

Article

Tree Stem and Height Measurements using Terrestrial Laser Scanning and the RANSAC Algorithm

Kenneth Olofsson *, Johan Holmgren and Håkan Olsson

Department of Forest Resource Management, Swedish University of Agricultural Sciences, SLU Skogsmarksgränd, Umeå 901 83, Sweden; E-Mails: johan.holmgren@slu.se (J.H.); hakan.olsson@slu.se (H.O.)

* Author to whom correspondence should be addressed; E-Mail: kenneth.olofsson@slu.se; Tel.: +46-90-786-8484.

Received: 3 March 2014; in revised form: 29 April 2014 / Accepted: 6 May 2014 /

Published: 12 May 2014

Abstract: Terrestrial laser scanning is a promising technique for automatic measurements of tree stems. The objectives of the study were (1) to develop and validate a new method for the detection, classification and measurements of tree stems and canopies using the Hough transformation and the RANSAC algorithm and (2) assess the influence of distance to the scanner on the measurement accuracy. Tree detection and stem diameter estimates were validated for 16 circular plots with 20 m radius. The three dominating tree species were Norway spruce (*Picea abies* L. Karst.), Scots pine (*Pinus sylvestris* L.) and birch (*Betula* spp.). The proportion of detected trees decreased as the distance to the scanner increased and followed the trend of decreasing visible area. Within 10 m from the scanner, the proportion of detected trees was 87% on average for the plots and the diameter at breast height was estimated with a relative root-mean-square-error (RMSE) of 14%. The most accurate diameter measurements were obtained for pine, which had a RMSE of 7% for all the full 20 m radius plots. The RANSAC algorithm reduced noise and made it possible to obtain reliable estimates.

Keywords: terrestrial laser scanning; stem; canopy; stem diameter; tree height; RANSAC

1. Introduction

Advanced models are currently being used to support sustainable management of forest resources with consideration to several values such as timber production, pulpwood, biodiversity, and bioenergy, for example in the forest management planning system Heureka [1]. The modelling is usually based on individual trees and the forest tree layer is projected to future time periods using growth models in order to evaluate the consequences of alternative forest management schedules at both the local and regional level. There are at least two limitations of the current manual field inventory methods used for retrieving tree data: (1) only a small portion of all trees are measured on sample plots; and (2) there are no detailed measurements of individual trees.

New ground-based sensors have recently been developed. Currently the most promising technology for automated measurements of tree stems is terrestrial laser scanning (TLS). Such systems are rapidly becoming less expensive and more compact while at the same time the measurement frequency is increasing, allowing for very detailed measurements of a field plot within a few minutes. A TLS system measures the distance to surrounding objects with mm-resolution based on the emission of laser light and the detection of reflected signals. The processing tasks include finding the positions of individual tree stems followed by measurements of the shape of the stems and preferably also the associated canopies. There are two common approaches when scanning a field plot: single scan and multiple scan. In the single scan approach, the sensor is positioned in the middle of the field plot and only scans one side of the trees. In the multiple scan approach, several positions of the field plot is scanned giving multiple views of the trees, and the data from the different views need to be rectified to a common coordinate system.

There are four main functions needed in an algorithm for automatic detection and modelling of single trees in TLS-data:

- (1) Extract a digital elevation model of the ground (DEM);
- (2) Delineate the data points for each single tree in the investigated area;
- (3) Classify the data points of each tree into different categories such as stem, branches and foliage;
- (4) Model the detected trees using the derived point sets of different classes.

1.1. Extraction of a DEM from TLS Data Points

Some algorithms for single tree detection and modeling of TLS data need a DEM as input in order to estimate the breast height level of the trees. Many of them create a raster where the height value of the lowest data point within each cell is recorded as the height of the ground in that position. Simonse *et al.* [2] for instance created a raster with $50 \times 50 \text{ cm}^2$ cell sizes. Numerous algorithms have been developed for creating digital elevationmodels from airborne laser scanning data [3]; some of them can be adapted for TLS-data.

1.2. Delineation of Single Trees in TLS-Data

There are a number of techniques to delineate single trees in a TLS-dataset of a forest field plot. One method is to find the part of the stem closest to the ground to get an initial position. Simonse *et al.* [2]

for instance used the Hough transform to find circles using the dataset cut out at 1.25–1.35 m above ground. Every detected circle was assumed to be a stem cross-section of a tree. The data points within a cylinder of the radius expanded by 10% was chosen and modelled. Bienert *et al.* [4] also used a 10 cm slice of the dataset at breast height to search for trees. A search window with a threshold detected clusters of points. If the clusters were separated by more than the window size they were treated as separate data sets (data points belonging to a single tree). Király and Brolly [5] used an iterative method of clustering data points in different height layers to delineate stems. Another method they tested was modified k-means clustering [6]. Lindberg *et al.* [7] used the Hough transform to get an initial position for the tree modeling algorithm.

Erikson and Vestlund [8] found line segments in the range image from the lidar data. Elongated and vertical structures were assumed to be stems. Forsman and Halme [9] and Pueschel *et al.* [10] also used range images to find trees.

One way to classify and delineate point clouds is to find the characteristics of local spatial point distributions. They can be classified as scattered, linear or flat using the eigenvalues. Lalonde *et al.* [11] for instance used a Gaussian Mixture Model and the EM algorithm to train the system to recognize and group points belonging to stems, ground and foliage. Liang *et al.* [12,13] used a similar technique where flat objects with the surface normal pointing in the horizontal plane were assumed to be stems. Similar flat objects positioned vertically were grouped. Rutzinger *et al.* [14] used the standard deviation of elevation and the point density ratio in a given height interval to separate trees from the ground vegetation.

Filling a three dimensional raster (a voxel space) with point density values has also been used to detect tree stems in TLS-data. Vonderach *et al.* [15] detected enclosed regions (the inside of a tree stem) by using a technique similar to tomography.

There are also examples of TLS equipment where the beam width is at the same scale as the width of the stems. Strahler *et al.* [16] modelled the response of a tree using such an instrument.

1.3. Classification of TLS Data Points of the Tree

Once the data points of the tree are delineated they can be further classified into stem, branches, twigs and foliage. Gorte and Pfeiffer [17] and Gorte and Winterhalder [18] used a voxel space of lidar point density and 3D mathematical morphology to find the stem and branches of a tree. Each data point was assigned to the closest branch. Clawges *et al.* [19] and Côté *et al.* [20] separated woody parts of the tree and foliage by the intensity of the lidar returns and Raumonon *et al.* [21] separated the point sets of the branches and twigs by a bifurcation recognition process.

1.4. Modeling the Tree

When extracting forestry parameters from the point clouds there is often a need to model the scanned tree. Different studies have different refinement of these models: from a simple measure of diameter at breast height (DBH) to a model with twigs and leafs. The most common technique is to fit circles or cylinders to data points belonging to the stem or the branches by non linear, or linearised,

least square adjustment [2,6,7,9–11,21–28]. Some studies use robust shape fitting of cylinders like Liang *et al.* [12,13] who used Tukey's estimator.

There are a number of techniques other than circle least square adjustment used, such as the crescent moon method by Király and Broly [5] and Wezyk *et al.* [29] who used the convex hull of the dataset. Binney and Sukhatme [30] build an iterative generative statistical model of trees. Rutzinger *et al.* [14] used alpha shapes to extract tree variables and Vonderach *et al.* [15] used the cross-section area to calculate the equivalent diameter as an approximation of the DBH. To get more details of the tree stem, some methods create triangular irregular networks [22,31] out of the data points or use splines [24] for instance, to model ovality.

