Ecological Questions 34(2023)2 http://dx.doi.org/10.12775/EQ.2023.017

Allometric relationships among tree-size variables under tropical forest stages in Gia Lai, Vietnam

Bui Manh Hung[1,*], Phung Van Khoa[2], Nguyen Thi Bich Phuong[3], Nguyen Van Quy[4], Bernard Dell[5]

[1]Department of Forest Inventory and Planning, Faculty of Forestry, Vietnam National University of Forestry, Xuan Mai, Chuong My, Hanoi, Vietnam
[2]Vietnam National University of Forestry, Xuan Mai, Chuong My, Hanoi, Vietnam
[3]Department of Soil Science, Faculty of Silviculture, Vietnam National University of Forestry, Xuan Mai, Chuong My, Hanoi, Vietnam
[4]Department of Forestry, Vietnam Forestry University-Second Campus, Dong Nai, Vietnam
[5]Murdoch University, 90 South St, Murdoch WA 6150, Australia

[*]Corresponding author e-mail: hungbm@vnuf.edu.vn

Received: 10 October 2022 / Accepted: 12 December 2022

Abstract. Allometric models play an undeniable role for estimating hard-to-measure quantities such as volume, biomass and carbon stock in forests. However, so far there has been limited model development for native forests in Vietnam. Therefore, this study was conducted to build and analyze the effectiveness of nonlinear and mixed models for secondary and old-growth forests in Gia Lai, Vietnam. The study measured diameter at breast height, total height, commercial height and crown width of forest trees in 20 plots (10 plots for each forest stage). The results showed that diameter had the strongest relationship with height. In the secondary forest, the Power, Korf and Ratskowky models were the best for pairs of variables, while Prodan, Weibull and Power models were the best fit in the old-growth forest. The nonlinear mixed-effect models were better than classic nonlinear models in both forest stages. Fixed and mixed models developed in this study are very valuable for estimating difficult-to-measure quantities and contribute to effective forest management in the study region.

Key words: Mixed-effect model; nonlinear model; old-growth; secondary forest.

1. Introduction

Allometric models are essential for estimating difficult-to-measure quantities against easily-measurable quantities of forest trees. In forest mensuration science, diameter at breast height is very easy to obtain, while total height, commercial height and crown diameter of forest trees are much harder to quantify, especially in natural forests (Chenge, 2021; P. X. Hoan & Ngu, 2003). This is because it is often difficult to accurately observe the apex of each forest tree to measure height (Chenge, 2021), and it is challenging to project the canopy perpendicularly to the ground to measure crown width (Brack, 1999; Hung, 2016). Therefore, allometric models are crucial for predicting these quantities, and for estimating volume, biomass and carbon stock accumulating in forests. The models provide the basis for auditing forest values (Miah et al., 2020; Sullivan et al., 2018).

Many studies have developed nonlinear allometric models for different types of forests and tree species. These studies have confirmed that there is a correlation with different levels of strength between tree-size variables such as diameter, height, commercial height and crown width (Abedi & Abedi, 2020; D. Avsar & Ayyıldız, 2005; M. D. Avsar, 2004; Kazmierczak et al., 2011; Maua et al., 2020; Scaranello et al., 2012). However, these models should only be applied to areas where trees are measured and they should not be implemented to other regions, because then there will be large errors (Chenge, 2021; Zhang, 1997). Nonlinear biometric patterns are often very different between regions, forest types, tree species, densities and nutrient conditions (Liu et al., 2017; Monteiro et al., 2016). In recent years, mixed nonlinear models have been proving to be better than classical nonlinear models to simulate relationships between variables. These models can examine the relationship between fixed variables and the effect of random factors on measured data (Chenge, 2021; Ogana et al., 2020; Özçelik et al., 2018; Petrás et al., 2014). However, current studies have applied mixed-effect models to the relationship between diameter and height; and relationships between diameter and commercial height, or diameter and crown width have not been analyzed.

In Vietnam, a number of studies have analyzed the relationship between diameter and height using classical nonlinear models (Linear, Logarithmic, Power, Inverse, Quadratic, Cubic, S, Compound, Growth, Exponential) for forests in Binh Dinh, Lao Cai, Bac Kan, Vinh Phuc, and Hanoi provinces (Hien et al., 2019; Huy & Hung, 2018; Phong, 2019; Trieu, 2017; Tuan, 2017; Van & Hien, 2018). The results varied widely between study areas. Diameter vs height of forests in Binh Dinh was best described by the Quadratic function (Van & Hien, 2018). Forests in Lao Cai, the Cubic and Power modes were found to be the best to simulate the relationship between diameter and height, while Cubic and Quadratic models were the best fit in Bac Kan, Vinh Phuc and Hanoi (Phong, 2019; Trieu, 2017; Tuan, 2017). This can be explained by dissimilarities in species, tree age, forest type and environmental conditions. Little use has been made of complex models such as Weibull, Prodan, Chapman-Richards, Korf, and Ratskowky in Vietnam yet in order to analyze the correlation between tree-size quantities. Also, nonlinear mixed models have not been developed for forest succession stages in Vietnam. Therefore, this study was carried out with the aim of: i) understanding tree-size variable characteristics between forest stages; ii) developing nonlinear models for secondary and old-growth forests in Gia Lai; and ii) developing and comparing the effectiveness of nonlinear mixed-effect models for these forest stages to assist in estimating difficult-to-measure quantities in the study area. Having working models will reduce effort and costs of field measurements in the future, and contribute to more effective management of forest resources in the national park.

