Tropical Peatland Vegetation Structure and Biomass: Optimal Exploitation of Airborne Laser Scanning

Chloe Brown 1,*, Doreen S. Boyd 1,†, Sofie Sjögersten 2, Daniel Clewley 3, Stephanie L. Evers 4,5,6 and Paul Aplin 7

1 School of Geography, University of Nottingham, Nottingham NG7 2RD, UK; doreen.boyd@nottingham.ac.uk
2 School of Biosciences, University of Nottingham, Nottingham NG7 2RD, UK; Sofie.Sjogersten@nottingham.ac.uk
3 Plymouth Marine Laboratory, Plymouth PL1 3DH, UK; dac@pml.ac.uk
4 School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool L3 3AF, UK; S.L.Evers@ljmu.ac.uk
5 School of Environmental and Geographical Sciences, University of Nottingham Malaysia Campus, Semenyih 43500, Selangor, Malaysia
6 Tropical Catchment Research Initiative (TROCARI), Semenyih 43500, Selangor, Malaysia
7 School of Geography, Edge Hill University, Ormskirk L39 4QP, UK; paul.aplin@edgehill.ac.uk
* Correspondence: chloe.brown@nottingham.ac.uk; Tel.: +44-115-846-7052

Received: 19 March 2018; Accepted: 24 April 2018; Published: 25 April 2018

Abstract: Accurate estimation of above ground biomass (AGB) is required to better understand the variability and dynamics of tropical peat swamp forest (PSF) ecosystem function and resilience to disturbance events. The objective of this work is to examine the relationship between tropical PSF AGB and small-footprint airborne Light Detection and Ranging (LiDAR) discrete return (DR) and full waveform (FW) derived metrics, with a view to establishing the optimal use of this technology in this environment. The study was undertaken in North Selangor peat swamp forest (NSPSF) reserve, Peninsular Malaysia. Plot-based multiple regression analysis was performed to establish the strongest predictive models of PSF AGB using DR metrics (only), FW metrics (only), and a combination of DR and FW metrics. Overall, the results demonstrate that a Combination-model, coupling the benefits derived from both DR and FW metrics, had the best performance in modelling AGB for tropical PSF ($R^2 = 0.77$, RMSE = 36.4, rRMSE = 10.8%); however, no statistical difference was found between the rRMSE of this model and the best models using only DR and FW metrics. We conclude that the optimal approach to using airborne LiDAR for the estimation of PSF AGB is to use LiDAR metrics that relate to the description of the mid-canopy. This should inform the use of remote sensing in this ecosystem and how innovation in LiDAR-based technology could be usefully deployed.

Keywords: tropical peat swamp; LiDAR; discrete return LiDAR; full waveform LiDAR; above ground biomass

1. Introduction

Tropical Peat Swamp Forests (PSF) are unique ecosystems; their health and functionality are dependent on the interactions between forest, peat and hydrology. Any unbalance in this sensitive environment from change in one can amplify across the ecosystem and impact the multitude of services they provide, creating positive feedback loops [1,2]. Tropical PSF are thus particularly vulnerable to human disturbances (as well as natural change). This is especially the case in South East Asia where there has been a long history of tropical PSF degradation [3,4]. At an estimated 88.6 Gt, tropical
PSF are one of the largest pools of terrestrial carbon [5,6], however, high volumes of deforestation and land conversion rates have transformed South East Asian PSF into one of the largest emitters of carbon globally [7–9]. Forest clearance and conversion not only decreases carbon stores directly due to reduction in above ground biomass (AGB), but also indirectly increases carbon emissions as a result of peat exposure [10,11]. Fire susceptibility in the landscape also increases [12–14], further compounding carbon losses. This, coupled with the need to understand the variability and dynamics of tropical ecosystems in general with respect to their role in the global carbon cycle [5,15], means there is an urgent need to track change in PSF via estimation of AGB over space and time. Not only will this afford understanding of carbon losses, this will also help guide tropical PSF restoration and sustainable management activities [16,17].

The structurally complex nature of tropical PSF, supporting high (50–70 m) to low (15–25 m) dense canopy forest vegetation, typically with buttresses and thick root mat floors [18], presents a challenging landscape for monitoring purposes. This is further complicated by degraded (and/or recovering) PSF generating even greater variation in forest structure, in response to disturbance events and changes in peat characteristics. Traditionally, forest AGB estimates have been based on field plot inventories, where plot measurements are used to infer AGB estimates through allometric equations [19]. This manual approach provides the most accurate AGB estimates [20], however it is labour intensive, time consuming and expensive [21]. Additionally, the extent and inaccessibility of PSF, coupled with the difficult logistics of field work posed by often waterlogged, swampy conditions, has restricted forest inventory data collection.

