Examining the effects of sampling point densities on laser canopy height and density metrics

by K. Lim[1], C. Hopkinson[2,3] and P. Treitz[4]

ABSTRACT

Forest resource managers rely on the information extracted from forest resource inventories to manage forests sustainably and efficiently, thereby supporting more precise decision-making. Light detection and ranging (LiDAR) is a relatively new technology that has proven to enhance forest resource inventories. However, the relationship between LiDAR sampling point density (which is directly related to acquisition and processing costs) and accuracy and precision of forest variable estimation has not yet been established across a range of forest ecosystems. In this study, 2 airborne LiDAR surveys using the same sensor, but configured with disparate parameters, were carried out over the York Regional Forest near Toronto, Canada producing 2 data sets characterized by different sampling point densities. The effects of 2 sampling point densities on 23 laser canopy height and density metrics typically used in forest studies at the plot level were examined with comparisons grouped by first and last return data. The minimum (h_{min}) and maximum (h_{max}) laser canopy heights were statistically different for first and last returns. The proportion of laser returns (i.e., canopy density) in the upper (d_1) and lower (d_{10}) range of laser canopy heights was statistically different for the first returns, whereas only a single canopy density metric was different for the last returns (d_9). These results suggest that changes in sampling point density (due to changes in scan angle and altitude) only affect laser canopy height and density metrics that are characterized by the small percentage of returns from the very top (h_{max}; d_1) and base of the canopy (h_{min}; d_{10}) (i.e., those metrics that characterize the tail ends of the distributions of laser canopy heights). Consequently, higher sampling point densities may add little value to current LiDAR forest research or operations at the stand level, as metrics derived from the canopy profile can be implemented for biophysical variable estimation. Implications for forest management are in terms of identifying which aspects of LiDAR project design: a) impact the quality and cost-effectiveness of derived FRI information; b) should be specified within a LiDAR request for proposals; or c) scrutinized within LiDAR project tender documentation.

Key words: LiDAR, airborne laser scanning, FRI, canopy profile, biophysical variables, sampling density

FRENCH TRANSLATION BEING DONE.

[1]Department of Geography, Queen's University, Kingston, Ontario K7L 3N6. (current address: Lim Geomatics Inc., 264 Catamount Court, Ottawa, ON K2M 0A9).
[2]Otterburn Geographic, RR3 Middleton, Nova Scotia B0S 1P0.
[3]Applied Geomatics Research Group, 50 Elliott Rd, Lawrencetown, Nova Scotia B0S 1M0.
[4]Department of Geography, Queen's University, Kingston, Ontario K7L 3N6. Corresponding author. E-mail: paul.treitz@queensu.ca

K. Lim

C. Hopkinson

P. Treitz

Introduction

Forest research at the scale of the individual tree, plot, and stand using airborne light detection and ranging (LiDAR) has progressed rapidly in recent years (e.g., Lefsky *et al.* 2002, Næsset *et al.* 2004, Andersen *et al.* 2005, Maltamo *et al.* 2005, Falkowski *et al.* 2006, Tickle *et al.* 2006, Bollandsås and Næsset 2007, Thomas *et al.* 2008). Parallel to these research efforts have been the continued advances in LiDAR technology with commercial sensors now capable of pulse repetition frequencies (PRFs) exceeding 167 kHz (Optech Inc. 2006) and waveform digitization (Hug *et al.* 2004).

Studies that have focused on the estimation and prediction of various forest biophysical variables (e.g., maximum canopy height, mean canopy height, Lorey's height, basal area, crown closure, diameter at breast height (DBH), biomass, leaf area index (LAI) and volume at the plot level using LiDAR have primarily relied on establishing empirical relationships between various laser canopy height and density metrics, and the biophysical variables of interest (e.g., Nelson 1997; Nelson *et al.* 1997, 2004; Magnussen and Boudewyn 1998; Lefsky *et al.* 1999; Magnussen *et al.* 1999; Means *et al.* 1999, 2000; Næsset and Bjerknes 2001; Næsset 2002; Lim *et al.* 2003a; Lim and Treitz 2004; Hopkinson *et al.* 2006, Thomas *et al.* 2006). Næsset and Bjerknes (2001) reported that the mean height of dominant trees could be estimated using the laser canopy height corresponding to the 90th percentile and a canopy density metric, which was defined as the ratio between the number of returns from the canopy and total laser pulses transmitted. The concept of canopy density metrics was extended by Næsset (2002) to include proportions of first and last returns at intervals throughout a forest canopy. Lim *et al.* (2003a) demonstrated that the maximum and mean laser height, and an average height based on LiDAR intensity, could be used to estimate 10 forest biophysical variables. In some studies, laser height metrics used in stepwise regression analyses are differentiated by their return type (i.e., only, first, or last return) (Næsset and Økland 2002, Holmgren 2004). The last return is accomplished by implementing last return logic in the LiDAR receiver electronics to sample and hold all returns after the first return but only record the final return. This last return can be representative of any part of the canopy profile below the first return. While the specific laser canopy height and density predictor variables used in these types of studies may vary, the general modelling approach and types of metrics employed are consistent (Lim *et al.* 2003b).

Of interest to these types of studies carried out at the plot level is how sampling point density and survey configuration affect laser canopy height and density metrics across different forest types. As the development of more technologically advanced LiDAR sensors continues, with advances manifesting most notably as increases in PRFs, the generation of spatially dense LiDAR data sets is becoming more prominent. However, high sample point densities still require low-altitude data acquisition and/or the adoption of a reduced scan angle. Flying with this type of configuration can increase acquisition time by several factors and therefore significantly increase acquisition costs. However, it remains unknown as to whether or not LiDAR data characterized by a high sampling point density are somehow "superior" with respect to the information content than those characterized by a lower sampling point density for forest studies. Recently, large-scale high-density commercial LiDAR remote sensing projects for forest resource inventory have been conducted in Norway. Næsset (2004a) reported on the validation of one of these projects and found that the accuracy of LiDAR models of various structural variables was accurate and consistent across the forest sampled. Validation revealed low standard errors for mean height (0.36–1.37 m); dominant height (0.70–1.55 m); basal area (2.38–4.88 m^2) and stand volume (13.9–45.9 m^3ha^{-1}). These results are encouraging for the implementation of LiDAR remote sensing for forest resource inventory. Unfortunately, studies comparing LiDAR sampling densities for predicting forest biophysical variables within the Canadian context are lacking. Thomas *et al.* (2006) compared a low-density to high-density LiDAR dataset for predicting variables related to tree height for a boreal mixedwood. Validation of the models against independent field plots revealed that low- and high-density LiDAR models were highly correlated with mean dominant height, basal area, and average aboveground biomass (low density: $R2$ = 0.90, 0.91, and 0.92, and high density: $R2$ = 0.84, 0.89, and 0.91). However, the low-density dataset was not able to accurately model variables related to canopy density (i.e., crown closure). These results would indicate that accurate predictions of many forest structural variables are possible with low-density LiDAR data at the plot and stand level, a requirement for the inclusion of LiDAR into cost-effective operational forest inventories. By furthering our understanding of the relationship between sampling point densities and laser canopy height and density metrics, resource managers and data providers will be better able to configure surveys to satisfy the specifications for given project goals and collect only data that are necessary to accurately estimate the biophysical properties of interest.