2. Materials and methods

2.1 Study site

The research was conducted in Kon Ka Kinh National Park, in Gia Lai province, Central Highlands, Vietnam. The geographical coordinates of the park run from 14°09'-14°30' N and 108°16'-108°28' E (Fig. 1). The study area's typical altitude ranges from 1200 to 1500 meters. The study area has two distinct seasons: rainy season (from May to November) and dry season. The average yearly temperature in the research area is between 21 and 25 degrees Celsius. The annual rainfall average is between 2000 and 2500 mm. The geological parent material of Kon Ka Kinh National Park comprises four rock groups: acid-volcanic granites; neutral-alkaline basalts; shales, consisting clay and mica schists; and accreting alluvium along stream lines (N. V. Hoan, 2013; Hung, 2016).

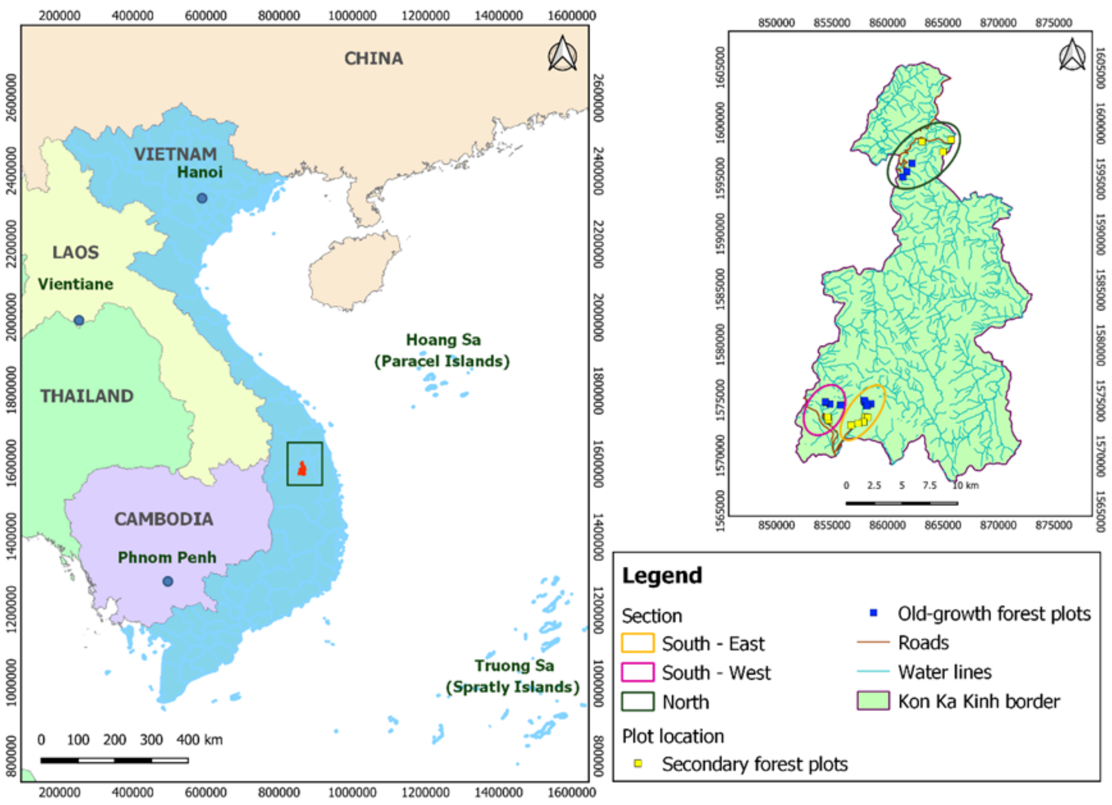


Figure 1. Study area. Maps of Vietnam (left) and Kon Ka Kinh national park (right).

2.2 Sampling design and data collection

Secondary and old-growth lowland tropical rainforests were selected to research and analyze the relationship between tree-size variables. Secondary forests developed after selective cutting which ended in 1986 (N. V. Hoan, 2013). This forest stage has a complex and unevenly-aged species composition, most trees <20 cm in diameter, and very few mother trees. The old-growth forests are primeval forests that have not been exploited. The old-growth stage has a stable structure, many storeys and many different diameter size classes (Trieu, 2017). Using a forest status map created by forest rangers, three areas of the national park (Fig. 1) were divided into secondary and old-growth forest stages and ten forest plots were randomly established in each. The plots were located in three areas of the park: South-East (4 plots in each), South-West (3 plots in each), and North-West (3 plots in each) (Fig. 1). The shortest distance between plots was greater than 100 m. Each plot was 2000 m^2 in size with sides of 50 and 40 m.

Trees with boles greater than 6 cm at breast height (1.3 m) were measured for diameter at breast height (DBH), total height (H), commercial height (CH), and crown width (CW) (P. X. Hoan & Ngu, 2003) (Fig. 2). This study built multi-species models, so species names were not considered. Calipers were used to measure the diameter over bark at 1.3 m above the ground. The total height was the distance from the ground to the uppermost part of the crown. Commercial height was the distance from the ground to the lowest living branch joining the main crown. A Blume Leiss was used to determine total height and commercial height. Crown width was determined by projecting the crown's edges to the ground and then measuring the distance from edge to edge along the widest and shortest axes (Brack, 1999; Philip, 1998).

