Remote Sensing of Environment 234 (2019) 111434

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Quantifying the contribution of spectral metrics derived from digital aerial photogrammetry to area-based models of forest inventory attributes

Piotr Tompalski[a,*], Joanne C. White[b], Nicholas C. Coops[a], Michael A. Wulder[b]

[a] Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC, V6T 1Z4, Canada
[b] Canadian Forest Service, (Pacific Forestry Center), Natural Resources Canada, 506 West Burnside Road, Victoria, BC, V8Z 1M5, Canada

ARTICLE INFO

Keywords:
Digital aerial photogrammetry
Forestry
Inventory
Remote sensing
Point cloud
Lidar
Airborne laser scanning

ABSTRACT

Digital aerial photogrammetry (DAP) has demonstrated utility across a range of forest environments as an alternative data source to airborne laser scanning (ALS) for estimating forest inventory attributes in an area-based approach. In this context, metrics are typically derived from the DAP point cloud in a manner analogous to that of ALS data. However, image matching algorithms also allow for spectral information from the image data to be transferred to the point cloud. Herein, we quantify the contribution of this spectral information to the area-based prediction of five forest inventory attributes: Lorey's mean height, quadratic mean diameter, basal area, gross volume per ha, and stems per ha in a highly productive coastal temperate rainforest on Vancouver Island, British Columbia, Canada. Using ground plots and ALS-derived area-based estimates as reference, we compare plot-level predictions generated using (i) DAP point cloud metrics, (ii) DAP spectral metrics, and (iii) combinations of DAP point cloud and spectral metrics. In addition to prediction accuracy, we assessed variable importance to identify those metrics that were most informative for the developed models. Our results indicated that for models generated using DAP data, prediction accuracy was greatest when the point cloud-based metrics were incorporated. Models that incorporated both point cloud and spectral information were only slightly more accurate than models based on point cloud metrics only. We found that the improvement in accuracy was not observed for all stand attributes. The highest increase in accuracy for models combining point cloud and spectral metrics was observed for quadratic mean diameter, basal area per hectare, and stem volume per hectare, with change in relative root mean square error of −1.3%, −1.75%, and −0.23%, respectively. Models derived with spectral metrics only had the lowest accuracy with R^2 values never exceeding 0.25. Analysis of the variable importance indicated that point cloud metrics are markedly more important than spectral metrics. We conclude that the benefit of the additional spectral information in this forest environment is negligible, and the effort to derive the spectral information cannot be justified for operational applications. Our results confirm those of other studies in other environments that have likewise found minimal benefit to the incorporation of DAP spectral information in area-based estimation.

1. Introduction

Airborne laser scanning (ALS) has been established as an accurate source of three-dimensional structural information for forest stand attribute estimation (Magnussen et al., 2010; White et al., 2017). Using area-based approaches (ABA; Næsset, 2002) either parametric or non-parametric predictive models can be developed (Penner et al., 2013; White et al., 2013a, 2017), which relate forest variables of interest including Lorey's height (HL), quadratic mean diameter (QMD), basal area (BA), or total volume (V) with ALS-based metrics as independent variables. The resulting spatial predictions represent a crucial component of enhanced forest inventories (EFI) (White et al., 2017), which can inform managers on the state of the current resources, as well as be linked to growth models to estimate future growth and yield (Tompalski et al., 2018).

Recently, digital aerial photogrammetry (DAP) has emerged as an additional technology capable of providing three-dimensional point clouds suitable for estimating forest stand attributes in an ABA. The advancement of DAP as a technology capable of providing high quality three-dimensional structural information of forests comes from rapid technological advancements and cost reductions in computer hardware, that have increased the automation of photogrammetric workflows (Leberl et al., 2010; Nolan et al., 2015). In particular, the use of image matching algorithms either area- or feature-based (Kukkonen et al.,

* Corresponding author.
E-mail address: piotr.tompalski@ubc.ca (P. Tompalski).

https://doi.org/10.1016/j.rse.2019.111434
Received 14 May 2019; Received in revised form 10 September 2019; Accepted 19 September 2019
Available online 31 October 2019

2017; Zitová and Flusser, 2003), have been successfully applied to generate three-dimensional point clouds for estimating structural attributes of vegetation (White et al., 2013b). In the case of ALS data, these point clouds represent the x, y, and z locations in the vertical canopy profile where laser energy is intercepted by a branch, foliage, etcetera. However, in the case of DAP, these point clouds represent x, y, and z locations of the same pixel that is matched on two or more digital images. In the case of DAP data, the object (pixel) must be visible in both (or multiple) images in order for a match to occur. Thus, while a certain amount of laser energy is able to reach the forest floor and enable characterization of the terrain surface under canopy, DAP data are limited to the outer canopy envelope. Moreover, the utility of these data in an area-based approach has been demonstrated across a range of forest environments (Bohlin et al., 2012; Hawryło et al., 2017; Pitt et al., 2014; Puliti et al., 2017a; Straub et al., 2013; White et al., 2015). The effectiveness of these image-matching algorithms in generating high-density point clouds has led to renewed interest in digital aerial photogrammetry (Remondino et al., 2014) due to their provision of very-high-resolution imagery and structural information at a lower relative cost than ALS (Rahlf et al., 2017; Stone et al., 2016). Moreover, the capacity to generate 3D point clouds from historical air photos in order to inform characterizations of forest stand growth and age provides additional data for these attributes that are challenging to characterize based on field plot data alone (Vastaranta et al., 2014; Véga and St-Onge, 2008).

In contrast to ALS however, image-based point clouds are limited to the top canopy surface with ground points located only when the ground can be observed on the imagery, such as in canopy gaps or exposed ground. The lack of sufficient ground representation requires an alternative digital terrain model in order to calculate the relative heights of points above ground (White et al., 2013b). In addition, image matching algorithms are confounded by shadows, which are common in forest canopies, as well as factors related to solar illumination (Baltsavias, 1999; St-Onge et al., 2008). Despite these challenges, the success with which DAP data have been used in area-based estimates of forest inventory attributes have resulted in them being considered as an effective data source for updating EFIs and for providing incremental forest stand growth information (Tompalski et al., 2018).

When applying an ABA for estimating forest inventory attributes, DAP data are typically used in a manner analogous to ALS data (White et al., 2017). A set of descriptive point cloud metrics are generated from the DAP data that characterize the vertical distribution of vegetation through the canopy profile. These metrics are statistical summaries of the point heights or z-values from the point cloud. As noted earlier, these DAP data primarily characterize the outer canopy envelope, and as a result, DAP point clouds are concentrated in the upper canopy. Thus, it is expected that values for DAP and ALS point cloud metrics will differ. White et al. (2015) demonstrated that although many of the ALS- and DAP-based point cloud metrics are statistically different, the similarity between metrics increases with increasing canopy cover and is likewise for those metrics that characterize the upper canopy. Unique to DAP point clouds is the potential to link the associated spectral information to z-heights. DAP data are typically acquired in the visible and near-infrared wavelengths as full colour or colour-infrared imagery. It is anticipated that this complementary spectral information associated with the DAP-derived three-dimensional clouds can provide additional information on forest species, land cover type, insights into the physiological condition of the stand, as well as potential insights into forest types and species. The integration of ALS point cloud metrics with spectral information derived from imagery has been demonstrated in several studies (e.g. Leckie et al., 2003; Hudak et al., 2006). In most cases, ALS data is merged with aerial imagery acquired simultaneously or at another time (Packalén and Maltamo, 2007; Niska et al., 2010).

Packalén and Maltamo (2007) used digitized analog air photos to derive spectral and textural plot-level predictors, with the objective of deriving species-specific area-based estimates of basal area, stem number, volume, and diameter and height of the basal area median tree. Textural features were based on grey-level co-occurrence matrix (GLCM) approach of Haralick et al. (1973), but were calculated for the entire plot without the use of a moving window. Spectral metrics were calculated as the mean and median intensities from each of the green, red, and near-infrared bands. Models were generated using a multivariate non-parametric k-MSN approach with 7 image-based and 18 ALS-based metrics. Image-based metrics were intended to improve separability between species, specifically between spruce and pine. The authors concluded that the ALS-based metrics were more effective for predicting structural attributes than the image-based metrics were for discriminating species. Niska et al. (2010) also assessed ALS- and image-based metrics for assessing species-specific plot volumes. Similar to Packalén and Maltamo (2007), the authors developed spectral and textural metrics from digitized analog air photos, which were used in concert with ALS metrics in various machine learning approaches. In that study, the relative contributions of the different metrics were not reported in depth.

In the case of DAP data, spectral information can be derived directly from the generated true-orthophoto, which is spatially aligned with the generated point cloud, thereby reducing issues associated with a geometric mismatch between the spectral information and point cloud data, or a temporal mismatch between the acquisition of the point cloud and the spectral data. To date, these DAP spectral metrics have rarely been used in area-based model development. Melin et al. (2017) analyzed the capability of DAP data to estimate forest canopy cover utilizing both point cloud and spectral metrics, demonstrating that the inclusion of spectral information resulted in a reduction in the relative RMSE by 1% (from 10.3% when only point cloud metrics were used, to 9.3% when spectral metrics were included). Melin et al. (2017) also demonstrated that the although the models based only on spectral metrics had slightly lower accuracy (relative RMSE of 12.4%), they were better at characterizing the extreme values of the canopy cover distribution than models based on point cloud metrics alone. Bohlin et al. (2012) used height-based texture metrics in addition to standard point cloud metrics, to improve accuracy of height, basal area, and stem volume predictions, reporting that models which incorporated texture metrics had lower RMSE for basal area and stem volume (e.g. RMSE was reduced from 14.6 to 13.1% for volume).

As interest in DAP approaches grow due to the potential cost savings associated with the use of DAP versus ALS, and given the opportunity to derive retrospective three-dimensional point clouds from previously acquired airborne imagery, additional studies are required to investigate the predictive power of these models compared to ALS-derived models. In this analysis, our goal is to advance understanding of how, if, and in what form spectral measures from DAP can be used to augment area-based models of forest inventory attributes generated from DAP point clouds. To meet this goal, we ask the following specific questions to guide our analysis:

What is the difference in prediction accuracy for common forest inventory attributes using conventional point cloud metrics derived from DAP and ALS data?

To address this question we examine the predictive capacity of models for common forest attributes using both DAP and ALS datasets, acquired in the same growing season, in the highly productive, dense canopy cover forests of northern Vancouver Island, BC, Canada. Using a machine learning approach, we examine the impacts of different structural point cloud metrics derived from ALS and DAP data on the prediction accuracies of area-based models of Lorey's mean height (HL), quadratic mean diameter (QMD), basal area (BA), gross volume per hectare (V), and stems per hectare (N). The objective is to establish a reference baseline or expected level of DAP model performance, building on the work of White et al. (2015).

