No. 2019_1_01

Institut de recherche sur les feuillus nordiques Inc.
Northern Hardwoods Research Institute Inc.

Regeneration Characterization from Remote Sensing

December 2019

Technical Note

Resource Characterization

INTRODUCTION

Remote sensing technologies such as satellite images and Light Detection and Ranging (LiDAR) have been used for a few decades in the field of forestry to compute different forest characteristics (Franklin 2001). However, forest attributes derived from remote sensing are mostly related to overstory due to the interference of canopy foliage on the detection of understory vegetation. Because of this, only a couple of studies have been done to characterize regeneration from remote sensing. Furthermore, they have been done mainly in areas with a low canopy cover.

On the other hand, the integration of two sensors (satellite or aerial images and LiDAR) has demonstrated to be more accurate than the utilization of only one sensor in the estimation of overstory tree attributes (Zald et al. 2014; Imangholiloo et al. 2019). However, a very few studies have assessed regeneration using combination of satellite or aerial images and LiDAR. Moreover, these studies have also been primarily conducted in areas with a low canopy cover (Su and Bork 2007; Korpela et al. 2008; Martin-Alcon et al. 2015).

HIGHLIGHTS

- **Models were created to estimate sapling density of 1) all species and 2) commercial species with an accuracy of +/- 2822 st/ha for all species and +/- 2807 st/ha for commercial species using LiDAR and environmental variables.**
- **LiDAR metrics alone are enough to have accurate estimates of sapling density.**
- **Canopy cover impacts the estimation of sapling density when using Sentinel-2 images along with environmental variables, but not when using LiDAR and the integration of LiDAR and Sentinel-2 images.**

The aim of this study was to create models to estimate regeneration density from remote sensing and understand the impact of canopy cover on the accuracy of the models. We also wanted to test if the integration of two sensors (satellite images and LiDAR) was more accurate than only one sensor.

METHODOLOGY

Field data

We calculated sapling (stem with a height ≥ 1 .3 m and a DBH ≥ 1 cm and ≤ 9 cm) density (st/ha) of 1) all species and 2) commercial species from the Continuous Land Inventory (CLI) plots from New Brunswick (n=813).

Sentinel Images

We used the four bands (blue, green, red and near-infrared) and the ARVI, EVI, NDVI and VARI index from Sentinel-2 images (10m x 10m resolution) obtained on July 1st 2016, July 9th 2017 and July 21st 2018.

LiDAR

LiDAR data was acquired during the summers of 2016, 2017 and 2018 at a point density of 6 points/m^2. We calculated LiDAR metrics for each plot using the package lidr in R (Table 1).

Site characterization

We estimated canopy cover of each plot using the proportion of LiDAR points over 7 m which is corresponding to the mean height of sampling. We also calculated the basal area of trees with a DBH ≥ 9.1 cm for each plot and the proportion of hardwood and softwood species based on basal area. We characterized sites using aspect and hillshade, ecoregion, ecodistrict and ecosite, the biomass growth index (BGI), depth to water table, soil type, and an index of site quality per species.

Statistical analyses

We built three models using random forest regression for each dependent variable (absolute density of sapling of 1) all species, and 2) commercial species). The first model included LiDAR metrics, spectral variables (bands and vegetation index) derived from Sentinel-2 images, and environmental variables, the second only LiDAR metrics and environmental variables, and the third included only spectral variables, and environmental variables. We selected variables in each models using the VSURF package in R based on their importance. We then calculated root mean square error (RMSE) with the random Forest package and obtained pseudo R-squared values were calculated with internal values of mean square error (MSE) and pseudo R-square of each tree respectively. Finally, to test the impact of canopy cover on the accuracy of all the models, we performed a Kruskal-Wallis test on the relative errors ((predicted-actual)/actual) of four different categories of canopy cover (low: 0 to 24%, medium: 25 to 49%, high: 50 to 74% and very high: ≥ 75%). Following the Kruskal-Walllis test, if significant, a Dunn test was done using the Bonferonni method to adjust the p-values.

