Canadian Journal of Forest Research
Revue canadienne de recherche forestière

Discrimination of canopy gaps and non-regenerating openings in old-growth temperate deciduous forests using airborne LiDAR data

Journal:	<i>Canadian Journal of Forest Research</i>
Manuscript ID	cjfr-2017-0340.R2
Manuscript Type:	Article
Date Submitted by the Author:	10-Apr-2018
Complete List of Authors:	Senécal, Jean-François; Université du Québec en Outaouais, Sciences naturelles; ISFORT; Université du Québec à Montréal Doyon, Frédéric; Université du Québec en Outaouais, St-Onge, Benoît; Department of Geography,
Keyword:	gap regime, gap fraction, canopy gap detection, micro-topography, disturbance regime
Is the invited manuscript for consideration in a Special Issue? :	N/A

SCHOLARONE™
Manuscripts

1 Discrimination of canopy gaps and non-regenerating openings in old-growth temperate
2 deciduous forests using airborne LiDAR data

3 Jean-François Senécal^{1,2}, Frédéric Doyon^{2,3} and Benoît St-Onge⁴

4 ¹Département des sciences biologiques, Université du Québec à Montréal, 141 Avenue
5 du Président-Kennedy, Montréal, Québec, H2X 1Y4, Canada

6 ²Institut des sciences de la forêt tempérée, 58, rue Principale, Ripon, Québec, J0V 1V0,
7 Canada

8 ³Département des sciences naturelles, Université du Québec en Outaouais, 283,
9 boulevard Alexandre-Taché, Gatineau, Québec, J8X 3X7, Canada

10 ⁴Département de Géographie, Université du Québec à Montréal, 1255 rue Saint-Denis,
11 Montréal, Québec, H2X 3R9, Canada

12

13 Jean-François Senécal: senecal.jean-francois.3@courrier.uqam.ca

14 Frédéric Doyon: frederik.doyon@uqo.ca

15 Benoît St-Onge: st-onge.benoit@uqam.ca

16 Corresponding author: Jean-François Senécal, ²Institut des sciences de la forêt tempérée,
17 58, rue Principale, Ripon, QC, J0V 1V0, Canada.

18

19 KEYWORDS: gap regime, gap fraction, canopy gap detection, micro-topography,
20 disturbance regime

21

22

Abstract

The use of LiDAR in the study of gap regimes has seen significant progress in recent years. Researchers have assumed that canopy gaps that are detected in this manner were ecologically equivalent to gaps sampled *in situ* by more traditional methods. However, those latter methods usually include canopy gaps only and ignore non-regenerating openings that are produced by causes limiting tree establishment. We developed a predictive model capable of discriminating between canopy gaps and non-regenerating openings using LiDAR-derived data. Selected predictive variables were related to conditions that limit tree establishment, such as zones of moisture accumulation and steep slopes, or to the resulting vegetation physiognomy. The model was applied to three old-growth forests to predict the fractions of canopy openings belonging to these two types. On average, non-regenerating openings represented 19.5% of the total area detected as canopy openings and occupied 1.37% of the sites. Canopy gaps formed 80.5% of the total area in canopy openings and covered 5.71% of sites that were studied. The non-regenerating opening seemed more frequent on thin surficial deposits. The canopy gap fraction was similar among study sites but had lower values than usually reported for temperate deciduous forest.

Résumé

L'utilisation du LiDAR pour l'étude des régimes des trouées a connu une importante progression au cours des dernières années. Généralement, les chercheurs ont adopté de façon implicite la prémisse que les trouées détectées ainsi étaient équivalentes, sur la base de leur écologie, aux trouées échantillonnées *in situ* selon des méthodes plus traditionnelles. Cependant, ces dernières tiennent seulement compte des trouées dans la canopée créées par la mort d'au moins un arbre mais non des ouvertures édaphiques qui sont produites par certaines conditions de site empêchant l'établissement et la croissance d'arbres. Nous avons produit un modèle prédictif capable de

47 distinguer les ouvertures édaphiques des trouées biotiques à partir de variables dérivées de
48 données LiDAR. Les variables prédictives retenues dans le modèle sont liées à des conditions de
49 site qui préviennent l'établissement d'arbres, comme les zones d'accumulation d'eau et les pentes
50 fortes, ou sont liées à la physionomie de la végétation qui en résulte. Le modèle a été appliqué à
51 trois sites érablières anciennes situés dans le sud-ouest du Québec. Les ouvertures édaphiques y
52 représentent 19.5% de la surface totale des trouées et occupent 1.37% de la surface des sites en
53 moyenne. Les trouées biotiques forment 80.5 % de la surface totale détectée en trouées et
54 couvrent 5.71% des sites étudiés. La proportion des sites en ouvertures édaphiques semble
55 influencée par le type de dépôt de surface. La proportion en trouées biotiques est similaire entre
56 les sites étudiés, mais est plus faible que les valeurs habituellement rapportées pour la forêt
57 feuillue tempérée.

