Forests **2014**, *5*, 1356-1373; doi:10.3390/f5061356

OPEN ACCESS

forests

ISSN 1999-4907

www.mdpi.com/journal/forests

Article

Mapping Above- and Below-Ground Biomass Components in Subtropical Forests Using Small-Footprint LiDAR

Lin Cao [1,2,*], Nicholas C. Coops [1], John Innes [1], Jinsong Dai [2] and Guanghui She [2]

[1] Department of Forest Resources Management, University of British Columbia, V6T 1Z4, Vancouver, BC, Canada; E-Mails: Nicholas.Coops@ubc.ca (N.C.C.); john.innes@ubc.ca (J.I.)

[2] College of Forest Resources and Environment, Nanjing Forestry University, 210037, Nanjing, China; E-Mails: Aaron-dai@163.com (J.D.); ghshe@njfu.edu.cn (G.S.)

* Author to whom correspondence should be addressed; E-Mail: ginkgocao@gmail.com; Tel./Fax: +86-25-8542-7327.

Received: 4 April 2014; in revised form: 8 May 2014 / Accepted: 10 June 2014 / Published: 16 June 2014

Abstract: In order to better assess the spatial variability in subtropical forest biomass, the goal of our study was to use small-footprint, discrete-return Light Detection and Ranging (LiDAR) data to accurately estimate and map above- and below-ground biomass components of subtropical forests. Foliage, branch, trunk, root, above-ground and total biomass of 53 plots (30 × 30 m) were modeled using a range of LiDAR-derived metrics, with individual models built for each of the three dominant forest types using stepwise multi-regression analysis. A regular grid covered the entire study site with cell size 30 × 30 m corresponding to the same size of the plots; it was generated for mapping each biomass component. Overall, results indicate that biomass estimation was more accurate in coniferous forests, compared with the mixed and broadleaved plots. The coefficient of determination (R^2) for individual models was significantly enhanced compared with an overall generic, or common, model. Using independent stand-level data from ground inventory, our results indicated that overall the model fit was significant for most of the biomass components, with relationships close to a 1:1 line, thereby indicating no significant bias. This research illustrates the potential for LiDAR as a technology to assess subtropical forest carbon accurately and to provide a better understanding of how forest ecosystems function in this region.

Keywords: biomass components; carbon; small-footprint LiDAR; subtropical forests; southeastern China

1. Introduction

Tropical forests are species-rich, carbon-dense and highly productive ecosystems, which not only play a significant role in maintaining the regional ecological environment, but also make important contributions to the global carbon balance [1]. In general, the tropics encompass the area between the Tropics of Cancer and Capricorn (23°27′ latitude north and south). The subtropics are differentiated from the tropics by thermal criteria, *i.e.*, the frost limit or the +18 °C isotherm of the coldest months in the lowlands, and consist of subtropical arid lands, the Mediterranean subtropics and the humid subtropics [2]. Within China, the area of subtropical forests is larger than in any other country and southeastern China contains the largest proportion of humid subtropical forests globally. As this area is located in China's most densely populated regions and these regions have the greatest economic growth, their subtropical forests are particularly important for conservation, wood production and the protection and improvement of the regional environment and carbon cycle. Despite their importance, there is still considerable uncertainty about carbon budgets within these subtropical forests, despite several national-scale studies using historical forest inventory data [3,4]. An understanding of the carbon in these forests is particularly important as China is considering a national forest carbon inventory and policy for carbon management that would require accurate methods of biomass estimation. In addition, estimation of biomass components, rather than just total biomass, can also provide additional information for forest management decisions. For example, information can be used for the estimation of stem and branch biomass for biofuels, as well as assessing crown biomass (*i.e.*, branch and foliage biomass) for predicting fire hazard [5,6].

Forest carbon stocks can be estimated using remote sensing from both satellite and airborne platforms, and a major benefit of this approach is the potential to provide "wall-to-wall" observation [7]. Remote-sensing methodologies have typically been more successful at measuring carbon stocks in boreal and temperate forests, or in young stands with lower forest carbon densities than in other forest types [8]. Tropical (and subtropical) forests are among the most structurally complex ecosystems in the world with a three-dimensional arrangement of canopy elements (e.g., leaves, branches, trunks) from the top of the canopy to the ground [9]. In such cases, the spectral response from remote-sensing instruments tends to saturate rapidly, which inhibits reliable forest carbon stock estimates in these ecosystems.

Light Detection and Ranging (LiDAR) is an active airborne remote-sensing laser technology capable of providing detailed, spatially explicit, three-dimensional information on vegetation structure [10]. LiDAR systems send out pulses of laser light and measure the signal return time to directly estimate the height and vertical structure of forests [11,12]. The saturation problem in biomass estimation over dense canopies can be overcome using LiDAR, as laser pulses penetrate the forest canopy, even through dense multi-layered canopies, and as a result there tends to be a strong correlation between LiDAR data and biomass. However, to date, much of the research related to

LiDAR biomass estimation has been conducted in temperate and boreal forests and the transition region of these forests [13–15]. Published studies from tropical (and subtropical) forests are few. Drake *et al.* (2003) [16] showed that metrics from Laser Vegetation Imaging Sensor (LVIS) large-footprint (25 m diameter) LiDAR waveforms were strongly correlated with mean stem diameter, basal area and above-ground biomass (AGB) in tropical moist forest (TMF) sites in Panama and a tropical wet forest (TWF) in Costa Rica. Simple linear regression using the height of median energy (HOME) to estimate AGB had an R^2 of 0.66 for TMF and of 0.89 for TWF. Saatchi *et al.* (2011) [17] mapped the above- and below-ground biomass of global tropical forests based on Geoscience Laser Altimeter System (GLAS) large-footprint (52 to 90 m) LIDAR waveforms (onboard the Ice, Cloud, and land Elevation Satellite, ICESat, acquired in 2003 and 2004). Indexes of total waveform extent and the height of the 10th and 90th percentiles of waveform height were used with least-squares regression to estimate Lorey's height, *i.e.*, plot-level means of basal area weighted tree heights. Allometric relations between Lorey's height and AGB of calibration plots differed significantly in Central and South America (R^2 of 0.89), in Africa (R^2 of 0.85), and in South-East Asia (R^2 of 0.79). Tsui *et al.*, (2012) [18] estimated above-ground component and total biomass for a coniferous temperate forest using LiDAR-derived metrics and backscatter coefficients from RADAR. The results of the LiDAR-based models indicated that stem biomass had the most accurate relationship with the LiDAR measurements, while crown biomass showed the lowest relationship (relative RMSE = 16%–22%). Næsset (2004) [19] estimated above-ground and below-ground biomass for young and mature coniferous forest using small-footprint discrete-return LiDAR in boreal forest ecosystems. The results demonstrated that above-ground biomass explained 92% of the variability, whereas the models for below-ground biomass explained 86%.

The research presented in this paper reports the use of a small-footprint LiDAR to estimate biomass components in subtropical forests in southeastern China. The objectives were (1) to build models for above- and below-ground biomass components for forest plots (30 × 30 m) in three subtropical forest types (*i.e.*, mixed, broadleaved and coniferous forests) using LiDAR data, and to assess the impact of forest type on the accuracy of the regression models; and (2) to map the above- and below-ground forest biomass components across the site using LiDAR data, and evaluate the accuracy using a suite of independent stand-level field inventory data.