The objective of this research is to compare laser canopy height and density metrics derived from LiDAR data collected at 2 different sampling point densities. This paper does not treat the topic of how these variables are used for the estimation of forest biophysical properties, but instead focuses on the types of variables typically employed and how sampling point density impacts those variables. Note that we do not attempt to treat how individual input survey parameters, such as PRF, scan frequency, scan angle, or beam divergence, affects laser canopy height and density metrics as in the case of Næsset (2004b) or Goodwin *et al.* (2006), but instead focus on the effects of sampling point density on these metrics. It is extremely important to understand the impact that sampling density has on estimating forest structural variables, in order that cost-effective surveys can be designed to maximize information content while minimizing acquisition costs.

Methods

Study site

The York Regional Forest (YRF) (44°05?N, 79°20?W) extends over 333 hectares of land near the Town of Whitchurch–Stouffville, which is located approximately 50 km north of Toronto, Ontario, Canada (Fig. 1). The YRF consists predominantly of managed red (*Pinus resinosa* Ait.) and white (*Pinus strobus* L.) pine plantations, and unmanaged deciduous stands largely composed of sugar maple (*Acer saccharum* Marsh.) and hickory (*Carya* spp.). The YRF has been used in previous studies to: (i) assess plot-level forest metrics (e.g., DBH and tree height) using ground-based LiDAR (Hopkinson *et al.* 2004a); (ii) map snowpack depth beneath forest canopies using airborne LiDAR (Hopkinson *et al.* 2004b); and (iii) estimate conifer canopy growth over a 5-year period (Hopkinson *et al.* 2008).

Sixty circular plots, representing stands of varying species composition, age, and forest structure, were systematically selected over the north tract of the YRF. A single plot located in an open area/clearing (i.e., non-forested) was excluded from the analysis reducing the number of sampling units to 59. The spatial configuration of plots was designed so as to allow a distance of 200 m between the eastings and 100 m between the northings of the plot centers (Fig. 1). The area of each plot was 0.04 ha (i.e., 400 m^2) in area. Although no specific ground reference data were collected for the plots, a recent forest resource inventory (FRI) was available to assist in interpreting the characteristics of the stand within which each plot was located. A total of 52 plots were located in conifer plantations. Of the remaining 7 plots, 4 were located in upland hardwood stands and 3 in mixed conifer stands. The species composition, age, and average tree height are summarized in Table 1. The adoption of a systematic spatial configuration of plots not only ensured that the study captured a range of different stands with respect to their age, species composition, and forest structure, but also allowed the study to account for any local spatial variability in sampling point density for any given survey.

LiDAR data

Two leaf-on LiDAR surveys of the YRF were carried out within a two-hour window on 24th September 2000 using an Optech ALTM 1225 (Optech Inc., Toronto, Ontario, Canada). The surveys were configured so as to yield 2 LiDAR data sets that differed with respect to sampling density by a ratio of approximately 3:1. The PRF for the ALTM 1225 is 25 kHz and for any given laser pulse, the first and last returns are recorded.

The input and output survey parameters for each of the 2 LiDAR surveys are summarized in Table 2. Although the PRF of 25 kHz and the aircraft velocity of 60 m/s were held constant for both surveys, the height aboveground level (AGL) at which the plane was flown, scan angle, scan frequency, and orientation of the flight lines differed. For the first survey, the LiDAR sensor was configured with a maximum scan angle of 12° and a scan frequency of 30 Hz with the aircraft travelling in a north–south direction at 800 m AGL. For the second survey, the LiDAR sensor acquisition settings included a maximum scan angle of 20° and a scan frequency of 21 Hz. The flight path was oriented in an east–west direction at 1200 m AGL. The flight line spacing for the first and second survey was 200 m and 550 m, respectively, which maintained close to 50% swath overlap for both surveys. As a result, the first survey was carried out at a lower altitude with a narrower scan angle and higher scan frequency than the second survey resulting in the production of a LiDAR data set with a higher average sampling point density.

The configuration of input parameters for the 2 surveys resulted in each survey having different laser footprint sizes in addition to other output survey parameters (Table 2). The concentrated energy of the laser pulse footprint diameters for the low- and high-altitude surveys were approximately 24 cm and 31 cm, respectively. Based on previous research, it was assumed that such a small difference between laser footprint sizes would not significantly affect measurements of laser canopy heights (Nilsson 1996, Næsset 2004b, Hopkinson 2007). The low-altitude survey had a swath of 340 m, whereas the high-altitude survey had a swath of 870 m. These survey configurations resulted in laser pulse postings in the across-

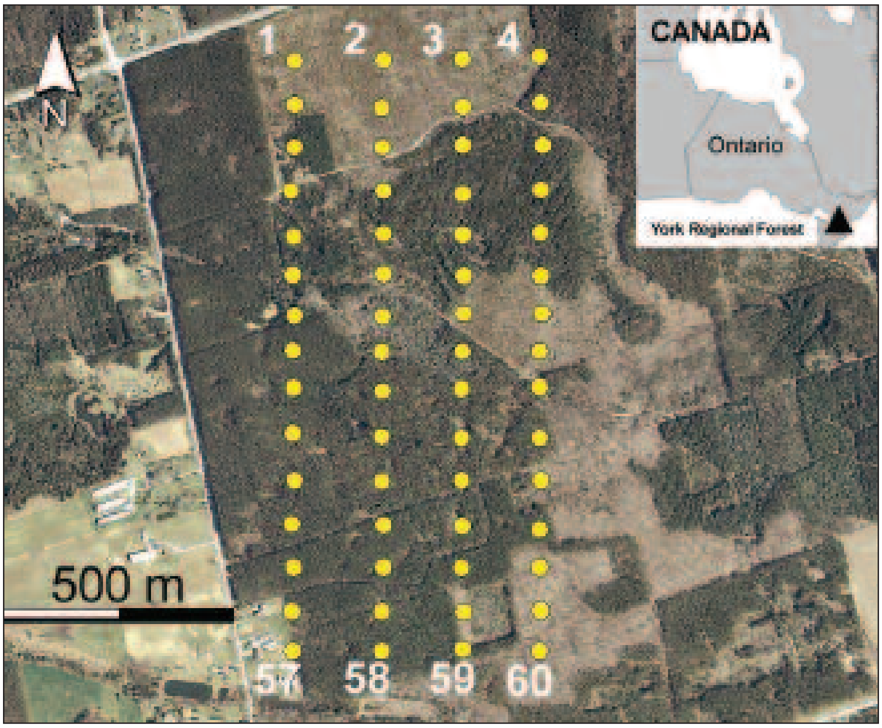

Fig. 1. The geographic location of the York Regional Forest (upper right inset) and the distribution of plots over the study area displayed against an aerial photograph acquired in the fall of 1999.

