FORECASTING FOREST INVENTORY FOR SPRUCE PLANTATIONS USING AIRBORNE LASER SCANNING DATA

by

Sean M. Lamb

BSc. Forestry and Environmental Management, University of New Brunswick, 2014

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Science in Forestry

in the Graduate Academic Unit of Forestry and Environmental Management

Supervisor: David A. MacLean, PhD, Forestry and Environmental Management

Examining Board: Fan-Rui Meng, PhD, Forestry and Environmental Management
Yun Zhang, PhD, Geodesy and Geomatics Engineering

This thesis is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

January 2018

© Sean Lamb, 2018

ABSTRACT

This thesis focuses on the imputation of tree-level inventory for light detection and ranging (LiDAR) grid cells across 83 000 ha of spruce (*Picea* sp.) plantations to permit forecasting forest inventory using a tree-list growth model. Grid cells with LiDAR-derived inventory predictions were matched with one of over 5500 spruce plantation sample plot measurements based on planted species and minimum sum of squared difference between total and merchantable basal area, total and merchantable volume, top height, and merchantable quadratic mean diameter. Imputed tree lists resulted in inventory variable values that were highly correlated with observed values and statistically equivalent basal area distributions. When input into a locally calibrated tree-list growth model, variable increments predicted using imputed tree lists were strongly correlated with those using measured tree lists, resulting in predicted volumes that were similar to observed volumes. Using forecasted inventory variables, annual commercial thinning treatments were planned at grid-cell resolution from 2018–2020.

DEDICATION

To my wife, Rosanna

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. David MacLean, and advisory committee, Dr. Chris Hennigar and Dr. Doug Pitt, for their guidance, support, and mentorship. I would like to thank the staff at J.D. Irving, Limited, particularly Mr. David Young, Mr. Danielle Arsenault, and Mr. Greg Adams, as well as Mr. Riley Côté-DeMerchant with Leading Edge Geomatics for their input and assistance with this project. I am particularly grateful to my wife for her unconditional support and encouragement.

This research was funded by the AWARE (Assessment of Wood Attributes using Remote sEnsing) Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development grant to a team led by Dr. Nicholas Coops and by J.D. Irving, Limited. I was also funded by an NSERC Canada Graduate Scholarship.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: GENERAL INTRODUCTION

1.1 Introduction

Active remote sensing technology for predicting forest inventory, specifically light detection and ranging (LiDAR), has vastly improved the resolution of forest inventory. LiDAR sensors emit pulses of laser light to the surface below from a fixed-wing aircraft and measure the roundtrip time from emission to return to calculate distance travelled to target. When the target is a forest canopy, some of the light will reflect off of the canopy (first return) while the remaining light may penetrate and come into contact with stems, branches, or leaves (intermediate returns) and finally, the ground (final return) (Reutebuch et al. 2005; Akay et al. 2009). Using a global positioning system (GPS), the geographic location of the aircraft is recorded during flight, while the three-axis orientation (pitch, roll, and yaw) is determined with an inertial measurement unit (Wehr and Lohr 1999). Precise vertical and horizontal location of each return is determined by mathematically integrating the LiDAR return data with GPS coordinates and the angular measurements from the inertial measurement unit to produce a single file referred to as a "point cloud" (Evans et al. 2006). Processing the point cloud to identify ground returns allows for the development of highly accurate digital elevation models (DEM) relative to sea level, with approximately 15-cm error in open areas (Pereira and Janssen 1999) and 31-cm under dense conifer forest cover (Reutebuch et al. 2003). Canopy heights are calculated by subtracting the ground elevation (from the DEM) from the elevations of the non-ground first returns, while the subtraction of the ground elevation from the intermediate returns provides the vertical distribution of the forest structure (Dubayah and Drake 2000; White et al. 2013).

Apart from the DEM and canopy heights that can be obtained directly from the LiDAR point cloud, additional variables can be predicted based on the vertical distribution

and density of LiDAR returns. Forest inventory variables can be predicted by determining the statistical relationship between variables measured in ground calibration plots and the corresponding LiDAR point cloud. Around the time of the flight, calibration plots are established to measure variables of interest (i.e., species, diameter at breast height [DBH], total height, etc.), which can then be used to calculate additional plot-level variables (i.e., total volume, total basal area, quadratic mean diameter, etc.). Plot locations are geo-referenced using a survey-grade GPS unit with sub-metre accuracy and selected to capture the range of forest structure variability (White et al. 2013). Using a geospatial processing program, (e.g., ArcGIS [ESRI, Redlands, CA]) the point cloud is "clipped" to the corresponding calibration plot locations. LiDAR point-cloud statistics are calculated for each associated calibration plot and include measures such as mean return height, return height percentiles, and standard deviation of return heights. Using parametric or non-parametric methods, the statistical relationship between inventory variables and point-cloud statistics is determined and used to predict variable values in non-sampled areas at resolutions matching the size of the calibration plots (e.g., 20 × 20 m) (Penner et al. 2013; White et al. 2013). This approach has been used for the prediction of forest inventory variables including average tree height (Næsset 1997; Magnussen and Boudewyn 1998), average tree DBH (Woods et al. 2008; Hayashi et al. 2014), basal area (Means et al. 2000; Holmgren et al. 2004), and volume (Nilsson 1996; Lim et al. 2003). The precision levels of these estimates have been shown to be comparable to those obtained from typical field inventories (i.e., sampling stands with several field plots) (Næsset and Bjerknes 2001; Næsset et al. 2005).

Quantifying forest inventory variables at 20 × 20 m resolution across entire forest estates is a vast improvement over existing inventory methods, which rely heavily on the sampling and the measurement of just one or two percent of the areas of interest. Typically, averaged estimates of stand volumes are obtained using temporary sample plots, which allow for the development of species- and management-based yield curves, or by using forest growth models to forecast development of stand structure and composition. Although this inventory method provides the means to obtain forecasted stand development over time, these estimates of future forest inventory are broad and do not capture between- and within-stand spatial variability. As a result, traditional inventory must be supplemented with costly and time consuming pre-harvest ground verification to ensure that stands meet wood supply requirements. However, LiDAR-derived inventory permits the quantification of current between- and within-stand spatial variability, and this high structural resolution offers the potential to then 'grow' the inventory spatially using existing stand growth forecasting methods. To take advantage of this potential, a method for forecasting LiDAR-derived inventory at grid-cell resolution is needed. The main requirement is accurate estimation, for each 20 × 20 m grid cell, of all variables required as input to forest growth models.

1.2 Overall goal and specific objectives

This study is part of the larger Assessment of Wood Attributes from Remote Sensing (AWARE) Natural Science and Engineering Research Council of Canada (NSERC) Collaborative Research and Development (CRD) funded project. The overall goal of the AWARE project is to use current remote sensing technologies to enhance forest

inventories and provide the fundamental information to improve forest management practices in Canada's forests. This specific study will assess the ability to forecast LiDAR-derived forest inventory to spatially plan annual commercial thinning treatments using LiDAR-derived inventory variables. Specific objectives are to:

1. Develop and test a method to impute tree-level inventory for individual LiDAR grid cells in spruce (*Picea* sp.) plantations using existing forest inventory plot measurements and a nearest neighbor matching algorithm (Chapter 2);
2. Assess the accuracy of forecasted forest inventory using tree-level inventory imputed for individual LiDAR grid cells and a locally calibrated tree-list growth model (Chapter 3);
3. Spatially plan annual commercial thinning treatments using a commercial thinning ranking system (Chapter 3).

1.3 Thesis structure

This thesis is presented in an "article" format. The body of the thesis consists of two distinct chapters (Chapters 2 and 3), which are described below. The last chapter of the thesis (Chapter 4) presents a general discussion of the conclusions from these chapters.

Chapter 2 presents methods for imputing tree-level inventory for LiDAR grid cells in spruce plantations using existing measurements of permanent sample plots in other spruce plantations throughout New Brunswick. This chapter has been published in the Canadian Journal of Remote Sensing (*Lamb, S.M., MacLean, D.A., Hennigar, C.R., and Pitt, D.G. 2017. Imputing tree lists for New Brunswick spruce plantations through nearest-neighbor*

matching of airborne laser scan and inventory plot data. Can. J. Remote Sens. 43(3): 269–285).

Chapter 3 evaluates the accuracy of imputed tree lists (trees by species and size) for LiDAR grid cells and a locally calibrated tree-list growth model to forecast inventory variables for 10 years in 1-year time steps and then demonstrates the application of this forecasted inventory for spatially planning annual commercial thinning treatments. This chapter is planned for submission for publication as: *Lamb, S.M., MacLean, D.A., Hennigar, C.R., and Pitt, D.G. 2018. Forecasting forest inventory using imputed tree lists for LiDAR grid cells and a tree-list growth model,* in the journal *Forests.*

I was the primary author for the manuscripts in the body of this thesis (Chapters 2 and 3). Dr. David A. MacLean, Dr. Chris R. Hennigar, and Dr. Douglas G. Pitt are included as co-authors on these manuscripts.

1.4 References

Akay, A.E., Oğuz, H., Karas, I.R., and Aruga, K. 2009. Using LiDAR technology in forestry activities. Environ. Monit. Assess. 151(1–4): 117–125.

Dubayah, R.O., and Drake, J.B. 2000. Lidar remote sensing for forestry. J. For. 98(6): 44–46.

Evans, D.L., Roberts, S.D., and Parker, R.C. 2006. LiDAR - A new tool for forest measurements? For. Chron. 82(2): 211–218.

Hayashi, R., Weiskittel, A., and Sader, S. 2014. Assessing the feasibility of low-density LiDAR for stand inventory attribute predictions in complex and managed forests of Northern Maine, USA. Forests 5(2): 363–383.

Holmgren, J. 2004. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. Scand. J. For. Res. 196(19): 543–553.

Lim, K., Treitz, P., Baldwin, K., Morrison, I., and Green, J. 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. Can. J. Remote Sens. 29(5): 658–678.

Magnussen, S., and Boudewyn, P. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. Can. J. For. Res. 28: 1016–1031.

Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., and Abstract, C.J.H. 2000. Predicting forest stand characteristics with airborne scanning lidar. Photogramm. Eng. Remote Sens. 66(11): 1367–1371.

Næsset, E. 1997. Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS J. Photogramm. Remote Sens. 52(2): 49–56.

Næsset, E., and Bjerknes, K.O. 2001. Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. Remote Sens. Environ. 78(3): 328–340.

Næsset, E., Bollandsås, O.M., and Gobakken, T. 2005. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. Remote Sens. Environ. 94(4): 541–553.

Nilsson, M. 1996. Estimation of tree heights and stand volume using an airborne lidar system. Remote Sens. Environ. 56(95): 1–7.

Pereira, L.G., and Janssen, L.L.F. 1999. Suitability of laser data for DTM generation: a case study in the context of road planning and design. ISPRS J. Photogramm. Remote Sens. 54(4): 244–253.

Penner, M., Pitt, D.G., and Woods, M.E. 2013. Parametric vs. nonparametric LiDAR models for operational forest inventory in boreal Ontario. Can. J. Remote Sens. 39(5): 426–443.

Reutebuch, S.E., Andersen, H.E., and McGaughey, R.J. 2005. Light detection and ranging (LIDAR): an emerging tool for multiple resource inventory. J. For. 103(6): 286–292.

Reutebuch, S.E., McGaughey, R.J., Andersen, H.E., and Carson, W.W. 2003. Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. Can. J. Remote Sens. 29(5): 527–535.

Wehr, A., and Lohr, U. 1999. Airborne laser scanning—an introduction and overview. ISPRS J. Photogramm. Remote Sens. 54: 68–82.

White, J.C., Wulder, M.A., Varhola, A., Vastaranta, M., Coops, N.C., Cook, B.D., Pitt, D., and Woods, M. 2013. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. Can. For. Serv., Can. Wood Fibre Cent. Inf. Rep. FI-X-010. 50p.

Woods, M., Lim, K., and Treitz, P. 2008. Predicting forest stand variables from LiDAR data in the Great Lakes - St. Lawrence forest of Ontario. For. Chron. 84(6): 827–839.

CHAPTER 2: IMPUTING TREE LISTS FOR NEW BRUNSWICK SPRUCE PLANTATIONS THROUGH NEAREST-NEIGHBOR MATCHING OF AIRBORNE LASER SCAN AND INVENTORY PLOT DATA

Paper published as:

Lamb, S.M., MacLean, D.A., Hennigar, C.R., and Pitt, D.G. 2017. Imputing tree lists for New Brunswick spruce plantations through nearest-neighbor matching of airborne laser scan and inventory plot data. Can. J. Remote Sens. 43(3): 269–285.

2.1 Abstract

Light detection and ranging (LiDAR) has greatly improved the spatial resolution and accuracy of operational forest inventories. However, a cost-effective method to impute species-specific tree-level inventory is needed, to be used as input to tree or stand growth models to project single-point-in-time LiDAR estimates. We evaluated a method to match stand structural variables estimated from LiDAR to those in a library of over 5500 sample plot measurements to impute tree lists for LiDAR grid cells across 83 000 ha of spruce (*Picea* sp.) plantations. Matches were determined based on planted species and minimum sum of squared difference between six inventory variables. Forest inventory variables obtained by the plot matches were highly correlated (r = 0.91–0.99) with those measured on 98 validation plots. Basal area distributions derived from plot matching were statistically equivalent to those observed on the validation plots 86% of the time (α = 0.05). When we aggregated the predictions for all validation plots, there was minimal difference between predicted and actual basal area distributions by planted species and species compositions were similar. Plot matching is a valid method to impute tree lists for LiDAR cells that combines the wealth of existing plot data with high resolution LiDAR-derived variables.

2.2 Introduction

Use of active remote sensing technology in industrial forestry, specifically airborne laser scanning (ALS, a.k.a. light detection and ranging [LiDAR]), has improved the ability of forest managers to make sound forest operations planning decisions with minimal field verification (Woods et al. 2011). LiDAR has improved the accuracy and resolution of digital elevation models compared to those derived from photogrammetry or topographic maps (Hodgson et al. 2003; Schiess and Krogstad 2003), and greatly assists in predicting forest hydrology (James et al. 2007; Murphy et al. 2008b), constructing accurate wet areas maps (Murphy et al. 2008a; White et al. 2012), identifying potential wildlife habitat (Hyde et al. 2005; Martinuzzi et al. 2009), and reducing costs of forest roads planning (Aruga et al. 2005).

LiDAR also provides a means to estimate forest inventory variables in a wall-to-wall fashion at high, previously unobtainable spatial resolutions (e.g., grid-cell predictions of 10–30 m), with stand-level errors comparable to those of typical field inventories (Næsset et al. 2005; Penner et al. 2015). Estimates are based on determining regression or non-parametric relationships between inventory variables collected from field-measured calibration plots and statistics describing the corresponding vertical distribution and density of LiDAR returns; commonly referred to as the area-based approach (e.g., White et al. 2013; Penner at al. 2013). The area-based approach has been used to predict mean height (Næsset 1997; Magnussen and Boudewyn 1998), basal area (Means et al. 1999; Holmgren 2004), volume (Means et al. 2000; Lim et al. 2003), quadratic-mean diameter (Woods et al. 2011; Hayashi et al. 2014), and basal area distribution (Maltamo et al. 2009; Gobakken and Næsset 2004).

The ability to quantify forest inventory variables at high spatial resolution across entire forest estates is a vast improvement over the traditional strategic inventory methods currently used in many of Canada's forests. Currently, sample plots from a range of stand types at varying stages of development are typically used to construct averaged age-based yield curves for broad species compositions and management types, which are then applied to photo-interpreted forest inventory polygons. This permits coarse estimation of current and future forest inventory variables, such as stand volume or basal area, suitable for strategic (long-term, 20-year) forest management decision making, but does not capture within-stand spatial variability for tactical (5-year) and operational (1-year) decision making. Tactical and operational planning must then be supplemented with costly pre-harvest ground verification or "operational cruising" to confirm that stands slated for harvest meet wood supply requirements. In contrast, the spatial resolution of LiDAR-derived forest inventory allows characterization of within- and between-stand variability, with minimal ground verification, and provides a stronger link between operational and strategic forest management (Næsset 2007; Woods et al. 2011).

Although LiDAR-derived forest inventory estimates are considered superior to conventional inventory methods, LiDAR acquisition and calibration is expensive and shares the problem with all inventory methods of being a single-point-in-time estimate. To extend the life of expensive LiDAR acquisition and forecast forest development at the grid-cell resolution, LiDAR-derived inventory variables (e.g., mean tree height, stem density, etc.) have been used as inputs in process-based forest models (Härkönen et al. 2013) and for matching grid cells with yield curves derived from size-class growth models (Tompalski et al. 2016). These growth models did not provide individual tree-level

forecasts and therefore do not provide the means to track important forestry and biodiversity variables such as species-specific product volumes and percent composition.

One alternative method to forecast LiDAR-derived variables and obtain more detailed inventory measures would be to use individual tree, distance-independent growth and yield models that predict development of individual tree characteristics, otherwise known as tree-list growth models (Crookston and Dixon 2005). These models combine known site productivity and measured tree-level inventory (i.e., species, diameter at breast height [DBH], and height) for the area of interest to predict forest inventory development over time and allow forest managers to quantify species-specific product returns, changes in species composition, and changes in wildlife habitat (Crookston and Dixon 2005; FORUS Research 2011). However, since tree-list growth models require tree-level inventory, there is a need for a practical means of bridging the gap between LiDAR-derived area-based estimates and tree lists.

Prior to the use of LiDAR in industrial forest management, non-parametric k-nearest neighbor imputation methods were used to impute complex stand variables, such as diameter distribution, for target data based on accurately measured reference data (Haara et al. 1997; Maltamo and Kangas 1998). These non-parametric methods have been effective in the prediction of LiDAR-derived species-independent diameter distribution (Gobakken and Næsset 2004; Maltamo et al. 2009, Shang et al. 2017), species-specific stand variables (Packalén and Maltamo 2006; 2007), and have been used to accurately predict species-specific diameter distributions (Packalén and Maltamo 2008; Peuhkurinen et al. 2008).

Although LiDAR-derived species-specific diameter distribution estimates could be used in a tree-list growth model, the substantial number of calibration plots required to capture a range of complex tree-level conditions may be prohibitive. For example, Packalén and Maltamo (2008) and Peuhkurinen et al. (2008) used nearly 500 sample plots to obtain accurate species-specific estimates for their respective 2000 ha and 1200 ha landbases. If species is ignored, basal area distribution may be accurately estimated for large areas with as few as 100 sample plots (Maltamo et al. 2009).

Using more than one observation to create an average estimate of tree-level inventory, as done by Packalén and Maltamo (2008) and Peuhkurinen et al. (2008) ($n = 5$), reduces estimation error, but ultimately creates a unique tree list for each LiDAR grid cell. For large forests, data storage and forecasting of each grid cell with a tree-list model may be problematic. Falkowski et al. (2010) showed that use of one nearest neighbor and a small sample of reference plots (88) resulted in good stand-level estimates for basal area and volume ($r > 0.88$), but only moderate to low estimates for density of trees > 23 cm and ≤ 23 cm (DBH) ($r = 0.55–0.64$ and 0.09–0.39, respectively) and species classification that was ≤ 31% better than random chance. The use of crown segmentation of LiDAR returns has also been shown to be an effective method for estimating tree-level inventory (Vauhkonen et al. 2010; Lindberg et al. 2013), however, the issue of a large number of unique tree lists remains. Additionally, a pulse density of 5–10 pulses·m^{-2} is typically required for accurate tree delineation (Hyyppä and Inkinen 1999; Persson et al. 2002), making this a more expensive option when compared with the pulse density required to make accurate inventory estimates using the area-based approach (as low as 0.1 pulses·m^{-2}) (Maltamo et al. 2006; Gobakken and Næsset 2008). Ultimately, a method that would

allow for accurate imputation of tree-level inventory for area-based LiDAR grid cells, without the need for large amounts of expensive calibration field data or high density LiDAR, would be an attractive option for forest managers.

Existing, historical permanent and temporary sample plot data could provide the variety of tree-level conditions needed to impute tree-level inventory for LiDAR grid cells, without additional cost. For example, over the past 30 years, New Brunswick Department of Energy and Resource Development (NB ERD) has established over 3600 fixed-area permanent sample plots, measured two to seven times (Porter et al. 2001), and over 65 000 variable-radius temporary sample plots (MacDonald 2008) distributed throughout a full range of management types and development stages. A nearest neighbor method that combines the accuracy of LiDAR-derived forest inventory with the wealth of stand structure data from existing sample plots may be a viable method to impute tree-level inventory.