2. Methods

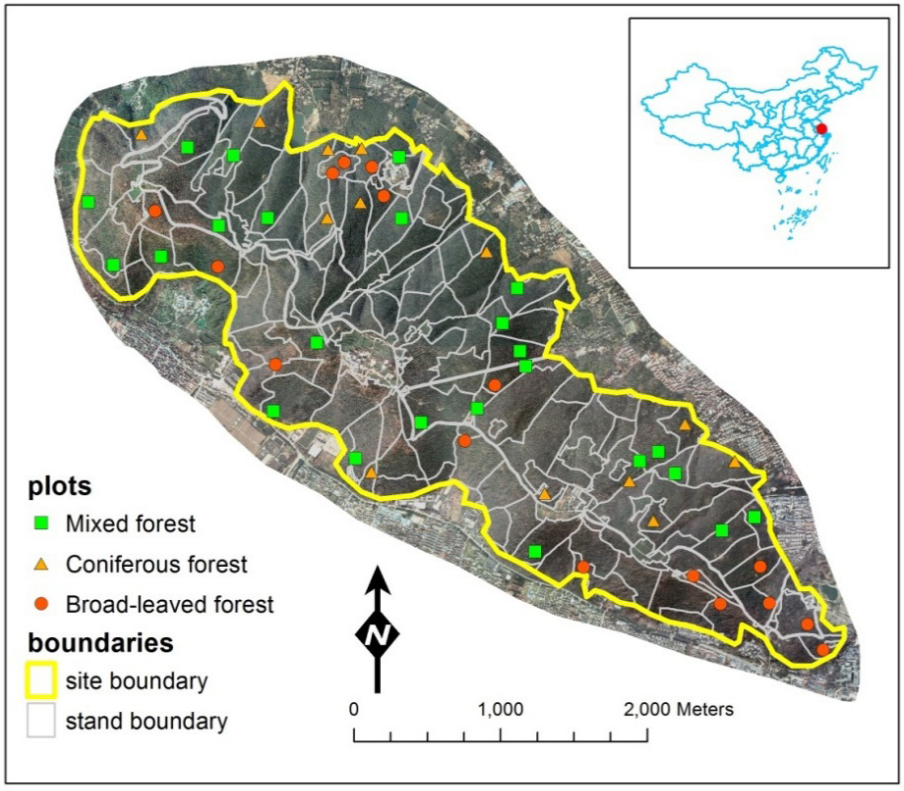
2.1. Study Site and Materials

2.1.1. Study Site

This study was conducted at Yushan Forest, a state-operated forest and national forest park, located near the town of Changshu in Jiangsu province, southeastern China (120°42′9.4″ E, 31°40′4.1″ N). The forest covers approximately 1,103 ha, with an elevation range of approximately 20–261 m above sea level (Figure 1). This site is situated in the north subtropical monsoon climatic region with an annual mean precipitation of 1062.5mm; the highest monthly precipitation occurs in June (171.3 mm) and July (147.0 mm). The forest in Yushan belongs to the north subtropical mixed secondary forest with three main forest types: coniferous-dominated, broadleaved dominated and mixed forests [20]. The

main conifers include Masson pine (*Pinus massoniana* Lamb.) and Chinese fir (*Cunninghamia lanceolata* (Lamb.) Hook.) as the apex species, and a small percentage of Slash pine (*Pinus elliottii* Engelm.) and Japanese Blackbark Pine (*Pinus thunbergii* Parl.). The primary broadleaved species are Sawtooth oak (*Quercus acutissima* Carruth.) and Chinese sweet gum (*Liquidambar formosana* Hance), associated with some evergreen broadleaved tree species belonging to the Fagaceae, Lauraceae and Theaceae.

Figure 1. Yushan forest study site boundaries and the locations of the plots of different forest types.



2.1.2. Plot Data

A total of fifty-three 30 × 30 m (900 m^2) plots were established within the study site, covering a range of species composition, age, and site index, according to a GIS-based historical forest inventory map (2005). These plots were divided into three forest types based on species compositions (this information was recorded in the attribute table of the forest type layer): 12 coniferous forest plots, all dominated by Masson pine and Chinese fir; 18 broadleaved forest plots dominated by Sawtooth oak and Chinese sweet gum, and 23 mixed species forest plots with a mixture of coniferous and broadleaved species were measured for this study (Table 1).

Plot data were collected in summer 2012 (June to July). Plot corners were located using Trimble GeoExplorer 3000 Handheld GPS units augmented with Satellite Based Augmentation Systems (SBAS) to improve the accuracy and integrity of the GPS data, resulting in submeter accuracy [21]. At each plot, all live trees with a DBH > 5cm, species, diameter, height, height to crown base, and crown width in both cardinal directions were measured. Small trees (DBH < 5cm) and dead wood were also

tallied for total stem number, but not used in biomass calculations. DBH was measured on all trees using a diameter tape. Heights of all trees were measured using a Vertex IV hypsometer. Crown widths were determined by measuring two crown radii, one being the major axis (to the tip of the outmost branch on the longest side of the crown) and the other being 90° (perpendicular) to the first axis.

Several plot-level forest variables were calculated based on the individual tree data, including basal area (m^2 ha^{-1}) and Lorey's height (*i.e.*, the basal area weighted height). DBH and height measurements were loaded into the Chinese Forest Carbon Measurement and Monitoring (CFC2M) database program for derivation of individual tree-based biomass components (*i.e.*, foliage, branch, trunk, root, above-ground and total biomass) using published, nationwide (categorized by provinces), species-specific, allometric equations [22]. Allometric equations developed from tree inventory data from local or nearby provinces were selected for this research. Tree-based biomass estimates were summed by plot to determine plot-level biomass components. A summary of the derived ground biomass estimates is shown in Table 1.

Table 1. Summary of the mean and range of the derived biomass attributes of the plots; $H_{\text{Lorey's}}$ (m): Lorey's height; DBH_{avg} (cm): average DBH; W_f (Mg ha^{-1}): foliage biomass; W_b (Mg ha^{-1}): live branch biomass; W_s (Mg ha^{-1}): trunk wood biomass; W_a (Mg ha^{-1}): total above-ground biomass; W_r (Mg ha^{-1}): root biomass; W_t (Mg ha^{-1}): total biomass.

Variables	***Coniferous forest*** (***n* = 12**)		***Broadleaved forest*** (***n* = 18**)		***Mixed forest*** (***n* = 23**)	
	Range	**Mean**	**Range**	**Mean**	**Range**	**Mean**
Biomass-related attributes						
$H_{\text{Lorey's}}$	4.50–14.18	10.54	7.70–18.52	11.96	7.79–14.83	10.63
DBH_{avg}	8.15–20.90	14.19	12.49–22.43	15.19	10.95–20.62	14.22
W_f (foliage)	1.04–23.57	11.69	2.13–8.67	5.12	3.18–19.93	7.89
W_b (branch)	1.62–25.12	12.36	9.72–44.46	19.44	6.42–25.33	13.90
W_s (trunk)	8.35–78.70	48.48	18.65–173.05	72.17	36.61–100.26	67.28
W_a (above-ground)	11.02–127.39	72.52	32.03–219.67	96.76	49.65–141.73	89.07
W_r (root)	6.25–39.42	22.60	10.31–45.62	29.06	15.70–43.05	27.26
W_t (total)	17.27–166.81	95.12	42.34–265.29	125.82	65.35–184.78	116.33
Species composition						
Chinese fir (%)	0–89	29	0	0	0–39	7
Pines (%)	0–90	53	0–29	13	19–52	40
Broadleaved (%)	2–29	18	71–100	87	27–67	53

