

The influence of flying altitude, beam divergence, and pulse repetition frequency on laser pulse return intensity and canopy frequency distribution

Chris Hopkinson

Abstract. Eight airborne light detection and ranging (lidar) data collections were carried out over a forested and agricultural study site in Nova Scotia during 2005. The influences of flying altitude, beam divergence, and pulse repetition frequency on laser pulse return intensities and vertical frequency distributions within vegetated environments were investigated. Experimental control was maintained by varying each survey configuration setting independently while keeping all other settings constant. The land covers investigated were divided into highway, tall vegetation (mature and immature mixed wood regeneration stands), and short vegetation (hay field and potato crop). Laser pulse return data for 24 tall and 18 short vegetation plots were extracted, and the quartile heights of each vegetation profile were compared for each configuration. Observed laser pulse intensity values were found to be linearly related (coefficient of determination $r^2 = 0.98$) to the peak pulse power concentration. A simple routine was developed to allow intensity data to be normalized and made comparable across datasets. By comparing the intensity and laser pulse return profiles it was found that reducing the peak pulse power concentration by widening the beam, increasing the flying altitude, or increasing the pulse repetition frequency tends to lead to (i) slightly reduced penetration into short canopy foliage by up to 4 cm, and (ii) increased penetration into tall canopy foliage (i.e., reduced maximum canopy return heights) by 15–61 cm. It is believed that a reduction in peak pulse power concentration delays pulse triggering within vegetation (i.e., increases penetration of the pulse into foliage) due to the need for increased surface area backscatter to raise the return pulse energy above some minimal threshold within the timing electronics of the sensor. Exceptions to these general observations were found in the high pulse repetition frequency data, where increased sample point density results in (i) increased noise and height range in the lidar distribution data, and (ii) increased likelihood of ground returns in the tall canopies sampled due to increased probability of pulses encountering canopy gaps. The implications of these results are that (i) laser pulse peak power concentration is the largest determinant of pulse return intensity and survey configuration based variations in canopy frequency distribution, and (ii) laser pulse height- and intensity-based models developed for vegetation structural or biomass assessment could be improved if they accounted for variations in peak power concentration.

Résumé. Huit campagnes d'acquisition de données lidar (« light detection and ranging ») aéroportées ont été réalisées au-dessus d'un site d'étude caractérisé par un couvert forestier et agricole en Nouvelle-Écosse, en 2005. Les influences de l'altitude de vol, de la divergence du faisceau et de la fréquence de récurrence des impulsions sur l'intensité des retours d'impulsion laser et les distributions de fréquence du balayage vertical au sein des environnements végétalisés ont été analysées. Un contrôle expérimental a été assuré en faisant varier de façon indépendante chaque réglage de configuration de relevé tout en maintenant les autres réglages constants. Les couverts analysés étaient constitués de route, ainsi que de végétation haute (peuplements matures et immatures de forêt mixte en régénération) et basse (foin et pomme de terre). Les données de retour d'impulsion laser ont été extraites pour 24 parcelles de végétation haute et 18 parcelles de végétation basse et les quartiles des hauteurs de chaque profil de végétation ont été comparés pour chacune des configurations. Les valeurs observées d'intensité d'impulsion laser se sont avérées linéairement corrélées (coefficient de détermination $r^2 = 0,98$) à la concentration de la puissance de crête des impulsions. Une routine simple a été développée pour permettre de normaliser les données d'intensité et les rendre comparables à travers les ensembles de données. En comparant les profils d'intensité et de retours d'impulsion laser, on a pu observer que la réduction de la concentration de la puissance de crête des impulsions, soit en élargissant la largeur du faisceau, en augmentant l'altitude de vol ou en augmentant la fréquence de récurrence des impulsions tend à entraîner : (i) un pouvoir de pénétration légèrement inférieur dans le feuillage de couvert bas pouvant atteindre jusqu'à 4 cm; (ii) un pouvoir de pénétration supérieur dans le feuillage de couvert haut (c.-à-d., hauteurs maximales des retours du couvert réduites) de 15 cm à 61 cm. Il y a lieu de croire qu'une réduction de la concentration de la puissance de crête des impulsions entraîne un délai dans le déclenchement de l'impulsion à l'intérieur de

Received 31 October 2006. Accepted 6 June 2007. Published on the *Canadian Journal of Remote Sensing* Web site at <http://pubs.nrc-cnrc.gc.ca/cjrs> on 23 August 2007.

C. Hopkinson. Applied Geomatics Research Group, Centre of Geographic Sciences, NSCC Annapolis Valley Campus, Middleton, NS BOS 1P0, Canada (e-mail: chris.hopkinson@nsc.ca).

la végétation (c.-à-d., augmente la pénétration de l'impulsion dans le feuillage) dû à la nécessité d'une plus grande surface de rétrodiffusion pour augmenter l'énergie de l'impulsion de retour au-dessus d'un seuil minimal à l'intérieur des composantes électroniques de chronométrage du capteur. On a noté des exceptions par rapport à ces observations générales dans les données de fréquence élevée de récurrence des impulsions où la densité accrue des points d'échantillonnage résulte en : (i) un accroissement du bruit et de l'amplitude des hauteurs dans les données de distribution lidar; et (ii) une probabilité accrue de retours terrestres dans les couverts hauts échantillonnés dû à la probabilité accrue que les impulsions rencontrent des trous dans le couvert. Les implications de ces résultats sont à l'effet que : (i) la concentration de la puissance de crête des impulsions laser constitue le plus grand déterminant des variations basées sur l'intensité du retour d'impulsion et la configuration du relevé dans la distribution des fréquences du couvert; et (ii) les modèles basés sur la hauteur et l'intensité de l'impulsion laser développés pour l'évaluation de la structure ou de la biomasse de la végétation pourraient être améliorés s'ils tenaient compte des variations de la concentration de la puissance de crête.

[Traduit par la Rédaction]

Introduction

Airborne light detection and ranging (lidar) mapping sensors combine four primary subsystems: (i) the global positioning system (GPS) to monitor the location of the airborne platform; (ii) an inertial measurement unit (IMU) to measure the orientation of the lidar sensor relative to the earth surface; (iii) a lidar unit that utilizes knowledge of the speed of light and precision timing electronics to measure the distance between the sensor and the ground surface; and (iv) a scanning system to redirect emitted laser pulses across the line of flight, resulting in a swath of laser pulse return (LPR) survey points beneath the aircraft. Current technology can collect multiple laser pulse returns at pulse repetition frequencies (PRF) up to and exceeding 150 kHz and can cover a ground swath greater than 3000 m depending on flying altitude and scan angle. The resultant LPR positional accuracy is typically at the decimetre level, with horizontal errors normally exceeding vertical errors and increasing with an increase in flying altitude (Baltasvias, 1999). For a calibrated airborne laser terrain mapper (ALTM) model 3100 (Optech Incorporated, Toronto, Ontario) operated during optimal GPS conditions and flown over both vertical and horizontal calibration features, the author typically observes vertical standard deviations between 6 and 10 cm over runway ground control points (GCPs) and horizontal standard deviations between 10 and 30 cm over break line control features.

Many studies using small footprint discrete return lidar data have demonstrated strong empirical relationships between LPR vertical frequency distribution metrics and vegetation height (see Lim et al., 2003a for a summary). Although much recent research has focused on individual tree height estimation, most attention has been on comparing plot-level tree heights with some LPR-derived height metric. For example, Naesset (1997) found that for conifer stands ranging in height from 8 to 24 m, maximum LPR heights above the ground level correlated well with Lorey's mean tree height for a given area. This work was expanded upon by Magnussen and Boudewyn (1998) by investigating a canopy LPR quantile-based approach for estimating height for conifer plots ranging in height from 15 to 27 m. Similar LPR metrics were combined with LPR intensity values to estimate height and other biometric properties of shade-tolerant hardwood plots of varying treatment and ranging in height from 10 to 30 m (Lim et al., 2003b).

