Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Upscaling coniferous forest above-ground biomass based on airborne LiDAR and satellite ALOS PALSAR data

Wang Li
Zheng Niu
Zengyuan Li
Cheng Wang
Mingquan Wu
Shakir Muhammad

Wang Li, Zheng Niu, Zengyuan Li, Cheng Wang, Mingquan Wu, Shakir Muhammad, "Upscaling coniferous forest above-ground biomass based on airborne LiDAR and satellite ALOS PALSAR data," *J. Appl. Remote Sens.* **10**(4), 046003 (2016), doi: 10.1117/1.JRS.10.046003.

Upscaling coniferous forest above-ground biomass based on airborne LiDAR and satellite ALOS PALSAR data

Wang Li,[a,*] Zheng Niu,[a] Zengyuan Li,[b] Cheng Wang,[c] Mingquan Wu,[a] and Shakir Muhammad[a]

[a]Institute of Remote Sensing and Digital Earth, The State Key Laboratory of Remote Sensing Science, CAS Olympic S&T Park No. 20, Da Tun Road, P. O. Box 9718, Beijing 100101, China
[b]Chinese Academy of Forestry, Research Institute of Forest Resource Information Techniques, Wanshoushanhou, Beijing 100091, China
[c]Chinese Academy of Sciences, Institute of Remote Sensing and Digital Earth, Laboratory of Digital Earth Sciences, No. 9 Dengzhuang South Road, Beijing 100094, China

Abstract. Forest above-ground biomass (AGB) is an important indicator for understanding the global carbon cycle. It is hard to obtain a geographically and statistically representative AGB dataset, which is limited by unpredictable environmental conditions and high economical cost. A spatially explicit AGB reference map was produced by airborne LiDAR data and calibrated by field measurements. Three different sampling strategies were designed to sample the reference AGB, PALSAR backscatter, and texture variables. Two parametric and four nonparametric models were established and validated based on the sampled dataset. Results showed that random stratified sampling that used LiDAR-evaluated forest age as stratification knowledge performed the best in the AGB sampling. The addition of backscatter texture variables improved the parametric model performance by an R^2 increase of 21% and a root-mean-square error (RMSE) decrease of 10 $\mathrm{Mg\,ha^{-1}}$. One of the four nonparametric models, namely, the random forest regression model, obtained comparable performance ($R^2 = 0.78$, $\mathrm{RMSE} = 14.95\ \mathrm{Mg\,ha^{-1}}$) to the parametric model. Higher estimation errors occurred in the forest stands with lower canopy cover or higher AGB levels. In conclusion, incorporating airborne LiDAR and PALSAR data was proven to be efficient in upscaling the AGB estimation to regional scale, which provides some guidance for future forest management over cold and arid areas. © *2016 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JRS.10.046003]

Keywords: airborne LiDAR; ALOS PALSAR; forest aboveground biomass; forest age; parametric and nonparametric models.

Paper 16369 received May 15, 2016; accepted for publication Sep. 27, 2016; published online Oct. 17, 2016.

1 Introduction

Above-ground biomass (AGB) is an important indicator for understanding the global carbon cycle. Remote sensing is gaining more importance in the mapping of forest AGB over regional and global scales due to its strong capacity for large-area monitoring with high-frequency visiting. It facilitates the quantification of carbon stock especially in inaccessible forest environments (e.g., tropical rain forest) where traditional field inventory can be hard to conduct.[1,2] Currently, many AGB estimation models are established based on the satisfactory response of forest structure to the remotely sensed spectral variables and synthetic aperture radar (SAR) backscatter. The SAR backscatter at different wavelengths (e.g., C and L band) outperformed other spectral variables due to its high-penetration depth through the forest canopy. However, the SAR backscatter often saturates at the AGB level of 60 to 100 $\mathrm{Mg\,ha^{-1}}$ among various forest types.[3–5] Thus, backscatter ratios and texture variables were added into the estimation models as these variables do not saturate quickly and can reduce topographic bias.[6–8]

---

*Address all correspondence to: Wang Li, E-mail: lwwhdz@sina.com

Airborne light detection and ranging (LiDAR) is another active remote sensing tool that shows a great deal of promise in overcoming the saturation problem. Forest parameters such as canopy height,[9] crown size,[10] canopy cover (CC), leaf area index (LAI),[11–13] and AGB[14–16] have been successfully estimated by airborne LiDAR. However, it is difficult to acquire a large-area of airborne LiDAR data due to its small spatial coverage and high cost. Therefore, an increasing number of studies incorporated the LiDAR-derived parameters (e.g., canopy height, CC, and AGB) with satellite image variables, which proved that LiDAR-derived parameters can be reliable reference data for large-area mapping of forest parameters.[17–20] Nevertheless, the incorporation should be based on the satisfactory correlation of LiDAR-derived parameters to the satellite image variables. Previous studies showed that the LiDAR-derived CC was closely related to the SPOT vegetation index and the field measurements.[21] In addition, high correlations were also found between the LiDAR-derived AGB and SPOT-derived texture variables[22] and SAR backscatter.[18,23] The response of LiDAR-derived forest attributes to the PALSAR backscatter was analyzed based on a nation-wide LiDAR dataset with a low point density in Denmark.[23] The spatial scales were fully investigated in their study and the backscatter was used to build the prediction models. However, the response of LiDAR-derived AGB to the SAR backscatter may vary in different forest areas according to many factors such as forest species, landscape pattern, data quality, and environmental conditions. For instance, point density is a very important factor that can influence the precision of LiDAR metrics and AGB estimation accuracy.[24] The point density can be affected not only by the acquisition pattern of LiDAR data during the airborne flight but also by the heterogeneity of the forest structure.

Based on the spatially explicit LiDAR-derived forest attributes, a large quantity of training dataset can be sampled through different sampling strategies.[18,22] Nevertheless, the performance of different sampling strategies may vary with the change of study site and data source. Therefore, it is necessary to compare the performance of different sampling strategies before modeling. Generally, a greater range of data variability can be obtained from random stratified sampling where prior knowledge for stratification is required.[25] For the AGB estimation, parametric models (e.g., stepwise multiple regression models) have been regarded as the mainstream estimation approaches.[26,27] As more remotely sensed variables appeared, the dimensionality of predictors increased, while multicollinearity and data redundancy also increased. In order to take full use of the high-dimensionality dataset, nonparametric models such as support vector machine (SVM) model,[28] Gaussian process (GP) model,[28] K-nearest neighbor (K-NN) model,[15,29] and random forest (RF) model[29,30] have been applied to AGB estimations. The RF model is an ensemble learning technique by combining a large set of decision trees with each tree built using a deterministic algorithm.[30] Furthermore, geostatistical models were also explored to estimate forest attributes by incorporating LiDAR-derived forest attributes with satellite data.[18,22,31] However, it is difficult to compare these models based on one dataset as many nonparametric models require a large training dataset, whereas the ground-based dataset in most cases is very small.

Therefore, this study aimed to incorporate airborne LiDAR and ALOS PALSAR data to estimate the AGB of a coniferous forest in a cold and arid environment. In order to achieve the objective, a first attempt was made by involving LiDAR-evaluated forest age into AGB sampling. A small-area of an AGB reference map was produced by LiDAR data and calibrated by field measurements. The response of LiDAR-derived AGB to PALSAR backscatter was comprehensively analyzed, which was further applied to mapping the AGB at a regional scale. The incorporation procedure may provide some insights into how to take full use of the two active remote sensing tools in large-scale mapping of forest AGB over cold and arid regions.

2 Materials

2.1 *Study Site*

The study area is a water conservation forest site called the Dayekou area, which is located in the Heihe River Basin, northwest China (Fig. 1). The Heihe River basin is the second largest inland river basin in China, located between 97°24′-102°10′E and 37°41′-42°42′N.[15] It covers three

