

Article

Effects of Pulse Density on Digital Terrain Models and Canopy Metrics Using Airborne Laser Scanning in a Tropical Rainforest

Endre Hofstad Hansen *, Terje Gobakken and Erik Næsset

Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ås, Norway; E-Mails: terje.gobakken@nmbu.no (T.G.); erik.naesset@nmbu.no (E.N.)

* Author to whom correspondence should be addressed; E-Mail: endre.hansen@nmbu.no; Tel.: +47-649-657-56.

Academic Editors: Lars T. Waser and Prasad S. Thenkabail

Received: 28 March 2015 / Accepted: 25 June 2015 / Published: 30 June 2015

Abstract: Airborne laser scanning (ALS) is increasingly being used to enhance the accuracy of biomass estimates in tropical forests. Although the technological development of ALS instruments has resulted in ever-greater pulse densities, studies in boreal and sub-boreal forests have shown consistent results even at relatively small pulse densities. The objective of the present study was to assess the effects of reduced pulse density on (1) the digital terrain model (DTM), and (2) canopy metrics derived from ALS data collected in a tropical rainforest in Tanzania. We used a total of 612 coordinates measured with a differential dual frequency Global Navigation Satellite System receiver to analyze the effects on DTMs at pulse densities of 8, 4, 2, 1, 0.5, and 0.025 pulses·m⁻². Furthermore, canopy metrics derived for each pulse density and from four different field plot sizes (0.07, 0.14, 0.21, and 0.28 ha) were analyzed. Random variation in DTMs and canopy metrics increased with reduced pulse density. Similarly, increased plot size reduced variation in canopy metrics. A reliability ratio, quantifying replication effects in the canopy metrics, indicated that most of the common metrics assessed were reliable at pulse densities >0.5 pulses·m⁻² at a plot size of 0.07 ha.

Keywords: ALS; airborne laser scanning; digital terrain model; DTM; LiDAR; reliability ratio; tropical rainforest

1. Introduction

Tropical forests store large amounts of carbon as biomass and regrowth of tropical forests sequester 1.85 ± 0.09 Pg of carbon annually [1]. However, the annual loss of forest carbon due to deforestation and forest degradation in tropical areas is estimated to 2.01 ± 1.1 Pg [1]. In response to the important role of forests both as a source of carbon emissions and as a potential sink source, the UNFCCC negotiations have resulted in a policy and economic incentive mechanism for creating economic incentives for reducing loss of forest biomass and increasing the net sink of carbon in forests [2]. A key component in the proposed mechanism is monitoring, which ensures reliable measurement, reporting, and verification of the reduced loss, or net increase of forest biomass. To provide precise estimates of forest biomass the use of airborne laser scanning (ALS) has become increasingly popular, as several studies have demonstrated its good performance in tropical areas [3–10].

Modern ALS instruments are able to emit pulses at a rate of up to around 800 kHz and are usually flown over the area of interest mounted on a small airplane. From each pulse the instrument is normally set to record up to 5–9 echoes per pulse, herein referred to as points. Acquisition of ALS data is costly compared to, for instance, optical satellite images or radar data. This cost is largely governed by the flight time, which can be reduced by flying higher and/or faster, resulting in cheaper, but fewer pulses per unit area.

Recent studies of biomass in tropical forests using ALS have been conducted using different pulse densities and plot sizes. Pulse densities from 25 pulses·m⁻² [11] down to about 1.5 pulses·m⁻² [3] have been used. The results from these studies are similar in terms of biomass prediction performance and show that great pulse density is not a requisite for estimation of forest biomass.

To study how sparse pulse densities (*i.e.* <1 pulse·m⁻²) affect the quality of the biomass estimates, controlled reduction of the ALS data has been used to assess the effect on model prediction accuracy [12–23]. Findings from such controlled experiments have resulted in reduced pulse density and lowered cost for ALS acquisitions for operational forest inventory purposes, making the technology more widely applied while providing biomass estimates at acceptable precision levels. Lowered costs and well-documented precision levels have resulted in sparse pulse density ALS missions covering large areas and even entire countries [24], providing valuable data for decision-making in forestry. Most studies of pulse density have been conducted in boreal coniferous forests [12–15,17,18] with examples also from temperate mixed-conifer forests [19,21] and subtropical pine plantation [16]. The main conclusions from the studies in coniferous forests have been that reducing pulse density down to, say, 0.1 pulses·m⁻² has hardly any effect on the accuracy and precision of timber volume prediction.

To our knowledge, the recent study by Leitold *et al.* [22] is the only study assessing effects of pulse density in tropical broadleaved forests. The study conducted in the Brazilian Atlantic forest reported a systematic effect of pulse density in the construction of the digital terrain model (DTM). Leitold *et al.* [22] concluded that this systematic error in the DTM was propagated to the canopy metrics, resulting in underestimation of forest biomass with reduced pulse density.

A prerequisite for successful application of ALS data for biomass prediction is a good quality terrain model, as information about vegetation height is derived relative to the predicted terrain surface. The quality of the terrain model is determined by the number of pulses that successfully reach

the ground. Dense vegetation obstructs the ALS pulses and results in fewer pulses reaching the ground and being available for DTM construction. The effect of vegetation on ALS-derived DTMs has been studied in different conditions and has resulted in both an over-prediction of terrain elevation [25–27] and under-prediction of terrain elevation [28,29]. Hodgson *et al.* [28] found that ALS-derived elevation was significantly under-predicted in all studied land cover classes. The under-prediction was largest in pine forest areas, by up to 0.24 m. Tinkham *et al.* [29] also found an under-prediction of 0.9–0.16 m in coniferous areas, when comparing two different ground classification algorithms. In their discussion of observed under-predicted terrain elevation in heavily vegetated areas, Hodgson *et al.* [28] suggested that the error was a result of point density, and/or the accuracy of correct classification of points as ground.

In addition to the number of pulses that reach the ground, the quality of the terrain model is dependent on a correct classification of these points as ground points. There are several algorithms for classifying individual points as either ground or vegetation points and they differ in performance under different terrain and vegetation conditions [29,30]. Algorithm parameter settings also affect the proportion of points classified as ground. Comparison of five vendors applying the same classification algorithm on the same ALS data showed a difference in the proportion of ground points of 17 percentage points [31]. This difference is most likely due to the parameter settings in the applied algorithm. Part of the result, however, may also be attributed to differences in data analysis routines not documented in the ALS data processing reports [31]. With greater biomass densities and dense broadleaved canopies, the proportion of points classified as ground is expected to be small in tropical rainforests. Furthermore, reduced ground point density, due to dense canopy, has been shown to result in increased error in the constructed DTM [32,33].