What is the additional predictive power provided by spectral information derived from the true orthophoto and are any particular spectral wavelengths more useful than others in increasing predictive accuracy?

Previous studies have demonstrated the predictive power of high spatial resolution spectral data in both the visible and near-infrared (NIR) wavelengths. Examples include conifer species composition and forest cover classification (Franklin et al., 2000), estimation of individual tree species and crown closure (Cohen et al., 2001), stem density and stand height (Franklin and McDermid, 1993), and volume (Wulder et al., 2012). In addition, information on the physiological capacity of vegetation such as leaf area index (LAI), biomass, and net and gross primary productivity (NPP, GPP) (Coops et al., 2010; Hilker et al., 2009; Turner et al., 2006) have also been successfully derived. Given these insights, do spectral metrics derived from DAP point clouds improve predictive capacity for the aforementioned suite of inventory attributes and if so, are there certain attributes for which the use of spectral metrics are particularly advantageous? In addition, are particular spectral bands or indices computed from these bands, more useful than others in improving the accuracy of the derived models? The objective of this component is to quantify any improvement to DAP area-based model performance through the use of additional spectral metrics.

2. Methods

2.1. Study area

Data used in our analyses were acquired over northern Vancouver Island, BC, Canada (NVI), and represent an area of approximately 52,000 ha, spanning three geographically separate areas (Fig. 1). The study area is a coastal temperate rainforest, with western hemlock (*Tsuga heterophylla* (Raf.) Sarg.), western red cedar (*Thuja plicata* Donn ex D. Don), and amabilis fir (*Abies amabilis* Douglas ex J. Forbes) as the most common tree species. Other tree species found in the study area include Douglas-fir (Pseudotsuga menziesii (Mirb.) Franco), red alder (*Alnus rubra* Bong.), yellow cedar (*Chamaecyparis nootkatensis* (D.Don) Spach), mountain hemlock (*Tsuga mertensiana* (Bong.) Carrière), and Sitka spruce (*Picea sitchensis* (Bong.) Carrière). Located primarily within the wetter Coastal Western Hemlock biogeoclimatic zone (CWH), the NVI study area is characterized by high annual precipitation (3000–5000 mm), mild winters (average temperature 0 °C to 2 °C), and cool summers (average temperature 18 °C to 20 °C). Elevation within the study area ranges from sea level to 1200 m, with an average slope of 23.6°. The average age of stands in the NVI was 144 years (standard deviation = 127 years).

2.2. Ground plot data

A total of 131 ground plots measured in summer of 2013 were used in this study, with plot centres measured with a Trimble GeoXH GPS receiver equipped with an external Tornado antenna. On average, more than 900 GPS measures were acquired per plot centre. Plot positions were differentially corrected to have sub-metre planimetric precision. Plots were circular in shape, with a radius of 14 m and an area of 615.75 m^2. Within each plot, all live standing trees with diameter at breast height (DBH) ≥12.0 cm were measured. Individual tree measures included DBH (cm), tree height (m), and species. Individual tree-based measurements of stem height and diameter were used to compute plot-level values of Lorey's mean height (HL), quadratic mean diameter (QMD), basal area (BA), gross volume per hectare (V), and stems per hectare (N). A summary of plot characteristics is presented in Table 1.

2.3. Point-cloud data

2.3.1. NVI DAP

Digital imagery was acquired for the study area using a Vexcel UltraCamX camera in 2012 (August 16, September 25, and October 4–6, between 10:00 and 15:00). The imagery was delivered as 4-band (RGB and NIR; red, green, blue, and near-infrared), with 8 bit radiometric resolution, a 0.30 m ground sampling distance (GSD) and was acquired along 6 flight lines, with a minimum 60% along-track and

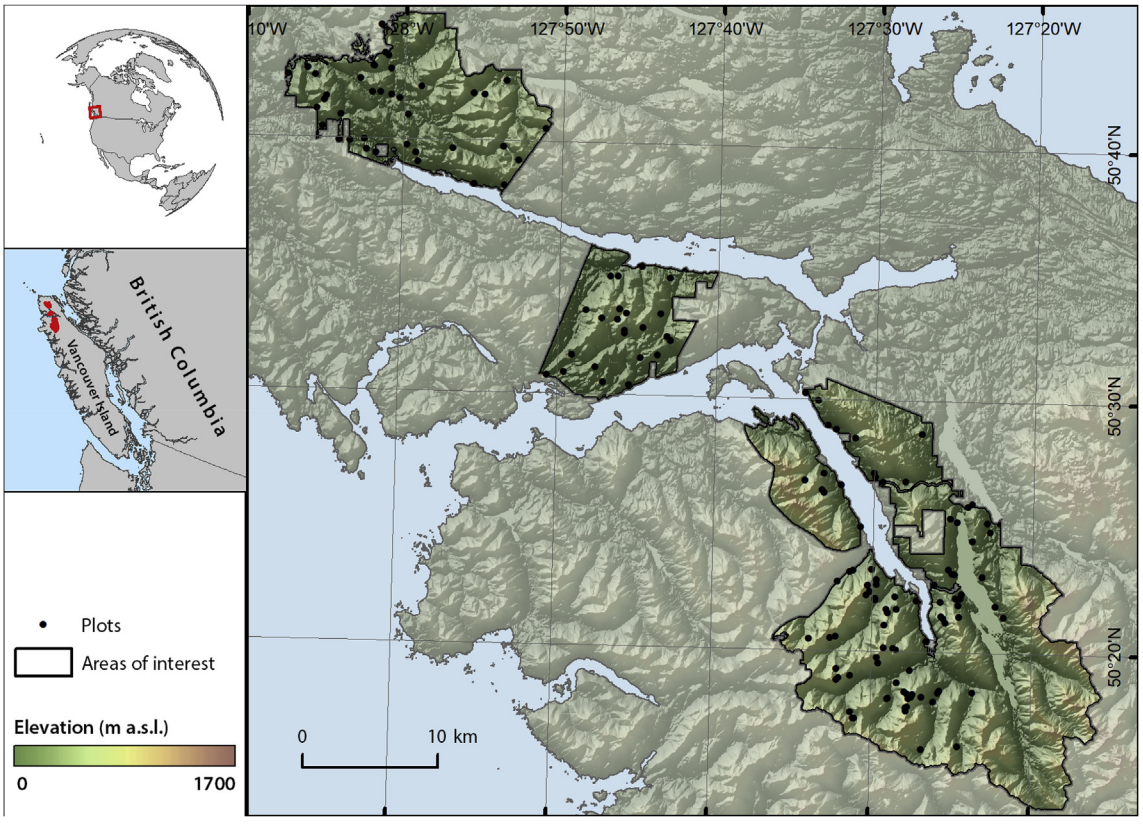

Fig. 1. Location and extent of the study area.

Table 1
Inventory field plot characteristics by species groups. Reported are mean values with standard deviation in parentheses. HW – Western hemlock, CW – Western red cedar, OC – other conifer, OD – other deciduous.

Species	count	HL [m]	QMD [cm]	BA [m^2/ha]	V [m^3/ha]	N
HW	74	32.7 (10.1)	38.3 (14.4)	64.0 (29.7)	913.4 (538.6)	690.4 (385.3)
CW	17	26.0 (8.3)	39.8 (15.8)	81.3 (38.9)	850.4 (495.4)	803.4 (475.6)
OC	32	33.0 (9.7)	45.1 (12.4)	88.4 (30.4)	1102.9 (488.7)	636.4 (261.7)
OD	8	23.2 (6.0)	28.0 (6.9)	34.00 (12.0)	373.7 (199.0)	676.0 (367.4)
Total	131	31.3 (10.0)	39.6 (14.3)	70.39 (33.3)	918.5 (528.7)	691.0 (370.4)

20% across-track overlap. We used Agisoft Metashape (Agisoft, 2019) to generate dense image point clouds and true-orthophotos. The initial image alignment as well as the dense point cloud generation settings were set to "highest" and "ultra high", respectively. The "depth filter" option during dense cloud generation was set to "mild". The resulting average point density of the image-based point cloud was 13.19 points/m^2. The R, G, B, and NIR values for each point were stored as additional attributes in the exported point clouds. A 4-band true-orthophoto was generated at 1 m resolution. The "mosaic (default)" option was used when generating the orthomosaic (Agisoft, 2019).

2.3.2. NVI ALS

ALS point clouds were acquired in August and September of 2012 using an Optech ALTM3100EA scanning system operated at an altitude of approximately 700 m above ground level (Table 2). The average return pulse density was 11.6 points/m^2 (Tompalski et al., 2015). A Digital Terrain Model (DTM) with a spatial resolution of 1 m was created using ground returns and standard pre-processing routines as per Axelsson (2000).

2.3.3. Data processing

The ALS and the DAP point clouds were processed following standard processing routines, which included filtering, tiling, and height normalization. Processing was performed using a combination of point cloud processing tools including LAStools (Isenburg, 2018), and the lidR package for R (R Core Team, 2019; Roussel and Auty, 2018). A tile size of 1000 × 1000 m and a 30 m buffer were used. Point cloud normalization of both ALS and DAP point clouds was performed using the ALS ground points and the lasnormalize function of the lidR package. The ALS points classified as ground were used to normalize the DAP point cloud elevations relative to the ground surface. The true-orthophotos and each generated point cloud dataset, were clipped using the outlines of the plots using lidR and raster packages for R (Hijmans, 2019).

2.3.4. Point cloud and CHM metrics

Point cloud metrics included measures of central tendency (mean, median), measures of dispersion (variance, standard deviation, interquartile range), percentiles, proportions, and densities of point heights above ground and were computed using lasmetrics function of the lidR package. A complete list of point cloud metrics can be found in Table 3.

Table 2
Airborne laser scanning (ALS) data acquisition parameters.

Acquisition parameter	Value
Year of acquisition	2012
Sensor	ALTM3100EA
Flying height	700 m AGL
Flight speed	130 kts
Pulse repetition rate	70 kHz
Scanning frequency	65 Hz
Scan angle	25 deg
Beam divergence	0.3 mrad
Average pulse density	11.6

Table 3
Point cloud metrics used as predictor variables.