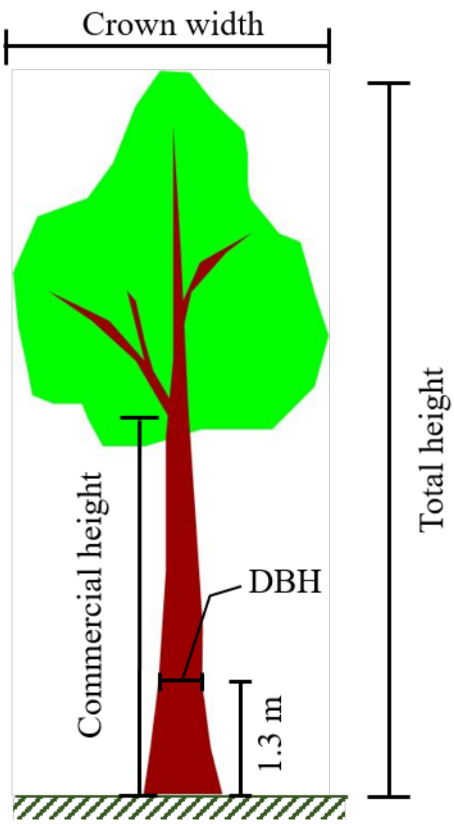


Figure 2. Measured variables of a tree.

The plot data were divided into two parts: 70% of the data (called fitting data) and 30% of the data (called validation data) (Anacioco et al., 2018; Thanh et al., 2019). The fitting data (7 plots in each of the two forest stages) were used to build nonlinear models, and the validation data (3 plots in each of the two forest stages) were used to validate the models. The detailed characteristics of the tree-size variables in each dataset are provided in the results.

2.3 Statistical analysis

Variable characteristics analysis. Principal component analysis was used to visualize differences in the data between the two forest stages, based on diameter at breast height, total height, commercial height and crown width variables. Permutational multivariate analysis of variance (Permanova) was used to compare tree-size variables between the two forest stages. Permanova uses distance matrices for partitioning sum of square. This study used 999 permutations for the Permanova test (Anderson, 2011; Hung, 2016). Pearson correlation coefficient was calculated for each pair of variables in each forest stage to provide a rough estimate of each relationship between variables. Then, descriptive statistics such as sample size (Count), mean, minimum value (Min), maximum value (Max) and standard deviation (SD) were calculated to show the characteristics of datasets (Hung, 2022; Zar, 2010).

Selected nonlinear growth models. The fitting data of each pair of variables were used to build ten selected models for each forest stage (Table 1).

Table 1. Ten selected growth models for each pair of variables.

Growth model	Equation	Reference
Power	$Y = aX^b$	(Fang & Bailey, 1998)
Naslund	$Y = 1.3 + (X / [a + bX])^c$	(Chenge, 2021)
Chapman-Richards	$Y = 1.3 + a(1 - \exp[-bX])^c$	(Scaranello et al., 2012)
Exponential	$Y = 1.3 + \exp(a + b / [X + c])$	(Anacioco et al., 2018)
Weibull	$Y = 1.3 + a(1 - \exp[-bX^c])$	(Sharma et al., 2016)
Logistic	$Y = 1.3 + a / (1 + b \exp[-cX])$	(Thanh et al., 2019)
Gompertz	$Y = 1.3 + a(\exp[-b \exp[-cX]])$	(Chenge, 2021)
Prodan	$Y = 1.3 + X^2 / (a+bX+cX^2)$	(Fang & Bailey, 1998)
Ratskowky	$Y = 1.3 + a(\exp[-b / (c+X)])$	(Chenge, 2021)
Korf	$Y = 1.3 + a(\exp[-bX^c])$	(Anacioco et al., 2018)

Note: In the above models, X was the easiest variable to measure (diameter at breast height). Y were variables more difficult to measure such as: total height, commercial height and crown width.

Nonlinear model selection. To compare and select the best model for each forest stage, this study used four criteria, with the following priority order: Akaike's information criterion (AIC), coefficient of determination (R^2), root of mean square error (RMSE) and model parameter significance. Of these, AIC was used to select the best model from selected models (the smaller, the better) (Chai et al., 2018; Hung, 2016); R^2 was used to understand the relative correlation and goodness of fit between the fitting data and expected values (the greater the value, the stronger the correlation among variables) (Li et al., 2015; Zar, 2010); RMSE was used to indicate the accuracy of the estimates (the model with the smaller RMSE was better) (Anacioco et al., 2018; Chai et al., 2018); and the significance of the parameters showed that the model was still significant for bigger sample sizes or other areas with similar conditions (where the p-value for t-test of the parameter was <0.05, the parameter was significant). The model with all significant parameters was preferred (Hung, 2022; Zar, 2010). The calculation formulas of these criteria are presented in Table 2.

Table 2. The evaluation statistics used to compare the performance of models.

