Drewno 2018, Vol. 61, No. 201
DOI: 10.12841/wood.1644-3985.D11.04

RESEARCH PAPERS

Krzysztof STEREŃCZAK, Marek LISAŃCZUK, Karolina PARKITNA, Krzysztof MITELSZTEDT, Piotr MROCZEK, Stanisław MIŚCICKI

THE INFLUENCE OF NUMBER AND SIZE OF SAMPLE PLOTS ON MODELLING GROWING STOCK VOLUME BASED ON AIRBORNE LASER SCANNING

Current forest growing stock inventory methods used in Poland are based on statistical methods using field measurements of trees on circular sample plots. Such measurements are carried out with traditional equipment, i.e. callipers and range finders. Nowadays, remote sensing based inventory techniques are becoming more popular and have already been applied in North America and some Scandinavian countries. Remote sensing based forest inventories require a certain amount of ground sample plots, which serve either as reference data used for model calibration and/or as a validation dataset for the assessment of the accuracy of modelled variables.

Using a set of 900 ground sample plots and Airborne Laser Scanner (ALS) from the Milicz forest district, a statistical model for the estimation of plot growing stock volume was developed. Next, the developed model was once again fitted to different variants of sample plot size and number of sample plots. Each variant was selected from a full 900 sample plot set. The selection started from 800, 700, 600, ..., down to 25 plots, respectively, and was carried out in proportion to the dominant tree age range. To account for the area effect, each plot number variant was similarly tested with various sample plot areas, i.e. 500, 400, ..., 100 m^2. Sampling in each variant was repeated in order to take into account the effect of a single selection. The results showed a strong relationship between obtained modelling errors and the size and number of used sample plots. It has been demonstrated that the number of sample plots has no influence on the accuracy of GSV estimation above about 300-400 sample plots (about 500 sample plots for bias), whereas sample plot size has a visible impact on estimation accuracy,

Krzysztof STEREŃCZAK✉ (*k.sterenczak@ibles.waw.pl*), Marek LISAŃCZUK (*m.lisanczuk@ibles.waw.pl*), Karolina PARKITNA (*k.parkitna@ibles.waw.pl*), Krzysztof MITELSZTEDT (*k.mitelsztedt@ibles.waw.pl*), Piotr MROCZEK (*p.mroczek@ibles.waw.pl*), Forest Research Institute, Laboratory of Geomatics, Sekocin Stary, Raszyn, Poland; Stanisław MIŚCICKI (*stanislaw_miscicki@sggw.pl*), Department of Forest Management, Geomatics and Forest Economics, Faculty of Forestry, Warsaw University of Life Sciences – SGGW, Warsaw, Poland

which reduces with decreasing sample plot size, regardless of the number of sample plots. If it is about precision, results showed that the influence of a single selection to be relevant only below 300-400 plots (about 500 for bias) and the same trend can be observed in each sample plot size variant. The results showed it is possible to strongly reduce the number of ground sample plots (minimum 300-400), while still maintaining decent accuracy and precision levels, at least in similarly investigated forest conditions.

Keywords: ALS, LIDAR, forest inventory, sampling intensity

Introduction

Forests are an important carbon storage reservoir [Lindner and Karjalainen 2007] and one of the most basic sources of renewable energy [McKendry 2002]. Forest growing stock volume (GSV) is one of the most important characteristic that describes available wood resources. Increasing human pressure on the environment and expanding demands for wood requires a need for more precise forest management, hence more accurate forest measurements [Tonolli et al. 2011].

The development of GSV inventory methods has resulted in the improvement of GSV estimation methods and tools. The appearance of remote sensing (RS) data on the market immediately led to them being used in the inventory of forest resources. Classical spectral data might be suitable for vast areas [Holmström et al. 2001]. However, the forest stand level GSV inventory (usually an area of 1-5 ha) requires more precise data. The development of Airborne Laser Scanning (ALS) has been a recent innovation in the vertical characterization of the tree and stand structures. Moreover, its features for modelling GSV are superior to those of spectral RS data [Maltamo et al. 2006]. In commercial forests, there is a strong proportional relationship between stem volume and biomass [Maltamo et al. 2016]; consequently, accurate ALS aided biomass estimates are also good volume estimates [Ruiz et al. 2014]. Thus experience in using ALS data for predicting biomass can be also used to some extent, for predicting GSV.

ALS data can be used to model GSV by analysing point clouds based on laser-derived height quantiles [Næsset 1997; Næsset 2002; Lim et al. 2003], Canopy Height Model (CHM) quantiles [Maltamo et al. 2006] or variables acquired from single tree detection results [Hyyppä et al. 2012; Miścicki and Stereńczak 2012]. Many studies use area based methods with laser-derived height quantiles, as these are some of the most accurate approaches [Næsset 1997; Lim et al. 2003].

Apart from a method chosen for GSV prediction, sample plot sizes and their number are important factors for the model development. The large number of studies described in Ruiz et al. [2014] used various sizes of sample plots from 0.01 ha up to 0.36 ha. The results of another study [Frazer et al. 2011] generally

indicated that larger plots significantly reduce the edge effect and co-registration errors. Ruiz et al. [2014] suggested that sample plots used for volume predictions should have a minimum size of 500-600 m^2. The same authors pointed out that larger plot sizes clearly increase the cost of fieldwork, while not significantly increasing the accuracy of the prediction models. The above--mentioned sample plot size is generally the maximum sample plot size used in forest inventories, at least in Poland.

The number of plots is another important aspect to consider, regarding GSV prediction models. To the best of our knowledge, only a limited number of studies have explored a minimum number of sample plots required to estimate the relationship between ALS metrics and GSV without systematic error. This could be due to the limited amounts of available field data. For example, in the study of Kallio et al. [2010] with the use of rather small samples (radii from 8-10 m), they found that the accuracy of estimated *Picea abies* volumes at the forest stand level did not decrease until the number of plots was reduced to below 200. On the other hand, the volume estimation accuracy of *Pinus sylvestris* and deciduous tree species decreased remarkably as the number of sample plots decreased. However, for all species, the total volume root mean square error (RMSE) were relatively high at the forest stand level: 74%, 91%, and 33% for pine, deciduous trees and spruce, respectively. Gobakken and Næsset [2008] pointed out that "average standard deviation increased when the number of sample plots was reduced".