When applying ALS data for biomass estimation following the commonly used area-based approach [34], a relationship between biomass measured on field plots on the ground and canopy metrics derived from ALS data from the corresponding area is modeled using statistical methods such as regression analysis, nearest neighbor classification, neural networks or ensemble learning [35–37]. It is possible to derive a large number of canopy metrics from the ALS data to be used in the modeling. However, since these metrics are highly correlated, only a few are usually selected in the final models. Common metrics for biomass estimation in tropical forests have been found to be the mean above ground elevation of the ALS points [11,38,39] and maximum above ground elevation of the ALS points [3,40]. Other metrics used include percentiles and variance of the above ground elevation of the ALS points [7,11]. In the present study, we therefore chose to assess canopy metrics similar to those used in the aforementioned studies.

In addition to pulse density, performance of ALS based biomass estimation is also affected by the field plot size [16,41,42]. Combined effects of pulse density and plot size were assessed by Watt *et al.* [16] who concluded that reduced pulse density and plot size had little effect on model fit for pulse densities >0.1 pulses \cdot m⁻² and plot sizes of >0.03 ha. As described in Vauhkonen *et al.* [43], the central issue is that the combination of plot area and pulse density includes enough points to reliably estimate the metrics used to predict forest biomass. Gobakken and Næsset [12] documented that canopy metrics have less variation on larger plots and that increasing plot size could compensate for sparse pulse density. The size of field plots applied in studies of biomass in tropical forests utilizing ALS data usually range between 0.1 and 1.0 ha [44]. To assess the effect of plot size on the mean values and

variation of ALS-derived canopy metrics we computed and analyzed metrics derived from plot sizes of 0.07, 0.14, 0.21, and 0.28 ha. The smallest plot size of 0.07 ha was chosen because it corresponds to the plot size used in the Tanzanian national forest inventory [45].

The objectives of the present study were to assess the effects of reduced ALS pulse densities on (1) the DTMs and (2) the canopy metrics derived from the ALS data used for biomass estimation in a dense tropical rainforest in Tanzania. We reduced the pulse density from an initial density of about 13 pulses·m⁻², to 8, 4, 2, 1, 0.5, and 0.25 pulses·m⁻². Following the pulse density reduction, we assessed the differences among the DTMs derived from ALS data and the elevation obtained from 612 ground coordinates measured using a survey-grade differential Global Navigation Satellite System (dGNSS) receiver. Furthermore, we produced standard canopy metrics commonly used for forest biomass prediction and compared the effect of reduced pulse densities on selected canopy metrics for four different field plot sizes (0.07, 0.14, 0.21, and 0.28 ha).

2. Materials and Methods

2.1. Study Area

The study area is located in eastern Tanzania in the Amani Nature Reserve (S 5°08', E 38°37', 200–1200 m above sea level) and covers around 8360 ha of tropical submontane rainforest. Most of the annual precipitation of around 2000 mm is received in two main rain seasons. Daily mean temperatures vary from about 16 to 25 °C. In an inventory carried out in 1986/87, about half of the nature reserve was classified as logged or covered with an non-indigenous tree species *Maesopsis eminii* as a result of logging [46]. The logging was stopped in the late 1980s and most of the nature reserve is now covered by closed forest. Situated in a mountainous area, most of the forest is located in sloped terrain. To characterize the terrain slope we constructed a DTM from the full ALS data. Terrain slope, averaged over a 1 m grid in a 100 m buffer area of each field plot area, ranged from 8.6° to 37.1° with a mean value of 21.4° and a standard deviation of 5.5°.

2.2. Field Data

For comparing the DTMs derived from ALS data to elevation obtained from ground coordinates measured using dGNSS, a total of 612 coordinates (x, y, and z positions) were collected in the corners of 153 rectangular field plots georeferenced in the period August 2011–April 2012. The distance between the plot corners was approximately 20 m in the north-south direction and 50 m in the east-west direction. The field plots were distributed on a regular grid of 900 m in the north-south direction and 450 m in the east-west direction. The coordinates were measured by means of dGNSS using two 40-channel survey-grade receivers (Topcon Legacy-E+). One receiver was placed at each plot corner (k) on a 2.9 m carbon rod for a minimum of 30 minutes, and a one second logging rate was used. A second receiver, acting as a base station, was placed on the roof of a house at the Amani Nature Reserve headquarters with a distance of <14 km from the plots. Prior to measuring the coordinates, the position of the base station antenna was determined with Precise Point Positioning with Global Positioning System and Global Navigation Satellite System data collected continuously for 24 hours according to Kouba [47]. Due to the sloped terrain and dense biomass conditions the

elevation measured at each coordinate with the dGNSS (z^{dGNSS}) had a mean precision (standard deviation) of 0.39 m reported from the post-processing using Pinnacle software [48]. Further details about the field data can be found in Hansen *et al.* [49].

2.3. ALS Data

Complete coverage ALS data were collected in January and February 2012 using a Leica ALS70 (Leica Geosystems AG, Switzerland) sensor mounted on a fixed wing aircraft. Average altitude and speed were 800 m above ground level and $70 \text{ m}\cdot\text{s}^{-1}$, respectively. The beam divergence was 0.28 mrad, which produced an average footprint size on the ground of about 22 cm. The sensor was operated using a laser pulse repetition frequency of 339 kHz and up to five echoes were recorded for each emitted pulse. Planned pulse density at acquisition was set to $10 \text{ pulses}\cdot\text{m}^{-2}$, but due to overlap between adjacent strips the average pulse density in the area of the field plots was $13 \text{ pulses}\cdot\text{m}^{-2}$.

2.4. Reduction of Laser Pulse Density

To study the effects of reduced pulse densities, a procedure was employed by which individual pulses were randomly discarded. Based on initial testing of the necessary number of repetitions, the random thinning procedure was repeated 50 times and the calculated statistics and metrics formed the basis for quantifying the effects of data reduction.

Table 1. Summary of the difference in elevation between the dGNSS measurements and the DTM elevation for different pulse densities.

