No. 2020_1_02

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

Preliminary evaluation at prediction gaps NBDNR-generated aerial-LiDAR based forest inventory (EFI)

Sept. 2020

Technical Note

Resource Characterization

INTRODUCTION

Enhanced Forest Inventory (EFI) derived using remote sensing technology (such as aerial LiDAR) are becoming used more and more in the Province of New Brunswick and are hoped to partially or totally replace traditional surveys in the future.

However, a previous study (Chouinard, C.-A., 2020) done by the NHRI showed that there were significant differences for some variables between traditional forest inventory and EFI.

HIGHLIGHTS

- **During recent years, the Province of New Brunswick-Canada has been a leader in the generation Enhanced Forest Inventory (EFI) derived from aerial LiDAR data (ALS) but there has been little effort spent to empirically evaluate their accuracy with independent surveys. Inputs needed:**
- **The NHRI has recently undertaken a significant initiative to evaluate their accuracy and found, in some instances, substantial gaps (Chouinard, 2020).**
- **The Northern Hardwood Research Institute (NHRI) developed different equations for nine essential forest inventory variables, modeling the differences between EFI and traditional forest inventory information, in order to narrow those differences and be able to use those EFI data in tactical planning and hopefully in operational planning.**
- **Relatively moderate and low correlations were found between all predicted and response variables. It oscillates between a minimum of 19.09% (r2) for live crown ratio and a maximum of 64.91% (r2) for stem density.**
- **Environmental variables were present in most equations developed underlining their importance when modeling EFI prediction gaps.**

Equations for nine essential forest variables were developed by modelling the difference between EFI and traditional forest inventory information and some dendrometric/environmental variables in order to find potential explanations for those differences.

METHODOLOGY

Data

To evaluate the accuracy of EFI data, we used data from traditional forest inventory plots realized for the 2019-2020 Groupe Savoie Inc. operational plan. Those plots (DBH ≥ 9.1 cm), situated in northern New Brunswick hardwoods dominated stands, were measured with a BAF2 angle gauge. From the 346 field plots used for this study, 45 plot centers were positioned at the exact same location as the center of the LiDAR's 20 meters by 20 meters cells and 301 field plot centers were located within a 10 meters radius from the LiDAR's cell centers.

The following 9 dendrometric variables were calculated for each field plot: average height (m), quadratic mean diameter (cm), gross merchantable volume (m3/ha), base to live crown (m), live crown ratio (%), stem density (trees/ha), average piece size (m3/tree), basal area (m2/ha) and percentage of hardwood composition (%). The following stand quality variables were also calculated from field plot even though they are not available from EFI: proportion of acceptable growing stock (%), proportion of poor form class (Pelletier et al., 2016) and proportion of good risk class (Pelletier et al., 2016).

We also used EFI's derived from aerial-LiDAR (ALS) data available from New Brunswick government (http://www.snb.ca/geonb1/). The same 9 dendrometric variables as field plots were obtained using EFI (2016) data and QGIS 3.8.1.

Same plot locations were used to extract environmental variables from GIS databases: elevation (m), aspect, soil, depth of the water table (m), biomass growth index (kg/ha/yr) (FORUS Research, 2016; Hennigar et al., 2016) and ecosite (Hennigar et al., 2016).

Data analysis

We used linear models (LMs) to determine the variables that could possibly affect accuracy of the EFI data. We used the difference between EFI and field data for each variable as the response variable. Environmental, stand quality and dendrometric field variables were used as explanatory variables. We used a model selection approach (Akaike's information criterion; Burnham and Anderson, 2002) to identify the combination of variables that best explains the differences between EFI and field data.

Particular attention was deployed to eliminate any multicollinearity that exists between response and explanatory variables. For example, gross merchantable volume was not included as an explanatory variable for the basal area equation because of the strong relation between them.

We also made sure that the conditions of homoscedasticity and normality of residuals were respected (Zuur et al., 2009) and then conducted the analysis. We performed statistical analyses using the R 3.5.1 software (R Core Team, 2018).