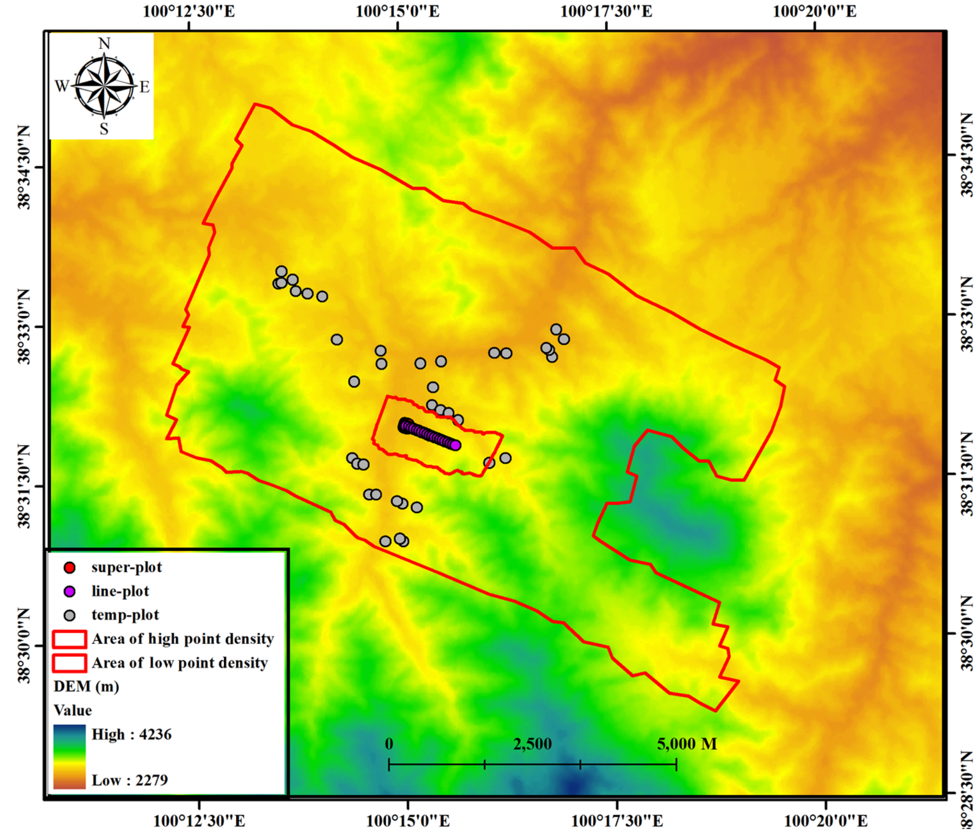


Fig. 1 The field plots, spatial extent of airborne LiDAR overlaid on the DEM.

vegetative climate zones and significantly affects soil/water conservation and biodiversity protection. The Dayekou area covers an area of 130 km^2 with an elevation of 2500 to 3300 m above sea level. The site was a core experimental site selected by the Watershed Allied Telemetry Experimental Research (WATER) project where simultaneous airborne satellite-borne remote sensing observations and ground-based measurements were carried out.[32,33] The climate at this site is a temperate continental mountainous climate with cold and dry weather in winters and drastic diurnal variations of temperature in summers.[15] The precipitation is mainly concentrated in the summer with a mean annual value of ~350 to 495 mm. The main species in the study area is *Picea crassifolia*, occupying 95% of the whole forestland over the study site.[34]

2.2 *LiDAR Data*

Small footprint full-waveform LiDAR data were acquired for the Dayekou area on June 23, 2008. The RIEGL LiteMapper-5600 airborne laser scanning system was used and configured to emit laser pulses of 3.5 ns at 50 kHz in the near-infrared band (1550 nm) with a beam divergence fixed with 0.5 mrad and a nominal height above ground of 700 to 800 m.[35] The average footprint diameter was 38 cm. The horizontal and vertical accuracies of the instrument were 0.1 and 0.3 m, respectively. For each pulse, the transmitted and returned pulses were recorded in the full-waveform data. Discrete point clouds provided by the flight vendor included two parts: one with low density [~1.89 pts m^{-2}] and another with high density [~3.43 pts m^{-2}] (Fig. 1). The high-density part (denoted as HDP, 1.7 km^2) was obtained by repetitive flights, and the spatial extent of the low-density part (denoted as LDP, 55 km^2) covered that of HDP.

In this study, only the original full-waveform LiDAR in HDP was used. A waveform decomposition approach was applied to decompose the original waveform data into another group of point clouds to acquire a higher point density.[36] It approximates the LiDAR waveform as a Gaussian mixture model. A stepwise strategy was used to implement waveform decomposition by defining the maximum iteration number and error expectations before decomposition.[36,37] The final point density of HDP was increased to 6.22 pts m^{-2} higher than those provided by the vendor. Then the digital elevation model (DEM), digital surface model (DSM), and canopy height model (CHM) with a spatial resolution of 0.5 m were derived from the decomposed point clouds. The heights of point clouds were normalized by the produced DEM. CHM pixels with values greater or equal to 1.5 m were classified as forest and used for subsequent sampling. More details on LiDAR data processing can be found in our previous study.[38]

2.3 *ALOS PALSAR Data*

Advanced Land Observing Satellite (ALOS) Phased Array L-band SAR (PALSAR) is an active remote sensing satellite which was launched in January 2006. The single look complex (SLC) image of PALSAR used in this study was acquired on May 12, 2008, which consisted of HH (horizontal transmit and horizontal receive) and HV (horizontal transmit and vertical receive) backscattering polarizations. There were no significant land use changes between the LiDAR and SAR data acquisition. The original intensity image was processed to obtain the backscattering coefficient using the NEST software tool developed by the European Space Agency (ESA). The SLC image was radiometrically corrected to make the SAR intensity truly represent the radar backscatter of the reflecting surface using the calibration equation in the following equation:

$$\sigma^0 = 10 * \log_{10}(I^2 + Q^2) + G - C, \tag{1}$$

where σ^0 is the backscatter coefficient, and I and Q are the real and imaginary components of SLC image. G and C are the calibration constants which can be found in the header file. The G value is -83.2 for HH polarization and -80.2 for HV polarization.[34] The C value equals 32.0 for both polarizations. To compensate for speckle noise, the Gamma map filter was applied to replace the central pixel with the average value in a moving window (3×3 pixels). Multilooking processing was applied to the SLC image with two looks in the range direction and ten looks in the azimuth direction.

As the SAR images are acquired using a side-looking geometry, distortions such as foreshortening, layover, and shadowing often exist in the images. In order to remove these distortions, terrain correction was used to geo-code and orthorectify the SAR intensity image using the one-second resolution ASTER DEM (30 m) (download from Ref. 39) using the software ASF MapReady. During the terrain correction, the intensity image was resampled to a pixel size of 30 m using bilinear interpolation. Radiometric terrain correction was conducted to correct the impact of topography on the SAR backscatter by multiplying a correction factor (a function of the incidence angle under a flat surface assumption and the actual incidence angel) with the above-terrain-corrected images.[40] The pixel brightness of SAR images was adjusted with respect to the observation geometry defined by incidence angle, the slope, and aspect of the local terrain.[41]

2.4 *Field Measurement*

Three types of field plots ($n = 69$) including 16 plots grouped into one plot (super-plot), 19 plots along a linear transect (line-plot), and 34 temporal plots (temp-plot) over the LDP were measured during the flight campaign (Fig. 1). The super-plot was a 100 m × 100 m area with a slope of >20 deg. It was divided into 16 subplots measuring 25 m × 25 m. The line-plot was a sample line consisting of 19 subplots measuring 20 m × 20 m distributed along the direction of airborne flight. The distance between the subplots in the line-plot was 50 m. The center location of each subplot was positioned using a differential global positioning system (DGPS) station. The DGPS mobile station was placed within the open space of the standing forest to ensure the positioning accuracy. The individual trees were located by the survey total stations with equipment models of TOPCON GTS-602 and TOPCON GPT-7002. Tree height (H) and first branch height (FBH) were measured using a laser hypsometer; diameter at breast height (DBH) was obtained after measuring the perimeter of the tree stem at the height of 1.2 m using a tape. Crown size (D) for each individual tree was calculated as the average of the crown size in two orthotropic directions

Table 1 Basic statistics of the field measurements.

Field measurement	Min	Max	Mean	Std
Height (m)	2.50	26.10	14.31	9.38
DBH (cm)	3.10	26.50	8.82	5.14
FBH (m)	1.21	19.00	3.32	2.44
D (m)	0.55	18.05	3.28	1.42
AGB ($\mathrm{Mg\,ha^{-1}}$)	36.88	174.88	108.85	29.94
AGE (year)	14	109	37	22

Note: H, tree height; DBH, diameter at breast height; FBH, first branch height; D, crown size; Max, maximum value; Mean, mean value; Std, standard deviation.

that was measured using a ribbon tape. The basic statistics of the field measurements are shown in Table 1.

The biomass of different forest components was calculated based on their DBH and H according to the empirical relative growth equations that were obtained from field measurements in the Qilian Mountains[42]

$$\text{stock biomass} = 0.0478 \times [(\mathrm{DBH})^2 \times H]^{0.8665}, \tag{2}$$

$$\text{branch biomass} = 0.0061 \times [(\mathrm{DBH})^2 \times H]^{0.8905}, \tag{3}$$

$$\text{leaf biomass} = 0.2650 \times [(\mathrm{DBH})^2 \times H]^{0.4701}, \tag{4}$$

$$\text{fruit biomass} = 0.0342 \times [(\mathrm{DBH})^2 \times H]^{0.5779}. \tag{5}$$

The AGB for each individual tree was calculated as the sum of all the biomass components (stock, branch, leaf, and fruit). Plot biomass is calculated as the sum of tree AGB within each plot. The subplot AGB in the super-plot was calculated using the measured trees within a radius of 20 m around the plot's center to be consistent with the plot size of other types of plots. The AGB for the 35 field plots within HDP obtained a mean value of 108.85 $\mathrm{Mg\,ha^{-1}}$ and a standard deviation of 29.94 $\mathrm{Mg\,ha^{-1}}$ (Table 1), which were used to produce the AGB reference map (Sec. 3.3). The additional 34 temp-plots distributed outside HDP were used to validate the estimated spatial distribution of AGB over the whole study area.

