Remote Sensing of Environment 156 (2015) 322–334

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Review

Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data

Marc Bouvier [a,*], Sylvie Durrieu [a], Richard A. Fournier [b], Jean-Pierre Renaud [c]

[a] UMR TETIS Irstea-Cirad-AgroParisTech/ENGREF, Maison de la Télédétection en Languedoc-Roussillon, 500, rue J.F. Breton BP 5095, 34196 Montpellier Cedex 05, France
[b] Centre d'Applications et de Recherches en Télédétection (CARTEL), Département de géomatique appliquée, Université de Sherbrooke, Sherbrooke, Québec J1K 2R1, Canada
[c] Office National des Forêts, 11 rue Ile de Corse, 54000 Nancy, France

ARTICLE INFO

Article history:
Received 22 July 2014
Received in revised form 2 October 2014
Accepted 3 October 2014
Available online xxxx

Keywords:
Canopy structure
Forest inventory
Airborne LiDAR
Model generalization
Area-based approach
Foliage density profile

ABSTRACT

This study proposed modifying the conceptual approach that is commonly used to model development of stand attribute estimates using airborne LiDAR data. New models were developed using an area-based approach to predict wood volume, stem volume, aboveground biomass, and basal-area across a wide range of canopy structures, sites and LiDAR characteristics. This new modeling approach does not adopt standard approaches of stepwise regression using a series of height metrics derived from airborne LiDAR. Rather, it used four metrics describing complementary 3D structural aspects of the stand canopy. The first three metrics were related to mean canopy height, height heterogeneity, and horizontal canopy distribution. A fourth metric was calculated as the coefficient of variation of the leaf area density profile. This fourth metric provided information on understory vegetation. The models that were developed with the four structural metrics provided higher estimation accuracy on stand attributes than models using height metrics alone, while also avoiding data over-fitting. Overall, the models provided prediction error levels ranging from 12.4% to 24.2%, depending upon forest type and stand attribute. The more homogeneous coniferous stand provided the highest estimation accuracy. Estimation errors were significantly reduced in mixed forest when separate models were developed for individual stand types (coniferous, mixed and deciduous stands) instead of a general model for all stand types. Model robustness was also evaluated in leaf-off and leaf-on conditions where both conditions provided similar estimation errors.

Contents

* Corresponding author. Tel.: +33 4 67 54 87 19; fax: +33 4 67 54 87 00.
E-mail address: marc.bouvier@teledetection.fr (M. Bouvier).

http://dx.doi.org/10.1016/j.rse.2014.10.004
0034-4257/

1. Introduction

Total aboveground woody volume (hereafter referred to as wood volume) and the stem merchantable volume of a tree (hereafter referred to as stem volume) are key forest inventory attributes that are required by forest managers. Reliable mapping of both volumes facilitates the implementation of sustainable management strategies and practices. These practices enable logging be optimized while contributing to forest ecosystem preservation and climate change mitigation (Picard, Saint-André, & Henry, 2012). Biomass of forest stands, including both aboveground and belowground components, is another attribute that is required to improve our knowledge of the carbon cycle. Wood volume, stem volume and aboveground biomass (AGB) are interdependent, and there are strong correlations among these three attributes (Brown & Lugo, 1984; Fang, Liu, & Xu, 1996). Most forest inventory programs estimate wood and stem volumes from field measurements of stem diameter from sampled trees. Tree heights are also measured, but generally only from a subset of individuals, as height measurements are more difficult and costly to collect than those for stem diameter (Avery & Burkhart, 2001; Kangas & Maltamo, 2006). Forest inventories generally use random or systematic sampling schemes (Scott & Gove, 2002). In both approaches, a limited number of plots are inventoried because this work is both costly and time-consuming. Furthermore, field measurements can only be performed in areas that are accessible to field crews. Remote sensing has the potential to provide quick and accurate measurements of stand attributes over large areas at a much lower cost than with traditional inventory practices. Using remote sensing data, coupled with a small number of field measurements, can thus be an effective solution for overcoming the aforementioned drawbacks of field measurements, while providing accurate and timely information on several key forest attributes.

Optical and radar remote sensing have been widely used to map forest structural attributes and biophysical parameters (Franklin, 2001; Le Toan, Beaudoin, Riom, & Guyon, 1992; Leboeuf, Fournier, Luther, Beaudoin, & Guindon, 2012). Assessment of wood volume and ABG remains challenging, especially for forests with high AGB levels (Cohen & Spies, 1992; Leboeuf et al., 2012). In a recent study, Zolkos, Goetz, and Dubayah (2013) analyzed a large set of published studies on AGB assessment from remote sensing data. The authors have brought to light the considerable variability in the accuracy of AGB predictions according to both forest environment and remote sensing data used.

LiDAR has emerged as one of the very promising technologies for forest applications and its potential for forest attribute estimation at several scales is widely acknowledged (Leeuwen & Nieuwenhuis, 2010; Næsset, 2004; Nelson, Krabill, & Tonelli, 1988). These active systems provide precise distance measurements that are based on elapsed time between the emission of a laser pulse and the measurement of the return signal. The spatial position of all the recorded returns upon the Earth's surface is calculated from both the position and orientation of the LiDAR system. These are measured using a differential global positioning system (DGPS) and an accurate inertial unit. The resulting point cloud is processed to assess diverse LiDAR metrics. LiDAR systems have either a small footprint (0.1–0.3 m) or a large footprint (8–70 m) (Lim, Treitz, Wulder, St-Onge, & Flood, 2003). Both of these systems have been used to estimate forest inventory attributes (Lefsky, Cohen, et al., 1999; Lefsky et al., 2002; Means et al., 2000; Næsset, 2002). Airborne Laser Scanner (ALS) is a LiDAR system combined with a scanning system. ALS is thus able to record data over a swath, the width of which depends upon both the scanning angle and flight altitude. These systems can be used to provide wall-to-wall coverage of areas of interest (Wulder et al., 2012). Two approaches can be applied using small-footprint discrete return ALS: (1) a tree-based approach (Li, Guo, Jakubowski, & Kelly, 2012; Maltamo, Eerikäinen, Pitkänen, Hyyppä, & Vehmas, 2004; Popescu, Wynne, & Nelson, 2003; Véga et al., 2014); or (2) an area-based approach (ABA) (Means et al., 2000; Næsset, 1997, 2002; Nelson et al., 1988). Current ALS that offer high sampling rates can be used to estimate single-tree attributes in a tree-based approach. However, individual tree segmentation algorithms frequently can be difficult to implement because they generate omission and inclusion errors when individualizing trees (Falkowski et al., 2006; Goerndt, Monleon, & Temesgen, 2010; Véga & Durrieu, 2011). Overlapping tree crowns can add confusion when identifying individual trees (Gleason & Im, 2011). Therefore, detection performance regarding individual trees is reduced in complex stands. ABAs are commonly used to generate maps of forest attributes in diverse forest biomes, temperate forests (Hall, Burke, Box, Kaufmann, & Stoker, 2005; Zhao & Popescu, 2009), boreal forests (Lim, Treitz, Baldwin, Morrison, & Green, 2003; Næsset & Gobakken, 2008), tropical forests (Asner, 2009; Kronseder, Ballhorn, Böhm, & Siegert, 2012), and savannah woodlands (Lucas et al., 2006). In such approaches, stand attribute estimations are computed from the statistical relationships between plot-level LiDAR metrics that are commonly extracted from point cloud data and stand attributes, which are derived from field plots. In general, numerous candidate metrics are derived from point height distributions at the plot-level, e.g., maximum and mean height values, percentiles of the distributions, and canopy densities (Næsset, 2002). Metrics that provide the greatest explanation are then selected, with only a few remaining in the final model (Hall et al., 2005; Lim & Treitz, 2004; Lim, Treitz, Wulder, et al., 2003; Patenaude et al., 2004). ABAs have been used to develop models for specific forested areas and for specific species or a specific group of species, which has led to a huge number of different models using diverse LiDAR metrics.

Currently, ABAs that are used to predict stand attributes have two major drawbacks, despite their proven usefulness for forest inventory and mapping (Næsset, 2002). The first drawback stems from the fact that candidate metrics generated from LiDAR data are known to be strongly inter-correlated (Chen, 2013). Furthermore, too many candidate metrics are generated, which complicates metric selection and the development of robust models (Hall et al., 2005; Khan, Van Aelst, & Zamar, 2007; Magnussen, Næsset, Gobakken, & Frazer, 2012). The second drawback is that metrics used to describe stand structure are generally derived from the vertical distribution of LiDAR returns. Indeed, these metrics do not sufficiently take into account several other canopy characteristics, including horizontal canopy heterogeneity. To overcome these drawbacks, new metrics have been identified to enhance the description of tree spatial distributions and, consequently, stand attribute predictions. For instance, canopy volume and canopy cover metrics have proved to be useful for AGB and volume estimation, in addition to height metrics (Chen, Gong, Baldocchi, & Tian, 2007; Hall et al., 2005; Kim et al., 2009; Næsset & Gobakken, 2008). These metrics take into account horizontal vegetation structures within model predictions. Other metrics have been estimated from the vegetation density profile. Profiles were integrated to estimate the distribution and total amount of foliage. These metrics have proved to be valuable for stand and tree attribute estimation (Allouis, Durrieu, Véga, & Couteron, 2013; Lefsky, Harding, Cohen, Parker, & Shugart, 1999). In such approaches, new metrics are generally identified, which are deemed to be unbiased and consistently meaningful (Magnussen & Boudewyn, 1998). The use of meaningful metrics that describe the 3D structural aspects of the stand canopy could help overcome the current limitations of ABA.

Model generalization is a key question that must be addressed to predict stand structural attributes (Chen, 2010). Few studies have assessed the potential of some LiDAR metrics for the prediction of diverse stand attributes or across diverse forest area types. Lefsky, Harding, et al. (1999) found that the quadratic mean canopy height (QMCH) improved estimates of AGB and basal area (BA). Lim, Treitz, Baldwin, et al. (2003) used mean laser height that was computed from filtered LiDAR returns, based upon the intensity return values, to estimate wood volume, AGB and BA. Magnussen et al. (2012) proposed a conceptual model for predicting tree-size-related forest attributes

from the height and variance of the canopy that was estimated from first returns. Lefsky et al. (2002) proposed a single equation for estimating AGB in three distinct biomes, which meant that when these models were applied to several forest types, they usually only required a calibration for each forest area type. The same metrics and model shape were maintained when moving from one area to another. A model generalization framework has yet to be developed using metrics that describe stand spatial heterogeneity, such as those reflecting both tree height and spatial distributions.

