Evaluation of Four Methods for Predicting Carbon Stocks of Korean Pine Plantations in Heilongjiang Province, China

Huilin Gao¹, Lihu Dong¹, Fengri Li¹*, Lianjun Zhang²

¹ Department of Forest Management, School of Forestry, Northeast Forestry University, Harbin, Heilongjiang, People’s Republic of China, ² Department of Forest and Natural Resources Management, College of Environmental Science and Forestry, State University of New York, Old Westbury, NY, United States of America

* fengrili@126.com

Abstract

A total of 89 trees of Korean pine (Pinus koraiensis) were destructively sampled from the plantations in Heilongjiang Province, P.R. China. The sample trees were measured and calculated for the biomass and carbon stocks of tree components (i.e., stem, branch, foliage and root). Both compatible biomass and carbon stock models were developed with the total biomass and total carbon stocks as the constraints, respectively. Four methods were used to evaluate the carbon stocks of tree components. The first method predicted carbon stocks directly by the compatible carbon stocks models (Method 1). The other three methods indirectly predicted the carbon stocks in two steps: (1) estimating the biomass by the compatible biomass models, and (2) multiplying the estimated biomass by three different carbon conversion factors (i.e., carbon conversion factor 0.5 (Method 2), average carbon concentration of the sample trees (Method 3), and average carbon concentration of each tree component (Method 4)). The prediction errors of estimating the carbon stocks were compared and tested for the differences between the four methods. The results showed that the compatible biomass and carbon stock models with tree diameter (D) as the sole independent variable performed well so that Method 1 was the best method for predicting the carbon stocks of tree components and total. There were significant differences among the four methods for the carbon stock of stem. Method 2 produced the largest error, especially for stem and total. Methods 3 and Method 4 were slightly worse than Method 1, but the differences were not statistically significant. In practice, the indirect method using the mean carbon concentration of individual trees was sufficient to obtain accurate carbon stocks estimation if carbon stocks models are not available.

Introduction

Carbon stocks estimation for the plot or the national level with reliable and verifiable techniques have always increased the social concerns [1–2], whereas the biomass calculation is the
basis [3]. Compared to the popularized technique of remote sensing in biomass/carbon stocks monitoring and estimation, the traditional statistical technique which expensively and labour-intensively depends on the destructive sampling is still widely accepted due to high accurate estimation on biomass/carbon stocks at both tree and forest levels [4–6]. Allometric model always showed good performance for components (i.e., stem, branch, foliage and root) biomass prediction of individual trees or forest and carbon stocks will be converted by the estimated biomass combining with carbon conversion factor [7]. However, the accuracy of this method is in great demand to the accuracy of the carbon conversion factor for that some of the conversion factors were only empirical (e.g., 0.5) or simply conducted by species-specific of the trees [8]. As a result, incorporating the comparison of biomass model into the carbon concentration analysis is meaningful to the carbon stocks estimation.

A common practice for estimating tree carbon stocks is that tree biomass (total and components) models are first developed based on field-sampled data, and then carbon stock is calculated by multiplying the model-predicted biomass by a conversion factor of carbon concentration. The acceptable conversion factors include (1) the carbon concentration of 50% for woody tissues and 50% for non-woody tissues, or (2) 50% for woody and 45% for non-woody tissues [9–15]. In addition, carbon stock models were also established to directly predict carbon stocks of tree and components and total [16]. However, some noteworthy issues and problems still existed in these methods. Some reported biomass and carbon models may not satisfy the additivity or compatibility property among tree components and total equations. The impact of the prediction errors of tree biomass on estimating carbon stock by using conversion factors is lack of quantification.

Over the last decades, numerous allometric biomass models have been developed for many tree species around the world [17–21]. To ensure the additivity or compatibility of tree total and component biomass models, a variety of parameter estimation methods have been proposed [22–24]. Generalized method of moments (GMM) has been proven the efficiency in parameter estimation without the specification for the nature of the heteroscedasticity [25–27]. Maximum likelihood (ML) [23] and two-stage error-in-variable model (TSEM) [24] has been also employed and given the introductions in the researches [27]. However, seemingly unrelated regressions (SUR) [22, 25] and non-linear seemingly unrelated regressions (NSUR) [23] have been proven more general and flexible in recent years [28].

However, there have been limited efforts of developing tree carbon models, compared to biomass models, using tree attributes such as tree diameter and height [16]. Therefore, the carbon stocks of tree components and total are computed as the tree biomass predictions multiplied by carbon conversion factors, either an acceptable common constant (e.g., 0.5) or an empirical constant based on available data. Although this method worked well in some cases [16], there is lack of effort for quantitatively evaluating the error sources of estimating carbon stocks due to the biomass models and different carbon conversion factors, i.e., the prediction error through a sequence of operations on tree carbon stocks estimations. Furthermore, to our best knowledge, the additive system of biomass models performed well in estimating biomass and carbon stocks of total and tree components [29].

Korean pine (Pinus koraiensis Sieb. et Zucc) forests mainly distribute in Xiaoxing'an Mountain, and Changbai Mountain in Northeast of China. It is one of the important tree species for plantations in the area. In addition to the good quality of its lumber for industrial uses, its seeds are extensively harvested and sold as pine nuts, which have been the most widely traded pine nut in international commerce. The objectives of this study were (1) to develop additive systems of tree biomass and carbon stock models, (2) to evaluate four methods for estimating the carbon stocks of tree total and components, including a direct method (carbon stock
models), and three indirect methods (i.e., 3 carbon conversion factors), (3) to compare and test the differences between the four methods.

**Data and Method**

**Biomass and carbon stock data**

The study area is located in Heilongjiang province in Northeast China (43°25' – 53°33' N, 121° 11’ – 135°05' E) in which the forest area is 19.62 million hectares and the forest coverage is 43.16%. 5 counties were included in the study area (Fig 1) and the mean annual rainfall is 485 – 577mm and the mean annual temperature of 1.69 – 3.60°C. A total of 17 plots were selected in the different age of forest stand with a total of 89 Korean pine trees were destructively sampled in our study. The characteristics of the stand level for the study area were listed in Table 1. The data collection process in each site has been permitted by the local authority of each location. The names of the six official organizations were Mengjiagang forest farm, Boli forest bureau, Acheng forest bureau, Qing'an forest bureau and Jidong forest bureau. The whole process didn’t damage and disturb the natural environment. All the location of our study was carried out in the state-owned forest land not the private forest.