Table 1: LiDAR metrics used to estimate sapling density

Abbreviation	Description
Zmax	Maximum height
Zmean	Mean height
Zsd	Standard deviation of height distribution
Zskew	Skewness of height distribution
Zkurt	Kurtosis of height distribution
Zentropy	Entropy of height distribution
Pzabovezmean	Percentage of returns above zmean
Pzabove2	Percentage of returns above 2m
Zqx	X percentile of height distribution
Zpcumx	Cummulative percentage of return in the xth layer according to Wood et al. 2008
itot	Sum of intensities for each return
Imax	Maximum intensity
Imean	Mean intensity
Isd	Standad deviation in intensity distribution
Iskew	Skewness of intensity distribution
Ikurt	Kurtosis of intensity distribution
Ipground	Percentage of intensity returned by points classified as ground
ipcumzqx	Percentage of intensity returned below the xth percentile of height
Pxth	Percentage sth returns
pground	Percentage of returns classified as ground

RESULTS

Variables selection and model accuracy

The estimation of sapling density of all species is more accurate than of commercial species despite the sensor used (both or only one: Table 2). Moreover, the models using the Sentinel-2 images are less accurate than the ones using LiDAR or both sensors and the accuracy of the models using LiDAR or both sensors is similar no matter the species group estimated (Table 2).

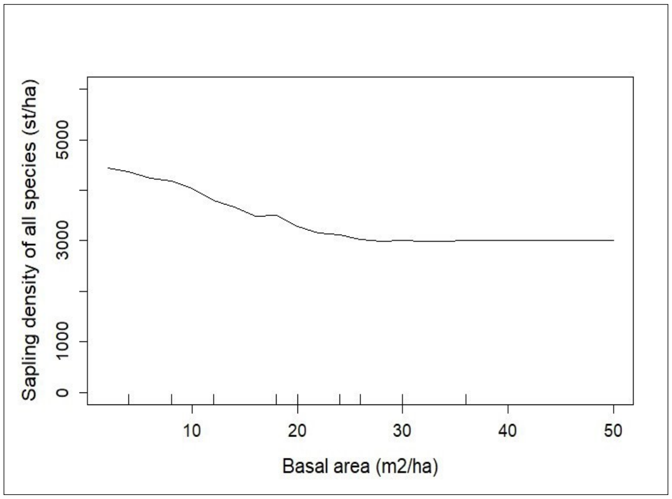
The integration of both sensors (LiDAR and Sentinel-2 images) increases the estimation accuracy of sapling density compared to Sentinel-2, but not to LiDAR.

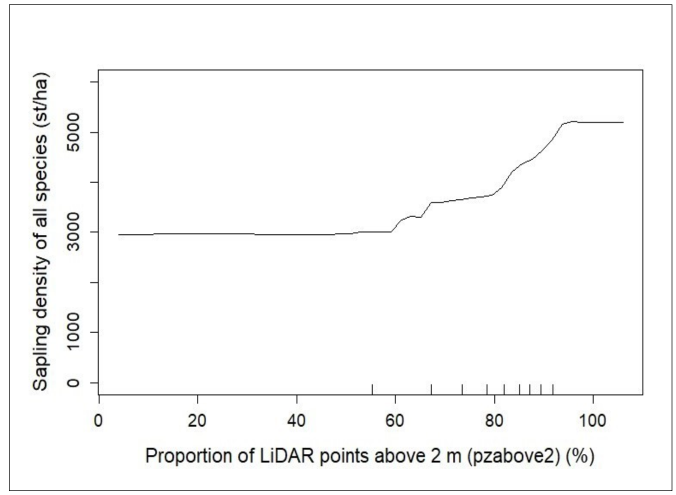
In regards to the variance explained, estimations of sapling density of all species yield higher pseudo R-squared than of commercial species (Table 2). Furthermore, the integration of both sensors increased the variance explained for commercial species estimation where it is similar to LiDAR when estimating sapling density of all species (Table 2). Again, the models using Sentinel-2 images have the lowest pseudo R-squared despite the species group estimated (Table 2).

