Construction of Compatible Volume Model for Populus in Beijing, China

Shan Wang 1, Zhichao Wang 2,*, Zhongke Feng 1,3,* 4, Zhuang Yu 1 and Jinshan Li 4

Abstract: The accurate assessment of tree volume is crucial for developing forest management plans, and this can be achieved using tree volume models. In this study, data on individual trees were collected and calculated, including the diameter at breast height (D), ground diameter (DG), tree height (H), and tree volume (V). A total of 400 Populus × tomentosa Carrière, 400 Populus tomentosa Carr, and 400 Populus × canadensis Moench trees were sampled. Two compatible volume model systems were established using corresponding methods. The models consisted of the following five types: V-DH, V-D, V-DG, H-D, and DG-D. In our calculations, before the horizontal line was the dependent variable, and behind the horizontal line was the independent variable. Variations in preferences for the H-D models were observed among the tree species, with the logistic function proving the most suitable for Populus × tomentosa Carrière, Chapman–Richard function for Populus tomentosa Carr, and power function for Populus × canadensis Moench. Among the three volume models, the V-DH model exhibited a superior performance, with its $R^2$ values ranging from 0.965 to 0.984 and mean estimated error (MPE) values ranging from 1.26% to 1.78%. Following this was the V-D model, with $R^2$ values between 0.9359 and 0.9704 and MPE values between 1.71% and 2.46%. The V-DG model ranked third, with $R^2$ values ranging from 0.8746 to 0.9501 and MPE values ranging from 2.33% to 3.16%. In the H-D model, the $R^2$ and MPE values ranged from 0.4998 to 0.7851 and from 1.31% to 1.45%, respectively. For the DG-D model, the $R^2$ values ranged from 0.9563 to 0.9868 and the MPE values ranged from 0.47% to 0.68%. Comparing both compatible methods, the nonlinear seemingly uncorrelated regression (NSUR) was more effective. The three volume models demonstrated high levels of accuracy and compatibility, providing a reliable scientific foundation for forest management and the formulation of harvesting plans in Beijing, with significant practical implications.

Keywords: tree volume; compatible model; forest management; forest harvesting

1. Introduction

Populus L. is, globally, the most widely distributed and adaptable tree species, with rapid growth and significant economic value [1]. Accurately assessing the wood resources of this species in Beijing is of considerable practical importance and will aid forest managers in timber valuation and the allocation of forest harvesting areas [2].

Undoubtedly, the most reliable approach for acquiring tree volume data is the harvesting of all trees in a forest. However, this method is characterized by its destructiveness, time intensiveness, and high cost, rendering it feasible only on a limited scale [3]. With advancements in modern science and technology, foresters have turned to precision measurement instruments for forest surveys. Certain studies have applied airborne laser radar to conduct large-scale measurements of vegetation structures, realizing accurate estimations of tree volumes, biomass, and carbon storage [4]. Nevertheless, precision equipment...
is expensive and involves technical complexities, limiting its widespread application in forestry investigations. Other studies have employed total station and electronic theodolite for measuring tree heights, volumes, and other basic tree factors (e.g., diameters at breast height) [5–7]. Total station equipment is costly, lacks portability, and faces challenges in measuring within complex stands, whereas electronic theodolites offer a comparatively lower cost, lighter weight, ease of transport, and operational simplicity.

Volume models offer a nondestructive method for estimating tree volume [8]. National-scale volume models have limited accuracy and applicability. Due to the variability in ecological conditions across large geographic areas, the predictive accuracy of these models decreases when they are applied to smaller scales [9–11]. For example, Case and Hall (2008) compared volume models at local, regional, and national scales for 10 tree species in central and western Canada [11]. As the generality of the model increased, the estimation error for local areas increased. In China, by the late 1970s, only three poplar volume tables specific to *Populus davidiana* Dode in the northeast, north, and northwest regions were enforced. A locally developed volume model offers a more precise estimate [12]. Therefore, establishing a poplar volume model for the Beijing area appears to be more feasible.

Most studies have utilized multivariate statistical methods to construct volume models, which are known for their clarity in expression. Conversely, some studies have employed machine learning techniques such as artificial neural networks, support vector machines, and random forests for volume modeling [13–15]. Compared with machine learning, a multivariate statistical model cannot easily fit large-scale data containing noise [16]. Although machine learning adeptly handles complex nonlinear relationships and demonstrates strong adaptability, it lacks explicit model expressions and practical applicability. Explicit multivariate statistical models are more appropriate for volume prediction in production scenarios than implicit machine learning approaches. In explicit models, tree volume estimation is achievable using either a tapered or volume model. A taper model facilitates integration to obtain trunk volume and volume at any given position [17]. For cases in which only trunk volume is of interest, a volume model is recommended.

In volume modeling, the selection of predictors holds particular significance. The diameter at breast height (*D*) reflects the radial growth of a tree, and tree height (*H*) represents vertical growth, both of which are deemed effective predictors [18]. Numerous allometric relationships have been established between tree volume (*V*), *D*, and *H*, exhibiting excellent performance and widespread application [3,19–22]. Due to the complexity of forest habitats, *D* can be more easily measured under field conditions, while *H* measurements encounter technical and operational challenges. It is common to measure the *D* values for all trees and the *H* values for a subset of sample trees and then utilize the *H*-*D* model to predict missing *H* values [23]. In addition, the *H*-*D* model finds application in stand growth models, particularly in determining the mean stand *H* based on the corresponding mean stand *D* [24]. In practice, instances of illegal tree felling arise, leaving only ground diameter (*DG*) information for felled trees. Hence, *DG* serves as a valuable predictor for accurately estimating tree-loss resources [25].

