

Airborne Laser Scanning Outperforms the Alternative 3D Techniques in Capturing Variation in Tree Height and Forest Density in Southern Boreal Forests

MIKKO VASTARANTA ^{1,2*}, TUOMAS YRTTIMAA ^{1,2}, NINNI SAARINEN ^{1,2}, XIAOWEI YU ³, MIKA KARJALAINEN ³, KIMMO NURMINEN ³, KIRSI KARILA ³, VILLE KANKARE ^{1,2}, VILLE LUOMA ¹, JIRI PYÖRÄLÄ ¹, SAMULI JUNTILA ¹, TOPI TANHUANPÄÄ ¹, HARRI KAARTINEN ^{3,4}, ANTERO KUKKO ³, EIJA HONKAVAARA ³, ANTONI JAAKKOLA ³, XINLIAN LIANG ³, YUNSHENG WANG ³, MATTI VAAJA ⁵, HANNU HYYPPÄ ⁵, MASATO KATOH ⁶, MICHAEL A. WULDER ⁷, MARKUS HOLOPAINEN ¹ AND JUHA HYYPPÄ ³

¹Department of Forest Sciences, University of Helsinki, P.O. Box 27, 00014 Helsinki, Finland

²School of Forest Sciences, University of Eastern Finland, P.O. Box 111, 80101 Joensuu, Finland

³Finnish Geospatial Research Institute, National Land Survey, Geodeetinrinne 2, 02431 Masala, Finland

⁴Department of Geography and Geology, University of Turku, FI-20014 Turku, Finland

⁵Research Institute of Measuring and Modelling for the Built Environment, Aalto University, P.O. Box 15800, 00076 Aalto, Finland

⁶Institute of Mountain Science, Shinshu University, 8304, Minamiminowa-Village, Kamiina-County, Nagano 399-4598, Japan

⁷Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, 506 West Burnside Road, Victoria, British Columbia V8Z 1M5, Canada

* Corresponding author: mikko.vastaranta@uef.fi; tel.: +358-50-512-7051

Vastaranta, M., Yrttimaa, T., Saarinen, N., Yu, X., Karjalainen, M., Nurminen, K., Karila, K., Kankare, V., Luoma, V., Pyörälä, J., Juntila, S., Tanhuanpää, T., Kaartinen, H., Kukko, A., Honkavaara, E., Jaakkola, A., Liang, X., Wang, Y., Vaaja, M., Hyypä, H., Katoh, M., Wulder, M.A., Holopainen, M. and Hyypä, J. 2018. Airborne Laser Scanning Outperforms the Alternative 3D Techniques in Capturing Variation in Tree Height and Forest Density in Southern Boreal Forests. *Baltic Forestry* 28(2): 268–277.

Abstract

The objective of this study is to better understand the relationship between forest structure and point cloud features generated from certain airborne and space borne sensors. Point cloud features derived from airborne laser scanning (ALS), aerial imagery (AI), WorldView-2 imagery (WV2), TerraSAR-X, and Tandem-X (TDX) data were classified as features characterizing forest height and density as well as variation in tree height. Correlations between these features and field-measured attributes describing forest height, density and tree height variation were investigated at plot scale. From the field-measured attributes, basal area (G) and the number of trees per unit area (N) were used as forest density indicators whereas maximum tree height (H_{max}) and standard deviation in tree height (H_{std}) were used as indicators for forest height and tree height variation, respectively. In the analyses, field observations from 91 sample plots (32 m × 32 m) located in southern Finland were used. Even though ALS was found to be the most accurate data source in characterizing forest structure, AI, WV2, and TDX were also capable of characterizing forest height at plot scale with correlation coefficients stronger than 0.85. However, ALS was the only data source capable of providing separate features for characterizing also the variation in tree height and forest density. Features related to forest height, generated from the other data sources besides ALS, also provided strongest correlation with the forest density attributes and variation in tree height, in addition to H_{max} . Due to these more diverse characterization capabilities, forest structural attributes can be predicted more accurately by using ALS, also in the areas where the relation between the attributes of interest is not solely dependent on forest height, compared to the other investigated 3D remote sensing data sources.

Keywords: airborne laser scanning; lidar; photogrammetry; radargrammetry; interferometry; tree; mapping

Introduction

Forest planning and wood procurement require information on forest structure from an area of interest (Bettinger et al. 2010, Kangas and Maltamo 2006). In forest inventories, forest structure is most often characterized by using attributes such as species composition, basal area (G), the number of trees per unit area (N), mean diameter at breast height (dbh) and mean height for each stand. Mean dbh and height attributes are most often weighted by G to better describe characteristics of dominant trees (Husch et al. 1982). Remote sensing has been used for decades to aid forest inventories often covering large and somewhat remote areas with limited access (Wulder and Franklin 2012). Since the 1960s, aerial imagery (AI) has been used for delineating stand boundaries and, in conjunction with sample plot data with measured forest structural attributes, visual interpretation and predicting forest attributes are undertaken for the entire area of interest (Poso et al. 1984, Wulder et al. 2012). If more accurate estimates are required, forest structural characteristics are determined in the field after stand boundaries have been delineated (e.g. Holopainen et al. 2010, Wulder et al. 2012). Over the previous decade, airborne laser scanning (ALS) and AI, together with data from accurately measured sample plots have emerged as information sources to produce forest inventories (Næsset et al. 2004, Hyypä et al. 2008, Wulder et al. 2013). ALS have been proven to be a rich information source for forest mapping as it can provide information on forest height and density which are strongly correlated with many required forest inventory attributes (White et al. 2013a, Bouvier et al. 2015).

The acquisition of detailed forest resource information may have a low expected return on investment in some jurisdictions due to large areas and extensive forest management practices or area of interest is a non-merchantable such as national parks or mountain ranges and thus, costs of the ALS-based inventory may persist as a limiting factor (Wulder et al. 2013). In countries applying intensive forestry, the cost of the ALS data acquisitions and logistics are, on the other hand, limiting the time interval between the inventories. Thus, although ALS has proven to be a valuable resource for forest inventory, there are also other alternative sensors capable of providing information on forest structure (see e.g. Rahlf et al. 2014, Yu et al. 2015). These alternatives have preferable characteristics for forest inventory as well as mapping and monitoring forest resources for large areas with required time interval. Stereo interpretation of aerial or satellite imagery could provide improved spatial coverage and spectral information in addition to point clouds with reduced costs (White et al. 2013b). The image-based techniques have also reached the same level, and in some

cases have even provided more accurate estimations than ALS for forest inventory attributes (see Yu et al. 2015). Point clouds derived from space-borne synthetic aperture radar (SAR) data can especially be cost-efficient in providing large data coverage and enhanced temporal resolution on forested areas (Karjalainen et al. 2012, Solberg et al. 2013, Karila et al. 2015). It should be pointed out that ALS is the only technique that can be used for deriving accurate digital terrain model (DTM) for vegetated areas (e.g. White et al. 2013b). Therefore, if ALS is not used, accurate DTM has to already be available.