Pulses·m ⁻²	\bar{D} (m)	S_D (m)	P50 _D (m)	P95 _D (m)	NMAD (m)	Parameter Settings ¹	
						g (m)	w (m)
0.25	1.77	3.20	0.90	7.72	2.15	-1.0	1.5
0.5	1.77	3.02	0.92	7.50	1.97	-1.5	2.0
1	1.79	2.93	0.94	7.34	1.85	-2.0	2.5
2	1.81	2.90	0.96	7.28	1.80	-2.5	3.0
4	1.81	2.88	0.95	7.20	1.75	-3.0	3.5
8	1.81	2.89	0.95	7.35	1.81	-3.5	4.0

Mean difference (\bar{D}), standard deviation (S_D), 50% quantile of the difference (P50_D), 95% quantile of the absolute value of the difference (P95_{|D|}) and the normalized median absolute deviation (NMAD).

¹ Settings of the g and w parameters in the applied ground classification algorithm.

The reduction of ALS data was executed on pulse level using the “ThinData” program in the FUSION toolkit [50]. ALS data were reduced from an initial density of about $13 \text{ pulses}\cdot\text{m}^{-2}$ to pulse densities of 8, 4, 2, 1, 0.5, and 0.25 pulses·m⁻². To mimic and maintain the fairly regular spatial distribution of ALS pulses inherent in the initial data, the reduction was performed on a grid size of 0.1, 0.2, 0.5, 1, 2, 10, and 20 m, for pulse densities of 8, 4, 2, 1, 0.5, and 0.25·m⁻², respectively. Following the reduction, a classification of ground points was conducted using the “GroundFilter” program in FUSION. The interpolating algorithm [51] implemented in “GroundFilter” initially makes an average surface based on all ALS points. Further, weights are given to all points based on their vertical distance to the initial surface. Low weight is given to points above the surface, and high weight

to points below. The weights are then used in re-fitting the surface. Two parameters in the algorithm can be adjusted to determine which points are given weights. Points located below the surface with a distance larger than parameter g are assigned the maximum weight value of 1.0, while points located above the surface with a distance larger than the parameter $w +$ the parameter g are assigned weights of 0.0 [50]. To adjust for the different pulse densities we controlled the two parameters while leaving the other parameters at the program default setting. The g and w parameter settings at different pulse densities are given in Table 1. Visual inspection of initial classifications of ground points showed large outliers and a smoothing filter of 3 m was applied to remove these outliers. From the points classified as ground, a 1 m gridded surface was created using the “TINSurfaceCreate” program in FUSION.

2.5. Assessing Effects of Pulse Density on DTM Quality

Reduction of the pulse density will affect the quality of the DTM since fewer ground points are available for constructing the DTM surface. To study the effects of reduced pulse densities on the DTMs we subtracted the elevation in the DTMs (z^{DTM}) of each plot corner (k), resulting from each pulse density (d), from the elevation of each plot corner measured with dGNSS (z^{dGNSS}) to get the difference for each coordinate (D_d^k , Equation (1)):

$$D_d^k = z_d^{\text{DTM}_k} - z^{\text{dGNSS}_k}. \quad (1)$$

The mean difference (\bar{D}) and standard deviation (S_D) of the differences (D_d^k) were calculated. To compare the \bar{D} at each pulse density level a t-test was performed using the Holm–Bonferroni procedure [52] for correction of p-values for multiple comparisons.

The conventional measures of accuracy, \bar{D} and S_D , assume no outliers and a normal distribution of errors. As pointed out by e.g. Zandbergen [53], errors in DTMs are often not normally distributed. We therefore checked for non-normality by inspecting a quantile-quantile (Q-Q) graph and calculated robust accuracy measures suited for characterization of non-normal distributions suggested by Höhle and Höhle [54]. The 50% sample quantile of the errors ($P50_D$, i.e., the median value) is a robust estimator for a systematic shift of the DTM [54]. The 95% quantile of the absolute value of the errors ($P95_{|D|}$) and the normalized median absolute deviation (NMAD, Equation (2)), a robust estimator for S_D , are estimators resilient to outliers [54].

$$\text{NMAD} = 1.4826 * \text{median}_i(|D_d^k - P50_D|). \quad (2)$$

2.6. Assessing Effects of Pulse Density on Canopy Metrics

After creating a gridded DTM, the elevation of the DTM was subtracted from the elevation for all ALS points resulting in an elevation above the ground for each individual ALS point. Together these points form a “cloud” of points, referred to as the point cloud. At the center of each of the 153 rectangular field plots we extracted the point clouds from four concentric circles of 0.07, 0.14, 0.21, and 0.28 ha. For each plot center, and for each plot size, a set of canopy metrics was computed using the “CloudMetrics” program in FUSION. Frequently used canopy metrics for biomass estimation in tropical forests include mean above ground elevation of the ALS points (E.mean) and maximum above ground elevation of the ALS points (E.max). Other commonly used metrics also include variance of the above ground elevation of the ALS points (E.var), percentiles of the above

ground elevation of the ALS points and canopy density. The canopy density metrics are commonly derived by dividing the canopy into 10 vertical parts of equal height and calculating the proportion of points above each vertical part. We selected the 10th and 90th percentile of elevation (E.10, E.90), the proportion of points above the ground (D.0) and above the mean above ground elevation of the ALS points (D.5), along with E.mean, E.max, and E.var for analysis.

From the repeated reduction of ALS pulse density described in Section 2.4, we calculated the mean (\bar{M}) and standard deviation of each canopy metric (S_M) on plot level across the 50 repetitions. Even though the canopy metrics are relatively unaffected by point density [55], reduced pulse density will increase S_M . As explained by Magnussen *et al.* [20], random factors affecting the canopy metrics suggest that the metrics should be considered as random variables instead of fixed, as is commonly the case. These random factors can be referred to as replication effects. Replication effects will weaken the fit of the biomass models with a factor termed as the reliability ratio [56]. By calculating the replication variance in the metrics, estimates of the reliability ratios for the metrics were calculated. The method was used by Magnussen *et al.* [20], in which the reliability ratio was calculated as the ratio of the variance of each metric among sample plots, to the total variance of the corresponding metric (Equation (3)):

$$\text{Reliability ratio} = (\hat{\sigma}_u^2)/(\hat{\sigma}_u^2 + \hat{\sigma}_w^2), \quad (3)$$

where $\hat{\sigma}_u^2$ is the estimated among-plot variance of the metric and $\hat{\sigma}_w^2$ is the estimated average within-plot variance. More within-plot variance in a metric compared to the variance among plots for the same metric results in a small reliability ratio, indicating that the metric is less reliable as a predictor.

3. Results

3.1. Effects of Pulse Density on DTM Quality

Effects of reduced pulse density resulted in a \bar{D} between the elevation of the 612 coordinates measurements recorded by the dGNSS and the elevation of the same coordinates in the ALS-derived DTM of 1.81 m for a pulse density of 8 pulses·m⁻² (Table 1). Thus, the elevation recorded by the dGNSS was higher than the ALS-derived DTM. Reduction of pulse density from 8 to 0.25 pulses·m⁻² gave no significant effect on the \bar{D} .