Some variables have been selected in almost all the models no matter the group of species estimated. Basal area of trees of DBH ≥ 9.1 cm, the higher quantiles of LiDAR height distribution (zqX), the proportion of LiDAR points near the ground (zpcum1), the percentage of LiDAR points over 2m (pzabove2) as well as the percentage of LiDAR points over the mean height have been selected in almost all the models. Predictions of sapling density of all species are higher when basal area and the value of the higher percentiles of height distribution are low, and contrarily the predictions are higher when the proportion of points above 2m is higher (Figure 1).

Table 2: Candidate models, variables selected, RMSE (st/ha), relative RMSE (%) and pseudo R-squared of models estimating sapling density of all species and commercial species using LiDAR metrics and/or spectral variables and environmental variables.

Candidate model	Variables selected	RMSE (st/ha)	Relative RMSE (%)	Pseudo R-squared
Sapling density of all species				
LiDAR + Spectral +	Zq80 + zq85 + zq75 + zq70 + zq60 + zqcum1 + zq55+ pzabove2 +	2807	82	0.34
LiDAR + Environmental	Zq80 + zq85 + zq75 + zq70 + zq60 + zpcum1 + zq95 + zq55 +	2822	83	0.33
Spectral + Environmental	Basal area	3321	97	0.08
Saplling density of commercial species				
LiDAR + Spectral + Environmental	Zq80 + zq75 + zq85 + zq90 + zpcum1 + pzabove2 + zq65 + zq95 + zq60 + zq30 + zq55 +zmean + zpcum2 + pzabovezmean + p1th	2724	97	0.27
LiDAR + Environmental	Zq80 + zq75 + zq85 + zpcum1 + zq90 + pzabove2 + zq70 + zq30	2807	100	0.22
Spectral + Environmental	Proportion of hardwood + basal area + EVI	3314	118	-0.09





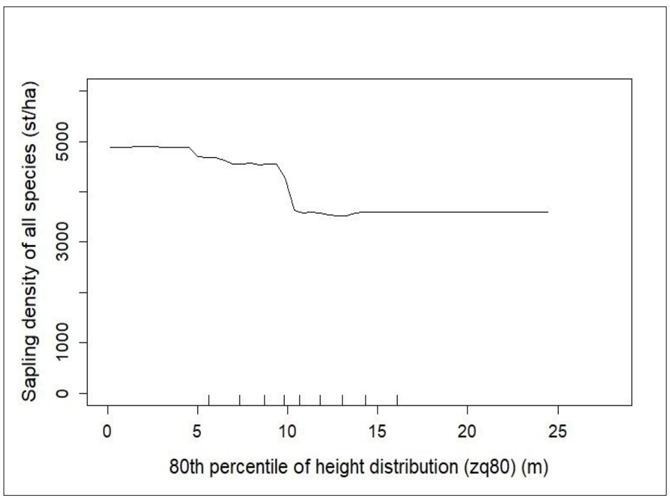
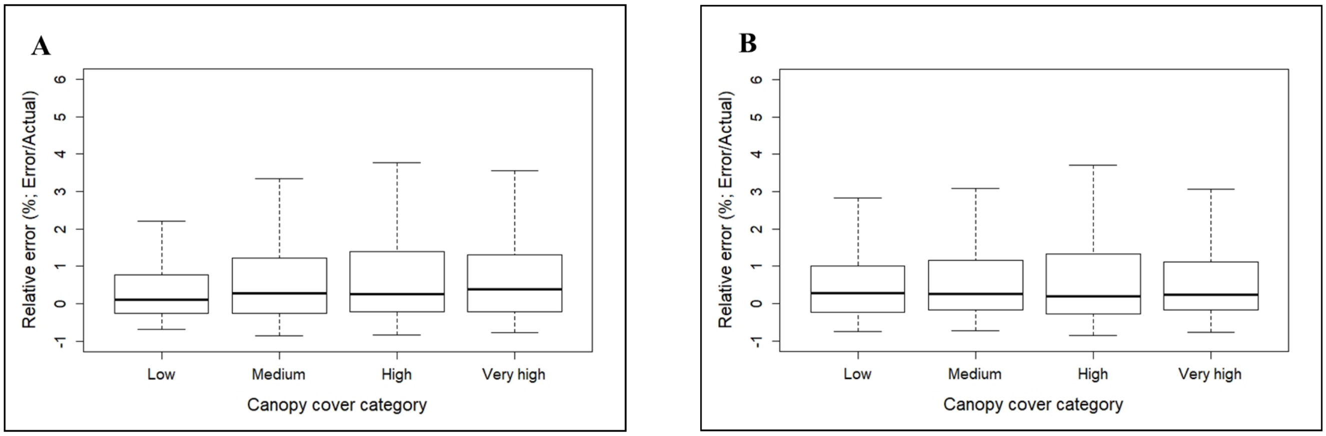