From a practical perspective, the number of sample plots and their sizes are the most crucial variables determining the costs of fieldwork during forest inventories. For this reason, we evaluated the effects of a sample plot size and the number of sample plots on GSV prediction at the sample plot level. Used variants consisted of 25 to 900 plots, selected from 900 available plots, which sizes varied from 0.01 to 0.05 ha.

This paper is arranged as follows: first, we describe the study area, LiDAR data, and the performed fieldworks. Next, we provide the methodology with all included variants of analysis, followed by the report, analysis, and discussion of the obtained results. Finally, we state the overall conclusions. In this work, the terms growing stock volume, standing volume, and volume are used interchangeably. Anywhere the expression "sample plot size" is used in the text, this refers to the area of a single sample plot. Anywhere the expression "number of sample plots" is used, this refers to the number of sample plots used for model calibration.

Materials and methods

Study area

The study was conducted in the Milicz Inspectorate as a part of the Milicz District. In this territory, forested areas extend over 8 500 ha (greyed outline in

fig. 1). Milicz is situated in south-west Poland, about 50 km north of Wrocław. Despite a relatively high species diversity, Scots pine (*Pinus sylvestris* L.) is the dominant tree species in the study area. Scots pine stands cover approximately 70% of the forest district area, while mixed pine-beech stands and pine-oak stands cover up to 10% each. There is also a minor covering of beech (*Fagus sylvatica* L.), common alder (*Alnus glutinosa* Gaertn.) and silver birch (*Betula pendula* Roth). The distribution of stand age classes is as follows: 40-60 years (about 30% of the district's area), 20-40 years and >120 years (about 15% each). The remaining age classes, i.e. 1-20, 60-80 and 80-100 years, have similar shares (about 10% each). The average standing volumes of pine, beech and oak (*Quercus* sp. L.), the main species in the study area, are 300 $m^3 \times ha^{-1}$ – pine and beech, and 275 $m^3 \times ha^{-1}$ – oak.

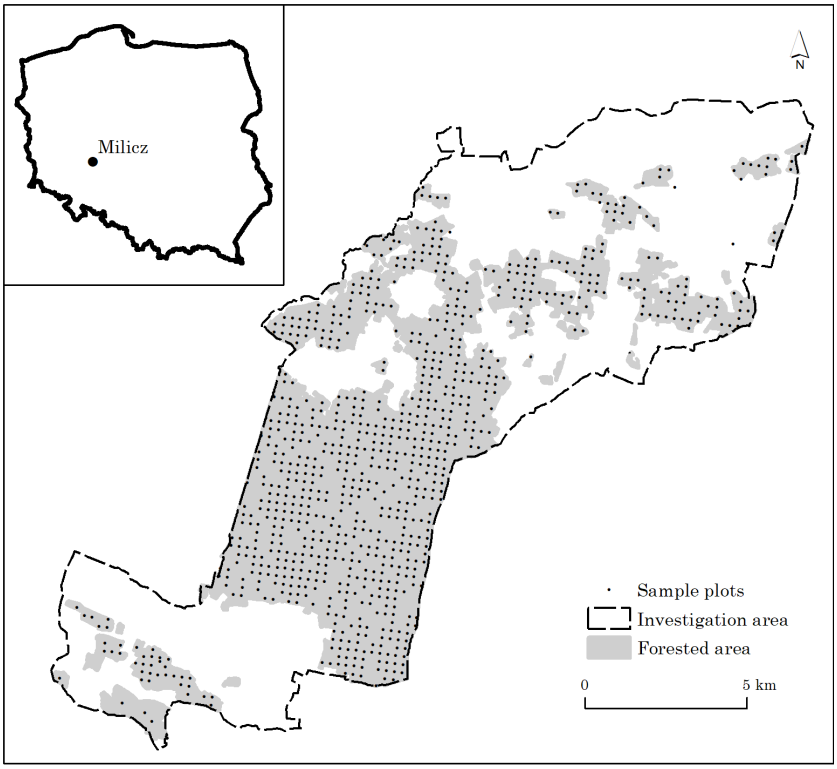


Fig. 1. Study area – Milicz forest district

Field reference data

There were 900 circular field sample plots with radii of 12.62 m (500 m^2), distributed over the study area in a regular grid pattern (350 m × 350 m). Field measurements were conducted in summer 2015. On each sample plot, all trees with a diameter at breast height of at least 7 cm were calipered, and their heights

measured with a Haglöf Vertex IV rangefinder. The volume of every single tree was computed based on formulas commonly used in Polish forest management planning [Bruchwald et al. 2000]. A general description of the data obtained from the sample plots is presented in table 1. The plot centre locations were measured with a Global Navigation Satellite System in Real Time Kinematic mode (RTK GNSS) using Virtual Reference Stations (VRS) technology. The horizontal precision of GNSS measurements was 0.044 m and vertical – 0.05 m.

Table 1. Descriptive statistics of growing stock volume (GSV) values and species proportion, based on data from the ground sample plots of 500 m^2

Plots	No. of sample plots	Mean no. of trees per sample plot	GSV [m^3 × ha^{-1}]				Species proportion [%]					
			Mean	Minimum	Maximum	SD	Pine	Oak	Beech	Birch	Spruce	Other
All sample plots	900	40	379	7	1302	165	64	13	9	3	2	8
Pine dominated	628	46	368	31	1002	138	91	1	2	2	1	3
Oak dominated	95	24	463	59	1206	223	3	86	3	1	2	5
Beech dominated	65	17	414	16	998	231	7	7	84	0	0	1
Birch dominated	12	36	231	42	350	104	9	4	2	68	0	17
Spruce dominated	7	33	345	169	584	160	7	0	15	7	69	2
Other	93	35	369	7	1302	182	13	11	8	8	4	56

ALS data

ALS data were collected in August 2015 using a Riegl LMSQ680i laser scanning system with a 360 kHz pulse rate frequency that resulted in point clouds with an average density of 10 pulses × m^{-2}. The mean flight altitude was 550 m and the field of view of the scanning system was 60 degrees. Along with the point clouds, the data provider generated a digital terrain model (DTM) with a spatial resolution of 0.5 m in TerraSolid software. This DTM was used to normalize all returns from the raw point clouds.