The objective of the study was to better understand the strengths and limitations of certain 3D remote sensing data sources for measuring forest height as well as capturing variation in tree height and forest density, which are the main characteristics that build the base for accurate forest attribute estimates (e.g. White et al. 2013a, Bouvier et al. 2015, White et al. 2016). To be able to examine the relationship between point cloud features from certain airborne and space borne sensors and forest structural attributes such as maximum tree height (H_{\max}), standard deviations in tree height (H_{std}), G and N, as well as correlations between them were investigated. By using simple statistical methods, we concentrated on testing the capabilities of the data itself and avoided the possibility of over fitting the prediction models. This study is a continuation of a study by Yu et al. (2015), where a basic suite of forest inventory attributes (e.g. volume and aboveground biomass) was predicted using ALS, AI and WorldView-2 satellite imagery (WV2), Tandem-X interferometry (TDX), and TerraSAR-X radar-grammetry (TSX). To enhance the existing knowledge, we investigated capabilities of these 3D remotely sensed data sources in producing point clouds for characterizing forest structural attributes and concentrating especially on forest height, variation in tree height, and forest density, in contrast to Yu et al. (2015).

Materials and Methods

Study area and field observations

The study area of ca. 2000 ha is located in the southern boreal forest zone (61.19° N, 25.11° E), approximately 145 km from Helsinki, Finland. Scots pine, *Pinus sylvestris* L., and Norway spruce, *Picea abies* (L.) H. Karst., stands dominate the undulating landscape with elevation varying from 125 m to 185 m above sea level. The stands are mainly even-aged and single layered, with an average stand size of slightly less than 1 ha. In total, 91 rectangular sample plots with 1024 m² area were mapped and measured in the field during summer of 2014. The sample plot locations were selected based on the existing ALS data collected in 2012 by National Land Survey of Finland to represent the variation in forest height and

densities in the study area. The sample plots represented a range of forest structural conditions (Table 1). The sample plot locations were determined using a total station (5602, Trimble, Sunnyvale, California, United States), which was oriented to local coordinate system using ground control points measured with virtual reference station global navigation satellite system (R8, Trimble, Sunnyvale, California, United States) in open areas close to each plot. From the sample plots, all trees having a dbh larger than 5 cm were measured with steel calipers from two directions perpendicular to each other and the dbh was calculated as the mean of these two. Tree height was measured using an electronic hypsometer (Vertex IV, Haglöf, Sweden) from all the trees as well. The precision of tree dbh and height measurements is expected to be approximately 0.3 cm and 0.5 m, respectively, based on our evaluations (Luoma et al. 2017). Then, the attributes of interest (i.e. H_{max} , H_{std} , G and N) were obtained from the tree data for the sample plots. H_{max} was simply the height of the tallest tree from each sample plot. H_{std} , G and N were calculated for each sample plot as follows:

$$H_{std} = \sqrt{\frac{\sum_{i=1}^n (h_i - \bar{h})^2}{n}} \quad (1)$$

$$G = \frac{\sum_{i=1}^n g_i}{A} \quad (2)$$

$$N = \frac{n}{A} \quad (3)$$

where n is the number of trees in a sample plot, h_i is the height for i^{th} tree, \bar{h} is the average tree height in the sample plot, g_i is the basal area of tree, and A is the area of the sample plot in hectares.

Table 1. Descriptive statistics, such as minimum (min), mean, maximum (max), and standard deviation (Sdev), for forest structural attributes on sample plots ($n = 91$). G = basal area, N = number of trees per unit area, H_G = basal-area weighted mean height, H_{max} = maximum tree height, H_{std} = standard deviation in tree height, and Vol = stem volume

Forest attribute	Min	Mean	Max	Sdev
G (m²/ha)	5.8	26.2	42.9	7.4
N (1/ha)	342	904	2881	530
H_G (m)	10.5	21.3	31.4	4.5
H_{max} (m)	15.1	26.8	37.5	5.2
H_{std} (m)	1.4	5.9	11.7	2.5
Vol (m³/ha)	31.4	266.5	507.7	106.1

Remotely sensed data and generation of features from remote sensing data

The remote sensing data sets were acquired from the study area during summer 2014 (Table 2). ALS data from 2500 m altitude (ALS-2500) were collected on May

22 and from 900 m altitude (ALS-900) on September 5. Both of the ALS data sets were collected using a Leica ALS70-HA SN7202 system (Leica Geosystems AG, Heerbrugg, Switzerland). The AI was acquired at 5 km altitude on May 22, 2014. The imaging sensor used was a Z/I Imaging Digital Mapping Camera (DMC) and images were captured with a stereoscopic forward overlap of 80% and side overlap of 64%. The area was covered by 24 images in total and stereo matching was done using BAE Systems Socet Set software. The same software was used for producing point clouds from a cloud-free WV2 image pair acquired on July 11, 2014 as well as from a six TSX images that were acquired in a time period of 10 days. TDX data acquired on June 5 were used for deriving interferometric point cloud using the SARscape 5.0 software package in the processing. All of the remote sensing data sets and the processing of the data are described more detailed in Yu et al. (2015). It should be noted that point cloud heights for all data sets were normalized as heights above ground using digital terrain model (DTM) derived from ALS-900 data.

Table 2. Remote sensing data sets used. ALS-900 = airborne laser scanning from 900 m altitude, ALS-2500 = airborne laser scanning from 2500 m altitude, AI = aerial imagery, WV2 = WorldView-2 satellite imagery, TDX = Tandem-X interferometry, and TSX = TerraSAR-X radargrammetry

	Acquisition date	Flying altitude, km	Realized mean point density in the sample plots, points/m ²
ALS-900	September 5, 2014	0.9	11.96
ALS-2500	May 22, 2014	2.5	1.18
AI	May 22, 2014	5.0	1.00
WV2	July 11, 2014	770	0.90
TSX	June 29-July 9, 2014	514	0.06
TDX	June 5, 2014	~500	0.05

The same set of features describing forest height, variation in tree height, and forest density were derived from each point cloud data and for each sample plot area to be used in further investigations. In Table 3, derived features and their definitions are presented. Most of the features were calculated from the normalized points above the 2-m threshold.

Investigating the relationship between point cloud features and forest structural attributes

Pearson correlation coefficients (r) were calculated between remote sensing-based features and field-observed forest structural attributes. The correlations were utilized when investigating the capability of remote sensing data sets to characterize forest structural attributes and when selecting explanatory variables for linear regression models to be generated.