Few studies have investigated the estimation of short (near ground surface) vegetation height from airborne lidar. The work of Davenport et al. (2000) and Cobby et al. (2001) demonstrated that crop vegetation up to approximately 1.2 m in height could be predicted from the standard deviation of topographically detrended laser pulse returns. Recent research has attempted to develop universally applicable (i.e., vegetation type, height, and data acquisition configuration independent) canopy height estimation techniques using the standard deviation of the vertical LPR canopy return frequency distribution (Hopkinson et al., 2006).

An obvious challenge to the development of vegetation structural and biomass models from LPR frequency distributions is that the shape of the distribution through the canopy can be influenced by a number of environmental and system-dependent factors. For example, it is known that canopy lidar returns tend to penetrate into foliage rather than backscatter from the outer extremities of the canopy surface (Gaveau and Hill, 2003). However, the amount of penetration and the overall distribution will be influenced by vegetation structural characteristics and lidar data acquisition factors such as pulse repetition frequency (Chasmer et al., 2006), flying altitude (Naesset, 2004; Yu et al., 2004; Goodwin et al., 2006), beam divergence (Andersen et al., 2006), and scan angle (Holmgren et al., 2003).

Further, given that researchers are beginning to explore the use of LPR intensity data to refine lidar-derived vegetation species, structural, and biomass models (e.g., Lim et al., 2003a; Holmgren and Persson, 2004), it is useful to investigate the controls on intensity. Some initial work on the radiometric correction of laser intensity based on atmospheric attenuation of the pulse, pulse geometry, and ground surface conditions has been presented by Coren and Sterzai (2006). However, although the surface reflectance and surface area encountered by an emitted laser pulse will exert a strong influence on the relative amount of energy returned to the sensor, the overall control on the magnitude of the return will be the strength of the pulse energy emitted. Sensor and flight settings directly control the energy and size of the emitted pulse, and the recorded intensity is directly related to the magnitude of the returned peak of the recorded backscatter (B. Smith, Optech Incorporated, Toronto, Ontario, personal communication).

This paper expands on previous research in three important ways: (i) by systematically investigating altitude, beam

divergence, and PRF so that the observations can be synthesized to create generally applicable rules; (ii) by investigating the systematic controls on LPR intensity; and (iii) by performing the analysis on both short (<2 m) and tall (>2 m) vegetation plots to better understand the pulse penetration properties in single versus multiple return foliage profile environments.

Study area

The study was conducted over a flat to rolling valley site (<50 m total elevation variation) near Nictaux in the Acadian forest ecozone of Nova Scotia (Figure 1). The study area was less than 1 km wide by approximately 2 km long and comprised a number of common land cover types for this region: predominantly Acadian mixed woodland (mostly yellow birch, *Betula alleghaniensis* Britton, with some mixed pine, *Pinus*, and mixed spruce, *Picea*, trees) and agricultural crops. The land covers investigated were divided into tall (>2 m) and short (<2 m) vegetation classes of mixed forest and birch regeneration and hay field and potato crop, respectively. Short and tall vegetation classes were divided because tall vegetation has the potential to trigger multiple returns, and short vegetation is typically within the range resolution (~2 m) for multiple return triggering, and consequently the shape of canopy LPR distributions can be substantially different. Using this criterion, short vegetation frequency distributions are typically unimodal, whereas those for tall vegetation are

typically bimodal or multimodal. See Hopkinson et al. (2005; 2006) for examples and further discussion of the LPR frequency distribution characteristics of tall and short vegetation classes.

This study area was chosen as it is within 5 km of the calibration site of the lidar sensor used in this study and it was therefore logistically straightforward to acquire multiple survey configurations within a short time period. Also, the site is the subject of ongoing lidar and agroforestry experiments, for which supplemental ground control, plot mensuration, and digital hemispherical photography (DHP) data exist. For example, Hopkinson et al. (2006) utilized vegetation height data from some of the same forested stands and agricultural fields illustrated in Figure 1 to develop a robust lidar canopy height modeling technique. A summary is provided of the canopy height data that were used in Hopkinson et al. and are representative of some of the plots used in this study (Table 1). Summertime leaf area index (LAI) values for the mixed wood and birch regeneration stands, as derived using DHP techniques (Leblanc et al., 2005), vary between 1.7 and 2.1 (unpublished field data collected in May and August 2006).

Methods

Ground control data

To validate the lidar data and to correct for the influence of airborne GPS trajectory bias in the data, over 100 GPS ground control points were collected on 20 July 2005. GCPs were

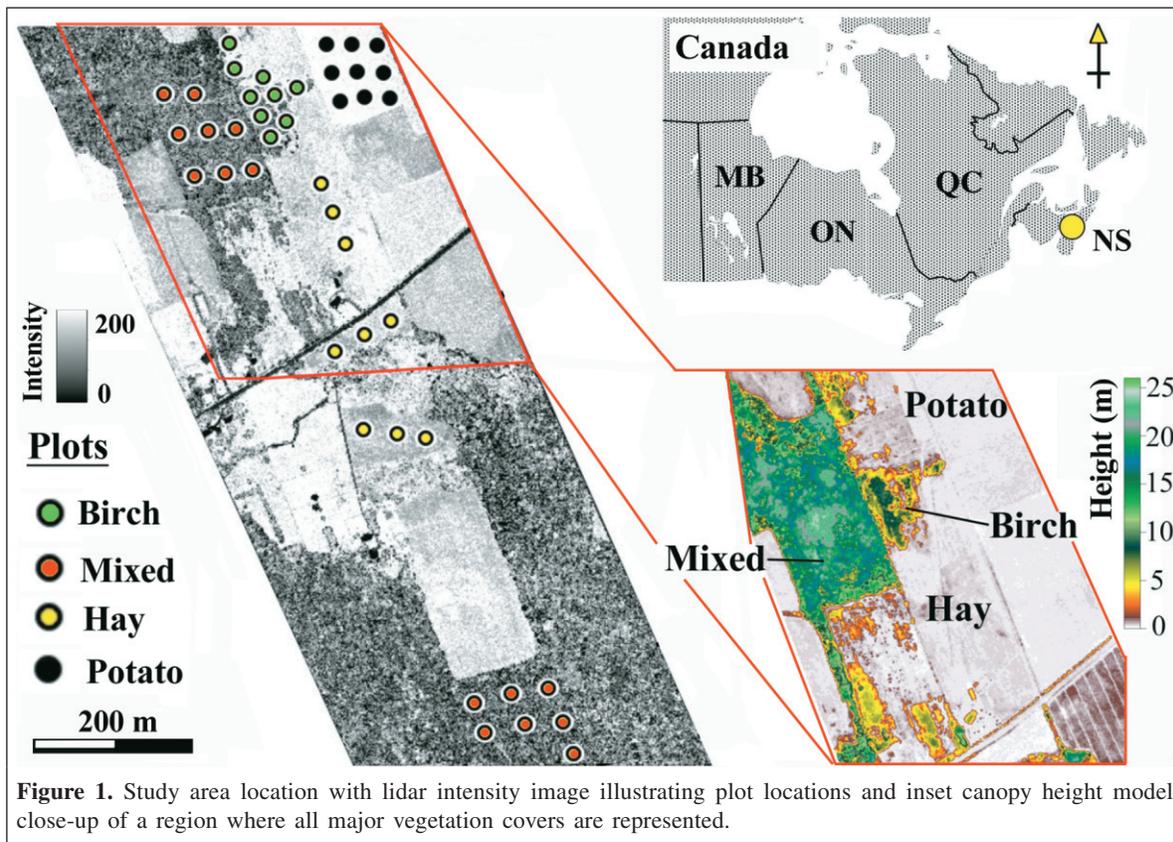


Figure 1. Study area location with lidar intensity image illustrating plot locations and inset canopy height model close-up of a region where all major vegetation covers are represented.