The goal of this study was to develop predictive models of forest stand attributes from ALS data using ABA, which offered high generalization potential. Therefore, the predictive models 1) should be applicable to several useful forest attributes in diverse forest types, 2) use the same LiDAR metrics, and 3) have a single model shape. Two specific objectives were identified. First, based on a conceptual approach, we identified four complementary LiDAR metrics for characterizing 3D structure of diverse stands. We assessed their relevance for implementation of prediction models. Second, we tested the generalizability of the model by evaluating its performance: (1) for the prediction of several variables, i.e., wood volume, stem volume, AGB and BA; (2) over a wide range of canopy structure; and (3) in leaf-off and leaf-on conditions.

2. Materials

2.1. Study sites

Three study sites were selected that represented the diversity of French forests. These included monospecific plantations of coniferous trees (coniferous site), a set of multi-layered deciduous stands (deciduous site), and a mountainous forest environment containing coniferous, deciduous and mixed stand types (mountainous site). Study site descriptions are summarized in Table 1.

The coniferous site was comprised of a 60 km^2 area that was located in the Landes region of southwestern France (44.40° N, 0.50° W). Climate of the region is oceanic (Joly et al., 2010). The Landes forest is characterized by nutrient-poor, sandy soil with a flat topography. The site is dominated by mono-specific stands of maritime pine (*Pinus pinaster* Aiton) in even-aged plantations.

The deciduous site was a 60 km^2 area that was located in the Observatoire Pérenne de l'Environnement in the Lorraine region of northeastern France (48.53° N, 5.37° E). Regional climate is semi-continental, and subject to an oceanic influence (Joly et al., 2010). The Lorraine forest is fragmented and highly managed. The forests in the selected areas are largely multi-layered and composed of many deciduous species, but are dominated by European beech (*Fagus sylvatic* L.), European hornbeam (*Carpinus betulus* L.), and Sycamore maple (*Acer pseudoplatanus* L.).

The mountainous site was comprised of a 1200 km^2 area that was located in the Vosges region in northeastern France (48.03° N, 7.08° E). The region's climate is semi-continental (Joly et al., 2010). The area is characterized by hilly topography, with elevations ranging from about 120 m to 1420 m. The Vosges forest is composed of pure coniferous stands, pure deciduous stands, and mixed stands. Stands of that area are typically heterogeneous and uneven-aged dominated by European beech, European silver fir (*Abies alba* Miller), and Norway spruce (*Picea abies* (L.) H. Karst).

2.2. Field plot data

Data were obtained for a series of field plots that were located in each study site for ground truthing; these values were later used to calibrate and validate the predictive models with ALS data. Field plot data were collected from 39 circular plots at the coniferous site from April to June 2011. For trees that had a diameter at breast height (DBH, 1.3 m) greater than 17.5 cm, 31 plots were established, each of which had a radius of 15 m. When trees were encountered that had a smaller diameter, the plot radius was reduced to 6 m; 8 of these plots were established. Species, tree status (i.e., dominant, co-dominant, dominant), height, and DBH were collected for all trees with DBH >7.5 cm. Field-plot center positions were measured using differential GPS (Leica GPS 120, Switzerland) with sub-meter accuracy. The GPS unit was placed in an opening adjacent to the plot and away from dense cover. A total station (Leica TS02, Switzerland) was used to measure the exact distance to each plot center.

Field plot data were collected in 28 circular plots (15 m radius) at the deciduous site from November 2009 to February 2010. Species, tree status and DBH were collected for all trees with DBH >17.5 cm and within a 6 m radius subplot for trees with a DBH between 7.5 cm and 17.5 cm. Six dominant tree heights were measured on each plot.

The same protocol was applied to 92 circular plots at the mountainous site with measurements taken from February to April 2012. Field plot center positions were measured using differential GPS (Trimble GeoXT 3000 or GeoXT 6000, USA) for both the deciduous and mountainous sites. Differential corrections were then applied using the closest fixed antenna for position accuracy, which was assumed to vary between 0 and 4 m. Plot position was further adjusted manually using LiDAR data.

Wood volume, stem volume, AGB and BA were either measured or estimated for each field plot. Field measurements and forest attributes are summarized in Table 1. BA (m^2/ha) was estimated as the cross-sectional area of an individual tree at breast height (1.3 m above the ground), as calculated from the measured DBH. Wood volume and stem volume were measured in m^3/ha. Wood volume is equal to the sum of total tree volumes for the stand including stumps and branches, whereas stem volume is equal to the sum of trunk volumes with DBH >7 cm. AGB is the dry mass (Mg/ha) of the summed tree elements that are above ground, including stems, branches, and leaves. Wood volume and AGB were derived using allometric equations that included tree DBH (cm) and height (H, m) measurements as explanatory variables for all the trees that were located within the plot and which were then rescaled to a per hectare basis. Heights were estimated for trees without height measurements using plot- or site-specific exponential regressions with DBH measurements. The allometric equations that were used to estimate wood volume were constructed for the

Table 1
Summary of the field plot data for wood volume, stem volume, AGB (Aboveground Biomass) and BA (Basal Area) that were used as reference values to calibrate and validate models.

	Coniferous site			Deciduous site			Mountainous site		
	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max
Slope (%)	1.1	3.3	12.2	2.1	9.2	29.4	2.2	38.7	76.4
Mean tree height (m)	5.0	18.4	29.5	9.9	17.3	39.5	5.5	17.5	35.4
Stem density (tree/ha)	142	464	1415	72	492	1691	85	740	2667
Gini coefficient	0.10	0.19	0.42	0.16	0.41	0.73	0.10	0.39	0.68
Shannon index	0	0.07	0.76	0	0.93	2.19	0	1.10	2.82
Wood volume (m^3/ha)	11.1	204.2	538.0	12.4	194.3	358.1	5.3	453.0	1027.5
Stem volume (m^3/ha)	4.8	187.8	508.1	7.9	136.5	247.9	2.8	396.0	965.9
AGB (Mg/ha)	9.5	84.3	193.6	6.9	123.8	221.3	2.6	205.9	401.1
BA (m^2/ha)	4.9	22.3	42.2	1.7	18.7	30.6	0.9	36.9	67.1

EMERGE project (Deleuze et al., 2013) and were species- and site-specific. Reference stem volumes were predicted from wood volumes using expansion factors that are provided by Longuetaud et al. (2013). Reference AGBs were predicted using specific equations that have been taken from Genet et al. (2011) and Shaiek et al. (2011) for European beech and maritime pine, respectively. Equations that were taken from Hounzandji, Jonard, Nys, Saint-André, and Ponette (2014) were used for pedunculate oak (*Quercus robur* L.) and sessile oak (*Quercus petraea* (Mattuschka) Lieblein). Species-specific equations were missing for other species present on the sites, and therefore, the global wood density database (Zanne et al., 2009) was used to predict reference AGBs from wood volumes.

In addition to measuring the stand attributes, the Gini coefficient was calculated from tree BAs for each plot in the three study sites. The Gini coefficient is used to measure tree size heterogeneity within a forest stand (Lexerød & Eid, 2006). This index has a minimum value of zero, expressing perfect equality when all trees are of equal size; it takes a theoretical maximum value of one, indicating greater size diversity when all trees except one have a value of zero. We also reported tree diversity with each plot with the Shannon index (Magurran, 2004), which accounts for both the abundance and richness of tree species. Shannon indices ranged from zero, when only a single species is present in a stand to $\log_2(S)$ (where S is the number of species in a stand), when each tree is a different species.

2.3. LiDAR data

Data were collected at all sites from small-footprint airborne LiDAR. The coniferous and deciduous sites were sampled using a LMS-Q560 (Riegl, Austria). ALS data were acquired in April 2011 on the coniferous forest with a pulse density of 8.1 pulses/m^2. Two acquisitions were made on the deciduous site. The first, under leaf-off conditions with a high pulse density of 17.9 pulses/m^2, was started in December 2009 and completed in March 2010. The second was acquired in October 2010 under leaf-on conditions, with a pulse density of 20.7 pulses/m^2. The mountainous site was sampled in March and April 2011 using an ALTM 3100 (Optech, Canada), with a lower pulse density of 3.4 pulses/m^2. Additional data specifications for the ALS data are given in Table 2.

Data pre-processing was performed for each study area by the data providers, Sintegra (France) and IGN (France), for the Riegl and Optech data, respectively. Ground points were classified using the *TIN*-iterative algorithm (Axelsson, 2000). They were used to produce raster digital terrain model (DTM) with a 1 m cell grid. The DTM was used to subtract the ground elevation from LiDAR point elevations to obtain their above-ground heights, thereby removing the effect of topography from the LiDAR metrics. The resulting LiDAR point clouds were extracted according to the spatial extent of each field plot.

3. Methods

We adopted an ABA to develop a generic model for stand attribute predictions. Data processing was performed in the R statistical environment (http://www.r-project.org/). We first defined four complementary ALS metrics that were linked to both the vertical and horizontal distributions of the forest elements. Our approach aimed to eliminate the search step that is commonly used to identify the best metric from amongst a multitude of potential LiDAR metrics. Four metrics were identified to construct a single equation. Regression analyses were applied for model calibration to predict diverse stand attributes, i.e., wood volume, stem volume, AGB and BA, in diverse forest types, i.e., the coniferous, deciduous and mountainous sites. Models were validated to evaluate the generalization capacity of the model by comparing its performances for each stand attribute that was predicted over each study site.

3.1. Selection of LiDAR metrics

Amongst the numerous height metrics that have been used for stand attribute predictions, heights that are measured by LiDAR have been shown to be closely correlated with the stand attributes upon which our study focuses (Dubayah, Knox, Hofton, Blair, & Drake, 2000; Lefsky, Harding, et al., 1999; Means et al., 2000). Individual tree heights are not identified in the ABAs. Instead, an average canopy height (μ_{CH}) was estimated from the LiDAR point cloud as the most important predictor variable (Asner et al., 2012; Lefsky et al., 2002). In general, μ_{CH} is calculated from LiDAR data using first return heights. Non-tree returns must be excluded from the point cloud to develop robust predictive models (Nilsson, 1996). Thus, the first selected metric is μ_{CH} and was calculated after applying a threshold value of 2 m for removing ground and understory laser hits prior to its calculation. Nevertheless, there may be other LiDAR metrics that could improve the estimation of stand attributes, as described below.

