Research Article

Site Specific Stem Volume Models for Pinus patula and Pinus oocarpa

Herbert Malata,¹ Elisha S. Ngulube,¹ and Edward Missanjo²

¹Department of Forestry, Faculty of Environmental Sciences, Mzuzu University, Private Bag 201, Luwinga, Mzuzu 2, Malawi
²Department of Forestry, Malawi College of Forestry and Wildlife, Private Bag 6, Dedza, Malawi

Correspondence should be addressed to Edward Missanjo; edward.em2@gmail.com

Received 1 July 2017; Revised 12 September 2017; Accepted 1 October 2017; Published 25 October 2017

Academic Editor: Qing-Lai Dang

Copyright © 2017 Herbert Malata et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sustainable management of timber forests requires availability and adequacy of models for accurate estimation of tree volumes. This study was conducted to develop site specific models for estimating individual total tree stem volume of Pinus patula and Pinus oocarpa at Chongoni Timber Plantation in Central Malawi. A total of 32 trees from Pinus patula compartment and 48 trees from Pinus oocarpa compartment were destructively sampled. Various predictors including diameter at breast height (dbh) and height (ht) were run against total stem volume using a nonlinear mixed-effects modelling approach. The results indicate that the developed site specific models showed a significant association between total stem volume and the predictors (dbh and ht). The developed volumetric models accounted for at least 99% of the total variation in the total stem volume data. This suggests that application of the developed site specific models is highly recommended when accurate results are required. The appropriateness of the developed models was also supported by the fact that the total relative errors (TRE) of these models were lower (range: −0.04% to 0.06%) than the TRE of some previously developed models (range: −12.40% to 41.70%) tested on the present data.

1. Introduction

Most of the commercial timber plantations in central Malawi are planted with tropical pine trees. The commonly planted timber species in these plantations include Pinus patula, P. kesiya, P. elliottii, and P. oocarpa [1, 2]. Volume is customarily used in these forest plantations as a standard measure on which timber pricing and yield are based. Estimation of standing timber volume is traditionally done by using allometric models. Common procedures used in allometry are the direct volumetric measurements of the stem. Volume and tariff tables are developed for the estimation of volumetric stem content. The implied conversion to some end-product procedure of volume determination (e.g., Doyle, Scriber, or International Log Rules) is less popular in Malawi’s commercial timber plantations.

Accurate estimation of biomass is a precursor to sustainable forest management [3–5]. Biomass estimates play an important decision support mechanism in commercial exploitation of timber, nontimber forest products, and other ecosystem services [5]. Models are vital for forest planning, yield estimation, projection, and regulation of harvests. To date, generalized allometric models are commonly used to estimate timber volume in Malawi. These models are designed to cover wider spatial dimensions or extended to conditions that are sometimes dissimilar from localities of their origin. However, it is important to note that different timber species exhibit different growth responses to different site conditions and associated management prescriptions [6]. Henceforth, the usefulness of generalized models is not without some limitations for effective decision support in forest management.

Although these models provide useful volume estimation insights into most or wider forest management applications [4], site and species specific volume modelling would precisely guide operational decisions affecting commercial timber plantation management in Malawi [7]. Development of localised models would not only help to maximise timber value but also help to identify key variable constructs of the model. Plausible advocacy for use of localised modelling has been reported in several studies [8–11] particularly where precision and revenues are priorities. This modelling approach
yields accurate and reliable volume estimates for sustainable decision support in the forest management [2, 12]. However, the quality of a model estimate is only as good as the quality of data variables from which the model is developed.

The main objective of this study was to develop site specific models for estimating total tree volume of pines at Chongoni Timber Plantation (CTP) in central Malawi, using mixed-effect modelling approach. Specific objectives of the study were to (i) develop total individual tree stem volume models for *P. patula* and *P. oocarpa* and (ii) compare the predictive performance between site specific and the generalized volume models at CTP.

### 2. Materials and Methods

#### 2.1. Study Site

The study was conducted at Chongoni Timber Plantation (CTP) in Dedza, Central Malawi. CTP covers an estimated area of 5270 ha located within Chongoni Forest Reserve. The reserve lies between the latitudes 14°10’S and 14°21’S and longitudes 34°09’E and 34°17’E. It receives a mean annual rainfall in the range of about 1200 mm to 1800 mm with mean annual temperatures ranging from 7°C to 25°C. The altitude of the reserves varies from 1570 m to 1690 m above the sea level [13]. The forest is dominated by ferruginous and humic ferrallitic soils [14] characterised by high organic matter, being leached, and occurring at high altitudes. The study was carried out in compartments 3A of *P. oocarpa* (16.4 ha, slope: 7%) and 19B of *P. patula* (10.4 ha, slope: 5%).