Table 1. Vegetation plot height statistics within the study area.

Height statistic	Hay field	Immature birch	Mixed wood
Avg. (m)	0.39	8.3	17.3
Max. (m)	0.72	13.5	29.6
SD (m)	0.12	1.8	3.7
<i>n</i>	150	250	203

Note: There are no field data for potato crops, and plots do not correspond exactly with lidar plots. Data originally presented in Hopkinson et al. (2006).

located midway between the centreline and road edge along both sides of the highway running across the entire width of the study polygon (**Figure 1**). Postprocess kinematic (PPK) ground GPS points were collected using a dual-frequency Leica SR530 rover receiver, which was differentially corrected to another Leica SR530 GPS base station receiver mounted on the Applied Geomatics Research Group (AGRG) calibration building less than 5 km from the highway. After differential correction, the standard deviations (σ) of all GCPs were within 1 cm. All LPRs falling within a 1 m radius of a GCP were outputted for elevation residual calculations and intensity assessment.

Airborne lidar data collection and processing

Three separate lidar acquisition flights were carried out for this study. The first flight occurred on 9 April 2005 during dry-ground and leaf-off conditions to provide a baseline ground digital elevation model (DEM). The lidar sensor used was an Optech Incorporated ALTM 3100 owned by the AGRG operating at a wavelength of 1064 nm (near infrared). This first dataset was collected to provide a good quality baseline bald earth DEM. The second flight took place on a clear afternoon during a 3 h period on 16 July 2005 during peak LAI conditions. Four separate configurations were flown: (i) control, (ii) wide-beam divergence, (iii) high altitude, and (iv) high PRF. For each survey, three flight lines with 50% swath sidelap covered the survey polygon (**Figure 1**).

The control survey configuration was flown at 1000 m agl, 0.3 mrad (1/e) beam divergence, 33 kHz pulse repetition frequency, and scan angle $\pm 20^\circ$ from nadir. For the wide-beam survey, the divergence was increased to 0.8 mrad (1/e) while all other survey settings remained constant. Laser pulse peak power was kept constant at 26.8 kW. The only difference this made to the data acquisition was that any pulse incident upon the ground had an increased diameter covering approximately seven times more area, and the associated pulse power concentration was reduced. The high-altitude survey was flown at 3000 m agl, and this had the dual consequence of increasing the footprint area by almost 10 times but also reducing the sampling point density at ground level. The high PRF configuration was set at 100 kHz, with all other settings equivalent to those of the control survey. By changing the PRF, the individual pulse footprint remained the same but sampling

point density increased while peak pulse power decreased (**Table 2**).

The airborne trajectory was differentially corrected to the AGRG GPS base station receiver less than 5 km from the centre of the survey site. Raw lidar ranges and scan angles were integrated with aircraft trajectory and orientation data using PosPAC (Applanix, Toronto, Ontario) and REALM (Optech Incorporated) software tools. The output from this processing procedure was a series of flight-line data files containing x - y - z - I (easting, northing, elevation, intensity) information for each laser pulse return collected.

Following lidar point position computation, the x - y - z - I data files were imported into the Terrascan (Terrasolid, Finland) software package for strip matching, removal of high and low outliers, classification into ground and nonground returns, and plot subsetting. All flight strips were independently compared with the highway validation data control points and adjusted accordingly to remove any systematic bias associated with suboptimal trajectory positioning within individual flight lines. Following LPR points classification, ground elevations determined from the leaf-off April data collection were subtracted from all point elevations to generate "normalized" heights above the ground surface. The density and accuracy of the classified lidar ground returns reduce with increasing vegetation cover (Hopkinson et al. 2005), so it was critical to use a good quality ground DEM collected during leaf-off conditions. Using separate ground classifications for each experimental configuration would have resulted in incomparable results, as the ground DEM (i.e., the datum to which all LPR distributions are referenced) would degrade in accuracy and resolution for the dense foliage datasets. By normalizing all LPR frequency distributions to the same high-quality DEM, the LPR foliage distributions could be directly compared. A minor consequence of uncertain elevation accuracy in the classified lidar ground dataset is that the ground or 0% level on the frequency distribution might be slightly offset from 0.00 m by up to a few centimetres (positive or negative). However, the shape of the distribution would be unaffected and any bias would be the same for both the control and test datasets.

Intensity comparisons

To investigate the influence of peak pulse power concentration on ground-level pulse intensity, a third flight was conducted to collect three additional datasets over the highway control surface in the same study polygon on 22 August 2005. The new sensor settings tested were 2000 m agl and 50 and 70 kHz PRFs. These additional settings allowed a greater spread of peak pulse power concentrations and intensity values to be compared. All highway ground conditions and remaining sensor settings were equivalent to those during the main data acquisition in mid-July. The highway provided a useful ground surface for this type of test, as it has no vertical structural component that could split the returns and possesses a relatively uniform reflectance in the near-infrared portion of the

Table 2. Laser pulse emission and return statistics.

Survey configuration	Peak pulse power (kW)	Footprint area (1/e) (m ²)	Power concentration (kW/m ²)	Avg. intensity	Point density (ppm ²)	Ground cover (%)
Control	21.8	0.07	311	98.0	1.5	12
Wide beam	21.8	0.50	44 (14)	21.0 (22)	1.5	78
High altitude	21.8	0.64	34 (11)	7.4 (8)	0.4	26
PRF	3.7	0.07	53 (17)	16.0 (16)	4.5	32

Note: The values in parentheses denote the return intensity proportion (%) relative to the control dataset.

spectrum. Therefore, most variation in LPR intensity should be due to internal changes to the data acquisition configuration.

The August datasets were comparable to the July datasets over the highway for the purpose of intensity assessment but could not be compared over the vegetated plots due to potential foliage growth and other canopy structural and reflectance changes during the intervening time period. However, the average intensity values for canopy returns from the four land-cover types tested were calculated for each of the July sensor configurations for the purpose of comparison. For mixed wood and immature birch plots, canopy returns were considered to be all those that were at least 2 m above the ground surface. For hay field and potato crop, average intensity was calculated from all returns, as it is difficult to accurately separate ground and foliage returns in areas of short vegetation when only a single return is possible.

Vegetation frequency distribution comparison

In total, 42 circular plots were extracted from each of the normalized lidar datasets, with 18 short vegetation plots (9 hay field and 9 potato crop) and 24 tall vegetation plots (9 immature birch regeneration and 15 Acadian mixed forest). The land covers were selected to represent dominant tall and short end member vegetation classes within the region. The plots were laid out using a semi-grid-based approach (**Figure 1**) and were evenly distributed at 50 m spacings throughout each of the fields or stands of interest. Many of the lidar plots were extracted from the same stands and fields for which field mensuration data already existed (Hopkinson et al., 2006).

Each LPR plot was 100 m² in area and contained between 20 and 500 returns. Frequency distributions were generated for each plot and each survey setting (168 in total), and the following distribution height statistics have been compared for the purpose of this analysis: minimum (0%), quartiles (25% and 75%), median (50%), maximum (100%), range (100%–0%), and average. An illustration of the quartile height comparison for two different point clouds with the associated canopy LPR distributions is provided in **Figure 2**.

