Generic additive allometric models and biomass allocation for two natural oak species in northeastern China

Shengwang Meng
Institute of Geographic Sciences and Natural Resources Research Chinese Academy of Sciences
Fan Yang
Academy of Forest Inventory and Planning, National Forestry and Grassland Administration
Haibin Wang
Planning & Design Institute of Forest Products Industry, National Forestry and Grassland Administration
Wei Wang
Academy of Forest Inventory and Planning, National Forestry and Grassland Administration
Sheng Hu
Planning & DeSign Institute of Forest Products Industry, National Forestry and Grassland Administration
Huimin Wang (wanghm@igsnrr.ac.cn)
Institute of Geographic Sciences and Natural Resources Research Chinese Academy of Sciences

---

Research

Keywords: Aboveground and belowground biomass, allometry, biomass allocation, multivariable model, heteroscedasticity

Posted Date: December 1st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-114961/v1

Generic additive allometric models and biomass allocation for two natural oak species in northeastern China

Meng Shengwang[1,2] Yang Fan[3] Wang Haibin[4] Wang wei[3] Hu Sheng[4] Wang Huimin[1*]
1 Qianyanzhou Ecological Research Station, Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
2 Zhongke-Ji'an Institute for Eco-Environmental Sciences, Ji'an 343016, China
3 Academy of Forest Inventory and Planning, National Forestry and Grassland Administration, Beijing 100714, China
4 Planning & Design Institute of Forest Products Industry, National Forestry and Grassland Administration, Beijing 100010, China
* Correspondence: wanghm@igsnrr.ac.cn

Abstract

Background: Accurate quantification of forest biomass through allometric equations is crucial for global carbon accounting and climate change mitigation. Current models for oak species could not accurately estimate biomass in northeastern China, since they were usually established limited to Mongolian oak (*Quercus mongolica*) on local sites, and specifically, no biomass models were available for Liaodong oak (*Quercus wutaishanica*). The goal of this study was, therefore, to develop generic biomass models for both oak species on large scale and evaluate biomass allocation patterns within tree components.

Results: The stem biomass accounts for about two-thirds of the aboveground biomass. The ratio of wood biomass holds constant and that of branch increases with increasing *D*, *H*, *CW*, *CL*, while a reverse trend was found for bark and foliage. The root-shoot ratio nonlinearly decreased with *D*, ranging from 1.06 to 0.11. Tree diameter proved to be a good predictor, especially for root biomass. Tree height is more prominent than crown size for improving stem biomass models, yet it puts negative effects on crown biomass models with non-significant coefficients. Crown width could help improve fitting results of branch and foliage biomass models.

Conslusion: We conclude that the selected generic biomass models for Mongolian oak and Liaodong oak will vigorously promote the accuracy of biomass estimation.

Keywords Aboveground and belowground biomass; allometry; biomass allocation; multivariable model; heteroscedasticity

Background

Mongolian oak (*Quercus mongolica*) and Liaodong oak (*Quercus wutaishanica*) are very similar in both morphological characteristics and geographical distribution. Both species are valuable and occupy the secondary broadleaved deciduous forest regions widely distributed in northeastern China [1]. Besides, the two oak species have strong adaptability with characteristics of resistance to drought and impoverishment, play an important role in biodiversity protection and sustainable forestry development. To assess forest productivity, carbon stocks and dynamics, matter and energy flows, forest ecosystem service, accurate quantification of forest biomass is of great importance [2-4]. The most accurate and direct way of quantifying tree biomass is to harvest all the trees in a specific area and weigh the dry weights of both stem and crown components. However, the use of destructive sampling is enormously time-consuming, expensive, and limited to small trees or small sample size; it is also not recommended in nature reserves and on endangered tree species [5]. To overcome these limitations, the indirect estimation of biomass based on allometric equations is an alternative and rapid method without destruction if accuracy is not compromised [6, 7]. To our knowledge, only a few studies have applied a destructive approach to establish allometric models based on diameter at breast height (*D*) and height (*H*) for estimating total and components biomass in Mongolian oak [8-10], but these studies were limited to specific regions and the sample sizes were relatively small. Furthermore, almost no biomass models existed for Liaodong oak, and especially no generic biomass equations were reviewed for the two species.

Tree biomass is generally divided into aboveground and belowground parts, and separating the aboveground biomass into four components including stem wood, bark, branch, and foliage. The biomass models of different components are available for elucidating the biomass variability within a tree and are useful for some other goals that need separate estimates such as wood production and bioenergy assessment provided by the crown [11]. When developing tree component biomass

regression equation, a desirable feature is that the sum of predicted values from separate equations for components should equal the prediction from the total tree biomass models, also known as additivity [12]. However, some published biomass models were non-additive [e.g., 10, 13], resulting in inconsistency between the sum of components estimates and the total estimates [14]. Besides, separately fitting the biomass models by applying the least square method might ignore the inherent correlations among different biomass components of an individual tree. Parresol [15] concluded that considering this inherent correlation by constructing an additive system of biomass equations can acquire higher statistical efficiency.

Many studies have discussed the issue of ensuring additivity among a set of equations for total tree and its biomass components [14, 16, 17], and the additivity can be solved by fitting total tree and components biomass models simultaneously as a model system. Some corresponding methods are as follows: the generalized method of moments (GMM), the adjustment in proportion and nonlinear joint estimate, the seemingly unrelated regression (SUR), and non-linear seemingly unrelated regression (NSUR) methods. Of which, SUR and NSUR methods are the most flexible and general method to ensure biomass additivity [12], and have become popular in recent years [5, 15, 18]. The two methods allow the additive system of biomass equations using different forms of mathematical equations for each component with specific explanatory variables [19]. In particular, heteroscedasticity could be addressed through NSUR by applying a specific weighting function for each component equation, and the different weighting functions can also remove the singularity problem in across-equation variance and covariance matrix [20].

The knowledge of belowground biomass and its dynamics is essential to fully understand the role of roots in carbon sequestration and climate change mitigation [21]. However, estimations of forest biomass have always been limited to the aboveground biomass by simply ignoring the belowground biomass [22] since the big workload in root excavation or the limited availability of allometric equations for belowground biomass [23]. Some other studies estimated belowground biomass by directly multiplying aboveground biomass by a root-to-shoot ratio (RS-ratio). But the RS-ratio varies considerably with tree age and environmental conditions [24, 25]. Since the large unexplored errors in belowground biomass estimation, an alternative option for quantifying the amounts of belowground biomass is allometric equations if the information on individual trees is available [2, 23].

In general, allometric biomass equations using *D* alone as an explanatory variable are commonly used because of its quick and easy measurement with high accuracy. Tree height has also been investigated and included as a main predictor to improve model availability across forest types [24, 26] as *D-H* allometry depends on environmental conditions [7]. Besides, crown width (*CW*) or crown area, wood density, and tree age were integrated into biomass estimations as well to enhance the predictive abilities of models [5, 22, 27, 28]. However, some variables measurements like wood density need additional work, extending project time, and increasing costs [29]. Furthermore, allometric models with multiple predictors require forest inventory datasets contain relevant factors.

The objective of this study was to develop generic biomass models for the two oak species in northeastern China. We focused on additive systems of biomass equations for aboveground components, but we also constructed biomass models for roots. Both aboveground and root biomass equations were validated through the leave-one-out cross-validation method and also compared to published models to explore the potential for application. Furthermore, we assessed the biomass allocation patterns and presented the trend of root-shoot ratios varied with diameter.

Methods

Study area description

This study was conducted in the Daxing'anling mountains (47°03′-53°33′N, 119°36′-127°00′E), the Xiaoxing'anling mountains (46°28′-49°21′N, 127°42′-130°14′E), and the Changbai mountains (43°31′-47°30′N, 127°40′-134°00′E), located in Inner Mongolia autonomous region and Heilongjiang province, northeastern China (Fig. 1). The three mountains have a distinct continental monsoon climate. In Daxing'anling mountains, the elevation varies between 425 m and 1760 m above sea level, the mean annual temperature is -2.8 ℃, and the annual precipitation is around 460 mm which is mainly distributed from May to October. In Xiaoxing'anling mountains, the elevation ranges from 500 m to 1000 m, the mean annual temperature is from -2.0 ℃ to 2.0 ℃, and the annual precipitation is from 550 mm to 670 mm, concentrated mostly in summer. In Changbai mountains, the elevation ranges from 800 m to 1800 m, the mean annual temperature is from -7.0 ℃ to 3.0 ℃, and the annual precipitation is from 700 mm to 1400 mm that mainly happens from June to October. The substrate soils are predominantly brown coniferous forest soil for the three regions. These cold temperate forests are dominated by Dahurian larch (*Larix* spp.), White birch (*Betula platyphylla*), Aspen (*Populus davidiana*), Korean pine (*Pinus koraiensis*), Amur linden (*Tilia amurensis*), Mongolian oak, Manchurian ash (*Fraxinus mandschurica*), Amur corktree (*Phellodendron amurense*), Maple (*Acer mono*) and Manchurian walnut (*Juglans mandshurica*), among others.

Biomass data

A total of 159 oak trees (120 and 39 trees for Mongolian oak and Liaodong oak, respectively) were selected with *D* between 1.5 cm and 33.0 cm. A destructive sampling method with direct field measurements was adopted to gather biomass data, stem wood (inside bark), bark, branch, and foliage components were measured separately. Since the difficulty in root excavation, 53 trees with *D* spanning 1.5-32.8 cm were selected for root biomass collection and used to establish allometric equations for root biomass estimation.

For each tree, the *CW* of the sampled trees was obtained from the average of two diameters measured along the North/South and East/West orientation. After felling with chain saws, *D* (1.30 m), *H*, and crown length (*CL*, the distance from the first live branch to the treetop) were immediately measured and recorded. A general positive correlation between *H* and *D* was observed as displayed in Fig. 2, in which the vertical bars indicate the sample distribution across different diameter classes.

For each felled tree, the live crown was evenly separated into three layers, and three sample branches with the representativeness of size and foliage amount for each layer were selected to separate and weigh the fresh weight of branches and foliage. Subsamples of branches and foliage were then randomly collected and weighed, and bagged for water content determination in the laboratory. Subsequently, the stem was equally divided into ten sections and each section was weighed directly in the field. At the position corresponding to 0.1, 0.3, and 0.7 of tree height, 3- to 5-cm-thick stem discs were cut and weighed, and to acquire the fresh weight of stem wood by removing bark. These discs and the removed barks were taken back to the laboratory for determining water content. For root extraction, the soil was manually and carefully removed to obtain as intact roots as possible, especially fine roots. All extracted roots were weighed and recorded. The tree stump was dug out completely at last to measure the fresh weight. Subsamples of different root classes and stump were randomly gathered and weighed, and taken back to the laboratory for determining water content.

