Estimation of Forest Biomass Patterns across Northeast China Based on Allometric Scale Relationship

Xiliang Ni 1,†, Chunxiang Cao 1,*, Yuke Zhou 2,†, Lin Ding 1, Sungho Choi 3, Yuli Shi 4, Taejin Park 3, Xiao Fu 5, Hong Hu 6 and Xuejun Wang 7

1 State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China; nixl@radi.ac.cn (X.N.); dinglin@radi.ac.cn (L.D.)
2 State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; zyk@lreis.ac.cn
3 Department of Earth and Environment, Boston University, 675 Commonwealth Avenue, Boston, MA 02215, USA; schoi@bu.edu (S.C.); parktj@bu.edu (T.P.)
4 School of Remote Sensing, Nanjing University of Information Science and Technology, Nanjing 210044, China; yshi.nuist@gmail.com
5 College of Applied Sciences and Humanities of Beijing Union University, Beijing 100083, China; fuxiao@buu.edu.cn
6 Haihe basin Soil and Water Conservation Monitor Centre, Tianjin 300171, China; huhong-2004@163.com
7 Survey Planning and Design Institute, State Forest Administration of China, Beijing 100714, China; wangxuejun320@126.com
*
† These authors contributed equally to this work.

Received: 30 June 2017; Accepted: 1 August 2017; Published: 8 August 2017

Abstract: This study develops a modeling framework for utilizing the large footprint LiDAR waveform data from the Geoscience Laser Altimeter System (GLAS) onboard NASA’s Ice, Cloud, and Land Elevation Satellite (ICESat), Moderate Resolution Imaging Spectro-Radiometer (MODIS) imagery, meteorological data, and forest measurements for monitoring stocks of total biomass (including aboveground biomass and root biomass). The forest tree height models were separately used according to the artificial neural network (ANN) and the allometric scaling and resource limitation (ASRL) tree height models which can both combine the climate data and satellite data to predict forest tree heights. Based on the allometric approach, the forest aboveground biomass model was developed from the field measured aboveground biomass data and the tree heights derived from two tree height models. Then, the root biomass should scale with the aboveground biomass. To investigate whether this approach is efficient for estimating forest total biomass, we used Northeast China as the object of study. Our results generally proved that the method proposed in this study could be meaningful for forest total biomass estimation ($R^2 = 0.699$, RMSE = 55.86).

Keywords: forest aboveground biomass; root biomass; tree heights; GLAS; artificial neural network; allometric scaling and resource limitation

1. Introduction

As the principal part of terrestrial ecosystems, forest ecosystems hold approximately 80% of the terrestrial aboveground and below-ground biomass, and play very important roles in the global carbon cycle and climate change [1–3]. Since the temperate and boreal forests play a crucial role as atmospheric CO2 sinks, more attention has been paid to climatic warming in mid-and high-latitudes
than in low-latitudes [4–6]. Many studies have shown that the lack of accurate forest biomass maps has generated a large uncertainty in forest carbon stocks of temperate and boreal forest regions [7–9]. The accurate estimation of forest biomass is not only necessary for improving the estimation of carbon pools, but also very important for forest management and understanding the response to climate change [10–13]. At the same time, the estimation of root biomass is suggested to be helpful for improving the estimation of terrestrial carbon stocks [14–16]. A large number of biomass estimation methods that involve extrapolation of biomass measurements in sufficient number of field sample plots can achieve very good estimation results for small forest areas [17,18]. However, there are some problems related to obtaining reliable forest biomass estimation results in large scale studies, because of the lack of field data, inconsistency of data collection methods, and less consideration of root biomass [19]. The forest biomass estimation based on satellite remote sensing is considered as a fairly advantageous approach in large scale studies because of wide and synoptic data [4,20,21]. As there is no remote sensing instrument developed to have the capacity of measuring biomass directly, the ground inventory data is required for building relationships between remote sensing parameters and biomass. Although passive optical satellite data can be used for biomass estimation across a variety of spatial and temporal scales due to its inexpensive technology, the low saturation level of vegetation spectral bands has had a serious effect on biomass estimation [22–24]. LiDAR is often more accurate in predicting forest structural parameters because its signal does not saturate in high-biomass forests [25,26]. Especially, airborne LiDAR can obtain the forest biomass estimation results with high precision. However, the high acquisition costs of large observation data bring limitations of its application at large scale region [27,28]. The Geoscience Laser Altimeter System (GLAS), onboard the Ice, Cloud, and land Elevation Satellite (ICESat), is a spaceborne LiDAR system which can provide the global recording of full waveforms over large footprints [29,30]. The GLAS data can be used for estimating forest tree heights and biomass by associating field plot data, because it has the capability of obtaining the vertical structural information. However, relatively sparse GLAS footprints could not provide the continuous forest biomass distribution image without the help of other remote sensing datasets [30–32]. Therefore, it is necessary that multi-sensor datasets synergy is considered to estimate forest biomass at the regional scale.

