Testing LiDAR models of fractional cover across multiple forest ecozones

Chris Hopkinson [a,*], Laura Chasmer [a,b]

[a] Applied Geomatics Research Group, NSCC Annapolis Valley Campus, Lawrencetown, NS, Canada B0S 1P0
[b] Department of Geography, Queen's University, Kingston, ON, Canada K7L 3M6

ARTICLE INFO

ABSTRACT

Four LiDAR-based models of canopy fractional cover (FC_{LiDAR}) have been tested against hemispherical photography fractional cover measurements (FC_{HP}) and compared across five ecozones, eight forest species and multiple LiDAR survey configurations. The four models compared are based on: i) a canopy-to-total first returns ratio ($FC_{LiDAR(FR)}$) method; ii) a canopy-to-total returns ratio ($FC_{LiDAR(RR)}$); iii) an intensity return ratio ($FC_{LiDAR(IR)}$); and iv) a Beer's Law modified (two-way transmission loss) intensity return ratio ($FC_{LiDAR(BL)}$). It is found that for the entire dataset, the $FC_{LiDAR(RR)}$ model demonstrates the lowest overall predictive capability of overhead FC (annulus rings 1–4) ($r^2 = 0.70$), with a slight improvement for the $FC_{LiDAR(FR)}$ model ($r^2 = 0.74$). The intensity-based $FC_{LiDAR(IR)}$ model displays the best results ($r^2 = 0.78$). However, the $FC_{LiDAR(BL)}$ model is considered generally more useful ($r^2 = 0.75$) because the associated line of best fit passes through the origin, has a slope near unity and produces a mean estimate of FC_{HP} within 5%. Therefore, $FC_{LiDAR(BL)}$ requires the least calibration across a broad range of forest cover types. The $FC_{LiDAR(FR)}$ and $FC_{LiDAR(RR)}$ models, on the other hand, were found to be sensitive to variations in both canopy height and sensor pulse repetition frequency (or pulse power); i.e. changing the repetition frequency led to a systematic shift of up to 11% in the mean $FC_{LiDAR(RR)}$ estimates while it had no effect on the intensity-based $FC_{LiDAR(IR)}$ or $FC_{LiDAR(BL)}$ models. While the intensity-based models were generally more robust, all four models displayed at least some sensitivity to variations in canopy structural class, suggesting that some calibration of FC_{LiDAR} might be necessary regardless of the model used. Short (<2 m tall) or open canopy forest plots posed the greatest challenge to accurate FC estimation regardless of the model used.

1. Introduction

1.1. Background and objective

Vegetation canopy cover exists at the interface of two important earth systems: the terrestrial and the atmospheric. The vegetative canopy acts to modify and control transfers of: i) energy in the form of radiant, sensible and latent heat; and ii) mass in the form of gas, liquid or solid, such as carbon dioxide (CO_2), nitrogen (N) and water (H_2O) (Chen et al., 2005; Law et al., 2002; Leuning et al., 2005). Information on canopy cover is essential for understanding spatial and temporal variability in vegetation biomass, local meteorological processes and hydrological transfers within vegetated environments. The ability to realistically model these transfers is dependent on accurate measurement of the canopy cover (e.g. Chen et al., 2007; Gower et al., 1999; Heinsch et al., 2003; Pomeroy and Dion, 1996). Canopy cover has been routinely monitored using satellite and airborne remote sensing instruments throughout the 20th century, providing extensive information on vegetation cover and biomass at local to global scales (Hall et al., 1991; Running et al., 2004; Tucker et al., 1986). Further, measurements of the radiation environment above and/or below the canopy (e.g. using hemispherical photography, quantum light sensors, etc.) have enabled validation of remote sensing estimates at local levels. Despite their use for pixel validation, the problems associated with local radiation measurements are two-fold: 1) they are time consuming and difficult to obtain over large spatial areas representative of remote sensing pixels; and 2) it is often expensive to obtain measurements several times throughout the growing season (except in the case of permanently logging radiation sensors) (Heinsch et al., 2006). The ability to map canopy structural attributes simultaneously over large areas and at the individual tree crown scale has, until recently, been limited either by the passive nature of the sensor employed (i.e. an inability to 'see within' the canopy) or low resolution (e.g. Tian et al., 2002; Fernandes et al., 2004).

Since the early to mid 1990s, airborne LiDAR (light detection and ranging) has demonstrated its potential to map canopy cover at the scale of tree crowns to stands by actively sampling the canopy environment at relatively high resolutions (Lefsky et al., 1999, 2005; Nelson et al., 1984; Popescu et al., 2003). The increasing availability of widespread airborne LiDAR data coverage will soon enable regional ecological, micro-meteorological and hydrological modelling and

* Corresponding author. Tel.: +1 902 825 5424.
  E-mail address: Chris.hopkinson@nscc.ca (C. Hopkinson).

assessment (e.g. Chasmer et al., in press, 2009; Kotchenova et al., 2004). In recent years, numerous studies have examined the use of small footprint discrete return airborne LiDAR data for extracting canopy structural parameters such as: gap fraction (*P*), fractional cover (FC), effective leaf area index (LAIe), fraction of incoming photosynthetically active radiation (FIPAR), transmittance (*T*), and extinction coefficient (*k*) (e.g. Barilotti et al., 2006; Hopkinson & Chasmer, 2007; Kusakabe et al., 2000; Lovell et al., 2003; Magnussen & Boudewyn, 1998; Morsdorf et al., 2006; Parker et al., 2001; Riaño et al., 2004; Solberg et al., 2006; Thomas et al., 2006; Todd et al., 2003). One of the limitations associated with drawing conclusions from such a body of literature is that the canopy structural models are generally representative of small research-based datasets that display minimal variation in either environmental or technical attributes. This constitutes a problem from the perspective of applying models at a regional scale, for we know that forest canopy structural attributes and light scattering properties are widely variable (e.g. Chen, 1996; Chen et al., 2006). It has also been shown that LiDAR data acquisition configuration can bias the canopy information extracted (Chasmer et al., 2006; Hopkinson, 2007; Naesset, 2004).

To address the above concern, we test published LiDAR-based canopy fractional cover (FC_LiDAR) models over several LiDAR datasets collected across a variety of forest ecozones and canopy structural classes. The objectives of this study are to: 1) compare FC models with hemispherical photography (HP) across a range of species types, forest stand ages, and structural canopy characteristics; 2) assess the influence of canopy structural characteristics on FC; and 3) determine if there is any sensitivity of the FC models to LiDAR sensor configuration. Four FC_LiDAR models are tested: i) the first return ratio (RR) model (e.g. Morsdorf et al., 2006); ii) the all return ratio (RR) model (e.g.; Barilotti et al., 2006; Morsdorf et al., 2006; Solberg et al., 2006); iii) a pulse return intensity ratio (IR) model (Hopkinson & Chasmer, 2007); and iv) a Beer's law modified (BL) version of in the IR model (Hopkinson & Chasmer, 2007). The paper concludes with an illustration of how LiDAR can be used to map canopy fractional cover at the scale of individual dominant canopy elements. This study, therefore, identifies the most accurate FC_LiDAR model to use (without site-specific calibration) within a range of species types, and LiDAR survey configurations.

1.2. LiDAR estimates of fractional cover

For every emitted laser pulse, there can be several reflecting surfaces along the travel path. Those backscatter elements that are strong enough to register a distinct amplitude of reflected energy at the sensor are known as 'returns'. For a discrete pulse return system such as the Airborne Laser Terrain Mapper (ALTM, Optech Inc., Toronto, Canada), the recorded ranges can be separated into single, first, intermediate and last returns. Single returns are those for which there is only one dominant backscattering surface encountered (e.g. a highway surface). For the ALTM, it is possible to also record two intermediate returns making a total of four possible returns from a single emitted pulse. While there is some slight loss of detection capability between adjacent returns (known as "dead time"), this multiple return capability means that there is a reasonable probability of sampling the dominant canopy and ground elements along the pulse travel path. The classification of the return type is encoded within the LAS binary data format (ASPRS, 2005) that is commonly used to store the raw laser point cloud data.

Laser pulses that are returned from within canopy environments have intercepted enough surface area of foliage to be recorded by the receiving optics within the LiDAR system. The remaining laser pulse energy from the same emitted pulse continues until it intercepts lower canopy vegetation, low-lying understory and/or the ground surface. Laser pulse returns from the ground surface have inevitably passed through canopy gaps both into and out of the canopy. Increasing numbers of gaps within the canopy will result in FC ap-

proaching zero, whereas fewer gaps within the canopy will result in FC approaching unity. LiDAR estimates of FC are generally based on the assumption that gap fraction (*P*) is equivalent to transmittance (*T*), and is the opposite of fractional cover (FC). From the Beer–Lambert Law:

$$1-FC = P = T = \frac{I_l}{I_o} = e^{-kLAIe} \tag{1}$$

Where I_o is open sky light intensity above canopy, I_l is the light intensity after travelling a path length (*l*) through the canopy and *k* is the extinction coefficient. The main geometric difference between the canopy interaction of solar and airborne LiDAR laser pulse radiation is that solar radiation can be incident across a wide range of zenith angles if its temporal and latitudinal distribution is considered, while laser pulses are typically incident only at near overhead (0 to 30 degrees) angles. Therefore, any direct LiDAR sampling of canopy FC will be biased towards overhead canopy elements and for a path length close to the height of the canopy. This has some advantages, as it implies that LiDAR estimates of FC can be used to directly estimate effective leaf area index (LAIe).

Studies that have examined the use of LiDAR for obtaining FC, *P*, and LAI assume that FC may be directly inferred by a pulse return ratio of the number of canopy-to-total returns (e.g. Barilotti et al., 2006; Morsdorf et al., 2006; Riaño et al., 2004; Solberg et al., 2006). Morsdorf et al. (2006) found that the best FC_LiDAR estimate was returned when only the first and single return data were used to predict FC_HP for the first two overhead HP annulus rings. Other authors have implicitly used all returns (i.e. first and last) in their computation of FC_LiDAR (e.g. Barilotti et al., 2006; Riaño et al., 2004). There are slightly different rationales behind each approach. In the first return ratio approach (FC_LiDAR(FR)), it is assumed that if a first return hits a gap, that it will be from ground, while if it is intercepted by foliage then it represents canopy. By inference, therefore, the ratio of canopy-to-total first returns provides a direct estimate of canopy cover:

$$FC_{Lidar(FR)} = \frac{\sum R_{Canopy(First)}}{\sum R_{Total(First)}} \tag{2}$$

where $R_{Canopy(First)}$ is the frequency of first returns above some height threshold, while $R_{Total(First)}$ is the frequency of total first returns throughout the canopy to ground profile. However, this method does not account for the fact that if an emitted pulse encounters only a small area of canopy foliage, there might be insufficient reflected energy to actually record a canopy first return; i.e. canopy could be under-estimated. Conversely, most canopy-level first returns have reflected from areas of foliage that represent less than 100% of the pulse area and thus the canopy cover would be over-estimated. There is no evidence to suggest that partial pulse reflection of first and single returns leads to a systematic under- or over-estimation, or that these effects cancel one another out. Bearing these factors in mind, there is a rationale that using all returns can provide just a good a predictor of FC; not least because using single, first, intermediate and last returns provides an increased sampling density of points throughout the canopy to ground profile. The second model tested is thus:

$$FC_{Lidar(RR)} = \frac{\sum R_{Canopy(All)}}{\sum R_{Total(All)}} \tag{3}$$

where $R_{Canopy(All)}$ is the frequency of all (single, first, intermediate and last) returns above some height threshold, while $R_{Total(All)}$ is the frequency of total returns throughout the canopy to ground profile.