Evaluation statistic	Formula
Akaike's information criterion (AIC)	$AIC = n\ln(RMSE) + 2k$
Coefficient of determination (R^2)	$R^2 = 1 - \left[\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}\right]$
Root of mean square error (RMSE)	$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}}$

Nonlinear mixed-effects model. The best nonlinear model for each pair of variables in each stage was used to build a nonlinear mixed-effects model. The mixed model was an extension of the nonlinear regression model. Mixed models analyze the relationship between a response variable and independent variables (Faraway, 2016). The response variable is continuous, such as total height, commercial height and crown width. The independent variables can be both continuous and categorical (West et al., 2015). There are two main parts in the mixed-effect model: the fixed factor and the random factor. In this study, the fixed variable was diameter at breast height and the random factor was section and plot. Random factors in our study were nested (Faraway, 2016). Therefore, the general equation of the mixed model for each pair of variables in each forest stage was as follows:

$$Y_{i,j} = f(X_{i,j}) + (b)Section/Plot + \varepsilon_{i,j} \quad (1)$$

where: $f(X_{i,j})$ is the nonlinear form.

b is parameters for the random effect.

$\varepsilon_{i,j}$ is error.

$X_{i,j}$ is diameter at breast height variable

$Y_{i,j}$ is total height, commercial height and crown width, respectively.

The maximum likelihood method was applied to achieve mixed-effects models parameters. AIC was also used to compare mixed models and nonlinear models.

Variance homogeneity of residuals was examined using a diagnostic plot of raw residuals against fitted values. If there was no heteroscedasticity, the plot displayed a random pattern and constant variability along the vertical axis (Gałecki & Burzykowski, 2013). In other words, the plot showed the same residual spread per stratum for some of the variables (Zuur et al., 2009). To deal with heteroscedasticity, this study used varExp function in the NLME package in R to weight variances differently between groups and make variances more homogeneous. After updating the mixed model with varExp, the model was compared with the original to see the improvements (Pinheiro et al., 2016).

QQplot was used to check and test the normality of residuals; the quantiles of ordered residuals were plotted against the corresponding values for the standard normal distribution. If all the scatter points are close to the reference line, the dataset follows the normal distribution (Gałecki & Burzykowski, 2013).

All computations were performed using R software version 4.0.2 (package: factoextra, GGally, ggplot2, nlme, vegan and nlme) (Team, 2021).

3. Results

3.1 Tree-size variable characteristics between forest stages

Data classification. The tree-size variable data of all plots in the secondary and old-growth forests were used for comparison by principal component analysis. The results showed that the data were statistically different between the two forest stages (Permanova, F=37,732, p-value=0.001). The data were separated into two distinct clusters (Fig. 3). The data plots in the same forest stage were relatively homogeneous.

Fitting and validation data description. In each forest stage, 7 plots were used to fit models and 3 plots were used to validate models. Data of diameter at breast height, total height, commercial height and crown width for the fitting and validation datasets were used to compute descriptive statistics (Table 3).

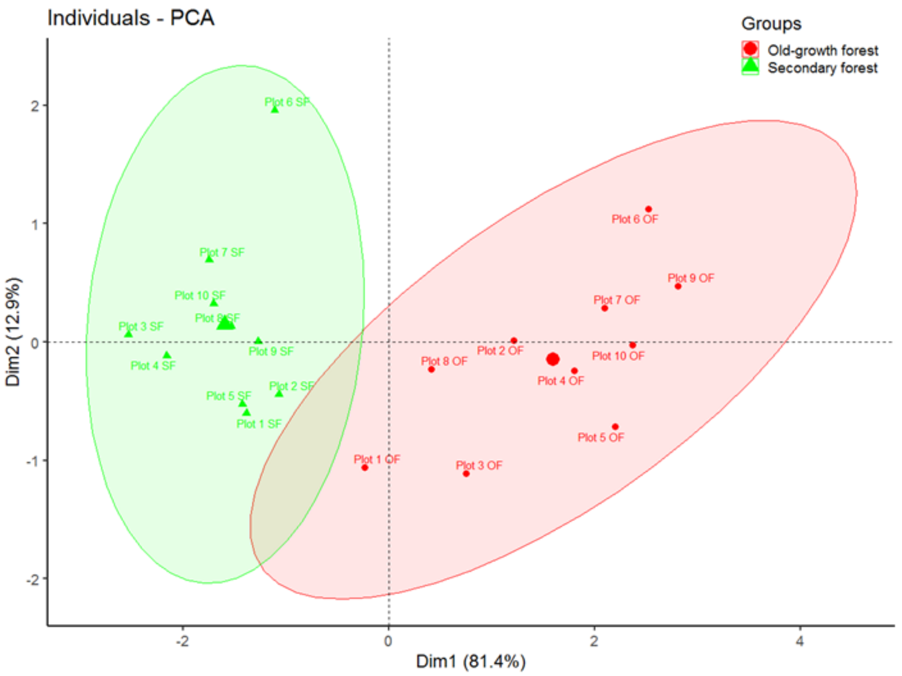


Figure 3. Principal component analysis for tree-size variables.