The objective of this paper was to develop and test a method using existing forest inventory plot measurements to impute a tree list for individual LIDAR grid cells in spruce (*Picea* sp.) plantations. We hypothesized that a LiDAR grid cell characterized by stand-level variables such as basal area, volume, and quadratic mean diameter, will share the tree-level inventory of a sample plot that has similar stand-level variables. The proposed method matches stand-variable predictions from LiDAR grid cells in spruce plantations of the target area with similar variables in ancillary plot data from spruce plantations throughout New Brunswick using a distance-independent, nearest neighbor matching algorithm (e.g., Crookston and Finley 2008). Matches would be determined based on planted species and by calculating the minimum cumulative difference between multiple

stand-level inventory variables, ignoring spatial location. Once a match is determined, the tree list from the selected plot could be used as a surrogate tree list for the LiDAR grid cell. Such imputed values, if acceptably accurate, would provide species-specific tree size-distribution information important for operational and tactical planning, as well as the foundational data for forecasting stand growth using a tree-list growth model.

2.3 Methods

2.3.1 Study area

The J.D. Irving, Limited Black Brook District, located in northwestern New Brunswick, Canada (47°9’51”N to 67°55’27”W) (Fig. 2.1a), is a 210 000 ha contiguous forest that contains some of the most intensively managed forest lands in Canada (MacLean et al. 2010). The District falls within Ecoregion 3 (Zelazny et al. 2008) and contains 65 000 ha of high-quality shade-tolerant hardwoods, managed for veneer and sawlogs using selection and group shelterwood harvesting systems, and 83 000 ha of spruce plantations, commercially thinned one to three times before final harvest. These plantations include 37 000 ha of black spruce (*Picea mariana* (Mill.) B.S.P.), 33 500 ha of white spruce (*Picea glauca* (Moench) Voss), 11 000 ha of Norway spruce (*Picea abies* (L.) Karst.), and 1500 ha of red spruce (*Picea rubens* Sarg.). Growth and yield of these plantations have been monitored using almost 400 permanent sample plots, each measured every 3–5 years since 1980.

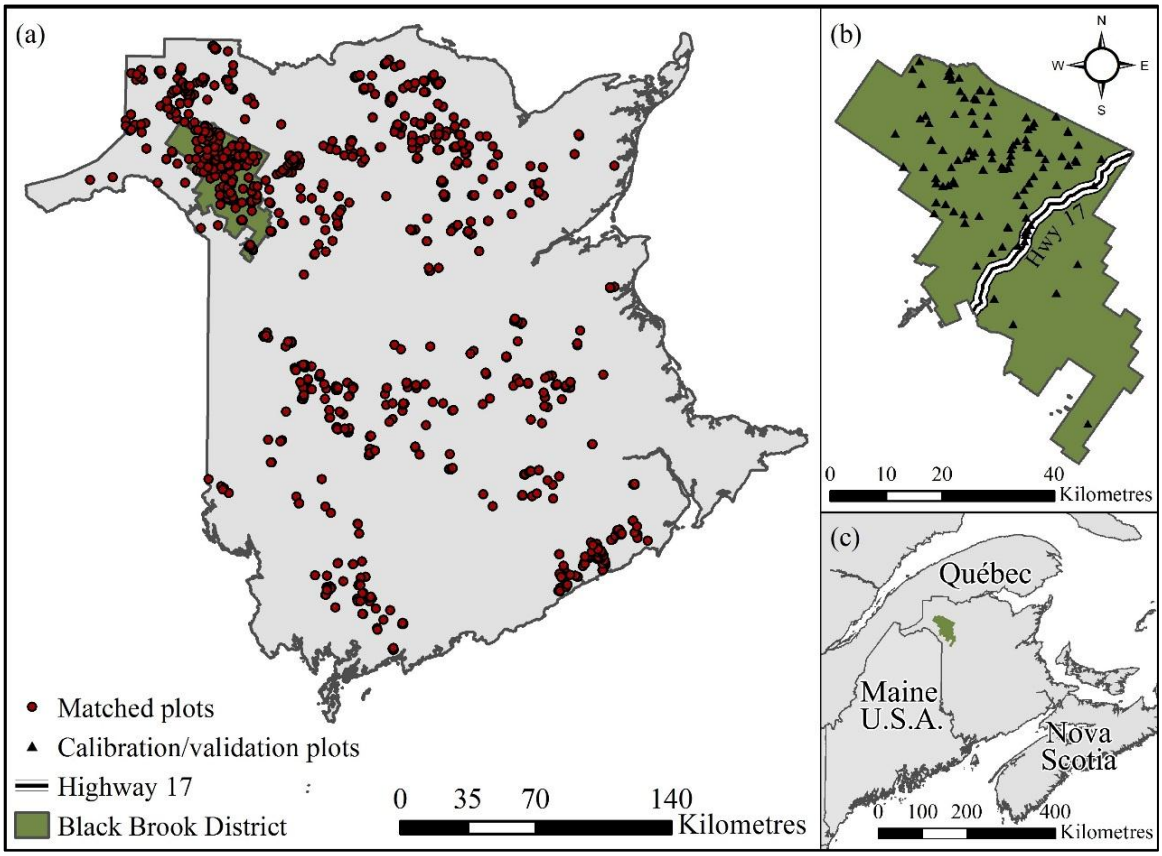

Fig. 2.1. Locations of (a) spruce plantation forest inventory plot measurements throughout New Brunswick that were matched with LiDAR grid cells, (b) LiDAR calibration/validation plots in spruce plantations located in the Black Brook District, and (c) the Black Brook District in northwestern New Brunswick. LiDAR was acquired in 2011 and 2013 for all areas north and south of Highway 17, respectively (b).

2.3.2 LiDAR data and calibration/validation plot data

High density LiDAR (6 pulses·m^{-2}, with up to 8 returns per pulse) was acquired for the northern portion (north of Highway 17, Fig. 2.1b) of the District in September 2011, by a private LiDAR contractor (Leading Edge Geomatics 2015). These data covered 60% of the District area. LiDAR data for the remaining southern portion of the area was acquired by the same contractor in July 2013, and scanning parameters were identical for both acquisitions (Table 2.1).

Table 2.1. LiDAR acquisition specifications for the 2011 and 2013 flights.

Parameter	Specification
Sensor	Riegl LMSQ-680i
Aircraft	Cessna 172 & Cessna 206
Altitude	670 m
Swath width	650 m
Overlap	50%
Field of view	40°–60°
Nominal pulse density	$6 \cdot m^{-2}$
Returns/pulse	8
Maximum pulse repetition frequency	150 000 Hz
Leaf on/off	On
Horizontal accuracy	~0.18 cm
Vertical accuracy	< 15 cm
Scan Rate	87 Hz

During the summer of 2012, 95 circular 400-m^2 plots were established by the Canadian Wood Fibre Centre, Natural Resources Canada, for use in calibrating LiDAR predictions: 41 in black spruce, 42 in white spruce, and 12 in Norway spruce plantations to capture a range of ages (5–48 years old) and management types (no commercial thinning, one commercial thinning, and two commercial thinnings) (Fig. 2.1b) (Leading Edge Geomatics 2015). All of these plots were located in the area flown in 2011. An additional 5 plots were established within the area flown in 2013 during the summer of 2014 to capture stand conditions not present in the original plots: 3 in white spruce and 2 in Norway spruce. Plot centres were geo-referenced using an SX Blue II GPS with a minimum of 300 GPS points recorded at each location to obtain sub-metre accuracy. In each plot, trees with DBH ≥ 3 cm were assessed for species and status (live or dead). DBH was measured on all trees and height was measured on three randomly selected trees within each 5-cm DBH class and species.

Stand variables including total and merchantable basal area (BA_t, BA_m), gross total and merchantable volume (VOL_t, VOL_m), top height (TopHt), and merchantable quadratic

mean diameter (QMD_m) were calculated from the tree measurements made on each plot using the formulae outlined in Table 2.2 and 9.1 cm DBH as the lower merchantability limit. Volumes were calculated using the NB ERD specifications, which include Li et al. (2012) taper equations for all commercial species and Honer et al. (1983) for all non-commercial species.

Table 2.2. Equations and descriptions of the inventory variables derived from LiDAR and the plot-matching approach. All variables include live trees only; total and merchantable variables include minimum DBH of 3 cm and 9.1 cm, respectively.

Inventory variable	Description
Basal area (BA_t or BA_m; $m^2 \cdot ha^{-1}$)	Fixed area plot: $\sum_{i=1}^{n} DBH_i^2 \times \frac{\pi(m^2)}{40\,000} \times TF$; where *TF* is tree factor for the ith tree based on plot size. Variable radius plot: $\sum_{i=1}^{n} BAF_i \times DBH_i^2 \times \frac{\pi(m^2)}{40\,000}$; where *BAF* is basal area factor tree factor for the ith tree of the angle gauge used during sampling.
Volume (VOL_t or VOL_m; $m^3 \cdot ha^{-1}$)	$\sum_{i=1}^{n} VOL \times TF$
Top height (*TopHt*; m)	Average height of 100 stems per hectare with the largest DBH.
Merchantable quadratic mean diameter (QMD_m; cm)	$\sqrt{\frac{\sum_{i=1}^{n} DBH_i^2}{n}}$; where *n* is the number of stems with $DBH \geq 9.1$ cm.

* Volume of individual tree calculated using regional taper equations (Honer et al. 1983; Li et al. 2012). Assumes a 15 cm stump height for total and merchantable volume and

2.3.3 LiDAR-derived forest inventory predictions

Using the area-based methods presented by Woods et al. (2011) and White et al. (2013), the LiDAR contractor produced LiDAR point cloud statistics (i.e., mean height, standard deviation of heights, height deciles, etc.) for each calibration plot. The contractor then used the R package (R Development Core Team 2012) to run Random Forests™ in regression mode (Breiman 2001) to build ensemble decision trees for the prediction of forest inventory variables (i.e., BA_t, BA_m, VOL_t, VOL_m, TopHt, and QMD_m). This method

was also used by the contractor to predict basal area within each 2-cm DBH class (≥ 9.1 cm). The resulting merchantable basal area distribution (BA_m distribution) was then scaled so that the sum of the distribution was equal to the LiDAR-derived BA_m. These decision trees were used to predict wall-to-wall estimates of LiDAR-derived inventory at 20 × 20 m grid-cell resolution for the Black Brook District.

For each Random Forests™ ensemble decision tree, the default parameters of 500 trees, 1/3 of the predictor variables used at each split, and a random 2/3 of the observations used to build any one tree were used. Over-fitting was not taken into account as Random Forests™ is robust against over-fitting (Breiman 2001). Since calibration plots were sampled half-way through the growing season after LiDAR was acquired, a slight over-prediction of the forest inventory was expected, but considered negligible. Upon completion of all inventory predictions by the contractor, the finalized LiDAR-derived enhanced forest inventory was delivered to J.D. Irving, Limited, which was then provided for our use in this study. LiDAR point cloud statistics for the calibration plots and 20 × 20 m grid were not made available by the contractor. This area-based stand-level forest inventory derived from point cloud statistics will be referred to herein as "LiDAR-derived".

To prepare the LiDAR data for plot matching, LiDAR grid cells that fell within spruce plantations were identified by intersecting cells with the J.D. Irving, Limited forest resource inventory polygons using ArcGIS (ESRI, Redlands, CA) and the planted species was determined from the inventory records. Cells were excluded where their centroid intersected roads or non-spruce-plantation polygons. Cells that fell on the boundary of differing species of spruce plantations were identified and removed to ensure that plot matching would be performed only on LiDAR grid cells that were completely within

spruce plantations of the same species. This resulted in over 1.7 million LiDAR grid cells in spruce plantations: 766 830 black spruce, 697 250 white spruce, 222 720 Norway spruce, and 32 430 red spruce.

2.3.4 Ancillary forest inventory plot data

Plots used as the library for the plot-matching approach included permanent, fixed-area plots installed in spruce plantations by the NB ERD (Porter et al. 2001) and J.D. Irving, Limited (including MacLean et al. [2010] and Omari and MacLean [2015]). All trees in the permanent sample plot database were measured for DBH and height in each measurement. A total of over 65 000 variable-radius plantation temporary sample plots from the NB ERD Forest Development Survey database (MacDonald 2008) were also used. These plots used a 3 $m^2 \cdot ha^{-1}$ basal area factor BAF prism to sample trees for species and DBH and a 10 $m^2 \cdot ha^{-1}$ BAF prism to sample trees for height. A 12.6 m^2 fixed area plot was located at plot centre to sample small trees (< 9.1 cm). The same suite of stand-level inventory variables calculated for the calibration/validation plots were generated from the tree data from each of these ancillary plots using the formulae provided in Table 2.2. Although ancillary plots do not share identical plot designs, variable-radius and fixed area sample plots provide forest inventory results that are not significantly different (Piqué et al. 2011). Grenier et al. (1991) and Lecomte et al. (1994) also showed that both inventory methods do provide significantly similar results with respect to stand-level diameter distribution.

Since the LiDAR-derived forest inventory did not include detailed species information, it was necessary that the plot-matching library contain plots with species

compositions similar to what was observed in the Black Brook District. To accomplish this, the 95th percentile of species compositions by basal area for spruce (≥ 60%), balsam fir (< 40%), hardwood (< 17%), and non-commercial species (< 5%) were calculated from 399 Black Brook spruce plantation plots with almost 1400 total measurements. All measurements with species composition that were outside of this range were omitted from the plot-matching library, including those within Black Brook.

In total, 4447 plot measurements contained species compositions that fell within the 95th percentile of observations from the Black Brook District. This included 2983 black spruce, 1184 white spruce, 63 Norway spruce, and 218 red spruce dominated plantation plot measurements. Given the low sample size of Norway spruce measurements and the presence of over 220 000 Norway spruce LiDAR grid cells, additional data for Norway spruce were needed. Since white spruce and Norway spruce are both planted on high quality sites in Black Brook and contain similar ranges of basal area, quadratic mean diameter, and diameter distribution, data from the 1184 white spruce plots were combined with the Norway spruce plot measurements to accommodate the small number of Norway spruce plots. Inventory variables for white spruce observations were recalculated using Norway spruce height-diameter relationships and taper equations. The assumption that white and Norway spruce plot measurements can be used interchangeably was tested by comparing BA_m distributions between Norway spruce plot matches and observed plot conditions, as discussed below.

2.3.5 Plot-matching approach

The matching of LiDAR grid-cell variables to those in the ancillary plot library used a nearest-neighbor matching algorithm built within the FORUS program (FORUS Research 2011). LiDAR grid-cell variables were matched with ancillary plot variables based on the smallest sum of squared difference between each inventory variable used, referred to as 'distance' (Equation 2.1). To remove the influence of different variable ranges (e.g., BA_t vs VOL_t) on matching, values were normalized by dividing by the maximum of observed plot values for each variable. Plot-cell match distance was calculated using:

[2.1] $$Distance = \sum_{v}^{V} w_v \left(\frac{c_v}{Max(p_{i...n})_v} - \frac{p_v}{Max(p_{i...n})_v} \right)^2$$

where *w* is variable weight, *c* represents a LiDAR cell value, and *p* represents ancillary plot value (*i*...*n*), for each variable *v*: planted species, BA_t, BA_m, VOL_t, VOL_m, TopHt, and QMD_m. As this matching algorithm did not permit use of categorical variables, planted species was converted to a numeric value (e.g., black spruce = 1, white spruce = 2, etc.). Because it was important that LiDAR grid cells and ancillary plot matches be for the same species, an arbitrarily high variable weight of 1000 was used for planted species. All other variables were given an equal weight of 1. LiDAR grid cells were matched with only one ancillary plot (k = 1) to ensure that the total number of unique tree lists was limited to the amount of ancillary plot data. Variables extracted from matched plots are referred to herein as "plot-matched variables".

2.3.6 Plot-level comparison of plot-matching results

Geo-registration of the 100 calibration/validation plots enabled comparison of the plot-matched variable values with 1) actual field-measured plot data, and 2) the LiDAR-derived predictions for these plots. The six stand-level inventory variables used in plot matching (BA_t, BA_m, VOL_t, VOL_m, TopHt, and QMD_m) were considered in this comparison. Since the calibration/validation plots were circular 400-m^2 plots, they did not fall completely within a single square 400-m^2 LiDAR grid cell. Plot-matched and LiDAR-derived variable values per calibration/validation plot were calculated by summing the product of proportion of the plot area intersecting a grid cell with the corresponding plot-matched and LiDAR-derived variable values. Plot-matched and LiDAR-derived variable values were plotted against observed calibration/validation plot inventory values for BA_t, BA_m, VOL_t, VOL_m, TopHt, and QMD_m and the Pearson correlation coefficient (r) and root mean squared error (RMSE) calculated for each case. To compare across variables, RMSE was calculated as a percentage by dividing it by the mean of each respective variable. This comparison was done using 98 of the LiDAR calibration/validation plots, as two of the plots fell outside of planted stand polygon boundaries.

To validate tree-list imputations, each imputed BA_m distribution was compared to the field-measured BA_m distribution for the 98 calibration/validation plots using a two-sample Kolmogorov-Smirnov (KS) test (Massey 1951) with a significance level of α = 0.05. LiDAR-derived BA_m distribution estimates were similarly compared with the field-measured BA_m distributions to assess the quality of these predictions relative to those obtained by plot matching. The KS test is often used when comparing forestry-related diameter distributions, as it does not assume normality (Little 1983; Westphal et al. 2006).

Since the two-sample KS test requires that samples compared both contain at least one observation within the distribution and LiDAR-derived distributions only include merchantable stems, 19 of the 98 calibration plots were omitted from the KS test as they only contained tree measurements with DBH from 3–9 cm. An additional KS test compared the basal area distribution across all diameter classes for the calibration/validation plots and the associated plot matches. Since the LiDAR-derived basal area distribution did not contain predictions for DBH < 9.1 cm, LiDAR-derived basal area distribution was not included in this test. To illustrate examples of good, average, and poor BA_m distributions derived from plot matching, calibration/validation plots from each planted species and the corresponding plot matching-derived BA_m distribution were graphed.

2.3.7 Plantation- or stand-level comparison of plot matching results

To fully validate tree-list imputations at the plantation or stand level would require a complete tree-level census of a number of plantations, something beyond the financial scope of this study. As an alternative, we calculated the difference between each plot-matched and LiDAR-derived inventory variable (BA_t, BA_m, VOL_t, VOL_m, TopHt, and QMD_m) for each grid cell and then averaged these differences for each plantation polygon. These averages were plotted against plantation/polygon size and tested for a correlation. Since grid cells were excluded from plot matching if they did not fall entirely within a spruce-plantation polygon, plantation area was calculated as the number of grid cells within a plantation polygon used in plot matching multiplied by grid-cell area (400 m^2). Mean bias for each inventory variable used in plot matching was calculated for all matches to determine the landscape-level bias that plot matching may introduce.

In the absence of having a complete tree-level census of a number of plantations, plot- matched BA_m distributions were aggregated by species and compared with BA_m distribution from the calibration/validation plots using a KS test with a significance level of $\alpha = 0.05$. The aggregated BA_m distributions were graphed to provide a visual comparison of the resulting BA_m distribution, using all 98 LiDAR calibration plots. Mean difference and range of difference were calculated for each BA diameter class per species. To compare species composition between the aggregated calibration/validation plots and plot matches, percent planted spruce species, balsam fir (*Abies balsamea* (L.) Mill.), and softwood, by basal area, were calculated. This comparison was conducted because harvest prescriptions in the Black Brook District, and many other Canadian forest lands, are currently scheduled based on aggregates of LiDAR grid-cell predictions rather than individual grid-cell estimates.

2.4 Results

2.4.1 Comparison of plot-matched and field-measured variables

Fifty-one percent of the 5630 ancillary plot measurements in spruce plantations throughout New Brunswick were matched with LiDAR grid cells in the Black Brook District while the remaining 49% were not used because they contained combinations of inventory values not represented in the LiDAR-derived estimates. Over one-half of the grid cells in Black Brook District spruce plantations were best matched with ancillary plot measurements located outside of the District (Fig. 2.1a). Although 1400 plot measurements were available within the Black Brook District, only 62% of these were used in plot matching, versus 43% of plot measurements originating outside the District being used.