Using the large footprint TLS equipment Echidna, Lovell *et al.* [32] modelled the response to get an estimate of the DBH. Côté *et al.* [20] used skeletonizing of the point cloud by the kNN method to get the branch structure and attractors to foliage points to build a detailed model of a tree. Bucksch *et al.* [33,34] also used skeleton algorithms to build advanced models of trees, as well as Eysn *et al.* [35].

The geometric tree parameters like DBH, volume and height are not the only ones that are modelled. For instance, leaf area index (LAI) [16,19,36–38], wood quality [39] and chlorophyll content [40] have been extracted.

1.5. Objectives

In order to use TLS for forest applications, algorithms should be developed and validated also for trees with heavy branching (e.g., spruce trees). The modified RANSAC algorithm validated in this study has the potential to produce reliable estimates based on laser returns that are from both tree stem and branches. From a production point of view it is also beneficial if there are as few measurements as possible. This makes it interesting to explore the potential of using a single scan setup in TLS field plot inventory. The objectives of this study were (1) to develop and validate a new method for the detection, classification and measurements of tree stems and canopies using the Hough transform and the RANSAC algorithm and (2) to assess the influence of distance to the scanner on the measurement accuracy.

2. Materials and Methods

2.1. Software

The software tools used for automatic measurements and analysis of the data were developed by the authors and implemented in C, Python and the R programming languages.

2.2. Study Area

The test site was located in a hemi-boreal forest at the Remningstorp estate in south-western Sweden (lat. 58°28'N, long. 13°38'E). The dominating tree species were Norway spruce (*Picea abies* L. Karst.), Scots pine (*Pinus sylvestris* L.) and birch (*Betula* spp.).

2.3. Field Data

Sixteen plots with 20 m radius measured between June and August 2011 were used in the validation (Table 1). The DBH was measured 1.3 m above ground level in two directions: towards the plot center and perpendicular to that direction. The mean value from these two directions was used for validation of the TLS measurements. Only trees with a DBH ≥ 40 mm were manually measured in field. The plot centers and tree positions were measured with a total station relative to reference points that were placed in open areas. The position of the reference points were measured with RTK-GPS. The height of 165 sample trees were measured with a scale hypsometer to be used in the evaluation.

Table 1. Descriptive statistics of the diameter at breast height (DBH) for the evaluation plots.

PLOT	Mean DBH (mm)	Std DBH (mm)	Min DBH (mm)	Max DBH (mm)	Density (stems/ha)	Pine (%)	Spruce (%)	Decid. (%)
1	157	67	44	284	716	0	98	2
2	271	121	44	484	660	0	100	0
3	321	105	42	559	509	9	91	0
4	181	72	42	348	1027	0	97	3
5	211	67	42	390	820	17	79	5
6	258	49	150	364	589	61	39	0
7	221	138	42	412	525	52	42	6
8	168	59	40	302	979	0	72	28
9	185	166	39	565	907	23	42	35
10	239	77	50	360	637	0	82	18
11	346	79	168	672	358	0	98	2
12	269	71	66	438	660	0	92	8
13	186	65	42	326	1042	0	93	7
14	316	127	45	516	509	30	58	12
15	191	62	46	362	1042	0	77	23
16	226	69	44	346	875	0	93	7

2.4. TLS Data

A Leica ScanStation C10 was used with the high scanning mode with a distance of 0.5 cm between measurements in horizontal and vertical direction, at a distance of 10 m. The scanner had a green laser, 532 nm, a beam divergence of 0.1 mrad, a scanning pulse rate up to 50 kHz, and horizontal and vertical field of view of 360 and 270 degrees, respectively. A single scan setup was used with the scanner in the center of the field plot.

2.5. Estimation of Ground Model

In order to measure the height above ground a DEM was first estimated from the TLS data. The algorithm used to acquire the DEM was based on the idea of sampling the laser data in two rasters with

different cell sizes in order to separate data points that belong to the ground from data points that belong to the canopy. For details see Lindberg *et al.* [7].

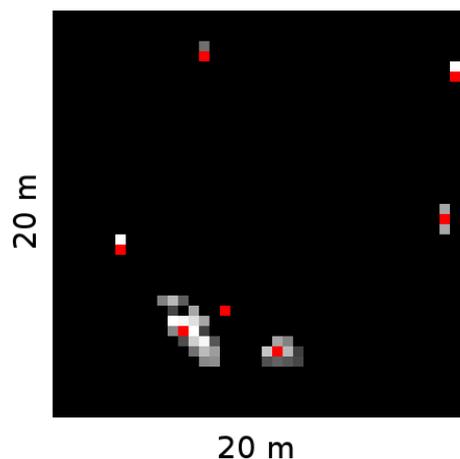
2.6. Detection of Tree Stem Positions

The assumption used for the tree stem detection algorithm was that the trunk of a tree is continuous in the vertical direction whereas branches and low vegetation is not continuous in the vertical direction. A raster in the x, y-plane was created with a cell size large enough to allow for several laser returns within each cell that originates from data points along a vertical tree stem. The cell size was set to 0.5 m. The z-value of each laser return was normalized using the DEM to obtain the vertical distance to the ground, *i.e.*, the height. Only laser returns within the height interval from 1 to 2 m were selected for the task of finding tree positions. This interval was further divided into 32 height intervals to describe the height distribution of laser returns. For each raster cell, the number of laser returns located within a height interval was counted and if there was at least one return within this interval, it was registered as filled. For each raster cell the number of filled sections and the total number of laser returns was registered. The total number of laser returns in each raster cell was normalized by the distance to the sensor.

The number of filled sections was normalized to be in the interval 0–1 (fill ratio) where 1 meant that all sections between 1 to 2 m were filled and 0 that no sections were filled. The fill ratio of each raster cell was multiplied by the normalized number of laser returns to obtain a stem probability factor. This stem probability factor is high where there are many laser returns and where they are continuously distributed along the height interval.

A threshold of 25% of the maximum stem probability factor of the raster was used to remove noise. All cells with values below the threshold was set to zero. In this raster of stem probability factors (Figure 1) there were contiguous regions of cells that had values larger than zero surrounded by areas with zero values. The assumption was that these contiguous regions of cells were located around stem positions. The stem probability factor cell with the maximum value in each region was chosen to be a stem position. If there were no other stem positions within a distance of two raster cells, this was set to be a final stem position. Otherwise the neighbour tree region with the highest stem probability factor was chosen to be the final stem position.

Figure 1. A raster of the stem probability factor in grey scale. The final tree stem positions are marked in red.



