

Article

# Deep Learning Based Retrieval of Forest Aboveground Biomass from Combined LiDAR and Landsat 8 Data

Linjing Zhang <sup>1</sup>, Zhenfeng Shao <sup>2,\*</sup>, Jianchen Liu <sup>1</sup> and Qimin Cheng <sup>3</sup>

<sup>1</sup> Geomatics College, Shandong University of Science and Technology, 579 Qianwangang Road, Qingdao 266590, China; zhanglinjing@sdust.edu.cn (L.Z.); liujianchen@sdust.edu.cn (J.L.)

<sup>2</sup> State key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China

<sup>3</sup> School of Electronics Information and Communications, Huazhong University of Science and Technology, 1037 Luoyu Road, Wuhan 430074, China; chengqm@hust.edu.cn

\* Correspondence: shaozhenfeng@whu.edu.cn; Tel.: +86-158-2718-8114

Received: 6 May 2019; Accepted: 13 June 2019; Published: 20 June 2019



**Abstract:** Estimation of forest aboveground biomass (AGB) is crucial for various technical and scientific applications, ranging from regional carbon and bioenergy policies to sustainable forest management. However, passive optical remote sensing, which is the most widely used remote sensing data for retrieving vegetation parameters, is constrained by spectral saturation problems and cloud cover. On the other hand, LiDAR data, which have been extensively used to estimate forest structure attributes, cannot provide sufficient spectral information of vegetation canopies. Thus, this study aimed to develop a novel synergistic approach to estimating biomass by integrating LiDAR data with Landsat 8 imagery through a deep learning-based workflow. First the relationships between biomass and spectral vegetation indices (SVIs) and LiDAR metrics were separately investigated. Next, two groups of combined optical and LiDAR indices (i.e., COLI1 and COLI2) were designed and explored to identify their performances in biomass estimation. Finally, five prediction models, including K-nearest Neighbor, Random Forest, Support Vector Regression, the deep learning model, i.e., Stacked Sparse Autoencoder network (SSAE), and multiple stepwise linear regressions, were individually used to estimate biomass with input variables of different scenarios, i.e., (i) all the COLI1 (ACOLI1), (ii) all the COLI2 (ACOLI2), (iii) ACOLI1 and all the optical (AO) and LiDAR variables (AL), and (iv) ACOLI2, AO and AL. Results showed that univariate models with the combined optical and LiDAR indices as explanatory variables presented better modeling performance than those with either optical or LiDAR data alone, regardless of the combination mode. The SSAE model obtained the best performance compared to the other tested prediction algorithms for the forest biomass estimation. The best predictive accuracy was achieved by the SSAE model with inputs of combined optical and LiDAR variables (i.e., ACOLI1, AO and AL) that yielded an  $R^2$  of 0.935, root mean squared error (RMSE) of 15.67 Mg/ha, and relative root mean squared error (RMSE<sub>r</sub>) of 11.407%. It was concluded that the presented combined indices were simple and effective by integrating LiDAR-derived structure information with Landsat 8 spectral data for estimating forest biomass. Overall, the SSAE model with inputs of Landsat 8 and LiDAR integrated information resulted in accurate estimation of forest biomass. The presented modeling workflow will greatly facilitate future forest biomass estimation and carbon stock assessments.

**Keywords:** LiDAR; Landsat 8 OLI; combined optical and LiDAR indices; synergy; Stacked Sparse Autoencoder network (SSAE); forest aboveground biomass (AGB)

## 1. Introduction

Reliable, up-to-date forest aboveground biomass (AGB) mapping is a prerequisite for understanding the relationship between AGB and climate change. Plot-based estimations of forest AGB, while typically of high accuracy, are costly and can only provide quality information for a limited number of stands at the landscape scale [1]. Consequently, numerous remote sensing techniques have been increasingly utilized to assist forest AGB estimation during the last few decades [2]. It is possible to provide updated, consistent, and spatially explicit assessment of forest biomass and its dynamics by using remote sensing images, particularly in large areas with limited accessibility [3,4].

Numerous statistical models have been explored in relating field-measured AGB to remotely sensed variables [5,6]. Multispectral vegetation indices (SVIs), defined with various combinations of visible, near-infrared (NIR) and shortwave reflectance, are the most widely used for retrieving biophysical and biochemical parameters. It is proved that SVIs have strong correlations with vegetation structure characteristics, such as AGB and leaf area index (LAI) [7–9]. However, vegetation indices tend to saturate for forests at high biomass levels [10–12]. The saturation point varies greatly depending on the source data and the vegetation type and ranges from 15 to 100 Mg/ha for different visible/NIR vegetation indices [13,14]. Additionally, optical remote sensing provides limited information on the vertical distribution of forest structure [15], and it is not always possible to compile a temporally and radiometrically consistent cloud-free datasets over large areas [3]. The past two decades have witnessed a large number of studies using Synthetic Aperture Radar (SAR), which could penetrate clouds and forest canopies with appropriate wavelength and polarization modes, for mapping forest AGB [16–19]. However, it remains problematic to estimate AGB using SAR backscattering signals due not only to the saturation at high biomass levels but also to the high sensitivity to soil conditions, including surface roughness and soil moisture [19,20].

Recently, light detection and ranging (LiDAR) has proven to be very powerful in estimating forest structure attributes, such as canopy height, leaf area index (LAI), and AGB [21–24]. The most commonly used LiDAR metrics are height metrics which can be directly measured by LiDAR and provide information related to the vertical structure of individual trees and forest stands. The prevalence of LiDAR data in forest studies is attributed to its penetration by which layered structural echoes through a certain canopy depth can be detected. However, LiDAR data cannot provide sufficient spectral characteristics of vegetation canopies since most LiDAR systems only work at a single wavelength [25]. Although hyperspectral LiDAR systems have emerged to capture spectral and structural information simultaneously [26–28], they are mainly tested inside the laboratory. It is, thus, not feasible to apply these new systems to normal data collection in large areas [26]. Therefore, the integration of LiDAR data with optical remote sensing imagery has been identified as the most promising approach to acquire structural and spectral information from forests simultaneously for biomass estimation. Combining both airborne LiDAR and hyperspectral data has shown great capability to map tree species in different forest areas [29,30]. In addition, Graham [31] found that retrieval of canopy LAI in a coniferous and broadleaf mixed forest can be improved by integrating LiDAR with WorldView-2 data.

The parametric models, such as multiple regression, are commonly used to develop relationships between forest attributes and remote sensing predictors [32,33]. In recent years, non-parametric machine learning models have become prevalent. Contrary to the linear regression model, many machine learning techniques (e.g., Random Forest (RF), Support Vector Regression (SVR), K-nearest Neighbor (KNN), and Deep learning (DL)) are able to reveal complicated non-linear patterns [9,34,35]. Additionally, machine learning models are able to address issues associated with data dimensionality [36,37] in fitting models with a large number of predictors. Few studies have integrated LiDAR data with optical remote sensing metrics to estimate forest structure parameters using non-parametric machine learning algorithms, especially the deep learning models. The deep learning algorithms have shown their effectiveness in object detection and image classification [38–43]. The deep learning models can automatically extract invariant and abstract features which have better discrimination than artificial features. As one of the deep learning algorithms, the Stacked Sparse Autoencoder network

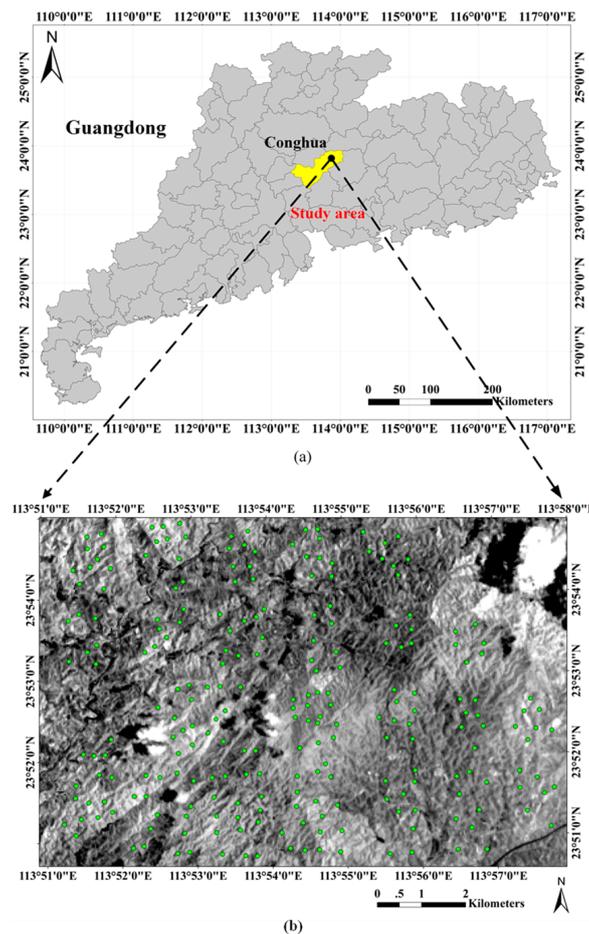
(SSAE) has been widely used in some fields, especially image classification [43]. Besides, few studies paid attention to the comparisons between deep learning and other machine learning algorithms in predicting forest parameters.

