Remote Sensing of Environment 334 (2026) 115223

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Large-scale tree-level mapping of forest structure including species type with remote sensing data and ground measurements

J. Kostensalo [a,*], P. Packalen [b], M. Kuronen [b], L. Mehtätalo [a], S. Tuominen [b], M. Myllymäki [b]

[a] Natural Resources Institute Finland (Luke), Yliopistokatu 6B, FI-80100 Joensuu, Finland
[b] Natural Resources Institute Finland (Luke), Latokartanonkaari 9, FI-00790 Helsinki, Finland

HIGHLIGHTS

- Structurally representative maps were created using RS data and ground truth.
- Species, DBH, height, volume, and location were predicted for each tree.
- Height, DBH, and volume distributions by species were accurately reproduced.
- The created maps can be reliably used to calculate biodiversity proxies.

ARTICLE INFO

Edited by Dr Marie Weiss

Keywords:
Airborne laser scanning
Hyperspectral imaging
Multispectral imaging
Species type detection
Digital twin

ABSTRACT

Remote-sensing based tree maps can be used to calculate various diversity indices, but the detection probability of trees depends on size and species. We propose a novel approach combining individual tree detection (ITD) with resampling corrections (+R) which aims to simultaneously correct the size, species, and spatial distribution of trees using scalable algorithms. Using airborne laser scanning, optical data, and ground measurements, we demonstrate the compatibility of ITD+R with two different types of forests and ITD algorithms, as well as its scalability to areas exceeding 3000 km^2. The tree maps were evaluated using plot-level variables and benchmarked against area-based k nearest neighbors (k-NN). The ITD+R improved ITD results for most studied metrics, with the Shannon index being an exception, and even outperformed k-NN in predicting dominant height in managed stands, though k-NN still outperformed for stem density and volume. The ITD+R approach was shown to be adaptable to various diversity indices which it has not been specifically trained on, with 254 m^2 plot-level predictions correlating at $r = 0.42$–0.91. While ITD trees could be classified with OA = 82.0%–86.6% to pine, spruce, and deciduous, further research is needed to account for rare tree species, as low prevalence results in a large number of false detections which cannot be sufficiently addressed with classification alone.

1. Introduction

Forest structure characterization can be done efficiently for large areas using remote sensing (RS) techniques like airborne laser scanning (ALS) (Naesset, 2002; Hansen et al., 2013; Lechner et al., 2020). Traditionally, the goal has been to map economically important variables using the area-based approach (ABA) which aims to capture stand-level characteristics (e.g., Magnussen et al., 2013). ABA can be used on ALS data with low pulse densities, but as these data are routinely collected with ever higher pulse densities (Kaartinen et al., 2012), this opens up the possibility for individual tree detection (ITD) (Hyyppä et al., 2001). There is a growing interest in using ALS for mapping biodiversity (e.g., Ørka et al., 2022; Toivonen et al., 2023), and an accurate tree-level map would be ideal, as it would allow for the flexible computation of *structural diversity* indicators, i.e., variation in tree size, species, and spatial distribution (Staudhammer and LeMay, 2001).

Since ABA focuses on predicting plot-level (or stand-level) variables, it tends to outperform ITD-methods for standing volume and stem density (e.g., Kotivuori et al., 2021), though results can be comparable for managed stands (Packalén et al., 2012), and in some studies ITD has outperformed ABA in predicting total volume (Yu et al., 2010).

* Corresponding author.
Email address: joel.kostensalo@luke.fi (J. Kostensalo).

https://doi.org/10.1016/j.rse.2025.115223
Received 3 February 2025; Received in revised form 19 December 2025; Accepted 28 December 2025
Available online 5 January 2026

However, both approaches perform poorly in predicting stand structure, i.e., structural diversity (Peuhkurinen et al., 2011). The main reason for ITD underperformance relates to the undetection of subdominant trees (Mehtätalo et al., 2022) biasing the size, species, and spatial distribution. In a typical ALS study about 40% of the trees are detected (see, e.g. Pitkänen et al., 2004; Ene et al., 2012). The missed understory layer also has silvicultural and ecological relevance (Pereira et al., 2013), and thus is desirable to correct for.

In Kostensalo et al. (2023) a new method was proposed for correcting the biases of ITD height and spatial distribution by resampling ground-measured trees which were undetected during model training. The method addresses both undetected trees and false detections, and provides as an output a tree map and not just a list of trees. The tree map aims to emulate a full wall-to-wall field inventory thus offering significant flexibility for the user. The method was shown to be computationally fast, and thus potentially applicable for country-wide mapping. The main limitation of the method as presented in Kostensalo et al. (2023) is that it does not include species in the tree maps. While many structural diversity indices, such as Gini-coefficients and standard deviations of height and diameter, can be calculated without species information, other important variables such as deciduous tree volume cannot. The ITD detection probabilities can vary by species, so the inclusion of species is important for improving false detection and undetection estimates as well. The species-dependence of detection probability necessitates a completely new formulation of the resampling procedure, but if successful, this could correct for the biases in size, species, and spatial distributions originating from undetected understory trees, while not actually detecting these trees.

In this work, we aim to extend the individual tree detection and resampling method to include species, resulting in tree-level predictions for height, diameter, species, volume, and location. We aim to create maps which can be used flexibly to calculate various structural diversity indices, while simultaneously capturing established inventory variables with precision as close as possible to a state-of-the-art ABA method.

2. Data

2.1. Study sites

This study was based on data from two separate areas in Southern Finland. The first one was in Mikkeli, approximately 61° 41' – 62° 11' N and 26° 25' – 27° 27' E, while the second one was in Evo, approximately 61° 09' – 61° 13' N and 25° 00' – 25° 13' E. The locations of the study areas and the field measurement plots are presented on a map in Fig. 1.

The Mikkeli area represents managed Southern boreal forest, which is predominantly in private ownership, with the dominant species being coniferous tree species, i.e., Norway spruce (*Picea abies* (L.) Karst.) and Scots pine (*Pinus sylvestris* L.), which are typical of Finnish managed forests. Deciduous trees are in the minority, with the most common broad-leaved species being downy birch (*Betula pubescens* Ehrh.) and silver birch (*Betula pendula* Roth).

The Evo study area covers mainly state-owned forest. The majority of the area is managed for timber production, but also contains some protected land areas, whose growing stock differs structurally from the managed forests. As in Mikkeli, the main tree species are Scots pine, Norway spruce, silver birch, and downy birch, but the managed part also has some Siberian and European larch stands. As all other species

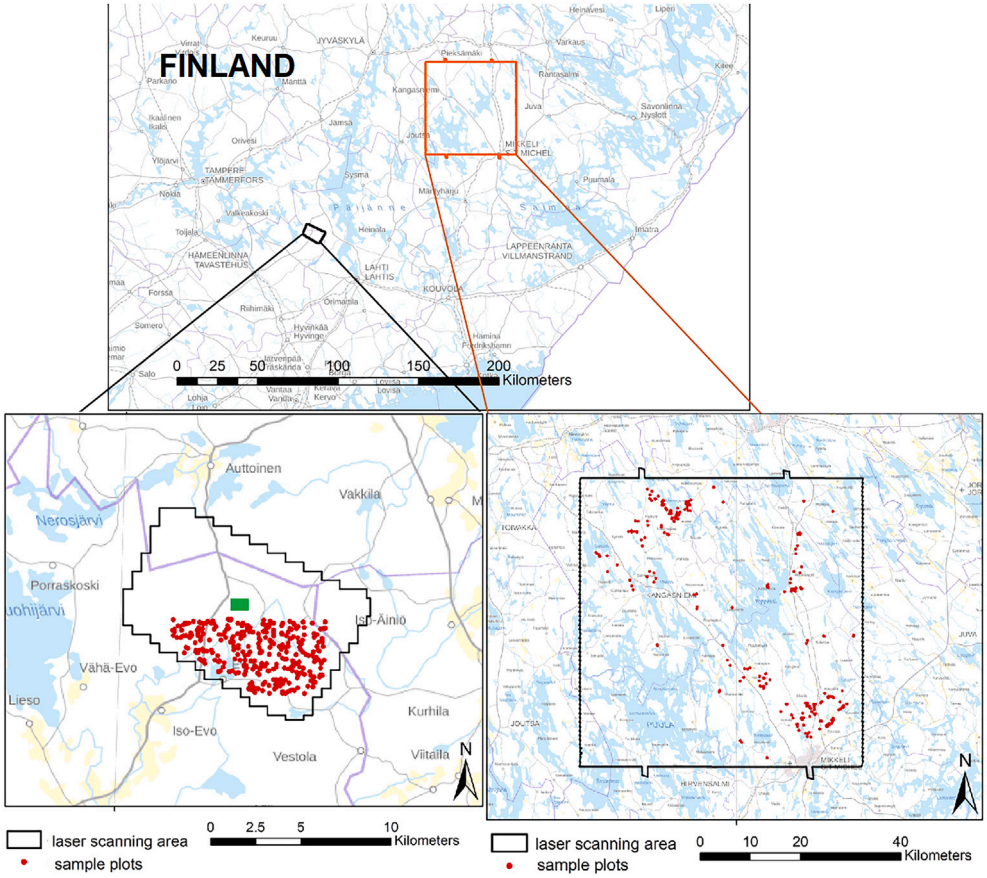

Fig. 1. The location of the two study regions: Mikkeli (right) and Evo (left). ALS and optical imagery data have been collected wall-to-wall (the gray area representing forests with remote sensing data available), while the red dots show the field plot locations. The background map ((c) National Land Survey of Finland 2022) shows rivers and lakes as blue. The green rectangular area in Evo marks the example area near lake Huhmari, for which the tree map is shown in Fig. 14.

Table 1
Characteristics of the two study sites used in this study. The base temperature for the annual effective temperature sum has been set at 5 °C. In both study areas, the topography is characterized by lakes, bogs, and low hills and ridges.

	Mikkeli	Evo
Forested area (scanned)	2221 km^2	88 km^2
Age of tally trees (mean, range)	46, 5–156	N/A
Vegetation zone	South-boreal	South-boreal
Dominant land cover type	Woodland	Woodland
Elevation, meters ASL	125–200	80–220
Annual rainfall, mm/year (mean 1991–2020)	601	634
Length of thermic growing season, days (mean 1991–2020)	170	170
Annual effective temperature sum, day-degrees (mean 1991–2020)	1280–1300	1270–1360

Table 2
Descriptive statistics of the ground measurements for the two study areas in Mikkeli (268 plots; 6407 trees) and Evo (279 plots; 10,898 trees) in Finland. DBH and height are given at the tree-level and other variables are calculated at the plot-level.

All plots	Mikkeli			Evo		
	Mean	SD	Range	Mean	SD	Range
DBH (cm)	16.2	8.7	[3.0, 54.2]	13.2	8.7	[4.5, 97.7]
Height (m)	15.5	6.9	[1.3, 34.8]	13.0	6.0	[3.7, 42.7]
Mean DBH (cm)	18.5	5.5	[7.5, 38.0]	15.4	5.9	[6.8, 38.6]
Mean height (m)	17.1	4.5	[7.1, 28.9]	14.4	4.4	[6.2, 28.6]
Stem density	940	570	[118, 3539]	1535	878	[197, 5227]
Total volume (m^3/ha)	254	138	[49, 817]	287	178	[18, 978]

besides pine, spruce, and birch represent only 0.96% of the standing volume in Mikkeli and 7.3% in Evo, we will consider deciduous trees here as a single category. Thus, the categorization in this work could be considered *species type detection* rather than true *species detection*. Relevant study site characteristics are listed in Table 1.

2.2. Field measurement data

In Mikkeli, all trees with DBH≥50 mm (or DBH 30–49 mm if their height was over half the mean stand height) were recorded from 268 circular plots with a radius of 9 m. For all tally trees, location, DBH, and species were recorded. Locations of the trees were based on the pseudolite-based positioning TerraHärps system implemented in Masser ExCaliper II callipers, with sixteen base stations for each plot. Height estimation as well as the location accuracy are described in Kostensalo et al. (2023). In Evo, all trees with DBH≥45 mm were measured from 279 circular plots with a radius of 9 m. The plots were allocated based on stratified sampling using species dominance, stem density, mean diameter, and basal area from previous inventory data as auxiliary information. Seedling stands were excluded from this study. The basal-area weighted median tree of each tree species was used for calibrating height predictions of the tally trees. Field inventories were carried out by experienced ground crews to ensure that the location, height, DBH, and species detection were done correctly. With species type classification pine/spruce/deciduous, there is little risk of misclassification is the ground data. Descriptive statistics are presented in Table 2.

2.3. Remote sensing data

2.3.1. ALS data

In Mikkeli, ALS data were collected by the scanning operator BSF Swissphoto AG as a part of the Finnish national campaign which covers the entire country every six years (https://www.maanmittauslaitos.fi/en/maps-and-spatial-data/expert-users/product-descriptions/laser-scanning-data-5-p). The ALS data were acquired in June 2020 using a RIEGL VQ-1560i laser scanner. In Evo, ALS data was acquired on July 16, 2018 using a Leica ALS70-HP scanner. The data were collected specifically from the study area in question. The scanning operator was Terratec AS. The parameters of ALS data collection in both Mikkeli and Evo are given in Table 3. ALS echoes were classified as ground and other hits similarly in both study areas. The ground classification was based on the method of Axelsson (2000), with Delaunay triangulation used for modeling the terrain. Echo heights were then normalized by subtracting the terrain level from the echoes. Examples of the ALS data are presented in Fig. 2.

Table 3
Details of the ALS devices used and the relevant parameters.

