Tree Inventory with LiDAR Data

Ilya A. Grishin, Elizaveta K. Sakharova, Sergey M. Ustinov, Anton I. Kanev and Valeri I. Terekhov

Bauman Moscow State Technical University, Moscow, Russia
ilia-grishin@mail.ru, elizabeth.saharova16@gmail.com, serega32br@mail.ru, aikanev@bmstu.ru, terekhow@bmstu.ru

Abstract. At present, the rational use of forest resources requires a constant assessment of the state and the implementation of a forecast of the dynamics of the forest fund. For these purposes, it is necessary to update existing data on forest areas in a timely manner. With the spread of technologies for remote data collection, there is a need to improve methods for measuring the taxation parameters of stands. Thus, the paper proposes a methodology based on various methods for calculating taxation parameters and obtained results, in some cases better than the results of existing solutions. The methods of fitting a circle for further estimation of the diameter of a tree trunk are compared: Least Square fit and HyperLS fit. A comparative analysis of the results of measurements of parameters with the results of measurements made using similar software in relation to the data collected by field measurement methods was carried out.

Keywords: LiDAR, tree inventory, machine learning, DBH, circle fitting.

1 Introduction

Forest inventory is a forestry knowledge management that studies the methods of comprehensive accounting of forest resources, determining the volume of the forest area, the timber stock of rare trees and identifications. Taxation is necessary for organizing, inventorying and drawing up a plan for the development of forestry.

The purpose of the statistical inventory of forests is to provide comprehensive information on the state and dynamics of forests for strategic and management planning. Tree inventory is used both for the assessment and calculation of the cost of wood, and for other reasons. For example, to carry out preventive activities and raise awareness, monitor the health of the forest, identify potential fire hazards. An individual tree is the main and forming object of study in forest inventory. According to natural signs, three main parts are distinguished in a tree: root, trunk and crown.

Indicators that characterize the quantitative and qualitative side of the plantation are called taxation parameters. These include the origin of the plantation, its shape, composition, age, good-quality of the growth medium. The main taxation parameters of a tree are diameter, height, length. Other indicators characterizing the longitudinal and transverse shape of the trunk, stem taper, cross-sectional area, crown span, are auxiliary.

With the advent of new technologies for data collection by remote methods, a huge number of data collection tools appear: high resolution satellite imagery, aerial photography (APS) from unmanned and manned aerial vehicles, as well as airborne (ALS) and terrestrial (TLS) laser scanning. LiDAR scanning – high-speed measurement of the distance to an object and registration of radiation directions, resulting in its three-dimensional model.

Much papers done to date are focused on forestry. Thus, in the study [1] a method for the automated determination of the species composition and characteristics of plantations is proposed. Research has been carried out using neural networks to classify tree species on laser scan data [2]. The authors of the study [3] use LiDAR data in conjunction with multispectral aerial photographs and determine the coverage and geometric shapes of trees to create maps of carbon sequestration. The authors of works [4-9] focus on determining the characteristics of forest plantations, such as tree height, crown volume and trunk diameter. The studies demonstrate the high accuracy of calculations, for example, the results of the study [5] show that the mean absolute error (MAE) of the diameter estimate is about 5 mm, the average error of the tree height estimate is about 4 cm. The root mean squared error (RMSE) of the diameter at breast height (DBH) estimate in the study [7] was 1.27 cm, and the RMSE of the tree height estimate is 0.24 m. In [8], two methods of fitting a circle for estimating the diameter of a tree trunk were compared: fitting by the Pratt method and fitting a circle using the least squares method. The RMSE results are 2.38 cm and 2.82 cm, respectively, for the methods.

With the advent and spread of technologies for simplified remote data collection, as well as the possibility of obtaining results with an accuracy exceeding the accuracy of measurements by traditional methods, it becomes necessary to study and improve methods that allow a comprehensive inventory of the forest fund and get the most complete picture of forest stands.

This study was carried out on tree species growing in Russia. The result of the work is the creation of open source software to improve the calculation of a number of taxation parameters of trees.

2 Methods

2.1 Preprocessing Data

Before processing dense point clouds, which are individual trees, 3D data is pre-processed. Thus, the data may include a part of the ground surface. The method for estimating surface normals in a point cloud allows to calculate the normals for each point, segment the input data, and separate the tree from the ground surface topography on which the tree is located. The method finds neighboring points and calculates the principal axis of neighboring points using analysis of covariance. Most often, the method of partitioning a plane based on a k-d-tree acts as an auxiliary method. Partitioning helps to implement the basis for finding the nearest neighbor.