Rather than compare entire frequency distributions, it was decided to compare individual distribution statistics and height quantiles because a commonly adopted method for lidar-based height, stem density, and biomass prediction is to use metrics derived from the frequency distribution or ratios of quantile heights as the model input variables (e.g., Magnussen and Boudewyn, 1998; Naesset and Bjerknes, 2001; Lim et al., 2003a). The wide-beam, high-altitude and PRF plot-level

frequency distribution statistics were individually compared with those of the control data to quantify any difference in distribution shape as manifested in the minimum, maximum, quartile, range, and average heights. After performing individual plot-level comparisons, the results were summarized for each of the tall and short vegetation classes, and the significance of any observed height difference was evaluated using a standard two-tailed *t* test. An alternative method for comparing differences between frequency distributions is to use a Kolmogorov–Smirnov (KS) test (e.g., Chasmer et al., 2006). However, this approach was not adopted here because KS is not sensitive to differences in the low-frequency portions of the distribution, i.e., at the distribution maximum or canopy surface level, where much of the difference in the distributions might be expected.

Results and discussion

Highway ground control

The results of the highway validation and control are summarized in **Table 3**. The overall variability in the elevation data, as illustrated by the standard deviation (σ), is between 5 and 11 cm for each of the survey settings, with 100 kHz PRF demonstrating the most variability relative to the control. For both the control and wide-beam divergence surveys, no vertical bias in the data was evident (i.e., σ and root mean square error are equivalent). However, for the high-altitude and high-PRF surveys, there were 7 and 3 cm of positive bias, respectively. The 7 cm mean offset for the high-altitude dataset ($\sigma = 6$ cm) was significantly different from zero at the 99% level of confidence, whereas the high PRF mean offset of 3 cm was not significantly different due to the large random error component ($\sigma = 11$ cm). The associated high-altitude and high-PRF flight strips were vertically registered to the control dataset by subtracting the observed height offset from all LPR elevations.

Laser pulse intensity relationships

From the survey settings and highway control intensity values in **Table 2** it is apparent that ground-level intensities are reduced for the test survey configurations relative to those of the control dataset. Also note that noise (standard deviation) in the data (**Table 3**) is lowest for the control data collection, which also displays the highest pulse power concentration and intensity values (**Table 2**). It is logical to infer, therefore, that as pulse power increases, ranging accuracy and reliability tend to

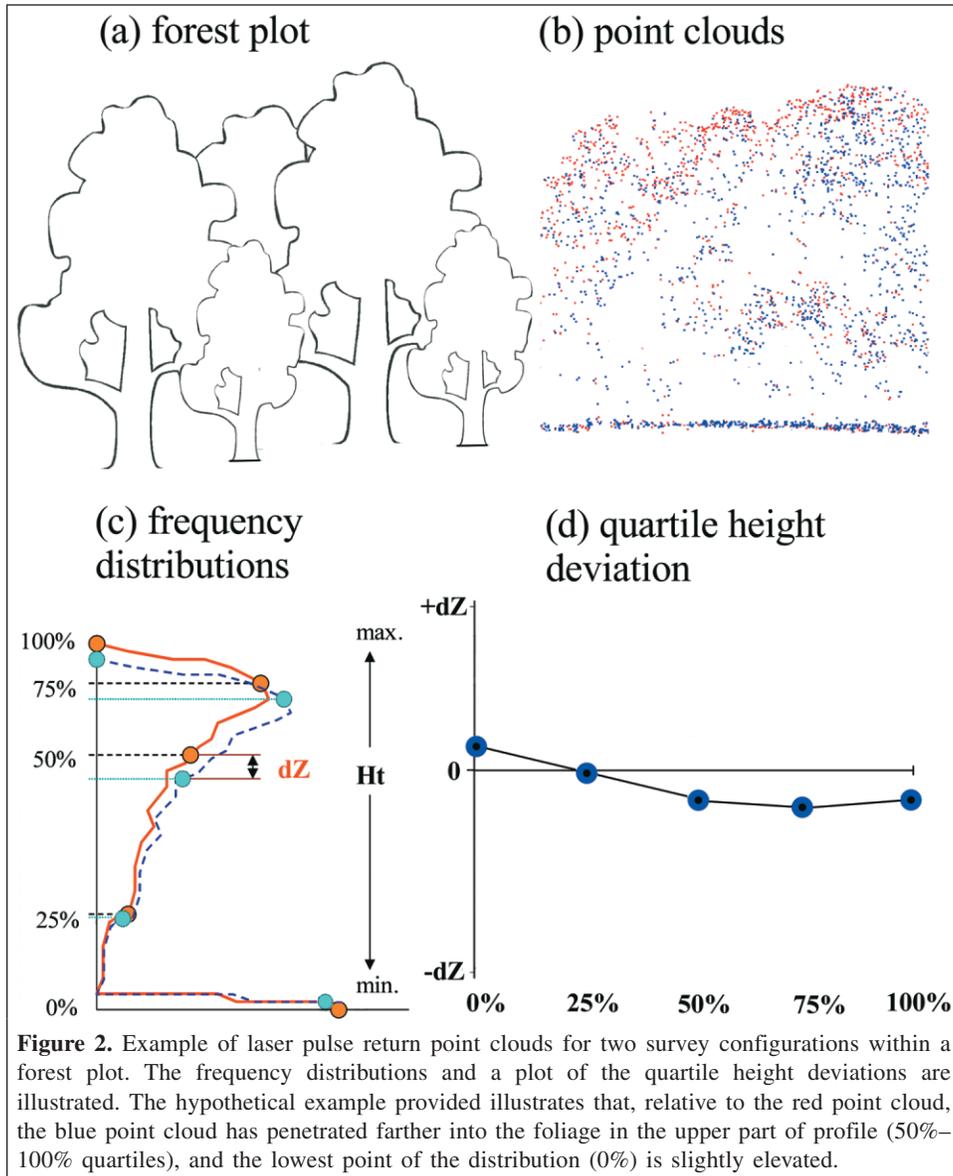


Table 3. Laser pulse return elevation (in m) offset statistics relative to ground control points collected over highway.

Survey configuration	Max.	Min.	Avg.	SD	RMSE
Control	0.12	-0.15	0	0.05	0.05
Wide beam	0.16	-0.21	0	0.07	0.07
High altitude	0.18	-0.10	0.07	0.06	0.12
High PRF	0.43	-0.29	0.03	0.11	0.14

improve as a result of increased intensity and a subsequent reduced signal-to-noise ratio observed at the electro-optical ranging device of the ALTM 3100. However, while it is interesting to consider ranging accuracy variations, the data presented in this regard are mostly anecdotal and insufficient to provide a statistically meaningful generalization of the relationship between intensity and accuracy. Of more immediate interest within the context of evaluating lidar survey

configuration settings on LPR characteristics is the consistency in intensity values across a range of configurations.