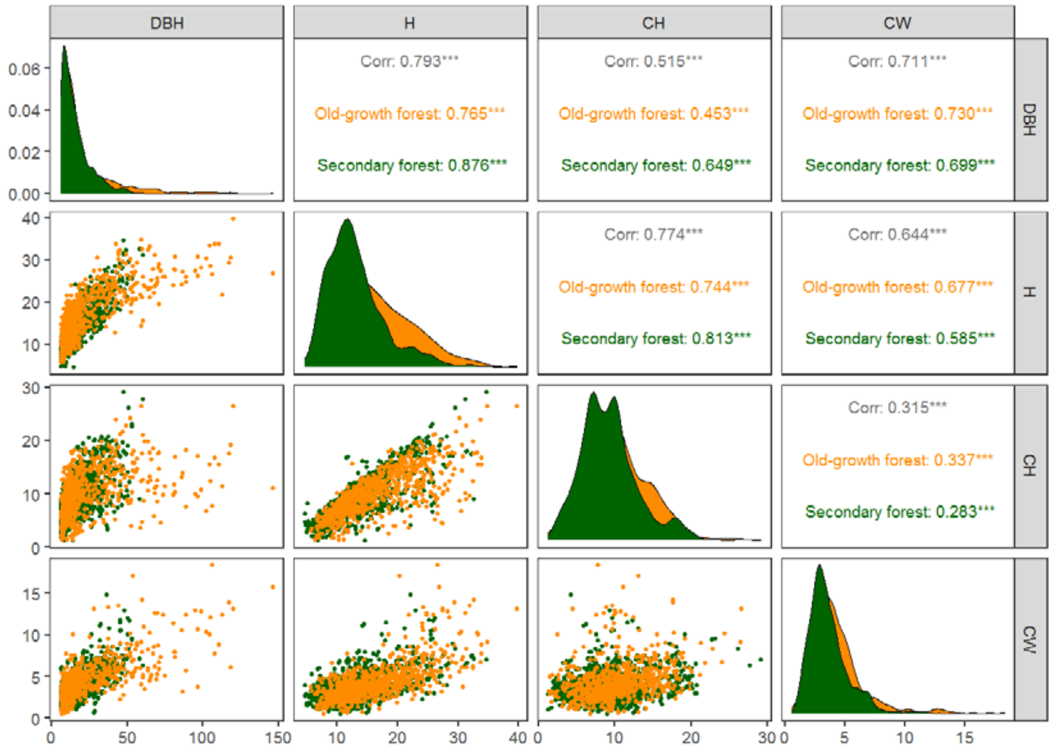
Table 3. Descriptive statistics for tree-size variables.

Data type	Forest type	Variable	Count	Mean	Min	Max	SD
Fitting (7 plots per forest type)	Secondary forest	DBH (cm)	1631	15.81	6.00	60.15	9.44
		H (m)	1631	12.91	4.60	34.60	4.61
		CH (m)	1631	9.22	1.30	29.10	3.70
		CW (m)	1631	3.62	0.60	14.85	1.71
	Old-growth forest	DBH (cm)	774	21.53	6.00	146.35	20.16
		H (m)	774	16.00	5.70	39.70	6.03
		CH (m)	774	10.04	1.40	26.50	4.06
		CW (m)	774	4.33	0.55	25.70	2.34
Validation (3 plots per forest type)	Secondary forest	DBH (cm)	407	14.81	6.05	63.30	8.71
		H (m)	407	12.14	5.20	34.80	4.17
		CH (m)	407	8.94	1.30	29.20	3.51
		CW (m)	407	3.46	0.80	12.20	1.49
	Old-growth forest	DBH (cm)	320	20.57	6.10	100.15	15.84
		H (m)	320	17.90	6.40	37.20	5.88
		CH (m)	320	11.24	2.50	31.10	4.25
		CW (m)	320	4.41	0.95	12.00	1.78

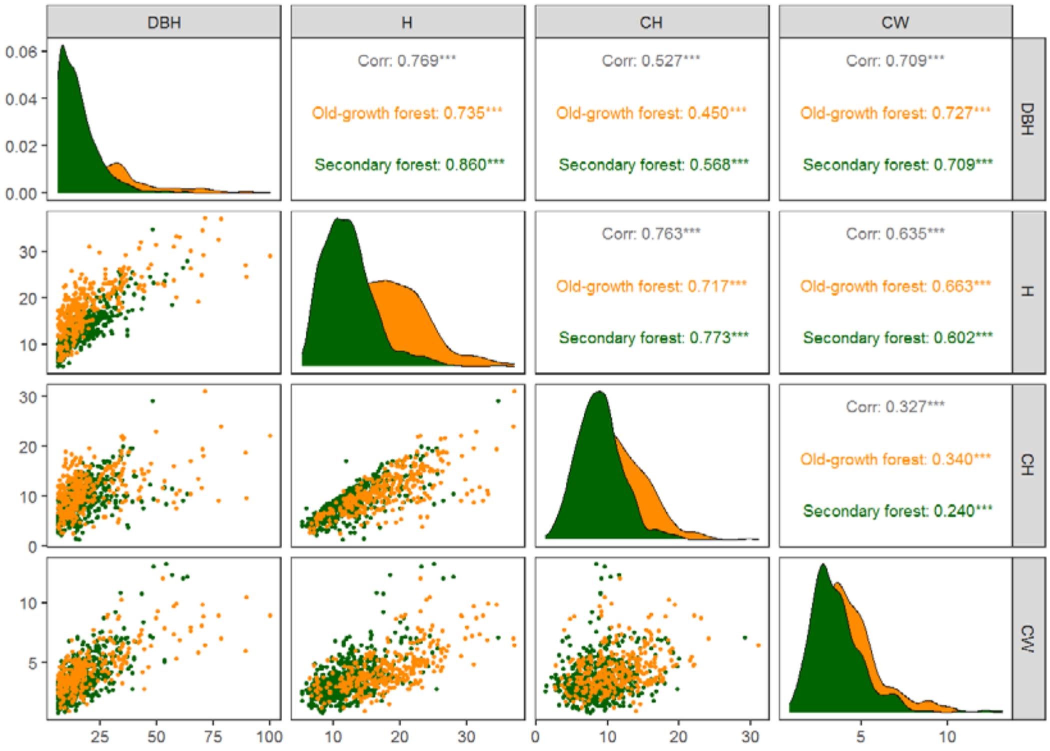
In general, the number of individuals with diameters >6 cm in secondary forests was higher than in old-growth forests in both the fitting data and validation data sets. The mean values of the measured variables of the secondary forest were smaller than those of the old-growth forest. In addition, secondary forests had smaller data dispersion than old-growth forests. Also, because the data were relatively homogeneous between the plots in each forest stage, the statistical descriptive values for the tree-size variables did not differ significantly between the fitting and the validation data sets.