The normals are returned as a normalized vector given by three coordinates. The direction of each normal vector can be set based on how the groups of neighboring points were obtained. Point filtering is combined with a filter that guarantees the proximity of points to the ground.

2.2 Cluster Analysis

In order to correctly determine the height of the tree, as well as the diameter of the tree, measured at a height of 1.3 meters from the collar root, it is necessary to perform segmentation to isolate the tree trunk, as well as possible branches and leaves, in order to eliminate the serious problem of the presence of additional objects in a dense a cloud of points, such as: thin branches of shrubs and tall grass adjacent to the trunk of a tree. Even in the complete absence of redundant objects in the point cloud, it is possible that the crown of the tree is located below the point where leaf-bearing branches are attached to the trunk. Such a case is common in trees with a weeping crown shape, as well as in trees with broken or weak branches.

The DBSCAN clustering algorithm does a good job of clustering data that has arbitrarily shaped clusters of different densities. As the main feature, the algorithm operates with data density. DBSCAN uses the following parameters: eps is the radius of the neighborhood in linear units of the spatial reference frame, i.e. the radius of the ε-neighborhood, and MinPts is the number of data points that must be contained in the neighborhood of the point to make it a base point, i.e. the minimum number of neighbors in order to determine the relation of a point to a group of points, which is considered a separate cluster.

The OPTICS clustering algorithm takes the basic ideas of the DBSCAN clustering algorithm. The OPTICS algorithm uses the distance between neighborhoods and the availability plot to separate clusters of different densities from ambient noise. However, OPTICS is computationally intensive. In this work, the OPTICS algorithm is tuned to obtain clusters of branches that can have different density and a clear separation from foliage and trunk points.

The longest calculation step is the stem and crown segmentation. Most parameters also depend on this clustering. The clustering execution time depends not only on the input clustering settings and the number of iterations, but also on the number of points. So, the time spent on the calculation can vary from a few seconds to hours.

It seems possible to reduce the number of input points, since changing other input control parameters and attributes will lead to a deterioration in the accuracy of clustering and calculations. The most effective method for solving this problem is to thin out the points of file. So, it was noticed that about 50,000 points per tree are enough to save the overall picture of the point cloud, which allows you to make calculations quickly and accurately. A further increase in the number of points does not allow adding sufficient accuracy of calculations, but only increases the running time of the algorithms.

Table 1 presents the results of calculating the parameters before and after thinning the points. There is a significant reduction in processing time while maintaining the accuracy of parameter calculations.

Table 1. Results of thinning calculations

Number of points in the cloud	403746	50469
Diameter at breast height, cm	24,38	24,35
Height, m	21,89	21,78
Length, m	21,90	21,78
Height to the live crown base, m	11,56	10,38
Area of crown convex hull, m^2	99,77	98,91
Volume of crown convex hull, m^3	128,48	122,00
Time spent on calculation, s	1437	27

To reduce the dependence of the segmentation quality on the supplied parameters of density-based clustering algorithms, which must be changed in each case individually, it is proposed to use additional attributes of point records, such as the point color (RGB), the intensity of the return beam from the point, i.e. the reciprocal strength of the laser beam, which depends on the composition of the surface of the object reflecting the laser beam (Intensity), as well as the calculated normal to the surface containing this point, namely the direction of the normal vector, i.e. coordinates of the normalized vector (Nx, Ny, Nz). The optimal parameters eps and MinPts of the DBSCAN segmentation method are found, equal to 0.35 and 100 for a tree point cloud not exceeding 100 thousand points. The optimal settings are used in conjunction with additional point record attributes such as Intensity and Nz.

2.3 Circle Fitting and Convex Hull

After the tree trunk has been allocated into a separate cluster using density-based clustering algorithms, the stem taper must be determined. To do this, it is necessary to determine the diameters of the circles describing the trunk at different heights. Fitting simple primitives to experimental data is one of the main problems in pattern recognition and computer vision. For given n points, the objective function is defined by equation (1).

$$F = \sum_{i=1}^{n}(\sqrt{(x_i - a)^2 + (y_i - b)^2} - R)^2 \quad (1)$$

where (x_i, y_i) are the coordinates of the circle points, (a, b) – coordinates of the center of the circle, R – radius of the circle. The problem is reduced to minimizing equation (1). In the case of using the least squares fit (LS) method for circles and circular arcs, the sum of squared distances to the center is minimized. Once the coordinates (a, b) are determined, the radius can be calculated directly.

