



# A new method for individual treetop detection with low-resolution aerial laser scanned data

Gergő Diószegi<sup>1</sup> · Vanda Éva Molnár<sup>2,3</sup> · Loránd Attila Nagy<sup>1</sup> · Péter Enyedi<sup>4</sup> · Péter Török<sup>3,5</sup> · Szilárd Szabó<sup>1</sup>

Received: 9 March 2024 / Accepted: 2 June 2024 / Published online: 13 June 2024  
© The Author(s) 2024

## Abstract

In the past decade, the use of three-dimensional forest information from airborne Light Detection and Ranging (LiDAR) has become widespread in forest inventories. Accurate Individual Treetop Detection (ITD) and crown boundary delineation using LiDAR data are critical for obtaining precise inventory metrics. To address this need, we introduced a novel growing tree region (GTR)-driven ITD method that utilizes canopy height models (CHM) derived from very low-resolution airborne LiDAR data. The GTR algorithm consists of three key stages: (i) preserving all height layers through incremental cutting and stacking of CHM; (ii) employing a three-layer concept to identify individual treetops; and (iii) refining the detected treetops using a distance-based filter. Our method was tested in five temperate forests across Central Europe and was compared against the widely-used local maxima (LM) search combined with an optimized variable window filtering (VWF) technique. Our results showed that the GTR method outperformed LM with VWF, particularly in forests with high canopy density. The achieved root mean square accuracies were 74% for the matching rate, 19% for commission errors, and 27% for omission errors. In comparison, the LM with the VWF method resulted in a matching rate of 71%, commission errors of 20%, and omission errors of 31%. To facilitate the application of our algorithm, we developed an R package called TREETOPS, which seamlessly integrates with the lidR package, ensuring compatibility with existing treetop-based segmentation methods. By introducing TREETOPS, we provide the most accurate open-source tool for detecting treetops using low-resolution LiDAR-derived CHM.

**Keywords** CHM-based treetop detection · Growing tree region · Local maxima · Variable window filtering · Low-resolution LiDAR · Central European forest · R

## Introduction

The use of Individual Treetop Detection (ITD) and crown delineation via active remote sensing technology, particularly Light Detection and Ranging (LiDAR) with Airborne Laser Scanning (ALS) data, has become essential for advancing our understanding of forest structure and ecology (Koch et al. 2006; White et al. 2016; Wulder et al. 2008). Over the past two decades, ALS technology has become essential for advancing our understanding of forest structure and ecology (Hyypä et al. 2008; Lim et al. 2003; White et al. 2016). The primary goal of forest management is to obtain precise individual tree metrics, including tree species, location, height, diameter at breast height (DBH), and crown dimensions (González-Ferreiro et al. 2012; Hu et al. 2014; Koch et al. 2006; Maltamo and Gobakken 2014). These

✉ Gergő Diószegi  
dioszegi.gergo@science.unideb.hu

<sup>1</sup> Department of Physical Geography and Geoinformatics, Faculty of Sciences and Technology, University of Debrecen, Egyetem Tér 1., Debrecen, Hungary

<sup>2</sup> ELKH-DE Anthropocene Ecology Research Group, University of Debrecen, Egyetem Tér 1., Debrecen, Hungary

<sup>3</sup> Department of Ecology, Faculty of Sciences and Technology, University of Debrecen, Egyetem Tér 1., Debrecen, Hungary

<sup>4</sup> Envirosense Hungary Ltd., Péchy M. U. 46, Debrecen, Hungary

<sup>5</sup> HUN-REN-UD Functional and Restoration Ecology Research Group, Egyetem Tér 1., Debrecen, Hungary

metrics are crucial for estimating forest characteristics, such as tree species composition, growing stock volume, canopy density, and mean basal area (Ene et al. 2012; Lee et al. 2013; Parkitna et al. 2021; Unger et al. 2014). However, the accuracy of these parameters is influenced by errors in various individual tree detection methods, including treetop detection, crown delineation, and tree segmentation. Consequently, developing the most accurate algorithms for these tasks remains a significant challenge.

Numerous studies have proposed various ITD methods for scientific and operational applications (Brososke et al. 2014; Coops et al. 2021; Popescu and Wynne 2004; Wulder et al. 2000). Technically, ITD and crown delineation methods follow two main concepts. The point-based concept utilizes laser scanned data with high and very high point densities to detect and segment individual trees directly from the point cloud by clustering points into objects (Duncanson et al. 2014; Li et al. 2012; Yao et al. 2014). The raster-based concept employs low-resolution LiDAR data (low point density). Initially, it used the point-to-raster algorithm to create a digitalized square unit (pixel) representation of the forest surface in a canopy height model (CHM) (Popescu et al. 2002; Popescu and Wynne 2004). Following this, treetops are detected using the local maxima (LM) algorithm with a fixed or variable window filtering (VWF) size (Pitkänen 2001; Popescu and Wynne 2004; Wulder et al. 2000). Finally, marker-controlled watershed segmentation or marker-controlled decision tree segmentation, utilizing treetops as markers, is applied for tree boundary delineation (i.e., crown) (Beucher and Meyer 1993; Dalponte and Coomes 2016; Lamar et al. 2005; Wang et al. 2004). The raster-based concept is essentially a CHM-based method.

Traditionally, computing the CHM involves subtracting the Digital Terrain Model (DTM) from the Digital Surface Model (DSM). However, it is crucial to approach CHM processing with care, as different retrieval methods can produce qualitatively distinct representations of tree heights compared to in-situ measured tree heights (Mielcarek et al. 2018). For optimal CHM accuracy, triangulation-based algorithms are recommended because they provide the most faithful representations of trees within CHMs (Mielcarek et al. 2018; Roussel et al. 2020). However, triangulation-based model generation may introduce empty or erroneous pixels, which are often referred to as pits. This can be effectively mitigated by employing various post-processing filters (Roussel et al. 2020; Stereńczak et al. 2020). Additionally, to reduce the number of Local Maxima (LM) a priori, Gaussian filtering with varying kernel window sizes, allowing for adjustment of the smoothing effect, can be applied to refine the model (Stereńczak et al. 2020). The processed CHM can then undergo ITD.

Eysn et al. (2015) conducted a thorough comparison of several CHM-based ITD methods using aerial laser scanned data collected in Central European alpine forests. They also

introduced an automated matching procedure with highly accurate evaluation metrics for ITD (Eysn et al. 2015; Stereńczak et al. 2020). Their findings revealed that the Local Maxima (LM) search with Variable Window Filtering (VWF) achieved the best results, with a matching rate of 60% for single-story coniferous stands and 47% for single-layered mixed forests with a 29% coniferous proportion. However, detecting co-dominant trees in dense forests poses challenges for all considered methods. Previous benchmarking studies (Kaartinen et al. 2012; Vauhkonen et al. 2012) have reported similar results. A novel ITD and segmentation approach using a CHM-based hierarchical transformation of height levels from highest to lowest demonstrated accuracies between 83 and 86% in various forest stands (Zhao et al. 2017). Furthermore, a sophisticated, self-calibrating segmentation algorithm (without ITD), tested in a Central European lowland forest, achieved accuracies of 85% in coniferous and deciduous stands and 75% in mixed forests (Stereńczak et al. 2020). However, these high-performance tools were not released under open-source licensing. Therefore, there is a critical demand for an ITD approach that (1) is freely accessible, (2) has the potential to surpass the LM with VWF and can be seamlessly integrated with different segmentation approaches, and (3) can be easily developed.

This study aimed to introduce a novel algorithm for ITD developed in Central European mixed forests, utilising low-resolution LiDAR-derived CHM. The specific objectives of this study were as follows:

- (i). Develop a straightforward yet flexible CHM-based algorithm for treetop detection.
- (ii). Create an open-source tool that seamlessly integrates this algorithm, which is compatible with the `lidR` package in the R computational environment.
- (iii). Assess and validate the accuracy of the proposed algorithm by utilizing reference data collected in situ from one forest and derived from combined laser scanning data from four other forests. This evaluation includes a performance analysis comparing our algorithm with the widely used treetop detection method that employs an LM search with an optimized VWF.

## GTR

### The algorithm

Our proposed method, called the Growing Tree Region (GTR), is a treetop localization algorithm based on the LiDAR-derived CHM. The algorithm was implemented in the R programming language and relied heavily on the C++-based `terra` (Hijmans 2023) and `sf` (Pebesma and Bivand 2023) packages. Therefore, the terminology used in this study was adjusted to align with the terminology used by those packages.

Similar to the LM algorithm using VWF, our algorithm conceptualizes tree crowns as mountains protruding from a three-dimensional canopy surface (i.e., CHM). The apex of a tree corresponds to the peak of a mountain. Height values decrease continuously starting from the treetop, following the slope of the mountain, which represents the tree crown. This descent stops at the valley, which can indicate either the boundary between the trees or the edge of the crown.

The algorithm comprises three major steps: (i) CHM cutting and storage, (ii) treetop location, and (iii) reduction in the number of treetops.

(i). CHM cutting and storage

The GTR algorithm employs a horizontal plane to vertically slice the CHM from top to bottom into multiple horizontal raster layers, where non-no data values are labelled as 1 (Fig. 1). The files were stored in a stacked format. The vertical cutting distance, referred to as the height increment, is set at 0.2 m. Consequently, each raster represents the CHM values (1) at a specific height, with the height information (Z) stored in the filename of the corresponding raster file within the stack.