	Mikkeli	Evo
ALS device	RIEGL VQ-1560i	Leica ALS70-HP
Data collection date	June 14–25, 2020	July 16, 2018
Flight altitude, AGL	1525 m	1300 m
Pulse repetition rate	1.4 MHz	0.42 MHz
Scanning frequency	148 Hz	66.3 Hz
Half-angle	20 deg	20 deg
Pulse density	7.2 points/m^2	10.2 points/m^2

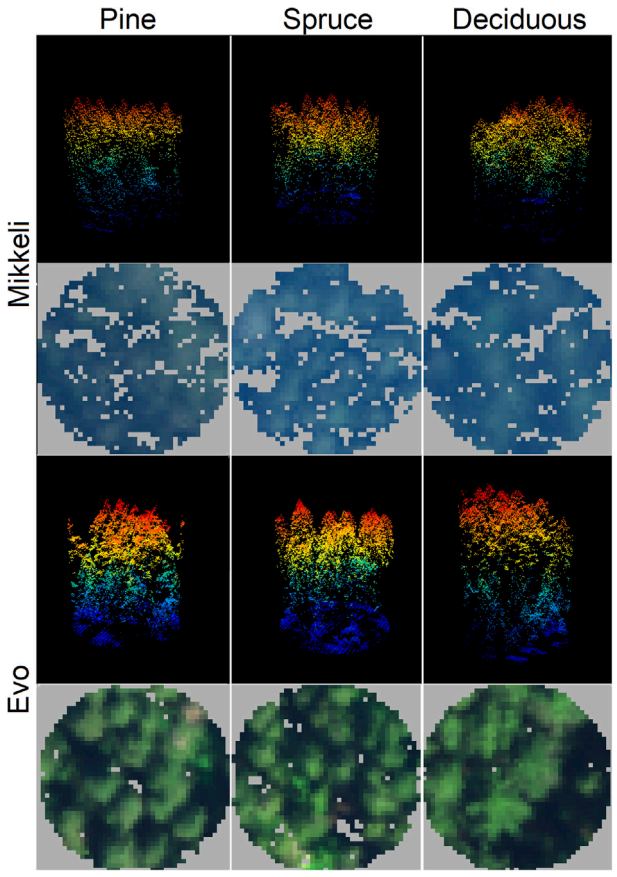

Fig. 2. Renderings of the ALS and optical imagery in Mikkeli and Evo for three different stands: ones dominated by pine, spruce, and deciduous trees, respectively. The images have a 15 m radius, i.e., they include a 6 m buffer around the 9 m plot. In Mikkeli, the ALS-density is lower and the mean value images have less contrast than those in Evo. The optical renderings here are based on RGB-bands from Mikkeli and on three bands of the hyperspectral images (ca. 450, 550, and 650 nm) for Evo.

2.3.2. Optical imagery

In Mikkeli, aerial images were acquired using the Vexcel UltraCam Eagle aerial camera on June 12, 2020. Flying altitude was 7700 m

above ground level, which resulted in a ground sampling distance of 120 cm in multispectral bands. We used only the original multispectral bands (R,G,B,NIR; processing level-2) without pan-sharpening or orthorectification. Aerial images and ALS data were fused in order to fetch pixel values of images to first ALS echoes. Unrectified images were used to avoid problems with relief displacement. External orientations were solved using the bundle block adjustment technique, in which unknown parameters were simultaneously solved using tie point and control point measurements (Mikhail et al., 2001). Internal orientation parameters were available in the camera calibration report. The location of the ALS echo in the single image plane was calculated by the collinearity equations based on internal and external orientation (Mikhail et al., 2001). Each echo is linked to multiple images because images overlap. Therefore, we computed the mean value of pixels by band from images for which an echo hit. In the end, each first echo of ALS data had pixel values for red, green, blue and NIR attached to it (see details in Packalén et al. (2009)).

In Evo, hyperspectral images were acquired using HySpex SWIR-384 and HySpex VNIR-1800 sensors on July 16, 2018 from an altitude of 1300 m above ground level. The spectral range of VNIR data was 406–995 nm (186 bands), and the spatial resolution of the imagery was 0.5 m. The spectral range of SWIR imagery was 956–2525 nm (288 bands), and the spatial resolution was 1.0 m. The field of view was ±8° for the SWIR sensor and ±8.5° for the VNIR sensor. Side overlap was approximately 23% for SWIR and VNIR imagery. Hyperspectral images were orthorectified and mosaicked, first echoes of ALS data were linked to hyperspectral data and bandwise intensity values were attached to each ALS echo. Examples of the optical data are presented in Fig. 2.

3. Methods

3.1. Recreating tree maps with remote sensing and resampling

In this section and the following subsections we will present the methodology for recreating tree maps using ground measurements in concert with remote sensing data. The methods are an extension of those presented in Kostensalo et al. (2023) where the framework of ITD optimization combined with the removal of false detections and resampling corrections for undetected trees was presented. We shall refer to this methodology presented in this work as individual tree detection and resampling (ITD + R).

The basic idea here is that the ground measurements can be represented by a matrix X encoding all key attributes of each tree: its location (x, y), height (h), DBH, volume (V), and tree species (type). The goal is then to create another matrix X' utilizing remote sensing data and ground measurements from training data such that

$$X' \approx X \tag{1}$$

meaning that the stem density, spatial pattern, and total volume as well as the height, DBH, and volume distributions (optimally both marginal and joint distributions) are approximately correct, both in aggregate as well as for each tree species. We include both understory and overstory trees as we are interested in every tree in this work. Thus, for each species type, we want

$$\#(X') \approx \#(X) \tag{2}$$

$$h_{X'} \overset{d}{\approx} h_X \tag{3}$$

$$\mathrm{DBH}_{X'} \overset{d}{\approx} \mathrm{DBH}_X \tag{4}$$

$$V_{X'} \overset{d}{\approx} V_X \tag{5}$$

here, # refers to the number of trees, and h_X, DBH_X, and V_X are the tree heights, DBHs, and volumes of the ground measurements, respectively, while $h_{X'}$, $\mathrm{DBH}_{X'}$, and $V_{X'}$ are the corresponding quantities for X'. In addition, the spatial structure should be well reproduced. In the present work structural diversity and inventory variables are seen as characteristics of a forest stand, and thus the tree level results are not of interest by themselves (Staudhammer and LeMay, 2001), though success in tree-level predictions is reflected in the quality of plot-level predictions.

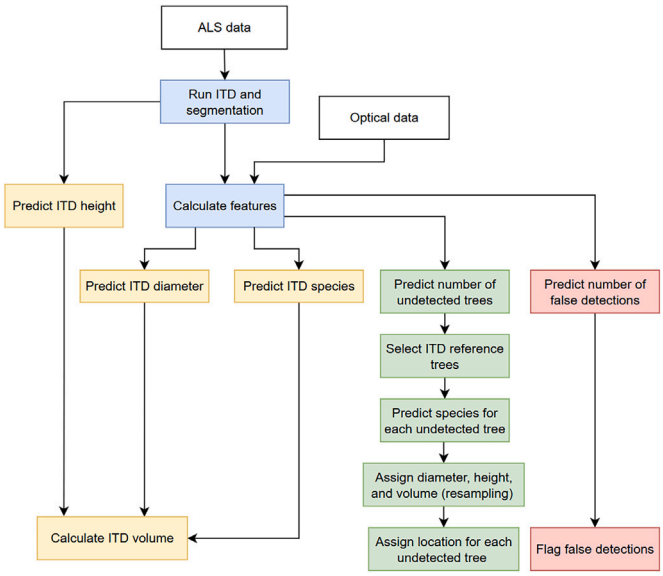

Fig. 3. Flowchart of the steps in the the ITD + R framework for building the matrix X' suggested in this work. Color-coding is used to highlight the steps mainly relating to all trees (blue), true detections (yellow), undetections (green), and false detections (red).

The ITD + R framework for building the matrix X' suggested in this work is illustrated in Fig. 3 and consists of these steps:

1. Detect and estimate the dominant trees
   (1) ITD and segmentation (Section 3.1.1)
   (2) Predict height, DBH, volume, and species type for dominant trees (Section 3.1.2)
2. Correct for undetected trees
   (1) Estimate the number of undetected trees and augment the detected trees with resampled trees (Section 3.1.3)
3. Correct for false detections
   (1) Remove likely false detections (Section 3.1.3)

In training and evaluating of the ITD algorithms, an algorithmic procedure for matching detected trees with the field measurements described in Kostensalo et al. (2023) was used. The algorithm has two parameters: *maximal xy-distance* and *maximal z-difference*, which limit how much the ITD and ground trees can differ in their location and height, respectively. The value of the first parameter was set by plotting ground-measured tree locations on top of RGBNIR (or hyperspectral) and lidar data for five plots with the smallest number of trees in each study area, and observing how much the locations of ITD treetops and ground measured trees differed. The second parameter was chosen by looking at the 2.5% and 97.5% percentiles of the distribution of maximum ground measured height to the highest ALS echo. Once these two parameters were set, the matching was done for each plot by

1. Creating a list with xy-distance and height difference for all possible pairs of ITD and ground measured trees
2. Removing all rows where either the xy-distance is over *maximal xy-distance*, or the height difference is over *maximal z-difference*
3. Sorting the list by xy-distance from smallest to largest

4. Keeping only the first appearance of each ground measured tree
5. Keeping only the first appearance of each ITD tree

Once the algorithm finishes, the rows remaining in the list correspond to matched trees. Step 2) ensures that the heights and locations do not differ more than can be explained by the accuracy of the data used and steps 4) and 5) ensure that each ground measured and ITD tree is not matched more than once.

3.1.1. Individual tree detection and segmentation

Local maximum filter

A canopy height model (CHM) was built based on the first returns of the ALS with a resolution of 0.5 m. We restricted the ITD approach to the first ALS echoes, as the later echoes cannot be reasonably linked to the optical imagery. The first method we used for individual tree detection was a local maximum filter algorithm *lmf* (see Popescu and Wynne, 2004) from the R (R Core Team, 2022) package lidR (Roussel et al., 2020). The minimum height parameter was fixed to 3.8 m and a circular detection window with a variable size was used. The minimum height was set based on Kostensalo et al. (2023), where an optimal minimum height of 3.6 and 4.0 m was reported. Since in that work the results were very similar for any value between 3 and 5 m, we did not optimize for this parameter separately. The window size was determined by a simple function being $ws_{\min}$ at ground level, $ws_{\max}$ at 30 m and growing linearly between 0–30 m and being constant above 30 m. The parameters $ws_{\min}$ and $ws_{\max}$ were optimized with the target being to maximize the F_1 score

$$F_1 = \frac{2TP}{2TP + FP + FN}, \tag{6}$$

where TP is the number of true positives, FN the number of false negatives, and FP being the number of false positives. The matching algorithm presented in Kostensalo et al. (2023) was used with the location and height having to match to an accuracy determined by graphical investigations of the data. The optimal values were found using a grid search for a wide range of $ws_{\min}$ and $ws_{\max}$ values. The grid cells were 0.1×0.1 m. The segmentation of the ALS point cloud was based on the algorithm presented in Silva et al. (2016) as implemented in the lidR package.

Deep-learning-based segmentation

As a second method, we used the *SegmentAnyTree*-algorithm (Wielgosz et al., 2024), which is a deep-learning-based method. The model uses a 3D convolutional neural network, and we used instance segmentation to detect trees. We used the implementation provided in the ForestSens-service (forestsens.com). Though ForestSens recommends using the algorithm with point densities above 50 pts/m^2, Wielgosz et al. (2024) tested the performance with point densities as low as 10 pts/m^2. While the performance was significantly worse than with high point densities, the method did not break down. As the Evo dataset had a point density of 10 pts/m^2, we decided to use this as a test case. The X–Y-location of the trees was calculated by averaging all the hits above 1.5 m and the height was registered as the maximum Z-coordinate. Unlike with the *lmf*-approach which was based on the first echoes, we gave the *SegmentAnyTree*-algorithm the full point cloud, as the method is in principle capable of segmenting trees below the canopy.

3.1.2. Predicting height, DBH, volume, and species type for dominant trees

Since the ITD tree heights may be biased, the tree heights were corrected using a basic linear regression model with the ITD tree height as the only covariate. DBH and species type classification models were based on both tree and stand-level features. The models were trained separately for each cross-validation fold with 90% of the data and tested on 10%.

The tree-level features were calculated from the ALS echoes assigned to the ITD tree by the segmentation algorithm. The tree-level features included the minimum, maximum, mean, median, lower and upper quartiles, standard deviation, inter-quartile range of the Z-coordinate from the ground-normalized segmented point clouds, crown diameter approximated as $0.5(\max(x) - \min(x) + \max(y) - \min(y))$, and the mean and standard deviation of the distance of the echoes on the X–Y-plane from the center of the segmented point cloud. Furthermore, the kurtosis and the skewness of the Z-coordinate were also included. Image features used in Mikkeli were ratios of the pixel values of the bands and the NDVI vegetation index. These were computed from the pixel values attached to the set of first ALS echoes (see Section 2.3.2). In Evo, the first five principal components of the hyper spectral data of the crown were used. We restricted the number to five, as these components were enough to explain 99.9% of the variation, with these components capturing approximately 91.9%, 5.9%, 1.9%, 0.1%, and 0.1% of the variation, with small differences between cross validation rounds. The stand-level features included the proportion of ground hits ($Z < 2$ m) and the number of ITD trees found and its square.

The DBH predictions for the ITD trees were based on a partial-least squares (PLS) approach. First, a model was fitted with all features and the most important features then selected using variable importance in projection (VIP) (Wold et al., 1993) with a cut-off value set to 0.25. The model was then re-fitted with these variables and finally, the number of components was set to 6 in Mikkeli and 5 in Evo based on the lowest RMSECV value for the first full CV fold of the model framework. A fourth-root transformation was used for DBH to ensure positive predicted values and to correct for heteroskedasticity. A second-order Taylor polynomial correction for bias was used when back-transforming the DBH prediction (e.g., Mehtätalo and Lappi, 2020).

The species type classification was performed using a random forest (RF) algorithm (Ho, 1995) with features selected from the same set of potential features as those used in the DBH model, but with the mean and standard deviation of the distance of the echoes on the X–Y-plane removed and R/NIR and G/NIR removed for Mikkeli, as these were highly correlated with other features. For all RF models 400 trees were used. In Evo, a single RF model was employed. In Mikkeli, the RF model had some trouble in separating the less common downy birch from spruces, and thus a second RF model was added on top of the first RF which was trained to only predict if a given tree is a downy birch or another species. For the final prediction, the predicted species type classification from the RF model with all species was used unless the downy birch model identified the tree as a downy birch. The number of split-variables was set to 5 in the full model for Mikkeli and to 6 in Evo and the downy birch model in Mikkeli.

