https://doi.org/10.26525/jtfs2019.31.3.312

ESTIMATING STAND-LEVEL STRUCTURAL AND BIOPHYSICAL VARIABLES OF LOWLAND DIPTEROCARP FOREST USING AIRBORNE LIDAR DATA

Muhamad-Afizzul M*, Siti-Yasmin Y, Hamdan O & Tan SA

Forest Research Institute Malaysia, 52109 Kepong, Selangor, Malaysia

**afizzul@frim.gov.my*

Submitted June 2018; accepted May 2019

Light Detection and Ranging (LiDAR) has been used in a wide range of applications including forestry. This study aims to investigate the potential use of airborne lidar scanning (ALS) data in estimating stand-level structural and biophysical variables of lowland dipterocarp forest. Five forest variables, namely mean height (Hm), basal area (BA), square mean diameter (Dg), stand density (S) and above ground biomass (AGB), were tested based on 40 field plots. A total of 34 ALS metrics were generated and tested for model development. A multiple linear regression approach was performed to generate the best model for estimating the variables. Models for BA and AGB gave strong precisions, with an adjusted-R^2 of 0.77 and 0.82 and RMSE of 5.45 m^2 ha^{-1} and 71.12 Mg ha^{-1}. The Hm and Dg gave moderate precisions, with R^2 of 0.61 and 0.44 and RMSE of 2.35 m and 6.07 cm, respectively, while S gave the lowest precision with an adjusted-R^2 of 0.27 and RMSE of 149.48 stem ha^{-1}. This study demonstrated that ALS data performs better in estimating stand-level structural and biophysical parameters of tropical forest, which is important for forest managers towards better monitoring, planning and managing their forests by using this technology.

Keywords: LiDAR, forest structure, biomass, multivariate linear regression, tropical forest

INTRODUCTION

Remote sensing technology is an essential resource in the forestry sector. The availibility of various types of space borne optical and radar data with various spatial, spectral and temporal resolutions provide options to the user in selecting and using the best data for monitoring, predicting and managing forest at larger landscape. The use of airborne system has shown significant growth in recent years with the increasingly available sensors on airborne platforms. Initially aerial photograph and hyperspectral images have been used in estimating forest attributes (Ustin & Trabucco 2000, Matejka 2009, Feret & Asner 2012, Shen et al. 2016). However, these data cannot provide the information under the forest canopy layer. Thus, the estimation of forest attributes can only be done at canopy level (Bucha et al. 2012, Feret & Asner 2012, McIntosh et al. 2012).

The use of LiDAR technology in the forestry sector has become prominent in recent years (Listopad et al. 2011, Kent et al. 2015, Leitold et al. 2015, Sato et al. 2016, Hansen et al. 2017). The capability of LiDAR sensor in capturing the information at all forest strata, with high positional and height accuracies, is a clear advantage over optical images, as LiDAR points can penetrate the top canopy layer. This superior advantage gives better results over orthophoto and hyperspectral images in predicting forest variables with high accuracy (Dandois and Ellis 2010, White et al. 2013, Wallace et al. 2016).

Most of the studies related to the estimation of forest attributes have been carried out with promising results in temperate forest and forest plantations (Stone et al. 2011, Treitz et al. 2012, Watt et al. 2013, Ruiz et al. 2014, Crespo-Peremarch et al. 2016). However, only few investigations are available for tropical forest ecosystems especially in Malaysia (Palace et al. 2015, Palace et al. 2016, Phua et al. 2016, Sadadi et al. 2016, Rahman et al. 2017, Wan-Mohd-Jaafar et al. 2017). Thus, more studies need to be carried out in order to explore and utilise ALS data in estimating attributes of tropical forest for management purpose.

Although similar approaches were adopted from previous studies, current investigations are needed on the utilisation of ALS data in estimating stand-level structural and biophysical variables of tropical forest. Variations between studies, such as differences in sensor specification during data acquisition, data acquisition method and ALS point's density, give huge differences in estimating the forest variables.

High precision of prediction on forest attributes, especially for tropical forest, provides added value in assessing the forest ecosystem. The common assessment of forest structure is often carried out by field inventory surveys. Traditional assessment usually requires a substantial amount of time, money and manpower. However, ALS data could be used to assist in assessing the structure of the forest without minimal time and manpower, to cover larger areas (Cauteron et. al. 2012, Thapa et. al. 2015). Thus, it is important to fully explore and utilise the ALS data in predicting forest attributes at highest precision, in order to assess forest structure at larger area with minimal error.

The main objective of this study is to explore the potential use of ALS data in estimating stand-level structural and biophysical variables of tropical forest ecosystem especially in lowland dipterocarp forest. Four structural variables, i.e., mean height (Hm), basal area (BA), square mean diameter (Dg) and stand density (S) were chosen in this study. Above ground biomass (AGB) variable was also selected in this study. Multiple linear regression approach was employed to develop the best prediction model for each variable.

MATERIALS AND METHODS

The study area

The study area is located at Port Dickson district in Negeri Sembilan, Malaysia. Sungai Menyala Forest Reserve is one of the few remaining patches of west coast lowland dipterocarp forest in Peninsular Malaysia. The area covers 1,305 hectares and is surrounded by oil palm plantations, farms and villages. The forest reserve is situated 5–6 km from the coastline with an average ground elevation of 20–40 m above sea level. Almost 15% of the total forest area is freshwater swamp forest. Sungai Menyala Forest Reserve is a virgin forests that has been designated as an eco-edutourism park in Peninsular Malaysia.