The sample trees were destructively felled at the ground level with tree diameter at breast height (D, cm), total tree height (H, m) and length of live crown (CL, m) measured and recorded in the field immediately. Each stem was cut into sections by the length of 1 m with each section weighed and recorded. Then, the live crowns were stratified into three layers along the stem. In each crown layer, 3 – 5 branches were selected as the samples and the branches and foliage were separated and weighed, respectively. We sampled about 50 – 100 g branches and foliage with the record of fresh weight and then took the samples to laboratory for moisture content determination. The tree roots were excavated using both a lifting machine and manual digging. The zone of excavating roots was limited to a circle of 3 m in radius, and the fine roots (< 5 mm) were not included. All roots of the sampled trees were divided into three categories: large roots (diameter ≥ 5 cm), medium roots (diameter 2 – 5 cm), and small roots (diameter < 2 cm). Each root class was sampled (about 100 – 200 g), weighed, and taken to laboratory for moisture content determination. All stem, branch, foliage, and root samples were oven-dried at 80°C and weighted. The dry biomass of each component was calculated by multiplying the fresh weight of each component by the dry / fresh ratio of each component. For each sample tree, the sum of branch dry biomass and foliage dry biomass yielded the crown dry biomass. The sum of crown dry biomass and stem dry biomass gave the aboveground biomass. The sum of aboveground dry biomass and root dry biomass produced the total tree biomass (Fig 2). For each oven-dried sample of stem, branch, foliage and root approximately 50 mg was used for measuring carbon concentration using a Multi N/C 3000 analyzer with 1500 Solids Module (Analytik Jena AG, Germany). The samples were then burned completely at 1200°C in a vial containing pure oxygen, and the carbon emitted was measured with a non-dispersion infrared ray (NDIR) analyzer. The carbon stock of each component was then calculated by multiplying the biomass of each tree component by the respective carbon concentration. Thus, the carbon stock of individual tree was obtained by summing the component estimates. The carbon concentration for the each component and total tree were showed as Fig 3.

**Model specification and estimation**

The data used to fit the models has been supported by the official authority and has no negative effect on the environment of human being and habitat of wildlife. Over the last two decades, researchers have tried different allometric equations for modeling tree biomass. The following
equation was considered excellent in performance and remarkable in flexibility [22, 30, 31].

\[ Y = aX^b \]

(1)

where \( Y \) is response variable (tree biomass or carbon stocks), \( X \) is independent variable, \( a \) and \( b \) are model parameters to be estimated. Diameter at breast height which was defined to be 1.3m
above the ground level, total tree height (H) and the combination forms were widely employed in the models. In our study, variable D and the combination form D^2H were used to develop biomass and carbon stock models so that X can be D or D^2H.

The additive systems of tree biomass and carbon stock models were formulated based on the methods by Tang and Li [32], in which the total tree biomass and carbon stocks were used as the constraints. The compatible biomass and carbon stock model which was derived from

| Study area | Number of plots | Plot size (ha) | Age (year) | Density (trees ha^-1) | Mean DBH (cm) | Slope (°) | Attitude (m) |
|------------|----------------|---------------|------------|-----------------------|---------------|-----------|--------------|
| Huanan     | 12             | 0.06–0.09     | 32–47      | 650–1650              | 15.5–22.0     | < 5       | 194–263      |
| Boli       | 2              | 0.27          | 35–44      | 2130–2135             | 12.6–12.7     | 12        | 382          |
| Acheng     | 1              | 0.27          | 27         | 1470                  | 12.7          | 0         | 269          |
| Qing'an    | 1              | 0.27          | 32         | 804                   | 17.4          | < 5       | 228          |
| Jingdong   | 1              | 0.27          | 37         | 1581                  | 16.6          | 12        | 430          |

doi:10.1371/journal.pone.0145017.t001

Fig 2. DBH (cm), total tree height (H, m) and components and total biomass (SB, BB, FB, RB and TB are stem, branch, foliage, root and total tree biomass in kg).

doi:10.1371/journal.pone.0145017.g002
Eq 1 was showed as Eq 2.

\[
\begin{align*}
Y_{si} &= a_i X_{bi} / (1 + r_{1i} X_{ki} + r_{2i} X_{bi}^{2} + r_{3i} X_{ri}^{2}) \\
Y_{bi} &= a_i r_{1i} X_{bi} / (1 + r_{1i} X_{ki} + r_{2i} X_{bi}^{2} + r_{3i} X_{ri}^{2}) \\
Y_{fi} &= a_i r_{2i} X_{bi} / (1 + r_{1i} X_{ki} + r_{2i} X_{bi}^{2} + r_{3i} X_{ri}^{2}) \\
Y_{ri} &= a_i r_{3i} X_{bi} / (1 + r_{1i} X_{ki} + r_{2i} X_{bi}^{2} + r_{3i} X_{ri}^{2})
\end{align*}
\]  

(2)

where \(Y_{si}, Y_{bi}, Y_{fi}, Y_{ri}\) are biomass or carbon stock for the stem, branch, foliage and root respectively. \(a_i, b_i, r_{1i}, r_{2i}, r_{3i}, k_{1i}, k_{2i}, k_{3i}\) and \(k_{3i}\) are model parameters to be estimated using the jackknifing technique. The subscript \(i = 1, 2\) were biomass models based on \(Y = aD^b\) and \(Y = a(D^2H)^b\) which were defined as BM1, BM2 respectively and \(i = 3, 4\) were carbon stock models based on \(Y = aD^b\) and \(Y = a(D^2H)^b\) which were defined as CM3, CM4. The detailed derivation process for Eq 2 was showed in the Appendix.