Data preparation

The first data preparation step was to create vector layers delineating the geographical extents of the circular sample plots for the five variants of sample plot size. The first variant was just the original and full dataset with plot characteristics computed on the basis of ground measurements. This dataset is introduced in the *Ground reference data* section (i.e. 900 sample plots, each

500 m^2 in size). Next, the geographical extents of 900 sample plots were delineated again; however, this time the area of each sample plot was limited to 400 m^2. Subsequently, these delineations of vector layers were performed again for the sample plots limited to 300, 200 and 100 m^2 (fig. 2A). Sample plots in each area variant were circular.

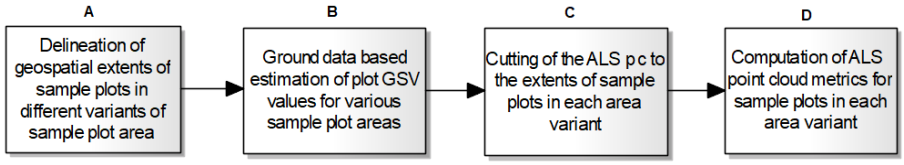


Fig. 2. The main steps in the data preparation workflow

The next step was to compute the growing stock volumes [m^3 × ha^{-1}] of every sample plot. The GSV for a given sample plot was the sum of the volumes of the single trees growing within the sample plot of corresponding sample size.

The third data preparation step was to cut the original ALS point cloud (ALS pc) to the previously prepared extent of sample plots, for each area variant (fig. 2C). Based on the truncated point cloud, a set of ALS metrics were computed for each sample plot in each area variant (fig. 2D). Computed statistics were mainly concerned with height values, as well as intensity and the distribution/quantity of particular echoes (in a specified height stratum), e.g.: mean height, height percentiles, and more complex metrics such as the number of first returns above the mean height value of all returns over a sample plot. The final dataset prepared for analysis consisted of 5 (area variants) × 900 (total number of sample plots). The following properties were determined for each sample plot: ground estimated GSV, area, age class and a set of ALS pc metrics. All computations were conducted in the R programming language based on original scripts.

Data analysis

The first data analysis step was to develop a relatively robust statistical model that allowed the estimation of sample plot GSVs from calculated point cloud metrics. This model was developed using a regression analysis of 900 sample plots 500 m^2 in size. Among the entire set of more than 100 point cloud metrics characterizing particular sample plots, only those which explained the underlying variation of the observed plot GSVs were selected, and those that were not auto-correlated with each other. In order to reach this, all variables which had a coefficient of variance less than 20% were discarded assuming that a low variability in predictor variable cannot sufficiently explain the variance of the dependent variable. Next, all those variables were discarded, which coefficient of correlation (r) was less than 0.5 in respect to the dependent

variable – ground GSV. Then, a stepwise regression was used to determine the final model. As the result of this procedure, the original set of ALS metrics was narrowed down to 19 predictor variables. An ordinary least square multiple regression method was used to estimate the final model parameters. Model selection was based on the following: (1) maximization of the $R^2_{adjusted}$ between the fitted and observed GSV values, and (2) the relative simplicity of the equation compared to other competitive models, without significant loss of $R^2_{adjusted}$. Equation 1 presents the general form of the final model:

$$GSV = a_0 + a_1(x_1 * x_2) + a_2 x_3^2 + a_3 x_4 \quad (1)$$

where: GSV – growing stock volume [$m^3 \times ha^{-1}$] of a sample plot,
a_0, a_1, a_2, a_3 – regression coefficients of the model,
x_1 – quadratic mean of point heights over a sample plot,
x_2 – ratio, number of 1st returns above the mean point heights, divided by the number of all 1st returns over a sample plot,
x_3 – ratio, number of all returns above the mean point heights, divided by the number of all returns over a sample plot,
x_4 – ratio, number of last returns above the 9th height stratum divided by the number of all last returns, where the height stratum consists of returns found in a specified height interval between 2 meters above ground and the 95th height percentile [Næsset and Gobakken 2005; Gobakken et al. 2012].

After the final model selection, an additional group of variants was added. The following number of sample plots per variant of sample plot size were used: 800, 700, 600, 500, 400, 300, 200, 100, 50, and 25. The selection of specific sample plots was based on the ground GSV distribution proportional to the dominant tree age of a given sample plot. Sampling with a replacement was used, as it was assumed that a proper number of iterations would exhaust variation among all the sample plot GSVs and corresponding point cloud metrics. It is worth mentioning, that sample plots were unique within each single sampling (iteration), however, the same sample plot could be selected again in subsequent iterations. Moreover, the number of iterations for each analysed number of sample plots variant, was required to be large enough to present a range of possible cases. Since the mean values of the estimates were not of high importance in this study (rather, the distribution of errors between repetitions) the influence of a large number of repetitions was not so relevant to the analysis and aims of the study. Table 2 presents the number of repetitions for each variant of the number of sample plots.

In order to analyse the effects of the number of sample plots and sample plot size (sample plot area) on the GSV estimation at the sample plot level, the

previously developed model was fitted to each variant of analysis, i.e. 800 sample plots of 500 m^2, then 800 sample plots of 400 m^2, …, down to 800

Table 2. Number of repetitions for each number of sample plots

Number of sample plots	Number of repetitions
800	100
700	200
600	300
500	400
400	500
300	600
200	700
100	800
50	900
25	1000

sample plots of 100 m^2. Subsequently, the model parameters were calibrated on 700 sample plots of 500 m^2, 400 m^2, …, down to 25 sample plots of 100 m^2. This process was repeated n times, according to a given variant of the number of sample plots (table 2). The model was evaluated at each repetition, by comparing the estimated and observed GSV values for 900 plots, independently for each variant of sample plot size. The following errors were calculated from the above described comparison for each variant of analysis and for each single sampling: relative root mean square error (%RMSE), relative mean absolute error (%MAE), relative BIAS (%BIAS) (equations: 2, 3, 4 respectively) and $R^2_{adjusted}$. Afterwards, medians and quartiles of error distribution from all repetitions for each variant were analysed in order to determine how GSV estimation accuracy and precision depends on the number and area of sample plots.