By combining the lidar data collected over the highway in July and August, it was possible to relate LPR intensity to the peak power concentration (**Figure 3**), which varied as a function of peak pulse power and footprint area. **Figure 3** demonstrates that when ground reflectance is constant for a given surface material, the pulse intensity (I) recorded by the ALTM 3100 is a linear function of the peak pulse power concentration (C), such that

$$I = f(C) = f\left(\frac{P}{A}\right) \quad (1)$$

where P is the emitted peak pulse power, and A is the Gaussian ($1/e$) footprint area illuminated on the ground. Therefore, if C is known, the observed intensity data (I_{obs}) from one survey can

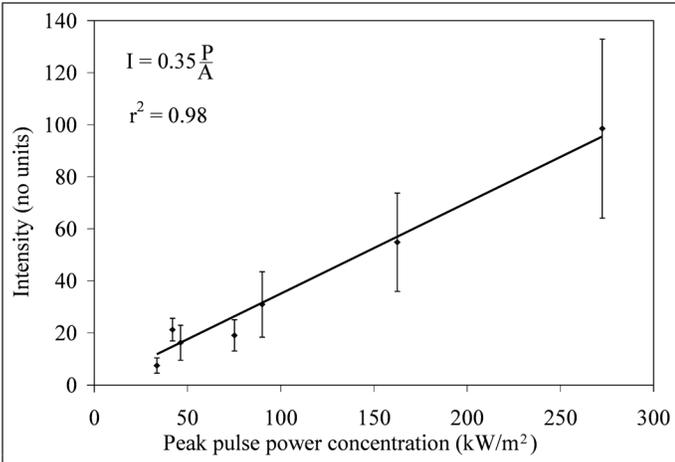


Figure 3. Variation in laser pulse intensity with peak pulse power concentration for seven different survey configurations collected over the highway control surface across the middle of the study site. Error bars illustrate one standard deviation in the observed intensity values.

be corrected (I_{cor}) (i.e., normalized) for comparison with a control or reference intensity dataset using the following equation:

$$I_{cor} = I_{obs} \frac{C_{ref}}{C_{obs}} \quad (2)$$

where C_{ref} and C_{obs} are the reference and observed peak pulse power concentration, respectively. If peak pulse power remains constant, then Equation (2) becomes

$$I_{cor} = I_{obs} \frac{A_{obs}}{A_{ref}} \quad (3)$$

where A_{ref} and A_{obs} are the reference and observed footprint area, respectively. Further, if range (R) or altitude varies for the intensity data to be compared, then footprint area will also vary as a function of range and divergence (D) such that Equation (3) becomes

$$I_{cor} = I_{obs} \frac{\pi \left(\frac{R_{obs} D_{obs}}{2} \right)^2}{\pi \left(\frac{R_{ref} D_{ref}}{2} \right)^2} \quad (4)$$

where R_{ref} and R_{obs} are the reference and observed range, respectively; and D_{ref} and D_{obs} are the reference and observed beam divergence, respectively. Assuming a constant divergence and simplifying, Equation (4) becomes

$$I_{cor} = I_{obs} \frac{R_{obs}^2}{R_{ref}^2} \quad (5)$$

Lastly, assuming natural surfaces are lambertian and the scan angle (θ) represents the mean angle of incidence at ground level, then a cosine correction can be applied such that

$$I_{cor} = I_{obs} \frac{R_{obs}^2}{(R_{ref} \cos \theta)^2} \quad (6)$$

The previous relationships explain how variations in altitude, PRF, divergence, and scan angle will systematically influence the peak pulse power concentration and subsequent return intensity of a laser pulse at ground level. Unlike the work of Coren and Sterzai (2006), the corrections applied here are relative and do not consider atmospheric attenuation, surface roughness, and the true angle of incidence. While inclusion of these parameters would no doubt improve the calibration of intensity data, it can be difficult to obtain accurate estimates of atmospheric attenuation coefficient, and computing the angle of incidence requires integration of the aircraft orientation, scan angle, and local ground slope. The aforementioned models are simple, provide a first approximation at normalizing the laser intensity, and can be computed directly from the lidar data without the need for terrain model iteration. However, the results presented in **Figure 3** and **Table 3** and the equations are based on data collected over flat unambiguous highway control surfaces with no vegetation cover to attenuate the emitted laser pulse energy and split the return backscatter.

In a transmissive canopy environment, it is expected that the intensity and location of laser pulse returns within the canopy will be influenced by pulse power concentration, canopy height, foliage density and orientation, pulse sampling density, and pulse geometry. In **Table 4** the increased near-infrared reflectance of vegetated surfaces is evident in the increased intensity values compared with the black-top highway results in **Table 2**. Short vegetation intensities are almost double those of highway results, whereas tall vegetation values are only approximately 50% greater. One reason for the difference between short and tall vegetation (other than natural variations in surface reflectance) is that tall vegetation attenuates more of the emitted pulse energy, effectively trapping it in the canopy. Also, for features taller than a threshold of approximately 2 m, the ALTM can record multiple returns. Therefore, the intensity of split returns from multiple small surfaces along the pulse path length will tend to be lower than that of single returns from a single larger backscatter surface.

Frequency distribution comparisons

A summary of the survey configuration frequency distribution comparisons for tall and short vegetation is presented in **Table 5**. The average heights above a constant ground reference datum for each distribution quartile, average, and range are given for the control and test data along with the height differences and tail probabilities. P values <0.1 are highlighted because they demonstrate a significant difference in the distribution heights at the 90% confidence level.

Table 4. Average foliage intensity statistics for each vegetation type and survey configuration.

Survey configuration	Short vegetation		Tall vegetation	
	Hay field	Potato crop	Immature birch	Mixed wood
Control	178	196	139	128
Wide beam	50 (28)	59 (30)	29 (21)	23 (18)
High altitude	18 (10)	20 (10)	6 (4)	16 (13)
High PRF	42 (24)	45 (23)	33 (24)	30 (23)

Note: Intensity is given a numeric value related to the voltage recorded within the ALTM timing electronics and has no units. The values in parentheses denote the proportion (%) of the control intensity.

Table 5. Vegetation LPR frequency distribution control dataset height statistics with the observed height deviations for each test dataset.

Class	Avg. plot height distribution statistics							
	Min.	25%	50%	75%	Max.	Avg.	Range	
Control								
Short	Height	-0.06	0.09	0.16	0.22	0.41	0.16	0.47
Tall	Height	1.3	9.3	11.3	13.2	16.7	11.1	15.4
Beam divergence								
Short	dZ	0.02	0.00	0.01	0.03	0.00	0.02	-0.02
	<i>P</i>	0.18	0.85	0.26	0.01	0.98	0.04	0.48
Tall	dZ	1.10	0.09	0.04	0.03	-0.32	0.06	-0.93
	<i>P</i>	0.10	0.89	0.58	0.72	0.05	0.65	0.20
Altitude								
Short	dZ	0.04	0.04	0.03	0.04	0.00	0.04	-0.04
	<i>P</i>	0.05	0.00	0.01	0.00	0.99	0.00	0.20
Tall	dZ	0.65	-0.83	-0.54	-0.47	-0.61	-0.75	-0.55
	<i>P</i>	0.50	0.01	0.00	0.00	0.00	0.00	0.19
PRF								
Short	dZ	-0.11	-0.03	-0.01	0.01	0.06	0.01	0.17
	<i>P</i>	0.00	0.01	0.13	0.40	0.00	0.25	0.00
Tall	dZ	-1.20	-0.45	-0.25	-0.22	-0.15	-0.32	1.10
	<i>P</i>	0.08	0.02	0.01	0.00	0.27	0.00	0.11

Note: *P* values in bold denote that the observed difference between the control and test profile heights is statistically significant at the 90% level of confidence.

From a summary inspection of **Table 5**, it appears that all survey configurations provide a similar LPR distribution description of the foliage profile, and the range of heights recorded is similar to those observed in the field (**Table 1**). At a superficial level, these observations bode well for the widespread application of LPR frequency distribution based models and in part confirm the findings of other related studies (e.g., Naesset, 2004; Goodwin et al., 2006). Of all the configurations tested, the wide-beam divergence results tend to be most closely matched to the control data. However, high-altitude and high-PRF results do demonstrate significant differences between distribution heights for the same plots, and this is likely in part due to the different sample point densities, i.e., increased for high PRF and reduced for high altitude (**Table 2**).

The differences observed in **Table 5** for the short-vegetation frequency distributions tend to be small and, in most cases, below the standard deviation of variability in the raw data. However, for both wide-beam (**Figure 4a**) and high-altitude

(**Figure 4b**) surveys, the average distribution height is shifted upwards relative to the control dataset (2 and 4 cm, respectively). Although these differences are small, they do appear to be consistent, systematic, and, for many of the quartiles tested, significant at the 90% level of confidence. (It is instructive to note that had the highway GCP vertical coregistration of all datasets not been performed, the observed differences for the high-altitude data would be 11 cm instead of the 4 cm reported in **Table 5**.)