Table 1. Species composition, age, average tree height, and the general classification of each plot based on the forest resource inventory data where available

Species	n	Age (years)			Average Tree Height[a] (m)			Category
		Min.	Max.	Mean	Min.	Max.	Mean	
Red Pine	41	11	72	38.7	15.2	26.0	20.6	Plantation
Red & White Pine Mix	4	13	72	52.5	20.0	28.0	25.3	Plantation
White Pine	1	71	71	71	N/A	N/A	N/A	Plantation
Jack Pine	6	73	73	73	15.6	17.5	16.2	Plantation
Pine/Cedar	3	70	70	70	26.0	26.0	26.0	Mixedwood
Sugar Maple	4	75	75	75	30.3	30.3	30.3	Upland Hardwood

[a]Average tree heights were not available for all plots.
Red Pine – *Pinus resinosa* Ait.; White Pine – *Pinus strobus* L.; Jack Pine – *Pinus banksiana* Lamb.; Cedar – *Thuja occidentalis* L.; Sugar Maple – *Acer saccharum Marsh*.

Table 2. Input and output parameters for the low- and high-altitude surveys

	Low-altitude Survey	High-altitude Survey
Input Parameters		
PRF (kHz)	25	25
Scanner Frequency (Hz)	30	21
Scan Angle (°)	12	20
Aircraft Velocity (ms^{-1})	60	60
Flying altitude (m AGL)	800	1200
Line spacing (m)	200	550
Output Parameters		
Across track pulse spacing (m)	0.8	1.5
Along track pulse spacing (m)	0.9	1.4
Footprint (cm)	24	36
Swath width (m)	340	870

and along-track directions of 0.8 m and 0.9 m (low altitude), and 1.5 m and 1.4 m (high altitude), respectively.

The canopy sampling point densities per 400 m2 for each plot across the 2 surveys and grouped by laser return class (i.e., first and last returns) are presented in Fig. 2. Although there is obvious variability in the density produced from each survey for each plot, it was consistently higher for the low-altitude survey, with the ratio typically ranging from 2 to 4 times.

Laser canopy height and density metrics

Classified ground laser returns acquired from a LiDAR survey of the YRF during deciduous leaf-off and minimum understory conditions in December 2000 were used to interpolate a digital elevation model (DEM). Snow was not present in the study area during the December LiDAR acquisition. These data were collected using a similar Optech ALTM sensor with all data registered to the same survey monument and GPS base station as the other surveys. During leaf-off conditions, there are more samples of the true ground surface, providing for a more accurate and precise DEM than could be achieved during leaf-on conditions. A triangular irregular network (TIN) was constructed from the classified ground laser returns and then converted using linear interpolation to a DEM with a cell resolution of 1 m. The accuracy of the DEM was not critical for this study as the same DEM was used across both surveys to derive vegetation laser heights. By using a common (i.e., baseline) DEM to calculate vegetation laser heights, any bias introduced into laser canopy heights as a function of DEM inaccuracies will be equally present in both surveys and therefore of no consequence to the tests performed in this study. The rationale for using a DEM generated from a leaf-off LiDAR survey (even for a conifer-dominated study site) was that foliage reduces the ability of LiDAR to accurately sample the "true" ground surface.

To obtain vegetation laser heights, the DEM z-values were subtracted from the corresponding z-value of each laser return at matching x–y coordinates. Vegetation laser heights that were less than 2 m were removed from the LiDAR data sets of each survey. A "hard" threshold of 2 m was selected to eliminate laser returns corresponding to the ground and low-lying vegetation (e.g., brush) and is an approach that has been adopted by others (e.g., Nilsson 1996, Næsset 2004b).

The minimum (h_{min}), maximum (h_{max}), and mean laser canopy height (h_{mean}), in addition to the deciles (i.e., $h_1 \dots h_9$) and coefficient of variation (h_{cv}) of laser canopy heights, were derived for each plot across both surveys producing 13 pairs of metrics (Ritchie *et al.* 1993, Magnussen and Boudewyn 1998). A set of 10 canopy density metrics (i.e., $d_1 \dots d_{10}$) were also derived (Næsset 2002). For these, the range of the laser canopy heights was divided into 10 equal intervals. Each of the 10 canopy density metrics corresponds to a proportion of laser canopy heights within a given interval. For example, the first canopy density metric would correspond to the proportion of laser canopy heights found in the first interval at the lower range. Conversely, the 10th canopy density metric would correspond to the proportion of laser canopy heights in the upper interval at the upper range. These metrics were selected for analysis as they have been demonstrated in previous studies, on an individual basis or in combinations, as possible predictors of various forest biophysical variables (Lim *et al.* 2003b). Inclusion of these metrics also provides a suitable characterization of the canopy height and density profiles. The characteristics of the 23 metrics considered in this study are summarized in Table 3.

