



Predicting forest stand variables from airborne LiDAR data using a tree detection method in Central European forests

Ivan Sačkov^{1*}, Lubomír Scheer² and Tomáš Bucha¹

¹ National Forest Centre - Forest Research Institute Zvolen, T. G. Masaryka 2175/22, SK – 960 01 Zvolen, Slovak Republic

² Technical University in Zvolen, Faculty of Forestry, T. G. Masaryka 24, SK – 960 53 Zvolen, Slovak Republic

Abstract

In this study, the individual tree detection approach (ITD) was used to estimate forest stand variables, such as mean height, mean diameter, and total volume. Specifically, we applied the multisource-based method implemented in reFLex software (National Forest Centre, Slovakia) which uses all the information contained in the original point cloud and a priori information. For the accuracy assessment, four reference forest stands with different types of species mixture and the area of 7.5 ha were selected and measured. Furthermore, independent measurements of 1 372 trees were made for the construction of allometric models. The author's ITD-based method provided slightly more accurate estimations for stands with substantial or moderate dominance of coniferous trees. However, no statistically significant effect of species mix on the overall accuracy was confirmed ($p < 0.05$). The root mean square error did not exceed 1.9 m for mean height, 3.0 cm for mean diameter, and 12.88 m³ ha⁻¹ for total volume.

Key words: forest inventory; airborne laser scanning; individual tree detection; multisource-based method

Editor: Peter Surový

1. Introduction

In recent years, airborne laser scanning (ALS, also called airborne LiDAR), has developed to an effective technology for predicting forest stand variables. This is primarily because a laser beam used by these systems can penetrate even through the dense and multi-layered forest canopies to the ground, and ALS data can be used for direct estimate of a spatially explicit three-dimensional canopy structure with sub-meter accuracy (Ginzler & Waser 2017).

A comprehensive overview of ALS-based applications in forestry was published by Maltamo et al. (2014) and current state-of-the-art was examined by Surový & Kuželka (2019). These reviews, as well as other studies, reported accuracy ranging between 32–89% for individual tree detection (Kaartinen et al. 2013; Jeronimo et al. 2018), 6–33% for tree heights estimation (Awaya & Takahashi 2017; Jin et al. 2018), 16–46% for tree diameters derivation (Holopainen et al. 2010; Preditis et al. 2012), and 10–42% for growing stock estimation (Hansen et al. 2017; Kandare et al. 2017).

This study focused on predicting forest stand variables from ALS data using individual tree detection approach (ITD). ITD-based techniques usually involve a sequence of tree detection, feature extraction, and esti-

mation of tree attributes. In general, tree tops/crowns and tree heights are extracted directly from ALS data or ALS-based derivatives, whereas other biophysical attributes are inferred indirectly through allometric models. However, the stand variables that are directly linked to the number of detected trees are often underestimated due to problems with detection of suppressed and understory trees. To solve this problem, semi-ITD approaches and several techniques for modelling understory trees have been proposed (e.g., Lahivaara et al. 2014; Melville et al. 2015; Kansanen et al. 2016).

Many algorithms related to the ITD approach have been developed to predict forest stand variables (e.g., Apostol et al. 2016; Dalponte et al. 2017). In this study, we applied our own multisource-based method implemented in reFLex software (National Forest Centre, Slovakia). The algorithm attempts to eliminate several shortcomings of the current ITD-based methods through the following improvements: (i) the algorithm uses the complete information contained in ALS data in all procedures of tree detection workflow, and optimizes the computationally demanding operations by tiling and thinning techniques applied on the original ALS data, (ii) treetops detection and tree crowns delineation is done iteratively, and each iteration includes tests for

*Corresponding author. Ivan Sačkov, e-mail: sackov@nlcsk.org, phone: +421 949 381 250

treetop verification based on tree allometry rules, aiming at reduction of incorrectly detected trees, (iii) users can modify a number of parameters and customize the algorithm for matching specific stand conditions and/or meeting specific objectives.

The overall objective of the study was to use an author's ITD-based method to estimate forest stand variables in heterogeneous Central European forests with various species mixtures. We focused on main forest stand variables, namely mean height, mean diameter, and total volume.

2. Materials and methods

The study area (Fig. 1) is a part of the University Forest Enterprise of the Technical University in Zvolen, located in central Slovakia (48°37'N, 19°05'E). Discrete airborne LiDAR data and two independent groups of field data were available for the whole study area: (i) ground calibration data for construction of allometric models and (ii) ground reference data for validation of method.