$$\%RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(GSV_{ALSi} - GSV_{REFi}\right)^2}{n}} / \overline{GSV_{REF}} * 100 \quad (2)$$

$$\%MAE = \frac{\sum_{i=1}^{n}\left|GSV_{ALSi} - GSV_{REFi}\right|}{n} / \overline{GSV_{REF}} * 100 \quad (3)$$

$$\%BIAS = \frac{\sum_{i=1}^{n} (GSV_{ALSi} - GSV_{REFi})}{n} / \overline{GSV_{REF}} * 100 \quad (4)$$

where: n – number of sample plots (for error computation it was always 900 sample plots regardless of the variant of analysis),

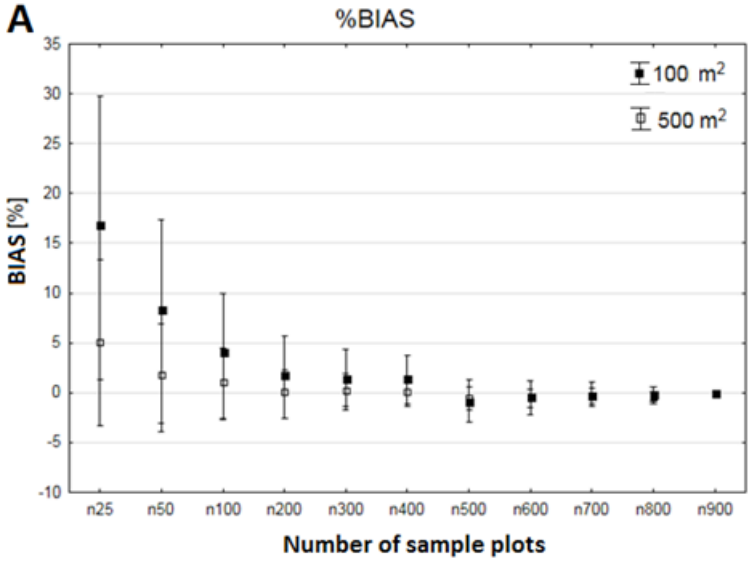
GSV_{ALSi} – growing stock volume of i-th sample plot, estimated for a given variant of analysis based on ALS point cloud metrics and the developed model,

GSV_{REFi} – growing stock volume of i-th sample plot, calculated on the basis of ground data for a given sample plot size variant,

$\overline{GSV_{REF}}$ – mean GSV of reference ground plot data.

Results

The ranges of %RMSE, %MAE, %BIAS and $R^2_{adjusted}$ values in each variant were different, according to the strength of the factor determining the given variant (either sample plot size and/or number of sample plots). For example; for 25 samples in 1000 iterations 99% of $R^2_{adjusted}$ values varied from -0.12 to 0.45 for 100 m^2 sample plots, and from 0.08 up to 0.74 for 500 m^2 plots. For 800 samples in 100 iterations corresponding $R^2_{adjusted}$ values varied from 0.471 to 0.472 for 100 m^2 sample plots, and from 0.749 to 0.750 for 500 m^2 plots.



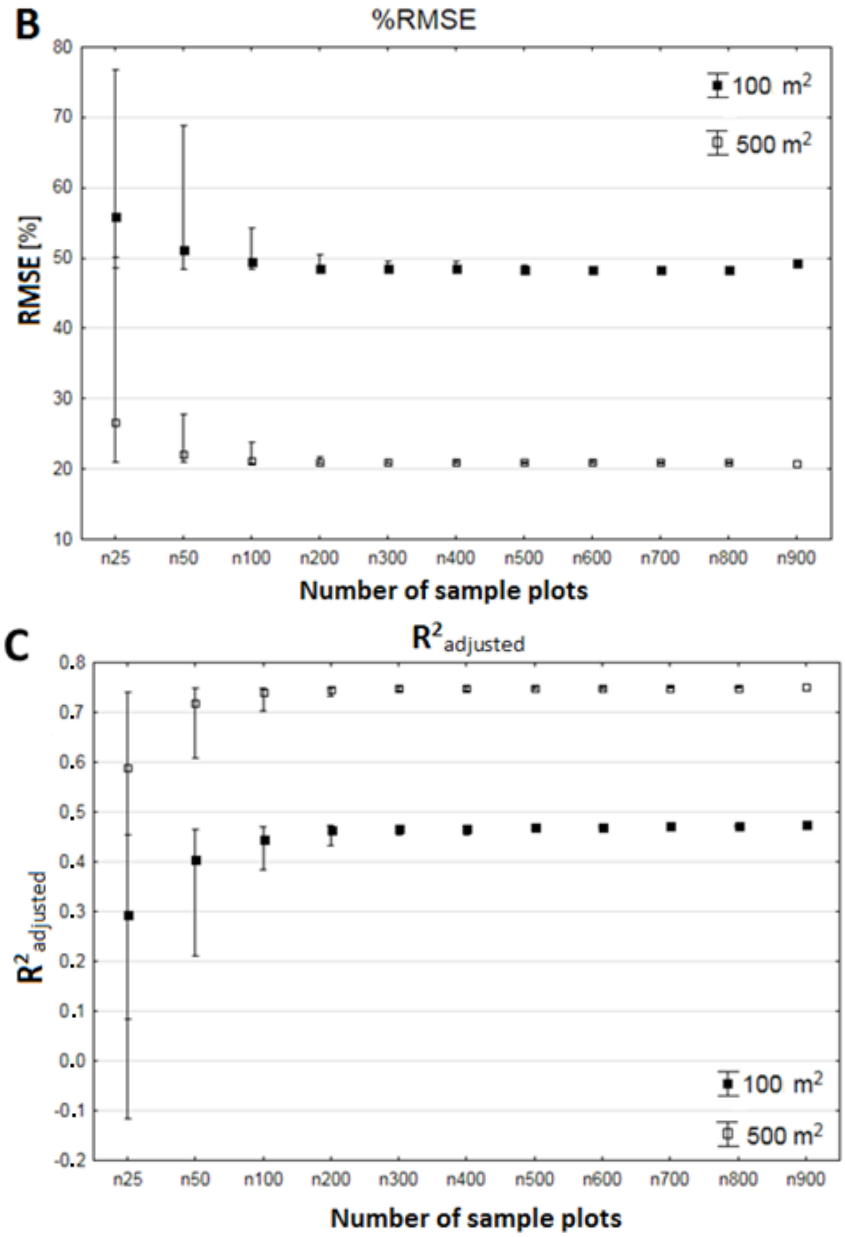


Fig. 3. Distribution of values for %BIAS (A), %RMSE (B) and $R^2_{adjusted}$ (C) for 100 and 500 m^2 sample plot size variants, for altering number of sample plots. Whiskers present range of obtained results as percentiles: 1-99%, squares presents medians from all iterations

The selected results are presented below (fig. 3). From the figures mentioned and the above quoted numbers, it is evident that the sample plot size affects the

ranges of possible cases (including extreme results) a lot more than the number of sample plots.