Data analysis

Paired *t*-tests were performed to determine if a canopy height or density metric from the low-altitude survey was statisti-

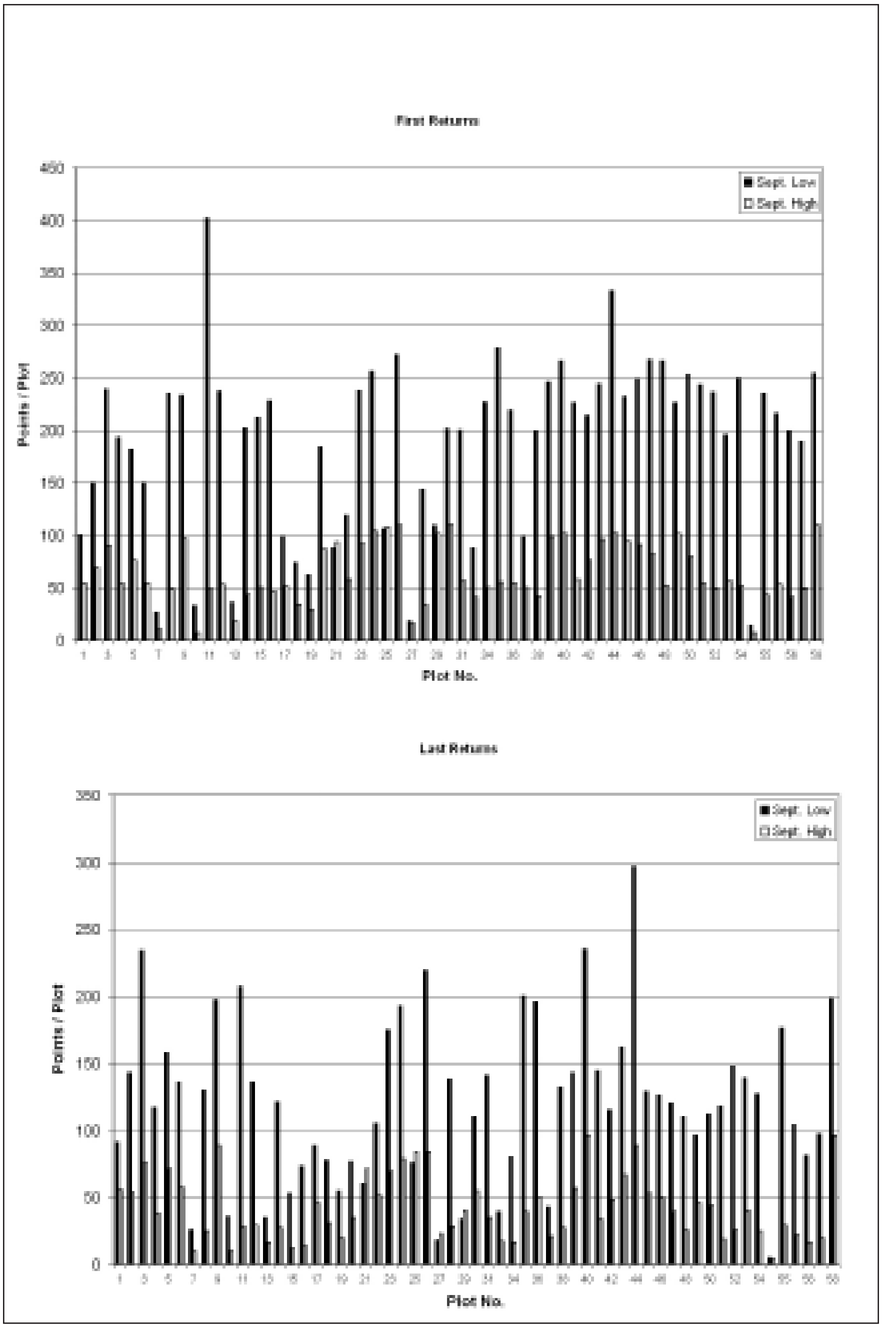

Fig. 2. Sampling point density per plot (area = 400 m^2, $z > 2$ m) from each survey and grouped by class of laser return.

cally different from its corresponding metric from the high-altitude survey. To strengthen the comparison, first and last return data for the low- and high-altitude surveys were compared separately, providing 2 samples by which sampling density could be tested. Twenty-three comparisons were carried out for each of the first- and last-return classes. Problems that may arise from multiple comparisons or simultaneous statistical inference are discussed by Miller (1981). In short, as the number of statistical tests performed increase, the probability of making a Type I error increases. For example, assuming the alpha value is 0.05, given a single statistical test, in no more than 1 in 20 tests will a significant difference result when in fact there is none. However, as the number of statistical tests increase to 5 and 10, the probability of making a Type I error increases to 0.22 and 0.40, respectively. Given the large number of paired *t*-tests performed in this study, Bonferroni corrections were applied to control for the increase in Type I errors. Following from the recommendations of Chandler (1995), i.e., to limit the probability of Type II errors, an experiment-wise error rate of 0.15 was applied, which resulted in a Bonferroni-corrected alpha value of 0.007 (i.e., 0.15/23) instead of 0.05.

Results

Canopy height

The results from the paired t-tests for the first and last returns are summarized in Table 4. A graphical representation of canopy height metrics is provided in Fig. 3 for LiDAR first return data. The maximum laser canopy heights (h_{max}) from the low- and high-altitude surveys for the first and last returns were statistically different from each other ($p = 0.0002$ and $p < 0.0001$, respectively). The mean difference (high altitude–low altitude) for the first returns was -0.04 m with a standard deviation of 0.84 m. For the last returns, the mean difference was -0.72 m with a standard deviation of 1.12 m. The minimum laser canopy heights (h_{min}) from each survey for the first and last returns were also statistically different ($p = 0.0015$ and $p = 0.0054$, respectively). The mean difference was 0.83 m for the first returns with a standard deviation of 1.91 m, whereas the last returns were characterized by a mean difference of 1.20 m and a standard deviation 3.20 m.

No significant differences were found for the mean laser canopy heights (h_{mean}) for the first ($p = 0.8733$) and last ($p = 0.5214$) returns. The mean differences were -0.01 m for the first returns and 0.08 m for the last returns. The standard deviations for this metric were 0.40 m and 0.92 m for the first and last returns, respectively. In the case of the coefficient of variation of laser canopy heights (h_{cv}), no significant differences were observed for the first and last returns ($p = 0.3377$ and $p = 0.0101$, respectively). The mean difference for the first returns was 0.61 m with a standard deviation of 4.89 m, whereas the mean difference for the last returns was -2.23 m with a standard deviation of 6.44 m.

