3. Virtual Trees

Several alternative solutions have been proposed for visualizing individual trees in virtual environments. On a general level, these can be divided into billboard approaches and 3D model approaches.

The billboard approach refers to visualizing a tree by using a textured plane that is set to rotate along its vertical axis so that the normal is always facing the camera. The

polygon count remains low as only a single plane is needed for each tree. The textures used for tree billboards can be obtained in several ways. In [8] real images from the field are edited to remove backgrounds. In [47] procedural modeling software is used to produce synthetic images. The use of synthetic textures reduces the amount of fieldwork required but requires procedural modeling tools that are able to produce the needed tree species.

The trees can also be represented in the virtual environment by 3D objects. In [7], simple 3D cones were used to denote trees. In [45] different 3D models were applied for different species, and their colorization was leveraged for data visualization. Currently, a similar approach of using simplistic 3D objects for denoting individual trees is commonly used in virtual globes, online 3D maps, and city model viewers (e.g., [51,52]).

In Han and Huang [53], more refined textured 3D models are applied separately for the evergreen and broadleaf trees. In [9], procedural modeling is applied to obtain synthetic 3D models of different tree species with specified heights and crown sizes. Procedurally generated models are beneficial for visualization: as each tree is unique, no unwanted repetitive effects occur. At the same time, the actual information stored per each tree remains compact. When using procedural modeling for texture generation, the model generators have to be able to represent the required species. Huang et al. [9] report having to combine tree models from two different procedural engines for sufficient species coverage. On game engine platforms, using procedurally generated tree models has become the most common approach for near-photorealistic visualization. The parameters for procedural tree model generation have also been extracted directly from a dense mobile laser scanning point cloud, allowing more detailed control of the procedural modeling [54]. Point clouds have also been suggested to directly represent trees and forests [55,56] and applied in some commercial examples of online visualization (e.g., [57]). Both of these directions are supported by the improvement of point cloud densities and classification (e.g., [58]).

Relying on detailed point clouds, methods for modeling single trees with a high level of detail have been proposed [59], with some open-source tools available: [60]. However, the approach of modeling individual trees in more detail has not been integrated with the visualization of larger areas, most likely due to the required computational power and data acquisition.

3. Materials and Methods

3.1. ITD Test Site and ALS Data

For ITD, the 5 points/m^2 point clouds from the ongoing Finnish national laser scanning programme were applied. Started in 2020, the programme aims to cover the entire Finland in six years. The classified laser scanning data are available (by permission) from the already surveyed areas as a file download, in 1 km^2 blocks in LAS format. The dataset is further described by the National Land Survey [61].

A single production area of the national laser scanning campaign was chosen as a test site for producing individual tree data (Figure 1). This dataset consisted of 3312 km^2 of ALS data, covering the area located between the cities of Hämeenlinna, Tampere, and Lahti. The chosen area also covers the site of SCAN FOREST (https://www.scanforest.fi/ (accessed on 12 September 2024)) research infrastructure, offering reference data, as in Section 3.7.1. According to metadata, the ALS mapping of the chosen area was performed in 2019 with the same specifications as the national laser scanning programme.

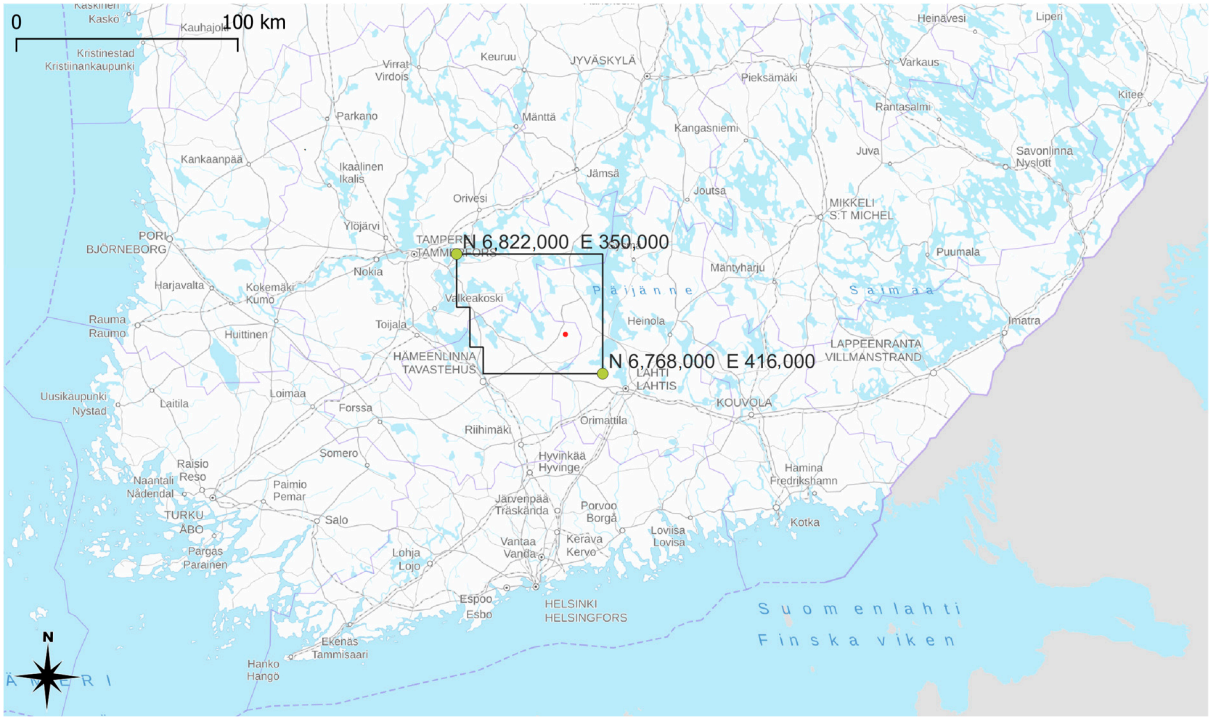


Figure 1. The area for which the ITD was performed with coordinates for the corner points given in ETRS-TM35FIN. The location of the virtual forest test site (within the ITD area) is denoted as red dot. The utilized sample plots of the SCAN FOREST research infrastructure are located in the same area as the virtual forest test site, on an approx 5 by 5 km area. Background map © National Land Survey of Finland.

For predicting tree species, crown base line, diameter, stem volume and biomass, the reference data from the SCAN FOREST infrastructure was utilized. The SCAN FOREST research infrastructure (https://www.scanforest.fi/), located in Evo, 130 km north of Helsinki, includes approximately 10,000 individual trees measured with conventional field inventory techniques and terrestrial laser scanning. The 5 km × 5 km research infrastructure is dominated by Scots pine (Pinus sylvestris; 40%), Norway spruce (Picea abies; 35%) and deciduous trees (mainly birch; 24%). Field measurements were conducted in 2019 with 37 sample plots of 32 m × 32 m, these data, including 3581 trees, were used in the training and testing of the machine learning method. Sample plots represent the range of forest conditions in the study area, including mature and young stands, as well as pine-, spruce-, birch-dominated, and mixed stands. For each sample plot, tree height and DBH have been manually measured for all trees with a DBH greater than 5 cm, using a hypsometer and caliper. Species were recorded by visual estimation. In addition, the living crown baseline was recorded. Tree locations were obtained from a tree map created with the TLS. The volume and biomass of each tree were calculated using national allometric models that are formulated based on diameter, height, and species.

3.2. Test Site for Virtual Forest

For the development of a virtual forest, a 1 km^2 test site (Figure 2) was defined from the larger area selected for ITD (Figure 1). The virtual forest test site is located near Evo campus shared by Häme University of Applied Sciences (HAMK) and Häme Vocational Institute. The selected area includes various forest and environment types (as described in Section 3.7.2).

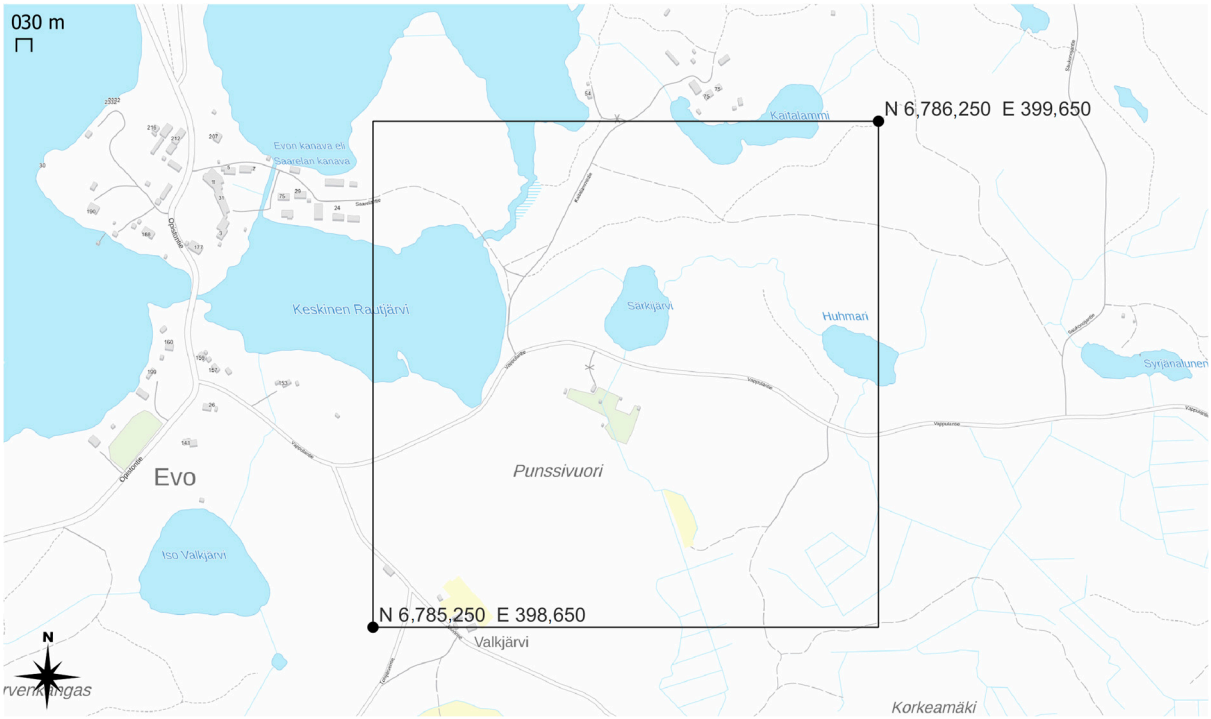


Figure 2. The chosen test area, with dimensions of 1 by 1 km, with coordinates for the corner points given in ETRS-TM35FIN. Background map © National Land Survey of Finland.

3.3. Supporting Datasets for Virtual Forest

In addition to the ITD data, open GIS datasets were utilized producing the virtual forest. A digital elevation model (DEM) with 2-m resolution [62] was utilized for producing the terrain. The DEM is based on laser scanning and offered as a set of GeoTIFF files from the National Land Survey of Finland (NLS) download service.

For describing the properties of forest to support ground texturization and undergrowth modeling, the Multi-source National Forest Inventory rasters of 2019 from the Natural Resources Institute Finland were applied. Out of the 45 thematic raster layers in the collection, the site fertility classes were used for this purpose. The site fertility index divides the land areas into ten vegetation classes according to their properties (Table 1). For a further description of the dataset and its production, the reader is advised to see: [63,64].

Table 1. Site fertility index classes.

	Class Description	Present in Test Data
1	Herb rich sites	Yes, applied
2	Herb rich heath forests	Yes, applied
3	Mesic forests	Yes, applied
4	Sub-xeric forests	Yes, applied
5	Xeric forests	Yes, applied
6	Barren forests	Yes, applied
7	Rocky and sandy soils, alluvial lands	Yes, applied
8	Summit and fjell land withsingle coniferous trees	No, not applied
9	Mountain birch dominated fjells	No, not applied
10	Open fjells	No, not applied

Aerial orthophotos (pixel size 0.5 m, see: [65]) were applied to tone the procedurally generated terrain textures and, thus, enhance the recognisability of the area from a distance. To populate the virtual environment with potentially recognizable features other than trees or landforms, a set of geospatial features from the national topographic database were applied to produce matching objects in the game engine visualization, namely lakes, roads, and large boulders.

3.4. Tree Detection from Airborne Laser Scanning

Two algorithms for ITD were experimented with. The first one was a raster-based method, one of the in-house developed methods mentioned in an earlier benchmarking study [14]. The algorithm starts with the calculation of the canopy height model with a resolution of 0.5 m $\times$ 0.5 m. Then, the tree tops are found by local maximum filtering followed by watershed transformation for crown delineation. The second one was a voxel-based version of the method in [18,66]. This is a top-down region growing method. The method segments trees iteratively from the tallest tree to the shortest based on a distance threshold and a minimum spacing rule. The setting for these two parameters should be adapted for the crown size and the tree spacing in the area. ITD was performed with both raster- and voxel-based methods for reference plots. We did not use deep learning, which is discussed more in Section 5.1. Only the raster-based method was applied to large area ITD due to its efficiency and the low point density of the source data. After detecting the individual trees, a set of tree features were estimated from the point cloud (Table 2).

Tree attributes (tree height, DBH, volume, and biomass) were derived for each segmented tree, based on estimation models developed using tree features (Table 2) as predictor variables and the field-measured tree attributes (Table 3), obtained from SCAN FOREST, as a response variable in a non-parametric regression for matched trees using random forests technique. Tree matching between detected trees and field-measured trees was carried out based on the method presented by Yu et al. [67]. A successful match is determined by the distance between a detected tree and a field-measured tree in an n-dimensional space. In this study, we applied the method in a 3-dimensional space (x, y, and height). The trees are considered to form a match if they are the closest pair formed by a field-measured tree and a detected tree within a distance threshold (e.g., 3 m). The methodology for building prediction models follows those described by Yu et al. [68]. Different estimation models were developed separately for different tree attributes. Among the 25 predictor variables (excluding tree ID, location, and intensity features), 10 best were selected to build estimation models based on the feature importance. This is a built-in measure of the random forest technique. Timber assortment estimates (amount of log wood, pulp, and energy wood) were obtained using existing conversion tables [69]. The value of the wood has been obtained from the 2022 yearly statistics of Natural Resources Institute Finland. Stored carbon is based on biomass and uses reference [70].

Table 2. The description of the tree features extracted from the laser point cloud.