Metric	Description
PTS_perc_all_above_threshold2	proportion of points above 2 m
PTS_perc_all_above_threshold5	proportion of points above 5 m
PTS_perc_all_above_mean	proportion of points above mean height
PTS_perc_all_above_median	proportion of points above median height
PTS_zmax	maximum height
PTS_zmean	mean height
PTS_zsd	standard deviation of point heights
PTS_zskew	skewness
PTS_zkurt	kurtosis
PTS_zq5, PTS_zq10, …, PTS_zq95	percentiles (5th - 95th)
PTS_zpcum1, PTS_zpcum2, …, PTS_zpcum9	deciles - cumulative percentage of return in xth layer (Woods et al., 2008). x = {1:9}
PTS_lad.cv	Coefficient of variation of the LAD profile (Bouvier et al., 2015)

Separate ALS and DAP canopy height models (CHMs) were generated for each plot at a 1 m resolution, and then were used to generate an additional set of metrics. The CHMs were generated using the point-to-raster method, i.e. for each pixel of the output raster, the height of the highest point is assigned (Roussel and Auty, 2018). The set of CHM-based metrics included measures of central tendency (mean, median), dispersion (sd, coefficient of variation), proportions (% of pixels above 2, 10, and 20 m), and a measure of CHM roughness (rumple index, Table 4) and were calculated using a combination of raster and lidR packages for R.

2.3.5. Spectral metrics

A number of spectral indices and band statistics (Table 5) were calculated based on the spectral information assigned to each point in the DAP point cloud as well as based on the true-orthophoto. First, the orthophoto was used to derive overall plot-level summary statistics that incorporated spectral values in each band, the Normalized Difference Vegetation Index (NDVI), as well as GLCM texture metrics derived for the NIR band. We used raster and glcm packages to calculate the plot-level summary statistics (Zvoleff, 2019; Hijmans, 2019). The pixel-level statistics were summarized at a plot level with mean and standard deviation. Second, using a custom function, the 3D spectral metrics

Table 4
Canopy height model (CHM) metrics used as predictor variables.

Metric	Description
CHM_rumple	rumple index (Kane et al., 2010)
CHM_chmmean	mean of CHM pixel values
CHM_chmsd	standard deviation of CHM pixel values
CHM_chmcv	coefficient of variation of CHM pixel values
CHM_CC2	proportion of CHM pixels above 2 m
CHM_CC10	proportion of CHM pixels above 10 m
CHM_CC20	proportion of CHM pixels above 20 m
CHM_CCmean	proportion of CHM pixels above chmmean
CHM_CCmeansd	proportion of CHM pixels above chmmean + chmsd

Table 5
Spectral metrics used as predictor variables in area-based models.

Metric	Description
SPC_x_Mean; SPC_x_Std.Dev	Average and standard deviation of x spectral band (orthophoto); x = {B, G, R, NIR}
SPC_NDVI_Mean; SPC_NDVI_Std.Dev	Mean and standard deviation of the point cloud-based NDVI
SPC_NDVI_P5, SPC_NDVI_P10, …, SPC_NDVI_P95	Percentiles of the point cloud-based NDVI
SPC_NDVI_L1, SPC_NDVI_L2, …, SPC_NDVI_L10	Average point cloud-based NDVI in L layer of the point cloud. Height of each layer is equal to 1/10 of the range of point heights. L = {1:10}.
SPC_NDVI_above_Z_P5, SPC_NDVI_above_Z_P10, …, SPC_NDVI_above_Z_P95,	Average point cloud-based NDVI calculated on a subset of points above each of the height percentiles.
TX_glcm_x_Mean; TX_glcm_x_Std.Dev …	Grey level co-occurrence matrix (GLCM)-based texture metrics derived using the NIR band. x = {Mean, Variance, homogeneity, Contrast, Dissimilarity, Entropy, Second_moment, Correlation (Haralick et al., 1973)

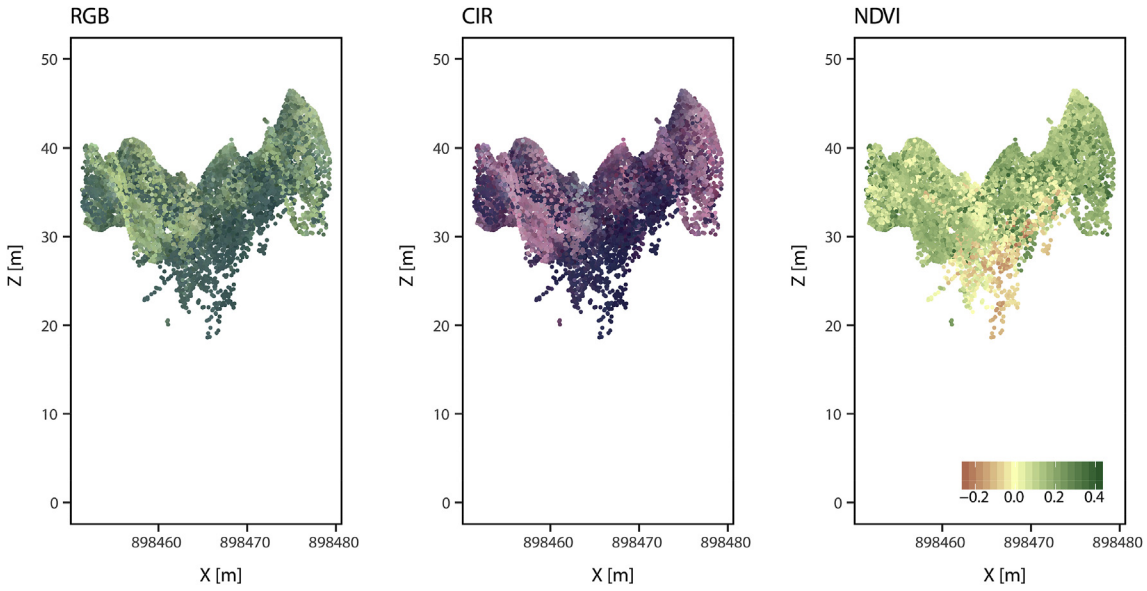

Fig. 2. A profile of a DAP point cloud for a field plot showing different color composites (RGB and CIR) as well as the NDVI values associated with every point in the cloud. CIR – colour-infrared. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

derived from the spectral information corresponding to each of the ALS returns were used to characterize the three-dimensional variability of the NDVI (Fig. 2). NDVI was first calculated for each point and then summarized with mean, standard deviation, coefficient of variation, and percentiles at the plot level. To characterize the vertical variability of the NDVI, we divided the point cloud into 10 equal vertical layers with layer height equal to 10% of the point heights range and calculated a mean NDVI for each of the layers. We also calculated the average NDVI above and below every height percentile.

2.4. Experimental design

To examine the differences between the prediction accuracy of ALS- and DAP-based metrics, as well as the role of spectral-based metrics, we designed an experimental framework that included different combinations of available predictors to model HL, QMD, BA, V, and N. The experimental design included four scenarios (Table 6). Overall, three classes of predictors were considered in model development: point cloud metrics, CHM metrics, and DAP spectral metrics. In the first scenario we created area-based models using ALS-derived metrics as predictors. Results from this scenario were then used as a baseline for a comparison with results from the three other scenarios. In the second scenario the area-based models were developed using DAP point cloud metrics. In the third scenario only DAP spectral metrics were used, while in the final, fourth scenario, models were developed using all available DAP metrics (i.e. both point cloud and spectral metrics). A random forest (RF) modelling approach was used to develop predictive models for each of the five forest stand attributes. RF was selected because it is computationally efficient, robust to overfitting, readily parameterized, and provides useful measures of variable importance (Breiman, 2001). Prediction accuracy was assessed by calculating an R^2 coefficient, bias, and RMSE (absolute and relative). Leave-one-out cross validation was used, and the accuracy measures were calculated based on the predictions for the left-out samples.

The R^2 was calculated using the following formula:

Table 6
Predictor sets used in each of the four model scenarios examined.

Area-based model scenario	Predictors included
S1 – Reference	ALS point cloud and CHM metrics (Tables 3 and 4); reference data for benchmarking
S2	DAP point cloud and CHM metrics (Tables 3 and 4)
S3	DAP spectral metrics only (Table 5)
S4	DAP point cloud, CHM, and spectral metrics (Tables 3–5)

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \quad (1)$$

where SSE is sum of square error, SST is total sum of squares, y is the observed value, $\hat{y}$ is the predicted value, and $\bar{y}$ is the mean of observed values.

Bias and RMSE were calculated as follows:

$$bias = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i) \quad (2)$$

$$bias\% = \frac{bias}{\bar{y}} * 100 \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \quad (4)$$

$$RMSE\% = \frac{RMSE}{\bar{y}} * 100 \quad (5)$$

where: N is the number of plots, y_i is the observed value, $\hat{y}_i$ is the predicted value, and $\bar{y}$ is the mean of the inventory variable.

In addition, we used statistical equivalence tests to assess the predictions. The standard statistical tests (e.g. *t*-test, Wilcoxon test) are designed to test for differences in the mean values (or mean ranks), i.e. the null hypothesis is posit that there are no differences between the model and the data (Robinson et al., 2005). Failure to reject the null hypothesis in a standard statistical test indicates we do not have enough evidence to suggest that the two variables are different. However, this result does not warrant that the variables are similar as the selected test may be weak (Fekety et al., 2018; Robinson and Froese, 2004). These tests are also sensitive to the sample size, as their ability to detect differences increases with increasing sample size (Robinson et al., 2005; Robinson and Froese, 2004). As opposed to the standard tests, equivalence tests reverse the usual null hypothesis – they postulate that the compared variables are different and use the data to prove otherwise, and are therefore more suitable for model validation (Robinson et al., 2005).

We used the *equiv.boot* function available in the equivalence R package (Robinson, 2016) to validate the models. This implementation of the test uses the bootstrapped, regression-based, two one-sided test (TOST) of equivalence. In the regression-based validation strategy a linear regression model is established between the two compared variables (e.g. observation and prediction), and predetermined regions of equivalence are established for the intercept and slope. The tests for intercept and slope are performed separately, and are based on determining if the confidence intervals are contained inside the regions of equivalence (Robinson et al., 2005). As described by Fekety et al. (2018) the test for intercept informs on the bias and determines if the means of the two compared variables are equal. The test for slope, which determines if the slope is equal to 1, informs on the proportionality of the observed and predicted variables. We used the default region of equivalence of ± 25% as per previous studies (Fekety et al., 2015, 2018; Penner et al., 2013).

The equivalence test was used to evaluate the predictions of each scenario with field data as reference, as well as to compare predictions derived using the various combinations of DAP data metrics (scenarios S2, S3, and S4) with ALS-based predictions (scenario S1) used as reference. These two variants of the test allowed us to use it not only for model validation, but also further investigate how different the DAP-based predictions are from the ALS-derived baseline when different sets of metrics are used. In addition we also used the test to determine if the results of scenarios S2 and S4 are equivalent.