The volume model selection is of great importance. Various forms of bivariate volume models exist, including the combined variable volume equation [26], the Schumacher and Hall volume equation [27], and comprehensive combined variable equations [28]. Among these, the Schumacher and Hall functions are the most commonly utilized [19,21,29,30]. For univariate volume models, linear, exponential, power, polynomial, sigmoid, and inverse proportional functions are frequently used [3]. Numerous studies have demonstrated the effectiveness of the power function for volume predictions [3,19,31]. Thakur Subedi tested six univariate volume models and observed that the power function performed the best, constructing a *Shorea robusta* volume model in western Nepal [3]. Similarly, Bruno O. Gimenez tested four univariate volume models, concluding that the power function yielded the best results and subsequently constructing an *Eschweileria* volume model in Manaus, Brazil [19]. Currently, univariate volume models, including *DG*, are relatively less developed.
The compatibility modeling method offers the distinct advantage of ensuring consistency in prediction results across multiple models and has found widespread application in various research fields. In certain studies, this method has been employed to guarantee the compatibility between biomass and volume models [29,30,32], whereas in others, it ensures compatibility between taper and volume models [33–35]. Additionally, some studies have utilized the compatibility modeling method to ensure compatibility between the total biomass and component biomass [10,36].

Nonlinear simultaneous equations are suitable for modeling when the dependent variable in one model is the independent variable in another or when there are shared parameters between models. Parameter estimation methods primarily include the nonlinear seemingly unrelated regression method (NSUR), the error-in-variable simultaneous equations method (EIV), and the generalized moment method (GMM) [37–39]. There is no consensus on the superiority of any particular parameter estimation method. NSUR and EIV have gained popularity in recent years [10,40–42]. Because they are more general and flexible and allow each component model to have its own independent variables, each component model can use a heteroscedasticity weighting function [37].

Based on measured data from three species, including *Populus × tomentosa* Carrière, *Populus tomentosa* Carr, and *Populus × canadensis* Moench, we established volume models using $D$, $H$, and $DG$ as predictors. The objective of this study was to develop a volume prediction model for three poplar species in Beijing. These models were designed to accurately estimate the volume, thus providing important data support to help forest managers develop scientific logging and conservation plans. To achieve this objective, a compatible volume model system was developed consisting of the following five models: $V-DH$, $V-D$, $V-DG$, $H-D$, and $DG-D$, with the $H-D$ and $DG-D$ models serving as bridges.

In our calculations, before the horizontal line was the dependent variable, and behind the horizontal line was the independent variable. Two compatible modeling methods were employed, and their advantages were compared and analyzed.

2. Materials and Methods

2.1. Study Site and Data

2.1.1. Study Site

Beijing (east longitude, 115°25′–117°35′ E and north latitude, 39°28′–41°03′ N) is situated at the intersection of the Inner Mongolia Plateau and the North China Plain, with an altitude of not more than 100 m and predominantly ranging between 30 and 50 m. The climate is characterized by a distinct warm, temperate, semi-humid continental monsoon pattern, featuring distinct seasons, with a hot and rainy summer and a cold and dry winter. The average annual temperature is approximately 11.5 °C, with a frost-free period lasting 5 to 6 months annually. The city receives an average annual precipitation of 585 mm, with approximately 74% occurring during the summer. In terms of zonal vegetation type, Beijing falls within a warm, temperate, deciduous, broad-leaved forest area.

2.1.2. Data Collection

Non-destructive sampling was used to obtain the samples. A total of 120 survey plots were established, each with a size of 1 hectare (Table S1 and Figure S1). Within these plots, trees situated in open areas with straight trunks were selected for investigation.

Table 1 presents descriptive statistics for the dataset, comprising measurements taken from 1200 populus trees. This sample included 400 Chinese white poplar (*Populus × tomentosa* Carrière), 400 fast-growing poplar (*Populus tomentosa* Carr), and 400 Canadian poplar (*Populus × canadensis* Moench) trees. The $DG$ and $D$ were measured using a diameter ruler (Pacific, nominal accuracy of ±1 mm). The $H$ and $V$ of each individual tree were non-destructively measured using a Southern electronic theodolite (DT-02) [5]. In the process of volume measurement, from a height of 1.3 m to the top of each tree, the horizontal angle when aiming at the left and right edges of a tree and the zenith distance when aiming at the right edge of a tree were measured in sequence. The height of each tree was calculated...
using the principles of triangulation. According to the roundtable accumulation method, the volume of a standing tree is determined using the sectional measurement method.

**Table 1.** Five statistical indicators (minimum, maximum, mean, standard deviation, and coefficient of variation) of the characteristics (D, DG, H, and V) of the three tree species (*Populus × tomentosa* Carrière, *Populus tomentosa* Carr, and *Populus × canadensis* Moench).

<table>
<thead>
<tr>
<th>Tree Species</th>
<th>Characteristics</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Coefficient of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Populus × tomentosa</em> Carrière</td>
<td>D/cm</td>
<td>6.6</td>
<td>69.5</td>
<td>25.0</td>
<td>13.2</td>
<td>52.9%</td>
</tr>
<tr>
<td></td>
<td>DG/cm</td>
<td>7.8</td>
<td>84.7</td>
<td>30.3</td>
<td>15.5</td>
<td>51.2%</td>
</tr>
<tr>
<td></td>
<td>H/m</td>
<td>9.6</td>
<td>36.8</td>
<td>21.6</td>
<td>6.4</td>
<td>29.8%</td>
</tr>
<tr>
<td></td>
<td>V/m³</td>
<td>0.0156</td>
<td>6.4576</td>
<td>0.7961</td>
<td>1.1344</td>
<td>142.5%</td>
</tr>
<tr>
<td><em>Populus tomentosa</em> Carr</td>
<td>D/cm</td>
<td>5.2</td>
<td>49.3</td>
<td>21.1</td>
<td>9.4</td>
<td>44.4%</td>
</tr>
<tr>
<td></td>
<td>DG/cm</td>
<td>6.9</td>
<td>56.3</td>
<td>24.8</td>
<td>10.3</td>
<td>41.4%</td>
</tr>
<tr>
<td></td>
<td>H/m</td>
<td>9.2</td>
<td>36.4</td>
<td>20.5</td>
<td>5.8</td>
<td>28.3%</td>
</tr>
<tr>
<td></td>
<td>V/m³</td>
<td>0.0114</td>
<td>2.0876</td>
<td>0.4470</td>
<td>0.4480</td>
<td>100.2%</td>
</tr>
<tr>
<td><em>Populus × canadensis</em> Moench</td>
<td>D/cm</td>
<td>8.5</td>
<td>73.1</td>
<td>39.3</td>
<td>12.9</td>
<td>32.8%</td>
</tr>
<tr>
<td></td>
<td>DG/cm</td>
<td>10.5</td>
<td>89.5</td>
<td>45.5</td>
<td>15.1</td>
<td>33.2%</td>
</tr>
<tr>
<td></td>
<td>H/m</td>
<td>10.3</td>
<td>35.0</td>
<td>24.1</td>
<td>4.6</td>
<td>19.1%</td>
</tr>
<tr>
<td></td>
<td>V/m³</td>
<td>0.0385</td>
<td>6.1957</td>
<td>1.4889</td>
<td>1.0187</td>
<td>68.4%</td>
</tr>
</tbody>
</table>