The Q–Q graph of the distribution of errors (Figure 1) showed non-normality and justified the presentation of robust measures of accuracy. Precision of DTM did not vary at pulse densities of >1 pulse·m⁻² (Table 1). Further reduction resulted in loss of precision, expressed by both the conventional measure of precision (S_D) and the more robust measure NMAD (Table 1).

3.2. Effects of Pulse Density on Canopy Metrics

Canopy metrics were derived from the ALS data from field plot sizes of 0.07, 0.14, 0.21, and 0.28 ha. Mean values from the repeated simulations showed that most of the assessed metrics were unaffected by the reduced pulse density (Table 2). E.max, however, decreased with reduced pulse density, and showed a significant difference ($p < 0.01$) of 0.58 m at 1 pulse·m⁻² compared to the value at 8 pulses·m⁻², at a plot size of 0.07 ha. This effect was reduced with increased plot size, but was still significant ($p < 0.01$) for 1 pulse·m⁻² at the plot size of 0.28 ha.

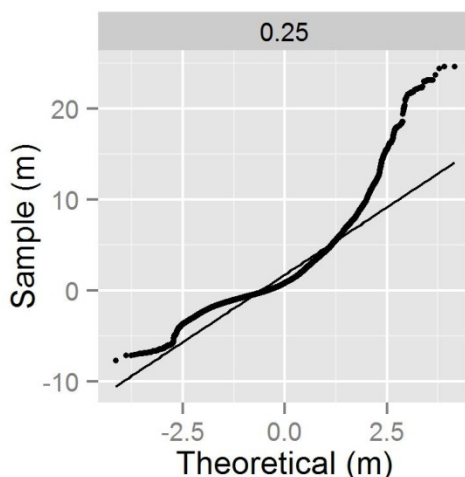


Figure 1. Normal Q-Q graph for the distribution of the difference in elevation between the elevation recorded by the dGNSS and the elevation of the corresponding ALS-derived DTM at pulse density of 0.25 pulses·m⁻². Higher pulse densities resulted in similar distributions with decreasing differences.

Reduced pulse density resulted in an increased variation in the canopy metrics on plot level (Figure 2). The standard deviation for the canopy metrics (S_M) showed that E.mean and E.90 were more stable than E.max and E.10. Reduction of the pulse density from 8 to 0.25 pulses·m⁻² increased the S_M for all variables. At a plot size of 0.07 ha the S_M increased from 0.09 to 0.90 m for E.mean and from 0.10 to 0.97 m for E.90. Correspondingly, the S_M increased from 0.16 to 1.03 m for E.max and from 0.19 to 1.50 m for E.10. Similarly, the S_M for D.0 and D.5 increased from 0.21 and 0.29 to 1.70 and 2.06 at a plot size of 0.07 ha.

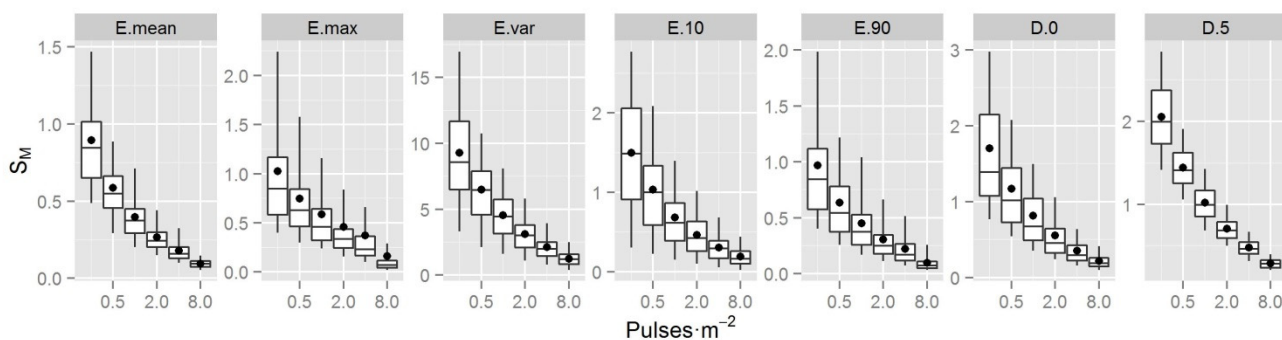


Figure 2. Box and whisker graphs of standard deviations (S_M) of canopy metrics at pulse densities of 0.25, 0.5, 1, 2, 4, and 8 pulses·m⁻² and a plot size of 0.07 ha. Whiskers at 5th and 95th percentile, mean value (black dot), and median value (black line). E.mean (mean above ground elevation of the ALS points), E.max (maximum above ground elevation of the ALS points), E.var (variance of the above ground elevation of the ALS points), E.10 and E.90 (10th and 90th height percentile of above ground elevation of the ALS points, respectively), and D.0 and D.5 (the proportion of ALS points above the ground and above the mean above ground elevation of the ALS points, respectively).

Table 2. Mean values (\bar{M}) of canopy metrics for plot sizes of 0.07, 0.14, 0.21, and 0.28 ha and pulse densities of 0.25, 0.5, 1, 2, 4, and 8 pulses·m⁻².

Plot Size (ha)	Pulse Density	\bar{M}						
		E.mean	E.max	E.var	E.10	E.90	D.0	D.5
0.07	0.25	24.59	41.42	107.59	10.06	36.15	89.99	55.18
0.07	0.5	24.49	41.75	108.79	9.79	36.18	90.01	55.25
0.07	1	24.47	41.98	108.93	9.72	36.16	90.09	55.39
0.07	2	24.49	42.22	109.17	9.70	36.19	90.17	55.47
0.07	4	24.59	42.36	108.72	9.82	36.21	90.29	55.68
0.07	8	24.69	42.56	108.17	9.93	36.26	90.33	55.73
0.14	0.25	24.59	43.44	112.27	9.48	36.64	90.07	54.67
0.14	0.5	24.47	43.70	113.24	9.28	36.60	90.07	54.74
0.14	1	24.45	43.92	113.60	9.21	36.60	90.15	54.80
0.14	2	24.49	44.11	113.67	9.24	36.63	90.23	54.89
0.14	4	24.60	44.23	113.15	9.40	36.68	90.38	55.05
0.14	8	24.70	44.39	112.60	9.52	36.74	90.44	55.14
0.21	0.25	24.60	44.38	114.04	9.33	36.93	90.16	54.35
0.21	0.5	24.47	44.61	114.91	9.12	36.86	90.14	54.42
0.21	1	24.43	44.81	115.22	9.04	36.84	90.18	54.50
0.21	2	24.49	45.03	115.35	9.07	36.88	90.28	54.59
0.21	4	24.59	45.17	114.83	9.22	36.93	90.40	54.72
0.21	8	24.70	45.33	114.22	9.35	37.00	90.48	54.82
0.28	0.25	24.59	45.01	115.87	9.13	37.05	90.01	54.31
0.28	0.5	24.45	45.23	116.54	8.93	36.97	89.97	54.36
0.28	1	24.44	45.47	116.95	8.87	36.98	90.04	54.45
0.28	2	24.48	45.68	116.93	8.91	37.00	90.12	54.52
0.28	4	24.59	45.85	116.40	9.05	37.05	90.24	54.66
0.28	8	24.69	45.98	115.84	9.18	37.12	90.33	54.75