Therefore, in this study, airborne LiDAR data were integrated with Landsat 8 imagery for mapping forest aboveground biomass. The objectives of the study were (i) to evaluate the potentials of optical spectral vegetation indices (SVIs) and LiDAR metrics, respectively, for estimating forest AGB; (ii) to determine if the combined LiDAR and Landsat 8 indices can enhance the AGB estimation; and (iii) to provide a comparison of five predictive models, i.e., K-nearest Neighbor (KNN), Random Forest (RF), Support Vector Regression (SVR), Stacked Sparse Autoencoder network (SSAE), and multiple stepwise linear regressions (MSLR) models, for estimating biomass.

## 2. Materials

### 2.1. Study Area

The study site is located in the northeast part of Conghua (23°52′35.65″N, 113°54′46.17″E, Figure 1), Guangdong province, China. The study site has an area of 100 km<sup>2</sup>. The elevation ranges from about 196 m in the northeast of the area to approximately 620 m in the southwest. This area has a subtropical monsoon climate. The mean annual temperature and rainfall are about 22.1 °C and 1952.5 mm, respectively. The area is characterized by the coniferous and broadleaf mixed forest, and *Schima superba*, *Castanopsis fissa*, *Cinnamomum porrectum*, and *Pinus massoniana* are the main tree species.



**Figure 1.** The study area. (a) The location of the study area in Conghua, Guangdong province, China; (b) Landsat 8 OLI near-infrared image of the study area and the plot positions (green dots) within the study area.

## 2.2. Field Data

Aboveground biomass values of 236 inventory plots distributed across the study area were collected in 2013. The locations of plots were selected according to the subjective sampling evaluations, and the size of each plot was 30 m × 30 m. The tree height (H) was measured using a laser hypsometer, and the tree diameter at breast height (DBH) was measured using a tape. The center of each plot was correctly determined using a GPS (Garmin MAP 60CS, accuracy ±3 m). Additionally, the species and type (i.e., evergreen or deciduous) of each tree were recorded. Trees with DBH more than or equal to 3 cm were measured in the survey. Subsequently, a plot would be classified into a coniferous (deciduous) plot, when it consisted of over 75% coniferous (deciduous) trees [44]. Otherwise, we identified the plot as the mixed forest [45]. The observed plots (n = 236) were randomly split into training (n = 177) and validation (n = 59) datasets at a ratio of 3:1.

In general, AGB of each individual tree in the plot can be derived by using species-specific allometric equations with inputs of the DBH and H [20,46,47]. However, no allometric equations are available for the study area. Accordingly, the methodology put forward by Fang et al. [48] was employed to calculate the biomass at the plot level. In this method, the relationship between total volumes and total AGB of each plot is shown in Equation (1). The total volumes could be obtained by adding up the volumes of all individual trees in the plot. The single tree volume was estimated based on a volume table with inputs of H and DBH.

$$TAGB = a \times TV + b \quad (1)$$

where TAGB and TV are separately the total AGB and total volumes in a plot, and the plot AGB needs to be finally computed using TAGB at a megagrams per hectare conversion unit. a and b are coefficients from [48] and they are different for diverse forest types. The biomass values were in the range from 17.464 to 313.918 Mg/ha with a mean value of 125.745 Mg/ha and a standard deviation of 71.13 Mg/ha.

## 2.3. Landsat 8 Acquisition and Preprocessing

The Landsat 8 image used in the study was acquired on the 3rd of August 2014. The image was acquired during an almost clear and sunny sky condition with 5% cloud cover, sun azimuth angle of 102.15 and sun elevation angle of 65.12. The Landsat 8 satellite contains two pushbroom instruments, i.e., the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) including two thermal infrared bands at a 100 m spatial resolution. The Landsat 8-OLI is composed of 30 m multispectral bands corresponding to band 1—coastal/aerosol (0.433–0.453 μm), band 2—blue (0.450–0.515 μm), band 3—green (0.525–0.600 μm), band 4—red (0.630–0.680 μm), band 5—NIR (0.845–0.885 μm), band 6—SWIR1 (1.560–1.660 μm), band 7—SWIR2 (2.100–2.300 μm), and band 9—Cirrus (1.360–1.390 μm), in addition to a 15 m panchromatic image (band 8, 0.500–0.680 μm). Preprocessing was performed using the ENVI 5.1 software. The Landsat-8 OLI image bands were converted from digital number format (DN) to Top-Of-Atmosphere (TOA) spectral radiance, following steps summarized on the USGS website. Subsequently, the Landsat 8 OLI image was atmospherically corrected to surface reflectance using the Fast Line-of-sight Atmospheric Analysis of Spectral Hypercube (FLAASH). The final step was to orthorectify the Landsat 8 image using the high-resolution DEM.

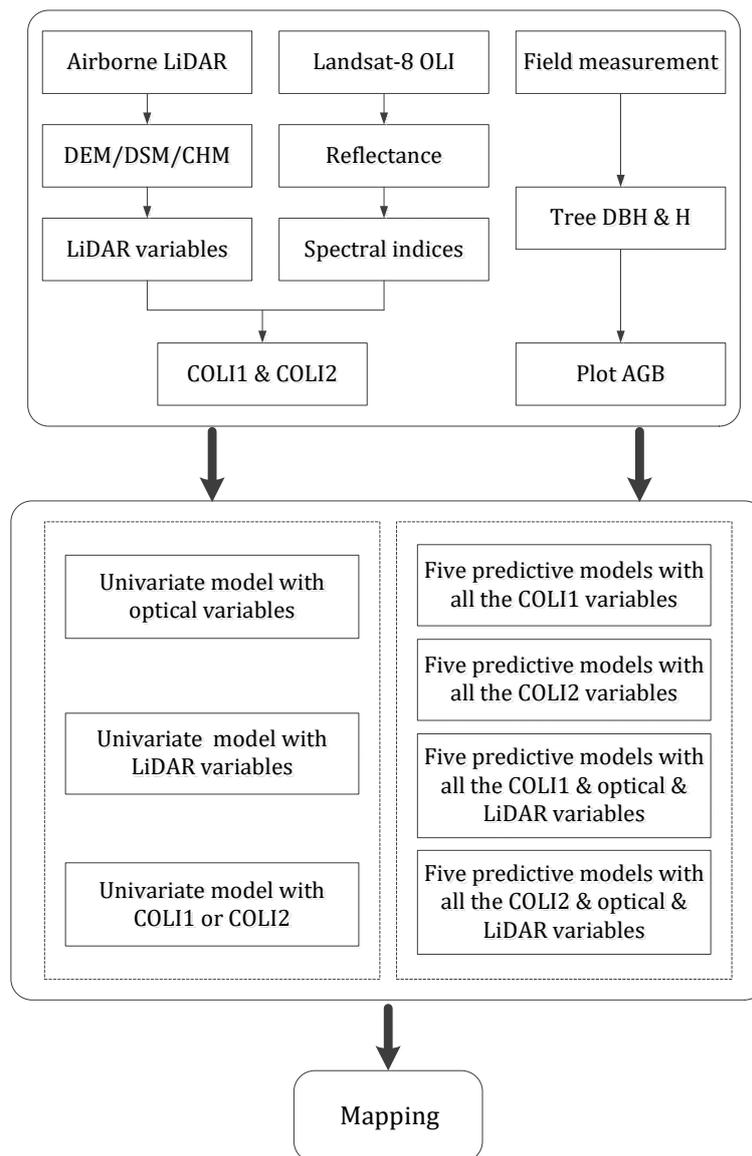
## 2.4. LiDAR Data Acquisition and Preprocessing

Airborne observations for the study area were conducted in June 2013 using a Leica ALS50-II laser scanning system. The system recorded both first and last return data for each laser pulse. The pulse frequency was 52.9 kHz and the flying speed was 80 knots, producing an average density of 0.8 pulses/m<sup>2</sup>. The LiDAR data was preprocessed using the Terrascan software (v4.006-Terrasolid, Helsinki, Finland). Firstly, the points were removed if they had only few neighbors and/or their elevations were higher than the median elevation of surrounding points. The points were then classified as ground and non-ground returns. The ground returns were interpolated to produce a

digital elevation model (DEM), and the first returns were interpolated to derive a digital surface model (DSM) with a resolution of 1 m. Finally, a Canopy Height Model (CHM) was generated by subtracting the DEM from the DSM. According to the minimum and maximum height of field-measured trees within the area, the CHM pixels with values ranging from 2 m to 35 m were extracted to ensure the understory vegetation and objects exceeding the tree height were excluded [21,49,50].