2.1.3. Stand Inventory Data

Additional set of stand inventory data was available to serve as independent validation data, measured in summer 2012 (June to August) following the Chinese National Forest Inventory (CNFI) guidelines. From the approximately 150 stands in the dataset, stands on steep slopes (>30°) dominated by non-timber species (e.g., fruits, bamboo and tea plantations), and close to anthropogenic activities within the forest (such as stands with large roads intersecting the forest), were removed. In addition, if one of the original 53 plots occurred within a stand, that stand was also not selected. The remaining

45 stands were then stratified by age class, site index, and tree species, and again classified as coniferous forest (11 stands), broadleaved forest (15 stands) and mixed forest (19 stands). Square plots which were distributed systematically according to a regular grid within each stand were used to provide tree level data. Stand-level biomass components were calculated from the species-specific allometric equations using the CFC2M database program as biomass per hectare. A summary of selected stand data is shown in Table 2.

Table 2. Summary of the calculated attributes of the selected stands; W_f (Mg ha^{-1}): foliage biomass; W_b (Mg ha^{-1}): live branch biomass; W_s (Mg ha^{-1}): trunk wood biomass; W_a (Mg ha^{-1}): total above-ground biomass; W_r (Mg ha^{-1}): root biomass; W_t (Mg ha^{-1}): total biomass.

Variables	*Coniferous forest* (*n* = 11)		*Broadleaved forest* (*n* = 15)		*Mixed forest* (*n* = 19)	
	Range	**Mean**	**Range**	**Mean**	**Range**	**Mean**
Biomass-related attributes						
W_f (foliage)	2.69–18.06	10.16	3.18–12.52	8.64	4.83–20.76	9.61
W_b (branch)	3.58–18.86	12.08	8.12–29.57	19.48	6.08–22.89	12.96
W_s (trunk)	10.34–67.30	47.32	32.19–135.85	78.75	34.33–95.01	57.40
W_a (above-ground)	16.56–95.26	73.64	54.51–183.45	108.13	50.02–131.73	85.29
W_r (root)	8.31–33.75	21.26	23.04–39.84	29.25	13.84–34.97	25.16
W_t (total)	23.81–146.15	94.18	93.24–216.20	140.96	75.44–160.28	116.37
Species composition						
Chinese fir (%)	0–90	21	0–10	2	0–20	4
Pines (%)	0–100	61	0–30	6	30–50	52
Broadleaved (%)	0–30	19	70–100	94	40–60	47

2.1.4. LiDAR Data

Airborne laser scanner data were acquired on August 17th in 2013 using a Riegl LMS-Q680i flown at 900 m above-ground level, with a flight line side-lap of ≥60%. The system acquired data at a 400 kHz pulse rate with a scanning angle of ±30° from nadir; pulses transmitted at scan angles that exceeded 15° were excluded from the final dataset The returned waveforms were recorded with a temporal sample spacing of 1 ns and beam footprint size around 0.25 m. This configuration yielded a pulse density of ≥10 points m^{-2}. Full-waveform decomposition has been performed by using the RIEGL RiAnalyze® software provided point clouds with XYZ coordinates and additional parameters [23]. A 1 m digital terrain model (DTM) was created in two steps from the point cloud data: first, the data was filtered to remove the above-ground returns, and secondly the DTM was created by calculating the average elevation from the remaining (ground) LiDAR returns within a cell (Cells that contain no points are filled by interpolation using neighboring cells). The normalized point clouds were then extracted for each plot using the coordinates of the lower left and upper right corners.

2.2. LiDAR Metrics

LIDAR metrics are descriptive structure statistics calculated from the raw LiDAR point cloud. The metrics for each plot within the 30 × 30 m area of the 53 plots were calculated using the procedures of Lim *et al.* 2003 [13], Naesset 2002 [24], Naesset 2004 [25] and Ferster *et al*. 2009 [26], and a summary of the LiDAR metrics with corresponding descriptions is shown in Table 3 [13,24–26].

LiDAR metrics were generated from first returns (*i.e.*, the first thing hit by a laser pulse, corresponding to the forest canopy) based on research from a number of studies that have found that first returns may have more predictive ability for forest biophysical attributes than all returns [27]. As a result, percentile heights and canopy return density metrics were computed from first returns with a height of 2 m above-ground-level, reducing potential biases due to low-lying vegetation [28,29].

Table 3. Summary of LiDAR metrics computed from LiDAR point cloud.

Metrics	Description
Percentile height (h_5, h_{10}, h_{20}, …, h_{95})	The percentiles of the canopy height distributions (5th, 10th, 20th . . . 95th) of first returns.
Canopy return density (d_0, d_1, d_2, …, d_9)	The canopy return density over a range of relative heights, *i.e.*, percentage (0%–100%) of first returns above the quantiles (0, 10, 20 . . . 90) to total number of first returns.
Mean height (h_{mean})	Mean height above ground of all first returns.
Maximum height (h_{max})	Maximum height above ground of all first returns.
Coefficient of variation of heights (h_{cv})	Coefficient of variation of heights of all first returns.
Canopy cover above 2 meters (CC_{2m})	Percentages of first returns above 2 m.
Canopy cover above mean (CC_{mean})	Percentages of first returns above the first return mean heights.

2.3. Statistical Analyses

All of the dependent variables (ground-truth data of the biomass components) and independent variables (LiDAR-derived plot-level metrics) were transformed using the natural logarithm and corrected for bias [30]. To ensure the independent variables were not highly correlated, collinearity was evaluated using Principal Component Analysis (PCA) based on the correlation matrix. Models with condition number (κ) lower than 30 were accepted to ensure that there was no serious collinearity in the selected models [31]. Individual regression models were developed for all of the biomass components. Stepwise selection was performed to select variables for the final models. Predictor variables were left in the model using an F-test with a $p < 0.05$ significance level. Prediction models were assessed using coefficient of determination (R^2), Root-Mean-Square Error (RMSE), which has been transformed back to original scale, and relative RMSE (rRMSE) defined as the percentage of the ratio of RMSE and the observed mean values.

Dummy variables (or class variables) were added to the selected models as the dependent variables to assess whether these models differ between forest types [32]. Three types of combinations of dependent variables, *i.e.*, Common, Intercept Only and Full, were chosen to evaluate the effects of the

intercepts and slopes of the final model fitting. "Common model" implies a reduced model with one common equation regardless of forest types. "Intercept Only" models implies reduced models with common slopes for all forest types, but different intercepts. "Full model" implies both intercepts and slopes varying by forest types. The partial-*F* test of Full *versus* Common (FVC) and Full *versus* Intercepts Only (FVIO) was applied to test whether only the intercepts or both intercepts and slopes are different among separated models of different forest types. Cross-validation was used to assess model accuracy. This is an iterative process were one plot is removed from the dataset at a time, and the selected models re-fitted to the n-1 remaining plots [24]. Once models were developed, differences between the field-measured stand-level biomass components and the predicted values were evaluated for the 45 selected stands. A 1:1 line was applied to assess accuracy.