Another method used in this work is the Hyper Least Squares (HyperLS, HLS) method [10]. The method maximizes accuracy by introducing a normalization that eliminates statistical error up to second-order noise terms. This was made possible by error analysis followed by subtraction of high-order bias terms. Hyperaccurate [11] was obtained using fourth-order Taylor expansions to compare different algebraic fits and develop better fits, and research [12] shows improved results compared to other circle fitting methods.

In the developed system, calculations are made for three layers at the following heights and ranges: 125 ± 2.5 cm, 130 ± 2.5 cm, 135 ± 2.5 cm, then one of the two methods is calculated, and the median value is taken as the final diameter.

The convex hull of the set of points S in n dimensions is the intersection of all convex sets containing S. For N points $p_1 \ldots p_N$ the convex hull C is given by equation (2).

$$C = \{\sum_{i=1}^{N} \lambda_i p_i : \forall i \in \overline{1,N}\ \lambda_i \geq 0, \sum_{i=1}^{N} \lambda_i = 1\} \quad (2)$$

A fast algorithm for finding the convex hull is the QuickHull algorithm, which can be naturally generalized to the case of arbitrary dimension. The result of the algorithm is a convex hull, which is a polyhedron whose vertices are connected to the original faces that created them. This algorithm is used to construct the minimum convex hull of the tree crown. The resulting hull allows one to calculate the area and volume of the crown surface.

To compare the results of the work with other studies, the following comparison metrics were chosen: mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE).

Fig. 1 shows the interface of the developed system with the results of clustering and calculations.

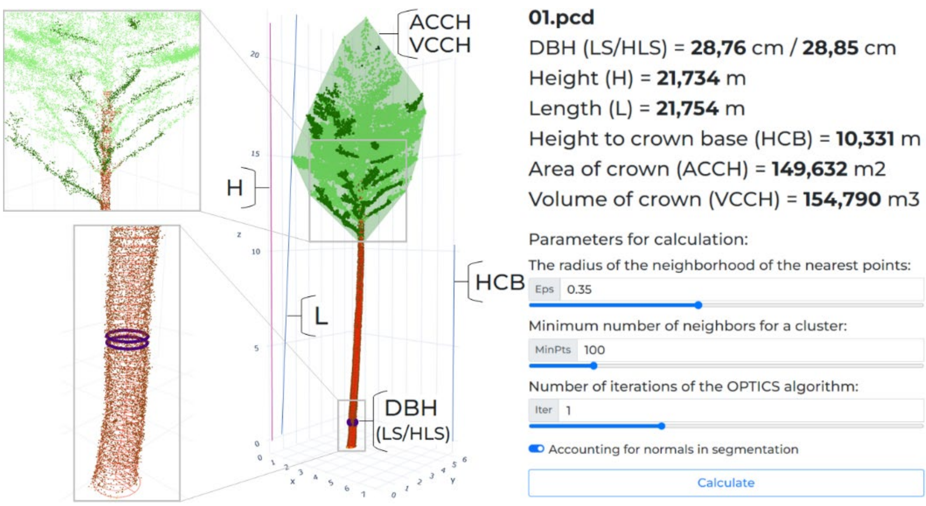

Fig. 1. Interface of the developed system

3 Experiment

To conduct a comparative study on the site, data of 86 pines were collected, obtained using field methods for measuring taxation parameters. The site was surveyed using ground laser scanning and manually segmented into individual trees. The results of calculations of parameters of trees of the developed system were compared with the results of calculations of parameters of plantations obtained using 3DForest, as well as with the results of measured parameters using traditional methods.

Fig. 2 shows the captured dense point cloud of the area selected as the study in this work, and its further segmentation.

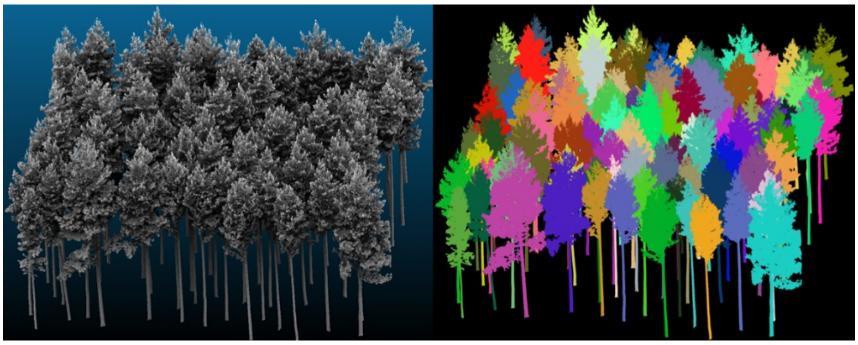

Fig. 2. Forest area selected for research (initial data and manual segmentation result)

The compared taxation parameters are: the height (H) and length (L) of the tree, the diameter of the tree (DBH – 1.3 m above the ground), the height to the live crown base (HCB), the volume (VCCH) and area (ACCH) of crown convex hull. The data that was obtained using field measurements contained information on such parameters as: tree diameter, tree height and height to the live crown base.