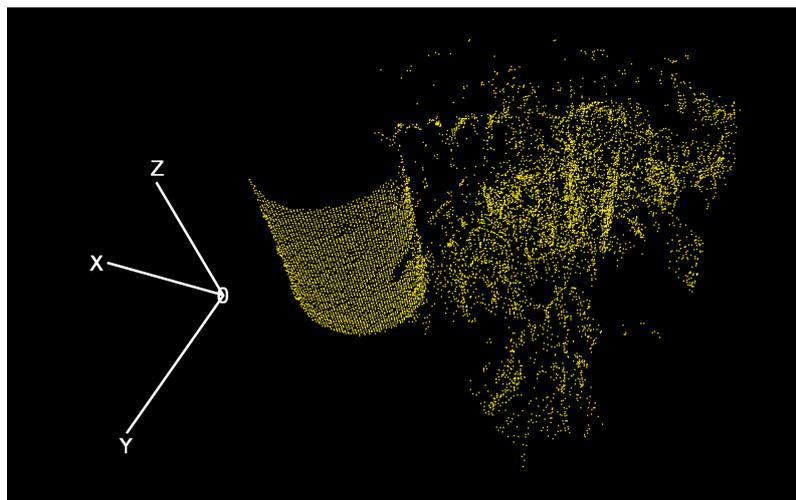
2.7. Classification of Laser Data Points to Tree Stem and Tree Canopy

A start coordinate that is close to the root of a tree was a precondition when classifying laser data points in this method. The stem positions found in the earlier stage were used for this purpose. The method was a two stage process. First, cross-section circles at different heights along the stem were found by searching upwards using a Hough transform based method. Second, the shape of the tree crown was found by searching downwards using the tree stem from the earlier stage. The assumption was that the trees had a main stem like most conifers and that they were approximately circularly symmetric. The maximum tree height in this study was assumed to be 30 m.

2.7.1. Finding Stem Cross Sections

To find the first stem cross-section, laser data points were cut out around the start position within a rectangular box with a width and depth of 1.2 m and a height interval 1–1.5 m above ground (Figure 2).

Figure 2. TLS data points cut out at 1.0–1.5 m above ground: To the left part of a tree stem and to the right part of a small tree with dense branches.

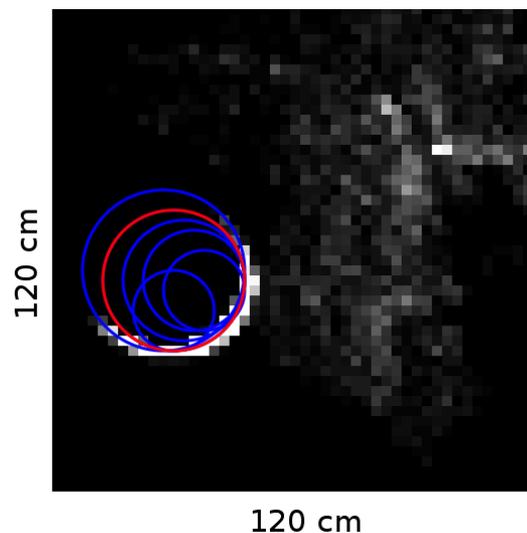


All TLS points in this subset were projected to a raster in the x, y-plane (where the z-axis is pointing from ground to sky). The resolution of the raster was 2.5 cm. Each data point was added to the corresponding raster cell giving a high pixel value where many data points are located. The pixel values of this raster were normalized to 0–255 and all values below the threshold of 64 were set to zero. The threshold level was chosen to suppress most of the noise. A stem in this raster looked like a semi-circle, since the TLS sensor only scanned one side of the tree (Figure 3).

In this raster 7 radius classes were searched for, using an increment of 2.5 cm, starting at 10 cm. Radii with values other than the classes were detected by the method but the class centers had the strongest signals. The smallest radius class was chosen to avoid false positives. Small circles tend to fit any pattern. The largest radius class was estimated to be appropriate for the investigated type of forest. The algorithm is based on the Hough transform and previously described by Lindberg *et al.* [7]. A pixel that belongs to the edge of a circle of a certain radius has a number of possible circle centres that lie along a semi-circle pointing away from the sensor. By adding every possible circle centre to a raster, for every radius class,

a probability image for a circle position was made. Positions that were inside compact objects could be removed since the sensor cannot see inside the tree stem. In addition, pixels not connected to other pixels could be removed assuming that the stem rim is continuous (for details see Lindberg *et al.* [7]). Local maxima in this probability image were assumed to be candidates to the circle center. The circle with the highest number of votes was chosen as the most probable choice. If no circle had the votes of more than 20% of what a tree stem of full intensity would give, no tree was detected for this position.

Figure 3. TLS data points projected to a raster in the x, y-plane. A high density of laser data points give a bright pixel. The coloured circles are the results from the stem finding algorithm. The red circle is the final chosen stem diameter by the algorithm.



Next TLS data were cut out from 1.5–2.0 m. The same procedure as previous was applied. If several circles were found, the one closest to the first was chosen. Now two circle centers at different heights could be connected to allow for leaning trees as well as varying thickness along the stem. If the second stem radius was 8 cm larger than the first, the stem was considered invalid and discarded. If the tree was leaning more than 45° the choice was also discarded. If the stem was positioned more than 30 cm away from the start position, it was discarded. The values of the settings were chosen after trials on test data separate from the evaluation dataset.

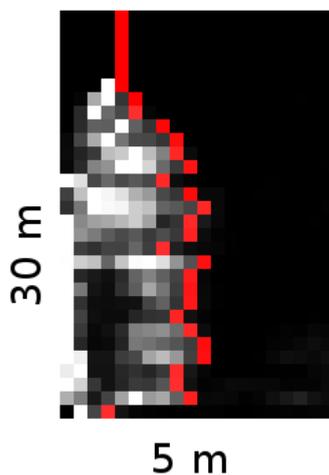
The two circles gave a new start position calculated from the leaning tree and the next height interval was processed, from 2.0–2.5 m and 2.5–3.0 m and so on. If only one or zero circles were found the whole interval 2.0–3.0 m was chosen with the assumption that twice the amount of datapoints will make it possible to fit a circle. If only one circle was fitted within the whole interval the tree was assumed to point along the z axis. The procedure continued until no more circles were found or the stem reached the maximum height 30 m.

Usually, a dataset is sparse high in the canopy making it more difficult to detect circles. To classify datapoints where the algorithm did not succeed to detect any circles, an average stem radius with a corresponding cylinder was calculated to be used where there was missing data. The circles found by the algorithm in the lower regions were averaged to estimate direction and size of the stem in the higher regions.

2.7.2. Finding Tree Crown Cross Sections

To detect the tree crown outline a radially symmetric image was produced (Figure 4). Every data point in the surroundings of the found tree stem was projected onto this image where the x-axis was the radial distance from the tree stem at the chosen height and the y-axis was the height from ground. The resolution of this image was 0.25 m in the radial direction and 1 m in the height direction. Every data point increased the corresponding pixel value by one in the image. To make data points closer to the tree stem more important, the values were scaled by dividing with the circular area at the chosen radius. Finally every height row was scaled from 0 to 1.

Figure 4. The outline of the tree crown in red. The TLS data points are projected onto a radially symmetric image. The x-axis is the radial distance from the tree stem, from left to right and the y-axis is the height from ground. The algorithm start the search from the top until the first canopy pixel is found.