Lastly, we used the RF-based method of characterizing variable importance to identify variables with the highest contribution for each of the models. Instead of calculating the variable importance as the mean decrease in accuracy if a given variable is randomly permuted, we used the conditional variable importance measure as calculated using the *cforest* function implemented in the party package for R (R Core Team, 2019; Strobl et al., 2008). Strobl et al. (2008) showed that the variable importance calculated in the original randomForest function (Breiman, 2001) assigns higher values of importance to correlated predictor variables and that the variable importance is affected by the number of categories and scale of measurement of the predictors (Strobl et al., 2007). As a result Strobl et al. (2008) suggested the conditional variable importance measure as a reliable alternative.

In this study we calculated the variable importance during every model run of the leave-one-out cross validation. Because it was possible that the variable importance values would be slightly different, variable importance values for every model run were saved and then averaged across all model runs. Before every model run we also evaluated correlations between predictor variables and kept only predictors that had a Pearson's correlation coefficient less than 0.8. Correlation between predictors has been demonstrated to influence variable importance measures in random forests, as well as to dilute the importance of key predictors (Gregorutti et al., 2017; Kuhn and Johnson, 2013; Strobl et al., 2008).

3. Results

Our results indicated that model prediction accuracy was highest when the ALS-derived point cloud metrics were used as predictors (S1; Table 7). For this scenario, models for each of the five forest stand attributes had higher R^2, and lower RMSE%. The proportion of explained variance was highest for models predicting HL and lowest for models predicting N, with R^2 values equal to 0.92 and 0.33, respectively. RF predictions of HL with ALS metrics had also the lowest bias% (−0.16%) and RMSE (9.23%). Predictions of QMD, BA, and V had R^2 values ranging from 0.43 to 0.65, and RMSE% ranging from 22.04% to 34.16% (Table 7). The results of the equivalence test (Fig. 4) showed that the predictions of all attributes were statistically equivalent in terms of bias, and that predictions of HL and V were statistically equivalent in terms of proportionality.

The results reported in S1 established a baseline for comparison with prediction accuracies in S2, S3, and S4 (Fig. 3). The most accurate predictions for DAP-derived metrics achieved for S2 (DAP point cloud metrics only) and S4 (DAP point cloud and spectral metrics combined), and lowest for S3 (DAP spectral metrics only). The results of the

Table 7
Accuracy of the five stand attributes (HL, QMD, BA, V, N) predicted under four different scenarios with different sets of metrics used as predictors: S1 – ALS metrics; S2 – DAP point cloud metrics; S3 – DAP spectral metrics; S4 – DAP point cloud and spectral metrics.

Attribute	Scenario	R^2	bias	Bias%	RMSE	RMSE%
HL	S1	0.89	0.21	0.68	3.34	10.65
HL	S2	0.75	0.17	0.56	4.93	15.74
HL	S3	0.14	0.03	0.11	9.25	29.52
HL	S4	0.76	0.17	0.53	4.90	15.63
QMD	S1	0.60	0.12	0.31	8.99	22.73
QMD	S2	0.43	0.34	0.86	10.72	27.10
QMD	S3	0.24	0.11	0.27	12.40	31.35
QMD	S4	0.49	0.15	0.39	10.21	25.80
BA	S1	0.42	−0.14	−0.20	25.21	35.81
BA	S2	0.24	0.61	0.87	28.95	41.13
BA	S3	0.16	0.41	0.58	30.48	43.30
BA	S4	0.30	0.65	0.93	27.72	39.38
V	S1	0.63	4.57	0.50	319.20	34.75
V	S2	0.50	7.66	0.83	374.17	40.74
V	S3	0.09	2.43	0.26	501.86	54.64
V	S4	0.50	8.42	0.92	372.07	40.51
N	S1	0.37	0.74	0.11	292.53	42.33
N	S2	0.20	−0.24	−0.03	330.78	47.87
N	S3	0.07	2.46	0.36	355.64	51.47
N	S4	0.19	2.15	0.31	331.24	47.94

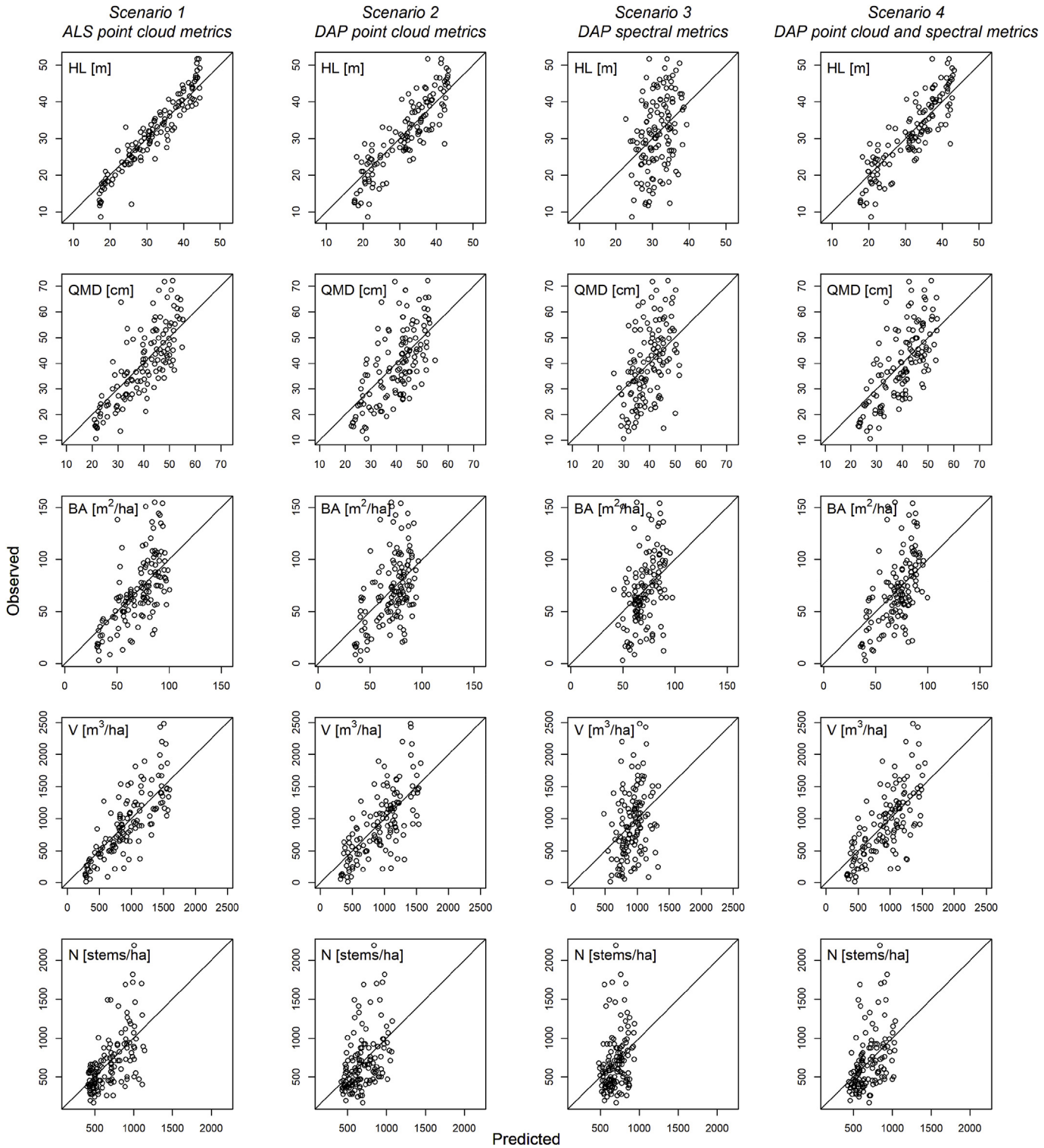

Fig. 3. Scatterplots of predicted vs observed values for the five modelled forest stand attributes (HL, QMD, BA, V, and N), under each scenario. HL – Lorey's height, QMD – quadratic mean diameter, BA – basal area, V – volume, N – number of stems.

equivalence test for these scenarios showed that the predictions of all attributes were statistically equivalent in terms of bias. In terms of proportionality, predictions of HL and BA were statistically equivalent for S2. In S4 only predictions of HL were statistically equivalent.

. For S3, the achieved prediction accuracy was low, with high RMSE % (e.g. 51.00% for V) and low R^2 (e.g. 0.15 for BA, 0.09 for V). The results of the equivalence test for this scenario showed that no predictions for any of the forest attributes considered were statistically equivalent in terms of proportionality.

The effect of using a combination of point cloud- and spectral-based metrics on the prediction accuracy is depended on the target attribute. Results for S2 and S4 were similar to each other, especially for HL and V (Fig. 4). Predictions of HL had the highest R^2 equal to 0.79 for both scenarios, with RMSE% of 14.59 and 14.67, respectively. Accuracy of QMD, BA, and N was slightly higher in S4 than in S2. For example, the R^2 for QMD increased from 0.47 to 0.52, and the RMSE% decreased