#### 2.2. History and Description of Forest Stands

Both stands were established from nursery seedlings after harvesting of the first rotation crop. *Pinus oocarpa* and *P. patula* seedlings were planted at the age of six months after prickling out (propagating plants sown on seed bed) in 1991. The mean height and root collar diameter of the seedlings were 26.3 cm and 5.4 mm, respectively. These seedlings were planted in 30 × 30 cm hoe dug pits at a spacing of 2.75 m × 2.75 m (approx. 1320 stems ha⁻¹). Thinning was carried out at the ages of 9, 15, and 19 years by removing 35%, 28%, and 39% of the trees in a stand, respectively. Pruning was done at the ages of 4, 9, and 15 years to half of the stem heights in each case. The stands were characterised by the following descriptive variables (Table 1).

#### 2.3. Field Sampling and Data Collection

A systematic sampling design with a randomly selected starting point was used. Thereafter, plots were spaced at uniform intervals of 200 m along strips throughout the compartments. Strips were spaced at 100 m intervals. This sampling design was chosen to uniformly distribute the sample units over the entire population [15].

The minimum required number of sample plots was determined using the following formula [16]:

$$ n = \frac{ap}{s}, \quad (1) $$

where *n* is the required number of sample plots, *a* is the area of the compartment in ha, *p* is the sampling intensity, and *s* is the plot size in ha. In this study a sampling intensity of 3% (0.03) and sample plot size of 0.04 ha were used. A total of 20 sample plots were established. Out of these, 12 sample plots were determined in *P. oocarpa* compartment and 8 sample plots in *P. patula* compartment. After determining the number of sample plots, square plots of 0.04 ha were systematically laid down at every 200 m throughout each compartment. In each sample plot, all standing trees were assessed for diameter at breast height (dbh) (1.3 m from the ground) and their corresponding heights using a caliper and a Suunto clinometer, respectively. Every 5th tree was selected for destructive sampling in each sample plot. A total of 4 trees in each plot were felled. These trees were cut at 30 cm above ground on the buttress using a chain saw. A total of 48 trees from *P. oocarpa* compartment and 32 trees from *P. patula* compartment were destructively sampled for log volume determination. The total height of trees was confirmed on felled trees by using a 50 m linear measuring tape. The felled trees were marked into logs approximately 3 m long each (where applicable) up to the minimum top (5 cm) diameter. Diameters at lower, middle, and top end of each log were measured using a caliper for the purpose of estimating log volume. The volume of each log was estimated using newton’s formula. This method was chosen in order to maximise precision and accuracy of the volume due to taper [17]. The equation used was

$$ V = \pi \left( \frac{d_l^2 + 4d_m^2 + d_u^2}{240000} \right) L, \quad (2) $$

where *V* is the volume in m³ of the log, *d_l* is the diameter at lower end of the log in cm, *d_m* is the diameter at the middle of the log in cm, *d_u* is the diameter at the upper end of the log in cm, and *L* is the length in m of each log. Total stem
volume was estimated as the sum of volumes of all log sections as follows:

\[ V = \sum_{i=1}^{t} v_i \]  

(3)

where \( i \) is the \( i \)th log section, \( v_i \) is the log volume in \( m^3 \), and \( t \) is the number of logs in that particular stem.

2.4. Statistical Analysis. Prior to analysis, data were tested for normality and homogeneity using Kolmogorov-Smirnov D and normal probability graphical tests. Thus, the dependent variable (volume) was plotted against each of the explanatory variables to examine the range and shape of the functional relationship and to assess the heterogeneity of the variance. The following general nonlinear model forms, which described individual tree volume as a function of the diameter at breast height (dbh) and total height (ht), were fitted:

\[ V = b_0 \times (dbh^2 \times ht)^{b_1} \]  

(4)

\[ V = b_0 \times ddbh^{b_1} \times ht^{b_2} \]  

(5)

\[ V = b_0 \times (dbh^2)^{b_1} \times ht^{b_3}, \]  

(6)

where \( V \) is total stem tree volume over bark (\( m^3 \)), \( dbh \) is the diameter at breast height (cm), \( ht \) is total tree height (m), and \( b_0, b_1, \) and \( b_2 \) are parameter estimates. The general nonlinear forms used in this study have adequate mathematical properties and have performed satisfactorily in previous studies [7, 18–20].