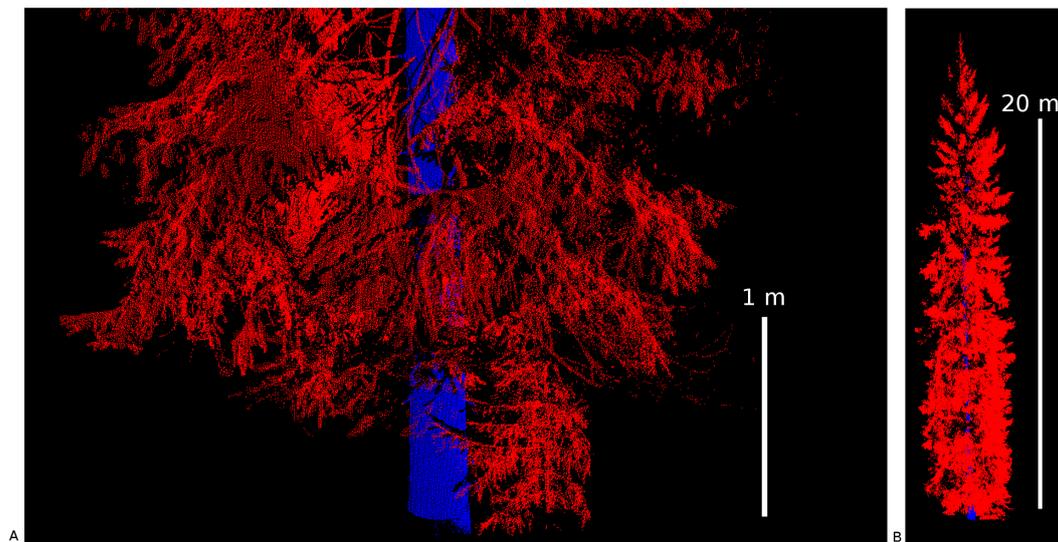
The search of the outline began at the highest point of this image at the radius 1 m and stopped when the first part of the tree top was found (a pixel value larger than the threshold 0.2 was assumed to be part of the canopy). From then on the algorithm searched down one pixel at the time. If no valid crown pixel was found (value less than the threshold) the algorithm searched towards the stem. If a valid crown pixel was found the algorithm searched radially outwards until the edge of the crown was found. No crown radius sections larger than 2.5 m were accepted in order to make the algorithm more robust. Otherwise a neighboring tree too close to the investigated one might be included in the canopy. When separating different tree canopies from each other the laser data point was assigned to the closest stem. To avoid situations where the algorithm does not find a tree top, the smallest accepted tree crown radius was 1 m at 30 m height above ground and linearly decreasing towards zero at the root of the tree. All laser points within this cone were guaranteed to belong to the tree. This way there were most often laser data from the treetop remaining for the height estimates.

2.7.3. Classifying Laser Data Points by Stem and Tree Crown Cross Sections

When all start positions were investigated a number of stem and tree crown outlines were available. To designate a TLS data point to a class and a single tree, all stem outlines were investigated. If a data

point was within a distance of (1.5 times the stem radius) from the stem center it was designated to that particular tree. If it was inside one or several tree crown outlines the closest stem gave which tree this canopy data point belonged to. All other data points were considered to be ground, low vegetation or measurement outliers (Figure 5).

Figure 5. Laser data points of a spruce classified into stem (BLUE) and canopy (RED).



2.8. The Use of Classified Laser Data Points to Extract Forest Variables

When the laser data points were designated to single trees and to the stem class or the crown class (or one of the other classes) it was possible to estimate stem attributes. In this investigation quite straightforward methods to estimate two of the variables of high interest for forestry were suggested but once the data points are classified more advanced techniques could also be used. The chosen forest variables for this study were stem diameter at breast height (*i.e.*, stem diameter at 1.3 m above ground level, DBH) and tree height.

2.8.1. Estimation of Tree Height by Percentiles

For each single tree the data points classified as tree crown were chosen. For this pointcloud the height percentiles from ground were calculated. The tree height was then estimated as the 100th height percentile.

2.8.2. Estimation of Stem Diameter by a Modified RANSAC Method

The Random Sample Consensus (RANSAC) method [41,42] is an iterative algorithm based on the fact that there are data points that belong to the model (inliers) and those that do not belong to the model (outliers). If a number of random points in the dataset is chosen, there is a probability of them all being part of the model. Equation (1) gives the number of iterations necessary to find at least one such case.

$$N = \frac{\log(1 - P)}{\log(1 - p^n)} \quad (1)$$

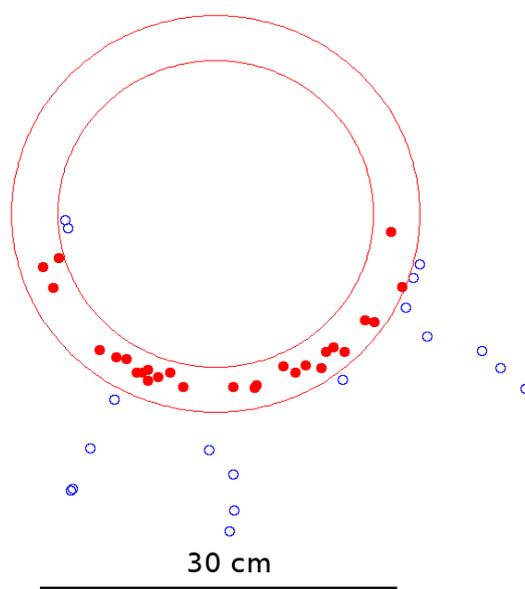
where N is the number of iterations used where at least one good model is found with the probability P . The number of chosen data points is n and p is the probability of one chosen point belonging to the model. The parameter p is given by Equation (2).

$$p = \frac{M}{D} \quad (2)$$

where M is the number of inliers in the model and D is the number of points in the dataset. Usually this ratio is not known and has to be estimated for the used dataset. By studying a typical dataset it is possible at least to obtain a rough estimate of this ratio. In the case of stem detection one would need to know approximately how many points that belong to the branches and surrounding vegetation rather than the stem itself.

Tree stems can often be modeled by circles or cylinders which are well suited forms for application of the RANSAC algorithm. When modeling a circle at least three data points are necessary. That makes parameter $n = 3$ in the RANSAC algorithm. Therefore three points were randomly chosen to fit a circle for one of the iterations. Each data point that was within 2.5 cm from this circle was considered to be an inlier (Figure 6).

Figure 6. In each iteration in the RANSAC algorithm inliers are found within a given tolerance of the chosen circle. This model is saved for further calculations if the circle is of a valid size, there are few points inside the trunk, and the number of inliers is larger than the previous chosen circle.



For the version of RANSAC modified for this study the number of data points that were inside the trunk was also calculated for each iteration, with the assumption that no measurements can penetrate the stem. If the randomly chosen circle for the iteration had more than 1% of the points inside the trunk the solution was considered invalid. If the randomly chosen circle had a smaller radius than 2 cm or a larger radius than 30 cm it was also considered invalid. In this version of the RANSAC method it was the valid iteration with the highest number of inliers that was chosen as the set to which a circle was fit.

By least square adjustment, a circle was then fitted to the inliers of the chosen iteration. The DBH of the tree was then estimated as the diameter of this chosen circle.

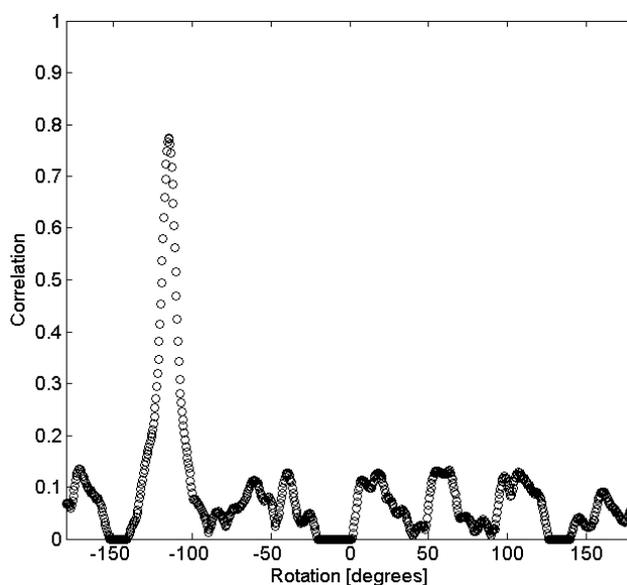
The parameters for this study were set to: $P = 0.99$, $p = 0.80$, and $n = 3$. The number of iterations was $100N$.