Most allometric equations are non-linear, including those used in this study dealing with the relationships between height and diameter, volume or AGB for a tree (Chave et al., 2005; Schumacher & Hall., 1933). When shifting to the stand-level, the predicted attributes correspond to the sum of individual tree attributes in the stand, whereas our first selected metric, μ_{CH}, is an average value of canopy height. Jensen (1906) stated that for a non-linear function, a function of an average value $f(\overline{x})$ does not equal the average of the same function $\overline{f(x)}$. This inequality implies that for any non-linear function, the variance has consequences that cannot be inferred from the average value alone. Ruel and Ayres (1999) have used Jensen's inequality in environmental issues and have argued that whenever biological systems involve non-linear responses, the description and interpretation of environmental data should include an explicit consideration of the variance. In our case, this amounts to saying that a stand

Table 2
Technical specifications for the ALS sensor and data that were acquired for the three test sites.

	Coniferous site	Deciduous site		Mountainous site
		Leaf-off	Leaf-on	
Date of survey	April 2011	Dec. 2009 and Mar. 2010	October 2010	March–April 2011
ALS sensor	Riegl LMS-Q560	Riegl LMS-Q560	Riegl LMS-Q560	Optech ALTM3100
Wavelength (nm)	1550.0	1550.0	1550.0	1064.0
Scan angle	29.5	29.5	29.5	16.0
Beam divergence (mrad)	0.5	0.5	0.5	0.8
Footprint diameter (cm)	27.5	27.5	27.5	120
Ground speed (m/s)	50.0	50.0	45.0	80.0
Recovery rate (%)	40.0	60.0	60.0	55.0
Repetition rate (Hz)	100.0	100.0	100.0	71.0
Pulse density (pulses/m^2)	8.0	18.0	20.5	3.5
Flight altitude (m)	550.0	550.0	550.0	1500.0

cannot be summarized as a set of "mean trees" when using height to assess stand attributes. Within the ABAs, an indicator of tree height heterogeneity at plot level thus should be used in addition to an indicator of mean height to link structural information that is derived from LiDAR at the plot-level with a sum of tree attributes that are assessed from field measurements. We therefore chose to use variation in canopy height (σ^2_{CH}) to characterize tree height heterogeneity. The σ^2_{CH} was calculated as the variance of first return heights, considering only returns above 2 m, thereby providing a second metric that was potentially useful for improving predictive models.

Tree attributes that were measured in the field are related to the whole-plot area, accounting for several stand densities in the model or for the presence of gaps, which entails evaluating the proportion of open area in the plot. A third metric could be a variable describing the whole-plot area. The proportion of open area in the plot accounts for stand densities. Therefore, a gap fraction (P_f) was calculated from the LiDAR data as the ratio of the number of first returns below a specified height threshold to the total number of first returns (Hopkinson et al., 2013; Morsdorf, Kötz, Meier, Itten, & Allgöwer, 2006). When calculated in this manner, P_f is related to penetration of light through the canopy and is also well correlated with fractional cover (Hopkinson & Chasmer, 2009). P_f remains a good indicator of the horizontal distribution of the canopy when linking structural LiDAR metrics to field measurements, and, as such, it was considered as another metric. This is the third on our list of metrics that was required to improve estimation models.

The three metrics discussed so far, i.e., μ_{CH}, σ^2_{CH} and P_f, only refer to the structural properties at the top of the canopy. Sub-dominant trees would not be taken into account in a model if only these three metrics were used. Due to the capacity of LiDAR to penetrate into the vegetation and measure the understory, information on sub-dominant trees is contained in LiDAR-derived vegetation profiles. A fourth metric was derived from these vegetation profiles, and it has the potential to further enhance the estimation models.

Leaf area index (LAI) has been found to be correlated with forest ecosystem productivity (Reich, 2012; Waring, 1983). LAI can be retrieved using light interception-based techniques (Bréda, 2003; Vose, Clinton, Sullivan, & Bolstad, 1995) and, as the LiDAR signal behaves in a manner similar to direct sunlight, some authors have developed approaches to assess LAI from LiDAR data (Hopkinson et al., 2013; Tang et al., 2012, 2014). We thus chose to use the properties of LAI, and in particular, its vertical profiles, as a fourth metric that was linked to vegetation stratification. To our knowledge, no study has yet used the measurements that can be derived from LAI profiles obtained from multi-echo ALS data as a predictor of wood volume, stem volume, AGB or BA assessments in ABAs. We subsequently explain how LAI profiles were computed to define an additional metric.

LAI and the gap fraction, P, can be linked using a probabilistic approach using Beer–Lambert theory for light penetration into the canopy (Campbell & Norman, 1989; Nilson, 1971, 1999; Ross, 1981). In the following equation, it can be shown that:

$$P(\theta) = e^{-G(\theta,\alpha)LAI/\cos\theta} \tag{1}$$

where θ is the view zenith angle, α is the leaf angle distribution, $P(\theta)$ is the gap fraction, and $G(\theta,\alpha)$ is the G-function, which corresponds to the fraction of foliage projected onto the plane normal to the zenith direction. Assuming that leaves are randomly distributed and that their individual size is small when compared to the canopy, the gap fraction can be considered to be equivalent to transmittance (Bréda, 2003). For direct-lighting conditions, cumulative LAI at a given canopy depth z, LAI_z, can be linked to the fraction of PAR (Photosynthetically Active Radiation) that is transmitted at canopy depth z, using the Beer–Lambert law expressed as (Bréda, 2003; Martens, Ustin, & Rousseau, 1993; Vose et al., 1995; Waring, 1983):

$$LAI_z = \frac{\ln\left(\frac{I_z}{I_0}\right)}{k} \tag{2}$$

where I_z is PAR transmittance at canopy depth z, I_0 is total incoming PAR, and k is the extinction coefficient. Vose et al. (1995) indicated that it is convenient to assume a random distribution of the foliage and a spherical distribution of leaf angles. The Beer–Lambert equation is relatively insensitive to violations of these assumptions. From Eq. (1), the expression for k should be:

$$k = k(\theta,\alpha) = \frac{G(\theta,\alpha)}{\cos(\theta)} \tag{3}$$

where k-values have been estimated comparing LAI direct measurements with light interception rate measurements that have been made by several authors. According to the values that have been previously reported (Bréda, 2003; Martens et al., 1993; Solberg, Næsset, Hanssen, & Christiansen, 2006; Vose et al., 1995), k can vary from 0.28 to 0.67, considering both coniferous and broad leaved stands. For a canopy with a spherical distribution of leaf inclination angles, k is well approximated by 0.5 (Martens et al., 1993). Tang et al. (2014) reported that this value allows an adequate representation of many real canopy types and assumes a constant projected foliage surface with respect to inclination angle. It has been used by several authors to retrieve LAI or LAI profiles from LiDAR data (Solberg et al., 2009; Tang et al., 2014). LAI at a given z can therefore be calculated using the following adaptation of the Beer–Lambert law (Hopkinson et al., 2013; Lovell, Jupp, Culvenor, & Coops, 2003):

$$LAI(z) = -\frac{\ln P(z)}{k} \tag{4}$$

where $P(z)$ is the gap fraction from the top of the canopy to a given height z. The gap fraction was estimated as previously described, by the ratio between the number of returns below a specified height threshold and the total number of returns, as suggested by Morsdorf et al. (2006) and Hopkinson et al. (2013).

LAI profiles were calculated using a k-value of 0.5 for all study sites. The choice of k-value may affect the LAI value. However, we assumed that this choice would not significantly affect the LAI profile shapes. Adopting this working hypothesis is practical as we intended to focus upon the analysis of vertical profile shapes and that of their relative changes from one area to another. Therefore, this approximation should not be a hindrance to the development of our model. LAI profiles were calculated from gap fractions at each height interval with a given thickness dz. This required assessing the number of laser beams that actually reached the layer $z + dz$ and those that went through the layer $[z; z + dz]$. Eq. (4) was thus modified, and P_i and LAI_i for the layer i between z and $z + dz$ were calculated as:

$$P_i = \frac{N_{[0;z]}}{N_{total} - N_{[0;z+dz]}} \tag{5}$$

$$LAI_i = -\frac{\ln(P_i)}{k} \tag{6}$$

with $N_{[0;z]}$ being the number of returns below z, N_{total} is the total number of returns, $N_{[0;z\,+\,dz]}$ is the number of returns below $z + dz$, and $k = 0.5$. Eq. (6) corrects for occlusion, providing a realistic vertical

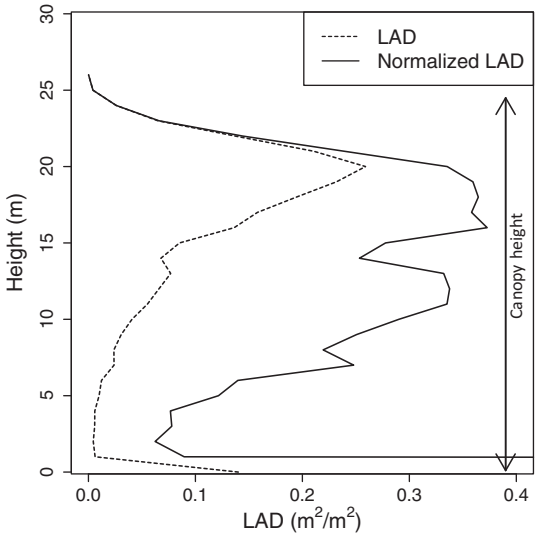

Fig. 1. LAD (leaf area density) profile (dashed line) and normalized LAD profile (solid line). The LAD profile was corrected from occlusion effects to produce a normalized LAD profile.

profile of LAI. We divided the resulting LAI_i values by dz, thereby obtaining a profile of leaf area density, LAD. Fig. 1 shows how efficient this correction was in providing enhanced information on the undergrowth and also shows its strong normalization power. Variation within the vertical LAD profile contains information on the vertical structure of a stand, and provides information on vertical heterogeneity and understory vegetation. A normalized measure of vertical dispersion of foliage density through the canopy is given by the coefficient of variation for LAD (Cv_{LAD}). We therefore adopted Cv_{LAD} as our fourth metric. Cv_{LAD} is defined as the ratio of the standard deviation to the mean of normalized *LAD*. Cv_{LAD} was calculated without LAD values under 2 m height. For example, a stand with a single layer of even-aged individuals, with most of the leaves located towards the upper canopy, leads to a LAD profile with a peak in its upper part and a high Cv_{LAD}. Conversely, a highly stratified stand leads to nearly constant vertical LAD values and a low Cv_{LAD}.