As for the deciles of laser canopy heights, there was not a single comparison of first and last returns that resulted in a statistically significant difference (i.e., all $p > 0.1044$ for first returns and all $p > 0.0122$ for last returns). For first returns,

Table 3. Summary of laser canopy height and density metrics grouped by return class

	First returns				Last returns			
	High-altitude Survey		Low-altitude Survey		High-altitude Survey		Low-altitude Survey	
Metric	Range	Mean	Range	Mean	Range	Mean	Range	Mean
h_{max} (m)	3.71–27.88	17.86	4.89–28.84	18.29	3.84–28.36	17.47	4.90–28.92	18.19
h_9 (m)	3.25–26.67	16.16	3.93–26.62	16.20	3.41–26.40	15.97	3.89–26.52	16.17
h_8 (m)	3.21–26.11	15.35	3.60–26.06	15.42	3.26–25.55	15.08	3.60–25.95	15.33
h_7 (m)	3.11–25.68	14.67	3.35–25.50	14.76	2.83–25.23	14.42	3.36–25.26	14.50
h_6 (m)	2.94–25.16	14.14	3.10–25.08	14.18	2.77–24.40	13.48	2.86–24.20	13.69
h_5 (m)	2.86–24.72	13.54	2.74–24.46	13.60	2.68–23.38	12.29	2.73–23.74	12.46
h_4 (m)	2.63–24.15	12.88	2.52–24.04	12.97	2.56–22.24	10.78	2.55–22.2	10.27
h_3 (m)	2.43–23.49	12.08	2.38–23.47	12.13	2.39–21.26	9.30	2.34–21.13	8.80
h_2 (m)	2.24–22.55	10.98	2.21–22.40	10.90	2.24–20.07	7.59	2.27–19.58	7.47
h_1 (m)	2.13–21.41	9.22	2.14–21.39	8.95	2.09–19.53	7.59	2.10–18.81	5.62
h_{min} (m)	2.01–14.75	4.74	2.00–14.65	3.91	2.01–16.66	3.70	2.00–15.63	2.50
h_{mean} (m)	2.78–24.28	13.03	2.98–24.16	13.04	2.87–21.85	11.53	2.97–21.88	11.45
h_{cv} (%)	9.14–56.35	26.52	7.49–55.59	25.90	4.00–72.92	37.06	5.71–77.39	39.29
d_1 (%)	0.93–28.57	7.67	0.00–45.62	11.30	1.12–50.00	14.80	0.67–42.03	15.23
d_2 (%)	0.00–28.57	6.28	0.00–24.32	6.09	0.00–44.83	9.38	0.00–28.41	8.96
d_3 (%)	0.00–29.35	7.21	0.00–28.57	7.26	0.00–30.00	7.65	0.00–34.55	7.31
d_4 (%)	0.00–28.57	7.21	0.00–29.41	7.28	0.00–25.64	7.14	0.00–24.57	7.40
d_5 (%)	0.00–25.49	8.02	0.00–22.45	7.51	0.00–26.09	6.74	0.00–25.00	6.15
d_6 (%)	0.00–26.00	8.58	0.00–21.89	9.66	0.00–23.08	5.85	0.00–15.73	5.36
d_7 (%)	0.00–24.05	11.57	0.99–31.13	13.03	0.00–30.00	8.09	0.00–29.41	6.25
d_8 (%)	0.00–45.74	16.08	0.00–41.22	17.91	0.00–50.00	11.05	0.00–38.98	11.32
d_9 (%)	0.00–54.17	16.68	0.00–39.83	16.17	0.00–54.55	14.35	0.00–54.55	18.85
d_{10} (%)	1.27–40.00	10.70	0.67–40.25	8.16	1.32–57.14	14.94	0.72–53.57	13.18

the mean differences ranged from -0.09 m to 0.27 m whereas the standard deviations for these differences ranged from 0.37 m to 1.67 m. The greatest mean differences and largest standard deviations for first returns tended to be observed for the lower deciles (e.g., h_2 and h_1). The range of mean differences for the last returns was -0.25 m to 0.51 m. The standard deviations of the mean differences for the last returns ranged from 0.61 m to 2.67 m. In general, the mean differences and standard deviations for the decile metrics were greater for the last return data than for the first return data.

Canopy density

For the first return canopy density metrics, comparisons revealed 2 metrics that were statistically different. The d_1 metric was statistically different across the 2 surveys ($p = 0.0001$) with a mean difference of -3.63% and a standard deviation of 6.97%. The second metric that was significantly different was the d_{10} metric ($p = 0.0069$). Its mean difference and corresponding standard deviation was 2.55% and 6.96%, respectively. No significant differences were observed for the remaining first-return canopy density metrics (i.e., $d_2 \ldots d_9$) (all $p > 0.0469$). The mean differences for the remaining canopy density metrics ranged from -1.83% to 0.51%, whereas the standard deviation of these differences ranged from 4.52% to 8.49%.

In the case of the last-return canopy density metrics, only the d_9 metric from the 2 surveys was significantly different ($p = 0.0048$). The mean difference of d_9 was -4.50% with a standard deviation of 11.77%. The remaining canopy density metrics (i.e., $d_1 \ldots d_8$ and d_{10}) were not significantly different for the 2 surveys (all $p > 0.0707$). The range of these remaining metrics was -0.43% to 1.85% and the standard deviation of the mean differences range from 4.92% to 13.39%.

Discussion

The issue of whether or not different sampling point densities influence distributions of canopy height measurements, specifically canopy height profiles, is a contentious one. Some would argue that it is counterintuitive to examine sampling point density because if it is found to influence how canopy height measurements are derived, then one is unable to identify whether it is sampling point density itself or one of the many survey parameters (e.g., pulse repetition frequency, scan angle, scan frequency, aircraft velocity), which collectively determine sampling point density, that is the contributing factor. What is often not realized is that there exists trade-offs between individual sensor and flight configuration-based data collection parameters. Such trade-offs make attempts to assess the effects of individual survey parameters on canopy height measurements challenging due to the need for experi-

Table 4. Mean differences (D; high altitude–low altitude) between laser canopy height and density metrics for high and low altitude surveys and the standard deviation of those differences grouped by return class

Metric[a]	*D*, first returns			*D*, last returns		
	Mean	Std. Dev.	*p*-value[b]	Mean	Std. Dev.	*p*-value[b]
h_{max} (m)	-0.04	0.84	0.0002*	-0.72	1.12	<0.0001*
h_9 (m)	-0.03	0.47	0.6100	-0.21	0.61	0.0122
h_8 (m)	-0.07	0.45	0.2490	-0.25	0.84	0.0271
h_7 (m)	-0.09	0.42	0.1044	-0.09	0.90	0.4697
h_6 (m)	-0.05	0.37	0.3125	-0.22	2.04	0.4165
h_5 (m)	-0.06	0.47	0.3077	-0.17	1.71	0.4558
h_4 (m)	-0.09	0.49	0.1650	0.51	2.37	0.1024
h_3 (m)	-0.04	0.64	0.5973	0.51	2.17	0.0791
h_2 (m)	0.08	1.17	0.6033	0.12	1.54	0.5606
h_1 (m)	0.27	1.67	0.2234	0.28	2.67	0.4207
h_{min} (m)	0.83	1.91	0.0015*	1.20	3.20	0.0054*
h_{mean} (m)	-0.01	0.40	0.8733	0.08	0.92	0.5214
h_{cv} (%)	0.61	4.89	0.3377	-2.23	6.44	0.0101
d_1 (%)	-3.63	6.49	0.0001*	-0.43	6.78	0.6304
d_2 (%)	0.19	5.61	0.7949	0.41	7.12	0.6571
d_3 (%)	-0.05	5.27	0.9428	0.34	6.61	0.6904
d_4 (%)	-0.08	4.90	0.9010	-0.26	6.49	0.7602
d_5 (%)	0.51	4.52	0.3855	0.58	6.16	0.4690
d_6 (%)	-1.08	4.81	0.0894	0.49	4.92	0.4429
d_7 (%)	-1.46	5.51	0.0469	1.85	7.70	0.0707
d_8 (%)	-1.83	8.49	0.1027	-0.27	9.33	0.8279
d_9 (%)	0.51	7.74	0.6164	-4.50	11.77	0.0048*
d_{10} (%)	2.55	6.97	0.0069*	1.76	13.39	0.3168