Development in remote sensing technology over the last two decades, specifically in Light Detection and Ranging (LiDAR) [22,23], has revolutionised the retrieval and accuracy of forest structural attributes, including AGB, across landscapes scales [24]. LiDAR data has been widely adopted in forest structural applications, both at the stand and individual tree level [25], due to its ability to characterise both vertical and horizontal structures. A common theme for many studies is the development of LiDAR-derived metrics related to height distribution, density, and intensity of returns [26–28]. Typically, in LiDAR AGB studies, models are developed from the relationships between field plot estimated AGB, calculated from tree allometry, and LiDAR-derived metrics [29–31]. Establishing these relationships between metrics and forest attributes are complex, in particular for dense, multi-layered forest canopies associated with the tropics. The majority of work conducted in this area has employed discrete return (DR) derived metrics [32], with focus on the utility of DR data for carbon accounting and environmental modelling [33,34]. Asner et al. [35] examined the effectiveness of DR LiDAR derived mean canopy height to estimate AGB in four tropical regions (Panama, Peru, Madagascar, and Hawaii). The results highlighted the value of DR LiDAR in generating high-resolution carbon maps across the tropics. However, the application of DR LiDAR is restricted by the number of returns which can penetrate through dense canopies down to the forest floor, and by the limited point measurements from the canopy and the understory vertical structure [36,37].

Full Waveform (FW) LiDAR data has the potential to provide a more detailed and comprehensive 3D view of forest canopy structure in comparison to DR sensors [38,39]. Consequently, a growing body of research has focused on developing FW metrics for forest assessment. Drake et al. [40,41] are early examples of the development of FW metrics, using large-footprint FW data (laser vegetation imaging sensor) to generate a list of metrics for AGB estimations. Height of medium energy (HOME) was demonstrated to correlate strongly with mean stem diameter, basal area and AGB. For small-footprint FW LiDAR, Wu et al. [38] explored the potential of a list of structural and statistics-based AGB predictor metrics. The results produced a high correlation between FW-derived metrics with individual tree height, foliar and biomass for a savanna environment. Pirotti et al. [42] also explored the potential for estimating AGB from small-footprint FW data, finding the median height of echo count (HOMTC) performed best at estimating AGB at plot level for a complex tropical forest.

Recently, several studies have evaluated the fusion of DR and FW metrics in forest AGB estimation [43–46]. Cao et al. [47,48] assessed the potential of DR and FW metrics, individually and in
combination, as predictors of total living biomass and its components (AGB, root, foliage, branch and trunk) for sub-tropical forests. The study found metrics related to canopy height to be the strongest predictors for AGB; additionally, the combination of DR and FW metrics, by multivariate linear regression, improved AGB estimation.

The benefits of LiDAR data for monitoring tropical environments has been well established, however with the structural complexity and unique interactions between forest and peat swamp, general tropical forest models may not be transferable. Limited research has been published on the application of LiDAR in tropical PSF [49,50]. Jubanski et al. [51] and Kronseder et al. [52] examined the use of LiDAR point cloud data in AGB estimation for varied levels of degraded lowland dipterocarp PSF in Central Kalimantan, Indonesia. Both studies concluded that AGB could be estimated successfully, though with varying levels of accuracy.

According to our knowledge, no previous study has examined the use of both DR and FW metrics or the fusion of DR and FW metrics in AGB estimation for tropical PSF. This paper will examine the relationship between the AGB of a tropical PSF in North Selangor, Peninsular Malaysia, and small-footprint airborne LiDAR (DR and FW) derived metrics. This forest is of particular importance since it is one of the few large areas of PSF remaining in peninsular Malaysia [53]. Three LiDAR-based approaches are investigated for modelling AGB: (1) DR-derived metrics (only); (2) FW-derived metrics (only); (3) and a combination of DR- and FW-derived metrics. The accuracy of the models is evaluated by comparing LiDAR-based AGB estimates with field (in situ plots) measured AGB, with a view to establishing the optimal model for estimating AGB of the North Selangor PSF.

2. Materials and Methods

2.1. Study Site

The North Selangor Peat Swamp Forest (NSPSF) is located on the flat coastal plains of northwest Selangor, Peninsular Malaysia (Figure 1). With a combined total of 81,304 hectares, NSPSF is one of the largest remaining PSF in the region; however, over 30 years of intensive timber harvesting has left the once pristine NSPSF degraded and vulnerable to further environmental threat [18,54]. Repeated cycles of logging activities have created a complex, multi-layered canopy structure, indicated by mixed stand density and height (Figure 1). Additionally, the extraction of timber from the forest has led to the formation of a 500 km network of canals [55], scarring the forest and accelerating the drainage and erosion of valuable peat soil stores (and the carbon they hold) [53].
Figure 1. (a) The location of the North Selangor Peat Swamp Forest in Peninsular Malaysia; (b) Magnified study area map at bottom left corner shows forest type (forest height (high, medium, low) and forest density (high, medium, low)) results from a forest management inventory undertaken in 2000 [56].

In response to the damaging impact of past logging activities on tropical PSF ecosystem health and the services they provide, the Selangor State Government agreed on a 25-year moratorium on timber harvesting for the State of Selangor, with all logging activities ceased in 2007 [55]. Despite this halt in logging activities, the reserve remains exposed to disturbance from the surrounding land uses, in particular accelerated drainage from Tanjong Karang Irrigation Scheme and land conversion to oil palm agriculture [57].