### 3. Methodology

Figure 2 shows the workflow of the biomass estimation processes, including the variables calculation, model calibrations using different data scenarios, accuracy assessment, and wall-to-wall biomass mapping using the calibrated model.



**Figure 2.** The workflow of the modeling process.

#### 3.1. Variables Calculation

##### 3.1.1. Landsat 8 OLI and LiDAR Variables

According to previous studies [7,9,51], seven SVIs were selected, including the Normalized Difference Vegetation Index (NDVI), Simple Ratio Vegetation Index (SR), Enhanced Vegetation Index

(EVI), Atmospherically Resistant Vegetation Index (ARVI), Soil Adjusted Vegetation Index (SAVI), Modified Soil Adjusted Vegetation Index (MSAVI), and Optimized Soil Adjusted Vegetation Index (OSAVI) (Table 1). A total of 16 LiDAR variables were computed at the plot level. The canopy cover (Cov) variable was derived using the proportion of the number of pulses returned from the canopy to all the returns. Six classes of canopy height metrics calculating the height distribution and height summary statistics of the canopy were extracted, including the mean ( $H_{\text{mean}}$ ), maximum ( $H_{\text{max}}$ ), standard deviation ( $H_{\text{std}}$ ), variance ( $H_{\text{var}}$ ), coefficient of variation ( $H_{\text{cv}}$ ), and percentiles with interval of 10% ( $H_p$ :  $p_{10}, p_{20}, \dots, p_{90}$ ). In addition, canopy relief ratio (CRR) which is a quantitative index of the canopy relative shape, describing the ratio of all returns above the mean heights was estimated (Equation (2)) [52].

$$CRR = \frac{H_{\text{mean}} - H_{\text{min}}}{H_{\text{max}} - H_{\text{min}}} \quad (2)$$

**Table 1.** The prediction variables derived from the Landsat 8 OLI image.

Vegetation Indices	Formula
NDVI	$\frac{NIR-R}{NIR+R}$
SR	$\frac{NIR}{R}$
EVI	$\frac{2.5 \times (NIR-R)}{1+NIR+6 \times R-7.5 \times B}$
ARVI	$\frac{NIR-(2 \times R-B)}{NIR+(2 \times R-B)}$
SAVI	$(1+0.5) \frac{NIR-R}{NIR+R+0.5}$
MSAVI	$NIR+0.5 - \sqrt{(NIR+0.5)^2 - 2 \times (NIR-R)}$
OSAVI	$(1+0.16) \frac{NIR-R}{NIR+R+0.16}$

### 3.1.2. Combined Optical and LiDAR Index (COLI)

As described in the Section 1, optical data can provide spectral information about vegetation canopy, while LiDAR data measures the vertical structure of the forest. Thus, the synergistic utilization of airborne LiDAR and optical remote sensing data is highly valuable for estimating forest biomass. Inspired by the SVIs of optical remote sensing, two types of new indices incorporating optical and LiDAR information (COLI1 and COLI2) were established by integrating the best-performing LiDAR variable (i.e., the LiDAR variable achieving the best AGB estimation accuracy) with each SVI. COLI1 and COLI2 can be written as:

$$COLI1 = SVI_i \times BLV \quad (3)$$

$$COLI2 = SVI_i_{BLV} = (BLV - SVI_i) / (BLV + SVI_i) \quad (4)$$

where BLV is the best-performing LiDAR variable and  $SVI_i$  is the optical spectral vegetation index. Prior to calculating COLI, the BLV and  $SVI_i$  were normalized to allow direct comparisons between optical and LiDAR variables, e.g., with different scales and dynamic ranges.

### 3.2. Regression Algorithms

In the study, five prediction techniques including KNN, RF, SVR, SSAE, and MSLR were used to estimate biomass. The RF technique is based on an ensemble of binary regression trees that are fitted to randomly selected subsets of the training data. In addition to resampling the observations to obtain multiple trees, the RF algorithm also selects a random subset of predictors in tree construction, which is particularly useful when a great number of possibly redundant predictors are available. It has been reported to be an efficient prediction approach for estimating vegetation attributes, especially when the number of predictors is very large [35,53,54]. In [35], the RF was compared with other prediction methods in regard to their predictive power of forest AGB, and RF showed a better estimation accuracy

than other algorithms. SVR identifies optimum hyperplanes using kernel functions to separate groups of input data with similar responses to predict a target variable. The SVR model has been successfully applied in biomass mapping and other remote sensing applications [55,56]. In KNN method, the distance from an unknown pixel of the target dataset to each of the known pixels in the reference dataset is computed. The unknown target pixel is assigned a weighted mean of the  $k$  most similar neighbours, which is computed based on their distance to the target pixel in the feature space [57]. Tian et al. [34] found the performance of KNN outperformed those of other algorithms for estimating forest biomass using Landsat Thematic Mapper (TM) data. SSAE is a kind of deep learning method that can automatically learn useful features layer by layer. The SSAE deep learning algorithm has been successfully used in several fields, such as image classification [39,43]. Nonetheless, as far as we know, this model was rarely applied to estimate forest parameters. Therefore, the SSAE algorithm was introduced in detail in the following. When SSAE model builds a deep neural network to extract deep features, it utilizes a hierarchical training strategy. The model comprises some sparse autoencoder networks (SAEs), and each SAE consists of two parts: an encoder and a decoder. After each SAE is trained, its decoded layer is removed and then an SSAE is established using the encoder parameters of all SAEs in a layer-by-layer manner. Afterwards, this network is linked to a regression model to implement the prediction work. That is to say, an SSAE model is comprised of the trained SAEs and the regression model. The MSLR was taken into consideration in order to compare the prediction performances of the more prevalent and complex machine learning models with that of the frequently-used parametric model.

We employed the R statistical package [58] to implement RF, SVR, KNN, and MSLR models. The Deep Learning toolbox, an open source library (<https://github.com/rasmusbergpalm/DeepLearnToolbox#deeplearntoolbox>), was modified to perform SSAE algorithm.

### 3.3. AGB Modeling and Validation

Table 2 showed the three experiments (Experiments 1–3) designed with different data scenarios. Experiment 1: univariate linear or nonlinear regression models were developed based on a single SVI or LiDAR variable; Experiment 2: each COLI was used to establish the univariate linear or nonlinear regression model; Experiment 3: four data combinations were individually fitted to all the five prediction models described in Section 3.2. The four data combinations included (i) all the COLI1 (ACOLI1), (ii) all the COLI2 (ACOLI2), (iii) ACOI1 and all the optical (AO) and LiDAR variables (AL), and (iv) ACOI2, AO and AL. Individual model predictions were validated against the same independent dataset ( $n = 59$ ). Three indicators were calculated, including coefficient of determination ( $R^2$ ), root mean squared error (RMSE), and relative root mean squared error ( $RMSE_r$ ), for accuracy assessment.