[a] h_{max} = max. laser canopy height (LCH); h_{min} = min. LCH; h_{mean} = mean LCH; h_1 through h_9 = deciles of LCH; h_{cv} = coefficient of variation of LCH; d_1 through d_{10} = canopies densities or proportion of all laser returns in each interval.

[b] Bonferroni-corrected · = 0.007 used for significance testing.

*Significant difference

mental control (e.g., Hopkinson 2007) and the cost of multiple data collections.

For example, in order to maintain an even sampling point density in the across- and along-track flight directions any increase in scan angle necessitates a decrease in scan frequency. Increases in flying altitude increase the laser footprint because laser footprint is a function of flying altitude and beam divergence. Perhaps more importantly, changes in pulse repetition frequency lead to changes in laser pulse power output and the distribution of laser power within the footprint (Chasmer *et al*. 2006, Hopkinson 2007). Clearly, synergies between individual survey parameters exists, but are not often acknowledged.

Evans *et al*. (2001) raise the issue of how the actual patterns of sample points (e.g., Z-shaped from a scanning mirror) can potentially miss the apex of trees in regularly-spaced plantations. Magnussen and Boudewyn (1998) indirectly state that scan angles greater than ±10° will have impacts on tree height measurement. Holmgren *et al*. (2003) simulated the effects of scan angle on tree height measurements. However, their results may be inconclusive, given that they were employing solid geometric models with zero transmittance and ray tracing without factoring for beam divergence. These earlier studies did not have the benefit of current LiDAR technologies engineered for high pulse repetition frequencies and operation at higher altitudes, nor do they consider how overlapping flight lines can compensate for these potential sources of error.

These studies have focused on tree height and not on the vertical distribution of canopy height. The emphasis has been on the direct measurement of the top of the canopy instead of those measurements found within the canopy. In contrast, the approach adopted in this study has been to conceptualize the problem as a statistical sampling problem. LiDAR data characterize canopy structure by sampling the vertical distribution of biomass through the canopy profile/volume. In statistical terms, canopy structure is the population parameter and the laser canopy height returns represent samples from this population. By collecting a sufficient number of laser canopy height returns it is possible to characterize the nature of the canopy height distribution. Increasing sample size will not improve the nature of the normal distribution, with the exception that the tail ends of the distribution may be better described. Analyses that focus on individual trees implicitly deal with the tails of the distributions of canopy height measurements, which are likely more susceptible to individual survey parameter effects.

Here, we adopt the premise that LiDAR is sampling canopy structure (i.e., the population). Previous research has suggested that canopy structure is the foliage component of the canopy sampled (Magnussen and Boudewyn 1998). The approach to the issue presented here is related to the fundamental question of how different sampling intensities affect the characterization of the population of interest. Distributions of canopy height or the collective set of samples can be

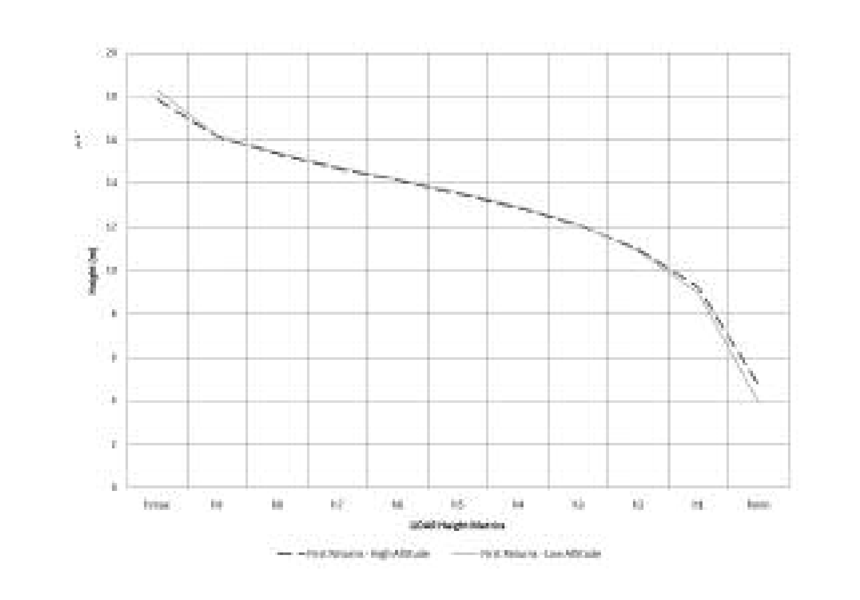

Fig. 3. Canopy height metrics for LiDAR first returns collected from 2 altitudes (maximum laser canopy height [h_{max}], decile laser canopy heights [i.e., $h_1 \ldots h_9$], and minimum laser canopy height [h_{min}]).

a function of several parameters, which we have chosen to amalgamate into the parameter "sampling point density." Finally, if a sample properly characterizes a population, increasing the sampling intensity will not provide any value-added information, with the exception that high sampling intensity may further characterize the tails of the distribution.

The results presented here support this framework and indicate that the majority of laser canopy height and density metrics from the first and last return classes are insensitive to variable laser sampling point densities produced from 2 different survey configurations and across different types of forest plots. The exceptions observed suggest that the tail ends of the distributions of laser canopy heights (i.e., h_{max} and h_{min}) in both return classes are affected by different sampling point densities. Likewise, the proportion of first returns at the upper and lower range of laser canopy heights (i.e., d_1 and d_{10}) is affected by laser sampling point density. Only the d_9 laser canopy density metric for the last returns was found to be affected by variable laser sampling point density.