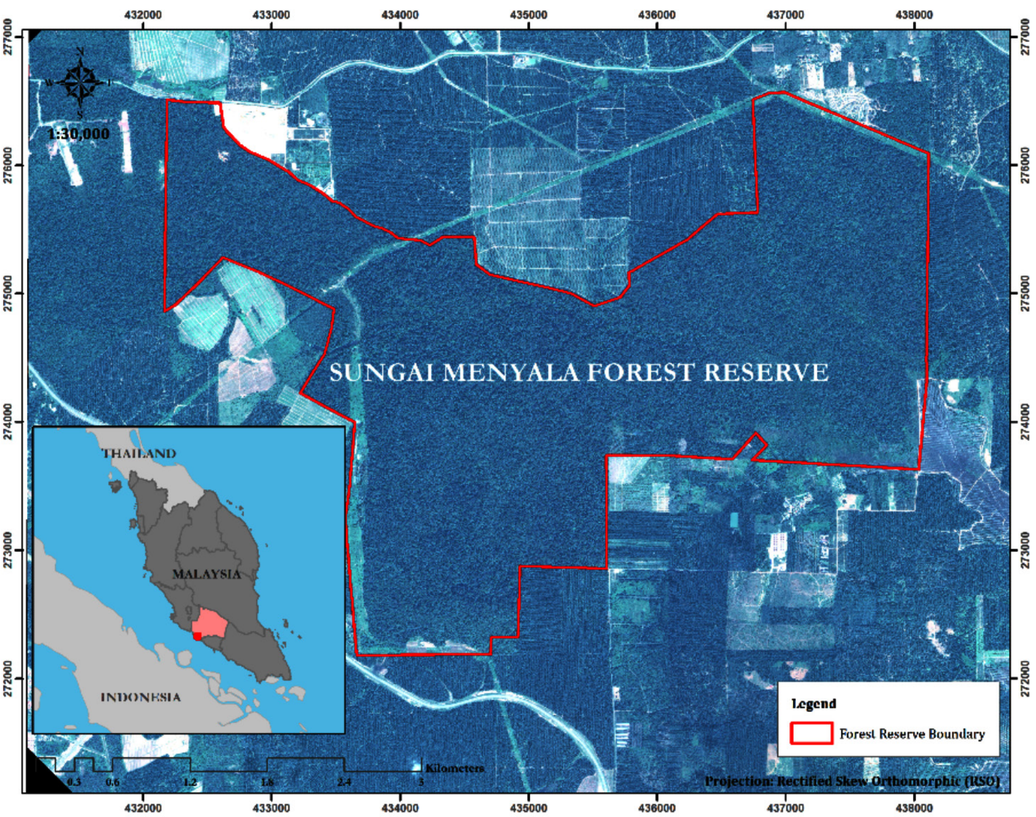

Figure 1 Location of study area overlaid with high resolution Pleiades image (band combination: 3, 2, 1)

Field plot data

A total of 40 forest inventory plots were established within the study area. These plots were used to develop and assess the models. Stratified random sampling technique was used to determine plot locations. The stratification is based on elevation, canopy height and tree cover information extracted from ALS data. Forest inventory activities were carried out from May to August 2016. The center of the circular plots (20 m radius) was located in the field using handheld GPS.

The sampling plot design was similar to Winrock International (Walker et al. 2012). However, some changes have been made to suit tropical forest conditions. The plot was designed in a circle with smaller nests inside. The nest sizes (A) were 20, 12 and 4 m as shown in Figure 2. Not all trees were measured within the nest plots. Each nest measured different sizes of trees, as summarised in Table 1. Diameter at breast height (DBH) of trees was measured using diameter tape. The total tree height and bole height were measured using a hypsometer. Other information such as tree species and position were also recorded during the inventory.

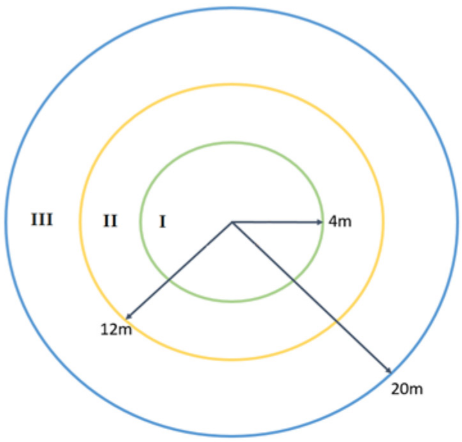

Figure 2 Circular sampling plot of 20 m radius with two nest sizes within the plot (12 and 4 m)

Table 1 Summary of living trees measurement in a plot

Nest radius (m)	Size	Tree size, DBH (cm)	Nest area, A (ha)
4	Small	10.0 – 19.9	0.00503
12	Medium	20.0 – 39.9	0.04524
20	Large	≥ 40.0	0.12566

Estimation of plot-based forest variables

Forest variables at plot level were calculated based on field inventory data. The Hm was calculated by the arithmetic mean of all tree heights in the plot. The BA of each plot was estimated based on total BA of all trees within the plot, whereby BA of each tree was calculated as follows:

$$BA_i = \pi * DBH_i^2 / 40000 \qquad (1)$$

where BA_i = basal area of an individual tree (m^2) and DBH_i = diameter at breast height of an individual tree (cm). The calculation of Dg of each plot was based on equation 2 as follows:

$$Dg = SQRT\ [SUM\ (DBH_i^2)/S] \qquad (2)$$

where Dg = square mean diameter (m^2), DBH_i = diameter of breast height of an individual tree and S = stand density. The S was estimated by taking the summation of total trees measured within the plot. The AGB of each tree was calculated based on allometric function established by Chave et al (2014), as shown in equation 3, before the calculation of total AGB of every plots.

$$AGB_i = \exp\ [-1.803 - 0.976E + 0.976\ln(p) + 2.763\ln(DBH_i) - 0.0299[\ln(DBH_i)]^2] \qquad (3)$$

where AGB_i = above ground biomass of individual tree (kg), DBH_i = diameter at breast height of an individual tree (cm), E = bioclimatic variable and p = wood specific gravity or wood density. The three forest variables, BA, S and AGB, estimated at plot level, was standardised to per-hectare-basis. The summary statistics for stand-structural and biophysical information of all plots are listed in Table 2.

LiDAR data

LiDAR data was acquired in April 2015 using RIEGL LMS-Q560 sensor. The average flying height during the acquisition was 450 m above mean sea level. The resulting ALS point density for all returns and last return were 12.39 and 7.31 point m^{-2} respectively. The final product was delivered in 1 x 1 km tiles containing x, y, z coordinates and classified to ground and non-ground points. The date was saved as LAS binary files format v 1.2 and projected as West Malaysia Rectified Skew Orthomorphic (Kertau RSO).