2.2. Field Data

To obtain field estimates of AGB, thirty circular field plots with a radius of 15 m were established within the reserve during 2017. This relatively small sample size, albeit statistically relevant, was determined by the challenging logistics of tropical PSF fieldwork, as previously mentioned. The forests plots were established to be representative of a secondary PSF ecosystem, where no obvious structural disturbances were observed since the LiDAR survey (2014)—for example major tree fall, logging activities or fire events—based on an existing forestry management plan [55], multispectral imagery and expert knowledge and experience from NSPSF Rangers. The location of the central tree, and thus the centre point of the plot, was recorded by a field GPS (Garmin GPSMAP 64), with an average horizontal error of <5 m when measurements were averaged over ten minutes. The positions of all further trees within the plot were noted relative to the centre (distance and angle) (Figure 2). For each plot, all live trees with a diameter at breast height (DBH) of >5 cm were measured and tree species recorded. Additionally, dependent on the top of tree visibility, selected trees of varying height classes were measured by a laser height meter.
An allometric equation developed by Chave et al. [19] was used to calculate AGB from field measured DBH, species (to estimate wood density), and additionally tree height when available (missing height was inferred from pantropical model). The AGB allometric equation used was:

\[
AGB = 0.0673 \times (\rho D^2 H)^{0.976}
\]

where \( \rho \) = wood density (g/cm\(^3\)), \( D \) = DBH (cm) and \( H \) = height (m).

The ‘BIOMASS’ package in R software (version 2.10.0) [58] was used to run AGB estimation. AGB values were calculated for individual trees within each 15 m radius field plot, and then summed to obtain a total AGB value for each plot.

2.3. LiDAR Data Acquisition

Airborne LiDAR data was flown by the UK’s Natural Environment Research Council (NERC) Airborne Research Facility (ARF) as part of its 2014 Malaysia Campaign. It is important to note the three-year time delay between the airborne LiDAR data collection (2014) and field plot inventory collection (2017). This may introduce error into the AGB models, however as previously mentioned, all field plots were established to be free from logging and fire, with no obvious disturbance within the three-year period. The data was captured in November 2014 from a Leica ALS50-II LiDAR system, a small footprint LiDAR sensor with the capacity to record both DR and FW LiDAR simultaneously. The Leica ALS50-II LiDAR system was on board a Dornier 228-201, flying at 132–136 knots at an altitude of 2145–2181 m. The sensor emits the laser beam at 1064 nm, additionally providing an intensity digital image from the return signal intensity data. The accuracy report that accompanied the LiDAR data estimated the vertical accuracy to be 0.023 m RMSE, where LiDAR elevation was assessed against ground control points. In total 29 flightlines were flown, surveying an area covering 690 km\(^2\) of the reserve and surrounding area. However, given the constraints of fieldwork and to ensure representation, a subset of the LiDAR flightlines (161.73 km\(^2\)), focused on the north-west of the NSPSF reserve (Figure 3), was extracted for optimal LiDAR-derived AGB model development and assessment. The dataset has an average point density of 2.03 m\(^2\) and a point spacing of 0.70 m. The LiDAR data, DR and FW, was pre-processed by NERC’s Data Analysis Node and delivered in LAS 1.3 format.
2.4. DR-Derived Metrics

A number of descriptive structural plot-level metrics were calculated from a DR height normalised point cloud. All data processing was carried out in LAStools software—a complimentary full academic LAStools licence was generously supplied through the LASmoons programme. Outliers, or noisy data, were first eliminated from the DR LAS files supplied, and the data was then filtered to identify laser ground and non-ground (unclassified) hits. The normalised above ground heights were calculated by subtracting the ground points, interpolated into a surface using the adaptive triangulated irregular network (TIN) filtering algorithm, from the non-ground (unclassified) points. Above ground height DR metrics were then calculated. DR plot metrics included: minimum height (min), maximum height (max), average height (avg), standard deviation (std), skewness (skw), kurtosis (kur), average square height (qav), height percentile (p01, p05, p10, p25, p50, p75, p90, P9, P99), height bincentiles (b10, b20, b30, b40, b50, b60, b70, b80, b90).

2.5. FW-Derived Metrics

Informed from the success in previous studies-based in the sub-tropics, three plot-level FW metrics were derived from the recorded waveform. All data processing was carried out in the Sorted Pulse Software Library (SPDlib), benefiting from the recent additions made to the metric library. Gaussian decomposition was applied to each waveform to derive point information (x, y, z). Points were classified into ground or non-ground using a combination of progressive morphology filtering and maximum curvature classification (MCC) algorithms to produce a Digital Terrain Model (DTM), to which heights were normalised. The FW plot metrics included: height of median energy (HOME), waveform distance (WD) and roughness of the outermost canopy (ROUGH). FW metrics were summarized as the mean and standard deviation for each plot.
2.6. Statistical Modelling and Analyses