All subsamples of stem wood, bark, branch, foliage, root, and stump were oven-dried at 105 ℃ for

72 h to constant weight to estimate the water content. The dry weights of different components were then computed by subtracting the weight of water from fresh weights, namely, the dry biomass of each component was calculated by multiplying the fresh weight by the respective ratio of dry mass to fresh mass. For each sampled tree, the sum of dry biomass of stem wood, bark, branch, and foliage yielded aboveground biomass. Root (belowground) biomass was regarded as the sum of root dry biomass and stump dry biomass. Summary statistics for basic variables and biomass components of all trees used for establishing biomass models are displayed in Table 1.

Data analysis

Aboveground biomass allocated to wood, bark, branch, and foliage was evaluated by calculating the biomass fractions for each component, i.e., the proportion of component biomass to aboveground biomass. Bivariate relationships between biomass fractions and *D*, *H*, *CW*, and *CL* were further graphically examined for each component. The root-shoot ratio (i.e., root biomass/ aboveground biomass) varied with diameter classes was graphically examined.

Allometric models specification

Allometry relationships between component biomass and explanatory variables are generally expressed as a power function such that:

$$W = \beta_0 X_1^{\beta_1} X_2^{\beta_2} \cdots X_n^{\beta_n} + \varepsilon, \quad \varepsilon \sim N\left(0, \sigma^2\right) \tag{1}$$

where, W represents the biomass (kg) and X_i (i=1, 2, …, n) represent easily measured explanatory variables which reflect tree dimensions, β_i (i=1, 2, …, n) are model coefficients, and ε is the error term which is assumed to be normally distributed. As biomass data are significantly heteroscedastic, some approaches should be therefore taken to eliminate heteroscedasticity before parameter estimation [30]. In the present study, weighted nonlinear regression was adopted, and the specific weighting function expressed as $w = 1/\sqrt{D^Q}$ with a parameter Q, was derived from the residuals of biomass models fitted separately by the ordinary least squares method [15, 31].

The number of explanatory variables used in allometric biomass models always ranges from one single predictor (*D*) to several or multiple predictors (e.g., *H*, *CW*, and *CL*) [32]. Since the tree-dimension information available from the oak datasets in this study were *D*, *H*, *CW*, and *CL*, we established basic biomass models based on *D* alone and tested the model fitting improvement by adding *H*, *CW*, and *CL* variables as additional predictors. Thus, the following seven candidate models were fitted to the data:

$$W = \beta_0 D^{\beta_1} + \varepsilon \tag{2}$$

$$W = \beta_0 D^{\beta_1} H^{\beta_2} + \varepsilon \tag{3}$$

$$W = \beta_0 D^{\beta_1} CW^{\beta_2} + \varepsilon \tag{4}$$

$$W = \beta_0 D^{\beta_1} CL^{\beta_2} + \varepsilon \tag{5}$$

$$W = \beta_0 D^{\beta_1} H^{\beta_2} CW^{\beta_3} + \varepsilon \tag{6}$$

$$W = \beta_0 D^{\beta_1} H^{\beta_2} CL^{\beta_3} + \varepsilon \tag{7}$$

$$W = \beta_0 D^{\beta_1} H^{\beta_2} CW^{\beta_3} CL^{\beta_4} + \varepsilon \tag{8}$$

where W represents the biomass (kg), β_i are coefficients, ε is the error term.

To ensure the additivity, tree aboveground biomass and its components were simultaneously fitted using the NSUR method based on the aforementioned seven nonlinear models (equations (2)-(8)) with D, H, CW, CL as explanatory variables. Due to the difference in the sample size of the root and aboveground biomass, root biomass models were separately developed using these equations. Following the model structure specified by Parresol [15], the additive model system constrains component biomass to equal the aboveground biomass can be written as:

$$\begin{aligned}
W_{\text{wd}} &= f_{\text{wd}}\left(X_{\text{wd}}, \beta_{\text{wd}}\right) + \varepsilon_{\text{wd}} \\
W_{\text{bk}} &= f_{\text{bk}}\left(X_{\text{bk}}, \beta_{\text{bk}}\right) + \varepsilon_{\text{bk}} \\
W_{\text{br}} &= f_{\text{br}}\left(X_{\text{br}}, \beta_{\text{br}}\right) + \varepsilon_{\text{br}} \\
W_{\text{fo}} &= f_{\text{fo}}\left(X_{\text{fo}}, \beta_{\text{fo}}\right) + \varepsilon_{\text{fo}} \\
W_{\text{ag}} &= f_{\text{wd}}\left(X_{\text{wd}}, \beta_{\text{wd}}\right) + f_{\text{bk}}\left(X_{\text{bk}}, \beta_{\text{bk}}\right) + f_{\text{br}}\left(X_{\text{br}}, \beta_{\text{br}}\right) + f_{\text{fo}}\left(X_{\text{fo}}, \beta_{\text{fo}}\right) + \varepsilon_{\text{ag}}
\end{aligned} \tag{9}$$

where W_m represent components biomass (kg), m = wd, bk, br, fo, and ag, represent wood, bark, branch, foliage, and aboveground, respectively; f_m (X_m, β_m) is a nonlinear equation of biomass component m, ε_m is an error term for biomass component m.

In the additive model system, all components and aboveground biomass models were jointly fit adopting weighted NSUR [15, 20]. Fitting was conducted using the PROC MODEL procedure of SAS 9.3 (SAS Institute Inc, Cary, NC, USA).

Model evaluation and selection

The aboveground additive biomass model systems and root biomass equations were fitted to the entire dataset. The model fitting was assessed by looking at three goodness-of-fit statistics: root mean square error (RMSE), and adjusted coefficient of determination (R^2_{adj}). The best equation with the most suitable predictors for each component was selected according to minimum RMSE and maximum R^2_{adj}. The statistical significance (p-value < 0.05) of the estimated coefficients was taken into consideration as well. The best additive model system was reconstructed by combining the selected best biomass equations with available predictors for each component. Since the diameter is an imperative factor in the field measurement and other factors may be missing in some forest inventory data, an additive model system using D as the only independent variable was considered as an alternative approach for predicting biomass. Both model systems were assessed by the aforementioned criteria. The leave-one-out cross-validation method was further used to test the selected aboveground and root biomass models. Graphical analyses of predicted vs. observed values were plotted for each component and the prediction performance was evaluated by two statistical indices: mean prediction error (MPE), and mean percent standard error (MPSE). The mathematical expressions of these statistics were expressed as follows:

$$\text{RMSE} = \sqrt{\sum\left(y_j - \hat{y}_j\right)^2 / \left(n-k\right)} \tag{10}$$

$$R^2_{\text{adj}} = 1 - \left(1 - \sum\left(y_j - \hat{y}_j\right)^2 / \sum\left(y_j - \bar{y}\right)^2\right) \times \left(n-1\right) / \left(n-k\right) \tag{11}$$

$$\text{MPE} = t_\alpha \times \sqrt{\sum\left(y_j - \hat{y}_{j,-j}\right)^2 / \left(n-p\right)} / \left(\bar{y} \times \sqrt{n}\right) \times 100 \tag{12}$$

$$\mathrm{MPSE} = \sum \left| \left(y_j - \hat{y}_{j,-j} \right) / \hat{y}_{j,-j} \right| / n \times 100 \tag{13}$$

where y_j is the j-th observed biomass value; $\hat{y}_j$ is the j-th biomass value estimated by the model; $\bar{y}$ is the mean of observed biomass value; n is the total number of samples; k is the number of parameters; $\hat{y}_{j,-j}$ is the j-th value predicted by the model fitting with n-1 remaining observations, excluding the j-th observation.

Comparison to existing biomass models

The measured biomass data of Mongolian oak in this study was used to compare relative error (RE) of the existing species-specific equations reported by these studies [8-10] which included D and H as predictors. These models are only valid within the range of the independent variables, and the measured data within the specific range are therefore used. Table 2 provides a summary of these existing biomass models.

$$\mathrm{RE} = \left(\hat{W}_i - W_i \right) / W_i \times 100 \tag{14}$$

where $\hat{W}_i$ and W_i represent the predicted and observed dry weight of tree i.

Results

Aboveground biomass partitioning

The fraction of different components account for aboveground biomass reflects a large difference. Aboveground biomass allocated a large proportion to stem (approximately 68%), with partitioning proportion of 53% and 15% for stem wood and bark, respectively. By contrast, crown biomass possesses a relatively small part of aboveground biomass with an average value of 32%, which can be separated into branch and foliage with the proportion of 25% and 7%, respectively. The relationships between component biomass fractions and diameter, height, crown length, and crown width were further examined graphically and displayed in Fig. 3. In general, the ratio of wood biomass to aboveground biomass remains constant versus these factors. With the increase of D, H, CW, and CL, both bark and foliage showed a decreasing percentage, while a reverse trend was observed for the branch component.

Root-shoot ratio

Among individual trees that sampled both aboveground and root biomass, the root-shoot ratio (RS-ratio, the amount of belowground biomass to aboveground biomass) varied from 0.11 to 1.06 and was related to diameter (Fig. 4). The RS-ratio nonlinearly decreased with increasing D. The mean RS-ratio of the 53 individuals was 0.37.

Biomass models

The non-linear trend in observed biomass of stem wood, bark, branch, foliage, aboveground, and root as a function of D, H, CL, CW is shown in Fig. 5. The fitting results of the aboveground additive systems and root biomass equations based on the seven-candidate models are summarized in Table 3. Regression coefficients of D-based models were statistically highly significant ($p < 0.001$) for all components, with R^2_{adj} ranging from 0.754 to 0.920. When adding other variables, aboveground components biomass models had a better fitting result with lower RMSE and higher R^2_{adj} based on significant model coefficients. While for root biomass, although both RMSE and R^2_{adj} showed little

improvements in several models, coefficients for the added variables were not significant. Equation with *D* as the only predictor was therefore most suitable for root biomass. For aboveground components, the stem wood biomass models showed best fitting effects when using *D*, *H*, *CW*, and *CL* as predictors, which had lower RMSE (18.548 kg) and higher R^2_{adj} (0.958). Bark biomass equations showed less error with *D*, *H*, and *CW* as predictors, and the corresponding RMSE and R^2_{adj} were 4.466 kg and 0.946, respectively. In comparison with the improvement of wood and bark biomass models using *H* as an additional predictor, branch and foliage equations yielded larger errors as well as non-significant coefficients when *H* was included. Both branch and foliage biomass were better predicted with smaller estimation errors when *D* and *CW* were used, the corresponding RMSE and R^2_{adj} were 19.081 kg and 0.907 for branch and 3.275 kg and 0.765 for foliage, respectively.

Allometric equations considered appropriate and available for biomass estimation are displayed in Table 4. Since *D* is the most important factor in forest inventory and a desirable predictor in tree growth models, the additive system of biomass equations using *D* as the single predictor was chosen for potential biomass prediction (system 1). The best additive model system of aboveground biomass equations was obtained through assembling the selected equation for each biomass component (system 2). Less obvious changes were observed for biomass equations in system 2 compared to that in their original systems. While in comparison with system 1, the fitting results showed improvement for both aboveground and its components, especially for stem wood with R^2_{adj} increased from 0.886 to 0.959 and RMSE decreased by 12.191 kg.