RESULTS

Preliminary analysis (Fig.1) was used to compare the mean differences between EFI and field data and to assess each variable's significance. The greatest mean differences are observed for merchantable tree density (50%), followed by the percentage of hardwood composition (-36%) and gross merchantable volume (31%). However, the most accurate variables studied were average height (-6%), followed by the basal area (14%) (Fig.1).

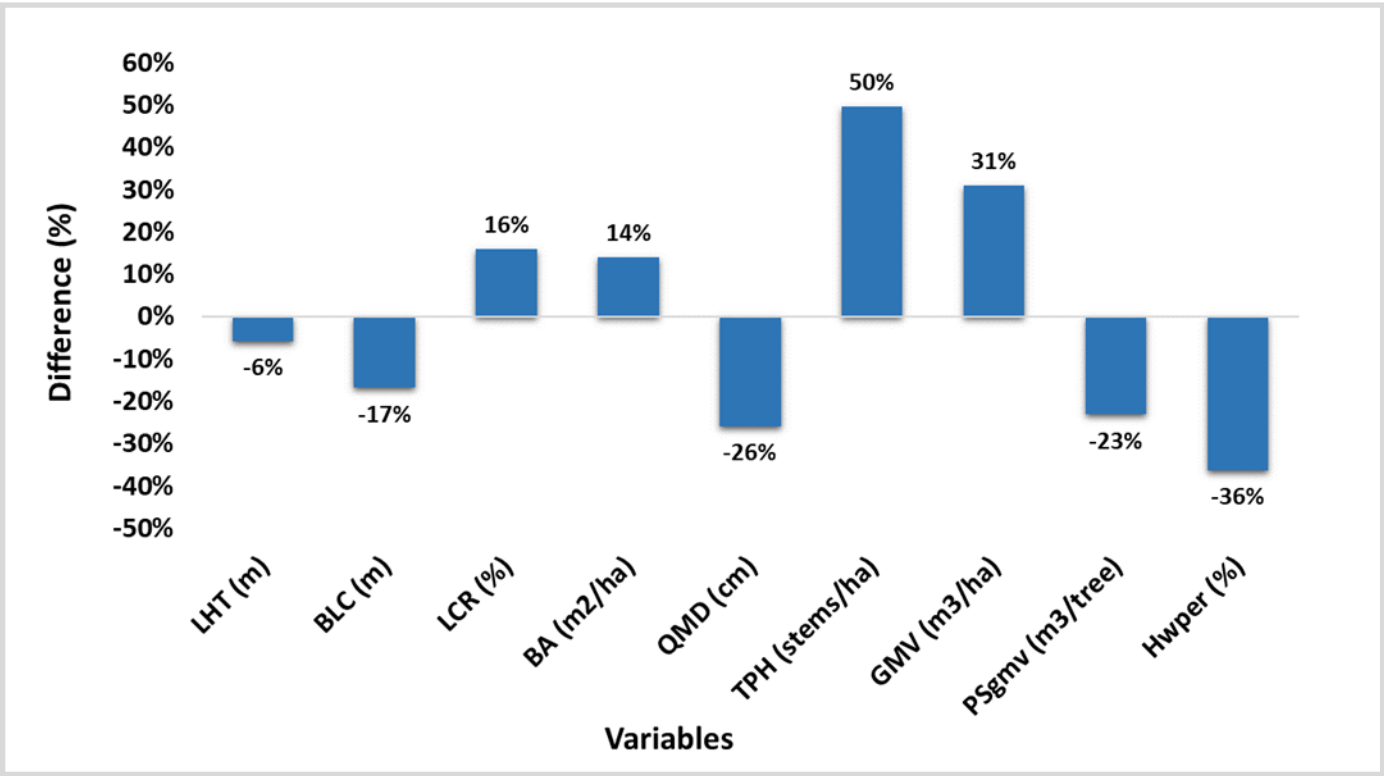

Figure 1. Mean difference comparisons between EFI and FI data for the nine variables.

Nine equations were developed for each EFI variables using environmental, stand quality and dendrometric field data. The fit statistics of the different equations are shown in Table 1. The highest correlation equation was recorded for merchantable tree density (r2=64.91%) which includes one environmental variable (elevation) and four dendrometric field variables (basal area, quadratic mean diameter, average piece size, percentage of hardwood composition) indicating a significant effect of these variables on merchantable tree density. Basal area equation also had a relative high correlation factor (r2=60.9%) with soil being the second most significant variable. At the other end, the lowest correlations were obtained for live crown ratio (r2=19.09%) and base to live crown (19.3%) equations.