58

Draft

Introduction

The natural disturbance regime in the temperate deciduous forest is characterized by gap dynamics, which is the process of continuous creation and subsequent filling of canopy gaps or developmental gaps (*sensu* Lertzman et al. 1996). Canopy gaps are canopy openings created by tree damage, either through tree death or branch loss (Runkle 1985). However, openings in the canopy are not necessarily canopy gaps. In the case of *edaphic openings*, the opening is usually associated with edaphic conditions that limit or prevent the establishment of tree regeneration (Lertzman et al. 1996). Tree growth and establishment in edaphic openings can be limited by soil quality (*e.g.* thickness, richness, compactness, and moisture) (Bartemucci et al. 2002) or by non-edaphic factors such as the presence of dense non-tree vegetation communities (Engelman and Nyland 2006; Royo and Carson 2006). For clarity, the term ‘Non-Regenerating Openings’ (NRO) will hereafter be used to designate canopy openings not caused by tree damage and having such edaphic and/or vegetation characteristics. Field-ecologists have often deliberately excluded NRO from their analyses because they don’t provide the same regeneration opportunities as usual canopy gaps. These can close by growth of trees from the bottom and by lateral inward extension of peripheral branches, while NRO can close exclusively by lateral inward extension unless the conditions limiting tree growth and establishment eventually wane. If NRO close from lateral growth of branches, they will likely remain devoid of tree regeneration for a long period of time despite appearing closed when imaged from above. It can be argued that closed spaces with limiting conditions have different dynamic than closed canopy gaps. Therefore, it is crucial to distinguish between these two types of canopy opening in order to fully and correctly understand and characterize gap dynamics.

Remote sensing has been used increasingly for automated canopy gap detection and characterization (St-Onge et al. 2014; Zhang 2008). Canopy gap detection with remote sensing presents a challenge, especially when gaps are small, because features that distinguish canopy

gaps from the rest of the canopy are mostly structural (Koukoulas and Blackburn 2004) . LiDAR (Light Detection And Ranging) sensors directly record information about the three-dimensional geometry of targets. This method is useful because of its capacity to measure the elevation of the canopy as well as that of the underlying terrain, both with great accuracy and high spatial resolution (Lim et al. 2003). These desirable properties explain the increase of LiDAR usage in recent years for the study of forest gap regimes (Vepakomma et al. 2010).

To the best of our knowledge, no attempts have been made to differentiate canopy gaps from NRO using LiDAR data. As a result, in past studies in which LiDAR was used to detect canopy gaps, an unknown proportion of NRO could likely have been counted as canopy gaps. This discrepancy between field studies and LiDAR studies might result in problems when applying the principles of ecosystemic management because the two wouldn't be readily comparable. As forest managers are increasingly relying on LiDAR data to plan their operations and produce reliable inventories, there is a need for new technologies to be compatible with knowledge derived from older procedures or data.

In the study area selected for the present study, we observed that NRO have structural features that differentiate them from canopy gaps. First, vegetation height is usually greater in canopy gaps than in NRO. This is true even in recently created canopy gaps because of the omnipresence of advance regeneration of sugar maple (*Acer saccharum* Marsh.) and American beech (*Fagus grandifolia* Ehrh.) under closed canopy (Senecal *et al.* 2018a). Subcanopy is often undamaged following canopy height loss and vegetation height within canopy gaps thus remains well above ground level (Senecal *et al.* 2018b). Second, most NRO with signs of soil moisture saturation are close to streams or lie within concave terrain features favoring water accumulation or conducive of poor drainage. It is expected that water features are a major cause of NRO (Lertzman et al. 1996). We expect these conditions to be limiting for tree establishment and growth within most forest biomes that are not made up of trees adapted to such conditions (*e.g.* mangroves). Third,

NRO caused by rock outcrops often present a convex shape or are at the upper edge of a steep slope. Such edaphic features are expected to create NRO (Bartemucci et al. 2002). These three characteristics can be theoretically detected by proper analysis of LiDAR data. With short term multi-temporal LiDAR analysis, canopy height loss between LIDAR acquisitions can be reliably attributed to tree damage (Senecal et al. 2018b). However, multi-temporal LiDAR data is unlikely to be available in most locations, thus diminishing the usefulness of methods requiring such data. Also, canopy openings that remain open throughout the time interval covered by a short term multi-temporal dataset cannot be ascribed with reasonable certainty to the canopy gap or NRO classes. There is therefore a need to distinguish NRO from canopy gaps using means readily available, on the basis of NRO and canopy gap characteristics. The goal of this paper was to propose a methodological approach for discriminating NRO from canopy gaps using mono-temporal LiDAR data, assess the accuracy of this automated opening classification, and to study the frequency and characteristics of NRO in the selected study area.

Methods

Study sites

The study area consists of three sites approximately 1 km² each in Southwestern Quebec, Canada. Two sites (A and B) are located in the Forêt-la-Blanche Ecological Reserve (45°44'N, 75°16'W), a protected area of 2,052 ha dedicated to the conservation of ecosystems and research. The third site (C) is located in the Lac-de-l'Écluse Exceptional Forest Ecosystem (45°52'N, 75°24'W), which is also under conservation. Sites A and B are 1.4 km apart, while site C is 16 km north of the first two. Half of site A is characterized by flat terrain covered by a fine-texture outwash surficial deposit, with the rest being slightly hilly terrain with thin to very thin undifferentiated till. Site B and C are more hilly and covered by undifferentiated till, which varies in thickness from moderate to very thin deposits over bedrock. Mean annual temperature varies from 2.5 to 5.0°C, with a growing season (daily average over 5.6°C) of 180-190 days (Robitaille and Saucier

1998). Mean annual precipitation ranges from 900 to 1000 mm, with a third of it falling as snow. The dominant tree species in the three sites is sugar maple (*Acer saccharum* Marsh.), often in combination with American beech (*Fagus grandifolia* Ehrh.). Other notable species are basswood (*Tilia americana* L.), Eastern hemlock (*Tsuga canadensis* (L.) Carrière), ironwood (*Ostrya virginiana* (Mill.) K. Koch), red maple (*A. rubrum* L.), striped maple (*A. pensylvanicum* L.), white ash (*Fraxinus americana* L.) and yellow birch (*Betula alleghaniensis* Britt.). The three study sites are old-growth forests that have no known catastrophic disturbances. The three sites were affected by an exceptionally severe ice storm in 1998, receiving 40-100 mm of freezing rain (Olthof et al. 2004) that damaged some trees. Furthermore, beech bark disease (Houston 1994) has been present for several years in the area. During our 2009 and 2010 field surveys, many of the large beeches in the study sites were dead or dying.