In discrete return scanning airborne LiDAR systems, intensity is recorded as a scaled index of the reflected pulse energy amplitude for each range measured by the LiDAR sensor. This information has implicitly been used in estimates of canopy gap fraction in the full

waveform LiDAR literature where the strength of the returned signal from within or below the canopy is considered to be directly related to the transmissivity of the canopy (Kotchenova et al., 2003; Lefsky et al., 1999; Parker et al., 2001). For example, in Lefsky et al. (1999), it was suggested that canopy fractional cover can be estimated as a function of the ratio of the power reflected from the ground surface divided by the total returned power of the entire waveform. This concept was examined in Hopkinson and Chasmer (2007) using discrete return LiDAR data, whereby two models of P (or $1 - FC$) were developed and tested utilising discrete return intensity information as an index of the pulse return power as it passed through the canopy.

For laser pulses encountering and returning from a forested canopy at near-nadir scan angles, we cannot easily estimate the incident pulse intensity as it enters the canopy; neither can we measure the transmitted intensity after it has passed through the canopy. However, by considering the total reflected energy from i) the canopy to ground profile and ii) the ground level, as being some proportion of the total available laser pulse intensity, we have a means of estimating canopy FC at near-nadir angles. Further we can assume that atmospheric transmission losses for all outgoing and returning laser pulses are similar and small in magnitude relative to canopy losses, and therefore can be ignored. By adapting Eq. (2), FC can be estimated from a discrete pulse return ratio of the sums of canopy to total return intensities. This approach is termed the 'intensity ratio' (IR) method:

$$FC_{Lidar(IR)} = \frac{\sum I_{Canopy}}{\sum I_{Total}} \qquad (4)$$

Where $\sum I_{Canopy}$ is an index of canopy power (the sum of all canopy level return intensity) and $\sum I_{Total}$ is an index of the total returned power (sum of all return intensity) for the entire canopy to ground profile.

A limitation of the IR model in Eq. (4) is that it does not explicitly account for potentially different probabilities associated with receiving a return signal from the ground or canopy level. For discrete return data, it is fair to assume that the majority of first and single returns have not incurred any appreciable transmission loss prior to being reflected back towards the sensor. However, intermediate or last returns are, by definition, a reflected component of the residual energy left over after a previous return was reflected from a surface encountered earlier in the travel path of the emitted pulse. From Beer–Lambert's Law and assuming uniform transmission losses per unit path length travelled, it is possible that a below canopy (ground level) return incurs a similar proportion of transmission loss during its exit from the canopy as it did on the way into the canopy. Further, assuming canopy pulse return intensity is largely a function of the surface reflecting area, it follows that the extinction of pulse energy into and out of the canopy will follow a power law reduction.

The following 'thought experiment' illustrates this point:

For a canopy of 40% fractional cover where all surfaces reflect primarily in proportion to their surface area (i.e. differences in surface reflecting area are greater than differences in surface spectral reflectance) and where we expect to lose an equal amount of energy both into and out of the canopy, we can expect to see the following: 1) a loss of 40% available pulse energy on the way in (i.e. 60% transmittance) and a further loss of 40% of what is left after ground reflection on the way out through the canopy; 2) Given the same proportional transmittance is expected in both directions, the transmittance is compounded; 3) we are therefore left with 60%×60%=36% of the total energy returned representing the ground level; 4) Therefore, to account for this two-way energy transmission loss we square root the proportion of the ground level energy reflectance (relative to the total energy profile) to reverse estimate the actual one-way transmittance of the canopy. 5) We subtract the transmittance value from unity to estimate the fractional cover.

The Beer's Law modification (BL) of Eq. (4) for intermediate and last returns from within or below the canopy is expressed below:

$$FC = 1 - \sqrt{\frac{\sum I_{Ground}}{\sum I_{Total}}} \qquad (5)$$

where $\sum I_{Ground}$ is below canopy power (the sum of all below canopy return intensity).

To account for the unequal transmission losses associated with each return type (i.e. first and single vs. intermediate and last), Eqs. (4) and (5) must be combined. The notation for this new Beer's Law modified fractional cover equation is as follows:

$$FC_{Lidar(BL)} = 1 - \left(\frac{\left(\frac{\sum I_{GroundSingle}}{\sum I_{Total}} \right) + \sqrt{\frac{\sum I_{GroundLast}}{\sum I_{Total}}}}{\left(\frac{\sum I_{First} + \sum I_{Single}}{\sum I_{Total}} \right) + \sqrt{\frac{\sum I_{Intermediate} + \sum I_{Last}}{\sum I_{Total}}}} \right) \qquad (6)$$

where each subscript refers to the class and/or sub-class of pulse return. In this model, first and single returns incur no reverse transmission loss through canopy and so are not square rooted, while intermediate and last returns could potentially lose similar proportions of energy due to interception on both incoming and outgoing transmission; i.e. a power function loss.

The analysis presented in this paper builds on previous research primarily by testing the models described over multiple datasets collected over several forest ecozones, and then elucidating some model sensitivities to important environmental and technical data attributes. The first model tested is the $FC_{LiDAR(FR)}$, which used the ratio of canopy first returns to total first returns (Eq. (2)). The second model tested is the $FC_{LiDAR(RR)}$, which used the ratio of all canopy returns to all total returns (Eq. (3)). The third model tested was the $FC_{LiDAR(IR)}$, and employed the ratio of cumulative canopy return intensity to cumulative total return intensity (Eq. (4)). The final model tested was the $FC_{LiDAR(BL)}$ (Eq. (6)), which was modified from Eqs. (4) and (5) to account for a potential two-way transmission loss of pulse energy for intermediate or last returns.

2. Study areas

The study was conducted over seven sites, across Canada between 2002 and 2007 (Fig. 1) and within five distinct Canadian forest ecozones: Boreal Forest (Wolf Creek, Baker Creek and Prince Albert National Park BERMS — Boreal Ecosystem Research and Monitoring Sites); Western Boreal Plains (Utikuma Lake); Canadian Rockies Montane Forest (Bow Summit); Southern Great Lakes (Vivian Forest); and Acadian Forest (Annapolis Valley). Each site has been surveyed with airborne LiDAR at least once, has been sampled with multiple analogue or digital hemispherical photographs (HP), and geo-located using survey-grade or handheld global positioning systems (GPS). Sites contain one to many different forest species types, varying ages, and canopy structural characteristics (Table 1). Sites also vary in topography, where some sites are flat (e.g. Annapolis Valley and Vivian Forest), gently rolling (e.g. Baker Creek and Lake Utikuma), or mountainous with steep terrain (e.g. Bow Summit and Wolf Creek). In total, 245 geolocated HP images were obtained over 80 spatially independent plots of corresponding LiDAR and HP data (Fig. 1).

In this study, four sites are within the boreal forest ecozone. Wolf Creek lies in a region of sub-arctic climate with vegetation zones ranging from boreal forest to alpine tundra and is located near Whitehorse, Yukon Territory (Pomeroy et al., 2005; Quinton et al., 2005). Baker Creek is located near Yellowknife in the Northwest Territories, is underlain by Canadian Shield bedrock and demonstrates a patch work of taiga woodland and boreal forest species. The third Boreal ecozone site (BERMS) is located, near Prince Albert National Park, Saskatchewan and is part of Fluxnet-Canada (2002–2007) and the Canadian Carbon Program (2007–2011) networks (Barr et al.,
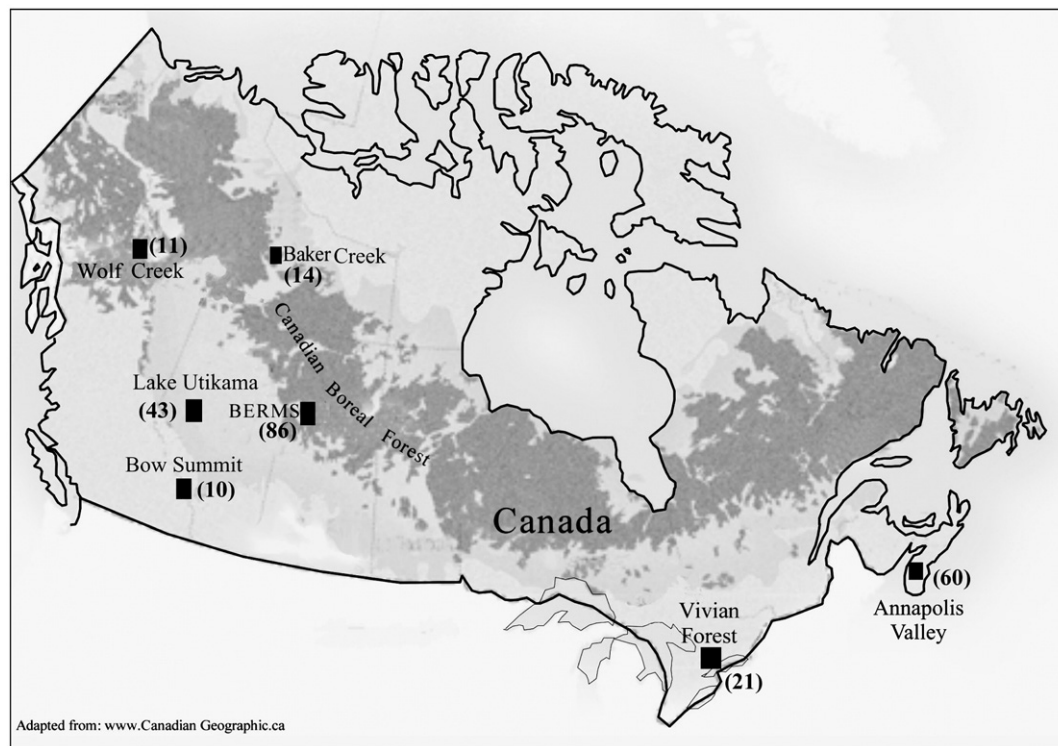
Fig. 1. Map of the seven study locations across Canada. Number in brackets denotes number of photo stations at each site containing corresponding LiDAR and HP data.

2002; Schwalm et al., 2006). The site possesses a mix of managed clear cut, harvested and naturally regenerating and mature boreal forest species examined in this study. The Utikuma Lake area, located in the Western Boreal Plains is approximately 300 km north of Edmonton in Alberta (Hopkinson et al., 2005).

The second ecozone examined is classified as Montane or sub-alpine and contains one site. The Bow Summit site is located about 200 km northwest of Calgary on the Icefields Parkway between Banff and Jasper on the border of Alberta and British Columbia (Hopkinson and Demuth, 2006; Hopkinson et al., 2006). The site traverses an elevational gradient from 1700 m a.s.l. to 2000 m a.s.l., terrain is steep, and canopies tend to be open and irregular.