2.9. Linking Detected Tree Stems to Manually Measured Reference Trees

The tree positions were manually measured relative to reference points which were measured with a global navigation satellite system (GNSS). The manual measurements of trees could therefore be geo-referenced. However, the TLS system was placed near the center of the field plots with no targets used for geo-referencing. Instead, an earlier developed algorithm (Olofsson *et al.* [43]) was applied to match tree stems detected from the laser data with the positions of manually measured trees. The algorithm used the center of the field plot as an estimated position of the scan center to start the search. From the start position, TLS single tree data from within the search area was collected as a list. This list contained the tree position x , y -coordinates, and the stem diameter. For the manually measured trees a similar list was required with the positions of the trees and stem diameter. The tree lists were used to create two tree position images. Within the image, each tree was displayed as a Gaussian surface where the x and y coordinates determined the position within the image, the tree size variable determined the amplitude of the Gaussian function, and the standard deviation was set to the expected tree position precision (1 m in this study). The rotation was $\pm 180^\circ$ with 0.5° steps and the translation was ± 2 m with 0.25 m steps. The rotation and translation which yielded the maximum correlation between the two tree position images was then used.

The correlations that were a result of different rotations of the scanners local coordinate system were plotted for different rotation angles (Figure 7). A sharp peak could be observed when a correlation value was much higher than for other rotation angles, indicating a rotation with a high probability of being correct.

Figure 7. The correlation for different rotation angles used to match TLS detected trees with manually measured tree positions for one example field plot.



Once the tree coordinates from the TLS were transformed to the global coordinates from the manual measurements, the trees detected by the TLS approach were linked to the manually measured trees. To limit the size of the calculations only trees detected by the TLS approach close enough to the manually measured trees were included for a possible search. A search radius of 1.5 m was used to include a possible link between trees. To measure the quality of a link a weight (w) was calculated based on the Euclidean horizontal distance between trees (d), Equation (3).

$$w = \frac{1}{(1 + d)^2} \quad (3)$$

In the tree link list the algorithm then searched for connected tree clusters, *i.e.*, a group of TLS detected and manually measured trees that were linked together. Each link only contained one TLS detected and one manually measured tree, but since each TLS detected tree could be linked to several manually measured trees and each manually measured tree could be linked to several TLS detected trees, a network of connections became a cluster of trees, a tree list graph. Since every TLS detected tree should only be connected to one manually measured tree, multiple links were removed from the list. The algorithm solved this by trying every possible combination of links in the tree clusters and combinations with multiple links were discarded. The link combination with the highest sum of weights was the solution that was chosen for each tree cluster and all other links were removed from the list.

The linked trees were used to evaluate the tree detection results and the estimation of the stem diameter. Also, estimation of tree heights was validated using the subset of trees with manually measured heights.

The proportion of detected tree stem positions was calculated as the number of TLS detected trees that was linked to a field tree divided by the total number of field trees. The proportion of detected stem cross sections was calculated as the number of TLS detected trees where the algorithm managed to estimate the stem diameter, and link to a field tree, divided by the total number of field trees.

2.10. Calculating the Shaded Area of Trees in a Plot

The shaded area behind a tree was calculated by Equation (4), where (r) is the stem radius, (R) is the distance from the sensor to the tree and (θ) is the angle of the circle sector, assuming ($R \gg r$). The shaded area starts behind the tree at the radial distance ($R + r$) and ends at the radius of the field plot. This is similar to the shaded areas in the Seidel and Ammer study [44].

$$\theta = 2 \arctan \frac{r}{R} \quad (4)$$

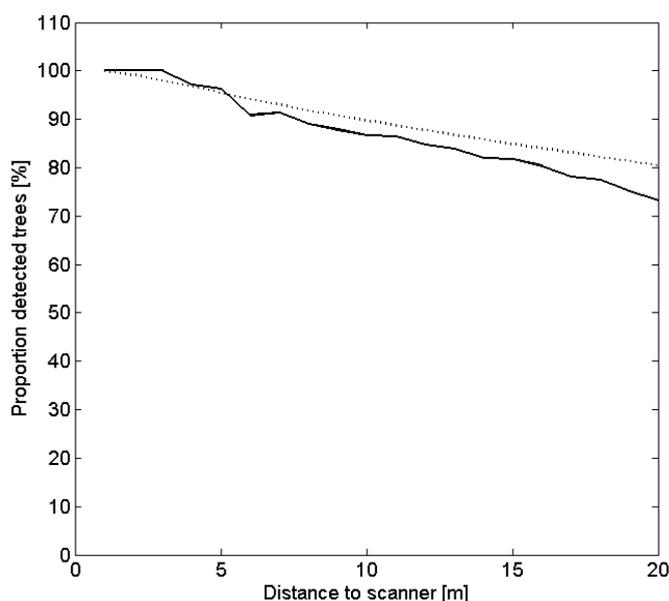
All shaded areas from the trees were saved to a raster with resolution 1 dm². If a pixel center was inside one or more of the shaded areas it was set as a shaded pixel. This way the shaded areas of the trees were allowed to partly cover each other. The number of shaded pixels inside a plot gave the total shaded area of the plot. The non shaded area of the plot was the complementary to the shaded area. The average of the non shaded areas of all plots was calculated to give the shading effect for the region.

3. Results

3.1. Detected Proportion of Tree Stem Positions

The detected proportion of tree stem positions decreased with increasing distance from the scanner (Figure 8). Because this trend might be caused by trees obscured by other trees, the non-shaded proportion of the field plot area was also plotted against the distance to the scanner (Figure 8).

Figure 8. The proportion of detected trees (solid line) and proportion of non-shaded area for all plots (dotted line) plotted against distance to the scanner.



3.2. Proportion of Detected Stem Cross Sections

The proportion of detected stem cross-sections decreased as a function of distance to the scanner. Larger trees were easier to detect. The proportion of detected cross-sections was lower compared with detection of tree stem positions. The difference was, however, small for large trees (Table 2).

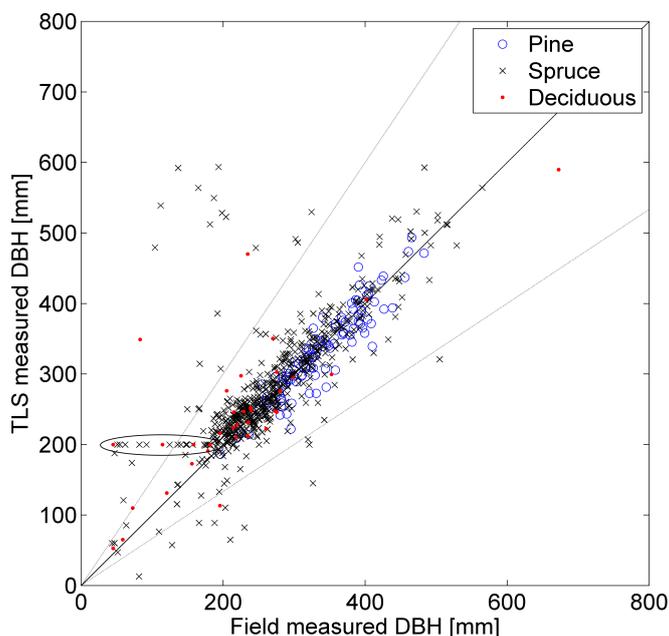
Table 2. Proportion (%) of detected trees at different distances for cross-sections (cs.) of different sizes and tree stem positions (stems).