**Table 2.** Predictor variables used in each data scenario for modeling AGB.

Data Scenarios	Data Source	Details	Experiment	
a single SVI or LiDAR variable	Landsat 8 OLI/ LiDAR	each optical or LiDAR variable	1	
a single COLI	Landsat 8 OLI, LiDAR	each COLI (Table 5)	2	
ACOLI1	Landsat 8 OLI, LiDAR	All the COLI1	i	
ACOLI2		All the COLI2	ii	
ACOLI1+AO+AL		All the COLI1, optical and LiDAR variables	iii	3
ACOLI2+AO+AL		All the COLI2, optical and LiDAR variables	iv	

## 4. Results

### 4.1. Univariate Model Performance

#### 4.1.1. Relationships between SVIs or LiDAR Metrics and AGB (Experiment 1)

The statistical regression analysis was used to model relationships between SVIs and biomass. Accuracy assessment of each modeling case is shown in Table 3. The best statistical model was found by using the OSAVI as the input variable ( $R^2 = 0.594$ ,  $RMSE = 37.097$  Mg/ha,  $RMSE_r = 27.004\%$ ), followed by using the NDVI ( $R^2 = 0.563$ ,  $RMSE = 37.702$  Mg/ha,  $RMSE_r = 27.445\%$ ). The EVI fitted model yielded the lowest  $R^2$  and highest RMSE and  $RMSE_r$  values (0.209, 52.012 Mg/ha and 37.861%, respectively). Overall, in the experiment 1, statistical models with the single SVI did not show satisfactory performance with the  $R^2 < 0.6$  and  $RMSE > 37$  Mg/ha. Table 4 shows moderate to good modeling performances for AGB estimation by using LiDAR metrics. The variables of  $p_{60}$  and Cov had the highest ( $R^2 = 0.696$ ,  $RMSE = 31.326$  Mg/ha,  $RMSE_r = 22.803\%$ ) and lowest correlations ( $R^2 = 0.463$ ,  $RMSE = 47.298$  Mg/ha,  $RMSE_r = 34.43\%$ ) with biomass, respectively. The model with the input variable of  $H_{mean}$  yielded the next best performance with an  $R^2$  of 0.684 and RMSE of 32.127 Mg/ha. It was noted that statistical models fitted by LiDAR metrics generally resulted in higher  $R^2$  and smaller RMSE values than those fitted by the optical variables. The differences in  $R^2$  and RMSE between the best LiDAR variable based model and the best optical variable based model were approximately 0.1 and 5.77 Mg/ha, respectively.

**Table 3.** Accuracy assessment of the models with input variables of SVIs.

Vegetation Indices	$R^2$	RMSE (Mg/ha)	$RMSE_r$ (%)
ARVI	0.505	41.216	30.002
SR	0.496	41.536	30.236
NDVI	0.563	37.702	27.445
EVI	0.209	52.012	37.861
SAVI	0.426	44.994	32.752
MSAVI	0.473	42.497	30.935
OSAVI	0.594	37.097	27.004

**Table 4.** Accuracy assessment of the models with input variables of LiDAR metrics.

Lidar Metrics	$R^2$	RMSE (Mg/ha)	$RMSE_r$ (%)
$H_{max}$	0.634	37.015	26.944
<b><math>H_{mean}</math></b>	<b>0.684</b>	<b>32.127</b>	<b>23.386</b>
$p_{10}$	0.553	40.76	29.671
$p_{20}$	0.614	38.92	28.331
$p_{30}$	0.631	37.351	27.189
$p_{40}$	0.660	34.812	25.341
$p_{50}$	0.679	33.037	24.049
<b><math>p_{60}</math></b>	<b>0.696</b>	<b>31.326</b>	<b>22.803</b>
$p_{70}$	0.669	34.059	24.792
$p_{80}$	0.658	35.673	25.968
$p_{90}$	0.679	33.351	24.277
$H_{std}$	0.479	46.581	33.908
$H_{cv}$	0.503	44.785	32.6
$H_{var}$	0.628	38.028	27.682
Cov	0.463	47.298	34.43
CRR	0.573	39.23	28.557

Note: The best ( $p_{60}$ ) and the next best ( $H_{mean}$ ) Lidar metrics were put in bold.

#### 4.1.2. Relationships between COLIs and AGB (Experiment 2)

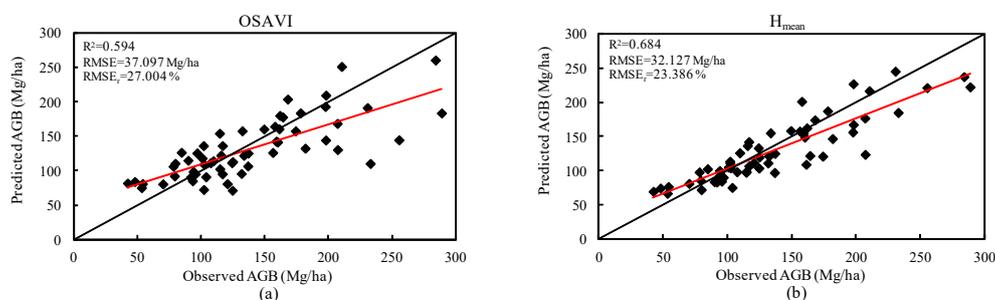
As shown in Table 4, the variable of  $H_{\text{mean}}$  had a relatively good correlation with biomass, following the variable of  $p_{60}$ . Previous studies have used the LiDAR height metrics  $H_{\text{mean}}$  as effective prior knowledge in the sampling design [59,60] because  $H_{\text{mean}}$  can be treated as a direct indicator of the forest structure and growth status. Therefore, two types of the combined optical and LiDAR indices (i.e., COLI1 and COLI2) were generated by integrating  $H_{\text{mean}}$  with each of the seven SVIs according to Equations (3) and (4), respectively. More details on COLI1 and COLI2 were provided in Section 3.1.2.

Table 5 shows AGB predictions could be improved by statistical models that have both  $H_{\text{mean}}$  and one of the SVIs as input variables, compared to those using either optical or LiDAR data alone. Among all the models in COLI1, the statistical model with input variable of  $\text{OSAVI} \times H_{\text{mean}}$  had the best modeling performance with an  $R^2$  of 0.788 and RMSE of 24.787 Mg/ha. In COLI2, the best statistical model ( $R^2 = 0.763$ , RMSE = 26.568 Mg/ha) was achieved by using  $\text{OSAVI}_H_{\text{mean}}$  as well. Additionally, it was observed that statistical models in COLI1 generally had slightly better performance than those in COLI2. The differences in  $R^2$  and RMSE between the model using the best COLI1 ( $\text{OSAVI} \times H_{\text{mean}}$ ) and the model using the best COLI2 ( $\text{OSAVI}_H_{\text{mean}}$ ) reached about 0.03 and 1.78 Mg/ha, respectively. Figure 3 shows AGB comparisons between field observations and model predictions using different data scenarios: (a) OSAVI, (b)  $H_{\text{mean}}$ , (c)  $\text{OSAVI} \times H_{\text{mean}}$ , and (d)  $\text{OSAVI}_H_{\text{mean}}$ . Obviously, the models with combined indices showed fewer deviations in the slopes of the fitted trend-lines (the red lines) away from the 1:1 line (the black lines) than those using either the optical-only or LiDAR-only data.

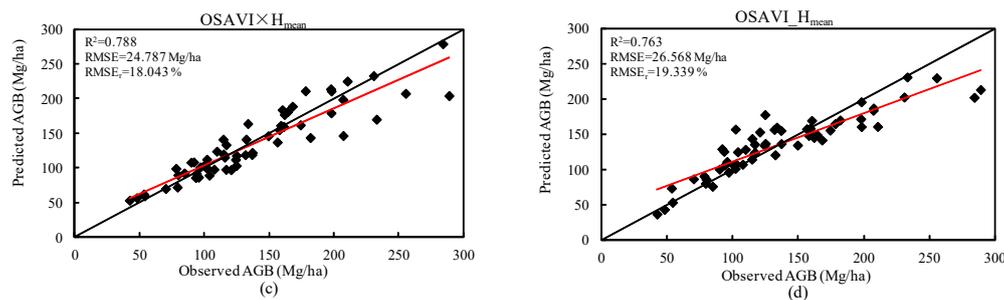
**Table 5.** Accuracy assessment of the models with input variables of COLIs.