The Southern Great Lakes ecozone contains the Vivian Forest site. This site is located approximately 50 km north of Toronto, Ontario (Hopkinson et al., 2004, 2008). While the Vivian Forest area is characterised by a variety of managed and natural forest coverage, the forest plots sampled in this study were exclusively within a Red pine (*Pinus resinosa* Ait.) plantation displaying a uniform canopy height with no understory.

Finally, three Acadian Forest ecozone sites located within the Annapolis Valley of Nova Scotia, approximately 150 km west of Halifax, were also studied: Two sets of plots were set up within uneven aged natural regeneration stands comprising predominantly Yellow birch (*Betula alleghaniensis* Britton) and White pine (*Pinus strobus* L.) with

Table 1
Forest plot descriptions and species type per site location

Site (dominant canopy species)	Location	Age (yrs)	Average height (m)	Average LAIe	Reference(s) on sites and methods
Mature Red Pine (*Pinus resinosa* Ait.)	Vivian Forest, Newmarket ON	50	23	4.2	Chasmer et al. (2006) Hopkinson et al. (2008)
Mature Sub-Alpine Fir (*Abies lasiocarpa* Nutt.)	Bow Summit, Banff National Park, AB	40	9	0.9	Hopkinson and Demuth (2006) Hopkinson et al. (2006)
Mature Trembling Aspen (*Populus tremuloides* Michx.)	Lake Utikuma, AB	–	16	0.8	Hopkinson et al. (2005) Hopkinson et al. (2006)
Mature Black Spruce (*Picea mariana* Mill.)	Lake Utikuma, AB	–	7	0.6	Hopkinson et al. (2005) Hopkinson et al. (2006)
Old Jack Pine (*Pinus banksiana* Lamb.)	Prince Albert National Park, SK	90	14	1.6	Chen et al. (2006) Schwalm et al. (2006)
Harvested 1975 Jack Pine (*Pinus banksiana* Lamb.)	Prince Albert National Park, SK	30	6.3	2.8	Chen et al. (2006) Schwalm et al. (2006)
Harvested 1994 Jack Pine (*Pinus banksiana* Lamb.)	Prince Albert National Park, Saskatchewan	11	1.6	1.1	Chen et al. (2006) Schwalm et al. (2006)
Mature Acadian Mixedwood (*Acer saccharum* Marsh., *Pinus strobus* L. *Betula alleghaniensis* Britt.)	Annapolis Valley, Nova Scotia	100	21	3.4	Hopkinson and Chasmer (2007)
Immature Birch (*Betula alleghaniensis* Britt.)	Annapolis Valley, Nova Scotia	12	5	2.5	Hopkinson and Chasmer (2007)
Mature Red Pine (*Pinus resinosa* Ait.)	Annapolis Valley, Nova Scotia	40	20	–	Hopkinson and Chasmer (2007)
Mature White Spruce (*Picea glauca* (Moench) Voss.)	Wolf Creek, Whitehorse, Yukon Territories	–	10	0.2	Pomeroy et al. (2005) Quinton et al. (2005)
Mature Trembling Aspen (*Populus tremuloides* Michx.)	Baker Creek, Yellowknife NWT	–	7	1.23	–
Mature Black Spruce (*Picea mariana* Mill.)	Baker Creek, Yellowknife NWT	–	9	0.18	–
Mature Birch (*Betula alleghaniensis* Britt.)	Baker Creek, Yellowknife NWT	–	5	2.31	–

(–represent unavailable data).

occasional Maple (*Acer*) and Spruce (*Picea*) trees interspersed throughout. A third set of plots was located in a nearby Red pine (*Pinus resinosa* Ait.) plantation. The mixed wood sites within the Annapolis Valley were used for the development of the laser pulse return intensity-based FC models developed in Hopkinson and Chasmer (2007). In order to ensure a fair test of the four models discussed above, the lidar and HP datasets used in Hopkinson and Chasmer (2007) are not used in the model comparison analysis presented here.

While a diversity of forest species were sampled at the seven sites (Table 1), it was decided to stratify all plots by canopy structural class (Fig. 2). The classes identified were: i) closed canopy conifer (e.g. pine stands: Vivian Forest Red and White pines; Annapolis Valley Red pine; BERMS mature Jack pine); ii) open or short canopy conifer (e.g. Black spruce associated with lowland environments or immature Jack pine regeneration: Prince Albert National Park (harvested Jack pine); iii) mixed wood (e.g. the mixed birch and spruce plots common to the Acadian Forest ecozone and mature Trembling aspen of Baker Creek). These three classes were chosen because they represent structural canopy end members within the sites sampled and can be expected to display different behaviours in terms of vertical LiDAR pulse return sampling within and below the canopy. Of the 80 independent test plots, 33 (64) were closed canopy conifer, 19 (86) were open canopy conifer, and 28 (95) were either mixed or hardwood (numbers in brackets represent the HP images acquired for each canopy class). There were seven plots that contained only hardwood species but given the small number, overlap in species composition and general similarity in canopy structure to the mixed wood plots (i.e. random foliage distribution and minimal clumping) the mixed and hardwood classes were combined.

3. Methods

3.1. Hemispherical photography data collection and analysis

Canopy FC was collected using analogue and digital HP methods at geo-located sites within representative forest types throughout each study area. Photographic plots were set up in two ways: a) as individual plots containing five photograph stations. One photograph was taken at the centre of the plot, and four were located 11.3 m from the centre along cardinal (N, S, E, and W) directions, determined using a compass and measuring tape; b) along transects with HP images acquired at between 15 m and 100 m intervals. The centres of the photo plots (method 'a') were geo-located using survey-grade, differentially corrected GPS receivers (Leica SR530, Leica Geosystems Inc. Switzerland; Ashtec Locus, Ashtec Inc., Hicksville, NY) to the same base station coordinate as was used to control the LiDAR surveys. Geo-

location accuracies vary from 1 cm to 1 m depending on canopy density at time of GPS data collection. The tape and bearing methods used to locate the four cardinal direction photo stations are believed to be accurate to within approximately 2 m.

Photographs collected along transects (method 'b') were located using WAAS-enabled (wide area augmentation service) handheld GPS (Trimble Inc. GeoExplorer, Idaho, USA). These photographs have a locational accuracy of between 2 m and 10 m depending on GPS satellite configuration and canopy cover at the time of data collection. Given the reduced geolocation accuracy and the fact that HPs collected along transects cannot be considered spatially independent if they are close to one another, we chose to average the results of any HP images that were within 25 m of one another. Sites that were revisited more than once using HP (i.e. Annapolis Valley) had permanent stakes for each photo station. Photographs at all sites except for the BERMS HJP94 (Regenerating Jack pine site harvested in 1994) were recorded at a height of 1.3 m above the ground level. At HJP94, photographs were taken at a height of 0.7 m due to the average tree height being around 2 m within this immature stand.

All photographs were collected within a week of the corresponding LiDAR survey and were taken either during diffuse daytime conditions, or within 30 min of dawn or dusk to reduce the influence of sun brightness and apparent foliage area reduction within the photograph (Zhang et al., 2005). Exposure settings were set one 'f' stop smaller than the automatic exposure reading to slightly under-expose the image and increase contrast between vegetation and sky. All digital HPs were collected at between 4 and 8 megapixels, while the 74 analogue HPs (collected in 2002 at Vivian Forest, Utikuma Lake and Bow Summit only) were digitized to 4 megapixels resolution at a commercial professional photography lab prior to analysis. To obtain estimates of FC each photograph was processed following sky and vegetation thresholding methods of Leblanc et al. (2005). Photograph channels were separated into red, green, and blue using the software *Paint Shop Pro* (Corel Inc. Ottawa, Canada). The blue channel is best able to differentiate between sky and foliage cover (Leblanc et al., 2005), and therefore was further processed for gap fraction using DHP (digital hemispherical photography) version 1.6.1 software (S. Leblanc, Canada Centre for Remote Sensing provided to L. Chasmer through the Fluxnet-Canada Research Network). Separation of sky and foliage components by DHP software was performed using two thresholds: one for separating canopy pixels and the second for separating sky pixels. The thresholds used were not the default thresholds of Leblanc et al. (2005). Instead, each photograph was assessed individually for the best thresholds to use. Detailed comparisons between fractional cover and effective LAI at the BERMS plots were made in Chasmer et al. (2008a) and Chasmer et al. (2008b) and compared with those of Chen
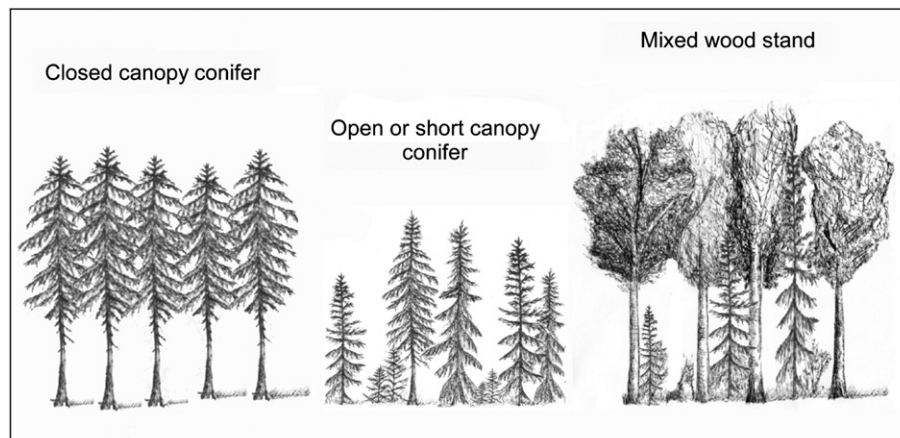

Fig. 2. Forest canopy classes assigned to each of the forest plots sampled.

Table 2
Hemispherical photography set up and geo-location at individual plots within each site

Survey site	Date(s) of photo survey	Camera model	Lens type	Height of photo (m)	Time of day	# of photos per plot	Coordinate location device	Positional accuracy (m)
Vivian Forest, ON	August 7, 2002	Nikon F-601 AF	Nikkor 8 mm 180° fisheye lens	1.3	Dusk	4	GPS survey of corners, tape and bearing to photo centres	<3
Bow Summit, AB	August 22, 2002	Nikon F-601 AF	Nikkor 8 mm 180° fisheye lens	1.3	Midday diffuse	5	GPS survey of corners, tape and bearing to photo stations	<3
Lake Utikuma, AB	August 25–30 2002	Nikon F-601 AF	Nikkor 8 mm 180° fisheye lens	1.3	Midday diffuse	5	GPS survey of corners, tape and bearing to photo stations	<3
BERMS, SK	August 10–20, 2005	Nikon Coolpix 8.0 megapixel	Nikon FC-E9 180° fisheye lens	1.3 (0.7 m at HJP94)	30 min before dusk/dawn	5	GPS survey of plot centre, tape and bearing to photos	<1
Annapolis Valley, NS	Sept. 24, 2007 Dec. 3, 2007	Nikon Coolpix 8.0 megapixel	Nikon FC-E9 180° fisheye lens	1.3	Midday diffuse Dusk	5	GPS survey of plot centre, tape and bearing to photos	<1
Wolf Creek, YT	Aug. 12, 2007	Nikon Coolpix 8.0 megapixel	Nikon FC-E9 180° fisheye lens	1.3	Midday diffuse	1	Handheld GPS of centre of single photo	<5
Baker Creek, NWT	Aug. 20, 2007	Nikon Coolpix 8.0 megapixel	Nikon FC-E9 180° fisheye lens	1.3	Midday diffuse	1	Handheld GPS of centre of single photo	<5

et al. (2006) using TRAC, Li-2000, and HP methods. The findings of Chasmer et al. (2008a,b) were similar to those of Chen et al. (2006), despite differences in the exact location of plots and transects. FC_{HP} data were output for each of the ten hemispherical annulus rings; where ring one represents the overhead zenith angles from 0°–9° and ring 10 represents the horizon zenith angles from 81°–90°. Table 2 summarises the dates of the survey and types of cameras and lenses used.