Simple linear regression analysis on field measured plot AGB values against LiDAR derived DR and FW metrics was conducted. Multiple linear regression models were developed independently from the DR metrics, FW metrics and DR and FW metrics in combination for estimating AGB. To ensure model variables were not highly correlated, a ‘data mining’ approach, outlined by Langton et al. [64], was first adopted to identify the key predictor variables for multiple regression analyses. This stage was required to avoid model over-fitting given the relatively small sample size ($n = 30$) and potential high collinearity of a number of LiDAR metrics with one another. The ‘MuMIn’ package for R software (version 1.9.5) [65] was used to run a modified Akaike Information Criterion (AIC), termed AICc [66], this comparing the quality of predictive multiple regression model inputs. An automated stepwise AICc was performed to select input variables for the final models. The delta-AIC value was assessed for each candidate model, a threshold value for delta-AIC was set to $<2$, models which met this requirement were determined to be within range of plausible models that best fit the observed values [66]. The top ranked AICc regression models were assessed based on a number of diagnostics tests (ANOVA, students $t$-test and variance inflation factor ($\leq 1$)) to produce final predictive models for DR metrics (only), FW metrics (only), and a combination of DR and FW metrics. The final predictive modes—DR-model, FW-model, and Combination-model—were fitted and assessed based on a combination of: R-square ($R^2$), adjusted R-square (Adj-$R^2$), Root Mean Square Error (RMSE), and the relative Root Mean Square Error (rRMSE).

To judge the reliability of the models, leave-one-out-cross validation (LOOCV) was performed. The LOOCV method for model evaluation was selected due to its proven effectiveness for models with a small sample size [67]. The Root Mean Square Error from the cross-validation analysis (RMSEcv) and bias (BIAScv) was calculated and presented as indicators of the three models’ predictive power. Finally, to investigate the significance of the addition of FW metrics in PSF AGB estimation, the Kruskal-Wallis H test (or one-way ANOVA on ranks) was used to determine if there were statistically significant differences between the three final models’ AGB plot predictions.

3. Results

Plot-based AGB values were calculated from field measured DBH, species and height data (when available). For the 30 plots, AGB ranged from 126.96 Mg/ha to 443.27 Mg/ha, with a mean value of 319.52 Mg/ha, with a standard deviation of 76.34 Mg/ha (Table 1).
Table 1. Plot measured forest parameters.

<table>
<thead>
<tr>
<th>Plot Measurements</th>
<th>Values: Mean (Std Dev)</th>
<th>Min/Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree number</td>
<td>89.33 (21.09)</td>
<td>39/128</td>
</tr>
<tr>
<td>DBH (cm)</td>
<td>46.39 (37.16)</td>
<td>17/375</td>
</tr>
<tr>
<td>AGB (Mg/ha)</td>
<td>319.52 (76.34)</td>
<td>126.96/443.27</td>
</tr>
</tbody>
</table>

The results from simple linear regression analysis between field-based AGB estimates and 31 LiDAR derived DR and FW metrics are presented in Tables 2 and 3. For DR metrics, the 50th height percentile (‘p50’) showed the highest correlation with field-based AGB ($R^2 = 0.68$, RMSE = 42.1), followed by average height (‘avg’) ($R^2 = 0.64$, RMSE = 45.2). Both minimum height (‘min’) and standard deviation (‘std’) displayed no significant relationship or sensitivity to field-based AGB. For FW metrics, the mean HOME (‘H_mean’) provided the highest correlation with field-based AGB ($R^2 = 0.72$, RMSE = 39.4) (Figure 5). The mean input variables calculated for all FW metrics displayed a moderate correlation and significant relationship with field-based AGB, however, the standard deviation input variables calculated for all FW metrics showed very low correlation, with standard deviation HOME (‘H_sd’) and standard deviation ROUGH (‘R_sd’) presenting no significant relationship to field-based AGB estimates. It is evident from the linear regression results that metrics representing mid-canopy levels, for example ‘p50’ and ‘H_mean’, are important predictors for AGB estimation in PSF. The variation in mid-canopy structure captured by these metrics can be seen in Figure 6. Where a waveform has been extracted from the centre point (1 m buffer) for each of the 30 field plots.

![Figure 5](image-url)  
Figure 5. Field measured AGB modelled by the mean value from the FW mean HOME metric (H_mean). Bands represent the 95% confidence envelope.
Figure 6. Cont.
Figure 6. Centre point (1 m buffer) plot vegetation profiles, each waveform height has been subtracted from LiDAR derived DTM; graphs are fitted Height (m)/Amplitude (DN). Height scale (y-axis) ranges from 0 to 40 m, and the amplitude (x-axis) from 0 to 120 DN. Graphs are accompanied with forest canopy structure description.
Table 2. Results from simple linear regression using DR metrics. Values underlined represent metrics of non-significance \( p > 0.05 \).