Model [2] comprises a system of four nonlinear equations in our study. NSUR was applied to estimate the model coefficients in these nonlinear simultaneous equations using SAS/ETS PROC MODEL procedure [26]. NSUR employs the nonlinear joint generalized least squares to...
explain the correlations in the system of nonlinear biomass and carbon stock models [23, 25, 28]. Besides, NSUR also allow each model of the system to specify its own weighting function to overcome heteroscedasticity. To overcome the heteroscedasticity in biomass and carbon stock model, weighted functions were multiplied at the two side of the non-linear function with 1/D^2 and (D^2H)^p so that this problem can be resolved. The weighting function of each tree component for BM1 and CM3 was Y_i = resid.Y_i/D^p (i = 1, 3 which has been defined above) and for BM2 and CM4 was Y_i = resid.Y_i/(D^2H)^p. Error variance model e_i = x_p was developed (e_i is the model residual of the unweighted model, x is D or D^2H) in which p was determined for each component [28].

Model assessment and validation

In this study, jackknifing technique was used to assess the model parameters and validation model performance. The entire data size by leaving one out (n-1, n is the sample size) were used to estimate the model parameters, thus the one left sample not used in the model fitting was employed to predict by the fitted model. The \( R^2 \) (Eq 3) for each component and the total and the prediction for each left out independent variable were calculated for each round of fitting. The mean and standard deviation for each parameter were calculated. \( R^2 \) was used to assess the goodness-of-fit of the model. Due to the good performance in incorporating both variance of prediction error and the bias of prediction, the root mean squared error of prediction (RMSEp) was used as the model validation criteria [33, 34]. As a result, a total of five statistics (Eqs 5–9) were used to validate the model performance.

Adjusted coefficient of determination:

\[
R^2_a = 1 - (1 - R^2) \left( \frac{n - 1}{n - p - 1} \right)
\]

where

\[
R^2 = 1 - \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_{i-1})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}
\]

Root mean squared error of prediction:

\[
RMSE_p = \sqrt{\frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_{i-1})^2}{n - p}}
\]

Mean prediction error:

\[
ME = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_{i-1})
\]

Mean prediction error percent:

\[
ME\% = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{Y_i - \hat{Y}_{i-1}}{Y_i} \right) \times 100
\]
Mean absolute error:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |Y_i - \hat{Y}_{i-1}|$$  \hspace{1cm} (8)

Mean absolute percent error:

$$\text{MAE\%} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Y_i - \hat{Y}_{i-1}}{Y_i} \right| \times 100$$  \hspace{1cm} (9)

Where $Y_i$ is the $i$th observed value of response variable (i.e., biomass or carbon stock), $\hat{Y}_{i-1}$ is the $i$th predicted value for the dependent variable without inclusion in the fitting dataset, $\bar{Y}$ is the mean of observed response variable, $n$ is sample size, and $p$ is the number of model parameters.

Comparison of carbon stock estimation by four methods

Four methods were used in this study to evaluate the carbon stocks of tree components and total: (1) the developed additive systems of carbon models (CM3 and CM4) were directly used to compute the carbon stocks of tree components and total, given tree D and H (Method 1); (2) the developed additive systems of biomass models (BM1 and BM2) were first used to estimate the biomass of tree components and total, and then the predicted biomass was multiplied by the carbon conversion factor 0.5 (Method 2); (3) the developed additive systems of biomass models were first used to estimate the biomass of tree components and total, and then the predicted biomass was multiplied by the average carbon concentration of total tree (i.e., 0.48; Fig 3) (Method 3); and (Eq 4) the developed additive systems of biomass models were first used to estimate the biomass of tree components and total, and then the predicted biomass was multiplied by the average carbon concentration of each tree component (i.e., stem 0.47, branch 0.48, foliage 0.49, and root 0.48; Fig 3) (Method 4). The last three methods were considered as the indirect methods.

Relative root mean squared error (RMSE$_r$) (Eq 10) was used to compare the prediction error of the four methods. Thus, the relative root mean squared error for carbon (RMSE$_{rc}$) was calculated and RMSE$_{rc}$ of indirect method was based on Eq 11.

$$\text{RMSE}_r = \sqrt{\frac{1}{n-p} \sum_{i=1}^{n} (Y_i - \hat{Y}_{i-1})^2} \cdot \frac{1}{\bar{Y}} \times 100$$  \hspace{1cm} (10)

$$\text{RMSE}_{rc} = \sqrt{\frac{1}{n-p} \sum_{i=1}^{n} (C_{mi} - \hat{C}_{mi - mi})^2} \cdot \frac{1}{\bar{C}_m} \times 100$$  \hspace{1cm} (11)

where $C_{mi}$ is the $i$th observed value of carbon stock for the $m$ tree components ($m = 1, 2, 3,$ and 4 denotes stem, branch, foliage and root, respectively), $\hat{C}_{mi - mi}$ is the $i$th predicted value of carbon stock without inclusion in the fitting dataset for the $m$ tree components. $\bar{C}_m$ is the mean of predicted value of carbon stock for the $m$ tree components with the 89 rounding processes, and the other symbols were defined as above.
Given Eq 11, RMSErc of indirect method was derived as Eqs 12 and 13.

\[
\text{RMSE}_{rc} = \frac{1}{\sqrt{n - p}} \cdot \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left( W_{mi} \cdot C_{mi}\% - \hat{W}_{mi} \cdot C_{mk}\% \right)^2} \times 100
\]  

(12)

\[
\text{RMSE}_{rc} = \frac{1}{\sqrt{n - p}} \cdot \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left( W_{mi} \cdot C_{mi}\% - \hat{W}_{mi} \cdot C_{mk}\% \right)^2} \cdot \frac{1}{W_m} \times 100
\]  

(13)

Where \( W_{mi} \) is the \( i \)th observed value of biomass for the \( m \) tree components (\( m = 1, 2, 3, \) and \( 4 \) denotes stem, branch, foliage and root, respectively), \( W_{mi-mu} \) is the \( i \)th predicted value of biomass without inclusion in the fitting dataset for the \( m \) tree components, \( \hat{W}_m \) is the mean of predicted value of biomass of the \( m \) tree components with the 89 rounding processes, \( C_{mi}\% \) is the \( i \)th observed value of carbon concentration for the \( m \) tree components, \( C_{mk}\% \) is the \( k \)th carbon conversion factor used in the three indirect methods for the \( m \) tree components (\( k = 1 \) denotes carbon conversion factor 0.5, \( k = 2 \) denotes the average carbon concentration of total tree 0.48, and \( k = 3 \) denotes the average carbon concentration of each tree component).