Mean above ground elevation of the ALS points (E.mean), maximum above ground elevation of the ALS points (E.max), variance of the above ground elevation of the ALS points (E.var), 10th and 90th percentile of above ground elevation of the ALS points (E.10 and E.90, respectively), and the proportion of ALS points above the ground (D.0) and above the mean above ground elevation of the ALS points (D.5).

Further, reduced pulse density resulted in decreased reliability ratio, the ratio of the estimated among-plot variance to the estimated total variance (Figure 3). At pulse densities ≥ 2 pulses·m⁻² and a plot size of 0.07 ha, the reliability ratio was >0.95 for all canopy metrics. At a pulse density of 0.5 pulses·m⁻² and a plot size of 0.07 ha, the reliability ratios of E.var and D.0 were reduced down to 0.60 and 0.90, respectively. At a pulse density of 0.25 pulses·m⁻² and a plot size of 0.07 ha, E.var, E.10, and D.0 had an estimated reliability ratio of <0.9 , while the rest of the metrics had a reliability ratio of >0.9 .

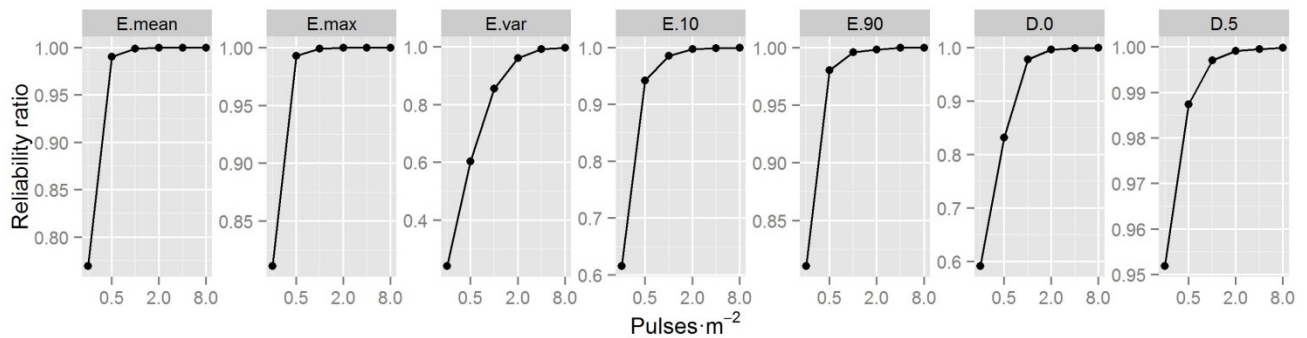


Figure 3. Graphs of reliability ratios for canopy metrics at pulse densities of 0.25, 0.5, 1, 2, 4, and 8 pulses·m⁻² and a plot size of 0.07 ha. E.mean (mean above ground elevation of the ALS points), E.max (maximum above ground elevation of the ALS points), E.var (variance of the above ground elevation of the ALS points), E.10 and E.90 (10th and 90th height percentile of above ground elevation of the ALS points, respectively), D.0 and D.5 (the proportion of points above the ground and above the mean above ground elevation of the ALS points, respectively).

3.3. Effects of Plot Size on Canopy Metrics

The statistics for the canopy metrics resulting from the simulations showed that the variation in the metrics was reduced with plot size for all metrics and at all pulse densities. Increasing the plot size means that the probability of including larger trees increases. As a result, we found that the maximum elevation (E.max) and the elevation of the top of the canopy (E.90) increased in value with increased plot size (Table 2). Increasing the plot size also means that more of the variability in elevation is captured by the plot and that this will result in an increased E.var. Metrics describing the elevation of the lowest part of the canopy (E.10) and the proportion of points above the mean above ground elevation of the ALS points (D.5), however, decreased in value with increasing plot size. E.mean and D.0 did not show any trend with increased plot size. The reliability ratio increased for all metrics with increasing plot size.

4. Discussion

The present study examined the effects of reduced ALS pulse density on the quality of the derived DTM and selected canopy metrics through a repeated reduction of pulse density. Pulse density was reduced from 13 pulses·m⁻² to densities of 8, 4, 2, 1, 0.5, and 0.25 pulses·m⁻². Because the accuracy of the dGNSS measurements is unknown (see discussion below), we compared the relative change in elevation of the dGNSS coordinates among pulse densities.

Our analysis showed that the mean DTM elevation was unaffected by the reduction in pulse density from 8 to 0.25 pulses·m⁻². This was in contradiction with results from other studies on reducing ALS data density [22,57,58]. In a tropical forest with similar conditions as in the present study, Leitold *et al.* [22] found an increased DTM elevation of 3.02 m at 1 pulse·m⁻², compared to a DTM from 20 pulses·m⁻². Leitold *et al.* [22] attributed the increased elevation to the algorithm [59] used to classify ground points. Hyyppä *et al.* [58], who used data collected in three separate flights, attributed an increased

DTM elevation to the beam size and sensitivity of the laser. In initial analysis of the present study the ground classification algorithm [51] was used with default settings and without the g and w parameter. This resulted in systematic reduction of DTM elevation with reduced pulse density. It is therefore important to assess if observed effects of reduced pulse density on the DTM are related to the applied ground classification algorithm.