2.1. Airborne LiDAR data

ALS data acquisition was performed in September 2011 using a Riegl L-680i scanner. The study area was scanned from an altitude of 700 m with a 60° field of view, 320 kHz laser pulse repetition rate (PRR), and 122 Hz sampling rate (SR). The resulting vertical standard error was 0.047 m and the average density of point cloud reached 5.1 points per m².

2.2. Ground calibration data

Ground calibration data were used for construction of diameter at breast height (DBH) models. A total of 1 372 trees with a DBH ≥ 7 cm within nine calibration plots covering a total area of 3.3 ha were assessed and measured for species, height, DBH, vitality, and social status. On the plots, a relevant ranges of slope gradients as well as of forest stands in different developmental stages and vertical structures were represented. Also a majority of tree species native to the region was present in the plots. The conifers were mostly composed of Norway spruce (*Picea abies* L. Karst) with 42% coverage, Silver fir (*Abies alba* Mill.) with 13% coverage, and European larch (*Larix decidua* Mill.) with 0.1% coverage. The broadleaves mostly consisted of European beech (*Fagus sylvatica* L.) with 34% coverage, Sessile oak (*Quercus petraea* Matusch) with 6% coverage, and European hornbeam (*Carpinus betulus* L.) with 3% coverage. The crown canopy closure in the plots ranged between 78 and 100%, and almost 75% of the measured trees were situated in the overstory.

2.3. Ground reference data

Ground reference data were used for validation of the method. These data were collected during the leaf-on season in 2013 from four reference stands with different types of species mixture (predominantly coniferous, predominantly broadleaved and mixed). These stands covered approximately 7.5 ha, and their borders were determined with GNSS measurements, resulting in a positional error of less than 1 m.

Variables from 2 203 measured trees with DBH ≥ 7 cm and heights ≥ 5 m were used for the purposes of

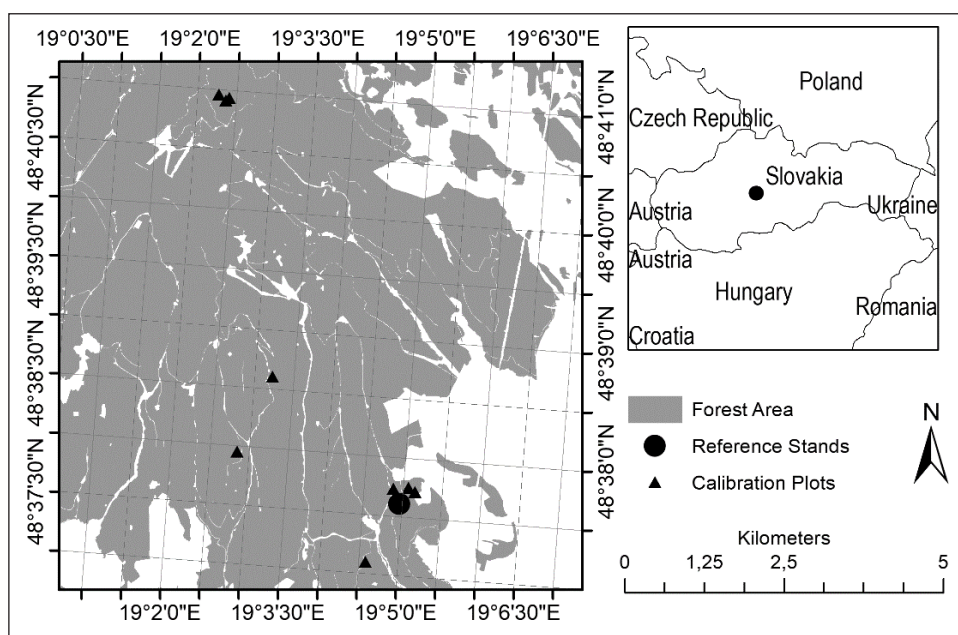


Fig. 1. Location of the study area, calibration plots, and reference stands.

this study. A minimum tree height parameter was set to reflect the commonly accepted forest definition by FAO FRA (Food and Agriculture Organization of the United Nations, 2018). The quadratic mean diameter (d_q) was calculated based on DBHs, and the quadratic mean height (h_{dq}) was calculated using a regression of d_q . The total volume (V) calculation was obtained by summing up the volumes of the measured trees. Here, the tree volume models introduced by Petráš et al. (1991) were used. An overview of the ground reference data is presented in Table 1.