Note: diameter at breast height (D), ground diameter (DG), tree height (H), and tree volume (V).

### 2.2. Statistical Analysis

#### 2.2.1. Model Form Selection

For the H-D model, the candidate model forms included the power function, Chapman–Richards function, and logistics function (Models 1–4). A linear function (Model 5) was employed for the DG-D model. The Schumacher and Hall volume equation (Model 6) was adopted for the V-DH model. The calculations for each model were as follows:

Model 1: \[ H = a_0 D^{a_1}, \]

Model 2: \[ H = a_0 \left(1 - e^{-a_1 D}\right)^{a_2}, \]

Model 3: \[ H = a_0 \left(1 + a_1 \cdot e^{a_2 D}\right), \]

Model 4: \[ H = D / (a_0 + a_1 D), \]

Model 5: \[ DG = b_0 + b_1 D, \]

Model 6: \[ V = c_0 D^{c_1} H^{c_2}. \]

Akaike information criterion (AIC) statistics were used to assess the model complexity and its goodness of fit, with preference being given to models with smaller AIC values. Using the AIC minimization criterion, the optimal model form for the H-D model was selected.

#### 2.2.2. Compatible Volume Model

In modeling V-DH, V-D, and V-DG volume models, it is typical to treat them independently, often disregarding their correlations. Fitting these models individually can lead to incompatibility of volume predictions. In this study, the H-D and DG-D models were employed as bridges, employing two compatible modeling methods to ensure the compatibility of the predicted results for all three of the volume models.

Compatibility method 1: The V-DH model was denoted as \( \hat{V} = f_1(D, H) \), the H-D model was denoted as \( \hat{H} = f_2(D) \), and the DG-D model was denoted as \( \hat{DG} = f_3(D) \). The
error-in-variable simultaneous equations method was employed alongside a nonlinear seemingly unrelated regression to estimate the parameters, as follows:

\[
\begin{align*}
V &= f_1(D, H) = c_0D^a H^b \\
\hat{H} &= f_2(D) \\
DG &= f_3(D)
\end{align*}
\]

Using the estimation of value of \( f_2(D) \), the \( V-D \) model was derived, i.e., \( \hat{V} = f_1(D, f_2(D)) \). The inverse function of \( f_3(D) \) was obtained. The \( V-DG \) model was then derived, i.e., \( \hat{V} = f_1\left(f_3^{-1}(DG), f_2\left(f_3^{-1}(DG)\right)\right) \).

Compatibility method 2: The value of \( V = aD^bH^c \) is fitted, and the parameters of the \( V-DH \) model were obtained. After the parameters of the \( V-DH \) model were determined, the \( H-D \) model, \( \hat{H} = g_1(D) \), was obtained by fitting the \( V-D \) model, i.e., \( V = aD^b\left(g_1(D)\right)^c \).

After fixing the parameters of the \( V-D \) model, the \( D-DG \) model, \( \hat{D} = g_2\left(DG\right) \), was obtained by fitting the \( V-DG \) model, i.e., \( V = a\left(g_2\left(DG\right)\right)^b\left(g_1\left(g_2\left(DG\right)\right)\right)^c \). The inverse function of the \( D-DG \) model was the \( DG-D \) model.

2.2.3. Model Evaluation and Validation

This study employed all samples for modeling without distinguishing between the modeling and test samples. The six indicators for evaluating the regression models included the coefficient of determination (\( R^2 \)), standard error of estimates (\( SEE \)), total relative error (\( TRE \)), mean systematic error (\( MSE \)), mean estimated error (\( MPE \)), and mean percentage standard error (\( MPSE \)), as follows:

\[
R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2},
\]

\[
SEE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n - p}},
\]

\[
TRE = \frac{\sum (y_i - \hat{y}_i)}{\sum \hat{y}_i} \times 100,
\]

\[
MSE = \frac{\sum (y_i - \hat{y}_i)}{\hat{y}_i} \times 100,
\]

\[
MPE = t_a \times \left(\frac{SEE}{\bar{y}}\right) / \sqrt{n} \times 100,
\]

\[
MPSE = \frac{1}{n} \sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{\hat{y}_i}\right| \times 100,
\]

where \( y_i \) and \( \hat{y}_i \) denote the measured and estimated values of the \( i \)-th sample, respectively; \( \bar{y} \) denotes the average of the measured value of all samples; \( n \) represents the number of samples; \( p \) represents the number of model parameters; and \( t_a \) represents the \( t \)-value of the \( n \)-\( p \) degrees of freedom and the 95% confidence level.

2.2.4. Correlation Analysis

The Pearson correlation coefficient was used to calculate the correlations between the variables. Nonmetric multidimensional scaling (NMDS) was used in this study, which simplifies multidimensional space to low-dimensional space. An analysis of similarities (ANOSIM) was used to test whether the differences in the trees between the groups were significantly greater than the differences in the trees within the groups based on Bray-Curtis distance.