Table 3. Features extracted from the normalized point clouds

Feature category	Feature	Definition
Forest height	h_{max}	Maximum of the normalized heights of all points
	h_{mean}	Arithmetic mean of normalized height of all points above 2-m threshold
	h_{mode}	Mode of normalized height of all points above 2-m threshold
	hp_{10} hp_{90}	10% to 90% percentiles of normalized height of all points above 2-m threshold with a 10% increment
Forest density	pr	Penetration ratio; $N_{h=2}/N_{total} * 100$, where $N_{h=2}$ is the number of all points, and $N_{h=2}$ the number of points below and equal to 2 m.
	cc_{1-CC10}	$cc_i = N_i/N_{total}$, where $i = 1$ to 10, N_i is the number of points within i^{th} layer when tree height was divided into 10 intervals starting from 2 m, N_{total} is the number of all points.
Variation in tree height	h_{std}	Standard deviation of normalized height of all points above 2-m threshold
	h_{cv}	Coefficient of variation calculated as h_{std}/h_{mean}

Linear regression models were developed for predicting forest structural attributes namely H_{max} , H_{std} , G and N , using R statistical software (R Core Team 2017). Prediction models were developed based on results of correlations analyses between the features derived from the point clouds from different sensors and observed forest structural attributes. In our prediction models, height-related forest structural attributes (H_{max} and H_{std}) were modelled with only a single remote sensing feature related to forest height or height variation as an explanatory variable (See feature categories in Table 3). Thus, the feature with the strongest correlation with the field-observed height attribute was included in the model. G and N were modelled with one feature used for characterizing forest density and another for forest height. At first, a feature related to forest density having the strongest correlation with the density-related forest attribute (i.e. N or G) was included in the model. Then, a height-related feature with the strongest correlation with field-observed G or N was included in a respective model, but only after considering the correlations between the possible predictors: the maximum correlation between the candidate features describing forest density and height used as explanatory variables in the model was limited in 0.9 (Hudak et al. 2008). Features generated from the point clouds from different sensors also describe variation in tree height. In addition to information on forest height and density, information about variation in tree height was assumed to bring additional explanatory power to the prediction models for G and N . Therefore, features describing tree height variation (either h_{std} or h_{cv}) were also tested as an additional explanatory variable. If residual plots revealed a dependence or trend for a specific explanatory variable, simple transformations were tested (e.g., different powers or logarithmic transformations). Statistical significance of each feature in the model was also considered, and only features with $p < 0.05$ were finally accepted as explanatory variables.

Accuracy assessment

The accuracy of the developed prediction models was evaluated with leave-one-out cross validation and by calculating bias, root-mean-square error (RMSE), correlation coefficient (r), and coefficient of determination (R^2) between the estimated and the observed values of forest structural attributes (i.e. H_{max} , H_{std} , G and N). Bias and RMSE were calculated as follows:

$$bias = \frac{\sum_{i=1}^n (\hat{X}_i - X_i)}{n} \tag{4}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{X}_i - X_i)^2}{n}} \tag{5}$$

where n is the number of plots, \hat{X}_i is the model-estimated value for plot i based on remotely sensed data, and X_i is the field-observed value for plot. The relative bias and RMSE were calculated as a relation to the mean of the field-observed value of $X(\bar{X})$:

$$bias (\%) = \frac{bias}{\bar{X}} \times 100 \tag{6}$$

$$RMSE (\%) = \frac{RMSE}{\bar{X}} \times 100 \tag{7}$$

The standard deviation in the estimated attributes was also computed and compared to the corresponding field-observed ones. Student's t -test was used for assessing the statistical significance of bias of the estimated attributes.

Results

Correlations between the field-observed N , G , H_{max} , and H_{std} , and the features derived from the remote sensing data sets are presented in Tables 4 and 5.

From the features classified as describing forest height, high percentiles (hp_{70} - h_{max}) derived from ALS had the strongest correlation with H_{max} ($r > 0.95$) (Figure 1). The same percentiles derived from the AI, WV2, and TDX provided correlation coefficients of 0.94, 0.92, and 0.87 for H_{max} , respectively. Height features derived from TSX reached weaker correlation coefficients ($r < 0.72$) with H_{max} .

Differences between the remote sensing data sets were more apparent when analysing correlations between the derived features describing variation in tree height as well as forest density and field-measured H_{std} , G and N (Figures 2 and 3; Tables 4 and 5). With both ALS data sets (i.e. with flying altitude of 900 m and 2500 m), the derived h_{std} correlated strongly ($r = 0.89$ - 0.91) with the field measured H_{std} . Standard deviation (h_{std}) derived with the point clouds from any other 3D data set, correlated only moderately with the field-measured H_{std} ($r < 0.60$).

Table 4. Correlations between the features generated from point clouds from different sensors and the field-observed forest structural attributes. The features are grouped by categories explained in Table 3. The darkness of grey indicates stronger correlation. N = Number of trees per unit area, G = basal area, H_{max} = maximum tree height, and H_{std} = standard deviation in tree height. ALS-900 = airborne laser scanning from 900 m altitude, ALS-2500 = airborne laser scanning from 2500 m altitude, AI = aerial imagery