Combined Indices		$R^2$	RMSE (Mg/ha)	RMSE <sub>r</sub> (%)
COLI1	$\text{ARVI} \times H_{\text{mean}}$	0.772	25.693	18.703
	$\text{SR} \times H_{\text{mean}}$	0.76	27.084	19.715
	$\text{NDVI} \times H_{\text{mean}}$	0.767	26.278	19.129
	$\text{EVI} \times H_{\text{mean}}$	0.743	27.92	20.324
	$\text{SAVI} \times H_{\text{mean}}$	0.76	26.933	19.605
	$\text{MSAVI} \times H_{\text{mean}}$	0.763	26.688	19.427
	<b><math>\text{OSAVI} \times H_{\text{mean}}</math></b>	<b>0.788</b>	<b>24.787</b>	<b>18.043</b>
COLI2	$\text{ARVI}_H_{\text{mean}}$	0.759	27.056	19.695
	$\text{SR}_H_{\text{mean}}$	0.689	32.034	23.319
	$\text{NDVI}_H_{\text{mean}}$	0.751	27.436	19.972
	$\text{EVI}_H_{\text{mean}}$	0.761	26.898	19.58
	$\text{SAVI}_H_{\text{mean}}$	0.744	28.961	21.082
	$\text{MSAVI}_H_{\text{mean}}$	0.737	29.33	21.35
	<b><math>\text{OSAVI}_H_{\text{mean}}</math></b>	<b>0.763</b>	<b>26.568</b>	<b>19.339</b>

Note: The best combined indices in COLI1 ( $\text{OSAVI} \times H_{\text{mean}}$ ) and COLI2 ( $\text{OSAVI}_H_{\text{mean}}$ ) were put in bold.



**Figure 3.** Cont.



**Figure 3.** Comparisons between field observations and model predictions by using different data scenarios: (a) OSAVI, (b)  $H_{\text{mean}}$ , (c)  $\text{OSAVI} \times H_{\text{mean}}$ , and (d)  $\text{OSAVI}_{H_{\text{mean}}}$ . Note: The red and black lines separately denoted the fitted trend-lines and the 1:1 line.

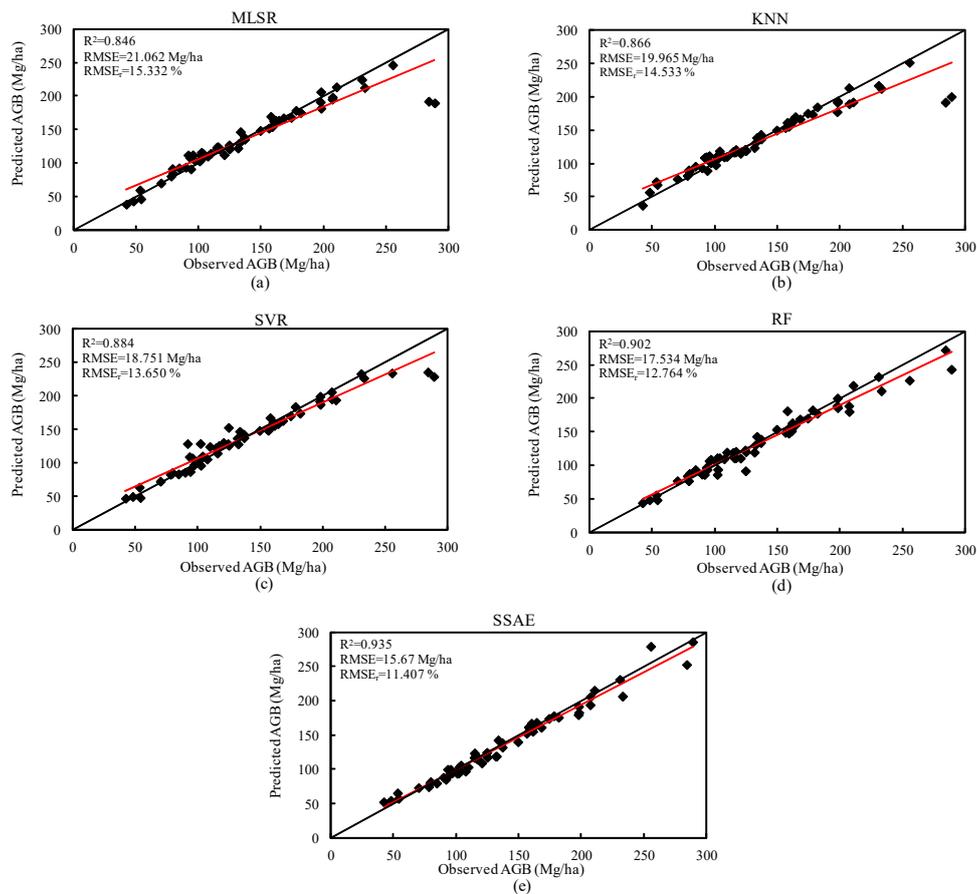
#### 4.2. Multivariate Model Performance (Experiment 3)

The MSLR, KNN, SVR, RF and SSAE algorithms were individually adopted to estimate biomass using ACOLI1, ACOLI2, the combination of ACOLI1, AO and AL, and the combination of ACOLI2, AO and AL. Table 6 shows that the SSAE model has the best modeling performance, followed by the RF model. The MSLR model, in most cases, produced the most error in AGB estimate with the lowest  $R^2$  and highest RMSE values. Figure 4 shows comparisons between field surveys and predicted AGB values derived from the five prediction models calibrated by variables of ACOLI1, AO and AL. Obvious differences in modeling performances among the five prediction algorithms could be identified. The SSAE model showed higher agreements between the model estimates and field observations than other models. Additionally, results showed that the prediction accuracy was improved with increased input variables, i.e., the models with input variables of ACOLI1 (ACOLI2), AO and AL achieved better performances than those with ACOLI1 (ACOLI2). Furthermore, slightly better prediction accuracies were observed for the models calibrated by ACOLI1, AO and AL integrated variables compared to those by ACOLI2, AO and AL. The best performing model, SSAE with ACOLI1, AO and AL integrated variables, reached a  $R^2$  of 0.935, a RMSE of 15.67 Mg/ha and a  $\text{RMSE}_r$  of 11.407%.

**Table 6.** Accuracy assessment for models in Experiment 3.

Algorithm	Predictors	$R^2$	RMSE (Mg/ha)	$\text{RMSE}_r$ (%)
MSLR	ACOLI1	0.812	22.752	16.562
	ACOLI2	0.801	23.569	17.156
	ACOLI1+AO+AL	0.846	21.062	15.332
	ACOLI2+AO+AL	0.834	21.593	15.718
KNN	ACOLI1	0.832	21.861	15.914
	ACOLI2	0.841	21.243	15.463
	ACOLI1+AO+AL	0.866	19.965	14.533
	ACOLI2+AO+AL	0.861	20.225	14.722
SVR	ACOLI1	0.848	20.895	15.21
	ACOLI2	0.836	21.629	15.744
	ACOLI1+AO+AL	0.884	18.751	13.65
	ACOLI2+AO+AL	0.882	18.908	13.764
RF	ACOLI1	0.875	19.215	13.987
	ACOLI2	0.871	19.433	14.146
	ACOLI1+AO+AL	0.902	17.534	12.764
	ACOLI2+AO+AL	0.898	17.93	13.052
SSAE	ACOLI1	0.908	17.058	12.417
	ACOLI2	0.903	17.415	12.677
	<b>ACOLI1+AO+AL</b>	<b>0.935</b>	<b>15.67</b>	<b>11.407</b>
	ACOLI2+AO+AL	0.928	16.158	11.762

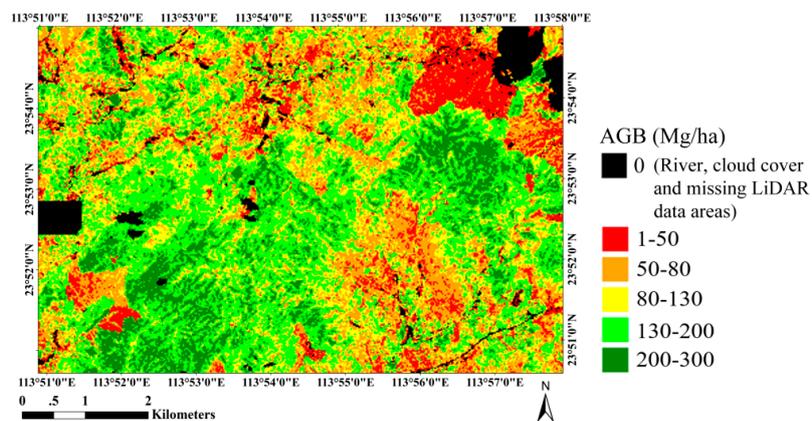
Note: The best performing model (i.e., SSAE with ACOLI1, AO and AL integrated variables) was put in bold.