Table 1: List of candidate models, the selected variables, the equations with the parameter estimates and the associated R-squared.

Candidate models	Variables selected	Equations and Parameter estimates	R squared (%)
Average height-LHT (m)	Ecosite + Elevation + BLC + LCR + BA + Hwper	LHT= -18.172 +2.160 ecosite4c + 0.903 ecosite5 - 0.032 ecosite5c+ 0.736 ecosite6 + 0.405 ecosite7+ 0.136 ecosite7c+2.674 ecosite8 +0.008 Elevation + 0.953 BLC + 19.267 LCR -0.032 BA - 1.896 Hwper	43.2
Base to live crown-BLC (m)	Ecosite + Elevation + LHT + BA	BLC= -4.112 +1.112 ecosite4c -1.042 ecosite5 - 1.111 ecosite5c - 3.136 ecosite6 - 0.809 ecosite7 - 0.772 ecosite7c - 0.879 ecosite8 + 0.005 Elevation + 0.227 LHT + 0.031 BA	19.3
Live crown-LCR (%)	Ecosite + Elevation + LHT + BA	LCR= -0.02 -0.056 ecosite4c +0.096 ecosite5+0.07 ecosite5c +0.231 ecosite6+0.073 ecosite7 +0.122 ecosite7c +0.156 ecosite8 - 0.00038 Elevation +0.004 LHT -0.0045 BA	19.09
Basal area-BA (m^2/ha)	Soil + QMD + TPH + PSgmv	BA= -34.185 -4.667 SoilGE- 0.664 SoilGF - 0.920 SoilHM-3.665 SoilKE -6.650 SoilLL -1.853 SoilMG + 3.517 SoilMV+ 2.256 SoilTH -6.234 SoilTT -7.057 SoilUN+ 1.271 SoilVI + 0.492 QMD + 0.019 TPH + 11.61 PSgmv	60.91
Quadratic mean diameter-QMD (cm)	Elevation + BA + TPH	QMD= -2.473 +0.036 Elevation + 0.198 BA -0.009 TPH	52.79
Density-TPH (stems/ha)	Elevation + BA + QMD + PSgmv + Hwper	TPH= 162.102-1.953Elevation +26.810 BA-16.736 QMD - 824.264 PSgmv + 249.553ter_Hwper	64.91
Gross merchantable volume-GMV (m^3/ha)	Soil + Elevation + QMD + TPH	GMV= -0.017 -0.06 SoilGE -0.294 SoilGF -0.240 SoilHM+ 11.22 SoilKE -40.52 SoilLL-25.22 SoilMG -12.55 SoilMV+ 18.47 SoilTH-20.79 SoilTT-125.3 SoilUN-22.49 SoilVI + 0.181 Elevation + 4.552 QMD+ 0.048 TPH	43.2
Average piece size-PSgmv (m^3/tree)	Soil + BA + TPH + GMV + Hwper	PSgmv= 0.1481 + 0.167 SoilGE+0.196 SoilGF + 0.061 SoilHM + 0.016 SoilKE + 0.10 1SoilLL + 0.073 SoilMG + 0.121 SoilMV+ 0.0473 SoilTH + 0.04 SoilTT + 0.008 SoilUN + 0.115 SoilVI +0.0012 BA- 0.0009 GMV -0.0003 TPH -0.059 Hwper	48.35
Percent of hardwood composition-Hwper (%)	Soil + Elevation+ LHT + BA + TPH + GMV	Hwper= -0.027 +0.237 SoilGE -0.213 SoilGF+0.039 SoilHM- 0.043 SoilKE-0.25 SoilLL-0.091 SoilMG+0.26 SoilMV-0.214 SoilTH+0.158 SoilTT+0.412 SoilUN -0.13 SoilVI -0.0002 Elevation+ 0.019 LHT +-0.019 BA +0.0001 TPH + 0.001 GMV	31.09

After average height, environmental variables were either the most or the second most equation significant variables (Fig. 2). Elevation explanatory variable is present at different levels of importance in 7 of the 9 equations but was particularly significant in quadratic mean diameter and merchantable tree density equations. Base to live crown and live crown ratio equations both had ecosite as a relative significant variable. Significant importance of soil is seen in gross merchantable volume and percentage of hardwood composition equations.