Distance (m)	Prop. cs. all	Prop. cs. > 100 mm	Prop. cs. > 200 mm	Prop. stems
0–5	64.5	72.5	95.4	96.3
0–10	58.8	67.3	87.0	86.6
0–15	55.8	63.3	82.9	81.8
0–20	50.8	58.8	74.8	73.3

3.3. Estimation of Stem Diameter

To see the effect of gross errors the data set was analyzed both with and without outliers removed. All datapoints above the line with slope 3/2 and below the line with slope 2/3 were considered to be outliers (Figure 9).

Figure 9. Scatterplot for TLS estimated versus field measured diameter at breast height (DBH). All points outside the dotted lines were denoted as outliers. For the points at DBH = 200 mm, inside the ellipse, the RANSAC algorithm was not able to give any results and therefore the values were set to the diameter class obtained by the Hough transform.



The estimate of the DBH had a relative RMSE of 12%–21% where trees close to the scanner had smaller errors (Table 3). However, the error was only 7.1% for Scots pine trees. The stem estimation errors were smallest for Scots pine trees and largest for deciduous trees (Table 4). There was a small negative bias (*i.e.*, underestimation) of stem diameter for Scots pine. However, the stem diameter was overestimated for Norway spruce, and to the highest degree for deciduous trees (Table 4). The errors decreased for most plots after the removal of outliers (Table 5).

Table 3. Stem diameter estimates at different distances, (I) complete dataset, and (II) dataset with the outliers removed.

Distance (m)	Set	RMSE (mm)	relRMSE (%)	BIAS (mm)	relBIAS (%)
0–5	I	33	11.9	9	3.5
0–5	II	26	9.6	9	3.2
0–10	I	39	14.2	10	3.6
0–10	II	27	9.9	7	2.7
0–20	I	59	21.4	13	4.7
0–20	II	28	10.4	6	2.2

Table 4. Stem diameter estimates for different tree species, (I) complete dataset, and (II) dataset with the outliers removed.

Species	Set	RMSE (mm)	relRMSE (%)	BIAS (mm)	relBIAS (%)
Pine	I	23.8	7.1	−6.0	−1.8
Pine	II	23.8	7.1	−6.0	−1.8
Spuce	I	61.8	23.2	15.4	5.8
Spuce	II	28.8	10.6	8.3	3.1
Deciduous	I	74.8	34.1	25.2	11.5
Deciduous	II	34.4	14.4	6.9	2.9

Table 5. Stem diameter estimates for each plot, (I) complete dataset, and (II) dataset with the outliers removed.

Plot	RMSE I (mm)	relRMSE I (%)	RMSE II (mm)	relRMSE II (%)
1	64	23.5	39	14.1
2	96	34.9	38	13.7
3	42	15.2	42	15.2
4	66	24.2	29	10.4
5	50	18.4	21	7.8
6	20	7.2	20	7.2
7	38	14.0	21	7.8
8	93	34.0	34	12.3
9	29	10.7	29	10.7
10	78	28.3	38	13.8
11	33	12.2	33	12.2
12	36	13.1	27	9.9
13	80	29.4	34	12.4
14	38	14.0	38	14.0
15	51	18.7	37	13.6
16	41	14.9	29	10.5

3.4. Estimation of Tree Height

A small portion of the tree heights were severely overestimated by the TLS measurements. However, the tree heights were on average underestimated for all sample trees (Table 6).

Table 6. Tree height estimates.

Distance (m)	RMSE (m)	relRMSE (%)	Bias (m)	relBias (%)
0–10	4.3	17.6	−0.1	−0.5
0–20	4.9	20.3	−0.1	−0.6

4. Discussion

This study suggests a method to automatically detect trees in a TLS-scanned field plot, classify the laser returns into stem or canopy for each of the detected trees, and extract forest variables from the classified data.

The proportion of detected tree stems decreases as a function of distance to the scanner and follows the proportion of non-shaded area. Thus, there is a high probability of detecting a tree with the proposed method if the tree is not obscured by other trees. The problems with obscured trees in single scans have also been observed in previous studies. For example, Thies and Spiecker [45] detected 52% of the trees in multiple scans but only 22% in single scans. Maas *et al.* [46] reported a success rate of 97.5% on plots with both multiple and single scans. However, the stem density on the four field plots used for validation was low with only 15–29 trees on 12–15 m radius plots. Liang *et al.* [13] validated an algorithm for detection of trees using single scan data in 9 plots and with higher stem densities: on average 1022 stems/ha compared with 741 stems/ha in this study. They obtained a detection accuracy of 73% on 10 m radius plots. This is similar to the number obtained in this study where the plot radius was 20 m. In this study, 87% of the trees within 10 m were detected. However, this higher detection rate as compared with the study by Liang *et al.* [13] is probably a result of different stem densities. The assumption that the detection rate follows the decreasing proportion of non-shaded area as the distance to the scanner increases could be confirmed in this study. That is probably the reason why detection rates close to 100% are reported in earlier studies where data from multiple scans have been used. Because the detection rate and proportion of non-shaded area are similar, one conclusion is that the detection rate in this study was high with almost 100% for non-shaded areas. Reported commission errors in earlier studies are usually small. In this study the commission error was only 5.5%. For example, Maas *et al.* [46] obtained a commission error of 6.1% (5 false tree detections and 82 field trees). Liang *et al.* [13] obtained a commission error of 14.5% (42 false tree detections and 289 field trees) but that was for a dense forest, probably with more branches and low vegetation.

In a second step, stem cross-sections were estimated using the Hough-transform. The detected proportion was low for small trees but higher for large trees. However, detection of stem cross-sections was a first step to obtain estimates of the entire tree stem and canopy. A complete description of the entire tree stem for large trees is what is of interest for forestry applications. These trees are used for wood assortments while small trees, which require less detailed data, are either not harvested or are used for bioenergy or pulp wood production.

The overall accuracy of stem diameter estimates was somewhat lower than reported from earlier studies using similar scanner data and circle fit techniques according to a review of recent studies by Pueschel *et al.* [10]. However, it should be noted that the stem diameter estimates in this study were performed using 16 plots with 20 m radius with a variety of site conditions, including heavily branched spruce trees, and high stem densities compared with other studies. In earlier studies, the stem diameter estimates have been obtained on a few plots with sparse, low growing vegetation and few branches at breast height. For example, Maas *et al.* [46] reported a RMSE of 1.4–3.25 cm for plots with clearly visible tree stems and in forest with a low stem density (15–29 trees on 12–15 radius plots). There are probably two main reasons for large errors for some trees in this study: (1) trees in dense

stands are sometimes partly shaded by other trees; and (2) branches could be confused with tree stems by the algorithm. The relative RMSE was only 7.1% for Scots pine trees which is in line with this assumption: Scots pine trees usually have few branches at breast height and the stands containing pine trees are usually sparser, resulting in less shaded parts of trees. A modified RANSAC algorithm was implemented in this study to overcome the problems with outliers caused by heavy branching. However, the RANSAC followed by the circle fit was only applied for a limited height section of the tree (at breast height). A further development could be to combine several estimates at different heights to increase the reliability of a single estimate, Pueschel *et al.* [10]

In this study, a positive bias was observed for Norway spruce and deciduous trees. One reason for this bias could be branches and low vegetation which was more common on plots dominated by these tree species. A small negative bias was observed for Scots pine trees that usually are rough-barked compared with trees in the two other tree species groups. Brolly and Kiraly [6] observed an underestimation for the rough-barked larch trees but not for the smoothed-barked beech trees and suggested that this could be because also rifts in the bark are measured by the laser.