3.2. Establishment of the model

The multiplicative power model form has been widely adopted for stand attribute predictions (Lim & Treitz, 2004; Lim, Treitz, Baldwin, et al., 2003; Means et al., 1999; Næsset, 2002; Zianis & Mencuccini, 2004). Most allometric equations that link structural variables to tree attributes, e.g., AGB and volumes, are also using multiplicative power models (Chave et al., 2005; Schumacher & Hall., 1933; Ung, Bernier, & Guo, 2008). Therefore, we adopted this model form to predict several stand attributes with the four selected metrics:

$$\hat{y} = \beta_0\ \mu_{CH}^{\beta_1}\ \sigma_{CH}^{2\,\beta_2}\ P_f^{\beta_3}\ Cv_{LAD}^{\beta_4} \tag{7}$$

where $\hat{y}$ is the fitted value of wood volume, stem volume, AGB or BA. The βs are the regression model parameters. A linear relationship amongst log-transformed variables was applied. Log-transformation of stand attributes and LiDAR metrics was used to accommodate non-linearity of response. This also allows normality of residuals and stabilized variances of stand attributes to be achieved. Models were fitted using maximum likelihood analysis. We back-transformed the final model by exponentiating both sides of the log-log regression model. As systematic bias is introduced by log-transformation, the application of a correction factor (CF) was necessary (Baskerville, 1972; Finney, 1941; Sprugel, 1983). Like the protocols that were used by Hall et al. (2005) and Ferster, Coops, and Trofymow (2009), the $\hat{y}$ values were finally multiplied by the following CF:

$$CF = \exp\left(\frac{SEE^2}{2}\right) \tag{8}$$

where SEE is the standard error of regression estimates, with p being the number of parameters in the final model and n the number of field plots:

$$SEE = \sqrt{\frac{\sum_{i=0}^{n}\left(\log y_i - \log \hat{y_i}\right)^2}{n-p}} \tag{9}$$

3.3. Model validation

For an unbiased assessment of the predictive capacity of a model, it is recommended that the reference dataset be to create a validation dataset that is independent of the calibration dataset (Snee, 1977). Following this recommendation, a subset of field plot data has been used to support regression model validation. In our study, insufficient training field plots were available to provide an independent validation dataset for each study area. Therefore, a leave-one-out cross-validation method (LOOCV) was applied to evaluate the accuracy of the predictive models (Picard & Cook, 1984). With this method, all but one field plot are used to calibrate the model, and the remaining field plot is used for validation. This process is repeated until all the field plots are used once as validation data. Model coefficients are averaged from all of the iterations to obtain the final model. Similarly, LOOCV enables the assessment of model goodness-of-fit by averaging statistical estimators of model accuracy that were computed at each step of the LOOCV. We used three estimators: the coefficient of determination (R^2); relative standard deviation (RSD); and the mean percentage error (MPE). R^2 is the fraction of variance that is explained by the model. RSD is a representative measure of overall forecast quality. RSD corresponds to the standard error divided by the mean and is expressed as a percentage, thereby facilitating model comparison. MPE indicates forecast bias and is also expressed as a percentage:

$$MPE = \frac{100}{n}\sum_{i=1}^{n}\frac{(y_i - \hat{y}_i)}{y_i} \tag{10}$$

where y_i is the reference value of wood volume, stem volume, AGB or BA, and $\hat{y}_l$ is the value predicted for each step i of the LOOCV.

3.4. Analysis of the generalization capacity of the model

The performance of the proposed models and their generalization were evaluated in several ways. First, model generalization capacity was assessed globally by examining model performances when predicting several stand attributes for a wide range of forest environments, i.e., for the three selected study sites. The relevance of the selected LiDAR metrics that were used to build the model was assessed by analyzing inter-variable correlations and their individual contributions to the model for all of the predicted stand attributes. Considering the *P*-value that was given by the partial Fisher tests (Fisher, 1935) helped us to determine the significance of the contributions made by the selected explanatory variables to the model. The individual contribution of each metric to the model was measured by comparing *P*-values between the model using all four LiDAR metrics and four alternative models that were based on three metrics, alternatively removing μ_{CH}, σ^2_{CH}, P_f, and Cv_{LAD}.

Second, model generalization was evaluated according to stand characteristics. The effects of stand complexity on the predictive

capacity of the model were further analyzed using the available datasets that include various species compositions and a gradient of structural complexity when considering the three study sites. A first assessment of stand complexity effects on model efficiency was made from the comparison of results that were obtained for the simplest structure with mono-specific homogeneous plantation stand, i.e., the coniferous site, to gradually more complex sites, i.e., the deciduous stand and the mountainous site. The mountainous site was used for additional useful comparisons, given that it contained coniferous, deciduous, and mixed stands. The number of field plots at this site (92 plots) enabled us to split the dataset into three subsets that were based on forest type: 33 plots with >75% coniferous cover were labeled as coniferous stands, 36 plots with >75% deciduous cover as deciduous stands, and the 23 remaining plots were considered as mixed stands. This classification was chosen to comply with the French NFI classification that is used to produce stand type maps (IGN, 2012). A general model for all stands in the mountainous site was first established. It was validated for the whole dataset and then applied to each data subset that was taken separately so as to validate the global model for each stand type. Also, three models were produced for this site, one for each stand type. Using these four models on the mountainous site, and the models that were obtained for the coniferous and the deciduous sites, several comparisons were made. Through these comparisons, we sought to determine if stand- or species-specific regression models are needed when various stand types or several species coexist on a study site. We compared the characteristics and performances of coniferous models to those of deciduous models and the characteristics of two coniferous models (coniferous site and the model made from coniferous stands on the mountainous site) and two deciduous models (leaf-on model of the deciduous site and the model that was constructed for deciduous plots on the mountainous site).

Third, model generalization was assessed according to the relative effects of leaf-off and leaf-on conditions in deciduous stands. Leaf-off and leaf-on acquisition conditions substantially affected the ability of the LiDAR signal to penetrate the canopy. We evaluated how this affected the predictive capacity of the models for the deciduous site.

4. Results

4.1. Validation of model shape through the global evaluation of model generalization level

Four stand attributes (wood volume, stem volume, AGB, or BA) were predicted for each study site, i.e., the coniferous, deciduous and mountainous sites with a variety of stand types, using the same model shape and the same four LiDAR metrics (μ_{CH}, σ^2_{CH}, P_f, and Cv_{LAD}).

The analysis of the correlations between the four metrics that were used to construct the prediction models revealed high between-site differences (Table 4). The correlation between σ^2_{CH} and Cv_{LAI} was not statistically significant in both deciduous and mountainous sites (r of −0.09 and 0.02, respectively). However, σ^2_{CH} and Cv_{LAD} were the most closely correlated metrics in the coniferous site ($r = 0.68$). Values of μ_{CH} and σ^2_{CH} in the coniferous site were also correlated (r = 0.67). Smaller correlations were observed between μ_{CH} and σ^2_{CH} in the deciduous and mountainous sites ($r = 0.43$ and $r = 0.49$, respectively). Considering each of the 6 pairs of metrics, excluding the pair (μ_{CH}, σ^2_{CH}), there is at least one site for which each pair exhibits a correlation with an absolute value less than 0.32. In addition, the analysis of P-values revealed the relevance of the four selected LiDAR metrics in stand attribute models (Table 5). The metric μ_{CH} appears to have the greatest explanatory power. A value of $P = 0.01$ for BA in the coniferous site was judged sufficiently low to consider μ_{CH} as the most significant variable. The metrics σ^2_{CH} and P_f were also highly significant in the coniferous site. However, P-values of Cv_{LAD} metric exceeded 0.1 (ranging from 0.15 to 0.24) for the four stand attributes, which suggests that this metric was not significant for the model of the coniferous site. In the deciduous site, the most significant metrics were μ_{CH} ($P < 0.001$), Cv_{LAD} ($P = 0.01$–0.02; $P = 0.04$ for BA) and σ^2_{CH} ($P = 0.03$–0.07; $P = 0.19$ for BA). P_f was not deemed significant for the deciduous site, given $P > 0.2$ (0.2–0.28). In the mountainous site, the four metrics had $P < 0.10$. To conclude, each metric was significant for the prediction of at least one stand attribute on one of the three study sites.

Table 4
Inter-variable correlations (r) between the four LiDAR metrics in the three study sites.

r		Coniferous site (39 field plots)	Deciduous site (leaf-on) (28 field plots)	Mountainous site (92 field plots)
μ_{CH}	σ^2_{CH}	0.67	0.43	0.49
μ_{CH}	P_f	0.51	0.55	0.06
μ_{CH}	Cv_{LAD}	0.43	−0.31	0.32
σ^2_{CH}	P_f	0.37	−0.01	−0.21
σ^2_{CH}	Cv_{LAD}	0.68	−0.09	0.02
P_f	Cv_{LAD}	−0.03	−0.61	−0.37

All of the stand attribute models that were constructed using the four LiDAR metrics and the defined model shape provided R^2 values that ranged from 0.59 to 0.98, with RSD ranging between 12.4% and 24.2% (Table 3). Wood volume, stem volume and AGB models provided higher R^2 than BA, with all R^2 being greater than 0.80, with the exception of for AGB on the mountainous site ($R^2 = 0.77$; Fig. 2). Negative biases (MPE) were found for all models, ranging from −1.39% to −11.57%. Standard errors (RSD) were similar for each study site, with variations of less than 4% amongst stand attributes.

4.2. Impact of stand complexity on model quality

Effects of stand structural complexity on the predictive capacity of the model were assessed for the three study sites using Gini coefficients. Mean Gini coefficients ranged from 0.19 for the coniferous site, to 0.41 for the deciduous site, and 0.39 for the mountainous site. The model that was developed on the coniferous site, which is the site with the simplest stand structure, provided the best predictions. Stand attributes were assessed for the coniferous site, with R^2 values ranging from 0.84 to 0.98, and RSDs ranging from 12.4% to 15.0%. Models that were applied to the deciduous site (leaf-on) led to prediction errors (RSD) ranging from 17.4% to 19.6% for the four stand attributes. Last, models that were developed to predict stand attributes in the mountainous site had the highest RSD, ranging from 21.9% to 24.2%, which was not surprising as this site had the highest stand complexity. Indeed, Gini

Table 3
Goodness-of-fit expressed with the error estimators, which were associated with the prediction models, and estimated from the LOOCV (leave-one-out cross-validation).

	Coniferous site(39 field plots)			Deciduous site (28 field plots)						Mountainous site (92 field plots)		
				Leaf-off			Leaf-on					
	R^2	RSD (%)	MPE (%)	R^2	RSD (%)	MPE (%)	R^2	RSD (%)	MPE (%)	R^2	RSD (%)	MPE (%)
Wood volume	0.98	12.42	−1.53	0.87	17.37	−7.96	0.86	19.36	−7.59	0.82	21.86	−8.77
Stem volume	0.95	14.58	−2.79	0.88	16.90	−7.40	0.89	17.08	−6.69	0.81	24.19	−11.57
AGB	0.94	12.86	−1.39	0.86	18.09	−7.40	0.85	19.43	−7.32	0.77	22.26	−8.12
BA	0.84	14.96	−1.85	0.81	19.61	−8.39	0.78	20.67	−8.11	0.59	23.74	−7.40

Table 5
Individual contribution of each metric to the model. *P*-values of the four metrics were assessed for the four stand attributes in the three study sites.