Nonlinear regression procedure (NLP) in SAS software [21] was used to fit the models parameters. The importance of using NLP is well explained by Mugasha et al. [7]. However, (4) to (6) were log transformed in order to address heteroscedasticity [22, 23]. Selection of the best models was based on high adjusted coefficient of determination \( R^2_{adj} \), low root mean square error (RMSE), and finally low Akaike Information Criterion (AIC). The expressions for these statistics are summarized as follows:

\[ R^2_{adj} = 1 - \frac{(n-1) \sum (y_i - \bar{y})^2}{(n-p) \sum (y_i - \bar{y})^2} \]  

(7)

\[ RMSE = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-p}} \]

\[ AIC = n \log \sigma^2 + 2k - \min (n \log \sigma^2 + 2k), \]

where \( y_i \) are observed values, \( \bar{y}_i \) are estimated values, \( \bar{y} \) is the mean value of the samples, and \( n \) is the number of samples, \( p \) is the number of parameters to be estimated, \( k = p + 1 \), and \( \sigma^2 \) is found by the following formula:

\[ \sigma^2 = \frac{\sum (y_i - \bar{y})^2}{n}. \]  

(8)

Deviation in the predicted outputs was minimised by multiplying the exponential of the outputs by a correction factor (CF) [22, 23]. CF was computed as follows:

\[ CF = \exp\left(\frac{MSE}{2}\right), \]  

(9)

where \( MSE \) is the mean square error of the regression.

Nonlinear mixed-effects method was then applied to the “best” models which were fitted using the NLMIXED procedure in SAS [21]. Parameters of the “best” model were estimated with and without the presence of random parameter(s). Nonlinear mixed-effects modelling approach is well explained by other researchers [13, 18, 20, 24].

The selected models were evaluated by computing the total relative error (TRE) and mean predictor error as follows [25]:

\[ TRE = \frac{\sum y_i - \sum \bar{y}_i}{\sum \bar{y}_i} * 100 \]  

(10)

\[ MPE = \frac{t_{\alpha} \sqrt{\sum (y_i - \bar{y})^2 / (n-p)}}{\sqrt{\bar{y}^2 / n}} * 100, \]

where \( t_{\alpha} \) is the \( t \) value at confidence level \( \alpha \) with \( n - p \) degrees of freedom. If the value of TRE was less than that of MPE, then the developed volume model was considered acceptable. Furthermore, a generic volume model developed by Ingram and Chipompha (see (11)) [26] for conifer species (cypress and pines) in Malawi was also evaluated on the data. In addition, Student’s \( t \)-test was run to determine statistical differences on total relative errors and mean tree volume between the developed volume models and the generic volume model.

\[ \ln (V) = -10.228 + \ln (dbh^3ht), \]  

(11)

where \( \ln \) is the natural logarithm, \( V \) is the volume in \( m^3 \), \( dbh \) is diameter at breast height in cm, and \( ht \) is the total height in m.

3. Results and Discussion

3.1. Model Selection. Parameter estimates, fit statistics, and developed total stem volume models for \( P. patula \) and \( P. oocarpa \) at CTP are presented in Table 2. The models showed a statistically significant association between total stem volume and predictors (diameter at breast height and height). Nonlinear model (see (4)) performed better for both \( P. patula \) and \( P. oocarpa \).

The model for \( P. patula \) and \( P. oocarpa \) accounted for at least 95% and 86% of the total variation in the total stem volume models, respectively. The variation for the models developed is in line with the fact that essentially volume of a tree is the function of diameter and height [27, 28]. Unfortunately, height measurements for standing trees are generally difficult, time consuming, and costly as compared to dbh [13]. It must be appreciated, however, that precision goes with a cost. For models that only use dbh, there must be well known precise conversion factors that translate the
specific relationship between dbh and height of the particular timber species. The lower values of CF (approximately 1) in the models indicate a significant relationship between variables; that is, lack of fit of the models is negligible [29]. The best selected models were then subjected to mixed-effect modelling approach.