Table 1. Ground reference data characteristics.

Stand	Group of Species	A [ha]	TPH	h_{dq} [m]	d_q [cm]	VPH [$\text{m}^3 \text{ha}^{-1}$]
C	Coniferous $\geq 70\%$	1.64	241	33.70	43.01	520.77
CB	Con./Broad. $\approx 60/40\%$	1.90	199	27.31	35.76	229.89
BC	Broad./Con. $\approx 60/40\%$	2.07	366	24.59	30.69	292.95
B	Broadleaved $\geq 70\%$	1.85	364	25.20	31.73	337.97

Note: A (ha): area in hectares; TPH: number of trees per hectare; h_{dq} (m): regression height of the tree with the quadratic mean diameter in metres; d_q (cm): quadratic mean diameter in centimetres; VPH ($\text{m}^3 \text{ha}^{-1}$): total volume per hectare in cubic metres.

2.4. Predicting forest stand variables from airborne LiDAR data

2.4.1. Individual tree and tree height detection

The treetops, tree crowns, and tree heights were detected using the reFLex algorithm (National Forest Centre, Zvolen, Slovakia). There, the initial procedures (i) divided the points into a three-dimensional regular mesh, (ii) calculated the absolute height above the ground for each point, and (iii) reduced the number of points in the input file by applying a minimum tree height threshold (5 m). These operations produced a point cloud that was further used in an iterative search for treetops and tree crowns by using a moving-window analysis. Since there is a reason to assume that a part of the local maxima identified in the previous operation may not indicate the real treetops, an additional geo-dendrometric (GD) test was applied. The GD test is linked to a priori information and includes two steps: (i) evaluation of height differences between the local maxima, and (ii) evaluation of horizontal and vertical distances between the local maxima. This way, the false treetops located in the crowns of other trees are removed. The final procedures were applied to delineate the tree crowns. After the treetop identification and crown delineation, tree heights were recorded and crown coverage was calculated. Finally, the outputs of all procedures were exported to point and polygon vector files in an ESRI shapefile format. All details of the algorithm are listed in the study of Sačkov et al. (2017).

2.4.2 Classification of tree species groups

The classification was carried out for two general groups of tree species: (i) broadleaves, and (ii) conifers. These groups were classified based on combination of intensity

raster and canopy height model (CHM) with pixel size of 1.0 m. First, the range of spectral information in the form of the backscatter intensity of the laser signal for each group of species was identified through visual interpretation, and the intensity raster with pixels representing broadleaves, conifers, and ground was created. Second, pixels smaller than 5 m CHM height were removed from the intensity raster and, thus, the final intensity raster included only the broadleaved and coniferous group. Finally, one of the species groups was assigned to detected trees using zonal statistical functions.

2.4.3 Tree diameter derivation

The DBHs of the detected trees were derived based on nonlinear regression models. The model predictor was tree height for the selected group of tree species ($DBH = f(h)$). The calibration dataset included 769 broadleaved trees and 603 coniferous trees (Section 2.2). The statistical significance of models was assessed using the F-test at a significance level of $\alpha = 0.05$. We found the exponential function to be most suitable for DBH derivation. The average accuracy for these models was 19% at the tree level.

2.4.4 Tree volume derivation

The volume for detected trees was derived based on the models introduced by Petráš et al. (1991). For each remotely detected tree with estimated height, assigned tree species group, and derived DBH, the volume calculation was applied using the adopted model. We used beech function for trees assigned to the broadleaved group and spruce function for trees assigned to the coniferous group because these species achieved a highest proportion of the total volume within forest stands.

2.4.5 Stand variables calculation

With respect to selected criteria (Section 2.3), only detected trees higher than 5 m were used for the purposes of this study. The stand height and stand diameter were calculated as the average of the tree data. The total volume was calculated as a sum of the tree volumes.

2.5. Accuracy assessment

The accuracy of tree detection was assessed through extraction rate (ER) as the difference between total number of ALS-extracted trees and total number of ground-measured trees. Thus, the commission or omission error were not calculated directly. This is because we were particularly interested in evaluating the fully autonomous prediction of stand variables that is achievable

without human intervention (e.g., application of positive or negative biases of tree detection methods). On the other hand, the under- or over-estimation of detection was indirectly assessed through the histogram of height intervals with relative frequencies of ALS-extracted trees and ground-measured trees.