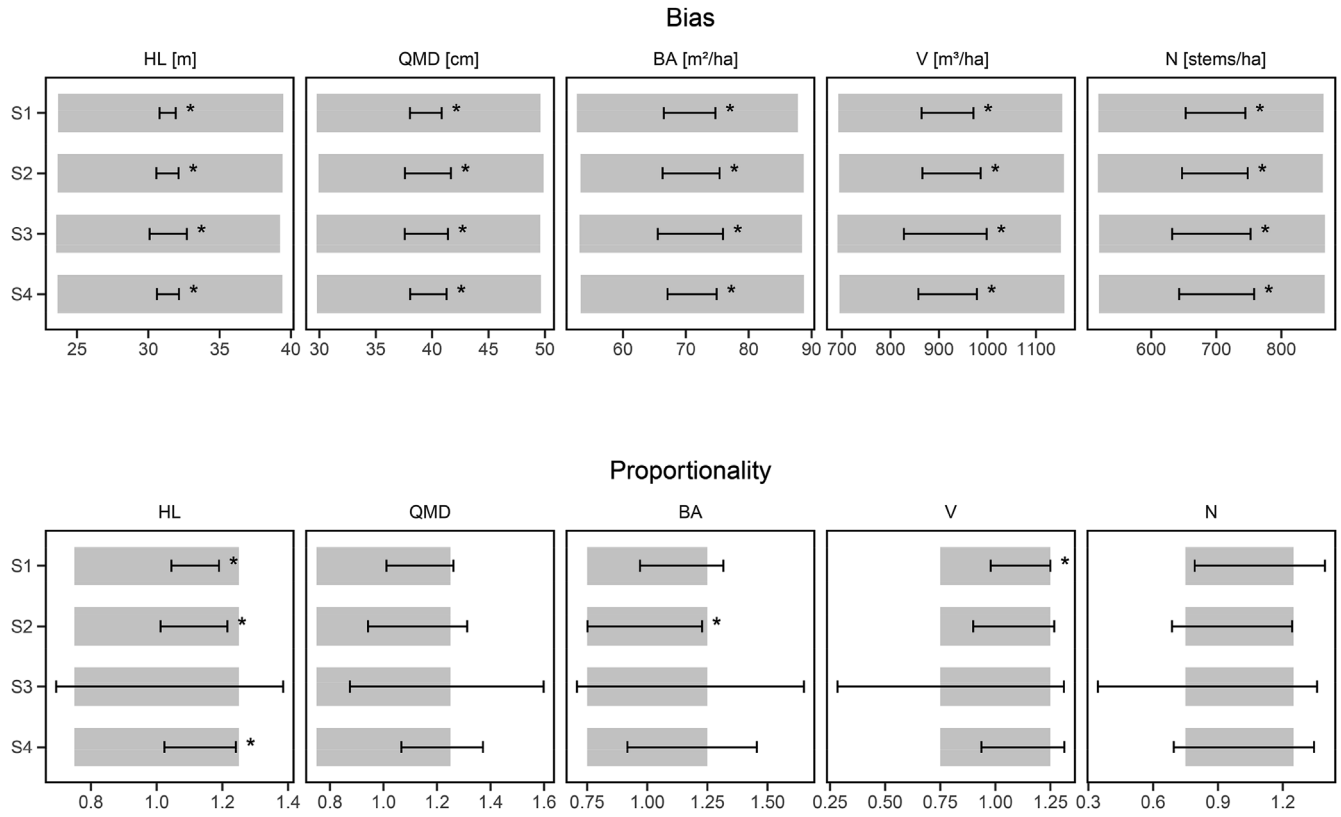

Fig. 4. Results of bootstrapped equivalence tests for all scenarios (S1 – S4) and all stand attributes. In each case (scenario + stand attribute) the equivalence test is performed for bias and for proportionality. The test for bias is centered on the mean value of the observed attribute, while the test for proportionality is centered on a slope value of 1. The grey rectangle indicates the region of equivalence, while the black crossbar depicts the 95% confidence interval. Asterisk indicates cases when the test is satisfied (i.e. confidence interval within the region of equivalence).

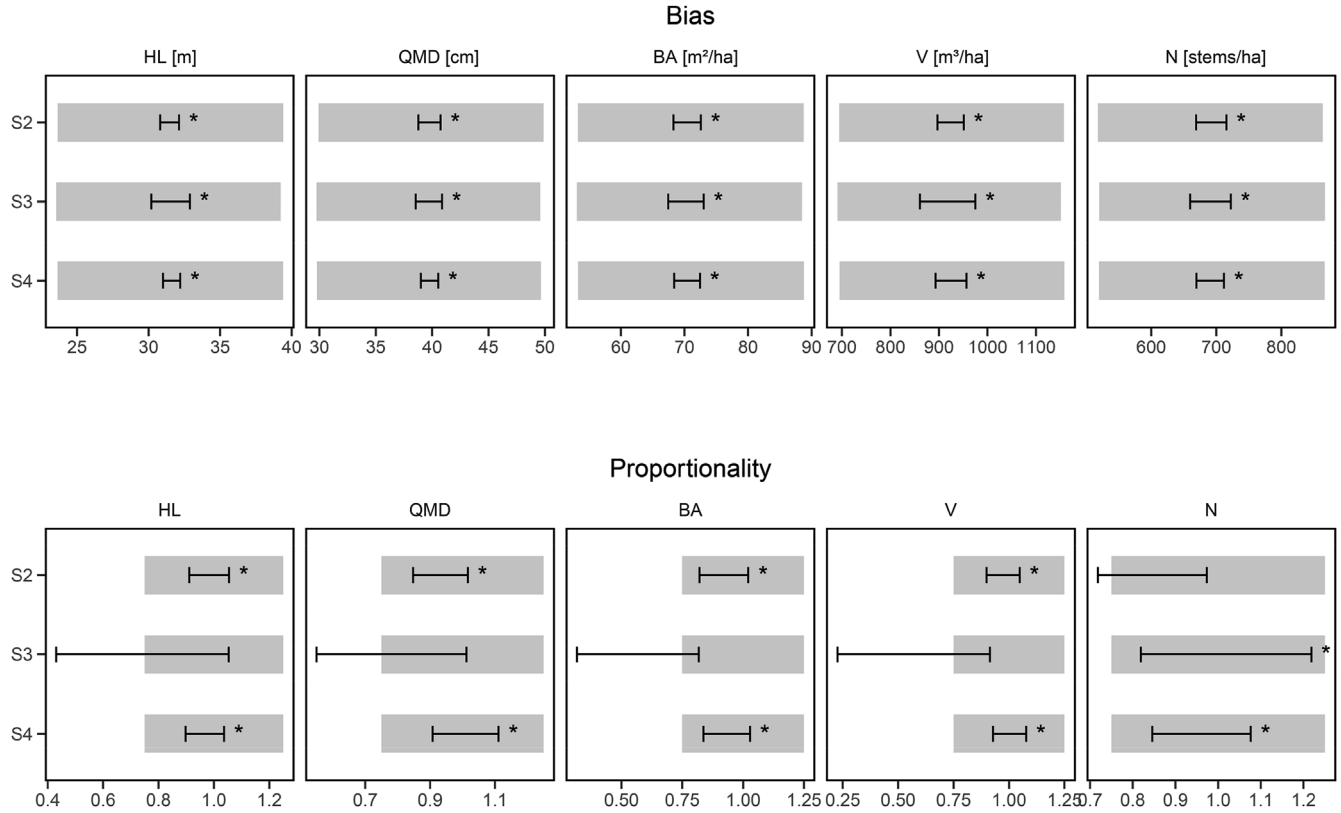

Fig. 5. Results of bootstrapped equivalence tests for scenarios that used DAP-derived metrics as predictors (S2 – S4) with ALS-based predictions (S1) used as reference. See Fig. 4 for details regarding figure interpretation.

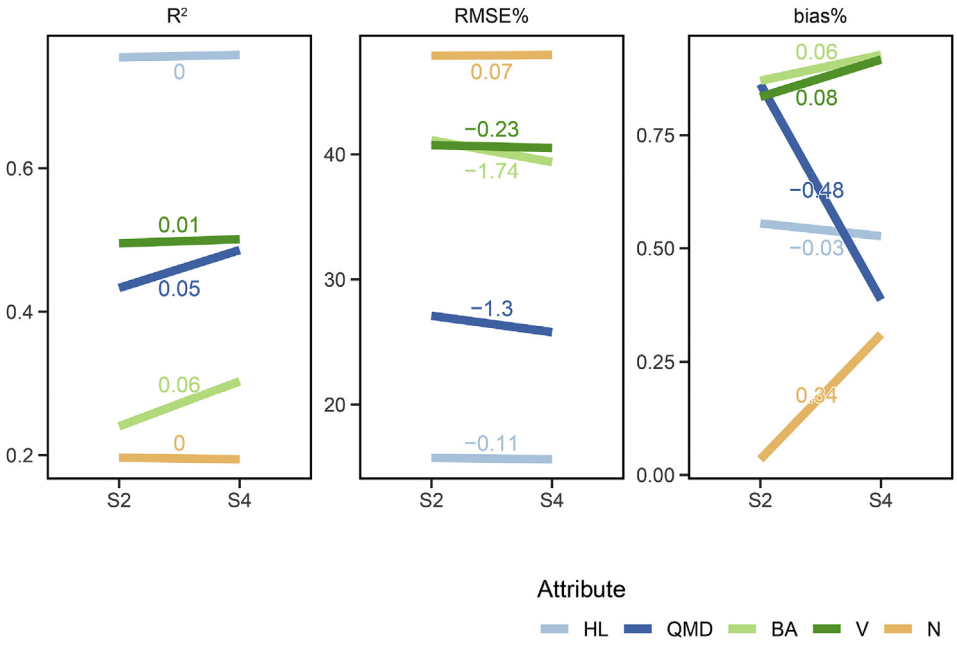

Fig. 6. Change in R^2, RMSE% and bias% between scenario 2 (S2) and scenario 4 (S4).

from 26.31% to 24.85%. The results of the equivalence test (Fig. 4) also showed similar outcome for these two scenarios. However, when using ALS-based predictions as a baseline (Fig. 5) the results of the equivalence test showed that in S4 all of the stand attributes were statistically equivalent for both bias and proportionality. These results were slightly better than for S2, in which the null hypothesis of the test for proportionality was not rejected for N.

Fig. 6 shows how the model accuracy changed when the DAP data was used without (S2) or with spectral metrics. The addition of spectral metrics never resulted in a decrease in R^2 or an increase in RMSE% for any of the attributes (Fig. 6). The influence that the addition of spectral metrics in S4 had on bias% was less consistent: for HL, QMD, and BA the bias% was reduced, whereas for V and N the bias% increased. By performing the equivalence test on the results of S2 and S4, we found that they were statistically equivalent for both bias and proportionality, for all stand attributes. The addition of spectral metrics in S4 provided no statistically significant increase in prediction accuracy for HL, QMD, BA, and V, relative to those of S2.

Analysis of the variable importance indicated that point cloud metrics are markedly more important than spectral metrics (Fig. 7). For scenarios that were based on point cloud metrics only (S1 and S2) the set of most important predictors was very similar, with height percentiles, CHM-based descriptive statistics, rumple, and proportions as the most important predictors. It is notable that the cover metrics (i.e. metrics describing the proportion of points above a threshold, e.g. PTS_perc_above_threshold2) and metrics describing the variability of point heights (e.g. PTS_zsd) were more important in S1, as the ALS point cloud better characterize vertical structure of the stand.

When the point cloud-based and spectral-based metrics were combined (S4), eight spectral-based metrics were among the top most important predictors and only 1 spectral metric was in the top 5. Indeed, the top 5 predictors for S2 and S4 were identical, with the exception of cumulative percentage of return in the 9th layer (PTS_zpcum9) being replaced by the standard deviation of the green band (SPC_G_Std.Dev). Overall, the important predictors for S4 were very similar to those for S2. For S3, the standard deviation in the green band was identified as the single most important metric for all modelled forest stand attributes, with markedly higher importance than any of the other predictors.

