Natural Resources Canada | Ressources naturelles Canada

CANADIAN FOREST SERVICE
CANADIAN WOOD FIBRE CENTRE
INFORMATION REPORT
FI-X-010

A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach

Joanne C. White, Michael A. Wulder, Andrés Varhola, Mikko Vastaranta, Nicholas C. Coops, Bruce D. Cook, Doug Pitt, and Murray Woods

Canada

The Canadian Wood Fibre Centre, Victoria, British Columbia

The Canadian Wood Fibre Centre brings together forest sector researchers to develop solutions for the Canadian forest sector's wood fibre related industries in an environmentally responsible manner. Its mission is to create innovative knowledge to expand the economic opportunities for the forest sector to benefit from Canadian wood fibre. The Canadian Wood Fibre Centre operates within the CFS, but under the umbrella of FPInnovations' Board of Directors.

FPInnovations is the world's largest private, not-for-profit forest research institute. With over 600 employees spread across Canada, FPInnovations unites the individual strengths of each of these internationally recognized forest research and development institutes into a single, greater force. For more information visit **http://www.FPInnovations.ca**

Additional information on Natural Resources Canada, the Canadian Forest Service, and Canadian Wood Fibre Centre research and publications is also available online at: **cfs.nrcan.gc.ca/regions/cwfc**. To download or order additional copies of this publication, see our online bookstore at: **cfs.nrcan.gc.ca/publications**.

A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach (Version 2.0)

Joanne C. White[1], Michael A. Wulder[1], Andrés Varhola[2], Mikko Vastaranta[1,3], Nicholas C. Coops[2], Bruce D. Cook[4], Doug Pitt[5], and Murray Woods[6]

[1]Natural Resources Canada, Canadian Forest Service,
Pacific Forestry Centre, Victoria, BC

[2]Integrated Remote Sensing Studio, Department of Forest Resources Management,
Faculty of Forestry, University of British Columbia, Vancouver, BC

[3]Department of Forest Sciences,
University of Helsinki, Helsinki, Finland

[4]NASA Goddard Space Flight Center,
Biospheric Sciences Laboratory, Greenbelt, MD, USA

[5]Natural Resources Canada, Canadian Forest Service,
Canadian Wood Fibre Centre, Sault Ste. Marie, ON

[6]Ontario Ministry of Natural Resources,
Southern Science & Information Section, North Bay, ON

Natural Resources Canada
Canadian Forest Service
Canadian Wood Fibre Centre
Information Report FI-X-010
2013

Natural Resources Canada
Canadian Forest Service
Canadian Wood Fibre Centre
506 West Burnside Road
Victoria, British Columbia
V8Z 1M5
Tel.: 250-363-0600

http://cfs.nrcan.gc.ca/regions/cwfc

Cover photos: top: D. Pitt, Canadian Wood Fibre Centre; bottom (l-r): G. Keith Douce, University of Georgia, Bugwood.org; Ronald F. Billings, Texas Forest Service; M. Woods, Ontario Ministry of Natural Resources

Printed in Canada

Library and Archives Canada Cataloguing in Publication

White, Joanne

A best practices guide for generating forest inventory attributes from airborne laser scanning data using the area-based approach / Joanne C. White, Michael A. Wulder, Andrés Varhola, Mikko Vastaranta, Nicholas C. Coops, Bruce D. Cook, Doug Pitt, Murray Woods.

(Information report ; FI-X-010)
Electronic monograph in PDF format.
Includes abstract in French.
Issued by: Canadian Wood Fibre Centre.
Issued also in print form.
Includes bibliographical references.

ISBN 978-1-100-22385-8
ISSN 1915-2264
Cat. no.: Fo148-1/10E-PDF

1. Forest surveys--Methodology. 2. Lasers in forestry. I. Canadian Forest Service II. Canadian Wood Fibre Centre III. Title. IV. Series: Information report (Canadian Wood Fibre Centre) FI-X-010

SD381.5 W54 2013 333.75'11 C2013-980063-8

Contents

Appendices

List of Tables

List of Figures

Acknowledgements

We thank the Canadian Wood Fibre Centre (CWFC) of the Canadian Forest Service, Natural Resources Canada for funding support that aided the preparation of this best practices guide. Thanks are also offered to the many staff of the CWFC and its partners who have spent countless hours supporting and advocating for the role of LiDAR in enhanced forest inventory across Canada. Special thanks to Roger Whitehead of the CWFC for his ongoing support of our work, especially for this best practices guide. Finally, thanks to the many individuals who provided useful and constructive feedback on an initial draft of this document: Chris Bater, Glenn Buckmaster, Chris Butson, Joan Luther, Daryl Price, Kim Rymer, Chris Vukovich, and Barry White. The contributions of Dr. Mikko Vastaranta were made possible by funding from the Academy of Finland ("Towards Precision Forestry").

Abstract

A best practice guide brings together state-of-the-art approaches, methods, and data to provide non-experts more detailed information about complex topics. With this guide, our goal is to inform and enable readers interested in using airborne laser scanning (ALS; also referred to as Light Detection and Ranging [LiDAR]) data to characterize, in an operational inventory context, large forest areas in a cost-effective manner. To meet this goal, we outline an approach to using ALS data that is based on (1) theoretical and technical applicability; (2) published or established heritage; (3) parsimoniousness; and (4) clarity. The best practices presented herein are based on more than 25 years of scientific research on the application of ALS data in forest inventory.

We describe the process required to generate forest inventory attributes from ALS data from start to finish, recommending best practices for each stage, from ground sampling and statistics, through to sophisticated spatial data processing and analysis. As the collection of ground plot data for model calibration and validation is a critical component of the recommended approach, we have placed appropriate emphasis on this section of the guide. Although many readers will not have the capacity—or need—to undertake all of the stages of this process themselves, we feel it is important for all readers to have some understanding of the various stages of the process. Such an understanding is necessary to make informed decisions when determining whether ALS is an appropriate data choice for a forest management area. Moreover, a minimum level of knowledge is useful when outsourcing or establishing collaborations for data acquisition, processing, or analysis, and when evaluating deliverables. To this end, we also provide some background information on ALS.

Keywords: airborne laser scanning, ALS, area-based approach, best practices, digital surface model, forest inventory, ground plot data, LiDAR, mapping, modelling, point cloud metrics.

Résumé

Ce guide de pratiques exemplaires comprend des approches, des méthodes et des données de pointe afin de fournir aux profanes des renseignements précis sur des enjeux complexes. Grâce à ce guide, nous espérons informer les lecteurs qui souhaitent utiliser des données par balayage laser aéroporté (BLA), aussi appelé détection et télémétrie par ondes lumineuses (LiDAR), et leur fournir les moyens de les utiliser en vue d'établir, tout en réduisant le coût, les caractéristiques de vastes zones forestières dans un contexte d'inventaire d'exploitation. Pour atteindre cet objectif, nous démontrons une façon d'utiliser des données par BLA fondées sur (1) les applications théoriques et techniques; (2) le patrimoine publié ou établi; (3) la parcimonie; (4) la clarté. Les pratiques exemplaires qui figurent dans ce guide sont fondées sur plus de 25 ans de recherches scientifiques sur la mise en application des données par BLA dans le contexte de l'inventaire forestier.

Dans ce guide, nous donnons un aperçu du processus requis pour générer, du début à la fin, des attributs d'inventaire forestier à partir de données par BLA, en recommandant des pratiques exemplaires à chaque étape, de l'échantillonnage du sol à la démarche statistique en passant par le traitement et l'analyse de données spatiales de pointe. Puisque la collecte de données sur les placettes au sol pour l'étalonnage et la validation du modèle est un élément essentiel à la démarche recommandée, cette partie du guide y accorde une attention particulière. Même si de nombreux lecteurs ne pourront suivre toutes les étapes de ce processus (ou n'auront pas besoin de le faire), nous considérons qu'il est important que tous les lecteurs comprennent les différentes étapes du processus. Une telle compréhension est essentielle à la prise de décisions éclairées au moment d'établir si le BLA constitue le bon choix quant à la collecte de données sur une zone d'aménagement forestier. De plus, il est utile de posséder un minimum de connaissances au moment de passer des marchés ou de signer des ententes de collaboration sur la collecte, le traitement ou l'analyse de données, et au moment de l'évaluation des produits livrables. Par conséquent, nous fournissons également certains renseignements de base sur le BLA.

Mots-clés : balayage laser aéroporté, BLA, démarche fondée sur la zone, pratiques exemplaires, modèle numérique de surface, inventaire forestier, données sur les placettes au sol, LiDAR, cartographie, modélisation, mesures en point de nuage.

1. Introduction

Interest continues to grow amongst a broad range of Canadian stakeholders, including provincial forest agencies, commercial forest companies, and forestry consultants, in the benefits of integrating data acquired from Airborne Laser Scanning (ALS)[1] into their business practices. Nevertheless, the capacity of these stakeholders for uptake of ALS-derived information in forest inventory applications is variable. In some jurisdictions in Canada, ALS data are acquired according to standardized specifications, sometimes over large areas, and in some cases via centralized, co-ordinated collection. Although these data may not have been acquired specifically for forest inventory applications, an opportunity exists for these data to provide useful information for forest inventory and management. Indeed, the successful integration of ALS data into the operational forest inventories of several forest management areas across the country has increased interest in acquiring ALS data specifically to support the development of enhanced forest inventories. The purpose of this guide is to increase the capacity for uptake of ALS data in operational forest inventory applications by synthesizing best practices for ALS acquisition, processing, analysis, and modelling.

Information needs for forest management are considered to be independent of the data sources used, providing the necessary information can be produced according to the required specifications for accuracy and precision and at an acceptable cost. The capability of ALS data to provide useful information for forest inventory has been demonstrated by more than 25 years of scientific research and through operational applications of ALS in other jurisdictions, primarily in Scandinavian countries (Næsset et al. 2004).

Once ALS data are acquired to an appropriate specification and processed to the point where the data can be used for analysis, a series of steps is required to transform measurements from the ALS point cloud to estimates of forest inventory attributes, such as mean tree height, basal area, and volume. For example, a modelling approach must be selected, ground plot measures collected, and descriptive statistics generated from the ALS data. As one might expect, several decision points are associated with this process, and different decisions are capable of affecting the quality, consistency, and comparability of the outcomes. In this guide, we take the reader through each step in the process, providing background and references from the scientific literature, or drawing on relevant operational experiences from Canada and other jurisdictions.

Our intention is to offer a rational and transparent approach for producing forest inventory attributes from ALS data. As it is not possible to address every possible contingency in a guide such as this, we hope readers will benefit from the suggestions offered and derive a sufficient understanding to know when our recommendation may not fit with their particular circumstances and to make appropriate adjustments. Often, what is considered as best practice is related to the particular context, questions of interest, or experience of the practitioner. In cases when optional approaches exist, we will identify these, indicate the nature of the differences, and then follow with a focus on a particular approach. Readers can then consider the options and use the approach best suited to their particular circumstances.

The use of ALS data in support of forest inventories is increasingly well understood by a growing community of users: no longer is this technology restricted to a small group of experts. Part of the reason that the community of users has increased is the availability of acquisition standards, tools, and analytical approaches, as well as an increasing number of ALS acquisitions. Most users no longer need to tackle the entire series of activities from acquisition through to the prediction of attributes. Commercial agents are often contracted to accomplish the initial ALS data collection, point classification, and surface generation. Acquisition standards promote transparent and consistent data collection and processing. Although the actual survey specifications will likely vary with the environment characterized, users now know what to ask for thematically and how to state this in a Request for Proposal (see also Appendix 1).

Basic background on ALS is provided in Section 1.2; however, many useful references provide a more detailed introduction to ALS and its application in a forestry context and users are encouraged to review these sources for more background information (e.g., Lim et al. 2003b; Reutebuch et al. 2005; Evans et al. 2006; Wulder et al. 2008; Hyyppä et al. 2008). The recommendations made in this guide are based on the application of an area-based approach for estimation of forest inventory attributes (Næsset 2002), which is described in detail in Section 2. In summary, the area-based approach involves the collection of ground measurements that are then linked, via modelling, with statistical and spatial generalizations of ALS data to enable area-wide estimation and mapping of forest inventory attributes (Figure 1). Most forest applications that use remotely sensed data to estimate an attribute require ground measurements to produce accurate and reliable

[1] Also referred to as Light Detection and Ranging, or LiDAR.

results (Wulder 1998). Having ground and remotely sensed data allows for the development of predictive relationships, with the ground measures acting as dependent variables. The number and quality of ground plot measurements made is a critical element of the process described in this guide, and issues associated with ground plot measurements are therefore given considerable emphasis. The generalization of classified (i.e., ground, non-ground) ALS point clouds to metrics enables the characterization of forest structure and the development of predictive models using the co-located ground measurements. The process of generating ALS metrics is described, as are options for building predictive models. Appendices to this best practices guide provide additional details regarding elements to include in Requests for Proposals and metric workflows.

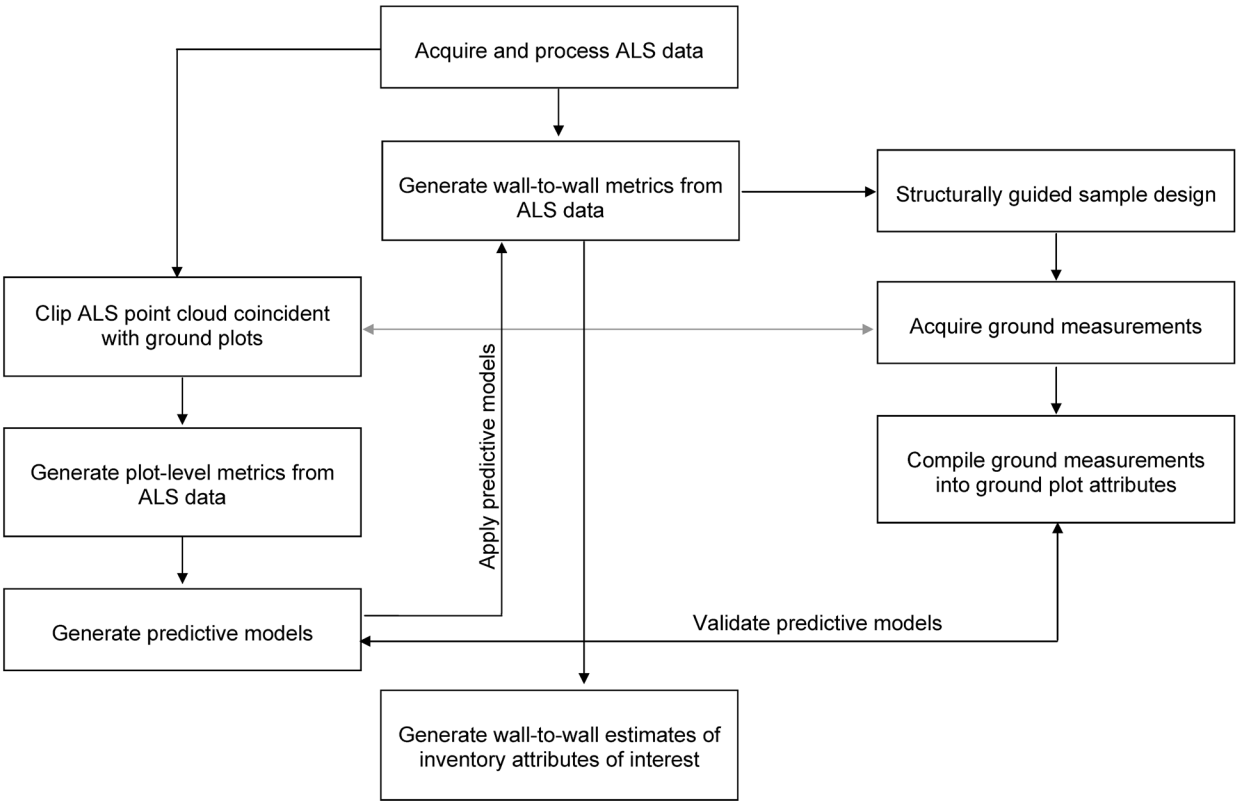

Figure 1. Overview of the approach used to generate estimates of forest inventory attributes using ALS data and ground plot measurements.

1.1 An Example Operational Context

Rather than propose entirely hypothetical situations, we have chosen to use real operational examples where possible. The experience in Alberta is informative: several of the elements that we focus on can be related to the experiences and lessons learned in Alberta. To date, over 28 million ha of ALS data have been purchased by the Government of Alberta. The forestry context in Alberta mirrors that of other jurisdictions in many ways, and certainly provides a case study for what could unfold in other jurisdictions in Canada or elsewhere. In Alberta, data have been collected to a defined specification by commercial vendors according to data acquisition and processing standards developed by the government. The government then licenses the data according to the following conditions:

> *Access is restricted to the Government of Alberta, its agencies, corporations and boards together with any contractors or sub-contractors working for the Licensee, for any purpose, or to any forestry company operating in the area covered by the Data and/or their contractors, solely for purposes related to activities intended to minimize damage related to the mountain pine beetle, providing such agencies, boards or corporations, forest companies or their contractors or sub-contractors agree in writing to*

> *handle, distribute and store the said Data in accordance with the terms of this Agreement* (Resource Information Management Branch 2011:26).

The acquisition specifications for Alberta's data were for a pulse density of 1.2 pulses or more per square metre, a scan angle < 25°, and a sidelap between flight paths of 50% or more. Accuracy requirements were for absolute vertical accuracies of 30 cm or less and horizontal accuracies of 45 cm or less. Deliverables included a digital elevation model (DEM) and a digital surface model (DSM), both with a 1-metre resolution, a classified point cloud, image intensity files, associated metadata, and various other calibration reports and supporting documentation (for more details, see Appendix 1). It is common for the Government of Alberta to distribute just the DEM and DSM; however, as discussed in Section 3.1, we recommend the use of the point cloud for the area-based approach.

The extensive spatial coverage of ALS data in Alberta has been used in a Wet Areas Mapping process, providing an information product that details local flow patterns, soil drainage, and soil moisture regimes (White et al. 2012). Such information is valuable for operational forest planning, as well as numerous other natural resource applications. It is worth noting that when the data acquisition specifications were initially developed for Alberta, the pulse rates of commercial laser systems were lower than those that are currently available. The minimum pulse density of 1.2 pulses per square metre was in keeping with the technology available at that time, and is appropriate for the area-based approach described here (in the forest environments of Alberta) (Treitz et al. 2012; Jakubowski et al. 2013). The widespread availability of ALS data in Alberta provides a significant opportunity for the development of enhanced forest inventories in the province. As a result, experiences and lessons learned in Alberta regarding large-area acquisitions and standards can aid in informing activities in other jurisdictions in Canada.

1.2 Background

LiDAR is an active remote sensing technology that uses the time-of-flight measurement principle to measure the range or distance to an object. A LiDAR sensor emits a laser pulse and measures the time it takes for the energy from that pulse to be reflected (returned or backscattered) to the instrument. The measure of time is then converted to a distance, or a range using the following equation:

$$\text{Range (m)} = (\text{Speed of Light} \times \text{Time of Flight}) \div 2$$

With the known position of the sensor and precise orientation of the range measurements between the sensor and the intercepting object, the position (*x, y, z*) of the object is defined. The principle of LiDAR measurements is the same regardless of the platform. In forest inventory mapping, the most commonly used platforms are fixed-wing aircraft or helicopters. Although these two airborne platforms have many similarities, helicopters are capable of flying lower and slower than fixed-wing aircraft, and can follow complex terrain; however, helicopters also have higher operating costs, which result in higher survey costs. In some areas (e.g., coastal forests of British Columbia), steep, variable terrain may require an aircraft to fly at an altitude that negatively affects survey parameters (e.g., pulse density, swath overlap). In those situations, a helicopter platform may be preferable and/or necessary to achieve the desired acquisition specifications.