Table 2 Summary statistics for forest stand-structural and biophysical information of all plots

Forest variable	Min	Max	Range	Mean	SD
Hm (m)	5.3	24.8	19.23	18.0	3.8
BA (m^2 ha^{-1})	3.2	49.8	46.6	25.4	10.5
Dg (m^2)	12.2	48.6	36.4	30.3	8.4
AGB (mg ha^{-1})	35.6	762.4	726.8	319.6	157.2
S (stem ha^{-1})	30.1	1113.2	1083.0	396.0	199.0

Hm = mean height, BA = basal area, Dg = square mean diameter, AGB = above ground biomass, S = stand density

ALS metrics were derived by FUSION v 3.6 open source software using normalised ALS data (McGaughey 2009). Normalised ALS data are cloud points where each point's height is converted from mean sea level (msl) to ground reference. Thus, the height of each cloud point refers to the height from ground, measured by substracting the z coordinate f non-ground points with digital terrain model (DTM), produced using ground points data. The normalised ALS points were clipped based on field inventory area before derivation of ALS metrics of each plot, performed by FUSION's software. A total of 34 ALS metrics were produced and used for modelling analysis.

Establishment of model

The selection of ALS metrics was performed before the model was built. To reduce the number of ALS metrics, only metrics with significant linear correlation with field plot data were selected, since multiple linear regression method was adopted to built the model. Only metrics that gave strong linear correlation with field plot data were chosen for model development. For this purpose, the simple linear regression approach was adopted. Since the distribution of field plot data was normal, after normality test, Spearman's rank correlation coefficient (p) was applied to ALS metrics to obtain the correlation value of each metrics with field plot data. Only metrics with significant correlation ($p < 0.05$) were chosen and used for the development of the model.

The development and establishment of the model in estimating stand-level structure and biophysical variables was done using multivariate linear regression approach. In general, this approach predicts the value of the dependent variable (forest variables) based on multiple independent variables (ALS metrics). Forward stepwise regression was applied to the selected ALS metrics. A maximum of four independent variables and variance inflation factors (VIFs) were set and used to minimise the effects of over-fitting and multicollinearity in the final model (Lin et. al. 2011).

Model validation

The best model of each forest variables was validated using leave-one-out cross-validation (LOOCV) technique. This technique assesses the accuracy of the predictive model by comparing the values estimated by the model with field plot data. This technique drops one field data and uses another 39 field data to develop the model. The model then estimates the variable of the dropped field data. This process is repeated 40 times until each field plot data is dropped once. An adjusted-R^2, root mean square error (RMSE) and bias were calculated and compared. The R statistical environment was used for processing and analysing and VIF's package was used to check the collinearity in the model (R Core Team, 2017).

RESULTS

A summary of the structural and biophysical estimation using field data is shown in Table 2. The range values show that there were significant variations in all of the parameters assessed, especially AGB and S. Thus, a minimum of 40 field plots, where 75% were used for model development while the remaining 25% were used for validation, are adequate to develop the model. Table 3 shows the Spearman's rank correlation coefficients between plot-derived metrics (Hm, BA, Dg, AGB and S) and ALS-derived metrics, respectively. Correlation of Hm to ALS metrics

Table 3 Correlation coefficients (p) describing the strength of linear relationships between plot-derived Hm, BA, Dg, AGB and S with ALS metrics

ALS metrics / forest variable	Hm	G	Dg	B	S
Canopy height percentile metrics (m)					
Elevation P01	0.20[ns]	0.27[ns]	0.13[ns]	0.33*	0.06[ns]
Elevation P05	0.34*	0.55**	0.28[ns]	0.58**	0.20[ns]
Elevation P10	0.42**	0.73**	0.36*	0.77**	0.25[ns]
Elevation P20	0.60**	0.82**	0.53**	0.85**	0.17[ns]
Elevation P25	0.61**	0.83**	0.56**	0.87**	0.15[ns]
Elevation P30	0.61**	0.84**	0.56**	0.88**	0.17[ns]
Elevation P40	0.59**	0.85**	0.56**	0.90**	0.17[ns]
Elevation P50	0.59**	0.85**	0.57**	0.90**	0.15[ns]
Elevation P60	0.60**	0.86**	0.58**	0.91**	0.14[ns]
Elevation P70	0.60**	0.84**	0.57**	0.89**	0.13[ns]
Elevation P75	0.62**	0.82**	0.56**	0.88**	0.13[ns]
Elevation P80	0.60**	0.81**	0.54**	0.87**	0.14[ns]
Elevation P90	0.62**	0.77**	0.55**	0.83**	0.12[ns]
Elevation P95	0.61**	0.75**	0.55**	0.82**	0.09[ns]
Elevation P99	0.59**	0.70**	0.55**	0.77**	0.06[ns]
Canopy height metrics (m)					
Elevation minimum	0.00[ns]	0.00[ns]	0.00[ns]	0.00[ns]	0.00[ns]
Elevation maximum	0.50[ns]	0.59**	0.51**	0.67**	0.02[ns]
Elevation mean	0.60*	0.84**	0.56**	0.90**	0.14[ns]
Elevation mode	0.47[ns]	0.65**	0.50**	0.71**	0.06[ns]
Canopy height variability metrics (m)					
Elevation SD	0.59**	0.62**	0.48**	0.68**	0.02[ns]
Elevation variance	0.59**	0.62**	0.48**	0.68**	0.02[ns]
Elevation CV	-0.45**	-0.72**	-0.33*	-0.72**	-0.29[ns]
Elevation IQ	0.41**	0.41**	0.32*	0.48**	-0.03[ns]
Elevation skewness	-0.50**	-0.75**	-0.44**	-0.75**	-0.21[ns]
Elevation kurtosis	0.18**	0.37*	0.21[ns]	0.36*	0.09[ns]
Canopy density metrics (%)					
First returns above 1.00 (%)	0.32**	0.48**	0.29[ns]	0.51**	0.16[ns]
All returns above 1.00 (%)	0.25*	0.42**	0.23[ns]	0.47**	0.13[ns]
First returns above mean (%)	0.43**	0.68**	0.45**	0.68**	0.22[ns]
First returns above mode (%)	-0.33*	-0.46**	-0.38*	-0.49**	-0.08[ns]
All returns above mean (%)	0.42**	0.67**	0.44**	0.68**	0.18[ns]
all returns above mode (%)	-0.35*	-0.48**	-0.41**	-0.51**	-0.07[ns]
All returns above 1.00 / total first returns (%)	0.50**	0.44**	0.56**	0.50**	-0.12[ns]
All returns above mean / total first returns (%)	0.52**	0.59**	0.60**	0.62**	-0.02[ns]
All returns above mode / total first returns (%)	-0.25[ns]	-0.43**	-0.29[ns]	-0.46**	-0.11[ns]

Hm = mean height, BA = basal area, Dg = square mean diameter, AGB = above ground biomass, S = stand density; **p < 0.01, *p < 0.05, ns = not significant (p > 0.05); SD = standard deviation, CV = coefficient variance, IQ = elevation interquartile distance was moderate. Only five ALS derived metrics were not passed at 0.05. The P75 and P90 gave the highest correlation to Hm with a strength of 0.62.