The accuracy of estimations of stand height, stand diameter and total volume was carried out by comparing of ALS-predicted stand variables and ground-measured stand variables. The mean difference (e) was calculated as the average of individual differences and was used as an indicator of underestimation or overestimation. The random error component (se) was used to assess the dispersion of differences around the mean difference. The root mean square error (RMSE) was used to aggregate both the systematic and random error components. The resulting RMSE should indicate the range of total accuracy for the whole study area at the 68% confidence level. The relative $e\%$, $se\%$, and $RMSE\%$ were calculated as the ratios of their absolute values and the arithmetic average of the reference data. Finally, we used the Mann-Whitney U test to assess the significance of differences (i) between predicted variables within forest stands, and (ii) between predicted and measured stand variables within forest stands ($p < 0.05$). This non-parametric test was selected primarily because a normal distribution of mean differences was not confirmed within the dataset ($p < 0.05$).

3. Results

The extraction rates for the C, CB, BC, B stand were 63%, 76%, 57%, and 63%, respectively. Specifically, Fig. 2

shows the relative frequencies of ground-measured and ALS-extracted trees within height intervals. The average underestimation achieved the value of -9.7% and the extreme was found within the height interval of 20–25 m (-14%). The average overestimation achieved the value of 5.7% and the extreme was found within the height interval of 25–30 m (15%). The differences in extraction rates between strata were not significant, thus the effect of tree species composition was marginal in these cases ($p < 0.05$).

The RMSE did not exceed 1.9 m (7.2 %) for mean height, 3.03 cm (8.6 %) for mean diameter, and $12.88 \text{ m}^3 \text{ ha}^{-1}$ (15.0 %) for total volume. The overview of all predictions and accuracies for all stands is presented in Fig. 3–5. The ALS-based predictions of all stand variables were statistically significantly different relative to the ground data ($p < 0.05$).

4. Discussion

In this study, we applied our own multisource-based method related to the ITD approach to estimate forest stand variables from airborne LiDAR data in Central Europe forests with several types of species mixture. Tree positions and tree heights were directly extracted from original point cloud. The DBHs and tree volumes were derived through allometric models. Finally, three main forest stand variables (i.e., mean height, mean diameter, and total volume) were calculated based on tree data.

A slightly higher accuracy of detection as well as prediction was achieved within stands, with substantial or moderate dominance of coniferous trees (C and CB).