2.4. Biomass Mapping

The second objective of this research was to map forest biomass components using the regression models estimated from the sample plots to improve our understanding of the spatial distribution of each component for the whole site. To do so, a regular grid covering the entire study site was generated, with a cell size of 30 × 30 m corresponding to the same size of the plots. The selected forest type-specific regression models were used to predict foliage biomass (W_f), live branch biomass (W_b), trunk wood biomass (W_s), above-ground biomass (W_a), root biomass (W_r) and total biomass (W_t) of each grid.

3. Results

All the biomass components were fitted against the predictor variables (LiDAR-derived metrics) derived from the first returns. Separate models were built for each of the forest types, *i.e.*, mixed, broadleaved and coniferous forests (Table 4). PCA analysis indicated that no serious collinearity problems existed in the models. All of the selected models included less than five predictor variables.

The fit for the common, generic model, for all stands, irrespective of dominant species was relatively low (R^2 = 0.26–0.63). In comparison, the relationship were generally improved for individual models based on dominant species (R^2 = 0.21–0.84). In general, the fitted models had higher correlations in coniferous forest (R^2 = 0.80–0.84) than in other forest types. The results for the biomass components indicated that root biomass (W_r) in coniferous forest (R^2 = 0.83) was well predicted, whereas models for root biomass (W_r) in broadleaved forests (R^2 = 0.53) and mixed (R^2 = 0.64) had a slightly lower level of significance. For the other individual components, namely foliage, branch, and stem biomass, the R^2 values ranged from 0.21–0.83. The poorest correlations were for foliage biomass (R^2 = 0.21) in coniferous forests and the highest rRMSE (43.22%) values also reported for the foliage biomass but in mixed forests. The coefficients of determination for the above-ground and total biomass ranged from 0.58 to 0.84.

Table 4. Summary of the biomass components prediction models and plot-level accuracy assessment results; W_f (Mg ha^{-1}): foliage biomass; W_b (Mg ha^{-1}): live branch biomass; W_s (Mg ha^{-1}): trunk wood biomass; W_a (Mg ha^{-1}): total above-ground biomass; W_r (Mg ha^{-1}): root biomass; W_t (Mg ha^{-1}): total biomass.

Dependent	Final models	R^2	RMSE	*r*RMSE (%)
Common models				
W_f	$\ln W_f = 3.590 + 2.334 \ln h_{cv} + 0.867 \ln h_{25} - 3.021 \ln CC_{mean} + 2.707 \ln d_4$	0.26	4.87	62.44
W_b	$\ln W_b = 1.198 - 0.907 \ln h_{10} + 2.635 \ln h_{25} - 0.633 \ln CC_{mean}$	0.56	4.55	29.47
W_s	$\ln W_s = 2.347 + 1.297 \ln h_{75} - 2.646 \ln CC_{2m} + 2.375 \ln d_2$	0.59	16.96	26.22
W_a	$\ln W_a = 2.464 - 0.634 \ln h_{10} + 1.997 \ln h_{25} - 0.279 \ln CC_{mean}$	0.60	20.74	23.58
W_r	$\ln W_r = 1.713 + 0.432 \ln h_{cv} + 1.036 \ln h_{25}$	0.59	4.85	18.10
W_t	$\ln W_t = 2.803 - 0.625 \ln h_{10} + 1.901 \ln h_{25} - 0.247 \ln CC_{mean}$	0.63	24.40	21.26
Coniferous forest				
W_f	$\ln W_f = -1.892 + 1.976 \ln h_{cv} + 2.902 \ln h_{25} - 3.059 \ln CC_{mean} + 3.055 \ln d_4$	0.80	4.10	35.07
W_b	$\ln W_b = -3.060 - 0.065 \ln h_{10} + 3.048 \ln h_{25} - 0.136 \ln CC_{mean}$	0.81	4. 51	36.46
W_s	$\ln W_s = -1.035 + 1.840 \ln h_{75} - 5.718 \ln CC_{2m} + 5.961 \ln d_2$	0.80	10.14	20.93
W_a	$\ln W_a = -0.735 + 0.228 \ln h_{10} + 2.166 \ln h_{25} + 0.076 \ln CC_{mean}$	0.83	18.52	25.54
W_r	$\ln W_r = 0.086 + 0.421 \ln h_{cv} + 1.799 \ln h_{25}$	0.83	4.46	19.74
W_t	$\ln W_t = 0.054 + 0.071 \ln h_{10} + 2.136 \ln h_{25} + 0.026 \ln CC_{mean}$	0.84	22.28	23.42
Broadleaved forest				
W_f	$\ln W_f = 4.980 + 1.227 \ln h_{cv} + 0.327 \ln h_{25} - 4.652 \ln CC_{mean} + 3.685 \ln d_4$	0.21	1.90	37.16
W_b	$\ln W_b = 3.104 + 0.882 \ln h_{10} + 0.173 \ln h_{25} - 0.605 \ln CC_{mean}$	0.71	4.79	24.67
W_s	$\ln W_s = 6.937 + 1.275 \ln h_{75} - 7.328 \ln CC_{2m} + 6.069 \ln d_2$	0.77	18.04	24.99
W_a	$\ln W_a = 3.407 - 0.336 \ln h_{10} + 1.622 \ln h_{25} - 0.471 \ln CC_{mean}$	0.62	24.69	25.52
W_r	$\ln W_r = 2.271 + 0.284 \ln h_{cv} + 0.689 \ln h_{25}$	0.53	4.76	16.37
W_t	$\ln W_t = 3.560 - 0.515 \ln h_{10} + 1.685 \ln h_{25} - 0.382 \ln CC_{mean}$	0.64	27.70	22.02
Mixed forest				
W_f	$\ln W_f = -1.925 + 2.595 \ln h_{cv} + 0.232 \ln h_{25} - 0.643 \ln CC_{mean} + 1.037 \ln d_4$	0.54	3.41	43.22
W_b	$\ln W_b = -1.802 - 1.228 \ln h_{10} + 2.391 \ln h_{25} + 0.398 \ln CC_{mean}$	0.62	4.25	30.60
W_s	$\ln W_s = 1.336 + 1.491 \ln h_{75} - 2.765 \ln CC_{2m} + 2.643 \ln d_2$	0.70	11.28	16.77
W_a	$\ln W_a = 2.295 - 0.451 \ln h_{10} + 1.588 \ln h_{25} - 0.084 \ln CC_{mean}$	0.58	17.10	19.20
W_r	$\ln W_r = 1.444 + 0.295 \ln h_{cv} + 1.091 \ln h_{25}$	0.64	4.27	15.66
W_t	$\ln W_t = 2.481 - 0.428 \ln h_{10} + 1.504 \ln h_{25} - 0.028 \ln CC_{mean}$	0.62	20.73	17.82

Cross-validation of the developed models indicated that mean differences between the estimated and field-measured values at the 53 sample plots were insignificant (Table 5). The standard deviations of the differences ranged from 11.12 to 29.15 Mg ha^{-1} for above-ground biomass (W_a) and from 16.60 to 24.83 Mg ha^{-1} for total biomass (W_t). For foliage (W_f), branch (W_b), trunk (W_s), and root biomass (W_r), the standard deviations ranged from 1.44 to 4.78 Mg ha^{-1} (18.3%–45.7%), 3.25 to

5.87 Mg ha^{-1} (23.4%–39.4%), 10.16 to 22.43 Mg ha^{-1} (15.1%–31.1%), and 3.25 to 6.50 Mg ha^{-1} (14.4%–22.4%), respectively.