LiDAR data processing and canopy openings detection

Airborne discrete return LiDAR data were acquired in September 2007 using an Optech ALTM 3100 instrument operating at an average altitude of 1300 m. Average point density was over two returns per m², while maximum half-scan angle was 20°, scan rate was 41 Hz, and line spacing was 750 m (for a targeted 50% overlap between strips). This resulted in a 2.5 to 4 first return points per m² because of flight line overlap. Preprocessing of the LiDAR data was accomplished with the Terrascan software package (Terrasolid, Leppävaara Finland); this process included data cleaning, atmospheric/bird hit removal, and classification of LiDAR points as ground level hits using the morphological properties of the point cloud. Ground hits density was between 0.2 to 0.3 points per m².

Data processing to produce height models and to detect canopy openings were carried out in ArcGIS Desktop (v. 9.1, Environmental Systems Research Institute Inc., Redlands, CA). A digital surface model (DSM) was initially created from first return LiDAR points (Lim et al. 2003). The DSM was produced by interpolation based on a TIN (Triangulated Irregular Network) rasterized

to a resolution of 50 cm. This resolution is considered appropriate in regard of the objectives of this study and the LiDAR point density of our dataset (Vepakomma et al. 2008). Using the same procedure, a digital terrain model (DTM) was produced using the points classified as ground returns. The canopy height model (CHM), a high-resolution raster representation of the height of vegetation above the ground, was then obtained by subtracting the DTM from the DSM. The LiDAR sensor was carefully calibrated before data acquisition. XYZ errors should be within the normal range for the ALTM3100.

Openings in the canopy were identified using an adaptive height thresholding algorithm. This algorithm was designed to detect depressions in the CHM, the edges of which are steep enough to represent a break compared to the surrounding canopy structure. Using this algorithm, a CHM pixel was classified as part of a canopy opening if its height was less than one of two height thresholds (see equations 1 and 2) that were computed using the pixels in a circular neighborhood of 0.25 ha:

$$[1] \quad H_i < [H_{1/4ha} - (1.5 * SD_{H_{1/4ha}})]$$

$$[2] \quad H_i < (0.20 * H_{max_{1/4ha}})$$

Where H_i is the CHM pixel height value being evaluated, $H_{1/4ha}$ is the height average of the CHM in an area of 0.25 ha around the pixel, and $SD_{H_{1/4ha}}$ and $H_{max_{1/4ha}}$ are the standard deviation and maximum value of the CHM in an area of 0.25 ha around the pixel, respectively. Contiguous pixels forming a distinct gap were then aggregated as canopy opening polygons. Canopy opening size was calculated from the total area of grouped pixels. Canopy openings measuring less than 4 m² were eliminated to remove artifacts caused by variations in LiDAR point density or resulting from the natural porosity of the canopy. The values of the constants in equations 1 and 2 (1/4 ha, 1.5, and 0.20), were determined empirically by exploring the value space until a satisfactory visual match with canopy openings in the CHM was achieved. The accuracy of canopy opening's boundaries was not assessed quantitatively *per se*, but visual observations

during field sampling indicated that it was sufficient for our purpose. Studies that have investigated this aspect obtained a good fit between canopy gaps detected by LiDAR and canopy gaps in the field (Vepakomma et al. 2008).

Field data collection

A subset of canopy openings detected using the detection algorithm were surveyed in the forests in October 2010. Because we surveyed gaps detected using the 2007 LiDAR data and a GPS, only gaps that opened prior to 2007 should have been sampled. Some gaps might have enlarged during the 3 years interval, although we did not see signs of recent gap enlargement during fieldwork. The subset contained openings with a mean CHM value lower than 5 m. It was assumed that most canopy openings with a mean height higher than 5 m would be canopy gaps rather than NRO. This was done to speed up fieldwork as we expected NRO to be fairly rare. Our operational definition of NRO was based on two conditions. First, the area with tree growth limiting features showed no signs of having recently supported canopy trees. Second, a substantial part (in this study, >30%) of its vertical projection had to be covered by features that limit the establishment of tree regeneration. Limiting factors found during the field survey were water (i.e. accumulation, stream, seepage), browsing by beaver (*Castor canadensis*), rocks (i.e., outcrop, boulders), and vegetation communities inhibiting plant succession (e.g., ferns (often *Athyrium felix-femina*), *Viburnum* spp., *Corylus cornuta* Marsh., *Carex* spp., *Polygonum* spp.) (Fig. 1). Browsing by beavers was included because it limits tree growth for a considerable period of time. Beaver population densities are high in the region where the study was conducted (Lafond et al. 2003). Each NRO could have more than one type of limiting factor (e.g., rocks and ferns).