Figure 1 illustrates the data collection and analysis process (Figure 1). All statistical calculations were conducted using R 4.3.1 [43], and parameter estimations were performed using the systemfit package [44].
Three tree species
- Populus × tomentosa Carrière
- Populus tomentosa Carr
- Populus × canadensis Moench

Tree factors
- Diameter at breast height (D)
- Ground diameter (DG)
- Tree Height (H)
- Tree Volume (V)

Data collection

Base model
- DG-D model
  \[ DG = b_1 + b_2D \]
- H-D model
  \[ H = a_D \]
- H-D model
  \[ H = a_H \]

Compatibility modeling method 1
- V-DH model
- DG-D model
- H-D model

Compatibility modeling method 2
- V-DH model
- DG-D model
- H-D model

Model evaluation

Figure 1. Flowchart of the data collection and analysis process. Note: diameter at breast height (D), tree height (H), ground diameter (DG), tree volume (V), coefficient of determination (R²), standard error of estimates (SEE), total relative error (TRE), mean systematic error (MSE), mean estimated error (MPE), and mean percentage standard error (MPSE).

3. Results
3.1. Correlation Analysis

The correlation analysis results are shown in Figure 2. DG, D, and H were positively correlated with V. The correlation between H and V was relatively low. The correlation coefficient for Populus × tomentosa Carrière ranged from 0.7 to 0.99. The correlation coefficient for Populus tomentosa Carr ranged from 0.85 to 0.99. The correlation coefficient for Populus × canadensis Moench ranged from 0.63 to 0.98 (Figure 3).

Figure 2. Correlation heat map of the four variables (diameter at breast height [D], tree height [H], ground diameter [DG], and tree volume [V]). Blue indicates positive correlations. The deeper the blue, the larger the positive correlation value.
For the two-dimensional distribution figure of the NMDS, the stress value was 0.0366, demonstrating that the figure could accurately reflect the real differences for each tree (Figure 3). The confidence ellipse was obtained with a confidence level of 95%, which showed that the confidence ellipses of the different species overlapped, demonstrating similarity between the groups. The analysis of similarities based on the Bray–Curtis distance suggested that the difference between the species was significantly greater than the difference within the groups ($R = 0.1827, p = 0.001$).

3.2. Independent Fitting Model Analysis

Model 3 exhibited the lowest AIC value (2018.225), indicating its superior performance in fitting the $H$-$D$ model for *Populus × tomentosa* Carrière. Similarly, Model 1 demonstrated the lowest AIC value (1976.589), signifying its effectiveness in fitting the $H$-$D$ model for *Populus tomentosa* Carr. For *Populus × canadensis* Moench, Model 2 achieved the lowest AIC value (2078.769), suggesting its optimal fit for the $H$-$D$ model (Table 2 and Figure 4). Additionally, the $R^2$ values for Model 5 were 0.9848, 0.9868, and 0.9601 for *Populus × tomentosa* Carrière, *Populus tomentosa* Carr, and *Populus × canadensis* Moench, respectively. These values indicated that the linear model successfully explained over 96% of the variation in the $DG$-$D$ relationship for all three tree species (Table 3 and Figure 5).
Table 2. Regression parameters (a₀, a₁, and a₂) and evaluation indicators (AIC and R²) for the four H-D candidate models (Models 1–4 in Section 2.2.1.) for the three tree species (Populus × tomentosa Carrière, Populus tomentosa Carr, and Populus × canadensis Moench).

<table>
<thead>
<tr>
<th>Tree Species</th>
<th>Model</th>
<th>a₀</th>
<th>a₁</th>
<th>a₂</th>
<th>AIC</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Populus × tomentosa Carrière</strong></td>
<td>Model 1</td>
<td>4.724</td>
<td>0.4830</td>
<td>-</td>
<td>2109.006</td>
<td>0.7285</td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>32.44</td>
<td>-0.06078</td>
<td>1.281</td>
<td>2020.619</td>
<td>0.7829</td>
</tr>
<tr>
<td></td>
<td>Model 3</td>
<td>30.83</td>
<td>4.505</td>
<td>-0.1091</td>
<td>2018.225</td>
<td>0.7842</td>
</tr>
<tr>
<td></td>
<td>Model 4</td>
<td>0.5527</td>
<td>0.02131</td>
<td>-</td>
<td>2042.76</td>
<td>0.7700</td>
</tr>
<tr>
<td><strong>Populus tomentosa Carr</strong></td>
<td>Model 1</td>
<td>4.080</td>
<td>0.5370</td>
<td>-</td>
<td>1976.589</td>
<td>0.7590</td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>192.4</td>
<td>-7.984 × 10⁻⁴</td>
<td>0.5415</td>
<td>1978.77</td>
<td>0.7583</td>
</tr>
<tr>
<td></td>
<td>Model 3</td>
<td>36.58</td>
<td>2.965</td>
<td>-0.06409</td>
<td>1990.518</td>
<td>0.7511</td>
</tr>
<tr>
<td></td>
<td>Model 4</td>
<td>0.5223</td>
<td>0.02185</td>
<td>-</td>
<td>1998.562</td>
<td>0.7454</td>
</tr>
<tr>
<td><strong>Populus × canadensis Moench</strong></td>
<td>Model 1</td>
<td>6.470</td>
<td>0.362</td>
<td>-</td>
<td>2104.811</td>
<td>0.4720</td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>28.32</td>
<td>-0.0601</td>
<td>1.214</td>
<td>2078.769</td>
<td>0.5065</td>
</tr>
<tr>
<td></td>
<td>Model 3</td>
<td>27.63</td>
<td>3.067</td>
<td>-0.08843</td>
<td>2079.17</td>
<td>0.5060</td>
</tr>
<tr>
<td></td>
<td>Model 4</td>
<td>0.5270</td>
<td>0.02696</td>
<td>-</td>
<td>2084.055</td>
<td>0.4987</td>
</tr>
</tbody>
</table>

Note: Akaike information criterion, AIC and coefficient of determination, R².