Description
tree ID
tree location (location of the max. height in canopy height model)
tree height (based on highest returns)
ground elevation at tree location (interpolated from ground points)
height of lowest branch (limited accuracy)
mean height of all returns
standard deviation of the height for all returns

Table 2. *Cont.*

Description
skewness of the height for all returns
the height at which the number of returns are greatest
crown height (difference between highest and lowest points)
crown area as 2D convex hull (based on points belonging to the tree segment)
crown volume as 3D convex hull
canopy height distributions at ith percentiles
Cumulative percentage of points in the ith layer which is divided equally between 2m and tree top.
mean value of third component of the normal vectors to crown surface triangulations
proportion of only returns to total returns
proportion of first returns to total returns
proportion of last returns to total returns
proportion of intermediate returns to total returns
minimum, mean and standard deviation of intensity for all points above ground.
minimum, mean and standard deviation of intensity for the points below 25th percentile of height and above ground.
minimum mean and standard deviation of intensity for the points above 80th percentile of height

Tree species were also classified using the random forest technique with the features derived from tree segments (Table 2) as predictor variables. To avoid overfitting, the best predictor variables were selected based on trial-and-error and feature importance. For species classification with random forest, we constructed 500 decision trees using the best 12 features. Only three main tree species (pine, spruce, and birch) were classified because of the small number of other species in the sample plots. Among the three main species, the proportion of the trees used in the study is 32.7%, 35.9%, and 25.1% for pine, spruce, and birch respectively, with other species being 6.3% in total.

Table 3. The descriptive statistics related to reference trees (N = 3581).

	Minimum	Maximum	Mean	Standard Deviation
H (m)	2.5	38.40	16.98	7.55
DBH (cm)	5.05	73.15	17.64	10.47
VOL (m^3)	0.01	5.0	0.36	0.56
Biomass (kg)	3.75	2229.80	175.51	241.08

The ITD from national ALS operated by processing 1 km^2 tiles individually. The program was parallelized to use all CPUs of the system by using MATLAB Engine API for Python and by running multiple MATLAB engines simultaneously, each processing 1 km^2 of ALS data at a time. The program was run on a Dell Precision 7920 Tower Workstation computer (Dell Inc., Round Rock, Texas, USA). The computer system has an Intel Xeon Gold 6234 processor with 8 cores, 16 threads, and 256 GB of RAM.

To improve the performance of the tree detection, the ITD algorithm was improved to leverage all the CPUs of the utilized hardware. Parallelization was achieved by processing multiple disjoint areas of forest simultaneously utilizing MATLAB Engine API for Python (version 23.2). MATLAB Engine API for Python provides a package for Python to call MATLAB as a computational engine and, therefore, multiple MATLAB engines could be run simultaneously, each of which would process one LAS file at a time. In addition to detecting the trees, the algorithm extracts a set of feature descriptors from the geometric properties of the segmented point cloud, such as the vertical distribution of points and intensity of the returns (Table 2) which are used for the determination of the tree characteristics.

It is worth mentioning that there would have been other available approaches to parallelize the ITD algorithm. For example, the Parallel Computing Toolbox developed by Mathworks could have been used to improve the performance of ITD. However, the toolbox is a proprietary software developed by Mathworks and the ITD algorithm does not depend on the toolbox. The performance improvement obtained with our solution was also enough to demonstrate the feasibility of nation-wide ITD. Therefore, utilizing MATLAB Engine API for Python was considered a viable solution.

3.5. Single-Tree Forest Inventory Data Storage and Distribution

As the ITD data consist of tree points with their associated parameters, several existing open software components are available for storing and visualizing it. Figure 3 illustrates the software architecture applied. PostgreSQL database server was used for storing the ITD tree points. The database is accessed by a self-developed Django HTTP REST API, run as a dockerized service. The API provides two endpoints for the 3D game engine (Section 3.6) for querying tree data. From the first endpoint, individual trees can be queried using a query geometry (a bounding box or GeoJSON polygon). The trees intersecting the query geometry are returned as a GeoJSON feature collection. From the second endpoint, tree counts can be queried for a bounding box that is divided into equal-sized grid cells. The response of this endpoint is a list of grid cells in JSON format and the tree count calculated separately for each cell.

In addition, the API provides an endpoint for querying statistical data of the trees within a user-defined query geometry. The statistical data includes the tree count, maximum height, average height, maximum diameter, average diameter, volume, biomass and basal area. The response also contains a height and diameter distribution, both represented as histograms. This API endpoint is utilized by the 2D browser-based viewer (Section 4.3), built with the open-source Leaflet.js library.

To facilitate efficient visualization of large tree amounts (e.g., when viewing large areas), a separate process was developed to pre-render 2D raster map tiles to an image pyramid (generated using the open-source Mapnik library). The image pyramid is utilized by the 2D browser-based viewer. The pre-rendered raster map tiles enable fast tree visualization even though the number of trees in the visible area of the map would be very large. An additional API endpoint was developed in the Django application to determine whether the user has clicked a tree or not and to return the attributes of the corresponding individual tree. This enables the 2D browser based viewer to display a popup element showing the attributes of a tree that the user has selected on the map.

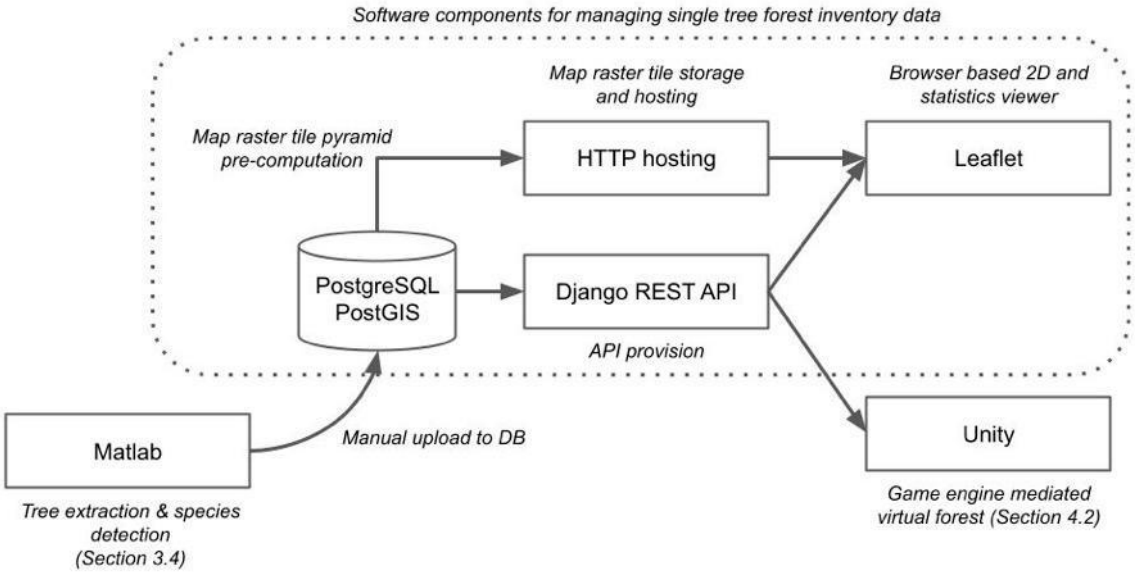


Figure 3. Overview of the software components applied.

PostgreSQL was selected as the data storage for storing the individual tree data processed by the ITD algorithm. PostgreSQL is an open-source object-relational database management system. PostGIS database extension was utilized to extend the features of PostgreSQL by adding support for storing, indexing, and querying geospatial data. The ITD data was uploaded into a database table which stored the predicted height, DBH, stem volume, biomass, and species in addition to the individual tree attributes presented in Table 2.

To ensure scalability in large forest areas, suitable database indexes needed to be created in the database. To speed up queries searching for trees with a height from a specific range, a B-Tree index was created on the tree height column. In addition, a GiST index was created on the tree coordinates column to enable efficient spatial queries over the tree records.

To facilitate efficient visualization of large forest areas in the 2D browser-based viewer, the ITD data needed to be post-processed into a tile pyramid. There are two main approaches for generating a tile pyramid, either using raster map tiles, or vector tiles. Raster map tiles were selected for generating the tile pyramid, even though vector tiles typically have some advantages over raster tiles, including lower bandwidth consumption and more advanced interactions in the client application. However, in the case of the ITD data, vector tiles turned out to consume significantly more bandwidth compared to raster tiles because of the large number of tree features to be encoded in the tiles. When visualizing an example forest area with 25,000 trees on the 2D browser-based viewer, utilizing vector tiles resulted in 495% higher bandwidth consumption compared to raster tiles.

Raster map tiles were generated using a separate tile rendering process that utilized the open-source Mapnik library (Figure 4). The process rendered raster tiles based on an XML configuration file that described how the ITD data should be rendered in the tiles. All trees taller than 5 m were rendered in the tiles for the highest zoom levels. For lower zoom levels, the tallest trees and a subset of the smaller trees were included in the tiles. TileMill, an open-source map design studio, was utilized together with the CartoCSS map design styling language to specify the rules for selecting which trees should be rendered for a specific zoom level. After specifying the rules, a Mapnik-compatible XML configuration file was exported from TileMill.

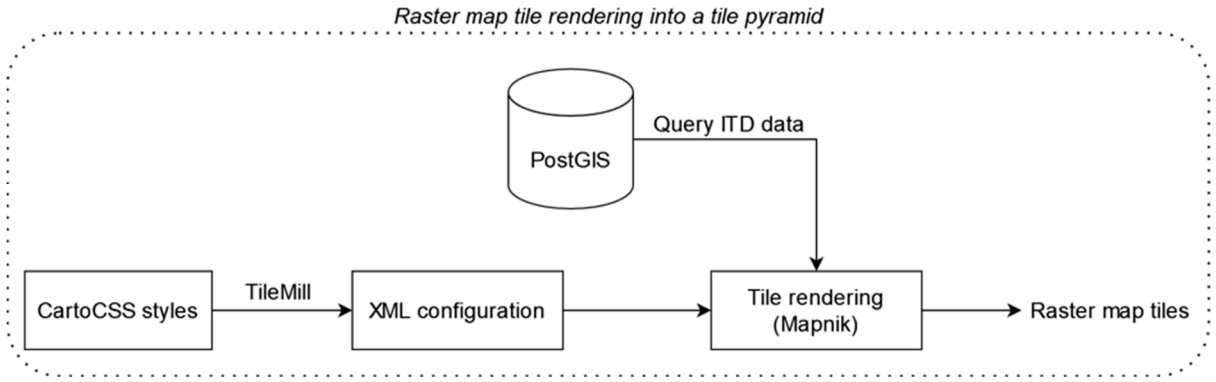


Figure 4. A pipeline for rendering raster map tiles into an image pyramid.

3.6. Game Engine Mediated Virtual Forest

The commercial 3D game engine, Unity 2020.3.18, was applied for producing the virtual forest on a Windows 10 platform. Later, during the development, the project was upgraded to Unity 2022.2.12.

The virtual scene is initiated by dividing the defined scene area into a two-dimensional grid of bounding boxes. The app connects to the tree database at the beginning of runtime and first queries the total tree count for the scene and each individual cell to reserve the needed memory for faster in-scene operations.

The terrain geometry can be produced from either an 8-bit or 16-bit grayscale image storing the DEM. For "populating" the terrain with suitable undergrowth vegetation and ground surface visualization materials, the Multi-source National Forest Inventory raster for site fertility class is used. To add a transition between different undergrowths and to smooth down the blending of the ground texture layers, the raster is rescaled to higher resolution with bilinear sampling, after which each pixel is compared by their approximated color distance to each site fertility color code. The virtual world location on the terrain is then populated randomly from a weighted pool of undergrowth vegetation models, linked to each specific color by the weight given by the distance. Aerial images are applied for fine-tuning the texture colorization of the ground. The terrain is divided into segments to facilitate optimization via frustum culling. For more distant terrain, the aerial image is applied directly as texture.

To attain a sufficient rendering performance, a level of detail system is applied for the tree objects. 2D billboard models are used for the farthest trees and 3D models are used for the more nearby trees. The tree models are generated as children of terrain grid segments to allow more efficient optimization.

A custom frustum culler was developed for the virtual camera to be used against the grid to determine the visible and the closest sectors of the area (Figure 5). The sectors are then prioritized by the distance and centrallness according to the viewer and placed on a loading stack. Stack size is dynamically increased or decreased by the performance in relation to the amount of trees in the cell after each successful load. The stack will be then emptied by requesting trees from the server for every individual grid cell in it. Trees are received in JSON format and are parsed to in-game data objects. For each cell, once the loading is done, one slot from the stack is freed and another one will take its place in a similar fashion. Data are interpreted and the trees are then generated from the closest match of prefabricated models (Section 3.6.1). The model scale is also modified to represent the trunk and height dimensions. After generation, the trees are placed on the terrain by translating the ETRS-TM35FIN coordinates to the virtual world coordinates and determining the height position by raycasting against terrain colliders masking all other objects out of the process.

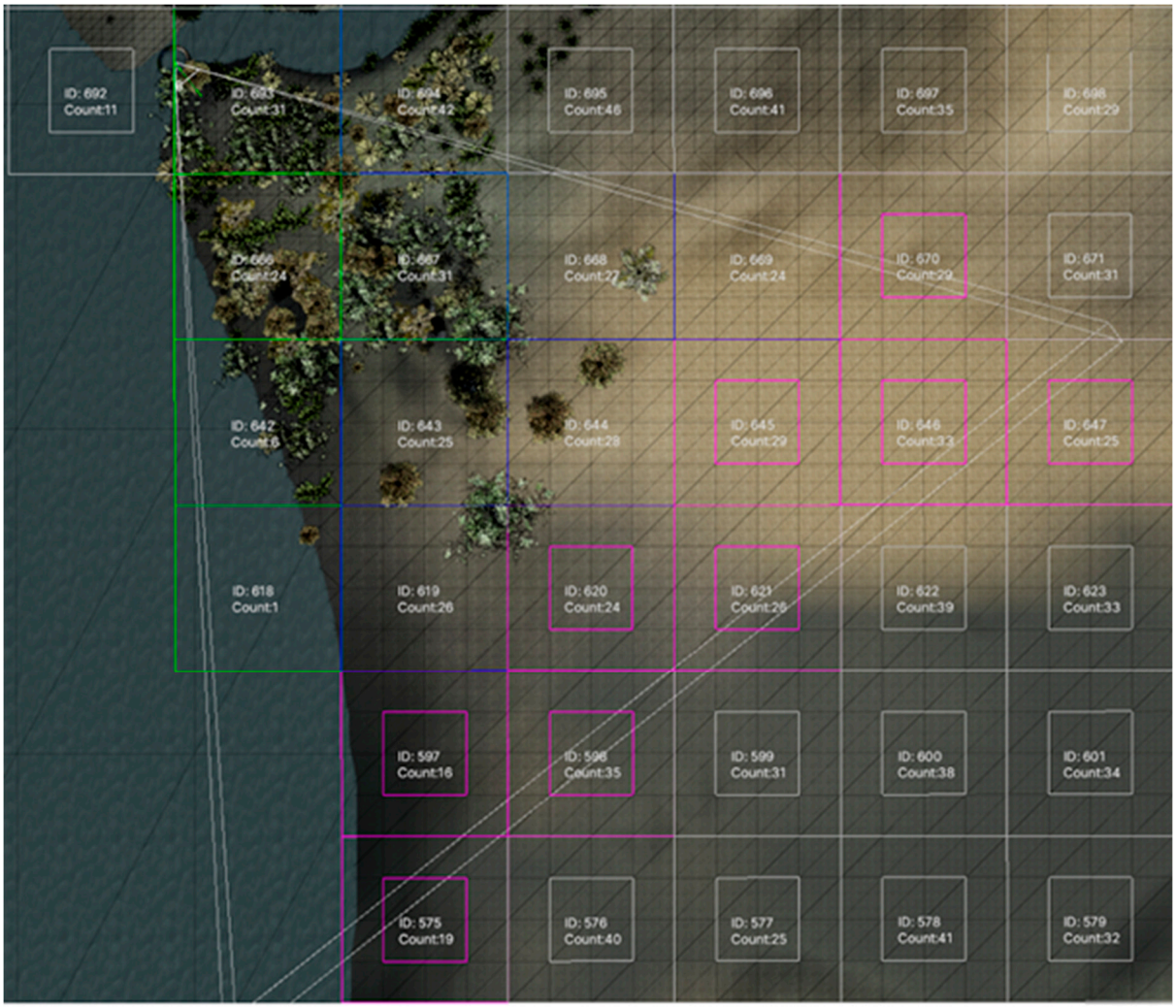


Figure 5. The terrain area is divided into a grid of bounding boxes (white) that receive in scene initiation their individual tree counts from the server. The user's (top left) current view is visualized by the frustum. The real time tree loading sequence has finalized the green outlined boxes while blue ones are currently being loaded and magenta ones wait in the stack.