The highest values of %BIAS (fig. 4), %RMSE (fig. 5), %MAE (fig. 6) and the smallest values of $R^2_{adjusted}$ (fig. 7) were found in the smallest analysed sample plot size variants. The area of the sample plot limits the possible accuracy, regardless of the number of sample plots used to calibrate the model.

For all variants for which the number of sample plots equaled 100 or more and sample plots were at least 200 m^2, median %BIAS values were in the range of ±1% of the mean ground GSV value. For the 100 m^2 sample plot size this was also true, but the number of samples used for model calibration had to be at least 400-500.

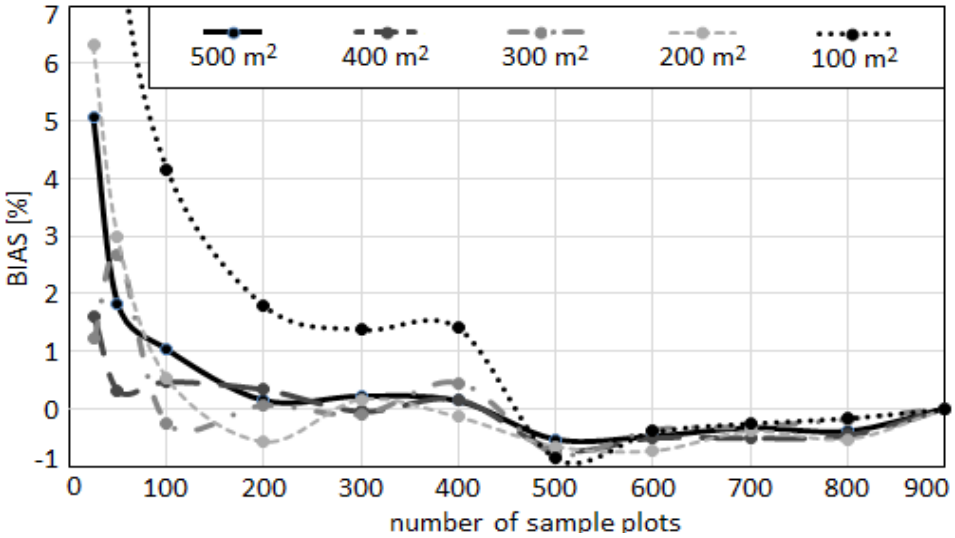


Fig. 4. Distribution of %BIAS values (medians of all iterations) according to sample plot size and number of sample plots

Regardless of the number of sample plots, the relative RMSE values increased with a decreasing sample plot size. The number of sample plots influenced the values of %RMSE only in the range of 25-200 samples, where %RMSE was higher in each sample plot size variant (fig. 3B, fig. 5) than at the 300-900 sample plots levels, where %RMSE maintained a practically constant level. The highest difference was noticed in the 25 sample variant, where %RMSE was in the range of 21.23-40.04% for 500 m^2, and 49.17-70.29% for 100 m^2 sample plot sizes. The smallest differences were noted for the 800 sample plot variant, where %RMSE was in the range of 20.88-20.91% for 500 m^2, and 48.32-48.40% for 100 m^2 sample plot size. In the above quoted values, there were the1st (lower bound) and the 99th (upper bound) percentiles of

obtained errors from all iterations in the given variant of analysis. The change of sample plot size from 500 to 100 m^2 caused %RMSE growth of about 30%.

The reduction of the number of sample plots per size variant from 900 to 25 enlarged the median error from 50% to 57% at 100 m^2, and from 21% to 27% at 500 m^2 sample plot size (fig. 5, fig. 8).

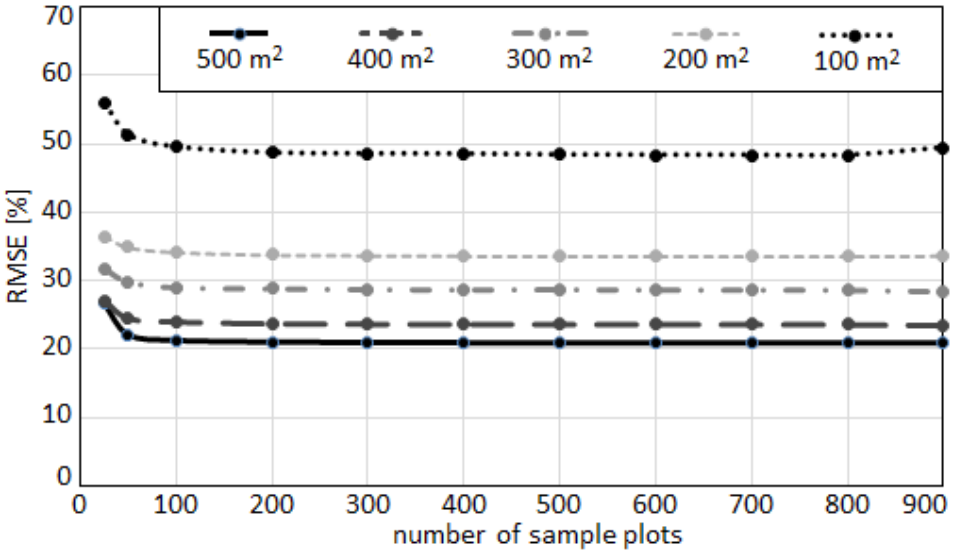


Fig. 5. Distribution of %RMSE values (medians of all iterations) according to sample plot size and number of sample plots

The trend of the %MAE distribution was fairly similar to the %RMSE distribution (fig. 5-6), i.e. apart from the 300 sample plots where we can observe a slight collapse, recorded errors were maintained at a very similar level (nearly straight trend lines). These lines show a deflection only below the 300-400 sample plots used for model calibration. Similar patterns were observed for the %MAE error distribution, where the number of sample plots only affected the values of %MAE in the range of the 25-300 plots. Enlarging the number of sample plots did not change the %MAE (and %RMSE) values. The highest difference was noted for the 25 sample plots variant, where %MAE values were in the range of 14.66-29.95% for 500 m^2, and 31.38-48.28% for 100 m^2. The smallest differences were noted for the 800 sample plots variant, where %MAE values were in the range of 14.35-14.44% for 500 m^2, and 31.12-31.53% for sample plots of 100 m^2 size. %MAE values had a similar trend to %RMSE values; however, they were generally about 30% lower than %RMSE (fig. 5-6). Moreover, the differences between %RMSE and %MAE indicated that there were some extreme values in the obtained results.