General relationship trends in data sets. Correlations between the pairs of variables for DBH-H, DBH-CH and DBH-CW were analyzed and are shown in Figure 4.

Pearson correlations revealed that the relationship between diameter and total height was strongest, followed by the relationship between diameter and canopy width, and finally the relationship between diameter and commercial height in both datasets. This trend occurred in the two forest stages. For all variables, values variation increased as the variables took on larger values. The graphs of forest tree density in Figure 4 by each pair of variables were mostly in the form of one peak and positive skew. The distributions of the secondary forest were often more skewed than those of the old-growth forest.



a)



b)

Figure 4. Correlation between pairs of variables in two datasets. a) for fitting data and b) for validation data.

3.2 Non-linear models

Model development. Parameter estimates for the best models are given in Table 4. The estimates were calculated based on the fitting data of the two forest states. Table 4 shows that the best equations varied widely between pairs of variables and between forest stages. The strongest relationship was between diameter and total height, followed by diameter and crown width. The weakest correlation was between diameter and commercial height. All parameters of the best equations were significant (Pr(>|t|) < 0.05).

Expected values of the best and the worst equations for pairs of variables are provided (Fig. 5). There was a significant difference between the best function and the worst function.

Model validation. After the models were built from the fitting data, they were evaluated with the validation data. The results of comparing the magnitudes of AIC, R^2 and RMSE are presented in Figure 6. The difference in goodness of fit statistics between the fitting and validation data sets was not great, especially with R^2 and RMSE (Fig. 6). All the goodness of fit statistics were smaller in the old-growth forest, except for the RMSE values of H-DBH and CH-DBH.

Table 4. Parameter estimates for tree-size variables for the best non-linear models.

<table>
<tr><th>Type</th><th>Relationship</th><th>Model</th><th>Parameter</th><th>Estimation</th><th>SE</th><th>t-value</th><th>Pr(>|t|)</th><th>R-square</th><th>AIC</th><th>RMSE</th></tr>
<tr><td rowspan="8">Secondary forest</td><td rowspan="2">H-DBH</td><td rowspan="2">Power</td><td>a</td><td>2.832</td><td>0.0666</td><td>42.55</td><td><2e-16</td><td rowspan="2">0.7753</td><td rowspan="2">6342.525</td><td rowspan="2">2.2494</td></tr>
<tr><td>b</td><td>0.567</td><td>0.0078</td><td>72.43</td><td><2e-16</td></tr>
<tr><td rowspan="3">CH-DBH</td><td rowspan="3">Korf</td><td>a</td><td>27.173</td><td>6.2534</td><td>4.35</td><td>1.49e-05</td><td rowspan="3">0.4530</td><td rowspan="3">6999.114</td><td rowspan="3">2.8320</td></tr>
<tr><td>b</td><td>4.576</td><td>0.3985</td><td>11.48</td><td>1.41e-06</td></tr>
<tr><td>c</td><td>0.504</td><td>0.1046</td><td>4.82</td><td>1.63e-06</td></tr>
<tr><td rowspan="3">CW-DBH</td><td rowspan="3">Ratskowky</td><td>a</td><td>33.650</td><td>9.7870</td><td>3.44</td><td>0.001</td><td rowspan="3">0.4892</td><td rowspan="3">4549.464</td><td rowspan="3">1.1961</td></tr>
<tr><td>b</td><td>137.322</td><td>34.2550</td><td>4.01</td><td>6.42e-05</td></tr>
<tr><td>c</td><td>34.877</td><td>7.1170</td><td>4.90</td><td>1.07e-06</td></tr>
<tr><td rowspan="8">Old-growth forest</td><td rowspan="3">H-DBH</td><td rowspan="3">Prodan</td><td>a</td><td>-0.112</td><td>0.4247</td><td>-0.27</td><td>0.791</td><td rowspan="3">0.7059</td><td rowspan="3">4038.728</td><td rowspan="3">3.2703</td></tr>
<tr><td>b</td><td>0.610</td><td>0.0490</td><td>12.47</td><td><2e-16</td></tr>
<tr><td>c</td><td>0.031</td><td>0.0010</td><td>30.74</td><td><2e-16</td></tr>
<tr><td rowspan="3">CH-DBH</td><td rowspan="3">Weibull</td><td>a</td><td>12.481</td><td>0.3644</td><td>34.25</td><td><2e-16</td><td rowspan="3">0.3325</td><td rowspan="3">4060.941</td><td rowspan="3">3.3176</td></tr>
<tr><td>b</td><td>0.074</td><td>0.0172</td><td>4.33</td><td>1.68e-05</td></tr>
<tr><td>c</td><td>1.040</td><td>0.1037</td><td>10.03</td><td><2e-16</td></tr>
<tr><td rowspan="2">CW-DBH</td><td rowspan="2">Power</td><td>a</td><td>1.055</td><td>0.0529</td><td>19.95</td><td><2e-16</td><td rowspan="2">0.5450</td><td rowspan="2">2848.357</td><td rowspan="2">1.5177</td></tr>
<tr><td>b</td><td>0.483</td><td>0.0145</td><td>33.34</td><td><2e-16</td></tr>
</table>