Feat.\Attr.	ALS-900				ALS-2500				AI				
	N	G	H_{std}	H_{max}	N	G	H_{std}	H_{max}	N	G	H_{std}	H_{max}	
Var.	h_{std}	-0.02	0.74	0.89	0.90	0.00	0.72	0.91	0.89	-0.21	0.02	0.23	0.31
	h_{cv}	0.40	0.43	0.55	0.45	0.41	0.44	0.61	0.48	-0.12	-0.56	-0.32	-0.37
Forest height	h_{mean}	-0.45	0.58	0.69	0.81	-0.48	0.58	0.67	0.82	-0.19	0.77	0.80	0.90
	h_{mode}	-0.20	0.42	0.47	0.44	-0.22	0.34	0.37	0.37	-0.16	0.62	0.65	0.69
	hp_{10}	-0.53	-0.04	-0.09	0.14	-0.52	-0.10	-0.26	0.02	-0.10	0.70	0.67	0.72
	hp_{20}	-0.60	0.08	0.08	0.28	-0.61	0.05	-0.03	0.20	-0.18	0.72	0.72	0.79
	hp_{30}	-0.57	0.29	0.37	0.47	-0.61	0.28	0.30	0.45	-0.22	0.73	0.77	0.85
	hp_{40}	-0.50	0.45	0.58	0.64	-0.52	0.46	0.56	0.65	-0.22	0.74	0.79	0.87
	hp_{50}	-0.42	0.57	0.72	0.77	-0.42	0.58	0.71	0.79	-0.21	0.75	0.80	0.89
	hp_{60}	-0.35	0.65	0.78	0.85	-0.35	0.66	0.79	0.87	-0.20	0.76	0.81	0.91
	hp_{70}	-0.32	0.69	0.81	0.89	-0.32	0.69	0.82	0.91	-0.20	0.76	0.81	0.92
	hp_{80}	-0.30	0.70	0.83	0.92	-0.30	0.70	0.83	0.94	-0.20	0.76	0.82	0.94
	hp_{90}	-0.29	0.71	0.84	0.95	-0.29	0.71	0.84	0.95	-0.20	0.75	0.82	0.94
	h_{max}	-0.21	0.72	0.80	0.97	-0.21	0.73	0.82	0.98	-0.22	0.67	0.81	0.92
	pr	-0.44	-0.87	-0.58	-0.66	-0.43	-0.86	-0.62	-0.67	-0.17	-0.59	-0.24	-0.37
	Forest density	cc_1	0.42	0.08	0.23	0.07	0.44	0.13	0.31	0.13	-0.13	-0.57	-0.30
cc_2		0.48	0.33	0.40	0.32	0.50	0.33	0.41	0.27	-0.07	-0.55	-0.33	-0.43
cc_3		0.33	0.15	0.19	0.24	0.32	0.21	0.30	0.29	0.03	-0.45	-0.26	-0.35
cc_4		0.17	0.04	0.02	0.11	0.16	0.14	0.19	0.24	0.06	-0.36	-0.25	-0.27
cc_5		0.03	-0.10	-0.16	-0.02	0.05	0.02	0.03	0.13	0.17	-0.18	-0.20	-0.15
cc_6		-0.09	-0.24	-0.42	-0.27	0.04	-0.13	-0.21	-0.08	0.13	-0.06	-0.04	0.03
cc_7		-0.23	-0.16	-0.36	-0.31	-0.13	-0.17	-0.45	-0.27	-0.02	-0.03	-0.03	0.10
cc_8		-0.27	-0.02	-0.05	-0.12	-0.25	-0.19	-0.35	-0.32	-0.10	0.20	0.05	0.11
cc_9		-0.31	0.06	0.26	0.13	-0.31	-0.02	0.07	-0.03	-0.02	0.32	0.27	0.18
cc_{10}		-0.28	0.06	0.36	0.22	-0.27	0.00	0.30	0.17	0.16	0.38	0.24	0.20

Table 5. Correlations between the features generated from point clouds from different sensors and the field-observed forest structural attributes. The features are grouped by categories explained in Table 3. The darkness of grey indicates stronger correlation. N = number of trees per unit area, G = basal area, H_{max} = maximum tree height, and H_{std} = standard deviation in tree height. WV2 = WorldView-2 satellite imagery, TDX = Tandem-X interferometry, and TSX = TerraSAR-X radargrammetry

Feat.\Attr.	WV2				TSX				TDX				
	N	G	H_{std}	H_{max}	N	G	H_{std}	H_{max}	N	G	H_{std}	H_{max}	
Var.	h_{std}	-0.30	0.09	0.42	0.48	-0.35	0.42	0.60	0.67	-0.21	0.25	0.45	0.55
	h_{cv}	-0.19	-0.52	-0.17	-0.22	-0.22	-0.17	0.08	0.02	-0.10	-0.30	-0.14	-0.10
Forest height	h_{mean}	-0.20	0.79	0.78	0.88	-0.15	0.55	0.55	0.65	-0.20	0.70	0.80	0.84
	h_{mode}	-0.13	0.61	0.63	0.64	0.23	0.25	0.16	0.13	0.01	0.56	0.54	0.54
	hp_{10}	-0.07	0.75	0.61	0.66	0.00	0.39	0.31	0.42	-0.14	0.67	0.71	0.73
	hp_{20}	-0.14	0.77	0.70	0.75	-0.04	0.45	0.34	0.43	-0.19	0.69	0.75	0.79
	hp_{30}	-0.18	0.78	0.72	0.82	-0.10	0.50	0.43	0.52	-0.21	0.70	0.77	0.81
	hp_{40}	-0.21	0.77	0.76	0.86	-0.11	0.52	0.49	0.58	-0.20	0.71	0.78	0.83
	hp_{50}	-0.22	0.77	0.78	0.88	-0.12	0.57	0.57	0.66	-0.19	0.70	0.80	0.84
	hp_{60}	-0.23	0.77	0.79	0.90	-0.14	0.58	0.60	0.69	-0.20	0.70	0.81	0.85
	hp_{70}	-0.23	0.76	0.81	0.91	-0.18	0.57	0.59	0.71	-0.21	0.70	0.81	0.86
	hp_{80}	-0.24	0.75	0.81	0.92	-0.22	0.56	0.61	0.72	-0.21	0.69	0.81	0.86
	hp_{90}	-0.25	0.73	0.82	0.92	-0.25	0.53	0.60	0.70	-0.22	0.67	0.79	0.87
	h_{max}	-0.30	0.69	0.81	0.91	-0.23	0.55	0.60	0.71	-0.23	0.63	0.77	0.85
	pr	-0.15	-0.56	-0.27	-0.36	-0.06	-0.17	-0.08	-0.07	-0.13	-0.27	-0.04	-0.05
	Forest density	cc_1	-0.17	-0.55	-0.21	-0.29	-0.06	-0.22	-0.03	-0.16	-0.12	-0.20	-0.10
cc_2		-0.12	-0.56	-0.23	-0.35	-0.16	-0.17	0.03	0.00	0.03	-0.25	-0.16	-0.25
cc_3		-0.02	-0.51	-0.30	-0.39	0.04	-0.07	0.07	0.03	-0.01	-0.28	-0.09	-0.10
cc_4		-0.03	-0.49	-0.30	-0.36	-0.17	-0.33	-0.35	-0.30	-0.08	-0.44	-0.22	-0.28
cc_5		-0.06	-0.32	-0.22	-0.20	-0.02	-0.23	-0.30	-0.30	0.01	-0.26	-0.20	-0.14
cc_6		-0.14	-0.18	-0.12	-0.08	0.10	0.12	-0.02	-0.03	-0.02	-0.40	-0.43	-0.30
cc_7		-0.11	0.13	0.08	0.15	0.12	0.27	0.14	0.22	-0.05	-0.03	0.02	0.10
cc_8		0.07	0.36	0.18	0.18	0.30	0.21	0.12	0.10	-0.11	0.07	0.08	0.07
cc_9		0.20	0.35	0.19	0.21	-0.14	0.15	0.15	0.20	0.05	0.38	0.23	0.18
cc_{10}		0.15	0.21	0.13	0.14	-0.15	0.17	0.20	0.23	0.15	0.40	0.31	0.23