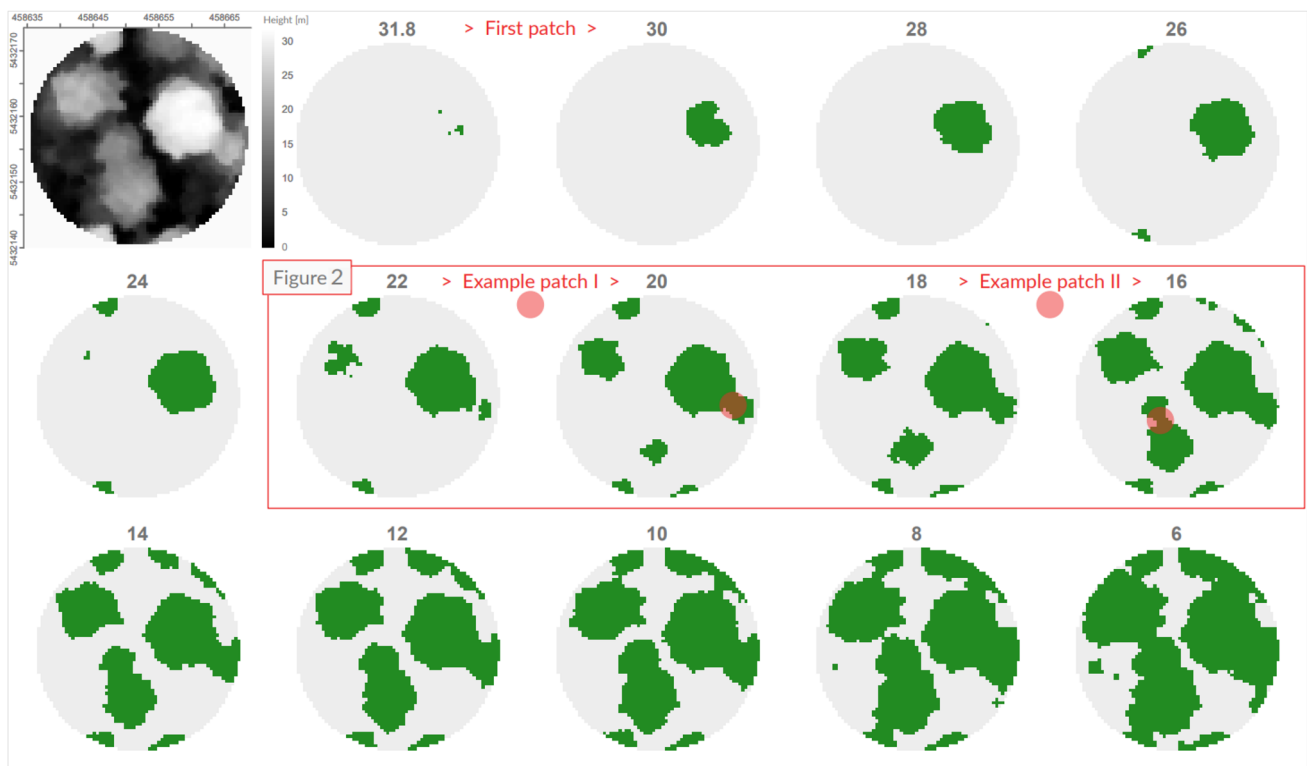
During the tree-growing process, sections of growth intersect, forming a new patch at a certain height. In our example, the first new patch (i.e., intersection) occurs between heights of 31.8 m and 30 m (Fig. 1).

(ii). Treetop location

GTR identification is a crucial step for locating treetops. The process is iterative and involves the selection of three consecutive layers (from top to bottom) at each iteration (Fig. 2). During one iteration, the algorithm searches for new intersections and identifies GTR. Treetops are extracted by determining the centroids of the GTR calculated from layers 2 and 3 and then filtering them based on the criterion of overlapping with the GTR from layers 1 and 2 (Fig. 2). The three-layer concept is explained below.

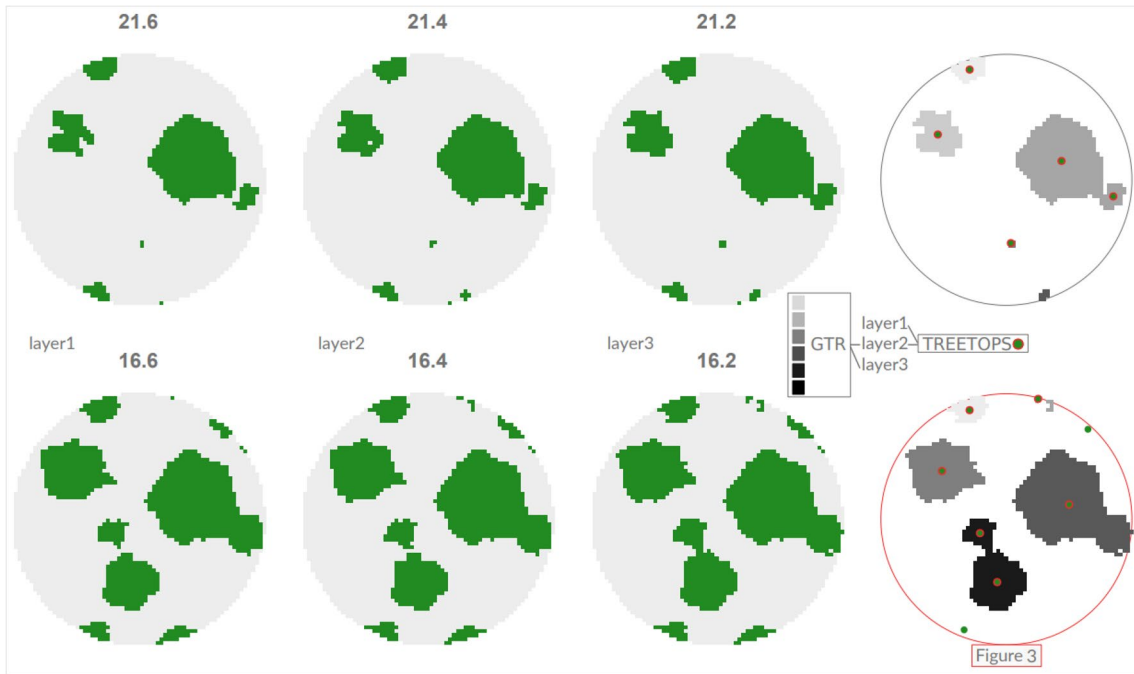
First, the first emerging tree region (FETR) is determined by subtracting layer1 from layer2 (FETR21), and the operation of layer3—layer2 results in FETR32 (Fig. 3). Subsequently, GTR21 and GTR32 were computed. GTR21 was calculated by utilizing layer2 and FETR21 under the following conditions (Eq. 1, Fig. 3):

$$layer2patch_i \supseteq FETR21patch_n \& N_{cell}layer2patch_i \neq N_{cell}FETR21patch_n \tag{1}$$

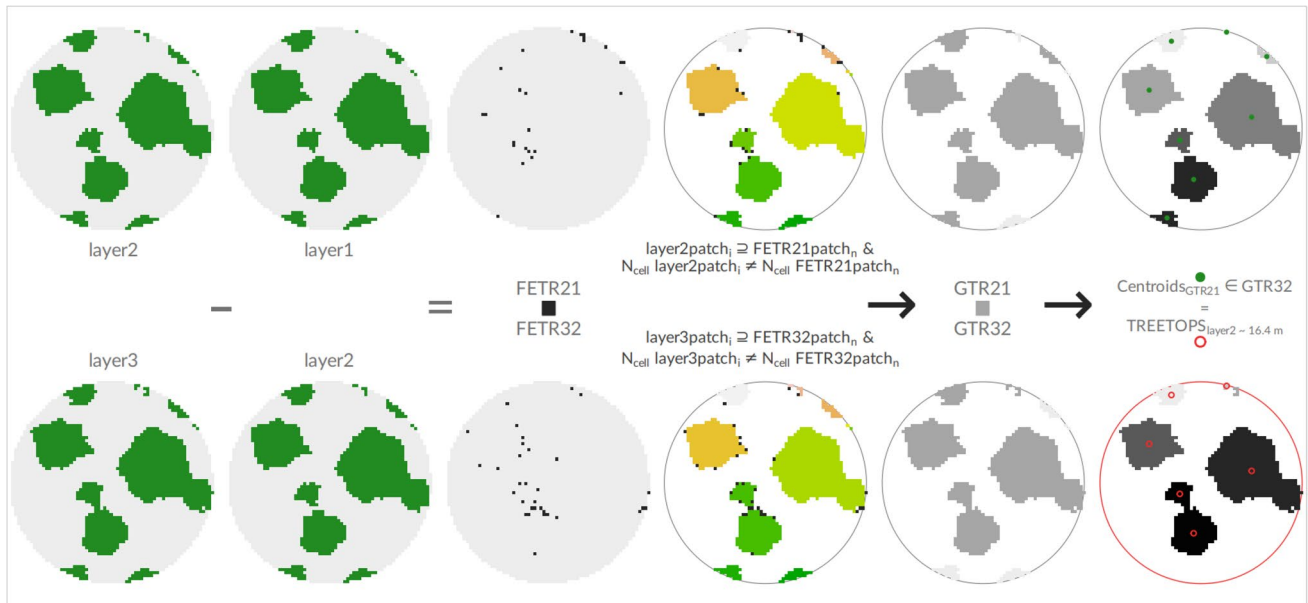


**Fig. 1** Circular example of a CHM with a pixel size of 0.5•0.5 m, displayed in decreasing order between 31.8 m (maximum height) and 6 m. The decrement interval is 2 m. In the top row on the left, the

initial example CHM is depicted, with green representing value 1 and grey indicating no data. The locations of the two example patches are highlighted by filled red circles with transparency



**Fig. 2** Example patches (Fig. 1) provide an overview of the three-layer concept. Green denotes value 1, whereas grey represents no data



**Fig. 3** The three-layer concept (Fig. 2), which is at the core of the GTR algorithm, is visualized and mathematically explained. In the first three columns, green represents value 1, whereas grey indicates no data. The distinct colours in the fourth column and shades of grey

in the sixth column represent the patch IDs. The red circle around the bottom-right raster highlights the ITD output (red unfilled dots) of the algorithm

where  $i$  in  $layer2patch_i$  denotes the  $i$ -th patch of layer2,  $\supseteq$  represents an improper superset and  $n$  in  $FETR21patch_n$  denotes the  $n$ -th patch of FETR21 (Fig. 3). Hence, this double-condition can be expressed as follows: if the  $i$ -th patch of layer2 is

a superset of the  $n$ -th patch of FETR21 whose pixels belong to layer2 AND the pixel number of the  $i$ -th patch of layer2 is not equal to the pixel number of the  $n$ -th patch of FETR21. The output GTR21 (and GTR32) is a binarized raster with

GTR and non-GTR values. Next, individual regions (i.e., patches) of both GTR21 and GTR32 were identified. Finally, the treetops were located using the following formula (Eq. 2, Fig. 3):

$$TREETOPS_{layer2} = Centroids_{GTR21} \in GTR32 \tag{2}$$

where  $TREETOPS_{layer2}$  indicates that the identified treetops in a given iteration are associated with the height value of layer2;  $\in$  represents the element-of symbol; centroids of GTR21 overlapping GTR32 are considered as the treetops.