The northeast part of China (NE China) is the most important forest region in China. NE China possesses the largest contiguous forest land area in China, accounting for 40% of total country forest biomass [33,34]. The forest of NE China was exploited much later than eastern and southern parts of the country. It has experienced drastic climatic warming since the 1980s, which brought a significant growth rate in forest biomass and productivity [35,36]. These changes also suggested that it is necessary to obtain the accurate forest biomass estimation.

In this study, we explored the capabilities of satellite remote sensing for mapping forest biomass in NE China region. The forest tree heights were estimated using artificial neural network (ANN) method and the allometric scaling and resource limitation (ASRL) approach by combining the climate data and satellite data, including ICESat/GLAS data, MODIS land surface products, and some ancillary datasets [37,38]. The potential information on root-shoot biomass allocation was studied using the field biomass measurements, and the best forest biomass estimation model was established.

2. Materials and Methods

2.1. Study Area

The research area of this study is NE China, which is defined here to extend longitudinally from 115°37’ E to 135°05’ E and latitudinally from 38°43’ N to 53°34’ N, with a total area of $1.66 \times 10^6$ km$^2$ (Figure 1). NE China in this study mainly includes Liaoning (LN), Jilin (JL), and Heilongjiang (HLJ) provinces and four leagues (an administrative division) in Eastern Inner Mongolia Autonomous region (IMA) [16].
The study area encompasses all the major forest types in Northeast Asia. The major forest types mainly include cold temperate coniferous forest, temperate coniferous and broad-leaved forests. The forested land area in NE China is about 50.5 million ha, accounting for the largest area of natural forests in the country. The elevation in most part of NE China is below 400 m (Figure 1). The DaXing’anling Mountain, Xing’anling Mountain and Changbai Mountain have relatively high elevation. The precipitation in NE China generally ranges from 300 mm to 1000 mm because of the monsoon climate of medium latitudes. The annual mean temperature ranged from −4 to 11.5 °C.

Three forest zones are covered from south to north in NE China, including warm temperate deciduous broadleaf forest zone, temperate coniferous and broadleaf mixed forest zone, and boreal forest zone [16,39].

![Figure 1](image.jpg) The study area and the biomass data collected in this study. These data are distributed well across the forest regions of the study area.

2.2. Datasets

2.2.1. Climate Data

The climatic variables used as the inputs of the two tree heights models in this study include precipitation, temperature, relative humidity, wind speed, and incoming solar radiation. The precipitation and temperature were the inputs of ANN tree heights model. These five climatic variables which determine the available and evaporative flow rate of trees were inputs of ASRL model. All these climate data were obtained from the China Meteorological Data Sharing Service System (CMDSSS) which provides the annual meteorological records [40,41].

In this study area, the observations spanning over recent 30 years (1978–2007) from total 118 meteorological stations were collected, comprised of annual total precipitation, average temperature, relative humidity, wind speed, and solar radiation time. We further interpolated these climatic variables to generate the continuous gridded maps of annual average climatic variables at 1km spatial resolution by using ordinary kriging [42]. Then, we averaged the continuous gridded maps of annual average precipitation and temperature from 2004 to 2006 in order to obtain mean precipitation and temperature maps during 2004–2006 as inputs of ANN tree heights model, respectively. Similarly, the mean values of five annual average climatic variables during 1978–2007 were obtained to be the inputs of ASRL model.
2.2.2. Land Surface Reflectance