P-values	Coniferous site				Deciduous site (leaf-on)				Mountainous site			
	μ_{CH}	σ^2_{CH}	P_f	Cv_{LAI}	μ_{CH}	σ^2_{CH}	P_f	Cv_{LAI}	μ_{CH}	σ^2_{CH}	P_f	Cv_{LAI}
Wood volume	<0.001	<0.001	<0.001	0.240	<0.001	0.042	0.258	0.016	<0.001	0.003	<0.001	0.088
Stem volume	<0.001	<0.001	<0.001	0.154	<0.001	0.027	0.279	0.018	<0.001	0.002	<0.001	0.050
AGB	<0.001	<0.001	<0.001	0.221	<0.001	0.073	0.199	0.014	<0.001	0.062	0.022	0.102
BA	0.011	0.006	<0.001	0.150	<0.001	0.194	0.218	0.041	<0.001	0.066	<0.001	0.092

P-values were classified between bounds 0, 0.001, 0.01, 0.05 and 0.10, yielding four intervals. $P \leq 0.001$: very highly significant; $0.001 < P \leq 0.01$: highly significant; $0.01 < P \leq 0.05$: significant; $0.05 < P \leq 0.1$: barely significant.

coefficients calculated for this site were in the same range as on the deciduous site, but the more varied stand compositions were an additional source of complexity when compared to the deciduous site. The corresponding goodness-of-fit statistics are summarized in Table 3. For the mountainous site, goodness-of-fit statistics were assessed individually for three subsets of plots, which corresponded respectively to mountainous coniferous, mixed, and deciduous stands (Table 6; Fig. 3). Both coniferous and mixed stands on the mountainous site had lower RSD (ranging from 16.3% to 20.9%) when compared with all stands on the mountainous site combined (RSD ranging from 21.9% to 24.2%). Deciduous stands at the mountainous site had the highest RSD, ranging from 23.2% to 31.4%. The latter previously analyzed model was developed and validated on the mountainous site using the combined plots. When models were developed and applied to three subsets of plots, the resulting models for each stand type and for the four stand attributes had higher R^2 and lower RSD values than when a single model was used (Table 6). Models that were produced for each stand type exhibited the effects of the latter on model quality. Best results were obtained in the mixed stands (RSD ranging from 15.6% to 18.9%), followed by the coniferous stands on the mountainous site (RSD ranging from 16.5% to 18.9%). Both mixed and coniferous stand models had similar R^2 values. The poorest performance was obtained with the model that had been developed for the deciduous stands on the mountainous site, where RSD values ranged from 19.5% to 22.8%. These values were slightly higher than those obtained with the model that had been developed on the deciduous site, but with lower R^2 (Table 6).

4.3. Effects of leaf-on versus leaf-off conditions on model quality

Regression models were developed for deciduous site from two LiDAR acquisitions under leaf-off and leaf-on conditions. Leaf-off conditions provided slightly better predictions than leaf-on conditions, with RSD values that were higher by 0.2% to 2.0% but similar R^2 values, when compared with the leaf-off condition, depending upon the stand attribute.

5. Discussion

We proposed a model generalization framework that was based on a set of complementary LiDAR metrics that could describe canopy structure. Metrics were used to develop models to predict stand attributes using ABA. Current predictive models of stand attributes using ALS metrics are usually based on a subset of height percentile and point density metrics from amongst a large initial set of metrics. This is a common approach, which has been used in many other studies (Lim & Treitz, 2004; Næsset, 2002, 2004). The use of the four metrics that we proposed (μ_{CH}, σ^2_{CH}, P_f, and Cv_{LAD}) improved the capacities of models to predict four stand attributes when compared with current predictive models. Indeed, according to Chen et al. (2007), metric selection from the stepwise regression of current predictive models usually suffers from over-fitting problems, whereas metric selection that is based on a priori knowledge might enhance model generalization capacities. This study explored how the use of only four ALS metrics, which were associated with different structural components, improved prediction models of diverse stand attributes. Previous LiDAR studies have reported considerable variability in the accuracy of stand attribute predictions. It is difficult to find studies justifying specific accuracy levels for these predictions. However, a small number of studies have defined accuracy requirements for AGB and have stated that AGB should be predicted within 20% of field estimates (F. G. Hall et al., 2011; Zolkos et al., 2013). Zolkos et al. (2013) conducted an analysis of reported AGB estimate accuracies that had been extracted from more than 70 published studies using different remote sensing platforms and systems and over diverse forest types; 34 studies provided AGB predictions from ALS data with an average RSD of 27%. Our models provided AGB predictions with a RSD ranging from 12.9% to 22.3%, depending upon the forest type.

Most studies regarding the development of stand attribute models from ALS data have focused upon coniferous forests, which are usually characterized by simple stand structures when compared with deciduous or mixed stands (Lim, Treitz, Wulder, et al., 2003). The model protocol that was developed in this study investigated a wide range of canopy structures from mono-specific coniferous stands that were composed of even-aged plantations to mixed stands that were located in mountainous terrain. The accuracy of AGB estimates that were analyzed by Zolkos et al. (2013) also applied to diverse forest types; 14 studies from temperate coniferous forests exhibited an average RSD of 28.7%, while 6 studies in temperate deciduous forests had an average RSD of 31.0%. For both our coniferous and deciduous sites, the results were significantly better than those reported by Zolkos et al. (2013), with RSD of 12.9% and 18.1%, respectively. Both sites had several coexisting species within a plot, as indicated by Shannon indices of up to 2.82 (Table 1). Since the relationships between stand structure and the biophysical attributes are species-dependent, it is not surprising that model prediction accuracy decreases in multispecies stands, particularly for the mountainous site.

The accuracy of our results was significantly improved when shifting from a general model to models that were based on the general stand type. When using AGB as an example, the general model, which pooled all stand types, generated a RSD of 22.3% on the mountainous site. A lower RSD was obtained for mixed stands (18.6%) when this was compared with RSD for the coniferous (21.8%) and deciduous stands (23.2%). The goodness-of-fit and estimation errors were clearly dependent upon stand type. The proportions of coniferous and deciduous trees were defined for all stands that were located on the mountainous site; such were available from the national forest inventory maps (IGN, 2012). It was therefore possible to apply one of the three predictive models (coniferous, deciduous, and mixed stand) to the whole study area. This approach reduced AGB estimation errors, especially for the deciduous stands, with RSDs that now could be estimated respectively at 16.2% (−2.4%), 16.5% (−5.3%), and 21.2% (−2.0%) for the mixed, coniferous and deciduous sites (Table 6). Even for the more complex mountainous site, our results compare very favorably with other studies (Zolkos et al., 2013). It is worth noticing the significant improvement in results for deciduous (9.8% decrease in RSD) and coniferous (12.2% decrease in RSD) stands that were obtained with our models when compared to mean results that were encountered in other studies, as reported by Zolkos et al. (2013). Discussion on model performance according to stand types was solely based on the results that were

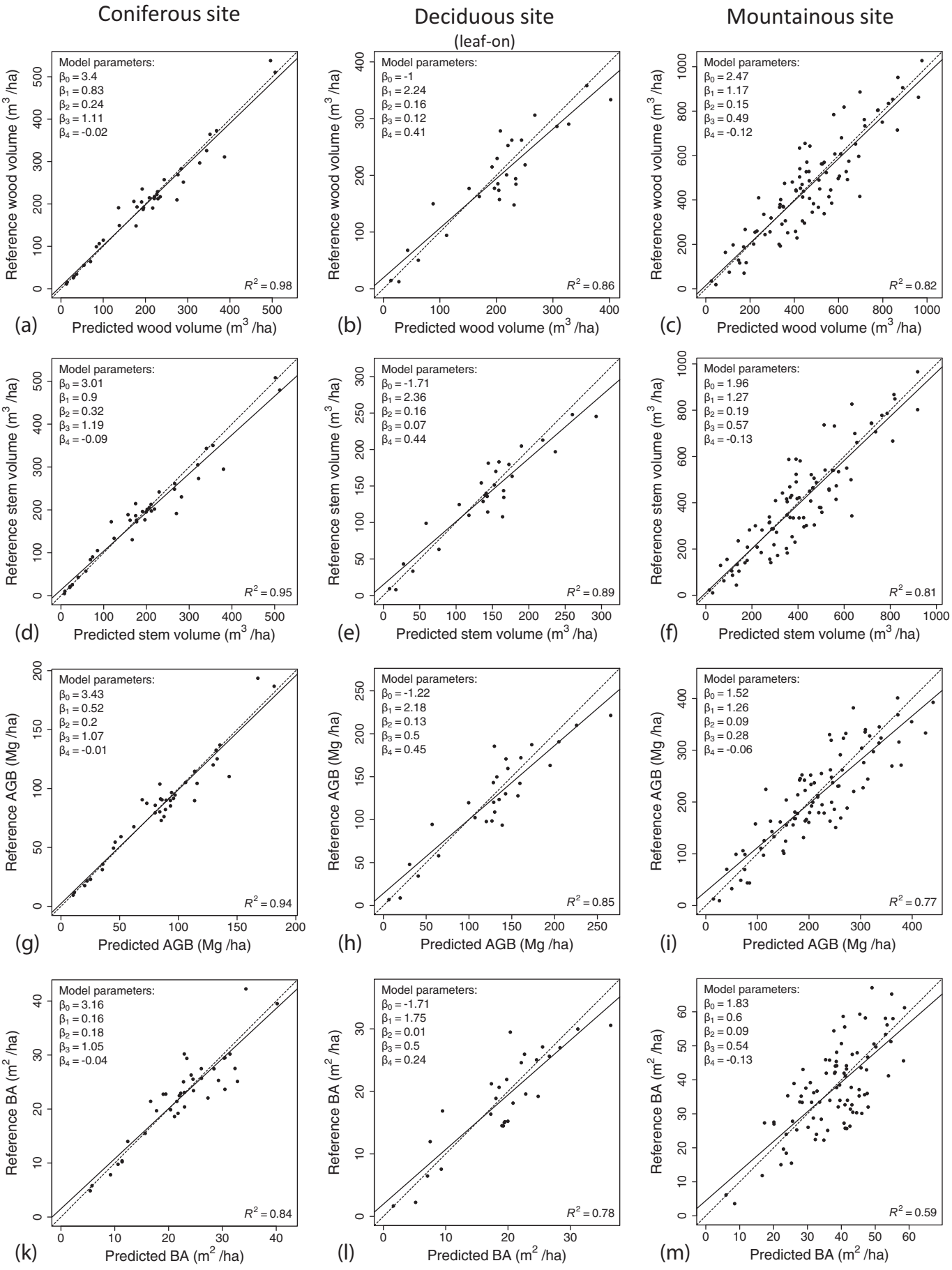

Fig. 2. Observed values of (a, b, c) wood volume, (d, e, f) stem volume, (g, h, i) AGB, and (j, k, l) BA versus their estimates for the three study sites (coniferous, deciduous and mountainous sites).