An ALS system is a package of instrumentation that includes a laser ranging unit; an opto-mechanical scanner; control, monitoring, and recording units; a kinematic global positioning system (GPS) receiver; and an inertial measurement unit (IMU) (Wehr and Lohr 1999) (Figure 2). The opto-mechanical scanner moves the laser across the flight path, within a user-specified angle. The GPS receiver is critical for accurately measuring the position of the platform (aircraft or helicopter), while the IMU measures the dynamic attitude (i.e., roll, pitch, and yaw) of the platform. The GPS and IMU provide the information necessary to accurately identify the location where the laser pulse intercepted an object. The ALS systems used for forest inventory purposes typically emit very short (3–10 ns), narrow-beam width (0.15–2.0 mrad), infrared (0.80–1.55 µm) laser pulses at near-nadir incidence angles (< 25°) with high pulse repetition frequencies (50–200 kHz). In general, when operated at flying altitudes of around 500–3000 m, ALS systems generate a dense sample pattern (0.5–20 pulses per square metre) with a small ground footprint (< 1 m).

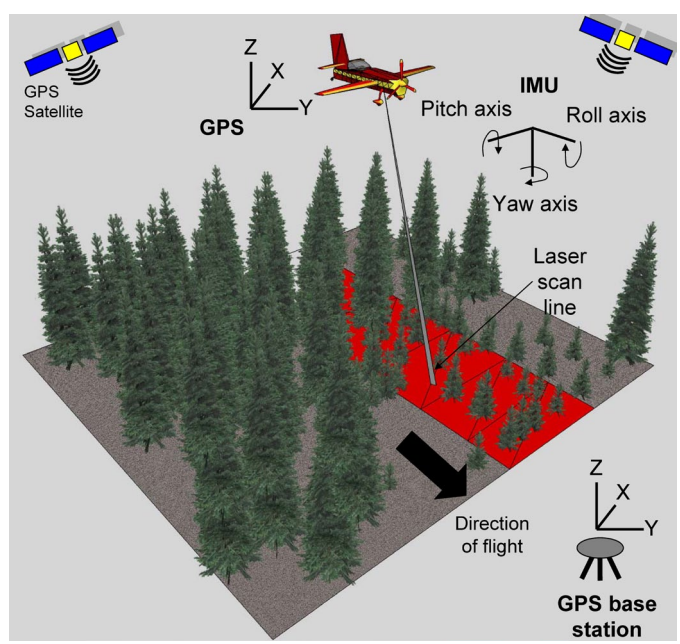

Figure 2. Schematic of an ALS system.

LiDAR instruments used in ALS systems are characterized as either full waveform or discrete return. Full waveform systems record the reflected or backscattered energy from each laser pulse as a single, continuous signal. Discrete return systems convert waveform data into return targets referenced in time and space. Discrete return systems are more commonly used in forest inventory applications and current instruments are capable of recording up to five returns for each laser pulse that is emitted. In the simplest case, when a laser pulse intercepts an object it cannot penetrate, such as a building, the ground, or a very dense forest canopy, only a single return of energy to the instrument will occur. In contrast, when the laser pulse intercepts an object through which it can penetrate, such as the forest canopy, some of the energy will be returned to the instrument (first return), and some will continue through the canopy and intercept stems, branches, and leaves before reaching the ground. This series of events may result in the recording of several returns for a single laser pulse, which are referred to as "multiple returns." First returns are assumed to come primarily from the top of the canopy, while last returns are assumed to originate from the ground or objects near the ground. Multiple returns produce useful information regarding forest vertical structure (Hyyppä et al. 2008).

The trunks, branches, and leaves of dense vegetation tend to cause multiple scattering or absorption of the emitted laser energy so that fewer backscattered returns are reflected directly from the ground (Harding et al. 2001; Hofton et al. 2002), and with fewer ground returns, it is more difficult to generate an accurate DEM. This effect increases when the canopy closure, canopy depth, and structural complexity increase because the laser pulse is greatly obscured by the canopy. In practice, the laser system specification and configurations also play an important role in how the laser pulse interacts with the forest canopy. For example, research shows that:

- small-footprint lasers tend to penetrate into the canopy before reflecting a signal that is strong enough to be recorded as a first return (Gaveau and Hill 2003);
- ground returns decrease as the scanning angle increases (Lovell et al. 2005; Disney et al. 2010);
- the penetration rate is affected by the laser beam divergence (Aldred and Bonner 1985; Næsset 2004);
- a higher flight altitude alters the distribution of laser returns from the top and within the tree canopies (Næsset 2004); and
- the distribution of laser returns through the canopy varies with pulse energy and the instrument's ability to detect and record multiple returns for a single laser pulse (Chasmer et al. 2006).

Furthermore, the sensitivity of the laser receiver, wavelength, laser power, and total backscattering energy from the tree tops are also factors that may influence the ability of laser pulses to penetrate and distribute laser returns from the forest canopy (Baltsavias 1999).

In post-processing of the data acquired by an ALS system, information from the ALS instrument, GPS, and IMU are brought together to create a precise, georeferenced, three-dimensional (*x*, *y*, *z*) location for each return into a single file referred to as a "point cloud." The point cloud is then processed to identify, at a minimum, ground and non-ground returns, followed by the generation of an accurate DEM from the classified ground returns and a DSM from the non-ground first returns (Figure 3). While a DEM represents heights of the ground surface relative to some reference (e.g., sea level), a DSM represents heights of objects above the ground surface, relative to the same reference. A Canopy Height Model (CHM) represents the height of the canopy above ground level, and is generated by subtracting the DEM from the DSM. In the same way, a DEM can be used to normalize the ALS point cloud to heights above ground level.

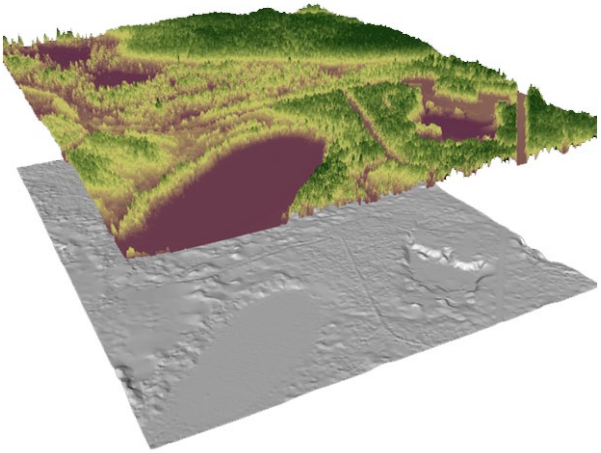

Figure 3. Basic products generated from the ALS point cloud: a Digital Elevation Model (DEM; shown as a grey hillshade) represents ground elevations relative to some reference, such as sea level; a Digital Surface Model (DSM; not shown) represents heights of objects above the ground surface, also relative to some reference; a Canopy Height Model (CHM; shown in colour) represents the normalized above-ground heights of non-ground objects. The CHM is generated by subtracting the DEM from the DSM.

2. Area-based Approach to Attribute Estimation

The area-based prediction of forest inventory attributes is based on a statistical dependency between predictor variables derived from ALS data and response variables measured from ground plots. An overview of the area-based approach is provided in Figure 1. The successful application of the area-based approach is predicated on accurate measurements of forest height and height variation from ALS data. The goal of the area-based approach is to generate wall-to wall estimates and maps of inventory attributes such as basal area or volume (Næsset 2002).

The area-based approach is accomplished in two stages. In the first stage, ALS data is acquired for the entire area of interest, tree-level measures are acquired from sampled ground plots, and predictive models are developed (e.g., regression or non-parametric methods). For the purposes of model development, the ALS point cloud is clipped to correspond to the area of each ground plot. Metrics (descriptive statistics) are calculated from the clipped normalized ALS point cloud and include measures such as mean height, standard deviation of height, height percentiles, and canopy cover (see Section 4.4 for a list of possible metrics). Attributes of interest are measured by ground crews (i.e., height, diameter) or modelled (i.e., volume, biomass) for each ground plot. Ground plots should represent the entire population and cover the full range of variability in the attributes of interest, which generally will require some form of stratified sampling approach, preferably with strata defined from the ALS metrics. Predictive models are then constructed using the ground plot attributes as the response variable and the ALS-derived metrics as predictors.

In the second stage of the area-based approach, the models are applied to the entire area of interest to generate the desired wall-to-wall estimates and maps of specific forest inventory attributes. The same metrics that are calculated for the clipped ALS point cloud (as described above) are generated for the wall-to-wall ALS data. The predictive equations developed from the modelling in the first step are then applied to the entire area of interest using the wall-to-wall metrics. The sample unit is a grid cell, the size of which relates to the size of the ground-measured plot (see Section 4.2). Once the predictive equations are applied, each grid cell will have an estimate for the attribute of interest (Figure 4).

The foremost advantages of the area-based approach compared with traditional stand-level forest inventories include having complete spatial knowledge of *X* and predicted *Y*, more precise predictions of certain forest variables, and the capability to calculate confidence intervals for estimates (e.g., Woods et al. 2011). In principle, ALS-based forest inventories do not depend on subjective, photo-interpreted stand boundaries; however, in practice, ALS estimates have typically been rolled-up to the stand level and integrated into

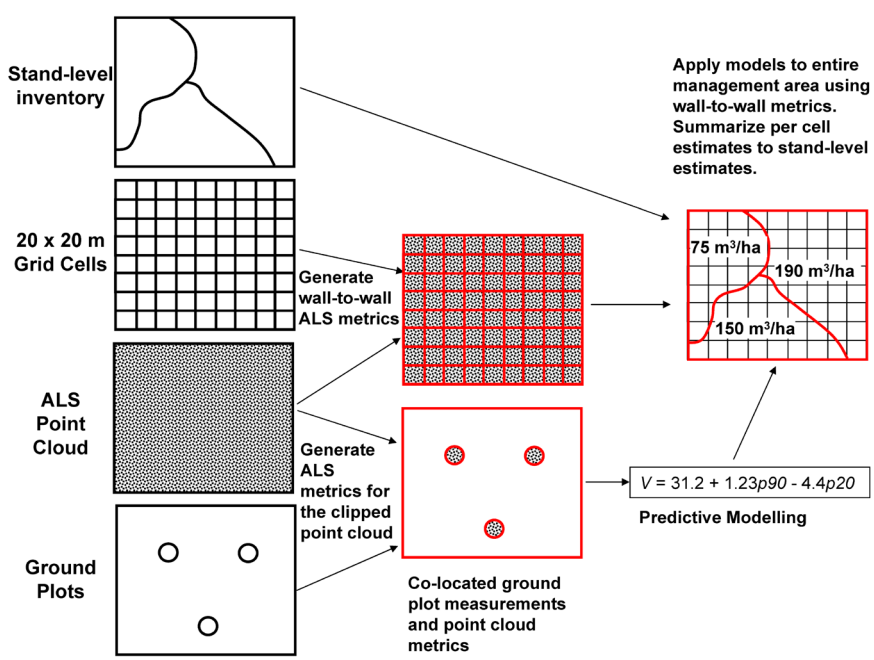

Figure 4. Schematic of the area-based approach.

existing inventory data. Forest attributes, such as biomass, stem volume, basal area, mean diameter, mean height, dominant height, and stem number, are predicted with better or comparable quality to traditional field inventories (e.g., Næsset et al. 2004). Forest management planning often requires species-specific information. Airborne laser scanning data with a resolution of less than 1 pulse per square metre does not provide much information regarding tree species composition. Optical data, such as aerial photography, can be used in addition to the ALS data to improve estimation of species-specific forest characteristics (Packalén and Maltamo 2007). In practice, species information (when required) is often derived from the existing stand-level forest inventory information.

3. Airborne Laser Scanning Data

In some cases, such as the Alberta context, ALS data may have already been acquired. In other cases, acquiring new ALS data for a forest management area may be necessary. In this section, we address both of these scenarios. We detail the basic ALS data products required to support the area-based approach and provide some recommended minimum survey specifications for data acquisition. These specifications will also provide a benchmark against which readers can compare the properties of pre-existing ALS data. Lastly, we provide some background information on the influence of ALS instrumentation over data acquisition.

3.1 Products

The minimum ALS products required for the area-based approach are the bare earth DEM and the classified (unfiltered) ALS point cloud, ideally delivered in standard LASer (.LAS) file format (American Society for Photogrammetry and Remote Sensing 2011). Currently, the Government of Alberta supplies end users with two products generated from the classified ALS point cloud: a DEM and a DSM (Section 1.2). Recall that the DSM is generated primarily from non-ground first returns, and although it may be useful for assessing how average tree height varies across a stand, it does not provide any information on subcanopy vertical forest structure. Airborne laser scanning pulses are able to penetrate the forest canopy and acquire additional returns at the subcanopy level, representing branches and understorey structure. This potentially rich source of information on vertical forest structure is lost if the end user is not provided with the full ALS point cloud. To obtain the most accurate results possible, all returns from each pulse are required. For this reason, we recommend that for forestry applications, end users should be provided with the bare earth DEM and the classified, unfiltered ALS point cloud. As noted by Gatziolis et al. (2010), the provision of the unfiltered point cloud enables options for future applications, wherein the point cloud may need to be reprocessed (i.e., to accommodate advances in ALS theory and data-processing techniques).

3.2 Data Quality Assessment

Forestry end users of ALS data will want to confirm the integrity of the data received before embarking on any analysis. Standard quality control and quality assurance procedures for ALS data are typically the responsibility of the data provider; data quality reports are often provided as contract deliverables and these reports should be made available to end users. Assuming the basic quality of the ALS data has been assessed, end users will want to confirm the following details.

- LiDAR instrument(s) used (could be more than one)
- Acquisition parameters (and documentation) (e.g., date[s], altitude)
- Completeness of trajectory data
- Environmental conditions during acquisition (specifically fog and precipitation)
- Processing methods (and documentation), including software and specific procedures followed
- Projection/datum information
- Spatial coverage of the LAS files and DEMs provided (i.e., complete spatial coverage provided, no gaps in acquisition)
- Adherence to fundamental, supplemental, and consolidated vertical accuracy requirements (Flood [editor] 2004)
- Content of the LAS files provided (i.e., are the returns classified appropriately and consistently? Is scan angle provided?)
- Reported pulse density
- Range of values in the LAS file (i.e., are there outliers?)

Different software tools are available to enable these basic quality assurance functions. FUSION, a freeware tool developed by the US Forest Service (McGaughey 2013), provides a useful function (*catalog*) to verify the quality of the point cloud files, the completeness of the spatial data coverage, and the return density.

Useful outputs from FUSION's *catalog* function include the following:

- An image file that shows the nominal coverage area for all data files included in the catalogue. Tiles that may contain outliers (e.g., the minimum, maximum, or

range of elevations are outside the range defined by the mean elevation ± 2 standard deviations) can also be flagged in the output image.

- A spreadsheet with each of its rows representing a separate LAS file in the area of interest. Columns report the extent of the tile, the minimum and maximum elevations, the number of returns by return type, the total number of returns, and the nominal return density.

3.3 Airborne Laser Scanning Data Acquisition Specifications

Depending on the information need, the design of an ALS survey for forest applications involves many trade-offs. Ultimately, these surveys are intended to provide data for a broad range of forest applications and, therefore, should be designed accordingly. Table 1 outlines the recommended specifications for several key acquisition parameters. Of particular importance is the requirement for more than 50% overlap between ALS swaths (Evans et al. 2006).

Pulse density is a function of pulse rate, instrument energy, receiver sensitivity, flying height and speed, and scan angles, among other considerations. Note that since the specification and initial collection of ALS data in Alberta, improved opportunities now exist for increased pulse rates and densities of ALS data. Since 2000, pulse rates have increased from approximately 10 000 pulses per second to the current capacity of 500 000 pulses per second (e.g., Optech 2013). Increased pulse rates allow data vendors to fly aircraft at higher altitudes and faster speeds to obtain data within a specified pulse density target (Laes et al. 2008). This ability to fly higher while still acquiring data within a target density can significantly reduce a vendor's acquisition costs; however, it can also increase footprint size and potentially reduce the number of returns recorded per pulse, neither of which are desirable for end users interested in characterizing forest structure. Although Table 1 indicates that a minimum of 1 pulse per square metre is necessary to characterize plot and stand-level models, greater pulse densities can improve the generation of bare earth DEMs, the precision of attribute estimates, and support individual tree work and future ALS applications (Jakubowski et al. 2013). We therefore recommend that ALS acquisition specifications be reviewed periodically to ensure currency with advances in technology. A data vendor should never have to discard data to meet an underspecified requirement.

Table 1. A summary of recommended ALS acquisition specifications for forestry applications (adapted from Reutebuch and McGaughey 2008).

Acquisition Parameter	Recommended Specification for Forestry Applications
Laser beam divergence	Narrow (e.g., 0.3 mrad; with "narrow" typically considered as 0.1–0.6 mrad). Influences the footprint size of the laser pulse on the ground. For example, a laser at an altitude of 1000 m with a beam divergence of 0.3 mrad will have a footprint that is approximately 30 cm in diameter. Thus, both beam divergence and flying altitude will influence footprint size.
Scan angle	< ± 12° (forest density can be used to guide scan angle, with more open canopies allowing for a greater scan angle).
Pulse repetition frequency	50 kHz to > 150 kHz (newer systems offer greater pulse repetition frequencies (e.g., 400 kHz)).
Pulse density per square metre	Research indicates that the point density required to support the area-based approach to forest attribute estimation can be as low as 0.5 pulses per square metre in some forest environments (Treitz et al. 2012). Other research shows that correlations between metrics such as tree height or total basal area are relatively unaffected by pulse density *until* pulse density drops below 1 pulse per square metre (Jakubowski et al. 2013). As a heuristic, consider a minimum of 1 pulse per square metre for stand-level canopy models and medium resolution DEMs (2 m). If interested in individual tree-canopy measures, greater pulse rates are required (i.e., > 4) with the size of the crown a primary consideration. Production of a high-resolution DEM under a dense canopy also indicates a need for a greater pulse density (i.e., > 4), regardless of terrain. Even greater pulse densities will be necessary in more complex terrain. Greater point densities enable an improved description of the forest canopy, increase the likelihood of obtaining ground returns in forested areas, and increase confidence in identifying ground returns in forested areas.
Returns per pulse	A sensor capable of returning a *minimum* of two returns per pulse for canopy and ground-surface measurements (first and last return), but four returns per pulse is well within the capacity of current LiDAR sensors.
Swath overlap	> 50% sidelap on adjoining swaths to prevent data gaps between swaths. Overlapping swaths enable higher pulse densities and multiple look angles, both of which increase the likelihood of ground returns in dense forest canopy.

Users should be mindful that ALS data collected by different vendors, different instruments, or at different times of the year—even with the same set of specifications—can still vary and may therefore require independent ground plots for model development. Reasons for the variation include different levels of instrument power (which determines the depth that ALS pulses can penetrate into the canopy and return sufficient energy for detection), the sensitivity of the instrument to return detection, the software and specific procedures used to post-process the ALS point cloud (ground/non-ground classification is a critical first stage that is often completed using proprietary algorithms), and forest conditions (i.e., leaf-on or leaf-off conditions) (Næsset 2009).

One of the implications of having multiple ALS surveys within an area of interest (and therefore the potential for varying return densities and characteristics) is that predictive relationships developed using ground measures and co-located ALS metrics may not be applicable over an entire area but only over smaller areas corresponding to specific ALS surveys. One way to address this is to treat areas with different surveys as distinct units. Within each unit, ground plots could be established, ALS metrics generated, and predictive models, unique to the area, developed. The necessity of this area-specific approach will depend on the nature and magnitude of the differences between the different ALS surveys.