Previous studies designated canopy openings as being edaphic when a ‘substantial’ part of them had an edaphic component detrimental to growth (Lertzman et al. 1996), without any quantitative threshold. We selected a 30% threshold after observing that limiting features in NRO were often

surrounded by a growth-inhibiting transition zone with small trees but no canopy trees. When accumulated, the limiting feature's area and the transition zone usually accounted for the major part of the opening area. Openings not fitting these criteria were labeled as canopy gaps. Very small openings (between 4 m² and 20 m² were not sampled in the field because of difficulties in positioning them precisely). Most of these very small openings were in high canopies and appeared to have been created by broken branches or were old canopy gaps in the process of closing laterally. For each valid canopy opening visited, we recorded the presence of establishment- and growth-limiting factors, as well as tree species composition of the surrounding canopy (Fig. 2). We collected data on 105 NRO and 107 canopy gaps.

Data preparation for NRO detection

We chose LiDAR-derived vegetation and terrain structural variables based on our observations of NRO features. CHM height statistics (minimum, maximum, median, 20th and 80th percentiles) were calculated within canopy openings outlines in SAGA-GIS version 2.0.5 (Conrad et al. 2015) in order to characterize intra-gap vegetation height distribution.

The shape of the terrain was described using a topographic position index (TPI) computed from the DTM (Jenness et al. 2005). The TPI is a measure of the height difference between a cell and the height average in a circular neighborhood of a given radius. It has a negative value for concave terrain and a positive value for convex terrain. We chose a radius of 5 m based on field observations of the size of topographic features related to NRO. Minimum, median and maximum TPI were calculated for each canopy openings.

The LiDAR DTM was used to calculate flow direction and flow accumulation datasets. From these observations, a flow accumulation threshold value was chosen to obtain the best prediction of the actual fine-scale stream network (Jenson and Domingue 1988). Based on this, a nearest-distance-to-nearest-water raster was then produced at 50 cm resolution and was used to calculate

distances (minimum, maximum and median) for each opening relative to the closest predicted stream. These three statistics can have different interpretations depending on canopy opening shape and orientation.

The LiDAR DTM was used to produce a slope raster and statistics for each opening were computed (minimum, maximum and median slope). Because of its nonlinear relationship with the probability of detecting NRO (Battles et al. 1995), the maximum slope variable was split into three categories according to the following scheme (low slope class 1 = $0.0 \leq \text{slope} \leq 24.8$ degrees, intermediate slope class 2 = $24.8 < \text{slope} \leq 39.8$ degrees, high slope class 3 = slope > 39.8 degrees).

The frequency of each surficial deposits classes by opening type was computed by extracting the surficial deposit thickness class of each canopy opening through a GIS overlay operation. Surficial deposit thickness classes were defined based on their description in the eco-forestry inventory map (Lord and Faucher 2003).

Modelling procedure

Statistical analyses were performed in R Version 2.12.1 (R Core Team 2016). We built a mixed-effects logistic regression model (Function *glmer* in the *lme4* package) of canopy opening types (NRO or canopy gap). The study site itself was set as a random effects variable. A large number of potential variables were chosen based on the hypotheses we formulated in regards of tree establishment limitation and vegetation physiognomy. Stepwise logistic regression was performed to select a subset of these variables to be included in the model. This method can be a useful data analysis tool when the association between the variables and the phenomenon under study are not well understood (Hosmer and Lemeshow 2000). Variables were selected according to Akaike information criterion (AIC) reduction and their statistical significance. The model was then trained with a randomly selected subset composed of 80% of the field data (n=169 or 47, 54

and 68 for sites A, B and C respectively). The remaining records served as validation data (n=43 or 15, 10, 18 for sites A, B and C respectively). We used the Hosmer-Lemeshow test to assess the model's goodness-of-fit (Hosmer and Lemeshow 2000). The model's discrimination ability was assessed with a receiver-operating characteristics (ROC) curve (Metz 1978; Pearce and Ferrier 2000). The kappa parameter curve (Guisan and Zimmermann 2000) was used to determine an optimal cut-off value for predicting opening type. Relation between each selected explanatory variable and opening type were explored with log-odds ratio plots.

Results

The model built using LiDAR-derived features succeeded in discriminating NRO from canopy gaps (Table 2). The analysis of the Kappa parameter showed that cut-off values between 0.40 and 0.72 offered similar classification accuracy (Fig. 3); *i.e.*, selecting a cut-off within this interval of values would balance increased classification performance of the model for one opening type with an equal decrease in the other. One-third of all canopy openings had a predicted probability between 0.40 and 0.72. We decided to retain the default cut-off value (0.5) in order to be unbiased.

Applying the model to the three study sites predicted that, NRO accounted for 19.5% of detected canopy openings area (15.1% to 33.5% for the 0.4 and 0.72 cut-off values) and occupied 1.37% of the study sites (Table 4). Canopy gaps are the most frequent of the canopy opening types, covering 5.71% of the study area. Canopy gap fractions were similar among study sites (5.51% to 5.96%), while the range in NRO fraction was greater (0.98% to 1.71%).

Log-ratios of the observed and predicted counts for the 10 deciles of the Hosmer-Lemeshow test indicated no lack-of-fit problem ($\chi^2 = 7.25$, $p = 0.51$) but they showed some volatility in the odds

when values were within the 0.40 to 0.72 interval of cut-off values (Fig. 4). Areas under the ROC curve were 93.29% and 82.26% for the validation and training data, respectively (Fig. 5). This indicated that the model has a discrimination ability between excellent and outstanding (Hosmer and Lemeshow 2000). Classification error was 17% for the validation data and 26% for the training dataset, at the 0.5 cut-off value. However, classification errors are not distributed evenly in the three study sites, with more commission errors for study site B (Table 3). The higher classification accuracy of the validation dataset can be explained by the lower proportion of randomly selected samples from study site B compared to the training dataset.