No significant differences in distribution height are observed for high altitude and wide-beam divergence at the maximum short vegetation heights, and most of the upshift in the distributions appears to occur from the minimum to the 75% quartile height (3–4 cm). This pattern is generally consistent for both wide-beam and high-altitude tests (**Figures 4a, 4b**) despite the large difference in sample point density (**Table 2**). For short vegetation, therefore, the increase in footprint size leads to a very slightly reduced range (2–4 cm) and poorer distribution representation near to the ground (i.e., reduced

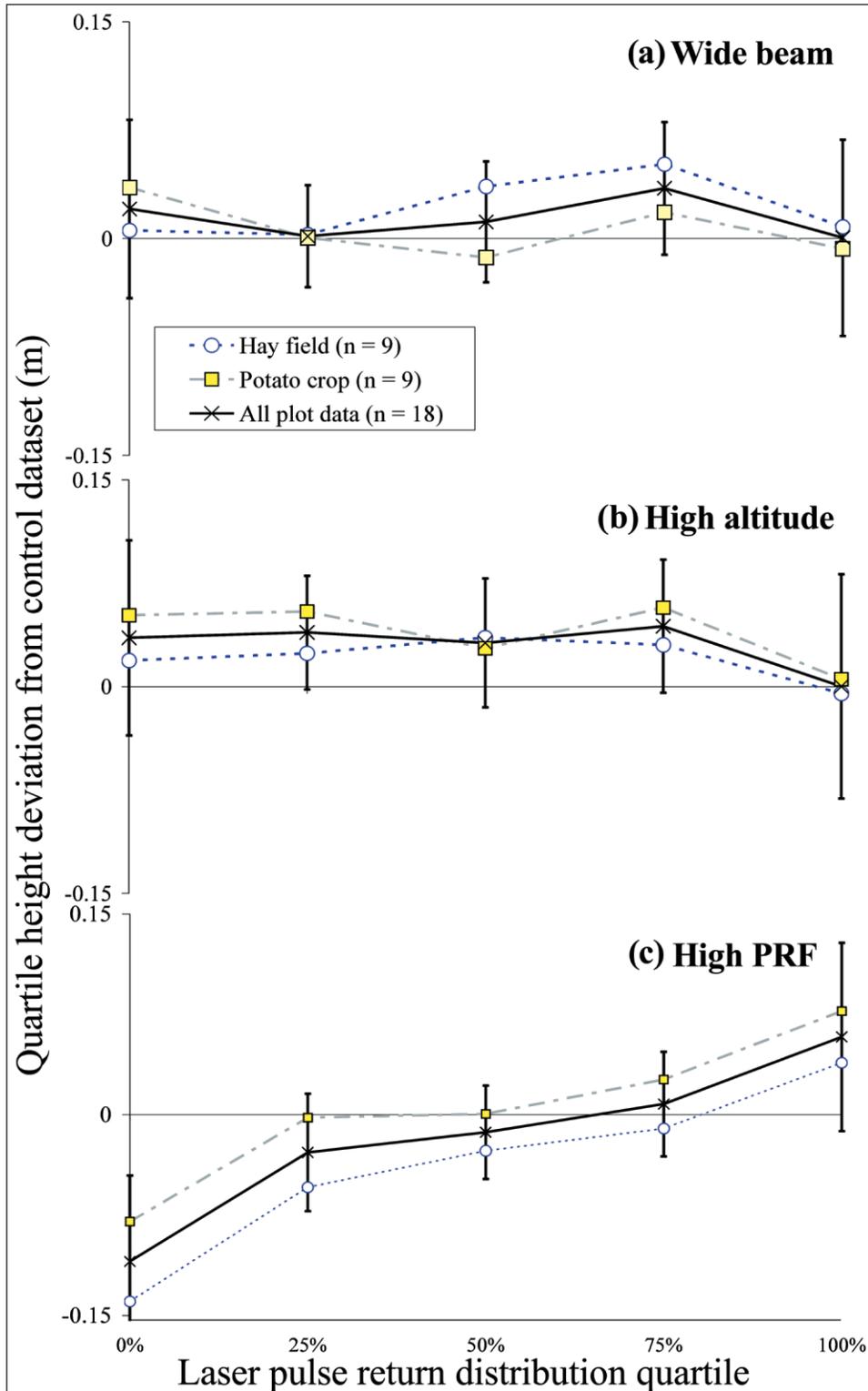


Figure 4. Short vegetation LPR frequency distribution quartile height deviations for each survey configuration tested. The x axis represents the quartile heights of the vertical LPR frequency distribution through the foliage, and the y axis represents the vertical height deviation at each quartile relative to the control dataset. A positive shift in the y axis means that the quartile height for the test data lies above that for the control data. Error bars represent one standard deviation for all plots.

short vegetation penetration capability) while apparently having little effect on the distribution maximum.

The high-PRF results for short vegetation (**Figure 4c**) differ from the high-altitude and wide-beam results in that there is no difference in the average distribution height, but there is a significant and systematic shift in both the upper and lower quartiles. Overall, the distribution range is increased by 17 cm for high-PRF data, and this is because of a reduction in the minimum (0%) height by 11 cm and an increase in the maximum (100%) height by 6 cm. This increased range could be partly related to the threefold increased sampling density associated with 100 kHz data and a commensurate increased representation of taller foliage elements. However, this is not thought to provide a satisfactory explanation, as laser pulses cannot penetrate ground, and from **Table 2** it is clear that while the 100 kHz data are three times denser than the control, the actual area of ground sampled (illuminated by laser pulses) is only 20% larger than that of the high-altitude data and less than half that of the wide-beam data. A simple explanation for the increased distribution range is found by examining the survey control data in **Table 3**. The high-PRF data display more than double the range (>0.7 m) and 50% more variability in elevation compared with any of the other survey configurations. Therefore, the 100 kHz data are more “noisy” (prone to random error) than any of the other acquisition settings, and this noise will tend to be amplified over short vegetation surfaces. Such increases in data noise are important to note because the range and standard deviation within LPR vegetation vertical frequency distributions have been used to predict short vegetation height in a number of studies (e.g., Davenport et al. 2000; Cobby et al., 2001; Hopkinson et al., 2006). It is clear from these observations, then, that range- or variance-based short vegetation canopy models developed using data collected at one PRF might not be valid for another PRF.

Differences in the tall-vegetation distributions are generally much larger due to the increased magnitude of the height distributions. However, as with short vegetation, not all of the differences observed are significant. For both the wide-beam and high-altitude experiments the only observed difference that is both consistent and statistically significant is that with increasing footprint size the maximum distribution height reduces by 32 and 61 cm for wide beam and high altitude, respectively (**Table 5; Figures 5a, 5b**). Given the high-altitude dataset has a reduced LPR point density compared with the wide-beam dataset, it might be fair to speculate that some of this difference could be due to reduced sampling of tree crown apices (e.g., St-Onge et al., 2003). However, from **Table 2** it is apparent that, relative to the control dataset, the actual LPR sample coverage is more than doubled due to the larger footprint area, and this despite the sample point density being reduced by almost four times. Therefore, it can be assumed that in both the wide-beam and high-altitude datasets, the reduced maximum distribution height is largely a function of increased pulse footprint area and reduced peak pulse power concentration.