Table 6
Analysis of goodness-of-fit by stand type for the general model of the mountainous site and for separate models that were developed for the three forest types. Prediction models were calibrated using data from all plots, while goodness-of-fit was expressed in terms of the error estimators derived from the LOOCV (leave-one-out cross-validation) for three subsets of plots.

		Coniferous stands (33 field plots)			Mixed stands (23 field plots)			Deciduous stands (36 field plots)		
		R^2	RSD (%)	MPE (%)	R^2	RSD (%)	MPE (%)	R^2	RSD (%)	MPE (%)
General model	Wood volume	0.82	19.94	3.68	0.83	16.31	−5.83	0.63	28.23	−22.86
	Stem volume	0.82	20.60	8.79	0.85	16.50	−7.04	0.65	31.41	−34.38
	AGB	0.83	21.78	−13.17	0.74	18.61	−7.41	0.71	23.17	−3.71
	BA	0.60	20.91	2.99	0.52	20.54	−8.45	0.29	27.75	−16.77
Separate models	Wood volume	0.85	17.97	−4.31	0.85	15.64	−3.64	0.82	19.47	−5.68
	Stem volume	0.87	18.01	−4.07	0.87	15.94	−3.90	0.85	19.66	−5.85
	AGB	0.87	16.50	−4.08	0.80	16.21	−3.41	0.80	21.23	−6.19
	BA	0.67	18.91	−4.77	0.60	18.89	−4.42	0.51	22.75	−5.75

obtained for AGB, but it could easily be expanded to include the other stand attributes. This decision was made for the sake of clarity, and also because it enabled us to compare our results with those reported in the study by Zolkos et al. (2013).

Prediction models of wood volume, stem volume and AGB were closely correlated with one another and they were all more accurate than BA estimates. This result was not surprising because LiDAR cannot directly assess DBH; rather, it but primarily provides metrics that are related to vegetation height. Consequently, tree DBH, which is one important explanatory variable contributing to BA, can only be assessed by inference from LiDAR. Thus, BA assessment that was based upon height measurements is marred by the uncertainty that is inherent to the relationship between height and DBH. Conversely, height is a parameter that is directly involved in determining volumes or AGB, which explains why heights that were measured using ALS data have been shown to be highly correlated with volumes and AGB (Dubayah et al., 2000; Means et al., 2000).

Our results on model performances from LiDAR acquisition under leaf-off and leaf-on conditions for the deciduous site did not allow us to draw any definitive conclusions. Canopy transmittance of deciduous stands evolves, depending upon the season, and this is an important feature that significantly affects LiDAR point cloud characteristics. Leaf-off conditions result in an improvement in the description of the ground, whereas leaf-on conditions improve capabilities for describing the top of the canopy and, therefore, to detect tree tops (Dupuy, Lainé, Tassin, & Sarrailh, 2013). Regardless of the differences between the two LiDAR datasets, the predictive ability of leaf-off models was slightly superior to those of leaf-on conditions, with an improvement in RSD that ranged from 0.2% to 2.0%, depending upon stand attributes. Further studies are needed in diverse deciduous forests to verify if leaf-off conditions can systematically improve predictions of stand attributes.

Stand attributes were predicted in diverse forest types using the same four metrics. Our study sites were characterized by a wide range of structures, with mean Gini coefficients ranging from 0.19 to 0.41, and mean Shannon indices ranging from 0.07 to 1.10. The results showed that the importance of each metric varied from one site to the next and according to the predicted stand attribute. The metric μ_{CH} had the greatest explanatory power in each study site, but σ^2_{CH} was also a significant contributor in each site. P_f was highly significant for the coniferous site and, conversely, barely significant for the deciduous site, given that canopy transmittance was quite low in the latter. Even if Cv_{LAD} is usually a minor metric it can be significant in specific cases. Cv_{LAD} was significant at the deciduous site, but not significant in the mono-layered coniferous site. Cv_{LAD} was obtained from vertical LAD profiles and it is assumed to be invariant to occlusion that affects most LiDAR data, which normally results in realistic normalized profiles. To our knowledge, this metric has not been used until now to develop prediction models. Finally, the use of the four metrics improves the stand attributes estimation models.

It is important to be aware of sources of error that affect the accuracy of the predictive models. As a first source of error, field measurement errors can contribute to uncertainties in the predictive models of forest stand attributes that are derived from LiDAR data (Wulder et al., 2012). In addition, for stand attributes that require allometric relationships (like wood volume, stem volume and AGB), field measurement errors compound the errors that originate from the allometric equations (a second source of error). Substantial bias can be introduced if inappropriate allometric equations are used for volumes and AGB estimates (Chave et al., 2004). In this study, species- and site-specific equations have been favored when available and appropriate. As a third source of error, edge effects may affect metric robustness (Frazer, Magnussen, Wulder, & Niemann, 2011). Edge

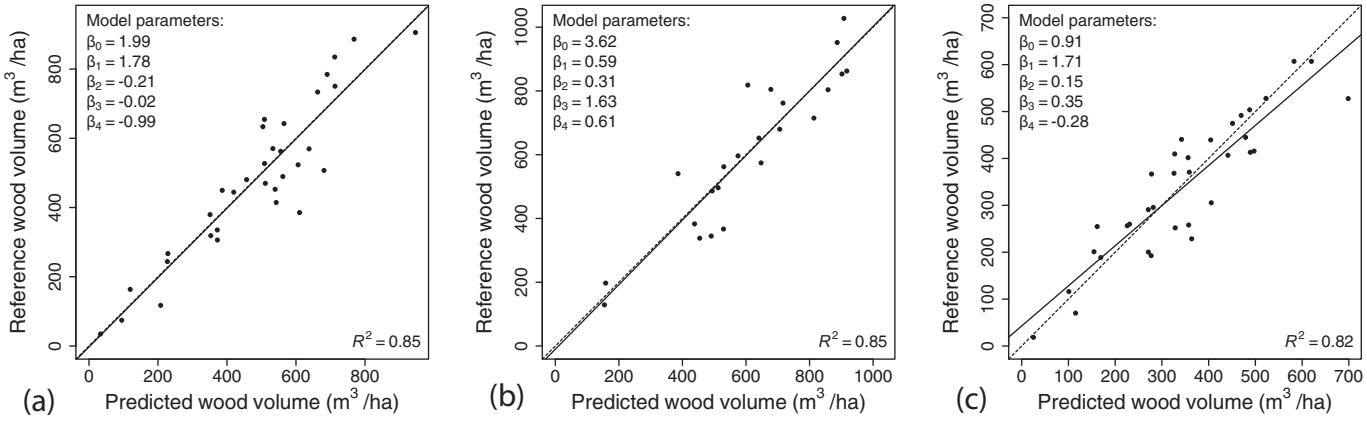

Fig. 3. Observed values of wood volume versus their estimates for the mixed mountainous forest, which was stratified into three forest types: (a) coniferous, (b) mixed and (c) deciduous stands.

effects were considered to be less problematic in coniferous sites that are composed of homogeneous stands. Edge effects were also likely to be less significant in dense stands and in large-sized plots (i.e., 15 m radius plots versus 6 m radius). As a fourth source of error that potentially can contribute to prediction errors in stand attributes, the influence of field plot positioning quality was also a potential source of discrepancies between the attribute estimates that were obtained from field measurements and predictions (Patterson & Williams, 2003). We expect a smaller effect of site position quality for the coniferous site for which positioning accuracy is likely to be at sub-meter scales. As a fifth source of error, it is important to decide if ALS metrics should be calculated with or without non-ground returns at heights under 2 m above the ground surface. The 2 m threshold has been suggested by several studies to reduce effects of understory vegetation and has resulted in more robust metrics (Frazer et al., 2011; Næsset, 2002; Nilsson, 1996). Studies have shown the relevance of applying a 2 m threshold in diverse forest types. However, other threshold values may be more suitable, depending upon forest type (Næsset, 2011; Nyström, Holmgren, & Olsson, 2012). A sixth source of error is related to the scan angle. Korhonen, Korpela, Heiskanen, and Maltamo (2011) found that oblique viewing angles can result in overestimation of the gap fraction (P) that was used in this model as a predictive metric. P has been estimated from data with two maximum scan angles: 16.5° in the mountainous site, and 29° in the other study sites. Therefore, P should be more strongly overestimated in the coniferous and deciduous sites, on average. Methods for correcting angle effects have been developed in previous studies (e.g., Korhonen et al., 2011), and might be slightly more efficient for assessing P. However, the comparison between results that were obtained at the coniferous and deciduous sites with those that were estimated for the coniferous and deciduous stands at the mountainous site suggests that factors other than that of the scan angle have a more significant effect on result accuracy. Indeed, results were poorer on the mountainous site, despite a lower scanning angle (with RSD ranging from 21.9% to 24.2%). Furthermore, the presence of substantial slopes in the mountainous site might have affected stand attribute predictions, as their presence leads to modifications in point cloud geometry (Véga et al., 2014). Sources of error were generally reduced in the coniferous site, thereby contributing to higher prediction accuracies when compared with results from other sites (deciduous or mountainous sites). However, these limiting factors are intrinsic to any model development project. One convenient way to evaluate how each of these sources of error affects the results is to perform a sensitivity analysis, such as those that were performed by Mutlu, Popescu, and Zhao (2008) and by Lu et al. (2012).

We recommend that further research concentrates on the four metrics μ_{CH}, σ^2_{CH}, P and Cv_{LAD}, as they are useful predictors of stand attributes. In particular, Cv_{LAD} metrics provide information on the vertical structure of a stand and, therefore, understory vegetation. The potential for further model generalization will have to be assessed by evaluating these metrics in diverse stands with a wide range of LiDAR characteristics. Our work has shown that the development of models related to stand type reduced estimation errors of stand attribute predictions. It is important, however, to assess how far one can specialize models according to stand type before they become less efficient than a more general model, which encompasses several stand types. Although we had four sites where data were collected at different pulse densities, our experimental setting was not appropriate for assessing the effects of pulse density on model quality. This is an important question that needs to be addressed, given that pulse density has a direct effect on LiDAR data acquisition costs.