(iii). Reduction in the number of treetops

Because the treetops are stored in a self-updating process after each iteration (Fig. 4a), it is necessary to reduce the number of treetops. Our solution is a highly flexible three-parameter distance-based treetop reduction method. Technically, it identifies all neighbours within a defined radius (first parameter), referred to as the *distance*, and retains the highest value (Fig. 4a and b). The other two parameters, *min\_H* (minimum height) and *max\_H* (maximum height), facilitate convenient threshold settings. This process enables the identification of lower trees in multilayered forests, as long as these trees are not under the canopy (Fig. 4b).

### TREETOPS

From the user’s perspective, the GTR algorithm is bundled into an R package called TREETOPS, which provides two main functions (Fig. 4c): *get\_TREETOPS()* and *finalize\_TREETOPS()*. The *get\_TREETOPS()* function executes the entire ITD process, including the CHM cutting, storage, and treetop location. The *finalize\_TREETOPS()* function applies the distance-based treetop-reduction method. Both

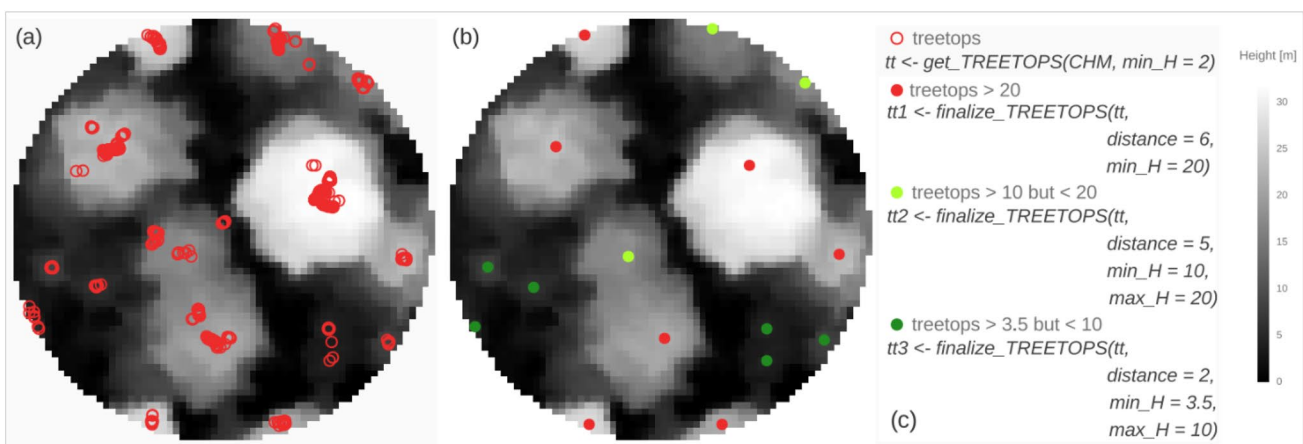
functions are user friendly and require a maximum of four parameters. To set the threshold, the user must set the *min\_H* and *max\_H* (height-defining) parameters within the *finalize\_TREETOPS()* function (Fig. 4c). The outputs of both functions are simple sf objects.

## Testing materials and methods

### Study sites

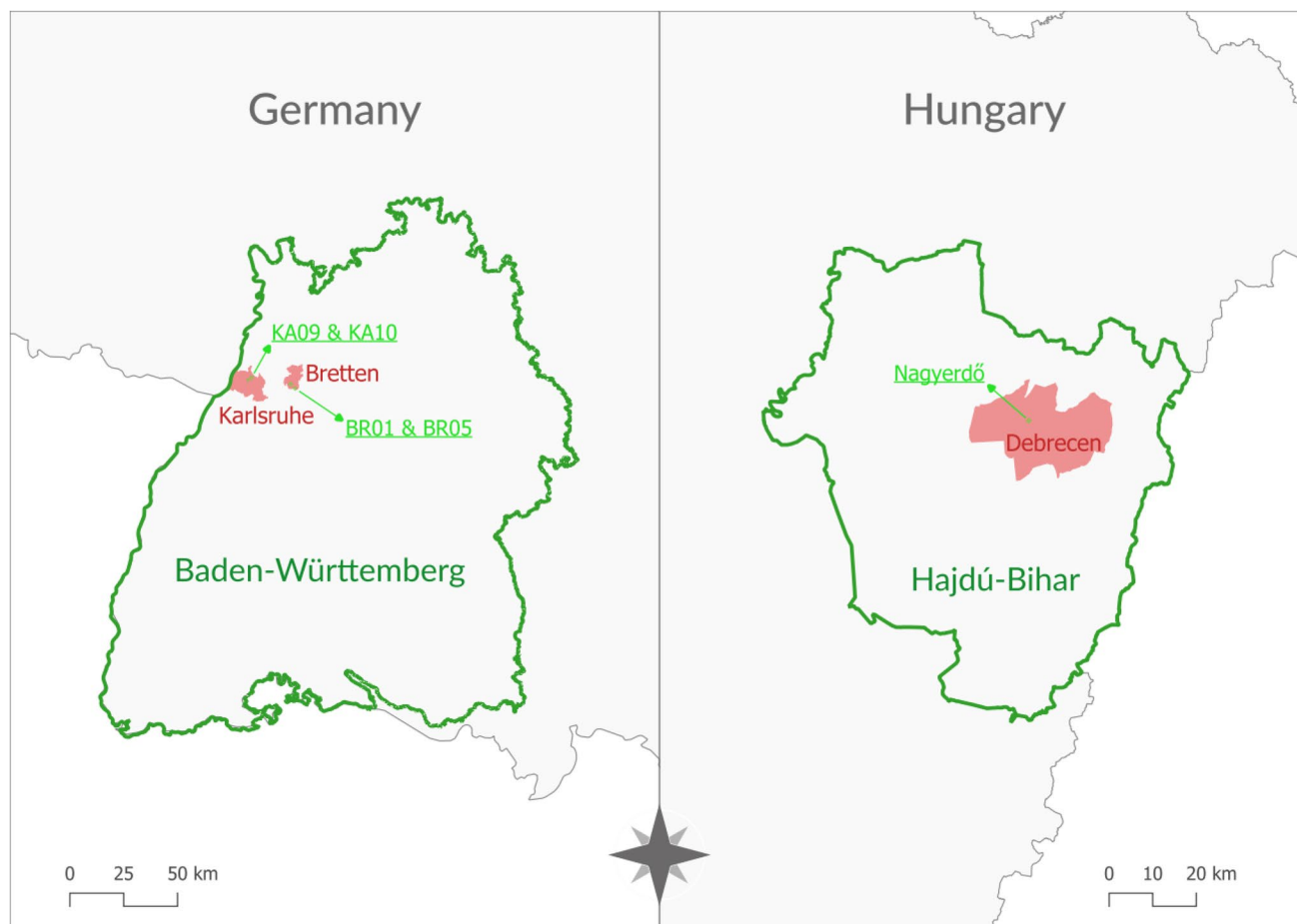
The research encompassed three distinct study areas: the Hardtwald forest in Karlsruhe, the Bretten municipal forest (both situated in the federal state of Baden-Württemberg, Germany), and the Nagyerdő forest in Debrecen (located in the federal state of Hajdú-Bihar, Hungary) (Fig. 5). Terrain characteristics varied across the study sites. The two Karlsruhe sites (KA09 and KA10) and the Nagyerdő forest were situated on flat terrain, whereas the two Bretten sites (BR01 and BR05) were characterized by a hilly landscape.

The studied managed German forests consist primarily of the following tree species: Norway spruce (*Picea abies*), Scots Pine (*Pinus sylvestris*), Douglas fir (*Pseudotsuga menziesii*), common oak (*Quercus robur*), red oak (*Quercus rubra*), sessile oak (*Quercus petraea*), European beech (*Fagus sylvatica*) and European hornbeam (*Carpinus betulus*). Scots pine, red oak, and European beech were dominant in KA09 and KA10. A more diverse tree species composition, including spruce, Douglas fir, European beech, oaks, and European hornbeam, characterises BR01 and BR02. The mixed species composition, dense canopy cover, and multiple layers describe the four forests. Detailed information on plot characteristics was obtained from a previous study (Weiser et al. 2022).



**Fig. 4** Example CHM (0.5•0.5 m pixel size) with treetops (a); example CHM with treetops after applying the distance-based treetop filter (b); (c) the two main functions, *get\_TREETOPS()* and *finalize\_TREETOPS()*

from the R package TREETOPS, and the implementation of threshold settings using the *min\_H* and *max\_H* parameters



**Fig. 5** Locations of the four German forest plots (left) in the federal state of Baden-Württemberg and the Hungarian Nagyerdő forest site (right) in the federal state of Hajdú-Bihar

Our selected Nagyerdő forest plot accommodates native trees like common oak, silver poplar (*Populus alba*), and stands of cultivated Austrian pine (*Pinus nigra*), Scots pine, common hackberry (*Celtis occidentalis*), red oak, and eastern American black walnut (*Juglans nigra*). The forest site is characterized by a very dense, mostly single-layer canopy structure and a composition dominated by deciduous tree species.