All of the above parameters were calculated using the developed system and 3DForest software, then the error values of the two systems were calculated relative to the original field measurements.

3DForest allows one to calculate the diameter at breast height (DBH) using two methods: randomized Hough transformation (RHT) [13] and least squares regression (LS).

The results of calculating the values of the errors in relation to field measurements, are presented in Table 2.

Table 2. Errors of the software

	Errors of the developed system				Errors of software 3DForest			
	DBH LS	DBH HLS	Height	HCB	DBH LS	DBH RHT	Height	HCB
MSE	0.69	0.58	2.81	3.29	0.54	1.18	2.88	4.25
RMSE	0.83	0.76	1.68	1.81	0.74	1.09	1.70	2.06
MAE	0.67	0.60	1.33	1.46	0.59	0.89	1.37	1.68
Error, %	-1.75	-1.26	0.87	-8.84	-1.42	-2.48	1.95	-11.44

Fig. 3 shows a comparative RMSE histogram of two software: the developed system and the 3DForest software.

Choosing the best methods in each system (where HLS is selected in the developed one, and LS in 3DForest), one can verify the similarity of the results (0.76 cm for HLS and 0.74 cm for LS) of the calculations. The LS methods of two different systems differ, but not in the very basis of the method, but in the algorithm for counting layers at a certain height.

Root mean square errors of diameters calculated by two methods in each system are measurement errors in centimeters, while height errors are presented in meters, however, for clarity, the units of measurement have not been reduced to general ones.

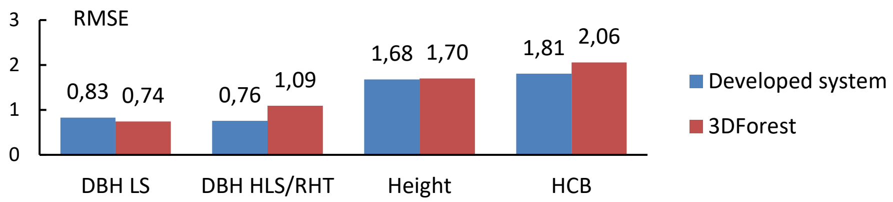

Fig. 3. Comparative RMSE histogram of two software

Fig. 4 shows the distribution graphs of tree parameters measured in different ways, in the box-and-whiskers diagram.

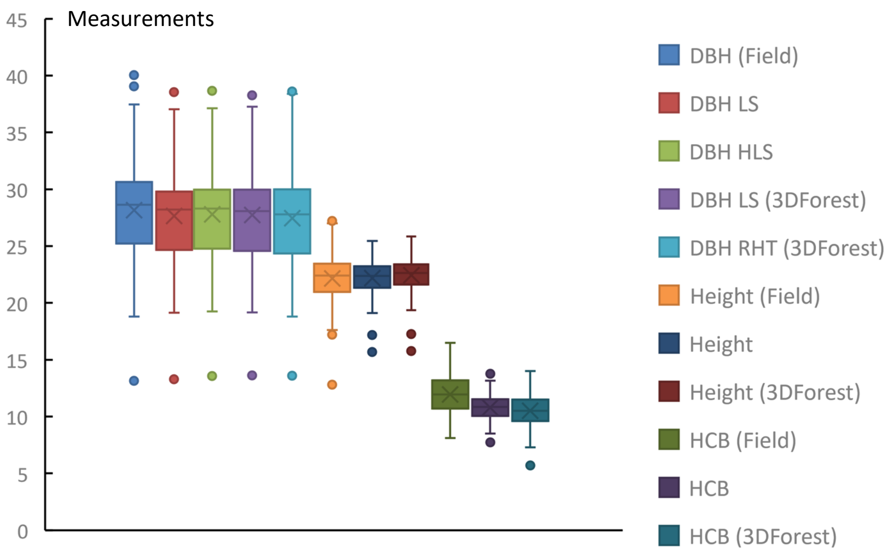

Fig. 4. Box-and-whiskers diagrams of tree diameters, tree heights and heights to the live crown base, measured by different methods

Height calculations are dependent on the initial quality of the survey of the forest area, since the tops of trees can be poorly captured during ground-based laser scanning, which can be mistaken for noise and not considered during processing.