To test whether the significant difference will be produced by the four methods for predicting carbon stocks of the tree components and total, the analysis of variance with randomized completed block design (RCBD) was used [35–37], in which the treatment was the four methods and blocking was each tree because the four methods were applied to each tree simultaneously. The least significant difference (LSD) was used for multiple mean comparisons. The PROC GLM procedure of SAS 9.3 was used for computation [38, 39], and the significance level was set at \( \alpha = 0.05 \).

**Results**

**Model fitting and validation**

The mean values and standard deviations of the each parameter for the compatible biomass models (i.e., BM1 and BM2) and the carbon stock models (i.e., CM3 and CM4) using jackknifing technique were listed in Table 2. The mean values and standard deviation of \( R_s^2 \) for compatible biomass and carbon stock models were listed in Table 3. Validation results based on jackknifing technique for compatible biomass models were showed in Table 4 and for compatible carbon stock models were showed in Table 5. The results indicated that the mean of \( R_s^2 \) of all the components and total for BM1 and BM2 had \( R_s^2 \geq 0.75 \), and \( R_s^2 \geq 0.73 \), respectively, while the jackknifing validation results were RMSEP \( \leq 22.10 \text{ kg} \), ME \( \leq \pm 0.28 \text{ kg} \) (ME% \( \leq \pm 7.24\% \)), MAE \( \leq 12.89 \text{ kg} \) (MAE \( \leq 28.01\% \)) for all components and total of BM1 and RMSEP \( \leq 23.89 \text{ kg} \), ME \( \leq \pm 0.31 \text{ kg} \) (ME% \( \leq \pm 8.19\% \)), MAE \( \leq 13.67 \text{ kg} \) (MAE \( \leq 31.01\% \)) for BM2, respectively. Similarly, the \( R_s^2 \) of tree components and total for CM3 and CM4 had \( R_s^2 \geq 0.74 \), and \( R_s^2 \geq 0.74 \), respectively, while the validation results were RMSEP \( \leq 10.59 \text{ kg} \), ME \( \leq \pm 1.21 \text{ kg} \) (ME% \( \leq \pm 13.99\% \)), MAE \( \leq 6.54 \text{ kg} \) (MAE \( \leq 28.49\% \)) for CM3 and RMSEP \( \leq 10.66 \text{ kg} \), ME \( \leq \pm 1.32 \text{ kg} \) (ME% \( \leq \pm 16.14\% \)), MAE \( \leq 6.89 \text{ kg} \) (MAE \( \leq 30.10\% \)) for CM4 respectively.

It was evident that BM2 fitted the data better (higher \( R_s^2 \)) than did BM1 for stem and root, indicating that adding tree H into the biomass model increased the explanation power of these two components. The jackknifing validation also proved this conclusion with significant larger RMSEP, ME and MAE for stem and root for BM1 compared to BM2. However, BM2 was much inferior to BM1 when considering the branch, foliage and the total. For CM3 and CM4, the
comparison results explained the similar conclusion to BM1 and BM2. However, BM2 and CM4 have much model parameters to be estimated and the correlations among these parameters could not be ignored. Thus, we preferred to specify the compatible biomass and carbon stock models using $D$ as the only independent variable to make methods comparison.

Comparison of four methods in carbon stock prediction

Another statistic, namely, relative root mean squared error for carbon ($\text{RMSE}_{rc}$) was derived to compare the carbon stock prediction of four methods [9, 16, 18, 40]. The $\text{RMSE}_{rc}$ of all tree components and total with the four methods was showed as Fig. 4. For tree stem, Method 1 was the best with the lowest $\text{RMSE}_{rc}$, followed by Method 4, Method 3, and Method 2. For tree branch, Method 1 was the best method with the lowest $\text{RMSE}_{rc}$, followed by Method 4, Method 2 and Method 3. As for foliage, Method 1 was also the best, followed by Method 4, Method 3 and Method 2. The $\text{RMSE}_{rc}$ of Method 1 for root was slightly larger than Method 2, and still lower than Method 3 and Method 4. For the total, Method 1 was notably lower than Method 2 but slightly larger than Method 3 and Method 4.

To test the differences among the four methods, the prediction errors were used as the response variable using analysis of variance (ANOVA) with least significant difference (LSD) for multiple mean comparisons. The test results were showed in Table 6. The F-test in ANOVA was statistically significant ($p$-value < 0.01) for stem, branch, foliage, root and total.

Table 2. The mean and standard deviation of the parameter for biomass and carbon stock models.

| Model parameter | BM1          | BM2          | CM3          | CM4          |
|-----------------|--------------|--------------|--------------|--------------|
|                 | Mean | Std | Mean | Std | Mean | Std | Mean | Std |
| $a_i$           | 0.0891 | 0.0042 | 0.0374 | 0.0042 | 0.0451 | 0.0009 | 0.0201 | 0.0034 |
| $b_i$           | 2.4560 | 0.0128 | 2.5120 | 0.0201 | 2.5001 | 0.0108 | 0.9814 | 0.0204 |
| $r_{1i}$        | 0.0009 | 0.0011 | 0.0751 | 0.0012 | 0.0043 | 0.0012 | 0.1021 | 0.0075 |
| $r_{2i}$        | 0.0240 | 0.0025 | 0.2210 | 0.0032 | 0.0302 | 0.0082 | 0.4441 | 0.0037 |
| $r_{3i}$        | 0.0892 | 0.0141 | 0.0848 | 0.0142 | 0.0714 | 0.0212 | 0.1541 | 0.0850 |
| $k_{1i}$        | 1.8254 | 0.0014 | 0.4701 | 0.0021 | 1.2801 | 0.0310 | 0.1021 | 0.0454 |
| $k_{2i}$        | 0.6090 | 0.0347 | -0.1302 | 0.0420 | 0.5205 | 0.0296 | -0.1213 | 0.0421 |
| $k_{3i}$        | 0.4901 | 0.0424 | 0.4924 | 0.0801 | 0.5713 | 0.0585 | 0.1091 | 0.0577 |