Ideally, ALS data will be acquired during the growing season and with leaf-on conditions, which varies by region but for most areas in Canada will be June through September; however, research shows that the estimation accuracy for forest attributes in mixed forest stands is unaffected by ALS data acquired during leaf-off conditions (Næsset 2005). Moreover, last returns are known to be affected more by leaf-off conditions than first returns, and canopy height measures have greater variability when estimated from leaf-off data (Næsset 2005). The suitability of leaf-off data for the area-based approach was also assessed and confirmed by Villikka et al. (2012), who further concluded that leaf-on and leaf-off data should not be combined in the area-based approach as it can lead to serious bias. Leaf-off ALS data consistently underestimated plot height for certain deciduous stands, but did not result in statistically significant differences for coniferous or mixed stands (Wasser et al. 2013). Differences in canopy penetration depth between leaf-on and leaf-off ALS data have been documented and will likely influence the accuracy of within-canopy vegetation structure characterization (e.g., canopy base height, understorey) (Hill and Broughton 2009; Ørka et al. 2010; Wasser et al. 2013). To summarize, we recommend the acquisition of ALS data during leaf-on conditions. Although use of leaf-off ALS data for the area-based approach may be acceptable, if a mix of leaf-on and leaf-off acquisitions exists for the same management area, separate models should be developed for each. Leaf-on models should not be applied to leaf-off data and vice versa.

Laes et al. (2008) summarized some of the potential cost savings for ALS data acquisitions. We have reproduced their information, with some modification in Table 2, as it provides an excellent synopsis of the consequences of changes to certain acquisition parameters.

3.4 Airborne Laser Scanning Instrument Considerations

Airborne laser scanning system selection and the various acquisition settings that can be utilized will affect the data collected. We advocate following a clear and consistent specification (as detailed in the previous section); however, specifying the actual sensor to be used may not be feasible. Use of the same sensor in different years could yield different results, owing to variations in instrument power or detection characteristics. An awareness of how different ALS systems or acquisition settings can affect a survey will enable the identification of issues having potential deleterious effects on survey outcomes. In addition, the rapid advances in ALS sensors over the last decade demands that researchers and practitioners alike remain up-to-date. Here we highlight some topics and considerations related specifically to ALS systems and instrumentation.

- GPS-Inertial Navigation System (INS): While we often focus on the ALS instrument, the quality and processing of the GPS-INS unit is equally important. Position, roll, pitch, and heading accuracy will tell you something about geo-location errors of the individual returns. Applanex POS-AV™ is a hardware and software system for direct georeferencing of airborne sensor data that is considered by many as the "gold standard."
- Scanning mechanism: The scanning mechanism determines how the laser pulses will be distributed on the ground. Scanning mirrors create a zigzag pattern on the ground, resulting in higher pulse densities at the point where the mirror slows down and changes direction. Rotating multi-faceted mirrors can provide more evenly spaced data on the ground.
- Pulse width: It is desirable to have a relatively narrow, along-beam pulse, which permits ranging with greater sensitivity. Unfortunately, the pulse width of some lasers (mostly older-generation lasers, and not so much for doped fibre-optic lasers) can be variable as a function of pulse repetition frequency and other environmental factors.
- Laser wavelength: This is a bit harder to assess, because trade-offs exist between laser wavelength, output energy, maximum ranging distance, and eye safety.

Table 2. Potential approaches to achieving cost savings for ALS data acquisitions and associated consequences (adapted from Laes et al. 2008).

Action	Consequences
Higher aircraft altitude	Produces a wider swath width and increases the footprint size at or near the surface. A wide footprint is less likely to penetrate dense vegetation, which can result in a less-accurate bare earth DEM. Also, in dissected terrain, pulses in deep valley bottoms often produce fewer returns due to atmospheric attenuation of the pulse energy.
Flying faster	Results in fewer pulses per nominal ground surface when flying height is constant. Projects applying the area-based approach for forest attribute estimation need a pulse density that is at least 1 pulse per square metre and may require a higher pulse density, depending on the complexity of the forest environment (see Table 1).
Increasing the scan angle	Provides a wider data swath, but along the edges of the data, fewer points will reach the ground in densely covered terrain. In open terrain, a wider scan angle has less effect on the number of ground points.
Less sidelap	More than 50% sidelap is recommended between swaths. Reducing sidelap to 30% may still result in sufficient coverage between adjacent flight lines to ensure no data gaps occur between swaths; however, it may affect pulse density and the likelihood of ground returns (see Table 1).
Accuracy	Emphasizing relative accuracy over absolute accuracy is crucial. For many applications, it may be more important to calibrate data from adjacent flight lines against each other (swath-to-swath matching) than verify that the *x*, *y*, *z* attributes of the return are within a certain specified range of real-world *x*, *y*, *z* co-ordinates. This is especially true for data that will be used in tabulated applications or statistical models; however, keep in mind the need for field verification when required.
Acquisition window	Provide the vendor as large a temporal window for acquisition as possible without compromising data needs (e.g., leaf-on conditions). A large acquisition window allows the vendor the opportunity to share aircraft mobilization, ferrying, and base-station setup costs across multiple projects.

Most commercial sensors operate at 1064 nm. While not common in the commercial domain, 1550 nm lasers have recently found favour as this wavelength is eye-safe at any distance (at much higher power levels than other wavelengths), thereby simplifying approval by the US Federal Aviation Authority. The development of multi-wavelength LiDARs specifically for vegetation purposes remains nascent. Gaulton et al. (2013) reported on a novel dual-wavelength system that is sensitive to vegetation moisture content. Vahukonen et al. (2013) used a new hyperspectral ALS to distinguish between spruce and pine trees.

- Footprint size: This is a function of beam divergence and flying altitude. In most cases, small footprints (i.e., < 20 cm) provide more information on canopy gaps and have greater ranging accuracy.
- Scan angle: Scanners with a wide field of view are acceptable for DEM and CHM development, but some ALS metrics are sensitive to off-nadir scanning data (Holmgren et al. 2003).
- Peak detection: Older, discrete-return ALS instruments often used analog peak thresholding techniques to produce binary output corresponding with peak height. Newer systems can perform onboard waveform processing. These instruments use fast analogue-to-digital electronics to convert analogue data to digital waveforms, which can be analyzed with more sophisticated signal processing algorithms (i.e., those that consider the area and shape of individual peaks) and produce more meaningful returns than thresholding techniques.
- Calibration: Some vendors (e.g., Riegl) now offer calibrated instruments for the output of reflectance products. Interpretation of these data requires caution, owing to cleanliness of the optics, atmospheric attenuation, partial interception of the beam, and optical geometry (i.e., target/scan angle and bidirectional reflectance distribution function). This is an important first step towards multi-wavelength ALS processing.

4. Generation of Airborne Laser Scanning Point Cloud Metrics

The ALS point cloud contains measurements in three-dimensional space (*x*, *y*, *z*), and descriptive statistics can be generated from these measurements to summarize the point cloud in a statistically and spatially meaningful way. These statistics are canopy height and density metrics that characterize the vertical forest conditions present in the canopy and include measures such as mean height, the 75th percentile of height, and the coefficient of variation of height. Note that the height values in the point cloud must be normalized to above-ground heights using the ALS-derived DEM prior to metric calculation.[2] The following section details the process of generating metrics from the ALS data (see also Appendix 2).

4.1 Software

Numerous software tools are available for manipulating ALS point clouds (e.g., LAStools: www.rapidlasso.com; Boise Center Aerospace Laboratory ALS tools: http://bcal.geology.isu.edu/tools/lidar) and for generating ALS point cloud metrics. Some users develop their own tools for calculating metrics in software packages such as R (R Core Team 2012) and while this customized approach provides the most flexibility, it does require specialized expertise. FUSION is a free software package for ALS data analysis and visualization developed by Robert McGaughey at the United States Forest Service's Pacific Northwest Research Station (McGaughey 2013; see also Appendix 3). FUSION was designed to facilitate the processing and extraction of information from large and (at times) unwieldy ALS data sets. End users are advised to consider the advantages and disadvantages of available tools in light of their particular application and information need. The advantages of FUSION are that it is relatively mature and stable software, and is free and easily accessible. Moreover, it is domain-specific (it was developed by a forester for forestry applications), and is maintained and periodically updated (http://forsys.cfr.washington.edu/fusion/fusionlatest.html).

4.2 Grid Cell Size

To enable the application of the area-based approach described in Section 2, an appropriate grid cell size must be selected. The selection of a grid cell size is determined primarily by the size of the ground plot (Section 5.1.1), as it is important that the area of the ground plot and the grid cell be as similar as possible (Magnussen and Boudewyn 1998; Næsset 2002). It may be easiest to demonstrate the importance of grid cell size through the presentation of an example.

Consider a grid cell size that is 5 x 5 m (25 m^2) (Figure 5). With such a small grid cell, only a portion of a tree crown would be captured and multiple cells would be required to characterize a single tree crown. Moreover, a circular ground plot with the same area would have a diameter of approximately 5.6 m and will likely contain too few trees (and have substantial edge effects, as discussed in Section 5.1.1). Therefore, larger grid cell sizes are preferred. Larger grid cells (and larger ground plots) will also contain a greater number of laser pulses, and are more likely to have a more uniform distribution of pulses (Næsset 2002). As indicated in McGaughey (2013:50):

> *To produce cover estimates that are meaningful, the (grid) cell size must be larger than individual tree crowns. With small cell sizes (less than 5 meters) the distribution of cover values of a large area tends to be heavy on values near 0 and 100 because each cell serves to test for the presence or absence of a tree instead of providing a reasonable sample area for assessing vegetation cover. For most forest types, cell sizes of 15-meters or larger produce good results.*

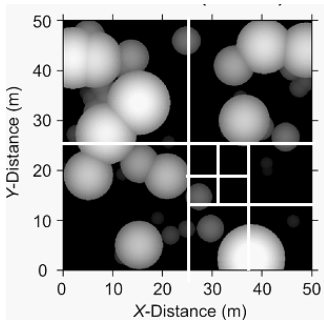

Figure 5. Implications of different grid cell sizes (adapted from Frazer et al. 2011b).

In Finland, a 16-m grid cell size is used operationally, relating to a ground plot with a 9-m radius. For work in Ontario, Woods et al. (2011) utilized a 20-m (400-m^2) grid cell size for metric calculation, supported by circular ground plots with a 11.28-m radius. In Alberta, a 25-m (625-m^2) grid cell size was used for the Hinton Forest Management Area (Frazer et al. 2011a), while a 20-m grid cell was used for the Grande Prairie Forest Management Area (Lim et al. 2013). In coastal rainforests of British Columbia, 20-m grid cells have been used for metric calculations, supported by circular ground plots with a 11.28-m radius. In Canada, we often store and analyze spatial data in a Universal Transverse Mercator (UTM) projection. It

[2] FUSION software uses both the ALS point cloud and DEM to calculate metrics, thereby normalizing the ALS point cloud heights "on the fly". Thus, depending on the software tool used to calculate metrics, it may not be necessary to normalize the ALS point cloud heights prior to metric calculation.

is, therefore, preferable if the grid cell size divides evenly into 100, enabling synergy and simplicity in integrating with other data sets in UTM. We recommend that the size of the grid cell match the size of the ground plot as closely as possible, and that the grid cell and ground plot be sufficiently large to accommodate the aforementioned considerations.

4.3 Tiling of Area of Interest

Numerous (likely hundreds) of LAS files will cover the area of interest. It is generally desirable (and often necessary from a file management point of view) to divide the area of interest into a manageable set of consistently sized tiles. Processing can then run on a per-tile basis. Users will often want to specify an origin for their tiling scheme (i.e., in the lower left, min x and min y) that enables the output ALS point cloud metrics to align with existing raster data sets (i.e., digital aerial photography, satellite imagery).

4.4 Metrics

Numerous forest height and density metrics can be generated from ALS data and many are known to be strongly intercorrelated. Generally, most forest applications will require some measure of height, variability of height, and amount of vegetation cover present (based on return density) (Lefsky et al. 2005). The FUSION *gridmetrics* command produces approximately 75 unique canopy and terrain metrics from ALS-measured heights, plus additional intensity and topographic metrics (Table 3).

Some studies used all returns to calculate metrics (e.g., Woods et al. 2011) and others calculated metrics separately for first and last returns (e.g., Næsset 2002); both methods of calculating metrics have produced robust, predictive models and to our knowledge, no studies in the peer-reviewed literature have rigorously examined whether one method is better than the other across a range of forest environments. Hawbaker et al. (2010) compared predictive models generated from first-return-only metrics and metrics generated from all returns in a mixed hardwood forest, and concluded that univariate models generated from first-return-only metrics explained more variability than models generated using all-return metrics; however, for multivariate models, the differences were small. It is worth noting that this study used leaf-off data, which may have affected the results for this forest type. Bater et al. (2011) demonstrated that metrics based on first returns in a coniferous-dominated coastal forest may be more stable across time and space. Therefore, an outstanding research question relates to the potential effect of metrics—calculated using first returns only or using all returns—on the accuracy of attribute estimates (e.g., basal area or volume) across a range of forest environments.

In the Hinton Forest Management Area, the FUSION metrics thought most useful for model building were identified and calculated (Table 4). From this subset, the most relevant metrics were identified using principal component analysis. The first three principal components extracted from a covariance matrix derived from the set of metrics presented in Table 4 accounted for 91.5% of the total variance found within the Hinton ALS data set. The first, PC1, explained 68.3% of the total variance, and was positively correlated with ALS canopy height (notably mean ALS canopy height). The second, PC2, accounted for 14.8% of the total variance, and was positively correlated with the vertical variability (dispersion) of ALS canopy height (i.e., the coefficient of variation of ALS canopy height). The third, PC3, was treated as the last non-trivial axis, and it accounted for 8.4% of the total variance; PC3 was positively correlated with canopy density (canopy cover). Previously published studies show that canopy height (PC1), the coefficient of variation of canopy height (PC2), and canopy density (PC3) are consistently reliable predictors of stand basal area, volume, and biomass (Lefsky et al. 2005; Li et al. 2008; Ni-Meister et al. 2010; Frazer et al. 2011b).

We recommend that end users follow an intuitive approach to metric selection, being mindful of their application and information needs. Principal component analysis can be used (as described above) to select a small set of relevant metrics. Although it may be tempting to generate all possible metrics and input these into a stepwise regression or similar approach to see what metrics emerge as significant predictors in model development, as we note above, the intercorrelation of many ALS point cloud metrics undermines this approach. To reiterate, in general, metrics informing on height, the variability in height, and the amount of vegetation present (as a minimum) should support a range of applications and can serve as a useful starting point for model development.

The strength of relationships between point cloud metrics and forest inventory attributes, such as volume and biomass, are predicated on the capability of ALS data to accurately characterize canopy height and density (Næsset 2011). Therefore, ensuring that non-canopy returns are separated from canopy returns is essential to develop robust predictive models. Nilsson (1996) was the first to exclude non-ground returns below a 2-m height threshold from the calculation of point cloud metrics and, subsequently, from model building and estimation. The 2-m threshold has since been applied in several studies that use the area-based approach (e.g., Næsset 2002; Andersen et al. 2005; Frazer et al. 2011a; Hyyppä et al. 2012; Wulder et al. 2012; Yu et al. 2013). Research indicates that the 2-m threshold is appropriate in the mature, boreal forest conditions where it was initially developed

Table 3. Airborne laser scanning metrics generated using FUSION *gridmetrics* command (see McGaughey 2013 for a full description of metrics).

Column	Elevation Metric	Column	Elevation Metric
1	Row	39	Return 1 count above htmin
2	Col	40	Return 2 count above htmin
3	Centre *X*	41	Return 3 count above htmin
4	Centre *Y*	42	Return 4 count above htmin
5	Total return count above htmin	43	Return 5 count above htmin
6	Elev minimum	44	Return 6 count above htmin
7	Elev maximum	45	Return 7 count above htmin
8	Elev mean	46	Return 8 count above htmin
9	Elev mode	47	Return 9 count above htmin
10	Elev stddev	48	Other return count above htmin
11	Elev variance	49	Percentage first returns above heightbreak
12	Elev CV	50	Percentage all returns above heightbreak
13	Elev IQ	51	(All returns above heightbreak) /(total first returns) * 100
14	Elev skewness	52	First returns above heightbreak
15	Elev kurtosis	53	All returns above heightbreak
16	Elev AAD	54	Percentage first returns above mean
17	Elev L1	55	Percentage first returns above mode
18	Elev L2	56	Percentage all returns above mean
19	Elev L3	57	Percentage all returns above mode
20	Elev L4	58	(All returns above mean) / (Total first returns) * 100
21	Elev L CV	59	(All returns above mode) / (Totalfirst returns) * 100
22	Elev L skewness	60	First returns above mean
23	Elev L kurtosis	61	First returns above mode
24	Elev P01	62	All returns above mean
25	Elev P05	63	All returns above mode
26	Elev P10	64	Total first returns
27	Elev P20	65	Total all returns
28	Elev P25	66	Elev MAD median
29	Elev P30	67	Elev MAD mode
30	Elev P40	68	Canopy relief ratio ((mean - min) / (max – min))
31	Elev P50	69	Elev quadratic mean
32	Elev P60	70	Elev cubic mean
33	Elev P70	71	KDE elev modes
34	Elev P75	72	KDE elev min mode
35	Elev P80	73	KDE elev max mode
36	Elev P90	74	KDE elev mode range
37	Elev P95		
38	Elev P99		

Table 4. Airborne laser scanning metrics generated using FUSION for the Hinton Forest Management Area.

Column	Elevation Metric	Name
8	Average of point heights > 2 m	LHMEAN
16	Average absolute deviation of point heights > 2 m	LHAAD
21	Second L-moment ratio (coefficient of variation) of point heights > 2 m	LHLCOV
22	Third L-moment ratio (coefficient of skewness) of point heights > 2 m	LHLSKEW
23	Fourth L-moment ratio (coefficient of kurtosis) of point heights > 2 m	LHLKURT
25	5th percentile of point heights > 2 m	LH05
26	10th percentile of point heights > 2 m	LH10
27	20th percentile of point heights > 2 m	LH20
28	25th percentile of point heights > 2 m	LH25
29	30th percentile of point heights > 2 m	LH30
30	40th percentile of point heights > 2 m	LH40
31	50th percentile of point heights > 2 m	LH50
32	60th percentile of point heights > 2 m	LH60
33	70th percentile of point heights > 2 m	LH70
34	75th percentile of point heights > 2 m	LH75
35	80th percentile of point heights > 2 m	LH80
36	90th percentile of point heights > 2 m	LH90
37	95th percentile of point heights > 2 m	LH95
50	% canopy density (cover) at 2 m	CC2M
56	% canopy density (cover) at mean canopy height	CCMEAN
57	% canopy density (cover) at modal canopy height	CCMODE

and applied, but that other thresholds are possible and may be more suitable in areas with different forest types or stand conditions (e.g., immature forests) (e.g., Nelson et al. 2004; Næsset 2011; Nyström et al. 2012). Studies in Ontario found that ALS metrics produced with no minimum height threshold resulted in better estimates than those that used a 2-m threshold (Woods et al. 2008; 2011); however, it should be noted that no alternative thresholds were tested in this study. We recommend that users consider applying a minimum height threshold when calculating ALS metrics. If the area of interest contains mature, coniferous forest, a 2-m threshold is likely suitable and is supported by peer-reviewed literature; otherwise, some experimentation with different threshold values is advised (e.g., Nyström et al. 2012), with the goal of ensuring that non-canopy or below-canopy returns are excluded from metric calculation.