Forest inventory variable values generated through plot matching were highly correlated with inventory variable values calculated from the field measurements in the geo-referenced calibration/validation plots, suggesting that good matches were found for the LiDAR grid cells in the ancillary plot library. Pearson correlation coefficients (*r*) ranged from 0.97–0.99 for BA_t, BA_m, VOL_t, VOL_m, and TopHt (Fig. 2.2a–e), and was 0.91 for QMD_m (Fig. 2.2f). Correlation coefficients of inventory variables derived by plot matching versus calibration/validation plot variables were nearly identical to those between LiDAR-derived variables and the calibration/validation plot variables (Fig. 2.2a–f). The weaker correlation of QMD_m resulted from the mathematical constraints of this variable; calibration/validation plots that had no merchantable stems (DBH ≥ 9.1 cm) had a QMD_m of 0 cm, whereas just one stem with DBH ≥ 9.1 cm resulted in a QMD_m of 9.1 cm or greater, and therefore LiDAR estimates of QMD_m < 9.1 cm were zeroed (Leading Edge Geomatics 2015). This resulted in a weaker correlation between LiDAR-derived QMD_m and field-measured QMD_m near the merchantable DBH limit (Fig. 2.2f).

Comparisons of the RMSEs of plot-matched predictions and LiDAR-derived predictions revealed that the plot-matching method had slightly smaller errors for BA_t, BA_m, and QMD_m, and slightly larger errors for TopHt, VOL_t, and VOL_m (Fig. 2.2). Mean percent RMSE across the six variables was similar at 14.8% for the plot-matching method and 14.1% for LiDAR-derivation. For both LiDAR-derivation and the plot-matching method, RMSE was highest (22.4% and 22.6%) for QMD_m and lowest (6.0% and 7.3%) for TopHt (Fig. 2.2).

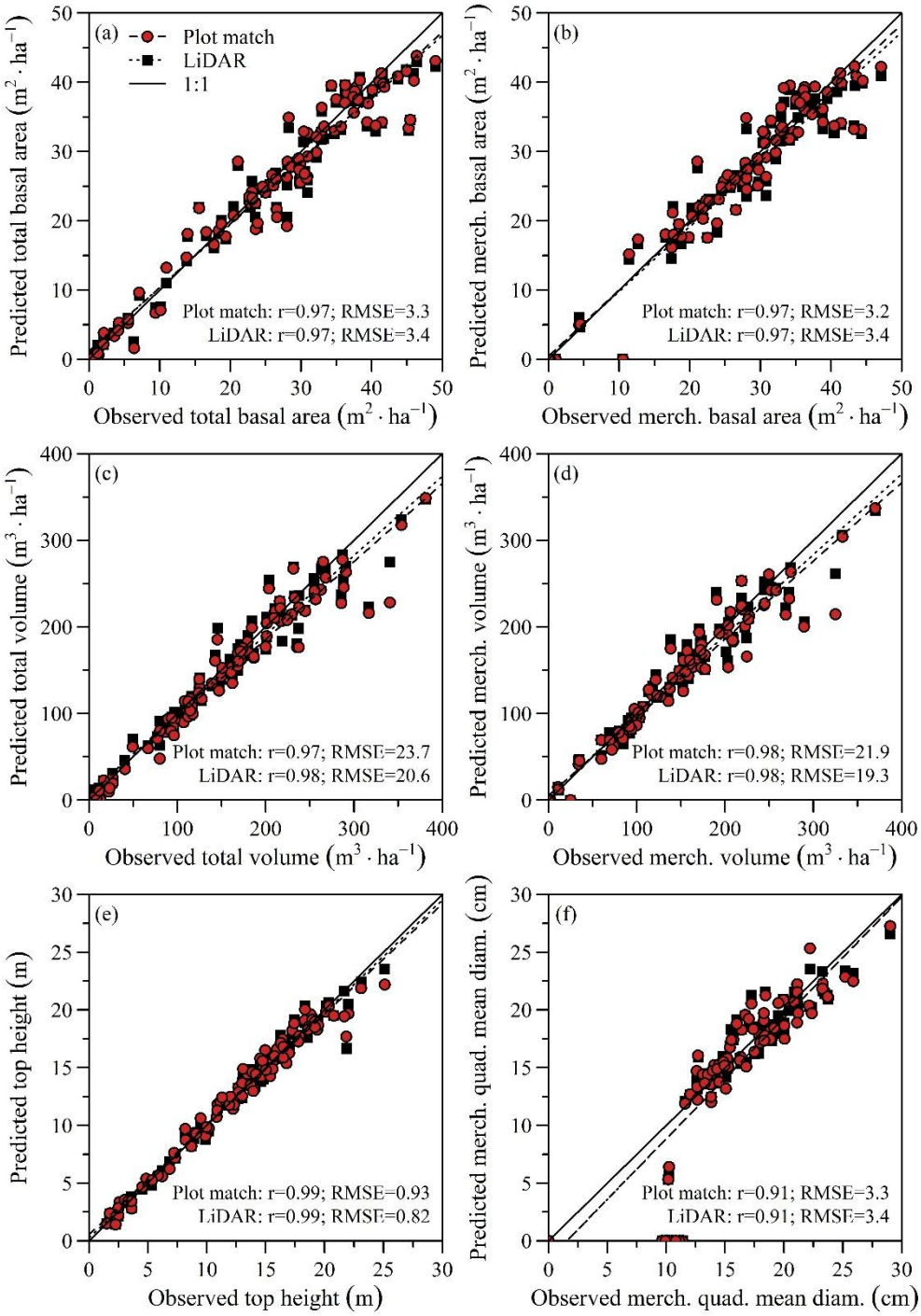

Fig. 2.2. Comparison of six inventory variables [(a)–(f): basal area, merchantable basal area, volume, merchantable volume, top height, and merchantable quadratic mean diameter] predicted by matching LiDAR grid-cell variables to ancillary plot data (red circles) or derived directly from LiDAR (closed squares) versus variables observed in 98 calibration/validation plots. The plot-matching method matched LiDAR grid-cell variables with individual plantation plot measurements throughout New Brunswick, based on planted species and the smallest sum of squared difference between six inventory variables. RMSE values presented in actual variable units.

Accurate description of inventory variables using the plot-matching method was largely dependent on the accuracy of LiDAR-derived inventory estimates. This was particularly apparent for calibration/validation plots with high basal area and volume. For the 27 plots with an observed $BA_t > 35\ m^2 \cdot ha^{-1}$, 20 had LiDAR-derived estimates that were 0.1–12.2 $m^2 \cdot ha^{-1}$ less than the observed BA_t (Fig. 2.2a). Similar results were observed with respect to plot-matched estimates, as 19 of those same 27 plots had plot-matched estimates of BA_t that were 0.1–11.8 $m^2 \cdot ha^{-1}$ less than the observed BA_t. This also occurred for volume, as 11 of 14 plots with > 250 $m^3 \cdot ha^{-1}$ VOL_t had LiDAR-derived VOL_t estimates that were 2.1–93.7 $m^3 \cdot ha^{-1}$ lower than observed (Fig. 2.2c), and 13 of those 14 plots had plot-matched VOL_t estimates that were 19.8–100.7 $m^3 \cdot ha^{-1}$ less than observed.

2.4.2 Validation of plot-level merchantable basal area distribution

The KS tests comparing plot-matched BA_m distributions to observed BA_m distributions indicated that plot matching is a valid method for tree-list imputation (Table 2.3) in the relatively homogeneous plantations studied. For 86% of the 79 calibration plots available for comparison, the plot-matched BA_m distributions did not differ from the observed distributions ($\alpha = 0.05$). By species, plot-matched BA_m distributions did not differ from observed for 91% of black spruce, 80% of white spruce, and 89% of Norway spruce plantations. These results were comparable to those of the LiDAR-derived BA_m distributions, where no differences were observed in 94% of black spruce, 89% of white spruce, 100% of Norway spruce, and 92% of all calibration/validation plots. When all diameter classes (≥ 3 cm) were considered, plot matching provided statistically equivalent BA_t distributions for 88 of the 98 calibration/validation plots. LiDAR-derived estimates

were not applicable for this comparison since LiDAR predictions were limited to DBH ≥ 9.1 cm.

Table 2.3. Number of calibration/validation plots where plot-matched and LiDAR-derived merchantable basal area distributions did not differ from observed distributions based on a two-sample Kolmogorov-Smirnov test (Massey 1951) (α = 0.05). Not testable refers to the number of calibration plots which could not be compared with a) plot matching or b) LiDAR predictions due to zero predicted merchantable basal area.

	Planted species	Significance (α = 0.05)		Not testable	Total
		No difference	Difference		
Plot match	Black spruce	32	3	5	**40**
	White spruce	28	7	9	**44**
	Norway spruce	8	1	5	**14**
	Total	**68**	**11**	**19**	**98**
LiDAR	Black spruce	33	2	5	**40**
	White spruce	31	4	9	**44**
	Norway spruce	9	0	5	**14**
	Total	**73**	**6**	**19**	**98**

A key factor contributing to accurate BA_m distribution from plot matching was the accuracy of LiDAR-derived QMD_m. In cases where plot matching provided good estimates of BA_m distribution (Fig. 2.3a–c), LiDAR-derived QMD_m did not differ by more than 0.5 cm. In contrast, where LiDAR-derived QMD_m differed from calibration/validation plots by 0.6–1.5 cm and > 1.5 cm, the resulting plot matching-derived BA_m distribution were average (Fig. 2.3d–f) and poor (Fig. 2.3g–i), respectively. Other inventory variables did not have as strong an effect on accuracy of plot matching BA_m distribution since errors for those variables fluctuated regardless of the accuracy of plot-matching derived BA_m distribution. For example, LiDAR-derived BA_m differed from the calibration plots by 2.5%–15.1% in good matches and -2.3%–16.9% in the poor matches.

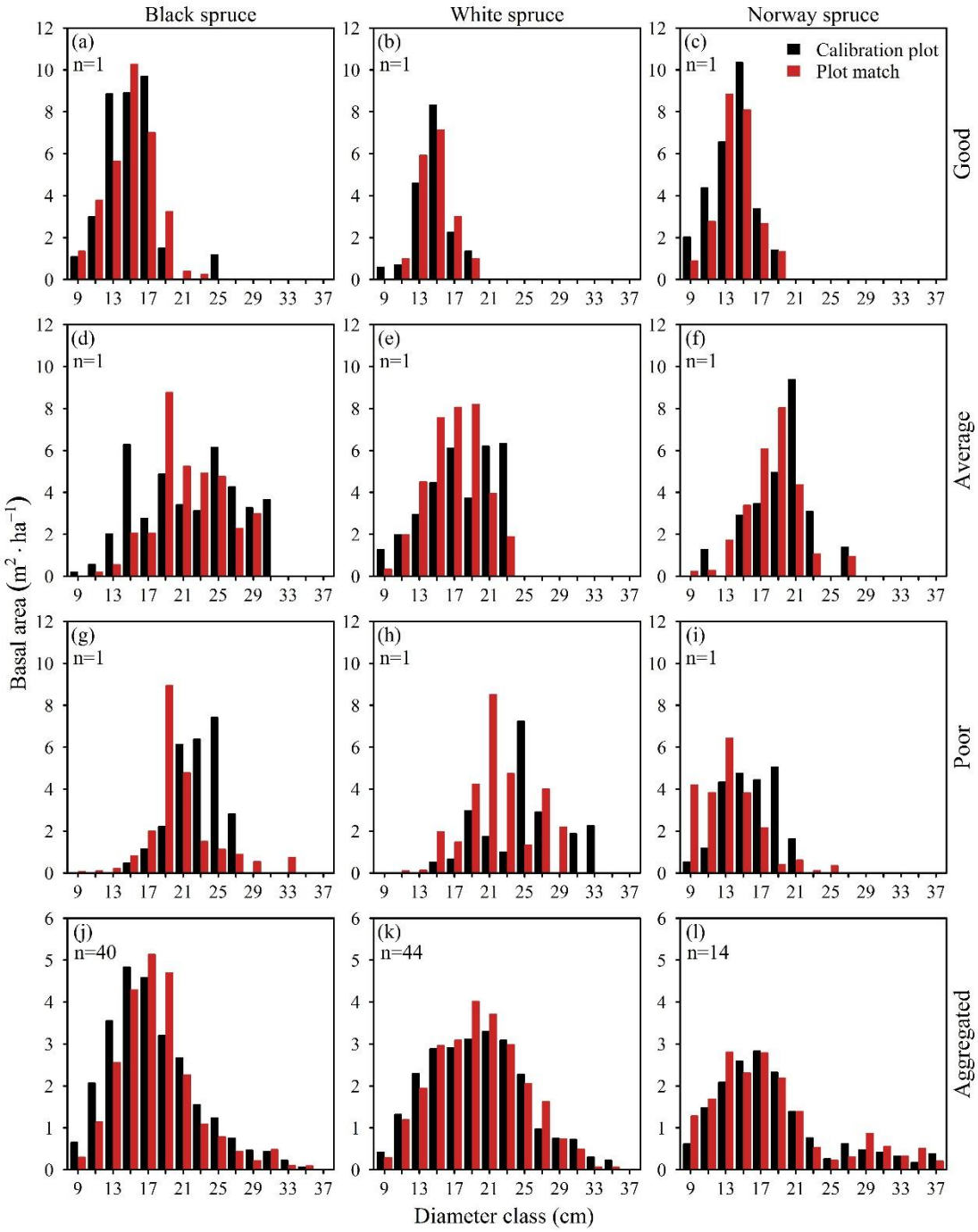

Fig. 2.3. Merchantable basal area distributions for calibration/validation plots (black) versus those estimated using the plot-matching approach (red). A sample calibration/validation plot from each of the three planted species (black, white, and Norway spruce) is presented to illustrate best (a)–(c), average (d)–(f), and worst (g)–(i) plot match-derived merchantable diameter distributions. Aggregated calibration/validation plots and the corresponding plot-matches for black spruce (j), white spruce (k), and Norway spruce (l) plantations (*n* = 40, 44, and 14 plots, respectively). The plot-matching method matched LiDAR grid-cell variables with individual plantation plot measurements, based on planted species and the smallest sum of squared difference between the six inventory variables.

The inclusion of white spruce plot measurements did not negatively affect estimates of BA_m distribution in Norway spruce calibration/validation plots. Of the 53 grid cells that fell in Norway spruce calibration/validation plots, 44 were matched with plot measurements that were originally measured in white spruce plantations. In fact, eight of the nine Norway spruce calibration plots that had no statistical difference in plot-matched BA_m distribution resulted from LiDAR grid cells being matched with white spruce plantation plot measurements.

2.4.3 Accuracy of plantation- or stand-level average matched estimates to LiDAR-derived estimates

As shown above, estimates of forest inventory variables derived from LiDAR and those derived from matching to a large sample of ancillary plot measurements resulted in plot- or cell-level estimates that are accurate enough for operational through strategic decision making. It follows that estimates for larger areas derived from the aggregation of individual cell estimates might also be expected to be accurate.

Plantation-level differences between plot-matched and LiDAR-derived inventory variables were dependent on plantation size (Fig. 2.4). Increasing plantation size resulted in progressively smaller differences between plot-matched and LiDAR-derived predictions. In total, 30% of plantations <1 ha, 42% of plantations 1–5 ha, 44% of plantations 5–20 ha, and 51% of plantations > 20 ha had zero difference (defined here as ±0.5 $m^2 \cdot ha^{-1}$ for BA_t and BA_m, ±5 $m^3 \cdot ha^{-1}$ for VOL_t and VOL_m,, ±0.5 m for TopHt, and ±0.5 cm for QMD_m) between plot-matched and LiDAR-derived predictions for all the inventory variables studied, accounting for 46% of all Black Brook District spruce

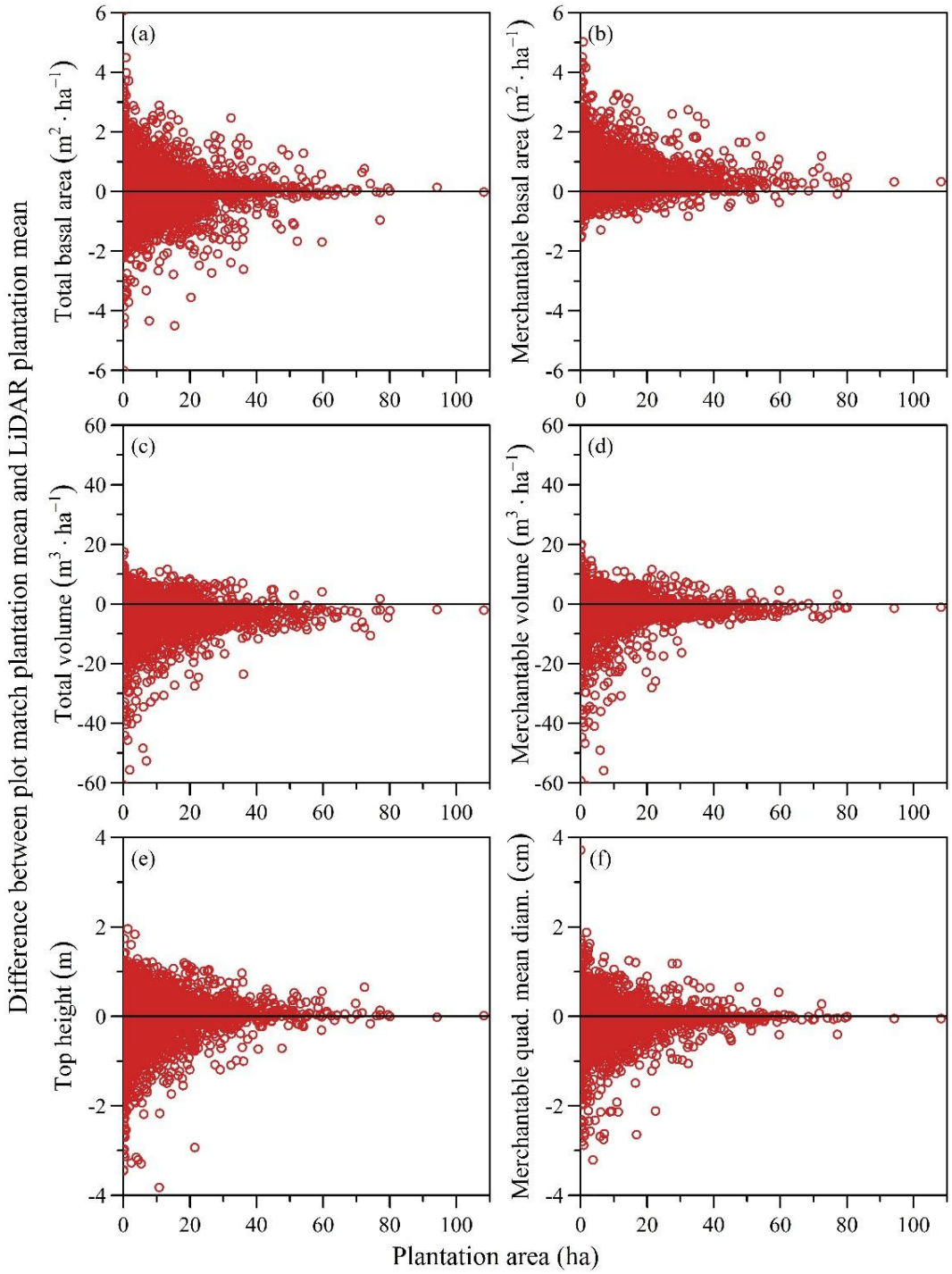

Fig. 2.4. Distribution of the differences between mean predictions from plot matching and LiDAR (*n* = 10 353) for six inventory variables, as a function of plantation area. The plot-matching method matched LiDAR grid-cell variables with individual, ancillary plantation plot measurements, based on planted species and the smallest sum of squared difference between the six inventory variables.

plantations by area. When considering plantations with all but one inventory variable having zero difference between plot-matched and LiDAR-derived inventory estimates, 24% of Black Brook spruce plantations fall within this level. In total, 70% of existing spruce plantations by area could be considered to have little difference between plot matching- and LiDAR-derived inventory estimates. Finally, when considering mean bias for all plot matches across the District (plot matching minus LiDAR), BA_t, BA_m, VOL_t, VOL_m, TopHt, and QMD_m differed by only -0.02 $m^2 \cdot ha^{-1}$, 0.34 $m^2 \cdot ha^{-1}$, -3.87 $m^3 \cdot ha^{-1}$, -1.44 $m^3 \cdot ha^{-1}$, -0.05 m, and -0.04 cm, respectively.

2.4.4 Plantation- or stand-level comparison of diameter distribution by species derived from plot matching

Plot-matched BA_m diameter distributions, by species, were very similar to those in the calibration/validation plots (Fig. 2.3j–l). KS tests showed no statistical differences between plot-matched and observed BA_m diameter distributions (α = 0.05). The mean difference (plot matched minus observed) in BA per diameter class was 0.1, -0.1, and 0.0 $m^2 \cdot ha^{-1}$ for black spruce, white spruce, and Norway spruce plantations, and ranges per species were -1.5–1.0, -0.7–0.3, and -0.9–0.3 $m^2 \cdot ha^{-1}$, respectively (Fig. 2.3j–l).