The weighting factors were $D^{2.33}$, $D^{2.01}$, $D^{1.67}$, $D^{2.19}$ for BM1, $(D^2H)^{0.77}$, $(D^2H)^{0.82}$, $(D^2H)^{0.60}$, $(D^2H)^{0.75}$ for BM2, $D^{2.22}$, $D^{2.13}$, $D^{1.93}$, $D^{2.28}$ for CM3 and $(D^2H)^{0.60}$, $(D^2H)^{0.67}$, $(D^2H)^{0.72}$, $(D^2H)^{0.81}$ for CM4 for stem, branch, foliage and root respectively.

doi:10.1371/journal.pone.0145017.t002

Table 3. The mean and standard deviation of $R^2$ for biomass and carbon stock models.

| Models | Statistics | Stem | Branch | Foliage | Root | Total |
|--------|------------|------|--------|---------|------|-------|
| BM1    | Mean       | 0.88 | 0.78   | 0.79    | 0.75 | 0.90  |
|        | Std        | 0.0021 | 0.0112 | 0.0074 | 0.0079 | 0.0019 |
| BM2    | Mean       | 0.91 | 0.73   | 0.76    | 0.81 | 0.89  |
|        | Std        | 0.0013 | 0.0114 | 0.0061 | 0.0094 | 0.0069 |
| CM3    | Mean       | 0.91 | 0.77   | 0.76    | 0.74 | 0.89  |
|        | Std        | 0.0011 | 0.0062 | 0.0051 | 0.0041 | 0.0010 |
| CM4    | Mean       | 0.93 | 0.75   | 0.74    | 0.77 | 0.88  |
|        | Std        | 0.0011 | 0.0058 | 0.0049 | 0.0047 | 0.0008 |

doi:10.1371/journal.pone.0145017.t003
tree carbon prediction. The blocking (trees) was also significant, indicating the differences among the sampled trees were also relatively large.

The multiple mean comparison results indicated that (1) for stem, Method 1 was significantly different from Method 2 and Method 3, but there was no difference between Method 3 and Method 4, as well as no difference between Method 1 and Method 4 (Table 6). One the other hand, for branch, foliage, and root, there was no difference between the four methods. For the total, there was no significant difference between Method 2 and Method 3, and no difference among Method 1, Method 3 and Method 4 (Table 6).

Discussion

To investigate which model is better for fitting the biomass and carbon stock data, tree diameter alone and the combination of diameter and height were used as the predictor variables for predicting the tree component and total biomass and carbon stocks in this study. Some researchers also used other tree variables such as crown width, stand age, wood density, diameter outside bark at the base of the live crown, crown height, and etc., which may performed well in some studies [18, 23, 41–44]. However, given the practical difficulties and costs of obtaining crown measurements, tree diameter and height are the most frequently used predictors in model construction. In our study, adding H into the biomass and carbon stock models only increased the $R^2$ of stem and root because the differences of tree height at the same tree diameter contributed more information for the biomass/carbon stocks of stem and root [45].

Table 5. Jackknifing validation for compatible carbon stock models.

| Model | Component | RMSE_p(kg) | ME(kg) | ME% | MAE(kg) | MAE% |
|-------|-----------|------------|--------|-----|---------|------|
| CM3   | Stem      | 6.38       | -0.18  | -3.18 | 4.33    | 13.10|
|       | Branch    | 5.29       | -0.05  | -13.99| 3.31    | 26.99|
|       | Foliage   | 2.20       | 0.02   | 8.63  | 1.51    | 18.66|
|       | Root      | 7.08       | 0.11   | 6.89  | 4.12    | 28.49|
|       | Total     | 10.59      | 1.21   | 4.30  | 6.54    | 16.19|
| CM4   | Stem      | 5.15       | 0.11   | 1.24  | 2.70    | 8.65 |
|       | Branch    | 5.38       | -0.08  | -16.14| 3.51    | 30.10|
|       | Foliage   | 2.65       | -0.03  | -9.20 | 1.63    | 18.75|
|       | Root      | 6.48       | 0.09   | 6.30  | 4.10    | 28.05|
|       | Total     | 10.66      | 1.32   | 4.57  | 6.89    | 17.15|

doi:10.1371/journal.pone.0145017.t005
and the estimation errors of other tree components (branch, foliage) were increased consistent with the decreased of $R^2_a$ after adding H into the models (Tables 3, 4 and 5). Therefore, we chose D as the sole predictor variable in this study to develop the biomass and carbon stock models. To ensure the additivity of the tree components and total, the compatible biomass and carbon stock models were constructed with the total tree biomass and carbon stock as the constraints based on the allometric equation [32]. Furthermore, the $R^2_a$ of some tree components might be sacrificed to ensure the additivity for the components and total compared to fit each component separately. Given NSUR is more flexible and has been widely accepted by forest modelers [25, 27, 28], we used NSUR as the only parameter estimation method in this study and did not compare NSUR against others parameter estimation methods.

Currently, the carbon stock calculation is based on either observed carbon concentrations [46] or a commonly accepted carbon conversion factor 0.5, which is multiplied by the estimated tree biomass [47]. However, the prediction error using indirect method is lack of evaluation. In this study we used RMAErc to compare and evaluate carbon stocks of tree components

![Fig 4. RMAErc of four methods for the carbon stock prediction of tree components and total. The error bars correspond to the lower and upper limits of the 95% confidence intervals.](image)

doi:10.1371/journal.pone.0145017.g004

Table 6. Comparison of treatment means between four methods of carbon stocks evaluation for each component and total, in which the same letters indicate the difference between treatments are not statistically significant.