A satisfactory relationship existed between the tree age (AGE) and DBH of *Picea crassifolia* [Eq. (6)] in Qilian Mountains.[43,44] Meanwhile, the field-measured DBH was found to be highly related to H based on the field measurements in this study [Fig. 2(a)]. Thus, the relationship between AGE and H can be approximately expressed as Eq. (7), which was further used to evaluate the forest age for the sampling stratification over the whole HDP (Sec. 3.4)

$$\mathrm{AGE} = 1.017 + 4.073 * \mathrm{DBH} \qquad R^2 = 0.944^{**}, \tag{6}$$

$$\mathrm{AGE} = 6.565 * H - 1.437. \tag{7}$$

3 Methods

3.1 *Crown Objects Segmentation*

In order to obtain forest age distribution over the whole HDP, watershed segmentation was conducted on the CHM to evaluate the forest age at crown object scale. The CHM was smoothed by a 3 m × 3 m (approximately to the mean of the field-measured crown size $\bar{D}$) Gaussian filter

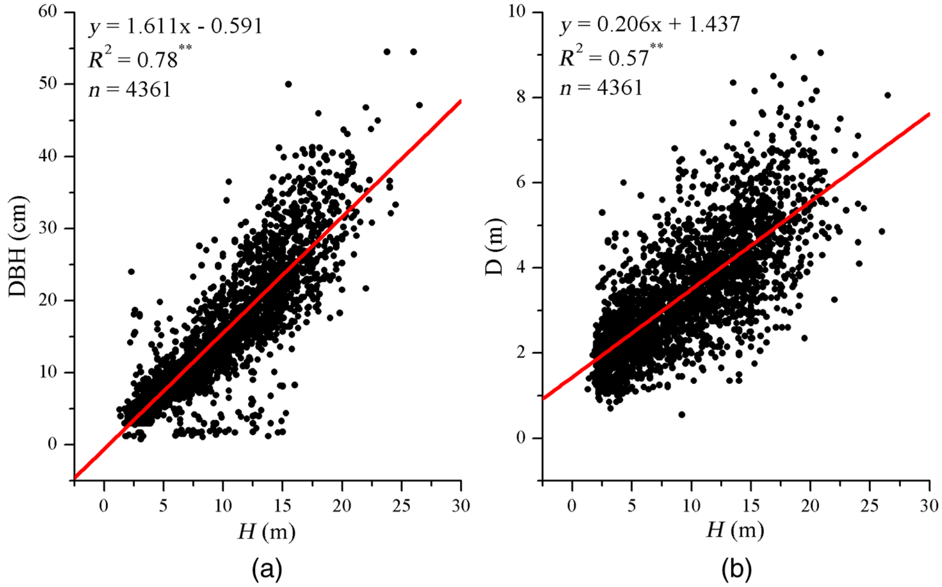


Fig. 2 Scatter plots for the field-measured (a) tree height (*H*) and DBH; (b) *H* and crown size (*D*). R^2 is the coefficient of determination. $^{**}p < 0.01$.

before segmentation. Afterward, the watershed algorithm was applied to the smoothed CHM.[45] It first inverted the CHM to a raster by converting peaks into depressions. Then the canopy basins were delineated on the depression raster by segmentation. Finally, the delineated basin raster was converted to a group of segments. A height threshold (1.5 m) was set to extract the tree crown objects from the obtained segments. Then the crown objects were transferred into polygonally shaped files. The highest pixel value obtained by zonal statistics within each crown polygon was assigned as the tree height and its location as the tree location. The age of each crown object was evaluated according to the local maximum canopy height according to Eq. (7). According to the evaluated age, we stratified the crown objects into three layers, namely, young (AGE ≤ 50 year), medium (50 year < AGE <= 100 year), and old (AGE > 100 year), which facilitated the subsequent analysis of response of AGB to SAR backscatter at different forest ages. In order to facilitate the sampling and spatial extrapolation of AGB over the whole study area, the segments of age layers with irregular shapes were attributed back into individual pixels (20 m). Therefore, all the variables and analysis in this study were conducted at the pixel level.

3.2 *Variables Generation*

Three groups of LiDAR metrics were calculated using the point clouds with normalized height for each field plot and the whole HDP, respectively. The LiDAR metrics are classified into canopy height metrics, canopy complexity metrics, and CC metrics. Canopy height metrics describe the height distribution of the canopy, including the maximum height ($H_{\max}$), mean height (H_{mean}), and percentile height (H_{p}). Canopy complexity metrics describe the height dispersion of the returns, including standard deviation of height (H_{std}), coefficient of height variation (H_{cv}), and variance of height (H_{var}). First echo cover index (FCI) was calculated to describe the fraction of canopy returns with respect to the total returns, which indicates the density, amount of foliage, and other woody components.[20] Subsequently, all LiDAR metrics over the HDP were calculated at a grid size of 20 m. The LiDAR metrics for all the field plots were used to produce the AGB reference map together with field measurements.

Two combined rasters, namely, backscatter ratio (Ratio) and difference (Diff), were generated based on the dual-polarization PALSAR backscatter. Ratio can reduce the systematic noise, whereas Diff is related to double bounce scattering mechanisms, which have been proven but do not saturate with AGB quickly and can reduce topographic bias.[6–8] Eight gray-level texture variables (mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation) were calculated for the corrected backscatter and corresponding combined raster

Table 2 Formulas of the gray level texture variables used in this study.[22]

Texture variables	Texture variables
Mean (ME) $= \sum_{i,j=0}^{N-1} iM_{i,j}$	Variance (Var) $= \sum_{i,j=0}^{N-1} iM_{i,j}(i - \mathrm{ME})^2$
Standard deviation (Std) $= \sqrt{\mathrm{Var}}$	Entropy (EN) $= \sum_{i,j=0}^{N-1} iM_{i,j}(-\ln M_{i,j})$
Dissimilarity (DI) $= \sum_{i,j=0}^{N-1} iM_{i,j}\lvert i - j\rvert$	Second moment (SM) $= \sum_{i,j=0}^{N-1} iM_{i,j}^2$
Homogeneity (HG) $= \sum_{i,j=0}^{N-1} i\frac{M_i}{1+(i-j)^2}$	Contrast (CT) $= \sum_{i,j=0}^{N-1} iM_{i,j}(i - j)^2$

Note: $M_{i,j}$, the normalized co-occurrence matrix such that SUM $(i, j = 0, N - 1)$ $(M_{i,j}) = 1$, i, j are the index of the element in the matrix; N, the number of elements in the matrix.

using a gray-tone spatial dependence matrix.[22] The matrix was calculated as an average of all the four possible directions (0, 45, 90, and 135 deg). The texture variables (Table 2) were calculated for each pixel in four different window sizes (e.g., 3×3, 5×5, 7×7, and 9×9 pixels) for the four PALSAR raster, because an appropriate window size is critical for the texture measurements. Usually, a small window can exaggerate the differences between the windows, whereas a large window will cause over-smoothing.[46] Therefore, there a total of 132 variables (2 backscatter + 2 combined variables + 128 texture variables) were obtained, which were further used as the predictor variables for the estimation models

$$\mathrm{Ratio} = \frac{\sigma_{\mathrm{HH}}}{\sigma_{\mathrm{HV}}} \tag{8}$$

$$\mathrm{Diff} = \sigma_{\mathrm{HH}} - \sigma_{\mathrm{HV}}. \tag{9}$$

3.3 *Production of Above-Ground Biomass Reference Map*

A spatially explicit AGB map over the HDP was produced using plot-level LiDAR metrics and field AGB. Only the LiDAR data of HDP were used to produce the AGB reference map due to its relative higher point density, which can ensure the precision and reliability of reference maps. In fact, other vegetation parameters aside from AGB that were derived from airborne LiDAR have also been used as a reliable reference data source for large-area mapping, e.g., canopy height and CC.[21,47] The mainstream approach in current AGB estimation, namely, stepwise multiple linear regression (SMLR) model, was established.[14,22,27,47–50] Both the field AGB and LiDAR metrics were log-transformed prior to regression and were back-transformed after regression. The 35 field AGB plots within the HDP were used as training and validation data. Considering the relatively small population of the field dataset within HDP, a leave-one-out cross-validation method (LOOCV) was used to assess the prediction accuracy.[51] The LOOCV approach made the field data from each plot act simultaneously as training and validation samples. Coefficients of final prediction models were obtained by averaging all the iterations, and goodness-of-fit of the model was assessed by averaging statistical estimators during LOOCV.[52]

3.4 *Sampling Design*

Three sampling strategies were designed to extract the independent and dependent variables from HDP. The three sampling strategies included: 1) regular sampling with a distance interval of 50 m (hereafter, denoted as S1); 2) random sampling, S2; and 3) random stratified sampling, S3. The LiDAR-evaluated forest age over HDP was used as the prior knowledge for the stratification of S3, dividing the whole HDP into three stratums (young, medium, old). The number of forest pixels for each stratum was counted, which was further used to obtain the sampling proportion for each stratum by dividing the total number of forest pixels. There were a total of 400 samples extracted by each sampling strategy which was determined according to the spatial extent of HDP. The reference AGB and corresponding PALSAR variables were extracted using the obtained sampling points and were subsequently used to analyze the response of the reference AGB to the PALSAR variables.