Comparison of species composition between aggregated calibration/validation plot data and plot-matched estimates showed some differences in percent planted spruce species, and balsam fir, but almost no difference in total percent softwood (Table 2.4). Plot matching resulted in slightly higher estimated percent white spruce and Norway spruce (5% and 6%, respectively), but the difference for black spruce was higher at 19%. The higher estimate of planted spruce from plot matches resulted in a lower estimate for percent

balsam fir, compared to observed data, as balsam fir content was under predicted by 8% to 18% (Table 2.4). Despite the differences in planted spruce species and balsam fir content, total softwood content did not differ by more than 1% between plot-matched estimates and actual data for calibration/validation plots.

Table 2.4. Percent composition of planted spruce, balsam fir, and softwood by basal area observed in 98 calibration/validation plots versus the proportions predicted by plot matching, for black spruce, white spruce, and Norway spruce plantations (aggregated data for *n* = 40, 44, and 14 plots, respectively).

	Planted species	**Percent composition**		
		Planted spruce	Balsam fir	Softwood
Calibration plots	Black spruce	70	25	99
	White spruce	79	16	100
	Norway spruce	77	20	100
Plot match	Black spruce	89	7	99
	White spruce	84	8	99
	Norway spruce	83	9	100

2.5 Discussion

LiDAR-derived forest inventories are an important tool for operational forest planning (Næsset 2007; Woods et al. 2011). However, the high cost of LiDAR acquisition and the need to forecast forest inventory for forest management planning would benefit from a method to estimate tree lists for area-based LiDAR grid cells, for use in tree-list growth models. The results of this investigation suggest that a nearest-neighbor approach that matches LiDAR-derived stand variables with ancillary plot data, by species, may be an effective means to impute tree lists.

Ancillary plot data from throughout the province of New Brunswick allowed tree lists to be accurately imputed for spruce plantation LiDAR grid cells from the 210 000 ha Black Brook District, as validated by ninety-eight 400-m^2 field-measured plots. The plot-

matching approach resulted in strong correlation and low RMSE for BA_t and BA_m, VOL_t and VOL_m, TopHt, and QMD_m, when compared with observed plot values. These results were similar to the correlation and RMSE between LiDAR-derived forest inventory predictions and the same observed data, suggesting that very good matches were found for the LiDAR grid cells in spruce plantations and the New Brunswick ancillary plot data used as a matching library. Since we restricted the plot-matching nearest neighbors to one, the number of unique tree lists generated was limited by a) the number of ancillary plots used, and b) the number of ancillary plots matched to LiDAR grid cells. In our study, 2870 ancillary plots were matched with LiDAR grid cells. Tree lists for these matched plots can be input into a tree-list growth model and forecast under the varying site conditions observed in the District. This will ultimately create a library of species-specific forecast yields that can be linked to each grid cell and result in only 2870 (plots) instead of 1.7 million unique individual LiDAR cells, reducing data storage and complexity.

As we hypothesized, using multiple inventory variables in plot matching resulted in BA_m diameter distributions that were similar to the observed BA_m distributions in plantation conditions. When plot- and cell-data were aggregated and compared, tests indicated that plantation- or stand-level BA_m distributions may be acceptably accurate for black spruce, white spruce, and Norway spruce. The accuracy of estimated plantation-level BA_m increased with increasing plantation size, and the difference between plot matching- and LiDAR-derived estimates diminished to what might be considered operationally insignificant values beyond about 20 ha. LiDAR-derived BA_m distributions were slightly better than those derived from plot matching, but this was likely due to the homogeneity of spruce plantations and the fact that the LiDAR-derived BA_m distribution and observed

distribution were based on the same plots used to build the Random Forests™ ensemble decision trees. The LiDAR-derived estimates are species-independent so are of limited use as input to tree-list growth models.

Restricting plot matches based on planted species, derived from dominant spruce species in plot measurements and from forest resource inventory polygons for LiDAR data, resulted in accurate estimates of species composition when LiDAR grid-cell predictions were aggregated. Percent planted spruce species by basal area derived from plot matching was overestimated 5%–6% when compared to calibration/validation plot data for white and Norway spruce plantations, and 19% higher in black spruce plantations. The overestimate of planted spruce resulted in underestimation of balsam fir by 8%–18% basal area. Plot-matched estimates of species composition were all sufficiently accurate (≥ 70% planted spruce) to be considered black spruce, white spruce, and Norway spruce dominated forest types (NB DNR 2012).

Given that harvest blocks are mapped by aggregating LiDAR grid cells and that aggregating plot matches provided sufficiently accurate species composition and reduced difference between LiDAR and plot matching inventory variables, there is arguably clear value for this method for operational and tactical planning. Matched plots provide tree-level inventory that can be used to: 1) calculate species-product volumes, 2) update future inventory using growth models, 3) calculate optimum rotation age at the stand level, and 4) substantially reduce the number of costly LiDAR inventory predictions as new stand metrics can be readily calculated from the imputed tree list.

As LiDAR point cloud statistics for the calibration plots and 20 × 20 m grid were not available for use, a *k*-nearest neighbor method for imputing tree-level inventory based

on LiDAR point cloud statistics was not possible for this study. This does presents a constraint in this study as it would have allowed for a direct comparison between the plot-matching method and the more commonly used *k*-nearest neighbor tree-list imputation methods (Packalén and Maltamo 2008; Peuhkurinen et al. 2008). Similar to the methods of Penner et al. (2015), the LiDAR contractor had originally attempted to predict BA_m distribution using yaImpute (Crookston and Finley 2008) in Random Forests™ classification mode to identify the nearest neighbor (k = 1) for each grid cell and impute species-independent BA_m distribution, however, this method ultimately resulted in weakly correlated diameter distribution. This suggests that the 100 calibration plots would not be effective in imputing species-specific tree-level inventory and that plot matching may produce better BA_m distributions in diverse stand conditions than can be adequately captured with a typical, 100 calibration plot sample size.

Previous studies that aimed to impute tree-level inventory have had varying success. Tree-list imputation based solely on photo-interpreted variables (i.e., percent crown closure and composition by species, height and age class, etc.) resulted in inaccurate estimates of diameter distributions and produced stand-level RMSE for basal area and volume of 17–28 $m^2 \cdot ha^{-1}$ and 93–280 $m^3 \cdot ha^{-1}$, respectively (Temesgen et al. 2003). Falkowski et al. (2010) imputed tree-level inventory based on LiDAR point cloud statistics and although this resulted in accurate estimates of basal area and volume, density of trees was only weakly to moderately correlated. This imputation method had species composition accuracies of 64%–69%, however the Kappa values were quite low (≤ 31%). Incorporation of spectral and textural features from aerial photo imagery with LiDAR point cloud statistics from 463 calibration plots for imputation of tree-level inventory resulted in

stand-level volume RMSE for Scots pine (*Pinus sylvestris* L.), Norway spruce, and hardwood of 28%, 33%, and 62%, respectively (Packalén and Maltamo 2008). In a similar study, Peuhkurinen et al. (2008) used 474 calibration plots to estimate species-specific diameter distributions by incorporating aerial photography with LiDAR point cloud statistics resulting in stand volume RMSE of 20% and 70% for dominant and minority species, respectively. Crown segmentation of LiDAR returns has improved tree-list imputation, with LiDAR point cloud and alpha shape metrics for individual trees classifying species as Scots pine, Norway spruce, or hardwood with 78% accuracy, and estimating DBH and height with a RMSE of 13% and 3%, respectively (Vauhkonen et al. 2010).

Although results for the plot-matching method are promising, estimates of tree-level inventory from plot matches are limited by the accuracy of the LiDAR-derived forest inventory. LiDAR-derived estimates of total volume were under predicted by as much as 94 $m^3 \cdot ha^{-1}$, resulting in plot matches that also under-predicted by up to 113 $m^3 \cdot ha^{-1}$. Since LiDAR-derived inventory variables are predicted independently, combinations of variables may be predicted that are outside the range of ancillary plot measurement data. Had the original LiDAR-derived inventory variables been predicted using non-parametric nearest neighbor methods, such as *k*-nearest neighbor, the relationship between variables in the calibration plots would have been retained and potentially resulted in more accurate matches. Because the plot-matching method used photo-interpreted plantation maps and known spruce plantation species composition from permanent sample plots to select plots suitable for plot matching, the plot selection process may not be applicable in some other areas. However, photo-interpreted species composition could be used to effectively

estimate species composition for each LiDAR grid cell and be added as additional variables during plot matching. As with the accuracy of LiDAR-derived inventory, the accuracy of photo-interpretation must be considered when assessing the plot-matching method. Classification accuracy of plot-level species composition using multi-spectral, high spatial resolution airborne images can be high in New Brunswick softwood dominated stands (74%–79%), but is less reliable in mixedwood and hardwood stand types (4%–63% accuracy) (Franklin et al. 2000). However, species composition may be better estimated with the incorporation of LiDAR data (Ørka et al. 2013). Current LiDAR-based species classification research underway by Perron and St-Onge (AWARE 2016), may improve plot matching by better classifying individual tree species and separating balsam fir within spruce plantations.

Although our plot-matching method was able to accurately impute tree lists in homogenous planted stands, modifications may be required to impute tree lists in more complex stand types. If available, additional LiDAR-derived inventory variables could be included. For example, incorporating LiDAR-derived percent hardwood or predicting variables that describe the diameter distribution (e.g., maximum DBH, standard deviation, Gini index, etc.) may improve results as we found that QMD_m was a key contributing factor in the accuracy of plot matching-derived BA_m distribution. To improve matches with respect to species composition, detailed species composition data could be included in matches rather than broad species types or known planted species. For example, NB ERD identifies percent species composition during photo interpretation for all species estimated to make up ≥ 10% of stand basal area. Incorporation of this level of species composition detail may improve species-level estimates from plot matches, but the accuracy of photo

interpretation must also be considered. Finally, the use of variable weights could be used to account for differences in LiDAR-derived inventory variable accuracy and possibly adjusted over multiple plot matches to optimize correlation between observed and estimated diameter distributions. Weighting variables by the inverse of LiDAR error was explored, but did not improve results; however, all variables chosen for matching had low error compared to other excluded variables like stem density, which had % RMSE = 48%. This method of matching existing forest inventory plots with LiDAR grid cells and incorporating photo-interpreted species composition was recently used outside of the study area by J.D. Irving, Limited for 825 000 ha of managed forest lands and has also been used in building the most recent forest management plan for the 9000 ha Acadia Research Forest in New Brunswick (Swift et al. 2006).

2.6 Conclusions

The plot-matching method tested herein effectively combined existing LiDAR-derived inventory variables and photo-interpreted cover-type information to match with plot data to accurately impute a species-specific tree list. This method required relatively few calibration plots in comparison to other LiDAR-based tree-list imputation methods and utilized existing plot measurements, which cover a wide range of diverse conditions. Individual product yields may be calculated and updated readily from the tree list and permit targeting particular stands given changing product markets. Data storage requirements and complexity are relatively low as the number of unique tree lists is limited by the number of plots that were matched. When the imputed tree list is used in a tree-list-based growth model, this method may provide a means to forecast stand development and

‘grow’ the LiDAR-based inventory forward for several decades. Future forest operations can be planned to target aggregated LiDAR grid cells that will provide the greatest financial return. This will prolong the time between expensive LiDAR acquisitions and reduce costs while retaining the spatial resolution that LiDAR provides. In addition, the outputs of these forecasts may help to describe changing species compositions and allow for additional resource planning such a biodiversity, carbon storage, and wildlife habitat. Although the plot-matching method used only a small number of variables for imputation, additional inventory variables derived from LiDAR or photo-interpretation may permit use of the method in more complex stand types.

2.7 Acknowledgements

This research was funded by the AWARE (Assessment of Wood Attributes using Remote sEnsing) Natural Sciences and Engineering Research Council of Canada Collaborative Research and Development grant to a team led by Dr. Nicholas Coops and by J.D. Irving, Limited. Sean Lamb was also funded by an NSERC Canada Graduate Scholarship. We thank the staff at J.D. Irving, Limited and Leading Edge Geomatics for their input and assistance with the project.

2.8 References

Aruga, K., Sessions, J., and Akay, A.E. 2005. Application of an airborne laser scanner to forest road design with accurate earthwork volumes. J. For. Res. 10(2):113–123.

AWARE. 2016. The Point Cloud, AWARE’s Electronic Newsletter. Vol. 2, No. 2. University of British Colombia. Available from <

http://aware.forestry.ubc.ca/newsletter/> [cited November 23, 2016].

Breiman, L. 2001. Random forests. Mach. Learn. 45(1): 5–32.

Crookston, N.L., and Dixon, G.E. 2005. The forest vegetation simulator: a review of its structure, content, and applications. Comput. Electron. Agric. 49(1): 60–80.

Crookston, N.L., and Finley, A.O. 2008. yaImpute: an R package for kNN imputation. J. Stat. Softw. 23(10): 1–16.

Falkowski, M.J., Hudak, A.T., Crookston, N.L., Gessler, P.E., Uebler, E.H., and Smith, A.M. 2010. Landscape-scale parameterization of a tree-level forest growth model: a k-nearest neighbor imputation approach incorporating LiDAR data. Can. J. For. Res. 40(2): 184–199.

FORUS Research. 2011. FORUS Slasher User Guide. FORUS Research, Fredericton, NB, Canada. 15p.

Franklin, S.E., Hall, R.J., Moskal, L.M., Maudie, A.J. and Lavigne, M.B. 2000. Incorporating texture into classification of forest species composition from airborne multispectral images. Int. J. Remote Sens. 21(1): 61–79.

Gobakken, T., and Næsset, E. 2004. Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. Scand. J. For. Res. 19(6): 529–542.

Gobakken T., and Næsset, E. 2008. Assessing effects of laser point density, ground sampling intensity, and field plot sample size on biophysical stand properties derived from airborne laser scanner data. Can. J. For. Res. 38(5): 1095–1109.

Grenier, Y., Blais, L., and Lavoie, É. 1991. Aire minimum d'échantillonnage ou nombre de points de prisme nécessaires pour établir la structure d'un peuplement inéquienne

[Minimum sampling area or number of prism points needed to establish the structure of an uneven-aged stand]. Can. J. For. Res. 21(11): 1632–1638.

Haara, A., Maltamo, M., and Tokola, T. 1997. The K-nearest-neighbour method for estimating basal-area diameter distribution. Scand. J. For. Res. 12(2): 200–208.

Härkönen, S., Tokola, T., Packalén, P., Korhonen, L., and Mäkelä, A. 2013. Predicting forest growth based on airborne light detection and ranging data, climate data, and a simplified process-based model. Can. J. For. Res. 43(4): 364–375.

Hayashi, R., Weiskittel, A., and Sader, S. 2014. Assessing the feasibility of low-density LiDAR for stand inventory attribute predictions in complex and managed forests of northern Maine, USA. Forests, 5(2): 63–383.

Hodgson, M.E., Jensen, J.R., Schmidt, L., Schill, S., and Davis, B. 2003. An evaluation of LIDAR-and IFSAR-derived digital elevation models in leaf-on conditions with USGS Level 1 and Level 2 DEMs. Remote Sens. Environ. 84(2): 295–308.

Holmgren, J. 2004. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. Scand. J. For. Res. 19(6): 543–553.

Honer, T.G., Ker, M.F., and Alemdag, I.S. 1983. Metric timber tables for the commercial tree species of central and eastern Canada. Can. For. Serv., Maritimes For. Res. Cent. Inf. Rep. M-X-140.

Hyde, P., Dubayah, R., Peterson, B., Blair, J.B., Hofton, M., Hunsaker, C., Knox, R., and Walker, W. 2005. Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. Remote Sens. Environ. 96(3): 427–437.

Hyyppä, J., and Inkinen, M. 1999. Detecting and estimating attributes for single trees using

laser scanner. Photogramm. J. Fin. 16(2): 27–42.

James, L.A., Watson, D.G., and Hansen, W.F. 2007. Using LiDAR data to map gullies and headwater streams under forest canopy: South Carolina, USA. Catena, 71(1): 132–144.

Leading Edge Geomatics. 2015. Data Dictionary: Black Brook planted spruce stands enhanced forest inventory description of data collection and processing. Unpublished report submitted to J.D. Irving, Limited. 40p.

Lecomte, H., Hébert, J., and Rondeux, J. 1994. Comparaison de plusieurs types d'unités d'échantillonnage dans la perspective d'un inventaire forestier régional. For. Chron. 70(3): 304–310.

Li, R., Weiskittel, A., Dick, A.R., Kershaw, J.A., and Seymour, R.S. 2012. Regional stem taper equations for eleven conifer species in the Acadian Region of North America: development and assessment. North. J. Appl. For. 29(1): 5–14.

Lim, K., Treitz, P., Baldwin, K., Morrison, I., and Green, J. 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. Can. J. Remote Sens. 29(5): 658–678.

Lindberg, E., Holmgren, J., Olofsson, K., Wallerman, J., and Olsson, H. 2013. Estimation of tree lists from airborne laser scanning using tree model clustering and k-MSN imputation. Remote Sens. 5(4): 1932–1955.

Little, S.N. 1983. Weibull diameter distributions for mixed stands of western conifers. Can. J. For. Res. 13(1): 85–88.

MacDonald, M. 2008. New Brunswick’s forest development survey (FDS) data collection manual (2003–2013 forest inventory cycle). Forest Management Branch, New

Brunswick Department of Natural Resources, Fredericton, NB. 105p.

MacLean, D.A., Amos-Binks, L., Adams G., Pelletier G., and Villard, M.-A. 2010. Legacy of the Sustainable Forest Management Network: outcomes of research collaborations among J.D. Irving, Limited, University of New Brunswick, and Université de Moncton. Sustainable Forest Management Network, Edmonton, AB, Canada. ISBN No. 978-1-55261-267-5. 52p.

Magnussen, S., and Boudewyn, P. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. Can. J. For. Res. 28(7): 1016–1031.

Maltamo, M., and Kangas, A. 1998. Methods based on k-nearest neighbor regression in the prediction of basal area diameter distribution. Can. J. For. Res. 28(8): 1107–1115.

Maltamo, M., Malinen, J., Packalén, P., Suvanto, A., and Kangas, J. 2006. Nonparametric estimation of stem volume using airborne laser scanning, aerial photography, and stand-register data. Can. J. For. Res. 36(2): 426–436.

Maltamo, M., Næsset, E., Bollandsås, O.M., Gobakken, T., and Packalen, P., 2009. Non-parametric prediction of diameter distributions using airborne laser scanner data. Scand. J. For. Res. 24(6):.541–553

Martinuzzi, S., Vierling, L.A., Gould, W.A., Falkowski, M.J., Evans, J.S., Hudak, A.T., and Vierling, K.T. 2009. Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. Remote Sens. Environ. 113(12): 2533–2546.

Massey, F.J. Jr. 1951. The Kolmogorov-Smirnov test for goodness of fit. J. Am. Stat. Assoc. 46(253): 68–78.

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. Photogramm. Eng. Remote Sens. 66(11): 1367–1372.

Means, J.E., Acker, S.A., Harding, D.J., Blair, J.B., Lefsky, M.A., Cohen, W.B., Harmon, M.E., and McKee, W.A. 1999. Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the Western Cascades of Oregon. Remote Sens. Environ. 67(3): 298–308.

Murphy, P.N., Ogilvie, J., Castonguay, M., Zhang, C.F., Meng, F.R., and Arp, P.A. 2008a. Improving forest operations planning through high-resolution flow-channel and wet-areas mapping. For. Chron. 84(4): 568–574.

Murphy, P.N., Ogilvie, J., Meng, F.R., and Arp, P. 2008b. Stream network modelling using lidar and photogrammetric digital elevation models: a comparison and field verification. Hydrol. Process. 22(12): 1747–1754.

Næsset, E. 1997. Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS J. Photogramm. Remote Sens. 52(2): 49–56.

Næsset, E. 2007. Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia. Scand. J. For. Res. 22(5): 433–442.

Næsset, E., Bollandsås, O.M., and Gobakken, T. 2005. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. Remote Sens. Environ. 94(4): 541–553.

NB DNR. 2012. New Brunswick integrated land classification system. Forest Management Branch, New Brunswick Department of Natural Resources, Fredericton, NB, Canada.

43p.