<table>
<thead>
<tr>
<th>DR Metric</th>
<th>( R^2 )</th>
<th>Adj-( R^2 )</th>
<th>RMSE (Mg/ha)</th>
<th>rRMSE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>0.04</td>
<td>0.01</td>
<td>73.4</td>
<td>23</td>
</tr>
<tr>
<td>Max</td>
<td>0.21</td>
<td>0.18</td>
<td>66.6</td>
<td>20.8</td>
</tr>
<tr>
<td>Avg</td>
<td>0.64</td>
<td>0.62</td>
<td>45.2</td>
<td>14.2</td>
</tr>
<tr>
<td>Qav</td>
<td>0.59</td>
<td>0.58</td>
<td>48</td>
<td>15</td>
</tr>
<tr>
<td>Std</td>
<td>0.1</td>
<td>0.07</td>
<td>71.2</td>
<td>22.3</td>
</tr>
<tr>
<td>Ske</td>
<td>0.61</td>
<td>0.6</td>
<td>46.7</td>
<td>14.6</td>
</tr>
<tr>
<td>Kur</td>
<td>0.23</td>
<td>0.21</td>
<td>65.7</td>
<td>20.6</td>
</tr>
<tr>
<td>P01</td>
<td>0.29</td>
<td>0.26</td>
<td>63.4</td>
<td>19.8</td>
</tr>
<tr>
<td>P05</td>
<td>0.29</td>
<td>0.26</td>
<td>63.4</td>
<td>19.8</td>
</tr>
<tr>
<td>P10</td>
<td>0.30</td>
<td>0.28</td>
<td>62.7</td>
<td>19.6</td>
</tr>
<tr>
<td>P25</td>
<td>0.58</td>
<td>0.57</td>
<td>48.5</td>
<td>15.2</td>
</tr>
<tr>
<td>P50</td>
<td>0.68</td>
<td>0.67</td>
<td>42.1</td>
<td>13.2</td>
</tr>
<tr>
<td>P75</td>
<td>0.58</td>
<td>0.56</td>
<td>48.9</td>
<td>15.3</td>
</tr>
<tr>
<td>P90</td>
<td>0.39</td>
<td>0.36</td>
<td>58.8</td>
<td>18.4</td>
</tr>
<tr>
<td>P95</td>
<td>0.30</td>
<td>0.28</td>
<td>62.7</td>
<td>19.6</td>
</tr>
<tr>
<td>P99</td>
<td>0.25</td>
<td>0.22</td>
<td>65</td>
<td>20.3</td>
</tr>
<tr>
<td>B10</td>
<td>0.25</td>
<td>0.22</td>
<td>65</td>
<td>20.3</td>
</tr>
<tr>
<td>B20</td>
<td>0.35</td>
<td>0.33</td>
<td>60.6</td>
<td>19</td>
</tr>
<tr>
<td>B30</td>
<td>0.49</td>
<td>0.48</td>
<td>53.4</td>
<td>16.7</td>
</tr>
<tr>
<td>B40</td>
<td>0.59</td>
<td>0.57</td>
<td>48.1</td>
<td>15.1</td>
</tr>
<tr>
<td>B50</td>
<td>0.62</td>
<td>0.61</td>
<td>46.3</td>
<td>14.5</td>
</tr>
<tr>
<td>B60</td>
<td>0.52</td>
<td>0.51</td>
<td>51.8</td>
<td>16.2</td>
</tr>
<tr>
<td>B70</td>
<td>0.39</td>
<td>0.37</td>
<td>58.6</td>
<td>18.4</td>
</tr>
<tr>
<td>B80</td>
<td>0.25</td>
<td>0.23</td>
<td>64.9</td>
<td>20.3</td>
</tr>
<tr>
<td>B90</td>
<td>0.16</td>
<td>0.13</td>
<td>68.9</td>
<td>21.5</td>
</tr>
</tbody>
</table>

Table 3. Results from simple linear regression using FW metrics. Values underlined represent metrics of non-significance \( p > 0.05 \).

<table>
<thead>
<tr>
<th>FW Metric</th>
<th>( R^2 )</th>
<th>Adj-( R^2 )</th>
<th>RMSE (Mg/ha)</th>
<th>rRMSE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H_mean</td>
<td>0.72</td>
<td>0.71</td>
<td>39.4</td>
<td>12.3</td>
</tr>
<tr>
<td>H_sd</td>
<td>0.00</td>
<td>-0.03</td>
<td>75</td>
<td>23.5</td>
</tr>
<tr>
<td>W_mean</td>
<td>0.57</td>
<td>0.55</td>
<td>49.3</td>
<td>15.4</td>
</tr>
<tr>
<td>W_sd</td>
<td>0.17</td>
<td>0.14</td>
<td>68.5</td>
<td>21.4</td>
</tr>
<tr>
<td>R_mean</td>
<td>0.61</td>
<td>0.59</td>
<td>47.1</td>
<td>14.7</td>
</tr>
<tr>
<td>R_sd</td>
<td>0.00</td>
<td>-0.04</td>
<td>75.1</td>
<td>23.5</td>
</tr>
</tbody>
</table>

The results of the best models for predicting PSF AGB (DR, FW and Combination) (Table 4) are summarised in Table 5 and visualized in Figure 7. The Combination-model, \( H_{mean}/B50 \), displayed the best agreement with field-based AGB values (\( R^2 = 0.77 \); RMSE = 36.4; rRMSE = 10.8%). Figure 8 displays an example of AGB (Mg/ha) estimates calculated from the Combination-model for a subset of the LiDAR flightlines in the northwest of the NSPSF reserve at a 5 m (0.0025 ha) spatial resolution.