Valbuena *et al.* [60] assessed the vertical accuracy of a dGNSS receiver (Topcon Hiperpro), similar to the receiver used in the present study, under pine canopies in Spain. By using true coordinates obtained in a total station traverse, they found the accuracy to be 1.18 m with a standard deviation of 1.55 m. It is therefore expected that our recordings under dense rainforest canopies with the reception of fewer satellite signals and more problems with multipath signals, result in lower accuracy. Thus, the \bar{D} of 1.8 m in the present study can be explained by the fact that, although the vertical precision of the dGNSS positions was reported to be 0.39 m (section 2.2), the accuracy remains unknown [60].

In biomass studies where the key information is the vegetation height relative to the terrain elevation derived from the same ALS data, a systematic shift in the modeled surface is not a problem in itself. Of greater concern is the random error of the modeled terrain elevation. The precision of the DTM elevation was relatively unaffected by the data reduction at pulse densities >1 pulse \cdot m $^{-2}$. As a consequence of having fewer ground points available for construction of the DTM the precision decreased at pulse densities <1 pulse \cdot m $^{-2}$. This decreased precision will directly translate into increased variation in the ALS-derived canopy metrics.

Analysis of commonly applied canopy metrics showed that the metrics were affected differently by pulse density. As previously documented by Gobakken and Næsset [12], E_{max} , which characterizes the maximum elevation of the canopy, decreases with decreasing pulse density. Mean values of the other metrics assessed in the present study were found to be stable. Reduced pulse density increased the variation in canopy metrics and will result in models with increased residual variance. The estimated reliability of the metrics as predictors, expressed by the reliability ratio, showed that all metrics were reliable (reliability ratio >0.9) at pulse densities of down to 2 pulses \cdot m $^{-2}$. Further reduction of pulse density resulted in some canopy metrics becoming less reliable although most metrics retain a reliability ratio of >0.9 at 0.5 pulses \cdot m $^{-2}$. In sparse pulse density conditions (<1 m $^{-2}$), and with use of predictors with a reliability ratio <0.9 Magnussen *et al.* [20] proposed the use of a model calibration procedure.

Prediction of forest biomass over large areas using ALS often relies on data collected using different sensors and flying altitudes. Thus, different areas will consequently differ in pulse density. Varying pulse densities within a single ALS mission will also be inevitable as a result of flight line overlap, where the pulse density will be roughly double the density in the rest of the area. In addition, flights covering rugged terrain and steep slopes will result in varying pulse densities depending on the scan angle and distance to the ground. Estimates for parts of the forest area could therefore be over or under-predicted, due to different pulse densities than those in areas used for model development. To prevent such effects, suitable strategies could be to reduce the pulse density down to the smallest density in the project area, to remove points from overlapping flight lines, to weight ALS points by the surrounding number of points [61] or to include pulse density as a predictor in the model [4]. Another strategy could be to let pulse densities be reflected in the sampling design of the field survey, for

example by using stratified sampling. That would require access to the ALS data in the design phase of the field survey.

Increasing the field plot size reduces the variation in ALS-derived metrics and could counter the effects in sparse pulse density missions, concurring with the results of Gobakken and Næsset [12]. However, large field plots are costly and finding the optimal balance between costs and desired accuracy has for decades been an issue of interest in designing forest inventories.

The present study aimed to simulate the effects caused by various flight elevations and speeds on the DTM and canopy metrics derived from ALS data. Some effects, however, were not simulated and studied. Increased flight elevation will result in increased footprint size. Studies of the effects of footprint size on derived tree heights have shown that increased footprint size reduces the derived tree height estimates [62]. This effect was stronger for trees with a narrow crown and in a tropical forest the effect is likely to be small for the derived canopy metrics. Larger footprint sizes will, however, also have less energy per unit area and be less able to penetrate through the canopy [58,63], resulting in a reduced proportion of ground points. The influence of footprint size and pulse energy is likely to be of importance and should be investigated in future studies.

5. Conclusions

The present study showed that canopy metrics derived from sparse pulse density ALS data can be used for biomass estimation in a tropical rainforest. Reducing the pulse density from 8 to 0.25 pulses·m⁻² increased the variation in the DTMs and canopy metrics. However, the replication effects, expressed by the reliability ratio, were not important (reliability ratio >0.9) at pulse densities of >0.5 pulses·m⁻². Increased plot size increased the reliability ratio of canopy metrics, and could be used to counter effects of variation in canopy metrics obtained from sparse pulse density ALS data. The combination of plot area and pulse density need to provide enough points for a reliable estimate of the metrics used to predict forest biomass.

Acknowledgments

The present study is part of the project “Enhancing the measuring, reporting and verification of forests in Tanzania through the application of advanced remote sensing techniques” funded by the Royal Norwegian Embassy in Tanzania as part of the Norwegian International Climate and Forest Initiative. We are grateful to Terratec AS, Norway, for conducting the ALS data acquisition.