Omari, K., and MacLean, D.A. 2015. Do biomass removal and structure-enhancing treatments influence deadwood characteristics following commercial thinning in spruce plantations in New Brunswick, Canada? Can. J. For. Res. 45(10): 1407–1418.

Ørka, H.O., Dalponte, M., Gobakken, T., Næsset, E., and Ene, L.T. 2013. Characterizing forest species composition using multiple remote sensing data sources and inventory approaches. Scand. J. For. Res. 28(7): 677–688.

Packalén, P., and Maltamo, M., 2006. Predicting the plot volume by tree species using airborne laser scanning and aerial photographs. For. Sci. 52(6): 611–622.

Packalén, P., and Maltamo, M. 2007. The k-MSN method for the prediction of species-specific stand attributes using airborne laser scanning and aerial photographs. Remote Sens. Environ. 109(3): 328–341.

Packalén, P., and Maltamo, M. 2008. Estimation of species-specific diameter distributions using airborne laser scanning and aerial photographs. Can. J. For. Res. 38(7): 1750–1760.

Penner, M., Pitt, D.G., and Woods, M.E. 2013. Parametric vs. nonparametric LiDAR models for operational forest inventory in boreal Ontario. Can. J. Remote Sens. 39(5): 426–443.

Penner, M., Woods, M., and Pitt, D.G. 2015. A comparison of airborne laser scanning and image point cloud derived tree size class distribution models in boreal Ontario. Forests, 6(11): 4034–4054.

Persson, A., Holmgren, J. and Söderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. Photogramm. Eng. Remote Sens. 68(9): 925–

932.

Peuhkurinen, J., Maltamo, M., and Malinen, J. 2008. Estimating species-specific diameter distributions and saw log recoveries of boreal forests from airborne laser scanning data and aerial photographs: a distribution-based approach. Silva Fenn. 42(4): 625–641.

Piqué, M., Obon, B., Condés, S. and Saura, S. 2011. Comparison of relascope and fixed-radius plots for the estimation of forest stand variables in northeast Spain: an inventory simulation approach. Eur. J. For. Res. 130(5): 851–859.

Porter, K.B., MacLean, D.A., Beaton, K.P., and Upshall, J. 2001. New Brunswick permanent sample plot database (PSPDB version 1.0): user's guide and analysis. Can. For. Serv., Atl. For. Cent. Inf. Rep. M-X-209. 65p.

R Development Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Schiess P., and Krogstad F. 2003. Lidar-based topographic maps improve agreement between office-designed and field-verified road locations. In 26th Annual Meeting of the Council on Forest Engineering, Bar Harbor, ME, USA. 7–10 September, 2003. 6p.

Shang, C., Treitz, P., Caspersen, J., and Jones, T. 2017. Estimating stem diameter distributions in a management context for a tolerant hardwood forest using ALS height and intensity data. Can. J. Remote Sens. 43(1): 79–94.

Swift, D.E., Kilpatrick, B., Murray, T., Toole, D., Henderson, J., and Pitt, C. 2006. Acadia Research Forest: a brief introduction to a living laboratory. Long-term Silvicultural

and Ecological studies: Results for Science and Management. Yale University, School of Forestry and Environmental Studies, Global Institute of Sustainable Forestry Research Paper, 5, pp.104–118.

Temesgen, H., LeMay, V.M., Froese, K.L., and Marshall, P.L. 2003. Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia. For. Ecol. Manage. 177(1): 277–285.

Tompalski, P., Coops, N.C., White, J.C., and Wulder, M.A. 2016. Enhancing forest growth and yield predictions with airborne laser scanning data: Increasing spatial detail and optimizing yield curve selection through template matching. Forests, 7(11): 255.

Vauhkonen, J., Korpela, I., Maltamo, M., and Tokola, T. 2010. Imputation of single-tree attributes using airborne laser scanning-based height, intensity, and alpha shape metrics. Remote Sens. Environ. 114(6): 1263–1276.

Westphal, C., Tremer, N., von Oheimb, G., Hansen, J., von Gadow, K., and Härdtle, W. 2006. Is the reverse J-shaped diameter distribution universally applicable in European virgin beech forests? For. Ecol. Manage. 223(1): 75–83.

White, B., Ogilvie, J., Campbell, D.M., Hiltz, D., Gauthier, B., Chisholm, H.K.H., Wen, H.K., Murphy, P.N., and Arp, P.A. 2012. Using the cartographic depth-to-water index to locate small streams and associated wet areas across landscapes. Can. Water Resour. J. 37(4): 333–347.

White, J.C., Wulder, M.A., Varhola, A., Vastaranta, M., Coops, N.C., Cook, B.D., Pitt, D., and Woods, M. 2013. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. Can. For. Serv., Can. Wood Fibre Cent. Inf. Rep. FI-X-010. 50p.

Woods, M., Pitt, D., Penner, M., Lim, K., Nesbitt, D., Etheridge, D., and Treitz, P. 2011. Operational implementation of a LiDAR inventory in Boreal Ontario. For. Chron. 87(4): 512–528.

Zelazny, V., Martin, G.L., Toner, M., Gorman, M., Colpitts, M., Veen, H., Godin, B., McInnis, B. Steeves, C., and Roberts, M. 2003. Our landscape heritage: the story of ecological land classification in New Brunswick. New Brunswick Department of Natural Resources, Fredericton, NB, Canada 359p.

CHAPTER 3: FORECASTING FOREST INVENTORY USING IMPUTED TREE LISTS FOR LIDAR GRID CELLS AND A TREE-LIST GROWTH MODEL

Paper planned for submission as:

Lamb, S.M., MacLean, D.A., Hennigar, C.R., and Pitt, D.G. 2018. Forecasting forest inventory using imputed tree lists for LiDAR grid cells and a tree-list growth model. Planned for submission to Forests.

3.1 Abstract

A method to forecast forest inventory variables derived from light detection and ranging (LiDAR) would increase the usefulness of such data in future forest management. We evaluated the accuracy of forecasted inventory from imputed tree lists for LiDAR grid cells in spruce (*Picea* sp.) plantations and tree growth predicted using a locally calibrated tree-list growth model. Tree lists were imputed by matching measurements from a library of sample plots with grid cells based on planted species and the smallest sum of squared difference between six inventory variables. Total and merchantable basal area, and total and merchantable volume, Lorey's height, and quadratic mean diameter increments predicted using imputed tree lists were highly correlated (0.75–0.86) with those from measured tree lists in 98 validation plots. Percent root mean squared error ranged from 12.8%–49.0%, but was much lower (4.9%–13.5%) for plots with ≤ 10% LiDAR-derived error for all plot-matched variables. When compared with volumes from 15 blocks harvested 3–5 years after LiDAR acquisition, average forecasted volume differed by only 1.5%. To demonstrate the novel application of this method for operational management decisions, annual commercial thinning was planned at grid-cell resolution from 2018–2020 using forecasted inventory variables and commercial thinning eligibility rules.

3.2 Introduction

The use of light detection and ranging (LiDAR) to produce forest inventories has significantly changed forest management (Næsset 2007; Woods et al. 2011). LiDAR-derived forest inventory using parametric and non-parametric methods can provide accurate estimates of average tree diameter at breast height (DBH) (Lim et al. 2003; Woods et al. 2008), average tree height (Næsset 1997; Holmgren et al. 2003), crown base height (Andersen et al. 2005; Popescu and Zhao 2008), basal area (Means et al. 2000; Hudak et al. 2006), and volume (Nilsson 1996; Holmgren 2004). Unlike traditional inventory methods, which rely on photo-interpreted forest polygons, LiDAR-derived inventory has high spatial resolution, with typical grid-cell predictions of 10–30 m. This allows precise quantification of both within- and between-stand variation and helps to ensure that scheduled harvests will meet expected volume returns (Næsset 2007; Wulder et al. 2012). Additionally, high resolution LiDAR-derived inventory permits silvicultural or harvest treatments to be prescribed at higher spatial resolution, minimizing the need for costly pre-harvest ground verification.

Commercial thinning is one silvicultural treatment that requires spatially accurate estimates of forest inventory. Commercial thinning is a partial harvest treatment, typically prescribed in immature managed stands, with the goal of removing merchantable stems to provide economic return and improve growing conditions for residual stems (Long 1985; Nyland 2007). While thinning does not typically increase annual stand volume accrual (Mäkinen and Isomäki 2004; Pelletier and Pitt 2008), it does increase light, soil moisture, and nutrient availability for the residual crop trees (Albaugh et al. 2004; Canham et al. 2004), increasing their size and value (Pelletier and Pitt 2008; Duchesne et al. 2013).

Accurate timing of commercial thinning entries, with respect to stand development, is vital to ensure effective response from the treatment and to realize maximum economic benefits. Optimally-timed commercial thinning that is conducted once density-dependent mortality is imminent will provide the greatest number of harvested stems with the highest average tree volume (Smith et al. 1997). Commercial thinning too early will result in lower average tree volume and reduced economic return (Nyland 2007), while thinning too late will result in lower volume due to tree mortality and lower live crown ratio (Pelletier and Pitt 2008). A live crown ratio of 35%–40% has been cited for a number of conifers as the minimum level before tree vigour and response to thinning is reduced (Smith et al. 1997; Nyland 2007). LiDAR can provide the inventory data required to delineate and prioritize thinning treatments (Watt et al. 2013), or to identify LiDAR grid cells that are eligible for thinning (Vastaranta et al. 2011).

While LiDAR-derived forest inventory can provide inventory data to make sound commercial thinning decisions, any forest inventory becomes less accurate and useful as the years since data acquisition increase. To increase the longevity of LiDAR-derived inventory and reduce the costs associated with LiDAR acquisition, processing, and inventory predictions, simplified process-based and size-class growth models have been used to forecast stand-level variables at grid-cell resolution (Härkönen et al. 2013; Tompalski et al. 2016). In addition, tree-level variables have been forecasted using a tree-list growth model and tree lists imputed for LiDAR grid cells (Falkowski et al. 2010). Tree-list growth models (e.g., Crookston and Dixon 2005; FORUS Research 2011) predict growth and yield of individual trees on an annual or periodic basis using species-specific diameter, height, mortality, and live crown ration increment equations to derive forecasted

inventory (Dale et al. 1985; Weiskittel et al. 2011). Providing forest managers with forecasted tree-level inventory for each LiDAR grid cell would allow users to track changes in species-specific products and identify grid cells that meet harvest requirements.

A major constraint of using tree-list growth models to forecast LiDAR inventory is the tree-level data (i.e., species, DBH, and height) required as model inputs. Predicting tree-level inventory for LiDAR grid cells requires a substantial number of calibration plots covering the full range of inventory conditions (White et al. 2013; Wulder et al. 2013). In studies by Peuhkurinen et al. (2008) and Packalén and Maltamo (2008), close to 500 calibration plots were needed to produce species-specific diameter distribution estimates for their respective 1200- and 2000-ha study areas. In similar stand types, species-independent diameter distributions can be produced with as few as 50–100 plots on much larger landbases (Gobakken and Næsset 2004; Maltamo et al. 2009). If tree-level inventory is produced for a landbase, incorporating it into a tree-list growth model introduces additional challenges. As each LiDAR grid cell would inevitably contain its own unique combination of estimated tree-level inventory, each cell would need to be forecasted individually. This becomes computationally challenging for large landbases.

A common method used for imputing forest inventory data is *k*-nearest neighbor imputation. This method was originally used to impute complex inventory data, such as basal area distribution, for forest stands (Haara et al. 1997; Maltamo and Kangas 1998) and more recently has proved to be a powerful tool to impute LiDAR-derived stand-level variables (McInerney and Nieuwenhuis 2009; Penner et al. 2013). Recently, Lamb et al. (2017) used non-parametric *k*-nearest neighbor methods to impute tree-level inventory for LiDAR cells in spruce (*Picea* sp.) plantations in New Brunswick, Canada. Six LiDAR-

derived variables (i.e., total and merchantable basal area, total and merchantable volume, top height, and merchantable quadratic mean diameter) and known planted tree species were used to match stand structural variables estimated from LiDAR to those in a library of over 5500 sample plot measurements to impute tree lists for LiDAR grid cells across 83 000 ha of spruce plantations. Matches were determined based on planted species and minimum sum of squared differences between six inventory variables. With this method, only 100 LiDAR calibration plots were required to predict the LiDAR-derived inventory variables for the 83 000 ha of spruce plantations. LiDAR grid cells were matched with only one plantation plot measurement (k = 1), which limited the number of unique tree lists to the 2870 spruce plantation plot measurements that were matched. This provides a framework for forecasting LiDAR grid cells at the tree level without the need for a large number of calibration plots, while minimizing computational challenges associated with forecasting tree lists for cells.

The objective of this study was to assess the accuracy of forecasted inventory using tree-level inventory imputed for LIDAR grid cells in New Brunswick spruce plantations based on the methods of Lamb et al. (2017). To do so, imputed tree lists were forecasted using a regional tree-list growth model (FORUS Research 2011) calibrated using growth and mortality data from permanent sample plots in spruce plantations from the study area. We hypothesized that inventory increments predicted using observed tree lists would be strongly correlated to increments predicted from imputed tree lists and the locally calibrated growth model. This would provide forest managers with a method to extend the longevity of LiDAR-derived inventory and allow use of the forecasted data to make operational-, tactical-, and strategic-management decisions.

3.3 Materials and methods

3.3.1 Study area

The Black Brook District is a 210 000 ha forest in northwestern New Brunswick, Canada (47°9’51”N to 67°55’27”W) (Fig. 3.1a), owned by J.D. Irving, Limited. The District is within Ecoregion 3 (Zelazny et al. 2008) and contains some of the most intensively managed lands in Canada (MacLean et al. 2010). The forest is comprised primarily of uneven-aged, shade-tolerant hardwoods, managed using selection and group shelterwood harvests (65 000 ha) and spruce plantations, commercially thinned one to three times before final harvest (83 000 ha). Spruce plantations include 37 000 ha of black spruce (*Picea mariana* (Mill.) B.S.P.), 33 500 ha of white spruce (*Picea glauca* (Moench) Voss), 11 000 ha of Norway spruce (*Picea abies* (L.) Karst.), and 1500 ha of red spruce (*Picea rubens* Sarg.). Close to 400 permanent sample plots, each measured every 3–5 years since 1980, have been installed in these spruce plantations to monitor growth and mortality.

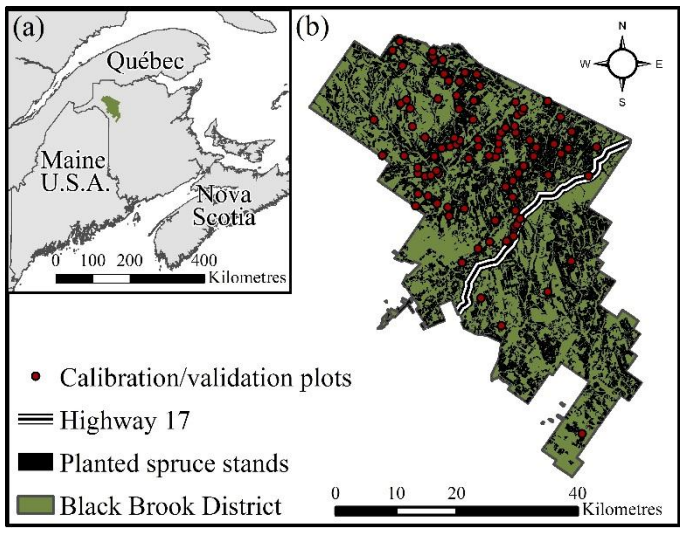

Fig. 3.1. Location of (a) Black Brook District in northwestern New Brunswick, and (b) LiDAR calibration/validation plots in spruce plantations. LiDAR was acquired in 2011 and 2013 for areas north and south of Highway 17, respectively.

3.3.2 LiDAR data and calibration/validation plot data

LiDAR data and calibration/validation plot data as described in Lamb et al. (2017) were used in this study. A private LiDAR contractor acquired high density LiDAR (6 pulses·m^{-2}, with up to 8 returns per pulse) in September 2011 for the northern portion (north of Highway 17, Fig. 3.1b) of the District and in July, 2013 for the southern portion. Scanning parameters described in Lamb et al. (2017) were identical in both acquisitions.

Ninety-five circular 400-m^2 plots were established in the summer of 2012 in the area flown in 2011 for use in calibrating LiDAR predictions (Fig. 3.1b). An additional five plots were established in 2014 in the area flown in 2013 (Fig. 3.1b). The plots were established in 41 black spruce, 45 white spruce, and 14 Norway spruce plantations to cover the range of inventory conditions observed across spruce plantations in the District (Leading Edge Geomatics 2015). Plot centers were geo-referenced using an SX Blue II global positioning system to obtain sub-meter accuracy. Species, status (live or dead), and DBH were measured for all trees with DBH ≥ 3 cm. Three randomly selected trees within each species and 5 cm DBH class were measured for height. For each plot, total and merchantable basal area (BA_t, BA_m), gross total and merchantable volume (VOL_t, VOL_m), top height (TopHt), and merchantable quadratic mean diameter (QMD_m) were calculated using the formulae described in Lamb et al. (2017), with 9.1 cm DBH as the lower merchantability limit. Volumes were calculated using taper equations from Li et al. (2012) for commercial species and from Honer et al. (1983) for non-commercial species.

3.3.3 LiDAR-derived forest inventory predictions

LiDAR point cloud statistics (i.e., mean height, standard deviation of heights, height deciles, etc.) were calculated for each calibration/validation plot. Using R (R Development Core Team 2012), the LiDAR contractor ran Random Forests™ (Breiman 2001) in regression mode to build ensemble decision trees to determine the statistical relationships between LiDAR point cloud statistics and forest inventory variables (BA_t, BA_m, VOL_t, VOL_m, TopHt, and QMD_m). Default parameters of 500 trees, one-third of the predictor variables used at each split, and a random two-thirds of the observations used to build any one tree were used for each ensemble decision tree. The resulting decision trees were used to predict estimates of forest inventory at 20 × 20 m grid-cell resolution for all Black Brook District spruce plantations. Forest inventory predictions derived from point cloud statistics will herein be referred to as "LiDAR-derived".

3.3.4 Ancillary forest inventory plot data

The library of plots used in the plot-matching approach was described in detail by Lamb et al. (2017) and only a summary will be given here. This included: 1) permanent, fixed-area plots installed in spruce plantations by the New Brunswick Department of Energy and Resource Development (NB ERD) (Porter et al. 2001) and J.D. Irving, Limited (including MacLean et al. [2010] and Omari and MacLean [2015]), with all trees measured for DBH and height in each measurement; and 2) variable radius plantation temporary sample plots from the NB ERD Forest Development Survey database (MacDonald 2008), which used a 3 $m^2 \cdot ha^{-1}$ basal area factor (BAF) prism to sample trees for species and DBH and a 10 $m^2 \cdot ha^{-1}$ BAF prism to sample trees for height. Overall, this included over 65 000

variable-radius and permanent, fixed-area plots installed in spruce plantations throughout New Brunswick. As with the LiDAR calibration/validation plots, stand-level inventory variables were calculated using the equations presented in Lamb et al. (2017). Since detailed species information was not available from the LiDAR-derived inventory, plots within the plot-matching library required species compositions representative of those observed throughout the District's spruce plantations. Almost 400 Black Brook spruce plantation permanent sample plots with close to 1400 total measurements were used calculate the 95th percentile of species compositions by basal area for spruce (≥ 60%), balsam fir (< 40%), hardwood (< 17%), and non-commercial species (< 5%). Only ancillary plot measurements that fell within the range of observed species composition were included in the plot-matching library. This included 2983 black spruce, 1184 white spruce, 63 Norway spruce, and 216 red spruce dominated plantation plot measurements. Since white spruce and Norway spruce shared similar ranges of forest inventory and are both planted on similar site conditions, the 1184 white spruce plots were combined with the Norway spruce plots to increase the sample size of Norway spruce plots. Inventory variables were recalculated for these plots using Norway spruce height-diameter relationships and taper equations. This was deemed necessary as over 270 000 LiDAR grid cells were located within Norway spruce plantations. In total, 5630 plot measurements were available for plot matching.

3.3.5 Plot-matching approach

A key prerequisite for successful application of the plot-matching method was that plot measurements and LiDAR grid cells share the same planted/dominant spruce species.

Planted species was identified for each LiDAR grid cell by intersecting the cell centroids with the J.D. Irving, Limited forest resource inventory polygons using ArcGIS (ESRI, Redlands, CA), which contained planted species and year of establishment for each plantation polygon. This resulted in almost 2.1 million LiDAR grid cells in spruce plantations: 927 360 black spruce, 843 360 white spruce, 272 130 Norway spruce, and 41 810 red spruce.