4. Discussion

Research has demonstrated that DAP data can provide results analogous to ALS data when used in an area-based approach to estimating forest inventory attributes; however, the results obtained with ALS data are always superior (Goodbody et al., 2019). In this work, we aimed to determine if the accuracy of DAP area-based predictions can be improved through the integration of spectral information. Different sets of ALS- and DAP-derived metrics were used to compare prediction accuracy of five forest stand attributes: HL, QMD, BA, V, and N. Using four different model scenarios with four different sets of predictor variables (Table 6) we analyzed the relative contribution of spectral metrics to improving the prediction accuracy of DAP area-based models. The relative importance of candidate predictors was also assessed to identify those predictors or spectral wavelengths that were most informative for the developed models.

ALS models were used to establish a baseline for the comparison and showed to have similar accuracy to previously reported results in the study area (White et al., 2015). The DAP-point cloud only models (scenario 2) were less accurate, which also corresponds to the findings reported in White et al. (2015), and to the accuracies reported in the literature (Goodbody et al., 2019; Noordermeer et al., 2019; Straub et al., 2013; Vastaranta et al., 2013; White et al., 2015). The addition of spectral metrics (S4) did not result in a statistically significant increase in prediction accuracy for any of the modelled attributes. The difference in model accuracy statistics was in fact minor and observed for only for 3 out of 5 modelled attributes (QMD, BA, V) and the maximum difference in R^2 coefficient between the two scenarios was 0.06. When compared to ALS-based predictions, the addition of spectral metrics showed to be important in only one case – predictions of N were statistically equivalent to predictions derived with ALS only when the point clouds and spectral metrics were combined.

The modeling scenario that incorporated only spectral metrics resulted in the least accurate estimates of every stand attribute. This confirms the importance of the three-dimensional information on the forest structure for predicting inventory attributes and demonstrates that the metrics used herein derived from the spectral values extracted from the point cloud are not suitable for this type of modeling. In particular, none of the metrics characterizing spectral values in different horizontal strata of the point cloud (and therefore taking the

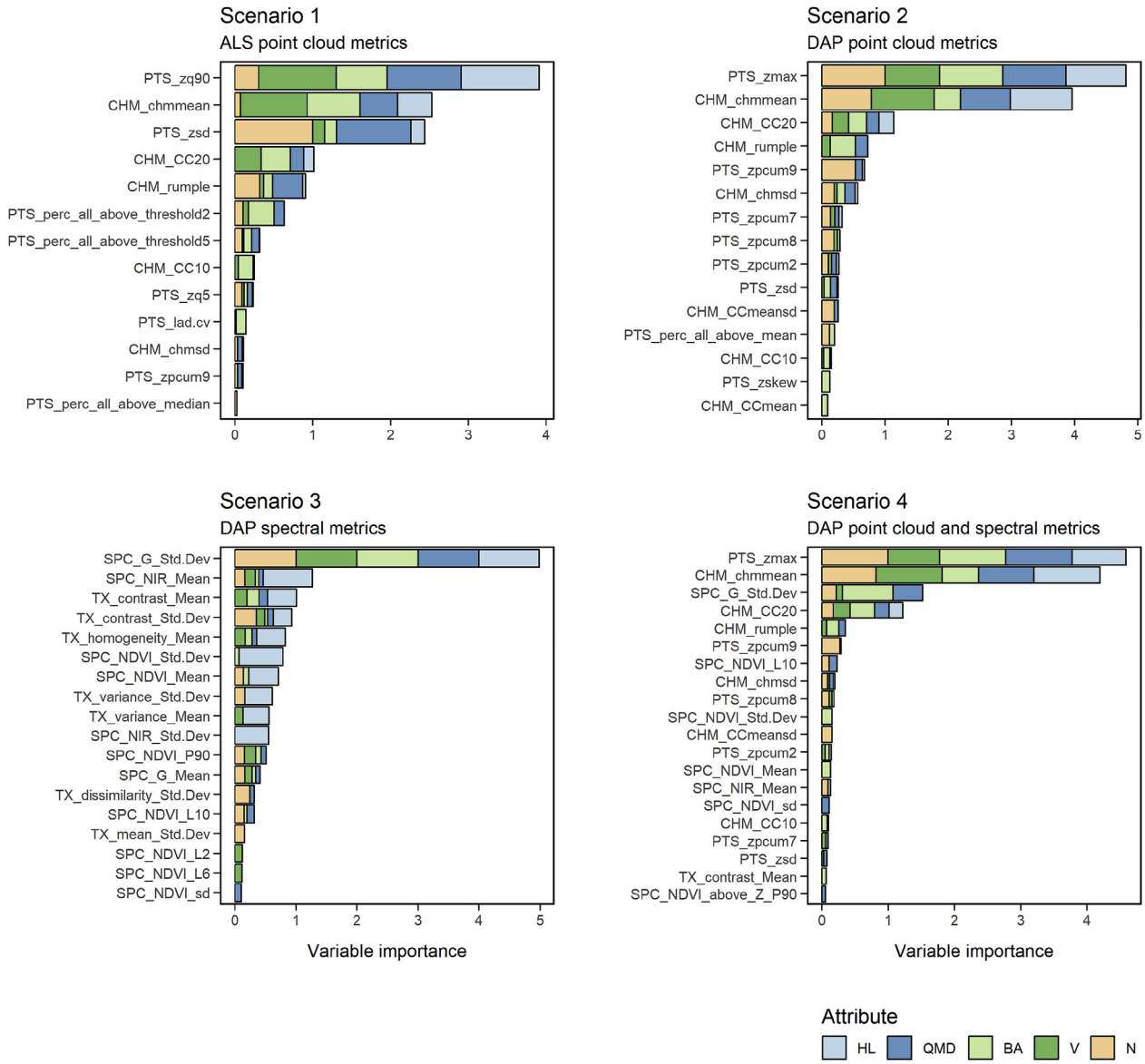

Fig. 7. Scaled variable importance values of the 10 most important metrics for each of the modelled forest stand attributes under each scenario. Each panel contains variable importance values for each of the scenarios. Within each panel, the stacked colored bars indicate the total importance of given predictor variable, with colors depicting the five modelled attributes. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

advantage of the combined spatial and spectral information of DAP point cloud), were found to be important for model development. The most important predictor was the standard deviation of the green band, derived from the orthophoto; this metrics was also the most important spectral variable when the point cloud and spectral metrics were combined. This finding corresponds to results reported by Puliti et al. (2015) who used a combination of UAV-acquired point cloud and spectral metrics to estimate stand attributes. They also found the green band-based metrics (mean and standard deviation) to be the most relevant spectral variables included in their models.

The most important and informative metrics for the estimation of HL, QMD, BA, V, and N using DAP data were those metrics derived from the point clouds. In models developed with the point cloud metrics only, and in models developed with the combined point cloud and spectral metrics, metrics characterizing canopy height and canopy cover were consistently identified as the most important variables for model development. This result corresponds to results reported in the literature for studies that used ALS point clouds integrated with aerial or satellite imagery (Erdody and Moskal, 2010; Hudak et al., 2006; Latifi et al., 2010; Tonolli et al., 2011). Here we confirm that the same applies to predictions based on DAP data – the benefit of the additional spectral information is negligible, with no clear and significant increase in accuracy observed for any of the modelled forest stand attributes. Similar outcomes were also reported by Puliti et al. (2015) and Kukkonen et al. (2017). In complementary research using spectral information in conjunction with ALS data, Melin et al. (2019) found that when building predictive models for avian habitat, the inclusion of hyperspectral data did not improve model performance over the use of ALS data alone.

Puliti et al. (2015) theorize that the limited improvement of prediction accuracy when the spectral variable were included might be caused by the time of their data acquisition (late fall) and often poor image quality (lack of spectral calibration of the UAS user-grade camera). In the case of our study, the timing of the image acquisition (August 16, September 25, and October 4–6, between 10:00 and 15:00) had less effect since the majority of the plots were conifer-dominated, and the acquisition was performed with calibrated large-format digital camera. The heterogeneity of the spectral values between individual

images that originates from changing topography, changing sun-surface-sensor configuration during the day (heterogeneity within individual acquisition) or time of year (heterogeneity between acquisition) may result in a lack of consistency between spectral-based metrics and therefore limit their suitability as predictor variables. However, White et al. (2015) explicitly tested the impact of acquisition date (for these same data) on area-based models of height, basal area, and volume, finding no significant difference in model estimates.

Given that tree crowns have internal spectral variability, even similar forest structures will have variable spectral response values resulting from viewing geometry and sensor-sun configurations (Leckie et al., 2005; Maschler et al., 2018; Wulder et al., 2008). This effect is likely to be more prominent in tall stands, where spectral variability increases through the vertical canopy profile and horizontally over a larger crown extent. As described by Leckie et al. (2005), in an old growth temperate conifer forest, the within-species spectral variability is often large because of illumination and view-angle conditions, openness of trees, natural variability, shadowing effects, and a range of crown health, while conversely, the spectral differences between species are often small. From investigations looking at individual tree crowns from high spatial resolution imagery, it is known that there are sunlit and shaded portions of a given crown (Coops et al., 2004; Leckie et al., 2005). As such, the two-dimensional expression of a tree crown is expected to have variability in spectral response from an image based perspective (Korpela et al., 2011; Wulder et al., 2004). When expanded to a three-dimensional environment (i.e. DAP point cloud with spectral information associated with every point), the impact of sun-surface-sensor viewing geometry resulting from variations in forest structure will be exacerbated with points located further from tree tops and into gaps between trees. This spectral variability likely confounds the utility of spectral information to relate meaningful information on forest structure. In contrast, as an active sensor, ALS measurements can be expected to consistently characterize the same structural conditions in the same way, when differences in instruments and acquisition parameters accounted for. Future studies aiming to link DAP spectral and structural information in a predictive sense would be improved by describing the expected relationship between the forest structural (and compositional) characteristics present in the area of interest and the spectral characteristics of the given sensor and resulting data.

As there was no marked improvement in prediction accuracy for models that incorporated spectral metrics versus those that did not, we cannot recommend incorporation of DAP spectral-based metrics as predictors for estimating HL, QMD, BA, V, or N with area-based approach. Moreover, the effort required to obtain and integrate the spectral information cannot be justified based on the results presented herein. Past studies that have attempted to integrate spectral and structural information have been motivated by a need to generate species-specific models (Niska et al., 2010; Packalén and Maltamo, 2007) in an environment with limited species complexity. Such information is less of a driver in this context, and the results of previous studies have indicated that the results for this application (distinguishing species) is not particularly compelling. In fact, based on the results, we cannot identify situations (forest types, stand attributes) that would always benefit from including the spectral metrics in the modeling.