The models were fitted using the R libraries PLS (Mevik and Wehrens, 2007) and ranger (Wright and Ziegler, 2017). With the height, DBH, and species type accounted for, the volume of the tree was calculated using the models for pine, spruce, and birch given in Kangas et al. (2022).

3.1.3. Correcting for undetected trees and false detections

The number of undetected trees in a plot was modeled with a PLS approach. In addition to the plot-level features listed in Section 3.1.2 (i.e., the proportion of ground hits ($Z < 2$ m), and the number of ITD trees found and its square), the maxima, means, and IQRs of the all the tree-level features used in the DBH and species type classification models for the dominant trees were used as possible predictors. A similar procedure for feature selection was used as with the DBH models with the VIP cut-off set to 0.25 and with nine components selected for the final model in Mikkeli and 13 in Evo. A fourth-root transformation was used here as well and the final result was obtained by back-transforming the prediction with a second-order Taylor polynomial bias correction and then rounding the predicted number of trees to the nearest integer.

Once the number of trees undetected by ITD was predicted for a plot, a corresponding number of trees were simulated inside the plot, one tree at a time. The trees were added using a resampling-type method, where the compensating trees are picked at random from the ground-measured trees which were not detected during the training of the models. In the training phase, the nearest neighbor distance distribution of the ground-measured tree locations is first calculated ($TDATA_{NN}$). Next, the set of

missed trees (height, diameter, volume, species) is obtained as those ground measured trees, which were not detected with the optimized ITD ($TDATA_{trees}$). The third and final dataset created using the training data is the species matrix ($TDATA_{species}$), which specifies the species distribution of undetected trees next to a detected tree of a given species (type).

The location and attributes of each simulated tree were then assigned using the following algorithm:

1. Select an ITD tree using random sampling with probabilities proportional to tree height.
2. Choose a random distance from the nearest-neighbor distance set $TDATA_{NN}$.
3. Choose a random direction (Unif($[0, 2\pi]$)) from the ITD tree selected in Step 1 and place the simulated tree at this distance from the ITD tree sampled in Step 2.
4. Sample the tree species (type) of the tree based on the probabilities from the column of the species matrix $TDATA_{species}$ corresponding to the tree species of the ITD selected in Step 1.
5. Choose a tree of the species sampled in Step 4 at random from $TDATA_{trees}$ which are smaller than the height of the ITD tree chosen in Step 1.

The location of each compensating tree is obtained in Steps 1–3, its species in Step 4, and its size (height, DBH, V) in Step 5. The species is determined for each compensating tree independently. Alternatively, the species of the compensating trees could be determined jointly. However, this would add another layer of complexity to the approach.

The number of false discoveries was predicted with a PLS model analogous to that for the undetected trees in Section 3.1.3, except that 14 components were selected in Evo, and a transformation of the response variable was not used. The final prediction was obtained by setting any negative predictions to zero and then rounding up to the nearest integer. In each plot the smallest trees were labeled as false detections, as based on the analysis performed in Kostensalo et al. (2023) the smallest ITD trees are the most likely to be something else than a treetop.

3.1.4. Benchmark method: area-based approach

We compared the accuracy of the proposed tree level approach to the accuracy obtained by an area-based k-NN approach. The variant of the area-based approach we used is described in Appendix A. This variant is similar to the approach used operationally in Finland (Maltamo et al., 2021) and is, therefore, a reasonable baseline for comparing new approaches.

3.2. Evaluation of tree maps

The quality of the ITD + R framework was evaluated at the plot level against ITD without corrections and the area-based k-NN predictions using 10-fold cross validation. In each CV fold, all the models were trained using 90% of the plots not in the fold starting from the optimization of the ITD, then fitting the height, DBH, and species type classification models, and creating the TDATA-datasets, which were then used to create tree maps for the 10% test set. For performance comparison dominant height, stem density (total and by species), total standing volume (total and by species), and basal-area weighted DBH (all trees) were considered. The primary metric used here was the root-mean squared error (RMSE), in addition to which the empirical bias and the coefficient of determination r^2 for the predicted and ground-truth values are also reported. When the RMSE value has been calculated using a test set independent of the training data the clarifying term RMSE of prediction (RMSEP) is used. R^2_{CV} refers to the coefficient of determination from cross validation. The analysis was carried out for Mikkeli and Evo separately.

As the ITD + R framework can also be used to map out proxy variables for structural diversity, we chose nine different indices from the literature (Storch et al., 2018; Toivonen et al., 2023): mean DBH^2, standard deviation of DBH and height, number of trees with DBH≥ 40 cm per hectare, Gini coefficient of DBH and height, Shannon index and Simpson's index, and the percentage of deciduous trees by volume. As none of these variables were considered during the development of the framework or fitting of the models, these indices serve to test how the maps could serve for applications for which they were not specifically optimized. In this work, the spatial pattern is not considered as the methodology remains unchanged from Kostensalo et al. (2023) in this regard.

3.3. Large-scale application

In order to demonstrate and test the large-scale applicability of the proposed methods, we applied the methods wall-to-wall to the entire forested areas shown in Fig. 1. First, all the models were fitted with full field data. The prediction of the number of false detections and undetected trees was done in a 16 × 16 m grid. As the 9 m circular plots with an area of 254.5 m^2 had almost exactly the same area as the 16 × 16 m rectangles (256 m^2) and the features used in the prediction (such as the number of ITD trees, proportion of ground hits) do not depend on the shape of the plots, the models were applied without any need for corrections. The entire areas were mapped wall-to-wall and then a 2 × 2 m forest mask based on land use type as recorded by National Land Survey Finland was applied to remove areas which do not represent forest. The deep-learning-based segmentation was much too computationally heavy to perform wall-to-wall mapping, and thus we have restricted the mapping to the *lmf*-approach.

4. Results

The results section is organized as follows. First, the georeferencing accuracy is addressed in Section 4.1 and the performance of the individual models making up the ITD + R framework is presented in Section 4.2, starting with ITD (Section 4.2.1), and followed by the models for predicting the number of false detections and undetections (Section 4.2.2). Next, the height and DBH results are presented (Section 4.2.3), and the section concludes with the species type classification results (Section 4.2.4). Next, the distributions of height, DBH, and volume are discussed for all trees and by species. In the distribution discussion the results are presented aggregated over each site since one plot has on average 23 trees in Mikkeli and 37 in Evo, but may have also as little as 3–6 trees. We finish the section with plot-level results in Section 4.4.

4.1. Georeferencing accuracy

In order to estimate the location accuracy for the matching algorithm and reveal any underlying issues with the data, we took five plots (approx. 20 trees in total) from each study area for graphical analysis. The plots with the smallest number of trees were selected, so that the tree locations could be accurately identified with the problem of overlapping canopies minimized.

The canopy height model, RGBNIR-values, and ground measured tree locations are displayed for two of these plots for Mikkeli and for Evo can be found in the Supplementary material. The canopy height models have been fitted with a resolution of 0.5 m, and the colors have been rasterized with a a pixel size of 0.7 m. For Evo where we have hyperspectral data, four bands corresponding to 450, 550, 650, and 850 nm were picked for the graphical investigations, as these correspond to the wavelengths of blue, green, red, and near infrared analogously to the RGBNIR-data from Mikkeli. The first plot in Mikkeli consists of three pines with heights of 23, 24, and 28 m. The second plot includes two large silver birches which are 28 and 30 m high, and one 5 m tall pine on the north-east side of the plot and one 6 m tall spruce on the south-west side of the plot. The Evo plots include five trees each. The first plot consists of three dominant pines with heights of 27, 29, and 30 m, and one 12 m tall downy birch on the north-east of the plot and one 10 m

tall on the east side of the plot. The small trees are completely covered by the canopies of the dominant trees. The second plot has three dominant trees: a 27 m tall pine on the eastern side of the plot and two spruces, 25 and 28 m tall on the west and south west. In addition, there are two spruces, 6 and 7 m tall below the canopy of one of the dominant spruces.

In both Mikkeli and Evo, the canopy height model based locations were within 0–3 pixels of the ground measurements for dominant trees, corresponding to a maximal error of 1.5 m. For the color bands, the locations matched within 0–4 pixels. As the pixel size was 0.7 m here, this corresponds to a location accuracy of 2.8 m, which was adopted for the matching algorithms. As an example, we judge the location of the color bands to be off by three pixels for the dominant pine center-north of the plot in the second Mikkeli plot (see Supp. material).

In Mikkeli, the 2.5% and 97.5% percentiles for the difference between the maximum ALS echo Z-coordinate and ground height were −2.7 and 0.3 m, with the mean difference being −1.2 m. In Evo, the corresponding numbers were −3.7 and 3.9 m, with the mean difference being −0.2 m. Thus, in both cases the ALS heights were slightly lower, as can be expected. Based on these, numbers, we limited the maximum height difference for the matching algorithm to 2.8 m in Mikkeli and to 3.9 m in Evo, so that it covers 95% of the observed differences.

4.2. Individual model performance

4.2.1. ITD performance

Local maximum filter. For the ITD algorithm the optimal minimum window size corresponding to ground level was 2.4 m, and the optimal maximum window size corresponding to points 30 m or higher was 3.4 m in all ten CV folds. In Evo the optimal minimal value varied between 1.4 and 1.5 m and the optimal maximum value between 2.4 and 2.6 m.

In Mikkeli, 45.6% of the trees in the training set were discovered with the optimal parameters, with the number of false detections being 6.7% of the number of ground measured trees. This corresponds to an F_1-value of 0.60. In Evo, 37.4% of the trees were discovered with an additional 14.5% of false detections, corresponding to an F_1-value of 0.49. In Mikkeli, the height differences between matched ITD and ground measured trees was on average −1.1 m (SD 0.8 m), and the xy-location difference was 0.52 m (SD 0.47 m). In Evo, the corresponding numbers were 0.2 m (SD 1.7 m), and 0.63 m (SD 0.55 m).

Deep learning. In the Evo dataset, the *SegmentAnyTree*-algorithm found 34.4% of the trees with an additional 5.3% of false detections, corresponding to an F_1-value of 0.49 as well. The z-difference was on average 1.4 m (SD 1.0 m), and the xy-distance 0.76 m (SD 0.52 m).

4.2.2. False detection, undetection, and stem density

Local maximum filter. The performance of the models for predicting the number of false detections and undetections, as well as the combined performance of the ITD and the two models is illustrated in Fig. 4 for Mikkeli and in Fig. 5 for Evo. In Mikkeli, the model for the number of false detections had $R^2_{CV} = 0.14$ with an RMSEP of 1.5 trees/plot. The model for undetected trees had $R^2_{CV} = 0.11$ with an RMSEP of 11.7 trees/plot. R^2_{CV} was 0.27 for the stem density, with an RMSEP value of 488 stems/ha.

In Evo, the model for the number of false detections had $R^2_{CV} = 0.48$ with an RMSEP of 3.1 trees/plot. The model for undetected trees had $R^2_{CV} = 0.26$ with an RMSEP of 15.8 trees/plot. R^2_{CV} was 0.47 for the stem density, with an RMSEP value of 640 stems/ha.

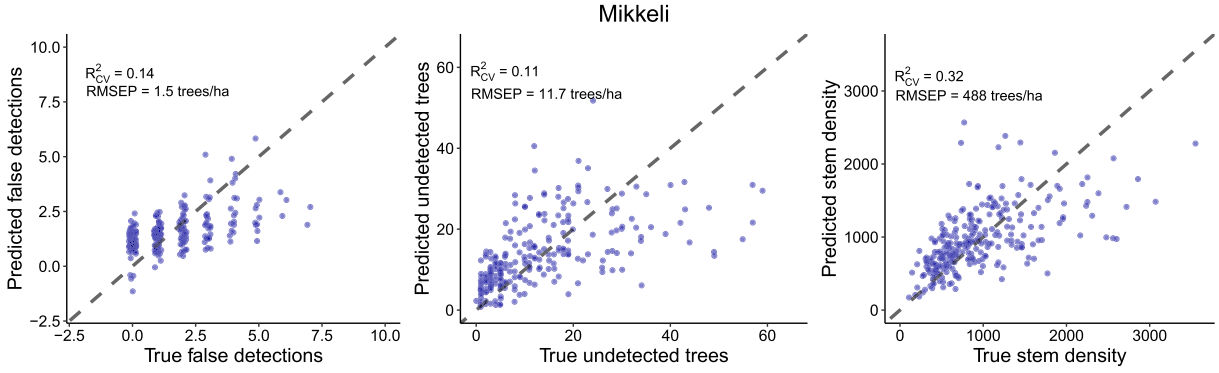

Fig. 4. Results from Mikkeli with the predicted number of false detections and the ground truth (left), the predicted number of undetected trees and the ground truth (center), and the total stem density (right). The first two plots are from one cross-validation fold (trees are not matched in the test set) while the stem density results are presented for the test plots with all ten folds included.

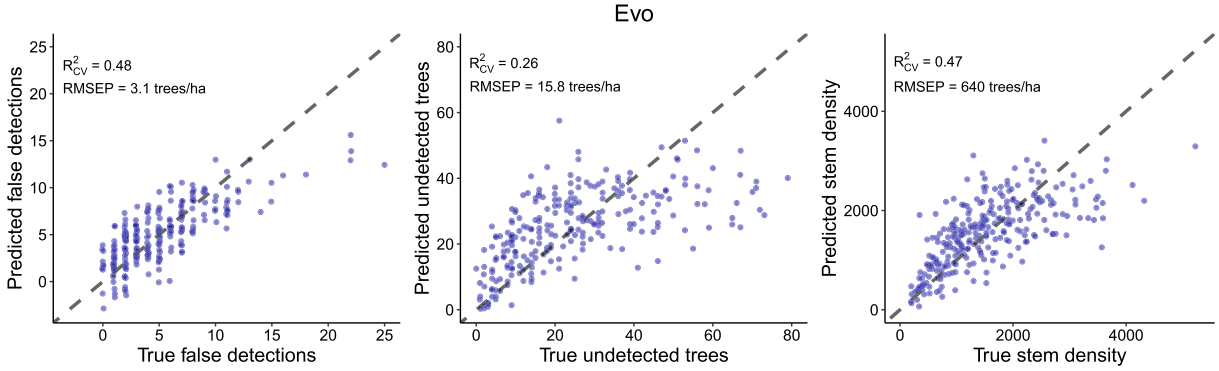

Fig. 5. Same as Fig. 4 for Evo.