Table 4. Model equations.

<table>
<thead>
<tr>
<th>Model Equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR ( =81.786 + 8.940 \times P75 - 78.063 \times Ske )</td>
</tr>
<tr>
<td>FW ( =101.457 + 19.579 \times H_{mean} - 12.555 \times W_{sd} )</td>
</tr>
<tr>
<td>Combination ( =171.3731 + 14.7802 \times H_{mean} - 1.5752 \times B50 )</td>
</tr>
</tbody>
</table>
Table 5. Results from the best preforming models for predicting AGB.

<table>
<thead>
<tr>
<th>Model</th>
<th>Variables</th>
<th>R²</th>
<th>Adj-R²</th>
<th>RMSE (Mg/ha)</th>
<th>RMSEcv (Mg/ha)</th>
<th>BIAScv (Mg/ha)</th>
<th>rRMSE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DR</td>
<td>P75 Ske</td>
<td>0.73</td>
<td>0.71</td>
<td>39.2</td>
<td>42.9</td>
<td>−0.83</td>
<td>12.3</td>
</tr>
<tr>
<td>FW</td>
<td>H_mean W_sd</td>
<td>0.76</td>
<td>0.74</td>
<td>36.9</td>
<td>40.9</td>
<td>−0.17</td>
<td>11.6</td>
</tr>
<tr>
<td>Combination</td>
<td>H_mean B50</td>
<td>0.77</td>
<td>0.75</td>
<td>36.4</td>
<td>39.8</td>
<td>0.38</td>
<td>10.8</td>
</tr>
</tbody>
</table>

Figure 7. Predicted vs field measured AGB from the best performing models: (a) DR-model; (b) FW-model; (c) Combination-model. Bands represent the 95% confidence envelope.
However, all three predictive models displayed a strong positive correlation and a significant relationship to field-based AGB estimates (DR-model—$R^2 = 0.73$; RMSE = 39.2; rRMSE = 12.3%) and (FW-model—$R^2 = 0.76$; RMSE = 36.9; rRMSE = 11.6%). Figure 9 displays the frequency distribution for model generated AGB (Mg/ha) values for the 30 plots, presented with model AGB (Mg/ha) summary statistics. The Kruskal-Wallis test showed that there was no statistically significant difference between the three models for PSF AGB prediction ($H(2) = 0.0038095; p = 0.99$). In summary, the synergistic use of DR- and FW-derived metrics displayed a stronger correlation with field-based AGB estimates in comparison to individual and multiple DR- and FW-models, however, there was no significant difference in tropical PSF AGB prediction between the three best multiple regression models (DR, FW and Combination).
Quantification of tropical PSF AGB cover is required to better understand and safeguard tropical peatlands’ ecosystem function and resilience to disturbance, both natural and anthropogenic [68]. AGB products provide accurate baseline estimates to support international climate change policies and programs such as the UN REDD+ (United Nations Reducing Emissions from Deforestation and Forest Degradation), as well as the development of sustainable management strategies for local stakeholders. Additionally, knowledge gained from the precise assessment of PSF AGB distribution will assist in Malaysia’s work towards meeting their Sustainable Development Goals (SDGs), specifically relating to SDG-13 ‘Climate action’. Using a remote sensing approach to acquiring AGB estimates is attractive in that large spatial extents are covered and monitoring can be frequent, particularly once a calibrated model has been established between forest inventory and remotely sensed data. In this case, the LiDAR campaign covered 53% of the NSPSF reserve. This is the first time such data had been captured for NSPSF and used in this way. The study demonstrates the utility of both DR and FW LiDAR data for tropical PSF AGB estimation and thus illustrates the real potential for a monitoring approach underpinned by laser scanning technologies.

The AGB range presented in the in-situ plots was 126.96–443.27 Mg/ha. AGB plot values are consistent with PSF estimates from the literature [52] and therefore considered reliable and reflective of a secondary tropical PSF, with a diverse structure and density (Figure 1). The investigation is limited by the relatively small sample size ($n = 30$) used for model development and validation [69]. However, the training plots were identified as being representative of the PSF conditions found in NSPSF; all field plots had good correspondence with LiDAR data providing detailed information on both forest canopy and ground (mean ground-classified hits = 8.3% per plot) relative to a densely closed and partially open canopy structure (Figure 6). The study recognises that a larger sample size is advisable in future work for model validation. Nevertheless, the plots sampled in this study cover

![AGB Distribution](image)

**Figure 9.** Frequency distribution for DR-, FW- and Combination-model generated AGB (Mg/ha) across the 30 field plots, presented with model generated plot minimum, maximum, mean and standard deviation AGB (Mg/ha).