Study site was not a significant random effect in the model (intercept variance = 9.93×10^{-13} ; and standard deviation = 9.96×10^{-7}). Four variables were selected to predict canopy opening type: the minimum value of TPI, the 20th percentile of CHM values, the maximum distance to the nearest modeled stream, and the maximum slope class (Table 2). Minimum TPI was the first variable selected to discriminate NRO from canopy gaps. Negative values indicated some form of local terrain depression within the opening (concavity) favoring moisture accumulation or soil saturation. The odds-ratio plot effectively showed that NRO were related to low values of minimum TPI (Fig. 5). Conversely, values of minimum TPI that were close to 0, which indicate flat terrain, did not limit tree establishment and, therefore, were indicative of canopy gaps. The departure from linearity of the log-ratios near the middle deciles of TPI was due to the heterogeneous distribution of values for NRO and canopy gaps. If values of TPI were distributed more homogeneously between the 4th and the 8th deciles, the log-ratios would have behaved in a linear fashion. Moreover, those deciles would not have had much explanatory power. The 20th percentile of CHM height had good explanatory power for distinguishing the two opening types; NRO are more associated with low values of the CHM while the converse was observed for canopy gaps (Fig. 6). Distance to water was the third variable selected (Table 2) and it served mainly to identify canopy gaps when openings are far from water (Fig. 6); openings that are close

to water were not necessarily NRO, but those that are far from water are more likely to be canopy gaps. The last variable selected was maximum slope class (Table 2), which had some explanatory power for identifying NRO on steep slopes (Fig. 6). Odds-ratios indicated that openings that are in the high maximum slope class were more likely to be NRO. High maximum slopes were found in only 10% of the sampled canopy openings.

The frequencies of limiting factors found during fieldwork in NRO were different in the three study sites (Fig. 1). Moisture accumulation and permanent water saturation in soil were the most common limiting factors in study sites A and B. Study site C had many NRO with more than one kind of limiting factor. Rock and succession inhibiting vegetation were frequently found together in NRO. The thickness of the surficial deposit seems to condition the spatial distribution of limiting factors related to the NRO in the three sites (Table 1). Rock was never a limiting factor on thick surficial deposits but in these conditions, succession inhibiting vegetation was a predominant cause. On thin deposits, excessive water was the most frequent limiting factor. Beaver damage was less frequent; they were found in only 15% of the NRO. Overall, most of the NRO had water accumulation zones and inadequate soil conditions (i.e., very thin soil, rock outcrops, etc.). Soil moisture saturation zones were aggregated and close to streams or scattered when lying within concave terrain features favoring water accumulation. Rock outcrops were located on convex features or on the edge of a steep slope. Features distinctive of NRO did not cover much area (10 m^2 to 30 m^2) and were rarely larger than the canopy opening itself. Mean vegetation height was usually greater within canopy gaps than in NRO, even in recently created openings, because of the omnipresence of sugar maple and American beech advance regeneration under closed canopy.

In general, total canopy opening fraction, as well as canopy gap fraction, increased as the surficial deposit becomes thinner (Table 4). If we exclude the data from the 8.6 ha of moderate surficial deposit thickness in site B, a corridor zone characterized by stream courses and water

accumulation, the canopy gap fractions were close to the general mean (5.71%), ranging from 5.35% to 6.33%. However, NRO fraction was more variable, particularly on thin surficial deposit (0.71% to 2.27%) and greater on moderately thick deposits.

Discussion

Model variables

In this study, LiDAR-derived information was shown to be a very good predictor for differentiating NRO from canopy gaps. Variables that were selected are related directly to conditions that made sites less suitable for tree establishment or to the resulting vegetation physiognomy. Small-scale terrain concavity or flatness, as expressed by the minimum TPI, was the variable with the highest predictive power in the model. Flat terrain was associated with canopy gaps and concavity with NRO. Soil moisture was also invoked as a limiting factor with the ‘maximum distance to water’ variable. The probability of being classified as a canopy gap increased with the distance between canopy openings and moisture accumulation zones. However, proximity to these wetter zones did not necessarily predict canopy opening type. This might indicate that water accumulation there sometimes crosses a threshold over which some trees can’t cope (Kozlowski 1985, 1986). The maximum slope parameter served to identify sites that are too steep for trees to establish or maintain themselves in those canopy openings. Topographic position, slope and elevation are known to influence canopy gap abundance (Battles et al. 1995), but their relationship with NRO’s frequency has yet to be researched.

We also expected that the absence of vegetation or the presence of resource-limited vegetation would exhibit a low height distribution within canopy openings. Among the variables related to this feature, the 20th percentile of the height distribution was the best selected indicator. Its odds-ratio profile (Fig. 4) clearly showed that NRO were identifiable when the 20th percentile of the height values is low. Other height deciles were correlated with the 20th percentile but were not as good. We had also expected that a measure of height variation within canopy openings would

help. However, the standard deviation was negatively correlated with the 20th percentile of height ($r = -0.51$, $p < 0.001$), and probably did not add enough new information to the model. The negative correlation indicated that canopy openings filled with tall vegetation have less height variation than openings with a low portion.