3.5 *Estimation Models and Validation*

SMLR models were built using the sampled AGB and SAR predictor variables. The predictor variables were split into two groups, namely, one without texture variables (G1) and another with texture variables (G2). Thus, two types of SMLR models (denoted as SMLR-I and SMLR-II) were established using G1 and G2. Comparison between these two models will help to test the necessity of adding the texture information into the models based on the dataset in this study. Since multicollinearity and over-fitting among the predictor variables may exist, four commonly used nonparametric models, namely, SVM model,[28] GP model,[28] K-NN model,[15,29] and random forest (RF) model,[29,30] were also established based on G2. These nonparametric models have been used to estimate many forest attributes including forest AGB, LAI, and canopy fuel.[15,28,53] More details on the mechanism of these models can be found in the above-cited literature and corresponding programming functions have been implemented in different library packages in the statistical software R.[54] In this study, most of the input parameters for the training models were settled as default (e.g., SVM and GP models). After trial and error, the number of neighbors (K) for the K-NN model input was set to 5. The random forest parameters ($m$try and $n$tree) were optimized to select a group of predictors with the best predictive power using a backward feature selection during the regression.[30] In this study, the random forest parameters $n$tree and $m$try were finally optimized to 1000 and 20, respectively. The sampled dataset was randomly split into two parts: 3/4 for model training ($n = 300$) and the left 1/4 for validation ($n = 100$). Statistical estimators were calculated to evaluate model accuracy including coefficient of determination (R^2), root-mean-square error (RMSE), relative root-mean-square error (rRMSE), and mean of absolute error (MAE). The model with highest accuracy was applied on the selected wall-to-wall predictor images. The estimation results were validated within and outside HDP using the independent reference AGB ($n = 100$) and the field-measured AGB from 34 temp-plots, respectively. To examine whether there were statistically significant differences among the estimated results from different models, a one-way analysis of variance (ANOVA) at the $p < 0.05$ level was conducted with the use of the Duncan *post hoc* test.

4 Results

4.1 *Sampled Above-Ground Biomass and Synthetic Aperture Radar Backscatter*

The LiDAR metrics selected by the SMLR model during the production of the AGB reference map included H_{mean}, H_{var}, H_{70}, and FCI. The SMLR model suggested that the selected LiDAR metrics can explain 78% of the variance of field-measured AGB with an RMSE of 24.21 $\mathrm{Mg\,ha^{-1}}$. Thus, the AGB reference map achieved satisfactory reliability and comparable accuracy to previous studies.[18,22] The spatial distribution of the sampled points from the three sampling strategies (Fig. 3) indicated that the S2 sampling obtained the lowest geographic representativeness compared with S1 and S3 since obvious clustering occurred in the northeast part of the HDP.

The sampled AGB and backscatter from the three strategies obtained different variation ranges (Table 3). Specifically, S3 obtained the largest range of AGB variability followed by S1 and S2. Similar variation ranges were found in the sampled HH and HV backscatter values. The mean and standard deviation of the sampled HV backscatter were always lower than those of the HH backscatter. The significantly variations at low AGB levels contributed to the dynamic range of HH and HV backscatter as shown in Fig. 4. In addition, higher correlations between the sampled AGB and backscatter were obtained by S3 followed by S2 and S1. These indicated that the representativeness of the sampled AGB and backscatter can be improved when the stratification of forest age was considered.

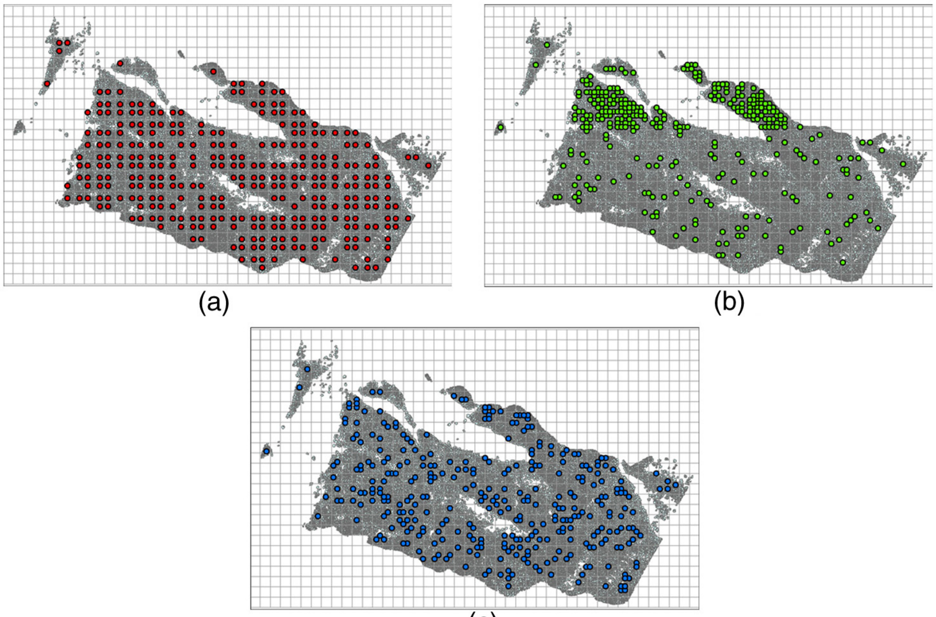

(a)
(b)
(c)

Fig. 3 The spatial distribution of the sampled points from (a) regular sampling S1; (b) random sampling S2; and (c) random stratified sampling S3.

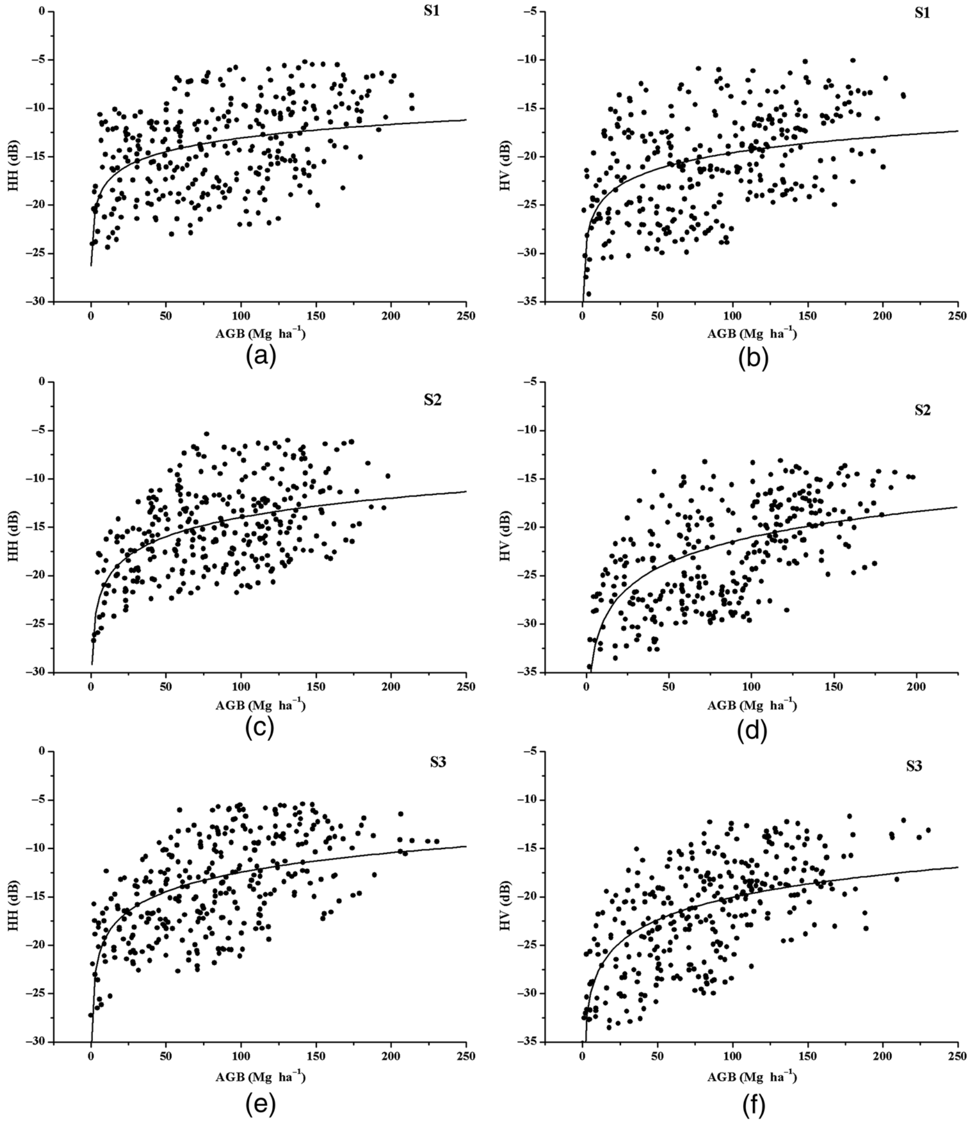
4.2 *Relationship Between Sampled Above-Ground Biomass and Synthetic Aperture Radar Backscatter*

The sampled AGB from the three sampling strategies showed moderate but significant correlation to the backscatter with a Pearson's correlation coefficient (R) below 0.7 (Table 3). The HV backscatter always showed higher correlation to AGB than HH backscatter. The highest R was found between the AGB and HV sampled by S3. Nevertheless, the saturation point of backscatter

Table 3 Basic statistics of the sampled variables from three sampling strategies.