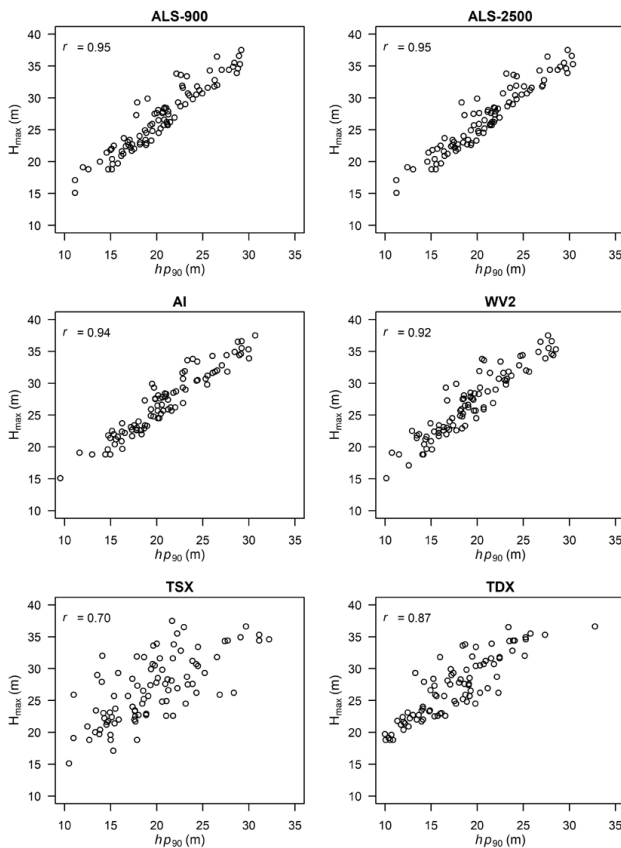


Figure 1. Relationship between the feature describing tree height (i.e. 90% percentile of normalized height of all points above 2-m threshold (hp_{90})) generated from point clouds from different sensors and field-measured maximum tree height (H_{max}). ALS-900 = airborne laser scanning from 900 m altitude, ALS-2500 = airborne laser scanning from 2500 m altitude, AI = aerial imagery, WV2 = WorldView-2 satellite imagery, TDX = Tandem-X interferometry, and TSX = TerraSAR-X radargrammetry

ALS-derived pr was the only feature that had a negative correlation coefficient greater than 0.85 with the field-measured G (Table 4). Features characterizing forest density derived from AI, WV2, TDX, and TSX provided only correlation coefficients of -0.59, -0.56, -0.44, and -0.33, respectively, with the G measured in the field (Tables 4 and 5). Also, ALS-derived cc_2 was the only forest density-related feature that had correlation coefficient close to -0.50 with the field-measured N (Table 4). Features characterizing forest density derived from AI, WV2, TDX, and TSX provided only correlation coefficients of 0.17, 0.20, 0.15, and 0.30, respectively, with the N measured in the field (Tables 4 and 5). In addition to H_{max} , features classified as representing forest height generated from the point clouds from AI, WV2, TSX, and TDX, also had the strongest correlations with the field-measured H_{std} , G and N when all possible features were compared. In

contrast, even considering all the features, ALS-based features provided the most strongly correlated features for each of the investigated forest structural attribute (Table 4).

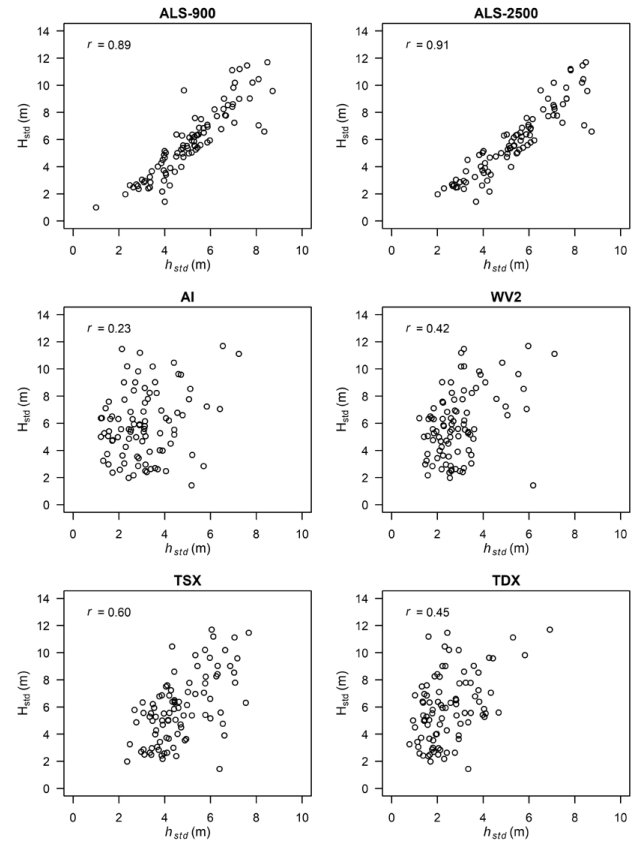


Figure 2. Relationship between the feature describing variation in tree height (i.e. standard deviation of normalized height of all points above 2-m threshold (h_{std})) generated from point clouds from different sensors and field-measured standard deviation of tree height (H_{std}). ALS-900 = airborne laser scanning from 900 m altitude, ALS-2500 = airborne laser scanning from 2500 m altitude, AI = aerial imagery, WV2 = WorldView-2 satellite imagery, TDX = Tandem-X interferometry, and TSX = TerraSAR-X radargrammetry

The forest height-related structural attributes (i.e. H_{max} and H_{std}) were predicted using one feature related to forest height or variation of tree height, derived from remotely sensed data sets (Table 6). H_{max} was the most accurately predicted structural attribute (Table 7). The differences in prediction accuracies between ALS, image- and radar-based point clouds were most obvious when H_{std} was predicted. With ALS data, H_{std} was modelled using the h_{std} -feature as the only explanatory variable as with other data sources (i.e., AI, WV2, TSX, TDX), forest height-related features were used for predicting H_{std} because of the lack of prediction power of the features related to variation in tree height derived from