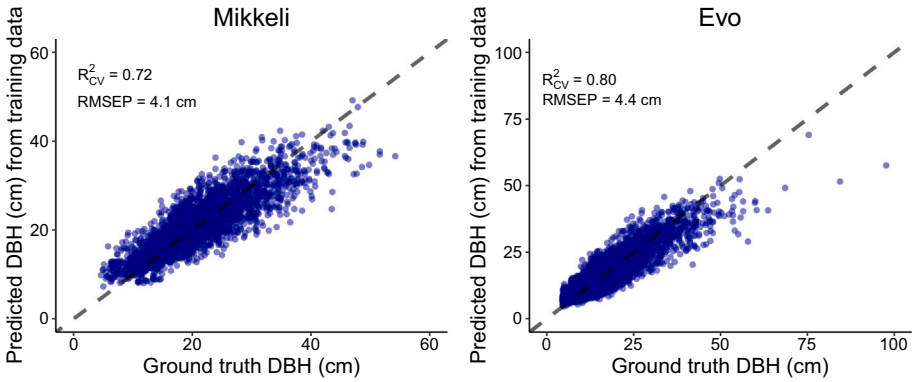

Fig. 6. The predicted and observed diameters at breast height (DBH) for all detected trees matched to ground-measured trees in the training set of one cross-validation fold for Mikkeli (left) and Evo (right).

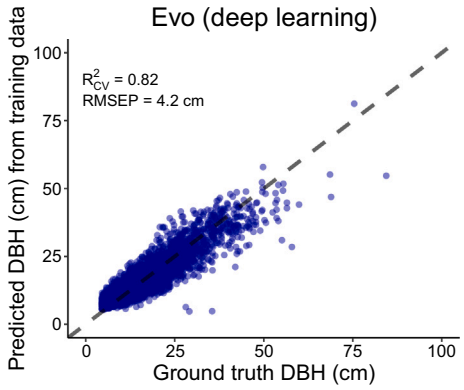

Fig. 7. Same as Fig. 6 for Evo with the *SegmentAnyTree*-algorithm.

Table 4
Confusion matrices for the random-forest based species type classification models for all ITD trees matched to ground-measured trees in the training set of one cross-validation fold for Mikkeli and Evo sites.

TRUE\PREDICTED	Mikkeli			Evo		
	Pine	Spruce	Deciduous	Pine	Spruce	Deciduous
Pine	89%	9%	2%	87%	10%	3%
Spruce	9%	81%	10%	14%	73%	13%
Deciduous	4%	5%	91%	5%	11%	84%

Table 5
The probabilities of an undetected tree being of a given species when the closest detected tree is of a given species (pine/spruce/deciduous) at the Mikkeli and Evo sites.

DETECTED\UNDETECTED	Mikkeli			Evo		
	Pine	Spruce	Deciduous	Pine	Spruce	Deciduous
Pine	34%	40%	26%	38%	46%	17%
Spruce	5%	81%	14%	5%	77%	18%
Deciduous	4%	46%	51%	3%	55%	41%

Deep learning. With the deep learning approach in Evo, the model for the number of false detections had R^2_{CV} = 0.17 with an RMSEP of 1.7 trees/plot. The model for undetected trees had R^2_{CV} = 0.48 with an RMSEP of 14.6 trees/plot. R^2_{CV} was 0.56 for the stem density, with an RMSEP value of 583 stems/ha (see Fig. 7).

Table 6
Confusion matrices for the random-forest based species type classification models for all ITD trees matched to ground-measured trees in the training set of one cross-validation fold in the Evo dataset based on segmentation using the *SegmentAnyTree*-algorithm.

TRUE\PREDICTED	Evo		
	Pine	Spruce	Deciduous
Pine	83%	14%	3%
Spruce	8%	83%	8%
Deciduous	11%	25%	64%

Table 7
The probabilities of an undetected tree being of a given species when the closest ITD tree is of a given species (pine/spruce/deciduous) in the Evo dataset based on segmentation using the *SegmentAnyTree*-algorithm.

DETECTED\UNDETECTED	Evo		
	Pine	Spruce	Deciduous
Pine	53%	21%	15%
Spruce	8%	68%	24%
Deciduous	5%	45%	49%

4.2.3. Height bias and DBH models

Local maximum filter. In Mikkeli, the ITD heights were systematically too short by 1.54 m, with the bias varying from −1.540 m for 2 m tall trees to −1.544 m for 36 m tall trees. The model had R^2 = 0.99 and a residual deviation of 0.58 m.

In Evo, the bias was non-trivial, varying from −0.68 m for 4 m tall trees to +1.32 m for 43 m tall trees. The model had R^2 = 0.93 and a residual deviation of 1.64 m. The ITD heights of true detection trees with a height around 17 m were unbiased.

The DBH results are presented in Fig. 6 for the true detections in both Mikkeli and Evo. In Mikkeli, the DBH model had R^2_{CV} = 0.72 with an RMSEP of 4.1 cm. In Evo, the DBH model had R^2_{CV} = 0.80 with an RMSEP of 4.4 cm. In both datasets, the most important predictors based on permutation for the DBH were the square of the number of ITD trees for the plot, the height of the tree, and the median, mean, and quartiles of the Z-coordinates of the segmented point cloud for the tree.

Deep learning. In Evo, the bias varied from +0.80 m for 4 m tall trees to −2.10 m for 43 m tall trees. The model had R^2 = 0.94 and a residual deviation of 1.64 m. The ITD heights of true detection trees with a height around 16 m were unbiased. The DBH model had R^2_{CV} = 0.82

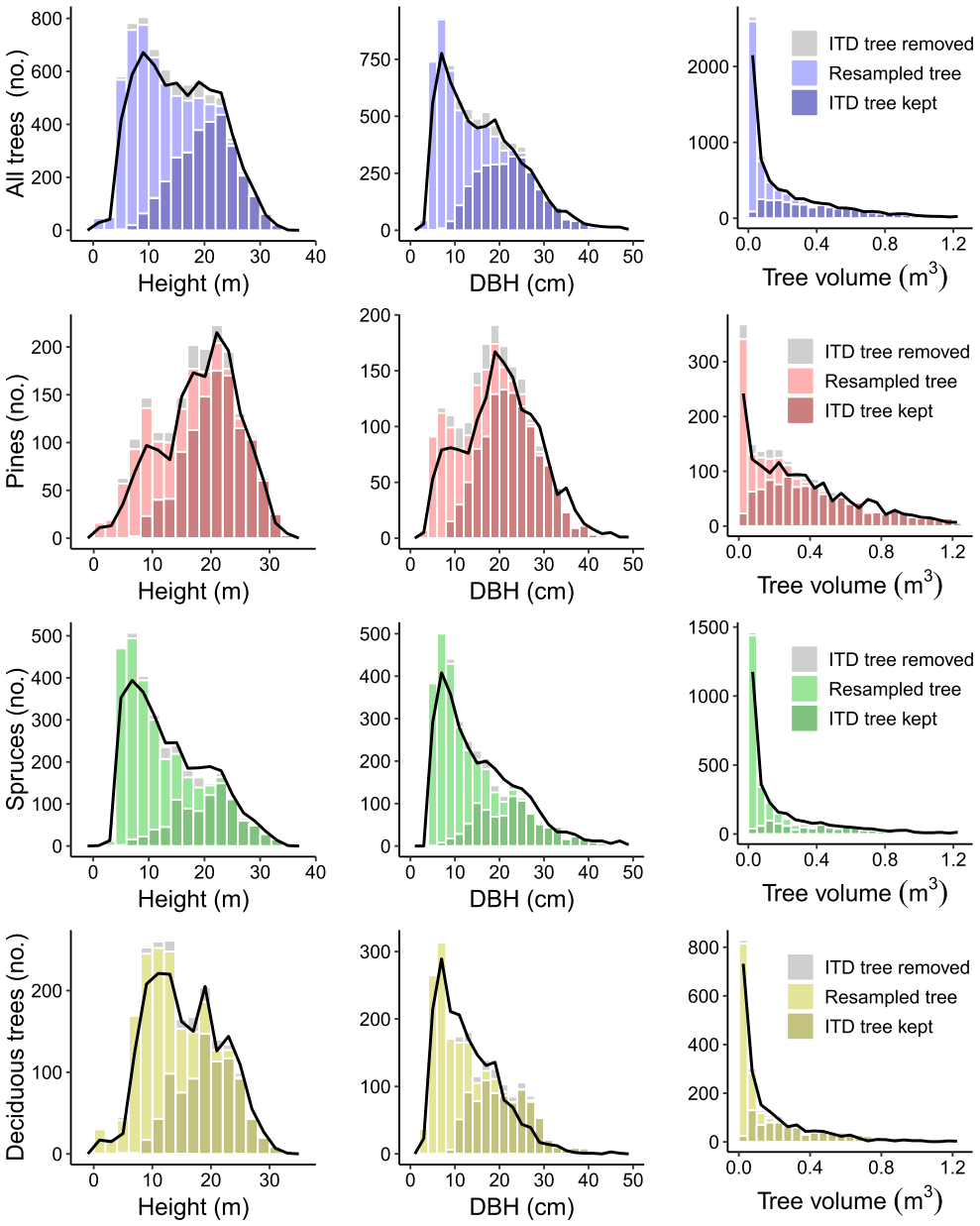

Fig. 8. The height, diameter at breast height (DBH), and volume distributions of trees in Mikkeli. The solid black line corresponds to the ground truth aggregated over all plots, and the histograms represent the ITD + R results (i.e., numbers of predicted trees in different bins) from a 10-fold cross-validation. The darker bars are the ITD trees kept, the lighter bars the resampled trees compensating for the missed trees, and the grey bars are the ITD trees removed as suspected false discoveries. The results are given for all trees (top row) and for by species (other rows).

with RMSEP of 4.2 cm with the same important predictor features as the *lmf*-based models.

4.2.4. Species type classification models for detected trees

Local maximum filter. The confusion matrices describing the random forest model performance for true detections (i.e., ITD tree successfully matched to a ground-truth tree) is presented in Table 4. In Mikkeli, the overall accuracy (OA) was 86.6%. The most important predictors were NDVI, R/NIR, and R/G. These were also the most important for the species-specific model for downy birch, though ITD height, proportion of ground hits and crown diameter also contributed here. Pines and deciduous trees were identified correctly 89% and 91% of the time, respectively, while the accuracy for spruces was 80%.

In Evo OA was 82.0%. There were two dominant contributors to the random forest model: the third and first PCA components of the hyperspectral data. PCA component 3 had strong positive loadings on 550–725 nm and negative loadings on 750–950 nm, while PCA 1 had approximately zero loading on visible light with a positive loading between 750–950 nm. Pines were identified correctly 87% of the time, while the accuracy for deciduous trees and spruces was 73% and 84%, respectively.

The species relationship matrix is presented in Table 5 for both locations. Undetected trees were most likely either spruce (40%–46%) or pine (34%–38%) if the closest ITD tree (i.e., dominant tree) was a pine. If the ITD tree was a spruce, then the nearest undetected trees were overwhelmingly spruces (77%–81%). If the ITD tree was deciduous, the undetected trees close to it were either spruces (46%–55%) or deciduous trees (41%–51%), but very unlikely pines (3%–4%).

Deep learning. In Evo the OA was 74.6% when the segmentation was based on the *SegmentAnyTree*-algorithm (see Table 6). There were two dominant contributors to the random forest model were PCA 3 and PCA

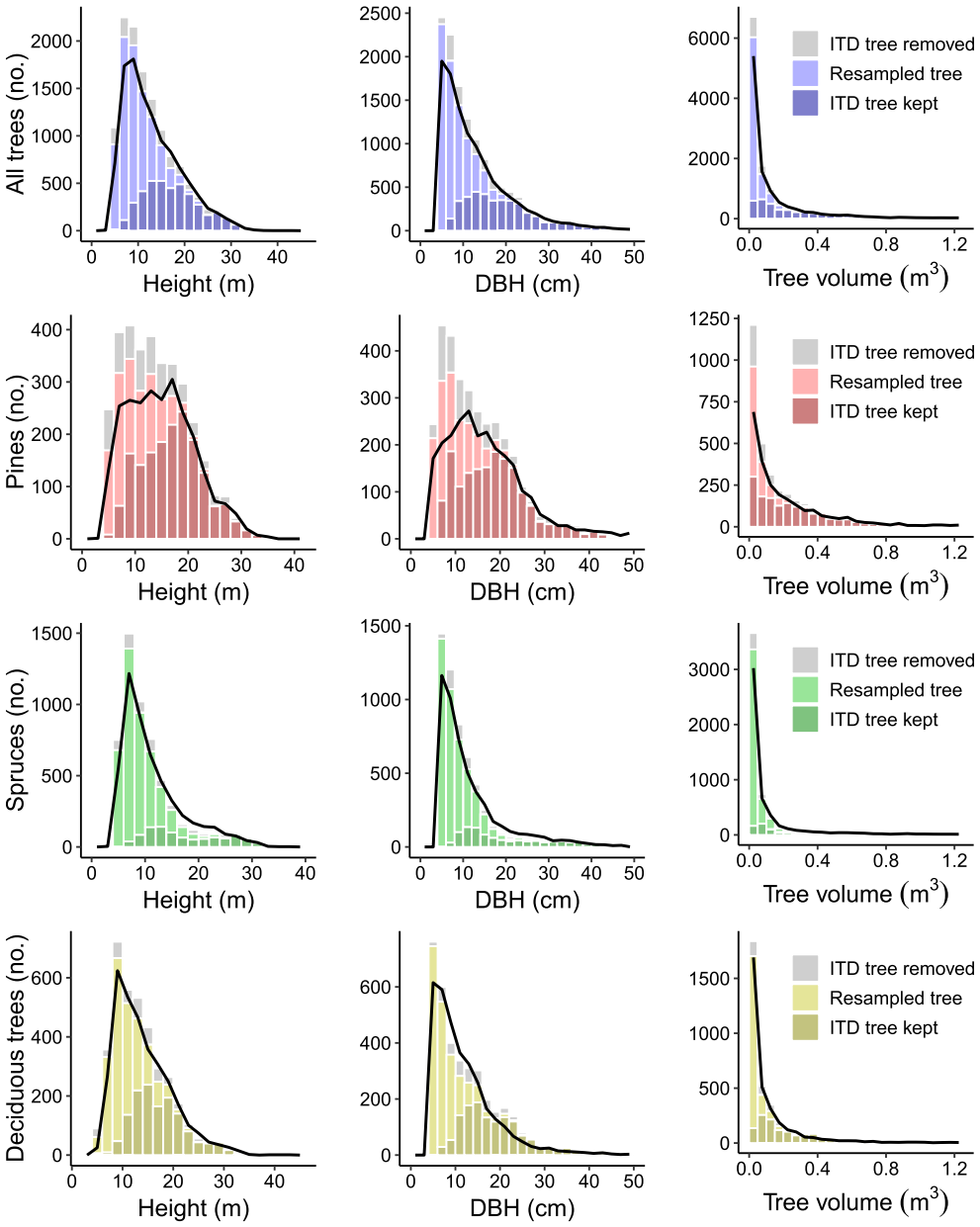

Fig. 9. Same as Fig. 8 for Evo with the *lmf*-based ITD + R.