Variable\statistic	Min	Max	Mean	Std	Range	R
S1						
AGB (Mg ha^{-1})	0.92	213.98	86.71	52.53	213.06	—
HH (dB)	−24.34	−3.32	−13.68	4.90	21.02	0.42
HV (dB)	−34.20	−9.16	−20.63	5.19	24.04	0.46
S2						
AGB (Mgha^{-1})	1.78	197.87	84.67	45.81	196.09	—
HH (dB)	−26.72	−5.37	−15.10	4.45	21.35	0.52
HV (dB)	−34.41	−10.66	−22.52	5.43	23.75	0.57
S3						
AGB (Mg ha^{-1})	0.04	230.38	84.67	49.33	230.34	—
HH (dB)	−27.23	−5.29	−13.76	4.93	21.94	0.57
HV (dB)	−35.16	−10.68	−21.08	5.85	25.48	0.63

Note: S1, regular sampling; S2, random sampling; S3, random stratified sampling; R, Pearson's correlation coefficient with respect to AGB.



(a) (b) (c) (d) (e) (f)

Fig. 4 The scatter plots of the sampled AGB and PALSAR backscatter from the three sampling strategies: (a) and (b) S1, (c) and (d) S2, and (e) and (f) S3.

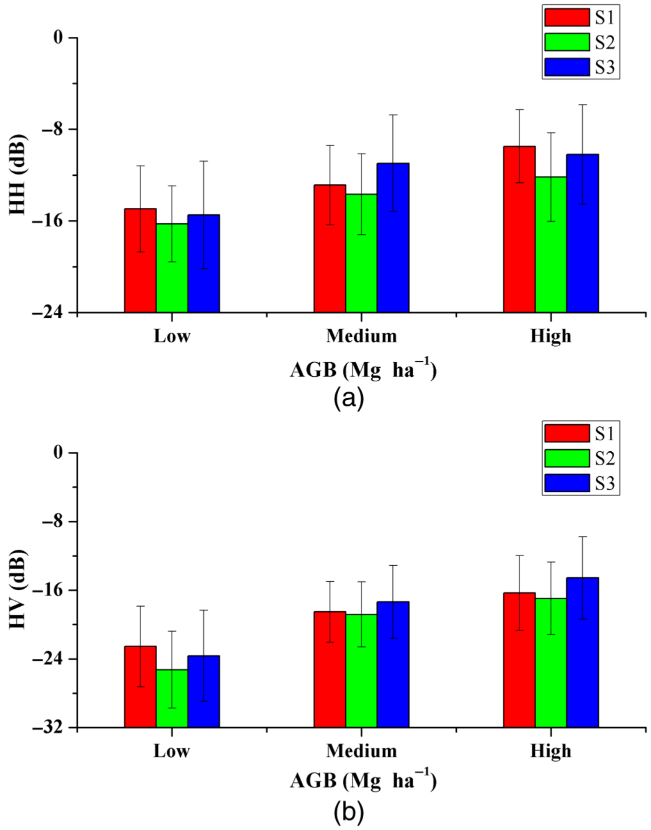
signal to AGB also occurred based on our dataset as many previous studies reported (Fig. 4). The saturation point of backscatter signals and corresponding AGB values were quantified using the parametric models demonstrated in the previous study.[55] The backscatter sampled by the three sampling strategies tended to saturate between −12 and −14 dB at HH and between −21 and −23 dB at HV. The saturation point of AGB was around 63 Mg ha^{-1} at HH and around 68 Mg ha^{-1} at HV. Considering the saturation problem, simple logarithmic models were fitted based on the sampled AGB-backscatter datasets (Table 4).

Figure 5 shows the backscatter for different AGB levels sampled from the three sampling strategies. The low AGB level ($\leq$50 Mg ha^{-1}) corresponded to vegetated areas where ground contributed the most to the backscatter and rapid attenuation of the signal occurred. The medium AGB level (50 Mg ha^{-1} < AGB $\leq$ 100 Mg ha^{-1}) corresponded to denser vegetated areas where volume scattering contributed the most to the signal. The high AGB level (>100 Mg ha^{-1}) corresponded to areas where signal saturation appeared. Generally, higher AGB corresponded to higher backscatter coefficients at both HH and HV polarizations. Slight variations of mean backscatter were found among the three sampling strategies. However, the HV backscatter showed higher standard deviations at different AGB levels than HH backscatter. Additionally, the backscatter sampled by S3 obtained larger standard deviations than the other two strategies especially

Table 4 The simple logarithm relationship between the sampled AGB and PALSAR backscatter, the saturation point of backscatter signal, and corresponding AGB point. $^{**}p < 0.01$.

	Fitted models	R^2	Saturation point (dB)	AGB point ($\mathrm{Mg\,ha^{-1}}$)
S1				
HH	$\mathrm{HH} = 2.193 * \ln(\mathrm{AGB}) - 22.80$	0.18^{**}	−13.68	63.96
HV	$\mathrm{HV} = 2.470 * \ln(\mathrm{AGB}) - 30.90$	0.21^{**}	−22.40	68.29
S2				
HH	$\mathrm{HH} = 2.875 * \ln(\mathrm{AGB}) - 27.21$	0.27^{**}	−12.96	62.30
HV	$\mathrm{HV} = 3.560 * \ln(\mathrm{AGB}) - 37.73$	0.32^{**}	−22.42	69.15
S3				
HH	$\mathrm{HH} = 2.864 * \ln(\mathrm{AGB}) - 25.67$	0.33^{**}	−13.76	63.58
HV	$\mathrm{HV} = 3.375 * \ln(\mathrm{AGB}) - 35.58$	0.37^{**}	−21.08	67.58

at low AGB level which contributed most to the backscatter dynamic ranges. As S3 obtained a larger range of AGB variability, variables sampled by S3 were finally selected as the basic dataset for the subsequent analysis. Aside from AGB and the backscatter, corresponding combined variables (Diff and Ratio) and the texture variables were also used to establish the estimation models.



(a)

(b)

Fig. 5 Mean of the PALSAR backscatter (HH, HV) by AGB levels (low: $\mathrm{AGB} \leq 50\ \mathrm{Mg\,ha^{-1}}$; medium: $50\ \mathrm{Mg\,ha^{-1}} < \mathrm{AGB} \leq 100\ \mathrm{Mg\,ha^{-1}}$; high: $\mathrm{AGB} > 100\ \mathrm{Mg\,ha^{-1}}$) sampled from three sampling strategies (S1, S2, S3). The vertical error bars represent the standard deviation.

Table 5 The parametric models established using the dataset without texture information (G1) and the dataset with texture information (G2). $^{**}p < 0.01$.

Estimation models	R^2	RMSE	rRMSE	MAE
SMLR-I				
$\text{AGB} = 210.254 + 6.83 * \text{HV} + 2.793 * \text{Diff}$	0.44^{**}	38.29	0.66	75.91
SMLR-II				
$\text{AGB} = 339.628 + 6.794 * \text{HV} + 2.677 * \text{Diff} + 1.562 * \text{HV\_Mean\_Win7} - 4.74 * \text{HH\_Disimilarity\_Win7} - 32.779 * \text{HV\_Entropy\_Win9} - 34.0466 * \text{Ratio\_Homogeneity\_Win3} + 20.216 * \text{Diff\_Correlation\_Win9}$	0.65^{**}	28.30	0.38	24.43

Note: Win3, Win7, and Win9 represent that the window size of texture variables is 3, 7, and 9, respectively.

4.3 *Biomass Estimation and Validation*

Two SMLR models were established using the sampled G1 and G2 dataset (Table 5). The results showed that the SMLR-II model obtained higher accuracy than the SMLR-I model with a much higher R^2 of 0.65, lower RMSE of 28.30 Mg ha^{-1}, and lower MAE of 24.34 Mg ha^{-1}. The robustness of SMLR-II is also higher than that of SMLR-I with a lower rRMSE ($0.38 < 0.66$). It indicated that the texture information enriched the explanatory power of the independent dataset and could improve model accuracy. The cross polarization backscatter HV and the backscatter difference Diff were selected by both of the models. The texture variables with different texture indications calculated for the three window sizes (3, 7, and 9) were selected by SMLR-II. The variables selected by the stepwise regression showed a low level of multicollinearity with high tolerance and low variable inflation factor (VIF) (Table 6). The added texture variables directly contributed to the increase of R^2 and decrease of error evaluations with a p value <0.05.