Only two metrics of BA, P01 and elevation minimum did not pass the correlation test at 0.05 significance level. The G is strongly correlated with canopy height percentile metrics with an average correlation coefficient value of eight. Only P01 (0.27) gave weaker correlation with BA while P60 showed the strongest correlation of 0.86. The ALS metrics responded moderately on Dg where maximum correlation coefficient values achieved were canopy density metrics of all returns above mean over total first returns (0.60). Seven ALS metrics did not pass the test at 0.05 significance level.

The AGB had similar tendency as BA. Almost all ALS metrics gave same or higher correlation coefficient values than BA. Only elevation mean gave non significance correlation with AGB. The AGB was strongly correlated with canopy height percentile metrics especially P60 with a correlation of 0.91. Finally, S showed weak correlation with all ALS metrics and none of the ALS metrics showed significant correlation with S. Canopy height variability metrics of elevation covariance gave the highest correlation with S. The correlation value achieved was -0.29; negative symbol denotes an inverse relationship. An exception was made for S where all metrics were included in the model's development in order to produce the best model for S.

Overall, all ALS metrics showed some correlation with Hm, BA, Dg, AGB and S except for elevation minimum metric. The AGB and BA showed strong correlation with ALS derived metrices. Only S showed weak and non-significant correlation to all metrics as compared to other variables. Table 4 shows the summary of the best model of each forest variable and the validation results based on LOOCV method. The Hm variable was best estimated by P70, P75, P80 and percentage of all returns above 1 m over total first returns, and percentage of all returns above mean over total first returns. The adjusted-R^2 was 0.36 with RMSE values of 2.92 m. The BA was best estimated using P40, P95, elevation interquatile distance (IQ) and percentage of all returns above mean with an adjusted-R^2 of 0.77. The RMSE of BA model is 5.45 m^2 ha^{-1}. The best model for Dg included P90, P99, the percentage of all returns above mode and percentage of all returns above 1 m over total first returns, with an adjusted-R^2 of 0.44 and RMSE of 6.07 m^2.

The AGB was best estimated using two ALS metrics, P60 and percentage of all returns above 1 m over total first returns. This model accounted for 82% of variance within the data, with RMSE of 71.12 mg ha^{-1}. Finally, S was best estimated using metrics from canopy density, which are percentage of first returns above 1 m, percentage of all returns above mean and percentage of all returns above 1 m over total first returns. The adjusted-R^2 for this model was 0.27. Canopy

Table 4 Summary of selected models and validation results for estimated variables

Forest variable	Predictive model	Fitting phase			Cross-validation		
		R^2	RMSE	Bias	R^2	RMSEcv	Bias
Hm	$-2.599(P_{70}) + 5.080(P_{75}) - 2.260(P_{80})$ + 0.078 [all returns above 1 m / total first returns (%)] – 3.090	0.61	2.35	0.00	0.64	2.38	0.06
BA	$2.230(P_{40}) - 0.952(P_{95})$ + 1.177(eElevation IQ) – 0.135 [all returns above mean (%)] + 12.227	0.77	5.45	0.00	0.79	5.53	-0.04
Dg	$-1.188(P_{90}) + 1.562(P_{99})$ – 0.063 [all returns above mode (%)] + 0.189 [all returns above 1 m / total first returns (%)] – 18.952	0.44	6.07	0.00	0.55	6.16	-0.04
AGB	$20.247(P_{60})$ – 1.192 [all returns above 1 m / total first returns (%)] + 36.316	0.82	71.12	0.00	0.84	72.10	-0.20
S	59.274 [first returns above 1 m(%)] + 5.847 [all returns above mean (%)] – 5.829 [all returns above 1 m / total first returns (%)] – 4739.695	0.27	149.48	0.00	0.30	151.60	-0.35

Hm = mean height, BA = basal area, Dg = square mean diameter, AGB = above ground biomass, S = stand density

density metrics seem important since all the best models used one of the ALS metrics in this category, as shown in Table 4.

Overall, the adjusted-R^2 was slightly improved during cross-validation phase as compared to fitting phase, indicating that the final model of each forest variable is not overfitted and can be used to estimate forest variables. The Dg's model showed a huge difference of adjusted-R^2 between fitting and cross-validation phases. The $RMSE_{CV}$ improved with increasing adjusted-R^2. The errors were relatively low for Hm, AGB and Dg. Plots predicted against reference values for Hm, BA, Dg and AGB showed little apparent bias (Figure 3). Bias values for Hm, BA, Dg, AGB and S obtained during cross-validation phase were 0.06 m, -0.04 m^2 ha^{-1}, -0.04 m, -0.20 mg ha^{-1} and -0.35 stem ha^{-1} respectively.

The use of ALS data in estimating forest structures and biophysical parameter in tropical ecosystems would become more apparent in the near future as this technology provides insight into the forest condition. Structure of the forest can be seen visually and statistically, as shown in Figure 4. Disturbed forest would have different stand structure as compared to mature and old growth forest. The ALS metrics generated by FUSION's software provides detailed information on forest structure at plot level, allowing better prediction on Hm, BA, Dg, AGB and S variables.