3.2. Mixed-Effects Modelling Approach. The mixed-effects model for (4) after transformation is

$$\ln V = \ln (b_0 + \theta_j) + \left( b_1 + \theta_j \right) \ln \left( \text{dbh}^2 \times \text{ht} \right) + \epsilon', \quad (12)$$

where \(b_0\) and \(b_1\) are fixed-effects regressions parameters; \(\theta_j\) and \(\theta_j\) are random parameters of the \(i\)th and \(j\)th variables, respectively; \(\epsilon' = \ln \epsilon\) and \(\epsilon\) is random error; \(V\), dbh, and ht are stem volume, diameter at breast height, and total height, respectively. The mixed-effects volume models for both \(P. patula\) and \(P. oocarpa\) were fitted using NLMIXED procedure in SAS [21], but the models failed to converge. Then, the number of random parameters was systematically reduced to achieve convergence. Convergence of the models for both \(P. patula\) and \(P. oocarpa\) was achieved by summing up parameter \(b_0\) to a random parameter \(\theta_1\). Therefore, the final mixed-effects model for both \(P. patula\) and \(P. oocarpa\) can be written as

$$\ln V = a + b_1 \ln \left( \text{dbh}^2 \times \text{ht} \right) + \epsilon', \quad (13)$$

where \(a = \ln (b_0 + \theta_1)\). Thus, the predicted volume has to be obtained by back-transforming equation (13) into original values.

The parameter estimates and fit statistics obtained by nonlinear mixed-effects modelling approach are presented in Table 3. The fit statistics \(R^2_{adj}\) increased by 2.31% while the RMSE and AIC decreased significantly by 10.5% and 18.1%, respectively, by performing a mixed-effects modelling approach without random parameters for \(P. patula\) model. Likewise, \(R^2_{adj}\) increased by 9.29% while the RMSE and AIC decreased significantly by 23.9% and 48.3%, respectively, by performing a mixed-effects modelling approach without random parameters for \(P. oocarpa\) model. This indicates that models developed by mixed-effects procedure without random parameters are superior to those developed by non-mixed-effects approach. Similar results were also reported by other researchers [20, 24].

Further analysis indicates that fit statistics (\(R^2_{adj}\), RMSE, and AIC values) of the mixed-effects model with random parameters were better than those of its fixed-effects counterpart. For instance, RMSE and AIC decreased significantly by 13.2% and 26.8%, respectively, while \(R^2_{adj}\) increased by 2.57% by including the random parameters for \(P. patula\) volume model. Similarly, RMSE and AIC decreased significantly by 31.4% and 42.2%, respectively, while \(R^2_{adj}\) increased by 4.07% by including the random parameters for \(P. oocarpa\) volume model. The present results are in agreement with those in literature [13, 20, 24, 30–33]. Missanjo and Mwale [13] and Guangyi et al. [20] reported that inclusion of random parameters in mixed-effects modelling approach increases predictive ability of the model.

3.3. Comparison between Developed and Generalized Tree Volume Models. Results on the comparison between developed volume models and the traditional (generalized) model are presented in Table 4. The generalized model (see (11)) [26] is underestimated by 41.7% and overestimated by 12.4% of mean tree volume in \(P. patula\) and \(P. oocarpa\), respectively, when applied at CTP. The TRE for the developed models for \(P. patula\) and \(P. oocarpa\) were not significantly different from zero and were less than their corresponding MPE. This indicates that the developed models are acceptable in the study area.

It is clear from the results that the traditional model was associated with larger TRE than the developed models. The present findings are in agreement with those in literature [27, 28]. For example, generalized model developed by Chave et al. [27] underestimated and overestimated volume in Rondo and Dindili forests in Tanzania, respectively, compared with the site specific models developed by Mugasha et al. [7]. Similarly, the generalized model developed by Hawkins [34] underestimated volume for \(Eucalyptus camaldulensis\) in Nepal than the site specific model developed by Mandal et al. [28].

### Table 2: Selected total stem volume models, parameter estimates, standard errors (SE), and statistic measures of fit for \(P. patula\) and \(P. oocarpa\).

<table>
<thead>
<tr>
<th>Species</th>
<th>(P. patula)</th>
<th>(P. oocarpa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated parameters</td>
<td>(b_0) = (-0.1466)</td>
<td>(b_0) = (-0.0455)</td>
</tr>
<tr>
<td>(\sigma)</td>
<td>0.0252</td>
<td>0.0166</td>
</tr>
<tr>
<td>(R^2_{adj})</td>
<td>0.951</td>
<td>0.869</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.076</td>
<td>0.046</td>
</tr>
<tr>
<td>AIC</td>
<td>92.07</td>
<td>45.83</td>
</tr>
<tr>
<td>CF</td>
<td>1.003</td>
<td>1.001</td>
</tr>
<tr>
<td>(P) values</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>(\ln(V))</td>
<td>(\ln(V) = -0.1437 + 7.1 \times 10^{-5} \ln(\text{dbh}^2 \times \text{ht}))</td>
<td>(\ln(V) = -0.0444 + 3.7 \times 10^{-5} \ln(\text{dbh}^2 \times \text{ht}))</td>
</tr>
</tbody>
</table>

Note: \(b_0\) is the intercept; \(b_1\) is the gradient of the predictor variable; \(R^2_{adj}\) is the adjusted coefficient of determination; CF(\(\sigma\)) is the correction factor; dbh is the diameter at breast height; and ht is the total tree height; ln is the natural logarithm; \(V\) is volume of a tree.
Table 3: Parameter estimates, standard errors, and values of the statistics for selected model fitted by use of mixed-effects modelling approach.