Author Contributions

Endre Hofstad Hansen was the main author of the manuscript, planned the study with the co-authors, carried out and supervised the fieldwork, and performed all data processing. Terje Gobakken planned the acquisition of remotely sensed data, co-authored and revised the manuscript. Erik Næsset designed parts of the study, co-authored and revised the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Grace, J.; Mitchard, E.; Gloor, E. Perturbations in the carbon budget of the tropics. *Glob. Change Biol.* **2014**, *20*, 3238–3255.
2. UNFCCC. Action taken by the conference of the parties at its sixteenth session. In Proceedings of the Sixteenth Session of the Conference of the Parties, Cancun, Mexico, 29 November–10 December 2010.
3. Asner, G.P.; Mascaro, J. Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. *Remote Sens. Environ.* **2014**, *140*, 614–624.
4. Jubanski, J.; Ballhorn, U.; Kronseder, K.; Franke, J.; Siegert, F. Detection of large above ground biomass variability in lowland forest ecosystems by airborne LiDAR. *Biogeosciences Discuss.* **2012**, *9*, 11815–11842.
5. Vincent, G.; Sabatier, D.; Blanc, L.; Chave, J.; Weissenbacher, E.; Pélissier, R.; Fonty, E.; Molino, J.F.; Coutron, P. Accuracy of small footprint airborne LiDAR in its predictions of tropical moist forest stand structure. *Remote Sens. Environ.* **2012**, *125*, 23–33.
6. Hou, Z.; Xu, Q.; Tokola, T. Use of ALS, Airborne CIR and ALOS AVNIR-2 data for estimating tropical forest attributes in Lao PDR. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 776–786.
7. Andersen, H.E.; Reutebuch, S.E.; McGaughey, R.J.; d'Oliveira, M.V. N.; Keller, M. Monitoring selective logging in western Amazonia with repeat LiDAR flights. *Remote Sens. Environ.* **2013**, *151*, 157–165.
8. Laurin, G.V.; Chen, Q.; Lindsell, J.A.; Coomes, D.A.; Frate, F.D.; Guerriero, L.; Pirotti, F.; Valentini, R. Above ground biomass estimation in an African tropical forest with LiDAR and hyperspectral data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *89*, 49–58.
9. Asner, G.P.; Powell, G.V. N.; Mascaro, J.; Knapp, D.E.; Clark, J.K.; Jacobson, J.; Kennedy-Bowdoin, T.; Balaji, A.; Paez-Acosta, G.; Victoria, E.; Secada, L.; Valqui, M.; Hughes, R.F. High-resolution forest carbon stocks and emissions in the Amazon. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 16738–16742.
10. Ioki, K.; Tsuyuki, S.; Hirata, Y.; Phua, M.H.; Wong, W.V. C.; Ling, Z.Y.; Saito, H.; Takao, G. Estimating above-ground biomass of tropical rainforest of different degradation levels in Northern Borneo using airborne LiDAR. *For. Ecol. Manage.* **2014**, *328*, 335–341.
11. d'Oliveira, M.V.N.; Reutebuch, S.E.; McGaughey, R.J.; Andersen, H.E. Estimating forest biomass and identifying low-intensity logging areas using airborne scanning LiDAR in Antimary State Forest, Acre State, Western Brazilian Amazon. *Remote Sens. Environ.* **2012**, *124*, 479–491.
12. Gobakken, T.; Næsset, E. Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Can. J. For. Res.* **2008**, *38*, 1095–1109.
13. Holmgren, J. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. *Scand. J. Forest Res.* **2004**, *19*, 543–553.
14. Maltamo, M.; Eerikainen, K.; Packalen, P.; Hyyppä, J. Estimation of stem volume using laser scanning-based canopy height metrics. *Forestry* **2006**, *79*, 217–229.
15. Magnusson, M.; Fransson, J.E. S.; Holmgren, J. Effects on estimation accuracy of forest variables using different pulse density of laser data. *For. Sci.* **2007**, *53*, 619–626.

16. Watt, M.; Adams, T.; Gonzalez Aracil, S.; Marshall, H.; Watt, P. The influence of LiDAR pulse density and plot size on the accuracy of New Zealand plantation stand volume equations. *J. For. Sci.* **2013**, *43*, doi: 10.1186/1179-5395-43-15.
17. Vauhkonen, J.; Tokola, T.; Maltamo, M.; Packalen, P. Effects of pulse density on predicting characteristics of individual trees of Scandinavian commercial species using alpha shape metrics based on airborne laser scanning data. *Can. J. Remote Sens.* **2008**, *34*, S441–S459.
18. Strunk, J.; Temesgen, H.; Andersen, H.E.; Flewelling, J.P.; Madsen, L. Effects of LiDAR pulse density and sample size on a model-assisted approach to estimate forest inventory variables. *Can. J. Remote Sens.* **2012**, *38*, 644–654.
19. Jakubowski, M.K.; Guo, Q.; Kelly, M. Tradeoffs between LiDAR pulse density and forest measurement accuracy. *Remote Sens. Environ.* **2013**, *130*, 245–253.
20. Magnussen, S.; Næsset, E.; Gobakken, T. Reliability of LiDAR derived predictors of forest inventory attributes: A case study with Norway spruce. *Remote Sens. Environ.* **2010**, *114*, 700–712.
21. Ruiz, L.; Hermosilla, T.; Mauro, F.; Godino, M. Analysis of the influence of plot size and LiDAR density on forest structure attribute estimates. *Forests* **2014**, *5*, 936–951.
22. Leitold, V.; Keller, M.; Morton, D.C.; Cook, B.D.; Shimabukuro, Y.E. Airborne lidar-based estimates of tropical forest structure in complex terrain: opportunities and trade-offs for REDD+. *Carbon Balance Manag.* **2015**, *10*, doi: 10.1186/s13021-015-0013-x.
23. Treitz, P.; Lim, K.; Woods, M.; Pitt, D.; Nesbitt, D.; Etheridge, D. LiDAR sampling density for forest resource inventories in Ontario, Canada. *Remote Sens.* **2012**, *4*, 830–848.
24. National Lidar Dataset. Available online: http://en.wikipedia.org/wiki/National_lidar_dataset (accessed on 25 February 2015).
25. Reutebuch, S.E.; McGaughey, R.J.; Andersen, H.E.; Carson, W.W. Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. *Can. J. Remote Sens.* **2003**, *29*, 527–535.
26. Töyrä, J.; Pietroniro, A.; Hopkinson, C.; Kalbfleisch, W. Assessment of airborne scanning laser altimetry (LiDAR) in a deltaic wetland environment. *Can. J. Remote Sens.* **2003**, *29*, 718–728.
27. Bowen, Z.H.; Waltermire, R.G. Evaluation of light detection and ranging (LiDAR) for measuring river corridor topography. *J. Am. Water Resour. Assoc.* **2002**, *38*, 33–41.
28. Hodgson, M.E.; Jensen, J.; Raber, G.; Tullis, J.; Davis, B.A.; Thompson, G.; Schuckman, K. An evaluation of LiDAR-derived elevation and terrain slope in leaf-off conditions. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 817–823.
29. Tinkham, W.T.; Huang, H.Y.; Smith, A.M. S.; Shrestha, R.; Falkowski, M.J.; Hudak, A.T.; Link, T.E.; Glenn, N.F.; Marks, D.G. A comparison of two open source LiDAR surface classification algorithms. *Remote Sens.* **2011**, *3*, 638–649.
30. Meng, X.L.; Currit, N.; Zhao, K.G. Ground filtering algorithms for airborne LiDAR data: A review of critical issues. *Remote Sens.* **2010**, *2*, 833–860.
31. Testprosjekt—Klassifisering av Laserdata. Available online: http://kartverket.no/Documents/Om%20Kartverket/Geovekst/Testprosjekt%20klassifisering%20av%20laserdata_2012.pdf (accessed on 30 June 2015).
32. Bater, C.W.; Coops, N.C. Evaluating error associated with LiDAR-derived DEM interpolation. *Comput. Geosci.* **2009**, *35*, 289–300.