3.2. LiDAR data collection and preparation

LiDAR data were collected at each site using two generations of Optech Inc. ALTMs operating at a wavelength of 1064 nm. Data collections in 2002 were planned by the authors, but collected and processed by Optech Inc. All LiDAR data collections from 2005 to 2007 were planned, collected, and processed by the authors using the Applied Geomatics Research Group (AGRG) ALTM 3100. During each data collection, at least one GPS base station was located over a known survey monument within 40 km of each study site for georegistration of the airborne trajectory and LiDAR pulse return data. Table 3 provides details of the LiDAR data collections and survey parameters used. All sites except for Wolf Creek were flown with 50% overlap of scan lines, ensuring that all objects on the ground were viewed from two directions and sample densities were as uniform as possible.

After data collection, the first data processing task was to differentially correct the aircraft GPS trajectories for each airborne survey to the respective GPS base station receiver/s on the ground. Raw laser pulse return ranges and scan angles were integrated with aircraft trajectory and attitude (ALTM sensor orientation) data using *PosPAC* (Applanix, Toronto) and *REALM* (Optech, Toronto) proprietary software tools. The outputs from these procedures were a series of flight line data files containing xyzi (easting, northing, elevation, intensity) information for each laser pulse return collected from the ground or canopy environment being sampled. All intensity data were normalised to a range of 1000 m to mitigate against geometric variations in intensity due to scan angle, terrain relief and aircraft altitude variation (see Hopkinson, 2007).

Following LiDAR point position computation, the xyzi data files were imported into the *Terrascan* (Terrasolid, Finland) software package for plot subsetting and to separate canopy and below canopy returns. All datasets were classified to extract ground returns (R_{Ground}) using the Terrascan morphological ground classification filter (e.g. Axelsson, 1999; Vosselman, 2000) to provide a digital elevation model (DEM) to which all laser pulse return heights could be normalised. After normalization, all elevations for all datasets were relative to the same ground level datum; i.e. possessed heights ranging from 0 m to 35 m. This allowed all returns to be divided into canopy (R_{Canopy}) and below canopy returns using a height threshold of 1.3 m (0.7 m in the

case of HJP94) to coincide with the height of the HP field data. A separate file (R_{Total}) was used to store the combination of both canopy and ground returns.

For each of the 74 analogue and 171 digital HPs collected across the seven LiDAR survey sites, all laser pulse return data were extracted within a circular radius of 11.3 m. This radius was chosen as it was: a) consistent with field mensuration practices; b) was close to the optimal radius of approximately 15 m observed in Morsdorf et al. (2006); and c) is sufficiently large to mitigate the effect of possible geolocation errors between LiDAR plots and photo stations. In addition to the R_{Canopy}, R_{Total} and R_{Ground} classes, the return data were further subdivided into four sub-classes related to the order of the return itself; i.e. single, first, intermediate and last returns. (These return classifications are explicitly encoded within the LAS binary format and do not need to be deduced in post-processing (ASPRS, 2005)). For the R_{Canopy} class, it is possible for a return to belong to any one of the four sub classes (provided the canopy is deep enough), however, R_{Ground} returns can only belong to either the last or single return sub class. These subdivisions of pulse return classes were necessary as this information is used within each of the models tested and is an indicator of whether or not the pulse has been split and therefore potentially susceptible to energy transmission losses on its way into and out of the canopy.

3.3. Testing the fractional cover models

FC estimates were extracted for each of the individual HP stations and related to plot-level LiDAR estimates derived from the four

Table 3
LiDAR survey configurations and collection dates for the FC_{LiDAR} test and for the survey configuration sensitivity analysis

Survey site location	Date(s) of survey	ALTM model	Flying height (m a.g.l)	Pulse frequency (kHz)	Scan angle (degrees)	Approx. spacing (m)
FC validation datasets:						
York Regional Forest, ON	July 29, 2002	2050	850	50	±12	0.9
Bow Summit, AB	Aug. 22, 2002	2050	1000–1500	50	±18	1.0
Lake Utikuma, AB	Aug. 30, 2002	2050	1200	50	±16	1.0
BERMS, SK	Aug. 12, 2005	3100	950	70	±19	0.50
Annapolis Valley, NS	Oct. 1, 2007	3100	1000	70	±20	0.8
	Nov. 26, 2007	3100	1000	70	±20	0.8
Wolf Creek, YT	Aug. 11, 2007	3100	1300–1600	33	±23	1.0
Baker Creek, NWT	Aug. 22, 2007	3100	1200	70	±25	0.5
Model sensitivity dataset						
Annapolis Valley, NS – Survey configuration test	May 15, 2006	3100	1000	70	±20	0.8
	May 15, 2006	3100	1000	33	±20	1.2

Table 4
Regression statistics of FC$_{HP}$/FC$_{LiDAR}$ for each of the four models tested

LiDAR Fractional Cover Model	Statistic	Hemispherical photo annulus rings and mean zenith angle								
		1–1 (4.5°)	1–2 (9°)	1–3 (13.5°)	1–4 (18°)	1–5 (22.5°)	1–6 (27°)	1–7 (31.5°)	1–8 (36°)	1–9 (40.5°)
First Return Ratio (FR)	Slope	0.76	0.77	0.78	0.79	0.78	0.77	0.76	0.76	0.71
	Intercept	−0.20	−0.18	−0.17	−0.15	−0.13	−0.10	−0.07	−0.04	0.00
	r^2	0.58	0.66	0.72	0.74	0.75	0.75	0.74	0.73	0.70
All Return Ratio (RR)	Slope	1.08	1.09	1.10	1.10	1.08	1.07	1.04	1.01	0.97
	Intercept	−0.24	−0.22	−0.20	−0.18	−0.16	−0.13	−0.10	−0.06	−0.02
	r^2	0.57	0.64	0.68	0.70	0.70	0.70	0.68	0.66	0.63
Intensity Ratio (IR)	Slope	0.89	0.89	0.88	0.88	0.86	0.84	0.81	0.78	0.75
	Intercept	−0.07	−0.05	−0.02	0.00	0.02	0.05	0.08	0.12	0.15
	r^2	0.66	0.74	0.77	0.78	0.76	0.75	0.72	0.69	0.65
Beer's Law Intensity Ratio (BL)	Slope	1.06	1.06	1.05	1.04	1.01	0.99	0.96	0.92	0.88
	Intercept	−0.06	−0.04	−0.02	0.01	0.03	0.06	0.09	0.13	0.16
	r^2	0.65	0.72	0.75	0.75	0.73	0.71	0.69	0.65	0.61

Shaded cells denote best-fit lines that display: a) a slope within 5% of unity; b) an intercept within 5% of the origin; c) a coefficient of determination exceeding 0.75; i.e. >75% explanation of sample variance.

models described above Eqs. (2), (3), (4), and (6). The modelled FC$_{LiDAR}$ estimates were compared with the FC$_{HP}$ observations for the first nine annulus rings (0°–81°) by starting with the first overhead ring (0°–9°) and sequentially adding the next ring until all nine had been compared. Regression analyses were performed to assess the correspondence between LiDAR and HP estimates of FC. Following the comparison of the models over all 80 test plots, the models were evaluated for their respective sensitivities to: i) variations in canopy height; ii) variations in canopy structure/openness due to species classification; iii) changes in LiDAR data acquisition settings.

In the case of i) and ii), this was possible by simply stratifying the results from the 80 test plots into canopy height and structural class, re-evaluating the regression results and performing a residual analysis for each model and each canopy class. To assess the sensitivity to data acquisition settings it was necessary to perform a new analysis. It is known that LiDAR survey configuration can have a marked influence on canopy penetration (Chasmer et al., 2006; Hopkinson, 2007; Naesset 2004). To test whether or not a change in pulse repetition frequency (hence also pulse power and sampling density) would influence FC$_{LiDAR}$, two surveys were performed over a conifer plantation

in the Annapolis Valley on the same day; the first at 70 kHz pulse frequency, 3.7 kW peak pulse power, 0.8 m point spacing and the second at 33 kHz, 21.8 kW peak pulse power, 1.2 m point spacing.