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimated parameters</th>
<th>SE</th>
<th>P value</th>
<th>$R^2_{adj}$</th>
<th>RMSE</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\ln(V)$ = $-12.2739 + 1.2631 \ln (\text{dbh}^2 \text{ht})$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$b_0$ $-12.420$</td>
<td>$1.7420$</td>
<td>$&lt;0.001$</td>
<td>0.973</td>
<td>0.068</td>
<td>75.40</td>
</tr>
<tr>
<td></td>
<td>$b_1$ $1.2792$</td>
<td>$0.0448$</td>
<td>$&lt;0.001$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\sigma^2_e$ $0.2718$</td>
<td>$0.0103$</td>
<td>$&lt;0.001$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$b_0$ $-12.987$</td>
<td>$0.2014$</td>
<td>$&lt;0.001$</td>
<td>0.998</td>
<td>0.059</td>
<td>55.21</td>
</tr>
<tr>
<td></td>
<td>$b_1$ $1.2631$</td>
<td>$0.0438$</td>
<td>$&lt;0.001$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\sigma^2_e$ $0.2557$</td>
<td>$0.0369$</td>
<td>$&lt;0.001$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Method 1 = selected model fitted without any random parameters. Method 2 = selected model fitted with a random parameter. $\sigma^2_e$ and $\sigma^2_\theta$ are variances for random-effect and residual error, respectively.

Table 4: Performance of developed models versus the generic model on the modelling data.

<table>
<thead>
<tr>
<th>Species</th>
<th>Model</th>
<th>Mean tree volume ($m^3$)</th>
<th>MPE (%)</th>
<th>TRE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. patula</td>
<td>$\ln(V) = -12.2739 + 1.2631 \ln (\text{dbh}^2 \text{ht})$</td>
<td>$0.376 \pm 0.058$</td>
<td>1.02</td>
<td>$-0.04^{**}$</td>
</tr>
<tr>
<td></td>
<td>$\ln(V) = -10.228 + \ln (\text{dbh}^2 \text{ht})$</td>
<td>$0.265 \pm 0.030$</td>
<td>31.93</td>
<td>41.70$^*$</td>
</tr>
<tr>
<td>P. oocarpa</td>
<td>$\ln(V) = -12.6562 + 1.2579 \ln (\text{dbh}^2 \text{ht})$</td>
<td>$0.224 \pm 0.018$</td>
<td>1.54</td>
<td>$0.06^{**}$</td>
</tr>
<tr>
<td></td>
<td>$\ln(V) = -10.228 + \ln (\text{dbh}^2 \text{ht})$</td>
<td>$0.252 \pm 0.016$</td>
<td>21.81</td>
<td>$-12.40^{**}$</td>
</tr>
</tbody>
</table>

Note. Mean volumes followed by the same letter within species significantly differ ($P < 0.05$); $^*$ significantly different from zero ($P < 0.05$); $ns$ = not significantly different from zero ($P > 0.05$); $^{**}$ whole dataset was used in modelling.

The magnitude of variations in errors may be due to differences in site characteristics, silvicultural practices, and the type of data used in the development of the allometry [29].

The significant deviation of the generalized model developed by Ingram and Chipompha [26] in estimating volume for both P. patula and P. oocarpa in the current area may be a clear indication that the model was used in conditions outside its generation bounds. Therefore, care must be taken in choosing volume models for application in different types of tree species and geographical areas [35] as this has the potential to affect the real value of timber resources.

4. Conclusion

The present study has developed total stem volume models for P. patula and P. oocarpa. The results of the statistics of fit were generally good, enabling one to use the models with confidence for estimation of total stem volume for P. patula and P. oocarpa in Chongoni Timber Plantation in Malawi. Due to deviation associated with volume estimates developed from the generalized volume model, application of the developed site specific models is highly recommended.

Conflicts of Interest

The authors declare that there are no conflicts of interest in any form regarding the publication of this article.

Acknowledgments

The authors are grateful to forestry staff at Dedza District Forestry Office for allowing the study to be conducted at Chongoni Timber Plantation and also for the assistance provided during data collection.

References