The results obtained in this study have shown the potential use of RIEGL LMS-Q560's product in predicting several variables, especially BA and AGB. The AGB was predicted with a high degree of precision (adjusted-R^2 of 0.82). The results were comparable with other studies conducted in tropical forest, whose adjusted-R^2 typically ranged from 0.74 to 0.90 (Clark et al 2011, Hernandez-Stefanoni et al. 2014, Loki et al. 2014, Vega et al. 2015, Kim et al. 2016, Coomes et al. 2017, El Hajj et al. 2017). The AGB was expected to have higher precision of prediction since most of the previous studies yielded best results not only in tropical but also in the temperate forest (Clark et al 2011, Crespo-Peremarch et al. 2016). However, several studies yielded moderate precision, whose

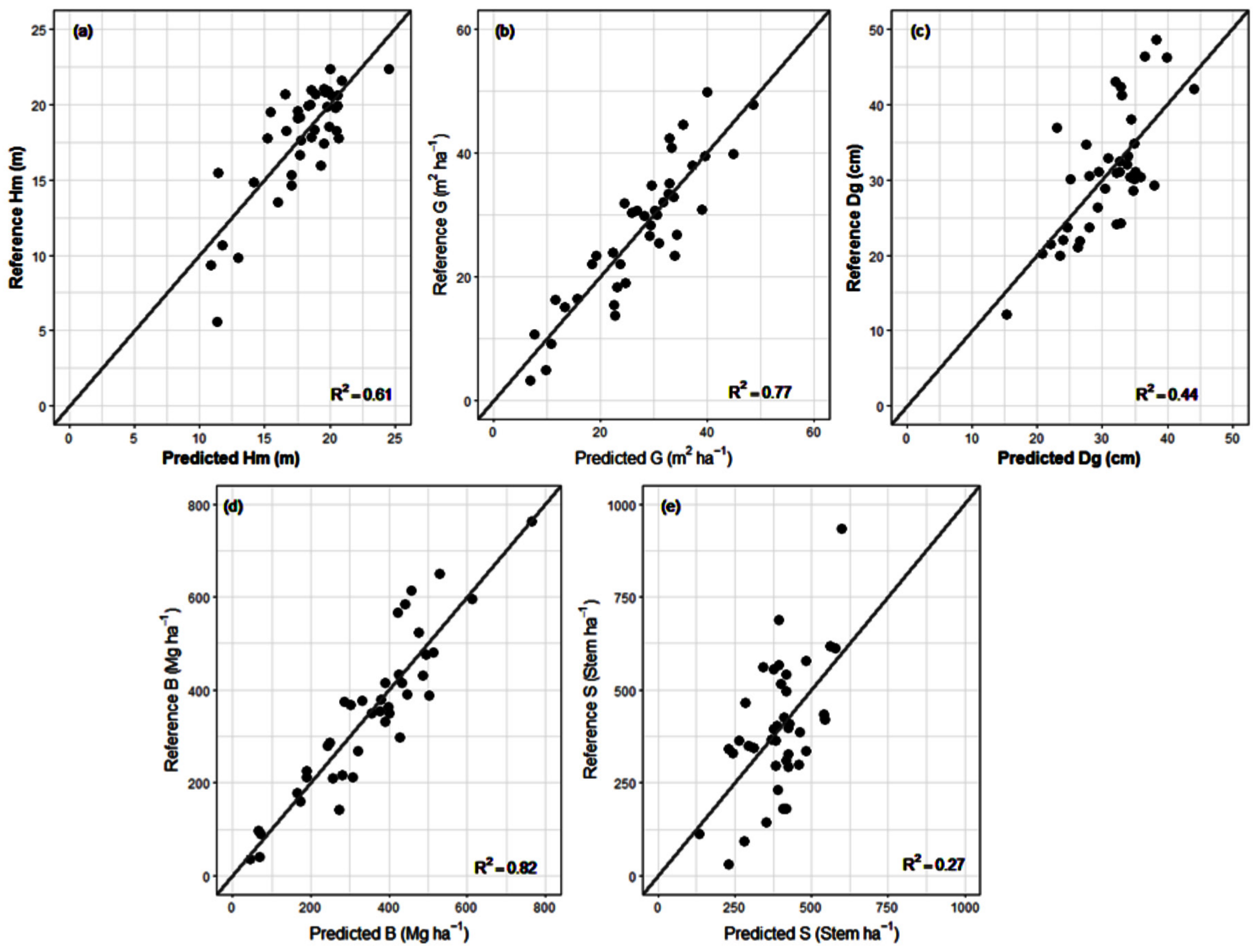

Figure 3 Relationship between predicted and reference (a) Hm, (b) BA, (c) Dg, (d) AGB and (e) S values