NRO

In situ determination of canopy opening type was not always straightforward. Two sources of uncertainty may have caused this. First, a gradient existed between NRO and canopy gaps rather than a simple dichotomy. Conditions that limit tree establishment will usually hinder growth in some portion of a canopy opening but not to the extent of preventing tree establishment entirely. Second, distinguishing the proportion of the canopy opening that was occupied by each of these two types was sometimes difficult. As such, the insensitivity of the model when choosing an optimal cut-off value might further illustrate the uncertainty of the NRO / canopy gap classification that we faced in the field. In spite of this issue, the model performed very well. A different cut-off might have been selected if our goal had been to optimize detection of canopy gaps or NRO. We chose to be unbiased, which means that we reported higher NRO frequencies of our goal had been to be sure of canopy gaps classification and much lower NRO frequency if the reverse goal had been sought. Future research is needed to further understand the importance of NRO on forest dynamic and productivity.

Despite their presence in forest ecosystems driven by gap dynamics, NRO have been excluded in most studies. One study that did explore the question of NRO was conducted in British Columbia. Bartemucci *et al.* (2002) reported neither NRO with growth inhibiting vegetation communities nor canopy openings that were partially edaphic/canopy gaps in their northern temperate study area. Only 0.2% of their study area was occupied by NRO; the values we estimated for our study area (0.88% to 1.35%) are several times higher. However, such differences can be expected since those areas are located in different biomes.

Limiting factors in NRO were clearly associated with surficial deposit thickness. For instance, 66% of the NRO on thin surficial deposit are related to water. On these deposits, small-scale terrain shape has more effects; the concavity of the underlying bedrock can generate small basins where water rapidly saturates the soil and accumulates. This might explain the higher proportion of NRO associated with water factors on thin surficial deposits. Inversely, NRO with succession inhibiting vegetation were more frequent on the thick deposits found in the lowlands. Beavers were a spatially segregated cause of NRO. Signs of beaver activity were found mostly in the lowlands close to the major streams and water bodies and were absent on thin deposits in the uplands.

Succession inhibiting vegetation was mostly composed of *Carex* spp., *Viburnum* spp., *Corylus cornuta* and ferns (mostly *Athyrium felix-femina*). These species have an ecological advantage in occupying NRO. We observed that succession inhibiting species shared certain functional traits, such as perennality, low nutrient requirements, sustainability under extreme and variable soil moisture conditions, and high light requirements compared to other common understory species. Trees that were able to establish themselves despite the limiting factors are likely to grow very slowly, making them vulnerable to browsing by white-tailed deer (*Odocoileus virginianus*), thus further slowing canopy closure (Horsley and Marquis 1983).

Conclusion

LiDAR was used to detect canopy openings and to predict with good accuracy whether they were canopy gaps or NRO using vegetation height, distance to moisture accumulation, topography concavity/convexity, and slope steepness. The fitness of our model was assessed using a validation dataset different from the training dataset, which ensured that we did not overfit the parameters. This is, to the best of our knowledge, the first time remote sensing has been successfully used for discriminating these canopy opening types. Small footprint airborne LiDAR data can measure not only vegetation height, but also small-scale terrain morphological features.

Ecologists have long known that those features condition the distribution of vegetation at the site-scale. Prior to the development of LiDAR, scaling up this information to the landscape-level was impossible with traditional remote sensing tools. The ability to distinguish between NRO and canopy gap is an important advantage when using LiDAR data to characterize gap regimes.

The general methodology used in this study might be applicable elsewhere. However, it is evident that the explanatory power of the selected variables was associated with specific characteristics of the study sites. The frequency of NRO in a landscape is hard to anticipate since they can be created by factors that vary over a relatively small spatial scale. Our results demonstrate that NRO should be investigated prior to canopy gap detection with LiDAR data. They also suggest that NRO can be quite common in some northern temperate forests.

Acknowledgements

We are grateful to Lauzon ressource forestières, SEPAQ (Reserve faunique Papineau-Labelle) and MDDEP (Réserve écologique de la Forêt-la-Blanche) for authorizing access to the study sites. The senior author thanks Claire Binnert and Audrey Girard-Miron for their help with the fieldwork. LiDAR data were acquired with funds from the Canadian Foundation for Innovation. We thank the Applied Geomatics Research Group (Centre of Geographic Sciences, Middleton, Nova-Scotia, Canada) and Professor Chris Hopkinson (Department of Geography, University of Lethbridge, Alberta, Canada) for helping with LiDAR data acquisition.

References

- Bartemucci, P., Coates, K.D., Harper, K.A., and Wright, E.F. 2002. Gap disturbances in northern old growth forests of British Columbia, Canada. *Journal of Vegetation Science* **13**(5): 685-696.
- Battles, J.J., Fahey, T.J., and Harney, E.M.B. 1995. Spatial Patterning in the Canopy Gap Regime of a Subalpine *Abies-Picea* Forest in the Northeastern United States. *Journal of Vegetation Science* **6**(6): 807-814.