### 4. Discussion

Quantification of tropical PSF AGB cover is required to better understand and safeguard tropical peatlands’ ecosystem function and resilience to disturbance, both natural and anthropogenic [68]. AGB products provide accurate baseline estimates to support international climate change policies and programs such as the UN REDD+ (United Nations Reducing Emissions from Deforestation and Forest Degradation), as well as the development of sustainable management strategies for local stakeholders. Additionally, knowledge gained from the precise assessment of PSF AGB distribution will assist in Malaysia’s work towards meeting their Sustainable Development Goals (SDGs), specifically relating to SDG-13 ‘Climate action’. Using a remote sensing approach to acquiring AGB estimates is attractive in that large spatial extents are covered and monitoring can be frequent, particularly once a calibrated model has been established between forest inventory and remotely sensed data. In this case, the LiDAR campaign covered 53% of the NSPSF reserve. This is the first time such data had been captured for NSPSF and used in this way. The study demonstrates the utility of both DR and FW LiDAR data for tropical PSF AGB estimation and thus illustrates the real potential for a monitoring approach underpinned by laser scanning technologies.

The AGB range presented in the in-situ plots was 126.96–443.27 Mg/ha. AGB plot values are consistent with PSF estimates from the literature [52] and therefore considered reliable and reflective of a secondary tropical PSF, with a diverse structure and density (Figure 1). The investigation is limited by the relatively small sample size ($n = 30$) used for model development and validation [69]. However, the training plots were identified as being representative of the PSF conditions found in NSPSF; all field plots had good correspondence with LiDAR data providing detailed information on both forest canopy and ground (mean ground-classified hits = 8.3% per plot) relative to a densely closed and partially open canopy structure (Figure 6). The study recognises that a larger sample size is advisable in future work for model validation. Nevertheless, the plots sampled in this study cover
advisable in future work for model validation. Nevertheless, the plots sampled in this study cover a wide range of forest types found across NSPSF, with plots positioned in high, medium, and low forest height stands with varying densities (Figure 1).

From the numerous DR and FW LiDAR metrics tested against this range in AGB measured in the in situ plots, three AGB prediction models were determined to have a high accuracy with an rRMSE <15%: one model using discrete return data only (DR-model = 12.3%), one model using full waveform data only (FW-model = 11.6%) and one model using a combination of DR and FW data (Combination-model = 10.8%). Overall, the Combination-model was best for explaining AGB, with an $R^2$ of 0.77 (RMSE = 36.4, rRMSE = 10.8%). There was no statistically significant difference between these three best models for PSF AGB prediction.

Simple linear regression revealed LiDAR-derived metrics relating to mid-story canopy levels—for example average height, 50th height percentile, 50th height bincentsile and HOME—to be the strongest predictors of AGB in tropical PSF. In this study, models which included FW metrics were dominated by the strong predictive capabilities of the mean HOME metric. Waveform derived HOME is the distance from the waveform centroid to the ‘ground’ (i.e., the ground position of the pseudo-waveform), HOME has previously been found to be sensitive to canopy openness (including tree density) and the vertical arrangement of a forest stand [40,41]. This suggests that for complex tropical PSF the often unobserved structural and compositional diversity in the lower portions of the forest canopy is crucial for accurate AGB model prediction. This observation is crucial for understanding how LiDAR technologies should be employed for optimal AGB estimation in these environments.

Focusing on the FW LiDAR data it can be seen that the mean values of all FW metrics displayed moderate to high correlations and significant relationships with in situ plot AGB values. These findings are consistent with past research in the use of FW LiDAR in forestry applications [70,71]. Reitberger et al. [72] found FW metrics to improve the detection of under and mid-story trees compared to first/last DR LiDAR by 10–15%, though the study was unable to detect 65% of mid-story trees, and 74% of understory trees in leaf on conditions. Indeed, FW LiDAR is still limited by the lack of significant transmission through dense forest leaves and stems, however its potential to improve our knowledge and understanding of multi-layered forest structures has been widely recognised by the remote sensing community [73]. Although in this study we found no statically significant enhancement (at 95% level of confidence) with the addition of FW metrics in the predictive models for AGB estimation, the full waveforms per plot do show discernible differences. Our focus here was on the use of FW LiDAR metrics common in related literature [41,42,47]. There may exist alternative FW metrics not tested here which could be implemented in future research. This could improve the already strong predictive AGB models, and also describe other structural variables in the forest. These include crown bulk density, canopy fuel weight and for larger dominant trees, crown width [74,75]. Thus, the continued development and application of FW LiDAR in future tropical PSF AGB research is recommended, with a focus on capturing and describing FW. Recent initiatives, such as the Global Ecosystems Dynamics Investigation (GEDI) LiDAR due to be deployed on the International Space Station, respond to the recognised need for vertical structure measurements of the Earth’s surface, as put forward by the host country’s (USA) National Academy of Sciences and NASA’s Science Mission Directorate [76]. The GEDI is just one example of the progress and development in LiDAR systems and platforms that should encourage the continued application of LiDAR data in future PSF applications. Other examples include advancements in small-scale sensor technology which has permitted the use of Unmanned Aerial Vehicles (UAVs) as a LiDAR sensor platform [77,78] thus profiting from potentially higher resolution data capture at a considerably lower surveying cost at local scales. Integration with hyperspectral systems could also be advantageous [79].