| Method | Stem | Branch | Foliage | Root | Total |
|--------|------|--------|---------|------|-------|
| Method 1 | A    | C      | A       | A    | B     |
| Method 2 | C    | A      | A       | A    | A     |
| Method 3 | B    | A      | A       | A    | A     |
| Method 4 | B    | C      | A       | A    | B     |

doi:10.1371/journal.pone.0145017.t006
of individual trees for the indirect method. The statistical tests (ANOVA with LSD) indicated that, for stem, the indirect method using the carbon conversion factor 0.5 produced significantly larger errors than the other two methods. Consequently, the larger prediction error for stem would surely lead to the large prediction error for the total biomass because it is known that more than 50% of the total biomass is located in the stem [16]. It seemed that the other two indirect methods, using either the average carbon concentration of total tree (0.48) (Method 3) or the average carbon concentration of each tree component (stem 0.47, branch 0.48, foliage 0.49, and root 0.48) (Method 4), were superior to the carbon conversion factor 0.5 (Method 2). Although the differences between Method 3 and Methods 4 was not significant, Method 4 would require considerably more time and effort to collect in the field. Thus, Method 3 (the average carbon concentration of total tree) should be sufficient to obtain good carbon stock prediction with the estimated tree biomass from the models. For branch, foliage and root, the prediction error among the four methods was not significant, this can be explained that the real carbon concentration of the three components are more close to the total tree concentration. Lehtonen et al. (2007) [48] analyzed the prediction error of biomass expansion factors and the error from variable measurements and regression functions. Their results showed that most of the error was due to the regression models. Therefore, it is crucial to develop and select the “best” biomass models as the foundation for computing carbon stocks using indirect methods. In our study, we used diameter at breast height and the combination of diameter at breast height and tree height to establish the biomass and carbon stock models. Based on goodness-of-fit and validation results for the different kinds of models, we selected the model which employed diameter at the breast height to make comparison among the four methods to decrease the error of indirect methods deriving from the biomass model as large as possible. There were no significant differences between Method 3 and Method 4 according to the total prediction error estimation indicating that using observed average carbon concentration of individual tree and / or tree components would lead to better prediction for carbon stocks. We concluded that it was not necessary to measure the carbon concentrations for tree components because they are time-consuming and costly. The average carbon concentration of individual tree should be sufficient to obtain accurate computation for carbon stocks given reliable and accurate estimation of biomass.

The results of our study were different from the study of [16], in which they used three methods: 1) estimated total biomass was multiplied by weighted mean carbon concentration; 2) estimated tree compartment biomass was multiplied by average carbon concentration of tree components; and 3) developing carbon stocks prediction model directly. Mello et al. (2012) [16] found that there was no statistical difference among the three methods. However, there were only 30 sample trees included in Mello et al. (2012) [16]. The small sample size might lead to an insignificant comparison. In addition, the biomass and carbon models developed in Mello et al. (2012) [16] did not hold the additivity or compatibility of tree components and total. In addition, the differences of prediction errors among the three methods at tree component levels were not clear in the study of Mello et al. (2012) [16].

Conclusion

Four methods were used to predict carbon stocks of tree components and total for Korean pine trees in the plantations of northeastern China. Based on the results of goodness-of-fit and model performance statistics of the biomass and carbon stock models, we chose the compatible systems of models with tree diameter as the sole independent variable in this study. The NSUR method was used to estimate the parameters in these models. The prediction errors of four methods for estimating carbon stocks of tree components and total were compared and tested.
The direct method using the compatible carbon stocks models (Method 1) was the best method for predicting the carbon stocks of stem and total. The indirect method using the average carbon concentration of tree (Method 3) and average carbon conversion of each tree component (Method 4) produced relative small prediction errors, while the indirect method using the carbon conversion factor 0.5 (Method 2) was the worse for estimating the carbon stocks for tree components and total of Korean pine in this study. However, there were no significant difference between Method 3 and Method 4. Thus, the average carbon concentration of individual tree rather than the average carbon concentration for each tree component is good enough to calculate the carbon stocks for tree components and total.

Appendix

The variable allomartie equation takes the form of:

\[ Y = aX^b \]  

(A1)

If the response variable \( Y \) is tree biomass and the independent variable \( X \) is tree diameter at breast height \( D \), Eq [A1] becomes:

\[ W = aD^b \]  

(A2)

In this study, our additive systems of tree biomass were formulated based on the methods by Tang and Li [32], in which the total tree biomass was used as the constraint as follows:

\[ W_t = W_s + W_b + W_f + W_r \]  

(A3)

where \( W_s, W_b, W_f, W_r \) and \( W_t \) are the biomass of stem, branch, foliage, root and total, respectively.

Thus, based on Eq [A2], the additive system of tree biomass is as follows:

\[
\begin{align*}
W_s &= a_1D^{b_1} \\
W_b &= a_2D^{b_2} \\
W_f &= a_3D^{b_3} \\
W_r &= a_4D^{b_4} \\
W_t &= W_s + W_b + W_f + W_r = a_0D^{b_0}
\end{align*}
\]  

(A4)

With the constraint of total biomass, the biomass of each tree component is further formulated as follows:

\[
\begin{align*}
W_s &= \frac{a_1D^{b_1}}{a_1D^{b_1} + a_2D^{b_2} + a_3D^{b_3} + a_4D^{b_4} \cdot a_0D^{b_0}} \\
W_b &= \frac{a_2D^{b_2}}{a_1D^{b_1} + a_2D^{b_2} + a_3D^{b_3} + a_4D^{b_4} \cdot a_0D^{b_0}} \\
W_f &= \frac{a_3D^{b_3}}{a_1D^{b_1} + a_2D^{b_2} + a_3D^{b_3} + a_4D^{b_4} \cdot a_0D^{b_0}} \\
W_r &= \frac{a_4D^{b_4}}{a_1D^{b_1} + a_2D^{b_2} + a_3D^{b_3} + a_4D^{b_4} \cdot a_0D^{b_0}}
\end{align*}
\]  

(A5)