Figure 1: Partial dependence plot of basal area (m^2/ha), 80th percentile of height distribution (zq80; m) and proportion of LiDAR points above 2 m (pzabove2; %) for random forest predictions of sapling density of all species (st/ha) using both sensors.

Impact of canopy cover

The relative errors of the models using only spectral and environmental variables are increasing with increasing canopy cover, although only the model estimating sapling density of all species is yielding significantly higher relative errors under a very high canopy cover than under a low and medium canopy cover (Figure 2 and 3).

Only the relative error of the model estimating sapling density of all species using Sentinel-2 images is significantly higher under a very high canopy cover (≥75%).



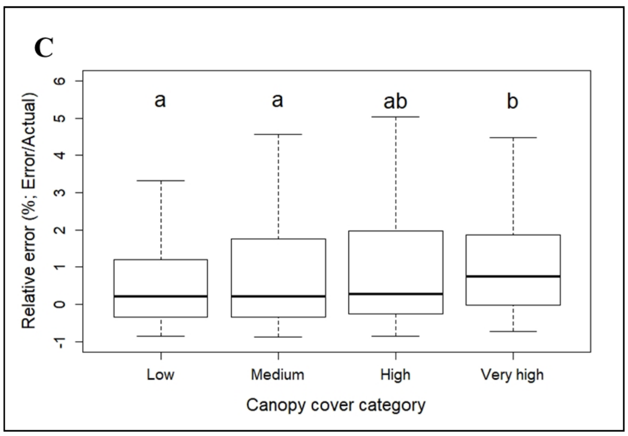
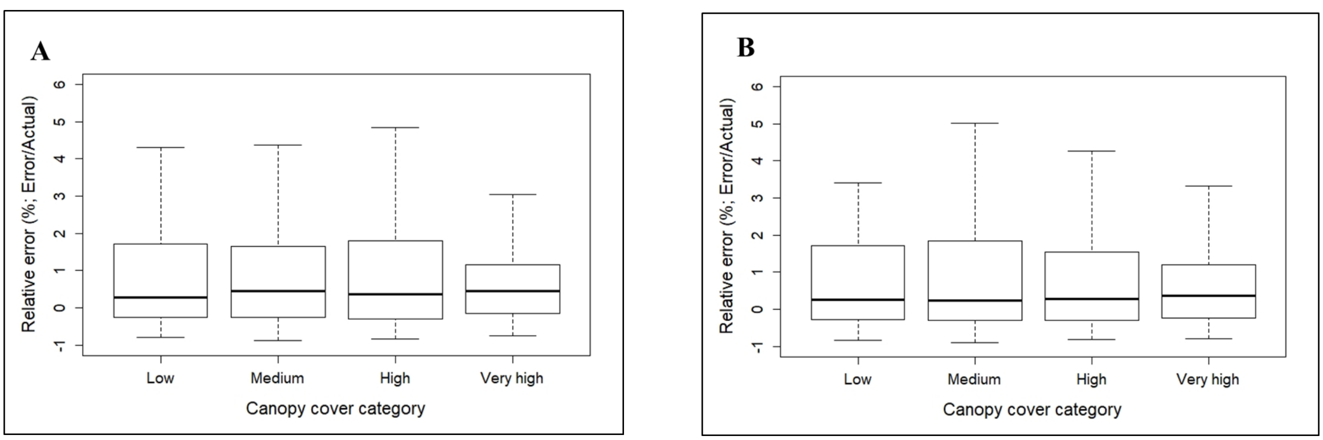