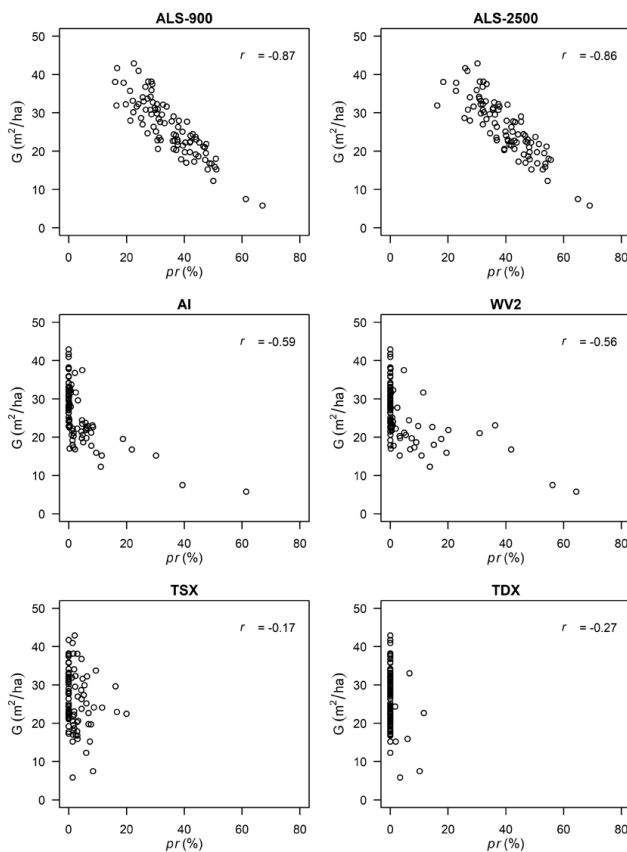


Figure 3. Relationship between feature describing forest density, (i.e. penetration ratio (*pr*)) generated from point clouds from different sensors and field-measured basal area (G). ALS-900 = airborne laser scanning from 900 m altitude, ALS-2500 = airborne laser scanning from 2500 m altitude, AI = aerial imagery, WV2 = WorldView-2 satellite imagery, TDX = Tandem-X interferometry, and TSX = TerraSAR-X radargrammetry

those point clouds. The relative RMSEs varied in the ranges of 4.1–13.7% for H_{max} , and 17.0–37.4% for H_{std} , depending on the remote sensing data used in the modelling.

Two or three selected features were used to predict G (Table 6). Most accurate G estimates with the lowest biases were obtained with ALS data followed by image- and radar-based point clouds. The relative RMSEs varied in the ranges of 12.9–23.4% for G depending on the remote sensing data used in the modelling. When ALS features were used for predicting G, the best performing model included three features; one describing forest height, one variation in tree height, and one forest density (Table 6). The image-based and the radar-based models for this attribute only used forest height and density features as explanatory variables. The TDX-based models predicted the forest structural attributes except N, more accurately than TSX counterparts, while neither of the SAR-based point clouds reached the performance of the ALS and image-based point clouds when considering estimation accuracy. With all of the developed models, N was predicted by using features related to forest height and density. The relative RMSEs varied in the ranges of 47.2–58.1%. The prediction models based on ALS data averaged the variation in forest structural attributes less than the models using other remote sensing data (Table 6). However, variation in estimates derived from the image-based models is close to the ALS-based models in predicted G, H_{max} , and H_{std} . ALS-based estimates for N include more variation compared to any other data sources. Variation in estimated attributes based on the TSX and TDX point clouds is generally lower than the estimates based on ALS and AI or WV2.

Table 6. Features selected as explanatory variables in the linear regression models, grouped by the feature category. Parameter estimates for model coefficients are provided with standard error estimates (in parenthesis). G = basal area, H_{max} = maximum tree height, H_{std} = standard deviation in tree height, and N = number of trees per unit area. ALS-900 = airborne laser scanning from 900 m altitude, ALS-2500 = airborne laser scanning from 2500 m altitude, AI = aerial imagery, WV2 = WorldView-2 satellite imagery, TDX = Tandem-X interferometry, and TSX = TerraSAR-X radargrammetry

		Model coefficients and parameter estimates grouped by feature category						
RS-data		(Intercept)	Variation in tree height		Forest height	Forest density		
$H_{max,m}$	ALS-900	1.77(±0.62)			+ 0.97(±0.02)	h_{max}		
	ALS-2500	2.51(±0.58)			+ 0.97(±0.02)	h_{max}		
	AI	6.55(±0.77)			+ 0.98(±0.04)	hp_{90}		
	WV2	6.64(±0.92)			+ 1.05(±0.05)	hp_{90}		
	TSX	12.69(±1.51)			+ 0.80(±0.08)	hp_{90}		
	TDX	12.16(±0.93)			+ 0.84(±0.05)	hp_{90}		
$H_{std,m}$	ALS-900	-1.65(±0.42)	+ 1.45(±0.08)	h_{std}				
	ALS-2500	-1.37(±0.36)	+ 1.34(±0.06)	h_{std}				
	AI	-2.60(±0.64)			+ 0.41(±0.03)	hp_{90}		
	WV2	-2.73(±0.65)			+ 0.45(±0.03)	hp_{90}		
	TSX	2.06(±0.75)			+ 0.31(±0.06)	hp_{40}		
	TDX	1.38(±0.51)			+ 0.40(±0.04)	hp_{10}		
G, m ² /ha	ALS-900	90.71(±8.64)	- 19.76(±5.71)	h_{cv}	+ 0.47(±0.09)	h_{max}	- 19.76(±1.77)	$\log(pr)$
	ALS-2500	97.74(±10.57)	- 18.08(±5.79)	h_{cv}	+ 0.51(±0.09)	h_{max}	- 21.28(±2.22)	$\log(pr)$
	AI	9.41(±2.31)			+ 1.00(±0.12)	h_{mean}	- 0.21(±0.07)	pr
	WV2	6.64(±2.51)			+ 1.26(±0.14)	h_{mean}	- 0.39(±0.19)	cc_2
	TSX	14.13(±3.07)			+ 0.90(±0.15)	hp_{70}	- 0.27(±0.12)	cc_4
	TDX	12.64(±2.09)			+ 1.03(±0.13)	hp_{40}	- 0.23(±0.11)	cc_4
N, 1/ha	ALS-900	1685.50(±300.19)			- 83.82(±19.58)	hp_{30}	+ 33.83(±15.10)	cc_2
	ALS-2500	1778.78(±288.03)			- 84.63(±17.66)	hp_{30}	+ 33.65(±15.47)	cc_2
	AI	1680.02(±261.34)			- 29.7(±10.25)	h_{max}	- 16.36(±6.49)	pr
	WV2	1494.53(±250.38)			- 33.27(±10.41)	h_{max}	+ 9.90(±4.50)	cc_9
	TSX	1197.81(±225.46)			- 32.53(±10.71)	hp_{90}	+ 27.32(±7.96)	cc_8
	TDX	1228.06(±185.93)			- 29.43(±10.59)	hp_{90}	+ 11.98(±5.25)	cc_{10}

Table 7. Accuracy assessment of the predicted forest structural attributes using various 3D remote sensing data sets. H_{max} = maximum tree height, H_{std} = standard deviation in tree height, G = basal area, N = number of trees per unit area. ALS-900 = airborne laser scanning from 900 m altitude, ALS-2500 = airborne laser scanning from 2500 m altitude, AI = aerial imagery, WV2 = WorldView-2 satellite imagery, TDX = Tandem-X interferometry, and TSX = TerraSAR-X radargrammetry