In this study, the RMSE increased with the distance from the scanner but this was caused by a small proportion of outliers. The number of trees increases proportional to the square of the distance and therefore the probability to include outliers is higher at long distances. There was no clear relationship between RMSE and distance after removal of the outliers with the above described boundaries. Also Lindberg *et al.* [7] and Pueschel *et al.* [10] observed no influence of the distance to the scanner. The cutouts of raw data for some of the outliers followed by visual inspection revealed the reason for the large DBH errors. The cutouts made by the Hough-transform were not from the tree stem but parts of the canopy. Thus, the cutout-algorithm should be improved.

One error source could also be the automatic linking applied in this study. The position and orientation of the scanner were calculated based on correlation between tree patterns and this was followed by one-to-one linking of TLS and manual measured trees. All trees were not detected in TLS data which makes the tree linking problem more difficult.

The tree height estimates are similar to what has been obtained in earlier studies. For example, Maas *et al.* [46] reported an RMSE of 4.55 m. In this study the overall height accuracy was decreased severely by a few estimates with overestimation, which was probably due to small trees growing close to taller trees. Hence, the crown delineation algorithm should be improved to also produce reliable results in dense forests. The main error sources in the TLS height estimates are that: the top of the tree crown can be shaded by itself or a neighbour tree, part of the crown of a higher neighbor tree is assigned as the top of the tree of interest, or the trees are not correctly linked in the evaluation.

The idea of classifying laser data points into stem and canopy is promising and a useful technique. Once classified data are accessible, advanced techniques for the extraction of forest variables are possible to apply, giving an opportunity to measure stem quality. Pfeiffer and Winterhalder [24] for instance distinguished detailed shape of the stems and measured properties like ovality and Thies *et al.* [25] implemented methods to access taper, sweep and lean from TLS-data. Terrestrial laser scanners could therefore be used where there is a need for detailed measurements of individual trees in field plots. Crown dimensions, competition and other parameters have also been useful in forest research based on isolated point clouds of trees, such as: Bayer *et al.* [47] and Seidel *et al.* [48].

5. Conclusions

This study suggests a method to automatically detect trees in a TLS-scanned field plot, classify the laser returns into stem or canopy for each of the detected trees, and extract forest variables from the classified data. The performance of the random sample consensus method (RANSAC) for stem detection in a forest from southern Scandinavia was shown. The assumption that the detection rate follows the decreasing proportion of non-shaded area as the distance to the scanner increases could be confirmed. The detection rate was almost 100% for non-shaded areas. A further development could be to combine several estimates of the stem diameter at different heights to increase the reliability of a single estimate.

Acknowledgments

This study was financed by Nils och Dorthi Troëdssons research foundation and Formas. The field data and TLS data were financed by Sven and Hildur Wingquist foundation for forest research.

Author Contributions

All co-authors contributed to the scientific content and authorship of this manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Wikström, P.; Edenius, L.; Elfving, B.; Eriksson, L.; Lämås, T.; Sonesson, J.; Öhman, K.; Wallerman, J.; Waller, C.; Klintebäck, F. The Heureka forestry decision support system: An overview. *Math. Comput. For. Nat.-Resour. Sci.* **2011**, *3*, 87–94.
2. Simonse, M.; Aschoff, T.; Spiecker, H.; Thies, M. Automatic Determination of Forest Inventory Parameters Using Terrestrial Laser Scanning. In Proceedings of the Scandlaser Scientific Workshop on Airborne Laser Scanning of Forests, Umeå, Sweden, 3–4 September 2003; pp. 251–257.
3. Meng, X.; Currit, N.; Zhao, K. Ground filtering algorithms for airborne LiDAR data: A review of critical issues. *Remote Sens.* **2010**, *2*, 833–860.
4. Bienert, A.; Scheller, S.; Keane, E.; Mullooly, G.; Mohan, F. Application of Terrestrial Laser Scanners for the Determination of Forest Inventory Parameters. In Proceedings of the ISPRS Commission V Symposium ‘Image Engineering and Vision Metrology’, Dresden, Germany, 25–27 September 2006.
5. Király, G.; Brolly, G. Tree height estimation methods for terrestrial laser scanning in a forest reserve. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2007**, *36*, 3, 211–215.
6. Brolly, G.; Király, G. Algorithms for stem mapping by means of terrestrial laser scanning. *Acta Silv. Lignaria Hung.* **2009**, *5*, 119–130.
7. Lindberg, E.; Holmgren, J.; Olofsson, K.; Olsson, H. Estimation of stem attributes using a combination of terrestrial and airborne laser scanning. *Eur. J. For. Res.* **2012**, *131*, 1917–1931.

8. Erikson, M.; Vestlund, K. Finding Tree-Stems in Laser Range Images of Young Mixed Stands to Perform Selective Cleaning. In Proceedings of the Scandlaser Scientific Workshop on Airborne Laser Scanning of Forests, Umeå, Sweden, 3–4 September 2003; pp. 244–250.
9. Forsman, P.; Halme, A. 3-D mapping of natural environments with trees by means of mobile perception. *IEEE Trans. Robot.* **2005**, *21*, 482–490.
10. Pueschel, P.; Newnham, G.; Rock, G.; Udelhoven, T.; Werner, W.; Hill, J. The influence of scan mode and circle fitting on tree stem detection, stem diameter and volume extraction from terrestrial laser scans. *ISPRS J. Photogramm. Remote Sens.* **2013**, *77*, 44–56.
11. Lalonde, J.F.; Vandapel, N.; Huber, D.; Hebert, M. Natural terrain classification using three-dimensional lidar data for ground robot mobility. *J. Field Robot.* **2006**, *23*, 839–861.
12. Liang, X.; Litkey, P.; Hyypä, J.; Kaartinen, H.; Vastaranta, M.; Holopainen, M. Automatic Stem Location Mapping Using TLS for Plot-Wise Forest Inventory. In Proceedings of the SilviLaser, College Station, TX, USA, 14–16 October 2009.
13. Liang, X.; Litkey, P.; Hyypä, J.; Kaartinen, H.; Vastaranta, M.; Holopainen, M. Automatic stem mapping U using single-scan terrestrial laser scanning. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 661–670.
14. Rutzinger, M.; Pratihast, A.K.; Oude Elberink, S.; Vosselman, G. Detection and modelling of 3D trees from mobile laser scanning data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2010**, *38*, 520–525.
15. Vonderach, C.; Voegtle, T.; Adler, P. Voxel-base approach for estimating urban tree volume from terrestrial laser scanning data. *Int. Arch. Photogr. Remote Sens. Spat. Inf. Sci.* **2012**, *XXXIX-B8*, 451–456.
16. Strahler, A.H.; Jupp, D.L.; Woodcock, C.E.; Schaaf, C.B.; Yao, T.; Zhao, F.; Yang, X.; Lovell, J.; Culvenor, D.; Newnham, G.; *et al.* Retrieval of forest structural parameters using a ground-based lidar instrument (Echidna). *Can. J. Remote Sens.* **2008**, *34*, S426–S440.
17. Gorte, B.; Pfeifer, N. Structuring laser-scanned trees using 3D mathematical morphology. *Int. Arch. Photogramm. Remote Sens.* **2004**, *35*, 929–933.
18. Gorte, B.; Winterhalder, D. Reconstruction of laser-scanned trees using filter operations in the 3D raster domain. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *36-8/W2*, 39–44.
19. Clawges, R.; Vierling, L.; Calhoun, M.; Toomey, M. Use of a groundbased scanning lidar for estimation of biophysical properties of western larch (*Larix occidentalis*). *Int. J. Remote Sens.* **2007**, *28*, 4331–4344.
20. Côté, J.F.; Widlowski, J.L.; Fournier, R.A.; Verstraete, M.M. The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. *Remote Sens. Environ.* **2009**, *113*, 1067–1081.
21. Raunonen, P.; Kaasalainen, M.; Åkerblom, M.; Kaasalainen, S.; Kaartinen, H.; Vastaranta, M.; Holopainen, M.; Disney, M.; Lewis, P. Fast automatic precision tree models from terrestrial laser scanner data. *Remote Sens.* **2013**, *5*, 491–520.
22. Aschoff, T.; Thies, M.; Spiecker, H. Describing forest stands using terrestrial laser-scanning. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *35*, 237–241.