Figure 2: Distribution of the relative error (%; error/actual) according to canopy cover category (Low: 0 to 24%, Medium: 25 to 49%, High: 50 to 74% and Very high: ≥ 75%) for models estimating sapling density (st/ha) of all species a) model integrating both sensors, b) model including LiDAR and environmental variables and c) model including spectral and environmental variables (Kruskal-Wallis test, a) $p = 0.25$, b) $p = 0.92$, and c) $p = 0.005$).



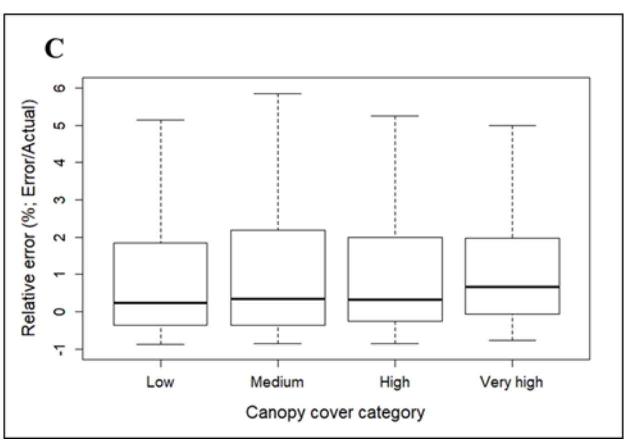


Figure 3: Distribution of the relative error (%; error/actual) according to canopy cover category (Low: 0 to 24%, Medium: 25 to 49%, High: 50 to 74% and Very high: ≥ 75%) for models estimating sapling density (st/ha) of commercial species a) model integrating both sensors, b) model including LiDAR and environmental variables and c) model including spectral and environmental variables (Kruskal- Wallis test, a) $p = 0.85$, b) $p = 0.99$, and c) $p = 0.11$).

DISCUSSION

Variables selection and model accuracy

The higher accuracy and variance explained of the estimation of sapling density of all species than of commercial species is probably due to the similarity of non-commercial species and commercial species of hardwood. Indeed, all the non-commercial species are hardwood species, and they are known to be harder than softwood species to differentiate via LiDAR (Collins et al. 2004) and satellite images (Weigel and Randolph 2013).

The models using spectral variables along with environmental variables yield the lowest accuracy and pseudo R- squared, meaning that LiDAR is more accurate than images to estimate regeneration characteristics such as sapling density. LiDAR is known to be better than images to penetrate through the canopy (Pesonen et al. 2008). Coops et al. 2004 also found that crown area was better estimated using LiDAR than multispectral images. Furthermore, LiDAR is known to be more accurate than satellite images to estimate density (Pearse et al. 2018), biomass (Hyde et al. 2006) and basal area (Hudak et al. 2006) of overstory trees.

Even if the accuracy of the models with both sensors are similar to the one using LiDAR for both species groups, the variance explained is higher when using both sensors to estimate commercial species. Furthermore, the green band was selected in the model estimating sapling density of commercial species with both sensors where none of the spectral variables were selected when estimating all species. The images are known to be more accurate than LiDAR to differentiate species (Zald et al. 2014). Moreover, the green band is better than the other bands to differentiate between over- and understory (Landry et al. 2018).

Basal area is an important variable in either case, estimation of sapling density of all species and commercial species and the use of both or one or the other sensors, predictions of sapling density of all species are higher when basal area is low. Under a dense canopy, there is a greater competition for the above and belowground resources, therefor sapling density is lower (Lundqvist and Fridman, 1996; Nolet et al. 2008).