Figure 4. Regression results for the four H-D candidate models (Models 1–4) of the three tree species (Populus × tomentosa Carrière (a), Populus tomentosa Carr (b), and Populus × canadensis Moench (c)).

Table 3. Regression parameters (b₀ and b₁) and evaluation indicators (AIC and R²) for the DG-D model for the three tree species (Populus × tomentosa Carrière, Populus tomentosa Carr, and Populus × canadensis Moench).

<table>
<thead>
<tr>
<th>Tree Species</th>
<th>Model</th>
<th>b₀</th>
<th>b₁</th>
<th>AIC</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Populus × tomentosa Carrière</strong></td>
<td>Model 5</td>
<td>1.188</td>
<td>1.164</td>
<td>1659.307</td>
<td>0.9848</td>
</tr>
<tr>
<td><strong>Populus tomentosa Carr</strong></td>
<td>Model 5</td>
<td>1.825</td>
<td>1.087</td>
<td>1273.561</td>
<td>0.9868</td>
</tr>
<tr>
<td><strong>Populus × canadensis Moench</strong></td>
<td>Model 5</td>
<td>0.4705</td>
<td>1.146</td>
<td>2023.708</td>
<td>0.9601</td>
</tr>
</tbody>
</table>

Note: Akaike information criterion, AIC and coefficient of determination, R².
3.3. Analysis of the Two Compatibility Models

Two compatibility methods were utilized to fit the following three models: DG-D, H-D, and V-DH. The results revealed minimal disparity in the parameters between the V-DH and DG-D models, whereas significant differences existed in the parameters of the DG-D models (Table 4). Based on the V-DH model, the H-D and DG-D models served as bridges to infer the results of the V-D and V-DG models, thereby ensuring compatibility among the three volume models.

Table 4. Parameter results ($a_0$, $a_1$, $a_2$, $b_0$, $b_1$, $c_0$, $c_1$, and $c_2$) for the three different models (V-DH model, H-D model, and DG-D model) for the three trees species (Populus × tomentosa Carrière, Populus tomentosa Carr, and Populus × canadensis Moench) under the two compatible modeling methods.

<table>
<thead>
<tr>
<th>Tree Species</th>
<th>Compatible Method</th>
<th>H-D Model</th>
<th>DG-D Model</th>
<th>V-DH Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Populus × tomentosa Carrière</td>
<td>(1)</td>
<td>30.92</td>
<td>4.313</td>
<td>-0.1063</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>32.50</td>
<td>2.564</td>
<td>-0.07401</td>
</tr>
<tr>
<td>Populus tomentosa Carr</td>
<td>(1)</td>
<td>4.096</td>
<td>0.5358</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>3.971</td>
<td>0.5454</td>
<td>-</td>
</tr>
<tr>
<td>Populus × canadensis Moench</td>
<td>(1)</td>
<td>28.37</td>
<td>-0.05930</td>
<td>1.193</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>27.54</td>
<td>-0.08152</td>
<td>2.047</td>
</tr>
</tbody>
</table>

Note: The models consisted of the following five types: V-DH, V-D, V-DG, H-D, and DG-D. Before the horizontal line was the dependent variable, and behind the horizontal line was the independent variable.

For Populus × tomentosa Carrière, the MPE values for the five models fell within the range of 0.62%–3.16%, suggesting an overall prediction accuracy exceeding 96.84%. For Populus tomentosa Carr, the MPE values for the five models varied from 0.47% to 2.57%, indicating an overall prediction accuracy of 97.43%. For Populus × canadensis Moench, the MPE values for the five models ranged from 0.65% to 2.40%, indicating an overall prediction accuracy exceeding 97.60% (Table 5).
Table 5. Precision index ($R^2$, SEE, TRE, MSE, MPE, and MPSE) results for the five different models (V-DH model, V-D model, V-DG model, H-D model, and DG-D model) for the three tree species (Populus × tomentosa Carrière, Populus tomentosa Carr, and Populus × canadensis Moench) under the two compatible modeling methods.