A nearest-neighbor matching algorithm built within the FORUS program (FORUS Research 2011) was used to match LiDAR grid cells to one (k = 1) ancillary plot based on the smallest sum of squared difference between each inventory variable used, referred to as 'distance' (Equation 3.1). The influence of different variable ranges (e.g., BA_t vs VOL_t) was removed from matching by normalizing each variable by dividing values by the maximum of observed plot values for each variable. Match distance was calculated (Lamb et al. 2017) as:

[3.1] $$Distance = \sum_{v}^{V} w_v \left(\frac{c_v}{Max(p_{i...n})_v} - \frac{p_v}{Max(p_{i...n})_v} \right)^2$$

where w is variable weight, c is LiDAR cell value, p is ancillary plot value ($i...n$), for each variable v: planted species, BA_t, BA_m, VOL_t, VOL_m, TopHt, or QMD_m.

Planted species was converted to a numeric value (e.g., black spruce = 1, white spruce = 2, etc.). To ensure that LiDAR grid cells and ancillary plot matches shared the same planted/dominant spruce, an arbitrarily high variable weight of 1000 was used for planted/dominant species, whereas all other variables (BA_t, BA_m, VOL_t, VOL_m, TopHt, and QMD_m) were given an equal weight of 1. Variables extracted from matched plots are referred to herein as "plot-matched variables". More details about the plot-matching

approach and plot-level and plantation-level accuracy of the approach are presented in Lamb et al. (2017).

3.3.6 Forest growth model and model calibration

The growth model used to forecast forest inventory for LiDAR grid cells was Open Stand Model (OSM) (Hennigar 2015). OSM is a tree-list growth simulation model calibrated for the Acadian Forest Region that predicts forest vegetation changes in response to stand development for a range of management types and site conditions. For each annual iteration of growth, forest inventory variables for individual trees (DBH and height) are altered using species-specific DBH and height increment and tree survival equations. Species DBH and height measurements are used to calculate volume using regional taper equations (Honer et al. 1983, Li et al. 2012). Growth rates are influenced by the Acadian biomass growth index (BGI) model, which takes into account climate, lithology, soils, and topography (Hennigar et al. 2017). The BGI is available as a 20 × 20 m spatial raster (http://www.forusresearch.com/bgi.php) and was used to determine BGI for each plot and grid-cell location.

While OSM is calibrated for intensively managed plantations in the Acadian Forest, the Black Brook District is known to have above average productivity and plantation management practices compared to most other areas in this region, so it was necessary to ensure growth derived from OSM would be representative of the District's plantations. To do this, the 399 permanent sample plots found in Black Brook spruce plantations were used to calibrate OSM. Plot measurements were used only if no treatments occurred between two measurements to remove the influence of treatments on observed growth and yield.

This resulted in 173 plots in plantations ranging from 13–45 years old, with 534 total measurements: 175 black spruce, 295 white spruce, 1 Norway spruce, and 63 red spruce.

Height increment calibration was done within the OSM software. Each of the 534 plot measurements were forecasted in OSM using their respective BGI to produce predicted height increments, and these were compared with observed increments. OSM contains methods for 1) validation of height and DBH increment models against local observations and 2) auto-adjustment of these models to correct bias through simple linear regression (FORUS Research 2011). Variables considered in model building included total plot basal area, DBH, and basal area of larger trees. Height increment models were built and used to calibrate growth for black spruce, white spruce, red spruce, and balsam fir (*Abies balsamea* (L.) Mill.). No height increment model was built for Norway spruce due to low sample size. DBH increment calibration models were not used, as they did not result in improved predictions of DBH increment.

In addition to height increment calibration, the need for mortality calibration was also assessed. The FORUS platform allows for the modification of background mortality rates by user-defined diameter classes, as well as the maximum relative density limit where self-thinning is to be applied. Plot-level observed and predicted annual mortality rates (stems·year^{-1}) were calculated for five diameter classes (3–10 cm, 10.1–20 cm, 20.1–30 cm, 30.1–40 cm, and > 40 cm) to compare mortality. Based on the predicted and observed mortality, mortality rates were reduced by 50% and 80% in the 10.1–20 cm and 20.1–30 cm DBH classes, respectively. Relative density where self-thinning was to be applied was increased from 85% to 100%.

3.3.7 Forecasting forest inventory for LiDAR grid cells

To obtain forecasted inventory for LiDAR grid cells, each plot used in matching was forecasted using OSM and a BGI value representative of the LiDAR grid cell to which it was matched. BGI for Black Brook spruce plantations ranged from 2251 $kg \cdot ha^{-1} \cdot year^{-1}$ to 4894 $kg \cdot ha^{-1} \cdot year^{-1}$. To reduce the combinations of plot and BGI values, and thus the need for individual calculations, three BGI values were used for forecasts: 3000, 3750, and 4500 $kg \cdot ha^{-1} \cdot year^{-1}$. The 5630 inventory plots used in plot matching were forecasted for 10 years in 1-year time steps using each of these three BGI values. BGI of each LiDAR grid cell was determined using ArcGIS (ESRI, Redlands, CA) and the regional BGI raster (http://www.forusresearch.com/bgi.php), and rounded to the nearest of the three BGI values used for forecasting. Each LiDAR grid cell was then linked with its forecasted matched plot, creating 10-year forecasted BA_t, BA_m, VOL_t, VOL_m, basal area weighted average height (Lorey's height; LHt), total quadratic mean diameter (QMD_t), and live crown ratio for each grid cell. Mean annual height increment was calculated for each grid cell by dividing its LHt by age of plantation at year of forecast.

3.3.8 Plot-level comparison of predicted inventory variable increments

The 100 calibration/validation plots were used to test the accuracy of inventory variable increments predicted using tree lists imputed from plot matching. Tree lists from the calibration/validation plots were input into OSM using the BGI for each plot location to obtain annual increments predicted from measured tree lists for six inventory variables (BA_t, BA_m, VOL_t, VOL_m, LHt, and QMD_t). Since these plots were 400-m^2 circular areas, they covered more than one single square 400-m^2 LiDAR grid cell. Plot match-derived

variable increments were calculated for each plot by summing the product of the variable increment multiplied by the corresponding proportions of the grid cell intersecting a plot area. Annual increment of the six variables predicted from tree lists imputed by plot matching were plotted against annual increment predicted from actual, measured plot tree lists.

Accuracy of LiDAR-derived inventory for the six plot-matched variables (BA_t, BA_m, VOL_t, VOL_m, TopHt, and QMD_m) was calculated to evaluate error between annual increments predicted from tree lists imputed by plot matching versus those predicted from measured plot tree lists. For each variable, error between LiDAR-derived versus measured calibration/validation plot values was calculated (LiDAR-derived variable value minus calibration/validation variable value). This error was divided by the measured inventory value for each plot and multiplied by 100 to calculate percent error for each plot-matched variable and plot. These percent error values were used to categorize each plot into two "goodness of fit" classes: "high accuracy" (≤ 10% error for all plot-matched variables) and "low accuracy" (> 10% error for one or more plot-matched variable). Pearson correlation coefficient (*r*), mean error, and root mean squared error (RMSE) were calculated for all plots as well as for the high accuracy and low accuracy plot groups. To compare across variables, mean error and RMSE were calculated as a percentage by dividing it by the mean of each respective variable and multiplying by 100. This comparison was done using 98 of the 100 calibration/validation plots, as two plots fell outside of spruce plantation boundaries and were excluded.

3.3.9 Plantation-level comparison of harvest volume and forecasted volume

Harvested volume data for 15 spruce plantation harvest blocks in the Black Brook District were compared with corresponding volumes forecast from tree lists imputed by plot matching. These blocks were located in six black spruce and nine white spruce plantations with areas ranging from 4 to 45 ha. All blocks were clearcut in 2016 with the same product specifications as used to calculate VOL_m. Four blocks were located in the 2011 LiDAR acquisition area, while the remaining 11 were located in the 2013 LiDAR area, equating to 5- and 3-year plot match-derived forecasted VOL_m, respectively. The 2016 plot match-derived VOL_m was intersected with the block polygons using ArcGIS (ESRI, Redlands, CA) to calculate forecasted VOL_m and this was plotted against corresponding measured harvest volumes. In addition, VOL_m derived from the original LiDAR predictions were plotted against measured harvest volumes to assess improvements when using plot match-derived forecasted VOL_m. Pearson correlation coefficients (r) and mean bias, both weighted by harvest block area, were calculated for each case. A paired t-test, weighted by block area, was used to determine if plot match-derived forecasted VOL_m and LiDAR-derived VOL_m were significantly different from measured harvest volumes.

3.3.10 Mapping commercial thinning eligibility for operational harvest planning

J.D. Irving, Limited commercial thinning eligibility rules were used to rank and map LiDAR grid cells that would qualify for commercial thinning in 2018 using plot match-derived forecasted inventory variables. These rules use “points” assigned to LiDAR- or plot match-derived inventory variables (i.e., VOL_m, average merchantable tree volume [VOL_m/stems per hectare], live crown ratio, and mean annual height increment

[$m \cdot yr^{-1}$]) as well as minimum/maximum age and planted species derived from forest inventory polygons. The sum of these points were used to rank cells for commercial thinning eligibility. Using plot match-derived forecasted variable values for 2018 and forest resource inventory polygons, commercial thinning eligibility points were calculated for each grid cell. Cells were then categorized based on their commercial thinning eligibility as high ranking (points ≥ 9.5), medium ranking (7.5–9.4), and low ranking (0.1–7.4). High ranking cells are expected to have the greatest response to release and highest financial return, in comparison to medium and low ranking cells.

To demonstrate the process for prioritizing annual commercial thinning treatments using plot match-derived forecasted inventory, the commercial thinning ranking system and a measure of block-level competition were used to spatially schedule commercial thinning for 2018, 2019, and 2020. The objective was to target high ranking blocks that would provide the greatest financial return and further prioritize to capture density-dependent mortality. Using the dissolve tool in ArcGIS (ESRI, Redlands, CA), all high-ranking cells sharing a border were dissolved together to create new contiguous block polygons. Block polygons less than 4 ha were removed from commercial thinning consideration as this was the smallest harvest block size observed from the 2016 harvest data. To quantify block-level competition, basal area was used as it is commonly used as a measure of competition (Curtis 1982; Biging and Dobbertin 1995). Mean BA_t was calculated for each block polygon using the 2018 plot match-derived forecasted BA_t. Block polygons were selected in order of decreasing mean BA_t until the sum of total area selected was ≤ 3000 ha (average total annual area commercially thinned in the District since 2014; David Young, Management Forester, J.D. Irving, Limited, personal communication,

September 1, 2017). These polygons represented the proposed 2018 commercial thinning blocks. To select commercial thinning blocks for 2019 and account for proposed treatments in 2018, all 2019 plot match-derived grid cells that were coincident with the proposed 2018 commercial thinning blocks were removed using the erase tool in ArcGIS (ESRI, Redlands, CA). Following the same process used for selecting 2018 commercial thinning blocks, ≤ 3000 ha of high ranking, high BA_t areas were selected from the 2019 plot match-derived forecasted inventory. This same process was repeated for the 2020 plot match-derived forecasted inventory. This resulted in proposed 2018, 2019, and 2020 commercial thinning blocks that were mapped to show the distribution across the District.

3.4 Results

3.4.1 Comparison of variable increments predicted from plot-matched versus field-measured tree lists

Inventory variable increments predicted with OSM using tree lists imputed by plot matching were strongly correlated with those using measured tree lists from 98 calibration/validation plots. For the six inventory variables compared, *r* ranged from 0.75–0.86 (Table 3.1 and Fig. 3.2). Percentage mean error and percentage RMSE ranged from -3.6%–0.0% and 12.8%–36.6% for BA_t, VOL_t, VOL_m, LHt, and QMD_t, but was 11.3% and 49.0% for BA_m, respectively (Table 3.1).

Comparison of variable increment predictions between the high accuracy and low accuracy plot groups revealed that plot matching performed best when accuracy of LiDAR-derived inventory was high (Fig. 3.2). Percentage RMSE was lowest (4.9%–13.5%) in high accuracy plots and highest (14.5%–56.2%) for low accuracy plots (Table 3.1). The

difference was greatest for BA_m where RMSE for high accuracy plots was 13.5% compared to 56.2% in low accuracy plots. Percentage mean error ranged from -0.5%–3.0% for high accuracy plots compared to -3.8%–14.4% for low accuracy plots (Table 3.1).

Table 3.1. Pearson correlation coefficient (*r*), mean error, and root mean squared error (RMSE) for annual increment of six inventory variables (basal area, merchantable basal area, volume, merchantable volume, Lorey's height, and quadratic mean diameter) forecasted for 5 years using Open Stand Model using tree lists imputed by plot matching versus tree lists measured in 98 calibration/validation plots. Plots were categorized based on accuracy of LiDAR-derived inventory variable values: high accuracy (≤ 10% error for all plot-matched variables) and low accuracy (> 10% error for one or more plot-matched variable).

Inventory Variable	Plot Group	n	*r*	Mean Error	Mean Error (%)	RMSE	RMSE (%)
Total basal area ($m^2 \cdot ha^{-1} \cdot yr^{-1}$)	All	98	0.82	-0.04	-3.6	0.41	36.6
	High accuracy	25	0.88	0.03	3.0	0.12	11.1
	Low accuracy	73	0.81	-0.04	-3.8	0.47	41.9
Merchantable basal area ($m^2 \cdot ha^{-1} \cdot yr^{-1}$)	All	98	0.75	0.12	11.3	0.54	49.0
	High accuracy	25	0.85	0.03	2.5	0.15	13.5
	Low accuracy	73	0.75	0.16	14.4	0.61	56.2
Total volume ($m^3 \cdot ha^{-1} \cdot yr^{-1}$)	All	98	0.78	-0.15	-1.6	2.21	23.4
	High accuracy	25	0.77	0.04	0.4	1.01	10.7
	Low accuracy	73	0.77	-0.22	-2.3	2.50	26.4
Merchantable volume ($m^3 \cdot ha^{-1} \cdot yr^{-1}$)	All	98	0.84	-0.01	-0.1	1.99	22.6
	High accuracy	25	0.77	0.04	0.5	1.03	11.7
	Low accuracy	73	0.83	-0.03	-0.3	2.23	25.3
Lorey's height ($m \cdot ha^{-1} \cdot yr^{-1}$)	All	98	0.86	-0.01	-1.5	0.05	12.8
	High accuracy	25	0.90	0.00	0.1	0.02	4.9
	Low accuracy	73	0.85	-0.01	-2.1	0.06	14.5
Quad. mean diameter ($cm \cdot ha^{-1} \cdot yr^{-1}$)	All	98	0.85	0.00	0.0	0.08	19.5
	High accuracy	25	0.92	0.00	-0.5	0.03	6.2
	Low accuracy	73	0.83	0.00	-0.1	0.09	22.3

[1] Percentage mean error and percentage root mean squared error calculated as a percentage by dividing it by the mean of each respective variable.

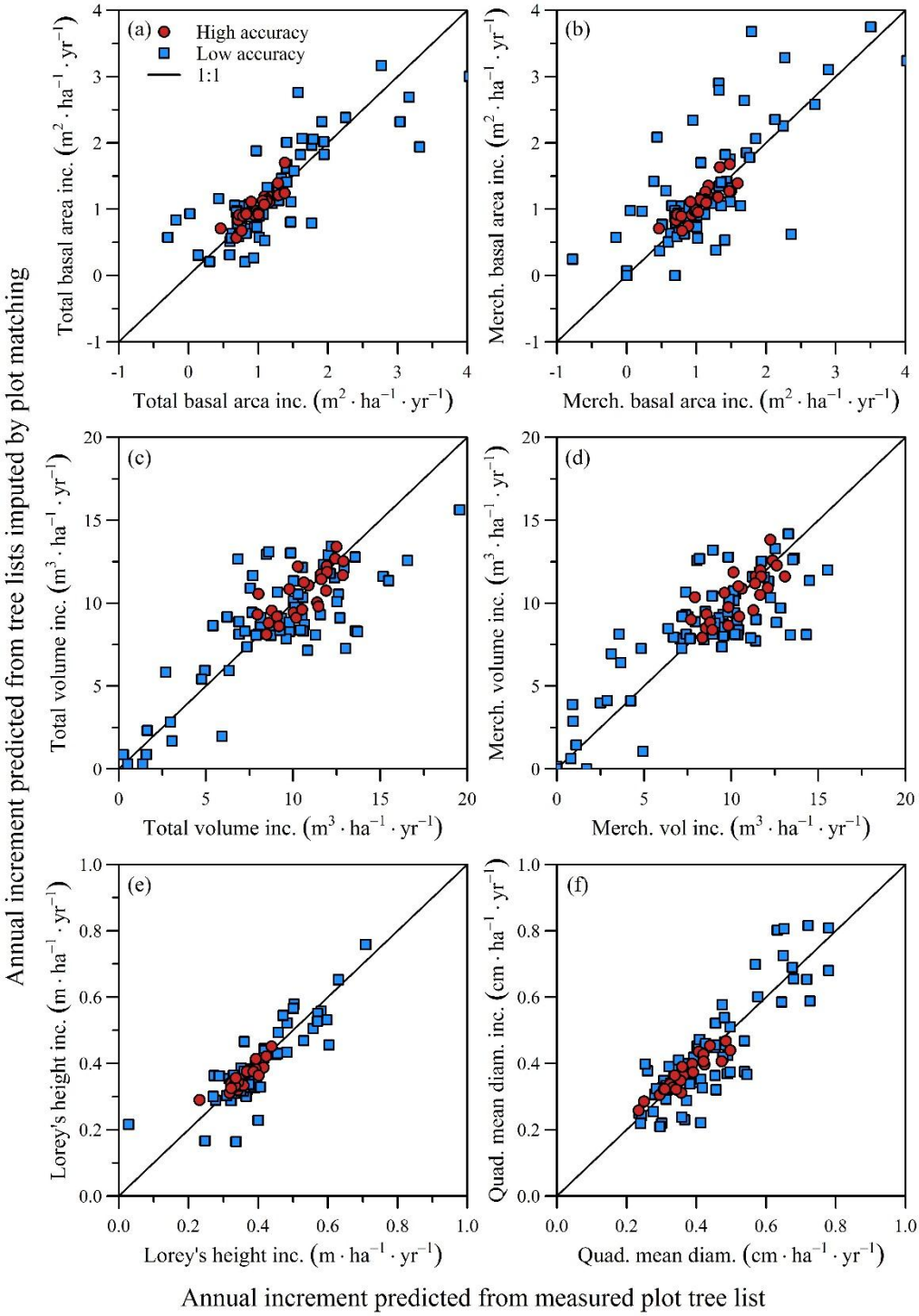

Fig. 3.2. Comparison of annual increment over 5 years predicted using Open Stand Model for six inventory variables [(a)–(f): basal area, merchantable basal area, volume, merchantable volume, Lorey's height, and quadratic mean diameter] for tree lists imputed by plot matching and measured in 98 calibration/validation plots. Plots were categorized based on accuracy of LiDAR-derived inventory variable values: high accuracy (≤ 10% error for all plot-matched variables; *n* = 25) and low accuracy (> 10% error for one or more plot-matched variable; *n* = 73).

3.4.2 Comparison of harvested volumes with predicted volumes derived from forecasted plot matches and LiDAR

Comparison of harvested volumes from 15 spruce plantations harvested in 2016 revealed that plot match-derived forecasted VOL_m were highly accurate (Table 3.2 and Fig. 3.3). Differences between plot match-derived forecasted VOL_m and harvested volumes ranged from -12.0%–14.4% (Table 3.2) with an overall mean difference of only 1.5% when weighted by block area. When year of LiDAR acquisition was considered, mean differences were -4.1% and 3.0% for harvest blocks in the 2011 and 2013 LiDAR acquisition areas, respectively. The greatest errors in this comparison resulted from inaccuracy of the LiDAR-derived VOL_m. For harvest blocks 11 and 12, with errors of 14.4% and 14.3%, LiDAR-derived VOL_m was 11.8 and 7.0 $m^3 \cdot ha^{-1}$ greater, respectively, than volumes harvested 3 years after LiDAR-derived predictions. With the exception of harvest blocks 3 and 13, forecasted plot matches resulted in larger VOL_m than the original LiDAR-derived predictions. This resulted in a stronger correlation and reduced mean bias between harvested volumes and plot match-derived forecasted VOL_m (0.92 and 2.5 $m^3 \cdot ha^{-1}$) when compared to LiDAR-derived VOL_m (0.83 and 11.0 $m^3 \cdot ha^{-1}$) (Fig. 3.3). In addition, the results from the paired *t*-test indicated that plot match-derived forecasted VOL_m was not significantly different from harvest volumes ($p = 0.54$), but that the unforecasted LiDAR-derived VOL_m was ($p = 0.032$).