The MODIS nadir bidirectional reflectance distribution function adjusted reflectance (NBAR) product which is named as MCD43B4 provides the land surface reflectance data at a spatial resolution of 1 km [43]. The NBAR product represents the best characterization of surface reflectance over a 16-day period. In this study, we selected the seven spectral bands of MCD43B4 data in growing season from 2004 to 2006 as the inputs of ANN tree heights model. The seven bands of MCD43B4 data used in this study include three visible bands (460, 555, and 659 nm) and four near-infrared bands (865, 1240, 1640, and 2130 nm). The MCD43B4 product can cover the study region by selecting the images with the scene orbiter h25v03, h25v04, h26v03, h26v04, h27v04, and h27v05.

2.2.3. Ancillary Data

In this study, there are two types of ancillary data. The first set of ancillary data was used as the inputs of ASRL-based tree height model, and consisted of the Advanced Space borne Thermal Emission and Reflection Radiometer (ASTER) Global DEM (GDEM V2, spatial resolution of 30 m), and the eight-day composites of the post-processed Moderate Resolution Imaging Spectro radiometer (MODIS) leaf area index (LAI) products (1 km spatial resolution) [43]. We further calculated the mean value of LAI in summer (June to August, the approximate growing season) for the time period from 2003 to 2006. The mean value of LAI was calculated by using the Equation (1):

\[
LAI_{\text{avg}} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} LAI_{8\text{-day},i,j}}{m \times n} \tag{1}
\]

Here, \(LAI_{\text{avg}}\) is the average LAI (June to August months from 2003 to 2006) and \(LAI_{8\text{-day},i,j}\) is the eight-day LAI product for \(i\)th week \((m = 12)\) of \(i\)th year \((n = 4)\).

In addition, the second set of ancillary data was used to identify forested lands, and consisted of land cover (LC) and vegetation continuous field (VCF) [44]. We selected the International Geosphere-Biosphere Programme (IGBP) of LC product (MCD12Q1, 500 m grid) and tree cover percent of VCF product (MOD44B, 250 m grid) of NE China for the year 2005 [45]. The land cover belonging to one of the five forest classes per the IGBP, and consisting of more than 50% tree cover per the VCF product were defined as forest land.

2.2.4. Tree Height and Biomass Measurements

GLAS-Derived Tree Heights

GLAS-Derived Tree Heights were used as the standard forest tree heights for training the ANN model and optimizing the parameters of ASRL based tree height model in this study.

The ICESat/GLAS is designed to obtain characteristics of the earth surface structures in three dimensions with high accuracy [26,38].

The Release-33 of GLAS laser altimetry data available from the National Snow and Ice Data Center [46] was used in this study. GLAS waveform data provide information on land elevation and vegetation cover within its ellipsoidal footprints about 72 diameters at about 170 m interval along the sub-satellite track [47,48]. National Snow and Ice Data Center released 15 GLAS products.

Amongst various altimetry products of the Release-33, we selected GLA14 product (GLAS Level-2 Land Surface Altimetry product, National Snow and Ice Data Center (NSIDC), Boulder, CO, USA) for the maximum tree height retrieval from May to October (2003–2006).

The GLAS waveform data are generally affected by the following factors: atmospheric forward scattering, signal saturation, background noise (low cloud), and the topographic slope gradient effects. In order to obtain the best GLAS data waveform for deriving accurate tree heights, the GLAS footprints screening is needed. At the same time, GLAS footprints over non-forest should be filtered. The GLAS footprints processing steps for selecting valid GLAS waveforms can be depicted as follows [37,38,47,48].
Firstly, the GLAS data of the approximate growing season from 2003 to 2006 were considered. Then, the GLAS data were further screened by applying the atmospheric forward scattering filters, signal saturation filters, background noise level correction filters, land cover mask conditions filters, and topographic slope gradient filter. Based on the procedure of GLAS footprints filtering, the best GLAS data waveform was selected.