4.4.1 Quality Assurance for Airborne Laser Scanning Metrics

Once produced, point cloud metrics require a quality assurance check. This initial post-processing task is necessary to ensure that a consistent "No Data" mask is applied to all the output rasters.

- "No Data" cells within each raster are identified and merged into a master "No Data" mask.
- Raster cells where maximum elevation is greater than a specified value are masked out (this value will likely vary according to what the "reasonable" maximum heights are for a given area).
- Raster cells with fewer than 70 canopy returns (e.g., returns > 2 m) are set to "No Data" (this is relevant for L-moments; see Guttman 1994).

This list of quality assurance tasks is by no means exhaustive. Output rasters should be examined for anomalous values and outliers. Simple queries (e.g., checking for values that are greater than two standard deviations from the mean) can provide some indication of potential outliers and indicate areas that may require further investigation.

5. Ground Plot Data

The collection of accurate ground information that describes relevant biophysical characteristics of forested plots is critical in the development of ALS-based predictive models of forest inventory attributes. Ground plots must represent the full range of variability of both response and predictor variables (Montgomery et al. 2006; Magnussen et al. 2010a; Frazer et al. 2011b), and be of sufficient size to avoid edge effects and minimize georeferencing errors (Gobakken and Næsset 2009; Frazer et al. 2011b). In this section, we describe various plot characteristics (size, shape, representativeness), sample size (number of plots), sampling design (distribution and location of plots), and procedures to measure or derive the forest attributes of interest (e.g., tree height, basal area, volume, biomass) from the ground plot measures. Based on a literature review and practical experience, this section summarizes best practices for acquiring ground data specifically to develop and validate predictive models derived from co-located ground measures and ALS data. Emphasis is placed on balancing accuracy and cost, and compliance with current, standard forest inventory procedures. While these recommendations generally apply to the installation of new ground plots, the information presented here is also valuable to determine whether existing plots from operational forest inventories or independent projects are useful to calibrate and validate ALS-based predictive models. As a general rule, the use of existing plot data (i.e., permanent sample plots) is not recommended for the area-based approach. The reasons for this include the potential for a substantial time lag between plot measurement and ALS data acquisition, unacceptable or unspecified (unknown) errors in plot positioning, plots with variable radii as opposed to fixed radii, and insufficient or inappropriate attribute information required for model development. These aspects are discussed in greater detail in the following sections.

5.1 Ground Plot Characteristics

5.1.1 Size

In general, ground plots used to support ALS-based forest estimation are larger than plots typically acquired for forest inventory or growth monitoring. Larger plots reduce the likelihood of edge effect, which occurs when canopy elements (i.e., tree crowns) found along the plot boundary are included inside the plot when these elements are actually outside the plot boundary, or (conversely), when excluded from the plot but actually occurring inside the boundary (Gobakken and Næsset 2009; Frazer et al. 2011b; Wulder et al. 2012) (Figure 6). In these cases, ground measurement crews should be mindful of how ALS data will characterize the plot, and know when it may be prudent to move (if possible) the plot centroid slightly to negate these edge effects.

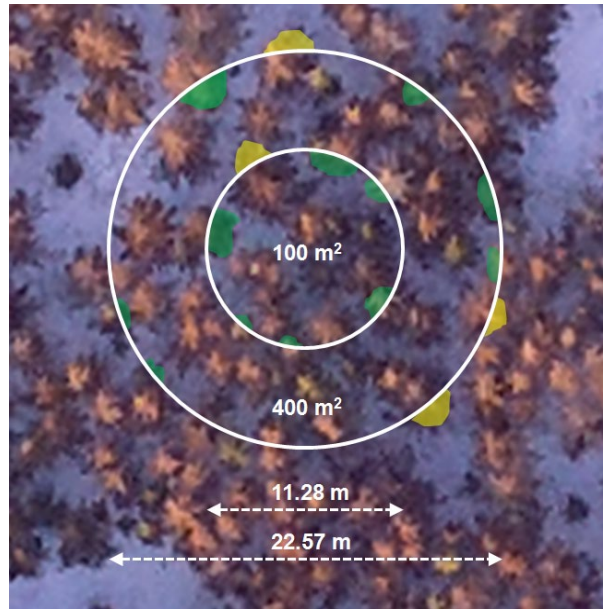

Figure 6. The concept of edge effect: green polygons represent portions of tree crowns found within the ground plot, where the majority of the tree crown (and likely the stem) is located outside the plot boundary. These trees would not have been measured on the ground and will not be included in the compilation of ground measures. The ALS will be clipped to be coincident with the ground plot and returns from these crowns will be included in the calculation of ALS plot metrics. Conversely, yellow polygons represent portions of tree crowns found outside the ground plot, where the majority of the tree crown (and likely the stem) is located inside the plot boundary. These trees would have been measured on the ground and included in the compilation of ground measures. When clipped to the plot boundary, the ALS point cloud will only include returns from that portion of the crown found within the plot boundary.

Similarly, when plots straddle two disparate conditions (i.e., boundary between two distinct vegetation types, or burned and unburned areas), it will make sense to move the plot to a position representing a uniform condition. Smaller plots have a larger perimeter-to-area ratio and therefore include more edge-related elements. This will translate into error in the calculated plot-level ALS metrics, resulting in metrics that are less accurate and less precise in describing a plot's vertical forest structure. The perimeter-to-area ratio

decreases exponentially with increasing plot size (Figure 7) and, therefore, the magnitude of the edge effect is reduced with larger ground plots (Frazer et al. 2011b). By maintaining more spatial overlap, larger plots also provide a buffer against co-registration errors between ground plots and ALS, which result from GPS positional errors (Gobakken and Næsset 2009; Frazer et al. 2011b). Assuming a fixed offset of 3 m between the ground plot centroid and the ALS plot centroid, Figure 8 illustrates that a 400 m^2 plot will maintain 19% more overlap area than a plot of only 100 m^2. Larger plots may also reduce undesirable between-plot noise and variance if each represents a relatively homogeneous section of a stand (Frazer et al., 2011b). Trade-offs between acceptable levels of cost and error will need to be considered in determining plot size.

A review of the scientific literature shows plot sizes ranging from 50 m^2 (8 m diameter) (Næsset and Økland 2002) to 2500 m^2 (56 m diameter) (Thomas et al. 2008), with an overall average of around 420 m^2 based on a sample of published articles (Nilsson 1996; Lefsky et al. 1999; Næsset and Bjerknes 2001; Popescu et al. 2002; Holmgren et al. 2003; Maltamo et al. 2004; Gatziolis et al. 2010). Although these studies provide no justification for choice of plot size, Gobakken and Næsset (2009) and Frazer et al. (2011b) specifically evaluated the combined effects of plot size and co-registration errors (analyzed in detail in section 5.3) on the relationships between ground-based and ALS-derived forest variables. In another similar study, Zhao et al. (2009) directly evaluated the effect of plot size on the estimation error of above-ground biomass from ALS.

Gobakken and Næsset (2009) used a 1.1 pulse per square metre discrete ALS data set to evaluate the effect of different plot sizes (200, 300, and 400 m^2) on the estimation of various attributes, including Lorey's height, basal area, and timber volume. Based on data from 132 concentric circular plots, they found a strong interaction between plot size and georeferencing error. Larger plots with the smallest georeferencing error provided the most accurate attribute estimates. Larger plots (300 or 400 m^2) were generally unaffected by positional errors of 5 m or less, while substantial biases in the estimation of Lorey's height, basal area, and volume could be introduced by even small errors in the position of 200 m^2 plots. Frazer et al. (2011b) expanded on the study by Gobakken and Næsset (2009) by evaluating a similar procedure applied to a broader range of plot sizes (314, 707, 1257, and 1964 m^2). The coefficient of determination (R^2) for the estimation of total above-ground biomass increased monotonically from 0.82 to 0.88 as plot size increased, with an asymptotic non-linear trend, suggesting that little improvement is expected for plots larger than 1257 m^2.

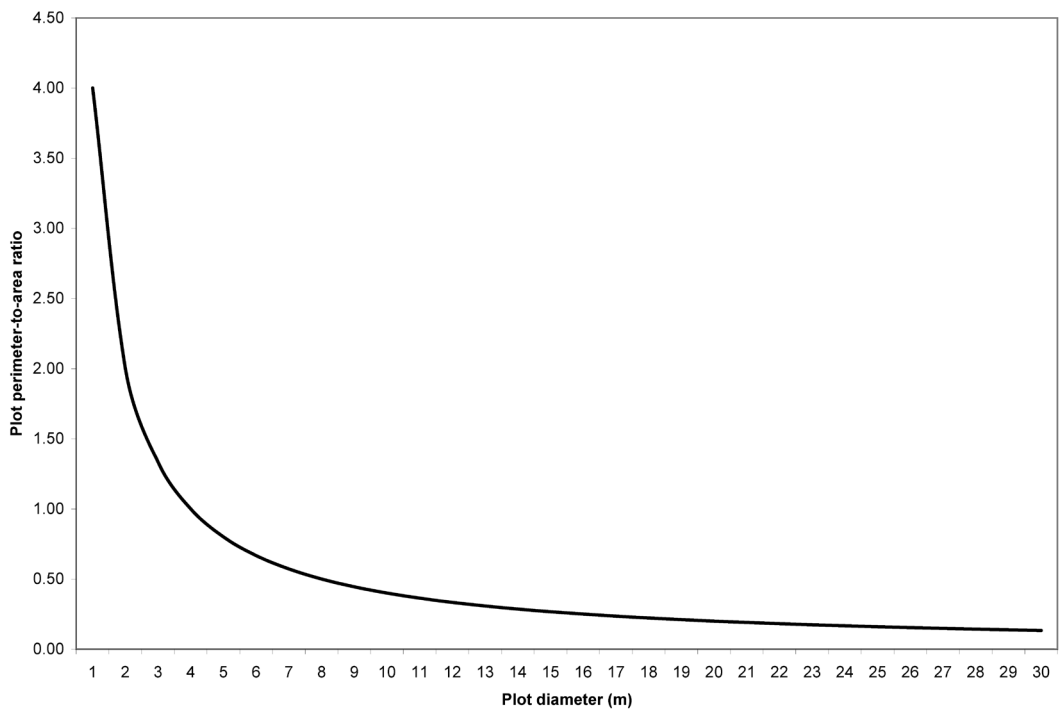

Figure 7. Plot perimeter-to-area ratio relative to plot diameter.

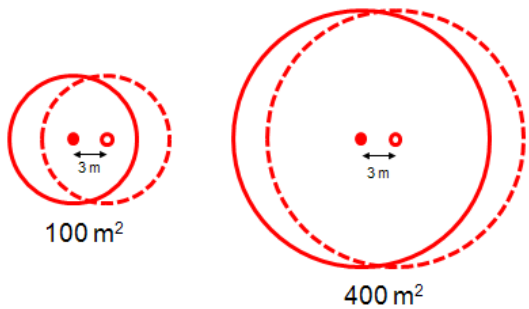

Figure 8. Effect of plot size on overlap between an actual plot (solid circles) and a plot georeferenced with a 3-m error (dashed circles); the overlap in this example is 63% for a 100 m^2 plot (left) and increases to 82% when plot size is 400 m^2 (right).

The results obtained by Gobakken and Næsset (2009) and Frazer et al. (2011b) indicate that it is not possible to recommend a *universal* optimum plot size for the modelling of forest inventory metrics with ALS, as the optimal plot size will depend on specific forest characteristics, plot georeferencing error, and (potentially) ALS return density. Nevertheless, the authors of these studies conclude that plot size is critical for minimizing estimation errors in ALS-based forest inventories by directly affecting the precision and accuracy of both ground-based and ALS-derived forest metric estimates and by reducing the ill-effects of plot georeferencing errors (Frazer et al. 2011b).

The optimum plot size for a particular forest area can be determined through preliminary sampling of the ALS data to determine the plot size at which estimates of height quantiles and/or canopy cover stabilize (Frazer et al. 2011b). In general, plot size can be reduced to optimize costs when:

- forest canopy conditions are homogeneous with low within-plot variability;
- plot georeferencing errors are small and consistent (e.g., with the use of GPS with differential correction); and
- variables of interest are less sensitive to spatial variability.

In relation to the latter point, Frazer et al. (2011b) and Gobakken and Næsset (2009) found that height metrics are less sensitive to plot size and georeferencing errors than density metrics.

To summarize, the literature suggests that the size of ground plots is fundamental to the accuracy of ALS-based forest inventories, with plot size and plot position having interactive effects. Frazer et al. (2011b) demonstrated the importance of plot size as a key sampling design parameter. Smaller plots are more sensitive to position issues and are overly influenced by edge effects. We recommend a ground plot size of 200–625 m^2 (i.e., with a plot radius of ~8–14 m) for use in the area-based approach. The selection of an optimum plot size from within this range should be determined by considering the need to:

- minimize edge effects;
- minimize planimetric co-registration error;
- maximize sampling efficiency; and
- maximize precision and accuracy of target and explanatory variables (Frazer et al. 2011b).

5.1.2 Shape

Both square and circular plots can be used. For example, Zhao et al. (2009) detected no notable differences in the accuracy of total above-ground biomass estimation from ALS when using square or circular plots; however, the choice of plot shape has practical implications. Authors estimating forest biophysical parameters have generally preferred circular (Næsset and Bjerknes 2001; Næsset 2002; Popescu et al. 2002; Holmgren et al. 2003; Inoue et al. 2004; Zhao et al. 2009; Gatziolis et al. 2010) over rectangular or square (Maltamo et al. 2004; Yu et al. 2004) plots. Circular plots are easier to establish in the field since only the centre needs to be registered as opposed to four corners (Adams et al. 2011). Additionally, circular plots have 13% less perimeter than their square counterparts of equal area, thus minimizing the negative impact of edge effects discussed previously, which introduce error into metric calculation (Wulder et al. 2012). As a better spatial correspondence between ground plots and co-located ALS subsets maximizes the correlations between response and predictor variables (Frazer et al. 2011b; Gobakken and Næsset 2009), we therefore recommend the use of fixed-area circular ground plots.

5.1.3 Other Ground Plot Considerations

Certain terrain and vegetation conditions can make the establishment of ground plots difficult. Because these same conditions also present challenges for accurate measurements of tree heights using ALS, ground plots should not be avoided when these conditions arise but rather extra care should be taken to ensure accurate ground measures.

Canopy Density and Surface Irregularity

Developing a reliable DEM with ALS data is fundamental to accurately calculating forest height metrics (Bater and Coops 2009). Dense canopies, especially when scanned by low-return density ALS systems, can drastically reduce the proportion of ground returns needed to develop accurate, detailed DEMs. Interpolation methods that connect sparse

ground returns can therefore miss important terrain features under these conditions and add bias to ALS-derived height metrics (Figure 9). Gatziolis et al. (2010) evaluated the effect of steep terrain and dense canopy cover on the accuracy of ALS-derived canopy heights in the temperate rainforests of the US Pacific Northwest. They concluded that tree height is substantially underestimated in areas with steep terrain and continuous, dense canopies where ALS penetration is poor. In addition, they found that tree lean, a common phenomenon in steep terrain, can further confound accurate height estimates. Dense forests and complex terrain will also be challenging for ground plot installation and extra care must be taken when measuring tree attributes (see Section 5.5) on these difficult sites.

Understorey Conditions

Under some circumstances, ALS returns intercepted by a dense understorey can be misclassified as forest canopy elements and therefore bias inventory metrics and/or minimize the proportion of ground returns needed for acceptable DEM creation (Haugerud et al. 2003). To address this issue, data processing methods have been developed to stratify vegetation layers and isolate understorey from tree canopies (Riaño et al. 2003; Maltamo et al. 2005). Sites with dense understorey are also challenging environments to establish ground plots and as with the case of complex terrain and dense canopies noted above, special care should be taken to ensure accurate tree measures at these sites.

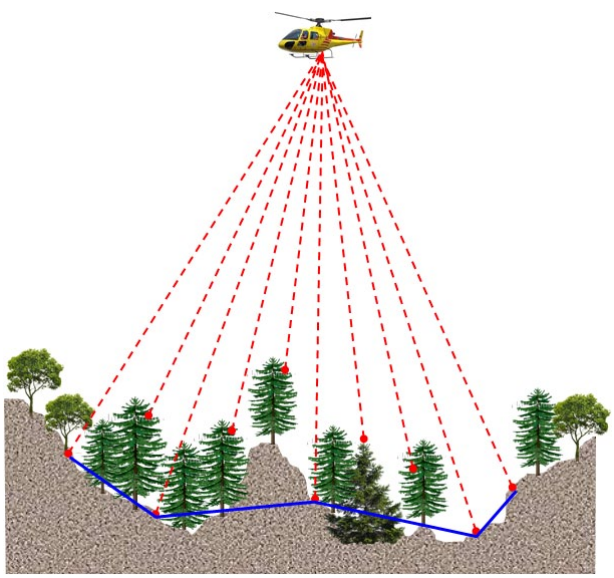

Figure 9. Errors in the creation of a DEM with ALS data (blue line) if return density is low or canopies are dense; the tree on the ridge top shown at the middle of the illustration will be attributed twice its real height owing to the lack of nearby ground returns; the height of other trees will be underestimated.

Timing

Three fundamental factors must be considered when obtaining ground plot data to support ALS-based forest inventories:

1. the advantages (and disadvantages) of collecting ground data before (or after) ALS acquisition;
2. the need to minimize the time elapsed between ground and ALS data collection; and
3. specific canopy conditions occurring at the time of both ground and ALS data collection (i.e., seasonal growth stages or leaf-on/leaf-off conditions where deciduous species are present).

If a gap of one or more growing seasons exists between the timing of the ALS acquisition and the ground plot measurements, Adams et al. (2011) suggested estimating growth differences and adjusting the ground or ALS measures before exploring correlations. This approach may be more relevant in areas where the annual growth increment is large. Gobakken and Næsset (2009) applied site-specific growth models to ground measures of forest height, basal area, and volume obtained 18 months before ALS data acquisition. Whether such an approach is applied will depend on the quality of the growth models available, the expected growth increment, and the length of time between the ground and ALS data acquisitions. In contrast, disturbances that have occurred since ALS or ground data acquisition (whether natural or anthropogenic in origin) should be accounted for. We recommend that ALS data be acquired before ground plots and, furthermore, that the ALS data be used to guide ground plot sample selection (see Sections 5.2 and 5.3, below). Ideally, ground plot information will be obtained within one growing season of the ALS acquisition.

5.2 Representativeness of Ground Plots

Numerous published studies show that regression models will have higher error rates if ground calibration data sets do not capture the full range of variability in the dependent variable (e.g., height, basal area, volume) (Demaerschalk and Kozak 1974; Hawbaker et al. 2009; Maltamo et al. 2011). This increase in prediction error occurs because models will perform best when operating within the bounds of the original calibration data (i.e., the smallest convex set containing all the design points; Cook 1975), and will perform poorly when forced to extrapolate beyond this region (Montgomery et al. 2006). Figure 10 illustrates the concept of plot representativeness in the context of a forest structure feature space defined by the second and third principal components generated from a set of ALS metrics. The full range of variability present in the area is indicated in red, whereas the variability

captured by the ground plots is shown in green. Two issues are apparent with the ground calibration data shown in this example:

1. the ground plots do not cover the full range of forest structural variability present in the area of interest (the red line); and
2. within the range of forest structural variability captured by the ground plots (the green line), the ground plots (green dots) are not evenly distributed in the feature space represented by the principal components but rather are clustered in a single area.

In the example presented in Figure 10, the ground plots are concentrated in younger, shorter stands. This often occurs when plots that exist for some other purpose (e.g., permanent sample plots for growth and yield) are used for model development. Therefore, we generally do not recommend that pre-existing plots be used for model development.