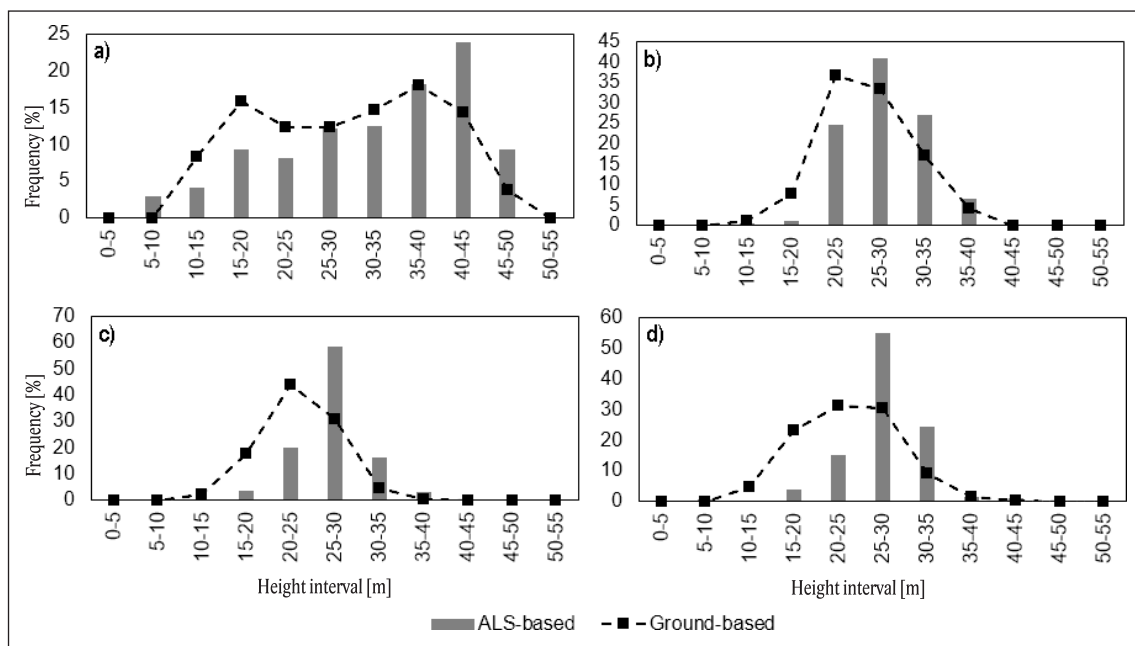


Fig. 2. Relative frequencies of ground-measured and ALS-extracted trees within height interval: (a) Coniferous stand, (b) Coniferous/Broadleaved stand, (c) Broadleaved/Coniferous stand, (d) Broadleaved stand.

However, the statistically significant effect of tree species mixture on the overall accuracy was not confirmed in this study ($p < 0.05$).

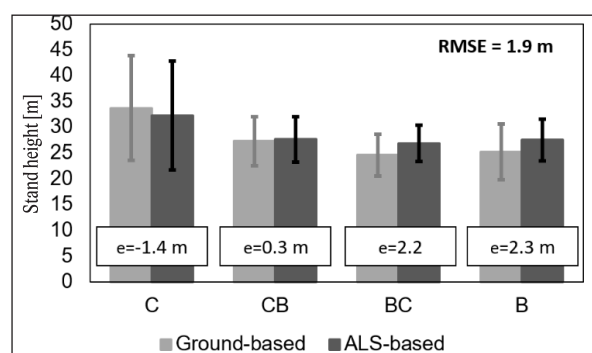


Fig. 3. Comparison of ground-measured and ALS-predicted stand heights.

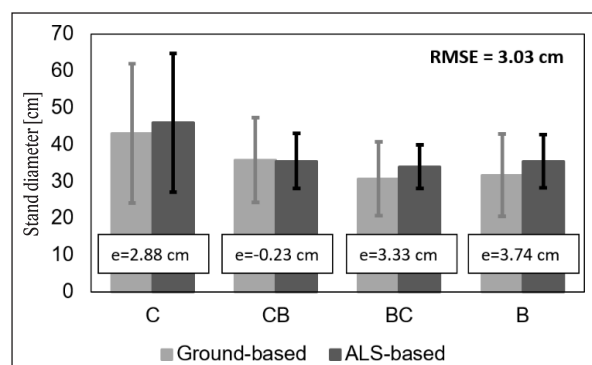


Fig. 4. Comparison of ground-measured and ALS-predicted stand diameters.

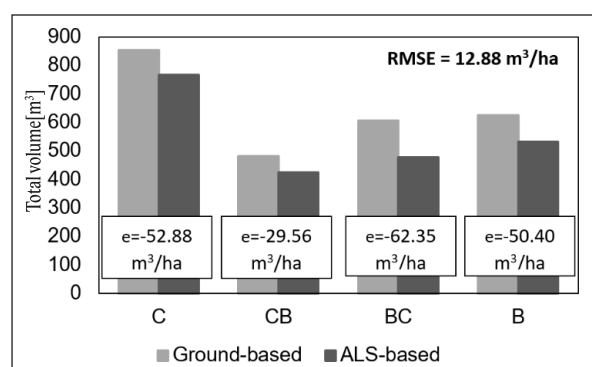


Fig. 5. Comparison of ground-measured and ALS-predicted total volumes.

The mean difference of ALS-predicted stand height ranged from -1.4 m to 2.3 m and decreased as all stand parameters increased. Using a similar approach and data format, Gaveau & Hill (2003) reported a mean difference that ranged from -0.91 m to -1.27 m for deciduous forest, while Sibona et al. (2017) achieved a mean difference of 0.95 – 1.13 m for coniferous forest. Chávez & Tullis (2013) evaluated stand height using ALS data and hyperspectral imagery over full-canopy oak-hickory

forests with a mean difference of 1.67 – 2.99 m and RMSE of 2.1 – 3.7 m.

The accuracy of ALS-predicted stand diameters increased with increasing mean heights, however, with increasing total volume and stocking density, the accuracy decreased. Although the models used for DBH derivation in this study were relatively simple, the mean difference was comparable with other studies. For example, Yu et al. (2011) and Wu et al. (2015) reported accuracies that ranged from -17.6 cm to 13.2 cm in terms of mean difference. Moreover, other models for DBH derivation often used more variables than just tree height (e.g., crown area, crown length, crown volume).