Table 5. Plot-level validation results: summary of the plot-level observed mean, the mean and standard deviation of the differences of the cross-validation results for the separated models.

Variables	Coniferous forest (n = 12)			Broadleaved forest (n = 18)			Mixed forest (n = 23)		
(Mg ha^{-1})	#OM	#MD	#SD	#OM	#MD	#SD	#OM	#MD	#SD
W_f	11.69	−1.09 NS	4.78 (40.9%)	5.12	−0.87 NS	2.34 (45.7%)	7.89	0.48 NS	1.44 (18.3%)
W_b	12.36	−0.51 NS	4.87 (39.4%)	19.44	−1.02 NS	5.87 (30.2%)	13.90	−0.21 NS	3.25 (23.4%)
W_s	48.48	−1.96 NS	11.96 (24.7%)	72.17	−0.42 NS	22.43 (31.1%)	67.28	−1.02 NS	10.16 (15.1%)
W_a	75.52	−0.13 NS	13.91 (18.4%)	96.76	−2.33 NS	29.15 (30.1%)	89.07	−1.76 NS	11.12 (12.5%)
W_r	22.60	−0.94 NS	3.25 (14.4%)	29.06	−1.12 NS	6.50 (22.4%)	27.26	−0.04 NS	3.32 (12.2%)
W_t	95.12	0.02 NS	16.60 (17.5%)	125.82	−4.98 NS	20.76 (16.5%)	116.33	−1.07 NS	24.83 (21.3%)

Level of significance: NS = not significant (>0.05); * $p < 0.05$, ** $p < 0.01$; #OM: The arithmetic mean of the observed value; #MD: Mean of the differences in cross-validation of the selected regression equations in Table 4; #SD: Standard deviation of the differences in cross-validation of the selected regression equations in Table 4.

The forest type-specific equations were applied to the 45 reference stands to predict the stand-level forest biomass components. The comparisons indicated that 9 of the 12 developed models in coniferous and broadleaved forests did not have significant mean differences between predicted and stand inventory measures ($p < 0.05$), whereas only one model for mixed forests was not significantly different (Table 6). For foliage biomass (W_f), a mean difference of 0.24 Mg ha^{-1} indicated an underestimation by LiDAR in broadleaved stands. In mixed and coniferous stands, the LiDAR-estimated foliage biomass (W_f) overestimated the field-measured values by 1.49–1.52 Mg ha^{-1}. The predicted branch biomass (W_b) overestimated the ground-truth by 3.01–3.05 Mg ha^{-1} in mixed and coniferous stands, but underestimated by 0.37 Mg ha^{-1} in broad-leaved stands, although these differences in broadleaved and coniferous stands were not significant. Stand-measured trunk biomass (W_s) values were significantly overestimated by 3.02–12.50 Mg ha^{-1}, which corresponded to a bias of 5.8%–18.6%, whereas an underestimation of between 9.1% and 17.4% was found for foliage biomass (W_f) in coniferous and mixed stands, and an overestimation of 6.5% for broadleaved stands. Above-ground biomass (W_a) and total biomass (W_t), were the only two models where differences were not significantly different for coniferous and broadleaved forests.

Table 6. Stand-level validation results: summary of the observed mean, the mean and standard deviation for the differences in stand predictions using the selected regression equations.

Variables	Coniferous forest (n = 11)			Broadleaved forest (n = 15)			Mixed forest (n = 19)		
(Mg ha^{-1})	#OM	#MD	#SD	#OM	#MD	#SD	#OM	#MD	#SD
W_f	7.08	1.52 NS	2.63 (37.1%)	9.25	−0.24 NS	2.15 (22.2%)	6.37	1.49 *	1.88 (29.5%)
W_b	7.83	3.01 NS	3.10 (39.6%)	19.48	−0.37 NS	6.18 (31.7%)	9.57	3.05 *	3.92 (41.0%)
W_s	38.93	10.54 *	12.93 (33.2%)	67.21	12.50 **	32.85 (48.9%)	52.32	3.02 NS	10.50 (20.1%)
W_a	59.48	9.90 NS	10.43 (17.5%)	94.12	15.81 *	31.55 (33.5%)	66.42	8.90 **	12.32 (18.5%)
W_r	24.94	−2.27 NS	2.28 (9.14%)	27.47	1.78 NS	7.43 (27.0%)	29.49	−4.33 **	3.35 (11.4%)
W_t	96.63	10.14 NS	15.70 (16.2%)	141.72	15.20 NS	19.03 (13.4%)	115.47	14.05 *	25.62 (22.2%)

Level of significance: NS = not significant (>0.05); * $p < 0.05$, ** $p < 0.01$; #OM: The arithmetic mean of the observed value; #MD: Mean of the differences between the model-predictions and the field inventory data; #SD: Standard deviation of the differences between the model-predictions and the field inventory data.

Figure 2 shows LIDAR-based biomass predictions *versus* the stand-based biomass estimates in the 45 reference stands. As indicated, overall model fit was good for most of the biomass components with relationships close to the 1:1 line. The mean value of the individual components for each sample plot, illustrated in Figure 3, were extrapolated across the entire site using the regression models estimated from the sample plots (Figure 4). As expected, relatively low biomass components were found on steep slopes, especially the mountain ridges of both southern and northern slopes. Higher values of biomass components were found in most valleys, especially on the northern slopes where the dominant species preferred cooler summer temperatures and sufficient water supply. In some of the north–central areas, very high value stands of both above-ground and root biomass were associated with old-growth broadleaved forests (e.g., Sawtooth oak), usually around old temples and abandoned buildings. Coniferous-dominated forests contributed high biomass in areas with good site quality (e.g., forest dominated by Masson pine in some valleys with thick soil layers) and relatively flat areas (e.g., stands dominated by Slash pine at the foot of southern part of the study area). There was a relatively large area of lower biomass, especially root biomass, in the northwest–central part of the study area, reflecting the presence of low-quality Chinese fir plantations on ridge tops with poor site quality and some agroforestry plantations (Chinese fir accompanied by tea) with large canopy gaps and low stem density.

Figure 2. Scatterplots and 1:1 line (dotted line) of biomass components between the field-measured and the model-estimated results of the selected stands (*i.e.*, (**a**) foliage biomass; (**b**) branch biomass; (**c**) stem biomass; (**d**) above-ground biomass; (**e**) root biomass; (**f**) total biomass). The models were built from the sample plots of three forest types.