Finally, we recognize that although the spectral information did not lead to increase in prediction accuracy in our case, it may be useful in forest stands of different structure or species composition. For example, Maltamo et al. (2006) found that adding spectral metrics may slightly improve the accuracy of the ALS-based volume predictions. They reported that the relative RMSE of volume decreased from 15.58% to 13.28% when data from aerial images was included. Puliti et al. (2017b) demonstrated the benefits of including the spectral information when predicting species proportions and species-specific volume. However, when compared to ALS, a combination of a DAP point cloud and spectral metrics did not result in higher prediction accuracy of plot-level total volume or species-specific volume (Puliti et al., 2017b). Spectral-based predictors may be also beneficial for individual tree based approach. As shown by St-Onge et al. (2015) when the point cloud and spectral data were integrated for individual tree approach, they were more accurate for discriminating species than ALS-derived structure and intensity metrics.

To summarize, as the influence of spectral metrics reported in the literature is not consistent, their role in predictions should be evaluated on a case by case basis with consideration of the effort required to generate them. Our results showed that among the spectral variables, the metrics based on individual green band were identified as most important, and, in general, similar spectral metrics were most informative in spectral-only models and in models that were based on integrated point cloud and spectral metrics. These spectral variables should be therefore among the first to incorporate in addition to point cloud metrics.

5. Conclusions

In this study we compared how different sets of predictors derived from DAP dataset influence the accuracy of area-based estimates of HL, QMD, BA, V, and N. We compared the results with the observed data (field plots) as well as with ALS-derived predictions. We found that even the best results derived with DAP data had lower accuracy than ALS-based estimates, although the difference was often small and acceptable for inventory applications. We found no benefit in including spectral metrics derived from the DAP-based ortho or DAP point cloud – the differences between model runs that included point cloud based predictors combined with spectral metrics, when compared to models developed using only point cloud predictors, were negligible. Given these results and caveats discussed herein, the level of effort required to extract the spectral metrics and incorporate them into area-based models is not presently justified for operational applications.

Acknowledgements

This research was supported by the Canadian Wood Fibre Centre (CWFC) of the Canadian Forest Service, Natural Resources Canada, with data for northern Vancouver Island generously supplied by Western Forest Products Inc. and BC Timber Sales. Support was also provided by a Natural Sciences and Engineering Research Council of Canada (NSERC) grant to Nicholas Coops. We thank the editors and three anonymous reviewers for their time and insightful comments that allowed us to improve our manuscript. Open Access supported by the Government of Canada.

References

Wulder, M.A., Hall, R.J., Coops, N.C., Franklin, S.E., 2004. High spatial resolution remotely sensed data for ecosystem characterization. Bioscience 54 (6), 511. https://doi.org/10.1641/0006-3568(2004)054[0511:hsrrsd]2.0.co;2.

Agisoft, 2019. Metashape. Version 1.5. Retrieved from. https://www.agisoft.com/.

Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. International Archives of Photogrammetry and Remote Sensing XXXIII, 110–117.

Baltsavias, E.P., 1999. A comparison between photogrammetry and laser scanning. ISPRS J. Photogrammetry Remote Sens. 54 (2), 83–94. Retrieved from. http://www.sciencedirect.com/science/article/pii/S0924271699000143.

Bohlin, J., Wallerman, J., Fransson, J.E.S., 2012. Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution DEM. Scand. J. For. Res. 27 (7), 692–699. https://doi.org/10.1080/02827581.2012.686625.

Bouvier, M., Durrieu, S., Fournier, R.a., Renaud, J.-P., 2015. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. Remote Sens. Environ. 156, 322–334. https://doi.org/10.1016/j.rse.2014.10.004.

Breiman, L., 2001. Random forest. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324.

Cohen, W.B., Maiersperger, T.K., Spies, T.A., Oetter, D.R., 2001. Modelling forest cover attributes as continuous variables in a regional context with Thematic Mapper data. Int. J. Remote Sens. 22 (12), 2279–2310. https://doi.org/10.1080/01431160121472.

Coops, N.C., Wulder, M.A., Culvenor, D.S., St-Onge, B., 2004. Comparison of forest attributes extracted from fine spatial resolution multispectral and lidar data. Can. J. Remote Sens. 30 (6), 855–866. Retrieved from. http://cat.inist.fr/?aModele=afficheN&cpsidt=16438849.

Coops, N.C., Hilker, T., Hall, F.G., Nichol, C.J., Drolet, G.G., 2010. Estimation of light-use efficiency of terrestrial ecosystems from space: a status report. Bioscience 60 (10), 788–797. https://doi.org/10.1525/bio.2010.60.10.5.

Erdody, T.L., Moskal, L.M., 2010. Fusion of LiDAR and imagery for estimating forest canopy fuels. Remote Sens. Environ. 114 (4), 725–737. https://doi.org/10.1016/j.rse.2009.11.002.

Fekety, P.A., Falkowski, M.J., Hudak, A.T., 2015. Temporal transferability of LiDAR-based imputation of forest inventory attributes. Can. J. For. Res. 45 (4), 422–435. https://doi.org/10.1139/cjfr-2014-0405.

Fekety, P.A., Falkowski, M.J., Hudak, A.T., Jain, T.B., Evans, J.S., 2018. Transferability of lidar-derived basal area and stem density models within a northern Idaho ecoregion. Can. J. Remote Sens. 44 (2), 131–143. https://doi.org/10.1080/07038992.2018.1461557.

Franklin, S.E., McDermid, G.J., 1993. Empirical relations between digital SPOT HRV and CASI spectral response and lodgepole pine (Pinus contorta) forest stand parameters. Int. J. Remote Sens. 14 (12), 2331–2348. https://doi.org/10.1080/01431169308954040.

Franklin, S.E., Hall, R.J., Moskal, L.M., Maudie, A.J., Lavigne, M.B., 2000. Incorporating texture into classification of forest species composition from airborne multispectral images. Int. J. Remote Sens. 21 (1), 61–79. https://doi.org/10.1080/014311600210993.

Goodbody, T.R.H., Coops, N.C., White, J.C., 2019. Digital aerial photogrammetry for updating area-based forest inventories: a review of opportunities, challenges, and future directions. Current Forestry Reports. https://doi.org/10.1007/s40725-019-00087-2.

Gregorutti, B., Michel, B., Saint-Pierre, P., 2017. Correlation and variable importance in random forests. Stat. Comput. 27 (3), 659–678. https://doi.org/10.1007/s11222-016-9646-1.

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics, SMC-3 6, 610–621. https://doi.org/10.1109/TSMC.1973.4309314.

Hawryło, P., Tompalski, P., Wężyk, P., 2017. Area-based estimation of growing stock volume in Scots pine stands using ALS and airborne image-based point clouds. Forestry, i 1–11. https://doi.org/10.1093/forestry/cpx026.

Hijmans, R.J., 2019. Raster: geographic data analysis and modeling. R package version 2.9-23. Retrieved from. https://cran.r-project.org/package=raster.

Hilker, T., Lyapustin, A., Hall, F.G., Wang, Y., Coops, N.C., Drolet, G., Black, T.A., 2009. An assessment of photosynthetic light use efficiency from space: modeling the atmospheric and directional impacts on PRI reflectance. Remote Sens. Environ. 113 (11), 2463–2475. https://doi.org/10.1016/j.rse.2009.07.012.

Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M., Gessler, P.E., Morgan, P., 2006. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. Can. J. Remote Sens. 32 (2), 126–138. https://doi.org/10.5589/m06-007.

Isenburg, M., 2018. LAStools. Version 180605, academic. Retrieved from. https://rapidlasso.com/.

Kane, V.R., Bakker, J.D., McGaughey, R.J., Lutz, J.A., Gersonde, R.F., Franklin, J.F., 2010. Examining conifer canopy structural complexity across forest ages and elevations with LiDAR data. Can. J. For. Res. 40, 774–787. https://doi.org/10.1139/X10-064.

Korpela, I., Heikkinen, V., Honkavaara, E., Rohrbach, F., Tokola, T., 2011. Variation and directional anisotropy of reflectance at the crown scale - implications for tree species classification in digital aerial images. Remote Sens. Environ. 115 (8), 2062–2074. https://doi.org/10.1016/j.rse.2011.04.008.

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer New York, New York, NY. https://doi.org/10.1007/978-1-4614-6849-3.

Kukkonen, M., Maltamo, M., Packalen, P., 2017. Image matching as a data source for forest inventory – comparison of Semi-Global Matching and Next-Generation Automatic Terrain Extraction algorithms in a typical managed boreal forest environment. Int. J. Appl. Earth Obs. Geoinf. 60, 11–21. https://doi.org/10.1016/j.jag.2017.03.012.

Latifi, H., Nothdurft, a., Koch, B., 2010. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. Forestry 83 (4), 395–407. https://doi.org/10.1093/forestry/cpq022.

Leberl, F., Irschara, A., Pock, T., Meixner, P., Gruber, M., Scholz, S., Wiechert, A., 2010. Point clouds. Photogramm. Eng. Remote Sens. 76 (10), 1123–1134. https://doi.org/10.14358/PERS.76.10.1123.

Leckie, D.G., Gougeon, F., Hill, D., Quinn, R., Armstrong, L., Shreenan, R., 2003. Combined high-density lidar and multispectral imagery for individual tree crown analysis. Can. J. Remote Sens. 29 (5), 633–649.

Leckie, D.G., Tinis, S., Nelson, T.A., Burnett, C., Gougeon, F.A., Cloney, E., Paradine, D., 2005. Issues in species classification of trees in old growth conifer stands. Can. J. Remote Sens. 31 (2), 175–190. Retrieved from. http://geog.uvic.ca/spar/papers/CJRS_05.pdf.

Magnussen, S., Næsset, E., Gobakken, T., 2010. Reliability of LiDAR derived predictors of forest inventory attributes: a case study with Norway spruce. Remote Sens. Environ. 114 (4), 700–712. https://doi.org/10.1016/j.rse.2009.11.007.

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

Maschler, J., Atzberger, C., Immitzer, M., 2018. Individual tree crown segmentation and classification of 13 tree species using airborne hyperspectral data. Remote Sens. 10 (8), 1218. https://doi.org/10.3390/rs10081218.