<table>
<thead>
<tr>
<th>Tree Species</th>
<th>Model</th>
<th>Compatibility Method</th>
<th>$R^2$</th>
<th>SEE</th>
<th>TRE</th>
<th>MSE</th>
<th>MPE</th>
<th>MPSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>V-DH</td>
<td>(1)</td>
<td>0.9840</td>
<td>0.1439</td>
<td>0.49%</td>
<td>4.75%</td>
<td>1.77%</td>
<td>10.75%</td>
</tr>
<tr>
<td></td>
<td>V-D model</td>
<td>(2)</td>
<td>0.9838</td>
<td>0.1448</td>
<td>−0.03%</td>
<td>1.37%</td>
<td>1.78%</td>
<td>8.97%</td>
</tr>
<tr>
<td>Populus × tomentosa Carrière</td>
<td>H-D model</td>
<td>(1)</td>
<td>0.7851</td>
<td>2.9979</td>
<td>−0.03%</td>
<td>−0.13%</td>
<td>1.36%</td>
<td>11.00%</td>
</tr>
<tr>
<td></td>
<td>V-D model</td>
<td>(2)</td>
<td>0.7566</td>
<td>3.1909</td>
<td>−1.52%</td>
<td>−2.57%</td>
<td>1.45%</td>
<td>12.06%</td>
</tr>
<tr>
<td></td>
<td>DG-D</td>
<td>(1)</td>
<td>0.9695</td>
<td>0.1996</td>
<td>1.08%</td>
<td>4.10%</td>
<td>2.46%</td>
<td>13.15%</td>
</tr>
<tr>
<td></td>
<td>V-DG</td>
<td>(2)</td>
<td>0.9704</td>
<td>0.1966</td>
<td>−0.02%</td>
<td>−1.37%</td>
<td>2.42%</td>
<td>11.89%</td>
</tr>
<tr>
<td></td>
<td>V-DG model</td>
<td>(1)</td>
<td>0.9501</td>
<td>0.2559</td>
<td>0.86%</td>
<td>6.57%</td>
<td>3.15%</td>
<td>21.54%</td>
</tr>
<tr>
<td></td>
<td>V-DG model</td>
<td>(2)</td>
<td>0.9500</td>
<td>0.2563</td>
<td>−0.17%</td>
<td>−3.65%</td>
<td>3.16%</td>
<td>19.27%</td>
</tr>
<tr>
<td></td>
<td>V-DH</td>
<td>(1)</td>
<td>0.9822</td>
<td>0.0599</td>
<td>−0.87%</td>
<td>−4.98%</td>
<td>1.31%</td>
<td>10.32%</td>
</tr>
<tr>
<td>Populus tomentosa Carr</td>
<td>V-D model</td>
<td>(2)</td>
<td>0.9812</td>
<td>0.0617</td>
<td>−0.06%</td>
<td>−0.89%</td>
<td>1.35%</td>
<td>9.04%</td>
</tr>
<tr>
<td></td>
<td>H-D model</td>
<td>(1)</td>
<td>0.7596</td>
<td>2.8486</td>
<td>−0.01%</td>
<td>0.01%</td>
<td>1.36%</td>
<td>10.91%</td>
</tr>
<tr>
<td></td>
<td>V-D model</td>
<td>(2)</td>
<td>0.7595</td>
<td>2.8495</td>
<td>0.15%</td>
<td>0.29%</td>
<td>1.36%</td>
<td>10.94%</td>
</tr>
<tr>
<td></td>
<td>DG-D</td>
<td>(1)</td>
<td>0.9590</td>
<td>0.0912</td>
<td>−0.77%</td>
<td>−4.84%</td>
<td>2.00%</td>
<td>15.63%</td>
</tr>
<tr>
<td></td>
<td>V-DG</td>
<td>(2)</td>
<td>0.9595</td>
<td>0.0908</td>
<td>0.00%</td>
<td>−0.53%</td>
<td>1.99%</td>
<td>14.76%</td>
</tr>
<tr>
<td></td>
<td>DG-D model</td>
<td>(1)</td>
<td>0.9868</td>
<td>1.1830</td>
<td>−0.01%</td>
<td>−0.06%</td>
<td>0.47%</td>
<td>3.71%</td>
</tr>
<tr>
<td></td>
<td>V-DG</td>
<td>(2)</td>
<td>0.9861</td>
<td>1.1217</td>
<td>0.41%</td>
<td>1.00%</td>
<td>0.48%</td>
<td>3.94%</td>
</tr>
<tr>
<td></td>
<td>DG-D model</td>
<td>(1)</td>
<td>0.9328</td>
<td>0.1172</td>
<td>−1.08%</td>
<td>−3.33%</td>
<td>2.57%</td>
<td>18.63%</td>
</tr>
<tr>
<td></td>
<td>V-DG</td>
<td>(2)</td>
<td>0.9326</td>
<td>0.1174</td>
<td>−0.06%</td>
<td>−1.69%</td>
<td>2.57%</td>
<td>18.46%</td>
</tr>
<tr>
<td>Populus × canadensis Moench</td>
<td>V-DH</td>
<td>(1)</td>
<td>0.9650</td>
<td>0.1914</td>
<td>−0.01%</td>
<td>−0.23%</td>
<td>1.26%</td>
<td>8.64%</td>
</tr>
<tr>
<td></td>
<td>V-D model</td>
<td>(2)</td>
<td>0.9650</td>
<td>0.1914</td>
<td>−0.07%</td>
<td>−0.50%</td>
<td>1.26%</td>
<td>8.70%</td>
</tr>
<tr>
<td></td>
<td>H-D model</td>
<td>(1)</td>
<td>0.5909</td>
<td>3.2326</td>
<td>−0.01%</td>
<td>0.00%</td>
<td>1.31%</td>
<td>10.98%</td>
</tr>
<tr>
<td></td>
<td>V-D model</td>
<td>(2)</td>
<td>0.4998</td>
<td>3.2627</td>
<td>0.45%</td>
<td>0.89%</td>
<td>1.33%</td>
<td>11.33%</td>
</tr>
<tr>
<td></td>
<td>DG-D</td>
<td>(1)</td>
<td>0.9360</td>
<td>0.2597</td>
<td>−0.13%</td>
<td>−0.30%</td>
<td>1.71%</td>
<td>12.88%</td>
</tr>
<tr>
<td></td>
<td>V-DG</td>
<td>(2)</td>
<td>0.9339</td>
<td>0.2599</td>
<td>−0.01%</td>
<td>0.12%</td>
<td>1.71%</td>
<td>12.79%</td>
</tr>
<tr>
<td></td>
<td>DG-D model</td>
<td>(1)</td>
<td>0.9602</td>
<td>3.0214</td>
<td>−0.01%</td>
<td>0.05%</td>
<td>0.65%</td>
<td>4.54%</td>
</tr>
<tr>
<td></td>
<td>V-DG</td>
<td>(2)</td>
<td>0.9563</td>
<td>3.1634</td>
<td>0.48%</td>
<td>1.67%</td>
<td>0.68%</td>
<td>5.58%</td>
</tr>
<tr>
<td></td>
<td>DG-D model</td>
<td>(1)</td>
<td>0.8746</td>
<td>3.645</td>
<td>−0.22%</td>
<td>1.66%</td>
<td>2.40%</td>
<td>16.98%</td>
</tr>
<tr>
<td></td>
<td>V-DG</td>
<td>(2)</td>
<td>0.8817</td>
<td>0.3539</td>
<td>0.37%</td>
<td>−1.31%</td>
<td>2.33%</td>
<td>17.24%</td>
</tr>
</tbody>
</table>

Note: The models consisted of the following five types: V-DH, V-D, V-DG, H-D, and DG-D. Before the horizontal line was the dependent variable, and behind the horizontal line was the independent variable. Note: coefficient of determination, $R^2$; standard error of estimates, SEE; total relative error, TRE; mean systematic error, MSE; mean estimated error, MPE; and mean percentage standard error, MPSE.