When the value of higher percentiles of height distribution is low it means that the points are mainly coming from low vegetation (Means, 2000; Wasser et al. 2013; Watt and Watt, 2013), which can represent a higher density of sapling and a sparse canopy. On the other hand, a higher proportion of points above 2 m yields higher sapling density of all species, sapling corresponds to a tree with a height greater than 1.3 m, thus a higher proportion of points above 2 m means that there is a greater proportion of vegetation in the sapling and over- story strata.

Impact of canopy cover

Models using only spectral and environmental variables are more affected by canopy cover than the ones using both sensors or only LiDAR and environmental variables. When the canopy cover increases, the accuracy of the models decreases, although only the model estimating sapling density of all species yields a significantly higher estimation when canopy cover is very high. LiDAR is better than images to penetrate through the cano- py (Pesonen et al. 2008). Moreover, satellite images are more related to canopy cover characteristics than of height (Hudak et al. 2006) explaining the sensibility of Sentitnel-2 to canopy cover.

CONCLUSION

As a result of this project, we can conclude that LiDAR only is able to estimate sapling density of all species and commercial species. Sentinel-2 images increase the variance explained when estimating sapling density of commercial species but not the accuracy. Moreover, LiDAR as for the integration of both sensors is less impacted by canopy cover than Sentinel-2 images. However, LiDAR coverage will only be renewed each 10 years in New Brunswick which can lead to false information about regeneration if a treatment is done between these cycles. Nevertheless, it is possible to generate points-cloud using other sensors like unmanned aerial vehicle (UAV). To increase the accuracy and the variance explained by the model, we suggest the use of leaf-off LiDAR and the integration of leaf-off LiDAR and satellite images from different timing in the phenology season. We also suggest to try using satellite images with an higher resolution than Sentinel-2 images.

REFERENCES

Collins, C.A., Parker, R.C. and Evans, D.L., 2004, May. Using multispectral imagery and multi-return Li DAR to estimate tree and stand attributes in a southern bottomland hardwood forest. In American Society of Photogrammetry and Remote Sensing Annual Conference Proceedings (pp. 23-28).

Coops, N.C., Wulder, M.A., Culvenor, D.S. and St-Onge, B., 2004. Comparison of forest attributes ex tracted from fine spatial reso- lution multispectral and lidar data. Canadian Journal of Remote Sensing, 30(6), pp.855-866.

Franklin, S.E., 2001. Remote sensing for sustainable forest management. CRC press.

Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M., Gessler, P.E. and Morgan, P., 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. Canadian Journal of Remote Sensing, 32(2), pp.126-138.

IHyde, P., Dubayah, R., Walker, W., Blair, J.B., Hofton, M. and Hunsaker, C., 2006. Mapping forest struture for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. Remote Sensing of Environment, 102(1-2), pp.63-73.

Imangholiloo, M., Saarinen, N., Markelin, L., Rosnell, T., Näsi, R., Hakala, T., Honkavaara, E., Holopainen, M., Hyyppä, J. and Vastaranta, M., 2019. Characterizing Seedling Stands Using Leaf-off and Leaf-on Photogrammetric Point Clouds and Hyperspectral Imagery Acquired from Unmanned Aerial Vehicle. *Forests*, *10*(5), p.415.

Korpela, I., Tuomola, T., Tokola, T. and Dahlin, B., 2008. Appraisal of seedling-stand vegetation with air borne imagery and discrete-return LiDAR–an exploratory analysis. *Silva Fennica*, *42*(5), pp.753-772.

Landry, S., St-Laurent, M.H., Nelson, P.R., Pelletier, G. and Villard, M.A., 2018. Canopy Cover Estimation from Landsat Images: Understory Impact onTop-of-canopy Reflectance in a Northern Hardwood Forest. *Canadian Journal of Remote Sensing*, *44*(5), pp.435-446.