1. Pines were identified correctly 83% of the time, while the accuracy for deciduous trees and spruces was 83% and 64%, respectively.

The species relationship matrix is presented in Table 7.

4.3. Aggregated height, diameter, and volume distribution

Local maximum filter. Tree height, DBH, and volume distributions are given for all trees and by species for Mikkeli in Fig. 8 and for Evo in Fig. 9. In both sites, the ground measured distributions for all trees are in good agreement with the ITD + R results, while the ITD only results are lacking about half of the trees with the smallest ones left undetected.

In Mikkeli, the ITD found most of the pines, while more than half of spruces and half of deciduous trees were left undiscovered. The suspected false discoveries were mostly pines, with some additional small spruces and deciduous trees. The ground truth species distribution is 25%/46%/28% (pine/spruce/deciduous) by number of trees and 38%/43%/18% by volume. The ITD-based numbers are 38%/33%/29% and 40%/36%/24%, while the ITD + R based numbers are 26%/45%/29% and 38%/38%/24%, respectively.

In Evo, most pines detected by the ITD, with only about one fourth of the spruces and about half of the deciduous trees detected. For all three species the false detection and resampling corrections of the ITD + R method managed to fix the biases of the site-level aggregated distribution. The ground truth species distribution is 23%/46%/30% (pine/spruce/deciduous) by number of trees and 37%/38%/25% by volume. The ITD-based numbers are 41%/27%/32% and 40%/31%/29%, while the ITD + R based numbers are 25%/45%/30% and 37%/35%/28%, respectively.

Deep learning. The *SegmentAnyTree*-based height, diameter, and volume distributions are given in Fig. 10. When compared with the *lmf*-based results, smaller portion of pines and deciduous trees were detected, while some more spruces were detected. However, the biggest difference appears to be the much smaller number of false detections

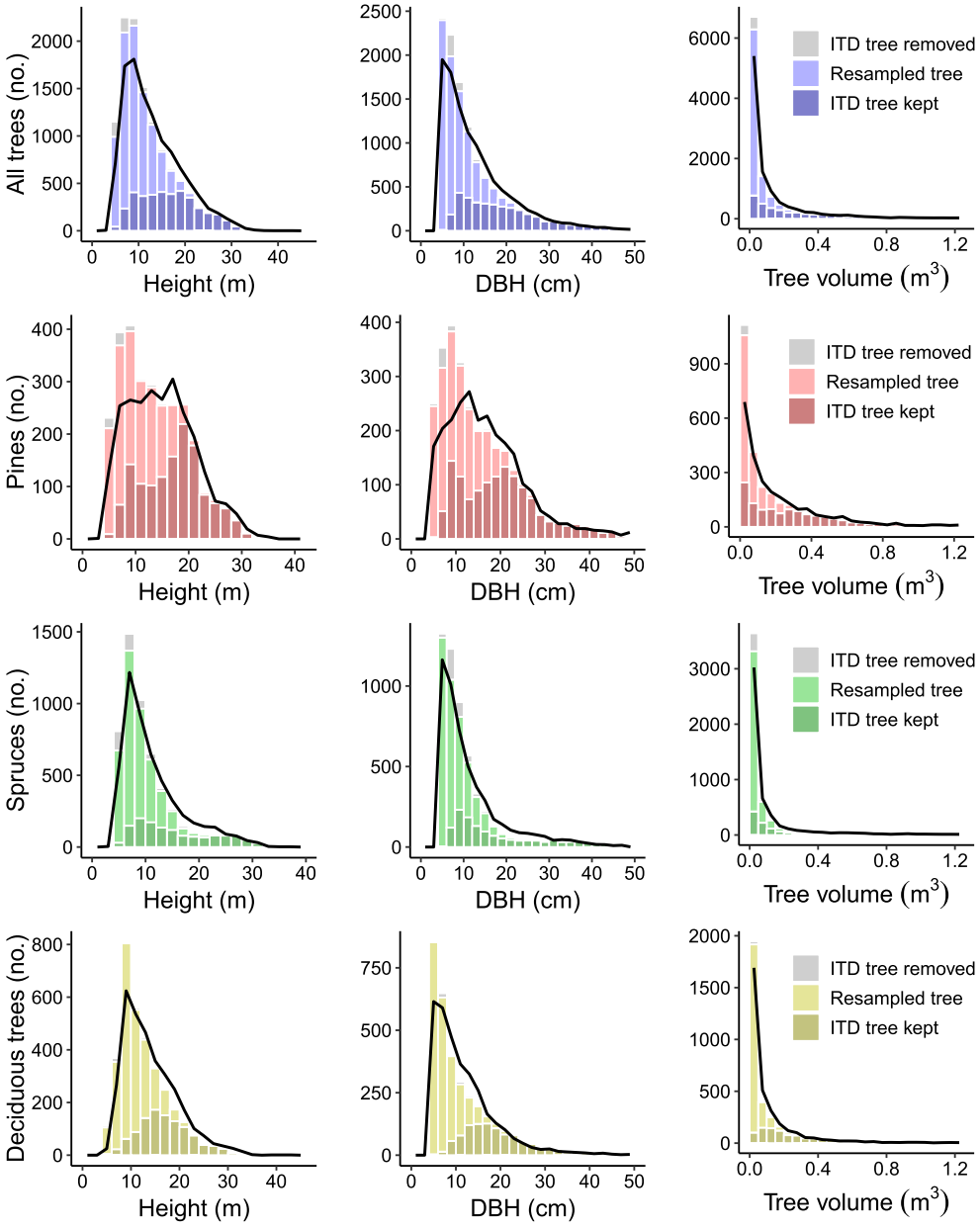

Fig. 10. Same as Fig. 8 for Evo with the deep-learning-based algorithm *SegmentAnyTree*.

with the *SegmentAnyTree*-algorithm, which lessens the need for corrections. The ITD-based species distributions are by number of trees 33%/41%/26% and 38%/37%/26% by volume, while the ITD + R based numbers are 25%/45%/30% and 37%/37%/26%, respectively.

4.4. Plot-level results

Local maximum filter. The ITD + R tree map reconstruction method is demonstrated for four representatively chosen plots from Mikkeli in Fig. 11. Looking at the ITD trees (solid points and empty circles), the number of trees is about one-half of the true number and the point-pattern is too regular, i.e., the clustering of trees seen in the ground truth map is missing. Furthermore, the smallest trees are systematically missing. The ITD + R corrects this, resulting in a more representative size distribution and spatial pattern. Also, the species distribution is better representative of the ground truth. In general, the species composition is well reproduced with the exception that in the first plot two small pines have been added even though the smaller trees are all spruces and deciduous trees according to the field measurements and one ITD tree has been misclassified as a pine in the third plot in addition to one simulated pine being added though the field measurements indicate that all trees were spruces. Also, in the second and fourth plot the species composition looks to be more even in the ITD + R map than in the ground truth. This would result in an overpredicting species diversity indices such as Shannon and Simpson's index, while a volume-based metric such as the proportion of deciduous trees by volume should be relatively robust, as the small simulated trees contribute relatively little to the total volume.

The tree maps in Fig. 11 highlight an important thing about the fundamental idea behind the ITD + R method: while the forest plot is reconstructed at the tree level, this does not mean that individual simulated trees would correspond to a specific tree in the forest. For example, in the first plot the ITD map only has large deciduous trees, though there are also small spruces and deciduous trees in the plot. The ITD + R correctly predicts the presence of such trees, and thus the resulting plot has

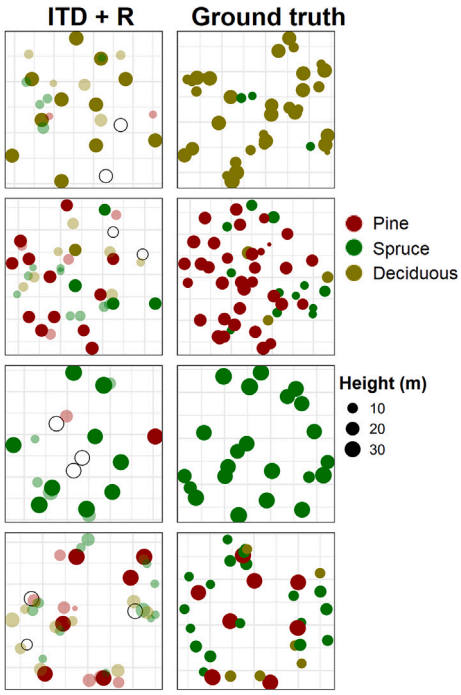

Fig. 11. Plot-level tree maps for four Mikkeli plots. The ITD + R maps are on the left-hand side with solid-colored points corresponding to ITD trees, empty circles to removed ITD trees, and transparent points to simulated trees.

some small deciduous trees and spruces. However, the five simulated small spruces are located mostly in the western side of the plot with one being right next to a large deciduous tree in the north-side of the plot, while in reality they are actually three spruces located at the center and eastern side of the plot. On one hand, for the goal of this paper, i.e., predicting plot-level metrics like the proportion of deciduous trees or mean DBH, errors in placement do not matter. It should be noted that even significant location errors, like rotating the plot by 90 or 180° or placing all trees at the center of the plot (though the latter would impact spatial summary statistics), has no impact on the stand metrics considered here. On the other hand, this means that maps constructed in this way are not suitable for tree-level management decisions.

Total standing volume and volume by species predictions of the ITD are presented in Fig. 12. In Mikkeli, the volumes are well predicted, but there is some systematic underpredicting of spruce volume and overpredicting of deciduous tree volume. In addition, there are some outliers for pine and spruce, indicating that a pine stand has been mistaken for a spruce stand and vice versa. In Evo, the total volume predictions are fine for most plots, but for the highest volume plots the errors are larger. While there are no clear systematic issues there is much more error in the species-level volumes than in Mikkeli, and there are more clear outliers for all species.

The results for the established inventory variables are presented in Table 8 for Mikkeli. In Mikkeli, ITD and ITD + R had an RMSEP of 0.9 m for dominant height, while the ABA method had 1.1 m. For stem density the ABA performed the best with an RMSEP of 384 stems/ha, while that of the ITD + R was 460 stems/ha and that of the ITD was 630 stems/ha. For the other inventory variables the ABA method performed the best. The RMSEP value for total standing volume was 78 m^3/ha for the ITD, 73 m^3/ha for ITD + R, and 58 m^3/ha for ABA. Regarding total volume by species, the results were similar for pine (RMSEP = 57 m^3/ha for ABA) but the RMSEP values for spruce and deciduous trees were about 10 m^3/ha larger. The ITD without corrections worked almost as well as the ITD + R for predicting total volume by species, but the stem density results were significantly worse.

The results for the established inventory variables are presented in Table 9 for Evo. In Evo, the ABA method performed the best for all metrics. The RMSEP value for total standing volume was 204 m^3/ha for the ITD, 128 m^3/ha for ITD + R, and 59 m^3/ha for ABA. For deciduous trees the uncorrected ITD outperformed the ITD + R method on stem density, though not on total volume.

The results for the proxy variables for structural diversity are given in Table 10 for Mikkeli and in Table 11 for Evo. In Mikkeli, the ITD based predictions had $r^2 = 0.15–0.81$ and ITD + R predictions $r^2 = 0.22–0.81$. In Evo the corresponding ranges were $r^2 = 0.12–0.82$ (ITD) and $r^2 = 0.18–0.80$ (ITD + R). Especially mean DBH^2, standard deviation of DBH, and percentage of deciduous trees by volume were well predicted in both locations. However, even the worst performing metric (Shannon index in Evo) for the ITD + R method $r^2 = 0.18$, meaning that the plot level predictions were still correlated at $r = 0.42$ with those calculated from the ground-measured tree maps.

Deep learning. The results for inventory variables using the deep-learning ITD *SegmentAnyTree* are given in Table 12 and Fig. 13. As was the case for the *lmf*-based results, the ABA approach did outperform the ITD and ITD + R results for these variables. While the ITD and ITD + R performed as well for dominant height (RMSEP 1.9 m) and the ITD slightly outperfomed the ITD + R for mean DBH (RMSEP 5.0 cm vs. 5.1 cm), the ITD + R was better for all other variables. The difference was largest for the stem densities, where the RMSEP values [trees/ha] for the ITD are 1233 (all), 355 (pine), 735 (spruce), and 575 (deciduous), while for the ITD + R 583 (all), 287 (pine), 508 (spruce), and 411 (deciduous). For total volume the RMSE is 105 m^3/ha with the ITD only but 80 m^3/ha with the ITD + R, and the bias was significantly smaller (−66 m^3/ha vs. −26 m^3/ha). With the uncorrected ITD the lowest r^2 value was 0.14 (for stem density of spruces), while for the ITD + R the corresponding value was 0.31. While *SegmentAnyTree*-based ITD + R outperformed the *lmf*-based one for total volume and spruce volume, it underperformed for pine and deciduous tree volume.