Table 3.2. Actual harvested versus forecasted merchantable volume derived from tree lists imputed by plot matching using Open Stand Model for 15 blocks harvested in 2016 in Black Brook District. Harvest blocks located in the 2011 and 2013 LiDAR acquisition areas represent 5- and 3-year plot match-derived forecasted merchantable volume, respectively. Averages were weighted by harvest block area. All blocks were clearcut with the same product specifications as was used to calculate plot match-derived merchantable volume.

Harvest Block	Planted Species	Area (ha)	Years Forecasted	Harvested volume ($m^3 \cdot ha^{-1}$)	Forecasted volume ($m^3 \cdot ha^{-1}$)	Percent difference (%)
1	Black spruce	4	3	127	133	4.6
2	Black spruce	4	3	154	157	2.1
3	Black spruce	8	3	181	186	2.9
4	Black spruce	10	5	165	145	-12.0
5	Black spruce	12	5	156	171	10.0
6	White spruce	12	5	248	221	-10.7
7	White spruce	14	3	179	177	-1.2
8	White spruce	17	5	203	199	-2.0
9	White spruce	17	3	230	248	8.0
10	White spruce	22	3	141	132	-6.4
11	White spruce	23	3	118	135	14.4
12	White spruce	24	3	129	147	14.2
13	White spruce	26	3	182	178	-2.3
14	Black spruce	29	3	184	172	-6.6
15	White spruce	45	3	160	172	7.5
Average	-	18	3.4	169	172	1.5

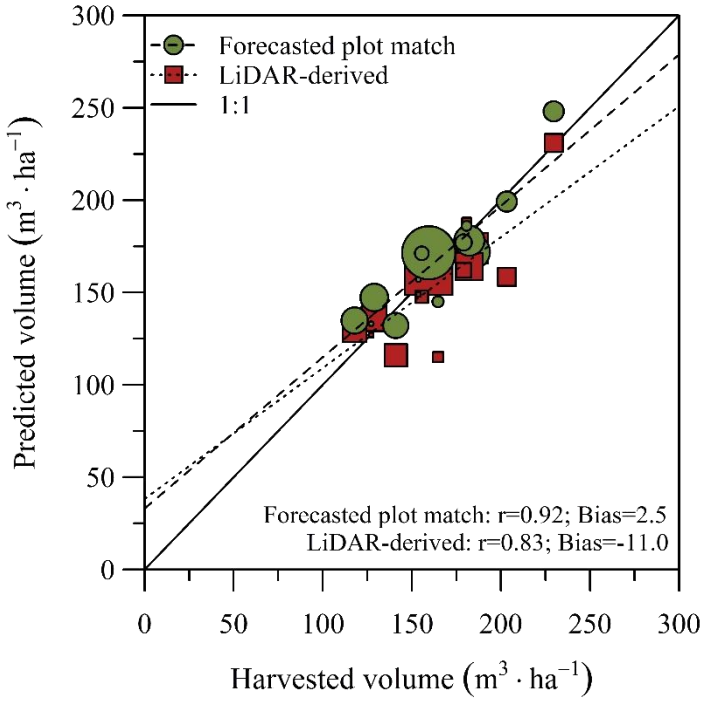

Fig. 3.3. Comparison of merchantable volume derived from tree lists imputed by plot matching and forecasted using Open Stand Model (green circles) and original, unforecasted LiDAR-derived predictions (red squares) with actual harvest volumes from 15 harvest blocks harvested in 2016. Time elapsed between LiDAR acquisition and harvest ranged from 3–5 years with an average of 3.4 years. Volume difference between forecasted plot matches and original LiDAR ranged from -2.0–40.8 $m^3 \cdot ha^{-1}$ with an average of 13.5 $m^3 \cdot ha^{-1}$. Point sizes represent relative harvest block size (3.9–45.0 ha). Bias is presented in variable units ($m^3 \cdot ha^{-1}$).

3.4.3 Mapping commercial thinning ranking and scheduling annual commercial thinning treatments for spruce plantations in the Black Brook District

Using the 2018 plot match-derived forecasted inventory, commercial thinning rankings were applied to individual grid cells using J.D. Irving, Limited commercial thinning eligibility rules. When grid cells were removed if they did not fall within the minimum and maximum age limits, 32 370 ha (809 250 cells) were “eligible” for commercial thinning (Fig. 3.4). Ranking individual cells resulted in 9870 ha, 12 780 ha, and 9720 ha of high, medium, and low ranking area (Fig. 3.4a). When bordering cells sharing the same commercial thinning rank were dissolved into larger polygons, the result

was 49 760 individual polygons. This number was greatly reduced when minimum polygon size was considered, as there were only 2840 and 1470 polygons ≥ 1 ha and ≥ 4 ha, representing 27 920 ha and 25 160 ha, respectively.

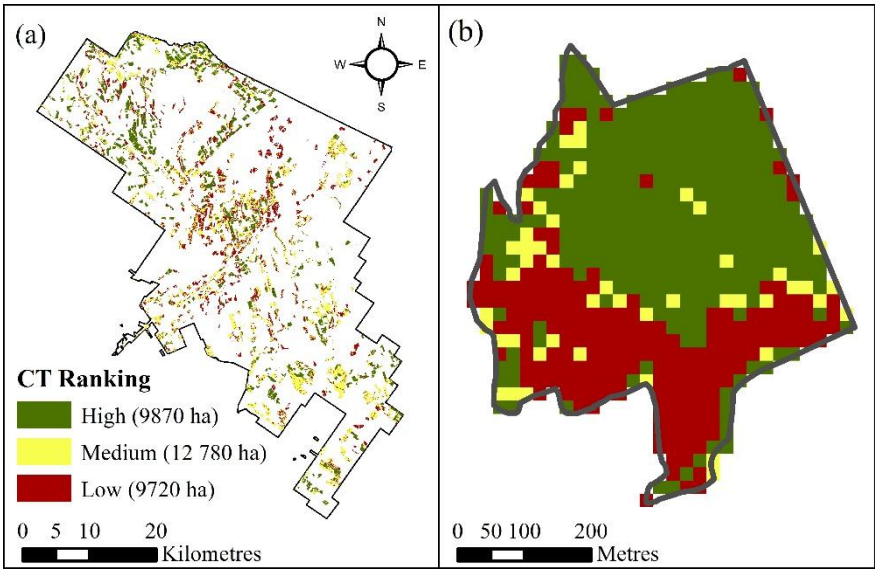

Fig. 3.4. Commercial thinning (CT) ranking for LiDAR grid cells in (a) Black Brook District and (b) one sample stand using inventory variables derived from tree lists imputed by plot matching and forecasted to 2018 using Open Stand Model. Cell ranking was determined using J.D. Irving, Limited commercial thinning eligibility rules for first commercial thinning, which uses species, age, gross merchantable volume, mean tree gross merchantable volume, mean annual height increment, and live crown ratio.

When annual commercial thinning was planned for high ranking blocks ≥ 4 ha using plot match-derived forecasted BA_t as a competition index, the result was 2970 ha, 2990 ha, and 2970 ha for years 2018, 2019, and 2020, respectively (Fig. 3.5). Minimum block size was 4 ha for all years, however, maximum size was 290 ha, 200 ha, and 100 ha, and mean size was 27 ha, 22 ha, and 19 ha for years 2018, 2019, and 2020, respectively. The number of blocks scheduled for years 2018, 2019, and 2020 was 112, 134, and 157, respectively.

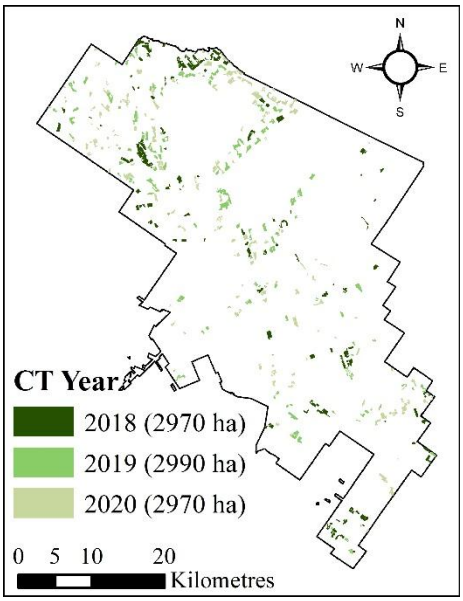

Fig. 3.5. Optimal year of first commercial thinning (CT) for 2018, 2019, and 2020 using inventory variables derived from tree lists imputed by plot matching and forecasted to 2018 using Open Stand Model. Proposed blocks include high ranking grid cells based on J.D. Irving, Limited commercial thinning eligibility rules as well as a block-level competition index (basal area). J.D. Irving, Limited commercial thinning eligibility rules use species, age, gross merchantable volume, mean tree gross merchantable volume, mean annual height increment, and live crown ratio. Priority was placed on blocks with greatest total basal area to capture density dependent mortality.

3.5 Discussion

LiDAR-derived forest inventory introduced a major shift in how forest management is practiced in Canada (Woods et al. 2011; White et al. 2013) and abroad (Næsset et al. 2004; Næsset 2014). With LiDAR, planning of harvest and silviculture treatments no longer need be determined by photo-interpreted stand polygons and strata-based yield curves. Instead, within- and between-stand variability can be determined using this high-resolution inventory, allowing for prescription development prior to site visits. Silviculture treatments and harvests can be prescribed to areas that will result in the greatest growth response to release and highest financial return. However, without a means of forecasting the inventory, forest managers would eventually be forced to return to stand-based yield curves for operational-, tactical-, and strategic-level planning between

inventory refresh cycles. Using tree-list imputation methods (Lamb et al. 2017), we have presented a framework for forecasting grid-cell inventory variables using imputed tree lists and a locally calibrated tree-list growth model. This extends the life of LiDAR-derived inventory and allows forest managers to continue high-resolution forest planning into the future.

Tree lists imputed for spruce plantation LiDAR grid cells from plot matches allowed for accurate inventory variable increment predictions when input into a locally calibrated growth model, as validated by comparing variable increments predicted using measured tree lists for 98 calibration/validation plots. Plot matches resulted in strong correlation, low mean bias, and low RMSE for BA_t, BA_m, VOL_t, VOL_m, LHt, and QMD_t increment predictions, compared with increments predicted from measured trees. When accuracy of LiDAR-derived plot-matched variable values were high (≤ 10% error for all plot-matched variables), percentage RMSE values were 9.6%–42.7% lower than for plots with lower accuracy (> 10% error for all plot-matched variables). Much of the error associated with variable increment predictions resulted from calibration/validation plots located in young plantations with merchantable values of 0. For example, of the 20 plots with plot match-derived BA_m increments that differed from measured tree-list increments by more than ± 0.5 $m^2 \cdot ha^{-1} \cdot year^{-1}$, 11 had LiDAR-derived merchantable values of 0. Without merchantable variables, plot matches were based largely on BA_t, VOL_t, and TopHt. This introduced the possibility that matched plots may not be representative of actual conditions, as the combination of BA_t, VOL_t, and TopHt was not sufficient to accurately describe density or average diameter. The addition of QMD_t as a plot-matched variable would ensure that plot matches are more representative of density and average

diameters, however, this LiDAR-derived variable was not predicted by the LiDAR contractor so was not available for matching. Matching to grid cells in unmerchantable stands has additional challenges in that few plots were in young stands. Of the 5630 plot measurements used in matching, only 9% had no merchantable tree measurements, compared with 35% of all LiDAR grid cells with merchantable value predictions of 0. Despite increased variance in the growth increment predictions of these plots, predictions remained unbiased suggesting that errors will average out when grid cells are aggregated.

When plantation harvest blocks were examined, plot match-derived forecasted VOL_m showed strong agreement with harvested volumes, suggesting that forecasted plot matches may provide estimates of stand-level volume that are acceptable for operational planning. Although plot match-derived forecasted VOL_m accuracy varied between blocks, overall, plot match-derived forecasted VOL_m was only 1.5% greater than harvested volume and was found to not be significantly different. The slightly higher average VOL_m may be the result of defects or cull not accounted for in the calculation of VOL_m. If available, species-specific cull rates for the area in question could be applied to imputed tree lists to calculate volumes that are more representative of ground conditions. In addition to the fact that imputed tree lists offer the ability to account for cull or defects, the benefits of plot matching and the resulting forecasted VOL_m were evident by the increased correlation and decreased bias when compared with 3–5-year-old LiDAR-derived VOL_m. For fast growing stands, such as the spruce plantations in the Black Brook District, the addition of plot match-derived forecasted inventory with LiDAR-derived estimates presents a significant advantage for operational forest management.

The use of LiDAR- or plot match-derived inventory variables to spatially map areas of interest may be the most beneficial advantage of this high-resolution inventory. Quantifying variability in inventory variables is of greater importance than knowing stand-level averages, as variability directly impacts operational costs, response to treatment, wildlife habitat, etc. (Woods et al. 2011). This is particularly true for silvicultural treatments such as commercial thinning, which have very exacting stand-condition requirements for positive cost-benefit outcomes Watt et al. (2013) concluded that LiDAR-derived estimates of average tree height and base to live crown were sufficiently precise to delineate and schedule commercial thinning in Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) stands. However, as time since acquisition of LiDAR-derived inventory increases, it becomes less representative of actual conditions. We have demonstrated how plot match-derived inventory variables, forecast for 3–5 years with a tree-list growth model, and commercial thinning ranking rules can be used to delineate high-ranking grid cells beyond the year of LiDAR acquisition. In addition, we presented a novel method for planning annual commercial thinning treatments using the commercial thinning ranking and a measure of competition. Although BA_t was used as the measure of competition, other competition indexes (e.g., stand density index or relative density [Weiskittel et al. 2011]) can be calculated from imputed tree lists and used to prioritize thinning treatments.

Forecasting forest inventory is an important issue for all inventory methods, including and LiDAR. Although methods to obtain forecasted inventory at grid-cell resolution have been proposed (Härkönen et al. 2013; Tompalski et al. 2016), the variables that can be obtained from the growth models used in these studies are limited as they did not use tree-level inventory. The advantage of the plot-matching method over other LiDAR

forecasting methods is the ability to generate tree lists and to calculate additional inventory variables that were not estimated from LiDAR, and to combine them to generate forecasts. Variables such as species-specific product volume (Peuhkurinen et al. 2008), biomass and carbon (Næsset and Gobakken 2008), and wildlife habitat (Vierling et al. 2008) have all been estimated using LiDAR, but each variable requires a unique set of calibration plots to cover the variability observed on the landbase (White et al. 2013). In contrast, these variables may be calculated from the plot match-imputed tree lists and forecasted, providing forest managers with the tools needed for high resolution management planning between inventory refresh cycles. As product specifications change, biomass equations improve, or wildlife habitat measures advance, these variables can be updated and forecasted, providing a very flexible and adaptable inventory dataset.

Despite the benefits that forecasted plot matches present, forecast inventory estimates are limited by several potential issues. Accuracy of plot match-derived inventory variables and subsequent forecasts are limited by the accuracy of the LiDAR-derived inventory (Lamb et al. 2017). Even with accurate LiDAR-derived inventory, plot match-derived forecasted inventory may still be inaccurate for stand conditions not represented in the plot measurement database, so, traditional age-based yield curves may be more accurate for rare forest cover types. Finally, the accuracy of the tree-list growth model used to forecast matches will influence results. Since our study area had a large number of permanent sample plot measurements available to calibrate height and mortality increments for OSM, this likely contributed to the accuracy of estimates of forecasted variable values.

3.6 Conclusions

We imputed tree-level inventory for 2.1 million LiDAR grid cells in spruce plantations using a nearest-neighbor matching algorithm and plot data from throughout the province of New Brunswick (Lamb et al. 2017). This method uses the library of existing plot measurements and a small number of LiDAR-derived inventory variables to impute complex tree-level inventory data for each grid cell. Using a locally calibrated tree-list growth model and BGI value for the cell location, imputed tree lists were forecast 10 years in 1-year time steps to obtain forecasted inventory variables (i.e., BA_t, VOL_t, LHt, etc.) at 20 × 20 m grid-cell resolution. Increments of variables predicted using tree lists imputed from plot matching were strongly correlated with those predicted from measured tree lists, suggesting that imputed tree lists were sufficiently accurate for forecasting inventory variables. This was supported by the comparison of plot match-derived forecasted VOL_m and harvested volume, which showed stronger correlation and lower bias when compared with 3–5-year-old LiDAR-derived VOL_m. Using plot match-derived forecasted inventory variables (e.g., VOL_m, mean annual height increment, live crown ratio, etc.) and commercial thinning eligibility rules, we demonstrated the ability to spatially plan annual silviculture treatments that were otherwise unobtainable with the older LiDAR-derived inventory.

The benefit of this approach is the use of a tree-list growth model to obtain high resolution forecasted inventory. Since these models track growth and yield of each tree, users are provided with the flexibility to calculate species-specific variables and manage for multiple objectives. Typically, forecasting forest inventory for a landbase at grid-cell resolution with a tree-list growth model would be computationally infeasible. However,

we have presented a novel and practical solution that allows forest managers to make detailed operational-, tactical-, and strategic-management plans, not otherwise available using traditional inventory methods.

3.7 Acknowledgements

We thank the staff at J.D. Irving, Limited and Leading Edge Geomatics for their input and assistance with this project. This research was funded by the AWARE (Assessment of Wood Attributes using Remote sEnsing) Natural Sciences and Engineering Research Council of Canada Collaborative Research and Development grant to a team led by Dr. Nicholas Coops and by J.D. Irving, Limited. Sean Lamb was also funded by an NSERC Canada Graduate Scholarship.

3.8 References

Albaugh, T.J., Allen, H.L., Dougherty, P.M., and Johnsen, K.H. 2004. Long term growth responses of loblolly pine to optimal nutrient and water resource availability. For. Ecol. Manage. 192(1): 3–19.

Andersen, H.E., McGaughey, R.J., and Reutebuch, S.E. 2005. Estimating forest canopy fuel parameters using LIDAR data. Remote Sens. Environ. 94(4): 441–449.

Biging, G.S., and Dobbertin, M. 1995. Evaluation of competition indices in individual tree growth-models. For. Sci. 41(2): 360–377

Breiman, L. 2001. Random forests. Mach. Learn. 45(1): 5–32.

Canham, C.D., LePage, P.T., and Coates, K.D. 2004. A neighborhood analysis of canopy tree competition: effects of shading versus crowding. Can. J. For. Res. 34(4): 778–787.

Crookston, N.L., and Dixon, G.E. 2005. The forest vegetation simulator: A review of its structure, content, and applications. Comput. Electron. Agric. 49(1): 60–80.

Curtis, O. 1982. A simple index of stand density for Douglas-fir. For. Sci. 28(1): 92–94.

Dale, V.H., Doyle, T.W., and Shugart, H.H. 1985. A comparison of tree growth models. Ecol. Model. 29(1–4): 145–169.

Duchesne, I., Pitt, D.G., and Tanguay, F. 2013. Effects of precommercial thinning on the forest value chain in northwestern New Brunswick: Part 4 - Lumber production, quality, and value. For. Chron. 89(4): 474–489.

Falkowski, M.J., Hudak, A.T., Crookston, N.L., Gessler, P.E., Uebler, E.H., and Smith, A.M.S. 2010. Landscape-scale parameterization of a tree-level forest growth model: a k- nearest neighbor imputation approach incorporating LiDAR data. Can. J. For. Res. 40(2): 184–199.

FORUS Research. 2011. FORUS Slasher User Guide. Fredericton, NB, Canada.

Gobakken, T., and Næsset, E. 2004. Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. Scand. J. For. Res. 19(6): 529–542.

Haara, A., Maltamo, M., and Tokola, T. 1997. The K-nearest-neighbour method for estimating basal-area diameter distribution. Scand. J. For. Res. 12(2): 200–208.

Härkönen, S., Tokola, T., Packalén, P., Korhonen, L., and Mäkelä, A. 2013. Predicting forest growth based on airborne light detection and ranging data, climate data, and a simplified process-based model. Can. J. For. Res. 43(4): 364–375.

Hennigar, C., Weiskittel, A., Allen, H.L., and MacLean, D.A. 2017. Development and evaluation of a biomass increment-based index for site productivity. Can. J. For. Res. 47(3): 400–410.

Hennigar, C.R. 2015. Open Stand Model version 1.03.2: Help documentation. Fredericton, NB, Canada.