Dendrometric variables are the most significant explanatory variables in 7 of the 9 equations. Merchantable tree density is the most frequent one being found in 3 equations (basal area, quadratic mean diameter and average piece size).

Except for the density and average piece size, all the equations present a negative slope, showing that all response variables are negatively associated with explanatory variables (Table1). The different statistics associated with the parameter estimates indicate that the equation coefficients are highly significant, especially soil and ecosite variables.

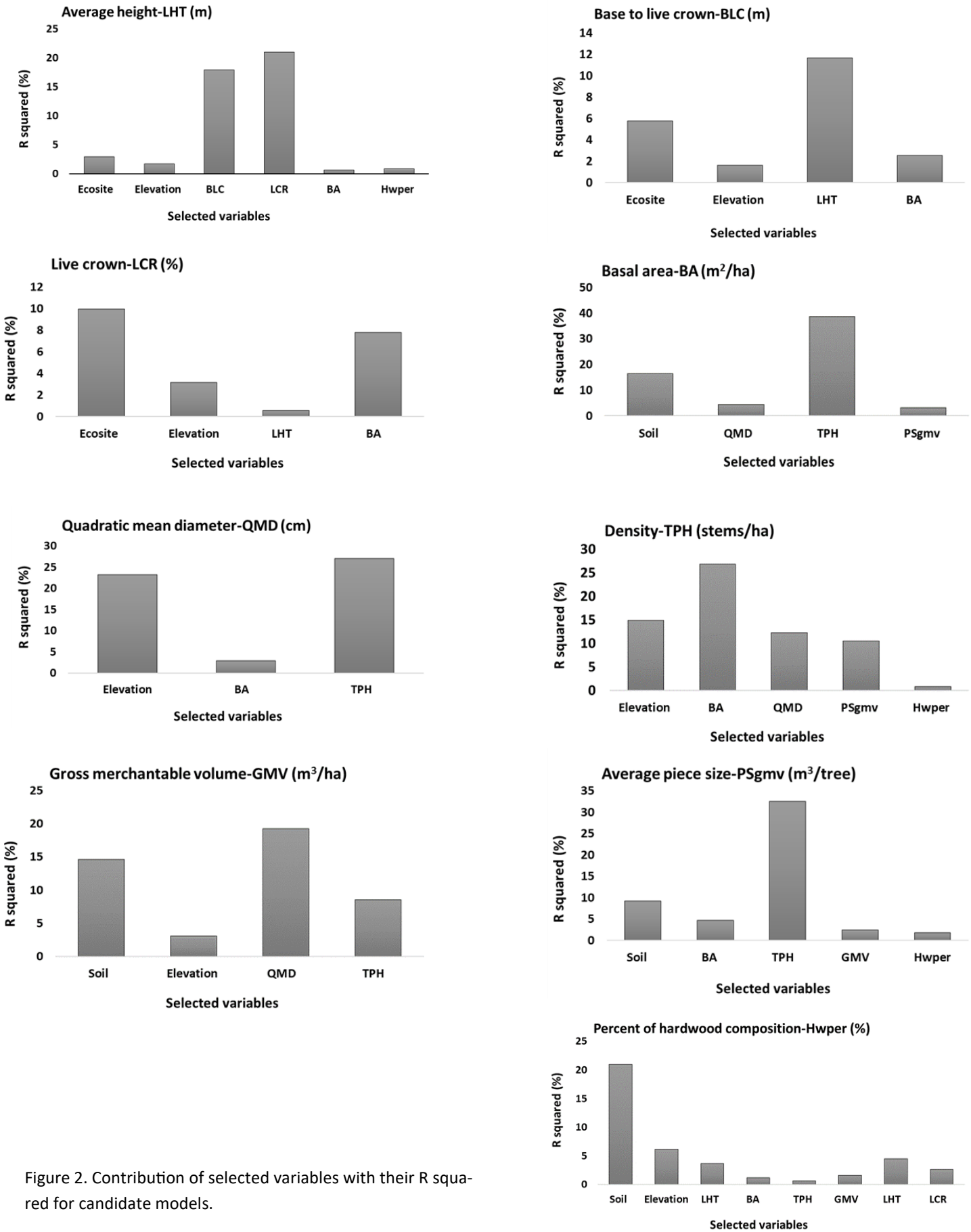

Figure 2. Contribution of selected variables with their R squared for candidate models.