Model cross-validation

The predicting performance of the appropriate allometric equations including system 1 (*D*-based) and system 2 (multivariable-based) for aboveground biomass components and *D*-based equation for root was assessed through the leave-one-out cross-validation method. The predicted biomass versus observed biomass for each component was displayed in Fig. 6. Biomass equations for wood, bark, and aboveground showed small prediction errors (the slopes were close to 1:1 line), whereas relatively large errors were obtained for the branch, foliage, and root equations. For each component, the linear fitting slopes for the scatter plot of predicted against observed biomass were essentially the same for the two additive systems of biomass equations. While the linear fitting effects based on system 2 were better than that based on system 1 for all aboveground components, with higher R^2_{adj}, lower MPE and MPSE (Fig. 7), indicating the predicted values of system 2 were closer to the observed values than system 1.

Discussion

Developing individual-based biomass models and estimating forest biomass has long become a prerequisite for studying productivity and in more recent times for assessing carbon accounting and its potential in mitigating climate change [33, 34]. The two additive systems of aboveground biomass equations (Table 4) and *D*-based root biomass equation obtained in the present study for two natural oak species (Mongolian oak and Liaodong oak) provide an operational approach for gaining precise biomass estimation in northeastern China.

Biomass partitioning

The share of biomass components out of aboveground biomass is commonly observed when making a comparison among species, diameter classes, and tree ages [5, 35]. The results of the present study indicated that wood biomass had the largest portion of aboveground biomass, and tree stem (both

wood and bark included) accounted for about 68% of the total aboveground, while the crown had relatively less aboveground biomass allocation and foliage represented the smallest portion. The aboveground biomass allocation characteristic is consistent with previous studies as for oak [8, 9]. The relatively steady contribution of the wood to the aboveground biomass and the increased ratio of branch biomass versus tree size suggesting that oak enhances the competitiveness of aboveground resources acquisition such as light perhaps mainly through promoting crown area growth to out-compete neighbor competitors. The foliage biomass proportion decreased against increasing tree size may be due to that foliage is borne on the relatively new branches, implying that the production of foliage biomass per unit of woody biomass decreases [5]. The increase of woody mass fractions including wood and branch components and decrease of foliage with increasing tree size obtained in the present study could be explained by the fact that woody mass usually accumulates at the expense of foliage mass [35, 36].

Root-aboveground biomass link

Belowground biomass is commonly missing in most biomass datasets and is often indirectly obtained through RS-ratio based on aboveground biomass [22]. The RS-ratio decreased significantly with an increase in diameter, namely tree growth, the trend is similar to relevant studies [24, 25]. The mean value of RS-ratio observed in this study (0.37) is comparatively higher than the global mean value (0.26) reported by Cairns et al [37] and the mean value (0.29) of broad-leaved species including oak in China [38]. While Mugasha et al [2] reported a mean value of 0.40 in miombo woodlands which is slightly higher. Since the nonlinear relationship between RS-ratio and diameter, the large difference among mean RS-ratios may attribute to sample size in different diameter classes. Although mean RS-ratios have frequently been accepted for root biomass estimation for purposes such as productivity quantification and carbon accounting, the nonlinear trend indicates that using a fixed RS-ratio to predict belowground biomass could introduce a bias [2, 39]. Therefore, if root allometric biomass equations are not available, the application of RS-ratios should be done with caution.

Allometric equations

Biomass additivity has long been recognized as a desirable and important property when developing models to predict components and whole tree biomass [40], since the additive property eliminates the logical inconsistency between the sum of estimated values for the tree components and the estimation provided by the total tree equation [14, 41]. However, many biomass models reported to date are non-additive and established using the ordinary least square method [10, 13]. The merits of NSUR make it an approach of selection to fit the system of biomass equations, considering the synergetic correlations for each component and correcting the inherent errors among the estimates of each model. In this case, the explanatory variables of the components biomass models constitute the predictors of the total tree biomass model and the additive property is thus guaranteed through setting parameter restrictions [15, 19]. Furthermore, an additive system of biomass equations when estimated by considering the contemporaneous correlations between biomass components has greater statistical efficiency and reduces the confidence and prediction intervals of the biomass estimates, resulting in lower variance [12, 15, 40].

Tree diameter, the most important index in forest inventories, is always used for developing biomass models, and tree height is often added to the models as an additional predictor [27, 41]. Besides, we tested two other commonly suggested explanatory variables (*CW* and *CL*) in explaining components biomass variability. Our results in Table 3 showed that the root biomass equation with *D* as the single explanatory variable performed best with significant model coefficients, although the goodness-of-fit

was not optimum. This was consistent with Djomo and Chimi [22] that *D* alone was a good predictor for the belowground biomass estimation. As confirmed in many studies, the aboveground biomass models using only *D* as a predictor also gave good fitting results with significant coefficients for all components [24, 42, 43]. Using *D* reduces the limitations in reliability of *H* and crown dimension (*CW* and *CL*) in most forest inventories since height and crown dimension measurements are difficult to measure in the field due to many factors such as dense canopy, taller trees, and obstruction hinders to see the top of the tree [22, 44].

Despite the difficulty in accurate measurements, including *H* as a second predictor in aboveground biomass models is often recommended as it improves fitting effects [45, 46]. In our case, the inclusion of *H* improved the stem wood and bark biomass estimations but reduced the fitting efficiency for branch and foliage biomass models, in which the coefficients for the *H* variable were nonsignificant and not always positive (Table 3). These results were in agreement with previous relevant studies [20, 47-49]. In general, the crown dimension varies greatly among species, reflecting different biomass allocation strategies [45]. This study found that inclusion of *CW* and/or *CL* slightly improved the model fits for aboveground components, especially for stem wood and bark, contradicting the results that adding crown radius could greatly improve model performance and reduce error [50].

Heteroscedasticity elimination

The heteroscedastic residual error is a common problem when developing non-linear biomass equations with additive error terms [41]. Logarithmic transformation of both response and explanatory variables is opted in most cases to deal with the heteroscedasticity of the original arithmetic data [e.g., 13, 22, 51]. However, the nonlinear transformation of converting arithmetic units to logarithms alters the relationship between response and explanatory variables and can lead to substantial bias when estimating the allometric model parameters on the original scale through logarithmic linear fitting [52, 53]. Also, a correction factor is usually selected to remove systematic bias generated by converting the estimated logarithmic value to the original scale [54]. While correcting bias, the additive property of biomass models is simultaneously tended to destroy. Weighted NSUR method with a unique weighting function for each equation is likely preferred to address heteroscedasticity during biomass model development [e.g., 30, 41, 55, 56]. In line with previous studies, our study also showed good performance of weighted NSUR for addressing the heteroscedasticity (Fig. 8).

Performance of existing equations

Through the literature review, some equations have been established for estimating components biomass applicable to this study [8-10]. The biomass measured in this study was used to compute REs to test the reliability of these models to the present study (Fig. 9). The four models all overestimated stem biomass (including both stem wood and bark), and only model type 2 could keep the mean relative error under 10%. While for branch biomass, all of these models gave underestimation, and the smallest mean RE was obtained by model type 1. Both model type 1 and 2 provided better estimations for foliage biomass with mean RE less than 5%, yet model type 3 and 4 yielded relatively large underestimations. Since stem account for a large portion of aboveground biomass, the distribution of REs for aboveground biomass estimation is similar to the stem component, namely model type 2 gave the smallest RE (slight underestimation). For belowground biomass, model type 1 provided overestimation while the other three models gained underestimation, and model type 4 provided the closest value to the measured biomass with a mean RE of -0.37%.

Although several equations performed well, most equations provided REs larger than 15%. The small sample sizes might affect the model predictive ability, especially for model type 3 and 4. Besides, biomass data for developing model type 3 was gathered in Heilongjiang province and that for model type 4 was obtained in Jilin province, while model type 1 and 2 were based on data collected in the two provinces (Table 2). However, in the western Daxing'anling mountains located in Inner Mongolia, no trees were sampled. Since allometric biomass equations vary with forest type, size, and stand condition, the application of biomass models not suitable to the specific environment may lead to large systematic errors from observed data [26, 27, 40].

Conclusion

Appropriate allometric biomass equations are critical in reducing uncertainties for obtaining reliable forest biomass estimation, especially for carbon accounting. In the present study, we developed two systems of oak tree biomass equations: one based on tree *D* only, another based on multivariable including *D*, *H*, *CW*, and *CL*. Both systems guarantee the additive property between tree components and aboveground biomass. The additive systems of equations were estimated using a nonlinear seemingly unrelated regression method with considering the inherent errors among the equations and addressed the heteroscedasticity through making each equation has its weighting function. Diameter alone proved to be a good predictor for all component biomass estimation, especially for root biomass. The system with additional crown dimension slightly improved the prediction of aboveground biomass components, but adding *CW* to branch and foliage biomass models could yield more improvements than stem wood and bark. The additive system of biomass equations with *D*, *H*, *CW*, and *CL* as predictors can be applied to natural oak species distributed in northeastern China for carbon accounting purpose, while the additive system of equations using *D* as a single predictor can be used when height and crown information is not available. Both systems of equations must be used with caution when predicting the biomass of trees outside the range of data used for model development and site conditions.

Acknowledgments

The authors would like to thank Chang Guangjun and those individuals involved in data collection for their kind technical support during tree sampling. We also would like to thank the second forestry monitoring and planning institute of the Inner Mongolia Autonomous region for their great help in material and logistics during fieldwork.

Funding

This work was supported by the Foundation of President of the Zhongke-Ji'an Institute for Eco-Environmental Sciences.

Availability of data and materials

The dataset used in the present study are available from the corresponding author for reasonable cases.

Ethics approval and consent to participate

Not applicable.

Authors' contributions

SWM and HMW, conceived and design the experiments; FY and SH, performed the experiments; WW, helped analyzed the data; SWM, wrote the draft manuscript. All authors read, improved, and approved the final manuscript.

Consent for publication

All authors agree the publication.

Competing interests

The authors declare no conflict of interest.