Another similarity between both the wide-beam and high-altitude datasets is that the minimum (0%) distribution heights

tend to be elevated (110 and 65 cm, respectively) relative to those of the control data and display an overall reduced distribution range. In general, therefore, the control survey configuration provides a better representation of the distribution tails than either the high-altitude or wide-beam configurations, i.e., the highest and lowest elements of the profile from ground up to canopy surface are better sampled. This is an interesting result because from **Table 2** we see that the actual laser pulse sample coverage is smaller for the control dataset, and it is reasonable to assume that with reduced spatial sample coverage, the tails (i.e., ground and canopy surface) of the height distribution might be truncated, leading to shorter canopy height estimates. In addition, the peak pulse power concentration and intensity values for both the control (**Table 2**) and test (**Table 5**) datasets illustrate that the pulse power available to trigger a return is at least five times greater for the control dataset relative to that for any other survey configuration tested. These observations indicate, therefore, that the differences in the high-altitude and wide-beam height distributions are not merely a function of the area sampled by laser pulses but are more a function of the peak pulse power concentration and the range triggering method within the sensor, i.e., with reduced pulse power concentration incident upon and traveling through foliage, the shape and amplitude of the backscatter signal will be altered and the range threshold potentially shifted farther into the canopy where larger and more numerous foliage elements are encountered. These observations are corroborated by the results of Andersen et al. (2006), who found that the underestimation of tree-top heights using wide-beam data was on average 0.4 m greater than the underestimation observed in narrow-beam data.

Overall, the high-PRF distribution is shifted downwards by an average 0.32 m for all distribution quartiles relative to the control dataset (**Figure 5c; Table 5**). The dominant downshift likely has two main contributing factors. First, **Table 5** illustrates that the control data for tall vegetation had an average minimum height of 1.3 m above the ground, whereas the average minimum height for the high-PRF data was 1.2 m below the control height, or very close to true ground level. This indicates that the high-PRF data were generally better at penetrating to ground level than any of the other configurations tested, and this is most likely due to the threefold increase in point density (also observed in Chasmer et al., 2006). Second, the downshift of the distribution at the upper quartiles by 15–22 cm can be considered counterintuitive, given the higher point density (i.e., increased probability of a pulse encountering canopy apices) and increased noise inherent in high-PRF data. However, this observation is similar to those for the high-altitude and wide-beam configurations and suggests that the cause might be the same. The peak power concentration and observed intensity values for the high-PRF data relative to the control (**Tables 2, 5**) display a fivefold reduction similar to that of the high-altitude and wide-beam data. It is therefore reasonable to assume that this penetration into the canopy observed at high PRF is also because of weakened LPR backscatter and a shifted range triggering threshold.

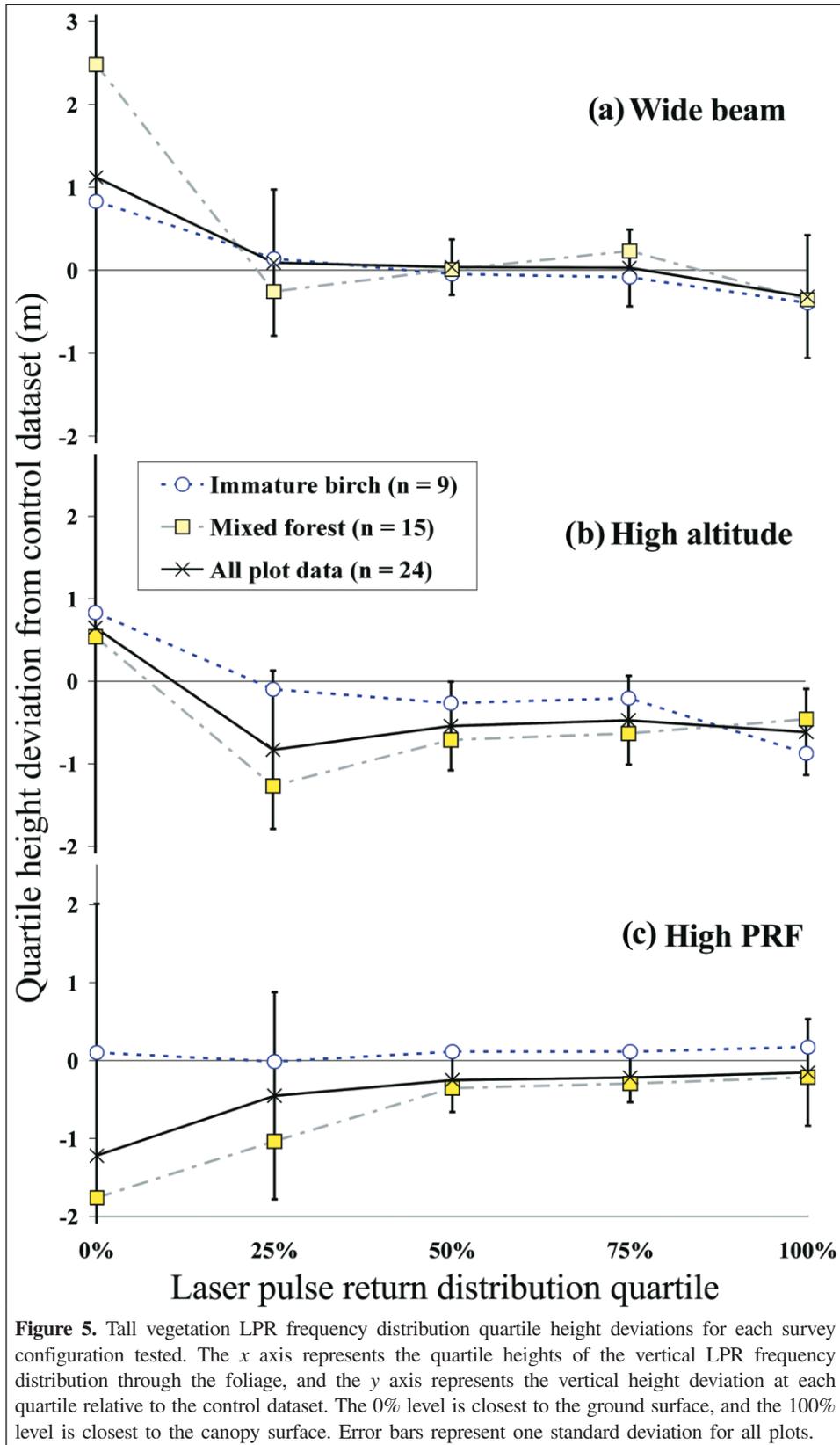


Figure 5. Tall vegetation LPR frequency distribution quartile height deviations for each survey configuration tested. The x axis represents the quartile heights of the vertical LPR frequency distribution through the foliage, and the y axis represents the vertical height deviation at each quartile relative to the control dataset. The 0% level is closest to the ground surface, and the 100% level is closest to the canopy surface. Error bars represent one standard deviation for all plots.

Of all the tests performed, the frequency distribution characteristics of the high-PRF data for tall vegetation are the only results that display a marked divergence between the two

candidate vegetation types. Although there is a predominant downshift in the combined birch and mixed wood high-PRF distribution, there appears to be no significant difference in the

birch data alone and, if anything, a slight upshift of around 0.1 m. Even though the birch data are close to the control data, it is important to note that the foliage pulse penetration behaviour differs between the two vegetation types, and this indicates that pulse penetration into foliage is not simply a function of data acquisition settings. This observation is corroborated by the work of Chasmer et al. (2006), who found that, although PRF did vary the level of penetration into conifer canopies, whether or not the magnitude of penetration increased with an increase in PRF was thought to depend on canopy openness, i.e., increased PRF was found to reduce penetration for closed conifer canopies but increased penetration for more open mixed wood canopies with understory. For the immature birch plots, it is worth noting that the overall reflectance is 9% higher than that for mixed wood (Table 4), and the canopy surface tends to be closed, with rounded canopy tops merging into one another. These observations suggest that differences in canopy reflectivity and structure may also play an important role in controlling the variable levels of lidar pulse penetration into foliage between survey configurations.