The nonparametric models also obtained a satisfactory performance with high R^2 and low errors (Table 7). However, not all the nonparametric models outperformed the parametric model (SMLR-II). The one-way ANOVA ($p < 0.05$) results revealed that statistically significant differences were found among the six models in the estimation of AGB (Fig. 6). The AGB estimated by K-NN and GP model was significantly higher than the AGB estimated by

Table 6 Statistics of collinearity and significance of the independent variables in the parametric models.

Variable\statistics	Tolerance	VIF	p
SMLR-I			
HV	0.746	1.340	<0.001
Diff	0.741	1.530	<0.001
SMLR-II			
HV	0.727	1.376	<0.001
Diff	0.729	1.372	<0.001
HV_Mean_Win7	0.969	1.032	<0.001
HH_Disimilarity_Win7	0.973	1.028	0.018
HV_Entropy_Win9	0.909	1.101	0.030
Ratio_Homogeneity_Win3	0.902	1.109	0.042
Diff_Correlation_Win9	0.895	1.117	0.020

Table 7 The nonparametric models established using the dataset with texture information (G2). $^{**}p < 0.01$.

Estimation models	R^2	RMSE ($\mathrm{Mg\,ha^{-1}}$)	rRMSE	MAE ($\mathrm{Mg\,ha^{-1}}$)
SVM	0.65^{**}	28.96	0.33	20.23
GP	0.42^{**}	42.82	0.62	35.94
K-NN	0.58^{**}	37.95	0.56	34.67
RF	0.81^{**}	14.85	0.28	19.90

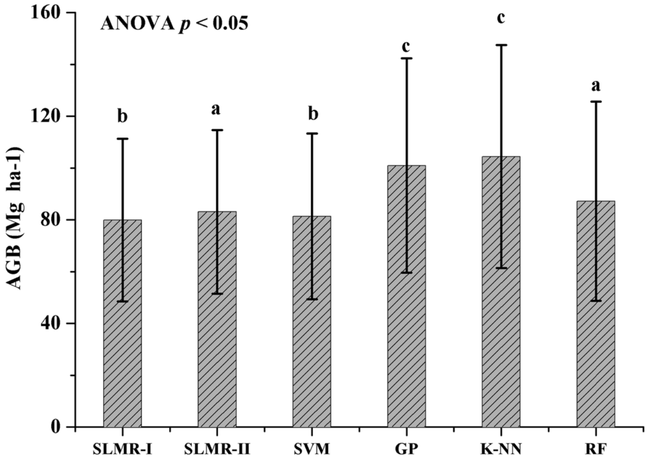


Fig. 6 The statistical significance among the six models for the estimated AGB. The histogram with bars indicated the means ± standard of error. Different letters indicate significant differences at the $p < 0.05$ level, according to one-way ANOVA.

other models. No statistically significant differences were found between the SLMR-II and RF model in the AGB estimation. The RF model achieved the highest model accuracy and robustness with an R^2 of 0.81 and rRMSE of 0.28. The final RF model also included the texture variables aside from the backscatter. The relative importance for the top 20 variables (out of 132) in the final random forest model is shown in Fig. 7. Finally, the RF model was selected as the model to map the spatial distribution of AGB over the study area.

The spatial distribution of AGB predicted by the RF model [Fig. 8(a)] was assessed using the field-measured AGB from 34 temp-plots distributed outside the HDP. The results showed that the predicted AGB was comparable to the field measurements, obtaining an R^2 of 0.62 and RMSE of 42.33 $\mathrm{Mg\,ha^{-1}}$ (Fig. 9). Thereafter, estimation error map for the HDP was generated by subtracting the reference map from the predicted map [Fig. 8(b)]. Approximately 58% of forest pixels obtained an MAE below 30 $\mathrm{Mg\,ha^{-1}}$ [Fig. 10(a)]. The overall MAE of AGB for all the forest pixels within HDP followed an approximately normal distribution [Fig. 10(a)]. The MAE tended to be lower in areas with higher CC which corresponded to the medium AGB levels [Fig. 10(b)]. Significant high MAE levels were found in areas which were dominated by young and old forest stands with low and high AGB intervals.

5 Discussion

5.1 *Utilization of LiDAR-Derived Prior Knowledge*

In this study, LiDAR-derived AGB was used as reference data for upscaling the estimation of AGB due to its high spatial resolution and comparable reliability to the field measurements. This

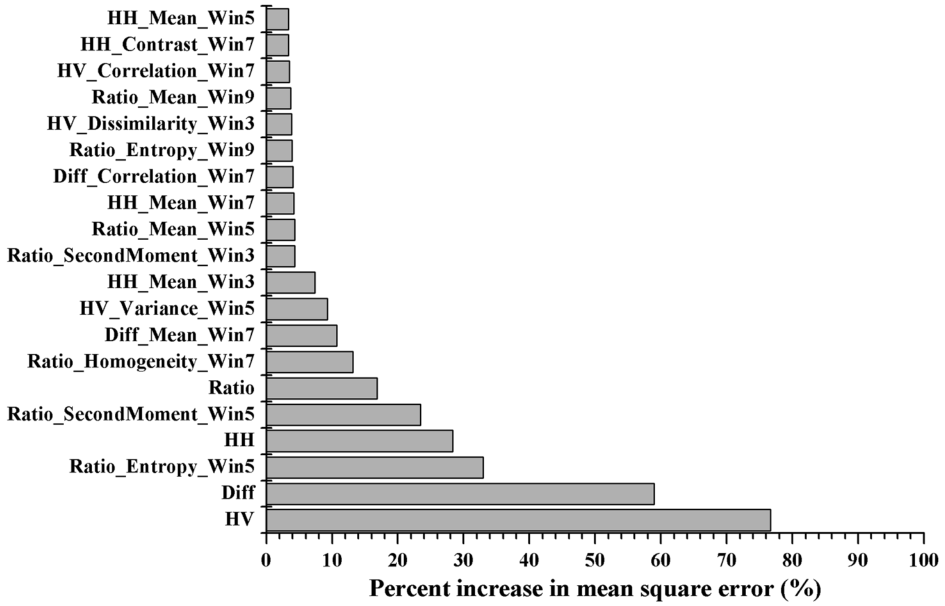
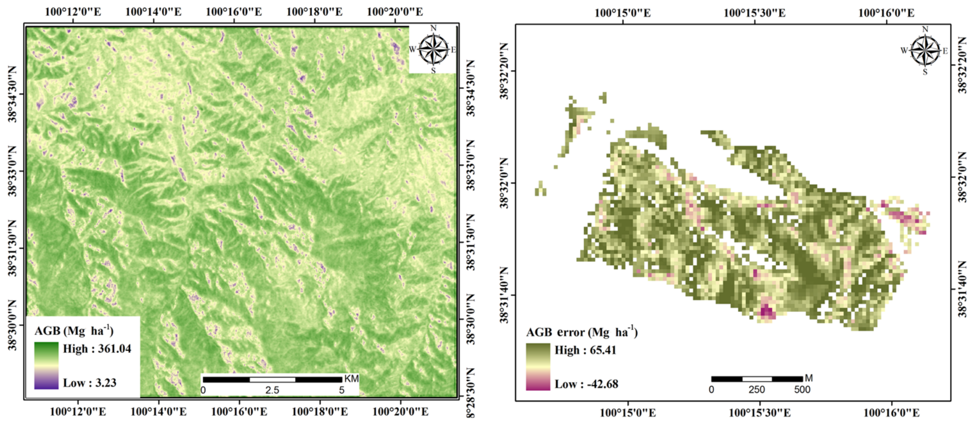


Fig. 7 The percent increase in mean square error without the variable in the random forest model for the top 20 predictor variables.



(a)

(b)

Fig. 8 The spatial distribution of (a) AGB over the whole study site estimated by RF model and (b) estimation error of AGB within HDP.

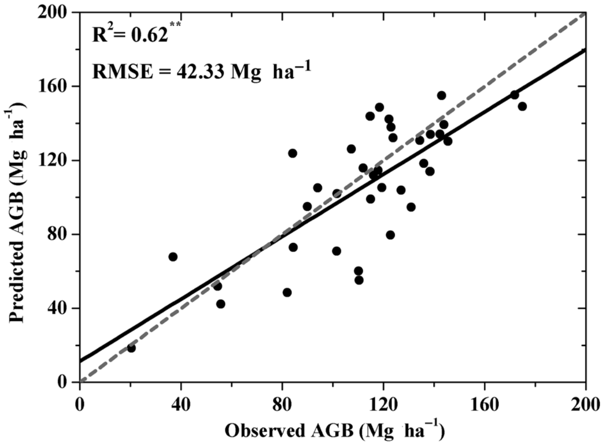


Fig. 9 The RF model predicted and field-measured AGB outside the HDP. $^{**}p < 0.01$.