Dividing \( a_1D^{b_1}, a_2D^{b_2}, a_3D^{b_3}, a_4D^{b_4}, \) respectively, to the numerator and denominator of each component equation in Eq [A5]. To simplify equation, we define: \( r_1 = \alpha_2/\alpha_1, r_2 = \alpha_3/\alpha_1, r_3 = \alpha_4/\alpha_1, k_1 = b_2 - b_1, k_2 = b_3 - b_1, k_3 = b_4 - b_1. \) Thus, we obtain the following
equation system:

\[
\begin{align*}
W_s &= \frac{a_0 D^b_0}{1 + \frac{a_2}{a_1} D^{b_2-b_1} + \frac{a_3}{a_1} D^{b_3-b_1} + \frac{a_4}{a_1} D^{b_4-b_1}} \\
W_b &= \frac{a_1 D^b_0}{a_2} D^{b_2-b_1} + 1 + \frac{a_3}{a_2} D^{b_3-b_1} + \frac{a_4}{a_2} D^{b_4-b_1} \\
W_f &= \frac{a_1 D^b_0}{a_3} D^{b_3-b_1} + \frac{a_2}{a_3} D^{b_2-b_1} + 1 + \frac{a_4}{a_3} D^{b_4-b_1} \\
W_r &= \frac{a_1 D^b_0}{a_4} D^{b_4-b_1} + \frac{a_2}{a_4} D^{b_2-b_1} + \frac{a_3}{a_4} D^{b_3-b_1} + 1 \\
W_s &= \frac{a_0 D^b_0}{1 + r_1 D^{k_1} + r_2 D^{k_2} + r_3 D^{k_3}} \\
W_b &= \frac{a_0 D^b_0}{1 + r_1 D^{k_1} + r_2 D^{k_2} + r_3 D^{k_3}} \\
W_f &= \frac{a_0 D^b_0}{1 + r_1 D^{k_1} + r_2 D^{k_2} + r_3 D^{k_3}} \\
W_r &= \frac{a_0 D^b_0}{1 + r_1 D^{k_1} + r_2 D^{k_2} + r_3 D^{k_3}} \\
\end{align*}
\]  

(A6)

Adding tree total height \( H \) to Eq [A2] results in another popular biomass model:

\[
Y = a(D^b H^h)
\]  

(A7)

Following the same derivation process above, the compatible biomass and carbon stocks models can be constructed as follows:

\[
\begin{align*}
W_s &= a_0 D^{b_0} H^h / (1 + r_1 D^{k_1} H^{k_1} + r_2 D^{k_2} H^{k_2} + r_3 D^{k_3} H^{k_3}) \\
W_b &= a_0 r_1 D^{k_1} H^{k_1} / (1 + r_1 D^{k_1} H^{k_1} + r_2 D^{k_2} H^{k_2} + r_3 D^{k_3} H^{k_3}) \\
W_f &= a_0 r_2 D^{k_2} H^{k_2} / (1 + r_1 D^{k_1} H^{k_1} + r_2 D^{k_2} H^{k_2} + r_3 D^{k_3} H^{k_3}) \\
W_r &= a_0 r_3 D^{k_3} H^{k_3} / (1 + r_1 D^{k_1} H^{k_1} + r_2 D^{k_2} H^{k_2} + r_3 D^{k_3} H^{k_3}) \\
\end{align*}
\]  

(A8)

where the model coefficients \( a_0, b_0, r_1, r_2, r_3, k_1, k_2 \) and \( k_3 \) are estimated from the tree biomass.

Supporting Information

S1 File. Biomass data used in the study. (XLS)

Acknowledgments

The authors would like to thank the staff and students of the Department of Forest Management, Northeast Forestry University (NEFU), P.R. China, who provided and collected the data for this study.

Author Contributions

Conceived and designed the experiments: FL HLG. Performed the experiments: HLG LHD. Analyzed the data: HLG. Contributed reagents/materials/analysis tools: HLG LHD FL LJZ. Wrote the paper: HLG. Provided detailed revised comments to the manuscript: LJZ LHD.

References

1. Picard N, Bosela FB, Rossi V. Reducing the error in biomass estimates strongly depends on model selection. Ann For Sci. 2015; 72: 811–823.
2. Gonzalez P, Battles J, Collins BM, Robards T, Saah DS. Aboveground live carbon stock changes of California wildland ecosystems, 2001–2010. For Ecol Manage. 2015; 348: 68–77.
3. Breugel MV, Ransijn J, Craven D, Bongers F, Hall JS. Estimating carbon stock in secondary forests: Decisions and uncertainties associated with allometric biomass models. For Ecol Manage. 2011; 262: 1648–1657.

4. Fassnacht FE, Hartig F, Latifi H, Berger C, Hernández J, Corvalán P, et al. 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.

5. Breidenbach J, Astrup R. Small area estimation of forest attributes in the Norwegian National Forest Inventory. Eur J Forest Res. 2012; 131: 1255–1267.

6. Zeng W. Development of monitoring and assessment of forest biomass and carbon storage in China. Forest Ecosystems. 2014; 1:1–10.

7. Planck NRV, MacFarlane DW. A vertically integrated whole-tree biomass model. Trees. 2015; 29: 449–460.

8. Jha KK. Carbon storage and sequestration rate assessment and allometric model development in young teak plantations of tropical moist deciduous forest, India. J For Res. 2015; 26(3): 589–604.

9. Zabek LM, Prescott CE. Biomass equations and carbon content of aboveground leafless biomass of hybrid poplar in Coastal British Columbia. For Ecol Manage. 2006; 229: 291–302.

10. Návar J. Allometric equations for tree species and carbon stocks for forests of northwestern Mexico. For Ecol Manage. 2009; 257: 427–434.

11. Kenzo T, Ichie T, Hattori D, Kendawang JJ, Sakurai K, Ninomiya I, et al. Changes in above- and belowground biomass in early successional tropical secondary forests after shifting cultivation in Sarawak, Malaysia. For Ecol Manage. 2010; 260: 875–882.

12. Gil MV, Blanco D, Carballo MT, Calvo LF. Carbon stock estimates for forests in the Castillay León region, Spain. A GIS based method for evaluating spatial distribution of residual biomass for bioenergy. Biomass Bioenergy. 2011; 35: 243–252.

13. Petersson H, Holma S, Ståhla G, Fridman J, Lehtonen A, et al. Individual tree biomass equations or biomass expansion factors for assessment of carbon stocks changes in living biomass-A comparative study. For Ecol Manage. 2012; 270: 78–84.

14. Ngo KM, Turner BL, Muller-Landau HC, Davies SJ, Larjavaara M, Hassan NFBN, et al. Carbon stocks in primary and secondary tropical forests in Singapore. For Ecol Manage. 2013; 296: 81–89.