While the study by Næsset (2004b) represents a study similar in nature to that presented here, the results from these 2 studies are not directly comparable. First, Næsset (2004b) did not intend to address the issue of how laser sampling point density affects various laser canopy and density metrics. Instead, his study examined how changing the flight altitude and hence, the laser footprint size, affected laser canopy height and density metrics equivalent to the h_{10}, h_{50}, h_{90}, h_{max}, h_{mean}, and h_{cv} metrics considered in this study. Second, while Næsset (2004a) applied Bonferroni corrections to control for Type I errors, the experiment-wise error rate was not adjusted to control for Type II errors as suggested by Chandler (1995), and without the *p*-values reported, direct comparisons of results from the 2 studies are not possible. Goodwin *et al.* (2006) compared normalized canopy height profiles obtained from different flying altitudes for a Eucalyptus forest in Australia. The authors found that at the plot level, there was no significant difference between the relative distribution of LiDAR returns for canopy height profiles derived from 3 different altitudes (1000 m, 2000 m and 300 m), indicating that flying altitude and footprint size do not appear to affect canopy height profile estimation.

The results from this study corroborate the finding of Goodwin *et al.* (2006) and suggest that for forest studies based on plot-level canopy data modelling, a higher laser sampling point density does not necessarily produce data that are "superior" with respect to information content (i.e., canopy height and density metrics). It is postulated that the primary advantage of using a higher laser sampling point density lies with a more precise characterization of the upper and lower forest canopy components and potentially for terrain surface modeling. However, for studies where this type of information is not as important, such as those forest studies that use deciles of laser canopy height and intermediate laser canopy density metrics for estimation of forest biophysical variables, data characterized by a lower laser sampling point density may suffice.

Combining the results of this study with existing related literature, it is becoming apparent that for many stand-level FRI applications sampling density specifications may be less critical than other elements of survey configuration. For example, while sampling density will undoubtedly influence the probability of sampling tree crown apices, parameters such as pulse power, beam divergence and swath overlap can influence the entire frequency distribution. The implications for the forest manager are in terms of identifying which aspects of LiDAR project design actually: a) impact the quality of derived FRI information; and therefore, b) need to be included within a data collection request for proposals or scrutinized within the follow-up tender document. The challenge ahead lies in determining what the optimal laser sampling point density should be for a given type of forest study and/or application.

Conclusions

This study examined how 2 sampling point densities across different types of forest plots affected 23 laser canopy height and density metrics typically used in studies that focus on the estimation of forest biophysical variables at the plot-level. Nonetheless, only 4 (i.e., h_{max}, h_{min}, d_1, d_{10}) of the 23 laser canopy height and density metrics were found to be significantly different across the 2 surveys when the first returns were considered. Likewise, for the last return data, only 3 laser canopy height and density metrics (i.e., h_{max}, h_{min}, d_9) were statistically different from one another. These laser canopy height and density metrics corresponded to metrics characterizing the tails of the distributions of laser canopy heights. This suggests that plot-level data characterized by a higher laser sampling point density do not necessarily equate to data that are richer in information content. Consequently, higher sampling point densities may add little value to current LiDAR forest research at the stand level as these studies tend to focus on LiDAR returns found within the canopy profile as opposed to at its extremes. Combined with research indicating that laser pulse power is a key parameter in controlling the representation of forest canopy attributes (e.g., Chasmer *et al.* 2006, Hopkinson 2007), the observations here suggest that sampling density alone is not a critical aspect of LiDAR survey design for operational stand-level FRI attribute extraction. Therefore, it may be feasible to obtain suitable estimates of forest inventory variables using LiDAR data collected at low sampling densities, and hence at reduced cost. This study helps support the need for a systematic collection and analysis of multiple sampling densities for a range of forest ecosystems to determine appropriate sampling rates for accurate and precise estimation of forest structural variables required by forest managers in government and industry. This work is currently being conducted by a team of researchers from government, industry and academia at study sites in the Great Lakes – St. Lawrence and Boreal forests of Ontario.

Acknowledgements

Optech Incorporated is thanked for carrying out the LiDAR surveys as well as Otterburn Geographic for providing logistical support. This research is a part of a larger project initiated by the Canadian Consortium for LiDAR Environmental Applications Research (C-CLEAR). Dr. P. Treitz of the Laboratory for Remote Sensing of Earth and Environmental Systems, Queen's University acknowledges research support from the Natural Sciences and Engineering Research Council (NSERC), the Premier's Research Excellence Award (PREA) and the Ontario Centres of Excellence for Earth and Environmental Technologies. Funding for K. Lim was provided by the NSERC and the Ontario Government in the form of a Postgraduate Scholarship (PGS) and an Ontario Graduate Scholarship in Science and Technology (OGSST). Mr. M. Sitar and Ms. L. Chasmer are acknowledged for their assistance with data processing and Mr. L. Munt, York Regional Forest Manager, for permitting research to be carried out in the Vivian Forest. The authors would like to thank Murray Woods, the anonymous reviewers and the Associate Editor for their valuable comments that improved the quality of this manuscript.

References

Andersen, H.-E., R.J. McGaughey and S.E. Reutebuch. 2005. Estimating forest canopy fuel parameters using LIDAR data. Remote Sensing of Environment 94: 441–449.

Bollandsås, O.M. and E. Næsset. 2007. Estimating percentile-based diameter distributions in uneven-sized Norway spruce stands using airborne laser scanner data. Scandinavian Journal of Forest Research 22: 33–47.

Chandler, C.R. 1995. Practical considerations in the use of simultaneous inference for multiple tests. Animal Behaviour 49: 524–527.

Chasmer, L., C. Hopkinson and P. Treitz. 2006. Examining the influence of changing laser pulse repetition frequencies on conifer forest canopy returns. Photogrammetric Engineering and Remote Sensing 72(12): 1359–1367.

Evans, D.L., S.D. Roberts, J.W. McCombs and R.L. Harrington. 2001. Detection of Regularly Spaced Targets in Small-footprint LIDAR Data: Research Issues for Consideration. Photogrammetric Engineering and Remote Sensing 67(10): 1133–1136.

Falkowski, M.J., A.M.S. Smith, A.T. Hudak, P.E. Gessler, L.A. Vierling and N.L. Crookston. 2006. Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of LiDAR data. Canadian Journal of Remote Sensing 32(2): 153–161.

Goodwin, N.R., N.C. Coops and D.S. Culvenor. 2006. Assessment of forest structure with airborne LiDAR and the effects of platform altitude. Remote Sensing of Environment 103: 140–152.

Holmgren, J. 2004. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. Scandinavian Journal of Forest Research 19: 543–553.