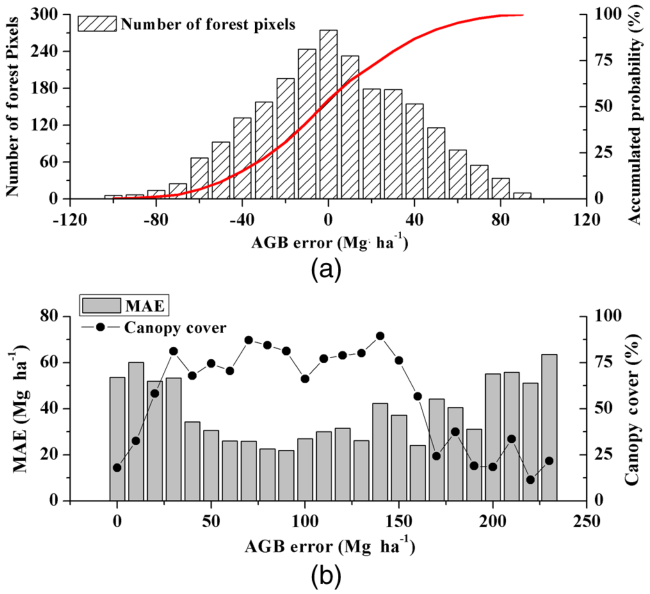


Fig. 10 The distribution of (a) number of forest pixels in different intervals of estimation errors and (b) MAE and CC in different AGB intervals.

can be helpful for cases where national forest inventories are unavailable since collecting a group of geographically and statistically representative field AGBs is of high cost and is inoperable. Thus, considering the sampling size of the training dataset, the utilization of LiDAR-derived AGB might be cost-efficient as the economic cost for a large quantity of field measurements sometimes can be even higher than the airborne campaign. A first attempt was made to use the LiDAR-evaluated forest age as the prior knowledge for random stratified sampling, which increased the representativeness and variation range of the sampled AGB. The forest age was directly evaluated based on the LiDAR-derived canopy height, which was supported by the previous study where the stand height proved to be the most important forest structural predictor.[56] Thus, the utilization of LiDAR-derived height to evaluate the stand age enriched the potential of airborne LiDAR on forest monitoring since most previous studies focused on using airborne LiDAR to predict height-related and structure parameters.

5.2 *Above-Ground Biomass Estimation and Uncertainty*

The saturated backscatter for different forest types were usually within a small variation range (e.g., between −8 and −11 dB at HH and between −11 and −15 dB at HV) at a given polarization and incidence angle.[5,57] In this study, the saturation point of HV backscatter was a bit lower than those in the cited studies.[5,57] This likely resulted from the dominating coniferous species and the mountainous topography in the Dayekou area. The better relationship of AGB to HV than HH backscatter is probably due to the fact that HV polarization is less influenced by soil, vegetation moisture, and topography.[58] This agrees well with previous studies where the HV backscatter of AIRSAR at C band were found to be saturated at a higher AGB level, and suggested that the L-band HV data from ALOS PALSAR should be considered when large-area mapping of forest AGB is needed.[59]

The estimation accuracies for the parametric models in this study were moderately high and acceptable compared with most previous studies where large quantities of field measurements were needed.[8] Only the RF model from the four nonparametric models obtained comparable performance to the SMLR-II model, which indicated that the estimation model of forest AGB is site-specific. Thus, comparisons among various models should be conducted before applying the model onto regional mapping. High correlation, found between the field-measured and predicted AGB outside the HDP ($R^2 = 0.62$, $\text{RMSE} = 42.33\ \text{Mg}\,\text{ha}^{-1}$), indirectly suggested that the SAR-predicted AGB have comparable accuracy to the field measurements. However, uncertainties still existed during the estimation procedure. We evaluated the forest age of HDP at crown object level and attributed it back to pixel level instead of directly evaluating it at pixel level. Although the crown-level evaluation procedure followed the natural logic but still involved some uncertainties during the scale shifting. Usually, the SAR backscatter tends to be saturated for dense CC. In this study, high errors occurred in the areas with high CC and AGB levels (e.g., old forest stands) and areas with lower CC (e.g., young forest stands). Some areas with higher CC also obtained lower MAEs due to the fact that the CC of the forest over the whole study site was moderately high and most of the forest stands were at a medium growth stage.[60] Thus, backscatter for young forest stand was highly affected by topography and background soil, whereas those for old forest tended to suffer more from signal saturation.

Most analyses in this study were conducted based on the raster with a spatial resolution of 20 m. The effects on AGB estimation from the spatial scale of field plots, LiDAR metrics, and SAR backscatter were not fully considered in this study. Previous studies reported that the correlation between AGB and backscatter will be improved when the pixel size increases,[61] because the effects from spatial heterogeneity decreased and local forest variations were smoothed. It is likely to take for granted that less noise or uncertainties will be smoothed when the sampling population becoming smaller. In fact, field plot size often varies in order to make it consistent with the spatial resolution of remote sensing images during the simultaneous ground-airborne or ground-satellite campaigns. In this study, the plot size of field measurements (20 m) was mainly designed for the airborne LiDAR campaign, which might be smaller than those used for satellite SAR images. Thus, the scale effect on AGB estimation from the field measurements should be further investigated in the future.

6 Conclusion

In this study, we tried to upscale the estimation of coniferous forest biomass using airborne LiDAR-derived biomass and satellite ALOS PALSAR data via parametric and nonparametric models over a mountainous environment in northwest China. Results showed that random stratified sampling that used LiDAR-evaluated forest age as stratification knowledge performed the best in the AGB sampling. The addition of backscatter texture variables improved the parametric model performance by an R^2 increase of ∼21% and an RMSE decrease of 10 $\mathrm{Mg\,ha^{-1}}$. Not all the nonparametric models, except for the RF model, obtained comparable performances ($R^2 = 0.78$, $\mathrm{RMSE} = 14.95\ \mathrm{Mg\,ha^{-1}}$) to the parametric model in our cases. Higher estimation errors occurred in the young forest stands with lower CC and older forest stands with higher AGB levels. In conclusion, the incorporation of these two active remote sensing tools was proven to be efficient in upscaling the AGB estimation to regional scale, which provides some guidance for future forest management over cold and arid areas.

Acknowledgments

This work was supported by China's Special Funds for Major State Basic Research Project of China (No. 2013CB733405), the National High Technology Research and Development Program of China (863 Program) (No. 2014AA06A511), the Special Foundation for Young Scientists from the director of the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, and the Hundred Talents Program of the Chinese Academy of Sciences. We thank the Chinese Academy of Sciences Action Plan for West Development Project WATER for providing the research data.

References

1. M. W. Palace et al., "Estimating forest structure in a tropical forest using field measurements, a synthetic model and discrete return lidar data," *Remote Sens. Environ.* **161**, 1–11, (2015).
2. G. P. Asner and J. Mascaro, "Mapping tropical forest carbon: calibrating plot estimates to a simple LiDAR metric," *Remote Sens. Environ.* **140**, 614–624, (2014).