3.6.1. Production of Tree Models for Visualization

Vegetation modeling software SpeedTree 7.1 was utilized for producing the game engine-compatible 3D tree models for the virtual forest. Speed Tree produces unique 3D tree models with procedural algorithms according to given parameters. As the generation of the tree models in real-time can be too resource intensive, a collection of tree models (Table 4) was generated beforehand to facilitate the visualization of trees of different ages.

Table 4. Counts of tree models in the tree 3D model collection used for the virtual forest.

Number of Tree Models Produced	Young Sapling	Young	Full Grown	Sparse Branched	Total
Pine	2	4	4	2	12
Spruce	2	4	4	2	12
Birch	2	2	4	2	10

Table 4. *Cont.*

Number of Tree Models Produced	Young Sapling	Young	Full Grown	Sparse Branched	Total
Aspen	2	4	4	2	12
Rowan	1	2	2	1	6
Total	9	16	18	9	52

3.7. Evaluation Methodology

3.7.1. Estimation of Tree Detection Performance

The performance of ITD was evaluated by comparing the SCAN FOREST reference trees measured in the field with the segmented trees. The tree detection rate (completeness) and correctness of the detected trees and F-score were calculated. The former two are also frequently referred to as recall and precision, respectively. The results from both raster- and voxel-based methods were evaluated for a comparison using the trees in the sample plots.

In applied random forest for tree attribute estimation and tree species, we only considered the matched trees in the evaluation. The method for determining the matches is described in Section 3.4. From the matched trees, two-thirds of the trees were used for teaching (1145) and one-third for testing (572) of the results for each decision tree which forms the random forest. With a bootstrap sampling strategy and out-of-bag prediction of the random forest, we could use the entire dataset for the evaluation. The evaluation was performed only on the ITD result of the raster-based method. Bias and root mean squared error (RMSE) were estimated for Height (H), diameter (DBH), volume (V), and above-ground biomass (AGB) by comparing the random forest estimates and reference values. For height and diameter, the values obtained by our method were compared with manual measurements of matched trees. For volume and biomass, the reference was formed by calculation based on height, DBH, and species [71–73]. Our method estimates the volume and above-ground biomass by correlating the teaching data with tree features (Table 2).

3.7.2. On-Site Reference for Evaluation of Virtual Forest Fidelity

For evaluating the verisimilitude of automatically produced virtual forest scenes (ITD data quality is addressed in Section 3.7.1), a collection of on-site reference data was obtained from a total of 16 test plots: (1) Potentially interesting terrain features from the 1 km^2 test site were identified from topographic maps. (2) A set of potential viewing points, allowing the study of these features, was drafted, with their characteristics and locations (Table 5). (3) The points were accessed and documented with a panoramic imaging-enabled TLS scanner (Faro Focus S350). A separate RTK-GNSS receiver (Topcon Hiper HR) was used to obtain coordinates for the scan position. The rotation of the TLS scan was fitted by manually rotating the captured point cloud and comparing it to the ALS point cloud from the same area.

Table 5. Locations and characteristics of obtained on-site reference.

ID	Forest/Terrain Type	GNSS Position	Appendix A Figure
A	Mixed-species forest by the river flowing into a lake	N = 6,786,021 E = 398,869	Figure A2
B	Uneven-aged mixed-species forest around a ditch and a forest road	N = 6,785,854 E = 398,919	Figure A3
C	An open field in the middle of a mixed forest	N = 6,785,778 E = 399,079	Figure A4
D	A hilly full-grown spruce forest	N = 6,785,759 E = 398,994	Figure A5

Table 5. *Cont.*

ID	Forest/Terrain Type	GNSS Position	Appendix A Figure
E	A hilly pine forest growing also young spruces	N = 6,785,700 E = 398,893	Figure A6
F	A thinned pine forest	N = 6,785,537 E = 398,805	Figure A7
G	A full-grown spruce forest	N = 6,785,617 E = 398,869	Figure A8
H	A pine-forest next to an overflown lake	N = 6,785,783 E = 399,513	Figure A9
I	Stony hill in a pine forest	N = 6,785,688 E = 399,538	Figure A10
J	A young thinned pine forest on a slope	N = 6,785,686 E = 399,364	Figure A11
K	A forest road going through a mixed species forest	N = 6,785,579 E = 399,550	Figure A12
L	Large glacially deposited rocks on a pine forest	N = 6,785,523 E = 399,537	Figure A13
M	Opening in a dense young spruce forest	N = 6,785,548 E = 399,603	Figure A14
N	An old abandoned field surrounded by mixed-species forests	N = 6,785,416 E = 399,275	Figure A15
O	A paved road running between overflowing lakes with dead trees	N = 6,785,774 E = 399,163	Figure A16
P	A river bank with a forest road and mixed species forest	N = 6,785,883 E = 398,913	Figure A17

The georeferenced panorama image was imported into the game engine and used as a texture for a sphere with inverted normals and rendered with an additional virtual camera copying the transformations of the main camera. As a result, images with identical camera properties (horizontal opening angle of 81.8 degrees, vertical 60 degrees) and orientation were produced both from the virtual environment and the real world. All test image pairs (1920 by 1080 pixels) are included in Appendix A. After completing the image pairs, the following parameters were estimated manually by a human operator for each image pair (Table 6). This was done to obtain some comparable parameters to estimate the verisimilitude of the virtual environment and the real test-site. Thus, the experiment also takes into account e.g., the suitability of the automatically produced tree models for visualizing the real forest environment. The trees chosen for the comparison are shown marked in the images of Appendix A.

Table 6. Comparison parameters.

Parameter	Unit	Description
Amount of matching trees	Pcs	Amount of trees with similar locations found from both virtual environment and real location
Trees with correct species	Pcs	Amount of trees with similar locations and correct species found from both virtual environment and real location
Amount of missing trees	Pcs	Amount of trees found from real location, but not from the virtual environment
Amount of additional trees	Pcs	Amount of trees found from the virtual environment, but not from real location
Horizontal error	Pixels	For matching trees, the average horizontal shift of stem base on image
Stem width error	percents	For matching trees, the average relative difference of the visualized stem widths, measured from images produced, from estimated breast height (mean((ref_d-virtual_d)/ref_d))
Vertical error	pixels	For matching trees, the average vertical shift of stem base on image

4. Results

4.1. Computation Time

From the selected test site, a total of 3896 LAS files were processed, and trees could be found from 3799 of the files. In total, 101,052,210 individual trees were detected from the ALS data in 4 days and 39 min using our raster-based method. This results in an average computation time of 1 min and 29 s per LAS file. On average, approximately 17,400 individual trees, or equivalently 0.57 km^2 of ALS data, could thus be processed per minute. In contrast, the computation time was 7.5 times higher on average with a voxel-based approach to process 1 km^2 of ALS data.

4.2. Tree and Species Detection Performance

The found tree counts using both raster- and voxel-based methods, denoted by 2D, and 3D, respectively, are reported in Figure 6, along with the tree counts from the reference data. Detection rates (completeness) for log-size trees (diameter 20 cm or wider) were 92% and 87% for methods 2D, 3D, respectively. The corresponding detection rates for trees having a diameter of 30 cm or wider were 97% and 96%, respectively. Corresponding detection rates for trees in diameter class 10–15 cm, 15–20 cm, and 30–35 cm were 25%, 60%, and 97% for raster-based method, and 35%, 61%, and 95% for the voxel-based method. The voxel-based method increased the detection rate of smaller trees, but Log-size trees were obtained with a slightly higher rate with a raster-based method. Thus, we continued with a raster-based method since our major aim was to derive the log-size trees as accurately as possible. The corresponding completeness (i.e., recall), correctness (i.e., precision) and F1-score value for all trees were 48.0%, 77.3%, and 59.2% for the raster-based method and 50.0%, 80.9%, and 61.8% for the voxel-based method. 85–86% of the total volume and biomass was found with these two approaches. The results of the tree species classification are reported in the confusion matrix (Table 7). Overall tree species classification accuracy for matched trees was 75.1%.

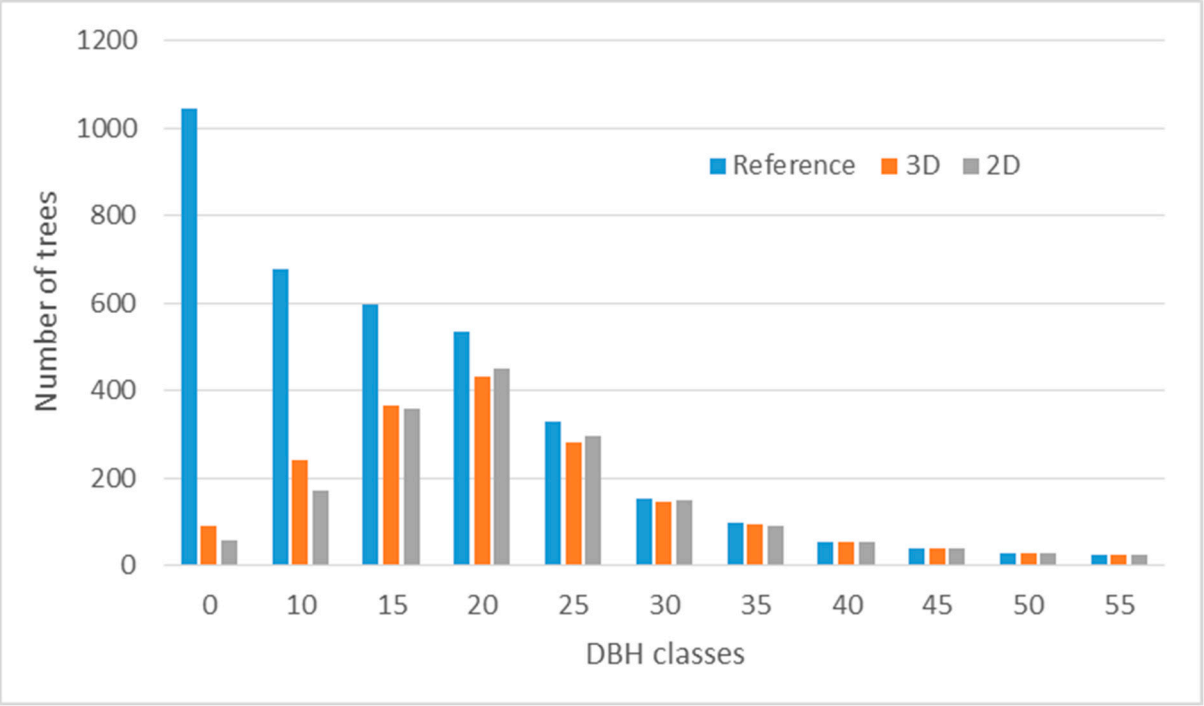


Figure 6. Found tree counts for both 2D and 3D methods as a function of DBH classes, compared to the reference.

Table 7. Confusion matrix of tree species classification using both point cloud and intensity features.

		Predicted Tree Species Class			Recall (%)
		Pine	**Spruce**	**Birch**	
Actual tree species class	Pine	437	49	75	77.9
	Spruce	33	492	92	79.7
	Birch	93	86	360	66.8
Precision (%)		77.6	78.5	68.3	overall = 75.1

Overall relative RMSE of 4.48%, 16.88%, 30.24%, and 29.01% were obtained for height, diameter, volume, and biomass (Table 8). Relative bias (%) and relative RMSE (%) for each tree diameter class, starting from trees having DBH bigger than 20 cm, were presented in Figure 7. Random forest prediction overestimated small trees (diameter, volume) whereas large trees were similarly underestimated. An overall estimation of relative bias (%) and relative RMSE (%) for different tree species is also given in Figure 8. Estimation of the volume of log-size spruces had the lowest errors, at the level of 25%. Estimation of the diameter of deciduous trees, on the other hand, had the largest.

Table 8. Bias, random error, and RMSE for tree attribute estimation using random forest models.

Tree Attribute	Bias	Relative Bias	Random Error	Relative Random Error	RMSE	Relative RMSE
H	−0.06 m	−0.25%	1.01 m	4.47%	1.01 m	4.48%
DBH	0.12 cm	0.49%	4.08 cm	16.87%	4.08 cm	16.88%
V	0.01 m^3	2.27%	0.18 m^3	30.19%	0.18 m^3	30.24%
AGB	3.79 kg	1.30%	84.35 kg	28.98%	84.44 kg	29.01%

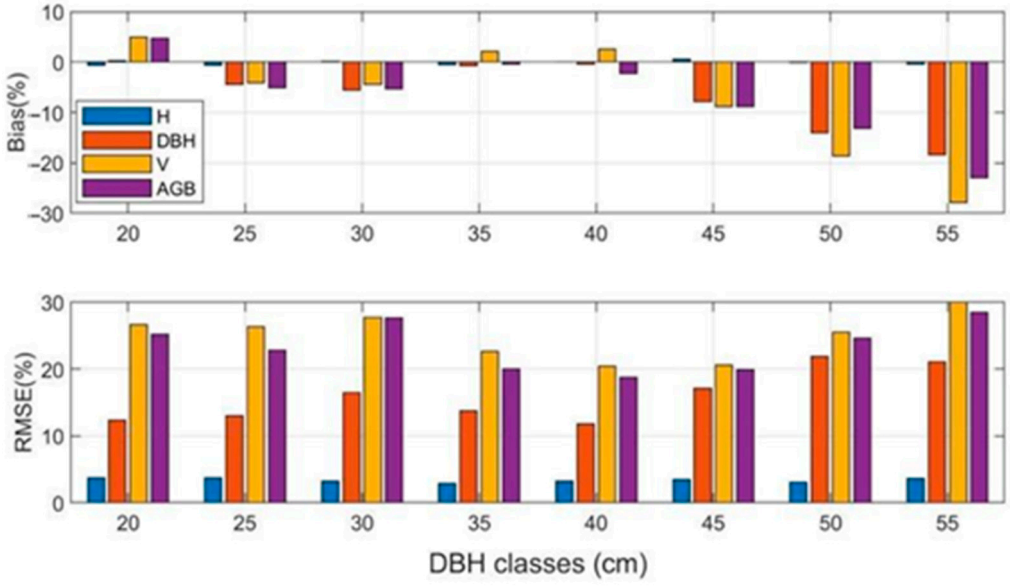


Figure 7. Relative bias (%) and relative RMSE (%) of attribute estimation as a function of DBH classes for trees over 20 cm in DBH, given for height (H), diameter (DBH), volume (V) and above ground biomass (AGB).