4. Results and discussion

4.1. Testing the FC$_{LiDAR}$ models

The best-fit regression results comparing the 9 ring FC$_{HP}$ with FC$_{LiDAR}$ for the 80 independent test plots across all seven sites are summarised in Table 4. It is observed that all three FC$_{LiDAR}$ models are significantly correlated ($P < 0.01$) with FC$_{HP}$ for all combinations of HP annulus rings from overhead zenith angles through to the average of the full hemisphere represented by rings 1–9. Generally, the FC$_{LiDAR(IR)}$ model illustrates the best predictive capability of FC$_{HP}$, with the nine combinations of annulus rings displaying r^2 values between 0.65 and 0.78 This is followed by FC$_{LiDAR(BL)}$ with all ring combinations displaying r^2 values between 0.61 and 0.75; then FC$_{LiDAR(FR)}$ with values between 0.58 to 0.75; and finally FC$_{LiDAR(RR)}$ with values between 0.57 and 0.70. The best explanation of variance is found between FC$_{LiDAR(IR)}$
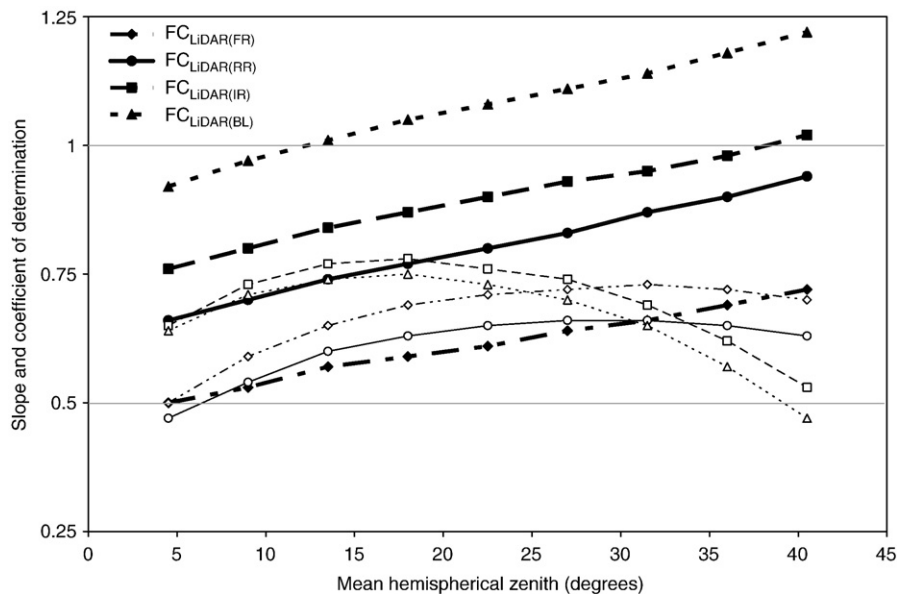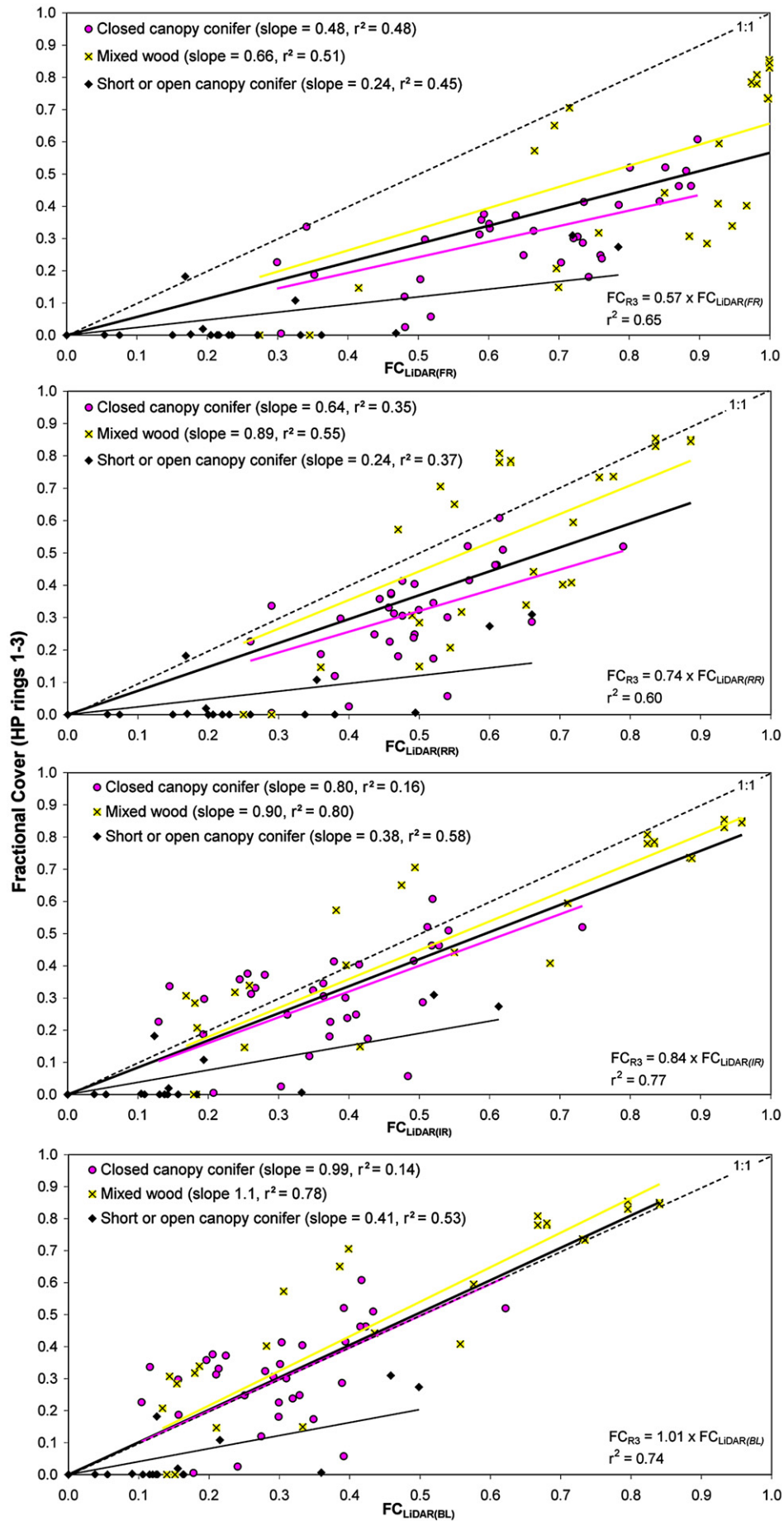

Fig. 3. Regression results by forcing all FC data through the origin. Best fit slope represented by solid symbols and coefficient of determination represented by hollow symbols.
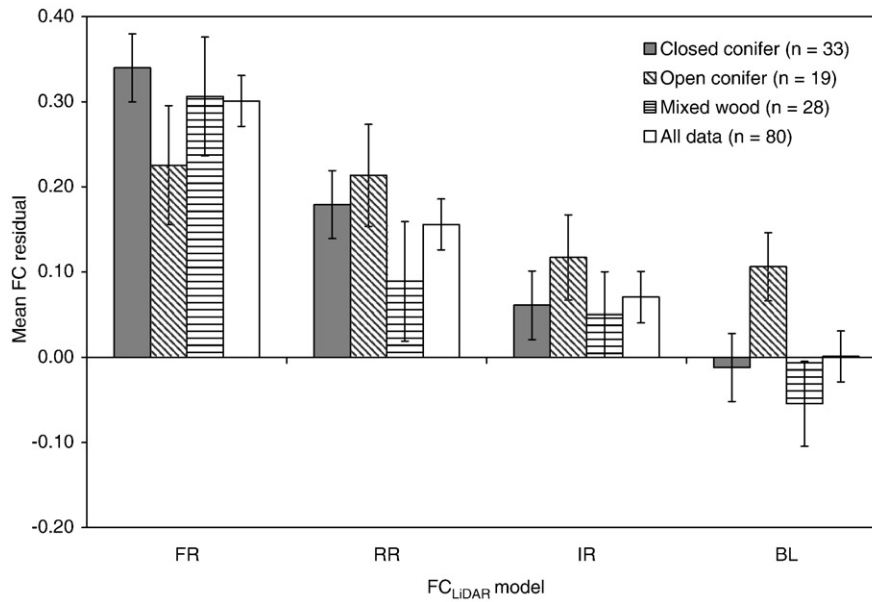
Figure: Four scatterplots of Fractional Cover (HP rings 1–3) versus FC_{LiDAR} for different return types.

Panel 1 ($FC_{LiDAR(FR)}$):
- Closed canopy conifer (slope = 0.48, $r^2 = 0.48$)
- Mixed wood (slope = 0.66, $r^2 = 0.51$)
- Short or open canopy conifer (slope = 0.24, $r^2 = 0.45$)
- $FC_{R3} = 0.57 \times FC_{LiDAR(FR)}$, $r^2 = 0.65$

Panel 2 ($FC_{LiDAR(RR)}$):
- Closed canopy conifer (slope = 0.64, $r^2 = 0.35$)
- Mixed wood (slope = 0.89, $r^2 = 0.55$)
- Short or open canopy conifer (slope = 0.24, $r^2 = 0.37$)
- $FC_{R3} = 0.74 \times FC_{LiDAR(RR)}$, $r^2 = 0.60$

Panel 3 ($FC_{LiDAR(IR)}$):
- Closed canopy conifer (slope = 0.80, $r^2 = 0.16$)
- Mixed wood (slope = 0.90, $r^2 = 0.80$)
- Short or open canopy conifer (slope = 0.38, $r^2 = 0.58$)
- $FC_{R3} = 0.84 \times FC_{LiDAR(IR)}$, $r^2 = 0.77$

Panel 4 ($FC_{LiDAR(BL)}$):
- Closed canopy conifer (slope = 0.99, $r^2 = 0.14$)
- Mixed wood (slope 1.1, $r^2 = 0.78$)
- Short or open canopy conifer (slope = 0.41, $r^2 = 0.53$)
- $FC_{R3} = 1.01 \times FC_{LiDAR(BL)}$, $r^2 = 0.74$

Fig. 5. Plot of mean residuals by canopy structure class and by FC_{LiDAR} model. Error bars indicate 95% confidence limits on the mean residual.

and FC_{HP} for annulus rings 1–4 ($r^2 = 0.78$). However, while the best-fit line passes close to the origin, suggesting generally good predictive capability throughout the range of values experienced, the $FC_{LiDAR(IR)}$ model tends to over-estimate fractional cover with the slope never exceeding 0.89 for the overhead annulus rings. In the case of the $FC_{LiDAR(FR)}$ model, the slope never exceeds 0.79 (rings 1–4) and the intercept approaches the origin only for the full hemisphere (rings 1–9) illustrating a tendency to over-estimate FC with weak predictive capability for overhead zenith angles. For the $FC_{LiDAR(RR)}$ model, the overall explanation of variance and intercept values do not deviate significantly from those of $FC_{LiDAR(FR)}$, however, the magnitude of the slope is always within 10% (i.e. close to unity) suggesting that the values predicted are closer in magnitude to actual FC_{HP} values. Of the four models tested, only $FC_{LiDAR(BL)}$ demonstrates a combination of high explanation of variance, a slope near to unity with an intercept near the origin; i.e. the $FC_{LiDAR(BL)}$ model demonstrates good predictive capability across the full range of values requiring no calibration.

By forcing all data through the origin (as it is impossible to have a negative FC) we see in Fig. 3, that the intensity-based FC_{LiDAR} models (IR and BL) tend towards better predictive capability for overhead zenith angles, while the pulse return ratio models (FR and RR) tend toward a better characterisation of the entire hemisphere. Further, while $FC_{LiDAR(FR)}$, $FC_{LiDAR(RR)}$ and $FC_{LiDAR(IR)}$ tend to over-estimate FC_{HP} for almost all zenith angles, we see that $FC_{LiDAR(BL)}$ demonstrates a near 1:1 predictive capability within the 5° to 20° range of mean zenith angles with optimal results for HP rings 1–3. These annulus rings are representative of overhead canopy conditions and represent zenith angles that are similar to the range of scan angles emitted in typical airborne LiDAR surveys.

The results displayed in Table 4 and Fig. 3 are consistent with the findings from the original intensity-based model development presented in Hopkinson and Chasmer (2007), however, the explanation of variance is up to 15% weaker for all models. One over-riding factor leading to these weaker regression results is that the datasets used in the original model development were collected over a relatively uniform canopy class (mixed wood and hardwood), using an equivalent survey configuration throughout. In the test dataset, the

vegetation species (Table 1), canopy attributes (Fig. 2) and survey configurations (Table 3) are highly variable and so an increased level of uncertainty or loss of predictive capability is to be expected. Some of the systematic factors that may influence the efficacy of each of the models are discussed below.