6. Conclusion

We developed a model presenting a high level of generalization from ALS data, following an ABA. Our approach aimed at providing a procedure that facilitates the development of estimation models of stand attributes using a specific group of four metrics. This new procedure improved the results relative to the classical approach, which consists of entering an large number of height-related metrics into a stepwise regression procedure to eliminate the non-significant metrics; this approach is known to result in over-fitting of the data. Four metrics were selected to describe 3D canopy structures: mean canopy height (μ_{CH}); height heterogeneity (σ^2_{CH}); horizontal canopy distribution (P); and a metric that was estimated from leaf area density profiles to provide information on total stand vertical structure (Cv_{LAD}). We formulated a conceptual model from these four LiDAR metrics that predicted four stand attributes (wood volume, stem volume, AGB, and BA) in diverse forest types using a single model shape.

Based on our results, this model development approach demonstrated its ability to predict diverse forest attributes under a wide range of canopy structures. The models provided prediction error levels ranging from 12.4% to 24.2%. By developing a separate model for coniferous, deciduous and mixed stands, the results were improved by 2.0% to 5.3%. The ability of the model to predict stand attributes in leaf-off and leaf-on conditions was also verified, as results were equally good under both conditions with an RSD ranging from 16.9% to 20.7%. We were thus unable to draw definitive conclusions regarding the best conditions and technical specifications that would be required. The height metric (μ_{CH}) remains undoubtedly the dominant metric in the predictive models. However, the use of the three other structural metrics significantly improved prediction accuracy. The metric σ^2_{CH} proved to be relevant in each study site. The P metric also had high explanatory power, except for dense canopies that were specific to deciduous forest. The capacity of Cv_{LAD} to inform on vertical heterogeneity and understory vegetation clearly demonstrates its potential for attribute prediction in complex stands. This study also highlights the positive contribution of new metrics to the development of generalized prediction models of stand attributes with LiDAR.

Acknowledgements

This research was supported by the French National Research Agency (ANR) Grant within the Framework of the FORESEE project (ANR-2010-BIOE-008). The authors would like to thank the Agence nationale pour la gestion des déchets radioactifs (Andra) for providing the LiDAR dataset for the deciduous site. We are thankful for field plot measurements that were provided by the Office national des forêts (ONF) of France. We thank W.F.J. Parsons (CEF) and Jarlath Slevin (AWAN) for improving the use of English in the manuscript. We are also indebted to the Ministère des Relations internationales, de la Francophonie et du Commerce extérieur in Quebec, and the Ministère des Affaires Etrangères et Européennes in France, for funding travel between Quebec and France, through the Samuel de Champlain program in 2012–2014.

References

Allouis, T., Durrieu, S., Véga, C., & Couteron, P. (2013). Stem volume and above-ground biomass estimation of individual pine trees from LiDAR data: Contribution of full-waveform signals. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6*(2), 924–934, http://dx.doi.org/10.1109/JSTARS.2012.2211863.

Asner, G.P. (2009). Tropical forest carbon assessment: Integrating satellite and airborne mapping approaches. *Environmental Research Letters, 4*(3), 034009, http://dx.doi.org/10.1088/1748-9326/4/3/034009.

Asner, G.P., Mascaro, J., Muller-Landau, H.C., Vieilledent, G., Vaudry, R., Rasamoelina, M., et al. (2012). A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia, 168*(4), 1147–1160.

Avery, T.E., & Burkhart, H.E. (2001). *Forest measurements* (5th ed.). Boston, MA: McGraw-Hill.

Axelsson, P. (2000). DEM generation from laser scanner data using adaptive TIN models. *International Archives of Photogrammetry and Remote Sensing, 33*(B4/1; PART 4), 111–118.

Baskerville, G.L. (1972). Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research, 2*(1), 49–53.

Bréda, N.J. (2003). Ground-based measurements of leaf area index: A review of methods, instruments and current controversies. *Journal of Experimental Botany, 54*(392), 2403–2417.

Brown, S., & Lugo, A.E. (1984). Biomass of tropical forests: A new estimate based on forest volumes. *Science*, *223*(4642), 1290–1293, http://dx.doi.org/10.2307/1692790.

Campbell, G.S., & Norman, J.M. (1989). The description and measurement of plant canopy structure. *Plant canopies: Their growth, form and function* (pp. 1–19). Cambridge: Cambridge University Press.

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., et al. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, *145*(1), 87–99, http://dx.doi.org/10.1007/s00442-005-0100-x.

Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., & Perez, R. (2004). Error propagation and scaling for tropical forest biomass estimates. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*(1443), 409–420.

Chen, Q. (2010). Retrieving vegetation height of forests and woodlands over mountainous areas in the Pacific Coast region using satellite laser altimetry. *Remote Sensing of Environment*, *114*(7), 1610–1627, http://dx.doi.org/10.1016/j.rse.2010.02.016.

Chen, Q. (2013). *Lidar remote sensing of vegetation biomass*. Taylor & Francis: Remote Sensing of Natural Resources. CRC Press, 399–420.

Chen, Q., Gong, P., Baldocchi, D., & Tian, Y.Q. (2007). Estimating basal area and stem volume for individual trees from lidar data. *Photogrammetric Engineering and Remote Sensing*, *73*(12), 1355.

Cohen, W.B., & Spies, T.A. (1992). Estimating structural attributes of Douglas-fir/western hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment*, *41*(1), 1–17, http://dx.doi.org/10.1016/0034-4257(92)90056-P.

Deleuze, C., Senga Kiessé, T., Renaud, J. -P., Morneau, F., Rivoire, M., Santenoise, P., et al. (2013). *Rapport final EMERGE sur les modèles de volumes. Projet ANR-08-BIOE-003, Programme BIOE 2008*.

Dubayah, R., Knox, R., Hofton, M., Blair, J.B., & Drake, J. (2000). Land surface characterization using lidar remote sensing. *Spatial information for land use management* (pp. 25–38). Singapore: International Publishers Direct.

Dupuy, S., Lainé, G., Tassin, J., & Sarrailh, J. -M. (2013). Characterization of the horizontal structure of the tropical forest canopy using object-based LiDAR and multispectral image analysis. *International Journal of Applied Earth Observation and Geoinformation*, *25*(0), 76–86, http://dx.doi.org/10.1016/j.jag.2013.04.001.

Falkowski, M.J., Smith, A.M.S., Hudak, A.T., Gessler, P.E., Vierling, L.A., & Crookston, N.L. (2006). Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of lidar data. *Canadian Journal of Remote Sensing*, *32*(2), 153–161.

Fang, J.Y., Liu, G.H., & Xu, S.L. (1996). Biomass and net production of forest vegetation in China. *Acta Ecologica Sinica*, *16*(5), 497–508.

Ferster, C.J., Coops, N.C., & Trofymow, J.A. (2009). Aboveground large tree mass estimation in a coastal forest in British Columbia using plot-level metrics and individual tree detection from lidar. *Canadian Journal of Remote Sensing*, *35*(3), 270–275.

Finney, D.J. (1941). On the distribution of a variate whose logarithm is normally distributed. *Supplement to the Journal of the Royal Statistical Society*, 155–161.

Fisher, R.A. (1935). *The design of experiments*. Edinburgh: Oliver & Boyd.

Franklin, S.E. (2001). *Remote sensing for sustainable forest management*. Boca Raton: CRC Press.

Frazer, G.W., Magnussen, S., Wulder, M.A., & Niemann, K.O. (2011). 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.

Genet, A., Wernsdörfer, H., Jonard, M., Pretzsch, H., Rauch, M., Ponette, Q., et al. (2011). Ontogeny partly explains the apparent heterogeneity of published biomass equations for Fagus sylvatica in central Europe. *Forest Ecology and Management*, *261*(7), 1188–1202.

Gleason, C.J., & Im, J. (2011). A review of remote sensing of forest biomass and biofuel: options for small-area applications. *GIScience & Remote Sensing*, *48*(2), 141–170.

Goerndt, M.E., Monleon, V.J., & Temesgen, H. (2010). Relating forest attributes with area- and tree-based light detection and ranging metrics for western Oregon. *Western Journal of Applied Forestry*, *25*(3), 105–111.

Hall, F.G., Bergen, K., Blair, J.B., Dubayah, R., Houghton, R., Hurtt, G., et al. (2011). Characterizing 3D vegetation structure from space: Mission requirements. *Remote Sensing of Environment*, *115*(11), 2753–2775.

Hall, S.A., Burke, I.C., Box, D.O., Kaufmann, M.R., & Stoker, J.M. (2005). Estimating stand structure using discrete-return lidar: An example from low density, fire prone ponderosa pine forests. *Forest Ecology and Management*, *208*(1), 189–209.

Hopkinson, C., & Chasmer, L. (2009). Testing LiDAR models of fractional cover across multiple forest ecozones. *Remote Sensing of Environment*, *113*(1), 275–288.

Hopkinson, C., Lovell, J., Chasmer, L., Jupp, D., Kljun, N., & van Gorsel, E. (2013). Integrating terrestrial and airborne lidar to calibrate a 3D canopy model of effective leaf area index. *Remote Sensing of Environment*, *136*, 301–314.

Hounzandji, A., Jonard, M., Nys, C., Saint-André, L., & Ponette, Q. (2014). Improving the robustness of biomass function: From empirical to functional approaches. *Annals of Forest Science*.

IGN (2012). La cartographie forestière. http://inventaire-forestier.ign.fr/spip/IMG/pdf/Guide_technique_cartographie.pdf

Jensen, J.L.W.V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, *30*(1), 175–193.

Joly, D., Brossard, T., Cardot, H., Cavailhes, J., Hilal, M., & Wavresky, P. (2010). Les types de climats en France, une construction spatiale. *Cybergeo: European Journal of Geography*, http://dx.doi.org/10.4000/cybergeo.23155.

Kangas, A., & Maltamo, M. (2006). *Forest inventory: Methodology and applications*, *Vol. 10*, Dordrecht: Springer.

Khan, J.A., Van Aelst, S., & Zamar, R.H. (2007). Robust linear model selection based on least angle regression. *Journal of the American Statistical Association*, *102*(480), 1289–1299.

Kim, Y., Yang, Z., Cohen, W.B., Pflugmacher, D., Lauver, C.L., & Vankat, J.L. (2009). Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park, USA using small-footprint lidar data. *Remote Sensing of Environment*, *113*(11), 2499–2510.

Korhonen, L., Korpela, I., Heiskanen, J., & Maltamo, M. (2011). Airborne discrete-return LIDAR data in the estimation of vertical canopy cover, angular canopy closure and leaf area index. *Remote Sensing of Environment*, *115*(4), 1065–1080.