DISCUSSION AND CONCLUSION

Being the first Canadian Province to complete aerial-LiDAR coverage with an Open Data License, New Brunswick is a pioneer in the development of LiDAR variables. The Northern Hardwoods Research Institute showed in a previous study that there were significant differences for some variables between traditional forest inventory and EFI.

NHRI undertook this preliminary research initiative expecting to find potential explanations for the differences observed between EFI and traditional forest inventory information using dendrometric and environmental variables.

In order to analyze those results adequately it is important to know how aerial-LiDAR metrics are converted into dendrometric information. In our case, it is normally done by modelling relationships observed between lidar plot metrics (not dendrometric variables) and field plot measurements (dendrometric variables) and then applying the models across the entire lidar acquisition area.

Another thing to consider when analyzing the results is the correlation level (R-squared) found in the different equations. Information like this found at https://people.duke.edu/~rnau/rsquared.htm should also be kept in mind: "...if the dependent variable is a properly "stationarized" series (e.g., differences or percentage differences rather than levels), then an R-squared of 25% may be quite good. In fact, an R-squared of 10% or even less could have some information value when you are looking for a weak signal in the presence of a lot of noise in a setting where even a very weak one would be of general interest. Sometimes there is a lot of value in explaining only a very small fraction of the variance, and sometimes there isn't."

The results of this study highlighted that some dendrometric and environmental variables were a key factor explaining the variance and increasing accuracy for all tested models. In fact, environmental variables such as ecosite and elevation were present in the different retained models underlining their important contribution in predicting tendencies.

EFI height underestimation result should be taken with caution. It could be explained by the fact that field inventories tend to overestimate height especially hardwood trees because of the crown shape. In this current case, we recommend that LiDAR-derived heights are usually more accurate.

Also, EFI seems to constantly underestimate some variables like quadratic mean diameter, average piece size and percentage of hardwood composition when compared to traditional field inventory data. This could be due to field measurements not being bounded by algorithm limits like EFI data are.

The percent of hardwood composition was present as a key variable in average height, stem density and average piece size models, showing that this variable has a significant effect on predicting tendencies.

These equations were developed to investigate if it's possible to correct the prediction error without recalculating EFI's from raw data. But, this study suggests that adding variables other than EFI metrics (point cloud data) alone in the analysis (models) could increase EFI data accuracy when producing them from scratch.

REFERENCES

-Baral, S.K., Danyagri, G., Girouard, M., Hébert, F., and Pelletier, G. 2016. Effects of suppression history on growth response and stem quality of extant northern hardwoods following partial harvests. Forest Ecology and Management. 372: 236–246.

-Burnham KP, Anderson DR. 2002. Model Selection and Multimodel Inference A Practical Information-Theoretic Approach.

- Chouinard, C.-A. 2020. Enhanced Forest Inventory vs Field Inventory: Preliminary Comparison of Inventory Results. Northern Hardwoods Research Institute, Edmundston, New Brunswick.

-FORUS, Research. 2016. http://www.forusresearch.com/bgi.php.

-Hennigar, C., Weiskittel, A., Allen, H.L., and MacLean, D.A. 2016. Development and evaluation of a biomass increment-based index for site productivity. Canadian Journal of Forest Research. 47: 400–410.

-Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M., Gessler, P.E. and Morgan, P., 2006. Regression modelling 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):126–138.

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

-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: 257–267.

-Pelletier, G., Landry, D., and Girouard, M. 2016. A tree classification system for New Brunswick. Northern Hardwoods Research Institute, Edmundston, New Brunswick.

FOR MORE INFORMATIONS, CONTACT:

info@hardwoodsnb.ca

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