In order to obtain the most accurate GLAS tree heights, the topographic correction approach of GLAS-derived tree heights was used according to the Equation (2) [37,38,41,47–50]:

$$H_{GLAS} = \left( D_{Sig\text{beg} Off} - D_{gp\text{cnt} \text{RegOff}} \right) - \frac{d_{GLAS} \times \tan \theta}{2}$$  \hspace{1cm} (2)

where $H_{GLAS}$ is the GLAS-derived tree heights, $D_{Sig\text{beg} Off}$ represents the location the GLAS full-waveform beginning signal, and $D_{gp\text{cnt} \text{RegOff}}$ represents the location of the ground peak, $d_{GLAS}$ is the GLAS footprint size, and $\theta$ is the topographic slope [40,48–52].

Field-Measured Tree Biomass

In this study, 515 field measured tree biomass measurement plots were compiled from NE China region. All these plots were compiled from literature by Wang [16]. For each plot, the total tree biomass, shoot (stem, branch and leaf) biomass and root biomass were calculated. In these plots, 85 plots were sampled by Wang [16]. The shoot and root biomass for the plots were estimated with DBH (and tree height) using allometric relationships developed by Zhu [53]. In addition, there were 161 plots obtained from the data base of Luo [54], and the rest of other plots were collected from 59 sources [16]. Since there are different numbers of effective records on the total biomass, shoot biomass and root biomass in Wang’s database, the plots which include all the useful records on the total biomass, shoot biomass, and root biomass were considered. Finally, 432 effective plots with valuable information of total biomass, shoot biomass and root biomass, were selected from the 515 plots to build and validate the biomass model. The detailed statistics information of these records is shown in Table 1.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Shoot Biomass (t/ha)</th>
<th>Root Biomass (t/ha)</th>
<th>Total Biomass (t/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum Value</td>
<td>369.1</td>
<td>106</td>
<td>432.4</td>
</tr>
<tr>
<td>Minimum Value</td>
<td>8.8</td>
<td>1.9</td>
<td>10.7</td>
</tr>
<tr>
<td>Mean Value</td>
<td>113.1044715</td>
<td>26.62682927</td>
<td>139.7052846</td>
</tr>
<tr>
<td>Variance</td>
<td>6088.720266</td>
<td>389.2960119</td>
<td>9226.870054</td>
</tr>
</tbody>
</table>

2.3. Methodology

We developed the approach of estimating forest biomass by combining multisource satellite data, meteorological data and field measured forest plot data.

Our analysis consisted of two main parts, (a) predicting the tree heights by using two tree heights models; (b) developing the biomass model for estimating forest biomass in study region (Figure 2).
This section describes tree height estimation methods used in this study and forest biomass modeling.

2.3.1. Tree Height Estimation Methods

Artificial Neural Network (ANN) Tree Heights Model Approach

The artificial neural network (ANN) tree heights model was proposed by Ni [38] for mapping the forest tree heights over China continent. We employed the ANN tree height model to obtain the tree heights in NE China. The inputs of ANN tree heights model consisted of 11 parameters, including climate variables (temperature, precipitation), ancillary data (land cover class, vegetation cover fraction), and seven multispectral bands of MODIS NBAR data. With the help of back-propagation (BP) process algorithm, the ANN tree height models was trained by using the GLAS-derived tree heights. The training and validation data pairs for each pixel of study region were selected according to the most similar climatic condition factor discipline according to described by Ni [38]. For each target pixel, one ANN-based tree heights model was trained by using 15 pairs of training data, while five pairs of validation data were used to prevent the over-fitting of the model. When the ANN-based tree heights model was trained well, it was used for estimating the forest canopy height over the target pixel.

Allometric Scaling and Resource Limitations (ASRL) Model Approach

We employed the ASRL model approach developed by Sungho et al. [41] to obtain the tree heights in NE China. The ASRL is a mechanism model based on a combination of the allometric scaling laws and local resource availability [41]. It can predict the potential maximum tree heights. A key hypothesis of the ASRL model is that a tree maintains its evaporative flow by collecting sufficient resources (water and light) while satisfying its basal metabolic needs, in turn, limiting maximum growth. The mechanism has a quantitative expression by the simple inequality \( Q_p \geq Q_e \geq Q_0 \). Here, the \( Q_p \) means the available flow rate, \( Q_e \) refers to the actual flow rate of a tree, and \( Q_0 \) represents the basic or required metabolic flow rate of a tree. Shi et al. [55] and Sungho et al. [41] proposed the parametric optimization methods of the ASRL model in order to obtain the actual tree heights. Shi et al. [55] recommend a parametric optimization, which can iteratively adjust three ASRL parameters to minimize disparities between the reference and modeled heights, and help improving the overall prediction accuracy. In this study, we applied the same processes to optimize ASRL model and obtain the actual tree heights in NE China.
2.3.2. Forest Biomass Modeling