3. W. Ni et al., "The penetration depth derived from the synthesis of ALOS/PALSAR InSAR data and ASTER GDEM for the mapping of forest biomass," *Remote Sens.* **6**, 7303–7319 (2014).
4. R. Avtar, R. Suzuki, and H. Sawada, "Natural forest biomass estimation based on plantation information using PALSAR data," *PLoS One* **9** (2014).
5. O. Hamdan, H. Khali, and I. Hasmadi, "L-band ALOS PALSAR for biomass estimation of Matang Mangroves, Malaysia," *Remote Sens. Environ.* **155**, 69–78 (2014).
6. J. E. Nichol and M. L. R. Sarker, "Improved biomass estimation using the texture parameters of two high-resolution optical sensors," *IEEE Trans. Geosci. Remote Sens.* **49**, 930–948 (2011).
7. M. Sarker et al., "Potential of texture measurements of two-date dual polarization PALSAR data for the improvement of forest biomass estimation," *ISPRS J. Photogramm. Remote Sens.* **69**, 146–166 (2012).
8. R. B. Thapa et al., "Potential of high-resolution ALOS-PALSAR mosaic texture for above-ground forest carbon tracking in tropical region," *Remote Sens. Environ.* **160**, 122–133 (2015).
9. E. Naesset and K. O. Bjerknes, "Estimating tree heights and number of stems in young forest stands using airborne laser scanner data," *Remote Sens. Environ.* **78**, 328–340 (2001).
10. S. C. Popescu and R. H. Wynne, "Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height," *Photogramm. Eng. Remote Sens.* **70**, 589–604 (2004).
11. F. Morsdorf et al., "Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction," *Remote Sens. Environ.* **104**, 50–61 (2006).
12. T. Sasaki et al., "Estimation of leaf area index and canopy openness in broad-leaved forest using an airborne laser scanner in comparison with high-resolution near-infrared digital photography," *Landscape Ecolog. Eng.* **4**, 47–55 (2008).
13. K. Zhao and S. Popescu, "Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA," *Remote Sens. Environ.* **113**, 1628–1645 (2009).
14. Y. Pang and Z. Y. Li, "Inversion of biomass components of the temperate forest using airborne Lidar technology in Xiaoxing'an Mountains, Northeastern of China," *Chin. J. Plant Ecol.* **36**, 1095–1105 (2013).
15. X. Tian et al., "Estimation of forest above-ground biomass using multi-parameter remote sensing data over a cold and arid area," *Int. J. Appl. Earth Obs. Geoinf.* **14**, 160–168 (2012).
16. E. Hansen et al., "Modeling aboveground biomass in dense tropical submontane rainforest using airborne laser scanner data," *Remote Sens.* **7**, 788–807 (2015).
17. G. Chen et al., "A multiscale geographic object-based image analysis to estimate lidar-measured forest canopy height using quickbird imagery," *Int. J. Geogr. Inf. Sci.* **25**, 877–893 (2011).
18. O. W. Tsui et al., "Integrating airborne LiDAR and space-borne radar via multivariate kriging to estimate above-ground biomass," *Remote Sens. Environ.* **139**, 340–352 (2013).
19. A. T. Hudak et al., "Integration of lidar and Landsat ETM+ data for estimating and mapping forest canopy height," *Remote Sens. Environ.* **82**, 397–416 (2012).
20. L. Korhonen et al., "Airborne discrete-return LIDAR data in the estimation of vertical canopy cover, angular canopy closure and leaf area index," *Remote Sens. Environ.* **115**, 1065–1080 (2011).
21. L. Korhonen et al., "Modelling lidar-derived boreal forest canopy cover with SPOT 4 HRVIR data," *Int. J. Remote Sens.* **34**, 8172–8181 (2013).
22. W. Li et al., "Geostatistical modeling using LiDAR-derived prior knowledge with SPOT-6 data to estimate temperate forest canopy cover and above-ground biomass via stratified random sampling," *Int. J. Appl. Earth Obs. Geoinf.* **41**, 88–98 (2015).
23. N. Joshi et al., "L-Band SAR backscatter related to forest cover, height and aboveground biomass at multiple spatial scales across Denmark," *Remote Sens.* **7**, 4442–4472 (2015).
24. W. Li et al., "Forest above-ground biomass estimation at plot and tree levels using airborne LiDAR data," *J. Remote Sens.* **19**, 669–679 (2015).
25. T. J. Hawbaker et al., "Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design," *J. Geophys. Res.* **114**, G00E04 (2009).

26. M. Garcia et al., "Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data," *Remote Sens. Environ.* **114**, 816–830 (2010).
27. W. Ni-Meister et al., "Assessing general relationships between aboveground biomass and vegetation structure parameters for improved carbon estimate from lidar remote sensing," *J. Geophys. Res.* **115**, G00E11 (2015).
28. K. Zhao et al., "Characterizing forest canopy structure with lidar composite metrics and machine learning," *Remote Sens. Environ.* **115**, 1978–1996 (2011).
29. S. L. Powell et al., "Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: a comparison of empirical modeling approaches," *Remote Sens. Environ.* **114**, 1053–1068 (2010).
30. O. Mutanga et al., "High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm," *Int. J. Appl. Earth Obs. Geoinf.* **18**, 399–406 (2012).
31. J. M. Galeana-Pizaña et al., "Modeling the spatial distribution of above-ground carbon in Mexican coniferous forests using remote sensing and a geostatistical approach," *Int. J. Appl. Earth Obs. Geoinf.* **30**, 179–189 (2014).
32. X. Li et al., "Watershed allied telemetry experimental research," *J. Geophys. Res.* **114**, D22103 (2009).
33. X. Li et al., "Observing and modeling the catchment scale water cycle," *Hydrol. Earth Syst. Sci.* **15**, 597–601 (2011).
34. Q. He et al., "Forest stand biomass estimation using ALOS PALSAR data based on LiDAR-derived prior knowledge in the Qilian Mountain, western China," *Int. J. Remote Sens.* **33**, 710–729 (2012).
35. C. Cao et al., "Retrieval of forest canopy attributes based on a geometric-optical model using airborne LiDAR and optical remote-sensing data," *Int. J. Remote Sens.* **33**, 692–709 (2012).
36. Y. Qin et al., "Stepwise decomposition and relative radiometric normalization for small footprint LiDAR waveform," *China Ser. D* **54**, 625–630 (2010).
37. Y. Qin et al., "Characterizing radiometric attributes of point cloud using a normalized reflective factor derived from small footprint LiDAR waveform," *IEEE J. Stars.* **8**, 740–749 (2015).
38. W. Li et al., "Correlating the horizontal and vertical distribution of LiDAR point clouds with components of biomass in a Picea crassifolia forest," *Forests* **5**, 1910–1930 (2014).
39. NASA and Japan ASTER Program, "The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM)," http://gdem.ersdac.jspacesystems.or.jp/.
40. J. J. Van Zyl et al., "The effect of topography on SAR calibration," *IEEE Trans. Geosci. Remote Sens.* **31**, 1036–1043 (1993).
41. L. M. Ulander, "Radiometric slope correction of synthetic-aperture radar images," *IEEE Trans. Geosci. Remote* **34**, 1115–1122 (1996).
42. J. Y. Wang et al., "Study on biomass of water conservation forest on North Slope of Qilian Mountains," *J. Fujian Coll. For.* **18**, 319–323 (1998).
43. M. Jin et al., "Spatial structure characteristic of Picea crassifolia in Qilian mountains," *Arid Land Geogr.* **35**, 587–593 (2012).
44. J. Q. Liu et al., "Structure and dynamic of Picea classifolia community in Qilian mountains forest area," *J. Northwest For. Univ.* **23**, 14–17 (2008).
45. L. Vincent and P. Soille, "Watersheds on digital spaces—an efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Anal.* **13**, 583–598 (1991).
46. L. R. Sarker and J. E. Nichol, "Improved forest biomass estimates using ALOS AVNIR-2 texture indices," *Remote Sens. Environ.* **115**, 968–977 (2011).
47. O. W. Tsui et al., "Using multi-frequency radar and discrete-return LiDAR measurements to estimate above-ground biomass and biomass components in a coastal temperate forest," *ISPRS J. Photogramm. Remote Sens.* **69**, 121–133 (2012).
48. K. S. Lim and P. M. Treitz, "Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators," *Scand J. For. Res.* **19**, 558–570 (2004).

49. R. Ahmed et al., “A study of forest biomass estimates from lidar in the northern temperate forests of New England,” *Remote Sens. Environ.* **130**, 121–135 (2013).
50. Q. He et al., “Above-ground biomass and biomass components estimation using LiDAR data in a coniferous forest,” *Forests* **4**, 984–1002 (2013).
51. R. R. Picard and R. D. Cook, “Cross-validation of regression models,” *J. Am. Stat. Assoc.* **79**, 575–583 (1984).
52. M. Bouvier et al., “Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data,” *Remote Sens. Environ.* **156**, 322–334 (2015).
53. A. D. Pierce et al., “Use of random forests for modeling and mapping forest canopy fuels for fire behavior analysis in Lassen Volcanic National Park, California, USA,” *For. Ecol. Manage.* **279**, 77–89 (2012).
54. R Development Core Team, “R: a language and environment for statistical computing,” R Foundation for Statistical Computing, Vienna, Austria, http://www.r-project.org/ (2008).
55. M. A. Tanase et al., “Sensitivity of L-band radar backscatter to forest biomass in semiarid environments: a comparative analysis of parametric and nonparametric models,” *IEEE Trans. Geosci. Remote Sens.* **52**(8), 4671–4685 (2014).
56. E. B. Racine et al., “Estimating forest stand age from LiDAR-derived predictors and nearest neighbor imputation,” *For. Sci.* **60**, 128–136 (2014).
57. T. Toan et al., “Relating radar remote sensing of biomass to modelling of forest carbon budgets,” *Clim. Change* **67**, 379–402 (2004).
58. A. Peregon and Y. Yamagata, “The use of ALOS/PALSAR backscatter to estimate above-ground forest biomass: a case study in Western Siberia,” *Remote Sens. Environ.* **137**, 139–146 (2013).
59. R. M. Lucas et al., “Empirical relationships between AIRSAR backscatter and LiDAR-derived forest biomass, Queensland, Australia,” *Remote Sens. Environ.* **100**, 407–425 (2006).
60. W. Li et al., “Analyzing and retrieving structure information of Picea crassifolia based on airborne light detection and ranging data,” *J. Remote Sens.* **6**, 1612–1626 (2013).
61. C. Robinson et al., “Impacts of spatial variability on aboveground biomass estimation from L-band radar in a temperate forest,” *Remote Sens. Basel* **5**, 1001–1023 (2013).

Wang Li received his BS degree in remote sensing from China University of Geosciences, Wuhan, China, in 2011, and his PhD in cartography and geographic information systems from the University of Chinese Academy of Sciences, Beijing, China, in 2015. He is currently a research associate at the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences (CAS), Beijing, China. His research interests include LiDAR remote sensing, vegetation biomass, structure, and biochemical parameter inversion.

Biographies for the other authors are not available.