Reference

1. Yin X, Zhou G, Sui X, He Q, Li R: **Dominant climatic factors of *Quercus mongolica* geographical distribution and their thresholds**. *Acta Ecol Sin* 2013, **33**(1):103-109. (In Chinese).
2. Mugasha WA, Eid T, Bollandsas OM, Malimbwi RE, Chamshama SAO, Zahabu E, Katani JZ: **Allometric models for prediction of above- and belowground biomass of trees in the miombo woodlands of Tanzania**. *Forest Ecol Manag* 2013, **310**:87-101.
3. He JS: **Carbon cycling of Chinese forests: From carbon storage, dynamics to models**. *Sci China Life Sci* 2012, **55**(2):188-190.
4. Alongi DM, Clough BF, Dixon P, Tirendi F: **Nutrient partitioning and storage in arid-zone forests of the mangroves Rhizophora stylosa and Avicennia marina**. *Trees* 2003, **17**(1):51-60.
5. Dimobe K, Mensah S, Goetze D, Ouédraogo A, Kuyah S, Porembski S, Thiombiano A: **Aboveground biomass partitioning and additive models for *Combretum glutinosum* and *Terminalia laxiflora* in West Africa**. *Biomass Bioenerg* 2018, **115**:151-159.
6. Sanaei A, Ali A, Ahmadaali K, Jahantab E: **Generalized and species-specific prediction models for aboveground biomass in semi-steppe rangelands**. *Journal of Plant Ecology* 2019, **12**(3):428-437.
7. Chave J, Rejou-Mechain M, Burquez A, Chidumayo E, Colgan MS, Delitti WBC, Duque A, Eid T, Fearnside PM, Goodman RC *et al*: **Improved allometric models to estimate the aboveground biomass of tropical trees**. *Global Change Biol* 2014, **20**(10):3177-3190.
8. Dong L, Zhang L, Li F: **Developing additive systems of biomass equations for nine hardwood species in Northeast China**. *Trees* 2015, **29**(4):1149-1163.
9. He H, Zhang C, Zhao X, Fousseni F, Wang J, Dai H, Yang S, Zuo Q: **Allometric biomass equations for 12 tree species in coniferous and broadleaved mixed forests, Northeastern China**. *PLoS ONE* 2018, **13**(1).
10. Wang CK: **Biomass allometric equations for 10 co-occurring tree species in Chinese temperate forests**. *Forest Ecol Manag* 2006, **222**(1-3):9-16.
11. Eker M, Poudel KP, Ozcelik R: **Aboveground Biomass Equations for Small Trees of Brutian Pine in Turkey to Facilitate Harvesting and Management**. *Forests* 2017, **8**(12).

12. Parresol BR: **Assessing tree and stand biomass: a review with examples and critical comparisons**. *Forest Sci* 1999, **45**(4):573-593.
13. Blujdea VNB, Pilli R, Dutca I, Ciuvat L, Abrudan IV: **Allometric biomass equations for young broadleaved trees in plantations in Romania**. *Forest Ecol Manag* 2012, **264**:172-184.
14. Kozak A: **Methods for ensuring additivity of biomass components by regression analysis**. *Forest Chron* 1970, **46**(5):402-405.
15. Parresol BR: **Additivity of nonlinear biomass equations**. *Can J Forest Res* 2001, **31**(5):865-878.
16. Tang S, Zhang H, Xu H: **Study on establish and estimate method of compatible biomass model**. *Sci Silvae Sin* 1999, **36**(supple. 1):19-27.
17. Chiyenda SS, Kozak A: **Additivity of component biomass regression equations when the underlying model is linear**. *Can J Forest Res* 1984, **14**(3):441-446.
18. Cao L, Li H: **Analysis of Error Structure for Additive Biomass Equations on the Use of Multivariate Likelihood Function**. *Forests* 2019, **10**(4).
19. Canga E, Dieguez-Aranda I, Afif-Khouri E, Camara-Obregon A: **Above-ground biomass equations for Pinus radiata D. Don in Asturias**. *Forest Systems* 2013, **22**(3):408-415.
20. Zhao D, Westfall J, Coulston JW, Lynch TB, Bullock BP, Montes CR: **Additive biomass equations for slash pine trees: comparing three modeling approaches**. *Can J Forest Res* 2019, **49**(1):27-40.
21. Helmisaari H-S, Derome J, Nöjd P, Kukkola M: **Fine root biomass in relation to site and stand characteristics in Norway spruce and Scots pine stands**. *Tree Physiology* 2007, **27**(10):1493-1504.
22. Djomo AN, Chimi CD: **Tree allometric equations for estimation of above, below and total biomass in a tropical moist forest: Case study with application to remote sensing**. *Forest Ecol Manag* 2017, **391**:184-193.
23. Xing DL, Bergeron JAC, Solarik KA, Tomm B, Macdonald SE, Spence JR, He FL: **Challenges in estimating forest biomass: use of allometric equations for three boreal tree species**. *Can J Forest Res* 2019, **49**(12):1613-1622.
24. Kachamba DJ, Eid T, Gobakken T: **Above- and Belowground Biomass Models for Trees in the Miombo Woodlands of Malawi**. *Forests* 2016, **7**(2).
25. Mokany K, Raison RJ, Prokushkin AS: **Critical analysis of root:shoot ratios in terrestrial biomes**. *Global Change Biol* 2006, **12**(1):84-96.
26. Mensah S, Veldtman R, Du Toit B, Glèlè Kakaï R, Seifert T: **Aboveground Biomass and Carbon in a South African Mistbelt Forest and the Relationships with Tree Species Diversity and Forest Structures**. *Forests* 2016, **7**(4):79.
27. Chave J, Andalo C, Brown S, Cairns MA, Chambers JQ, Eamus D, Folster H, Fromard F, Higuchi N, Kira T *et al*: **Tree allometry and improved estimation of carbon stocks and balance in tropical forests**. *Oecologia* 2005, **145**(1):87-99.
28. Xu Y, Zhang J, Franklin SB, Liang J, Ding P, Luo Y, Lu Z, Bao D, Jiang M: **Improving allometry models to estimate the above- and belowground biomass of subtropical forest, China**. *Ecosphere* 2015, **6**(12).
29. Alvarez E, Duque A, Saldarriaga J, Cabrera K, de Las Salas G, del Valle I, Lema A, Moreno F, Orrego S, Rodriguez L: **Tree above-ground biomass allometries for carbon stocks estimation in the natural forests of Colombia**. *Forest Ecol Manag* 2012, **267**:297-308.
30. Zeng W, Zhang L, Chen X, Cheng Z, Ma K, Li Z: **Construction of compatible and additive individual-tree biomass models for Pinus tabulaeformis in China**. *Can J Forest Res* 2017, **47**(4):467-475.

31. Zeng W, Duo H, Lei X, Chen X, Wang X, Pu Y, Zou W: **Individual tree biomass equations and growth models sensitive to climate variables for Larix spp. in China**. *Eur J For Res* 2017, **136**(2):233-249.
32. de-Miguel S, Pukkala T, Assaf N, Shater Z: **Intra-specific differences in allometric equations for aboveground biomass of eastern Mediterranean Pinus brutia**. *Ann Forest Sci* 2014, **71**(1):101-112.
33. Mensah S, Veldtman R, Seifert T: **Allometric models for height and aboveground biomass of dominant tree species in South African Mistbelt forests**. *South Forests* 2016, **79**(1):19-30.
34. Bi H, Murphy S, Volkova L, Weston C, Fairman T, Li Y, Law R, Norris J, Lei X, Caccamo G: **Additive biomass equations based on complete weighing of sample trees for open eucalypt forest species in south-eastern Australia**. *Forest Ecol Manag* 2015, **349**:106-121.
35. Mensah S, Kakaï RG, Seifert T: **Patterns of biomass allocation between foliage and woody structure: The effects of tree size and specific functional traits**. *Ann For Res* 2016, **59**(1):1-12.
36. Pajtik J, Konopka B, Lukac M: **Individual biomass factors for beech, oak and pine in Slovakia: a comparative study in young naturally regenerated stands**. *Trees* 2011, **25**(2):277-288.
37. Cairns MA, Brown S, Baumgardner HGA: **Root biomass allocation in the world's upland forests**. *Oecologia* 1997, **111**(1):1-11.
38. Zeng W: **Developing one-variable individual tree biomass models based on wood density for 34 tree species in China**. *Forest Resources Management* 2017(06):41-46. (In Chinese).
39. George G: **Prediction Bias and Response Surface Curvature**. *Forest Sci* 1991, **37**(3):755-765.
40. Riofrio J, Herrero C, Grijalva J, Bravo F: **Aboveground tree additive biomass models in Ecuadorian highland agroforestry systems**. *Biomass Bioenerg* 2015, **80**:252-259.
41. Bi HQ, Turner J, Lambert MJ: **Additive biomass equations for native eucalypt forest trees of temperate Australia**. *Trees* 2004, **18**(4):467-479.
42. Meng S, Jia Q, Liu Q, Zhou G, Wang H, Yu J: **Aboveground Biomass Allocation and Additive Allometric Models for Natural Larix gmelinii in the Western Daxing'anling Mountains, Northeastern China**. *Forests* 2019, **10**(2):150.
43. Kuyah S, Sileshi GW, Rosenstock TS: **Allometric models based on bayesian frameworks give better estimates of aboveground biomass in the miombo woodlands**. *Forests* 2016, **7**(2).
44. Abich A, Mucheye T, Tebikew M, Gebremariam Y, Alemu A: **Species-specific allometric equations for improving aboveground biomass estimates of dry deciduous woodland ecosystems**. *Journal of Forestry Research* 2019, **30**(5):1619-1632.
45. Tetemke BA, Birhane E, Rannestad MM, Eid T: **Allometric Models for Predicting Aboveground Biomass of Trees in the Dry Afromontane Forests of Northern Ethiopia**. *Forests* 2019, **10**(12):1114.
46. Rutishauser E, Noor'An F, Laumonier Y, Halperin J, Rufi'ie, Hergoualc'H K, Verchot L: **Generic allometric models including height best estimate forest biomass and carbon stocks in Indonesia**. *Forest Ecol Manag* 2013, **307**(Complete):219-225.
47. Lambert MC, Ung CH, Raulier F: **Canadian national tree aboveground biomass equations**. *Can J Forest Res* 2005, **35**(8):1996-2018.
48. Meng S, Liu Q, Zhou G, Jia Q, Zhuang H, Zhou H: **Aboveground tree additive biomass equations for two dominant deciduous tree species in Daxing'anling, northernmost China**. *J Forest Res-JPN* 2017, **22**(4):233-240.
49. Wagner RG, Ter-Mikaelian MT: **Comparison of biomass component equations for four species of northern coniferous tree seedlings**. *Ann Forest Sci* 1999, **56**(3):193-199.
50. Goodman RC, Phillips OL, Baker TR: **The importance of crown dimensions to improve tropical tree biomass estimates**. *Ecol Appl* 2016, **24**.

51. Xiao X, White EP, Hooten MB, Durham SL: **On the use of log-transformation vs. nonlinear regression for analyzing biological power laws**. *Ecology* 2011, **92**(10):1887-1894.
52. Packard GC: **On the use of logarithmic transformations in allometric analyses**. *J Theor Biol* 2009, **257**(3):515-518.
53. Packard GC: **On the use of log-transformation versus nonlinear regression for analyzing biological power laws**. *Biol J Linn Soc* 2014, **113**(4):1167-1178.
54. Sprugel DG: **Correcting for bias in log-transformed allometric equations**. *Ecology* 1983, **64**(1):209-210.
55. Zhao D, Kane M, Markewitz D, Teskey R, Clutter M: **Additive Tree Biomass Equations for Midrotation Loblolly Pine Plantations**. *Forest Sci* 2015, **61**(4):613-623.
56. Wang X, Zhao D, Liu G, Yang C, Teskey RO: **Additive tree biomass equations for Betula platyphylla Suk. plantations in Northeast China**. *Annals of Forest Ence* 2018, **75**(2):60.