For the structural diversity indices listed in Table 13, the results are mixed. For mean DBH^2, Gini-coefficient of DBH, and percentage of deciduous trees by volume the ITD + R outperformed the ITD as measured by the RMSE. However, for all the other variables (except for the number of trees with DBH≥40 cm where the performance is the same) the uncorrected ITD did better. The biggest advantage of the ITD seems to be with the Shannon index and the Simpson's index, where the ITD + R appears to bias these values upwards. The differences in size variation (Gini, SD) predictions are relatively small. Furthermore, the SD of height and DBH appear to have remarkably small errors with the ITD alone, and thus it is no surprise that the resampling does not offer further improvements, unlike in the *lmf*-case, where the results were significantly biased downwards due to missing small trees.

4.5. Large-scale application

Fig. 14 shows a small portion of the Evo map around the pond Huhmari (61.1935° N, 25.1324° E). The area here is 184 hectares with a total of 251,535 trees in the tree map. The upper-left panel shows an orthophotograph with the upper-right panel showing the locations of mapped trees overlaid on the image. The lower panels demonstrate how the 16 × 16 m pixels look on the map with total volume on the bottom left-hand side and standard deviation of DBH on the bottom right-hand side as examples.

The full Mikkeli area includes 110.1 million trees with an forested area (squares with at least one tree) of 1839.4 km^2, while the full Evo area map includes 11.7 million trees with a forested area of 77.7 km^2. Six different forest structure indices are shown in Figs. 15 and 16 as examples: stem density (DBH≥5 cm), Shannon index (with three species), and mean squared DBH for Mikkeli and number of trees with DBH≥40 cm, volume of deciduous trees, and standard deviation of DBH for Evo. The full mapping procedure took about one day of computation time on a laptop computer for Mikkeli and a few hours for Evo.

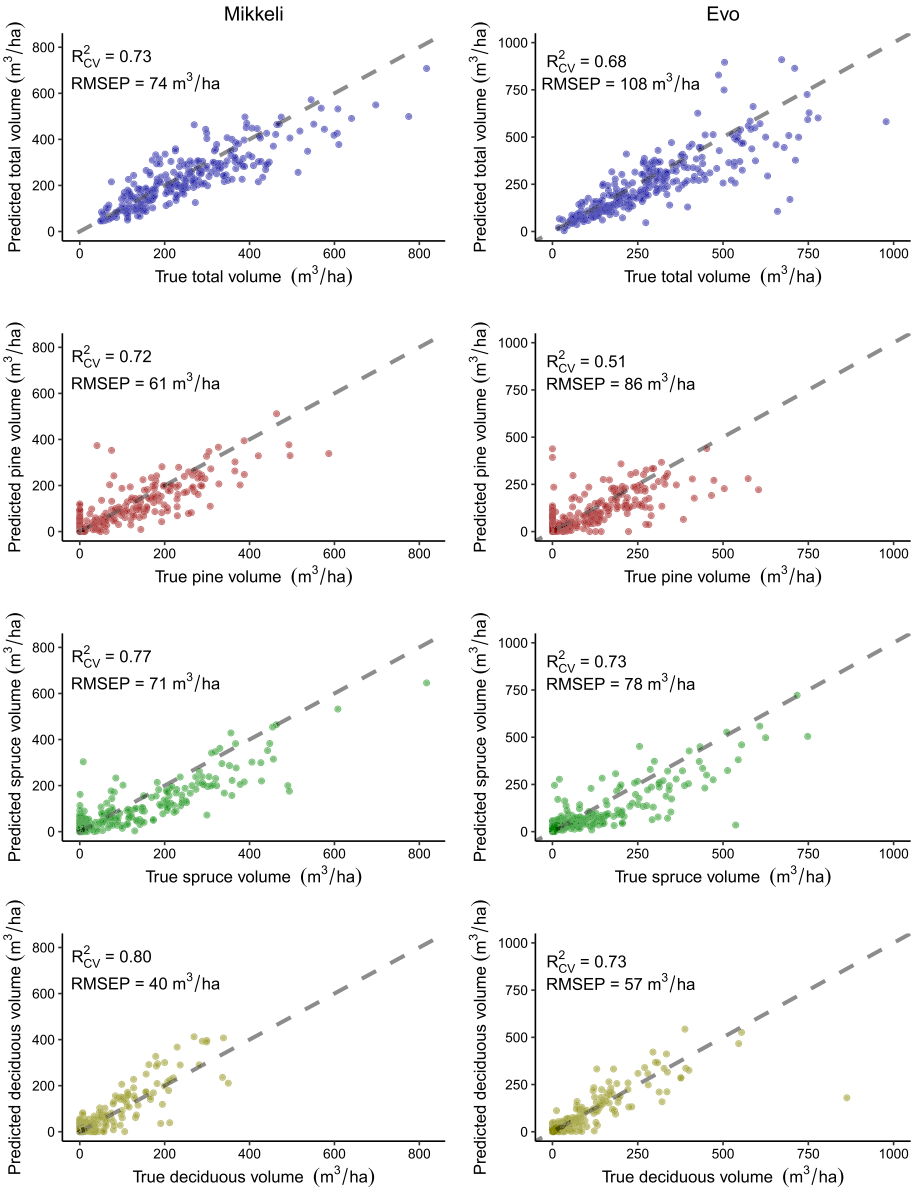

Fig. 12. Plot-level standing volumes for Mikkeli (left) and Evo (right) sites as predicted by the ITD + R (lmf) framework compared to the ground truth. The results are given for all trees (top row) and for by species (other rows).

Table 8
The r^2 values, empirical biases, and RMSEP values for various inventory variables: dominant height, stem density (all trees and by species), basal-area weighted diameter at breast height, and total volume(all trees and by species) for the Mikkeli site. The results are given for the uncorrected ITD (ITD), ITD and resampling framework presented in this work (ITD + R), and the benchmark *k*-NN area-based approach (ABA). The best performing model (smallest RMSEP) is in bold for each metric.

Species	Variable	Mean	ITD			ITD + R			ABA		
			r^2	bias	RMSEP	r^2	bias	RMSEP	r^2	bias	RMSEP
All	Dom. height (m)	23.4	**0.97**	**−0.1**	**0.9**	**0.97**	**−0.1**	**0.9**	0.95	0.0	1.1
	Stem density (#/ha)	939	0.27	−449	675	0.32	28	488	**0.55**	**18**	**385**
	Mean DBH (cm)	23.8	0.74	0.2	3.4	0.72	−1.4	3.8	**0.83**	**0.1**	**2.8**
	Total volume (m^3/ha)	254	0.75	−39	79	0.73	−20	74	**0.82**	**1**	**58**
Pine	Stem density (#/ha)	236	0.56	−50	224	0.53	14	221	**0.72**	**3**	**170**
	Total volume (m^3/ha)	97	0.72	−11	62	0.72	−8	61	**0.75**	**0**	**57**
Spruce	Stem density (#/ha)	437	0.37	−275	462	0.33	0	373	**0.65**	**15**	**267**
	Total volume (m^3/ha)	110	0.81	−33	71	0.77	−21	71	**0.81**	**0**	**59**
Decid.	Stem density (#/ha)	266	0.52	−123	320	0.48	14	297	**0.65**	**0**	**245**
	Total volume (m^3/ha)	47	0.81	6	40	0.80	9	40	**0.80**	**0**	**33**

Table 9
Same as Table 8 for Evo.

Species	Variable	Mean	ITD (*lmf*)			ITD + R			ABA		
			r^2	bias	RMSEP	r^2	bias	RMSEP	r^2	bias	RMSEP
All	Dom. height (m)	22.7	0.93	−0.8	1.7	0.90	−0.8	2.1	**0.93**	**0.1**	**1.6**
	Stem density (#/ha)	1535	0.16	−741	1094	0.47	21	640	**0.76**	**−9**	**428**
	Mean DBH (cm)	23.6	0.74	−0.9	5.1	0.76	−2.4	5.4	**0.83**	**−0.1**	**4.1**
	Total volume (m^3/ha)	287	0.68	−39	114	0.68	−34	108	**0.89**	**0**	**59**
Pine	Stem density (#/ha)	359	0.62	−33	298	0.70	28	264	**0.73**	**−2**	**249**
	Total volume (m^3/ha)	105	0.44	−7	96	0.51	−13	86	**0.68**	**−2**	**68**
Spruce	Stem density (#/ha)	711	0.12	−496	759	0.27	−5	525	**0.65**	**7**	**360**
	Total volume (m^3/ha)	109	0.76	−33	76	0.73	−21	78	**0.83**	**2**	**60**
Decid.	Stem density (#/ha)	465	0.73	−213	427	0.66	−2	373	**0.74**	**−14**	**317**
	Total volume (m^3/ha)	72	0.71	0	62	0.73	0	57	**0.77**	**1**	**52**

Table 10
The r^2 values, empirical biases, and RMSEP values for various proxy variables of structural diversity: mean DBH^2, standard deviation of DBH and height, number of trees with DBH ≥ 40 cm per hectare, Gini coefficient of DBH and height, Shannon index and Simpson's index, and the percentage of deciduous trees by volume for the Mikkeli site. The results are given for the uncorrected ITD (ITD) and ITD and resampling framework presented in this work (ITD + R). The best performing model (smallest RMSEP) is in bold for each metric. Note that due to non-linearity, the relative RMSEP of Mean DBH^2 may seem large, though it is entirely consistent with the errors in mean DBH.

Variable	Mean	ITD (*lmf*)			ITD + R		
		r^2	bias	RMSEP	r^2	bias	RMSEP
Mean DBH^2 (cm^2)	428	0.48	147	262	**0.53**	**−69**	**178**
SD DBH (cm)	7.1	0.37	−3.8	4.6	**0.45**	**0.5**	**2.5**
SD height (m)	5.3	0.16	−2.8	3.8	**0.31**	**1.0**	**2.5**
#DBH ≥ 40 cm (#/ha)	**11**	**0.20**	**−7**	**25**	**0.20**	**−7.5**	**25**
GC DBH	0.22	0.16	−0.14	0.17	**0.23**	**0.03**	**0.09**
GC height	0.18	0.15	−0.11	0.14	**0.16**	**0.05**	**0.10**
Shannon index	0.52	**0.13**	**−0.09**	**0.39**	0.20	0.25	0.41
Simpson's index	0.33	0.11	−0.04	0.27	**0.18**	**0.15**	**0.26**
Deciduous % (by vol.)	0.23	0.82	0.03	0.16	**0.83**	**0.01**	**0.14**

Table 11
Same as Table 10 for Evo.

Variable	Mean	ITD (*lmf*)			ITD + R		
		r^2	bias	RMSEP	r^2	bias	RMSEP
Mean DBH^2 (cm^2)	346	0.49	150	314	**0.60**	**−77**	**187**
SD DBH (cm)	7.8	0.57	−3.0	4.0	**0.60**	**−0.3**	**2.6**
SD height (m)	5.0	0.25	−1.7	2.7	**0.57**	**0.5**	**1.7**
#DBH ≥ 40 cm (#/ha)	26	**0.43**	**−9**	**41**	**0.44**	**−9**	**41**
GC DBH	0.27	0.28	−0.14	0.16	**0.35**	**0.02**	**0.08**
GC height	0.20	0.12	−0.09	0.12	**0.26**	**0.03**	**0.07**
Shannon index	0.63	**0.16**	**−0.04**	**0.34**	0.22	0.24	0.36
Simpson's index	0.40	**0.15**	**−0.02**	**0.23**	**0.20**	**0.15**	**0.23**
Deciduous % (by vol.)	0.26	0.83	0.03	0.15	**0.82**	**0.02**	**0.13**

From Fig. 15 we see that the Mikkeli area is quite evenly distributed w.r.t. stem density. Stem density is for the vast majority of plots between 0–1500 trees/ha with only some individual pixels exceeding 2000 trees/ha. Shannon index is also geographically evenly distributed with values ranging between 0–1.3. Single species stands (Shannon index 0) appear to be quite rare, which may reflect the positive bias seen in the Shannon index predictions. For mean DBH^2, the overwhelming majority of plots has values between 0–1500 cm^2, with only few individual pixels going as high as 4600 cm^2.

In Evo (Fig. 16), there is much clearer spatial variation in the indices. There are large areas with hundreds of trees with DBH≥40 cm in the east and north-east of the mapping area. The volume of deciduous trees is generally low (below 200 m^3/ha) with some individual pixels reaching 600 m^3/ha. The areas with a large number of trees with DBH≥40 cm appear to also be areas with the largest standard deviation of DBH. In general, the standard deviations appear to be either quite very low, or about 20 cm, with individual pixels reaching as high as 40 cm.