### LiDAR data

LiDAR acquisition for German forests was conducted using a RIEGL VQ-780i sensor (RIEGL Laser Measurements Systems, 2019). The flight was conducted on July 5, 2019. For Nagyerdő, the LiDAR data was acquired on May 22, 2020, using the RIEGL VQ780ii system. For a detailed description of the German LiDAR acquisition, refer to Weiser et al. (2022). LiDAR data characteristics for all study areas are summarized in Table 1. Note that the Nagyerdő point cloud has a mean point density of 10 pts/m<sup>2</sup>, which is considered low-resolution.

### Reference data

#### Tree stem positions and tree heights

Reference tree stem positions in the German forests were determined using DTMs derived from ALS point clouds, complemented with data from terrestrial laser scanning (TLS) or unmanned laser scanning (ULS) point clouds (Weiser et al. 2022). Tree heights were extracted from ALS data because of the reported non-visibility of treetops caused by dense canopy conditions (Weiser et al. 2022). In Nagyerdő, 60 reference stem coordinates were collected using the Stonex S9i RTK GNSS system (<https://www.stonex.it/project/s9i-gnss-receiver/>) to ensure accuracy under the canopy. For trees with problematic coordinates caused by a low received signal owing to dense canopy filtering, 15 repeated measurements were conducted, and the coordinates were calculated by averaging the results. Tree heights were measured using the Haglöf EC II D-R electronic tool (<https://haglofsweden.com/project/ec-ii-d-r>) employing the reference height mark control method. Each tree was marked

**Table 1** Characteristics of LiDAR data used for study sites

Characteristics	KA09, KA10, BR01 and BR05	Nagyerdő
ALS sensor	RIEGL VQ-780i	RIEGL VQ-780ii
Flying altitude AGL	650 m	1077 m
Off-nadir scan angle	$\pm 30^\circ$	$\pm 30^\circ$
Laser beam divergence	0.25 mrad	0.18 mrad
Pulse repetition frequency	1000 kHz	2000 kHz
Flight line distance	175 m	240 m
Flight line overlap	76%	20%
Point density	> 40 pts/m <sup>2</sup> (max: 103 pts/m <sup>2</sup> )	> 2 pts/m <sup>2</sup> (mean: 10 pts/m <sup>2</sup> )

at 2 m using a tape measure, and during measurement, the EC II D-R was aligned to ensure that it displayed 2 m when aimed at the mark. This accuracy control process resulted in the determination of the 60 reference tree heights.

### Filtering trees based on CHM-derived heights

Nagyerdő forest was selected as the target site, whereas the four German forests served as test sites. Therefore, Nagyerdő forest will undergo individual tree segmentation in a subsequent study, guided by the treetops identified by the CHM-based treetop detector presented in this study. Additionally, the LiDAR data acquired over Nagyerdő has considerably lower resolution compared to the German aerial laser scanned data (Table 1). We aimed to develop a raster-based ITD method using low-resolution LiDAR data. It should also be noted that the resolution of the raw LiDAR data did not affect the information quality obtained from the rasterized topological surface of the canopy. However, acquiring low-resolution ALS data is significantly less expensive.

Because German forests consist of multiple layers and raster-based ITD is not capable of capturing the treetops of suppressed or under-canopy trees (Eysn et al. 2015), CHM-derived forest height distributions were visualized to understand the similarities and differences between the target and the four test sites (Fig. 6). The differences between the two highest peaks of these distributions were calculated to define the CHM-based layering of the respective forests. While Nagyerdő presented a height difference of 1.93 m, suggesting a single-layer forest, the German sites revealed height differences ranging from 3.42 m to 12.17 m with multiple peaks (Fig. 6). Overall, the tree heights in the Nagyerdő forest were considerably lower. Moreover, the distributional line segments of BR01 and BR05 showed similarities over their 20-m ranges in the Nagyerdő height profile (Fig. 6). This phenomenon suggests that the CHM-induced one-to-one comparability between the trees of BR01 and BR05 and the trees of Nagyerdő is only meaningful for trees, especially of the BR01 and BR05 forests, which are higher than 20 m. Therefore, the reference trees of the four German forests

were filtered, and only trees higher than 20 m were considered in this study.

### Number of reference trees and tree species distributions

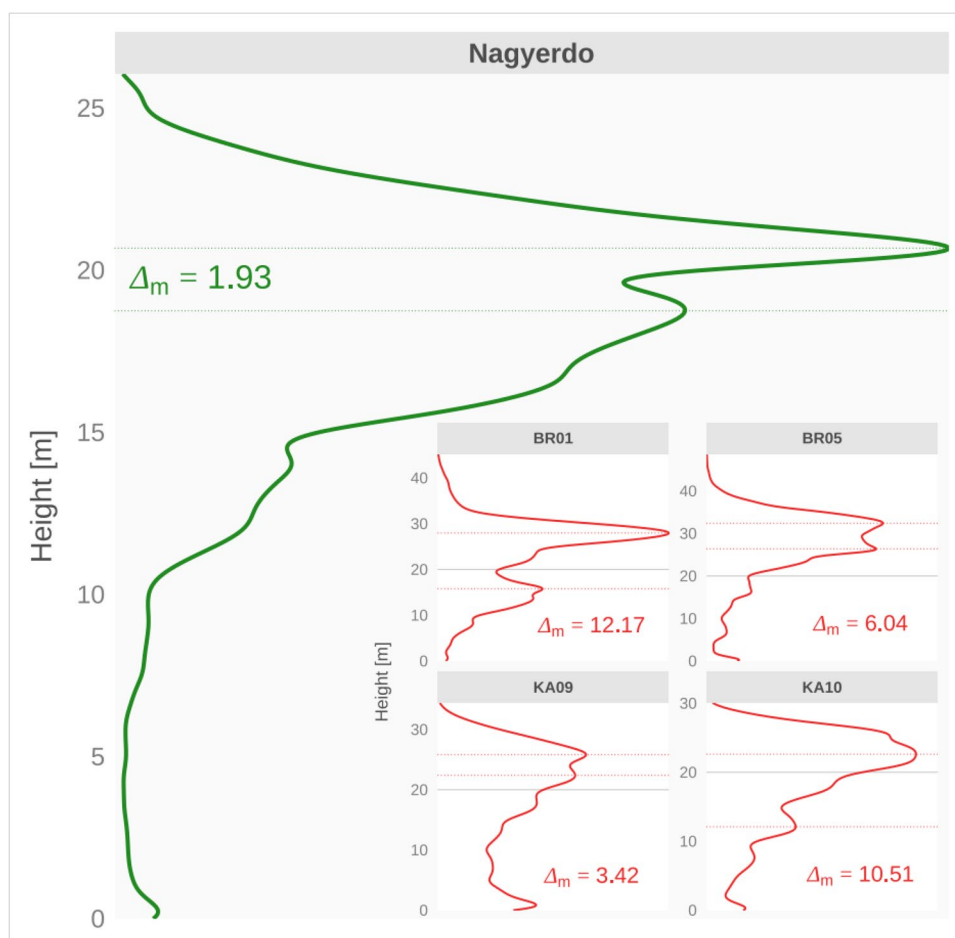
After height filtering, the number of reference trees in the five study sites: BR01, BR05, KA09, KA10, and Nagyerdő, became 153, 173, 106, 32 (Fig. 7a), and 60 (Fig. 7b), respectively. Concerning the area of the four test forests, BR01 accommodated 2.02 ha and BR05 1.89 ha, while KA09 and KA10 were left with 1.80 ha and 1.22 ha (Fig. 7c). Our target forest, Nagyerdő was the only 1 ha square area because it was earmarked for future research. The encircling cuts from the respective CHMs of the four test forests ensured that their area boundaries were close to the test reference trees, minimising their areas. Smaller areas facilitated faster ITD processing.

The selection of these German forests for testing was also motivated by the similarity between their tree species composition and that of Nagyerdő forest. From an ALS perspective, the ratio of coniferous to deciduous trees was of primary interest (Eysn et al. 2015; Stereńczak et al. 2020; Zhao et al. 2017). Therefore, the ratios of the reference tree species and coniferous-to-deciduous ratios of each site were plotted and visually compared (Fig. 8). The reference trees of BR01 and KA10 showed high similarity, those of BR05 showed low similarity, and those of the KS09 forest were dissimilar to the Nagyerdő reference trees (Fig. 8c).

### LiDAR data preprocessing

Three preprocessing steps were conducted using the lidR package (Roussel et al. 2020) in the R computational environment (R Core Team 2022). First, the raw LiDAR point cloud was classified into ground and surface points using cloth simulation filtering (Zhang et al. 2016). Next, the CHM was derived directly from the classified above-ground and ground points using a pit-free algorithm (Khosravipour et al. 2014). However, the resulting CHMs of the BR01, BR05, and KA10 test sites exhibited two, one, and three pits, respectively (Fig. 7a). These pits did not influence the

**Fig. 6** CHM-derived height distribution of the target forest (shown in green). The inset graphs show the vertical height profiles of the BR01, BR05, KA09, and KA10 test forests (depicted in red). In the insets, the horizontal dashed lines indicate the highest two peaks at each site, with the differences expressed using the symbol  $\Delta_m$ . The grey horizontal lines in the inset graphs represent the 20-m comparability threshold



ITD and were therefore not subjected to further processing. Finally, Gaussian filtering with a window size of  $3 \times 3$  pixels was applied for image smoothing and noise removal.