**Figure 4.** Comparisons between field observations and model predictions derived from the combination of ACOLI1, AO and AL for the five models: (a) MSLR, (b) KNN, (c) SVR, (d) RF, (e) SSAE. Note: The red and black lines separately denoted the fitted trend-lines and the 1:1 line.

### 4.3. Wall-to-Wall Predictions

The spatial distribution of AGB was mapped by applying the SSAE model (calibrated by ACOLI1, AO and AL integrated variables) to the whole study area (Figure 5). The predicted values varied from 24.64 Mg/ha to 299.206 Mg/ha, with a mean value of 113.456 Mg/ha. We have masked out the cloud and river areas. The non-forested regions were considered to have low biomass levels. In view of the raw LiDAR information located in the black rectangle of Figure 5 was missing, this part did not participate in the AGB retrieval.



**Figure 5.** Spatial distributions of AGB estimated by the SSAE model with ACOLI1, AO and AL integrated variables.

## 5. Discussion

### 5.1. Biomass Estimation Using COLI

In the study, the combined optical and LiDAR indices, i.e., COLI1 and COLI2, were developed by integrating the best-performing LiDAR predictor (i.e.,  $H_{\text{mean}}$ ) with each of the SVIs derived from the Landsat 8 OLI image. Both COLIs (i.e., COLI1 and COLI2) improved the biomass estimation accuracy compared to either the optical-only or LiDAR-only variable. This was in accordance with earlier studies [32,35,61,62]. Most of these studies used other approaches to combine optical and LiDAR data for estimating vegetation parameters. Usually, the common approach is the loosely coupled design [32,35,61–64]. For example, Kulawardhana et al. [32] found that the multiple regression models with input variables from the fusion of LiDAR metrics and multispectral vegetation indices had better performances than those calibrated by either the sole LiDAR or multispectral data for vegetation biomass predictions. In Li et al. [62], the statistical model using combined LiDAR and GF-1 data achieved a satisfactory performance for the maize AGB estimation. In this study, new indices were designed, and this approach presented a novel synergistic way of effectively estimating forest biomass. The satisfactory performances of COLIs can be attributed to the ability of multispectral images to provide surface information about canopy density and cover and the ability of LiDAR data to measure forests information about the branches and stems of the trees. Additionally, although COLI1 fitted models generally achieved smaller RMSE and higher  $R^2$  values than COLI2 fitted models, differences in model predictions between them were slight. It was revealed that  $\text{OSAVI} \times H_{\text{mean}}$  and  $\text{OSAVI}_{H_{\text{mean}}}$  could be used for accurate estimates of forest biomass, along with the fact that the other combined indices were also valuable (Table 5). Overall, it was concluded that the proposed COLIs could improve forest biomass predictions since weaknesses and limitations of one predictor might be compensated through the other one. However, the COLIs should experience further investigations and validations before the extensive application.

### 5.2. Biomass Estimation Using Different Prediction Models

To the best of our knowledge, the applications of deep learning models in estimating vegetation parameters are much less common, although they have been investigated by previous studies for many other fields, such as object detection and image classification [38–43]. The estimation performance of SSAE deep learning technique outperformed the other four models. This may be attributed to the fact that SSAE can automatically learn deep spatial sparse features of data layer by layer. In SSAE model, the input variables were mapped into another feature space, where the AGB can be accurately predicted. The overcomplete sparse features were derived due to the sparsity in the hidden layer, and the deep learning features were more representative because of stacked layers of neural networks. In future research, the SSAE modeling should be applied separately for different forest types to improve AGB estimation accuracy. Additionally, the RF algorithm achieved more outstanding predictive accuracy than other models besides SSAE, and its satisfactory performance has been demonstrated in previous studies [35,53,54]. The relatively outstanding performance of the RF model depends on its distinctive regression technique and conceptual design which results in the robustness and flexibility with respect to noise and outliers [65]. It should be noted that RF may not work effectively using a small amount of samples on account of the employment of a soft linear boundaries combination [66]. In the study, the prediction accuracy of SVR was worse than that of RF. This result may be attributed to the “noise” variables which might bias the optimal hyperplane in the SVR and thus weaken the model performance. In all the data scenarios, the results show that the SVR achieved a higher accuracy than the KNN. SVR was more powerful at dealing with high-dimensional and nonlinear problems than KNN, which has been suggested by previous studies [35,53]. Besides, the relatively poor performance of KNN might be related to the lack of an effective procedure to weight the prediction strength of the variables in the implementation process of KNN. In the majority of cases, MLR performed slightly worse than the other tested methods possibly due to its inherent limitations, e.g., the modeling accuracy no longer

changes when additional variables are added [67] because not all variables are linearly correlated with biomass.

### 5.3. Overall Performance of the Biomass Estimation

The biomass prediction result of the SSAE model with inputs of ACOLI1, AO and AL integrated variables ( $R^2 = 0.935$ ,  $RMSE = 15.67$  Mg/ha,  $RMSE_r = 11.407\%$ ) in our study showed a high accuracy comparable to or better than other similar studies. Li et al. [62] integrated airborne LiDAR with satellite GF-1 data for estimating maize AGB during peak growing season with an  $R^2$  of 0.69 and  $RMSE_r$  of 39%. Kulawardhana et al. [32] evaluated the performance of data fusion based on LiDAR metrics and multispectral vegetation indices using multiple regression models for vegetation AGB estimation with an  $RMSE_r$  of 25.9%. Fassnacht et al. [35] investigated the potential of hyperspectral and airborne LiDAR integrated dataset for forest biomass estimation with a mean  $R^2$  of 0.72 and  $RMSE_r$  of 28.46%. In Laurin et al. [61], the Partial Least Square Regression (PLSR) with inputs of the combined airborne LiDAR metrics and hyperspectral bands achieved the best performance ( $R^2 = 0.7$ ,  $RMSE_r = 35.83\%$ ) for the forest biomass estimation.

There is the possibility of improving model accuracy in some effective ways. For example, the methods to select the predictor variables should be introduced, and separate estimates for different forest types should be implemented to further improve the AGB prediction performance in future studies.

## 6. Conclusions

In this research, univariate regression models using combined optical and LiDAR indices (i.e., COLI1 and COLI2) significantly improved the biomass estimation accuracy compared to those using either data type alone. Therefore, the proposed COLIs provide an avenue to synergistically use optical and LiDAR data for mapping biomass and other vegetation parameters. However, these COLIs should be expanded to other forest types for further validations and wide applications. In addition, it should be noted that the presented results are the first step towards the integration of remotely sensed spatial and spectral information for a precise and non-destructive estimation of forest biomass. Other data fusion methods may further increase the prediction power. The deep learning model (i.e., SSAE), which has previously been rarely explored in mapping forest attributes, was found to be superior over other prediction models for estimating forest AGB. This conclusion may facilitate further applications of deep learning models for mapping vegetation structure parameters. The SSAE prediction model calibrated by variables of ACOLI1, AO and AL yielded the best biomass estimation accuracy. Overall, the deep learning model with inputs of structural and spectral integrated information can become a powerful tool for applications in precision forest biomass monitoring.

**Author Contributions:** All authors made great contributions to the work. L.Z. and Z.S. conceived this research. L.Z. performed the experiments and J.L. wrote the paper. Z.S. analyzed data, provided advice and modified the paper. Q.C. modified the English language.

**Funding:** This research was funded by grants from the National key R & D plan on strategic international scientific and technological innovation cooperation special project (Grant No.2016YFE0202300); the National Natural Science Foundation of China (Grants No. 41801385, 61671332, 41771452, 51708426 and 41771454); ShanDong Provincial Natural Science Foundation, China (Grant No. ZR2018BD004, No. ZR2019QD010); Scientific Research Foundation of Shandong University of Science and Technology for Recruited Talents (Grant No. 2017RCJJ076); the Natural Science Fund of Hubei Province in China (Grant No.2018CFA007); the Independent Research Projects of Wuhan University (Grant No.2042018kf0250).