H_{max} (m)	RMSE	RMSE %	Bias	Bias %	r	Sdev est.	Sdev.
ALS-900	1.15	4.29	0.00	0.00	0.97	5.03	5.20
ALS-2500	1.11	4.13	0.00	0.00	0.98	5.04	
AI	1.68	6.28	0.00	0.00	0.94	4.88	
WV2	1.98	7.41	0.00	0.00	0.92	4.76	
TSX	3.66	13.65	0.01	0.02	0.70	3.70	
TDX	2.65	9.90	0.02	0.07	0.86	4.55	
H_{std} (m)	RMSE	RMSE %	Bias	Bias %	r	Sdev est.	Sdev.
ALS-900	1.11	18.88	0.00	0.00	0.89	2.21	2.50
ALS-2500	1.00	16.95	0.00	0.00	0.91	2.27	
AI	1.40	23.79	0.00	0.00	0.82	2.04	
WV2	1.41	24.00	0.00	0.00	0.82	2.03	
TSX	2.19	37.35	-0.01	-0.12	0.46	1.21	
TDX	1.77	30.19	0.00	-0.08	0.69	1.75	
G (m ² /ha)	RMSE	RMSE %	Bias	Bias %	r	Sdev est.	Sdev.
ALS-900	3.37	12.88	0.02	0.06	0.89	6.68	7.40
ALS-2500	3.71	14.17	0.05	0.20	0.87	6.61	
AI	4.68	17.87	-0.05	-0.19	0.78	6.21	
WV2	4.59	17.55	-0.01	-0.02	0.78	5.96	
TSX	6.11	23.36	0.02	0.07	0.56	4.50	
TDX	5.24	20.02	0.01	0.06	0.70	5.42	
N (1/ha)	RMSE	RMSE %	Bias	Bias %	r	Sdev est.	Sdev.
ALS-900	441.54	48.84	1.67	0.18	0.56	327.61	530.35
ALS-2500	426.29	47.15	0.22	0.02	0.60	341.82	
AI	513.55	56.81	-4.20	-0.47	0.27	193.72	
WV2	509.92	56.40	-0.14	-0.02	0.29	201.27	
TSX	499.65	55.27	-0.96	-0.11	0.34	220.81	
TDX	525.08	58.08	0.68	0.08	0.20	176.46	

Discussion and Conclusions

The prediction of forest structural attributes was undertaken using simple statistical methods, a limited number of predictors, and a sample plot size of 1024 m² used for model development and validation. The correlation analyses strengthened the knowledge that ALS outperforms image- and radar-based techniques in deriving a point cloud to describe the variation in tree height as well as forest density (e.g. White et al. 2013b, Vastaranta et al. 2013, Rahlf et al. 2014). Furthermore, based on our analyses most of the predictive power of image- and radar-based point clouds is coming from features describing height. When compared to the other methods, ALS-based point clouds are more evenly distributed in the canopy both horizontally and vertically.

The correlations were similar among the two ALS data sets (i.e. ALS-900 and ALS-2500) acquired with the same sensor although the measurement density varied from 12 points/m² (ALS-900) to 1.2 points/m² (ALS-2500) and flying altitude from 900 m to 2500 m. Moreover, correlations between features derived from point clouds based on AI and WV2 imagery and forest structural attributes were rather similar despite the different data acquisition parameters. Although the realized mean point density was similar in TSX and TDX data, due to different measurement principles, bistatic interferometry tends to provide point clouds that concentrate close to the envelope of the canopy as the realized radargrammetric points seem to fluctuate more within canopy. The reason for this fluctuation originates from the radar speckle, which makes the stereo tie-point matching challenging

in forested area, and the precision of elevation measurement is somewhat inferior than in case of TDX bistatic interferometric point clouds. Thus, the correlations with forest structural attributes calculated using TSX features were generally weaker than the ones calculated for features from TDX and other sensors.

The accuracy of the predicted forest structural attributes based on both ALS-900 and ALS-2500 data sets were similar regardless of the different measurement densities. In the same way, the accuracy of estimates based on both AI and WV2 imagery did not remarkably differ from each other. However, ALS outperformed other remote sensing materials consistently, but differences between the ALS and imagery-derived point clouds were rather minor in our study area. Similar findings have also been reported by Nurminen et al. (2013) and Vastaranta et al. (2013) for ALS and aerial stereo imagery, and Straub et al. (2013) for ALS and satellite stereo imagery. Compared to Yu et al. (2015), in our study ALS provided consistently the most accurate estimates followed by image-based and radar-based prediction models. In this study, difference between the ALS and image-based data sets was most notable when predicting tree height variation (i.e. standard deviation of tree height).

The point-clouds based on SAR could not describe the forest structure at plot level as accurately as ALS or optical imagery, consequently models for forest structural attributes based on either TSX or TDX produced less accurate estimates for forest structural attributes. This is in line with the findings by Karjalainen et al. (2012), Karila et al. (2015), and Yu et al. (2015). In addition, the method to derive a point cloud from radar

data affected the prediction accuracy: the estimates for forest attributes based on TDX were generally more accurate than the estimates based on TSX data, which was expected based on the correlation analysis. However, it should be kept in mind that the measurement densities in TDX and TSX data sets were only ~ 0.05 points/m². Due to the lower point density and different measurement principles, it cannot be even expected that TDX or TSX point clouds could capture the variation in tree height and forest density as detailed as ALS. However, the point cloud features based on satellite radar data, especially on TDX, correlate with the forest height and can be used consequently in the prediction of forest structural attributes even at plot-level at reasonable accuracy.

Based on our analyses, when forest structural attributes are predicted using the features derived from image- or radar-based point clouds, most of the prediction power is coming from the features describing forest height. In addition to features related to forest height, ALS is capable of capturing variation in tree height as well as forest density which is important in complex forest environments. The capabilities to characterize forest height and density as well as variation in tree height are limited with image-based and radar-based point clouds although the attributes that are associated with forest height can be predicted rather accurately for the purposes of the forest planning and wood procurement in stands that are mainly even-aged and single layer. However, users of the image-based or radar-based point clouds should be aware, that the estimates of forest inventory attributes that are not directly related to forest height (e.g. G or N) may have limited accuracy in the complex forest environments.

In this study, the capabilities of point clouds derived from ALS, AI, WV2, TSX, and TDX data to characterize variation in forest structure were investigated. Features derived from ALS, AI, WV2, and TDX correlated strongly ($r > 0.85$) with maximum tree height whereas the correlations calculated from TSX features remained close to 0.7 after the point clouds were normalized with detailed ALS-based DTM. ALS was the only data source capable of providing separate features for characterizing also the variation in tree height and forest density. With AI, WV2, TSX, and TDX, features describing forest height were also the most strongly correlated with the variation in tree height and forest density. This finding indicates that forest inventory attributes that are correlated with height can be predicted rather well also with AI, WV2, and TDX, but if it is important to characterize also the forest density and tree height variation in area of interest, ALS data should be used. When considering forest inventory using any of the investigated data sources first time, ALS is required, because all the

alternative methods require detailed DTM for height normalization and ALS is the only option for mapping ground elevation and forest height simultaneously.