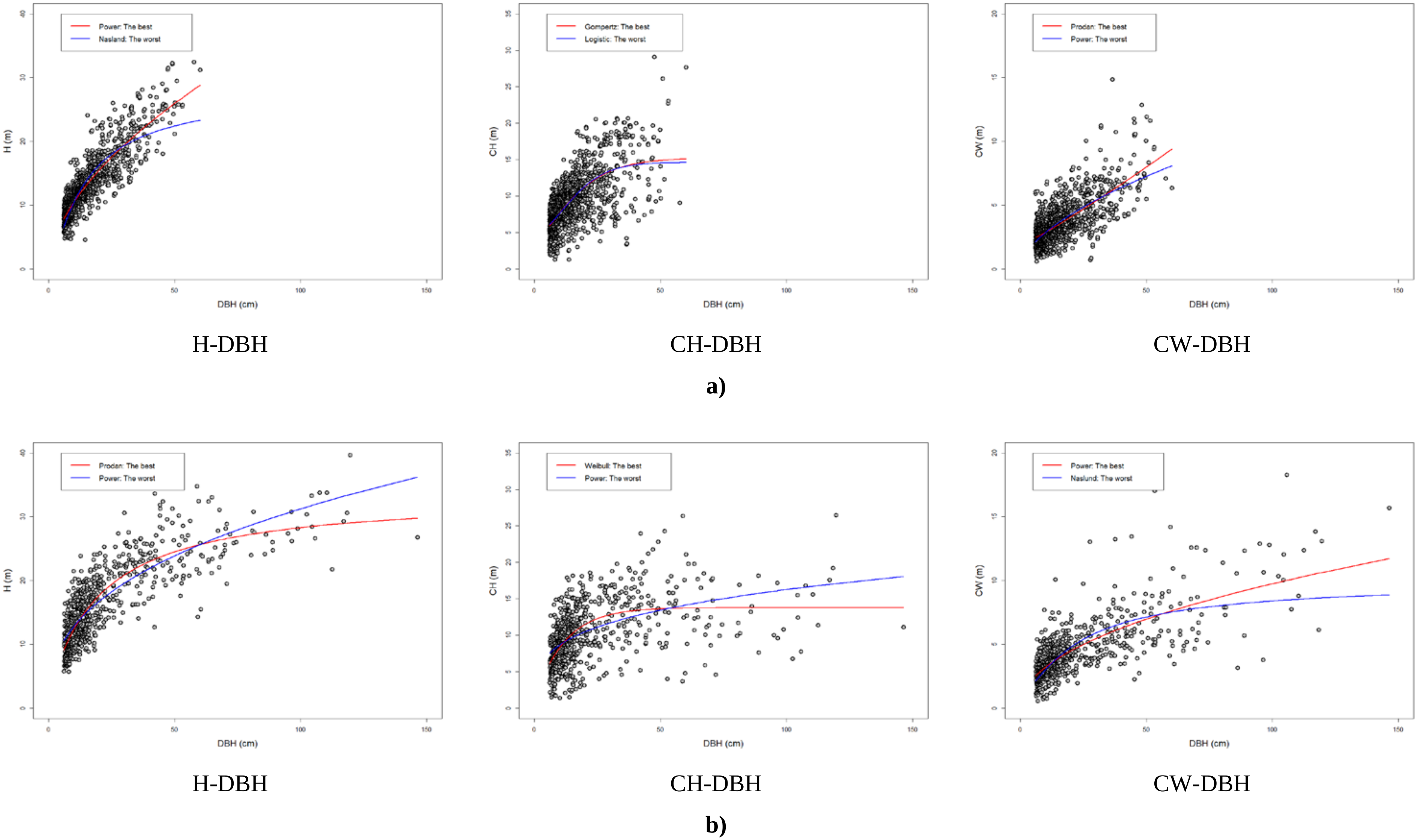
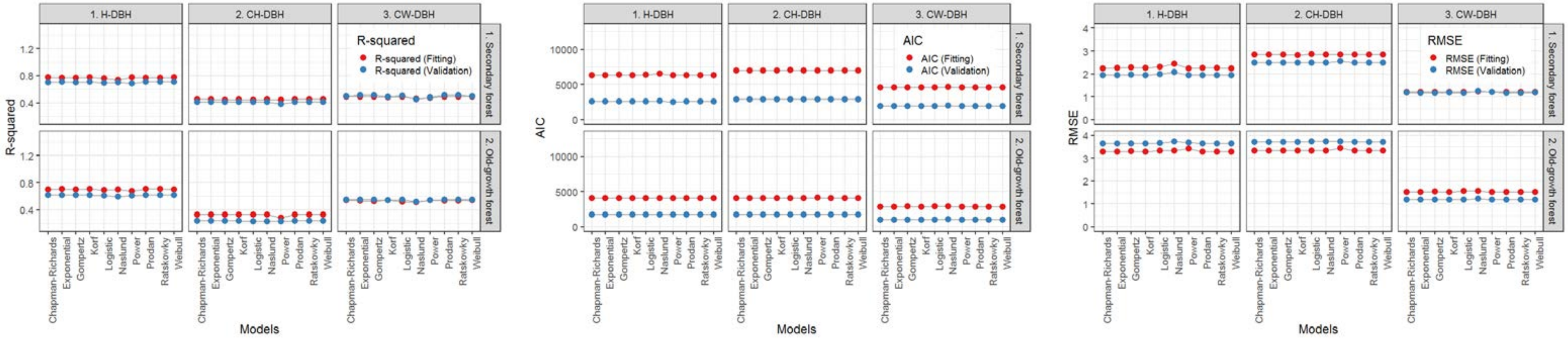