Among the three volume models, the V-DH model demonstrated the highest performance level, with $R^2$ values ranging from 0.965 to 0.984 and MPE values ranging from 1.26% to 1.78%. The V-D model exhibited $R^2$ values between 0.9359 and 0.9704, with MPE values ranging from 1.71% to 2.46%. Finally, the V-DG model showed $R^2$ values ranging from 0.8746 to 0.9501 and MPE values ranging from 2.33% to 3.16%. For the tree H-D model, the $R^2$ values ranged from 0.4998 to 0.7851, with MPE values of between 1.31% and 1.45%. Conversely, the DG-D model showed $R^2$ values that ranged from 0.9563 to 0.9868 and MPE values that ranged from 0.47% to 0.68% (Table 5).

A comparative analysis revealed that compatibility method 2 outperformed compatibility method 1 among the three models. However, among the remaining 12 models, compatibility method 1 exhibited the best performance (Table 5, Figures 6 and 7). Thus, the nonlinear seemingly uncorrelated regression (NSUR) was more effective.
Figure 6. Fitting results for the measured volume values and predicted volume values for the five models (V-DH model, V-D model, V-DG model, H-D model, and DG-D model) for the three tree species (Populus × tomentosa Carrière, Populus tomentosa Carr, and Populus × canadensis Moench) under the two compatible modeling methods. The x axis is the measured value and the y axis is the predicted value of the model.
Figure 7. Radar map of the accuracy results for the five models (V-DH model, V-D model, V-DG model, H-D model, and DG-D model) for the three tree species (Populus × tomentosa Carrière, Populus tomentosa Carr, and Populus × canadensis Moench) under the two compatible modeling methods. Note: standard error of estimates, SEE; total relative error, TRE; mean systematic error, MSE; mean estimated error, MPE; and mean percentage standard error, MPSE.
4. Discussion

We observed differences in the preferences for the H-D model among the three species, with the logistic function proving more suitable for the H-D fitting of *Populus × tomentosa* Carrière, the Chapman–Richards function for *Populus tomentosa* Carr, and the power function for *Populus × canadensis* Moench. Various forms of the H-D model have existed, including the Chapman–Richards, logistic, Korf, and Gompertz functions [45,46], and there is no consensus on the optimal selection. Numerous studies have highlighted the superior performance of the Chapman–Richards model [47,48]. Given the fixed form of the H-D model, it would have been difficult to depict all possible H-D relationships across the different trees. In many permanent plots, although the D values for all the sampled trees could be commonly measured, only the H value of a subsample was recorded. Consequently, the H-D model could be frequently used to estimate the heights of other sampled trees. Hence, it may be essential to compare and select a suitable model form, develop an appropriate H-D model, and, subsequently, apply it to forest surveys.

Some studies have incorporated factors, such as site conditions and stand structures, into predictors within the H-D model, demonstrating better predictive accuracy than simple functions [49]. Nonlinear mixed-effects models offer greater efficiency and flexibility than least-squares regression by addressing the nested structure of data. For instance, Sharma et al. developed an H-D mixed-effect model utilizing stand dominance height, stand basal area, and stand density as site characteristics, achieving better fitting than the Chapman–Richards model [50]. Similarly, Bronisz et al. employed a mixed-effect model to investigate the impacts of stand density, productivity, and age on the H–D relationship [51]. Hence, there are plans to sample trees with varying site conditions and stand structures over a period of 3–5 years to establish an H-D mixed-effect model.

A robust volume model requires a substantial number of data samples to ensure its accuracy. When the sample size of a dataset is small, the accuracy of the model can be reduced. In this study, the dataset comprised 1200 samples, which was sufficient for constructing a volume model for the three tree species. Typically, volume model studies use sample sizes ranging from tens to hundreds [34,35]. For example, Zhang et al. developed a volume-biomass compatibility model based on data from 80 samples [29]. Subedi et al. derived a volume model from 48 sample trees in the western lowlands of Nepal [3]. Gimenez et al. employed non-destructive climbing techniques to measure 64 trees and established a volume model [19]. Similarly, Boczniewicz et al. developed a volume model by destructively sampling seventy-four trees across eight experimental sites in New Zealand [33].

Previous studies have typically fit univariate and bivariate volume models independently [52,53]. Each model can be estimated in isolation, ignoring the inherent correlations between tree factors. The drawback of this approach lies in the divergence of the predictions yielded by different volume models. For instance, previous studies have separately fitted taper and volume models, leading to significant discrepancies in trunk volumes [54].

An internationally recognized approach utilized by researchers involves integrating volume and taper models. A taper model predicts the diameter at a specified height, facilitating trunk volume determination through established formulas, such as Huber, Smalian, and Newton. When the integral of a taper model yields the same volume as that of a volume model, the compatibility between the two is confirmed [55]. For example, Boczniewicz established taper models for outer bark, inner bark, and heartwood to ensure compatibility with their respective volume models [33]. Similarly, Zhao et al. (2019) developed a fully compatible system for taper and volume models [56]. The advantage of employing a taper model is its ability to provide trunk volume and volume at any position through integration. However, this method requires tree harvesting and utilizes diameter data from various heights [33–35].

The compatible modeling method facilitates the synchronization of multiple response variables and finds widespread applications across various domains. However, its precise interpretation may vary across contexts. For instance, Zhang et al. (2016) developed a com-
compatibility model system for aboveground biomass and volumes of poplar trees in Jiangsu Province, China [29]. Zeng and Tang (2012) constructed a compatibility model system for the aboveground biomass and volumes of Masson pine in China [30]. Zeng et al. (2011) devised a compatible biomass model for Pinus massoniana at different scales, ensuring that the sums of estimates from the small-scale model matched those of the large-scale model [57].