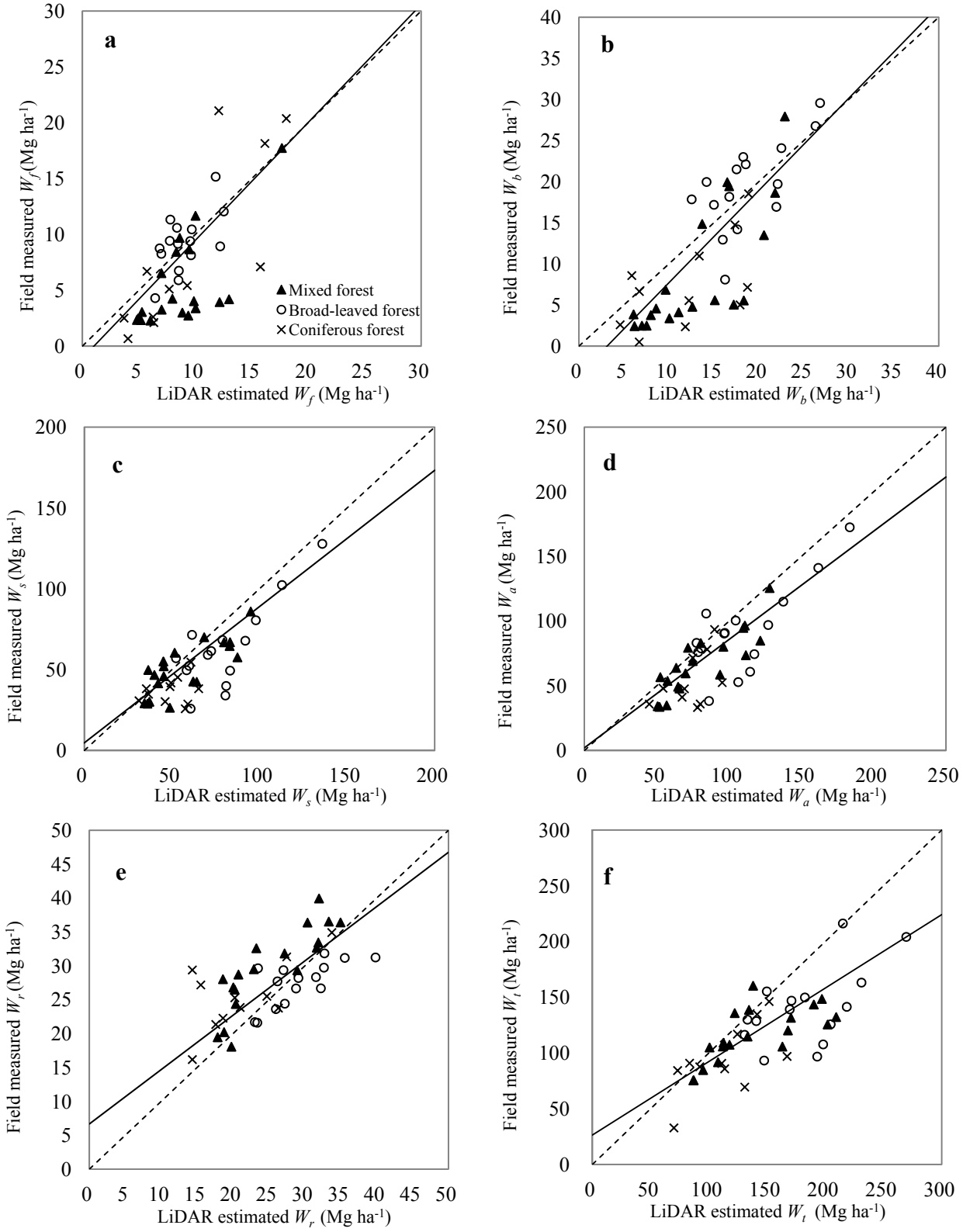


Figure 3. Mean value of the biomass components (Mg ha^{-1}) of the 53 sample plots from (1) field measurements; (2) estimations of the dummy variable including stepwise regression model; and (3) predictions of the general models of three forest types, *i.e.*, coniferous, broadleaved and mixed forests.

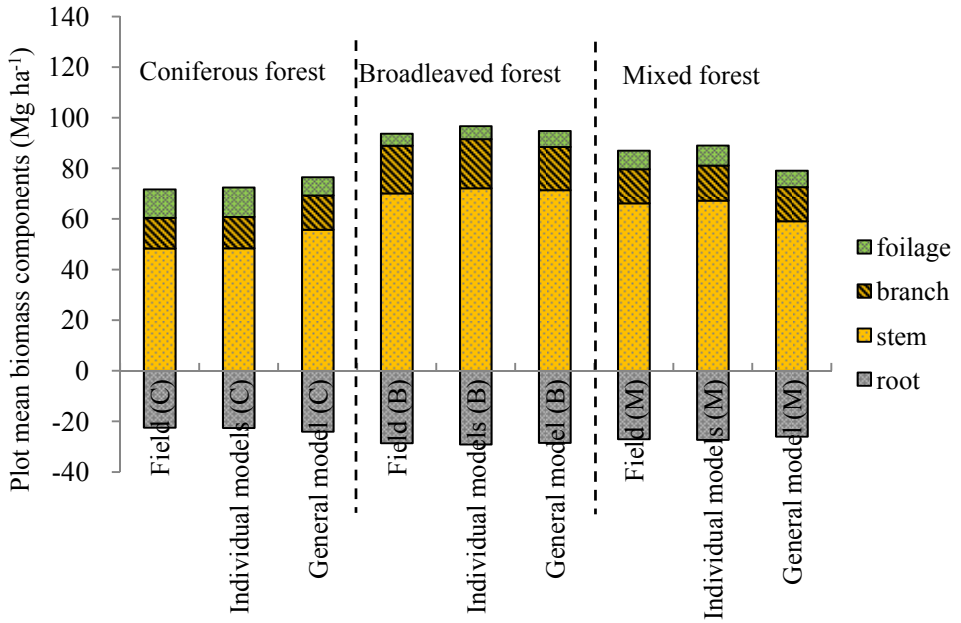
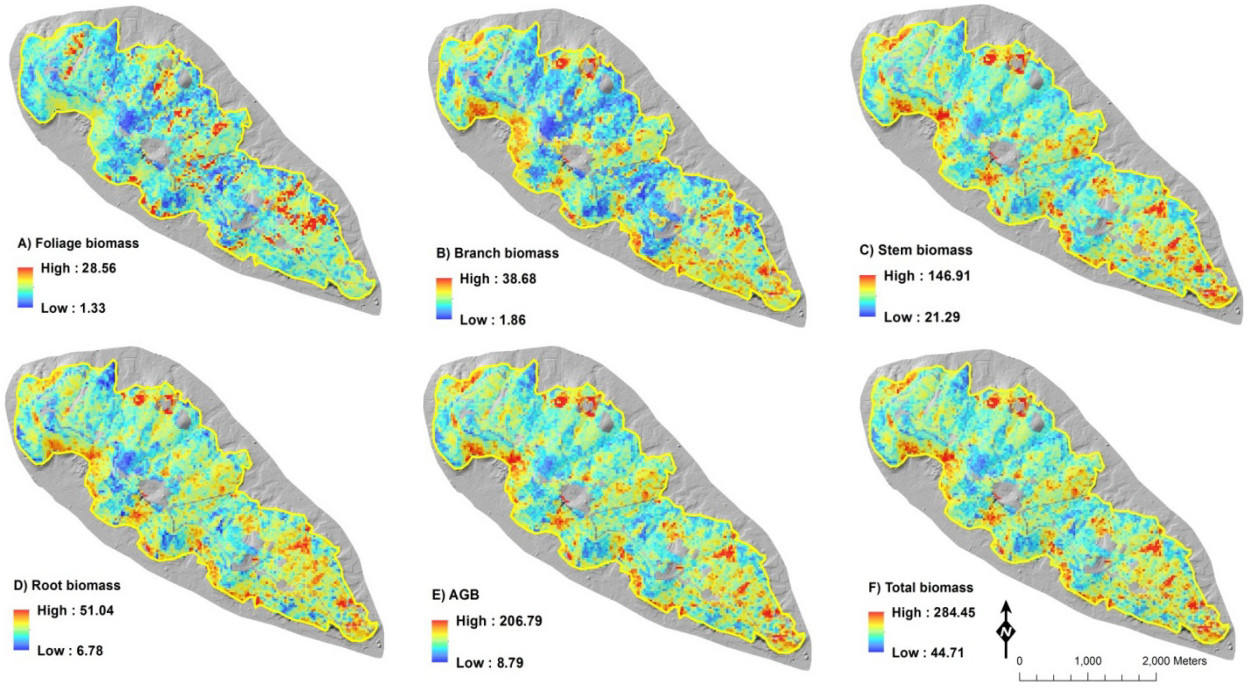


Figure 4. Maps of biomass components (Mg ha^{-1}) across the research site (pixel size = 30m), which calculated from the individual models; (**A**) map of the estimated foliage biomass; (**B**) branch biomass; (**C**) stem biomass; (**D**) root biomass; (**E**) above-ground biomass (AGB); (**F**) total biomass.