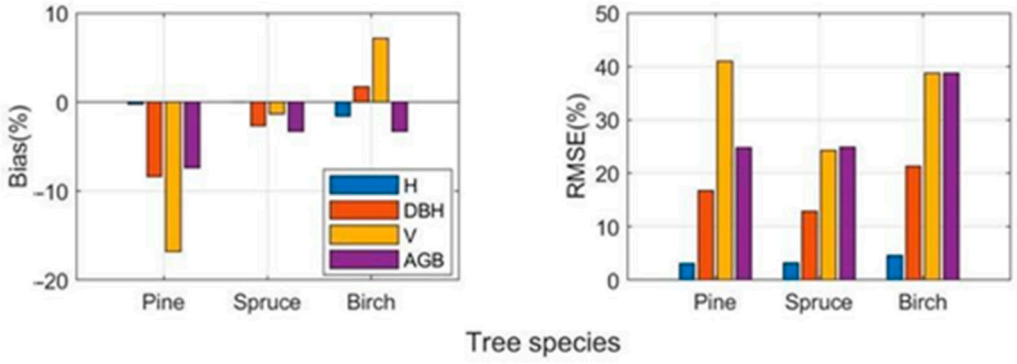


Figure 8. Relative bias (%) and relative RMSE (%) of attribute estimation as a function of tree species for trees over 20 cm in DBH, given for height (H), diameter (DBH), volume (V), and above ground biomass (AGB).

4.3. 2D Online Viewer for Single-Tree Forest Inventories

The developed browser-based 2D viewer enables the user to study the single-tree forest inventory on an online map (Figure 9). All detected trees with an estimated height of more than 5 m are visualized on the map when viewed on the highest zoom levels. Each tree is denoted with a circle marker, with larger and darker markers used for taller trees. Different tree species are denoted by colors. Thus, the regional distributions of different heights and species can be observed very clearly.

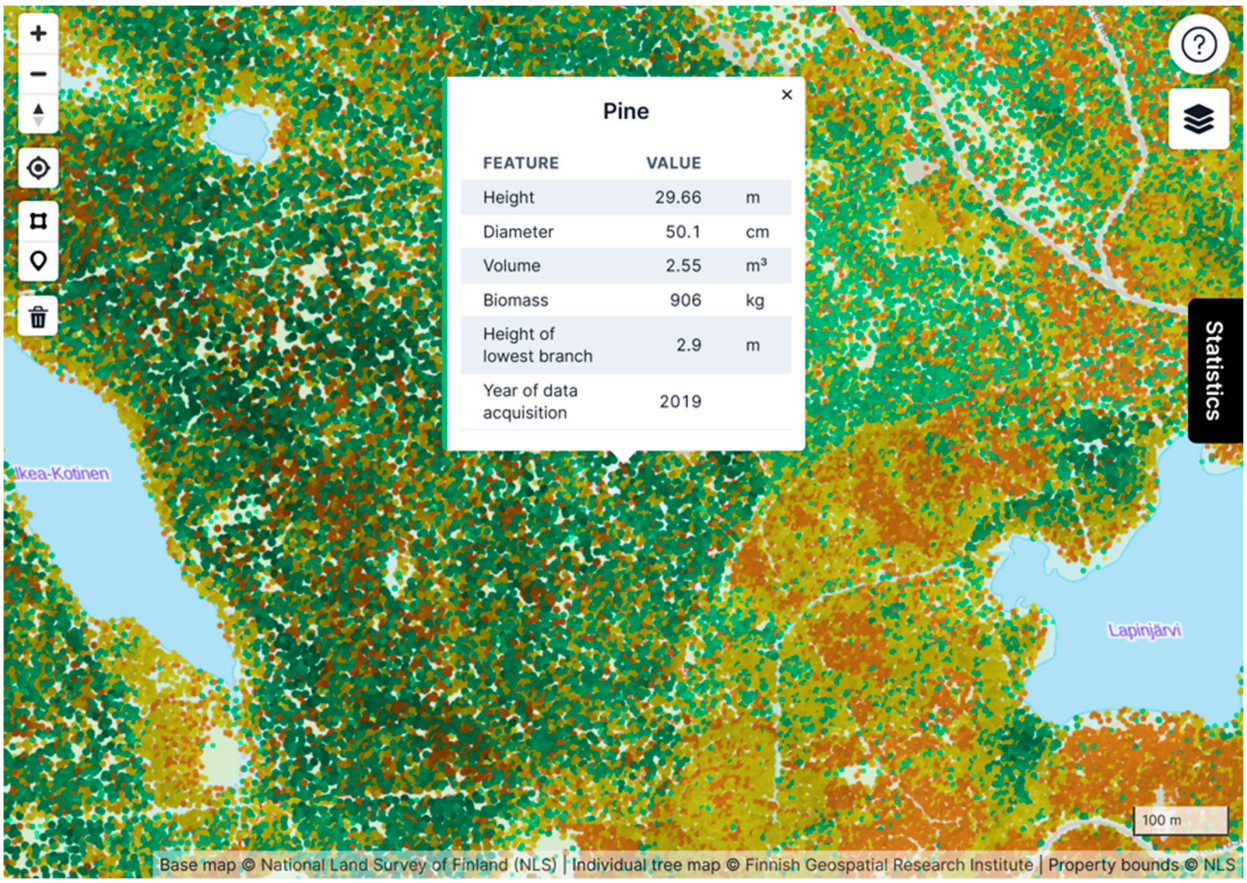


Figure 9. Individual trees drawn on the map with the following color codes: pine = brown, spruce = green and deciduous tree = yellow. The taller the tree, the darker the color of the circle marker. By clicking the tree, the most important characteristics of the tree, such as the species, height, diameter, volume, and biomass, are given in the info box.

Users can also query statistical descriptions of the tree population from an arbitrary polygonal area (Figure 10). The statistical data are visualized with a dashboard which contains a number of statistics calculated from the trees located in the query polygon (Figure 11). On the left in Figure 11, descriptive statistics from key attributes, including mean diameter, total volume, biomass, basal area, and the number of stems per hectare, have been visualized in a table. The statistics have been calculated for all trees and also separately for pines, spruces, and deciduous trees. Furthermore, the dashboard gives an automatically calculated estimation of the value of the forest, which is calculated based on the volumes of log wood, pulp, and energy wood. The 2D viewer also estimates the amount of carbon dioxide the trees have absorbed during their lifespan and presents height and diameter distributions using histograms (on the right in Figure 11).

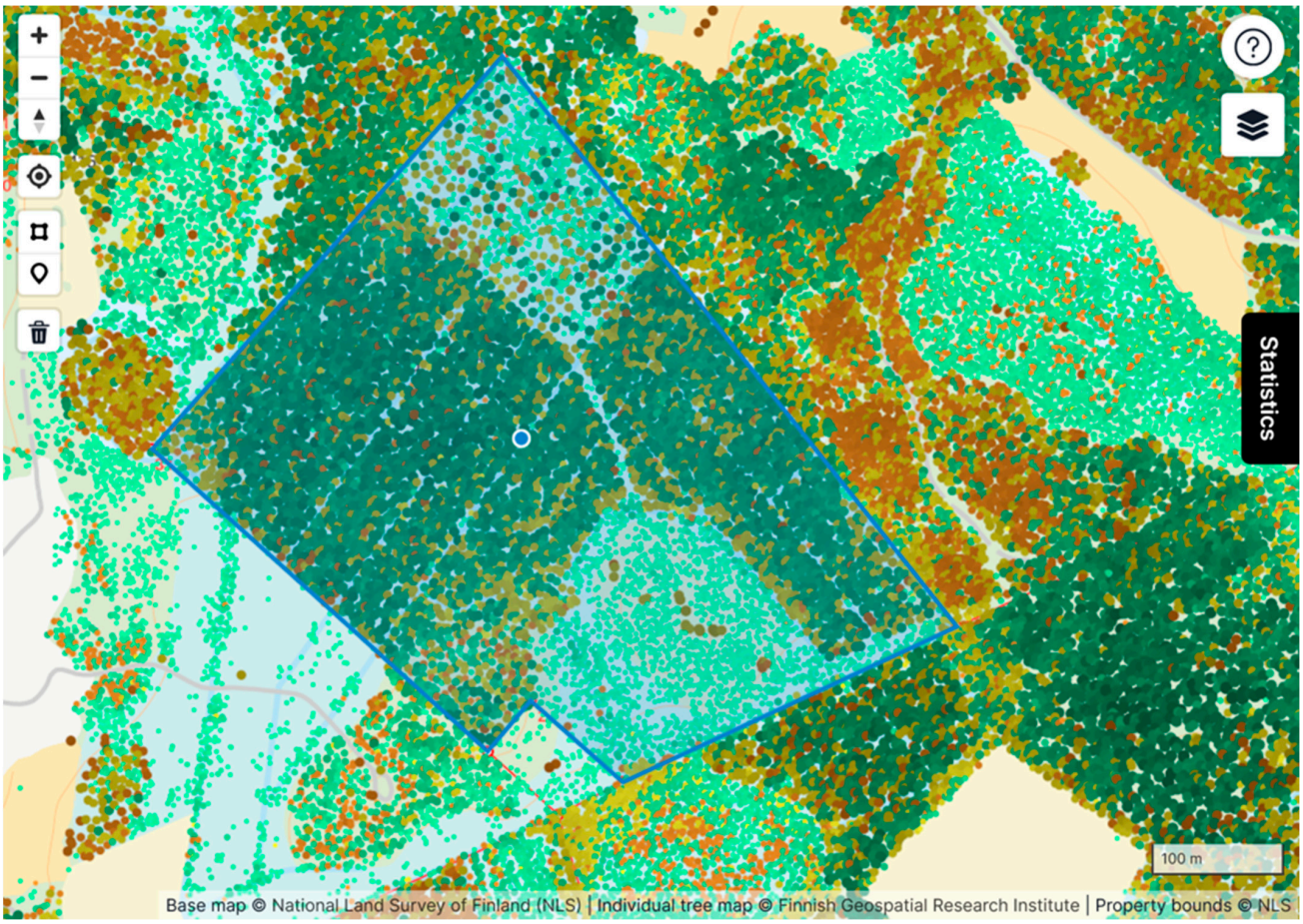


Figure 10. The user has selected a property on the map to query statistical data of the trees belonging to the property. The property has been selected by placing a blue marker on the map after which the boundaries of the selected property have become highlighted with blue color. The topographic map underneath the trees was obtained from the API of the National Land Survey of Finland. Individual trees drawn on the map with the following color codes: pine = brown, spruce = green and deciduous tree = yellow. The taller the tree, the darker the color of the circle marker.

In addition to the tree data, a topographic map is visualized in the background. Property boundaries (obtained from the API of the NLS) and forest stand boundaries (from the API of the Forest Center) can also be visualized in the viewer to offer additional context.

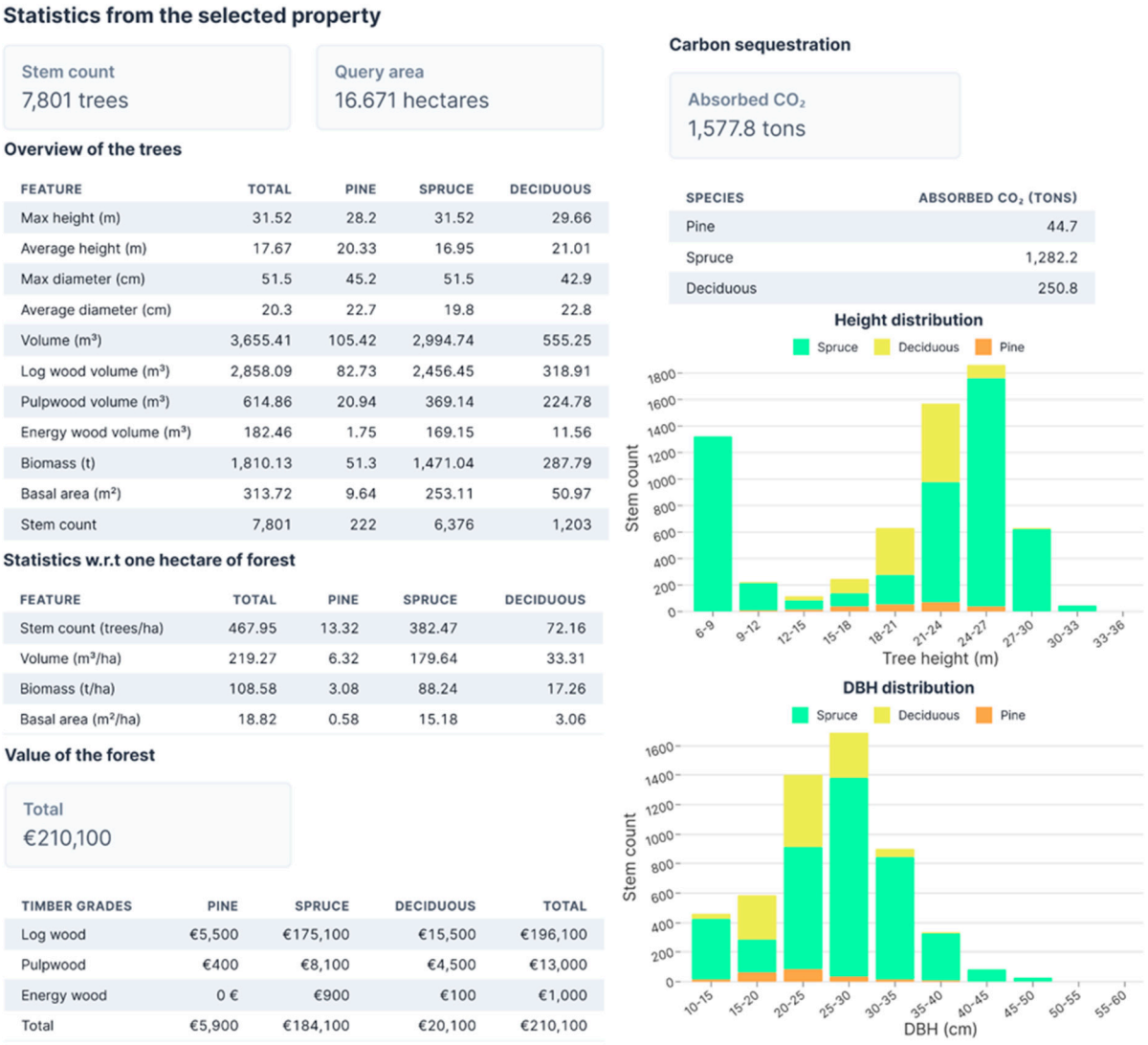


Figure 11. Dashboard visualizing the statistics of the trees from the selected area. Height and DBH histograms of the trees in the selected area are displayed on the right in addition to an estimation of the amount of carbon dioxide the trees have absorbed during their lifespan. Descriptive statistics for key attributes, such as mean diameter, total volume, biomass, basal area and the number of stems per hectare, have been presented on the left with tables in addition to an estimation of the value of the forest.