Holmgren, J., M. Nilsson and H. Olsson, 2003. Simulating the effects of LiDAR scanning angle for estimation of mean tree height and canopy closure. Canadian Journal of Remote Sensing 29(5): 623–632.

Hopkinson, C. 2007. The influence of flying altitude and beam divergence on canopy penetration and laser pulse return distribution characteristics. Canadian Journal of Remote Sensing 33(4): 312–324.

Hopkinson, C., L. Chasmer and R. Hall. 2008. Monitoring inter-annual red pine conifer growth using airborne LiDAR. Remote Sensing of Environment 112: 1168–1180.

Hopkinson, C., L. Chasmer, K. Lim, P. Treitz and I. Creed. 2006. Towards a universal LiDAR canopy height indicator. Canadian Journal of Remote Sensing 32(2):139–152.

Hopkinson, C., L.E. Chasmer, C. Young-Pow and P. Treitz. 2004a. Assessing plot-level forest metrics with a ground-based scanning LiDAR. Canadian Journal of Forest Research 34: 573–583.

Hopkinson, C., M. Sitar, L. Chasmer and P. Treitz. 2004b. Mapping snowpack depth beneath forest canopies using airborne LiDAR. Photogrammetric Engineering & Remote Sensing 70: 323–330.

Hug, C., A. Ullrich and A. Grimm. 2004. LiteMapper-5600 – A waveform-digitizing LiDAR terrain and vegetation mapping system. Proceedings of the ISPRS Working Group VIII/2 Laser-Scanners for Forest and Landscape Assessment, Freiburg, Germany (International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXVI, Part 8/W2). pp. 24–29.

Lefsky, M., W. Cohen, G. Parker and D. Harding. 2002. LiDAR remote sensing for ecosystem studies. Bioscience. 52: 19–30.

Lefsky, M., D. Harding, W. Cohen, G. Parker and H. Shugart. 1999. Surface LiDAR remote sensing of basal area and biomass in deciduous forests of Eastern Maryland, USA. Remote Sensing of Environment 67: 83–98.

Lim, K. and P. Treitz. 2004. Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. Scandinavian Journal of Forest Research. 19: 558–570.

Lim, K., P. Treitz, K. Baldwin, I. Morrison and J. Green. 2003a. LiDAR remote sensing of biophysical properties of tolerant northern hardwood forests. Canadian Journal of Remote Sensing 29: 658–678.

Lim, K., P. Treitz, M. Wulder, B. St-Onge and M. Flood. 2003b. LiDAR remote sensing of forest structure. Progress in Physical Geography 27: 88–106.

Magnussen, S. and P. Boudewyn. 1998. Derivation of stand heights from airborne laser scanner data with canopy-based quantile estimators. Canadian Journal of Forest Research 28: 1016–1031.

Magnussen, S., P. Eggermont and V. LaRiccia. 1999. Recovering tree heights from airborne laser scanner data. Forest Science 45: 407–422.

Maltamo, M., P. Packalén, X. Yu, K. Eerikäinen, J. Hyyppä and J. Pitkänen. 2005. Identifying and quantifying structural characteristics of heterogeneous boreal forests using laser scanner data. Forest Ecology and Management 216: 41–50.

Means, J., S. Acker, J. Brandon, M. Renslow, L. Emerson and C. Hendrix. 2000. Predicting forest stand characteristics with airborne scanning LiDAR. Photogrammetric Engineering and Remote Sensing 66: 1367–1371.

Means, J., S. Acker, D. Harding, J. Blair, M. Lefsky, W. Cohen, M. Harmon and W. McKee. 1999. Use of large-footprint scanning airborne LiDAR to estimate forest stand characteristics in the western Cascades of Oregon. Remote Sensing of Environment 67: 298–308.
Miller, R. 1981. Simultaneous statistical inference. Springer-Verlag, New York.
Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80: 88–99.
Næsset, E. 2004a. Accuracy of forest inventory using air borne laser scanning: evaluating the first Nordic full-scale operational project. Scandinavian Journal of Forest Research 19(6): 554–557.
Næsset, E. 2004b. Effects of different flying altitudes on biophysical stand properties estimated from canopy height and density measured with a small-footprint airborne scanning laser. Remote Sensing of Environment 91: 243–255.
Næsset, E. and K.O. Bjerknes. 2001. Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. Remote Sensing of Environment 78: 328–240.
Næsset, E. and T. Økland. 2002. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. Remote Sensing of Environment 79: 105–115.
Næsset, E., T. Gobakken, J. Holmgren, H. Hyyppä, J. Hyyppä, M. Maltamo, M. Nilsson, H. Olsson, Å. Persson and U. Söderman. 2004. Laser scanning of forest resources: the Nordic experience. Scandinavian Journal of Forest Research 19: 482–499.
Nelson, R. 1997. Modeling forest canopy heights: the effects of canopy shape. Remote Sensing of Environment 60: 327–334.
Nelson, R., R. Oderwald and G. Gregoire. 1997. Separating the ground and airborne laser sampling phases to estimate tropical forest basal area, volume, and biomass. Remote Sensing of Environment 60: 311–326.
Nelson, R., A. Short and M. Valenti. 2004. Measuring biomass and carbon in Delaware using an airborne profiling LiDAR. Scandinavian Journal of Forest Research 19: 500–511.
Nilsson, M. 1996. Estimation of tree heights and stand volume using an airborne LiDAR system. Remote Sensing of Environment 56: 1–7.
Optech Inc. 2006. Optech's premier Multipulse ALTM Gemini now delivering 167 kHz! [online]. Available at http://www.optech.ca/press_ALTM_Gemini-167.htm [accessed 21 June 2007].
Ritchie, J.C., D.L. Evans, D. Jacobs, J.H. Everitt and M.A. Weltz. 1993. Measuring canopy structure with an airborne laser altimeter. Transactions of the ASAE 36: 1235–1238.
Thomas, V., P. Treitz, J.H. McCaughey and I. Morrison. 2006. Mapping stand-level forest biophysical variables for a mixedwood boreal forest using LiDAR: an examination of scanning density. Canadian Journal of Forest Research 36: 34–47.
Thomas, V., P. Treitz, J.H. McCaughey, T. Noland, and L. Rich. 2007. Canopy chlorophyll concentration estimation using hyperspectral and LiDAR data for a boreal mixedwood forest in northern Ontario, Canada. International Journal of Remote Sensing 29(4):1029–1052.
Tickle, P.K., A. Lee, R.M. Lucas, J. Austin, and C. Witte. 2006. Quantifying Australian forest floristics and structure using small footprint LiDAR and large scale aerial photography, Forest Ecology and Management. 223: 379–394.