4. Discussion

Previous studies have demonstrated that in forests with highly variable species composition, habitat-based models are required to ensure accurate biomass predictions [24,33,34]. Likewise in our study, the inclusion of class variables based on dominant species addressed this issue successfully and improved estimations significantly. For foliage, branch, and stem biomass, most of the R^2 values of the individual species models were above 0.54. The coefficients of determination for above-ground and total biomass ranged from 0.58 to 0.84, and the R^2 for root biomass in coniferous forest was as high as 0.83. Lim *et al.* (2003) [13] estimated biomass equations in hardwood forests in Canada by small-footprint LiDAR. The highest proportion of explained variability in the estimated log-transformed linear models was 85%, which is similar to the results in our study. Our results indicate that biomass predictions for coniferous stands were more accurate than broadleaved and mixed forest types. In coniferous forests, 80%–84% of the variability could be explained, whereas, 21%–77% and 54%–70% could be explained in broadleaved and mixed forests separately. Næsset (2004) [19] reported that forest type did not have any significant impact on the estimated biomass models. However, in our study, we assumed the complex forest conditions in the subtropical forests contained a much larger diversity, making the effects of tree-species composition (classified as forest types) significant.

The dominated coniferous species in our research site are Masson pine and Chinese fir (Table 1), which, in this forest, have a relatively even-aged tree structure and higher homogeneous composition, *i.e.*, less variation in tree height and density at the plot level. In contrast, the broadleaved forest dominated by Sawtooth oak and Chinese sweet gum, have more variability in tree height and density and as a result more variable LiDAR-derived metrics within plots. Most of the selected models comprised three independent variables (except the foliage biomass models, which include four). All of the models contained at least one variable related to canopy height (h_{cv}, h_{10}, h_{25}, h_{75}) and most of them had one variable related to canopy cover (CC_{mean}, CC_{2m}) derived from the first returns, indicating sensitivity and high potential for predicting biomass components. LIDAR data allows estimation of the 3-D structural characteristics of the forest canopy, which provides a good foundation for strong relationships between the LIDAR-derived metrics and forest biomass. Since most LiDAR returns are from dominant trees, the LiDAR upper percentile heights (e.g., h_{75}) likely represent the height of overstory trees. The inclusion of variables which describe height variations (e.g., h_{cv}) and lower percentile heights (e.g., h_{10} and h_{25}), account for intermediate tree crowns in the midstory and suppressed trees in the understory, therefore potentially improving the characterizing of vertical structure.

Fu *et al.* (2011) [33] estimated above-ground biomass components using the lower hit density (distance=1.08 m) small-footprint LiDAR data in 78 smaller circle plots (radius = 7.5m or 15m) in southern China. In similar forest types, 43% and 68% of the variability in above-ground biomass was explained within broadleaved and coniferous forests, respectively, but there was no significant correlation between any of the biomass components and the LiDAR-derived metrics in mixed forests. Both the increased plot size and the higher hit density may be reasonable for the more significant correlations found in this study. Studies predicting root (below-ground) biomass from small-footprint LiDAR especially in subtropical forest are rare. Næsset (2004) reported a $R^2 = 0.86$ for

below-ground biomass prediction in boreal forests [19]. Pang and Li (2012) [34] fitted regression models for biomass components in temperate forests in northeastern China and explained 85% of the variability in root (below-ground) biomass in conifers, which is higher than reported in this study (53%–83%). The reason for the lower accuracy in our study may be explained by the complexity of subtropical forest systems with root biomass of some less common tree species calculated from generalized allometric equations. Additionally, in complex multilayered sub-tropical forests, especially the old-growth forests, there is generally far greater variability in stem height and density compared to boreal and temperate forests. Further research should be done to apply state-of-the art allometric equations for below-ground biomass collected from large samples from a greater number of sites to cover a wide variety of site qualities. The development of these equations is underway as part of a nationwide initiative.

5. Conclusions

Overall, small-footprint LiDAR can be used efficiently to estimate the amount of biomass components in subtropical forests. Forest type-specific models represent an improved performance than the general model. Results vary between forest types; in this study, the most accurate results were obtained for coniferous stands in homogeneous forest conditions. In addition, some of canopy height and cover metrics applied (previously used in temperate, boreal and tropical forests) appear to be useful in estimating biomass in subtropical forests such as those in southeastern China. These methods may prove important in supporting management and polices addressing carbon stocks in China and elsewhere.

Acknowledgments

This research is supported by a research grant named "Adaptation of Asia–Pacific Forests to Climate Change" (Project #APFNet/2010/PPF/001) funded by the Asia–Pacific Network for Sustainable Forest Management and Rehabilitation. We also acknowledge grants from a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The authors gratefully acknowledge the foresters in Yushan National Forest Park for their assistance with data collection and sharing their rich knowledge and working experiences of the local forest ecosystems.