### LM with VWF versus the GTR

To ensure logical comparability between the two methods, we proposed an equation based on our target study site to adjust the window size of the VWF depending on the tree height embedded in the CHM. This function was defined as follows (Eq. 3, Fig. 9):

$$y = 5.7 \cdot (-e^{-0.08 \cdot (x-5)} - 1) + 1 \quad (3)$$

where  $y$  represents the window size for the VWF and the *distance* parameter of *finalize\_TREETOPS()* and  $x$  represents the height values stored by the CHM. A sensible minimum height threshold of 5 m was chosen, below which no trees could be found in the single-layered Nagyerdő forest. The delinearised green curve became linear at 20 m (Fig. 9, x-axis), indicating that the window size was fixed at 5 m for

the B01, B05, KA09, and KA10 test forests. Throughout this study, the LM with VWF was optimised and computed using the *lidR* package.

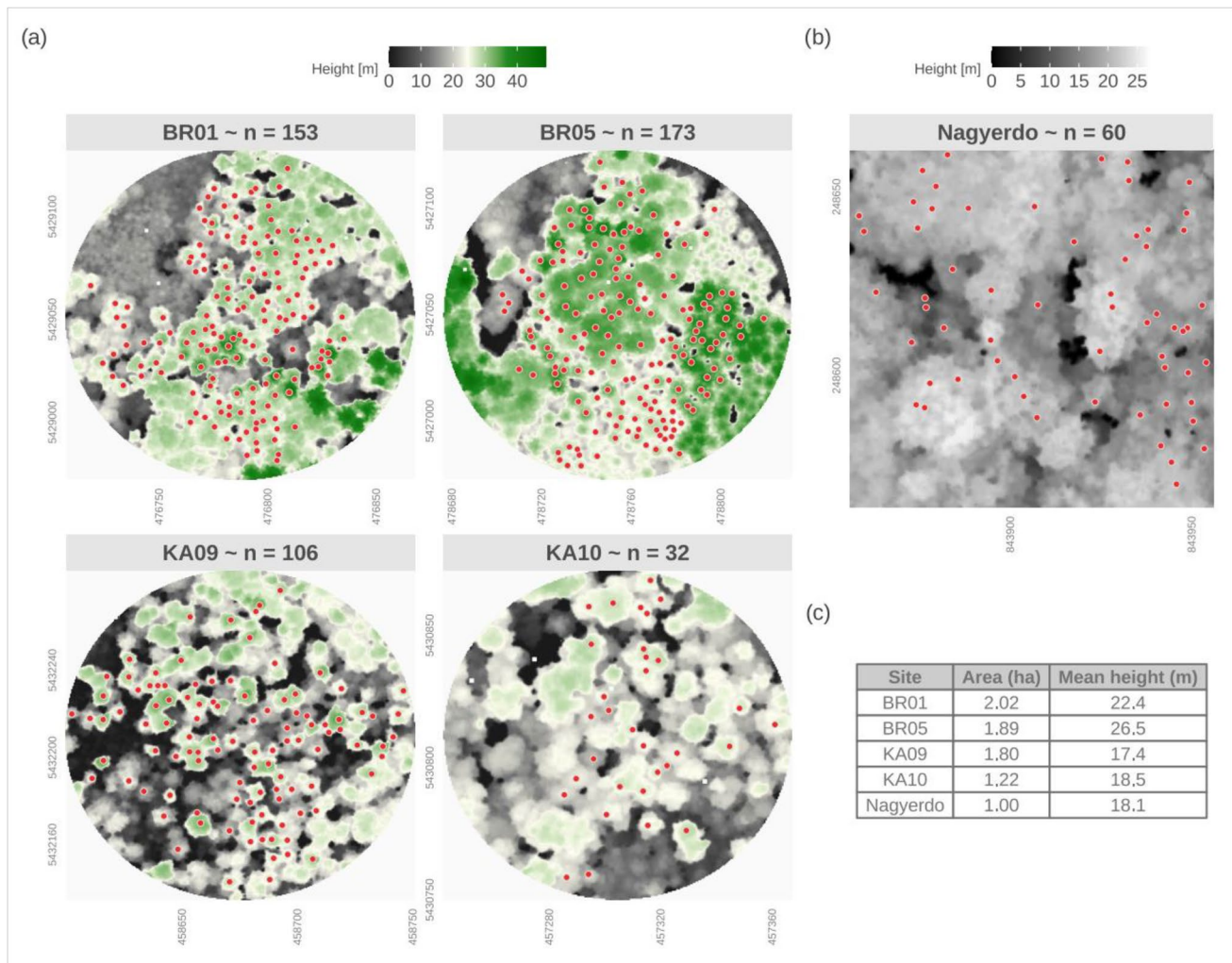
The GTR distances (i.e., the window sizes of VWF) were calculated for the following thresholds: 5.00 to 9.99 m, 10.00 to 20.00 m, and above 20.01 m, using the following formulae (Eq. 4–6, Fig. 9):

$$4 = VWS_{height>20} - 1 \quad (4)$$

$$2.93 = (VWS_{height=10} + (VWS_{height=20} - VWS_{height=10})/2) - 1 \quad (5)$$

$$0.94 = (VWS_{height=5} + (VWS_{height=10} - VWS_{height=5})/2) - 1 \quad (6)$$

where  $VWS_{height>}$  and  $VWS_{height=}$  denote the variable window size greater than and equal to a given height ( $y$ -axis in Fig. 9), respectively. The distance for treetop reduction for treetops taller than 20 m corresponds to 4 m, whereas for treetops between 10 and 20 m, it is 2.93 m. For treetops smaller than 10 m, this distance was negligible (Fig. 9).



**Fig. 7** Study sites and reference tree coordinates of the four German forests (a) and of Nagyerdó forest (b). Site-specific reference tree characteristics in tabular form (c). The mean height were derived from the respective CHMs

### Accuracy assessment of individual treetop detection

A verification method developed by Eysn et al. (2015) has established itself in studies focussing on Central European alpine and lowland forests (Eysn et al. 2015; Stereńczak et al. 2020). The key algorithm of this approach is called the Matching Algorithm (MA), which links treetop detection results to reference data. During the matching process, the MA produces various qualitative and quantitative statistical parameters (Eysn et al. 2015). A modified version of MA was applied to assess the accuracy of the ITD results. This had to be altered because our particular interest lay in a comparative analysis between the LM with VWF and the GTR on a one-to-one basis. More importantly, the German and the Nagyerdó reference trees did not provide reference information for every tree standing within the study area boundaries. Thus, the number of trees extracted using the two methods was not reliably linked to the number of reference trees.

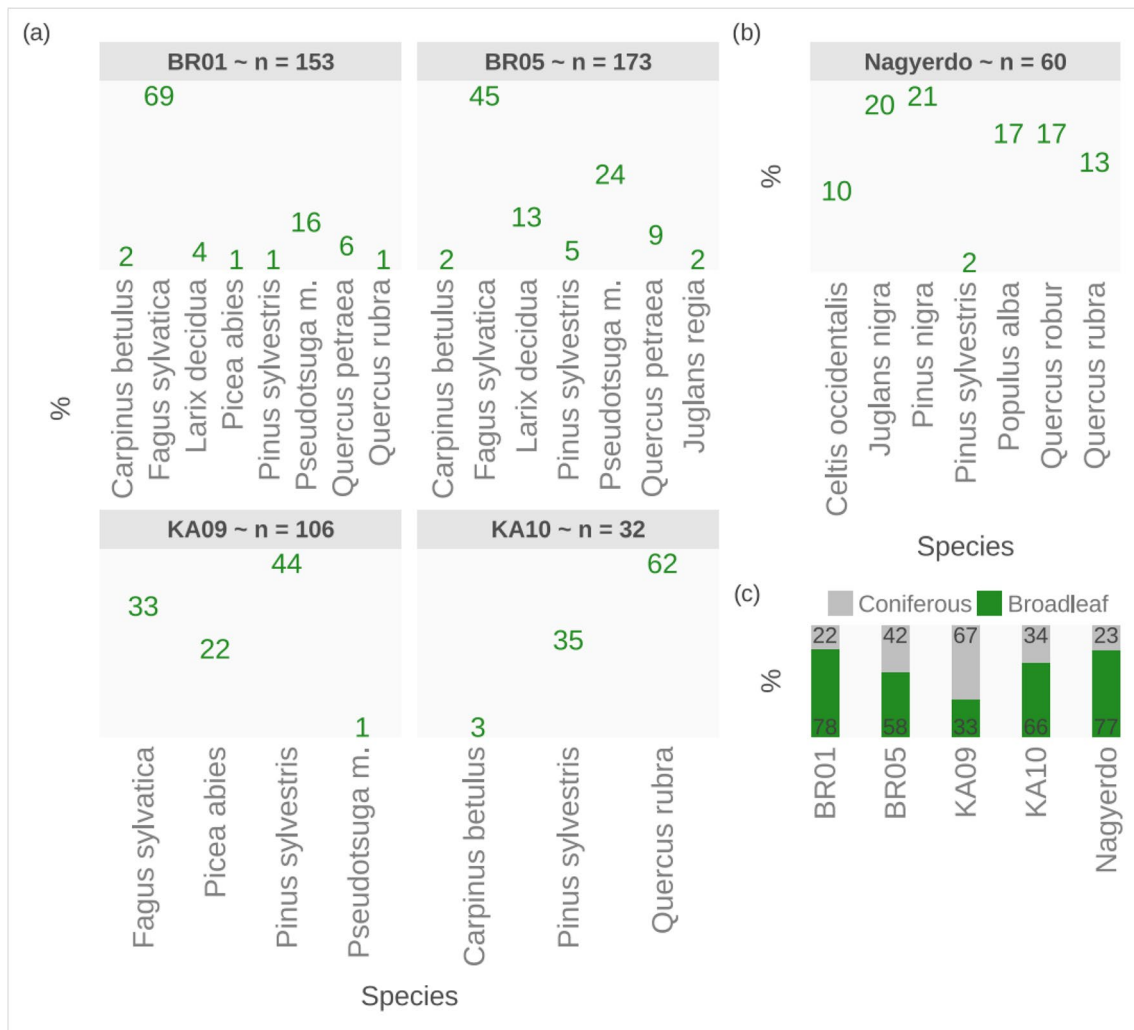
This phenomenon technically hindered obtaining the correct extraction and commission rates. To address this issue, both the extraction and commission rates were mathematically adjusted using Eq. 7:

$$adjustment = \log_{10}(N_{MTest}) \cdot 20 \tag{7}$$

where the  $N_{MTest}$  parameter is the Modified Number of Extracted trees described below.