The $R^2_{adjusted}$ values for each variant of the number of sample plots stabilized around 300-400 sample plots. Even for the 100 sample plots variant, the final

median value was very similar to the full variant, i.e. 900 sample plots (fig. 3C, fig. 7).

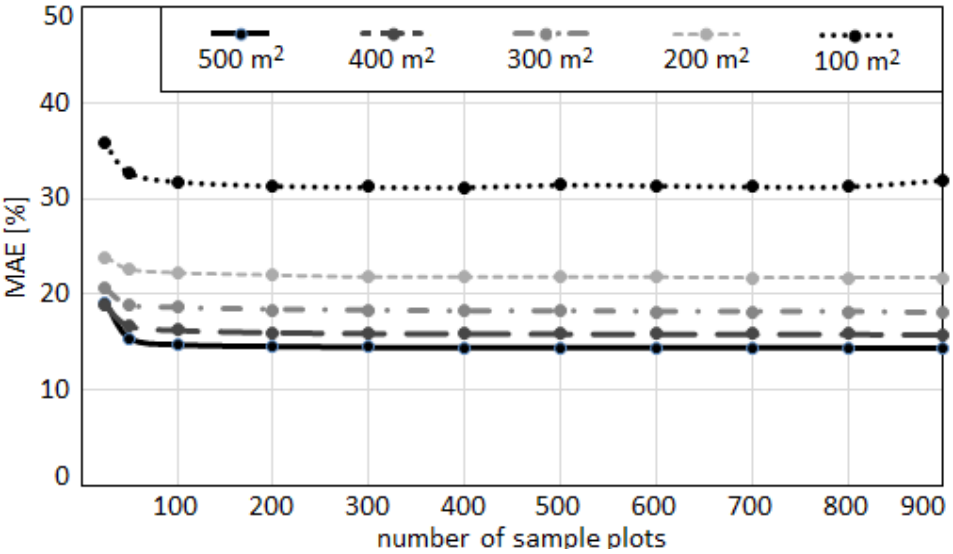


Fig. 6. Distribution of %MAE values (medians of all iterations) according to sample plot size and number of sample plots

Table 3. Distribution of $R^2_{adjusted}$ ranges according to the analyzed variants. The range considered the difference between the highest and the lowest $R^2_{adjusted}$ values among all iterations (repetitions) for given variants of analysis

n_samples	n_iter	100 m^2	200 m^2	300 m^2	400 m^2	500 m^2
25	1000	0.46	0.54	0.62	0.61	0.70
50	900	0.37	0.23	0.16	0.16	0.19
100	800	0.15	0.09	0.10	0.06	0.08
200	700	0.05	0.04	0.03	0.02	0.02
300	600	0.03	0.02	0.02	0.01	0.01
400	500	0.03	0.01	0.01	0.01	0.01
500	400	0.01	0.01	0.01	0.01	0.00
600	300	0.01	0.01	0.01	0.00	0.00
700	200	0.00	0.00	0.00	0.00	0.00
800	100	0.00	0.00	0.00	0.00	0.00

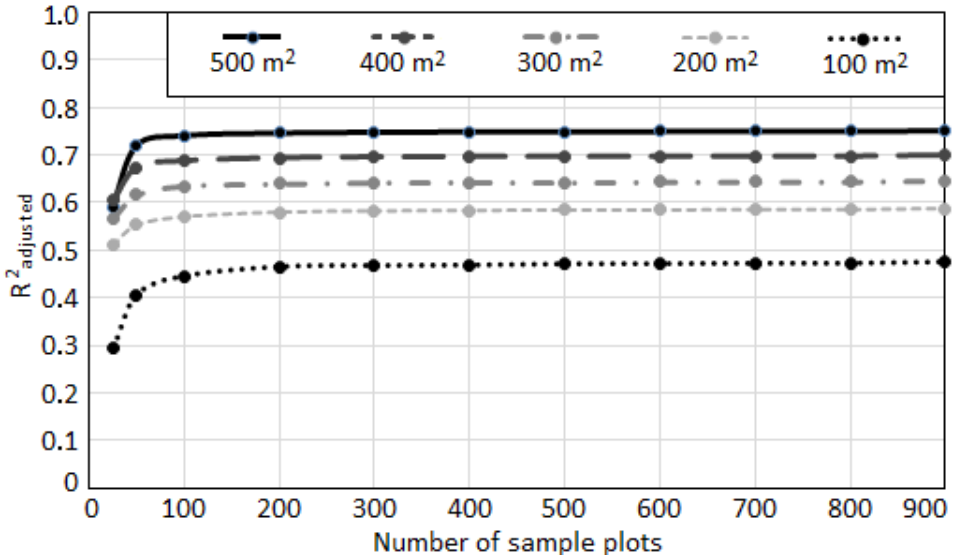


Fig. 7. Distribution of $R^2_{adjusted}$ values (medians of all iterations) according to sample plot size and number of sample plots