15. Chaturvedi RK, Raghubanshi AS. Assessment of carbon density and accumulation in mono- and multi-specific stands in Teak and Sal forests of a tropical dry region in India. For Ecol Manage. 2015; 339: 11–21.

16. Mello AAD, Nutto L, Weber KS, Sanquette CE, Monteiro de Matos JL, Becker G, et al. Individual biomass and carbon equations for Mimosa scabrella benth. (bracatinga) in southern Brazil. Silva Fenn. 2012; 46: 333–343.

17. Jenkins JC, Chojnacky DC, Heath LS, Birdsey RA. National-scale biomass estimators for United States tree species. For Sci. 2003; 49: 12–35.

18. Zhang Y, Borders BE, Will RE, Posadas HDLS. A model for foliage and branch biomass prediction for intensively managed fast growing loblolly pine. For Sci. 2004; 50: 65–80.

19. Lambert M-C, Ung C-H, Raulier F. Canadian national tree aboveground biomass quations. Can J For Res. 2005; 35: 1996–2018.

20. Socha J, Wozny P. Allometric equations for estimating the foliage biomass of Scots pine. Eur J Forest Res. 2007; 126: 263–270.

21. Alvarez E, Duque A, Saltarriaga J, Cabrera K, Salas GDL, Valle ID, et al. Tree above-ground biomass allometries for carbon stocks estimation in the natural forests of Colombia. For Ecol Manage. 2012; 267: 297–308.

22. Parresol BR. Assessing tree and stand biomass: a review with examples and critical comparisons. For Sci. 1999; 45: 573–593.

23. Parresol BR. Additivity of nonlinear biomass equations. Can J For Res. 2001; 31: 865–878.

24. Tang S, Li Y, Wang Y. Simultaneous equations, error-in-variable models, and model integration in systems ecology. Ecol Model. 2001; 142: 285–294.

25. Bi H, Turnier J, Lambert MJ. Additive biomass equations for native eucalypt forest trees of temperate Australia. Trees. 2004; 18: 467–479.

26. SAS Institute, Inc. SAS/ETS 9.3. User’s Guide. SAS Institute, Inc., Cary, NC; 2011.

27. Dong L, Zhang L, Li F. A compatible system of biomass equations for three conifer species in Northeast, China. For Ecol Manage. 2014; 329: 306–317.
28. Dong L, Zhang L, Li F. A three-step proportional weighting system of nonlinear biomass equations. For Sci. 2015; 61: 35–45.

29. Dong L, Zhang L, Li F. Allometry and Partitioning of Individual Tree Biomass and Carbon of Abies nephrolepis Maxim in Northeast China. Scand J Forest Res. 2015; doi: 10.1080/02827581.2015.1060257

30. Vallet P, Dhôte J-F, Moguédec GL, Ravart M, Pignard G. Development of total aboveground volume equations for seven important forest tree species in France. For Ecol Manage. 2006; 229: 98–110.

31. Zianis D. Predicting mean aboveground forest biomass and its associated variance. For Ecol Manage. 2008; 256: 1400–1407.

32. Tang S, Li Y, Fu L. Statistical foundation for biomathematical models. Higher Education Press; 2015 (in Chinese).

33. Wallach D, Goffinet B. Mean squared error of prediction as a criterion for evaluating and comparing system models. Ecological Modelling. 1989; 44: 299–306.

34. Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE)?--Arguments against avoiding RMSE in the literature. Geoscientific Model Development. 2014; 7: 1247–1259.

35. Gezan SA, White TL, Huber DA. Comparison of experimental designs for clonal forestry using simulated data. For Sci. 2006; 52: 108–116.

36. Fiedler CE, Metlen KL, Dodson EK. Restoration treatment effects on stand structure, tree growth, and fire hazard in a Ponderosa Pine/Douglas-Fir forest in Montana. For Sci. 2010; 56: 18–31.

37. Iglay RB, Leopold BD, Miller DA. Vegetation responses to fire and herbicide in intensively managed, mid-rotation pine. For Ecol Manage. 2014; 328: 69–78.

38. SAS Institute, Inc. SAS/STAT 9.3. User’s Guide. SAS Institute, Inc., Cary, NC; 2011.

39. McRoberts RE, Westfall JA. Effects of uncertainty in model predictions of individual tree volume on large area volume estimates. For Sci. 2014; 60: 34–42.

40. Zianis D, Mencuccini M. Aboveground biomass relationships for beech (Fagus moesiaca Cz.) trees in Vermio Mountain, Northern Greece, and generalized equations for Fagus sp. Ann For Sci. 2003; 60: 439–448.

41. Bond-Lamberty B, Wang C, Gower ST. Aboveground and belowground biomass and sapwood area allometric equations for six boreal tree species of northern Manitoba. Can J For Res. 2002; 32: 1441–1450.

42. Wang C. Biomass allometric equations for 10 co-occurring tree species in Chinese temperate forests. For Ecol Manage. 2006; 222: 9–16.

43. Green C, Tobin B, O’Shea M, Farrell EP, Byrne KA. Above- and belowground biomass measurements in an unthinned stand of Sitka spruce (Picea sitchensis (Bong) Carr.). Eur J Forest Res. 2007; 126: 179–188.

44. Bastien-Henri S, Park A, Ashton M, Messier C. Biomass distribution among tropical tree species grown under differing regional climates. For Ecol Manage. 2010; 260: 403–410.

45. António N, Tomé M, Tomé J, Soares P, Fontes L. Effect of tree, stand, and site variables on the allometry of Eucalyptus globulus tree biomass. Can J For Res. 2007; 37: 895–906.

46. Cheng JZ, Lee X, Theng BKG, Zhang L, Fang B, Li F, et al. Biomass accumulation and carbon sequestration in an age-sequenced of Zanthoxylum bungeanum plantations under the Grain for Green Program in karst regions, Guizhou province. Agr Forest Meteorol. 2015; 203: 88–95.

47. Park A. Carbon storage and stand conversion in a pine-dominated boreal forest landscape. For Ecol Manage. 2015; 340: 70–81.

48. Lehtonen A, Cienciala E, Tatarinov F, Mäkipää R. Uncertainty estimation of biomass expansion factors for Norway spruce in the Czech Republic. Ann For Sci. 2007; 64: 133–140.