Acknowledgements

This work was supported by the Academy of Finland through Grants 293389 and 272195 as well as the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement 606971. The authors are grateful to Lic.Sc. Risto Viitala and Häme University of Applied Sciences for all the help with practical arrangements at Evo.

References

- Bouvier, M., Durrieu, S., Fournier, R.A. and Renaud, J-P. 2015. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sensing of Environment* 156: 322–334. doi: <https://doi.org/10.1016/j.rse.2014.10.004>.
- Holopainen, M., Vastaranta, M., Rasinmäki, J., Kalliovirta, J., Mäkinen, A., Haapanen, R., Melkas, T., Yu, X. and Hyypä, J. 2010. Uncertainty in timber assortment estimates predicted from forest inventory data. *European Journal of Forest Research* 129: 1131–1142. doi: <https://doi.org/10.1007/s10342-010-0401-4>.
- Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E. and Falkowski, M.J. 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment* 112: 2232–2245. doi: <https://doi.org/10.1016/j.rse.2007.10.009>.
- Hyypä, J., Hyypä, H., Leckie, D., Gougeon, F., Yu X. and Maltamo, M. 2008. Review of methods of small footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing* 29: 1339–1366. doi: <https://doi.org/10.1080/01431160701736489>.
- Karila, K., Vastaranta, M., Karjalainen, M. and Kaasalainen, S. 2015. Tandem-X interferometry in the prediction of forest inventory attributes in managed boreal forests. *Remote Sensing of Environment* 159: 259–268. doi: <https://doi.org/10.1016/j.rse.2014.12.012>.
- Karjalainen, M., Kankare, V., Vastaranta, M., Holopainen, M. and Hyypä, J. 2012. Prediction of plot-level forest variables using TerraSAR-X stereo SAR data. *Remote Sensing of Environment* 117: 338–347. doi: <https://doi.org/10.1016/j.rse.2011.10.008>.
- Luoma, V., Saarinen, N., Wulder, M.A., White, J.C., Vastaranta, M., Holopainen, M. and Hyypä, J. 2017. Assessing Precision in Conventional Field Measurements of Individual Tree Attributes. *Forests* 8: 38. doi: <https://doi.org/10.3390/f8020038>.
- Nurminen, K., Karjalainen, M., Yu, X., Hyypä, J. and Honkavaara, E. 2013. Performance of dense digital surface models based on image matching in the estimation of plot-level forest variables. *ISPRS Journal of Photogrammetry and Remote Sensing* 83: 104–115. doi: <https://doi.org/10.1016/j.isprsjprs.2013.06.005>.
- Næsset, E., Gobakken, T., Holmgren, J., Hyypä, H., Hyypä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, Å. and Söderman, U. 2004. Laser scanning of forest resources: the Nordic experience. *Scandinavi-*

- an Journal of Forest Research* 19: 482–499. doi: <https://doi.org/10.1080/02827580410019553>.
- Poso, S., Häme, T. and Paananen, R.** 1984. A method for estimating the stand characteristics of a forest compartment using satellite imagery. *Silva Fennica* 18(3): 261–292. doi: <https://doi.org/10.14214/sf.a15398>.
- R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: <https://www.r-project.org/> [Cited 3 Jan 2018].
- Rahlf, J., Breidenbach, J., Solberg, S., Næsset, E. and Astrup, R.** 2014. Comparison of four types of 3D data for timber volume estimation. *Remote Sensing of Environment* 155: 325–333. doi: <https://doi.org/10.1016/j.rse.2014.08.036>.
- Solberg, S., Astrup, R., Breidenbac, J., Nilsen, B. and Weydahl, D.** 2013. Monitoring spruce volume and biomass with InSAR data from TanDEM-X. *Remote Sensing of Environment* 139: 60–67. doi: <https://doi.org/10.1016/j.rse.2013.07.036>.
- Straub, C., Tian, J., Seitz, R. and Reinartz, P.** 2013. Assessment of Cartosat-1 and WorldView-2 stereo imagery in combination with a LiDAR-DTM for timber volume estimation in a highly structured forest in Germany. *Forestry* 86: 463–473. doi: <https://doi.org/10.1093/forestry/cpt017>.
- Vastaranta, M., Wulder, M.A., White, J.C., Pekkarinen, A., Tuominen, S., Ginzler, C., Kankare, V., Holopainen, M., Hyyppä, J. and Hyyppä, H.** 2013. Airborne laser scanning and digital stereo imagery measures of forest structure: comparative results and implications to forest mapping and inventory update. *Canadian Journal of Remote Sensing* 39: 382–395. doi: <https://doi.org/10.5589/m13-046>.
- White, J.C., Wulder, M.A., Varhola, A., Vastaranta, M., Coops, N.C., Cook, B.D., Pitt, D. and Woods, M.** 2013a. A best practice guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. *The Forestry Chronicle* 89: 722–723. doi: <https://doi.org/10.5558/tfc2013-132>.
- White, J.C., Wulder, M.A., Vastaranta, M., Coops, N.C., Pitt, D. and Woods, M.** 2013b. The utility of image-based point clouds for forest inventory: A comparison with airborne laser scanning. *Forests* 4: 518–536. doi: <https://doi.org/10.3390/f4030518>.
- White J.C., Coops N.C., Wulder M.A., Vastaranta M., Hilker T. and Tompalski P.** 2016. Remote sensing technologies for enhancing forest inventories: A review. *Canadian Journal of Remote Sensing* 42: 619–641. <https://doi.org/10.1080/07038992.2016.1207484>.
- Wulder M.A., White J.C., Nelson R.F., Næsset E., Ørka H.O., Coops N.C., Hilker T., Bater C.W. and Gobakken, T.** 2012. Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment* 121: 196–209. doi: <https://doi.org/10.1016/j.rse.2012.02.001>.
- Wulder, M.A., Coops, N.C., Hudak, A.T., Morsdorf, F., Nelson, R., Newnham, G. and Vastaranta, M.** 2013. Status and prospects for LiDAR remote sensing of forested ecosystems. *Canadian Journal of Remote Sensing* 39: S1–S5. doi: <https://doi.org/10.5589/m13-051>.
- Yu, X., Hyyppä, J., Karjalainen, M., Nurminen, K., Karila, K., Vastaranta, M., Kankare, V., Kaartinen, H., Holopainen, M. and Honkavaara, E.** 2015. Comparison of laser and stereo optical, SAR and InSAR point clouds from air- and space-borne sources in the retrieval of forest inventory attributes. *Remote Sensing* 7: 15933–15954. doi: <https://doi.org/10.3390/rs71215809>.