Concluding remarks

For the data presented in this study, observed laser pulse intensity values were found to be linearly related ($r^2 = 0.98$, $n = 7$) to the peak pulse power concentration at the point of contact between the emitted laser pulse and the surface encountered. Peak pulse power concentration varies systematically with survey configuration such that observed intensity will reduce for an increase in flying height, scan angle, beam divergence, and PRF. Using this knowledge it was possible to derive a simple intensity correction model that can be used to normalize the intensity values either across a single data collection of variable peak pulse power concentration (e.g., in the case of variable terrain and (or) scan angle) or for datasets collected using different survey configurations.

Increased footprint size either by widening the emitted laser pulse beam divergence or by increasing the platform flying altitude leads to reduced peak pulse power concentration, which leads to (i) reduced penetration into short canopy foliage (i.e., an upwards bias in ground level height data), (ii) increased penetration into tall canopy foliage (i.e., reduced maximum canopy return heights), and (iii) reduced representation of LPR distribution tails. Increased laser pulse repetition frequency is associated with a reduction in peak pulse power and an increase in sample point density. This results in (i) increased noise and range in the height distribution data, (ii) increased penetration to ground level in canopies that are not completely closed due to the increased probability of a pulse encountering a gap in the foliage, (iii) increased foliage penetration at upper height quartiles in canopies that are not completely closed due to a reduced signal-to-noise ratio, and (iv) potentially decreased foliage penetration in closed canopies. These are important observations because the tails of the distributions for many vegetation assessment and ground classification applications

provide the most useful information, as they correspond most closely to ground and canopy surface heights.

For tall vegetation containing some canopy gaps, reduced peak pulse power concentration means that a pulse has to penetrate farther into the foliage before sufficient surface area is encountered to backscatter enough energy towards the lidar sensor timing electronics to trigger a first return. Multiple returns cannot be discerned for short vegetation classes, and there appears to be no systematic variation in the upper distribution heights other than those associated with increased noise in high-PRF data. In general, reduced pulse energy concentration appears to reduce pulse penetration through foliage to ground level due to increased pulse attenuation in the canopy causing a slight upward bias in last return ground elevations. The important exception to this observation is in high-PRF data, where reduced pulse power concentration effects are overshadowed by increased noise in the data (short vegetation) and increased sample point density (tall vegetation).

The implications of these results are that (i) laser pulse peak power concentration is the single largest determinant of both LPR intensity values and systematic variations in canopy frequency distribution quantile heights, and (ii) LPR frequency distribution and intensity based models developed for vegetation characterization (e.g., canopy height, stem density, biomass estimation) should account for variations in laser pulse peak power concentration if they are to be transferable across datasets collected using different survey settings. Alternatively, if this is not possible or if the same sensor–survey settings cannot be used, a sensitivity analysis should be performed to estimate model imprecision due to footprint and pulse power variability.

Acknowledgements

Infrastructure and operating funding from the Canada Foundation for Innovation (CFI) and funding from the Natural Sciences and Engineering Research Council of Canada (NSERC) under the College and Community Innovation Program were used to partially support this project. Laura Chasmer is gratefully acknowledged for assisting with data collection and proofreading this manuscript. Chris Beasy and Tristan Goulden are acknowledged for assistance with ground control data collection.

References

- Andersen, H.E., Reutebuch, S.E., and McGaughey, R.J. 2006. A rigorous assessment of tree height measurements obtained using airborne lidar and conventional field methods. *Canadian Journal of Remote Sensing*, Vol. 32, No. 5, pp. 355–366.
- Baltsavias, E.P. 1999. Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 54, pp. 199–214.
- Chasmer, L.E., Hopkinson, C., Smith, B., and Treitz, P. 2006. Examining the influence of changing laser pulse repetition frequencies on conifer forest canopy returns. *Photogrammetric Engineering & Remote Sensing*, Vol. 17, No. 12, pp. 1359–1367.

- Cobby, D.M., Mason, D.C., and Davenport, I.J. 2001. Image processing of airborne scanning laser altimetry data for improved river flood modelling. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 56, pp. 121–138.
- Coren, F., and Sterzai, P. 2006. Radiometric correction in laser scanning. *International Journal of Remote Sensing*, Vol. 27, pp. 3097–3104.
- Davenport, I.J., Bradbury, R.B., Anderson, G.Q.A., Hayman, G.R.F., Krebs, J.R., Mason, D.C., Wilson, J.D. and Veck, N.J. 2000. Improving bird population models using airborne remote sensing. *International Journal of Remote Sensing*, Vol. 21, pp. 2705–2717.
- Gaveau, D., and Hill, R.A. 2003. Quantifying canopy height underestimation by laser pulse penetration in small footprint airborne laser scanning data. *Canadian Journal of Remote Sensing*, Vol. 29, No. 5, pp. 650–657.
- Goodwin, N.R., Coops, N.C., and Culvenor, D.S. 2006. Assessment of forest structure with airborne LiDAR and the effects of platform altitude. *Remote Sensing of Environment*, Vol. 103, pp. 140–152.
- Holmgren, J., and Persson, Å. 2004. Identifying species of individual trees using airborne laser scanning. *Remote Sensing of Environment*, Vol. 90, pp. 415–423.
- Holmgren, J., Nilsson, M., and Olsson, H. 2003. Simulating the effects of lidar scanning angle for estimation of mean tree height and canopy closure. *Canadian Journal of Remote Sensing*, Vol. 29, No. 5, pp. 623–632.
- Hopkinson, C., Chasmer, L.E., Sass, G., Creed, I., Sitar, M., Kalbfleisch, W., and Treitz, P. 2005. Vegetation class dependent errors in lidar ground elevation and canopy height estimates in a boreal wetland environment. *Canadian Journal of Remote Sensing*, Vol. 31, No. 2, pp. 191–206.
- Hopkinson, C., Chasmer, L.E., Lim, K., Treitz, P., and Creed, I. 2006. Towards a universal lidar canopy height indicator. *Canadian Journal of Remote Sensing*, Vol. 32, No. 2, pp. 139–152.
- Leblanc, S.G., Chen, J.M., Fernandes, R., Deering, D., and Conley, A. 2005. Methodology comparison for canopy structure parameters extraction from digital hemispherical photography in boreal forests. *Agricultural and Forest Meteorology*, Vol. 129, pp. 187–207.
- Lim, K., Treitz, P., Wulder, M., St-Onge, B., and Flood, M. 2003a. Lidar remote sensing of forest structure. *Progress in Physical Geography*, Vol. 27, pp. 88–106.
- Lim, K., Treitz, P., Baldwin, K., Morrison, I., and Green, J., 2003b. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Canadian Journal of Remote Sensing*, Vol. 29, No. 5, pp. 658–678.
- Magnussen, S., and Boudewyn, P. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research*, Vol. 28, pp. 1016–1031.
- Naesset, E. 1997. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 52, pp. 49–56.
- Naesset, E. 2004. 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*, Vol. 91, pp. 243–255.
- Naesset, E., and Bjerknæs, K.O. 2001. Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. *Remote Sensing of Environment*, Vol. 78, pp. 328–340.
- St-Onge, B., Treitz, P., and Wulder, M. 2003. Tree and canopy height estimation with scanning lidar. In *Remote sensing of forest environments: concepts and case studies*. Edited by S. Franklin and M. Wulder. Kluwer Academic Publishers, Dordrecht, The Netherlands. pp. 489–509.
- Yu, X., Hyypä, J., Kaartinen, H., and Maltamo, M. 2004. Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. *Remote Sensing of Environment*, Vol. 90, pp. 451–462.