Statistics describing model fit will only illustrate the expected prediction accuracy within the convex hull defined by the calibration data (the green line in Figure 10) (Cook 1975). Outside this convex hull, parametric models (e.g., least-squares regression) are considered to operate in extrapolation mode, making prediction accuracies far less certain (Demaerschalk and Kozak 1974; Montgomery et al. 2006). Non-parametric models (e.g., kNN, Random Forests), which depend on the proximity of nearest neighbours in a reference set to impute plausible predictions, cannot extrapolate. Indeed, because kNN cannot predict a value less than (or greater than) the smallest (or largest) measured value of *Y* (i.e., ground plot measures), the method is inherently biased (Magnussen et al. 2010b). For both parametric and non-parametric approaches, extrapolation errors are often restricted to rare forest structural types (i.e., those located on the margins of feature distribution space); however, rare forest types can also contain a disproportionate amount of wood volume, biomass, and favourable habitat, owing to the presence of unusually large trees and/or more spatially complex stand structures (Frazer et al. 2011b).

If the majority of ground plots are concentrated within a few stand structural types (e.g., shorter, younger, and structurally homogeneous stands), rather than being uniformly distributed across the range of structural variability present in the study area (and captured by the wall-to-wall ALS coverage), these areas of high sample concentration will tend to have a large influence on model fit. This can lead to larger prediction errors and bias for sparsely sampled rare forest types. The

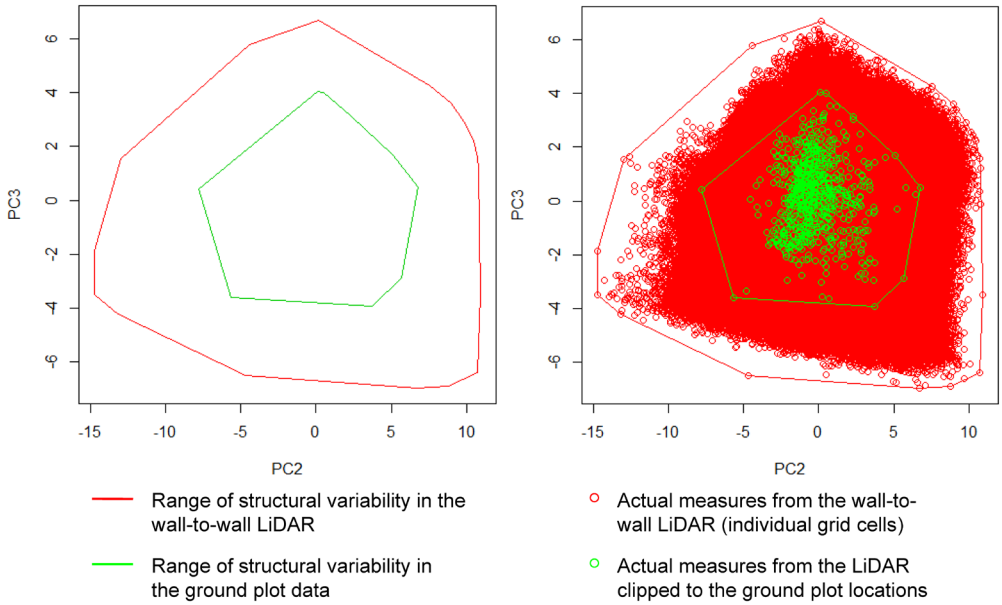

Figure 10. Representativeness of ground plots assessed against the full range of variability present in ALS data acquired for an area of interest. The feature space is defined by the second principal component (PC2, *x*-axis) and the third principal component (PC3, *y*-axis) generated from a set of ALS metrics. The left panel shows the convex hull for the full range of forest structural variability in the area of interest (as captured by the wall-to-wall ALS coverage) compared to the convex hull of the ALS data corresponding to the locations of the ground plot measurements. The right panel shows the location of actual measures in the feature space.

clustering of calibration observations within a few canopy structure types has the potential to distort least-squares model fit away from the true population model. For example, a high concentration of ground calibration observations within shorter stands will have an overwhelming influence on the final model specification (i.e., selection of predictors, data transformation, model form). Consequently, prediction models are arbitrarily and unduly weighted towards shorter stands. Such models are therefore more likely to produce substantially higher prediction errors and bias when applied to taller and more spatially complex stand structures.

Although sampling is typically undertaken to make inferences for a population, in the context of the area-based approach, sampling is carried out to develop more robust predictive models. As such, sampling is no longer focussed solely on accurately and precisely estimating the population mean and standard error, but rather on ensuring that the full range of forest structural variability is captured with ground plot measurements. Once acquired, wall-to-wall ALS data can be used to stratify the area of interest according to the forest structure present, and thereby guide the development of experimental designs that optimize sampling parameters and plot size (Frazer et al. 2011b). Structurally guided sampling based on ALS ensures that the full range of forest conditions present in the area of interest are captured, and that these samples are evenly distributed within the range of variability (Hawbaker et al. 2009; Maltamo et al. 2011). A structurally guided sampling approach would involve stratification of the area of interest using ALS metrics. Since it is known that ALS canopy height and density metrics are strongly intercorrelated (Lefsky et al. 2005), we recommend the use of a few key ALS metrics for stratification (Maltamo et al. 2011; see Section 4.4 for a more details).

5.3 Selection of Ground Plot Locations

A sample design guided by the structural variability present in the area of interest (as characterized by the ALS data) should be used to select appropriate sample locations. Theoretically, the centroids of all grid cells within the area of interest are potential sample locations. If any issues are associated with the ALS survey, then these locations may be pre-screened. For example, several different ALS surveys may be conducted over a very large area. If the parameters of these surveys (e.g., the timing of ALS acquisitions, the instrumentation, or pulse density) are significantly different, then the development of separate sample designs and models may be desirable for the different survey areas (see Section 3.3). The structurally guided sample would use a few key ALS metrics to stratify the area of interest. The available samples would be assigned to a stratum and then the required number of samples would be randomly selected for each stratum. On a practical note, since the objective of this ground sampling is to develop robust relationships between ground measures and the ALS metrics, crews may need to move plot locations under certain conditions (see Section 5.1.1).

5.4 Ground Plot Positioning

Accurate georeferencing of ground plots is fundamental to maximize the predictive power of the developed models. Although increasing plot size can help mitigate the impact of georeferencing error (Figure 8), a direct improvement in georeferencing is preferred (Figure 11). Ground plot position is measured using GPS technology. Forests are challenging environments for recording accurate GPS positions because trees and terrain can obstruct clear views of the sky (Bolstad et al. 2005). The low power signals of GPS satellites have difficulty penetrating forest canopies, and the moisture contained in canopy elements further reduces GPS signal strength. This results in a low signal-to-noise ratio under forest canopies, with the magnitude of this effect increasing with increasing canopy density (Edson and Wing 2012). Given an unobstructed view of the sky, a GPS receiver is able to calculate which combination of four GPS satellites (of the > 24 GPS satellites orbiting the earth) provides the best geometry at any given time to obtain the most accurate position. The effect of satellite geometry on the accuracy of GPS positions is measured using several dilution-of-precision indices. Clear-sky views are typically limited by the forest canopy and a GPS unit will often only receive signals from satellites directly overhead (Johnson and Barton 2004). Satellites that are clustered together provide redundant information, resulting in less accurate GPS positions. This effect is measured as the positional dilution-of-precision index, which is often high in forest environments. When a GPS signal intercepts an object, it may be reflected, a phenomenon known as multipath (Wing 2008). Many objects can reflect GPS signals in forest environments (e.g., tree branches, stems), making it difficult for a GPS unit to discriminate between true signals and reflected ones. This multipath phenomenon is further exacerbated in wet conditions.

The accuracy of a GPS position also depends on the quality of the GPS receiver used. Receivers are categorized as recreation, mapping, or survey grade. Recreation-grade GPS receivers do not provide sufficient accuracy for forestry applications, and survey-grade receivers are prohibitively expensive. Mapping-grade receivers are further classified into commercial grade units capable of achieving horizontal accuracies of 3 m or greater, or differential grade units achieving sub-metre accuracy (Edson and Wing 2012). Using mapping-grade GPS receivers, positional errors associated with ground plot locations in forested environments typically range between 1 and 5 m (Deckert and Bolstad 1996; Næsset 1999; Næsset and Jonmeister 2002; Bolstad et al. 2005; Wing and Karsky 2006; Wing and Eklund 2007; Wing et al. 2008;

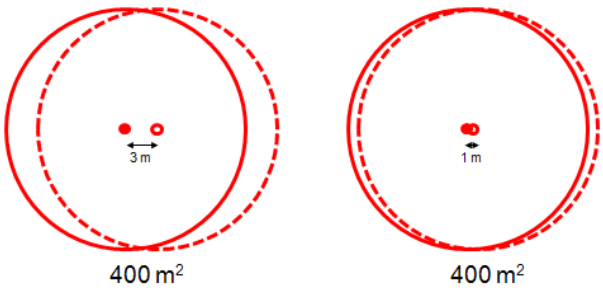

Figure 11. Effect of co-registration error on the overlap between an actual plot (solid circles) and a plot with a georeferencing error (dashed circles). The overlap in this example is 82% for a 3 m error (left) and increases to 96% when the error is reduced to 1 m (right).

Edson and Wing 2012). Although positional errors are also associated with the ALS data, these errors tend to be minor (i.e., < 0.5 m) in comparison (Gatziolis and Andersen 2008; Edson and Wing 2012).

Frazer et al. (2011b) and Gobakken and Næsset (2009) evaluated the combined effect of plot size and horizontal co-registration spatial errors on the relationships between ground- and ALS-based forest metrics. Both studies were based on simulations generating ALS metrics in hundreds of spatial subsets covering the same area as ground plots but randomly varying their azimuthal location around their true centres. These iterations were performed for different fixed distances from the correct position of ground plots, varying by 0.5–20 m in Gobakken and Næsset (2009) and 1–5 m in Frazer et al. (2011b), who argued that typical GPS measurements are within this range. Gobakken and Næsset (2009) concluded that differences between ground-truth values and the median of iteratively predicted values of Lorey's height, basal area, and volume were not substantially affected by georeferencing errors of less than 5 m. In all cases, the use of larger plots minimized the potential biases introduced by positioning errors, but the authors emphasized that increasing plot size and improving the accuracy of GPS measurements would increase the costs of ground surveys.

As most modern GPS systems are capable of achieving positional accuracies of less than 5 m under forest canopies (with differential correction), the study by Frazer et al. (2011b) suggested that standard field plot installation procedures based on mapping-grade GPS receivers are sufficiently accurate. Differential correction involves the use of two GPS receivers, one at a known (fixed) position (i.e., the base station) and another used by the ground crew for measurement (i.e., the roving unit). Signals are acquired by both receivers concurrently; the base station calculates its position relative to its known location, and the difference is used for correction. Furthermore, since the base and roving units acquire data simultaneously, both will experience similar atmospheric errors. Differential correction is therefore a post-processing correction, with base station data used to correct information gathered by the roving unit. Data from permanent base stations are also available for differential correction. Offering precise ephemeris data from GPS satellites, information from these stations is typically available 12 days after collection (Edson and Wing 2012). Adams et al. (2011) suggested using the most up-to-date, professional GPS equipment that a user can afford to buy or rent.

We recommend that users collect a minimum of 500 location measurements at each point of interest (centre for circular plots; four corners for square plots), at a rate of 1 point per second, and apply differential post-processing correction procedures to maximize plot positional accuracy. The use of antenna extensions can also improve the accuracy of measurements in the forest (Edson and Wing 2012). The influence of planimetric errors is lessened when forest stands are homogeneous, larger plot sizes are used, and variables of interest are less sensitive to horizontal spatial variability (e.g., Frazer et al. [2011b] and Gobakken and Næsset [2009] found that height metrics were less sensitive to plot size and georeferencing errors than density metrics).

5.5 Tree Measures

Individual tree measures or attributes derived from individual tree measures are key response variables in the development of ALS-based predictive models of forest inventory attributes. Therefore, the accuracy with which tree measurements are made is fundamental for the success of the models and subsequent stand-level predictions. Only a few basic direct measurements need to be made at the plot level, with additional attributes derived or compiled from the tree measures. This section describes the basic attributes that are either directly measured at the ground plot-level, or derived/compiled from ground plot measures. Methods of measurement and derivation/compilation are recommended. Field procedures to obtain ground measures are generally simple and straightforward (e.g., Lim et al. 2003b; Maltamo et al, 2004; Næsset 2004; Parker and Evans 2004; Reutebuch et al. 2005; Wulder et al. 2012). Table 5 provides some examples of measured and compiled attributes from the scientific literature.

Table 5. Examples of ground measurements and methodological notes from literature.

Study	Ground Measures	Methodological Notes	Compiled Ground Plot Attributes	ALS-estimated Attributes
Næsset (2004)	• species • dbh • height	Minimum dbh thresholds were > 4 cm on sites classed as young forest, and > 10 cm on sites classed as mature forest. Tree heights were measured on sample trees (selected with probability proportional to stem basal area). Sampling intensity on each plot was proportional to inverse of stand basal area to ensure equal number of sample trees per plot. Heights of trees that were not measured were calculated from diameter-height relationships.	Lorey's mean height (mean height weighted by basal area); dominant height (arithmetic mean height of sample trees corresponding to 100 largest trees per hectare, by diameter); mean plot diameter (mean diameter by basal area); stem number computed as number of trees per hectare; plot basal area (basal area per ha). Volume for each tree calculated using volume equations, with measured (or modelled) height and dbh as predictors. Total plot volume = sum of individual tree volumes.	• mean tree height • dominant height • mean diameter • stem number • basal area • timber volume
Hawbaker et al. (2009)	• species • dbh • height	Minimum dbh threshold = 12.7 cm (and tree within 17 m of plot centre). Total tree height was measured for a subset of trees (selected using a 2.3 m^2/ha basal-area factor prism).	Mean dbh, mean tree height, basal area summarized at the plot level. Total tree biomass derived from species-specific allometric equations.	• dbh • basal area • mean tree height • biomass
Woods et al. (2011)	• species • tree status (live or dead) • crown class • height to crown base • height[2]	All ground variables measured in trees with dbh > 9.1 cm. Crown class = dominant, co-dominant. Lag between ALS and ground data acquisition, disturbed areas excluded from ground sampling. Height of deciduous species were measured in leaf-off conditions to obtain most accurate height possible. A minimum of 300 GPS points per post; post-processed against base station.	Top height estimated as the mean height of largest 100 trees (dbh) per hectare; average height of all trees dbh > 9.1 cm. Density is the number of live trees with dbh > 9.1 cm. Quadratic mean dbh (cm) is derived from dbh and the number of stems per plot. Basal area is derived from dbh. Gross total volume (inside bark) is derived from species-specific dbh-height equations. Gross merchantable volume (inside bark) derived from total volume. Above-ground biomass calculated with an allometric equation based on dbh (with different model coefficients for each species).	• top height • average height • density (stems/ha) • quadratic mean dbh • basal area • gross total volume • gross merchantable volume • above-ground biomass

5.5.1 Measured Attributes

Species, Status, and Crown Class

Species of all measured trees in the ground plot must be recorded in the field. Accurate species identification in the field is fundamental because the equations used to estimate certain forest attributes can be species (or forest type) specific. The status of the tree (live or dead) should also be recorded, along with the crown class (e.g., dominant or codominant).

Diameter at Breast Height

Diameter at breast height (dbh; cm) is the most basic and common tree measurement and is defined as the outside bark stem diameter of a tree at a point on the stem that is 1.3 m above the ground (Avery and Burkart 2002). Typically, a minimum threshold for measurement is specified, ranging from 5 to 10 cm, and only those trees in the plot with a dbh greater than this minimum threshold are measured. The dbh of all trees that meet the threshold should be measured, regardless of status (live or dead). The most common methods for measuring dbh in the field are to use a diameter tape that is calibrated in units of diameter or to use callipers. A diameter tape is commonly used for permanent sample plots in Canada because it is more consistent for repeat measures (Binot et al. 1995). The diameter tape must be oriented perpendicularly to the stem's vertical axis, and several anomalous conditions must be considered, such as:

- in steep slopes, the 1.3 m for dbh recording should be measured in the upslope side of the tree;

- when a tree forks below 1.37 m, a dbh measurement can be taken for each stem (i.e., each stem is considered an individual tree);
- if branches or other anomalies are located at 1.3 m, dbh can be measured upwards in a location arbitrarily selected to better represent the tree diameter;
- if inventories are carried out when snow is on the ground, the field crew must make sure that the bare ground is used as a reference to measure dbh.

Regardless of the method selected to measure dbh, the same method should be used for all ground plots.

Height

Tree height is the distance between the base of a tree and the top of the tree. The measurement of height in ground plots is time consuming, so most studies have opted to measure only a subsample of tree heights in the plot (e.g., Holmgren et al. 2003; Næsset 2004; Andersen et al. 2005; Gobakken and Næsset 2009) rather than all trees in the plot (Popescu et al. 2002; Maltamo et al. 2004). If only a subsample of tree heights are measured, then the selected sample must be representative of the dbh frequency distribution within a plot. The heights of the unmeasured trees can then be estimated through species-specific dbh–height regression models developed from the measured subsample (e.g., Richards 1959; Sharma and Parton 2007). Clinometers and hypsometers are the most common instruments to measure individual tree height. These instruments are based on simple trigonometric relationships between the known planimetric distance from the instrument to the tree and the angles from the instrument to the base and top of the tree, which must be clearly visible when performing measurements (Andersen et al. 2006). A Vertex™ hypsometer (Haglöf, Långsele, Västernorrland, Sweden) is one of the most popular instruments for measuring tree height because it is efficient, easy to calibrate, and automatically measures the distance to the tree.

Stem Number

The stem number is an enumeration of the total number of measured trees in the plot.

5.5.2 Derived or Compiled Attributes

Basal Area

Basal area is the common term used to describe the average amount of area (usually in square metres) occupied by tree stems (Avery and Burkart 2002). Specifically, it is the sum of the cross-sectional area of individual trees at breast height, expressed per unit of land area (i.e., m^2/ha). Basal area is calculated from dbh (measured in centimetres) using the following equation:

$$\text{Basal area } (m^2) = 0.00007854 \times \text{dbh}^2$$

Plot Height

Plot-level variations of height are derived from individual tree height measures, including mean tree height, which is the arithmetic mean of all tree heights in the plot; Lorey's height, which is the arithmetic mean height weighted by the basal area of the trees in the plot; and dominant height (or top height), which is the arithmetic mean of the 100 largest trees per hectare (by diameter).

Volume

Timber volume, measured in cubic metres per hectare (m^3/ha) is a key inventory variable. Individual tree volume is typically modelled and not directly measured. Taper equations are used to predict whole stem volume and merchantable volume from species, height, and dbh measurements. Taper equations are regression functions derived from data acquired via destructive sampling of individual trees that represent the size classes of interest, and can be species-specific, and even site- and age-specific for improved estimates of volume (Adams et al. 2011). Merchantable volume is derived from total volume and is the whole stem volume minus deductions for bark, tree stumps, and tops (Avery and Burkhart 2002). For eastern Canadian species, the equations based on dbh and height derived by Zakrzewski (1999) and Honer (1964) are commonly used.