Kronseder, K., Ballhorn, U., Böhm, V., & Siegert, F. (2012). Above ground biomass estimation across forest types at different degradation levels in Central Kalimantan using LiDAR data. *International Journal of Applied Earth Observation and Geoinformation*, *18*, 37–48.

Le Toan, T., Beaudoin, A., Riom, J., & Guyon, D. (1992). Relating forest biomass to SAR data. *Geoscience and Remote Sensing, IEEE Transactions on*, *30*(2), 403–411.

Leboeuf, A., Fournier, R.A., Luther, J.E., Beaudoin, A., & Guindon, L. (2012). Forest attribute estimation of northeastern Canadian forests using QuickBird imagery and a shadow fraction method. *Forest Ecology and Management*, *266*, 66–74.

Leeuwen, M., & Nieuwenhuis, M. (2010). Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research*, *129*(4), 749–770, http://dx.doi.org/10.1007/s10342-010-0381-4.

Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., & Harding, D. (1999a). Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment*, *70*(3), 339–361.

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 and Biogeography*, *11*(5), 393–399.

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

Lexerød, N.L., & Eid, T. (2006). An evaluation of different diameter diversity indices based on criteria related to forest management planning. *Forest Ecology and Management*, *222*(1), 17–28.

Li, W., Guo, Q., Jakubowski, M.K., & Kelly, M. (2012). A new method for segmenting individual trees from the lidar point cloud. *Photogrammetric Engineering and Remote Sensing*, *78*(1), 75–84.

Lim, K.S., & Treitz, P.M. (2004). Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest Research*, *19*(6), 558–570.

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, http://dx.doi.org/10.1191/0309133303pp360ra.

Longuetaud, F., Santenoise, P., Mothe, F., Senga Kiessé, T., Rivoire, M., Saint-André, L., et al. (2013). Modeling volume expansion factors for temperate tree species in France. *Forest Ecology and Management*, *292*, 111–121.

Lovell, J.L., Jupp, D.L., Culvenor, D.S., & Coops, N.C. (2003). Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests. *Canadian Journal of Remote Sensing*, *29*(5), 607–622.

Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang, M., et al. (2012). Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates. *International Journal of Forestry Research*, http://dx.doi.org/10.1155/2012/436537 (2012).

Lucas, R.M., Cronin, N., Lee, A., Moghaddam, M., Witte, C., & Tickle, P. (2006). Empirical relationships between AIRSAR backscatter and LiDAR-derived forest biomass, Queensland, Australia. *Remote Sensing of Environment*, *100*(3), 407–425.

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*(7), 1016–1031.

Magnussen, S., Næsset, E., Gobakken, T., & Frazer, G. (2012). A fine-scale model for area-based predictions of tree-size-related attributes derived from LiDAR canopy heights. *Scandinavian Journal of Forest Research*, *27*(3), 312–322.

Magurran, A.E. (2004). *Measuring biological diversity*. Oxford: Blackwell Publishing, 285–286.

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.

Martens, S.N., Ustin, S.L., & Rousseau, R.A. (1993). Estimation of tree canopy leaf area index by gap fraction analysis. *Forest Ecology and Management*, *61*(1), 91–108.

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

Means, J.E., Acker, S.A., Harding, D.J., Blair, J.B., Lefsky, M.A., Cohen, W.B., et al. (1999). Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the Western Cascades of Oregon. *Remote Sensing of Environment*, *67*(3), 298–308.

Morsdorf, F., Kötz, B., Meier, E., Itten, K.I., & Allgöwer, B. (2006). Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. *Remote Sensing of Environment*, *104*(1), 50–61.

Mutlu, M., Popescu, S.C., & Zhao, K. (2008). Sensitivity analysis of fire behavior modeling with LIDAR-derived surface fuel maps. *Forest Ecology and Management*, *256*(3), 289–294.

Næsset, E. (1997). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *52*(2), 49–56, http://dx.doi.org/10.1016/S0924-2716(97)83000-6.

Næsset, E. (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, http://dx.doi.org/10.1016/S0034-4257(01)00290-5.

Næsset, E. (2004). Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, *19*(2), 164–179.
Næsset, E. (2011). Estimating above-ground biomass in young forests with airborne laser scanning. *International Journal of Remote Sensing*, *32*(2), 473–501.
Næsset, E., & Gobakken, T. (2008). Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment*, *112*(6), 3079–3090, http://dx.doi.org/10.1016/j.rse.2008.03.004.
Nelson, R., Krabill, W., & Tonelli, J. (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, *24*(2), 247–267.
Nilson, T. (1971). A theoretical analysis of the frequency of gaps in plant stands. *Agricultural Meteorology*, *8*, 25–38.
Nilson, T. (1999). Inversion of gap frequency data in forest stands. *Agricultural and Forest Meteorology*, *98*, 437–448.
Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment*, *56*(1), 1–7, http://dx.doi.org/10.1016/0034-4257(95)00224-3.
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*(0), 271–279, http://dx.doi.org/10.1016/j.rse.2012.03.008.
Patenaude, G., Hill, R., Milne, R., Gaveau, D.L.A., Briggs, B.B.J., & Dawson, T.P. (2004). Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sensing of Environment*, *93*(3), 368–380, http://dx.doi.org/10.1016/j.rse.2004.07.016.
Patterson, P.L., & Williams, M.S. (2003). Effects of registration errors between remotely sensed and ground data on estimators of forest area. *Forest Science*, *49*(1), 110–118.
Picard, R.R., & Cook, R.D. (1984). Cross-validation of regression models. *Journal of the American Statistical Association*, *79*(387), 575–583.
Picard, N., Saint-André, L., & Henry, M. (2012). *Manual for building tree volume and biomass allometric equations: from field measurement to prediction.* Manual for building tree volume and biomass allometric equations: from field measurement to prediction and Centre de Coopération Internationale en Recherche Agronomique pour le Développement (http://www.fao.org/docrep/018/i3058e/i3058e.pdf).
Popescu, S.C., Wynne, R.H., & Nelson, R.F. (2003). Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing*, *29*(5), 564–577.
Reich, P.B. (2012). Key canopy traits drive forest productivity. *Proceedings of the Royal Society Series B: Biological Sciences*, *279*(1736), 2128–2134.
Ross, I. (1981). *The radiation regime and architecture of plant stands*, *Vol. 3*, The Hague: Dr. W. Junk Publishers.
Ruel, J.J., & Ayres, M.P. (1999). Jensen's inequality predicts effects of environmental variation. *Trends in Ecology & Evolution*, *14*(9), 361–366, http://dx.doi.org/10.1016/S0169-5347(99)01664-X.
Schumacher, F.X., & Hall., D.S. (1933). Logarithmic expression of timber-tree volume. *Journal of Agricultural Research*, *47*, 719–734.
Scott, C.T., & Gove, J.H. (2002). Forest inventory. *Encyclopedia of Environmetrics* (pp. 814–820). Chichester: Wiley.
Shaiek, O., Loustau, D., Trichet, P., Meredieu, C., Bachtobji, B., Garchi, S., et al. (2011). Generalized biomass equations for the main aboveground biomass components of maritime pine across contrasting environments. *Annals of Forest Science*, *68*(3), 443–452, http://dx.doi.org/10.1007/s13595-011-0044-8.
Snee, R.D. (1977). Validation of regression models: Methods and examples. *Technometrics*, *19*(4), 415–428.
Solberg, S., Brunner, A., Hanssen, K. H., Lange, H., Næsset, E., Rautiainen, M., et al. (2009). Mapping LAI in a Norway spruce forest using airborne laser scanning. *Remote Sensing of Environment*, *113*(11), 2317–2327.
Solberg, S., Næsset, E., Hanssen, K. H., & Christiansen, E. (2006). Mapping defoliation during a severe insect attack on Scots pine using airborne laser scanning. *Remote Sensing of Environment*, *102*(3), 364–376.
Sprugel, D.G. (1983). Correcting for bias in log-transformed allometric equations. *Ecology*, *64*(1), 209–210.
Tang, H., Brolly, M., Zhao, F., Strahler, A.H., Schaaf, C.L., Ganguly, S., et al. (2014). Deriving and validating Leaf Area Index (LAI) at multiple spatial scales through lidar remote sensing: A case study in Sierra National Forest, CA. *Remote Sensing of Environment*, *143*, 131–141.
Tang, H., Dubayah, R., Swatantran, A., Hofton, M., Sheldon, S., Clark, D.B., et al. (2012). Retrieval of vertical LAI profiles over tropical rain forests using waveform lidar at La Selva, Costa Rica. *Remote Sensing of Environment*, *124*, 242–250.
Ung, C. -H., Bernier, P., & Guo, X. -J. (2008). Canadian national biomass equations: New parameter estimates that include British Columbia data. *Canadian Journal of Forest Research*, *38*(5), 1123–1132.
Véga, C., & Durrieu, S. (2011). Multi-level filtering segmentation to measure individual tree parameters based on Lidar data: Application to a mountainous forest with heterogeneous stands. *International Journal of Applied Earth Observation and Geoinformation*, *13*(4), 646–656, http://dx.doi.org/10.1016/j.jag.2011.04.002.
Véga, C., Hamrouni, A., El Mokhtari, S., Morel, J., Bock, J., Renaud, J. -P., et al. (2014). PTrees: A point-based approach to forest tree extraction from lidar data. *International Journal of Applied Earth Observation and Geoinformation*, *33*, 98–108, http://dx.doi.org/10.1016/j.jag.2014.05.001.
Vose, J.M., Clinton, B.D., Sullivan, N.H., & Bolstad, P.V. (1995). Vertical leaf area distribution, light transmittance, and application of the Beer–Lambert Law in four mature hardwood stands in the southern Appalachians. *Canadian Journal of Forest Research*, *25*(6), 1036–1043.
Waring, R.H. (1983). Estimating forest growth and efficiency in relation to canopy leaf area. *Advances in Ecological Research*, *13*, 327–354.
Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., et al. (2012). Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, *121*, 196–209.
Zanne, A.E., Lopez-Gonzalez, G., Coomes, D.A., Ilic, J., Jansen, S., Lewis, S.L., et al. (2009). *Global wood density database. Dryad. Identifier*, *235*, (http://hdl.handle.net/10255/dryad).
Zhao, K., & Popescu, S. (2009). Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA. *Remote Sensing of Environment*, *113*(8), 1628–1645.
Zianis, D., & Mencuccini, M. (2004). On simplifying allometric analyses of forest biomass. *Forest Ecology and Management*, *187*(2), 311–332.
Zolkos, S.G., Goetz, S.J., & Dubayah, R. (2013). A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sensing of Environment*, *128*(0), 289–298, http://dx.doi.org/10.1016/j.rse.2012.10.017.