Although all ALS-extracted trees were used for the stand volume prediction, the total volume was underestimated within all reference forest stands. The mean difference ranged from -29.56 m³ ha⁻¹ to -62.53 m³ ha⁻¹ and the accuracy increased as all stand parameters increased. There were three possible reasons for this. First, the omission error caused by undetected understorey and suppressed trees, mostly resulting in underestimations of biophysical attributes when aggregated at the plot level. Second, the volume functions only for two species groups were used as a compromise. Finally, the density of ALS point cloud was relatively low for ITD approach. Despite these findings, our results were still within the range of achievable accuracy indicated by other studies. Kandare et al. (2017) reported the RMSE of 152.18 m³ ha⁻¹ and mean difference of 132.37 m³ ha⁻¹ using an ITD approach, the RMSE of 102.78 m³ ha⁻¹ and mean difference of -1.59 m³ ha⁻¹ using a semi-ITD approach, and the RMSE of 182.75 m³ ha⁻¹ and mean difference of -15.88 m³ ha⁻¹ using an area-based approach for heterogeneous forests in Italy. Ullah et al. (2016) achieved the RMSE of 66.31 – 66.67 m³ ha⁻¹ through multiple linear regression, k-Nearest Neighbour, and support vector machine for heterogeneous forest in Germany.

5. Conclusions

Results show that an author's method that is based on ITD approach can be used for predicting forest stand variables, such as stand height, stand diameter, and total volume. With respect to other studies, our findings also indicated that airborne LiDAR data are suitable mainly for the prediction of stand heights (Smreček et al. 2018). The method has been less successful for indirectly derived stand variables (e.g., stand diameter, total volume). However, the prediction of these variables typically depends on point density across a point cloud, input variables, and allometric models (Kamińska et al. 2019).

A significant improvement to the method would, therefore, be achieved mainly via using more complex techniques for tree species classification and more precise allometric models. Further research and development of reFLex algorithm is also needed to implement

the presented approach into forest inventory practice in various ecosystems.

Acknowledgments

This research was supported by the Slovak Research and Development Agency in the framework of the project “Innovations in the forest inventories based on progressive technologies of remote sensing, ProMon” (APVV-15-0393) and by the Research and Development Operational Programme funded by the EU in the framework of the project “Using forest dynamics models for decision support systems, DYNAMO” (NFP313010T597).

Ivan Sačkov acknowledges the receipt of a fellowship from the OECD Co-operative Research Programme: Biological Resource Management for Sustainable Agricultural Systems in 2019.