Fig. 1 The location of the study area and sampling sites in Northeastern China.

Fig. 2 The scatter points depict height against diameter and the vertical bars represent the sample distribution across 4-cm diameter class.

Fig. 3 The fraction of aboveground biomass allocated in wood, bark, branch, and foliage varied with diameter, height, crown length, crown width. Red solid lines and grey areas are loess smoothing with 95% confidence interval.

Fig. 4 Root-shoot ratio against diameter (n = 53). The red line represents the fitted power function, and the blue short dashed line represents the mean value of the root-to-shoot ratio.

Fig. 5 Relationships between stem wood, bark, branch, foliage, aboveground, and root biomass of the sampled trees and diameter, height, crown length, crown width.

Fig. 6 Scatterplots of the predicted (y-axis) and the observed (x-axis) values for the wood, bark, branch, leaf, aboveground, and root biomass using the leave-one-out cross-validation method.

Fig. 7 Model predictive ability tested by the leave-one-out cross-validation method.

Fig. 8 The heteroscedastic and homoscedastic residual errors of the original biomass models (unweighted) and the biomass models with weighting functions (weighted).

Fig. 9 Distributions of relative error of published Mongolian oak biomass models. The length of the box indicates the 25th-75th percentile, the horizontal lines attached to the box represent the smallest and maximum values, the dot inside the box represents the mean value, the horizontal line inside the box represents the median, outliers are marked as asterisks. The violin represents a probability density.

Table 1 Summary statistics of *D* (cm), *H* (m), *CW* (m), *CL* (m), stem wood, bark, branch, foliage, aboveground, and root biomass (kg) for the sampled oak trees.

Statistics	***n***	**Mean**	**Min**	**Max**	**SD**
D	159	15.1	1.5	33	10.0
H	159	9.11	1.4	20.5	4.94
CW	159	4.22	0.45	9.70	3.51
CL	159	6.13	0.60	13.61	3.51
Wood	159	73.55	0.14	336.27	90.67
Bark	159	16.26	0.06	74.66	19.13
Branch	159	45.19	0.03	298.32	62.55
Foliage	159	5.59	0.03	26.57	6.75
Aboveground	159	140.59	0.33	668.12	173.45
Root	53	37.66	0.23	153.71	4.08

n, sample size; Min, minimum; Max: maximum; SD, standard deviation.

Table 2 Comparison of existing models using the measured biomass data in this study

Study	Model type	Predictor	*D* range (cm)	Sample size	Site	*N*
Dong et al [8]	Type 1	*D*	4.2-37.1	64	HLJ, JL	100
	Type 2	*D*, *H*				
Wang [10]	Type 3	*D*	4.3-57.1	10	HLJ	99
He et al [9]	Type 4	*D*	4.2-41.2	10	JL	100

N is the number of measured trees used for testing the existing equations. HLJ: Heilongjiang province; JL: Jilin province.

Table 3 Regression coefficients and statistics of the aboveground additive systems and root biomass equations with a specific weighting function.

Component	Model Predictor	Regression coefficients					Fitting Statistics		Weight function
		β_0	β_1	β_2	β_3	β_4	RMSE (kg)	R^2_{adj}	
Wood	*D*	0.050***	2.473***				30.602	0.886	$D^{4.9241}$
	D, *H*	0.033***	1.761***	1.043***			19.138	0.955	$D^{4.2932}$
	D, *CW*	0.061***	2.252***	0.271***			30.605	0.886	$D^{3.3857}$
	D, *CL*	0.059***	1.845***	0.826***			25.121	0.923	$D^{3.6802}$
	D, *H*, *CW*	0.034***	1.689***	1.031***	0.121*		18.833	0.957	$D^{4.3493}$
	D, *H*, *CL*	0.030***	1.834***	1.191***	-0.235***		18.968	0.956	$D^{5.1990}$
	D*, *H*, *CW*, *CL	**0.034*****	**1.723*****	**1.173*****	**0.174*****	**-0.252*****	**18.548**	**0.958**	$\boldsymbol{D^{4.3589}}$
Bark	*D*	0.022***	2.249***				5.397	0.920	$D^{4.2996}$
	D, *H*	0.018***	1.916***	0.491***			4.571	0.943	$D^{4.1216}$
	D, *CW*	0.022***	2.291***	-0.076^{ns}			5.354	0.922	$D^{4.2709}$
	D, *CL*	0.024***	1.975***	0.367***			4.835	0.936	$D^{4.0130}$
	D*, *H*, *CW	**0.017*****	**2.022*****	**0.467*****	**-0.146***		**4.466**	**0.946**	$\boldsymbol{D^{4.1201}}$
	D, *H*, *CL*	0.018***	1.939***	0.565***	-0.107^{ns}		4.623	0.942	$D^{4.1430}$
	D, *H*, *CW*, *CL*	0.017***	2.003***	0.560***	-0.090^{ns}	-0.109^{ns}	4.561	0.943	$D^{4.1967}$
Branch	*D*	0.011***	2.772***				23.320	0.861	$D^{4.2891}$
	D, *H*	0.012***	2.716***	0.035^{ns}			23.440	0.860	$D^{4.3426}$
	D*, *CW	**0.018*****	**2.090*****	**0.939*****			**19.081**	**0.907**	$\boldsymbol{D^{4.9263}}$
	D, *CL*	0.014***	2.504***	0.290**			23.052	0.864	$D^{5.0652}$
	D, *H*, *CW*	0.020***	2.128***	-0.110^{ns}	0.964***		19.060	0.907	$D^{4.8475}$
	D, *H*, *CL*	0.019***	2.572***	-0.538**	0.692***		22.606	0.869	$D^{3.4718}$
	D, *H*, *CW*, *CL*	0.033***	1.976***	-0.655***	0.985***	0.607***	19.426	0.904	$D^{4.6120}$
Foliage	*D*	0.019***	1.964***				3.350	0.754	$D^{3.7889}$
	D, *H*	0.025***	1.981***	-0.126^{ns}			3.429	0.742	$D^{2.9305}$
	D*, *CW	**0.023*****	**1.626*****	**0.473*****			**3.275**	**0.765**	$\boldsymbol{D^{3.9640}}$
	D, *CL*	0.017***	2.029***	-0.050^{ns}			3.370	0.751	$D^{3.4793}$
	D, *H*, *CW*	0.020***	1.747***	-0.052^{ns}	0.441**		3.296	0.762	$D^{3.0790}$
	D, *H*, *CL*	0.025***	1.930***	-0.053^{ns}	-0.018^{ns}		3.401	0.746	$D^{2.9473}$
	D, *H*, *CW*, *CL*	0.043***	1.692***	-0.704***	0.590***	0.323*	3.664	0.706	$D^{3.9980}$
Root	***D***	**0.086*****	**2.072*****				**14.843**	**0.887**	$\boldsymbol{D^{4.2914}}$
	D, *H*	0.077***	2.118***	-0.008^{ns}			14.725	0.888	$D^{3.6076}$
	D, *CW*	0.087***	2.033***	0.058^{ns}			14.991	0.884	$D^{4.2293}$
	D, *CL*	0.076***	2.091***	0.044^{ns}			14.518	0.892	$D^{3.3309}$
	D, *H*, *CW*	0.075**	2.112***	0.012^{ns}	-0.002^{ns}		14.769	0.888	$D^{3.4907}$
	D, *H*, *CL*	0.081**	2.099***	-0.076^{ns}	0.083^{ns}		14.881	0.886	$D^{3.5074}$
	D, *H*, *CW*, *CL*	0.081**	2.104***	-0.069^{ns}	-0.008^{ns}	0.081^{ns}	15.008	0.884	$D^{3.4890}$

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, ^{ns} non-significant; aboveground weight function: $D^{4.5621}$, $D^{4.6374}$, $D^{4.5688}$, $D^{4.0230}$, $D^{4.0223}$, $D^{5.2550}$, $D^{5.4720}$ for equation (2)-(8), respectively. The selected models for each component are given in bold.

Table 4 Aboveground additive systems of biomass models based on *D* and selected predictors

Model system	Component	Model expressions	RMSE (kg)	R^2_{adj}
System 1	Wood	$W_{wd} = 0.050D^{2.473}$	30.602	0.886
	Bark	$W_{bk} = 0.022D^{2.249}$	5.397	0.920
	Branch	$W_{br} = 0.011D^{2.772}$	23.320	0.861
	Foliage	$W_{fo} = 0.019D^{1.964}$	3.350	0.754
	Aboveground	$W_{ag} = 0.050D^{2.473} + 0.022D^{2.249} + 0.011D^{2.772} + 0.019D^{1.964}$	45.478	0.931
System 2	Wood	$W_{wd} = 0.028D^{1.740}H^{1.299}CW^{0.140}CL^{-0.312}$	18.411	0.959
	Bark	$W_{bk} = 0.017D^{2.010}H^{0.502}CW^{-0.132}$	4.473	0.945
	Branch	$W_{br} = 0.017D^{2.156}CW^{0.870}$	19.005	0.908
	Foliage	$W_{fo} = 0.023D^{1.645}CW^{0.438}$	3.280	0.764
	Aboveground	$W_{ag} = 0.028D^{1.740}H^{1.299}CW^{0.140}CL^{-0.312} + 0.017D^{2.010}H^{0.502}CW^{-0.132} + 0.017D^{2.156}CW^{0.870} + 0.023D^{1.645}CW^{0.438}$	31.021	0.968

All coefficients are significant.