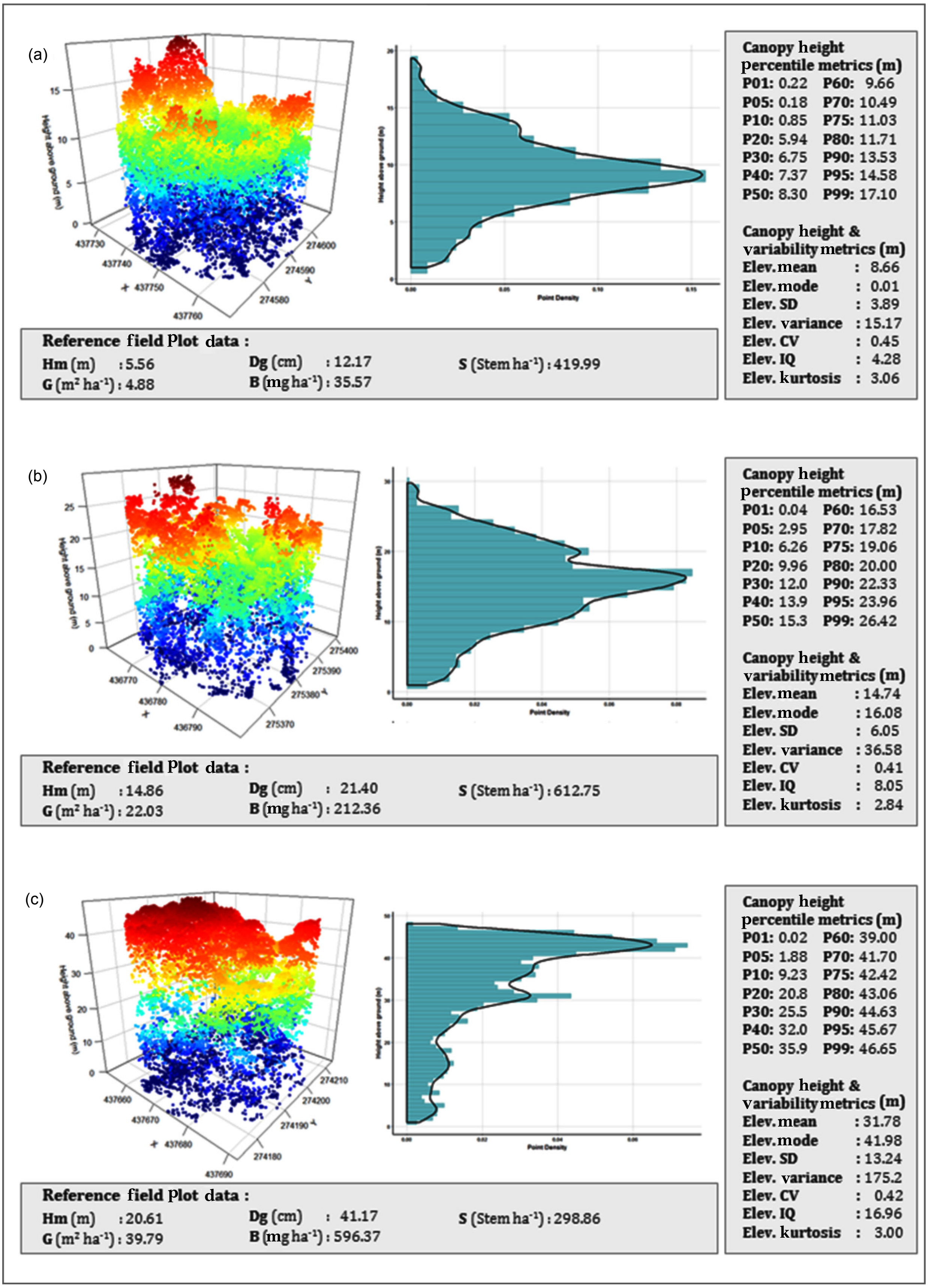

Figure 4 Metrics associated with the vertical distribution of ALS returns in three selected field plots (a) disturb forest, (b) mature forest and (c) old growth forest; Hm = mean height, BA = basal area, Dg = square mean diameter, AGB = above ground biomass, S = stand density; P = elevation height, Elev = elevation; SD = standard deviation, CV = coefficient variance, IQ = interquartile distance

adjusted-R^2 ranged from 0.46 to 0.63 (Cao et al. 2016, Ferraz et al. 2016). This is due to the different approaches and parameters used in their studies.

Model prediction for BA had good precision with adjusted-R^2 of 0.77, which was higher than the study by Palace et al. (2015). However, the result was comparable with their study using TLS data in predicting BA (adjusted-R^2 of 0.75). The result was also comparable to studies carried out in the temperate forest with adjusted-R^2 values ranging from 0.55 to 0.93 (Ruiz et al. 2014, Crespo-Peremarch et al. 2016, Ferraz et al. 2016). A good prediction of BA gives an overview of the standing trees in the area. Higher BA indicates that larger trees dominate the area. This provides valuable information for forest management in estimating timber volume for harvesting. Mature and old forest generally would have higher BA values compared to young forest.

The Dg was predicted with lower precision (adjusted-R^2 of 0.44) as compared to other studies carried out in pine forest (adjusted-R^2 of 0.85), Douglas-fir plantation (adjusted-R^2 of 0.86) and tropical forest (adjusted-R^2 of 0.51). However, the differences within a tropical forest is minimal. Large differences were observed between tropical forest with pine and plantation forest, due to different structure and composition of tropical forest as compared to pine and plantation forests (Watt et al. 2013, Palace et al. 2015, Montealegre et al. 2016).

As for Hm, the model had a moderate precision with an adjusted-R^2 of 0.61. The result is lower compared to other studies conducted in pine forest, whose adjusted-R^2 typically ranged from 0.75 to 0.96 (Stone et al. 2011, Gonzalez-Ferreiro et al. 2012, Treitz et al. 2012, Watt et al. 2013, Ediriweera et al. 2014, White et al. 2015). This may be due to variations of tree heights in dense tropical forest as opposed to pine forest that makes the prediction of Hm a lot harder. The accuracies in the estimation of total height during field inventory activities may also introduce error in this model due to difficulties in allocating top height of the trees.

The S was predicted with the lowest degree of precision as compared to other variables (adjusted-R^2 of 0.27). The model precision is lowered when compared to studies carried out by Palace et al (2015) (adjusted-R^2 of 0.43) and Palace et al (2016) (adjusted-R^2 of 0.50). The use of ALS metrics in predicting S may not be suitable. A single tree delineation approach may produce better results in predicting S. However, higher points density is also needed in order to capture understorey trees and produces better estimates.

DISCUSSION

The modified design of field sampling from Winrock International works well, especially for BA and AGB. The model for AGB was expected to perform well since the inventory technique was originally designed for estimating biomass. Sampling design may not be one of the factors affecting the precision of predicting Hm, Dg and S. Other studies conducted in the tropical forest also show similar accuracy although their approaches in field inventory differ from the approach taken in this study (Omar and Misman, 2018).

A time delay (i.e. ~1 year) between ALS data acquisition and field inventory did not affect the predictions of Hm, BA, Dg, AGB and S, as the study site did not differ considerably during the period. No logging activities were carried out in the study area as it was gazetted as an eco-edutourism park. Several aspects can be considered in order to improve the prediction of these variables. The use of higher point density ALS data may give a better prediction. Higher point density would give higher chances of ALS points to penetrate the canopy layer resulting in better information on the structure of the understorey layer.

Another aspect of improving the prediction of these variables is the processing approaches. Exploring other methods in predicting these variables would be of great interest in order to obtain the highest accuracy possible. Several studies have been carried out using other approaches in predicting Hm and S (Dandois and Ellis 2010, St-Onge et al. 2015, Wallace et al. 2016, Heinzel and Huber 2017). The use of other techniques such as non-linear regression, random forest, machine learning and others can be explored further in order to improve the prediction of forest variables. A similar study should also be conducted in other places such as hill, upper hill and montane dipterocarp forest with differences in structure and composition. The results may vary from the results obtained in this study.