33. Liu, X.; Zhang, Z.; Peterson, J.; Chandra, S. The effects of LiDAR data density on DEM accuracy. In Proceedings of 2007 International Congress on Modeling and Simulation, Christchurch, New Zealand, 31 December 2007.
34. Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* **2002**, *80*, 88–99.
35. Næsset, E.; Bollandsås, O.M.; Gobakken, T. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. *Remote Sens. Environ.* **2005**, *94*, 541–553.
36. Li, Y.Z.; Andersen, H.E.; McGaughey, R. A comparison of statistical methods for estimating forest biomass from light detection and ranging data. *West. J. Appl. For.* **2008**, *23*, 223–231.
37. Fassnacht, F.E.; Hartig, F.; Latifi, H.; Berger, C.; Hernández, J.; Corvalán, P.; Koch, B. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sens. Environ.* **2014**, *154*, 102–114.
38. Asner, G.P. Tropical forest carbon assessment: integrating satellite and airborne mapping approaches. *Environ. Res. Lett.* **2009**, *4*, doi:10.1088/1748-9326/4/3/034009.
39. Mascaró, J.; Detto, M.; Asner, G.P.; Muller-Landau, H.C. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sens. Environ.* **2011**, *115*, 3770–3774.
40. Asner, G.P.; Knapp, D.E.; Martin, R.E.; Tupayachi, R.; Anderson, C.B.; Mascaró, J.; Sinca, F.; Chadwick, K.D.; Higgins, M.; Farfan, W.; Llactayo, W.; Silman, M.R. Targeted carbon conservation at national scales with high-resolution monitoring. *Proc. Natl. Acad. Sci.* **2014**, *111*, E5016–E5022.
41. Hernández-Stefanoni, J.; Dupuy, J.; Johnson, K.; Birdsey, R.; Tun-Dzul, F.; Peduzzi, A.; Caamal-Sosa, J.; Sánchez-Santos, G.; López-Merlín, D. Improving species diversity and biomass estimates of tropical dry forests using airborne LiDAR. *Remote Sens.* **2014**, *6*, 4741–4763.
42. Mauya, E.W.; Hansen, E.H.; Gobakken, T.; Bollandsås, O.M.; Malimbwi, R.E.; Næsset, E. Effects of field plot size on prediction accuracy of aboveground biomass in airborne laser scanning-assisted inventories in tropical rain forests of Tanzania. *Carbon Balance Manag.* **2015**, *10*, doi: 10.1186/s13021-015-0013-x.
43. Vauhkonen, J.; Maltamo, M.; McRoberts, R.; Næsset, E. Introduction to forestry applications of airborne laser scanning. In *Forestry Applications of Airborne Laser Scanning*; Maltamo, M., Næsset, E., Vauhkonen, J., Eds.; Springer Netherlands: Dordrecht, Netherlands, 2014; Vol. 27, pp. 1–16.
44. Zolkos, S.G.; Goetz, S.J.; Dubayah, R. A meta-analysis of terrestrial aboveground biomass estimation using LiDAR remote sensing. *Remote Sens. Environ.* **2013**, *128*, 289–298.
45. Tomppo, E.; Malimbwi, R.; Katila, M.; Mäkisara, K.; Henttonen, H.M.; Chamuya, N.; Zahabu, E.; Otieno, J. A sampling design for a large area forest inventory: case Tanzania. *Can. J. For. Res.* **2014**, *44*, 931–948.
46. Hamilton, A.C.; Bensted-Smith, R. *Forest Conservation in the East Usambara Mountains, Tanzania*; International Union for Conservation of Nature: Gland, Switzerland.
47. Kouba, J. A Guide to Using International GNSS Service (IGS) Products. Available online: <https://igsceb.jpl.nasa.gov/components/usage.html> (accessed on 30 June 2015).

48. Pinnacle User's Manual: Available online: <http://folk.uio.no/treiken/GEO4530/pinnacle.pdf> (accessed on 30 June 2015).
49. Hansen, E.; Gobakken, T.; Bollandsås, O.; Zahabu, E.; Næsset, E. Modeling aboveground biomass in dense tropical submontane rainforest using airborne laser scanner data. *Remote Sens.* **2015**, *7*, 788–807.
50. McGaughey, R.J. FUSION/LDV: Software for LiDAR Data Analysis and Visualization. Available online: http://forsys.cfr.washington.edu/fusion/FUSION_manual.pdf (accessed on 30 June 2015).
51. Kraus, K.; Pfeifer, N. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS J. Photogramm. Remote Sens.* **1998**, *53*, 193–203.
52. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.
53. Zandbergen, P.A. Positional accuracy of spatial data: Non-normal distributions and a critique of the national standard for spatial data accuracy. *Trans. GIS* **2008**, *12*, 103–130.
54. Höhle, J.; Höhle, M. Accuracy assessment of digital elevation models by means of robust statistical methods. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 398–406.
55. Lim, K.; Hopkinson, C.; Treitz, P. Examining the effects of sampling point densities on laser canopy height and density metrics. *For. Chron.* **2008**, *84*, 876–885.
56. Fuller, W.A. *Measurement Error Models*. John Wiley and Sons: New York, NY, USA, 1987; p. 440.
57. Anderson, E.S.; Thompson, J.A.; Crouse, D.A.; Austin, R.E. Horizontal resolution and data density effects on remotely sensed LiDAR-based DEM. *Geoderma* **2006**, *132*, 406–415.
58. Hyyppä, J.; Yu, X.; Hyyppä, H.; Kaartinen, H.; Honkavara, E.; Rönholm, P. Factors affecting the quality of DTM generation in forested areas. In Proceedings of 2005 ISPRS Workshop on Laser Scanning. Enschede, Netherlands, 12–14 September 2005.
59. Zhang, K.; Chen, S.C.; Whitman, D.; Shyu, M.L.; Yan, J.; Zhang, C., A progressive morphological filter for removing nonground measurements from airborne LiDAR data. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 872–882.
60. Valbuena, R.; Mauro, F.; Rodriguez-Solano, R.; Antonio Manzanera, J. Partial least squares for discriminating variance components in global navigation satellite systems accuracy obtained under scots pine canopies. *For. Sci.* **2012**, *58*, 139–153.
61. Nyström, M.; Holmgren, J.; Olsson, H. Prediction of tree biomass in the forest-tundra ecotone using airborne laser scanning. *Remote Sens. Environ.* **2012**, *123*, 271–279.
62. Andersen, H.-E.; Reutebuch, S.E.; McGaughey, R.J. A rigorous assessment of tree height measurements obtained using airborne LiDAR and conventional field methods. *Can. J. Remote Sens.* **2006**, *32*, 355–366.
63. Goodwin, N.R.; Coops, N.C.; Culvenor, D.S. Assessment of forest structure with airborne LiDAR and the effects of platform altitude. *Remote Sens. Environ.* **2006**, *103*, 140–152.