Figures

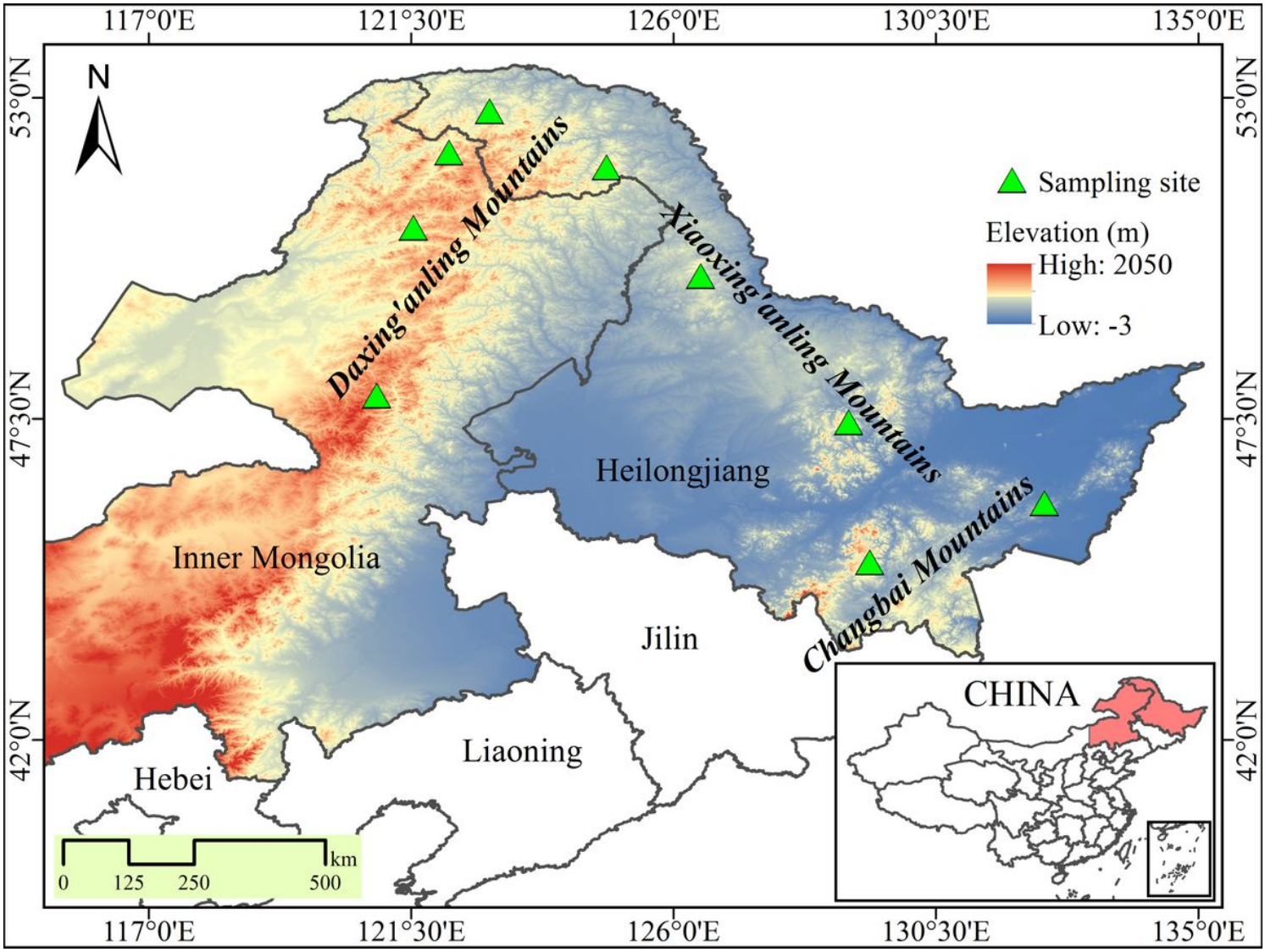

Figure 1

The location of the study area and sampling sites in Northeastern China. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

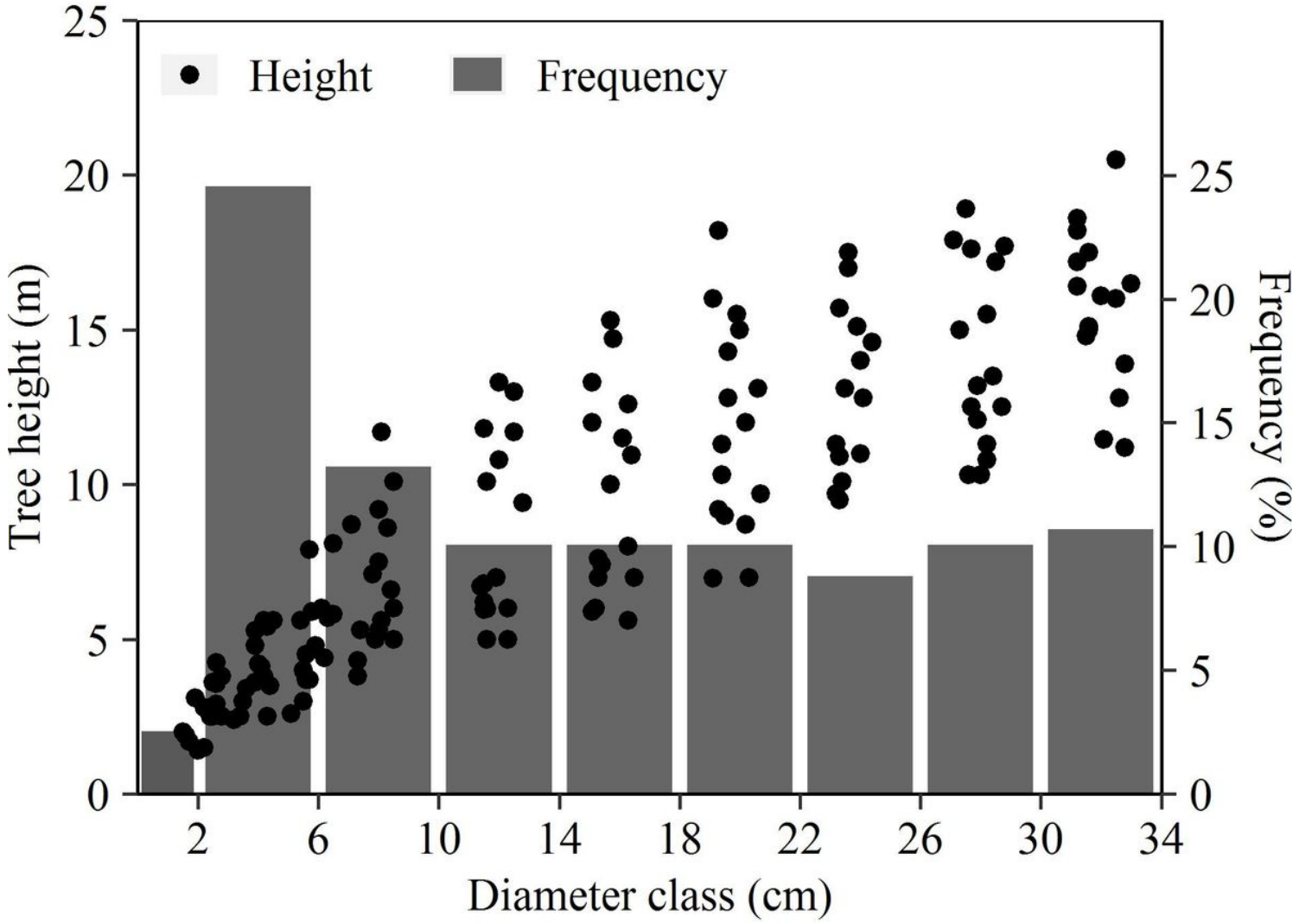

Figure 2

The scatter points depict height against diameter and the vertical bars represent the sample distribution across 4-cm diameter class.

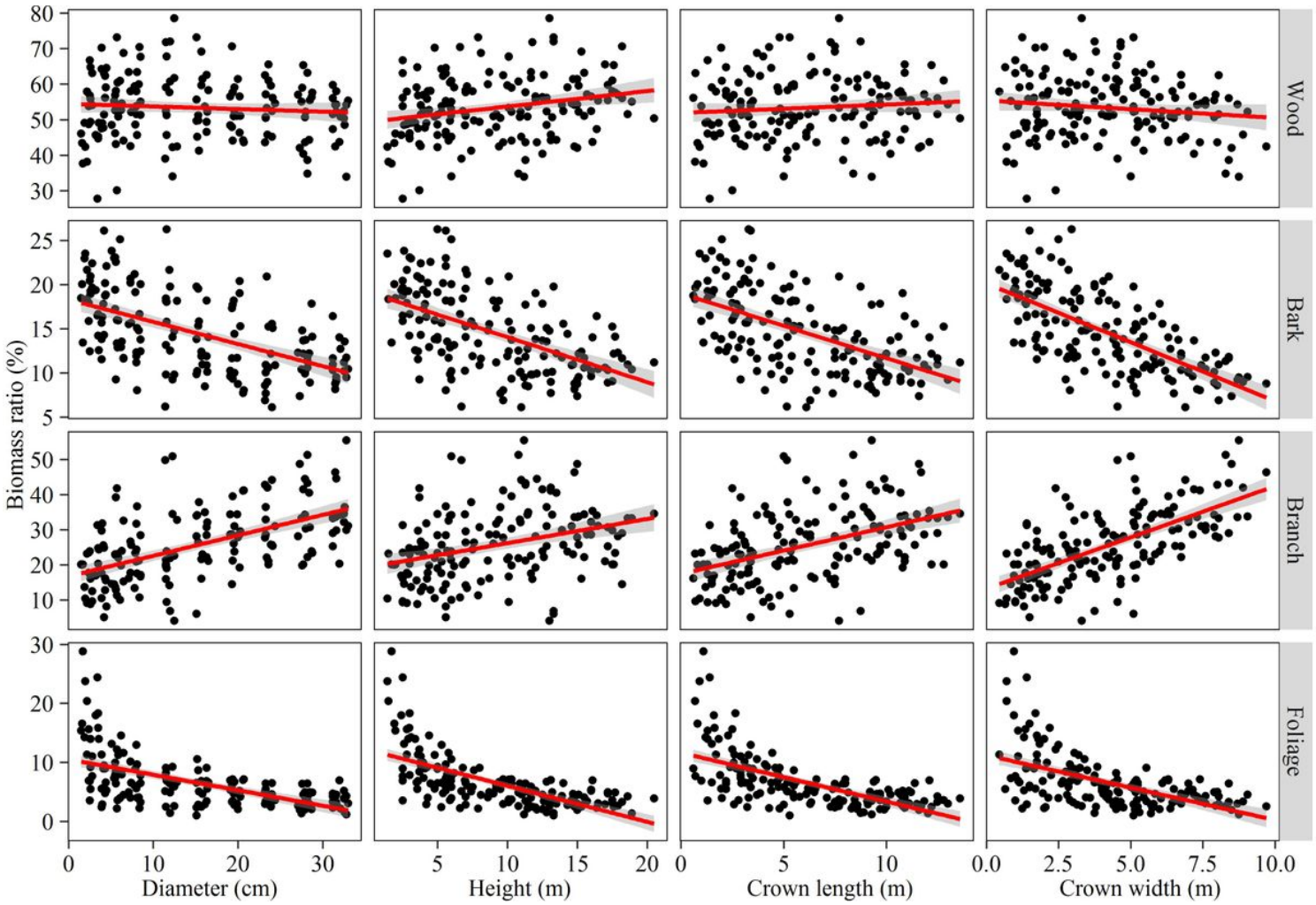

Figure 3

The fraction of aboveground biomass allocated in wood, bark, branch, and foliage varied with diameter, height, crown length, crown width. Red solid lines and grey areas are loess smoothing with 95% confidence interval.