The following validation parameters were output by the MA:

- $N_{Test}$  → Number of Extracted trees,
- $N_{Ref}$  → Number of Reference trees,
- $R_{Ref}$  → Rate of Reference trees  $\sim N_{Ref} / (N_{Ref} + N_{Test})$ ,
- $N_{MTest}$  → Modified  $N_{Test} \sim$  if  $R_{Ref} < 0.3 \rightarrow N_{MTest} = N_{Test} \cdot (1 - R_{Ref}) - N_{Ref}$
- else →  $N_{MTest} \quad N_{Test}$



**Fig. 8** Reference tree species distributions (%) of the four German forests (a) and of the Nagyerdő forest (b). Panel (c) shows the ratios of coniferous to broadleaf tree species in the respective forest plots

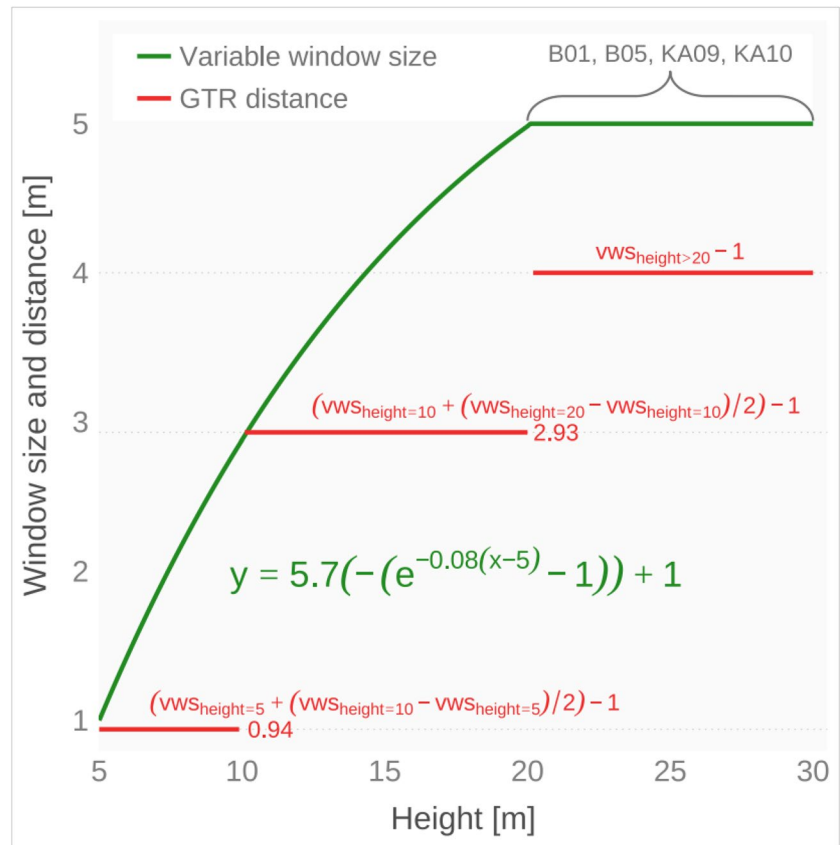
$N_{Match}$	→	Number of Matched trees,	$N_{Match}/N_{Ref}$	→	Matching rate,
$N_{Com}$	→	Number of Extracted trees that could not be matched $\sim N_{MTest} - N_{Match}$ ,	$N_{om}/N_{Ref}$	→	Omission rate, Reference trees that could not be matched.
$N_{Om}$	→	Number of Reference trees that could not be matched $\sim N_{Ref} - N_{Match}$ ,			
$N_{MTest}/N_{Ref}$	→	Modified Extraction rate,			
$N_{Com}/N_{MTest}$	→	Modified Commission rate, Extracted trees that could not be matched.			
$H_{Test}$	→	Extracted Height of Matched Test trees,			
$H_{Ref}$	→	Height of Reference trees,			
$D_{Hor}$	→	Horizontal distance between Matched Test and Reference trees,			

Using the values of  $H_{Test}$  and  $H_{Ref}$ , the mean absolute error (MAE) was determined for height difference verification between Matched Test and Reference trees by using Eq. 8:

$$MAE_{height[m]} = (1/N_{Ref}) \cdot \sum |H_{Ref_i} - H_{Test_i}| \quad (8)$$

where  $i$  indicates the  $i$ -th value regarding Reference and Test heights and the modulus is denoted by the  $||$  symbol. Furthermore, the MA output parameters described above were used to define the following root mean square (RMS) parameters:

**Fig. 9** Calibrated variable window size for the Nagyerdő forest (depicted by the green line) and the adjusted GTR distances (represented by the red line) in relation to height (x-axis). The main equation is displayed in green, while equations and values of distances are written in red;  $VWS_{\text{height} \geq}$  signifies a variable window size greater than or equal to a given height (y-axis)



- $RMS_{Match} \rightarrow$  Root Mean Square of Matching rates,
- $RMS_{Extr} \rightarrow$  Root Mean Square of Extraction rates,
- $RMS_{Com} \rightarrow$  Root Mean Square of Commission rates,
- $RMS_{Om} \rightarrow$  Root Mean Square of Omission rates.

## Results

### ITD results at site per method

The highest matching rate was found in the KA10 forest, where both methods showed 81% matching. Interestingly, the lowest rate (55%) was produced by LM with VWF in the BR01 forest (Fig. 10a). Regarding our target forest, the GTR method (80%) slightly outperformed the LM with VWF (77%). Overall, a minor performance advantage was observed in favour of the GTR method.

Regarding the accuracy of height matching, the GTR method displayed Mean Absolute Error (MAE) values between 0.73 and 2.04 m, while the LM with VWF evidenced MAE values between 1.01 and 2.50 m (Fig. 10a). Across all forest sites, the GTR achieved slightly better MAE results. Spatially, it is noteworthy that the two types

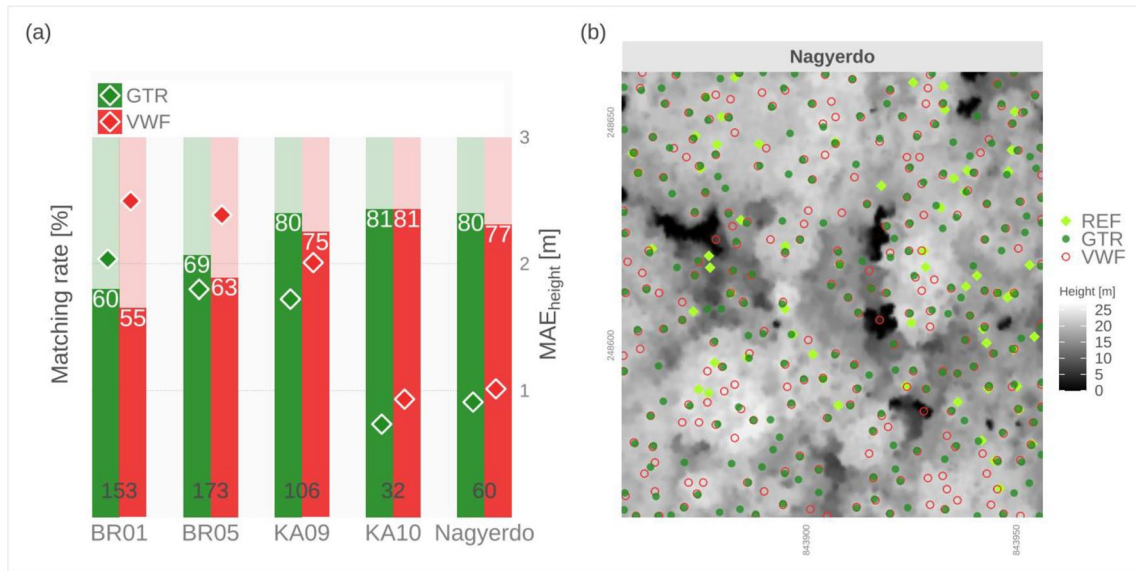
of matched trees (those located by the LM with VWF and the GTR methods matching a particular reference tree) were situated within a distance of less than 4 m from each other in every case where both methods matched the same reference tree (Fig. 10b).

Concerning commission errors, no differences were observed between the LM with VWF and the GTR methods at the KA09 site (Fig. 11). While the LM with VWF obtained a 2% lower value in KA10, the three remaining forests—BR01, BR05, and Nagyerdő—witnessed the GTR having 2%, 1%, and 2% lower rates, respectively (Fig. 11).

The highest omission rate of 45% was found with the LM using VWF in BR01, whereas the lowest rate (19%) was observed with both algorithms in the KA10 forest (Fig. 11). For BR01, BR05, KA09, and Nagyerdő, the GTR method showed lower omission error rates by 5, 6, 5, and 3%, respectively.