4.4. Game Engine Mediated Virtual Forest

Combining the ITD data with open GIS datasets, an interactive virtual scene, a virtual forest, was established on a commercial game engine. The scene is automatically generated from the given datasets. The HTC Vive VR HMD was used to study the scene.

Users can interact with individual trees or tree groups determined by filters based on the provided tree data. Trees can be filtered, e.g., by species or height (Figures 12 and 13). The HTC Vive controller is used to point and select trees via raycasting. Once selected, tree information is displayed on a floating pane next to the selected trees that are sufficiently close to the camera. When displaying the tree information, the UI guides attached to the controller support the user in defining different filters and setting their values. The chosen trees can be then removed or marked for future procedures. For moving in the virtual environment, the controller is used to point and click to "walk" to the desired location. The user can also move short distances physically, as supported by the headset's tracking system.



Figure 12. User querying tree properties while filtering the virtual environment to show only trees less than 15 m tall.



Figure 13. User querying tree properties while filtering the virtual environment to hide birches.

4.5. Visual Comparison of Virtual Forest and On-Site Reference

When comparing the virtual forest visually against on-site images from the same locations (see Appendix A), the estimated parameters (see Table 5) are provided for each image pair in Table 9. In short, species classification accuracy was 71%, and there were 16% omissions and 6% commissions in the tree detection.

Table 9. Results from comparing on-site images and those obtained from the game engine-mediated virtual forest scene visualization, applying the metrics defined in Section 3.7.2.

Parameter	Site ID															
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**
Matching trees (pcs)	9	10	16	13	5	13	15	14	12	13	13	13	4	8	12	10
Correct species (pcs)	5	6	10	10	4	12	10	12	11	12	7	11	4	4	7	3
Correct species of matching trees (%)	56	60	63	79	80	92	67	86	92	92	54	85	100	50	58	30
Missing trees, omissions (pcs)	0	2	0	1	3	3	1	2	2	2	0	2	6	1	8	2
Additional trees, commissions (pcs)	3	2	0	1	1	0	0	1	1	0	1	1	0	0	1	1
Horizontal error (px)	88	17	2	28	118	16	76	46	11	25	28	49	90	20	33	115
Stem width error (%)	−15	−24	1	34	4	−6	19	−5	11	−17	−6	−2	7	−12	0	−8
Vertical error (px)	82	81	26	111	35	34	58	70	10	71	22	10	32	32	7	31

The results of comparing the game engine-mediated visualization to on-site images reflect the quality of datasets used in generating the virtual scene. The highest numbers of missing trees were found from sites with an abundance of smaller trees that generally are not found by the ITD algorithm. On site M, the number of small trees that could be pinned for evaluation was limited due to the extremely dense young forest. In terms of missing trees, the best results were obtained from sparse, homogenous, grown forests (example in Figure 14). These also produced the best results in species detection correctness, with the worst results coming from mixed forests. It should be noted that as the species correctness was estimated for matching trees only, it mostly reflects the correctness of larger trees (that are found by the ITD algorithm). The characteristics of ITD-based tree detection (e.g., the assumption of vertically growing trees) are also reflected in horizontal error, which was worst on hilly sites and riverbanks: As there are several leaning trees, the canopy derived positions deviate from the stem positions, thus increasing the error (example in Figure 15). The vertical error also contains the inaccuracy of the DTM used for terrain. Here, the differences in undergrowth vegetation complicate the evaluation, as dense vegetation may hide the lowest parts of the stem.



Figure 14. Test site F on site (**left**) and as virtual scene (**right**). Correctly represented dominant trees almost begin to function as landmarks.



Figure 15. Test site P on site (**left**) and as virtual scene (**right**). Absence of smaller trees in the virtual scene, species errors, and position errors of leaning trees lead to poor visual correspondence between the real and virtual forest. The waterbody and road however act as potential landmarks.

5. Discussion

Our aim was to produce, as a prototype, the entire pipeline for country-level individual tree data, ranging from data production to user interfaces applying the produced "individual tree maps". For this, we implemented the parallelized ITD with a high degree of automation, DB for information storage, APIs, and two example UIs (browser-based map, and the "virtual forest" for VR HMD). While individual parts of the process have been demonstrated in prior research, the authors have not been able to find a similar system implementation. In the following, the limitations and potential future development of the presented work are discussed.

5.1. Limitations and Future Development for National Single-Tree Data

National laser scanning programs provide a potential starting point for national single-tree maps: they are sufficiently dense for the detection of individual trees and provide coverage. This could resolve the long-standing data availability issue in applying individual tree data [7]. The experiment of producing an individual tree forest inventory from ALS clearly demonstrated that in terms of computation time and resources, there are no major obstacles to implementing nationwide systems. With an eight-core workstation, the attained computing speed of 0.57 km^2/min would imply that with 10 such workstations (or virtual machines) the ITD could be performed for the entire Finland (surface area of 337,030 km^2 excluding 10% of water cover) in roughly 41 days, assuming that the point clouds are available and prepared.

In the presented work, ITD was produced once, from a single set of non-overlapping ALS point clouds. While this is sufficient as a first step, further development is needed to allow updating ITD data, by relying on the new ALS point cloud, regional mapping e.g., by drone, or terrestrial mapping. To support this, the temporal dimension should be included in ITD data, at minimum by including the time stamp for each tree. For end users, knowledge of growth, tree damage due to storm, snow, or insects, are needed. There are challenges in measuring the height growth of individual trees, as matching of trees from multitemporal surveys (airborne and field) is prone to errors [74]. If acquisition season changes from leaf-on to leaf-off or to vice-versa, the used algorithms are also affected. Change information, especially for harvested and fallen trees, could also be presented with raster layers.

The ITD data produced were also limited in quality. Firstly, the applied algorithm is unable to detect smaller trees that are obscured by the dominant ones. Detection of smaller trees with more dense point clouds is feasible when using the latest point cloud processing techniques, but it is challenging to find a higher number of smaller trees with point densities between 1 and 10 pts/m^2 using raster-based techniques. In this case, deep learning was not applied for ITD, as deep learning methods typically work well for data where human interpretation can be used as an annotated reference. However, at least prior

studies, namely by Vastaranta et al. [13] and Kaartinen et al. [14], found that visual, manual annotation using 8–10 pts/m^2 produces inferior results compared to automatic techniques. As even lower point densities (5 pts/m^2) are utilized in our study, manual annotations and deep learning would likely fail to improve the results. For higher point densities, possibly even with 20 pts/m^2, we expect deep learning [75] to be useful. The possibility of calculating different rasters based on the penetration of hits proposed by Hyyppä et al. [26] is an option. In [26], the number of small trees could be increased by 10%-units, which can actually double the number of small trees found. Following this approach, a different type of representation might be more suited for dense, small trees: they could be represented by a set of raster layers describing their occurrence and density and height, instead of aiming to detect and model individual stems.

For the larger trees (DBH over 20 cm), the estimated detection rate was over 90%. However, this corresponds to the commercial value of 95% of all trees in the stands. Thus, the resulting dataset is skewed towards the larger trees. Secondly, the species detection is limited, especially for deciduous trees. Finally, there are apparent false detections present in the ITD data, originating from man-made features such as utility poles and power lines, falsely identified as large trees. Some of them could likely be removed relying on existing GIS data.

These shortcomings limit the usability of resulting single-tree data. Due to the lack of reliable species detection for deciduous trees, the applicability of resulting ITD "maps" for biodiversity applications requires classification capacity to detect aspens and dead trees, since they represent the key biodiversity of boreal forests. For biodiversity applications detection of aspens and dead trees would be highly beneficial. These could be detected using intensity data and improved point density. In [76], laying deadwood was identified with a 52% detection rate and standing trees with a 71% detection rate using a point density of about 25 transmitted pulses per m^2. It is likely that the third national laser scanning programme in Finland, starting in 2026, will produce similar point densities. Thus, the ITD systems could also become a platform for maintaining "deadwood maps", with the potential to support biodiversity studies. The lack of small trees in the data hinders estimating forest regeneration. On the other hand, some forest companies are already planning to increasingly utilize dense remote sensing data in operational forest management (see e.g., [77]).

When comparing the results with recent close-range under-canopy drone studies, the obtained errors for tree attributes are relatively low. In [78], the under-canopy drone collecting point density of around 100,000 pts/m^2 could map DBH with a 5.18 cm (22.49%) RMSE and 4.16 cm (18.31%) bias. In another study [79] in the same area (and using part of our plots) as in this study, above-canopy UAV with manual processing and Mobile Laser Scanning (MLS) were applied for easy, medium-difficult, and difficult plots. UAV data of 5000 points/m^2 and manual processing resulted in 17%, 29%, and 26% relative RMSE for DBH, for easy, medium-difficult, and difficult plots, respectively, with a relative bias of over 10% in all cases. Similar results obtained with MLS having point density around 10,000 pts/m^2 were 11%, 23%, and 32%, respectively. Most of the MLS techniques reviewed in [80] resulted in more than 10% relative RMSE for DBH and most of the results have been achieved in easy forests (here, the relative bias was not available for all studies). The comparable relative RMSE achieved by our methods was 17% for DBH, with a very low bias, and our study includes all kinds of forests (easy, medium-difficult, and difficult). Comparing our volume estimation errors, we have obtained an estimated relative RMSE of 30% at individual tree level and 25–28% for log-size trees, again with a low relative bias (2%). By manually measuring each tree's DBH on the field, the stem volume can be estimated with about 20% error using national allometric models [71]. Thus, our results appear to have reasonably low deviations from reference data, especially when taking the low-point cloud density into account.

The achieved relatively modest species overall classification accuracy is likely due to low-density point cloud used. Distinguishing species based on these data is challenging, as

demonstrated by Hakula et al. [39], who achieved 73% overall accuracy with high-density (1400 pts/m^2) laser scanning data and geometry-based features. Our approach, which also incorporates intensity features, yielded a 2% improvement despite significantly lower point density (5 pts/m^2).

However, new techniques exist that are feasible for providing high-quality references for ALS processing from above-canopy drones, under-canopy drones, hand-held scanners, and other mobile laser scanners. As our application requires a reference with errors of 2–10% for DBH, the reference production could be based on, e.g., using a matching algorithm correcting the drifts that are present in the point cloud after Simultaneous Localization and Mapping, presented in [36] for backpack laser scanning, in [27,81] for other under-canopy systems and in [33] for above-canopy drone data. All of the mentioned methods also utilized stem curve extraction and interpolation of DBH from a measured stem curve, achieving ~4–6% error in sparse, homogenous forests. A similar approach could also be used to obtain wood quality estimates from ALS/TLS-derived features, as demonstrated by Pyörälä et al. [20]. Integration of wood quality estimates with national-level ITD could yield significant economic benefits in forestry. Tree health can be better provided through change detection using multispectral ALS surveys and/or dense point clouds, but such data are not yet currently available in national laser scanning programmes. However, such information could be produced with increasing point density. But high-class processing of data is a requisite since we have obtained relatively low errors taking into account the properties of the data we are using.

The work has demonstrated that large-scale ITDs could already be feasible and relatively accurate relying on current national LS data. Early results of this study have already led to a pre-operational project in Finland, where individual tree inventory will be integrated into country-level area-based inventory by the Finnish Forestry Centre.

5.2. Limitations and Future Development for Online Services for Single-Tree Data

By developing an online service for studying the single-tree forest data, we have successfully demonstrated that by relying on existing components (such as PostgreSQL) and development frameworks (such as Django), it is possible to develop an operational system for storing and distributing single-tree data from very large geographical areas. This was demonstrated for 3896 km^2. In addition to the further developments that are directly related to data (discussed in Section 5.1), there are several potential development directions for the online service. Firstly, additional data sources could be brought to the UI, to better support use cases such as silvicultural planning. These could include data related to relevant infrastructure elements such as ditches, forest roads, turning and loading sites, and power lines. Secondly, by relying on such additional data layers, tools could be developed for actual forest planning. The online environment would thus become a platform for precision forestry. Thirdly, more analysis functionalities could be introduced, either by relying on ITD alone, or ITD supported by integration with other data sources. Biomass, timber value, and plot densities could be estimated mostly relying on the ITD. By adding complementary data sources, such as a DTM, and models of roads and other infrastructure, parameters like the ease of harvesting or transport distances could be estimated, even for individual trees. Estimation of such parameters has also been suggested in prior research (e.g., [10]).

An additional major development direction would be the support for user-generated data. The users could, e.g., improve the ITD data with manual measurements from particularly important sites, or update the data following forestry operations by marking removed trees. As the online viewer and its back-end services were developed as demonstrators for ITD data, this would likely require some reconsideration of its technical architecture and the introduction of additional concepts such as different levels of authorization. While these are interesting, they go beyond the scope of building a demonstrator for the produced data. The possibility of adding the GNSS position of the user to the viewer and adding crowdsourcing information for improving the reference data is also an attractive future

development direction. Manual measurements (diameters), visually recorded data (tree species, biodiversity), and data collected with personal sensors, such as smartphones, can be added in this way. Data curation can be done by simple statistical analysis

It is also possible to imagine applications other than forestry that might benefit from ITD data. For example, habitat management or recreational forest use such as hiking and berry picking could benefit from new, more detailed maps and new parameters of forest areas, such as the species profile.