4.2. The influence of canopy structure

In an attempt to represent overhead canopy conditions only, the remainder of the analysis compares FC_{LiDAR} to FC_{HP} for annulus rings 1–3. The improvement in FC_{LiDAR} model capability from the return ratio models (RR and FR) to the intensity ratio (IR) to the Beer's Law modified intensity ratio (BL) is observed in Fig. 4. Further, when stratified by canopy structural class, each of the four models demonstrates different behaviour. For all FC_{LiDAR} models, the mixed wood/hardwood canopy class consistently displays the best results ($r^2 = 0.51$ to 0.80). For both conifer canopy classes, the regression results are generally weak for all four FC_{LiDAR} models ($r^2 = 0.14$ to 0.58). Furthermore, the nature of the FC_{LiDAR} model error differs, with the closed canopy conifer class tending towards a slight over-estimation in the FR, RR and IR models, with the open and short canopy conifer class tending toward a consistent and large over-estimate of FC_{HP} for all four models (Fig. 4).

The improvement in the magnitude of predictive capability from the $FC_{LiDAR(FR)}$ to $FC_{LiDAR(RR)}$ to $FC_{LiDAR(IR)}$ to $FC_{LiDAR(BL)}$ models is most clearly illustrated by plotting the magnitudes of the mean residuals. In Fig. 5 we observe that the mean overestimate in $FC_{LiDAR(FR)}$ for all data is 30%, for $FC_{LiDAR(RR)}$ it is 16%, for $FC_{LiDAR(IR)}$ it is 6%, while there is a small but not significant ($p = 0.05$) overestimate of <1% for $FC_{LiDAR(BL)}$. While the mean residuals for all canopy classes are positive for the FR, RR and IR models, the opposed behaviour of the three canopy classes is clear in the BL model. Therefore, by including these three conifer canopy classes in the analysis they have effectively compensated one another. This is important to note because it demonstrates that while the $FC_{LiDAR(BL)}$ model results are generally superior to the other three models when applied across a broad range of canopy classes, it may be prone to bias if applied exclusively to any distinct forest canopy class.

Fig. 4. Regression plots of FC_{HP} against FC_{LiDAR} models for each canopy structural class. Coloured lines represent the best fit regression line that passes through the origin for each of the canopy classes. The thick black line represents all data.
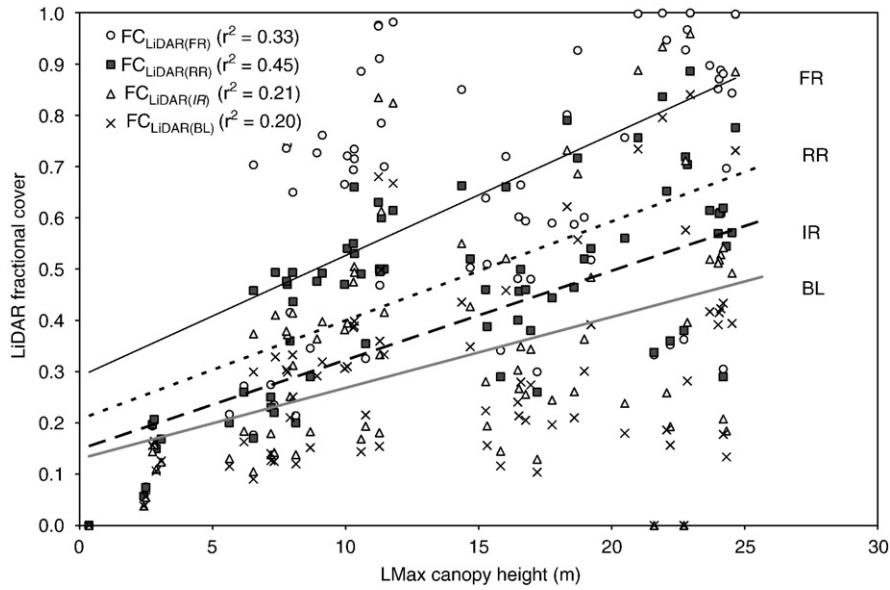
Fig. 6. The relationship between FC_{LiDAR} and LMax canopy height for all four models.

4.3. Influence of canopy height

Previous studies have indicated that LiDAR height and canopy cover are related for some ecosystems (e.g. Magnussen and Boude-wyn, 1998; Riaño et al., 2004; Thomas et al., 2006). Therefore, it is worth investigating whether or not plot-level canopy height influences the LiDAR-based models of FC. Before testing this, it was first necessary to establish whether or not there is any relationship between canopy height and FC_{HP}. Regression analyses were performed between plot-level mean maximum laser pulse return height above ground (LMax) and FC_{HP} for all annulus rings for all 245 HP photo stations within the 80 test plots. No strong relationships were found for any of the rings from 1 through to 9. The strongest relationship was with the average FC_{HP} rings 1 to 5 ($r^2 = 0.22$, $p<0.01$). It can be concluded, therefore, that height has at best a minimal autocorrelation with FC_{HP} measurements. This slight autocorrelation might be ex-

pected because as canopies mature, there are processes of self thinning, which allows trees to grow upwards and outwards until crown closure is reached (e.g. Chapin et al., 2002; Riaño et al., 2004). However, due to the wide variability in canopy structural classes and species mix, and limited variability in stand maturity displayed within the 80 test plots, this relationship is generally weak. Consequently, canopy height can not be assumed to be a good indicator of canopy fractional cover across the range of forest types examined in this study.

LMax height was regressed directly against the results of each of the four FC_{LiDAR} models to test the sensitivity of each model to variations in canopy height (Fig. 6). In all four cases, there was a statistically significant ($p<0.01$) relationship between LMax and FC_{LiDAR}. The two intensity-based FC_{LiDAR} models were only minimally sensitive to height with r^2 values of 0.21 and 0.20 for the Intensity Ratio and Beer's Law methods, respectively. This level of explanation of variance within the predicted FC_{LiDAR} data is similar to that
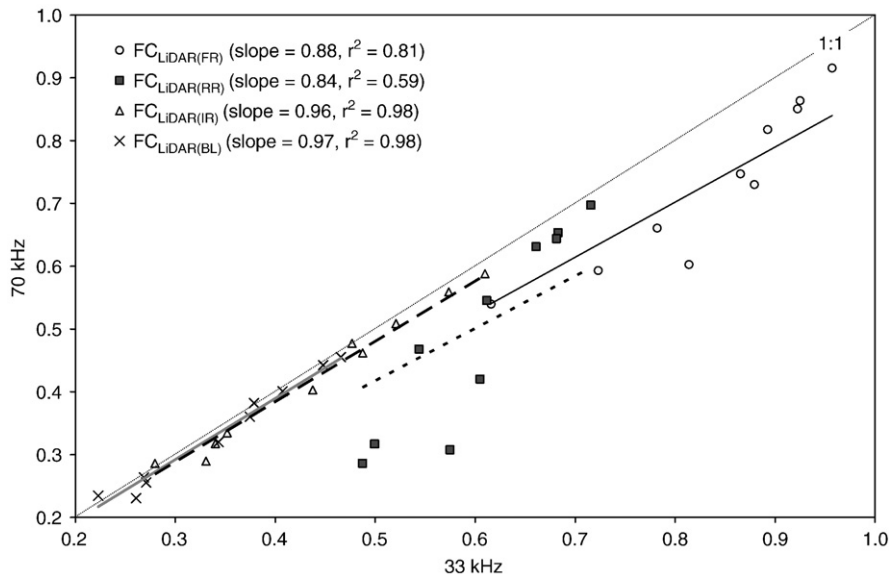

Fig. 7. The sensitivity of FC_{LiDAR} models to sensor pulse repetition frequency. Data collected in May 2006 over a red pine conifer plantation in the Annapolis Valley of Nova Scotia.

observed between LMax and the actual FC_{HP} values (max $r^2 = 0.22$). This similarity suggests that the intensity-based methods have a similar level of sensitivity to canopy height as exists in reality. Conversely, the pulse return ratio methods demonstrate increased sensitivity to variations in height, (with an r^2 values of 0.33 (FR) and 0.45 (RR)), which suggest that the incorporation of a height term might potentially improve these models. This enhanced sensitivity is likely due to the reduced probability of obtaining ground level returns in tall canopies; i.e. more returns are 'trapped' in taller canopies than shorter canopies. The intensity-based methods are less sensitive to height in this way because the differences in probability associated with taller and shorter canopies are essentially directly accounted for by the variations observed in the recorded intensity.

4.4. Influence of sensor configuration

The comparative FC_{LiDAR} results obtained from the 33 kHz and 70 kHz sensor pulse frequency configurations for 10 plots within a red pine plantation are illustrated in Fig. 7. By comparing the sample means, it was found that the difference in pulse power and sampling density associated with these two configurations produced signifi-

cantly different results for $FC_{LiDAR(RR)}$ (mean difference = 0.11, $p = 0.07$) and $FC_{LiDAR(FR)}$ (mean difference = 0.08, $p = 0.10$), while there is no significant difference for either of the intensity-based methods ($p = 0.72$ for $FC_{LiDAR(IR)}$ and $p = 0.81$ for $FC_{LiDAR(BL)}$). The regression plot in Fig. 7 illustrates that both intensity-based FC_{LiDAR} models are almost completely insensitive to pulse frequency. Meanwhile, both return ratio models (FR and RR) illustrate a tendency toward comparatively higher estimates of FC at the lower frequency or higher pulse power configuration of 33 kHz. Further the increased scatter in the FR and RR model observations in Fig. 7 implies that there is random behaviour inherent in the return ratio approaches that is mitigated by the inclusion of intensity data. This also suggests that one of the likely reasons for reduced coefficient of determination results in Table 4 and Fig. 3 for 'all data' in the $FC_{LiDAR(RR)}$ and $FC_{LiDAR(FR)}$ models are due to the variability of sensors and survey configurations used (Table 3).