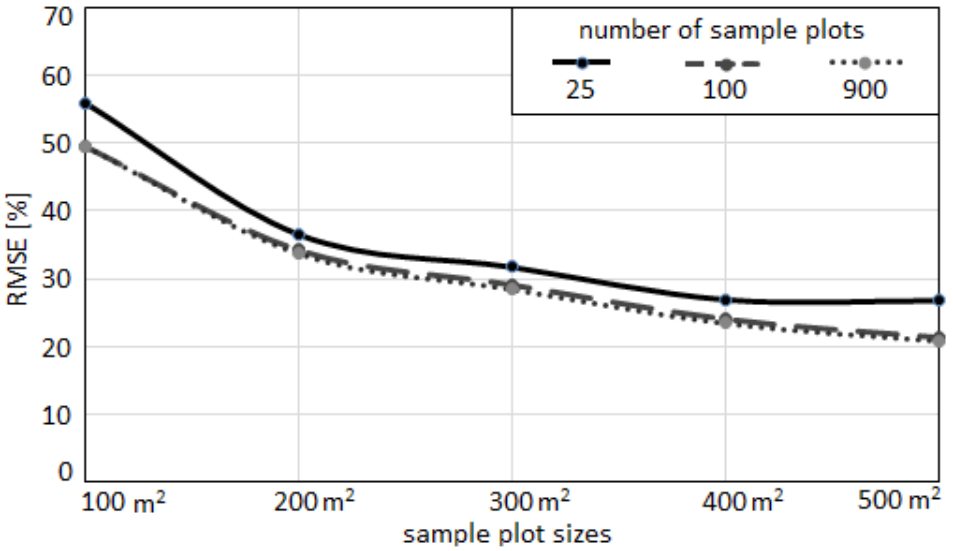


Fig. 8. Distribution of received %RMSE values according to all sample plot sizes and selected variants of sample plot number

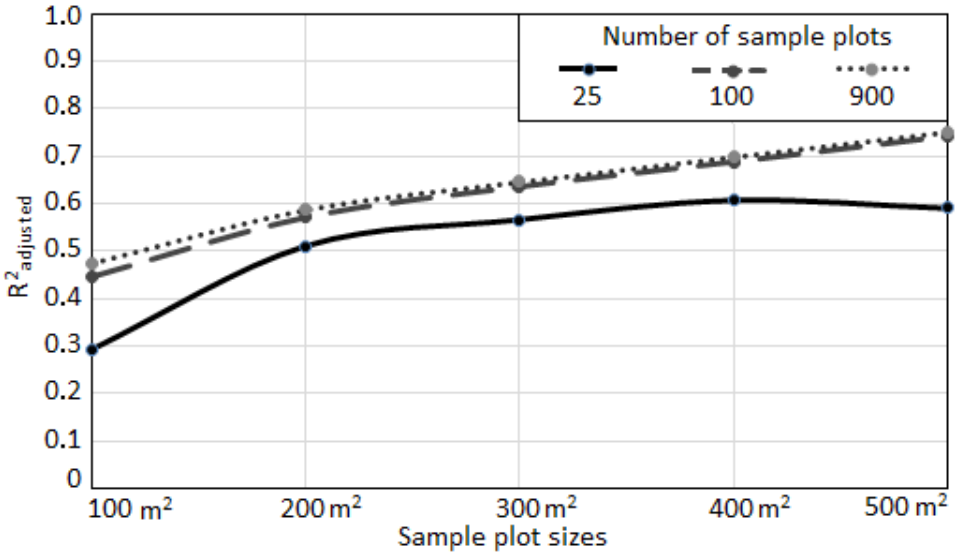


Fig. 9. Distribution of received $R^2_{adjusted}$ values according to all sample plot sizes and selected variants of sample plot number

On the basis of table 3, it could be said that the $R^2_{adjusted}$ ranges decrease with the number of sample plots used for model calibration. Approximately, in up to 300-400 sample plots, the range of obtained $R^2_{adjusted}$ values between the best and worst score did not exceed 1%. Higher differences were observed below that number of sample plots. With respect to sample plot size, broader $R^2_{adjusted}$ ranges were observed for smaller sample plots. In some instances, the obtained ranges were lower for smaller sample plot size. This was probably connected with a shift in both minimal and maximal $R^2_{adjusted}$ values among subsequent variants of analysis. This means that for 25 sample plots of 500 m^2 size, the range between min and max $R^2_{adjusted}$ was much broader than for 25 sample plots of 100 m^2. Apart from the broader range, the maximum value of $R^2_{adjusted}$ was also higher for 25 sample plots of 500 m^2, and there were more scores (sampling results) observed in the higher interval of $R^2_{adjusted}$ values (fig. 3C, table 3).

Discussion and conclusions

The modeling results contain all kinds of errors: starting from the field measurements and ALS data acquisition, through to the series of different data processing steps, up to all the errors related to statistical modelling. Using a linear regression model, we noted that a relatively small number of sample plots was enough to stabilize the relationship between ALS metrics and growing stock volume at the sample level. In the presented study, about 300-400 sample

plots were enough to keep %BIAS, %RMSE, %MAE and $R^2_{adjusted}$ at relatively similar levels to the higher numbers of sample plots. Furthermore, starting from about 300-400 sample plots, the effect of a single sampling should not be of great importance, as the range of obtained errors (precision) was very close to the median error value providing that the applied sample plot selection method covers the whole distribution of sampled stand ages and GSVs. Regarding the issues connected with sample plot size, one can state that the increase in a sample plot size (up to a certain area) leads to a better accuracy of ALS based GSV. Such a statement is true at least at the level of sample plot and for the investigated study area.

Since only one form of the model was used, other possible sources of errors were excluded. However, it is possible that suboptimal models were used for some variants of analysis. Nevertheless, the obtained accuracy of the general model introduced in this paper was comparable with other studies, which have been mainly carried out in Nordic countries [Næsset 1997; Holmgreen et al. 2003; Kallio et al. 2010; Maltamo et al. 2016].

The selection of sample plots in this study almost fully represents the possible variability in plot GSV values of populations. It is clear that regression models enable the extrapolation of the modelled variable. Since all the sample plots per variant were available to be selected for model calibration, multiple sampling provided information on the model variation. The results show a very narrow range of received %BIAS, %RMSE, %MAE and $R^2_{adjusted}$ values for variants above 200 sample plots (fig. 3). It is in agreement with the findings of Gobakken and Næsset [2008] who claim that the average standard deviation increases when the number of sample plots is reduced. Furthermore, it means that in general, the modelling procedure was stable and extreme errors were presented in only a small percentage of cases.