Zianis and Mencuccini [56] developed an allometric approach to estimate aboveground forest biomass (M) by regarding M and corresponding tree diameter at breast height (DBH). The method can be described from the following allometric equation [34] (Equation (3)):

\[ M = a \times DBH^b \]  

(3)

where \( a \) is the allometric intercept and \( b \) is the allometric exponent. Based on the scale relationship between tree height (\( H \)) and DBH, the aboveground tree biomass estimation equation can be rewritten as Equation (4):

\[ M = a \times H^b + c \]  

(4)

where \( c \) is a constant, \( a \) is the allometric intercept and \( b \) is the allometric exponent.

In this study, we have used the field-measured biomass data and forest tree heights which were calculated by averaging the ANN-derived tree heights and ASRL tree heights to build the least-square regression model to obtain the coefficients of Equation (4).

In order to obtain the forest root biomass and the total biomass, we developed the allometric relationship equation based on the field-measured tree aboveground biomass and root biomass. Figure 3 shows the fitted allometric relationship equations described in Equations (3) and (4).

![Figure 3. The allometric relationships of forest tree heights, aboveground biomass and root biomass. (a) The allometric relationship between field aboveground biomass and height; and (b) the allometric relationship between the field aboveground biomass and root biomass.](image)

3. Results

3.1. Canopy Heights Mapping in NE China by Two Tree Height Methods

Following the artificial neural network (ANN) tree height model, a contiguous map of canopy heights with a spatial resolution of 1 km in NE China region was generated. Figure 4a shows the tree height map derived from the ANN model in NE China.

Similarly, a contiguous map of the tree heights over NE China was generated from the optimized ASRL mode by using the gridded climate and ancillary data mentioned in Section 2.2. This map (Figure 4b) closely represents the spatial distribution of tree heights (mean = 24.37, SD = 9.35) in NE China region with the spatial resolution of 1 km, showing a high correspondence to the spatial pattern of ANN tree heights (mean = 26.66 m, SD = 10.13). In order to investigate the accuracy of the two tree heights maps, we tested the tree heights by comparing with the GLAS-derived tree heights according...
The validation methods from Ni [37] and Ni [38], respectively. The validation results ($R^2 = 0.811$, RMSE = 4.79; $R^2 = 0.664$, RMSE = 5.24) demonstrate the effectiveness of the ANN- and ASRL-based tree heights models.

From the estimation result of the ANN and ASRL tree heights in NE China, we can see that relatively tall trees were growing in the eastern research areas. The trees distributed in Northwestern China show obviously lower heights than those in eastern regions.

**Figure 4.** The forest tree heights based on ANN and ASRL methods respectively. (a) ANN tree heights; (b) ASRL tree heights; (c) ANN based tree heights validation; (d) ASRL based tree heights validation.

### 3.2. Forest Biomass Estimation

We firstly estimated the forest aboveground biomass by using the allometric approach described in Section 2.3.2. Based on the forest aboveground biomass model described in Figure 3a, a contiguous map of forest aboveground biomass at a spatial resolution of 1 km was generated (Figure 5a).

In addition, the forest root biomass was estimated according to the relationship equation described as Figure 3b. The total biomass was calculated by summing the forest aboveground biomass and root biomass. The root biomass and total biomass map can be seen from the Figure 5c,e.

The estimation result shows that the forest aboveground biomass, root biomass, and total biomass have the same density distribution trend in the research region. The forests in the eastern research region have relatively large forest biomass density. The low biomass density was distributed in northwestern research region. The parts of the southeastern NE China (Eastern Liaoning Province) show the largest biomass density in the research area.
Figure 5. The estimation and validation results of forest aboveground biomass, root biomass, and total biomass in NE China region. (a) Forest aboveground biomass; (b) estimation validation of aboveground biomass; (c) forest root biomass; (d) estimation validation of root biomass; (e) forest total biomass; and (f) estimation validation of total biomass.