- 429 Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann,
430 V., and Böhner, J. 2015. System for Automated Geoscientific Analyses (SAGA). *Geosci.*
431 *Model Dev.* **8**(7): 1991-2007. doi: 10.5194/gmd-8-1991-2015.
- 432 Engelman, H.M., and Nyland, R.D. 2006. Interference to hardwood regeneration in northeastern
433 North America: assessing and countering ferns in northern hardwood forests. *Northern*
434 *Journal of Applied Forestry* **23**(3): 166-175.
- 435 Guisan, A., and Zimmermann, N.E. 2000. Predictive habitat distribution models in ecology.
436 *Ecological Modelling* **135**(2-3): 147-186. doi: Doi: 10.1016/s0304-3800(00)00354-9.
- 437 Horsley, S.B., and Marquis, D.A. 1983. Interference by weeds and deer with Allegheny
438 hardwood reproduction. *Canadian journal of forest research* **13**: 61-69.
- 439 Hosmer, D.W., and Lemeshow, S. 2000. Applied logistic regression. Wiley-Interscience,
440 Hoboken, N.J. pp. 392.
- 441 Houston, D.R. 1994. Major new tree disease epidemics: beech bark disease. *Annual Review of*
442 *Phytopathology* **32**(1): 75-87.
- 443 Jenness, J., Brost, B., and Beier, P. 2005. Land Facet Corridor Designer: Extension for ArcGIS,
444 Jenness Enterprises.
- 445 Jenson, S.K., and Domingue, J.O. 1988. Extracting topographic structure from digital elevation
446 data for geographic information system analysis. *Photogrammetric engineering and remote*
447 *sensing* **54**(11): 1593-1600.
- 448 Koukoulas, S., and Blackburn, G.A. 2004. Quantifying the spatial properties of forest canopy
449 gaps using LiDAR imagery and GIS. *International Journal of Remote Sensing* **25**(15): 3049-
450 3072.
- 451 Kozłowski, T. 1985. Soil aeration, flooding, and tree growth. *Journal of Arboriculture*.
- 452 Kozłowski, T. 1986. Soil aeration and growth of forest trees. *Scandinavian Journal of Forest*
453 *Research* **1**(1-4): 113-123.

- 454 Lafond, R., Leblanc, Y., Pilon, C., and faune, S.d.l.f.e.d.p.d.Q.D.d.d.d.l. 2003. Bilan du plan
455 d'inventaire aérien des colonies de castors au Québec (1989-1994). Société de la faune et des
456 parcs du Québec, Direction du développement de la faune, Québec.
- 457 Lertzman, K.P., Sutherland, G.D., Inselberg, A., and Saunders, S.C. 1996. Canopy gaps and the
458 landscape mosaic in a coastal temperate rain forest. *Ecology* **77**(4): 1254-1270.
- 459 Lim, K., Treitz, P., Wulder, M., St-Onge, B., and Flood, M. 2003. LiDAR remote sensing of
460 forest structure. *Progress in Physical Geography* **27**(1): 88.
- 461 Lord, G., and Faucher, A. 2003. Normes de cartographie écoforestière: troisième inventaire
462 écoforestier. Ministère des ressources naturelles, de la faune et des parcs, Forêt Québec,
463 Direction des inventaires forestiers.
- 464 Metz, C.E. 1978. Basic principles of ROC analysis. *In* *Seminars in Nuclear Medicine*. Elsevier.
465 pp. 283-298.
- 466 Olthof, I., King, D.J., and Lautenschlager, R.A. 2004. Mapping deciduous forest ice storm
467 damage using Landsat and environmental data. *Remote Sensing of Environment* **89**(4): 484-
468 496. doi: <http://dx.doi.org/10.1016/j.rse.2003.11.010>.
- 469 Pearce, J., and Ferrier, S. 2000. Evaluating the predictive performance of habitat models
470 developed using logistic regression. *Ecological Modelling* **133**(3): 225-245.
- 471 R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for
472 Statistical Computing, Vienna, Austria.
- 473 Robitaille, A., and Saucier, J.-P. 1998. Paysages régionaux du Québec méridional. Publications
474 du Québec, Québec.
- 475 Royo, A.A., and Carson, W.P. 2006. On the formation of dense understory layers in forests
476 worldwide: consequences and implications for forest dynamics, biodiversity, and succession.
477 *Canadian journal of forest research* **36**(6): 1345-1362. doi: 10.1139/x06-025.

- 478 Runkle, J.R. 1985. Disturbance regimes in temperate forests. *In* The ecology of natural
479 disturbance and patch dynamics. *Edited by* S.T.A. Pickett and P.S. White. Academic Press,
480 Inc., San Diego.
- 481 Sénécal, J.-F., Doyon, F., and Messier, C. 2018a. Management implications of varying gap
482 detection height thresholds and other canopy dynamics processes in temperate deciduous
483 forests. *Forest Ecology and Management* **410**: 84-94. doi:
484 <https://doi.org/10.1016/j.foreco.2017.12.029>.
- 485 Sénécal, J.-F., Doyon, F., and Messier, C. 2018b. Tree Death Not Resulting in Gap Creation: An
486 Investigation of Canopy Dynamics of Northern Temperate Deciduous Forests. *Remote*
487 *Sensing* **10**(1): 121.
- 488 St-Onge, B., Vepakomma, U., Sénécal, J.-F., Kneeshaw, D., and Doyon, F. 2014. Canopy Gap
489 Detection and Analysis with Airborne Laser Scanning. *In* *Forestry Applications of Airborne*
490 *Laser Scanning. Edited by* M. Maltamo and E. Næsset and J. Vauhkonen. Springer
491 Netherlands. pp. 419-437.
- 492 Vepakomma, U., Kneeshaw, D., and St Onge, B. 2010. Interactions of multiple disturbances in
493 shaping boreal forest dynamics: a spatially explicit analysis using multi temporal lidar data
494 and high resolution imagery. *Journal of Ecology* **98**(3): 526-539.
- 495 Vepakomma, U., St-Onge, B., and Kneeshaw, D. 2008. Spatially explicit characterization of
496 boreal forest gap dynamics using multi-temporal lidar data. *Remote Sensing of Environment*
497 **112**: 2326-2340.
- 498 Zhang, K. 2008. Identification of gaps in mangrove forests with airborne LIDAR. *Remote*
499 *Sensing of Environment* **112**(5): 2309-2325.