Biomass

Biomass refers to the amount of vegetation matter per unit of area and is commonly measured in megagrams per hectare, where 1 Mg is equivalent to 106 g or one metric tonne. Although tree roots represent considerable biomass, they are not easily measured and therefore estimates are confined to the total above-ground biomass. Biomass is calculated as the dry weight of tree elements above ground, including stems, branches, and leaves. Since procedures to obtain this true biomass are destructive and expensive (Houghton 2005), substantial research has been dedicated to the development of species-specific allometric equations to derive biomass estimates from ground measurements of dbh (Frazer et al. 2011b). Examples include Ter-Mikaelian and Korzukhun (1997), Lambert et al. (2005), and Ung et al. (2008). Because biomass components are estimated from these generalized equations (rather than directly measured), we are almost certain to underestimate the error with which subsequent predictions are made from the ALS data. Therefore, such estimates of biomass should always be presented with a cautionary note.

5.5.3 Summary and Recommendations

Experimental Design

In a perfect scenario, ALS data would be acquired before ground plots. Metrics would be generated from the ALS and used to stratify the area of interest and develop a structurally guided sampling protocol. Ground plots would then be

acquired to develop and validate models of forest inventory attributes. Ideally, both ALS and ground data would be acquired in the same growing season.

Plot Characteristics

There is no universally optimum plot size. The optimum plot size will depend on stand variability, plot positioning errors, and available budget. Our recommended range of ground plot sizes for use in the area-based approach is 200–625 m^2 (8–14 m radius). The selection of an optimum plot size from within this range should be determined by the need to minimize edge effects and planimetric co-registration error and to maximize sampling efficiency, precision, and accuracy of target and explanatory variables. Circular plots are preferred over other shapes, owing to the ease and accuracy with which these can be installed and spatially referenced with a GPS, and their lower perimeter-to-area ratio reduces edge effects. The time elapsed between ground measurements and ALS acquisition should be minimized, and special care must be given to the measurement of ground plots located in areas where canopies are dense, understorey is thick, and/or terrain is irregular.

Plot Positioning

Modern mapping-grade GPS systems should result in plot positioning errors of 1–5 m (Wing et al. 2008). To lessen positional error, collect a minimum of 500 GPS points per plot (Adams et al. 2011), use a second GPS unit as a base station and an external antenna on the roving unit, and apply differential post-processing correction after acquisition. If plot positioning errors are large and differential correction is not available, increase plot size to ensure sufficient overlap with the ALS. Avoid plot locations that are close to stand boundaries.

Tree Measures

As most ground-based inventory variables are derived from dbh and height measurements, efforts should focus on measuring these two variables as accurately as possible. Selection of the most appropriate allometric equation for estimating volume and biomass is also critical as substantial bias can be introduced by using the inappropriate equations. Species, status, and crown class should also be recorded.

6. Modelling

To enable predictions of attributes of interest, relationships must be established between the ALS metrics and the co-located ground measures. Models may be developed for species or species groups (Woods et al. 2011) or, more broadly, for forest types (Frazer et al. 2011a). Different approaches may be used to build predictive models, but the most common methods used are either parametric (e.g., Means et al. 2000; Næsset 2002; Holmgren 2004; Næsset et al. 2011; Woods et al., 2011) or non-parametric (e.g., Packalén and Maltamo 2007; Hudak et al. 2008; Frazer et al. 2011a; Järnstedt et al. 2012; Vastaranta et al. 2012) approaches. Parametric approaches generate models that can be defined or parameterized by a finite number of parameters and make several *a priori* assumptions about relationships between response and predictor variables (i.e., errors are normally distributed, independent, and have a constant variance [homoscedastic]). Conversely, non-parametric approaches make no such assumptions.

Numerous parametric regression approaches have been used to build predictive models of forest inventory attributes. Næsset et al. (2005) compared ordinary least-squares regression, seemingly unrelated regression, and partial-least-squares regression but concluded that none of these approaches outperformed any of the others. Because ordinary least-squares is simple to apply and the results are readily interpreted, Næsset et al. (2005) suggested that ordinary least-squares regression should be considered as the approach of choice for practical forest inventories. Woods et al. (2011) used seemingly unrelated regression to predict a suite of attributes for an operational inventory in Ontario, Canada, selecting a maximum of two logical predictor variables, which were strongly correlated with the inventory attribute for which a predictive model was being developed, and which had low variance inflation factors. Transformation of ALS metrics (X) or ground plot measures (Y) may be necessary if regression-based approaches are used (Frazer et al. 2011b). Some studies have applied log transformations to both X and Y (Hudak et al. 2006; Næsset 2002), or only Y (Li et al. 2008), and others have applied squared transformations of X (Lefsky et al. 2002). Some studies have applied natural log transformations to both X and Y (Lim et al. 2003a; Holmgren 2004; Hawbaker et al. 2009). Frazer et al. (2011b) applied a Box-Cox transformation to Y.

Random Forests is the most common non-parametric approach applied for ALS-based forest inventories. It is a regression-based decision tree approach (Breiman 2001)—a virtual forest of regression trees built from bootstrapped training data. Each random selection (with replacement) of training data is used to build a separate regression tree and the overall performance of the final model depends on the prediction accuracy of each individual tree and the correlation between the residuals of individual trees (Segal 2004). Overall performance is optimized by two strategies:

1. Randomization is used to reduce the potential for correlation between residuals of individual trees, both

in the selection of training data used to grow the tree and in the number of predictor variables that are used at each node for splitting.

2. Individual trees are allowed to grow to their maximum depth to improve the prediction accuracy of individual trees (Breiman 2001).

Segal (2004) posited that allowing trees to grow to their maximum depth can help to control bias in the predictions from individual trees, but it will not control the variance. The Random Forests approach does not require feature selection and is reported to excel in situations where many more predictors exist than samples (Breiman 2001). A significant advantage of this approach over parametric approaches is the use of categorical variables as predictors (and the ability to predict categorical variables).

Parametric regression approaches make certain *a priori* assumptions and data must be tested to ensure that these assumptions are not violated. As a result, parametric approaches have more analytical overhead associated with them than non-parametric approaches: it takes more time to test for violations of assumptions, apply an appropriate transformation to *X* and/or *Y* variables, and develop robust models. If new ALS data are acquired, assumptions must again be tested and new models developed. In contrast, the Random Forests approach may be faster to develop and implement. Also, because it does not make *a priori* assumptions, if new ALS data are acquired, previously developed models could theoretically be applied to the new data, provided that metrics generated from the new data are within the same range of variability as the metrics used to develop the original models. The apparent advantages of this approach (e.g., ease and speed of implementation, portability of models) are not without trade-offs, with the primary one being the "black box" nature of the algorithm and the lack of transparency in model outcomes. Moreover, the application of Random Forests in the context of ALS-based forest inventory is relatively new in the scientific literature, and a rigorous evaluation or benchmarking of it has not yet been completed. Parametric approaches enable confidence intervals and significance testing, as well as sample size determination for given accuracy and precision requirements. The relative advantages and disadvantages of parametric regression and the Random Forests approach are summarized in Table 6.

Regardless of whether a parametric or non-parametric approach is selected, the temptation is often to include all possible ALS metrics for model development to see which ones emerge as significant predictors. Given that many of the ALS metrics are known to be correlated, we recommend that users select appropriate metrics for model building based on their information need or by using a method such as principal component analysis, or some other feature selection approach (Li et al. 2008; Stephens et al. 2012). In a parametric regression-based approach, failure to carefully select appropriate input metrics can result in models that have a large number of predictor variables, which in turn, can result

Table 6. Relative advantages and disadvantages of parametric regression and Random Forests approaches to modelling (in the context of the area-based approach).

	Parametric Regression	Random Forests
Advantages	• Transparent, easy to understand • Model is an equation that clearly quantifies the relationship between the predictors and the variable being predicted • Sample size determination is possible for given accuracy and precision requirements	• Categorical variables may be predicted and/or used as predictors • Faster and simpler to develop (does not require sophisticated statistical expertise) • Does not require individual strata-based models to be developed, provided calibration data represent the different strata involved • Does not require a pre-existing polygon-based inventory to implement strata-based models
Disadvantages	• Transformation of ALS metrics (X) or ground plot measures may be necessary to meet the assumptions of regression-based approaches, complicating interpretation and implementation • More statistical expertise and time are required to create the models • With strata-specific models, pre-existing stratification across the entire forest (i.e., an existing inventory layer) becomes prerequisite to implementation • Prediction errors will occur within polygons when individual grid cells do not match the overall strata assignment (e.g., pockets of aspen within a "spruce" polygon)	• "Black box" nature of the models • No equation output that is analogous to parametric regression • More critical to ensure that the full range of conditions are sampled, as this approach does not extrapolate like regression

in unstable predictions stemming from multicollinearity among two or more of explanatory variables. Although the capability to include all possible predictors is often touted as one of the advantages of the Random Forests approach, research indicates that Random Forests is also affected by correlated predictors (Strobl et al., 2008; Toloşi and Lengauer 2011; Adjorlolo et al. 2013). Finally, predictions from parametric and non-parametric models must be constrained to the observed range of data used to calibrate the model (see Section 5.2). Both parametric and non-parametric approaches require representative ground plot data for robust model development. Since a Random Forests model cannot extrapolate predictions, it requires ground samples to be evenly distributed across both *X* and *Y* space.

The selection of the best approach for modelling will depend on several factors. These include:

- the characteristics of the forest in question (e.g., is the forest composed of homogenous, managed stands or very heterogeneous, unmanaged stands?);
- the nature of the ground plot and ALS data available (e.g., do the ground plots represent the variability of the forest structure found in the area of interest?); and
- the information needs of the end user (e.g., are there specific accuracy and precision requirements?).

Selection of a modelling approach will also depend on the availability of statistical expertise to develop the predictive models.

Finally, the validation of model estimates is a critical step in model development. Typically, a certain proportion of ground plots are reserved for model validation (e.g., see Woods et al. 2011). When insufficient ground plots are available, cross-validation methods have also been applied (e.g., Næsset 2002, 2004, 2009). Ultimately, the best validation data are generated from areas as they are harvested. Scaling data, tracked by location, offer excellent opportunities to validate average tree size and volume predictions from ALS. Today, harvesters in many jurisdictions are equipped with on-board computers that record stem size, product recovery data, and GPS location for each tree that is cut—data that will be of great importance in the validation of all ALS predictions, including diameter and volume distributions.

7. Mapping

Regardless of the modelling approach used, once validated, predictive models can be applied to the entire management area using the wall-to-wall ALS metrics. Before applying the predictive equation, ensure that the common "No Data" mask has been applied to the necessary metric rasters (see Section 4.4.1). Note also that if models were developed for specific forest types or species, then the models must be applied to the appropriate stands as identified in the forest inventory. Similarly, any non-forest areas in the inventory should be masked out, as should any areas unrepresented by the ground sampling (i.e., immature forests). Once generated, these wall-to-wall rasters can be integrated into existing stand-level forest inventories. Additional stand-level generalizations and confidence intervals can then be calculated.

8. Summary

This guide summarizes best practices for the application of the area-based approach to producing an operational ALS-based enhanced forest inventory.

- The minimum ALS products required for the area-based approach are the bare earth DEM and the classified (unfiltered) ALS point cloud. The point cloud should contain *all* valid returns.
- Airborne laser scanning data that are appropriate for generating forest inventory attributes may be characterized by:
  - Small scan angles (< ±12°);
  - A *minimum* of 1 pulse per square metre (> 4 pulses per square metre for dense forests on complex terrain). Greater pulse densities may be required in more complex forest environments. Note that greater pulse densities can improve the generation of bare earth DEMs, the precision of attribute estimates, and support individual tree work and future ALS applications;
  - A sensor capable of recording a *minimum* of 2 returns per pulse (current instruments are capable of recording up to 4–5 returns per pulse); and
  - 50% overlap of adjacent flight swaths.

- Airborne laser scanning acquisition specifications should be reviewed periodically to ensure that they are in keeping with technological advances.
- Ideally, ALS data should be acquired for the entire area of interest in a single survey, at the same time, with the same instrumentation. Otherwise, special considerations are required for model development and implementation.
- It is recommended that ALS data be acquired during leaf-on conditions; however, leaf-off data may be appropriate for the area-based approach. It is important that models developed using leaf-on data not be applied to leaf-off data, and vice versa.
- A basic understanding of those factors associated with ALS instrumentation that can affect the quality of the acquired ALS data is helpful to ensure that the acquired ALS data is appropriate for this application.
- Airborne laser scanning metrics are descriptive statistics generated from the ALS point cloud.
  - Various software tools are available or users can develop their own tools in freeware packages such as R. FUSION is a recommended freeware tool developed by the US Forest Service.
  - It is recommended that the size of the grid cell used in the area-based approach match the size of the ground plots and be sufficiently large to enable the development of robust predictive models. Typical grid cell sizes are 20 x 20 m (400 m^2) or 25 x 25 m (625 m^2).
  - An area of interest, particularly a large management area, will need to be partitioned into units of "manageable size" (i.e., 5 x 5 km tiles) to enable processing of ALS metrics (which can be CPU intensive).
  - Hundreds of metrics are possible, many of which are intercorrelated. Tools such as FUSION produce a standard set of metrics. The scientific literature indicates that metrics related to height, the coefficient of variation of height, and the density of cover are the most commonly used in predictive models.
  - It is recommended that users consider applying a minimum height threshold before calculating ALS metrics to ensure non-canopy or below-canopy returns are not included in metric calculation.
  - Calculated metrics should be subject to quality assessment to generate a consistent "No Data" mask that is applied to all outputs, and to identify anomalous values/outliers.
- The collection of ground plot data is critical to building robust predictive models from the ALS.
  - It is recommended that ground plots be purpose-acquired for model development.
  - Ground plot size is very important to the area-based approach.
  - The recommended range of ground plot size for use in the area-based approach is 200–625 m2 (8–14 m radius). The optimum ground plot size from within this range can be determined by considering the need to minimize edge effects, minimize planimetric co-registration error, maximize sampling efficiency, and maximize precision and accuracy of desired attributes.
  - The use of fixed-area, circular plots is recommended.
  - Special care must be taken when measuring ground plots in areas where canopy and understorey are dense, or terrain is complex.
  - Modern mapping-grade GPS systems should result in plot positioning errors that are well within 5 m of the true location. To minimize positional error, collect a minimum of 500 GPS points per plot, use a second GPS unit as a base station and an external antenna on the roving unit, and apply differential post-processing correction after acquisition.
  - Ground plots must represent the full range of forest structural variability present in the area of interest (as captured by the ALS data).
  - It is recommended that ALS data is acquired before ground plots and that ALS metrics (typically a height metric, a canopy cover metric, and a metric that captures variability in height) are used to stratify the area of interest and develop a structurally guided sampling design.
  - Time between ALS and ground plot acquisition should be minimized and, if at all possible, will take place within the same growing season.
  - Accurate and consistent measurement of dbh and height in ground plots is essential. Species, status, and crown class should also be recorded.
- Approaches to building predictive models for forest inventory attributes from co-located ALS and ground plot measures have been based either on parametric regression or non-parametric methods such as Random Forests. Both approaches have advantages and disadvantages; the selection of the best method for a given area depends somewhat on the complexity

of the forests in the area of interest and the statistical expertise (either internal or via outsourcing) available for model development.

- Rather than using all possible metrics to build predictive models, the use of a feature-selection approach is recommended to obtain a reasonable subset of metrics for model building.
- Once developed, models should not be applied outside the range of values for which they were generated. This recommendation applies regardless of the method used to build the model (i.e., parametric or non-parametric).

9. Outlook

The best practices presented herein focus on the estimation of a basic suite of forest inventory attributes for which there is some agreement within the ALS community regarding best approaches and the quality of the results produced. Many other potential applications exist that are perhaps more experimental, less well established, or that have not yet been applied to a broad range of forest types. These include attributes associated with the prediction of species, site characteristics and quality, diameter distributions, leaf area index, age, and tree health. Moreover, in contrast to the area-based approach presented in this guide, an approach based on the identification and characterization of individual trees (i.e., the individual tree approach) is maturing in concert with increasing ALS pulse densities and computing capacities. All of this speaks to increasing opportunities for ALS in forest inventories of the future.

Interest is also growing in the generation of point clouds and DSMs from very high resolution digital aerial imagery. Semi-Global Matching is a technique used to generate an ALS-like, three-dimensional point cloud that defines the canopy surface from a stereo pair of digital images. The production of an accurate image-based point cloud requires a very precise, high-resolution, bare-earth DEM; such DEMs in forest environments are typically only available from ALS data. It has been suggested that image-based point clouds and DSMs could replace ALS in the area-based approach described in this guide. However, if the production of an image-based point cloud is predicated on the need for an accurate bare-earth DEM, then their use may be limited to those areas in which ALS data has already been acquired. The other limitation to the use of image-based point clouds is their inability to characterize the vertical structure of the forest between the outer canopy envelope and the ground. Recall that ALS pulses are capable of penetrating the forest canopy and thereby acquire additional returns at the subcanopy level, representing branches and understorey. This potentially rich source of information on vertical forest structure is not captured in the image-based point cloud, which only characterizes the outer canopy envelope. Research into the relevance and impact of the different ways in which ALS- and image-based point clouds characterize the vertical structure of forests is still in its early stages and is not yet conclusive.

10. Literature Cited

Adams, T.; Brack, C.; Farrier, T.; Pont, D.; Brownlie, R. 2011. So you want to use LiDAR? A guide on how to use LiDAR in forestry. New Zealand Journal of Forestry 55(4):19–23.

Adjorlolo, C.; Cho, M.A.; Mutanga, O.; Ismail, R. 2013. Optimizing spectral resolution for the classification of C3 and C4 grass species, using wavelengths of known absorption features. Journal of Applied Remote Sensing 6(1). doi:10.1117/1.JRS.6.063560.

Aldred, A.; Bonner, G. 1985. Application of airborne lasers to forest surveys. Agriculture Canada, Ministry of State for Forestry, Petawawa National Forestry Institute, Chalk River, Ont. Information Report PI-X-51.

American Society for Photogrammetry and Remote Sensing. 2011. LAS specification, Version 1.4–R12. 10 June 2012. http://asprs.org/a/society/committees/standards/LAS_1_4_r12.pdf (Accessed May 2013).

Andersen, H.-E.; McGaughey, R.J.; Reutebuch, S.E. 2005. Estimating forest canopy fuel parameters using LiDAR data. Remote Sensing of Environment 94(4):441–449.

Andersen, H.-E.; Reutebuch, S.E.; McGaughey, R.J. 2006. A rigorous assessment of tree height measurements obtained using airborne lidar and conventional field methods. Canadian Journal of Remote Sensing 32(5):355–366.

Avery, T.E.; Burkhart, H.E. 2002. Forest measurements. McGraw Hill, Boston, Mass.

Baltsavias, E.P. 1999. Airborne laser scanning: Basic relations and formulas. ISPRS Journal of Photogrammetric Engineering & Remote Sensing 54:199–214.

Bater, C.W.; Coops, N.C. 2009. Evaluating error associated with lidar-derived DEM interpolation. Computers & Geosciences 35(2):289–300.

Bater, C.W.; Wulder, M.A.; Coops, N.C.; Nelson, R.F.; Hilker, T.; Næsset, E. 2011. Stability of sample-based scanning LiDAR-derived vegetation metrics for forest monitoring. IEEE Transactions on Geoscience and Remote Sensing 49(6):2385–2392.

Binot, J-M.; Pothier, D.; Lebel, J. 1995. Comparison of relative accuracy and time requirement between the caliper, the diameter tape, and an electronic tree measuring fork. The Forestry Chronicle 71(2):197–200.

Bolstad, P.; Jenks, A.; Berkin, J.; Home, K.; Reading, W.H.2005. A comparison of autonomous, WAAS, real-time, and post-processed Global Positioning Systems (GPS) accuracies in northern forests. Northern Journal of Applied Forestry 22:5–11.

Breiman, L. 2001. Random forests. Machine Learning 45(1):5–32.