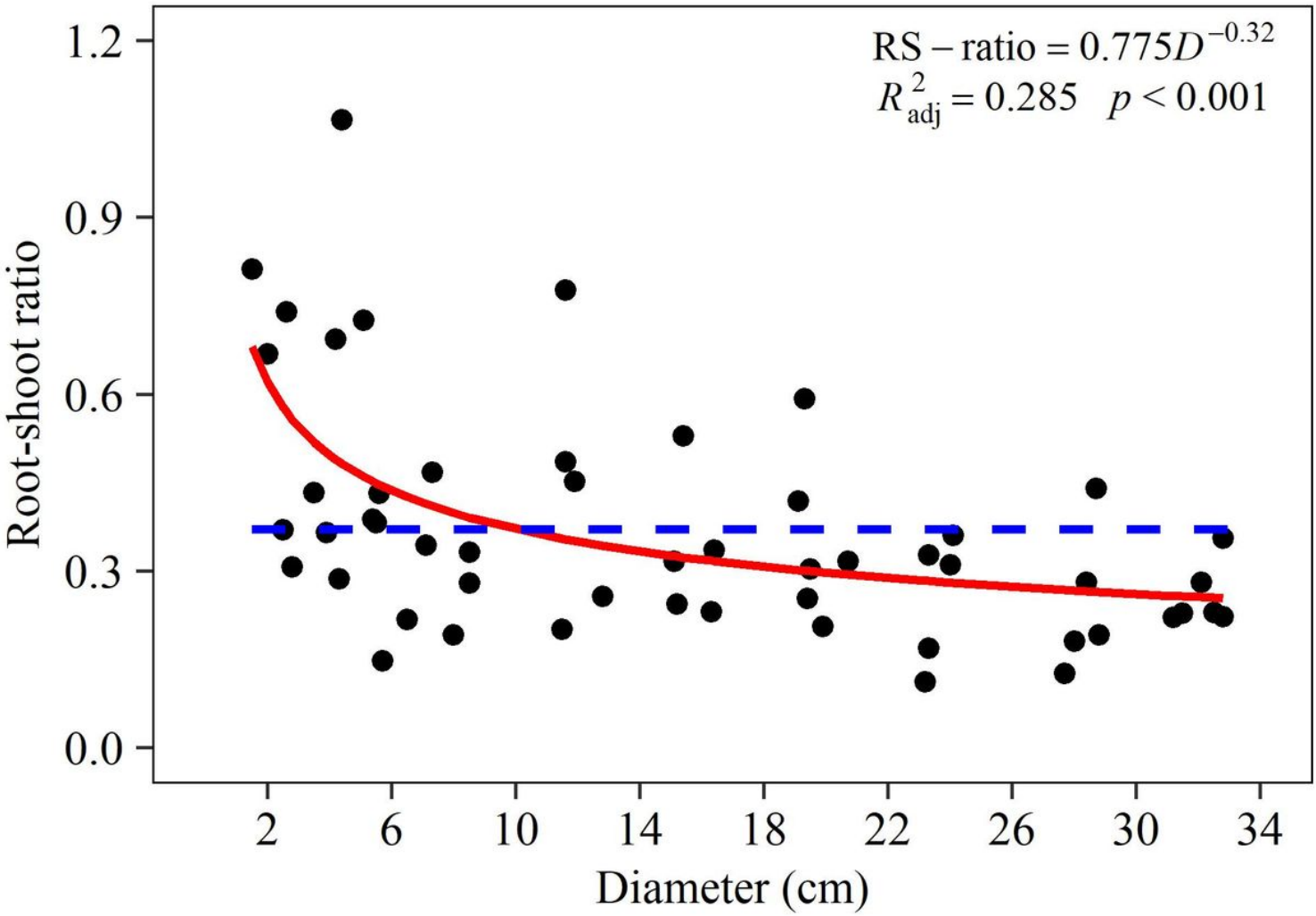

Figure 4

Root-shoot ratio against diameter (n = 53). The red line represents the fitted power function, and the blue short dashed line represents the mean value of the root-to-shoot ratio.

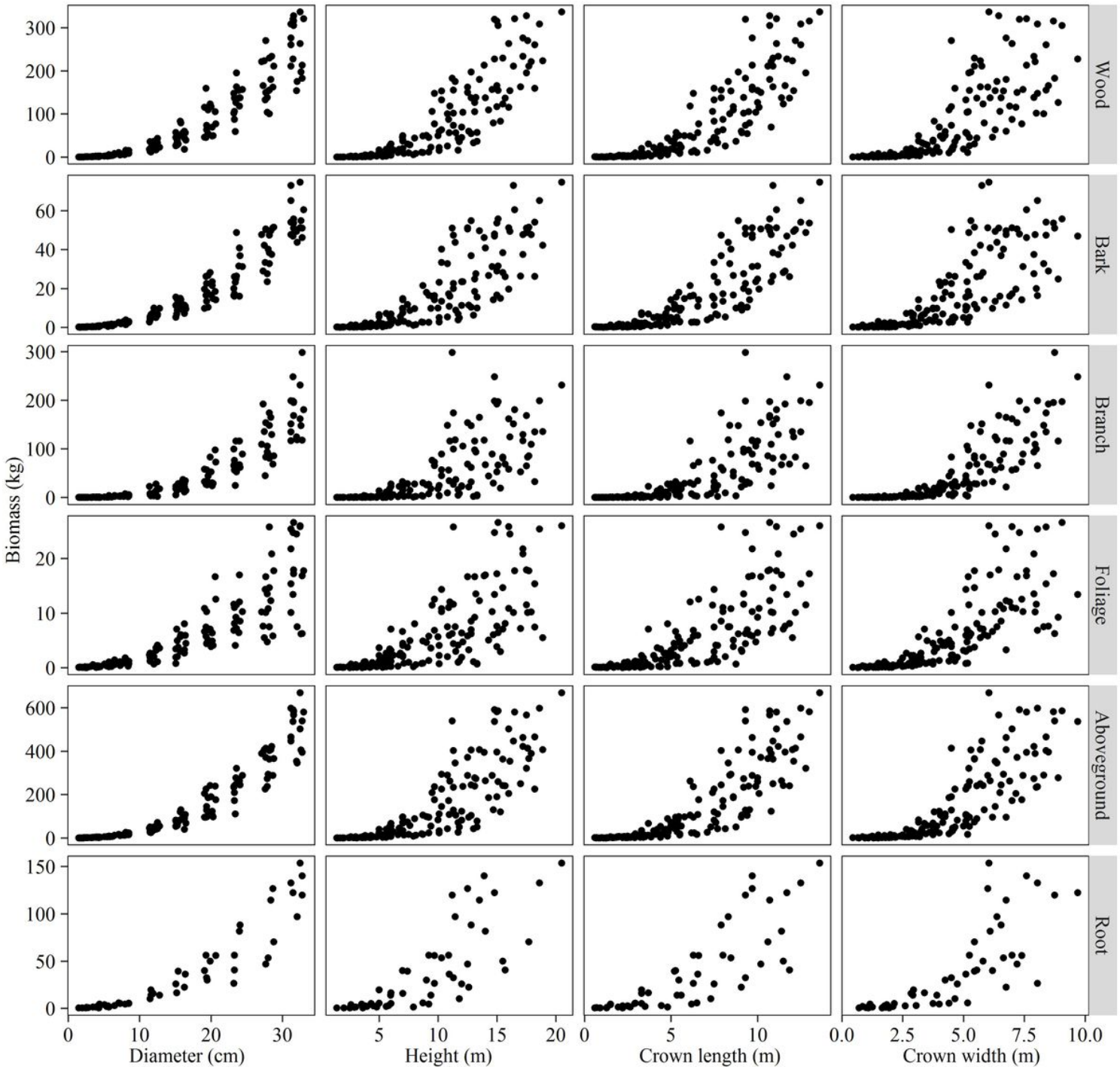

Figure 5

Relationships between stem wood, bark, branch, foliage, aboveground, and root biomass of the sampled trees and diameter, height, crown length, crown width.

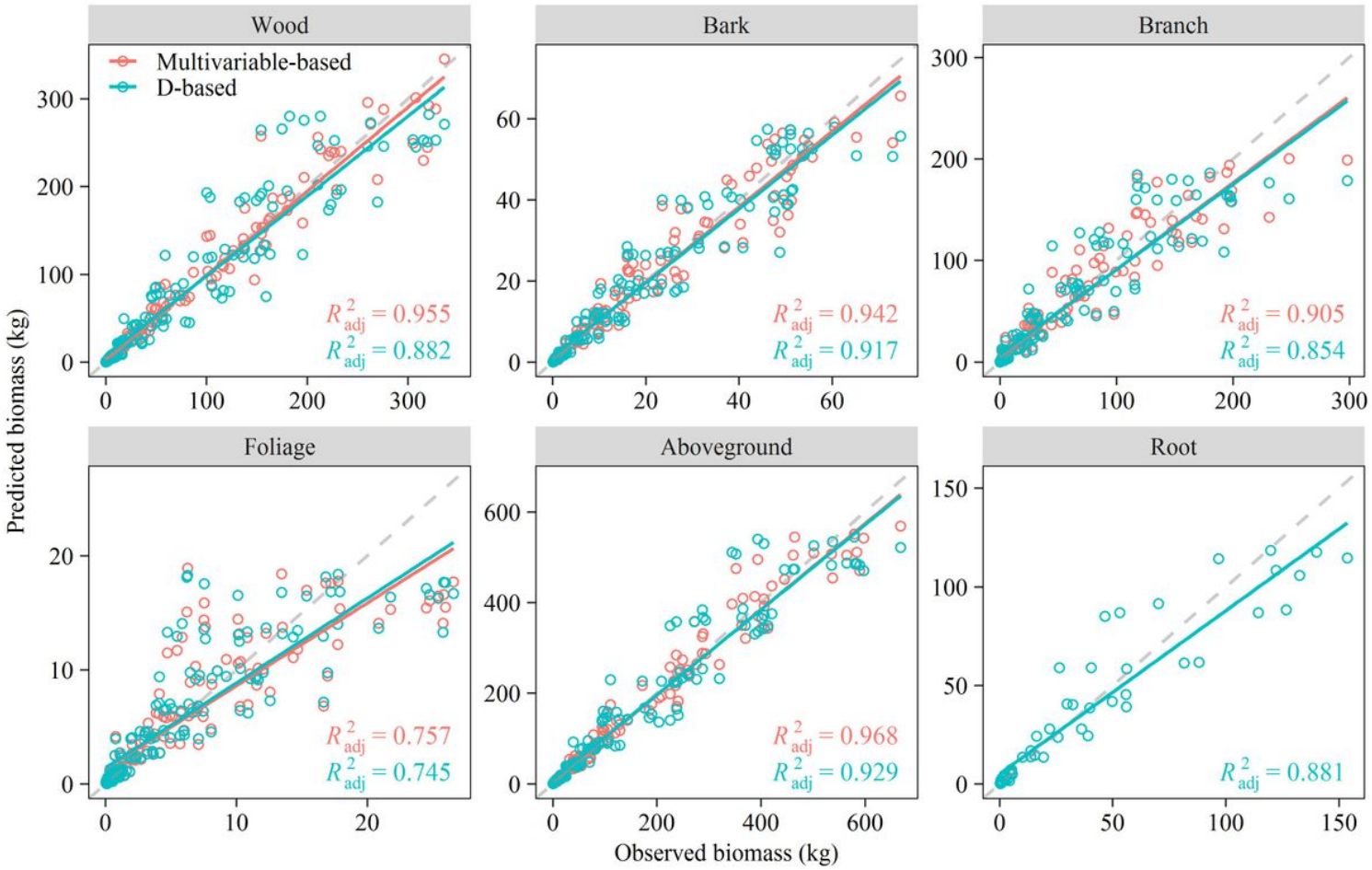

Figure 6

Scatterplots of the predicted (y-axis) and the observed (x-axis) values for the wood, bark, branch, leaf, aboveground, and root biomass using the leave-one-out cross-validation method.