5.3. Limitations and Future Development for Game Engine Mediated Single-Tree Visualization

Currently, the forest visualization in a game engine is typically realized with procedurally generated, game engine-compatible tree models that are positioned randomly according to plot characteristics [8,9]. Therefore these systems are not feasible for visualization or decision-making pertaining to individual trees. By connecting the game engine-mediated visualization to the ITD online service, we have demonstrated that a game engine visualization can also be efficiently realized with real tree data, while still maintaining the automated generation of the virtual scene. As open GIS assets and national laser scanning point clouds were used, this level of visualization could already be produced for large areas of Finland. For the future, the system provides the option to use either prefabricated "master tree"-models instead of simpler ones. Master tree models' tree parts are divided into separate objects all the way up to each individual branch, which can then be used for more precise data representation but with an increased memory and performance cost.

With current ITD data, the most truthful visualizations are produced from sparse, grown forests with clearly dominant tree species, low species variation, and clear landmarks that are found in the topographic data (such as water bodies, roads, or major boulders). In these cases, the virtual scenes could likely support identifying real-life locations and making decisions pertaining to individual major trees, for example in precision forestry [10]. On sites that have very few landmarks, the tree data alone are sufficient for identifying locations perhaps in the very best cases only (Figure 14). On sites with high species variability, complex terrain, or dense groups of small trees, the virtual scenes are not identifiable, even if the major trees are still represented. Confirming the identifiability of virtual forest scenes would require dedicated user tests with, e.g., navigation tasks [82]. As virtual forests are being suggested for a variety of applications [83] and the user test results may be affected by the professional background of the participants (as found in [82]), some user profiling would also be required for the experiments.

To progress from a demonstrator to an actual tool, several development tasks still remain: While the tree data are already retrieved over the developed API from an online storage, the other data assets used (DEM, aerial imagery, forest inventory data, etc.) are applied as pre-processed files. Thus, the virtual environment does not scale beyond the 1 km^2 boundary used in the demonstrator. As suitable APIs exist for accessing geospatial raster data (e.g., WMS, WMTS, or WCS) and geospatial features (e.g., WFS; OGC API Features), we feel confident that creating a background system for provisioning these data assets over an API, and its utilization in the game engine could be implemented.

The most obvious improvement of the developed virtual forest would be the inclusion of small trees in visualization. This could be realized by a set of raster layers derived from lidar point clouds, describing the occurrence, density, and height of smaller trees, based on which a procedural visualization could be accomplished, placing the smaller trees at random following their properties. A second major track for potential development relates to building the tools needed for use cases of ITD data, such as forest planning, education, or electronic tree trade. While the implementation of these is beyond the scope of building a demonstrator, such steps would still be needed to turn game engine-mediated virtual forest into a useful tool.

6. Conclusions

The potential of using forest inventories based on individual tree data is well known. However, the difficulty of producing precise and efficient ITD data has limited the practical applications. The lack of large-scale ITD data has also hindered the application of "virtual forests", i.e., interactive 3D visualizations that support decision-making in landscape planning and forestry management. Our study addresses these limitations by developing a system for producing ITD data on a large scale and proposing software infrastructure for storing, distributing, and viewing the ITD data. Key achievements include:

- National scale ITD data production: Leveraging the Finnish national laser scanning point cloud (5 points/m^2), we successfully detected over 100 million trees from an area of 3896 km^2 in less than 5 days of computing time on a single workstation. This demonstrates the feasibility of our system for national ITD production.
- High precision of key metrics: Our system achieved a 92% detection rate for log-size trees. Height, diameter, volume, and biomass were extracted with RMSE tolerable in operational forestry. Notably, the obtained RMSE are comparable to or even lower than some of the recent close-range remote sensing studies for reference measurements, and not far from field references based on using DBH and allometric models. Further, our estimates have very little bias.
- Software infrastructure for ITD data: We designed and implemented software infrastructure for storing, distributing, and online visualization of large quantities of ITD data. This demonstrates the feasibility of distributing and applying ITD data for entire countries. By also coupling the game engine-based visualization to the same system, we highlight the potential of building different types of visualizations and user interfaces for the ITD data.
- Automatic generation of "virtual forests": By combining the ITD data with open GIS datasets, a game engine-based immersive visualization was produced with a high degree of automation. From the 1 km^2 virtual forest test site, the visualization results were compared with georeferenced panoramic images. The virtual forest scenes were found to correspond best with the real-life sites in fairly sparse, grown forests having a clear dominant tree species and clear landmarks, such as rivers, roads, or fields. In these cases, it was possible to produce identifiable virtual scenes automatically. Game engine visualization could thus offer a starting point for decision-making and landscape visualization in precision forestry.
- Real-World Impact: Our work has already led to the decision by the Finnish Forestry Center to include ITD data in operational stand-level Finnish forest inventory, starting 2026, relying on national laser scanning data.

The quality of the individual tree data is currently limited by the species classification accuracy and the poor detection rate for smaller trees, especially diameter classes 0–10 cm, from the 5 pts/m^2 point cloud. These limitations affect especially the virtual forests, restricting them mainly to log-size trees. The inclusion of smaller trees would likely require introducing additional data sources, such as more dense ALS point clouds or imagery. Turning the current online viewers or game engine-mediated virtual forests into real tools would also require further development and collaboration with potential users.

Author Contributions: This work has two first authors, Matti Hyyppä and Tuomas Turppa. Matti Hyyppä developed the ITD data production and distribution system, and the browser-based viewer, and processed the ITD data. Tuomas Turppa developed the game engine mediated visualization. Heikki Hyyti performed the on-site comparison data acquisition and comparison. Xiaowei Yu developed the tree extraction algorithm and species detection, and analyzed the ITD data. Hannu Handolin supported the game engine visualization development. Juho-Pekka Virtanen drafted the manuscript as senior author and supported ITD production pipeline and browser-based viewer development. Juha Hyyppä and Antero Kukko helped finalize the manuscript, with contributions from all authors. Juha Hyyppä arranged the funding and was PI for the project. All authors have read and agreed to the published version of the manuscript.

Funding: We gratefully acknowledge for financial support the Ministry of Agriculture and Forestry project "Future forest information system at individual tree level" (VN/3482/2021), and the following Research Council of Finland projects "Forest-Human-Machine Interplay" (decision 357908), "Capturing structural and functional diversity of trees and tree communities for supporting sustainable use of forests" (348644), "High-performance computing allowing high-accuracy country-level individual tree carbon sink and biodiversity mapping" (359203), Collecting Accurate Individual Tree Information for Harvester Operation Decision Making (359554) and "Mapping of forest health, species and forest fire risks using Novel ICT Data and Approaches" (344755, Chistera Eranet).

Data Availability Statement: The open geospatial datasets applied in the research are available from their original providers, as referenced in the materials and methods section. The ALS datasets applied for ITD are subject to separate permissions for use, due to security reasons. Inquiries related to these datasets should be directed to the National Land Survey of Finland. The full ITD datasets presented in this article are not readily available because they cover extensive areas of privately owned forests, and thus contain information related to private property on a high level of detail. Inquiries related to the datasets should be directed to the corresponding author. A 9 km^2 sample dataset of ITD data is released as open data from the Evo area, Finland. This covers the 1 km^2 area used as the test site for virtual forest production (Figure 2) and the eight surrounding 1 km^2 tiles. The dataset is available at: https://doi.org/10.5281/zenodo.14203015.

Acknowledgments: Research Council of Finland Research infrastructure (346382) was applied for the study.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Comparison Image Pairs

In all images in this Appendix, the following marking (Figure A1) is applied for comparing images from the virtual scene (Figures A2–A17) (upper) with the on-site pictures (lower), marking the trees used in the analysis by indicating the type of match (correct species, incorrect species, or no match found) and marking the point for stem width comparison measurement.

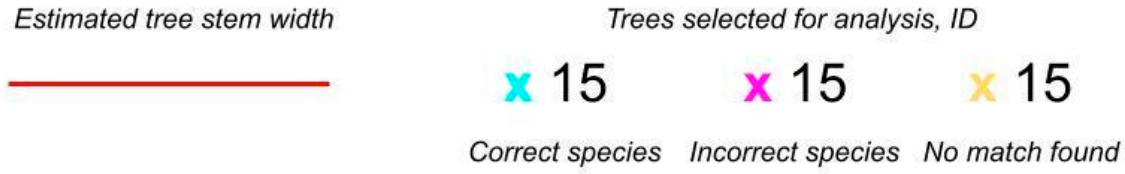


Figure A1. Legend used in all comparison image pairs in this Appendix.



Figure A2. Test site A, a mixed-species forest by the river flowing into a lake, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,786,021 E = 398,869.



Figure A3. Test site B, an uneven-aged mixed-species forest around a ditch and a forest road, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,854 E = 398,919.



Figure A4. Test site C, an open field in the middle of a mixed forest, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,778 E = 399,079.



Figure A5. Test site D, a hilly full-grown spruce forest, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,759 E = 398,994.

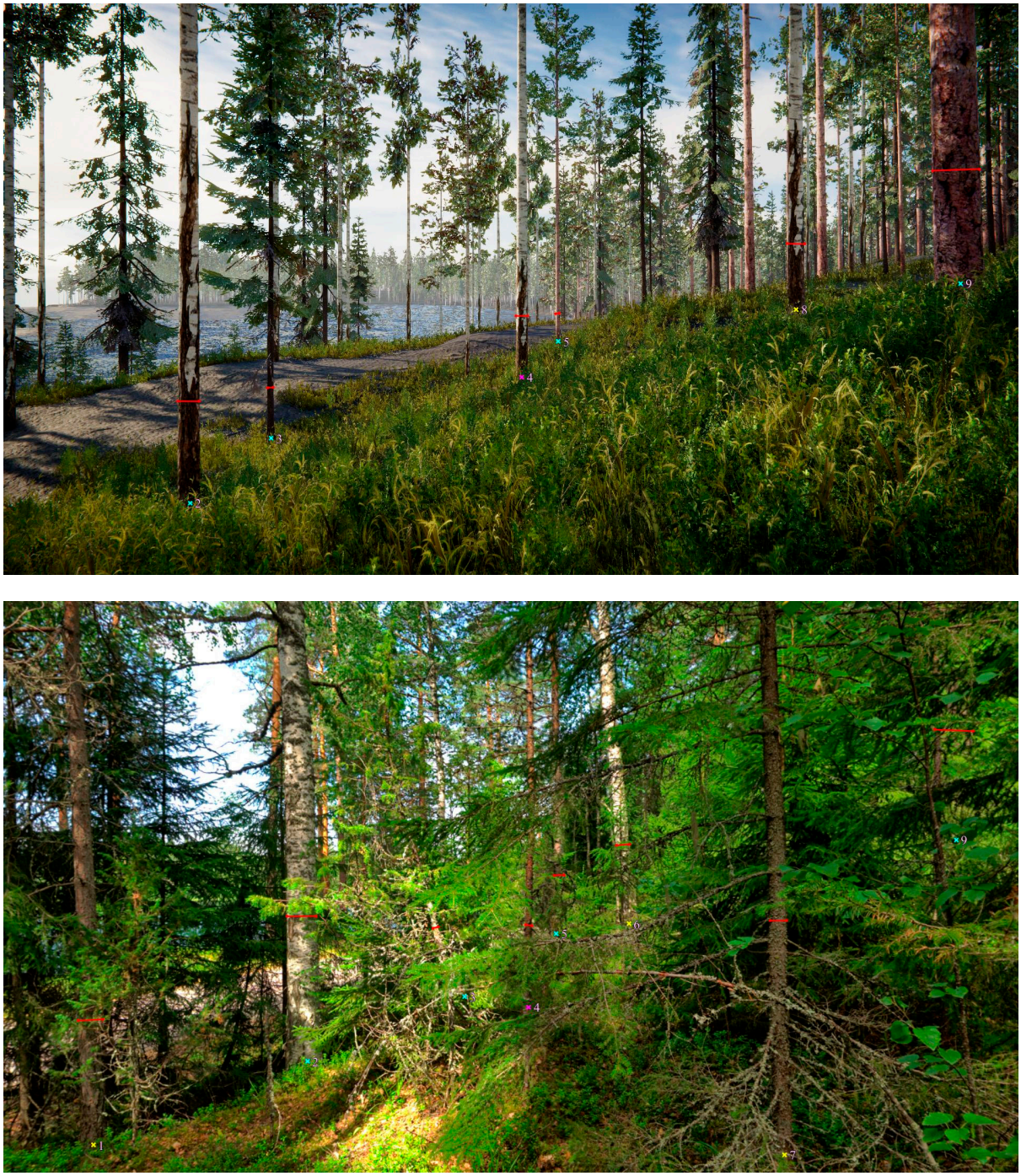


Figure A6. Test site E, a hilly pine forest growing also young spruces, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,700 E = 398,893.



Figure A7. Test site F, a thinned pine forest, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,537 E = 398,805.



Figure A8. Test site G, a full-grown spruce forest, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,617 E = 398,869.

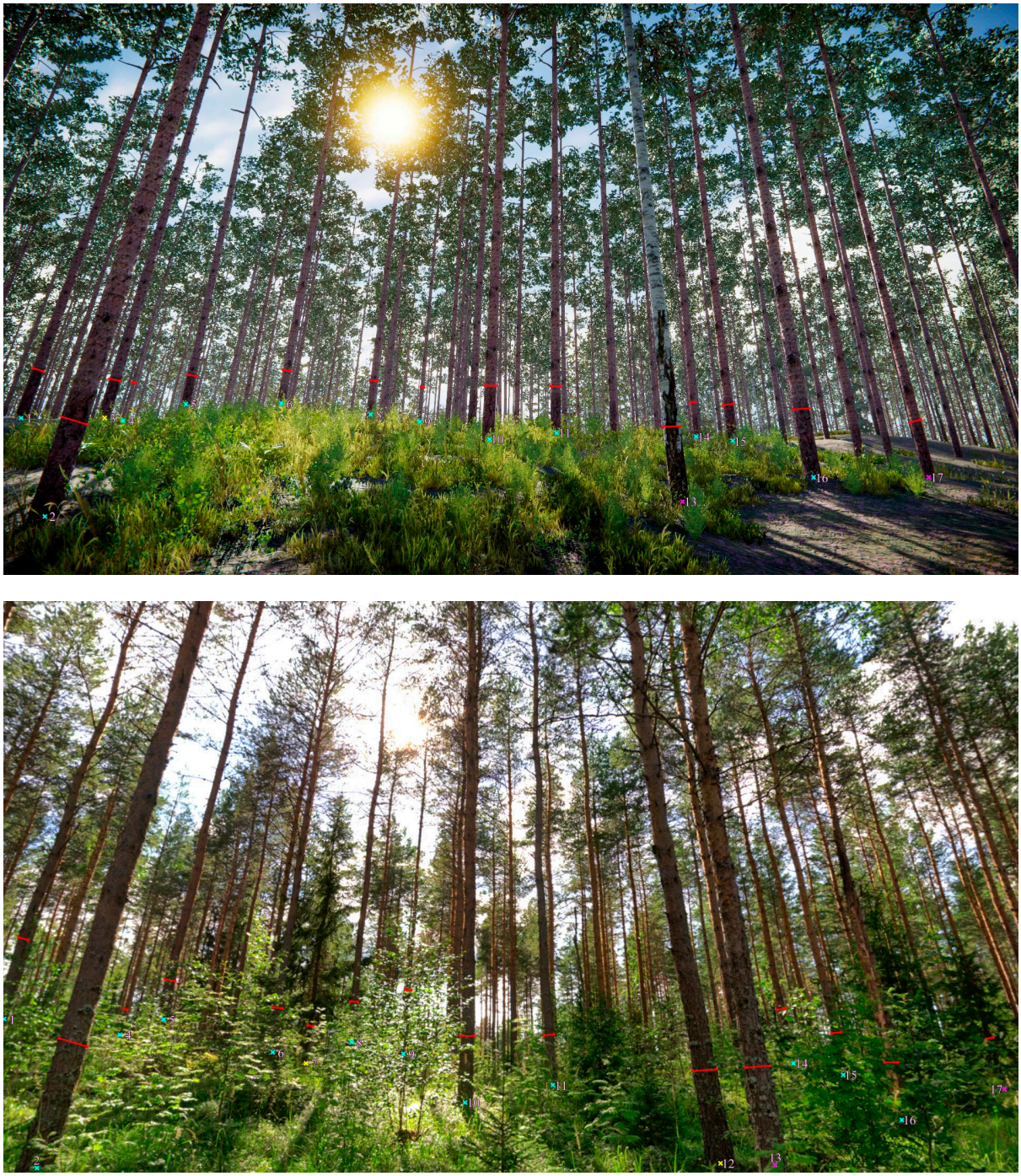


Figure A9. Test site H, a pine-forest next to an overflown lake, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,783 E = 399,513.