Chasmer, L.; Hopkinson, C.; Treitz, P. 2006. Assessing the three-dimensional frequency distribution of airborne and ground-based LiDAR data for red pine and mixed deciduous forest plots. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI-8/W2:66–70.

Cook, R.D. 1975. Influential observations in linear regression. Journal of the American Statistical Association 74(365):169–174.

Deckert, C.; Bolstad, P.V. 1996. Forest canopy, terrain, and distance effects on global positioning system point accuracy. Photogrammetric Engineering & Remote Sensing 62(3):317–321.

Demaerschalk, J.P.; Kozak, A. 1974. Suggestions and criteria for more effective regression sampling. Canadian Journal of Forest Research 4:341–348.

Disney, M.I.; Kalogirou, V.; Lewis, P.; Prieto-Blanco, A.; Hancock, S.; Pfeifer, M. 2010. Simulating the impact of discrete-return lidar system and survey characteristics over young conifer and broadleaf forests. Remote Sensing of Environment 114:1546–1560.

Edson, C.; Wing, M.G. 2012. Tree location measurement accuracy with a mapping-grade GPS receiver under forest canopy. Forest Science 58(6):567–576.

Evans, D.L.; Roberts, S.D.; Parker, R.C. 2006. LiDAR: A new tool for forest measurements. The Forestry Chronicle 82(2):211–218.

Flood, N. (editor). 2004. ASPRS guidelines: vertical accuracy reporting for Lidar data. Version 1. ASPRS Lidar Committee, American Society for Photogrammetry and Remote Sensing, Bethesda, Md. www.asprs.org/a/society/committees/lidar/Downloads/Vertical_Accuracy_Reporting_for_Lidar_Data.pdf (Accessed May 2013).

Frazer, G.W.; Hobart, G.; White, J.C.; Wulder, M.A. 2011a. Predictive modeling of forest inventory attributes using airborne LiDAR and ground-reference measurements derived from Hinton Wood Products' Permanent Growth Sample (PGS) program. Final report for Canadian Forest Service, Canadian Wood Fibre Centre, Pacific Forestry Centre, Victoria, B.C., and Hinton Wood Products, Edmonton, Alta.

Frazer, G.W.; Magnussen, S.; Wulder, M.A.; Niemann, K.O. 2011b. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. Remote Sensing of Environment 115(2):636–649.

Gatziolis, D.; Andersen, H.-E. 2008. A guide to LIDAR data acquisition and processing for the forests of the Pacific Northwest. U.S. Department of Agriculture Forest Service, Pacific Northwest Research Station, Portland, Oreg. General Technical Report PNW-GTR-768.

Gatziolis, D.; Fried, J.S.; Monleon, V.S. 2010. Challenges to estimating tree height via LiDAR in closed-canopy forests: a parable from western Oregon. Forest Science 56(2):139–155.

Gaulton, R.; Danson, F.M.; Ramirez, F.A.; Gunawan, O. 2013. The potential of dual-wavelength laser scanning for estimating vegetation moisture content. Remote Sensing of Environment 132:32–39.

Gaveau, D.L.A.; Hill, R.A. 2003. Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. Canadian Journal of Remote Sensing 29(5):650–657.

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

_____. 2009. Assessing effects of positioning errors and sample plot size on biophysical stand properties derived form airborne laser scanner data. Canadian Journal of Forest Research 39(5):1036–1052.

Guttman, N.B. 1994. On the sensitivity of sample L moments to sample size. Journal of Climate 7(6):1026–1029.

Harding, D.J.; Lefsky, M.A.; Parker, G.G.; Blair, J.B. 2001. Laser altimeter canopy height profiles: methods and validation for closed-canopy, broadleaf forests. Remote Sensing of Environment 76:283–297.

Haugerud, R.A.; Harding, D.J.; Johnson, S.Y.; Harless, J.L.; Weaver, C.S. 2003. High-resolution lidar topography of the Puget Lowland, Washington: A bonanza for earth science. GSA Today 4–10.

Hawbaker, T.J.; Gobakken, T.; Lesak, A.; Trømborg, E.; Contrucci, K.; Radeloff, V.C. 2010. Light detection and ranging-based measures of mixed hardwood forest structure. Forest Science 56(3):313–326.

Hawbaker, T.J.; Keuler, N.S.; Lesak, A.A.; Gobakken, T.; Contrucci, K.; Radeloff, V.C. 2009. Improved estimates of forest vegetation structure and biomass with LiDAR-optimized sampling design. Journal of Geophysical Research 114. doi:10.1029/2008JG000870.

Hill, R.A.; Broughton, R.K. 2009. Mapping the understorey of deciduous woodland from leaf-on and leaf-off airborne LiDAR data: a case study in lowland Britain. ISPRS Journal of Photogrammetry and Remote Sensing 64(2):223–233.

Hofton, M.A.; Rocchio, L.E.; Blair, J.B.; Dubayah, R. 2002. Validation of vegetation canopy lidar sub-canopy topography measurements for a dense tropical forest. Journal of Geodynamics 34(3–4):491–502.

Holmgren, J. 2004. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. Scandinavian Journal of Forest Research 19:543–553.

Holmgren, J.; Nilsson, M.; Olsson, H. 2003. Simulating the effects of lidar scanning angle for estimation of mean tree height and canopy closure. Canadian Journal of Remote Sensing 29(5):623–632.

Honer, T.G. 1964. The use of height and squared diameter ratios for the estimation of merchantable cubic foot volume. The Forestry Chronicle 40(3):324–331.

Houghton, R.A. 2005. Aboveground forest biomass and global carbon balance. Global Change Biology 11:945–958.

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

Hudak, A.T.; Crookston, N.L.; Evans, J.S.; Hall, D.E.; Falkowski, M.J. 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sensing of Environment 112:2232–2245.

Hyyppä, J.; Hyyppä, H.; Leckie, D.; Gougeon, F.; Yu, X.; Maltamo, M. 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. International Journal of Remote Sensing 29(5):1339–1366.

Hyyppä, J.; Yu, X.; Hyyppä, H.; Vastaranta, M.; Holopainen, M.; Kukko, A.; Kaartinen, H.; Jaakkola, A.; Vaaja, M.; Koskinen, J.; Alho, P. 2012. Advances in forest inventory using airborne laser scanning. Remote Sensing 4(5):1190–1207.

Inoue, A.; Yamamoto, K.; Mizoue, N.; Kawahara, Y. 2004. Calibrating view angle and lens distortion of the Nikon fish-eye converter FC-E8. Journal of Forest Research 9:177–181.

Jakubowski, M.K.; Guo, Q.; Kelly, M. 2013. Tradeoffs between lidar pulse density and forest measurement accuracy. Remote Sensing of Environment 130:245–253.

Järnstedt, J.; Pekkarinen, A.; Tuominen, S.; Ginzler, C.; Holopainen, M.; Viitala, R. 2012. Forest variable estimation using high-resolution digital surface model. ISPRS Journal of Photogrammetry and Remote Sensing 74:78–84.

Johnson, C.E.; Barton, C.C. 2004. Where in the world are my field plots? Using GPS effectively in environmental field studies. Frontiers in Ecology and the Environment 2(9):475–482.

Laes, D.; Reutebuch, S.; McGaughey, R.J.; Maus, P.; Mellin, T.; Wilcox, C.; Anhold, J.; Finco, M.; Brewer, K. 2008. Practical lidar acquisition considerations for forestry applications. U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center, Salt Lake City, Utah. RSAC-0111-BRIEF1.

Lambert, M.-C.; Ung, C.-H.; Raulier, F. 2005. Canadian national tree aboveground biomass equations. Canadian Journal of Forest Research 35:1996–2018.

Lefsky, M.A.; Cohen, W.B.; Harding, D.J.; Parker, G.G.; Acker, S.A.; Gower, S.T. 2002. Lidar remote sensing of above-ground biomass in three biomes. Global Ecology & Biogeography 11:393–399.

Lefsky, M.A.; Harding, D.; Cohen, W.B.; Parker, G.; Shugart, H.H. 1999. Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. Remote Sensing of Environment 67(1):83–98.

Lefsky, M.A.; Hudak, A.T.; Cohen, W.B.; Acker, S.A. 2005. Patterns of covariance between forest stand and canopy structure in the Pacific Northwest. Remote Sensing of Environment 95:517–531.

Li, Y.; Andersen, H.-E.; McGaughey, R. 2008. A comparison of statistical methods for estimating forest biomass from light detection and ranging data. Western Journal of Applied Forestry 23(4):223–231.

Lim, K.; Duhnam, V.; Lakusta, N. 2013. Weyerhaeuser Grande Prairie FMA LiDAR Project and AFRIDS. Presented to 2013 Southern Interior Silviculture Committee Winter Workshop. March 4, 2013. Kamloops, B.C. http://cif-ifc.org/uploads/Website_Assets/Lim_-_Kamloops.pdf (Accessed May 2013).

Lim, K.; Treitz, P.; Baldwin, K.; Morrison, I.; Green, J. 2003a. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. Canadian Journal of Remote Sensing 29(5):658–678.

Lim, K.; Treitz, P.; Wulder, M.; St-Onge, B.; Flood, M. 2003b. LiDAR remote sensing of forest structure. Progress in Physical Geography 27(1):88–106.

Lovell, J.L.; Jupp, D.L.B.; Newnham, G.J.; Coops, N.C.; Culvenor, D.S. 2005. Simulation study for finding optimal lidar acquisition parameters for forest height retrieval. Forest Ecology and Management 214(1–3):398–412.

Maltamo, M.; Bollandsås, O.M.; Næsset, E.; Gobakken, T.; Packalén, P. 2011. Different plot selection strategies for field training data in ALS-assisted forest inventory. Forestry 84(1):23–31.

Maltamo, M.; Eerikäinen, K.; Pitkänen, J.; Hyyppä, J.; Vehmas, M. 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. Remote Sensing of Environment 90(3):319–330.

Maltamo, M.; Packalén, P.; Yu, X.; Eerikäinen, K.; Hyyppä, J.; Pitkänen, J. 2005. Identifying and quantifying structural characteristics of heterogeneous boreal forests using laser scanner data. Forest Ecology and Management 216(1–3):41–50.

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

Magnussen, S.; Næsset, E.; Gobakken, T. 2010a. Reliability of LiDAR derived predictors of forest inventory attributes: a case study with Norway spruce. Remote Sensing of Environment 114(4):700–712.

Magnussen, S.; Tomppo, E.; McRoberts, R.E. 2010b. A model-assisted k-nearest neighbour approach to remove extrapolation bias. Scandinavian Journal of Forest Research 25(2):174–184.

McGaughey, R.J. 2013. FUSION/LDV: Software for LiDAR data analysis and visualization. Version 3.30. U.S. Department of Agriculture Forest Service, Pacific Northwest Research Station, University of Washington, Seattle, Wash. http://forsys.cfr.washington.edu/fusion/FUSION_manual.pdf (Accessed May 2013).

Means, J.E.; Acker, S.A.; Fitt, B.J.; Renslow, M.; Emerson, L.; Hendrix, C.J. 2000. Predicting forest stand characteristics with airborne scanning ALS. Photogrammetric Engineering & Remote Sensing 66(11):1367–1371.

Montgomery, D.C.; Peck, E.A.; Vining, G.G. 2006. Introduction to linear regression analysis. John Wiley & Sons, New York, N.Y.

Næsset, E. 1999. Point accuracy of combined pseudorange and carrier phase differential GPS under forest canopy. Canadian Journal of Forest Research 29:547–553.

_____. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80(1):88–99.

_____. 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. Scandinavian Journal of Forest Research 19:164–179.

_____. 2005. Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data. Remote Sensing of Environment 98(2–3):356–370.

_____. 2009. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. Remote Sensing of Environment 113(1):148–159.

_____. 2011. Estimating above-ground biomass in young forests with airborne laser scanning. International Journal of Remote Sensing 32(2):473–501.

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

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

Næsset, E.; Gobakken, T.; Holmgren, J.; Hyyppä, H.; Hyyppä, J.; Maltamo, M.; Nilsson, M.; Olsson, H.; Persson, Å.; Söderman, U. 2004. Laser scanning of forest resources: the Nordic experience. Scandinavian Journal of Forest Research 19:482–499.

Næsset, E.; Gobakken, T.; Solbert, S.; Gregoire, T.G.; Nelson, R.; Ståhl, G.; Weydahl, D. 2011. Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: a case study from a boreal forest area. Remote Sensing of Environment 115(12):3599–3614.

Næsset, E.; Jonmeister, T. 2002. Assessing point accuracy of DGPS under forest canopy before data acquisition, in the field, and after postprocessing. Scandinavian Journal of Forest Research 17:351–358.

Næsset, E.; Økland, T. 2002. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. Remote Sensing of Environment 79(1):105–115.

Nelson, R.; Short, A.; Valenti, M. 2004. Measuring biomass and carbon in Delaware using an airborne profiling LiDAR. Scandinavian Journal of Forest Research 19:500–511.

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

Ni-Meister, W.; Lee, S.; Strahler, A.H.; Woodcock, C.E.; Schaaf, C.; Yao, T.; Ranson, K.J.; Sun, G.; Blair, J.B. 2010. Assessing general relationships between aboveground biomass and vegetation structure parameters for improved carbon estimate from lidar remote sensing. Journal of Geophysical Research 115. doi:10.1029/2009JG000936.

Nyström, M.; Holmgren, J.; Olsson, H. 2012. Prediction of tree biomass in the forest–tundra ecotone using airborne laser scanning. Remote Sensing of Environment 123:271–279.

Optech. 2013. PEGASUS HD 500: summary specification sheet. www.optech.ca/pdf/ALTM-Pegasus-SpecSheet-120924-WEB.pdf (Accessed May 2013).

Ørka, H.O.; Næsset, E.; Bollandsås, O.M. 2010. Effects of different sensors and leaf-on and leaf-off canopy conditions on echo distributions and individual tree properties derived from airborne laser scanning. Remote Sensing of Environment 114(7):1445–1461.

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

Parker, R.C.; Evans, D.L. 2004. An application of LiDAR in a double-sample forest inventory. Western Journal of Applied Forestry 19(2):95–101.

Popescu, S.C.; Wynne, R.H.; Nelson, R.F. 2002. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. Computers and Electronics in Agriculture 37(1–3):71–95.

R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org (Accessed May 2013).

Resource Information Management Branch. 2011. Part 7. General specifications for acquisition of ALS data. Sustainable Resource Management, Government of Alberta, Edmonton, Alta. Unpublished report.

Reutebuch, S.E.; Andersen, H.-E.; McGaughey, R.J. 2005. Light detection and ranging (LiDAR): an emerging tool for multiple resource inventory. Journal of Forestry 103:286–292.

Reutebuch, S.E.; McGaughey, R.J. 2008. LiDAR: an emerging tool for multiple resource measurement, planning, and monitoring. Western Forester 53(2):1–2.

Riaño, D.; Meier, E.; Allgöwer, B.; Chuvieco, E.; Ustin, S.L. 2003. Modeling airborne laser scanning data for the spatial generation of critical forest parameters in fire behavior modeling. Remote Sensing of Environment 86:177–186.

Richards, F.J. 1959. A flexible growth function for empirical use. Journal of Experimental Botany 10(2):290–301.

Segal, M.R. 2004. Machine learning benchmarks and random forest regression. Center for Bioinformatics & Molecular Biostatistics, University of California, San Francisco, Calif. Technical Report. www.escholarship.org/uc/item/35x3v9t4 (Accessed May 2013).

Sharma, M.; Parton, J. 2007. Height–diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. Forest Ecology and Management 249(3):187–198.

Stephens, P.R.; Kimberley, M.O.; Beets, P.N.; Paul, T.S.H.; Searles, N.; Bell, A.; Brack, C.; Broadley, J. 2012. Airborne scanning LiDAR in a double sampling forest carbon inventory. Remote Sensing of Environment 117:348–357.

Strobl, C.; Boulesteix, A.-L.; Kneib, T.; Augustin, T.; Zeileis, A. 2008. Conditional variable importance for random forests. BMC Bioinformatics 9. doi:10.1186/1471-2105-9-307.

Ter-Mikaelian, M.T.; Korzukhun, M.D. 1997. Biomass equations for sixty-five North American tree species. Forest Ecology and Management 97(1):1–24.

Thomas, V.; Oliver, R.D.; Lim, K.; Woods, M. 2008. LiDAR and Weibull modeling of diameter and basal area. The Forestry Chronicle 84(6):866–875.

Toloşi, L.; Lengauer, T. 2011. Classification with correlated features: unrealibility of feature ranking and solutions. Bioinformatics 27(14):1986–1994.

Treitz, P.; Lim, K.; Woods, M.; Pitt, D.; Nesbitt, D.; Etheridge, D. 2012. LiDAR sampling density for forest resource inventories in Ontario, Canada. Remote Sensing 4(4):830–848.

Ung, C.-H.; Bernier, P.Y.; Guo, X. 2008. Canadian national biomass equations: new parameter estimates that include British Columbia data. Canadian Journal of Forest Research 38:1123–1132.

Vahukonen, J.; Hakala, T.; Suomalainen, J.; Kaasalainen, S.; Nevalainen, O.; Vastaranta, M.; Holopainen, M.; Hyyppä, J. 2013. Classification of spruce and pine trees using active hyperspectral LiDAR. IEEE Geoscience and Remote Sensing Letters. doi:10.1109/LGRS.2012.2232278.

Vastaranta, M.; Kankare, V.; Holopainen, M.; Yu, X.; Hyyppä, J.; Hyyppä, H. 2012. Combination of individual tree detection and area-based approach in imputation of forest variables using airborne laser data. ISPRS Journal of Photogrammetry and Remote Sensing 67:73–79.

Villikka, M.; Packalén, P.; Maltamo, M. 2012. The suitability of leaf-off airborne laser scanning data in an area-based forest inventory of coniferous and deciduous trees. Silva Fennica 46(1):99–110.

Wasser, L.; Day, R.; Chasmer, L.; Taylor, A. 2013. Influence of vegetation structure on lidar-derived canopy height and fractional cover in forested riparian buffers during leaf-off and leaf-on conditions. PLoS ONE 8(1): e54776. doi:10.1371/journal.pone.0054776.

Wehr, A.; Lohr, U. 1999. Airborne laser scanning: an introduction and overview. ISPRS Journal of Photogrammetry and Remote Sensing 54:68–82.

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

Wing, M.G. 2008. Consumer-grade Global Positioning Systems (GPS) receiver performance. Journal of Forestry 106(4):185–190.

Wing, M.G.; Eklund, A. 2007. Performance comparison of a low-cost mapping grade global positioning systems (GPS) receiver and consumer grade GPS receiver under dense forest canopy. Journal of Forestry 105(1):9–14.

Wing, M.G.; Eklund, A.; Sessions, J.; Karsky, R. 2008. Horizontal measurement performance of five mapping-grade global positioning system receiver configurations in several forested settings. Western Journal of Applied Forestry 23(3):166–171.

Wing, M.G.; Karsky, R. 2006. Standard and real-time accuracy and reliability of a mapping-grade GPS in a coniferous western Oregon forest. Western Journal of Applied Forestry 21(4):222–227.

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

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

Wulder, M.A. 1998. Optical remote sensing techniques for the assessment of forest inventory and biophysical parameters. Progress in Physical Geography 22(4):449–476.

Wulder, M.A.; Bater, C.W.; Coops, N.C.; Hilker, T.; White, J.C. 2008. The role of LiDAR in sustainable forest management. The Forestry Chronicle 84(6):807–826.

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

Yu, X.; Hyyppä, J.; Kaartinen, H.; Maltamo, M. 2004. Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. Remote Sensing of Environment 90(4):451–462.