23. Pfeifer, N.; Gorte, B.; Winterhalder, D. Automatic Reconstruction of Single Trees from Terrestrial Laser Scanner Data. In Proceedings of 20th ISPRS Congress, Istanbul, Turkey, 12–23 July 2004; pp. 114–119.
24. Pfeifer, N.; Winterhalder, D. Modelling of tree cross sections from terrestrial laser scanning data with free-form curves. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *36-8/W2*, 76–81.
25. Thies, M.; Pfeifer, N.; Winterhalder, D.; Gorte, B.G.H. Three-dimensional reconstruction of stems for assessment of taper, sweep and lean based on laser scanning of standing trees. *Scand. J. For. Res.* **2004**, *2004*, 571–581.
26. Hopkinson, C.; Chasmer, L.; Young-Pow, C.; Treitz, P. Assessing forest metrics with a ground-based scanning lidar. *Can. J. For. Res.* **2004**, *34*, 573–583.
27. Watt, P.J.; Donoghue, D.N.M. Measuring forest structure with terrestrial laser scanning. *Int. J. Remote Sens.* **2005**, *26*, 1437–1446.
28. Henning, J.G.; Radtke, P.J. Detailed stem measurements of standing trees from ground-based scanning lidar. *For. Sci.* **2006**, *52*, 67–80.
29. Wezyk, P.; Koziol, K.; Glista, M.; Pierzchalski, M. Terrestrial laser scanning versus traditional forest inventory first results from the polish forests. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2007**, *3*, 424–429.
30. Binney, J.; Sukhatme, G.S. 3D Tree Reconstruction from Laser Range Data. In Proceedings of the IEEE International Conference on Robotics and Automation, Robotics and Automation (ICRA'09), Kobe, Japan, 12–17 May 2009; pp. 1321–1326.
31. Aschoff, T.; Spiecker, H. Algorithms for the automatic detection of trees in laser scanner data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *36-8/W2*, 71–75.
32. Lovell, J.; Jupp, D.; Newnham, G.; Culvenor, D. Measuring tree stem diameters using intensity profiles from ground-based scanning lidar from a fixed viewpoint. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 46–55.
33. Bucksch, A.; Lindenbergh, R. CAMPINO—A skeletonization method for point cloud processing. *ISPRS J. Photogramm. Remote Sens.* **2008**, *63*, 115–127.
34. Bucksch, A.; Lindenbergh, R.; Menenti, M. SkelTre: Robust skeleton extraction from imperfect point clouds. *Vis. Comput.* **2010**, *26*, 1283–1300.
35. Eysn, L.; Pfeifer, N.; Ressel, C.; Hollaus, M.; Graf, A.; Morsdorf, F. A practical approach for extracting tree models in forest environments based on equirectangular projections of terrestrial laser scans. *Remote Sens.* **2013**, *5*, 5424–5448.
36. Hosoi, F.; Omasa, K. Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning lidar. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 3610–3618.
37. Moorthy, I.; Miller, J.R.; Hu, B.; Chen, J.; Li, Q. Retrieving crown leaf area index from an individual tree using ground-based lidar data. *Can. J. Remote Sens.* **2008**, *34*, 320–332.
38. Jupp, D.L.; Culvenor, D.; Lovell, J.; Newnham, G.; Strahler, A.; Woodcock, C. Estimating forest LAI profiles and structural parameters using a ground-based laser called ‘Echidna(R)’. *Tree Physiol.* **2008**, *29*, 171–181.

39. Schütt, C.; Aschoff, T.; Winterhalder, D.; Thies, M.; Kretschmer, U.; Spiecker, H. Approaches for Recognition of Wood Quality of Standing Trees b Based on Terrestrial Laser Scanner Data. In Proceedings of the 2004 WG VIII/2 Laser-Scanners for Forest and Landscape Assessment, Freiburg, Germany, 3–6 October 2004; Volume 36.
40. Eitel, J.U.; Vierling, L.A.; Long, D.S. Simultaneous measurements of plant structure and chlorophyll content in broadleaf saplings with a terrestrial laser scanner. *Remote Sens. Environ.* **2010**, *114*, 2229–2237.
41. Chum, O. Two-View Geometry Estimation by Random Sample and Consensus. Ph.D. Thesis, Czech Technical University, Prague, Czech, 2005.
42. Choi, S.; Kim, T.; Yu, W. Performance evaluation of RANSAC family. *Br. Mach. Vis. Assoc.* **2009**, *81*, doi:10.5244/C.23.81.
43. Olofsson, K.; Lindberg, E.; Holmgren, J. A Method for Linking Field-Surveyed and Aerial-Detected Single Trees using cross Correlation of Position Images and the Optimization of Weighted Tree List Graphs. In Proceedings of the 8th SilviLaser, Edinburgh, UK, 17–19 September 2008.
44. Seidel, D.; Ammer, C. Efficient measurements of basal area in short rotation forests based on terrestrial laser scanning under special consideration of shadowing. *iFor.-Biogeosci. For.* **2014**, doi:10.3832/ifor1084-007.
45. Thies, M.; Spiecker, H. Evaluation and future prospects of terrestrial laser scanning for standardized forest inventories. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2004**, *2*, 192–197.
46. Maas, H.; Bienert, A.; Scheller, S.; Keane, E. Automatic forest inventory parameter determination from terrestrial laser scanner data. *Int. J. Remote Sens.* **2008**, *29*, 1579–1593.
47. Bayer, D.; Seifert, S.; Pretzsch, H. Structural crown properties of Norway spruce (*Picea abies* [L.] Karst.) and European beech (*Fagus sylvatica* [L.]) in mixed *versus* pure stands revealed by terrestrial laser scanning. *Trees* **2013**, *27*, 1035–1047.
48. Seidel, D.; Leuschner, C.; Mller, A.; Krause, B. Crown plasticity in mixed forests quantifying asymmetry as a measure of competition using terrestrial laser scanning. *For. Ecol. Manag.* **2011**, *261*, 2123–2132.