Our results confirm that there is a strong relationship between the size of the sample plots and the accuracy of GSV prediction at the sample plot level [Gobakken and Næsset 2008; Frazer et al. 2011; Ruiz et al. 2014]. However, in the presented case we utilized a much higher number of available sample plots, (900) compared to earlier studies. We were able to capture any trends that depended on the number of samples used in modelling. We noticed that below the quantity of about 200-300 sample plots, a downtrend of modelling results ($R^2_{adjusted}$) accelerate. Thus, it is possible to indicate a minimum number of sample plots that guarantees the best possible accuracy and precision, for a specified sample plot size. This is in agreement with the previous study of Kallio et al. [2010].

The strongest effect on ALS based GSV estimation accuracy comes from the sample plot size. From a practical perspective, there is a specified (maximum) accuracy limit, which is related to the number of sample plots. There is an easily visible pattern on all presented graphs in this paper, where each sample plot size

variant reaches its best %BIAS, %RMSE, %MAE and $R^2_{adjusted}$ values at around 300-400 (or more) sample plots, and does not change significantly with higher numbers of sample plots. There is no need to collect more field data because we do not expect to improve the model used in the study. Our results could possibly reduce the cost of field inventory campaigns [Ruiz et al. 2014]. The reduction by half of field work costs, compensates for the costs of ALS data acquisition and data processing.

Future investigations should include similar analyses, based on other modelling methods and other possible ALS metrics derived from canopy height models (CHM) or individual tree detection (ITD) results. Additionally, artificially reduced point clouds densities or alternative ways of calculating ALS metrics [Hayashi et al. 2015] should be evaluated. Going forward, the obtained results could also be analysed separately for specific species, age and/or site classes in different locations, in order to check which stands cause the highest errors, and/or to develop site-independent ALS data based models for predicting not only GSVs values, but also estimating different (single) forest stand characteristics.

References

Bruchwald A., Dudek A., Michalak K., Rymer-Dudzińska T., Wróblewski L., Zasada M. [2000]: Wzory empiryczne do określania wysokości i pierśnicowej liczby kształtu grubizny drzewa (Empirical formulae for defining height and dbh shape figure of thick wood). Sylwan 144 [10]: 5-13

Frazer G.W., Magnussen S., Wulder M.A., Niemann K.O. [2011]: Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR--derived estimates of forest stand biomass. Remote Sensing of Environment 115 [2]: 636-649

Gobakken T., Næsset E. [2008]: Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. Canadian Journal of Forest Research 38 [5]: 1095-1109

Gobakken T., Næsset E., Nelson R., Bollandsås O.M., Gregoire T.G., Ståhl G., Holm S., Ørka H.O., Astrup R. [2012]: Estimating biomass in Hedmark County, Norway using national forest inventory field plots and airborne laser scanning. Remote Sensing of Environment 123 [2012]: 443-456

Hayashi R., Kershaw J.A., Weiskittel A.R. [2015]: Evaluation of alternative methods for using LiDAR to predict aboveground biomass in mixed species and structurally complex forests in northeastern North America. Mathematical and Computational Forestry and Natural-Resource Sciences 7 [2]: 49-62

Holmgren J., Nilsson M., Olsson H. [2003]: Estimation of tree height and stem volume on plots using airborne laser scanning. Forest Science 49 [3]: 419-428

Holmström H., Nilsson M., Ståhl G. [2001]: Simultaneous estimations of forest parameters sing aerial photograph interpreted data and the k nearest neighbour method. Scandinavian Journal of Forest Research 16 [1]: 67-78

Hyyppä J., Yu X., Hyyppä H., Vastaranta M., Holopainen M., Kukko A., Kaartinen H., Jaakkola A., Vaaja M., Koskinen J., Alho P. [2012]: Advances in forest inventory using Airborne Laser Scanning. Remote sensing 2012 [4]: 1190-1207

Kallio E., Maltamo M., Packalén P. [2010]: Effect of sampling intensity on the accuracy of species-specific volume estimates derived with aerial data: A case study on five privately owned forest holdings. Proceedings of: 10th International Conference on LiDAR Applications for Assessing Forest Ecosystems, 14-17 September 2010. Freiburg, Germany: 169-178

Lim K., Treitz P., Baldwin K., Morrison I., Green, J. [2003]: Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. Canadian Journal of Remote Sensing 29 [5]: 648-678

Lindner M., Karjalainen T. [2007]: Carbon inventory methods and carbon mitigation potentials of forests in Europe: a short review of recent progress. European Journal of Forest Research 126 [2]: 149-156

Maltamo M., Eerikäinen K., Packalén P., Hyyppä J. [2006]: Estimation of stem volume using laser scanning-based canopy height metrics. Forestry 79 [2]: 217-229

Maltamo M., Bollandsäs O.M., Gobakken T., Næsset E. [2016]: Large-scale prediction of aboveground biomass in heterogeneous mountain forests by means of Airborne Laser Scanning. Canadian Journal of Forest Research 46 [9]: 1138-1144

McKendry P. [2002]: Energy production from biomass (part 1): overview of biomass. Bioresource Technology 83 [1]: 37-46

Miścicki S., Stereńczak K. [2012]: A two-phase inventory method for calculating standing volume and tree-density of forest stands in central Poland based on Airborne Laser Scanning data. Forest Research Papers 74 [2]: 127-136

Næsset E. [1997]: Estimating timber volume of forest stands using Airborne Laser Scanner data. Remote Sensing of Environment 61 [2]: 246-253

Næsset E. [2002]: Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80 [1]: 88-99

Næsset E., Gobakken T. [2005]: Estimating forest growth using canopy metrics derived from airborne laser scanner data. Remote Sensing of Environment 96: 453-465

Ruiz L.A., Hermosilla T., Mauro F., Godino M. [2014]: Analysis of the influence of plot dize and LiDAR density on forest structure attribute estimates. Forests 5 [5]: 936-951

Tonolli S., Dalponte M., Vescovo L., Rodeghiero M., Bruzzone L., Gianelle D. [2011]: Mapping and modelling forest tree volume using forest inventory and airborne laser scanning. European Journal of Forest Research 130 [4]: 569-577

Acknowledgements

This study has been supported by the project REMBIOFOR (Remote sensing based assessment of woody biomass and carbon storage in forests), supported by The National Centre for Research and Development under the BIOSTRATEG program, agreement no. BIOSTRATEG1/267755/4/NCBR/2015.

Submission date: 27.10.2017

Online publication date: 19.04.2018