Figure A10. Test site I, a stony hill in a pine forest, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,688 E = 399,538.



Figure A11. Test site J, a young thinned spruce forest on a slope, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,686 E = 399,364.



Figure A12. Test site K, a forest road going through a mixed species forest, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,579 E = 399,550.

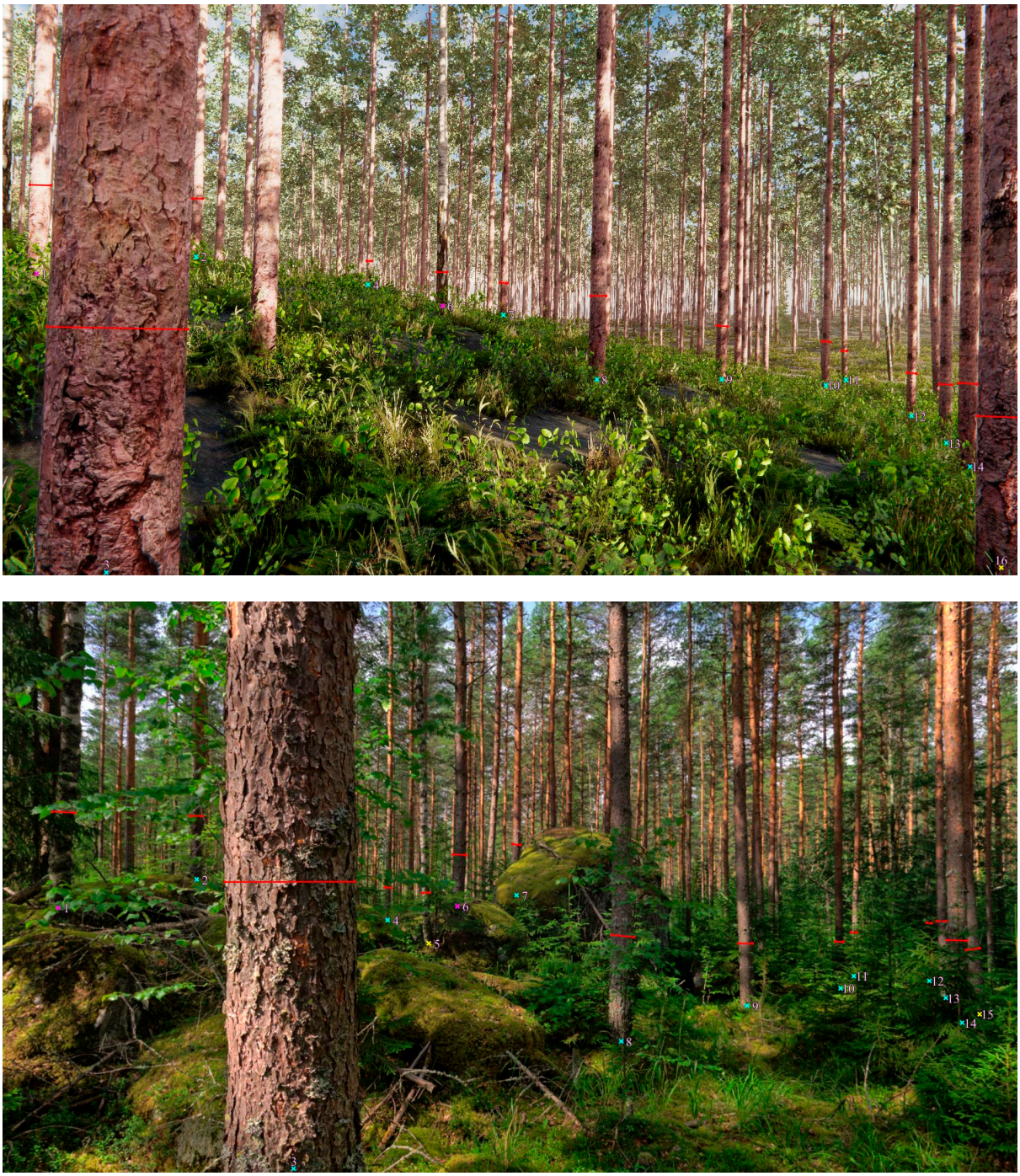


Figure A13. Test site L, large glacially deposited rocks on a pine forest, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,523 E = 399,537.



Figure A14. Test site M, an opening in a dense young spruce forest, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,548 E = 399,603.



Figure A15. Test site N, an old abandoned field surrounded by mixed-species forests, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,416 E = 399,275.



Figure A16. Test site O, a paved road running between overflowing lakes with dead trees, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,774 E = 399,163.



Figure A17. Test site P, a river bank with a forest road and mixed species forest, showing the virtual scene (**top**) and on-site image (**bottom**). GNSS Position (TM-GK35FIN): N = 6,785,883 E = 398,913.

References

1. EEA European Environment Agency, European Forest Ecosystems—Key Allies in Sustainable Development, Publications Office of the European Union. 2023. Available online: https://data.europa.eu/doi/10.2800/572787 (accessed on 8 November 2024).
2. FAO. *The State of the World's Forests 2024—Forest-Sector Innovations Towards a More Sustainable Future*; FAO: Rome, Italy, 2024. [CrossRef]
3. Stokland, J.N.; Tomter, S.M.; Söderberg, U. Development of dead wood indicators for biodiversity monitoring: Experiences from Scandinavia. *Monit. Indic. For. Biodivers. Eur.—Ideas Oper.* **2004**, *51*, 207–226.
4. Goheen, D.J.; Hansen, E.M. Effects of pathogens and bark beetles on forests. In *Beetle-Pathogen Interactions in Conifer Forests*; Schowalter, T.D., Filip, G.M., Eds.; Academic Press: San Diego, CA, USA, 1993; pp. 175–196.

5. Kangas, A.; Haara, A.; Holopainen, M.; Luoma, V.; Packalen, P.; Packalen, T.; Ruotsalainen, R.; Saarinen, N. Kaukokartoitukseen Perustuvan Metsävaratiedon Hyötyanalyysi: Metku Project Final Report. 2019. Available online: http://urn.fi/URN:ISBN:978-952-326-707-7 (accessed on 8 November 2024).
6. Uusitalo, J.; Orland, B. Virtual Forest Management: Possibilities and Challenges. *Int. J. For. Eng.* **2001**, *12*, 57–66. [CrossRef]
7. Orland, B.; Radja, P.; Su, W. SmartForest: An Interactive Forest Data and Modeling and Visualization Tool. In Proceedings of the Fifth Forest Service Remote Sensing Applications Conference, Portland, OR, USA, 11–15 April 1994. Available online: https://books.google.fi/books?hl=en&lr=&id=fsJkEr14QA4C&oi=fnd&pg=PA283&dq=related:56NBRsHRx3gJ:scholar.google.com/&ots=jmouQ_wwc5&sig=QW_SPV7OawsVEYt3zx5BfyXErfE&redir_esc=y#v=onepage&q&f=false (accessed on 12 September 2024).
8. Cristal, I.; Ameztegui, A.; González-Olabarria, J.R.; Garcia-Gonzalo, J. A Decision Support Tool for Assessing the Impact of Climate Change on Multiple Ecosystem Services. *Forests* **2019**, *10*, 440. [CrossRef]
9. Huang, J.; Lucash, M.S.; Scheller, R.M.; Klippel, A. Visualizing Ecological Data in Virtual Reality. In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 23–27 March 2019; pp. 1311–1312. [CrossRef]
10. Holopainen, M.; Vastaranta, M.; Hyyppä, J. Outlook for the Next Generation's Precision Forestry in Finland. *Forests* **2014**, *5*, 1682–1694. [CrossRef]
11. Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* **2002**, *80*, 88–99. [CrossRef]
12. Hyyppä, J.; Inkinen, M. Detecting and estimating attributes for single trees using laser scanner. *Photogramm. J. Finl.* **1999**, *16*, 27–42.
13. Vastaranta, M.; Kankare, V.; Holopainen, M.; Yu, X.; Hyyppä, J.; Hyyppä, H. Combination of individual tree detection and area-based approach in imputation of forest variables using airborne laser data. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 73–79. [CrossRef]
14. Kaartinen, H.; Hyyppä, J.; Yu, X.; Vastaranta, M.; Hyyppä, H.; Kukko, A.; Holopainen, M.; Heipke, C.; Hirschmugl, M.; Morsdorf, F.; et al. An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. *Remote Sens.* **2012**, *4*, 950–974. [CrossRef]
15. Wang, Y.; Hyyppä, J.; Liang, X.; Kaartinen, H.; Yu, X.; Lindberg, E.; Holmgren, J.; Qin, Y.; Mallet, C.; Ferraz, A.; et al. International benchmarking of the individual tree detection methods for modeling 3-D canopy structure for silviculture and forest ecology using airborne laser scanning. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5011–5027. [CrossRef]
16. Eysn, L.; Hollaus, M.; Lindberg, E.; Berger, F.; Monnet, J.M.; Dalponte, M.; Kobal, M.; Pellegrini, M.; Lingua, E.; Mongus, D.; et al. A benchmark of lidar-based single tree detection methods using heterogeneous forest data from the alpine space. *Forests* **2015**, *6*, 1721–1747. [CrossRef]
17. Vauhkonen, J.; Ene, L.; Gupta, S.; Heinzel, J.; Holmgren, J.; Pitkänen, J.; Solberg, S.; Wang, Y.; Weinacker, H.; Hauglin, K.M.; et al. Comparative testing of single-tree detection algorithms under different types of forest. *Forestry* **2012**, *85*, 27–40. [CrossRef]
18. Parkan, M.J. *Combined Use of Airborne Laser Scanning and Hyperspectral Imaging for Forest Inventories (No. THESIS)*; EPFL: Lausanne, Switzerland, 2019.
19. Reitberger, J.; Schnörr, C.; Krzystek, P.; Stilla, U. 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 561–574. [CrossRef]
20. Pyörälä, J.; Saarinen, N.; Kankare, V.; Coops, N.C.; Liang, X.; Wang, Y.; Holopainen, M.; Hyyppä, J.; Vastaranta, M. Variability of wood properties using airborne and terrestrial laser scanning. *Remote Sens. Environ.* **2019**, *235*, 111474. [CrossRef]
21. Krzystek, P.; Serebryanyk, A.; Schnörr, C.; Červenka, J.; Heurich, M. Large-Scale Mapping of Tree Species and Dead Trees in Šumava National Park and Bavarian Forest National Park Using Lidar and Multispectral Imagery. *Remote Sens.* **2020**, *12*, 661. [CrossRef]
22. Ma, Q.; Lin, J.; Ju, Y.; Li, W.; Liang, L.; Guo, Q. Individual structure mapping over six million trees for New York City USA. *Sci. Data* **2023**, *10*, 102. [CrossRef]
23. Hyyppa, J.; Kelle, O.; Lehikoinen, M.; Inkinen, M. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 969–975. [CrossRef]
24. Persson, A.; Holmgren, J.; Soderman, U. Detecting and measuring individual trees using an airborne laser scanner. *Photogramm. Eng. Remote Sens.* **2002**, *68*, 925–932.
25. Rahman, M.Z.A.; Gorte, B. Tree filtering for high density airborne LiDAR data. In Proceedings of the International Conference on LiDAR Applications in Forest Assessment and Invertory, Edinburgh, UK, 17–19 September 2008. Available online: http://resolver.tudelft.nl/uuid:efd5d40c-3062-4d98-9f71-bbc431b9cd65 (accessed on 12 September 2024).
26. Hyyppä, J.; Yu, X.; Hyyppä, H.; Vastaranta, M.; Holopainen, M.; Kukko, A.; Kaartinen, H.; Jaakkola, A.; Vaaja, M.; Koskinen, J.; et al. Advances in Forest Inventory Using Airborne Laser Scanning. *Remote Sens.* **2012**, *4*, 1190–1207. [CrossRef]
27. Hyyppä, E.; Yu, X.; Kaartinen, H.; Hakala, T.; Kukko, A.; Vastaranta, M.; Hyyppä, J. Comparison of backpack, handheld, under-canopy UAV, and above-canopy UAV laser scanning for field reference data collection in boreal forests. *Remote Sens.* **2020**, *12*, 3327. [CrossRef]
28. Persson, Å.; Holmgren, J.; Söderman, U.; Olsson, H. Tree species classification of individual trees in Sweden by combining high resolution laser data with high resolution near-infrared digital images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *36*, 204–207. Available online: https://www.isprs.org/proceedings/xxxvi/8-w2/PERSSON.pdf (accessed on 12 September 2024).