**Acknowledgments:** The authors would like to extend their appreciation to Lei Wang, Peng Fu and Li Yin for their support and help during the experimental and writing period.

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

## References

1. Song, C.; Schroeder, T.A.; Cohen, W.B. Predicting temperate conifer forest successional stage distributions with multitemporal Landsat Thematic Mapper imagery. *Remote Sens. Environ.* **2007**, *106*, 228–237. [[CrossRef](#)]

2. Mcroberts, R.E.; Ståhl, G.; Vidal, C.; Lawrence, M.; Tomppo, E.; Schadauer, K.; Chirici, G.; Bastrupbirk, A. National Forest Inventories: Prospects for Harmonised International Reporting. In *National Forest Inventories*; Springer: Dordrecht, The Netherlands, 2010; pp. 33–43.
3. Avitabile, V.; Baccini, A.; Friedl, M.A.; Schmullius, C. Capabilities and limitations of Landsat and land cover data for aboveground woody biomass estimation of Uganda. *Remote Sens. Environ.* **2012**, *117*, 366–380. [[CrossRef](#)]
4. Herold, M.; Johns, T. Linking requirements with capabilities for deforestation monitoring in the context of the UNFCCC-REDD process. *Environ. Res. Lett.* **2007**, *2*, 045025. [[CrossRef](#)]
5. Wu, C.; Han, X.; Niu, Z.; Dong, J. An evaluation of EO-1 hyperspectral Hyperion data for chlorophyll content and leaf area index estimation. *Int. J. Remote Sens.* **2010**, *31*, 1079–1086. [[CrossRef](#)]
6. Gonsamo, A. Leaf area index retrieval using gap fractions obtained from high resolution satellite data: Comparisons of approaches, scales and atmospheric effects. *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, 233–248. [[CrossRef](#)]
7. Dube, T.; Mutanga, O.; Dube, T.; Mutanga, O. Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in uMgeni catchment, South Africa. *ISPRS J. Photogramm. Remote Sens.* **2015**, *101*, 36–46. [[CrossRef](#)]
8. Zandler, H.; Brenning, A.; Samimi, C. Quantifying dwarf shrub biomass in an arid environment: Comparing empirical methods in a high dimensional setting. *Remote Sens. Environ.* **2015**, *158*, 140–155. [[CrossRef](#)]
9. Chen, B.; Wu, Z.; Wang, J.; Dong, J.; Guan, L.; Chen, J.; Yang, K.; Xie, G. Spatio-temporal prediction of leaf area index of rubber plantation using HJ-1A/1B CCD images and recurrent neural network. *ISPRS J. Photogramm. Remote Sens.* **2015**, *102*, 148–160. [[CrossRef](#)]
10. Labrecque, S.; Fournier, R.A.; Luther, J.E.; Piercey, D. A comparison of four methods to map biomass from Landsat-TM and inventory data in western Newfoundland. *For. Ecol. Manag.* **2006**, *226*, 129–144. [[CrossRef](#)]
11. Duncanson, L.I.; Niemann, K.O.; Wulder, M.A. Integration of GLAS and Landsat TM data for aboveground biomass estimation. *Can. J. Remote Sens.* **2010**, *36*, 129–141. [[CrossRef](#)]
12. Mutanga, O.; Skidmore, A.K. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *Int. J. Remote Sens.* **2004**, *25*, 3999–4014. [[CrossRef](#)]
13. Zheng, G.; Chen, J.M.; Tian, Q.J.; Ju, W.M.; Xia, X.Q. Combining remote sensing imagery and forest age inventory for biomass mapping. *J. Environ. Manag.* **2006**, *85*, 616–623. [[CrossRef](#)] [[PubMed](#)]
14. Lu, D. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* **2006**, *27*, 1297–1328. [[CrossRef](#)]
15. Wulder, M. Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters. *Prog. Phys. Geogr.* **1998**, *22*, 449. [[CrossRef](#)]
16. Yong, W.; Kasischke, E.S.; Melack, J.M.; Davis, F.W.; Christensen, N.L., Jr. The effects of changes in loblolly pine biomass and soil moisture on ERS-1 SAR backscatter. *Remote Sens. Environ.* **1994**, *49*, 25–31.
17. Hamdan, O.; Aziz, H.K.; Rahman, K.A. Remotely sensed l-band sar data for tropical forest biomass estimation. *J. Trop. For. Sci.* **2011**, *23*, 318–327.
18. Sandberg, G.; Ulander, L.M.H.; Fransson, J.E.S.; Holmgren, J.; Toan, T.L. L- and P-band backscatter intensity for biomass retrieval in hemiboreal forest. *Remote Sens. Environ.* **2011**, *115*, 2874–2886. [[CrossRef](#)]
19. Castel, T.; Beaudoin, A.; Stach, N.; Stussi, N.; Toan, T.L.; Durand, P. Sensitivity of space-borne SAR data to forest parameters over sloping terrain. Theory and experiment. *Int. J. Remote Sens.* **2001**, *22*, 2351–2376. [[CrossRef](#)]
20. He, Q.; Cao, C.; Chen, E.; Sun, G.; Ling, F.; Pang, Y.; Zhang, H.; Ni, W.; Xu, M.; Li, Z. Forest stand biomass estimation using ALOS PALSAR data based on LiDAR-derived prior knowledge in the Qilian Mountain, western China. *Int. J. Remote Sens.* **2012**, *33*, 710–729. [[CrossRef](#)]
21. Hyyppä, J.; Yu, X.; Hyyppä, H.; Vastaranta, M.; Holopainen, M.; Kukko, A.; Kaartinen, H.; Jaakkola, A.; Vaaja, M.; Koskinen, J. Advances in Forest Inventory Using Airborne Laser Scanning. *Remote Sens.* **2012**, *4*, 1190–1207. [[CrossRef](#)]
22. Li, W. Correlating the Horizontal and Vertical Distribution of LiDAR Point Clouds with Components of Biomass in a *Picea crassifolia* Forest. *Forests* **2014**, *5*, 1910–1930. [[CrossRef](#)]
23. Zhao, K.G.; Popescu, S.; Nelson, R. Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sens. Environ.* **2009**, *113*, 182–196. [[CrossRef](#)]