Melin, M., Korhonen, L., Kukkonen, M., Packalen, P., 2017. Assessing the performance of aerial image point cloud and spectral metrics in predicting boreal forest canopy cover. ISPRS J. Photogrammetry Remote Sens. 129, 77–85. https://doi.org/10.1016/j.isprsjprs.2017.04.018.

Melin, M., Hill, R.A., Bellamy, P.E., Hinsley, S.A., 2019. On bird species diversity and Remote sensing — utilizing lidar and hyperspectral data to assess the role of vegetation structure and foliage characteristics as drivers of avian diversity. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 1–9. https://doi.org/10.1109/JSTARS.2019.2906940.

Niska, H., Skon, J.-P., Packalen, P., Tokola, T., Maltamo, M., Kolehmainen, M., 2010. Neural networks for the prediction of species-specific plot volumes using airborne laser scanning and aerial photographs. IEEE Trans. Geosci. Remote Sens. 48 (3), 1076–1085. https://doi.org/10.1109/TGRS.2009.2029864.

Nolan, M., Larsen, C.F., Sturm, M., 2015. Mapping snow-depth from manned-aircraft on landscape scales at centimeter resolution using Structure-from-Motion photogrammetry. Cryosphere Discuss. 9 (1), 333–381. https://doi.org/10.5194/tcd-9-333-2015.

Noordermeer, L., Bollandsås, O.M., Ørka, H.O., Næsset, E., Gobakken, T., 2019. Comparing the accuracies of forest attributes predicted from airborne laser scanning and digital aerial photogrammetry in operational forest inventories. Remote Sens. Environ. 226 (March), 26–37. https://doi.org/10.1016/J.RSE.2019.03.027.

Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. 80 (1), 88–99. https://doi.org/10.1016/S0034-4257(01)00290-5.

Packalén, P., Maltamo, M., 2007. The k-MSN method for the prediction of species-specific stand attributes using airborne laser scanning and aerial photographs. Remote Sens. Environ. 109 (3), 328–341. https://doi.org/10.1016/j.rse.2007.01.005.

Penner, M., Pitt, D.G., Woods, M.E., 2013. Parametric vs. nonparametric LiDAR models for operational forest inventory in boreal Ontario. Can. J. Remote Sens. 39 (5), 426–443. https://doi.org/10.5589/m13-049.

Pitt, D.G., Woods, M., Penner, M., 2014. A Comparison of point clouds derived from stereo imagery and airborne laser scanning for the area-based estimation of forest inventory attributes in boreal Ontario. Can. J. Remote Sens. 40 (3), 214–232. https://doi.org/10.1080/07038992.2014.958420.

Puliti, S., Ørka, H.O., Gobakken, T., Næsset, E., 2015. Inventory of small forest areas using an unmanned aerial system. Remote Sens. 7 (8), 9632–9654. https://doi.org/10.3390/rs70809632.

Puliti, S., Ene, L.T., Gobakken, T., Næsset, E., 2017a. Use of partial-coverage UAV data in sampling for large scale forest inventories. Remote Sens. Environ. 194, 115–126. https://doi.org/10.1016/j.rse.2017.03.019.

Puliti, S., Gobakken, T., Ørka, H.O., Næsset, E., 2017b. Assessing 3D point clouds from aerial photographs for species-specific forest inventories. Scand. J. For. Res. 32 (1), 68–79. https://doi.org/10.1080/02827581.2016.1186727.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria Retrieved from. http://www.r-project.org/.

Rahlf, J., Breidenbach, J., Solberg, S., Næsset, E., Astrup, R., 2017. Digital aerial photogrammetry can efficiently support large-area forest inventories in Norway. Forestry 90 (5), 710–718. https://doi.org/10.1093/forestry/cpx027.

Remondino, F., Spera, M.G., Nocerino, E., Menna, F., Nex, F., 2014. State of the art in high density image matching. Photogramm. Rec. 29 (146), 144–166. https://doi.org/10.1111/phor.12063.

Robinson, A., 2016. Equivalence: provides tests and graphics for assessing tests of equivalence. R package version 0.7.2. Retrieved from. https://cran.r-project.org/package=equivalence.

Robinson, A.P., Froese, R.E., 2004. Model validation using equivalence tests. Ecol. Model. 176, 349–358 doi101016/jecolmodel200401013.

Robinson, A.P., Duursma, R.A., Marshall, J.D., 2005. A regression-based equivalence test for model validation: shifting the burden of proof. Tree Physiol. 25 (7), 903–913. https://doi.org/10.1093/treephys/25.7.903.

Roussel, J.-R., Auty, D., 2018. lidR: airborne LiDAR data manipulation and visualization for forestry applications. Retrieved from. https://github.com/Jean-Romain/lidR.

St-Onge, B., Vega, C., Fournier, R.A., Hu, Y., 2008. Mapping canopy height using a combination of digital stereo-photogrammetry and lidar. Int. J. Remote Sens. 29 (11), 3343–3364. https://doi.org/10.1080/01431160701469040.

St-Onge, B., Audet, F.A., Bégin, J., 2015. Characterizing the height structure and composition of a boreal forest using an individual tree crown approach applied to photogrammetric point clouds. Forests 6 (11), 3899–3922. https://doi.org/10.3390/f6113899.

Stone, C., Webster, M., Osborn, J., Iqbal, I., 2016. Alternatives to LiDAR-derived canopy height models for softwood plantations: a review and example using photogrammetry. Aust. For. 79 (4), 271–282. https://doi.org/10.1080/00049158.2016.1241134.

Straub, C., Stepper, C., Seitz, R., Waser, L.T., 2013. Potential of UltraCamX stereo images for estimating timber volume and basal area at the plot level in mixed European forests. Can. J. For. Res. 43 (8), 731–741. https://doi.org/10.1139/cjfr-2013-0125.

Strobl, C., Boulesteix, A.L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinf. 8. https://doi.org/10.1186/1471-2105-8-25.

Strobl, C., Boulesteix, A.L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinf. 9, 1. https://doi.org/10.1186/1471-2105-9-307.

Tompalski, P., Coops, N.C., White, J.C., Wulder, M.A., 2015. Augmenting site index estimation with airborne laser scanning data. For. Sci. 61 (5), 861–873. https://doi.org/10.5849/forsci.14-175. (28).

Tompalski, P., Coops, N., Marshall, P., White, J., Wulder, M., Bailey, T., 2018. Combining multi-date airborne laser scanning and digital aerial photogrammetric data for forest growth and yield modelling. Remote Sens. 10 (3), 347. https://doi.org/10.3390/rs10020347.

Tonolli, S., Dalponte, M., Neteler, M., Rodeghiero, M., Vescovo, L., Gianelle, D., 2011. Fusion of airborne LiDAR and satellite multispectral data for the estimation of timber volume in the Southern Alps. Remote Sens. Environ. 115 (10), 2486–2498. https://doi.org/10.1016/j.rse.2011.05.009.

Turner, D.P., Ritts, W.D., Cohen, W.B., Gower, S.T., Running, S.W., Zhao, M., Costa, M.H., Kirschbaum, A.A., Ham, J.M., Saleska, S.R., Ahl, D.E., 2006. Evaluation of MODIS NPP and GPP products across multiple biomes. Remote Sens. Environ. 102 (3–4), 282–292. https://doi.org/10.1016/j.rse.2006.02.017.

Vastaranta, M., Wulder, M.A., White, J.C., Pekkarinen, A., Tuominen, S., Ginzler, C., Kankare, V., Holopainen, M., Hyyppa, J., Hyyppa, H., 2013. Airborne laser scanning and digital stereo imagery measures of forest structure : comparative results and implications to forest mapping and inventory update. Can. J. Remote Sens. 39 (5), 382–395.

Vastaranta, M., Saarinen, N., Kankare, V., Holopainen, M., Kaartinen, H., Hyyppä, J., Hyyppä, H., 2014. Multisource single-tree inventory in the prediction of tree quality variables and logging recoveries. Remote Sens. 6 (4), 3475–3491. https://doi.org/10.3390/rs6043475.

Véga, C., St-Onge, B., 2008. Height growth reconstruction of a boreal forest canopy over a period of 58 years using a combination of photogrammetric and lidar models. Remote Sens. Environ. 112 (4), 1784–1794. https://doi.org/10.1016/j.rse.2007.09.002.

White, J.C., Wulder, M.A., Varhola, A., Vastaranta, M., Coops, N.C., Cook, B.D., Pitt, D., Woods, M., 2013a. A Best Practices Guide for Generating Forest Inventory Attributes from Airborne Laser Scanning Data Using an Area-Based Approach.

White, J.C., Wulder, M.A., Vastaranta, M., Coops, N.C., Pitt, D., Woods, M., 2013b. The utility of image-based point clouds for forest inventory: a comparison with airborne laser scanning. Forests 4 (3), 518–536. https://doi.org/10.3390/f4030518.

White, J., Stepper, C., Tompalski, P., Coops, N., Wulder, M., 2015. Comparing ALS and image-based point cloud metrics and modelled forest inventory attributes in a complex coastal forest environment. Forests 6 (10), 3704–3732. https://doi.org/10.3390/f6103704.

White, J.C., Tompalski, P., Vastaranta, M., Wulder, M.A., Saarinen, S., Stepper, C., Coops, N.C., 2017. A Model Development and Application Guide for Generating an Enhanced Forest Inventory Using Airborne Laser Scanning Data and an Area-Based Approach. Canadian Forest Service, Pacific Forestry Centre, Victoria, BC, Canada, pp. 38. https://doi.org/10.5558/tfc2013-132. CWFC Information Report FI-X-018.

Woods, M., Lim, K., Treitz, P., 2008. Predicting forest stand variables from LiDAR data in the great lakes - St. Lawrence forest of ontario. For. Chron. 84 (6), 827–839. https://doi.org/10.5558/tfc84827-6.

Wulder, M.A., Ortlepp, S.M., White, J.C., Coops, N.C., 2008. Impact of sun-surface-sensor geometry upon multitemporal high spatial resolution satellite imagery. Can. J. Remote Sens. 34 (5), 455–461. https://doi.org/10.5589/m08-062.

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 Sens. Environ. 121, 196–209. https://doi.org/10.1016/j.rse.2012.02.001.

Zitová, B., Flusser, J., 2003. Image registration methods: a survey. Image Vis Comput. 21 (11), 977–1000. https://doi.org/10.1016/S0262-8856(03)00137-9.

Zvoleff, Alex, 2019. Glcm: calculate textures from grey-level Co-occurrence matrices (GLCMs). R package version 1.6.4. Retrieved from. https://cran.r-project.org/package=glcm.