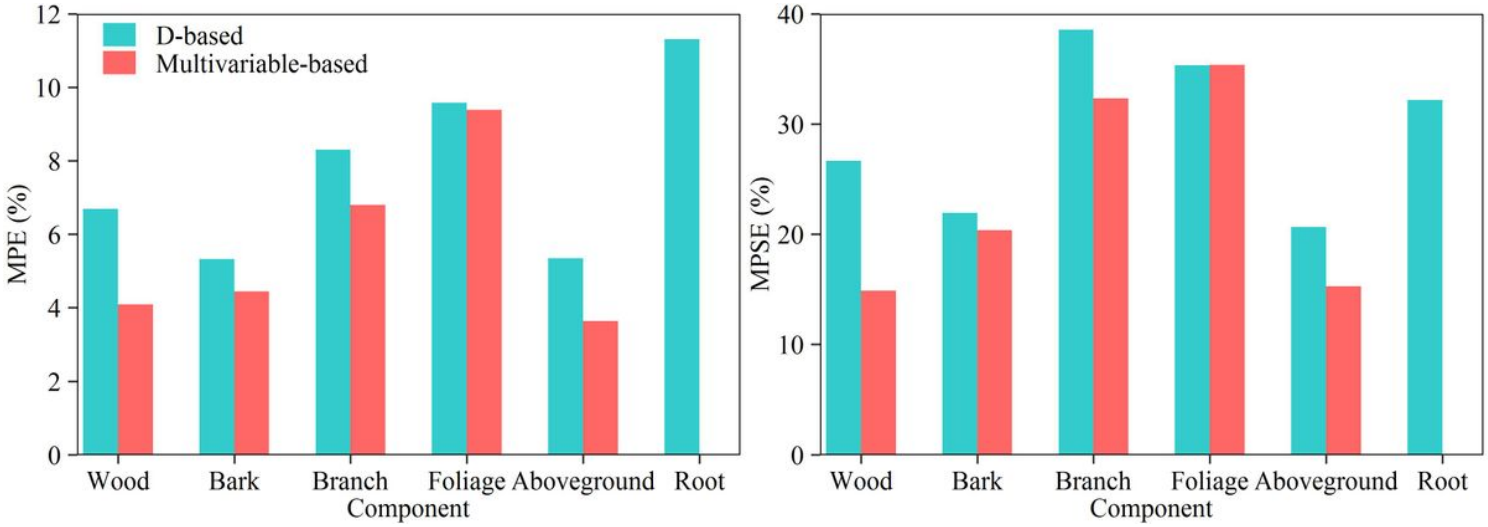

Figure 7

Model predictive ability tested by the leave-one-out cross-validation method.

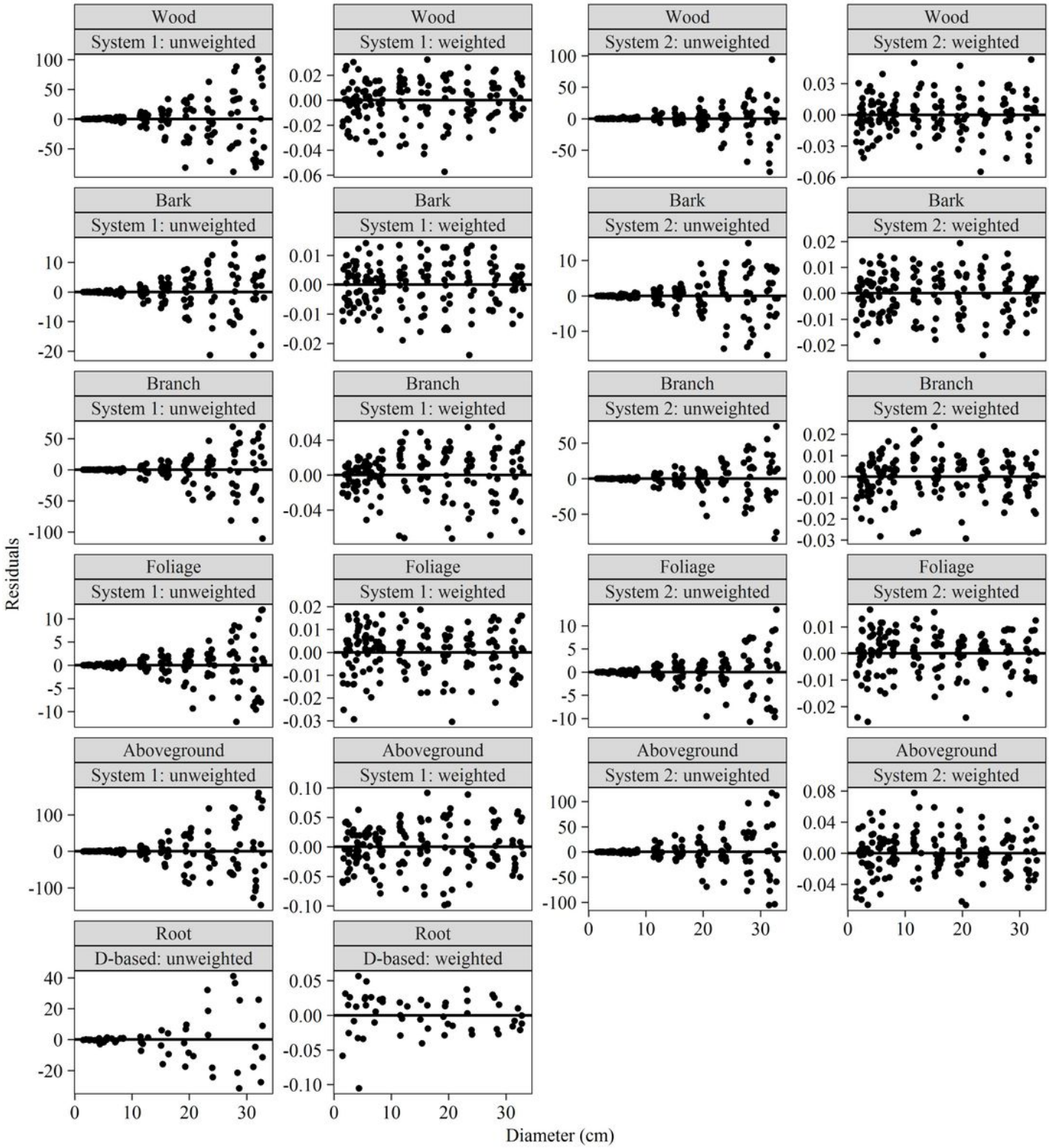

Figure 8

The heteroscedastic and homoscedastic residual errors of the original biomass models (unweighted) and the biomass models with weighting functions (weighted).

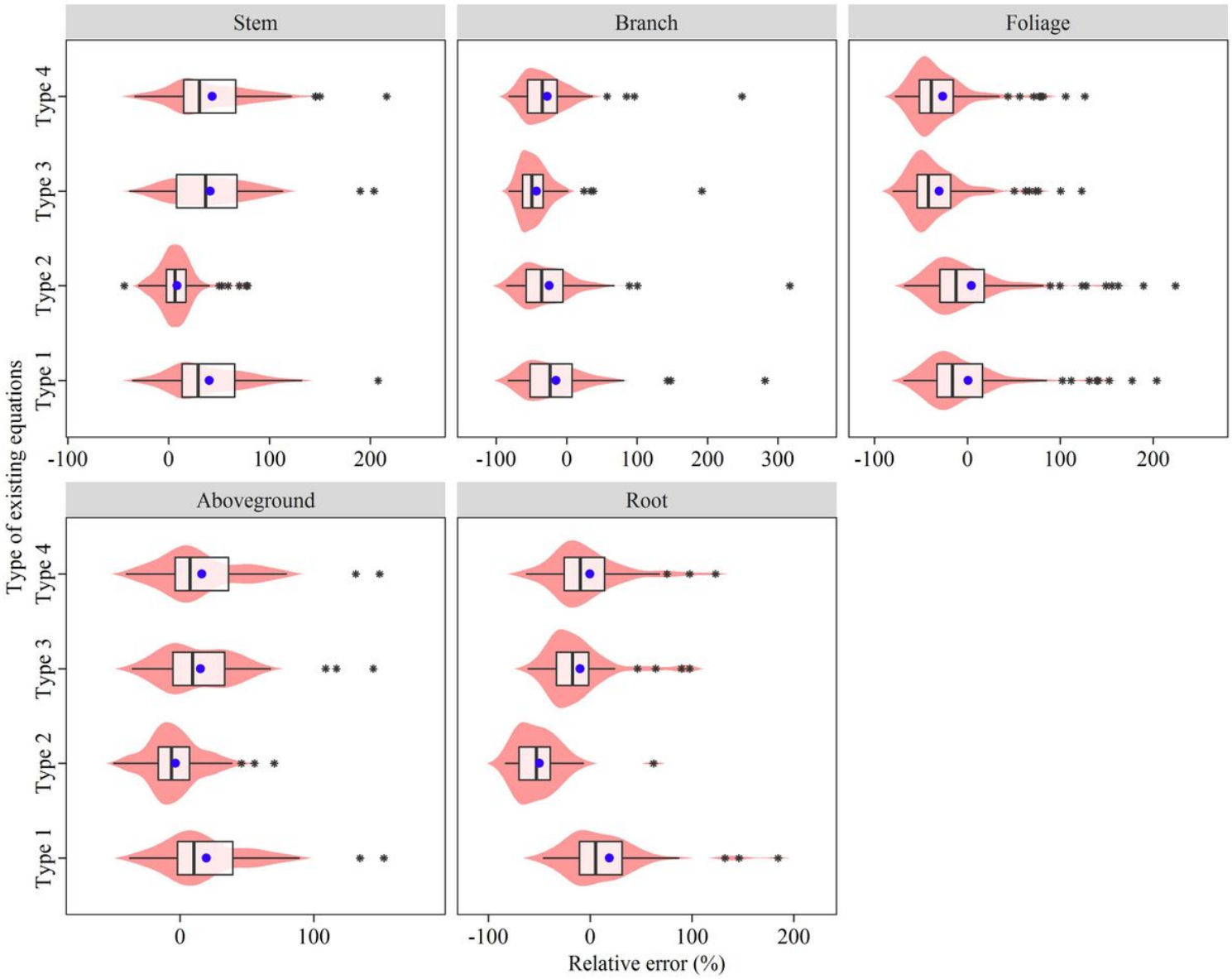

Figure 9

Distributions of relative error of published Mongolian oak biomass models. The length of the box indicates the 25th-75th percentile, the horizontal lines attached to the box represent the smallest and maximum values, the dot inside the box represents the mean value, the horizontal line inside the box represents the median, outliers are marked as asterisks. The violin represents a probability density.