Although some of the models developed showed better performance in estimating stand-level structural and biophysical variables of lowland dipterocarp forests, there are several limitations of using ALS data to estimate these variables. First, limited number of sampling points (40 field plots) were used to develop the model. Thus, the variations of forest parameters may be smaller than anticipated. The developed models usually perform poorly if the value that the model needs to predict is outside the range. Increasing the number of sampling points may give more variations of forest variables and improve the performances of the model.

Second, the developed model is limited to predicting lowland forests with similar conditions to test site, which are ALS data specification, forest types and logging history. Since ALS data was captured using a specific sensor and height during data acquisition, the model developed in this study can only be applied to ALS data from other sites with similar specifications. Furthermore, structure and composition of lowland forest differ from hill, upper hill and montane forests, which makes the model non-applicable for these types of forests. Besides, logging history also plays an important role as it affects the structure and composition of the forest. There are 14 stratas, comprising of various structure and composition, which can be divided, for inland forest of Peninsular Malaysia (FDPM 2014). The chosen study site only comprised one strata, which is strata 1 (virgin forest: lowland and hill forest). Thus, the model developed in this study may not give accurate results when estimating forest variables of other stratas. A robust model that caters to all these variations should be developed for better predictions of forest variables. However, it will require a substantial amount of time, manpower and money to achieve it.

CONCLUSION

The use of airborne LiDAR data has proven its potential in estimating stand-level structural and biophysical variables of lowland dipterocarp forests. The final model for BA and AGB performed well with adjusted-R^2 of 0.77 and 0.82 while models for Dg and Hm performed moderately with adjusted- R^2 of 0.44 and 0.61. However, ALS data did not perform well in estimating S with adjusted-R^2 of 0.27. Although each model performed differently, the results were comparable to previous studies, but low in comparison with temperate and plantation forests. Many factors contributed towards such differences such as forest types, LiDAR point intensity and approaches in model development. Further improvement can be made for better estimation of stand-level structural and biophysical variables of lowland dipterocarp forests by catering for the above factors. Besides field inventory method, ALS data can provide an alternative way in estimating forest structure and biophysical variables at larger coverage, improving the efficiency of monitoring and managing the forest.

ACKNOWLEDGEMENTS

The study was funded by the government of Malaysia under Eleventh Malaysia Plan (2016–2020). Special gratitude to the Forestry Department of Peninsular Malaysia and State Forestry Department of Negeri Sembilan for providing ancillary data and access to Sungai Menyala Forest Reserve for field data collection. Deepest thanks to anonymous reviewers who gave constructive comments.

REFERENCES

Bucha T, Juris M & Vladovic J. 2012. Derivation of 3-D forest canopy model and detection of tree crowns from high-resolution aerial photos. *Forest Science* 57: 225–236.

Cao L, Coops NC, Innes JL, Sheppard SRJ, Fu L, Ruan H & She G. 2016. Estimation of forest biomass dynamics in subtropical forest using multi-temporal airborne LiDAR data. *Remote Sensing of Environment.* 178: 158–171.

Chave J, Rejou-Mechain M, Burquez A, et al. 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology* 20: 3177–3190.

Clark ML, Roberts DA, Ewel JJ & Clark DB. 2011. Estimation of tropical rain forest aboveground biomass with small-footprint LiDAR and hyperspectral sensors. *Remote Sensing of Environment* 115: 2931–2942.

Coomes DA, Dalponte M, Jucker T, et al. 2017. Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asean Forests from airborne laser scanning data. *Remote Sensing of Environment* 194: 77–88.

Couteron P, Barbier N, Proisy C, Pelissier R & Vincent G. 2012. Linking remote-sensing information to tropical forest structure: The crucial role of modelling. *Earthzine* 4: 8.

Crespo-Peremarch P, Ruiz LA & Balaguer-Beser A. 2016. A comparative study of regression methods to predict forest structure and canopy fuel variables from LiDAR full-waveform data. *Revista de Teledeteccion*. 45: 27–40.

Dandois JP & Ellis EC. 2010. Remote sensing of vegetation structure using computer vision. *Remote Sensing*. 2: 1157–1176.

Ediriweera S, Pathirana S, Danaher T & Nichols D. 2014. LiDAR remote sensing of structural properties of sub tropical rainforest and eucalypt forest in complex terrain in North-Eastern Australia. *Journal of Tropical Forest Science* 26: 397–408.

El Hajj M, Baghdadi N, Fayad I, Vieilledent G, Bailly JS & Minh DHT. 2017. Interest of integrating spaceborne LiDAR data to improve the estimation of biomass in high biomass forested area. *Remote Sensing* 9: 213.

Feret JP & Asner GP. 2012. Semi-supervised methods to identify individual crowns of lowland tropical canopy species using imaging spectroscopy and LiDAR. *Remote Sensing* 4: 2457–2476.

Ferraz A, Saatchi S, Mallet C & Meyer V. 2016. LiDAR detection of individual tree size in tropical forests. *Remote Sensing of Environment*. 183: 318–333.

FDPM (Forest Department Peninsular Malaysia). 2014. *Laporan Inventori Hutan Nasional Kelima (IHN-5)*. Forestry Department Peninsular Malaysia, Kuala Lumpur.

Gonzalez-Ferreiro E, Dieguez-Aranda U & Miranda D. 2012. Estimation of stand variables in Pinus radiata D. Don plantations using different LiDAR pulse densities. *Forestry* 85: 281–292.

Hansen EH, Ene LT, Mauya EW, Patocka Z, Mikita T, Gobakken T & Naesset E. 2017. Comparing empirical and semi-empirical approaches to forest biomass modelling in different biomes using airborne laser scanner data. *Forests* 8: 170.

Heinzel J & Huber MO. 2017. Detecting tree stems from volumetric TLS data in forest environments with rich understory. *Remote Sensing* 9: 9.

Hernandez-Stefanoni JL, Dupuy JM, Johnson KD, et al. 2014. Improving species diversity and biomass estimates of tropical dry forests using airborne LiDAR. *Remote Sensing*. 6: 4741–4763.

Kent R, Lindsell JA, Laurin GV, Valentini R & Coomes DA. 2015. Airborne LiDAR detects selectively logged tropical forest event in an advanced stage of recovery. *Remote Sensing* 7: 8348–8367.