The return ratio models demonstrate increased sensitivity to pulse repetition frequency due to the difference in the information content between a pulse return frequency distribution and an intensity power distribution. The ratio of canopy-to-total returns (models (2) and (3)) provides a direct quantification of the frequency of points that were reflected from the canopy vs. the total frequency of points reflected in
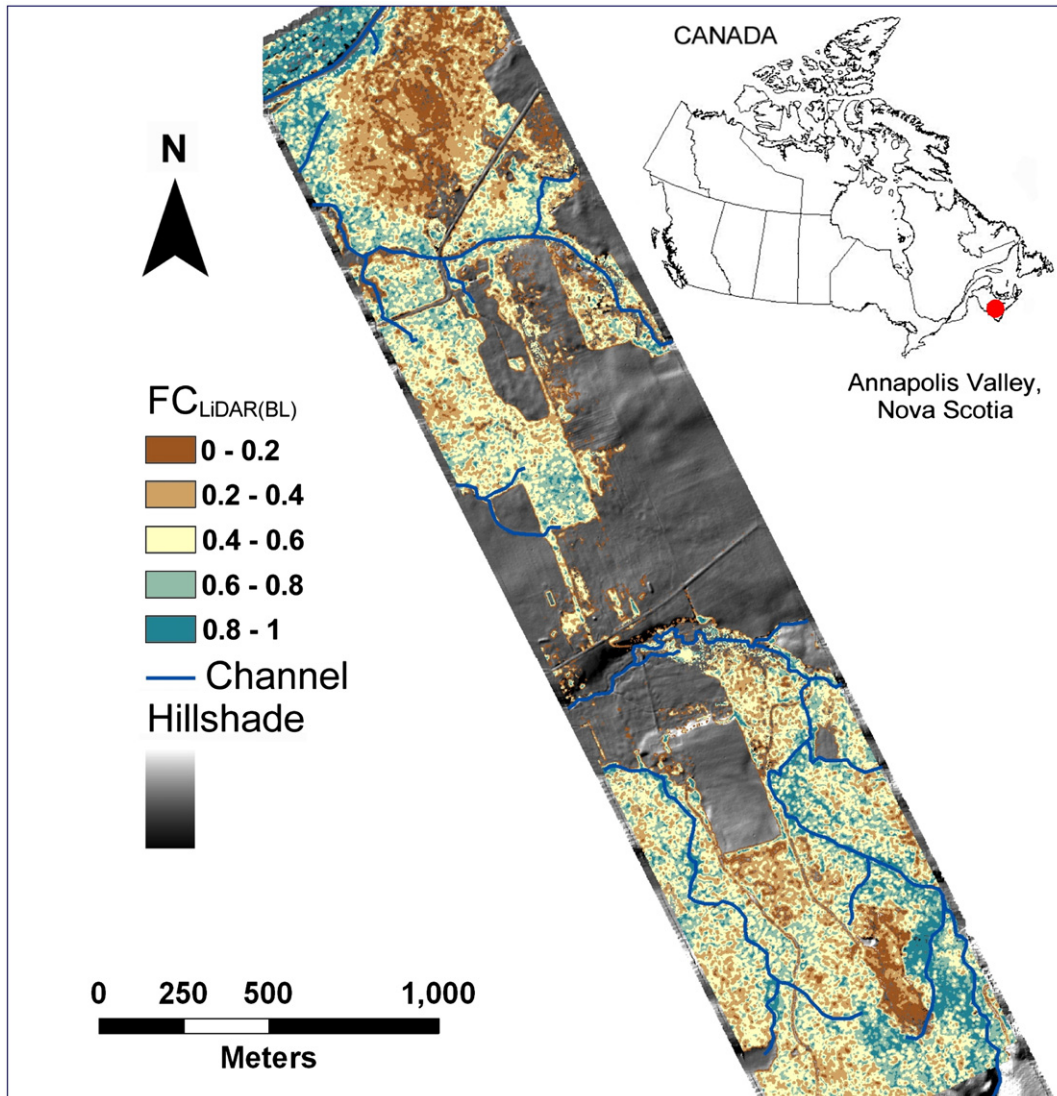

Fig. 8. A 1 m grid resolution image of $FC_{LiDAR(BL)}$ at the scale of individual canopy elements exceeding 1.3 m in height. Map illustrates a combined agricultural and forested Acadian mixed wood landscape in the Annapolis Valley of Nova Scotia during August of 2006.

the entire vertical profile. The results here indicate, therefore, that when operating at 33 kHz, the sensor systematically records a higher proportion of returns within the conifer canopy, than it does at 70 kHz, while the intensity power distribution remains virtually unchanged.

Statistically, there is no reason to assume that a lower sampling rate will lead to a shift of the frequency distribution toward the canopy, as all a lower sampling rate should do is reduce the density of points while maintaining the overall shape of the frequency distribution. However, the other important difference between these two configurations is that the peak pulse power at 33 kHz is almost six times greater than it is at 70 kHz (Hopkinson, 2007). This difference in pulse power directly influences the 'detectibility' of foliage elements. With increased pulse power (i.e. greater illumination), more of the small elements in the canopy that might otherwise go unnoticed at lower levels of power are able to reflect sufficient energy to register a return (either single, first, intermediate or last) at the sensor (Chasmer et al., 2006). At ground level, the surface area available for pulse reflection is effectively 100%, and so provided a pulse of sufficient energy makes it into and out of the canopy, there is a high likelihood that the same pulse will register a return at the sensor regardless of the actual pulse power. Therefore, at higher pulse power, the probability of increasing the number of first and intermediate returns from the canopy increases more than it will for ground-level returns; thereby systematically shifting the ratio toward increased fractional cover. By incorporating intensity into the model, these shifts in the distribution of pulse return probabilities are implicitly accounted for because the return intensity is primarily a function of pulse power and the contact area of the reflecting surface (Hopkinson, 2007).

A secondary influence on the absolute return intensity is spectral reflectance of the contact surface at the 1064 nm wavelength of the laser. However, although green foliage might be twice as reflective as soil at NIR wavelengths (Lillesand & Kiefer, 1994), the difference in laser pulse return contact area between a minimally detectable element of foliage (e.g. 35 cm^2) and the ground surface (e.g. 700 cm^2 at 1000 m a.g.l flying height) could easily exceed an order of magnitude. Further, it is not necessarily fair to assume that the sub canopy ground cover is bare soil, or that all foliage elements are green leaves.

Due to the sensor configuration test being applied only to a red pine plantation, the observations cannot be assumed equally applicable to all canopy types. However, it is clear that in at least this type of conifer plantation environment, estimates of FC derived from the widely adopted LiDAR pulse return ratio methods might be sensitive to the sensor configuration. In itself, this does not invalidate the FC$_{LiDAR(RR)}$ and FC$_{LiDAR(FR)}$ model approaches; it merely limits the ability to perform comparative analyses across LiDAR datasets of disparate sensor or data acquisition configurations. Alternatively, it implies that such models should be calibrated independently for each change in survey configuration.

5. Model application

To illustrate the utility of the Beer's Law modified intensity ratio model, Eq. (6) was applied to a LiDAR data set collected over the Annapolis Valley test sites to generate a map of canopy fractional cover. The LiDAR pulse return classes defined in Eq. (6) were rasterised at a 1 m grid resolution by summarising the laser pulse return intensity statistics within a moving window of 3 m radius. FC$_{LiDAR(BL)}$ was then mapped by integrating each of the raster layers according to Eq. (6) within ArcGIS (ESRI, Redlands, Calif.) (Fig. 8). The field FC$_{HP}$ estimates from 30 HP images collected coincident with the LiDAR dataset ranged from 0.2 to 0.95 with a mean of 0.58. No meaningful comparison can be made between the HP estimates and the LiDAR model, as these HP data were already used in the initial test of the model.

The area illustrated in Fig. 8 is a rural agricultural landscape that has undergone substantial anthropogenic modification. Areas of vegetation lower than 1.3 m in height are not represented here and in their

place the underlying grey-scale hill shade is visible. However, in the remaining 'tall' canopy environments FC heterogeneity is clearly mapped. Trails are evident within the wooded area at the south of the image as lines of low FC. The impacts of both selective and small area clearcut logging are visible both on the north and south ends of the map. Further, the high FC associated with riparian buffers are also visible and it is possible to directly quantify the extent and foliage cover within these zones. While mapping FC over agricultural landscapes like this may not be of immediate obvious practical application, the ability to map these features at this resolution and at the levels of accuracy implied by the tests above is powerful because it provides opportunities for better understanding canopy-level ecological processes and interactions at the landscape scale.

6. Concluding remarks

The results of the tests performed over multiple ecozones and datasets elucidate the sensitivities and range of application of LiDAR-based models of canopy fractional cover. Published canopy-to-total return ratio (FC$_{LiDAR(RR)}$ and FC$_{LiDAR(FR)}$) methods have demonstrated strong correlations with canopy fractional cover for single datasets collected over homogeneous canopy types. However, it is shown here that significant improvements to these model approaches can be achieved by incorporating the pulse return intensity information (FC$_{LiDAR(IR)}$). Moreover, if the intensity power distribution is modified using a turbid medium Beer's Law approach (FC$_{LiDAR(BL)}$) to account for secondary return two-way pulse transmission losses within the canopy, the resultant model requires no calibration and provides a 1:1 direct estimate of overhead fractional cover for a range of canopy types and sensor configurations.

Of the four models tested, the two return ratio methods (FC$_{LiDAR(RR)}$ and FC$_{LiDAR(FR)}$) were most sensitive to canopy height, canopy structural class and sensor configuration. In all cases, the intensity-based methods proved less sensitive and generally more stable. This is because the intensity methods implicitly provide some quantification of the surface area that the pulse is interacting with in the form of the reflectance amplitude (intensity). The position of the return in the canopy and its proximity to other canopy and ground returns contains no direct information regarding the amount of foliage it has reflected off and thus a barely detectable return that has reflected off a twig or leaf carries the same weight as a highly detectible return from an area of dense foliage or ground.

These observations do not invalidate use of the pulse return ratio models in situations where the intensity information might not be available but it is recommended that it should be independently calibrated whenever different configurations, canopy structural classes, or canopy heights are encountered. Within homogeneous canopies, any of the four models tested could supply accurate FC results with appropriate calibration. If intensity and return sequence information is available to the end user, then it is recommended that FC$_{LiDAR}$ models utilise this information to ensure most accurate results. However, caution would still need to be exercised where heterogeneous canopies are encountered given the intensity methods demonstrate some sensitivity to canopy structure. For the Beer's Law modified intensity method (FC$_{LiDAR(BL)}$), it might be reasonable to assume a randomly foliated mixed wood canopy acts something like a turbid medium but this is obviously an invalid assumption for an open and highly clumped conifer canopy typical of some of the black spruce plots sampled in this study. It is probably also invalid to assume that the closed canopy conifer plots would behave like a turbid medium, as many of these sample plots were within plantation environments for which there is a distinct single storey canopy with no under-storey. Nonetheless, of the models tested, the Beer's Law modified intensity ratio approach provides the most robust first approximation across a range of canopy classes and is the only model tested that can potentially provide an accurate estimate of fractional cover without the need for training data.

Acknowledgements

References

ASPRS (2005). *LAS specification version 1.1 March 07, 2005.* Data format documentation created by the LiDAR sub-committee of the American Society of Photogrammetry and Remote Sensing Available online: http://www.asprs.org/society/committees/lidar/lidar_downloads.html 11 pp. Last accessed on July 1st, 2008.

Axelsson, P. (1999). Processing of laser scanner data algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing, 54*, 138–147.

Barilotti, A., Turco, S., & Alberti, G. (2006). LAI determination in forestry ecosystems by LiDAR data analysis. *Workshop on 3D Remote Sensing in Forestry, 14–15/02/2006, BOKU Vienna.*

Barr, A. G., Griffis, T. J., Black, T. A., Lee, X., Staebler, R. M., Fuentes, J. D., et al. (2002). Comparing the carbon balances of boreal and temperate deciduous forest stands. *Canadian Journal of Forest Research, 32*, 813–822.

Chasmer, L., Barr, A., Hopkinson, C., McCaughey, H., Treitz, P., & Black, A. (2009). Scaling and assessment of GPP from MODIS using a combination of airborne LiDAR and eddy covariance measurements over jack pine forests. *Remote Sensing of Environment, 113*, 82–93.

Chasmer, L., Hopkinson, C., Smith, B., & Treitz, P. (2006). Examining the influence of changing laser pulse repetition frequencies on conifer forest canopy returns. *Photogrammetric Engineering and Remote Sensing, 17*(12), 1359–1367.