Author Contributions

Lin Cao designed the study, analyzed the data and wrote the paper. Nicholas C. Coops and John Innes helped in project design, paper writing and analysis. Jinsong Dai and Guanghui She undertook field works and helped with data analysis.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Dixon, R.K.; Brown, S.; Houghton, R.A.; Solomon, A.M.; Trexler, M.C.; Wisniewski, J. Carbon Pools and Flux of Global Forest Ecosystems. *Science* **1994**, *263*, 185–190.
2. Pancel, L. *Tropical Forestry Handbook*; Springer-Verlag: Berlin, Germany, 1993.
3. Fang, J.; Chen, A. Dynamic Forest Biomass Carbon Pools in China and Their Significance. *Acta Bot. Sin.* **2001**, *43*, 967–973.
4. Wang, X.; Feng, Z.; Ouyang, Z. The Impact of Human Disturbance on Vegetative Carbon Storage in Forest Ecosystems in China. *For. Ecol. Manag.* **2001**, *148*, 117–123.
5. Lambert, M.C.; Ung, C.H.; Raulier, F. Canadian National Tree Aboveground Biomass Equations. *Can. J. For. Res.* **2005**, *35*, 1996–2018.
6. Saatchi, S.; Halligan, K.; Despain, D.G.; Crabtree, R.L. Estimation of Forest Fuel Load from Radar Remote Sensing. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1726–1740.
7. Gibbs, H.K.; Brown, S.; Niles, J.O.; Foley, J.A. Monitoring and Estimating Tropical Forest Carbon Stocks: Making REDD a Reality. *Environ. Res. Lett.* **2007**, *2*, 1–13.
8. Rosenqvist, A.; Shimada, M.; Igarashi, T.; Watanabe, M.; Tadono, T.; Yamamoto, H. Support to Multi-National Environmental Conventions and Terrestrial Carbon Cycle Science by ALOS and ADEOS-II-the Kyoto & Carbon Initiative. In Proceedings of the IGARSS 2003, IEEE International Geoscience and Remote Sensing Symposium, VOLS I-VII, Learning from Earth's Shapes and Sizes, New York, NY, USA, 21–25 July 2003; pp. 1493–1495.
9. Richards, P.W. *The Tropical Rain Forest: An Ecological Study*; Cambridge University Press: Cambridge, UK, 1996.
10. Lefsky, M.A.; Cohen, W.B.; Parker, G.G.; Harding, D.J. LiDAR Remote Sensing for Ecosystem Studies. *BioScience* **2002**, *52*, 19–30.
11. Dubayah, R.; Drake, J. LiDAR Remote Sensing for Forestry. *J. For.* **2000**, *98*, 44–46.
12. Patenaude, G.; Hill, R.A.; Milne, R.; Gaveau, D.L.A.; Briggs, B.B.J.; Dawson, T.P. Quantifying Forest above Ground Carbon Content Using LiDAR Remote Sensing. *Remote Sens. Environ.* **2004**, *93*, 368–380.
13. Lim, K.; Treitz, P.; Baldwin, K.; Morrison, I.; Green, J. LiDAR Remote Sensing of Biophysical Properties of Tolerant Northern Hardwood Forests. *Can. J. Remote Sens.* **2003**, *29*, 658–678.
14. Naesset, E.; Gobakken, T. Estimation of above- and below-Ground Biomass Across Regions of the Boreal Forest Zone Using Airborne Laser. *Remote Sens. Environ.* **2008**, *112*, 3079–3090.
15. Huang, W.L.; Sun, G.Q.; Dubayah, R.; Cook, B.; Montesano, P.; Ni, W.J.; Zhang, Z.Y. Mapping Biomass Change after Forest Disturbance: Applying LiDAR Footprint-Derived Models at Key Map Scales. *Remote Sens. Environ.* **2013**, *134*, 319–332.
16. Drake, J.B.; Knox, R.G.; Dubayah, R.O.; Clark, D.B.; Condit, R.; Blair, J.B.; Hofton, M. Above-Ground Biomass Estimation in Closed Canopy Neotropical Forests Using LiDAR Remote Sensing: Factors Affecting the Generality of Relationships. *Glob. Ecol. Biogeogr.* **2003**, *12*, 147–159.

17. Saatchi, S.S.; Harris, N.L.; Brown, S.; Lefsky, M.; Mitchard, E.T.A.; Salas, W.; Zutta, B.R.; Buermann, W.; Lewis, S.L.; Hagen, S.; Petrova, S.; White, L., Silman, M.; Morel, A. Benchmark Map of Forest Carbon Stocks in Tropical Regions across Three Continents. In Proceedings of the National Academy of Sciences of the United States of America, Irvine, CA, USA, December 2011; doi:10.1073/pnas.1019576108.
18. Tsui, O.W.; Coops, N.C.; Wulder, M.A.; Marshall, P.L.; McCardle, A. Using Multi-Frequency Radar and Discrete-Return LiDAR Measurements to Estimate Above-Ground Biomass and Biomass Components in a Coastal Temperate Forest. *ISPRS J. Photogr. Remote Sens.* **2012**, *69*, 121–133.
19. Næsset, E. Estimation of Above- and Below-Ground Biomass in Boreal Forest Ecosystems. In *Laser-Scanners for Forest and Landscape Assessment*; ISPRS Archives; Thies, M., Koch, B., Spiecker, H., Weinacker, H., Eds.; International Society for Photogrammetry and Remote Sensing: Freiburg, Germany, 2004; pp. 145–148.
20. Pan, B.H.; Zhang, Z.M.; Cao, T.R. Natural Vegetation of Yusan National Forest Park in Jiangsu Province, China. *J. Cent. South. Univ. For. Technol.* **2007**, *27*, 123–128.
21. Bastide, F.; Alexander, K. Satellite-Based Augmentation Systems (SBAS) Combined Performance. Presented at International Committee on GNSS (ICG-4), Working Group A, St. Petersburg, Russia, 14–18 September 2009.
22. Feng, Z.W.; Wang, X.K.; Wu, G. *Biomass and Productivity of Chinese Forest Ecosystem*; Science Press: Beijing, China, 1999.
23. Anonymous. *RIEGL Software Manual: RiANALYZE for RIEGL Airborne Laser Scanners LMS-Q680*, 5th ed.; RIEGL: Horn, Austria, 2008.
24. Naesset, E. Predicting Forest Stand Characteristics with Airborne Scanning Laser Using a Practical Two-Stage Procedure and Field Data. *Remote Sens. Environ.* **2002**, *80*, 88–99.
25. Naesset, E. Effects of Different Flying Altitudes on Biophysical Stand Properties Estimated from Canopy Height and Density Measured with a Small-Footprint Airborne Scanning Laser. *Remote Sens. Environ.* **2004**, *91*, 243–255.
26. Ferster, C.J.; Coops, N.C.; Trofymow, J.A. Aboveground Large Tree Mass Estimation in a Coastal Forest in British Columbia Using Plot-Level Metrics and Individual Tree Detection from LiDAR. *Can. J. Remote Sens.* **2009**, *35*, 270–275.
27. Kim, Y.; Yang, Z.Q.; Cohen, W.B.; Pflugmacher, D.; Lauver, C.L.; Vankat, J.L. Distinguishing between Live and Dead Standing Tree Biomass on the North Rim of Grand Canyon National Park, USA Using Small-Footprint LiDAR Data. *Remote Sens. Environ.* **2009**, *113*, 2499–2510.
28. Kraus, K.; Pfeifer, N. Determination of Terrain Models in Wooded Areas with Airborne Laser Scanner Data. *ISPRS J. Photogr. Remote Sens.* **1998**, *53*, 193–203.
29. Naesset, E.; Bjerknes, K.O. Estimating Tree Heights and Number of Stems in Young Forest Stands Using Airborne Laser Scanner Data. *Remote Sens. Environ.* **2001**, *78*, 328–340.
30. Sprugel, D.G. Correcting for Bias in Log-Transformed Allometric Equations. *Ecology* **1983**, 64, 209–210.
31. Weisberg, S. *Applied Linear Regression*, 2nd ed.; Wiley: New York, NY, USA, 1985.
32. Kutner, M.; Nachtsheim, C.; Neter, J.; Li, W. *Applied Linear Statistical Models*, 5th ed.; McGraw-Hill/Irwin: New York, NY, USA, 2004.

33. Fu, T.; Pang, Y.; Huang, Q.F.; Liu, Q.W.; Xu, G.C. Prediction of Subtropical Forest Parameters Using Airborne Laser Scanner. *J. Remote Sens.* **2011**, *15*, 1092–1104.
34. Pang, Y.; Li, Z.Y. Inversion of Biomass Components of the Temperate Forest Using Airborne LiDAR Technology in Xiaoxing' an Mountains, Northeastern of China. *Chin. J. Plant Ecol.* **2012**, *36*, 1095–1105.