Kim E, Lee WK, Yoon M, Lee JY, Son Y & Salim KA. 2016. Estimation of voxel-based above-ground biomass using airborne LiDAR data in an intact tropical rain forest, Brunei. *Forests* 7: 259.

Leitold V, Keller M, Morton DC, Cook BD & Shimabukuro YE. 2015. Airborne lidar-based estimates of tropical forest structure in complex terrain: opportunities and trade-offs for REDD+. *Carbon Balance and Management* 10: 13.

Lin D, Foster DP & Ungar LH. 2011. VIF-Regression: a fast regression algorithm for large data. *Journal of the American Statistical Association* 106: 232–247.

Listopad CMCS, Drake JB, Masters RE & Weishampel JF. 2011. Portable and small airborne footprint LiDAR: forest canopy structure estimation of fire managed plots. *Remote Sensing* 3: 1284–1307.

Loki K, Tsuyuki S, Hirata Y, et al. 2014. Estimating above ground biomass of tropical rainforest of different degradation levels in Northern Borneo using airborne LiDAR. *Forest Ecology and Management*. 328: 335–341.

Ustin LS & Trabucco A. 2000. Using hyperspectral data to assess forest structure. *Journal of Forestry* 98: 47.

Matejka K. 2009. Assessment of tree layer biomass and structure using aerial photos in lake catchments of Sumava Mts. *Journal of Forest Science* 55: 63–74.

Mcgaughey R. 2009. *FUSION/LDV: Software for LiDAR Data Analysis and Visualization*. Forest Service, Pacific Nortwest Research Station, US Department of Agriculture, Oregon.

Mcintosh ACS, Gray AN & Garman SL. 2012. Estimating canopy cover from standard forest inventory measurements in Western Oregon. *Forest Science* 58: 154–167.

Montealegre AL, Lamelas MT, De La Riva J, Garcia-Martin A & Escribano F. 2016. Use of low point density ALS data to estimate stand-level structural variables in Mediterranean Aleppo pine forest. *Forestry* 89: 373–382.

Omar H & Misman MA. 2018. Time-series maps of aboveground biomass in dipterocarps forests of Malaysia from PALSAR and PALSAR-2 polimetric data. *Carbon Balance and Management* 13: 19.

Palace M, Sullivan FB, Ducey M & Herrick C. 2016. Estimating tropical forest structure using a terresterial lidar. *PloS ONE* 11: 1–19.

Palace MW, Sullivan FB, Ducey MJ, Treuhaft RN, Herrick C, Shimbo JZ & Mota-E-Silva J. 2015. Estimating forest structure in a tropical forest using field measurements, a synthetic model and discrete return lidar data. *Remote Sensing of Environment* 161: 1–11.

Phua MH, Hue SW, Ioki K, et al. 2016. Estimating logged-over lowland rainforest aboveground biomass in Sabah, Malaysia using airborne LiDAR data. *Terrestrial, Atmospheric and Oceanic Science* 27: 481–489.

R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.

Rahman MZA, Bakar MAA, Razak KA, et al. 2017. Non-destructive, laser-based individual tree aboveground biomass estimation in a tropical rainforest. *Forests* 8: 86.

Ruiz LA, Hermosilla T, Mauro F & Godino M. 2014. Analysis of the influence of plot size and LiDAR density on forest structure forest attribute estimates. *Forests* 5: 936–951.

Sadadi O, Hussin YA, Kloosterman H & Ismail MH. 2016. Tropical rain forest tree height measurements using ALS and TLS for estimating forest biomass and carbon stock in Ayer Hitam Forest, Malaysia. Paper presented at the *37th Asian Conference on Remote Sensing (ACRS 2016)*. Colombo.

Sato LM, Gomes VCF, Shimabukuro YE, et al. 2016. Post-fire changes in forest biomass retrieved by airborne LiDAR in Amazonia. *Remote Sensing* 8: 839.

Shen X, Cao L & She G. 2016. Subtropical biomass estimation based on hyperspectral and highresolution remotely sensed data. *Journal of Remote Sensing* 20: 1446–1460.

Stone C, Penman T & Turner R. 2011. Determining the optimum model for processing LiDAR data at the plot-level: results for a *Pinus radiata* plantation in

New South Wales, Australia. *New Zealand Journal of Forest Science* 41: 191–205.

St-Onge B, Audet FA & Begin 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: 3899–3922.

Thapa RB, Watanabe M, Motohka T, Shiraishi T & Shimada M. 2015. Calibration of abovegroundforest carbon stock models for major tropical forests in Central Sumatra using airborne LiDAR and field measurement data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8: 661–673.

Treitz P, Lim K, Woods M, Pitt D, Nesbit D & Etheridge D. 2012. LiDAR sampling density for forest resource inventories in Ontario. *Canadian Journal of Remote Sensing* 4: 830–848.

Vega C, Vepakomma U, Morel J, et al. 2015. Aboveground-biomass estimation of a complex tropical forest in India using lidar. *Remote Sensing* 7: 10607–10625.

Walker SM, Pearson TRH, Casarim FM, et al. 2012. *Standard Operating Procedures for Terrestrial Carbon Measurement: Version 2014.* Winrock International, Arkansas.

Wallace L, Lucieer A, Melenovsky Z Turner D & Vopenka P. 2016. Assessment of forest structure using two UAV techniques: a comparison of airborne laser scanning and structure from motion (SFM) point clouds. *Forests* 7: 62.

Wan-Mohd-Jaafar WS, Woodhouse IH, Silva CA, Omar H & Hudak AT. 2017. Modelling individual tree aboveground biomass using discrete return LiDAR in lowland dipterocarp forest of Malaysia. *Journal of Tropical Forest Science* 29: 465–484.

Watt MS, Meredith A, Watt P & Gunn A. 2013. Use of LiDAR to estimate stand characteristics for thining operations in young Douglas-firr plantations. *New Zealand Journal of Forest Science* 43: 18.

White JC, Stepper C, Tompalski P, Coops NC & Wulder MA. 2015. Comparing ALS and image-based point cloud metrics and modelled forest inventory attributes in a complex coastal forest environment. *Forests* 6: 3704–3732.

White JC, Wulder MA, Vastaranta M, Coops NC, Pitt D & Woods M. 2013. The utility of image-based point clouds for forest inventory: a comparison with airborne laser scanning. *Forests* 4: 518–536.