Holmgren, J. 2004. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. Scand. J. For. Res. 196(19): 543–553.

Holmgren, J., Nilsson, M., and Olsson, H. 2003. Estimation of tree height and stem volume on plots using airborne laser scanning. For. Sci. 49(3): 419–428.

Honer, T.G., Alemdag, I.S., and Ker, M.F. 1983. Metric timber tables for the commercial tree species of Central and Eastern Canada. Can. For. Serv., Maritimes For. Res. Cent. Inf. Rep. M-X-140.

Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M.S., 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. Can. J. Remote Sens. 32(2): 126–138.

Lamb, S.M., MacLean, D.A., Hennigar, C.R., and Pitt, D.G. 2017. Imputing tree lists for New Brunswick spruce plantations through nearest-neighbor matching of airborne laser scan and inventory plot data. Can. J. Remote Sens. 43(3): 269–285.

Leading Edge Geomatics. 2015. Data Dictionary: Black Brook planted spruce stands enhanced forest inventory description of data collection and processing. Unpublished report submitted to J.D. Irving, Limited.

Li, R., Weiskittel, A., Dick, A.R., Kershaw, J.A., Seymour, R.S.S.R.S., and Kershaw Jr., J.A. 2012. Regional stem taper equations for eleven conifer species in the Acadian Region of North America: development and assessment. North. J. Appl. For. 29(1): 5–14.

Lim, K., Treitz, P., Baldwin, K., Morrison, I., and Green, J. 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. Can. J. Remote Sens. 29(5): 658–678.

Long, J.N. 1985. A practical approach to density management. For. Chron. 61(1): 23–27.

MacDonald, M. 2008. New Brunswick's forest development survey (FDS) data collection manual (2003–2013 forest inventory cycle). Forest Management Branch, New Brunswick Department of Natural Resources, Fredericton, NB. 105 p.

MacLean, D.A., Amos-Binks, L.J., Adams, G., Pelletier, G., and Villard, M.-A. 2010. Legacy of the Sustainable Forest Management Network: outcomes of research collaborations among J.D. Irving, Limited, University of New Brunswick, and Université de Moncton. Sustainable Forest Management Network Center of Excellence, University of Alberta.

Mäkinen, H., and Isomäki, A. 2004. Thinning intensity and growth of Norway spruce stands in Finland. For. 77(4): 349–364. doi:10.1093/forestry/77.4.349.

Maltamo, M., and Kangas, A. 1998. Methods based on k-nearest neighbor regression in the prediction of basal area diameter distribution. Can. J. For. Res. 28: 1107–1115.

Maltamo, M., Næsset, E., Bollandsås, O.M., Gobakken, T., and Packalén, P. 2009. Non-parametric prediction of diameter distributions using airborne laser scanner data. Scand. J. For. Res. 24(6): 541–553.

McInerney, D.O., and Nieuwenhuis, M. 2009. A comparative analysis of kNN and decision tree methods for the Irish National Forest Inventory. Int. J. Remote Sens. 30(19): 4937–4955.

Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., and Abstract, C.J.H. 2000. Predicting forest stand characteristics with airborne scanning lidar. Photogramm. Eng. Remote Sens. 66(11): 1367–1371.

Næsset, E. 1997. Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS J. Photogramm. Remote Sens. 52(2): 49–56.

Næsset, E. 2007. Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia. Scand. J. For. Res. 22(5): 433–442.

Næsset, E. 2014. Area-based inventory in Norway – from innovations to operational reality. In Forestry Applications of Airborne Laser Scanning, Concepts and Case Studies. Edited by M. Maltamo, E. Næsset, and J. Vauhkonen. Springer, Dordrecht, Netherlands. pp. 215–240.

Næsset, E., and Gobakken, T. 2008. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. Remote Sens. Environ. 112(6): 3079–3090.

Næsset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, Å., and Söderman, U. 2004. Laser scanning of forest resources: the nordic experience. Scand. J. For. Res. 19(6): 482–499.

Nilsson, M. 1996. Estimation of tree heights and stand volume using an airborne lidar system. Remote Sens. Environ. 56(95): 1–7.

Nyland, R.D. 2007. Silviculture: Concepts and Applications. In 2nd edition. Waveland Press, Inc., Long Grove, IL, USA.

Omari, K., and Maclean, D.A. 2015. Do biomass removal and structure-enhancing treatments influence deadwood characteristics following commercial thinning in spruce plantations in New Brunswick, Canada? Can. J. For. Res. 45(10): 1407–1418.

Packalén, P., and Maltamo, M. 2008. Estimation of species-specific diameter distributions using airborne laser scanning and aerial photographs. Can. J. For. Res. 38(7): 1750–1760.

Pelletier, G., and Pitt, D.G. 2008. Silvicultural responses of two spruce plantations to midrotation commercial thinning in New Brunswick. Can. J. For. Res. 38(4): 851–867.

Penner, M., Pitt, D.G., and Woods, M.E. 2013. Parametric vs. nonparametric LiDAR models for operational forest inventory in boreal Ontario. Can. J. Remote Sens. 39(5): 426–443.

Peuhkurinen, J., Maltamo, M., and Malinen, J. 2008. Estimating species-specific diameter distributions and saw log recoveries of boreal forests from airborne laser scanning data and aerial photographs: A distribution-based approach. Silva Fenn. 42(4): 625–641.

Popescu, S.C., and Zhao, K. 2008. A voxel-based lidar method for estimating crown base height for deciduous and pine trees. Remote Sens. Environ. 112(3): 767–781.

Porter, K.B., MacLean, D.A., Beaton, K.P., and Upshall, J. 2001. New Brunswick permanent sample plot database (PSPDB version 1.0): user's guide and analysis. Can. For. Serv., Atl. For. Cent. Inf. Rep. M-X-209. 65p.

R Development Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.r-project.org/.

Smith, D.M., Larson, B.C., Kelty, M.J., and Ashton, P.M.S. 1997. The practice of silviculture: applied forest ecology. In 9th edition. John Wiley and Sons, Inc., New York, NY, USA.

Tompalski, P., Coops, N.C., White, J.C., and Wulder, M.A. 2016. Enhancing forest growth and yield predictions with airborne laser scanning data: Increasing spatial detail and optimizing yield curve selection through template matching. Forests 7(11): 1–20.

Vastaranta, M., Holopainen, M., Yu, X., Hyyppä, J., Hyyppä, H., and Viitala, R. 2011. Predicting stand-thinning maturity from airborne laser scanning data. Scand. J. For. Res. 26(2): 187–196.

Vierling, K.T., Vierling, L. a., Gould, W. a., Martinuzzi, S., and Clawges, R.M. 2008. Lidar: Shedding new light on habitat characterization and modeling. Front. Ecol. Environ. 6: 90–98.

Watt, M.S., Meredith, A., Watt, P., and Gunn, A. 2013. Use of LiDAR to estimate stand characteristics for thinning operations in young Douglas-fir plantations. NZ J. For. Sci. 43: 1–10.

Weiskittel, A.R., Hann, D.W., Kershaw, J.A., and Vanclay, J.K. 2011. Forest Growth and Yield Modeling. John Wiley & Sons, Ltd., West Sussex, UK.

White, J.C., Wulder, M.A., Varhola, A., Vastaranta, M., Coops, N.C., Cook, B.D., Pitt, D., and Woods, M. 2013. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. For. Chron. 89(6): 722–723.

Woods, M., Lim, K., and Treitz, P. 2008. Predicting forest stand variables from LiDAR data in the Great Lakes - St. Lawrence forest of Ontario. For. Chron. 84(6): 827–839.

Woods, M., Pitt, D., Penner, M., Lim, K., Nesbitt, D., Etheridge, D., and Treitz, P. 2011. Operational implementation of a LiDAR inventory in boreal Ontario. For. Chron. 87(4): 512–528.

Wulder, M.A., Coops, N.C., Hudak, A.T., Morsdorf, F., Nelson, R., Newnham, G., and Vastaranta, M. 2013. Status and prospects for LiDAR remote sensing of forested ecosystems. Can. J. Remote Sens. 39: S1–S5.

Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., and Gobakken, T. 2012. Lidar sampling for large-area forest characterization: A review. Remote Sens. Environ. 121: 196–209.

Zelazny, V., Martin, G.L., Toner, M., Gorman, M., Colpitts, M., Veen, H., Godin, B., McInnis, B., Steeves, C., and Roberts, M. 2008. Our landscape heritage: the story of ecological land classification in New Brunswick. NB Department of Natural Resources, Fredericton, NB, Canada.

CHAPTER 4: GENERAL DISCUSSION AND CONCLUSIONS

4.1 Summary of results

This thesis examined methods to impute tree lists for individual light detection and ranging (LiDAR) grid cells in spruce (*Picea* sp.) plantations, in order to forecast inventory variables using a locally calibrated tree-list growth model. I evaluated a method to match stand structural variables estimated from LiDAR to those in a library of over 5500 sample plot measurements to impute tree lists for LiDAR grid cells across 83 000 ha of spruce plantations. Matches were determined based on planted species and minimum sum of squared difference between total and merchantable basal area, total and merchantable volume, top height, and merchantable quadratic mean diameter. Forest inventory variables obtained by the plot matches were highly correlated (r = 0.91–0.99) with those measured on 98 validation plots. Basal area distributions derived from plot matching were statistically equivalent to those observed on the validation plots 86% of the time (α = 0.05). When the predictions for all validation plots were aggregated, there was minimal difference between predicted and actual basal area distributions by planted species and species compositions were similar. Overall, results indicated that tree-level inventory could be imputed for LiDAR grid cells using existing forest inventory plot measurements and suggested that the tree-level inventory was acceptably accurate for updating future forest inventory using a tree-list growth model.

To evaluate the accuracy of tree-level inventory projections, I then compared inventory variable increments predicted with a locally calibrated tree-list growth model (FORUS Research 2011) using tree lists imputed by plot matching with those using measured tree lists from 98 validation plots. I found that total and merchantable basal area, total and merchantable volume, Lorey's height, and quadratic mean diameter increments

were highly correlated (0.75–0.86) with percent root mean squared error ranging from 13%–49%. For validation plots with ≤ 10% LiDAR-derived error for all plot-matched variables, percent root mean squared error was much lower (5%–13%). When compared with volumes from 15 blocks harvested 3–5 years after LiDAR acquisition, average forecasted volume differed by only 1.5% and was not significantly different ($p = 0.54$). In contrast, unforecasted LiDAR-derived volume was significantly different ($p = 0.032$) from actual harvested volumes, which showed added value in forecasting LiDAR-derived inventory using the plot-matching method.

To demonstrate the novel application of this method for operational management decisions, annual commercial thinning was planned at grid-cell resolution from 2018–2020 using forecasted inventory variables and commercial thinning eligibility rules.

4.2 Application of thesis results

LiDAR-derived forest inventory has been shown to be an effective tool for operational forest management (Næsset et al. 2007; Woods et al. 2011) because of its accuracy and high spatial resolution. However, without a means to forecast it forward in time, any inventory, whether traditional or LiDAR-derived, will become biased and less useful for forest management planning with time. As demonstrated in this thesis (Chapter 3), unadjusted LiDAR-derived volumes can differ measurably from harvest volumes in spruce plantations just 3–5 years after acquisition. This underscores a need to regularly refresh an inventory to maintain its accuracy. Using the methods presented in this thesis, tree-level inventory can be imputed for LiDAR grid cells and incrementally forecasted, thus delaying complete inventory refresh. For example, the New Brunswick Department

of Energy and Resource Development aims to acquire LiDAR for the entire province on a 10-year cycle; methods such as those tested herein will play an important role in maintaining the inventory within each of these cycles.

The plot-matching method also offers some additional benefits with respect to predicting several forest-related variables. In addition to common inventory variables predicted with LiDAR (i.e., volume, basal area, mean tree height, etc.), numerous studies have demonstrated the ability to predict non-timber value measures from LiDAR point clouds; e.g., wildlife habitat (Martinuzzi et al. 2009; Melin et al. 2013), carbon stocks (Patenaude et al. 2004; García et al. 2010), and forest canopy fuels (Andersen et al. 2005; Erdody and Moskal 2010). Such variables have become increasingly important for forest management; however, each additional variable calculated directly from LiDAR incurs additional costs and possible delays if the LiDAR contractor or expert is not available. On the other hand, if tree lists have been matched to individual inventory grid cells, then calculation of these additional forest inventory measures is straightforward. In essence, any variable that can be calculated from tree-level inventory becomes readily available for computation by using the tree-list imputation methods presented in this thesis.

Although this thesis focused on the application of forecasted LiDAR-derived inventory for planning commercial thinning in spruce plantations, these methods are also applicable to planning other harvest and silviculture treatments. Pre-commercial thinning and plantation cleaning are scheduled based on stem density and mean height (Nyland 2007), and these variables can be determined at the grid-cell level and used to plan annual treatment schedules across stands and within stands. Treatments and treatment response simulations can also be conducted for each stand to optimize prescription performance.

Finally, volumes derived from matching tree lists with LiDAR cells and forecasting them will allow forest managers to estimate expected harvest block volumes that more closely match actual harvest volumes (Chapter 3).

4.3 Limitations and future work

Despite advantages associated with the plot-matching method and the accuracy of the resulting forecasted inventory variables presented in this thesis, the methods do not come without some caveats and limitations. The plot-matching approach relies on having existing plot measurements that cover the full range of conditions observed in the target area. Any gaps in the range of conditions sampled could introduce error through an inability of the algorithm to locate a proper match. LiDAR grid cells containing uncommon species or unique structure (e.g., a shelterwood stand with a single dominant tree in the cell) may result in plot matches that are not representative of actual conditions if no similar reference plots are available. In these cases, plot matching may not be suitable until additional plots are collected that represent these conditions. However, with each additional LiDAR calibration plot and temporary or permanent sample plot measurement, the plot-matching library increases and improves the chances that a more representative plot match will be found. Therefore, my methods should help to focus sampling on stand types with inadequate existing inventory plots.

The plot-matching approach currently relies on manual photo-interpreted species composition at the stand polygon level to match individual 20 × 20 m LiDAR grid cells within the polygon. This generally works well in intensively managed plantations where species composition is largely homogenous throughout the stand, but is not expected to

work without error in more heterogeneous stands, where individual grid cells may not comply with the polygon species label. It is expected that this shortcoming will be circumvented through some of the current research aimed at automated species classification of individual trees within grid cells (e.g., Holmgren et al. 2008; Yao et al. 2012). Recently, Perron and St-Onge (AWARE 2016) have classified tree species, diameter at breast height, and total height for overstory trees across the Black Brook District. These predictions are being used to estimate overstory species composition, weighted by crown area, to provide accurate species composition at the grid-cell level. This improved species resolution, presents significant opportunity for further research of the tree-list imputation method.

Finally, the accuracy of inventory variables forecasted from the tree-list growth model must be considered. Even if the tree list matched to the LiDAR cell accurately represents actual conditions within that cell, forecasting the inventory variables may be inaccurate if the growth model relationships do not accurately predict actual growth and yield. In this thesis, I used local permanent sample plot measurements in plantations on the same landbase to calibrate tree-level growth and yield relationships in OSM, in order to improve forecasted inventory estimates. If similar plot data are not available for the study area in question, the user would be forced to rely on the accuracy of a regional tree-list growth model. Given advancements in soil attribute mapping and digital elevation models using LiDAR, the biomass growth index (Hennigar et al. 2017) used in this thesis (Chapter 3) should permit improvements in growth and yield predictions. Additionally, current research by Furze and Arp (AWARE 2016) aims to use LiDAR-derived variables to better map site productivity to improve growth projections.

4.4 References

Andersen, H.E., McGaughey, R.J., and Reutebuch, S.E. 2005. Estimating forest canopy fuel parameters using LIDAR data. Remote Sens. Environ. 94(4): 441–449.

AWARE. 2016. The Point Cloud, AWARE's Electronic Newsletter. Vol. 2, No. 1. University of British Colombia. Available from < http://aware.forestry.ubc.ca/newsletter/> [cited November 25, 2017].

AWARE. 2017. The Point Cloud, AWARE's Electronic Newsletter. Vol. 3, No. 1. University of British Colombia. Available from < http://aware.forestry.ubc.ca/newsletter/> [cited November 25, 2017].

Erdody, T.L., and Moskal, L.M. 2010. Fusion of LiDAR and imagery for estimating forest canopy fuels. Remote Sens. Environ. 114(4): 725–737.

FORUS Research. 2011. FORUS Slasher User Guide. Fredericton, NB, Canada.

García, M., Riaño, D., Chuvieco, E., and Danson, F.M. 2010. Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. Remote Sens. Environ. 114(4): 816–830.

Hennigar, C., Weiskittel, A., Allen, H.L., and MacLean, D.A. 2017. Development and evaluation of a biomass increment-based index for site productivity. Can. J. For. Res. 47(3): 400–410.

Holmgren, J., Persson, Å., and Söderman, U. 2008. Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. Int. J. Remote Sens. 29(5): 1537–1552.

Martinuzzi, S., Vierling, L.A., Gould, W.A., Falkowski, M.J., Evans, J.S., Hudak, A.T., and Vierling, K.T. 2009. Mapping snags and understory shrubs for a LiDAR-based

assessment of wildlife habitat suitability. Remote Sens. Environ. 113(12): 2533–2546.

Melin, M., Packalén, P., Matala, J., Mehtätalo, L., and Pusenius, J. 2013. Assessing and modeling moose (*Alces alces*) habitats with airborne laser scanning data. Int. J. Appl. Earth Obs. Geoinf. 23(1): 389–396.

Næsset, E. 2007. Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia. Scand. J. For. Res. 22(5): 433–442.

Nyland, R.D. 2007. Silviculture: Concepts and Applications. In 2nd edition. Waveland Press, Inc., Long Grove, IL, USA.

Patenaude, G., Hill, R.A., Milne, R., Gaveau, D.L.A., Briggs, B.B.J., and Dawson, T.P. 2004. Quantifying forest above ground carbon content using LiDAR remote sensing. 93: 368–380.

Woods, M., Pitt, D., Penner, M., Lim, K., Nesbitt, D., Etheridge, D., and Treitz, P. 2011. Operational implementation of a LiDAR inventory in Boreal Ontario. For. Chron. 87(4): 512–528.

Yao, W., Krzystek, P., and Heurich, M. 2012. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. Remote Sens. Environ. 123: 368–380.

APPENDIX 1: ARTICLE COPYRIGHT PERMISSION

Chapter 2: Taylor and Francis Group

Please review: http://taylorandfrancis.com/contact/rights-and-permissions/journals/

Taylor & Francis is pleased to offer reuses of its content for a thesis or dissertation free of charge contingent on resubmission of permission request if work is published.

CURRICULUM VITAE

Candidate's full name: Sean Lamb

Universities attended:
University of New Brunswick, Fredericton, Canada (2014–2017)
Master of Science in Forestry

University of New Brunswick, Fredericton, Canada (2012–2014)
Bachelor of Science in Forestry and Environmental Management

Articles published in refereed journals:

Lamb, S.M., MacLean, D.A., Hennigar, C.R., and Pitt, D.G. 2017. Imputing tree lists for New Brunswick spruce plantations through nearest-neighbor matching of airborne laser scan and inventory plot data. Can. J. Remote Sens. 43(3): 269–285. (This thesis, Ch.2).

Selected non-refereed contributions (posters, presentations):

Lamb, S.M. 2017. LiDAR-derived forest inventory. Oral presentation at Remsoft User's Group. June 14. Fredericton, NB, Canada.

Lamb, S.M. 2017. Improving pre-harvest ground verification of commercial thinning in spruce plantations using a nearest neighbor tree-list imputation method. Oral presentation at AWARE Annual General Meeting. May 17. Edmundston, NB, Canada.

Lamb, S.M. 2017. Improving pre-harvest ground verification of commercial thinning in spruce plantations using a nearest neighbor tree-list imputation method. Oral presentation at New Brunswick Enhanced Inventory Project – Team Meeting. January 11. Fredericton, NB, Canada.

Lamb, S.M. 2016. Tree-list imputation of LiDAR cells in New Brunswick softwood plantations. Oral presentation at AWARE Annual General Meeting. May 25. Cornerbrook, NL, Canada.

Lamb, S.M. 2016. Forecasting development of planted stands in New Brunswick using LiDAR. Oral presentation at New Brunswick Enhanced Inventory Project – Team Meeting. January 19. Fredericton, NB, Canada.

Lamb, S.M. 2015. Forecasting development of planted stands in New Brunswick using LiDAR. Oral presentation at New Brunswick Enhanced Inventory Project – Team Meeting. June 18. Edmundston, NB, Canada.