Of those alternative metrics which should be implemented in future tropical PSF AGB research, those relating to LiDAR intensity [80–82] are worth considering. Small-footprint LiDAR intensity data (amount of energy reflected back to the sensor) has been increasingly adopted in forestry applications [80,83]. Top of canopy intensity metrics, in combination with LiDAR derived height
values, have been proven to contribute significantly to improved forest AGB estimations [84] and have been linked with tree species number at mid-story canopy level [85]. For this particular investigation, LiDAR intensity metrics were not included in AGB forest estimation because of uneven atmospheric conditions during the flight, therefore the resulting intensity data could not be reliably corrected for or compared between flightlines. This represents a challenge to be tackled in the future.

The results in this study demonstrate the advantage of using both DR and FW LiDAR data in tropical PSF AGB inventory programmes, which in turn provide a more detailed and accurate account of PSF structural cover and above ground carbon allocation across landscape scales. Apart from discerning the role for LiDAR in estimating AGB of these PSFs, this study has also provided in situ plot AGB data for the NSPSF, providing new PSF AGB plot data for an ecosystem underrepresented by global forest inventory databases [86,87]. This is important for subsequent research on how these forests function. As stated earlier the airborne LiDAR data does not fully cover the whole of the NSPSF reserve, in such cases satellite data (i.e., Sentinel-2 or very high-resolution systems) may be used to scale up AGB estimates across larger landscape scales [88,89]. Indeed, given how well the LiDAR estimates AGB in this environment, forward monitoring could be based on the passive satellite systems with LiDAR technologies used in a sampling capacity to provide calibration and validation data [24].

5. Conclusions

This study examined the relationship between tropical PSF AGB and small-footprint airborne LiDAR DR- and FW-derived metrics with a view to establishing the optimal use of this technology in this environment. Plot-based multiple regression analysis was performed to establish the best predictive models of PSF AGB using DR metrics (only), FW metrics (only), and a combination of DR and FW metrics. All three final models presented a high correlation with AGB estimates, with an $R^2 \geq 0.73$ and no statistical difference between the rRMSE of these models.

Overall, the results demonstrate that a Combination-model, coupling the benefits derived from both DR and FW metrics, had the best performance in modelling AGB for tropical PSF ($R^2 = 0.77$, RMSE = 36.4, rRMSE = 10.8%). The study found that perhaps of greater importance is metric selection, as opposed to LiDAR data type (DR, FW, Combination) for assessing AGB prediction. In particular, the top ranked metrics in the best performing models related to mid-canopy order height metrics, this indicating the importance of suppressed tree crowns and forest structural diversity in the mid-level portions of PSF for reliable AGB assessment.

Overall, the utility of LiDAR for tropical PSF AGB mapping and monitoring is enormous. The ever-increasing role of LiDAR in supporting forestry applications is encouraged by the recent development and expansion in sensor technology and platforms, with LiDAR data more accessible and at a lower cost than ever before. Published studies on tropical PSF AGB are few, in particular for Peninsular Malaysia, thus as a result this study contributes significantly to tropical PSF ecology. The study underpins AGB estimation across the NSPSF reserve going forward and has direct applications for sustainable tropical PSF management, supporting decision and policy makers in the preservation and protection of South East Asia’s vulnerable tropical peatlands.

Author Contributions: C.B. is a PhD student and under the supervision of D.S.B. and S.S. conceived and executed the research. C.B. was assisted in the field by S.L.E. and with data processing by D.C. C.B. undertook analyses and C.B. and D.S.B. wrote the manuscripts with input from S.S. and P.A.

Acknowledgments: We would like to thank the Airborne Research Facility of the UK’s Natural Environment Research Council for providing the airborne LiDAR datasets (to PA and DB). This work was supported by the Natural Environment Research Council [NE/L002604/1], under the Envision DTP (to PA, SS, and DB). Additional support from the University of Nottingham. We are very grateful to Martin Isenburg for generously granting the project with a complimentary full academic LAStools licence, under the LASmoons programme. We would like to thank Lahirus S. Wijedasa for his assistance and guidance with fieldwork at the pilot stage of the project. Special thanks go to the Selangor State Forestry Department for granting reserve access and providing field ranger support for data collection, and sharing their expert knowledge and experiences working in the reserve.

Conflicts of Interest: The authors declare no conflict of interest.
References


29. He, Q.; Chen, E.; An, R.; Li, Y. Above-ground biomass and biomass components estimation using LiDAR data in a coniferous forest. *Forests* 2013, 4, 984–1002. [CrossRef]


48. Cao, L.; Coops, N.C.; Innes, J.; Dai, J.; She, G. Mapping above-and below-ground biomass components in subtropical forests using small-footprint LiDAR. *Fores* 2014, 5, 1356–1373. [CrossRef]


54. Tonks, A.J.; Aplin, P.; Beriro, D.J.; Cooper, H.; Evers, S.; Vane, C.H.; Sjøgersten, S. Impacts of conversion of tropical peat swamp forest to oil palm plantation on peat organic chemistry, physical properties and carbon stocks. *Geoderma* 2017, 289, 36–45. [CrossRef]


78. Shang, X.; Chazette, P. Interest of a full-waveform flown UV lidar to derive forest vertical structures and aboveground carbon. *Forests* 2014, 5, 1454–1480. [CrossRef]


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).