Table 1. Percentage of limiting factors in Non-Regenerating Openings (NRO) by surficial deposit thickness class.

Surficial deposit thickness class	Water	Rock	Succession inhibiting vegetation	Beaver	Number of sampled NRO		
					Site A	Site B	Site C
Thick	47.1%	0.0%	41.2%	11.8%	14	*	*
Moderate	32.2%	33.9%	25.4%	8.5%	*	7	35
Thin	65.5%	24.1%	8.6%	1.7%	12	39	2

* No surficial deposit of the thickness class was observed in the study site.

Draft

Table 2. Logistic regression model of canopy opening types with the minimum and maximum value of each variable (AIC = 189, AIC reduction = 106.88, -2LL reduction = 87.47, Nagelkerke’s $R^2 = 0.41$).

Predictor	Min-Max value	Coefficient (B)	Std. Error	Wald value	<i>p</i>	ln (B)
Intercept	-	0.030	0.650	0.046	0.963	-
20 th percentile of height	0.04 – 4.65	0.535	0.174	3.079	0.002	1.708
Maximal distance to water	3.20 – 92.98	0.036	0.012	2.855	0.004	1.036
Minimal value of TPI	-3.59 – 0.29	1.250	0.342	3.654	0.000	3.490
Intermediate slope class 2	0 – 1	-0.657	0.425	-1.544	0.122	0.519
High slope class 3	0 – 1	-1.743	0.669	-2.605	0.009	0.175

Draft

Table 3. Confusion matrix of all records in the dataset, by study site and overall (Percentage of total in parentheses).

Predicted	Observed		Commission error
	Site A	Site B	
	NRO	Canopy gaps	
	NRO	19 (30.6%) 6 (9.7%)	24.0%
	Canopy gaps	8 (12.9%) 29 (46.8%)	21.6%
	Omission error	29.6% 17.1%	
	Accuracy		77.4%
	Site B	Site C	
	NRO	Canopy gaps	
	NRO	28 (43.8%) 8 (12.5%)	22.2%
	Canopy gaps	13 (20.3%) 15 (23.4%)	46.4%
	Omission error	31.7% 34.8%	
	Accuracy		67.2%
	Site C	Overall	
	NRO	Canopy gaps	
	NRO	28 (32.6%) 7 (8.1%)	20.0%
	Canopy gaps	9 (10.5%) 42 (48.8%)	17.6%
	Omission error	24.3% 14.3%	
	Accuracy		81.4%
	Overall		
	NRO	Canopy gaps	
	NRO	75 (35.4%) 21 (9.9%)	21.9%
	Canopy gaps	30 (14.2%) 86 (40.6%)	25.9%
	Omission error	28.6% 19.6%	
	Accuracy		76.0%

Table 4. NRO and canopy gap percentages of the forest cover for the three study sites by surficial deposit thickness class.

Site	Surficial Deposit Thickness Class	Forest cover area (ha)	Area % in canopy opening type			Canopy opening type ratio (NRO/G %)
			Canopy gaps	NRO	Total	
A	Thick	37.5	5.36	1.34	6.69	25
B	Moderate	8.6	3.60	2.34	5.94	65
C	Moderate	51.4	5.64	1.46	7.10	26
	Total	60	5.35	1.59	6.94	30
A	Thin	44.9	6.00	0.71	6.71	12
B	Thin	79.2	6.33	1.37	7.70	22
C	Thin	22.7	5.60	2.27	7.88	41
	Total	146.8	6.11	1.31	7.42	21
A		85.8	5.51	0.98	6.49	18
B		90.0	5.96	1.45	7.41	24
C		74.0	5.63	1.71	7.34	30
Total		249.8	5.71	1.37	7.08	24

Figure captions

Figure 1. Percentage of NRO affected by one or more limiting factor(s) in each study site.

Figure 2. Examples of 2007 canopy height models of a portion of sites that were sampled during fieldwork of A-B) canopy gaps, C) Non-regenerating opening due to vegetation communities inhibiting plant establishment and D) Non-regenerating opening due to water accumulation. The pixels have a 50 cm resolution and the represented area is 30 x 30 meters. Dark pixels are low height areas and light pixels are higher heights. The tone-height relation is the same in all images. White polygons outlines are the canopy opening outlines detected using LiDAR data.

Figure 3. Sensitivity, specificity, and the kappa parameter for the logistic regression model that differentiated NRO and canopy gaps, and which was computed using the validation data, as a function of the cut-off value.

Figure 4. Log-ratios of the observed and predicted counts for the 10 bins of the Hosmer-Lemeshow test.

Figure 5. Receiver operating characteristic (ROC) curve of the model differentiating NRO and canopy gaps, produced from the validation and training datasets.

Figure 6. Log-ratios of the ten deciles of within opening values for the minimum topographic position index (TPI), 20th percentile of CHM values, maximum distance to water, and maximum slope class (Note: different X-axis scales per variable).

Figure 1

501

Draft

Figure 2

A) Canopy gap**B) Non-Regenerating opening**
Due to beavers **C) Non-Regenerating opening**
Due to vegetation communities**D) Non-Regenerating opening**
Due to water accumulation

Figure 3

Figure 4

Figure 5

Figure 6