Lundqvist, L. and Fridman, E., 1996. Influence of local stand basal area on density and growth of regeneration in uneven‐aged Picea abies stands. *Scandinavian Journal of Forest Research*, *11*(1-4), pp.364-369.

Martín-Alcón, S., Coll, L., De Cáceres, M., Guitart, L., Cabré, M., Just, A. and González-Olabarría, J.R., 2015. Combining aerial LiDAR and multispectral imagery to assess postfire regeneration types in a Mediterranean forest. *Canadian Journal of Forest Research*, *45*(7), pp.856-866.

Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L. and Hendrix, C.J., 2000. Predicting forest stand characteristics with airborne scanning lidar. *Photogrammetric Engineering and Remote Sensing*, *66*(11), pp.1367-1372.

Nolet, P., Bouffard, D., Doyon, F. and Delagrange, S., 2008. Relationship between canopy disturbance history and current sapling density of Fagus grandifolia and Acer saccharum in a northern hardwood landscape. *Canadian journal of forest research*, *38*(2), pp.216-225.

Pearse, G.D., Dash, J.P., Persson, H.J. and Watt, M.S., 2018. Comparison of high-density LiDAR and satellite photogrammetry for forest inventory. *ISPRS journal of photogrammetry and remote sensing*, *142*, pp.257-267.

Pesonen, A., Maltamo, M., Eerikäinen, K. and Packalèn, P., 2008. Airborne laser scanning-based prediction of coarse woody debris volumes in a conservation area. *Forest Ecology and Management*, *255* (8-9), pp.3288-3296.

Su, J.G. and Bork, E.W., 2007. Characterization of diverse plant communities in Aspen Parkland rangeland using LiDAR data. *Applied Vegetation Science*, *10*(3), pp.407-416.

Wasser, L., Day, R., Chasmer, L. and Taylor, A., 2013. Influence of vegetation structure on lidar-derived canopy height and fractio- nal cover in forested riparian buffers during leaf-off and leaf-on condi tions. *PLoS One*, *8*(1), p.e54776.

Watt, P. and Watt, M.S., 2013. Development of a national model of Pinus radiata stand volume from LiDAR metrics for New Zealand. *International Journal of Remote Sensing*, *34*(16), pp.5892-5904.

Weigel, D.R. and Randolph, J.C., 2013. Spectral reflectance of five hardwood tree species in southern Indiana. In *In: Miller, Gary W.; Schuler, Thomas M.; Gottschalk, Kurt W.; Brooks, John R.; Grushecky, Shawn T.; Spong, Ben D.; Rentch, James S., eds. Proceedings, 18th Central Hardwood Forest Conference; 2012 March 26-28; Morgantown, WV; Gen. Tech. Rep. NRS-P-117. New- town Square, PA: US Department of Agriculture, Forest Service, Northern Research Station: 276-282.* (Vol. 117, pp. 276-282).

Zald, H.S., Ohmann, J.L., Roberts, H.M., Gregory, M.J., Henderson, E.B., McGaughey, R.J. and Braaten, J., 2014. Influence of lidar, Landsat imagery, disturbance history, plot location accuracy, and plot size on accuracy of imputation maps of forest composition and structure. *Remote Sensing of Environment*, *143*, pp.26-38.

ACKNOWLEDGEMENT:

I would like to thank the stakeholder committee, composed of Pierre LeBel, Dale Wilson, Adam Dick, Dr. Chris Hennigar and Dr. Joanne White, for their advice as well as DNR and J.D. Irving Ltd. for providing data.

FOR MORE INFORMATION CONTACT:

info@hardwoodsnb.ca
Researcher: Stéphanie Lebel-Landry

Institut de recherche sur les feuillus nordiques Inc.
Northern Hardwoods Research Institute Inc.

ADDRESS

165, BOULEVARD HÉBERT

EDMUNDSTON, N.-B.

E3V 2S8

PHONE

1 506 737-4736

FAX

HARVEST KNOWLEDGE, **PROMOTE** GROWTH