References

- Apostol, B., Lorent, A., Petrila, M., Gancz, V., Badea, O., 2016: Height Extraction and Stand Volume Estimation Based on Fusion Airborne LiDAR Data and Terrestrial Measurements for a Norway Spruce (*Picea abies* [L.] Karst.) Test Site in Romania. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 44:313–323.
- Awaya, Y., Takahashi, T., 2017: Evaluating the Differences in Modeling Biophysical Attributes between Deciduous Broadleaved and Evergreen Conifer Forests Using Low-Density Small-Footprint LiDAR Data. *Remote Sensing*, 9:572.
- Dalponte, M., Frizzera, L., Gianelle, D., 2017: How to map forest structure from aircraft, one tree at a time. *Ecology and Evolution*, p. 1–8.
- Food and Agriculture Organization of the United Nations, 2018: Global Forest Resources Assessment 2020, Guidelines and specifications. Rome.
- Gaveau, D. L. A., Hill, R. A., 2003: Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. *Canadian Journal of Remote Sensing*, 29:650–657.
- Ginzler, Ch., Waser, L., 2017: Entwicklungen im Bereich der Fernerkundung für forstliche Anwendungen. *Schweizerische Zeitschrift für Forstwesen*, 3:118–126. (In German).
- Hansen, E. H., Ene, L. T., Mauya, E. W., Patočka, Z., Mikita, T., Gobakken, T. et al., 2017: Comparing Empirical and Semi-Empirical Approaches to Forest Biomass Modelling in Different Biomes Using Airborne Laser Scanner Data. *Forests*, 8:170.
- Holopainen, M., Mäkinen, A., Rasinmäki, J., Hyypä, J., Hyypä, H., Kaartinen, H. et al., 2010: Effect of tree-level airborne laser-scanning measurement accuracy on the timing and expected value of harvest decisions. *European Journal of Forest Research*, 129:899–907.
- Chávez, J. S., Tullis, J. S., 2013: Deciduous Forest Structure Estimated with LiDAR-Optimized Spectral Remote Sensing. *Remote Sensing*, 5:155–182.
- Jeronimo, S. M. A., Kane, V. R., Churchill, D. J., McGaughey, R. J., Franklin, J. F., 2018: Applying LiDAR Individual Tree Detection to Management of Structurally Diverse Forest Landscapes. *Journal of Forestry*, 116:336–346.
- Jin, S., Su, Y., Gao, S., Hu, T., Liu, J., Guo, Q., 2018: The Transferability of Random Forest in Canopy Height Estimation from Multi-Source Remote Sensing Data. *Remote Sensing*, 10:1183.
- Kaartinen, H., Hyypä, J., Yu, X., Vastaranta, M., Hyypä, H., Kukko, A. et al., 2013: An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, 4:950–974.
- Kamińska, A., Lisiewicz, M., Stereńczak, K., Kraszewski, B., Sadkowski, R., 2018: Species-related single dead tree detection using multi-temporal ALS data and CIR imagery. *Remote Sensing of Environment*, 219:31–43.
- Kandare, K., Dalponte, M., Ørka, H.O., Frizzera, L., Næsset, E., 2017: Prediction of Species-Specific Volume Using Different Inventory Approaches by Fusing Airborne Laser Scanning and Hyperspectral Data. *Remote Sensing*, 9: 400.
- Kansanen, K., Vauhkonen, J., Lahivaara, T., Mehtatalo, L., 2016: Stand density estimators based on individual tree detection and stochastic geometry. *Canadian Journal of Forest Research*, 46:1359–1366.
- Lähivaara T., Seppänen A., Kaipio J. P., Vauhkonen J., Korhonen L., Tokola T. et al., 2014: Bayesian approach to tree detection based on airborne laser scanning data. *IEEE Transactions on Geoscience and Remote Sensing*, 52:2690–2699.
- Maltamo, M., Naesset, E., Vauhkonen, J., 2014: *Forestry Application of Airborne Laser Scanning: Concept and Case Studies*. Springer Netherlands, Dordrecht, Netherlands, 460 p.
- Melville, G., Stone, C., Turner, R., 2015: Application of LiDAR data to maximise the efficiency of inventory plots in softwood plantations. *New Zealand Journal of Forestry Science*, 45:9.
- Petráš, R., Pajtík, J., 1991: Sústava česko-slovenských objemových tabuliek drevín. *Lesnícky časopis*, 37:49–56.
- Preditis, G., Šmits, I., Arhipova, I., Dagis, S., Dubrovskis, D., 2012: Tree Diameter Models from Field and Remote sensing data. *Int. Journal of Mathematical Models and Methods in Applied Science*, 5:707–714.
- Sackov, I., Hlásny, T., Bucha, T., Juriš, M., 2017: Integration of tree allometry rules to treetops detection and tree crowns delineation using airborne lidar data. *iForest*, 10:459–467.

- Sibona, E., Vitali, A., Meloni, F., Caffo, L., Dotta, A., Lingua, E. et al. 2017: Direct Measurement of Tree Height Provides Different Results on the Assessment of LiDAR Accuracy. *Forests*, 8:7.
- Smreček, R., Michnová, Z., Sačkov, I., Danihelová, Z., Levická, M., Tuček, J., 2018: Determining basic forest stand characteristics using airborne laser scanning in mixed forest stands of Central Europe. *iForest* 11: 181–188.
- Surový, P., Kuželka, K., 2019: Acquisition of Forest Attributes for Decision Support at the Forest Enterprise Level Using Remote-Sensing Techniques – A Review. *Forests*, 10:273.
- Ullah, S., Dees, M., Datta, P., Adler, P., Koch, B., 2017: Comparing airborne laser scanning, and image-based point clouds by semi-global matching and enhanced automatic terrain extraction to estimate forest timber volume. *Forests*, 8:215.
- Wu, J., Yao, W., Choi, S., Park, T., Myneni, R.B., 2015: A comparative study of predicting DBH and stem volume of individual trees in a temperate forest using airborne waveform LiDAR. *IEEE Geoscience and Remote Sensing Letters*, 12:2267–2271.
- Yu, X., Hyyppä, J., Vastaranta, M., Holopainen, M., 2011: Predicting individual tree attributes from airborne laser point clouds based on random forest technique. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66:28–37.