5. Discussion

In this work, we extended the ITD + R framework proposed in Kostensalo et al. (2023) for creating structurally realistic tree maps by including tree-level predictions for species, diameter, and volume in addition to the previously developed methods for recreating tree locations and heights. Species composition measured both by the number of individuals as well as by volume is a crucial aspect of biodiversity. We employed a fast *lmf* ITD algorithm as well as a deep-learning based one. Both approaches performed identically in Evo, as measured by the F_1-score. While the deep-learning based approach outperformed for total volume and stem density, the fast ITD resulted in better predictions for deciduous tree volume. Both ITDs were shown to significantly benefit from the resampling corrections when predicting inventory variables, as well as most structural diversity indicators. Specifically, the ITD-based species-distribution was highly biased due to the different detection probabilities of the tree species. In the aggregate, the resampling corrections resulted in a large reduction of bias, though the predictions for individual plots can have significant errors. The resampling corrections also reduced the bias for the volume-based species distribution, but not to the same extent as when counting individual trees, since the smaller trees which the resampling corrects for contribute relatively little to the total volume. The accuracy of stem number prediction was worse for spruce than for pine and deciduous trees. The likely reason for this is that because spruce is a shade tolerant tree species, it is more common in the understory than other species. In the future, ABA could be used when predicting the species of the missing trees, but this would need a complete rethinking of the species prediction process. Though resampling generally corrected the species distribution, for Shannon index it systematically biased the results. This reflects the fact that the method could benefit from introducing a more elaborate correlation structure. ITD + R does not depend on how exactly the ITD algorithm works, and the models for the number of false detections and undetected trees will be always fitted with the chosen ITD in mind. The resampling corrections should work no matter what type of mistakes the chosen ITD algorithm is inclined to make, since the resampled trees are a representative sample of the trees undetectable by the algorithm. It should be stressed, that while structural indices based on ITD + R may have little bias, the location of the simulated trees can in no way correspond to actual trees in the forest and thus should not be used for making, e.g., tree-level management decisions. The methodology aims for a lack of size, species, and spatial bias in large-scale mapping, with sufficient local accuracy for stand-level management decisions. For smaller areas much higher density laser scanning data (TLS, MLS, drone-based ALS) can be used reducing the number of missed trees. However, the need for resampling

Table 12
Same as Table 8 for Evo.

Species	Variable	Mean	ITD (*SegmentAnyTree*)			ITD + R			ABA		
			r^2	bias	RMSEP	r^2	bias	RMSEP	r^2	bias	RMSEP
All	Dom. height (m)	22.7	0.92	−1.1	1.9	0.92	−1.0	1.9	**0.93**	**0.1**	**1.6**
	Stem density (#/ha)	1535	0.16	−924	1233	0.57	16	583	**0.76**	**−9**	**428**
	Mean DBH (cm)	23.6	0.78	1.0	5.0	0.77	−1.3	5.1	**0.83**	**−0.1**	**4.1**
	Total volume (m^3/ha)	287	0.80	−66	105	0.83	−26	80	**0.89**	**0**	**59**
Pine	Stem density (#/ha)	359	0.69	−156	355	0.65	35	287	**0.73**	**−2**	**249**
	Total volume (m^3/ha)	105	0.46	−23	94	0.50	−8	89	**0.68**	**−2**	**68**
Spruce	Stem density (#/ha)	711	0.14	−461	735	0.31	−21	508	**0.65**	**7**	**360**
	Total volume (m^3/ha)	109	0.83	−28	66	0.82	−13	63	**0.83**	**2**	**60**
Decid.	Stem density (#/ha)	465	0.60	−307	575	0.55	2	411	**0.74**	**−14**	**317**
	Total volume (m^3/ha)	72	0.60	−15	70	0.60	−5	69	**0.77**	**1**	**52**

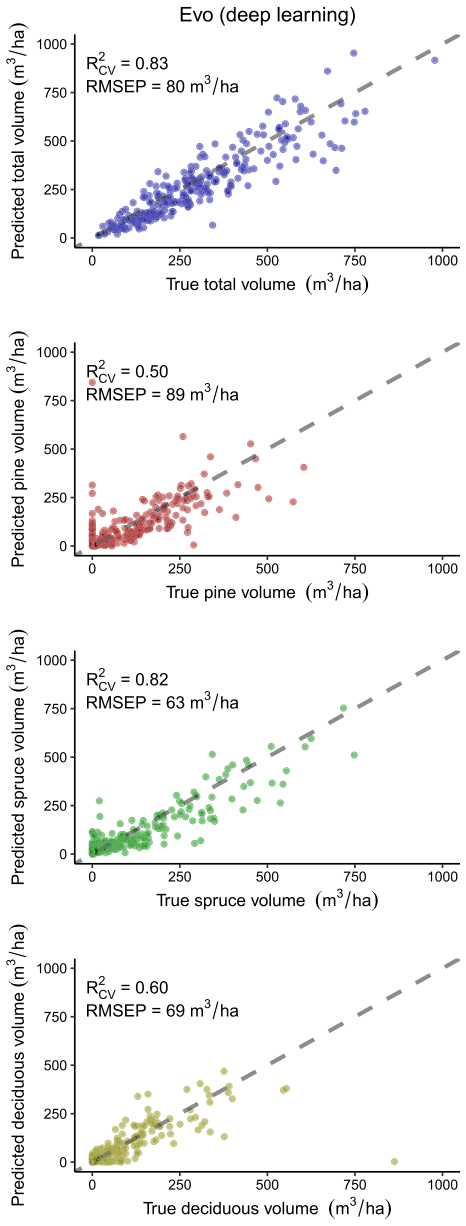

Fig. 13. Plot-level standing volumes for Mikkeli (left) and Evo (right) sites as predicted by the ITD + R (*SegmentAnyTree*-based) framework compared to the ground truth. The results are given for all trees (top row) and for by species (other rows).

Table 13
Same as Table 10 for Evo using the deep-learning-based ITD *SegmentAnyTree*.

Variable	Mean	ITD (*SegmentAnyTree*)			ITD + R		
		r^2	bias	RMSEP	r^2	bias	RMSEP
Mean DBH^2 (cm^2)	346	0.69	−169	287	**0.68**	**41**	**163**
SD DBH (cm)	7.8	**0.74**	**0.8**	**2.4**	0.65	−0.2	2.6
SD height (m)	5.0	**0.71**	**0.3**	**1.3**	0.61	−0.3	1.4
#DBH ≥ 40 cm (#/ha)	26	**0.41**	**−1**	**46**	**0.41**	**−1**	**46**
GC DBH	0.27	0.62	0.08	0.10	**0.33**	**−0.02**	**0.08**
GC height	0.20	**0.52**	**0.03**	**0.06**	0.28	−0.03	0.07
Shannon index	0.63	**0.26**	**0.12**	**0.25**	0.11	0.29	0.40
Simpson's index	0.40	**0.28**	**0.05**	**0.19**	0.13	0.17	0.25
Deciduous % (by vol.)	0.26	0.79	0.04	0.16	**0.77**	**0.01**	**0.14**

corrections cannot be completely solved with higher pulse densities and TLS, as even with over 1000 pulses/m^2 Wielgosz et al. (2024) report omission rates exceeding 20% and commission errors exceeding 10%.

ABA performed equally well or better than ITD + R for inventory variables and should be preferred when specific attributes and the exact grid in which they are to be calculated are known. However, ITD + R offers valuable flexibility when predicting forest structure indices (Rajala et al., 2025). For example, with a tree map it is easy to compute variables for stands which are of arbitrary shape. Furthermore, the mapping approach can be valuable for conducting complex simulations of, e.g., wind and insect damages or thinning operations.

It should be noted that the ground measured tree locations are based on the location of the trunk at breast height, while as the lmf-based tree locations indicate the location of the tree top. Further uncertainties relate to the accuracy of the field equipment. With the equipment used in Mikkeli, the empirical RMSEP is 0.325 m, with a maximal observed deviation of 1.9 m (Luopa, 2022). Thus, in principle a stricter requirement of, e.g., 2 m could be used when matching trees, but in Evo the location accuracy was worse, as only about 12% of the trees were matched with such a strict limit. The tree location matching limit should be thought of as a hyperparameter: by opting for a stricter limit, one risks labeling true trees as false discoveries, while a looser limit has the opposite problem. The optimal value will depend on the quality of the models predicting the number of false detections and undetected trees. One additional problem is displacement due to relief in the hyperspectral data due to which combining ALS data and hyperspectral data at the tree level is not exact. Alternatively, one could use multispectral lidar, which eliminates the data fusion issue.

ITD + R was applied in this work to two areas with the larger one having approximately 110 million trees, covering over 3000 km^2, and taking about one day to compute on regular laptop computer. Thus, the methods are easily scalable for country-wide mapping. In Finland, practically the entire land area has been recently mapped with 5 pts/m^2 ALS, and there are a lot of field data, e.g., national forest inventory

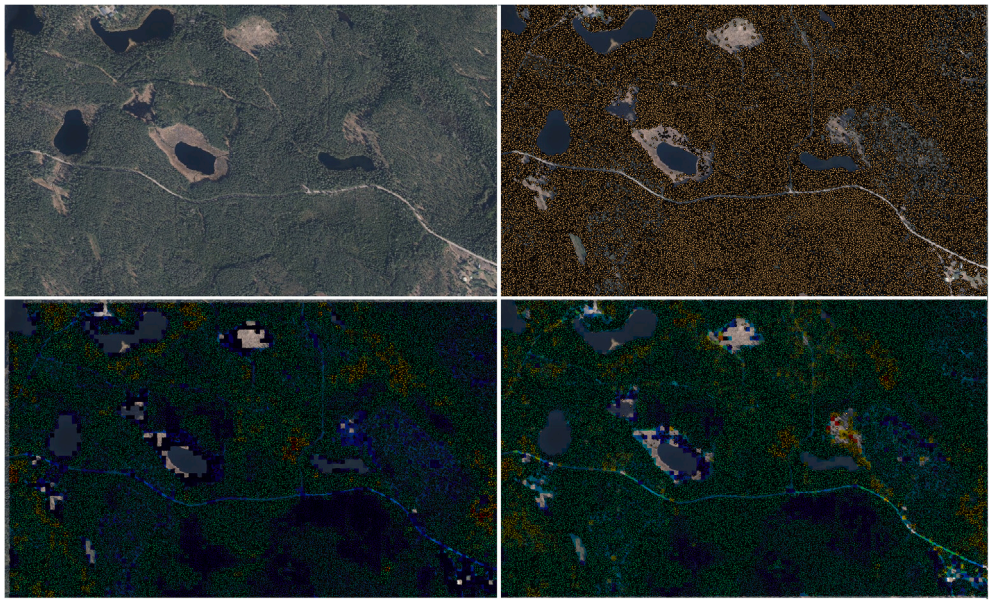

Fig. 14. A small portion of the Evo map around the pond Huhmari (61.1935° N, 25.1324° E). Orthoimage of the area [upper-left], overlaid with locations of the trees [upper-right], overlaid by a 16 × 16 m raster for total volume (m^3/ha) [bottom-left], and a 16 × 16 m raster for SD of DBH (cm) [bottom-right].

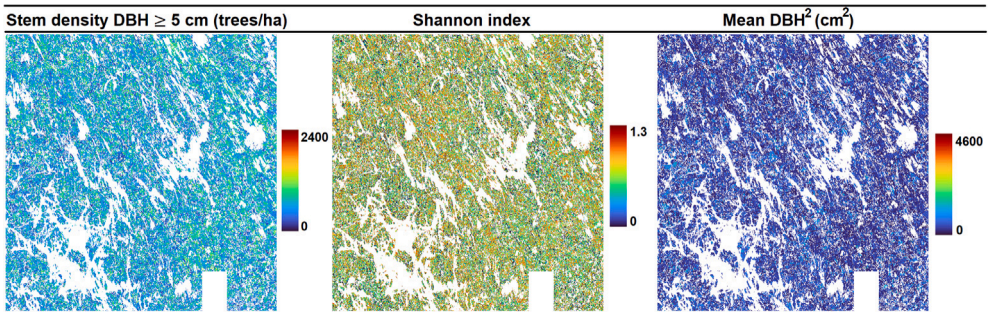

Fig. 15. 16 × 16 m rasters for forest structure variables for the full Mikkeli area. The subfigures show the stem density [left], Shannon index [center], and mean of squared DBH [right].

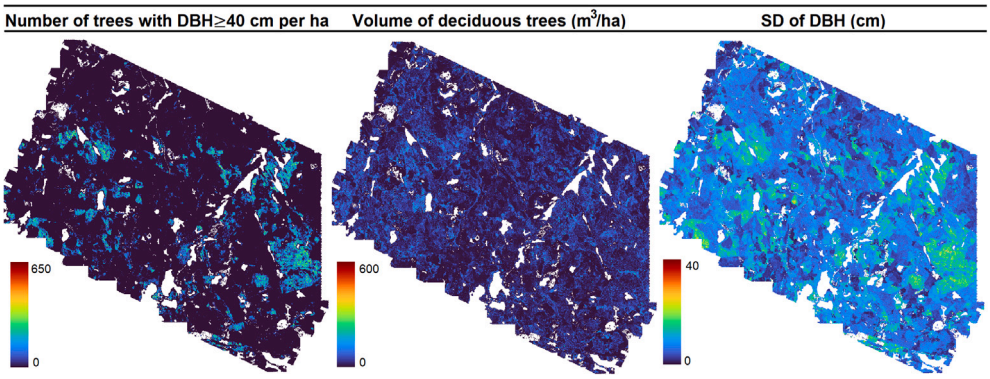

Fig. 16. 16 × 16 m rasters for forest structure variables for the full Evo area. The subfigures show the number of trees with DBH≥40 cm [left], the standing volume of deciduous trees [center], and the standard deviation of DBH [right].

(NFI) plots. The methods should work in the Boreal zone more generally, if data of similar quality are available, but may be unreliable in different climate conditions, as much higher species richness and denser forests may complicate things. In Evo, with an average stem density over 1500 trees/ha the results were poorer than in Mikkeli where the stem density is 900 trees/ha, even though the ALS pulse density was higher and hyperspectral data were available. Thus, in climatic conditions resulting in very dense forests performance may suffer. In our data sets, the canopy is not very dense, and especially Norway spruces have distinctive tops, which are easy to detect. If the canopy were extremely dense and even, ALS-based ITD might fail in ways which are not easily correctable with resampling. ITD + R requires field data with location,

species, height, and DBH measured for each tree, but some inventory campaigns do not collect only a subset of these data. Species and DBH are usually collected, and height is measured for a subset of trees and estimated for the rest to a sufficient accuracy using localized models. However, locations are not always recorded, and when they are, accuracy may vary. For example, the positioning system used in Mikkeli gives highly accurate locations, but such systems are not currently common. However, it would be possible to include such measurements in NFIs, at least for some subset of the plots.