Chasmer, L., McCaughey, H., Barr, A., Black, A., Shashkov, A., Treitz, P., et al. (2008a). Investigating light use efficiency (LUE) across a jack pine chronosequence during dry and wet years. *Tree Physiology, 28*, 1395–1406.

Chasmer, L., Hopkinson, C., Treitz, P., McCaughey, H., Barr, A., & Black, A. A. (2008b). A lidarbased hierarchical approach to assessing MODIS fPAR. *Remote Sensing of Environment, 112*, 4344–4357.

Chasmer, L., Kljun, N., Barr, A., Black, A., Hopkinson, C., & McCaughey, H. (in press). Vegetation structural and elevation influences on CO₂ uptake within a mature jack pine forest in Saskatchewan, Canada. *Canadian Journal of Forest Research.*

Chapin, F. S., III, Matson, P. A., & Mooney, H. A. (2002). *Principles of Terrestrial Ecosystem Ecology.* New York: Springer–Verlag Inc. 436 pp.

Chen, J. M. (1996). Optically-based methods for measuring seasonal variation in leaf area index of boreal conifer forests. *Agricultural and Forest Meteorology, 80*, 135–163.

Chen, J. M., Chen, X., Ju, W., & Geng, X. (2005). A remote sensing-driven distributed hydrological model: Mapping evapotranspiration in a forested watershed. *Journal of Hydrology, 305*, 15–39.

Chen, B., Chen, J. M., Mo, G., Yuan, K., Higuchi, K., & Chan, D. (2007). Modeling and scaling coupled energy, water, and carbon fluxes based on remote sensing: An application to Canada's landmass. *Journal of Hydrometeorology, 8*, 123–143.

Chen, J. M., Govind, A., Sonnentag, O., Zhang, Y., Barr, A., & Amiro, B. (2006). Leaf area index measurements at Fluxnet-Canada forest sites. *Agricultural and Forest Meteorology, 140*, 257–268.

Fernandes, R. A., Miller, J. R., Chen, J. M., & Rubinstein, I. G. (2004). Evaluating image based estimates of leaf area index in boreal conifer stands over a range of scales using high-resolution CASI imagery. *Remote Sensing of Environment, 89*, 200–216.

Gower, S. T., Kucharik, C. J., & Norman, J. M. (1999). Direct and indirect estimation of leaf area index, f_APAR, and net primary production of terrestrial ecosystems. *Remote Sensing of Environment, 70*, 29–51.

Hall, F. G., Botkin, D. B., Strebel, D. E., Woods, K. D., & Goetz, S. J. (1991). Large scale patterns of forest succession as determined by remote sensing. *Ecology, 72*(2), 628–640.

Heinsch, F. A., Reeves, M., Bowker, C. F., Votava, P., Kang, S., Milesi, C. et al., 2003. User's Guide, GPP and NPP (MOD17A2/A3, NASA MODIS Land Algorithm, Version 1.2. www.ntsg.umt.edu/modis/MOD17UsersGuide.pdf

Heinsch, F. A., Zhao, M., Running, S. W., Kimball, J. S., Nemani, R. R., Davis, K. J., et al. (2006). Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations. *IEEE Transactions on Geoscience and Remote Sensing, 44*(7), 1908–1925.

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., & Chasmer, L. E. (2007). Using discrete laser pulse return intensity to model canopy transmittance. *The Photogrammetric Journal of Finland, 20*(2), 16–26.

Hopkinson, C., & Demuth, M. D. (2006). Using airborne LiDAR to assess the influence of glacier downwasting to water resources in the Canadian Rocky Mountains. *Canadian Journal of Remote Sensing, 32*(2), 212–222.

Hopkinson, C., Chasmer, L. E., & Hall, R. J. (2008). The uncertainty in conifer plantation growth prediction from multitemporal LiDAR datasets. *Remote Sensing of Environment, 112*(3), 1168–1180.

Hopkinson, C., Chasmer, L. E., Lim, K., Treitz, P., & Creed, I. (2006). Towards a universal LiDAR canopy height indicator. *Canadian Journal of Remote Sensing, 32*(2), 139–153.

Hopkinson, C., Chasmer, L. E., Zsigovics, G., Creed, I., Sitar, M., Kalbfleisch, W., et al. (2005). Vegetation class dependent errors in LiDAR ground elevation and canopy height estimates in a Boreal wetland environment. *Canadian Journal of Remote Sensing, 31*(2), 191–206.

Hopkinson, C., Sitar, M., Chasmer, L. E., & Treitz, P. (2004). Mapping snowpack depth beneath forest canopies using airborne LiDAR. *Photogrammetric Engineering and Remote Sensing, 70*(3), 323–330.

Kotchenova, S., Shabanov, N., Knyazikhin, Y., Davis, A., Dubayah, R., & Myneni, R. (2003). Modeling LiDAR waveforms with time-dependent stochastic radiative transfer theory for estimations of forest structure. *Journal of Geophysical Research, 108*(D15), 4484. doi:10.1029/2002JD003288, 2003.

Kotchenova, S., Song, X., Shabanov, N., Potter, C., Knyazikhin, Y., & Myneni, R. (2004). LiDAR remote sensing for modeling gross primary production of deciduous forests. *Remote Sensing of Environment, 92*, 158–172.

Kusakabe, T., Tsuzuki, H., Hughes, G., & Sweda, T. (2000). Extensive forest leaf area survey aiming at detection of vegetation change in subarctic-boreal zone. *Polar Bioscience, 13*, 133–146.

Law, B. E., Falge, E., Gu, L., Baldocchi, D., Bakwin, P., Berbigier, P., et al. (2002). Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation. *Agricultural and Forest Meteorology, 113*, 97–120.

Leblanc, S. G., Chen, J. M., Fernandes, R., Deering, D., & Conley, A. (2005). Methodology comparison for canopy structure parameters extraction from digital hemispherical photography in boreal forests. *Agricultural and Forest Meteorology, 129*, 187–207.

Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., Spies, T. A., & Harding, D. (1999). LiDAR remote sensing of the canopy structure and biophysical properties of Douglas-fir Western hemlock forests. *Remote Sensing of Environment, 70*, 339–361.

Lefsky, M. A., Turner, D., Guzy, M., & Cohen, W. (2005). Combining LIDAR estimates of aboveground biomass and Landsat estimates of stand age for spatially extensive validation of modeled forest productivity. *Remote Sensing of Environment, 95*, 549–558.

Leuning, R., Cleugh, H. A., Zegelin, S. J., & Hughes, D. (2005). Carbon and water fluxes over a temperate Eucalyptus forest and a tropical wet/dry savannah in Australia: measurements and comparison with MODIS remote sensing estimates. *Agricultural and Forest Meteorology, 129*, 151–173.

Lillesand, T. M., & Kiefer, R. W. (1994). *Remote Sensing and Photo Interpretation*, 3rd. ed. New York: John Wiley & Sons 750 pp.

Lovell, J., Jupp, D., Culvenor, D., & Coops, N. (2003). Using airborne and groundbased ranging LiDAR to measure canopy structure in Australian forests. *Canadian Journal of Remote Sensing, 29*(5), 607–622.

Magnussen, S., & Boudewyn, P. (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research, 28*, 1016–1031.

Morsdorf, F., Kotz, B., Meier, E., Itten, K. I., & Allgower, B. (2006). Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. *Remote Sensing of Environment, 104*(1), 50–61.

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, 91*, 243–255.

Nelson, R., Krabill, W., & Maclean, G. (1984). Determining forest canopy characteristics using airborne laser data. *Remote Sensing of Environment, 15*, 201–212.

Parker, G. G., Lefsky, M. A., & Harding, D. J. (2001). Light transmittance in forest canopies determined using airborne laser altimetry and in-canopy quantum measurements. *Remote Sensing of Environment, 76*, 298–309.

Pomeroy, J. W., & Dion, K. (1996). Winter radiation extinction and reflection in a boreal pine canopy: measurements and modelling. *Hydrological Processes, 10*, 1591–1608.

Pomeroy, J. W., Granger, R. J., Hedstrom, N. R., Gray, D. M., Elliott, J., Pietroniro, A., & Janowicz, J. R. (2005). The process hydrology approach to improving prediction of ungauged basins in Canada. In C. Spence, J. Pomeroy, & A. Pietroniro (Eds.), *Predictions in Ungauged Basins: Approaches for Canada's Cold Regions.* Canada: Canadian Society for Hydrological Sciences, Environment pp. 67–100.

Popescu, S., Wynne, R., & Nelson, R. (2003). Measuring individual tree crown diameters with LiDAR and assessing its influence on estimation forest volume and biomass. *Canadian Journal of Remote Sensing, 29*(5), 564–577.

Quinton, W. L., Shirazi, T., Carey, S. K., & Pomeroy, J. W. (2005). Soil water storage and active-layer development in a sub-alpine tundra hillslope, southern Yukon Territory, Canada. *Permafrost and Periglacial Processes, 16*, 369–382.

Riaño, D., Valladares, F., Condes, S., & Chuvieco, E. (2004). Estimation of leaf area index and covered ground from airborne laser scanner (LiDAR) in two contrasting forests. *Agricultural and Forest Meteorology, 124*(3–4), 269–275.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M. C., & Hashimoto, H. (2004). A continuous satellite-derived measure of global terrestrial primary production. *BioScience, 54*, 547–560.

Schwalm, C. R., Black, T. A., Amiro, B. D., Arain, M. A., Barr, A. G., & Bourque, C. P. A. (2006). Photosynthetic light use efficiency of three biomes across an east-west continental-scale transect in Canada. *Agricultural and Forest Meteorology, 140*, 269–286.

Solberg, S., Næsset, E., Hanssen, K. H., & Christiansen, E. (2006). Mapping defoliation during a severe insect attack on Scots pine using airborne laser scanning. *Remote Sensing of Environment, 102*, 364–376.

Thomas, V., Treitz, P., McCaughey, J. H., & Morrison, I. (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*(1), 34–47.

Tian, Y., Wang, Y., Zhang, Y., Knyazikhin, Y., Bogaert, J., & Myneni, R. (2002). Radiative transfer based scaling of LAI retrievals from reflectance data of different resolutions. *Remote Sensing of Environment, 84*, 143–159.

Tucker, C. J., Fung, I. Y., Keeling, C. D., & Gammon, R. H. (1986). Relationship between CO_2 variation and a satellite-derived vegetation index. *Nature*, *319*, 195−199.

Todd, K. W., Csillag, F., & Atkinson, P. M. (2003). Three-dimensional mapping of light transmittance and foliage distribution using LiDAR. *Canadian Journal of Remote Sensing*, *29*(5), 544−555.

Vosselman, G. (2000). *Slope based filtering of laser altimetry data*, *Vol. XXXIII*. (pp. 935−942) Amsterdam, The Netherlands: ISPRS Part B3.

Zhang, Y., Chen, J., & Miller, J. (2005). Determining digital hemispherical photograph exposure for leaf area index estimation. *Agricultural and Forest Meteorology*, *133*, 166−181.