Yu, X.; Liang, X.; Hyyppä, J.; Kankare, V.; Vastaranta, M.; Holopainen, M. 2013. Stem biomass estimation based on stem reconstruction from terrestrial laser scanning point clouds. Remote Sensing Letters 4(4):344–353.

Zakrzewski, W.T. 1999. A mathematically tractable stem profile model for jack pine in Ontario. Northern Journal of Applied Forestry 16(3):138–143.

Zhao, K.; Popescu, S.; Nelson, R. 2009. Lidar remote sensing of forest biomass: a scale-invariant estimation approach using airborne lasers. Remote Sensing of Environment 113:182–196.

Appendix 1. Airborne Laser Scanning Data Acquisition

Details Required for Requests for Proposals

I. Input Specifications

A. System Calibrations

- Copy of manufacturer's calibration for entire system (laser unit, IMU, GPS).
- Evidence of calibration of entire unit within 6 months of data acquisition, using same parameters as for the designated survey to identify and correct systematic errors and confirm horizontal and vertical accuracies.

B. GPS Base Stations

- Two base stations within 30 km of each other and within 40 km of survey area.
- A report detailing the number and accuracy of base stations (must include number, location, published co-ordinates of active GPS reference stations/ monuments).

C. Flight Parameters

- Selected to satisfy the required point density (pulse repetition frequency, flying altitude) and minimize occurrence of data voids.
- Data voids are areas with no laser pulses within an area that is four times greater than the post-spacing of the data. Unacceptable causes of data voids are system malfunctions, data dropouts, flight-line data gaps. Acceptable data voids are those caused by water bodies.
- Overlap between flight lines must be > 50%.
- One cross-flight line flown at same altitude for quality-control purposes.
- Scan angle must not exceed ± 12° from nadir.

II. Acquisition Specifications

A. Airborne Laser Scanning Data Collection

- Data shall be collected at a density which meets the requirements of the end-use application.
- Surveys are required to collect, at a minimum, the *x*, *y*, and *z* position in metres, such that these positions can be delivered in the required co-ordinate system and horizontal and vertical datums, intensity, return number/number of returns, and GPS time for each point.
- Atmospheric conditions shall be clear and free of mist, fog, low cloud cover, and smoke.
- The ground shall be free of snow, unless otherwise specified.
- The collection area must be buffered by a minimum of 100 m.
- Spatial distribution of points is expected to be uniform and free from clustering.

B. Data Accuracy

- The recommended methodology for determining and reporting vertical and horizontal accuracies for ALS data is as per the American Society for Photogrammetry and Remote Sensing guidelines (Flood [editor] 2004): www.asprs.org/a/society/committees/lidar/Downloads/Vertical_Accuracy_Reporting_for_Lidar_Data.pdf

C. Flight Data

- Flight date and time
- Flight altitude(s)
- Airborne laser scanning system scan angle, scan rates, and pulse rates
- Times for activating/deactivating the ALS system
- Position dilution of precision (PDOP)
- Height of instrument (before and after flight)
- On-board antenna offsets
- Any site obstructions at GPS base station(s)
- Airborne and ground site GPS receiver types and serial numbers
- Ground site GPS station monument names and stability (survey control ID card)
- Flight staff

D. Data Processing

- All file labelling and cross-indexing
- Analyst name responsible for processing and product generation
- List of auxiliary information used during processing of ALS to generate products delivered
- List of software used for processing the laser, GPS, and IMU data

- Combined separation plot of kinematic GPS aircraft positioning used for blending GPS/IMU data, including standard deviation, minimum and maximum separation of the forward/reverse Kinematic GPS solution
- Post-processed Kinematic GPS solution plot showing fixed integer/float solution
- Post-processed PDOP plot during data acquisition time
- All computations done in double precision
- The data produced shall be referenced horizontally to the North American Datum of 1983 (Canadian Spatial Reference System, with epoch defined) (NAD83[CSRS-epoch]). Vertically, the data shall be referenced to the Canadian Geodetic Vertical Datum of 1928 (CGVD28). Data shall be projected in UTM, native zone, unless otherwise specified.

E. Bare Earth Model

- The DEM shall represent the bare-earth surface.
- The contractor shall remove elevation points on bridges, buildings, and other structures and on vegetation from the ALS data to create the bare-earth DEM.
- The contractor shall produce a bare-earth DEM with the minimum grid spacing, no greater than 1 m, in eastings and northings. The number of points per square metre shall be equal to or greater than the specified pulse density.
- The method of ground classification shall be reported.
- The method for interpolating the DEMs shall be reported.
- Model/software used shall be reported, as well as any modifications to work flow required to account for specific vegetation or terrain conditions.
- The Contractor shall classify raw data points from sidelap and overlap fields of separate flight lines.

F. Quality Assurance/Quality Control

- Quality assurance and control includes reviews of flight assignments and completeness of supporting data.
- The Contractor shall provide a report on the QA/QC procedure as well as any kinematic GPS data collected for the purpose of calibration/quality control. This data shall include co-ordinate information, date and time information, and PDOP.
- In addition, the quality assurance and control procedure shall include a results section that describes how the specifications were met for the delivery area. This section must contain a statistical analysis of the point densities, as well as the ASCM RMSE analysis, anomalies report, and neighbouring ALS coverage evaluation, which are further described in "Quality Control Products." These must be summarized and referenced to the "General Specifications for Acquisition of ALS Data."

III. Output Specifications

A. Data Quality

1. Artefacts
    - Artefacts are regions of anomalous elevations or oscillations and ripples within the DEM data, and result from systematic errors or environmental conditions. They may result from malfunctioning sensors, poorly calibrated instrumentation, adverse atmospheric conditions, or processing errors.
    - When present, the contractor shall provide an analysis of the effects of the artefacts on DEM accuracy. The analysis shall include a description of the causes (contributing sources) of the artefacts and a description of the steps to eliminate them.
2. Adjoining/Overlapping Existing DEMs/DSMs
    - Elevations of bare earth DEMs and full feature DSMs along the edge of the project area that are adjoining or overlapping existing ALS DEMs/DSMs shall be consistent with those of the existing ALS DEMs/DSMs. The elevations of the existing DEMs/DSMs shall be held fixed and the new DEMs and DSMs adjusted to fit the existing DEMs/DSMs.

B. Data Format and Naming Convention

- Specifies tiling convention (i.e., by NTS mapsheet) and number of tiles per mapsheet
- Naming convention (e.g., PD_20KNO_YYYMMDD.EXT) where:

  PD = product (e.g., BE for bare earth DEM, PC for point cloud)

  20K = 1:20 000 tile

  NO = tile number

  YYYYMMDD = date of acquisition

  EXT = file extension
- The DEMs and DSMs shall be georeferenced and submitted in ArcInfo Grid ASCII format, as well as ESRI GRID format.

- The shaded relief images derived from the DEM and DSM shall be georeferenced and be submitted in uncompressed Geo TIFF format.
- Airborne laser scanning intensity image shall be georeferenced and uncompressed in Geo TIFF format.
- The point cloud data shall be submitted in LAS Specification, version 1.0 or higher, American Society for Photogrammetry and Remote Sensing in LAS format (http://asprs.org/a/society/committees/standards/LAS_1_4_r12.pdf).

C. Media

- Specifications for data delivery; typically, a hard drive of a specified minimum capacity (e.g., 250 GB).

D. Metadata

- Specification for provision of metadata (e.g., in XML format).
- Templates are typically provided to the contractor.

Collection of the metadata shall be in compliance with Federal Geographic Data Committee Content Standards for Digital Geospatial Metadata (CSDGM FGDC-STD-001-199: www.fgdc.gov/metadata/documents/workbook_0501_bmk.pdf)

E. Reports

1. System Calibration Report
    - Procedure
    - Map/diagram
    - Date flown
    - Software used
    - Roll/pitch/heading/linear biases/mirror scale
    - Residual plot showing standard deviation of ALS points in both flight lines
    - Accuracy of the ALS co-ordinates in horizontal and vertical
2. Flight Report
    - Mission date
    - Time
    - Flight altitude
    - Airspeed
    - Scan angle and rate
    - Laser pulse rates
    - Airborne laser scanning beam dispersion: lateral point density along the swath, forward point density at swath ends, forward point spacing at nadir position, nominal point density
    - Times for activating/deactivating the ALS system
    - Position dilution of precision
    - Height of instrument (before and after flight)
    - On-board antenna offsets
    - Weather conditions
    - Ground conditions
    - Information about GPS-derived flight tracks
    - Detailed description of final flight-line parameters
    - Include ground truth and complementary data
    - Other pertinent information
3. Ground Control Report
    - GPS station monument names and stability
    - Methodology used
    - Geodetic ties
    - Accuracy of ground control
    - Geoid model used for deriving orthometric elevations
    - Any site obstructions at GPS base station(s)
    - Airborne and ground site GPS receiver types and serial numbers
4. Airborne Laser Scanning System Data Report
    - Field of view
    - Scan rate
    - Number of returns recorded
    - Intensity
    - Swath overlap
    - Data methods used including the treatment of artefacts
    - Final ALS pulse and scan rates
    - Scan angle
    - Capability for multiple returns from single pulses
    - Accuracy and precision of the ALS acquired
    - Accuracy and precision of the GPS/IMU solution obtained
    - Accuracy of the topographic surface products
    - Companion imagery if acquired during the mission
    - Data processing procedures for selection of posting and orthometric values of *x, y,* and *z* co-ordinates for ALS returns
    - Any other data deemed appropriate

F. Projection

- All deliverables shall conform to the projection, datum, and co-ordinate system specified in the contract. Each file shall be organized to facilitate data manipulation and processing.
- Horizontal datum shall be NAD83(CSRS); vertical datum shall be CGVD28.

G. Deliverables

1. Reports (as specified above)
2. Aircraft trajectories (SBET files)
    - Aircraft position (easting, northing, elevation) and attitude (heading, pitch, roll), and GPS time recorded at regular intervals of 1 second or less. May include additional attributes.
    - SBET.OUT, ASCII, and ESRI shapefile (.shp) format
3. Airborne laser scanning point cloud
    - Classified point data collected during the ALS survey, which must include (at a minimum) information for each point about the 3-dimensional position (*X* and *Y* in metres and elevation in metres, the classification value (ASPRS standard), intensity value, return number, and GPS time
    - Must include all valid returns
    - Data should be classified so that it is possible to identify the exact points used in the creation of the bare earth DEM and DSM (i.e., ground and non-ground)
    - Data files in LAS format (version to be reported)
    - Each file ≤ 2 GB
    - All returns (first, second, third, etc., and last) in each laser pulse as recorded
4. Bare earth DEM
    - Evenly spaced grid of points with elevations derived from where the *X*,*Y* position transects the TIN (Triangulated Irregular Network) created from the classified ground measured points (vegetation and artificial structures removed). Tiles are to be delivered in 32-bit GeoTIFF format, with floating point values to at least millimetre precision and associated world (.tfw) files, in addition to ESRI tm ASCII grid format.
5. Metdata (see above)

H. Optional Deliverables

1. Digital Surface Model (DSM) or full feature DEM
    - Evenly spaced grid of points with elevations derived from the highest (typically the first) laser return within each grid cell. This is the highest captured surface feature (i.e., trees and shrubs in vegetated areas, top of buildings and ground level in open areas).
2. Intensity image (or *normalized* intensity image)
3. Shaded relief
4. Contours

For additional information, please see:

Evans, J.S.; Hudak, A.T.; Faux, R.; Smith, A.M.S. 2009. Discrete return LiDAR in natural resources: recommendations for project planning, data processing, and deliverables. www.fs.fed.us/rm/pubs_other/rmrs_2009_evans_j002.pdf (Accessed May 2013).

Haugerud, R.; Curtis, T.; Madin, I.; Martinez, D.; Nelson, S.; Nile, E.; Reutebuch, S. 2008. A proposed specification for LiDAR surveys in the Pacific Northwest. http://pugetsoundlidar.ess.washington.edu/proposed_PNW_lidar_spec-1.0.pdf (Accessed May 2013).

Laes, D.; Reutebuch, S.; McGaughey, B.; Maus, P.; Mellin, T.; Wilcox, C.; Anhold. J.; Finco, M.; Brewer, K. 2008. Practical LiDAR-acquisition considerations for forestry applications. U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center, Salt Lake City, Utah. RSAC-0111-BRIEF1. Available online www.fs.fed.us/eng/rsac/lidar_training/pdf/0111-Brief1(3).pdf (Accessed May 2013).

Watershed Sciences Inc. 2010. Minimum LiDAR data density considerations for the Pacific Northwest. www.oregongeology.org/sub/projects/olc/minimum-lidar-data-density.pdf (Accessed May 2013).

Appendix 2. Airborne Laser Scanning Point Cloud Metrics

This appendix details elements that should be considered for inclusion in a Request for Proposals (RFP) for the generation of ALS metrics.

1. Software
    - Either identify a specific tool for metric generation (i.e., FUSION) or provide detailed specifications for metrics (see point #4 following and Section 4.4).
    - The advantage of using an off-the-shelf tool such as FUSION is a standardized set of metrics and transparent methods of calculating them.
2. Grid cell size
    - Specify the required grid cell size, which is typically determined by the size of the ground plot. Ideally, the ground plot and the grid cell will be of a similar size. Grid cells are typically 20 × 20 m (400 m^2) or 25 × 25 m (625 m^2) (see Section 4.2).
3. Tiling schema
    - Given the size of ALS point cloud files, most metric processing software will require the area of interest to be divided into a series of tiles to facilitate metric generation (e.g., 5 x 5 km tiles).
    - Each tile will be processed separately, and the resulting metrics will then be recombined to provide complete coverage of the area of interest.
    - Specify the origin of tiling schema, if it is critical that the generated metrics align with existing data. Tiling schema can be supplied to the contractor as a shapefile.
    - See Section 4.3 of this guide.
4. Metrics
    - If using a standard tool such as FUSION to produce metrics, it is easy to specify which metrics should be generated.
    - Otherwise, a detailed list (including equations) should be provided to the contractor so that no ambiguity exists concerning how metrics are to be calculated.
    - Specify a minimum height threshold for points used to compute metrics (note that in FUSION, all points are used to generate ***density*** metrics even if a minimum height threshold is specified. Minimum height thresholds are used for generating height metrics).
5. Quality assurance for metrics
    - Generation and application of a consistent "No Data" mask (see Section 4.4.1).
    - Methods for identifying and reporting anomalous values/outliers (see Section 4.4.1).

Appendix 3. A Sample FUSION Workflow for Metric Calculation

This appendix details a sample FUSION workflow for metric calculation.

1. Compile ALS point cloud files (LAS format) for the area of interest into a single working directory.
2. Convert all bare earth DEMs into .DTM format for use in FUSION.
    - FUSION reads surface models stored in PLANS DTM format. This is a binary format that requires less storage and can be read faster than an ASCII equivalent. Also, the binary format enables the consistent calculation of the byte position of the elevation for a specific grid point in the model file.
    - Requests for proposals should specify that ALS bare earth DEMs be delivered by the contractor in ASCII format.
    - FUSION has a tool called ASCII2DTM that will do the conversion.
3. Compile all of the bare earth DEM .DTM files into the same working directory as the ALS point cloud files (point #1 previous).
4. Run FUSION *Catalog* command over the entire directory and index all LAS files.
5. Get *X* and *Y* extents of processing tiles (from tiling schema) and convert to a script for batch processing (need: tile ID, xmin, ymin, xmax, ymax).
6. Using the information from the tiling schema generated (point #5 above), generate a script to batch process the metrics using the *GridMetrics* command:

    GridMetrics computes a series of descriptive statistics for a ALS data set. Output is an ASCII text file

with comma separated values (CSV format). Field headings are included and the files are easily read into database and spreadsheet programs. Each row in the file represents the metrics for an individual grid cell. By default, *GridMetrics* computes statistics using both elevation and intensity values in the same run.

Syntax (Table A3.1): *gridmetrics [switches] groundfile heightbreak cellsize outputfile datafile1 datafile2 …*

Example: *gridmetrics /minht:2 /nointensity / gridxy:XMIN,YMIN,XMAX,YMAX *.DEM 2 25 tile001 *.las*

7. Revise the ASCII header generated by the *Gridmetrics* command:
   a. change "xllcenter" to "xllcorner" and "yllcenter" to "yllcorner";
   b. replace the negative infinity values that get calculated for the L-moment metrics to 0; and
   c. both a and b can be automated using text manipulation software (i.e., Perl).
8. Convert the CSV files generated by the *GridMetrics* command into ASCII format rasters using the CSV2Grid command in FUSION.
9. Merge the output ASCII files into a single output using the *MergeRaster* command in FUSION.
10. Convert the merged ASCII files into ArcGIS GRID format (if working in an ArcGIS environment).
11. Set the projection within ArcGIS.
12. Generate a master "No Data" mask from metrics and apply to all output (see Section 4.4.1).

Table A3.1. Detailed syntax for the FUSION *gridmetrics* command.

Function	Value	Description
switches	/minht:2	A 2 m minimum height is used to compute metrics.
	nointensity	Do not compute metrics using intensity values (only elevation values).
	/gridxy:XMIN,YMIN,XMAX,YMAX */gridxy:507675,5835200,511675,5839200*	Force the output grid to have the extent specified (corresponds to tiling schema).
groundfile	*.DTM	Name for bare earth DEM. Use a wildcard and FUSION will identify those .DTM files that correspond to the extent of the processing tile (as specified by switch /gridxy).
heightbreak	2	A 2 m height break for cover calculation.
cellsize	25	Desired grid cell size in the same units as ALS data (25 m).
outputfile	tile001	Base name for output file (output = tile001.csv).
datafile	*.las	Input LAS file. Use a wildcard and FUSION will identify which LAS files correspond to the extent of the processing tile (as specified by switch/gridxy).

For more information about the Canadian Forest Service, visit our website at **cfs.nrcan.gc.ca** or contact any of the following Canadian Forest Service establishments

cfs.nrcan.gc.ca

Canadian Forest Service Contacts

1 Atlantic Forestry Centre
P.O. Box 4000
Fredericton, NB E3B 5P7
Tel.: (506) 452-3500 Fax: (506) 452-3525
cfs.nrcan.gc.ca/regions/afc

Atlantic Forestry Centre – District Office
Sir Wilfred Grenfell College Forestry Centre
University Drive
Corner Brook, NF A2H 6P9
Tel.: (709) 637-4900 Fax: (709) 637-4910

2 Laurentian Forestry Centre
1055 rue du P.E.P.S., P.O. Box 3800
Sainte-Foy, PQ G1V 4C7
Tel.: (418) 648-5788 Fax: (418) 648-5849
cfs.nrcan.gc.ca/regions/lfc

3 Great Lakes Forestry Centre
P.O. Box 490 1219 Queen St. East
Sault Ste. Marie, ON P6A 5M7
Tel.: (705) 949-9461 Fax: (705) 759-5700
cfs.nrcan.gc.ca/regions/glfc

4 Northern Forestry Centre
5320-122nd Street
Edmonton, AB T6H 3S5
Tel.: (403) 435-7210 Fax: (403) 435-7359
cfs.nrcan.gc.ca/regions/nofc

5 Pacific Forestry Centre
506 West Burnside Road
Victoria, BC V8Z 1M5
Tel.: (250) 363-0600 Fax: (250) 363-0775
cfs.nrcan.gc.ca/regions/pfc

6 Headquarters
580 Booth St., 8th Fl.
Ottawa, ON K1A 0E4
Tel.: (613) 947-7341 Fax: (613) 947-7396
cfs.nrcan.gc.ca/regions/nrc

Canadian Wood Fibre Centre
A virtual research centre of
the Canadian Forest Service,
Natural Resources Canada
cfs.nrcan.gc.ca/subsite/cwfc