The next steps for ITD + R are to include a more exhaustive species detection and integrate status prediction (dead/alive). The species relation matrix allows for an arbitrarily many species, and a similar approach could be extended for status. However, for rare species false detection rate is a problem that needs to be carefully considered. Further tree attributes could also be easily included, as long as they can be predicted for the ITD trees and are available for all ground measured trees.

6. Conclusions

In this work, we propose a species-type-specific version of a resampling-based framework, ITD + R, for creating structurally less biased large-scale tree maps (millions or even billions of trees), which retain sufficient local accuracy at the stand level. The resampling-based approach has been shown to significantly reduce bias in the species composition of ITD tree maps, though more research is needed to better account for the spatial correlation of tree species as well as rare species.

With point-densities 5–10 echoes/m^2, which are at the upper end of what is used for complete national campaigns, the bottleneck is not in the tree detection algorithms, but rather in the scarcity of below-canopy data, which necessitates compensation for small trees. The proposed methodology needs ground measurements with precise tree locations, heights and other key attributes, which may not currently be available in other geographic areas. Thus, we encourage the gathering of such data as a part of remote sensing campaigns. When applying to areas with higher species richness, grouping species which are difficult to discern from remote sensing data may be necessary.

CRediT authorship contribution statement

J. Kostensalo: Writing – original draft, Visualization, Validation, Software, Methodology, Funding acquisition, Formal analysis, Conceptualization. **P. Packalen:** Writing – review & editing, Validation, Methodology, Formal analysis. **M. Kuronen:** Writing – review & editing, Software, Resources, Data curation. **L. Mehtätalo:** Writing – review & editing. **S. Tuominen:** Writing – review & editing, Visualization, Data curation. **M. Myllymäki:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research has been funded by the European Union – NextGenerationEU in the Research Council of Finland's project (Grant numbers 348154, 352782) under the Research Council of Finland's flagship ecosystem for Forest–Human–Machine Interplay—Building Resilience, Redefining Value Networks and Enabling Meaningful Experiences (UNITE) (357909). This research has also received funding through the Research Council of Finland's grant number 361209. The Finnish Forest Centre provided the field data in Mikkeli. Acquisition of remote sensing and field data in the Evo region was financed by Finland's Ministry of Agriculture and Forestry / key project "Wood on the move and new products from forests".

Appendix A. Area-based prediction using *k*-NN imputation

Response variables

Plot level attributes were used as response variables. Volume, stem number, basal area weighted mean diameter at breast height and dominant height were predicted by pooling all tree species together (total attributes). Volume and stem number were also predicted by tree species type for pines, spruces and deciduous trees as a species group.

Predictor variables

Corresponding to response variables, predictor variables were computed at the plot level. ALS metrics were computed similarly in both study areas. In Mikkeli study area, spectral metrics were computed from aerial images whereas in Evo study area, spectral metrics were computed from hyperspectral images.

ALS metrics were computed separately for categories: first echoes ("first of many" + "only"), last echoes ("last of many" + "only") and all echoes ("first of many" + "last of many" + "intermediate" + "only"). Metrics were means, medians, standard deviations, kurtoses, and skewnesses of height, height quantiles (20/40/50/60/70/80/90/95) and covers (2 m, 5 m, 10 m, 15 m, 20 m). The quantiles were computed as explained in Hyndman and Fan (1996) and cover metrics were computed as the ratio of the numbers of echoes over certain height threshold, divided by the total number of echoes in the same echo category. Before computing any metric negative heights were set to zero. We also computed mean and standard deviations of lidar intensities for the first echoes category and the proportion of echoes by echo category.

Image metrics were computed from unrectified aerial images. First pixel values were linked to ALS echoes using known internal and external orientation parameter values (see Section 2.3.2 of the article) and then metrics were computed from pixel values of echoes above 2 m above the ground level. Only first echoes were used at this stage. Calculated spectral metrics were mean and standard deviation for each band, ratios between the bands and normalised difference vegetation index (NDVI).

Hyperspectral metrics were computed from orthorectified hyperspectral images. Calculated metrics were bandwise mean values (186 VNIR bands and 288 SWIR bands) and three vegetation indices: simple ratio (SR), difference vegetation index (DVI) and NDVI. Bands used to compute vegetation indices are given in Sun et al. (2021).

Variable selection

Predictor variables were selected to k-NN models using a heuristic algorithm VSSA proposed by Packalén et al. (2012). The principle is to minimize the squared error loss by solving the k-NN model repeatedly with a slightly modified set of predictor variables. The VSSA minimizes the mean of relative RMSE values over all response variables. It uses simulated annealing for approximating the global optimum (Kirkpatrick et al., 1983).

Numerous values of hyperparameters were tested in optimization using the principle of grid search. The number of predictor variables in the k-NN model varied between 5–12 and the initial temperature varied between 0.5 and 1.5 (see Packalén et al., 2012). 1/3 of the set of predictor variables was initially modified. The number of iterations (niter parameter in VSSA) was set to 3000 in univariate cases and to 30,000 in multivariate cases.

Nearest neighbor imputation

Tree species-specific attributes were predicted in a multivariate manner, i.e. all species-specific values were taken from the same nearest neighbors. This was done, however, separately for volume and stem number. Total attributes were predicted variable-by-variable in a univariate manner. The distance metric was MSN (Moeur and Stage, 1995) in multivariate imputation and a variant of weighted Euclidean distance

(NNSA in Packalén et al., 2012) in univariate case. We evaluated models using varying number of nearest neighbors (k). In final imputation models k varied between 4 and 6. The imputed value was the average of the values of k nearest neighbors. After preliminary testing, the initial temperature was set to 0.5 and the number of iterations to 150 in NNSA Packalén et al. (see details in 2012).

Prediction

Prediction was implemented using 10-fold cross-validation as in the ITD + R approach. Species-specific predictions of volume and stem number were calibrated such that they correspond with the separately predicted total values:

$$\hat{Y}_{a,s} = \frac{\hat{y}_{a,\text{total}}}{\hat{y}_{a,\text{pine}} + \hat{y}_{a,\text{spruce}} + \hat{y}_{a,\text{deciduous}}} \times \hat{y}_{a,s} \quad (7)$$

where $\hat{Y}_{a,s}$ is the calibrated value of attribute a (volume or stem number) and tree species type s (pine, spruce, deciduous trees), $\hat{y}_{a,s}$ is the original predicted value.

Appendix B. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.rse.2025.115223.

Data availability

Data will be made available on request.

References

Axelsson, P., 2000. DEM generation from laser scanner data using adaptive tin models. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 33, 110–117.

Ene, L., Næsset, E., Gobakken, T., 2012. Single tree detection in heterogeneous boreal forests using airborne laser scanning and area-based STEM number estimates. Int. J. Remote Sens. 33 (16), 5171–5193.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. Science 342 (6160), 850–853.

Ho, T.K., 1995. Random decision forests. In: Proceedings of 3rd International Conference on Document Analysis and Recognition, vol. 1, IEEE, pp. 278–282.

Hyndman, R.J., Fan, Y., 1996. Sample quantiles in statistical packages. Am. Stat. 50 (4), 361–365. Publisher: [American Statistical Association, Taylor & Francis, Ltd].

Hyyppä, J., Kelle, O., Lehikoinen, M., Inkinen, M., 2001. A segmentation-based method to retrieve STEM volume estimates from 3-D tree height models produced by laser scanners. IEEE Trans. Geosci. Remote Sens. 39 (5), 969–975.

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B.M., Wu, J.-C., 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. Remote Sens. 4 (4), 950–974.

Kangas, A., Pitkänen, T.P., Mehtätalo, L., Heikkinen, J., 2022. Mixed linear and non-linear tree volume models with regional parameters to main tree species in Finland. For. Int. J. For. Res. 96 (2), 188–206.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220 (4598), 671–680. Publisher: American Association for the Advancement of Science.

Kostensalo, J., Mehtätalo, L., Tuominen, S., Packalen, P., Myllymäki, M., 2023. Recreating structurally realistic tree maps with airborne laser scanning and ground measurements. Remote Sens. Environ. 298, 113782.

Kotivuori, E., Maltamo, M., Korhonen, L., Strunk, J.L., Packalen, P., 2021. Prediction error aggregation behaviour for remote sensing augmented forest inventory approaches. For. Int. J. For. Res. 94 (4), 576–587.

Lechner, A.M., Foody, G.M., Boyd, D.S., 2020. Applications in remote sensing to forest Ecology and management. One Earth 2 (5), 405–412.

Luopa, T., 2022. Puuston Sijaintitarkkuus Metsäkeskuksen Puukarttakoealoilla.

Magnussen, S., Næsset, E., Gobakken, T., 2013. Prediction of tree-size distributions and inventory variables from cumulants of canopy height distributions. For. Int. J. For. Res. 86 (5), 583–595.

Maltamo, M., Packalen, P., Kangas, A., 2021. From comprehensive field inventories to remotely sensed wall-to-wall stand attribute data — a brief history of management inventories in the nordic countries. Can. J. For. Res. 51 (2), 257–266. Publisher: NRC Research Press.

Mehtätalo, L., Lappi, J., 2020. Biometry for Forestry and Environmental Data: With Examples in R, Applied environmental statistics, CRC Press/Taylor & Francis Group.

Mehtätalo, L., Yazigi, A., Kansanen, K., Packalen, P., Lähivaara, T., Maltamo, M., Myllymäki, M., Penttinen, A., 2022. Estimation of forest stand characteristics using individual tree detection, stochastic geometry and a sequential spatial point process model. Int. J. Appl. Earth Obs. Geoinf. 112, 102920.

Mevik, B.-H., Wehrens, R., 2007. The PLS package: principal component and partial least squares regression in R. J. Stat. Softw. 18 (2), 1–23.

Mikhail, E., Bethel, K., McGlone, J., 2001. Introduction to Modern Photogrammetry. John Wiley & Sons, Ltd, New York.

Moeur, M., Stage, A.R., 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. For. Sci. 41 (2), 337–359.

Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. 80 (1), 088–99.

Ørka, H.O., Jutras-Perreault, M.-C., Næsset, E., Gobakken, T., 2022. A framework for a forest ecological base map – an example from Norway. Ecol. Indic. 136, 108636.

Packalén, P., Suvanto, A., Maltamo, M., 2009. A two stage method to estimate species-specific growing stock by combining als data and aerial photographs of known orientation parameters. Photogramm. Eng. Remote Sens. 75, 1451–1460.

Packalén, P., Temesgen, H., Maltamo, M., 2012. Variable selection strategies for nearest neighbor imputation methods used in remote sensing based forest inventory. Can. J. Remote Sens. 38 (5), 557–569. Publisher: Taylor & Francis _eprint: https://doi.org/10.5589/m12-046

Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C., Coops, N.C., Dulloo, E., Faith, D.P., Freyhof, J., Gregory, R.D., Heip, C., Höft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P.W., Stuart, S.N., Turak, E., Walpole, M., Wegmann, M., 2013. Essential biodiversity variables. Science 339 (6117), 277–278.

Peuhkurinen, J., Mehtätalo, L., Maltamo, M., 2011. Comparing individual tree detection and the area-based statistical approach for the retrieval of forest stand characteristics using airborne laser scanning in Scots pine stands. Can. J. For. Res. 41 (3), 583–598. Publisher: NRC Research Press.

Pitkänen, J., Maltamo, M., Hyyppä, J., Yu, X., 2004. Adaptive methods for individual tree detection on airborne laser based canopy height model. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 36, 187–191.

Popescu, S., Wynne, R., 2004. Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogramm. Eng. Remote Sens. 70, 589–604.

R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rajala, T., Kangas, A., Myllymäki, M., 2025. Computing maps of forest structural diversity: aggregate late. Ecol. Indic. 178, 114046.

Roussel, J.-R., Auty, D., Coops, N.C., Tompalski, P., Goodbody, T.R.H., Meador, A.S., Bourdon, J.-F., de Boissieu, F., Achim, A., 2020. Lidr: an r package for analysis of airborne laser scanning (ALS) data. Remote Sens. Environ. 251, 112061.

Silva, C.A., Hudak, A.T., Vierling, L.A., Loudermilk, E.L., O'Brien, J.J., Hiers, J.K., Jack, S.B., Gonzalez-Benecke, C., Lee, H., Falkowski, M.J., Khosravipour, A., 2016. Imputation of individual longleaf pine (pinus palustris mill.) tree attributes from field and lidar data. Can. J. Remote Sens. 42 (5), 554–573.

Staudhammer, C.L., LeMay, V.M., 2001. Introduction and evaluation of possible indices of stand structural diversity. Can. J. For. Res. 31 (7), 1105–1115.

Storch, F., Dormann, C.F., Bauhus, J., 2018. Quantifying forest structural diversity based on large-scale inventory data: a new approach to support biodiversity monitoring. For. Ecosyst. 5 (34).

Sun, G., Jiao, Z., Zhang, A., Li, F., Fu, H., Li, Z., 2021. Hyperspectral image-based vegetation index (HSVI): a new vegetation index for urban ecological research. Int. J. Appl. Earth Obs. Geoinf. 103, 102529.

Toivonen, J., Kangas, A., Maltamo, M., Kukkonen, M., Packalen, P., 2023. Assessing biodiversity using forest structure indicators based on airborne laser scanning data. For. Ecol. Manag. 546, 121376.

Wielgosz, M., Puliti, S., Xiang, B., Schindler, K., Astrup, R., 2024. SegmentAnyTree: a sensor and platform agnostic deep learning model for tree segmentation using laser scanning data. Remote Sens. Environ. 313, 114367.

Wold, S., Johansson, E., Cocchi, M., et al., 1993. PLS: partial least squares projections to latent structures. In: 3D QSAR in Drug Design: Theory, Methods and Applications. Kluwer ESCOM Science Publisher, pp. 523–550.

Wright, M.N., Ziegler, A., 2017. Ranger: a fast implementation of random forests for high dimensional data in C++ and R. J. Stat. Softw. 77 (1), 1–17.

Yu, X., Hyyppä, J., Holopainen, M., Vastaranta, M., 2010. Comparison of area-based and individual tree-based methods for predicting plot-level forest attributes. Remote Sens. 2 (6), 1481–1495. Number: 6 Publisher: Molecular Diversity Preservation International.