4. Discussion

A method to assess the forest biomass in NE China forests region was developed by combining satellite LiDAR data, optical remote sensing data, and field measurements. Due to the great difficulties in measuring root biomass for forest total biomass estimation, it is the primary method for estimating root biomass by building the allomeric relationship with aboveground biomass [54].
Based on the field measured forest aboveground biomass and root biomass, we obtained a relatively good allometric model for predicting root biomass ($R^2 = 0.761$) in the study area. However, there are some factors that influence the precision of root biomass prediction from forest aboveground biomass, including climate factors, and forest types. Especially at large scale, varying climate factors and forest types might cause the different allometric relationships between aboveground biomass and root biomass. Similarly, allometric models of aboveground biomass estimation from tree heights are also influenced by the climate factors and forest types.

In this case, different allometric models are necessary in the ecological zones which differ according to the climate factors and forest types. Generally, these allometric models would be optimized in these ecological zones. Due to the lack enough fields measured forest biomass and low resolution forest types data, in this study, we just used the united model which ignored the climate factors and forest types.

As shown in Figure 5, most of predicted and field biomass has good linear relationship, while there is a small quantity of points scattered out across the isoline. Although the validation result of forest biomass shows that our allometric metrics are efficient, some uncertainties occurred in the validation (Figure 5). The scattered points might be caused from the combination of several influenced factors, such as forest biomes, different measurement rules of field biomass data from various studies, and the propagation of errors from tree heights and biomass estimation models. The following points could be improved for higher precision of the forest biomass estimation in the future. Firstly, more effective geostatistics interpolation methods should be used for obtaining finer climate factor raster data, especially in the mountain areas which lack enough meteorological stations. Secondly, it is necessary to collect more field measured forest plots data, including the tree heights, aboveground biomass, root biomass, and total biomass. Finally, the allometric models for estimating aboveground biomass from tree heights and predicting root biomass from aboveground biomass should be respectively developed in the ecological zones which are divided based on the climate data and the forest type data.

5. Conclusions

We used two tree height models to obtain the forest tree heights map of NE China by combining GLAS data, meteorological data and optical remote sensing data. Based on the collected field measured forest plots data, we developed the allometric models for predicting the aboveground biomass and root biomass. Then, the forest total biomass map was generated by summing the aboveground biomass and root biomass.

The assessment of the three forest biomass map using field measured forest plots data was performed separately. At the forest sites level, high correlation appeared in the estimation results of aboveground biomass and total biomass ($R^2 = 0.695$, RMSE = 50.94; $R^2 = 0.699$, RMSE = 55.86). The estimation results of root biomass showed, relatively, a slightly lower correlation with the field measured data ($R^2 = 0.695$, RMSE = 12.22). In summary, this study demonstrated that forest aboveground biomass and root biomass can be estimated with relatively high precision by combining satellite LiDar data, MODIS data, and climate data in regional scale.

This study highlights the potential to estimate forest total biomass in conjunction with novel and efficient methods, such as ANN tree heights metrics, ASRL-based tree height models and allometric scaling relationships among tree heights, forest aboveground biomass, and root biomass. Combinations of these techniques were able to quantify the forest structure parameters and biomass. Application of the approach proposed in this study at the national scale would provide an opportunity to understand carbon sinks of all forest land in China. These relatively fine-scaled, spatially-explicit forest biomass maps can provide critical information for forest carbon cycle studies and forest resource management.

Acknowledgments: The authors would like to thank the three anonymous reviewers whose comments significantly improved this manuscript. This study was partially funded by the Special Fund for Forest Scientific Research in the Public Welfare (grant no. 201504323), the National Key Research and Development Program of China (grant no. 2016YFB0501505), the Special Fund for the Ecological Assessment of Three Gorges Project
Author Contributions: The analysis was performed by Xiliang Ni. All authors contributed with ideas, writing, and discussions.

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

References