In this study, a compatible model system was established between the V-DH, V-D, and V-DG models to ensure consistency in the predicted values across the three volume models. Unlike other studies, the focus was on utilizing the tree H-D model and DG-D model as bridges, effectively linking the three volume models. Therefore, we emphasized the importance of prioritizing compatibility modeling to address potential incompatibility issues in future research, thereby enhancing the applicability and stability of the model.

The compatible volume model developed in this study demonstrated high prediction accuracy, with MPE values ranging from 1.26% to 3.16% for the V-DH, V-D, and V-DG models, indicating an overall prediction accuracy exceeding 96%. Numerous studies have employed MPE as a key metric to assess model accuracy [30, 57]. The volume model established herein meets the precision standards necessary for compiling volume tables and can reliably predict the poplar volume in Beijing.

We observed that the V-DH model outperformed the V-D model, with the V-DG model ranking last. Numerous studies have shown that incorporating H as a predictor in volume models can significantly enhance model performance [22, 29]. However, contrary to our findings, Zhang et al. discovered that adding H to a biomass model did not yield the anticipated improvements in model performance, particularly for branch and leaf biomass [29]. Moreover, diameters at various heights can exhibit varying capabilities in explaining volume variation, with overall volume variations better explained by D than DG [14, 21, 47]. Hence, for accurate volume estimation, researchers are advised to prioritize obtaining D, followed by H and DG.

We observed that the modeling procedures for the two compatible methods differed significantly but yielded similar parameter fitting results. Notably, the parameters of the DG-D model exhibited considerable variations. A precision comparison revealed that compatible method 1 outperformed compatible method 2. Method 1 utilized NSUR, which considered inter-model correlation and ensured compatibility, thereby facilitating effective parameter estimation. Numerous field studies have corroborated the methodological advantages of using NSUR [10, 37, 58]. Thus, it is recommended that NSUR be prioritized for practical applications.

The division of sample data into modeling and testing sets remains controversial. Some studies have suggested that evaluating model prediction solely based on modeling data accuracy results lacks persuasiveness, and they have advocated for testing the models used [59]. Conversely, Kozak et al. contended that testing samples for suitability might sacrifice modeling information without providing additional insights [60]. Zeng et al. further suggested combining test and modeling samples to prevent sample loss [61]. This study employed all samples for modeling without distinguishing between modeling and test samples. The model evaluation and validation were performed on full datasets. This can help to avoid the loss of sample amount during modeling.

Volume models provide accurate tree volume estimates, which are essential for the sustainable management of forest resources for middle- and long-term periods of time. The application of a volume model can predict tree volume in small-scale areas and forest volume stocks in large-scale areas, which is helpful for making reasonable cutting and tree species conservation plans. It is possible to apply the volume model developed for the three poplar species to other forest species, but it needs to be handled carefully. Firstly, the similarity between the new tree species and poplar in morphology would need to be evaluated. If the similarity is high, the applicability of the model will be strong. Secondly, field data collection and validation must be carried out to evaluate the accuracy of the model and to adjust the model’s parameters. This can save resources and time.
using existing models while contributing to the unified management of forest resources of multiple species, forming a comprehensive database to support ecological research.

5. Conclusions

Volume modeling research is crucial in forest tree conservation because it provides an accurate estimate of tree volume and helps in the development of scientific cutting and conservation plans. These models enable forest managers to determine the optimal times and methods for harvesting, ensuring that tree removal is sustainable. The volume models developed for the three populus species demonstrated both high accuracy and compatibility, offering valuable support for poplar wood resource assessments in Beijing. The introduction of the H-D model notably reduced the fieldwork labor costs. Our research revealed variations in the H-D model preferences across different tree species, with the logistic model proving optimal for the H-D fitting of *Populus × tomentosa* Carrière, with the Chapman–Richard model being optimal for *Populus tomentosa* Carr and the power function being optimal for *Populus × canadensis* Moench. Hence, practical applications should consider selecting H-D models based on tree species. We utilized two compatible modeling methods to construct five models, employing the H-D and DG-D models as bridges to ensure compatibility among the volume models. Among the three volume models, the V-DH model emerged as the superior choice, followed by the V-D and V-DG volume models, respectively. Therefore, the V-DH model can be considered preferable for practical applications. Additionally, our comparison of the compatibility methods revealed that the use of NSUR yielded a higher fitting accuracy.

This study employed error-in-variable simultaneous equations to address compatibility issues between volume models, with the parameters treated as fixed effects. Mixed-effects models incorporating both fixed and random effects have been popular. Future enhancements to volume model accuracy could involve integrating mixed-effect models with simultaneous equations and accounting for inter-plot variation.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/f15061059/s1, Figure S1: Map of 120 sample plots in Beijing, China; Table S1: Location and basic factors of 120 sample plots.

**Author Contributions:** S.W.: conceptualization, formal analysis, writing—original draft preparation; Z.F.: conceptualization, formal analysis, writing—reviewing and editing; Z.W.: conceptualization, formal analysis, writing—reviewing and editing; Z.Y.: investigation, writing—reviewing and editing; J.L.: investigation, writing—reviewing and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the 5·5 Engineering Research & Innovation Team Project of Beijing Forestry University (BLRC2023A03), the Natural Science Foundation of Beijing (8232038, 8234065), and the Key Research and Development Projects of Ningxia Hui Autonomous Region (2023BEG02050).

**Data Availability Statement:** The data used to support the findings of this study are available from the corresponding author upon request.

**Acknowledgments:** The authors would like to thank all the reviewers who participated in the review.

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

**References**


5. Yao, Q.; Wang, J.; Zhang, J.; Xiong, N. Error Analysis of Measuring the Diameter, Tree Height, and Volume of Standing Tree Using Electronic Theodolite. *Sustainability* 2022, 14, 6950. [CrossRef]


17. Sharma, M. Increasing Volumetric Prediction Accuracy—An Essential Prerequisite for End-Product Forecasting in Red Pine. *Forests* 2020, 11, 1050. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.