### ITD results per method

Quantitatively, both ITD approaches yielded high RMS values: 74% by the GTR and 71% by the LM with VWF. The RMS differences concerning the modified commission



**Fig. 10** Matching rates concerning forest sites and the dual methods (a: diamonds indicate MAE values, grey numbers on bars represent the number of reference trees for each forest site, white numbers show the matching error rates for each forest depending on the methods:

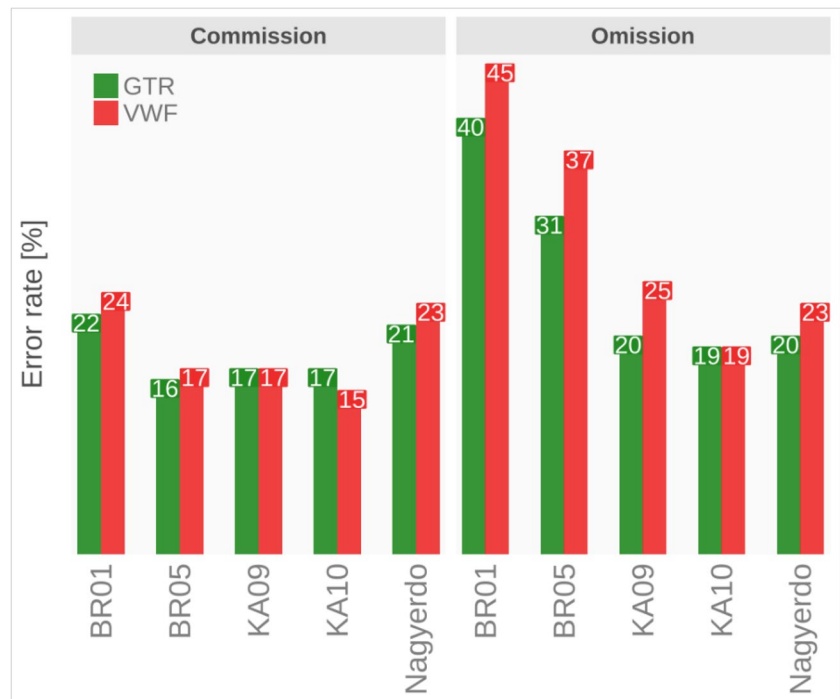
GTR, Growing Tree Region, and VWF, Variable Window Filtering). The spatial locations of the 60 reference (REF) trees, GTR-extracted treetops, and Local Maxima with VWF-extracted treetops of the Nagyerdo forest are shown in (b)

and omission rates were 1% and 4%, respectively, providing an edge to the GTR (Fig. 12). The modified extraction rates were high for both methods: 140% for GTR and 134% for LM with VWF (Fig. 12).

### Discussion

The GTR method described in this study is a simple CHM-based treetop detection algorithm that was developed for Central European forests. It was provided in a ready-to-use

**Fig. 11** Modified commission and omission error rates for the studied forest sites with regard to the two methods (GTR: Growing Tree Region; VWF: Variable Window Filtering)



form for both forest managers and the research community. Its open access availability, an issue that other researchers often overlook (Stereńczak et al. 2020; Zhao et al. 2017), makes it indisputably useful. We offer a flexible user-friendly tool that outperforms the widely used LM with VWF, and can be adjusted to various forest ecosystems depending on their vertical structure. The five tested Central European forests were mixed and partly dominated by deciduous trees. According to previous studies, similar forest communities make ITD highly demanding (Eysn et al. 2015; Vauhkonen et al. 2012). Based on prior knowledge (Eysn et al. 2015; Stereńczak et al. 2020), we selected possibly visible trees from the low-resolution ALS perspective using the canopy layer of multi-story forests. Moreover, our approach to accuracy assessment is derived from a method established in ITD performance evaluation, focusing on Central European alpine and lowland forests (as described in Section [Accuracy assessment of individual treetop detection](#)). This enabled our findings to be directly linked to those presented by Eysn et al. (2015). A segmentation-centered study that conducted ITD before delineation in mixed and deciduous forests, similar to the ones in this research, also provided comparable results to our findings (Zhao et al. 2017).

An important aspect of the comparative analysis between the GTR and the LM with VWF is that the VWF was calibrated specifically for the Nagyerdő forest (Section [LM with VWF versus the GTR](#)). Notably, the GTR was mathematically adjusted for logical comparison with the LM with VWF, which means that the threshold-setting capability

(Section [The algorithm](#), (iii) Number of treetop reduction) of the GTR tool was not fully utilized.

Eysn et al. (2015) analysed the performance of eight different ITD methods and reported RMS matching rates between 66 to 82% for trees above 20 m. One of the best-performing methods, considering extraction, commission, and omission error rates, was the LM with VWF, which achieved 72%. Although their research areas were located in the Alpine region, our results (74% revealed by GTR and 71% by the VWF) aligned with theirs. This phenomenon suggests, unlike statements from other studies (González-Ferreiro et al. 2013; Khosravipour et al. 2014), that dominant and co-dominant treetops can be equivalently well extracted independently of the topographic characteristics of the forest. In general, the declared overall performance was comparable. While the article published by Eysn et al. (2015) described 47% as the best matching rate, our study reported 52% for both methods. These correspondences underline the plausibility of our proposed method for commission and extraction adjustments.

At the Nagyerdő site, the GTR matching rate was 80%. Similar ITD results with percentages between 80 and 87 were reported in temperate, less dense boreal deciduous forests (Zhao et al. 2017). This analogy can be explained by the fact that both algorithms applied a theoretically similar top-down CHM-cutting approach. Although GTR outperformed the calibrated LM with VWF, the obtained RMS differences in matching, modified commission, and omission rates are marginal. While the VWF has been used throughout this study in its calibrated form based on Nagyerdő forest, the GTR was adjusted to conduct a logically founded comparative analysis. Thus, it can be stated that the GTR is a competitive alternative to the LM with VWF, and it should also be optimized by using its threshold-setting feature without limitations. However, a calibrated VWF can match the robustness of the non-optimized GTR method in lowland mixed and deciduous forests.

Three further issues need to be addressed. First, considering the CHM as a topographic surface makes it difficult for the GTR to detect suppressed trees and trees with flatly-shaped crowns because such objects lack the necessary protrusions emerging from the CHM. This issue becomes relevant if treetop distances in the same height class (as defined by threshold setting) of the studied forest vary, and the output of the *finalize\_TREETOPS()* function would be erroneous. Second, the lack of abundant reference trees covering the target (Nagyerdő forest) study area led us to develop a strategy for modified commission and extraction rates, which were used to evaluate the performance of both algorithms on the test sites (BR01, BR05, KA09, and KA10). While plausible, this strategy is a prone-to-error replacement for commission and extraction rates obtained using forest inventory data covering the entire study area. Third, the GTR was



**Fig. 12** Overall performance of the compared ITD methods: Root Mean Square (RMS) of matching, modified extraction, modified commission, and omission error rates for the two methods (GTR: Growing Tree Region; VWF: Variable Window Filtering) were computed across all study sites

tested in temperate Central European mixed- and deciduous-dominated forests; thus, the robustness of the algorithm in other forested ecosystems is unknown. It is reasonable to assume that the algorithm would perform better in a coniferous environment, as the ITD in coniferous forests has been reported to achieve higher accuracy. The robustness of the GTR method should be evaluated in the future using other types of LiDAR forest data combined with area-covering forest inventory data from different forested ecosystems.

## Conclusions

We introduced a novel and robust method for detecting tree-tops, termed the Growing Tree Region (GTR) algorithm, which leverages canopy height models (CHM) derived from low-resolution LiDAR data. Our algorithm surpasses the widely employed local maxima (LM) search technique with calibrated variable window filtering (VWF) in deciduous-dominated forests in Central Europe. To enhance the user-friendly implementation of our algorithm, we developed an R package named TREETOPS that seamlessly integrates into the lidR package specialised for LiDAR applications in forestry.

In contrast to the automated search of LM with VWF, TREETOPS provides users with a three-parameter-controlled threshold-setting option to effectively detect lower tree-tops in multilayered forests. While this study did not specifically address the performance of the threshold setting, it is likely influenced by the limitations of the CHM, particularly in the detection of trees beneath the canopy layer. TREETOPS has the capability to accurately detect younger (lower) trees that are visible from an airborne laser scanning (ALS) perspective and are not concealed by the canopy. This suggests that TREETOPS may outperform LM with VWF in detecting tree-tops of younger trees in multilayered forests, although comprehensive testing is essential to validate this assertion.

It is strongly recommended that future studies focus on evaluating the effectiveness of the GTR method across diverse forest types worldwide. Given the tool's open accessibility, exploring the combination of GTR with various segmentation techniques within the lidR programming framework is also encouraged. Additionally, a viable two-stage procedure for tree detection in multi-layered forests could involve i) CHM-based tree-top detection and ii) tree crown delineation utilizing the original point cloud in conjunction with the tree-tops identified by TREETOPS.

Acknowledging the low-resolution of LiDAR data, there is an ongoing effort to develop point-cloud-based tree

segmentation. This development relies on the assignment of laser points to detected tree-tops generated by TREETOPS. This approach aims to enhance the accuracy of tree segmentation despite data limitations (i.e., low-resolution LiDAR), emphasizing the potential of TREETOPS in contributing to comprehensive forest analysis methodologies.

**Acknowledgements** The authors were supported by the NKFI K138079 and the KKP 144068 projects during manuscript preparation. We are grateful to Envirosense Hungary Ltd. for providing the LiDAR data used in this study.

**Author contributions** *Gergő Diószegi* conceived the ideas, conceptualized, preprocessed data, conducted programming and visualising, wrote original draft and finalized draft; *Vanda Éva Molnár* and *Loránd Attila Nagy* collected reference data; *Péter Enyedi* provided LiDAR data; *Péter Török* reviewed, edited original draft and finalized draft; *Szilárd Szabó* led writing, reviewed and edited original draft and finalized draft. All authors contributed to the research and gave final approval for publication.