29. Holmgren, J.; Persson, Å.; Söderman, U. Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. *Int. J. Remote Sens.* **2008**, *29*, 1537–1552. [CrossRef]
30. Wang, Y.; Weinacker, H.; Koch, B.; Sterenczak, K. Lidar point cloud based fully automatic 3D single tree modelling in forest and evaluations of the procedure. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *37*, 45–51. Available online: https://www.isprs.org/proceedings/xxxvii/congress/6b_pdf/08.pdf (accessed on 12 September 2024).
31. Reitberger, J.; Heurich, M.; Krzystek, P.; Stilla, U. Single tree detection in forest areas with high-density LiDAR data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2007**, *36*, 139–144. Available online: https://www.isprs.org/proceedings/xxxvi/3-w49/partb/papers/139_pia07b.pdf (accessed on 12 September 2024).
32. Morsdorf, F.; Meier, E.; Allgöwer, B.; Nüesch, D. Clustering in airborne laser scanning raw data for segmentation of single trees. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2003**, *34*, W13. Available online: https://www.isprs.org/proceedings/xxxiv/3-w13/papers/morsdorf_alsdd2003.pdf (accessed on 12 September 2024).
33. Hyyppä, E.; Kukko, A.; Kaartinen, H.; Yu, X.; Muhojoki, J.; Hakala, T.; Hyyppä, J. Direct and automatic measurements of stem curve and volume using a high-resolution airborne laser scanning system. *Sci. Remote Sens.* **2022**, *5*, 100050. [CrossRef]
34. Windrim, L.; Bryson, M. Detection, segmentation, and model fitting of individual tree stems from airborne laser scanning of forests using deep learning. *Remote Sens.* **2020**, *12*, 1469. [CrossRef]
35. Liang, X.; Kankare, V.; Hyyppä, J.; Wang, Y.; Kukko, A.; Haggrén, H.; Yu, X.; Kaartinen, H.; Jaakkola, A.; Guan, F.; et al. Terrestrial laser scanning in forest inventories. *ISPRS J. Photogramm. Remote Sens.* **2016**, *115*, 63–77. [CrossRef]
36. Hyyppä, E.; Kukko, A.; Kaijaluoto, R.; White, J.C.; Wulder, M.A.; Pyörälä, J.; Liang, X.; Yu, X.; Wang, Y.; Kaartinen, H.; et al. Accurate derivation of stem curve and volume using backpack mobile laser scanning. *ISPRS J. Photogramm. Remote Sens.* **2020**, *161*, 246–262. [CrossRef]
37. Luoma, V.; Saarinen, N.; Wulder, M.A.; White, J.C.; Vastaranta, M.; Holopainen, M.; Hyyppä, J. Assessing Precision in Conventional Field Measurements of Individual Tree Attributes. *Forests* **2017**, *8*, 38. [CrossRef]
38. Straker, A.; Puliti, S.; Breidenbach, J.; Kleinn, C.; Pearse, G.; Astrup, R.; Magdon, P. Instance segmentation of individual tree crowns with YOLOv5: A comparison of approaches using the ForInstance benchmark LiDAR dataset. *ISPRS Open J. Photogramm. Remote Sens.* **2023**, *9*, 100045. [CrossRef]
39. Hakula, A.; Ruoppa, L.; Lehtomaki, M.; Yu, X.; Kukko, A.; Kaartinen, H.; Taher, J.; Matikainen, L.; Hyyppä, E.; Luoma, V.; et al. Individual tree segmentation and species classification using high-density close-range multispectral laser scanning data. *ISPRS Open J. Photogramm. Remote Sens.* **2023**, *9*, 100039. [CrossRef]
40. Hauglin, M.; Rahlf, J.; Schumacher, J.; Astrup, R.; Breidenbach, J. Large scale mapping of forest attributes using heterogeneous sets of airborne laser scanning and National Forest Inventory data. *For. Ecosyst.* **2021**, *8*, 65. [CrossRef]
41. Wagner, F.H.; Roberts, S.; Ritz, A.L.; Carter, G.; Dalagnol, R.; Favrichon, S.; Hirye, M.C.M.; Brandt, M.; Ciais, P.; Saatchi, S. Sub-meter tree height mapping of California using aerial images and LiDAR-informed U-Net model. *Remote Sens. Environ.* **2024**, *305*, 114099. [CrossRef]
42. Tolan, J.; Yang, H.-I.; Nosarzewski, B.; Couairon, G.; Vo, H.V.; Brandt, J.; Spore, J.; Majumdar, S.; Haziza, D.; Vamaraju, J.; et al. Very high resolution canopy height maps from RGB imagery using self-supervised vision transformer and convolutional decoder trained on aerial lidar. *Remote Sens. Environ.* **2024**, *300*, 113888. [CrossRef]
43. Schaller, C.; Ginzler, C.; Van Loon, E.; Moos, C.; Seijmonsbergen, A.C.; Dorren, L. Improving country-wide individual tree detection using local maxima methods based on statistically modeled forest structure information. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *123*, 103480. [CrossRef]
44. Hyyppä, J.; Mielonen, T.; Hyyppä, H.; Maltamo, M.; Yu, X.; Honkavaara, E.; Kaartinen, H. Using individual tree crown approach for forest volume extraction with aerial images and laser point clouds. In Proceedings of the ISPRS Workshop Laser Scanning 2005, Enschede, The Netherlands, 12–14 September 2005. Available online: https://www.isprs.org/proceedings/xxxvi/3-w19/papers/144.pdf (accessed on 12 September 2024).
45. Uusitalo, J.; Kivinen, V.-P. Implementation SmartForest forest visualization tool on PC environment. In Resource Technology '98 Nordic International Symposium on Advanced Technology in Environmental and Natural Resources; (Metsäntutkimuslaitoksen tiedonantoja; No. 791). Metla. 2000, pp. 1–7. Available online: https://jukuri.luke.fi/bitstream/handle/10024/529318/metla-2015102210538.pdf (accessed on 12 September 2024).
46. Fabrika, M. Virtual forest stand as a component of sophisticated forestry educational systems. *J. For. Sci.* **2003**, *49*, 419–428. Available online: https://www.agriculturejournals.cz/pdfs/jfs/2003/09/04.pdf (accessed on 12 September 2024). [CrossRef]
47. Chou, C.-Y.; Song, B.; Hedden, R.L.; Williams, T.M.; Culin, J.D.; Post, C.J. Three-Dimensional Landscape Visualizations: New Technique towards Wildfire and Forest Bark Beetle Management. *Forests* **2010**, *1*, 82–98. [CrossRef]
48. Wang, H.; Chen, X.; Polys, N.; Sforza, P. A Web3D forest geo-visualization and user interface evaluation. In Proceedings of the 22nd International Conference on 3D Web Technology (Web3D '17), Brisbane, QLD, Australia, 5–7 June 2017; pp. 1–9. [CrossRef]
49. Mattila, O.; Korhonen, A.; Pöyry, E.; Hauru, K.; Holopainen, J.; Parvinen, P. Restoration in a virtual reality forest environment. *Comput. Hum. Behav.* **2020**, *107*, 106295. [CrossRef]
50. Xiong, Q.; Huang, X. Speed Tree-Based Forest Simulation System. In Proceedings of the 2010 International Conference on Electrical and Control Engineering, Wuhan, China, 25–27 June 2010; pp. 3033–3036. [CrossRef]
51. Bundesamt Für Landestopografie Swisstopo, Die Schweiz in 3D. Available online: https://www.swisstopo.admin.ch/de/die-schweiz-in-3d (accessed on 7 April 2024).

52. Gemeente Rotterdam, Rotterdam 3D. Available online: https://www.3drotterdam.nl/#/ (accessed on 7 April 2024).
53. Han, J.H.; Huang, X.Y. Design of 3D game based on forestry science popularization. In Proceedings of the 2010 3rd International Conference on Computer Science and Information Technology, Chengdu, China, 9–11 July 2010; IEEE: Piscataway, NJ, USA, 2010; Volume 7, pp. 577–581. [CrossRef]
54. Rutzinger, M.; Pratihast, A.K.; Oude Elberink, S.J.; Vosselman, G. Tree modelling from mobile laser scanning data-sets. *Photogramm. Rec.* **2011**, *26*, 361–372. [CrossRef]
55. Nebiker, S.; Bleisch, S.; Christen, M. Rich point clouds in virtual globes—A new paradigm in city modeling? *Comput. Environ. Urban Syst.* **2010**, *34*, 508–517. [CrossRef]
56. Prandi, F.; Panizzoni, G.; Magliocchetti, D.; Devigili, F.; De Amicis, R. WebGL virtual globe for efficient forest production planning in mountainous area. In Proceedings of the 20th International Conference on 3D Web Technology (Web3D '15), Heraklion, Greece, 18–21 June 2015; Association for Computing Machinery: New York, NY, USA, 2015; pp. 143–151. [CrossRef]
57. Jyväskylä, Kunta 3D. Available online: https://jyvaskyla.kunta3d.fi/Map.html (accessed on 7 April 2024).
58. Roynard, X.; Deschaud, J.-E.; Goulette, F. Paris-Lille-3D: A large and high-quality ground-truth urban point cloud dataset for automatic segmentation and classification. *Int. J. Robot. Res.* **2018**, *37*, 545–557. [CrossRef]
59. Raumonen, P.; Kaasalainen, M.; Åkerblom, M.; Kaasalainen, S.; Kaartinen, H.; Vastaranta, M.; Holopainen, M.; Disney, M.; Lewis, P. Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. *Remote Sens.* **2013**, *5*, 491–520. [CrossRef]
60. Hackenberg, J.; Spiecker, H.; Calders, K.; Disney, M.; Raumonen, P. SimpleTree—An Efficient Open Source Tool to Build Tree Models from TLS Clouds. *Forests* **2015**, *6*, 4245–4294. [CrossRef]
61. National Land Survey of Finland, Laser Scanning Data 5p. 2024. Available online: https://www.maanmittauslaitos.fi/en/maps-and-spatial-data/expert-users/product-descriptions/laser-scanning-data-5-p (accessed on 7 April 2024).
62. National Land Survey of Finland, Digital Elevation Model. 2024. Available online: https://www.maanmittauslaitos.fi/en/research/interesting-topics/digital-elevation-model (accessed on 7 April 2024).
63. Tomppo, E.; Haakana, M.; Katila, M.; Peräsaari, J. Multi-source national forest inventory—Methods and applications. In *Managing Forest Ecosystems*; Springer: Berlin/Heidelberg, Germany, 2008; Volume 18, p. 374, ISBN 978-1-4020-8712-7. [CrossRef]
64. Mäkisara, K.; Katila, M.; Peräsaari, J. *The Multi-Source National Forest Inventory of Finland—Methods and Results 2015*; Natural resources and bioeconomy studies 8/2019; Natural Resources Institute Finland: Helsinki, Finland, 2019; 57p. Available online: http://urn.fi/URN:ISBN:978-952-326-711-4 (accessed on 7 April 2024).
65. National Land Survey of Finland, Orthophotos. 2024. Available online: https://www.maanmittauslaitos.fi/en/maps-and-spatial-data/expert-users/product-descriptions/orthophotos (accessed on 7 April 2024).
66. Li, W.; Guo, Q.; Jakubowski, M.K.; Kelly, M. A New Method for Segmenting Individual Trees from the Lidar Point Cloud. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 75–84. [CrossRef]
67. Yu, X.; Hyyppä, J.; Kukko, A.; Maltamo, M.; Kaartinen, H. Change detection techniques for canopy height growth measurements using airborne laser scanning data. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1339–1348. [CrossRef]
68. Yu, X.; Hyyppä, J.; Vastaranta, M.; Holopainen, M.; Viitala, R. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 28–37. [CrossRef]
69. Rantala, S. *Tapion Taskukirja (Finnish Handbook for Foresters)*, 25th ed.; Metsälehti: Helsinki, Finland, 2008; pp. 284–286.
70. Ecotree. 2023. Available online: https://ecotree.green/en/how-much-co2-does-a-tree-absorb (accessed on 28 January 2024).
71. Laasasenaho, J. Taper curve and volume functions for pine, spruce and birch. *Commun. Instituti For. Fenn.* **1982**, *108*, 74.
72. Repola, J. Biomass equations for birch in Finland. *Silva Fenn.* **2008**, *42*, 605–624. [CrossRef]
73. Repola, J. Biomass equations for Scots pine and Norway spruce in Finland. *Silva Fenn.* **2009**, *43*, 625–647. [CrossRef]
74. Soininen, V.; Kukko, A.; Yu, X.; Kaartinen, H.; Luoma, V.; Saikkonen, O.; Holopainen, M.; Matikainen, L.; Lehtomäki, M.; Hyyppä, J. Predicting Growth of Individual Trees Directly and Indirectly Using 20-Year Bitemporal Airborne Laser Scanning Point Cloud Data. *Forests* **2022**, *13*, 2040. [CrossRef]
75. Yun, T.; Li, J.; Ma, L.; Zhou, J.; Wang, R.; Eichhorn, M.P.; Zhang, H. Status, advancements and prospects of deep learning methods applied in forest studies. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *131*, 103938. [CrossRef]
76. Huo, L.; Strengbom, J.; Lundmark, T.; Westerfelt, P.; Lindberg, E. Estimating the conservation value of boreal forests using airborne laser scanning. *Ecol. Indic.* **2023**, *147*, 109946. [CrossRef]
77. Laitinen, O.; Hämäläinen, J. Wood Supply Goes Digital. Available online: https://www.metsagroup.com/news-and-publications/others/campaigns/intelligent-forest/wood-supply-goes-digital/ (accessed on 17 June 2024).
78. Liang, X.; Yao, H.; Qi, H.; Wang, X. Forest in situ observations through a fully automated under-canopy unmanned aerial vehicle. *Geo-Spat. Inf. Sci.* **2024**, *4*, 983–999. [CrossRef]
79. Liang, X.; Wang, Y.; Pyörälä, J.; Lehtomäki, M.; Yu, X.; Kaartinen, H.; Kukko, A.; Honkavaara, E.; Issaoui Nevalainen, O.; Vaaja, M.; et al. Forest in situ observations using unmanned aerial vehicle as an alternative of terrestrial measurements. *For. Ecosyst.* **2019**, *6*, 1–16. [CrossRef]
80. Balenović, I.; Liang, X.; Jurjević, L.; Hyyppä, J.; Seletković, A.; Kukko, A. Hand-held personal laser scanning–current status and perspectives for forest inventory application. *Croat. J. For. Eng. J. Theory Appl. For. Eng.* **2021**, *42*, 165–183. [CrossRef]
81. Hyyppä, E.; Hyyppä, J.; Hakala, T.; Kukko, A.; Wulder, M.A.; White, J.C.; Pyörälä, J.; Yu, X.; Wang, Y.; Virtanen, J.-P.; et al. Under-canopy UAV laser scanning for accurate forest field measurements. *ISPRS J. Photogramm. Remote Sens.* **2020**, *164*, 41–60. [CrossRef]

82. Zhao, J.; Sensibaughm, T.; Bodenheimer, B.; McNamara, T.; Nazareth, A.; Newcombe, N.; Minear, M.; Klippel, A. Desktop versus Immersive Virtual Environments: Effects on Spatial Learning. *Spat. Cogn. Comput.* **2020**, *20*, 328–363. [CrossRef]
83. Murtiyoso, A.; Holm, S.; Riihimäki, H.; Krucher, A.; Griess, H.; Griess, V.; Schweier, J. Virtual Forests: A Review on Emerging Questions in the Use and Application of 3D Data in Forestry. *Int. J. For. Eng.* **2023**, *35*, 29–42. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.