24. Zhao, K.G.; Popescu, S. 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.* **2009**, *113*, 1628–1645. [[CrossRef](#)]
25. Bouvier, M.; Durrieu, S.; Fournier, R.A.; Renaud, J.P. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sens. Environ.* **2015**, *156*, 322–334. [[CrossRef](#)]
26. Li, W.; Sun, G.; Niu, Z.; Gao, S.; Qiao, H. Estimation of leaf biochemical content using a novel hyperspectral full-waveform LiDAR system. *Remote Sens. Lett.* **2014**, *5*, 693–702. [[CrossRef](#)]
27. Hakala, T.; Suomalainen, J.; Kaasalainen, S.; Chen, Y. Full waveform hyperspectral LiDAR for terrestrial laser scanning. *Opt. Express* **2012**, *20*, 7119–7127. [[CrossRef](#)]
28. Chen, Y.; Räikkönen, E.; Kaasalainen, S.; Suomalainen, J.; Hakala, T.; Hyyppä, J.; Chen, R. Two-channel Hyperspectral LiDAR with a Supercontinuum Laser Source. *Sensors* **2010**, *10*, 7057–7066. [[CrossRef](#)]
29. Ghosh, A.; Fassnacht, F.E.; Joshi, P.K.; Koch, B. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 49–63. [[CrossRef](#)]
30. Dalponte, M.; Ørka, H.O.; Ene, L.T.; Gobakken, T.; Næsset, E. Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. *Remote Sens. Environ.* **2014**, *140*, 306–317. [[CrossRef](#)]
31. Pope, G.; Treitz, P. Leaf Area Index (LAI) Estimation in Boreal Mixedwood Forest of Ontario, Canada Using Light Detection and Ranging (LiDAR) and WorldView-2 Imagery. *Remote Sens.* **2013**, *5*, 5040–5063. [[CrossRef](#)]
32. Kulawardhana, R.W.; Popescu, S.C.; Feagin, R.A. Fusion of lidar and multispectral data to quantify salt marsh carbon stocks. *Remote Sens. Environ.* **2014**, *154*, 345–357. [[CrossRef](#)]
33. Sheridan, R.; Popescu, S.; Gatzliolis, D.; Morgan, C.; Ku, N.W. Modeling Forest Aboveground Biomass and Volume Using Airborne LiDAR Metrics and Forest Inventory and Analysis Data in the Pacific Northwest. *Remote Sens.* **2014**, *7*, 229–255. [[CrossRef](#)]
34. Tian, X.; Li, Z.; Su, Z.; Chen, E.; Tol, C.V.D.; Li, X.; Guo, Y.; Li, L.; Ling, F. Estimating montane forest above-ground biomass in the upper reaches of the Heihe River Basin using Landsat-TM data. *Int. J. Remote Sens.* **2014**, *35*, 7339–7362. [[CrossRef](#)]
35. Fassnacht, F.E.; Hartig, F.; Latifi, H.; Berger, C.; Hernández, J.; Corvalán, P.; Koch, B. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sens. Environ.* **2014**, *154*, 102–114. [[CrossRef](#)]
36. Hudak, A.T.; Crookston, N.L.; Evans, J.S.; Hall, D.E.; Falkowski, M.J. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sens. Environ.* **2008**, *112*, 2232–2245. [[CrossRef](#)]
37. Durbha, S.S.; King, R.L.; Younan, N.H. Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sens. Environ.* **2007**, *107*, 348–361. [[CrossRef](#)]
38. Yu, Y.; Li, J.; Guan, H.; Wang, C. Automated Detection of Three-Dimensional Cars in Mobile Laser Scanning Point Clouds Using DBM-Hough-Forests. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4130–4142. [[CrossRef](#)]
39. Chen, Y.; Lin, Z.; Zhao, X.; Wang, G.; Gu, Y. Deep Learning-Based Classification of Hyperspectral Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2094–2107. [[CrossRef](#)]
40. Zhang, L.; Shi, Z.; Wu, J. A Hierarchical Oil Tank Detector With Deep Surrounding Features for High-Resolution Optical Satellite Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4895–4909. [[CrossRef](#)]
41. Liang, H.; Li, Q. Hyperspectral Imagery Classification Using Sparse Representations of Convolutional Neural Network Features. *Remote Sens.* **2016**, *8*, 99. [[CrossRef](#)]
42. Yu, Y.; Li, J.; Guan, H.; Jia, F.; Wang, C. Learning Hierarchical Features for Automated Extraction of Road Markings From 3-D Mobile LiDAR Point Clouds. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 709–726. [[CrossRef](#)]
43. Zhang, L.; Ma, W.; Zhang, D. Stacked Sparse Autoencoder in PolSAR Data Classification Using Local Spatial Information. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1359–1363. [[CrossRef](#)]
44. Singh, K.K.; Chen, G.; McCarter, J.B.; Meentemeyer, R.K. Effects of LiDAR point density and landscape context on estimates of urban forest biomass. *ISPRS J. Photogramm. Remote Sens.* **2015**, *101*, 310–322. [[CrossRef](#)]

45. Godwin, C.; Chen, G.; Singh, K.K. The impact of urban residential development patterns on forest carbon density: An integration of LiDAR, aerial photography and field mensuration. *Landsc. Urban Plan.* **2015**, *136*, 97–109. [[CrossRef](#)]
46. Kattenborn, T.; Maack, J.; Faßnacht, F.; Enßle, F.; Ermert, J.; Koch, B. Mapping forest biomass from space – Fusion of hyperspectral EO1-hyperion data and Tandem-X and WorldView-2 canopy height models. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *35*, 359–367. [[CrossRef](#)]
47. Chave, J.; Andalo, C.; Brown, S.; Cairns, M.A.; Chambers, J.Q.; Eamus, D.; Fölster, H.; Fromard, F.; Higuchi, N.; Kira, T. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* **2005**, *145*, 87–99. [[CrossRef](#)]
48. Fang, J.; Liu, G.; Xu, S. Biomass and net production of forest vegetation in china. *Acta Ecol. Sin.* **1996**, *16*, 497–508.
49. Wulder, M.A.; White, J.C.; Nelson, R.F.; Næsset, E.; Ørka, H.O.; Coops, N.C.; Hilker, T.; Bater, C.W.; Gobakken, T. Lidar sampling for large-area forest characterization: A review. *Remote Sens. Environ.* **2012**, *121*, 196–209. [[CrossRef](#)]
50. Shao, Z.; Zhang, L.; Wang, L. Stacked Sparse Autoencoder Modeling Using the Synergy of Airborne LiDAR and Satellite Optical and SAR Data to Map Forest Above-Ground Biomass. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 5569–5582. [[CrossRef](#)]
51. Güneralp, İ.; Filippi, A.M.; Randall, J. Estimation of floodplain aboveground biomass using multispectral remote sensing and nonparametric modeling. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *33*, 119–126. [[CrossRef](#)]
52. Chirici, G.; Mcroberts, R.E.; Fattorini, L.; Mura, M.; Marchetti, M. Comparing echo-based and canopy height model-based metrics for enhancing estimation of forest aboveground biomass in a model-assisted framework. *Remote Sens. Environ.* **2016**, *174*, 1–9. [[CrossRef](#)]
53. Shao, Z.; Zhang, L. Estimating Forest Aboveground Biomass by Combining Optical and SAR Data: A Case Study in Genhe, Inner Mongolia, China. *Sensors* **2016**, *16*, 834. [[CrossRef](#)] [[PubMed](#)]
54. Latifi, H.; Nothdurft, A.; Koch, B. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: Application of multiple optical/LiDAR-derived predictors. *Forestry* **2010**, *83*, 395–407. [[CrossRef](#)]
55. Verrelst, J.; Muñoz, J.; Alonso, L.; Delegido, J.; Rivera, J.P.; Camps-Valls, G.; Moreno, J. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. *Remote Sens. Environ.* **2012**, *118*, 127–139. [[CrossRef](#)]
56. Tuia, D.; Verrelst, J.; Alonso, L.; Perez-Cruz, F. Multioutput Support Vector Regression for Remote Sensing Biophysical Parameter Estimation. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 804–808. [[CrossRef](#)]
57. Latifi, H.; Koch, B. Evaluation of most similar neighbour and random forest methods for imputing forest inventory variables using data from target and auxiliary stands. *Int. J. Remote Sens.* **2012**, *33*, 6668–6694. [[CrossRef](#)]
58. The R Development Core Team. *R: A Language And Environment For Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2013; Volume 1, pp. 12–21.
59. Li, W.; Niu, Z.; Liang, X.; Li, Z.; Huang, N.; Gao, S.; Wang, C.; Muhammad, S. 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.* **2015**, *41*, 88–98. [[CrossRef](#)]
60. Hawbaker, T.J.; Keuler, N.S.; Lesak, A.A.; Terje, G.; Kirk, C.; Radeloff, V.C. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. *J. Geophys. Res. Atmos.* **2009**, *114*, 363–369. [[CrossRef](#)]
61. Laurin, G.V.; Chen, Q.; Lindsell, J.A.; Coomes, D.A.; Frate, F.D.; Guerriero, L.; Pirotti, F.; Valentini, R. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *89*, 49–58. [[CrossRef](#)]
62. Li, W.; Niu, Z.; Wang, C.; Huang, W.; Chen, H.; Gao, S.; Li, D.; Muhammad, S. Combined Use of Airborne LiDAR and Satellite GF-1 Data to Estimate Leaf Area Index, Height, and Aboveground Biomass of Maize During Peak Growing Season. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4489–4501. [[CrossRef](#)]
63. Schaefer, M.; Lamb, D. A Combination of Plant NDVI and LiDAR Measurements Improve the Estimation of Pasture Biomass in Tall Fescue (*Festuca arundinacea* var. Fletcher). *Remote Sens.* **2016**, *8*, 109. [[CrossRef](#)]

64. Tilly, N.; Aasen, H.; Bareth, G. Fusion of Plant Height and Vegetation Indices for the Estimation of Barley Biomass. *Remote Sens.* **2015**, *7*, 11449–11480. [[CrossRef](#)]
65. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
66. Krahwinkler, P.; Rossman, J. Using decision tree based multiclass support vector machines for forest mapping. In Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 24–29 July 2011; pp. 307–318.
67. Hame, T.; Rauste, Y.; Antropov, O.; Ahola, H.A. Improved Mapping of Tropical Forests With Optical and SAR Imagery, Part II: Above Ground Biomass Estimation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 92–101. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).