**Funding** Open access funding provided by University of Debrecen.

**Data availability** TREETOPS package is available at <https://github.com/DijoG/TREETOPS>, and can be installed by executing the following line:

```
devtools::install_github("DijoG/TREETOPS").
```

## Declarations

**Conflict of 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.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Beucher S, Meyer F (1993) Segmentation: The Watershed Transformation. *Mathematical Morphology in Image Processing*. *Opt Eng* 34:433–481
- Broszofski KD, Froese RE, Falkowski MJ, Banskota A (2014) A Review of Methods for Mapping and Prediction of Inventory Attributes for Operational Forest Management. *Forest Science* 60:733–756. <https://doi.org/10.5849/forsci.12-134>
- Coops NC, Tompalski P, Goodbody TRH, Queinnec M, Luther JE, Bolton DK, White JC, Wulder MA, van Lier OR, Hermosilla T (2021)

- Modelling lidar-derived estimates of forest attributes over space and time: A review of approaches and future trends. *Remote Sens Environ* 260:112477. <https://doi.org/10.1016/j.rse.2021.112477>
- Dalpono, M, Coomes, D (2016) Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. *Methods in Ecol Evol* 7. <https://doi.org/10.1111/2041-210X.12575>
- Duncanson LI, Cook BD, Hurtt GC, Dubayah RO (2014) An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosystems. *Remote Sens Environ* 154:378–386. <https://doi.org/10.1016/j.rse.2013.07.044>
- 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* 33:5171–5193. <https://doi.org/10.1080/01431161.2012.657363>
- Eysn L, Hollaus M, Lindberg E, Berger F, Monnet J-M, Dalpono M, Kobal M, Pellegrini M, Lingua E, Mongus D, Pfeifer N (2015) A Benchmark of Lidar-Based Single Tree Detection Methods Using Heterogeneous Forest Data from the Alpine Space. *Forests* 6:1721–1747. <https://doi.org/10.3390/f6051721>
- González-Ferreiro E, Diéguez-Aranda U, Miranda D (2012) Estimation of stand variables in *Pinus radiata* D. Don plantations using different LiDAR pulse densities. *Forestry: An Int J Forest Res* 85:281–292. <https://doi.org/10.1093/forestry/cps002>
- González-Ferreiro E, Diéguez-Aranda U, Barreiro-Fernández L, Buján S, Barbosa M, Suárez JC, Bye IJ, Miranda D (2013) A mixed pixel- and region-based approach for using airborne laser scanning data for individual tree crown delineation in *Pinus radiata* D. Don plantations. *Int J Remote Sens* 34:7671–7690. <https://doi.org/10.1080/01431161.2013.823523>
- Hijmans R (2023) terra: Spatial Data Analysis. R package version 1.6-47. <https://github.com/rspatial/terra>
- Hu B, Li J, Jing L, Judah A (2014) Improving the efficiency and accuracy of individual tree crown delineation from high-density LiDAR data. *Int J Appl Earth Obs Geoinf* 26:145–155. <https://doi.org/10.1016/j.jag.2013.06.003>
- Hyypä J, Hyypä H, Leckie D, Gougeon F, Yu X, Maltamo M (2008) Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *Int J Remote Sens* 29:1339–1366. <https://doi.org/10.1080/01431160701736489>
- Kaartinen H, Hyypä J, Yu X, Vastaranta M, Hyypä H, Kukko A, Holopainen M, Heipke C, Hirschmugl M, Morsdorf F, Næsset E, Pitkänen J, Popescu S, Solberg S, Wolf BM, Wu J-C (2012) An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. *Remote Sensing* 4:950–974. <https://doi.org/10.3390/rs4040950>
- Khosravipour A, Skidmore AK, Isenburg M, Wang T, Hussin YA (2014) Generating Pit-free Canopy Height Models from Airborne Lidar. *Photogramm Eng Remote Sensing* 80:863–872. <https://doi.org/10.14358/PERS.80.9.863>
- Koch B, Heyder U, Weinacker H (2006) Detection of Individual Tree Crowns in Airborne Lidar Data. *Photogramm Eng Remote Sensing* 72:357–363. <https://doi.org/10.14358/PERS.72.4.357>
- Lamar WR, McGraw JB, Warner TA (2005) Multitemporal censusing of a population of eastern hemlock (*Tsuga canadensis* L.) from remotely sensed imagery using an automated segmentation and reconciliation procedure. *Remote Sens Environ* 94:133–143. <https://doi.org/10.1016/j.rse.2004.09.003>
- Lee SJ, Kim JR, Choi YS (2013) The extraction of forest CO<sub>2</sub> storage capacity using high-resolution airborne lidar data. *Gisience & Remote Sensing* 50:154–171. <https://doi.org/10.1080/15481603.2013.786957>
- Li W, Guo Q, Jakubowski M, Kelly M (2012) A New Method for Segmenting Individual Trees from the Lidar Point Cloud. *Photogrammetric Eng Remote Sens* 78:75–84. <https://doi.org/10.14358/PERS.78.1.75>
- Lim K, Treitz P, Wulder M, St-Onge B, Flood M (2003) LiDAR remote sensing of forest structure. *Progress in Physical Geography: Earth and Environment* 27:88–106. <https://doi.org/10.1191/0309133303pp360ra>
- Maltamo, M, Gobakken, T (2014) Predicting Tree Diameter Distributions, in: Maltamo, M., Næsset, E., Vauhkonen, J. (Eds.), *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies, Managing Forest Ecosystems*. Springer Netherlands, Dordrecht, pp. 177–191. [https://doi.org/10.1007/978-94-017-8663-8\\_9](https://doi.org/10.1007/978-94-017-8663-8_9)
- Mielcarek M, Stereńczak K, Khosravipour A (2018) Testing and evaluating different LiDAR-derived canopy height model generation methods for tree height estimation. *Int J Appl Earth Obs Geoinf* 71:132–143. <https://doi.org/10.1016/j.jag.2018.05.002>
- Parkitna K, Krok G, Miścicki S, Ukalski K, Lisańczuk M, Mitelsztedt K, Magnussen S, Markiewicz A, Stereńczak K (2021) Modelling growing stock volume of forest stands with various ALS area-based approaches. *Forestry An Int J Forest Res* 94:630–650. <https://doi.org/10.1093/forestry/cpab011>
- Pebesma E, Bivand R (2023) *Spatial Data Science: With applications in R*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429459016>
- Pitkänen J (2001) Individual tree detection in digital aerial images by combining locally adaptive binarization and local maxima methods. *Can J for Res* 31:832–844. <https://doi.org/10.1139/x01-013>
- 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. *Photogrammetric Eng Remote Sens* 70:589–604. <https://doi.org/10.14358/PERS.70.5.589>
- Popescu S, Wynne R, Nelson R (2002) Estimating plot-level tree heights with LiDAR: local filtering with a canopy-height based variable window size. *Comput Electron Agric* 37:71–95. [https://doi.org/10.1016/S0168-1699\(02\)00121-7](https://doi.org/10.1016/S0168-1699(02)00121-7)
- R Core Team (2022) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Roussel J-R, Auty D, Coops NC, Tompalski P, Goodbody TRH, Meador AS, 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. <https://doi.org/10.1016/j.rse.2020.112061>
- Stereńczak K, Kraszewski B, Mielcarek M, Piasecka Ż, Lisiewicz M, Heurich M (2020) Mapping individual trees with airborne laser scanning data in an European lowland forest using a self-calibration algorithm. *Int J Appl Earth Obs Geoinf* 93:102191. <https://doi.org/10.1016/j.jag.2020.102191>
- Unger DR, Hung I-K, Brooks R, Williams H (2014) Estimating number of trees, tree height and crown width using Lidar data. *Gisience & Remote Sensing* 51:227–238. <https://doi.org/10.1080/15481603.2014.909107>
- Vauhkonen J, Ene L, Gupta S, Heinzl J, Holmgren J, Pitkänen J, Solberg S, Wang Y, Weinacker H, Hauglin KM, Lien V, Packalén P, Gobakken T, Koch B, Næsset E, Tokola T, Maltamo M (2012) Comparative testing of single-tree detection algorithms under different types of forest. *Forestry: An Int J Forest Res* 85:27–40. <https://doi.org/10.1093/forestry/cpr051>
- Wang L, Gong P, Biging GS (2004) Individual Tree-Crown Delineation and Treetop Detection in High-Spatial-Resolution Aerial Imagery.

- Photogramm Eng Remote Sensing 70:351–357. <https://doi.org/10.14358/PERS.70.3.351>
- Weiser H, Schäfer J, Winiwarter L, Krašovec N, Fassnacht FE, Höfle B (2022) Individual tree point clouds and tree measurements from multi-platform laser scanning in German forests. *Earth Syst Sci Data* 14:2989–3012. <https://doi.org/10.5194/essd-14-2989-2022>
- White JC, Coops NC, Wulder MA, Vastaranta M, Hilker T, Tompalski P (2016) Remote Sensing Technologies for Enhancing Forest Inventories: A Review. *Can J Remote Sens* 42:619–641. <https://doi.org/10.1080/07038992.2016.1207484>
- Wulder M, Niemann KO, Goodenough DG (2000) Local Maximum Filtering for the Extraction of Tree Locations and Basal Area from High Spatial Resolution Imagery. *Remote Sens Environ* 73:103–114. [https://doi.org/10.1016/S0034-4257\(00\)00101-2](https://doi.org/10.1016/S0034-4257(00)00101-2)
- Wulder MA, Bater CW, Coops NC, Hilker T, White JC (2008) The role of LiDAR in sustainable forest management. *For Chron* 84:807–826. <https://doi.org/10.5558/tfc84807-6>
- Yao W, Krull J, Krzystek P, Heurich M (2014) Sensitivity Analysis of 3D Individual Tree Detection from LiDAR Point Clouds of Temperate Forests. *Forests* 5:1122–1142. <https://doi.org/10.3390/f5061122>
- Zhang W, Qi J, Wan P, Wang H, Xie D, Wang X, Yan G (2016) An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sensing* 8:501. <https://doi.org/10.3390/rs8060501>
- Zhao Y, Hao Y, Zhen Z, Quan Y (2017) A Region-Based Hierarchical Cross-Section Analysis for Individual Tree Crown Delineation Using ALS Data. *Remote Sensing* 9:1084. <https://doi.org/10.3390/rs9101084>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.