The methods of measuring the height to the base of the crown show the least accuracy, since the calculation algorithm is closely related to the quality of tree crown segmentation. Segmentation of points near the junction of the crown with the trunk is the most vulnerable step of segmentation of the tree as a whole, since in this place there is a junction of two clusters.

Table 3. Results of comparing the developed system and 3DForest software

	Length	Area of crown convex hull	Volume of crown convex hull
MSE	0.08	78.72	47.64
RMSE	0.28	8.87	6.90
MAE	0.24	7.07	5.90

Table 3 shows the results of comparing the developed system and 3DForest software by three parameters of trees, the data of which were not available in manual measurements, so the parameters were compared relative to 3DForest.

Thus, it can be concluded that the results of the developed system are comparable with the results of 3DForest in all parameters, and an increase in the accuracy of calculations was found for a number of parameters. The obtained accuracies are approximately equal, and in some cases exceeding, the accuracies obtained in studies [5, 7, 8].

4 Conclusion

The paper explores various methods for calculating the taxation parameters of trees, considers the procedure for segmenting parts of a tree, suggested optimal parameters, proposes a methodology and obtains results, in some cases better than the results of existing solutions. The methods of fitting a circle for further estimation of the diameter of a tree trunk are compared: Least Square fit and HyperLS fit, where the HyperLS method is more preferred. An open library has been created to improve the results of calculating a number of taxation parameters of trees. The results show that measurements of DBH, height and height to the live crown base using the developed system provide RMSE of 0.76 cm, 1.68 m and 1.81 m, respectively, while the errors of the same parameters using 3DForest are 0.74 cm, 1.70 m and 2.06 m.

References

1. Eroshenkova D. A., Terekhov V.I., Khusnetdinov D.R., Chumachenko S.I.: Automated determination of forest-vegetation characteristics with the use of a neural network of deep learning. In: Kryzhanovsky, B. et al. (eds.) International Conference on Neuroinformatics, vol. 856, pp. 295-302. Springer, Cham (2019).

2. Sakharova, E.K., Nurlyeva, D.D., Fedorova, A.A., Yakubov, A.R., Kanev, A.I.: Issues of Tree Species Classification from LiDAR Data Using Deep Learning Model. In: Kryzhanovsky, B. et al. (eds.) International Conference on Neuroinformatics, vol. 856, pp. 319-324. Springer, Cham (2021).
3. Klein, L.J., Zhou, W., Albrecht, C.M.: Quantification of carbon sequestration in urban forests. arXiv preprint arXiv:2106.00182. (2021).
4. McDaniel, M.W., Nishihata, T., Brooks, C.A., Salesses, P., Iagnemma, K.: Terrain classification and identification of tree stems using ground-based LiDAR. Journal of Field Robotics, 29(6), 891-910. (2012)
5. Dong, W., Isler, V.: Tree morphology for phenotyping from semantics-based mapping in orchard environments. arXiv preprint arXiv:1804.05905. (2018).
6. Proudman, A., Ramezani, M., Fallon, M.: Online Estimation of Diameter at Breast Height (DBH) of Forest Trees Using a Handheld LiDAR. In: 2021 European Conference on Mobile Robots (ECMR), pp. 1-7. IEEE. (2021).
7. Wang, P., Bu, G., Li, R., Zhao, R.: Automated Low-cost Terrestrial Laser Scanner for Measuring Diameters at Breast Height and Heights of Forest Trees. arXiv preprint arXiv:1702.02235. (2017).
8. Pierzchała, M., Giguère, P., Astrup, R.: Mapping forests using an unmanned ground vehicle with 3D LiDAR and graph-SLAM. Computers and Electronics in Agriculture, 145, 217-225. (2018)
9. Andersen, H.E., Reutebuch, S.E., McGaughey, R.J.: A rigorous assessment of tree height measurements obtained using airborne lidar and conventional field methods. Canadian Journal of Remote Sensing, 32(5), 355-366. (2006)
10. Rangarajan, P., Kanatani, K., Niitsuma, H., Sugaya, Y.: Hyper least squares and its applications. In: 2010 20th International Conference on Pattern Recognition, pp. 5-8. IEEE. (2010)
11. Kanatani, K.: Ellipse fitting with hyperaccuracy. IEICE Transactions on Information and Systems, 89(10), 2653-2660. (2006)
12. Al-Sharadqah, A., & Chernov, N.: Error analysis for circle fitting algorithms. Electronic Journal of Statistics, 3, 886-911. (2009)
13. Xu, L., Oja, E.: Randomized Hough transform (RHT): basic mechanisms, algorithms, and computational complexities. CVGIP: Image understanding, 57(2), 131-154. (1993)