Figure 5. Regression charts for the best (red line) and the worst (blue line) models for a) secondary forest, and b) old-growth forest.



a) b) c)

Figure 6. Goodness of fit statistics for 10 nonlinear models in two forest stages. a) for R2, b) for AIC and c) for RMSE**.**

3.3 Nonlinear mixed-effect models

The best nonlinear model for the pairs of variables was selected for inclusion in the nonlinear mixed-effect models. The results of parameter estimation for fixed effects, standard deviation for random effects and AIC for mixed models are presented in Table 5.

The results indicated that most of the mixed models were better than the classical nonlinear models. Only in the relationship between canopy width and diameter in the old-growth forest, the mixed model was not relevant. The rates of AIC reduction in secondary forests were greater than in old-growth forests. The distributions of the standardized residuals had no departures from the normal distribution. In addition, residuals' variances were quite homogeneous after applying the varExp function.

Table 5. Parameter estimates for mixed models.

Type	Relationship	Mixed model		Parameter	Estimation	AIC	Change
Secondary forest	H-DBH	Mixed Power	Fixed effects	a	3.05180	2440.975	-5.15%
				b	0.52362		
			Random effects	SD: Section	0.05775		
				SD: Plot in Section	0.00002		
	CH-DBH	Mixed Korf	Fixed effects	a	13.35266	2753.467	-4.36%
				b	7.62258		
				c	1.03289		
			Random effects	SD: Section	0.42823		
				SD: Plot in Section	0.00042		
	CW-DBH	Mixed Ratskowky	Fixed effects	a	28.12362	1848.224	-5.02%
				b	116.08720		
				c	30.68818		
			Random effects	SD: Section	0.00010		
				SD: Plot in Section	0.00128		

Old-growth forest	H-DBH	Mixed Prodan	Fixed effects	a	-0.14882	1732.928	-0.47%
				b	0.52656		
				c	0.02951		
			Random effects	SD: Section	0.49580		
				SD: Plot in Section	0.49580		
	CH-DBH	Mixed Weibull	Fixed effects	a	14.34438	1741.282	-0.75%
				b	0.15805		
				c	0.72385		
			Random effects	SD: Section	0.93070		
				SD: Plot in Section	0.93057		
	CW-DBH	Mixed Power	Fixed effects	a	6.00201	1753.222	70.50%
				b	0.37887		
			Random effects	SD: Section	0.00001		
				SD: Plot in Section	0.00001		

Note: SD is standard deviation, a, b, c are parameters for fixed effects. Change was (AIC mixed models-AIC nonlinear models)/AIC nonlinear models*100%.

4. Discussion

4.1 Tree-size variable characteristics and general correlations between them

Data and forest communities are often analyzed and classified into homogeneous groups in terms of tree-size variables, environmental factors, or a combination of both. After being classified, the groups are often used for later analysis or as a basis for proposing silvicultural, planning and management measures (Maua et al., 2020; Thormann et al., 2011; Truax et al., 2015; Tunçkol et al., 2020). An example of forest community classification is a study in Québec, Canada where the authors classified a research site into 11 forest communities based on species name, diameter and ecological factors such as soil, forest age, slope, and elevation (Truax et al., 2015). Another study in South Nandi forest, Kenya, divided survey plots into two plant communities based on diameter, height and species composition; one with 8 species and the other with 58 species (Maua et al., 2020). In this study, forest plots were also divided into two distinct groups. The plots were surveyed from 2 different forest stages: secondary forests and old-growth forests. The secondary and old-growth forests were statistically different (Fig. 3) and they should be kept separate for further analysis. The homogeneity of the data in each forest stage is a convincing basis for dividing the data into fitting and validation data sets.

In the secondary forest, the density of forest trees was higher (Table 3), but the average size of trees was smaller than in the old-growth forest. The density of forest trees was higher in the secondary forest because the forest is in the recovery stage after logging. As a result of selective harvest, gaps were created as the canopy was opened up, allowing trees to regenerate at high density (N.V. Hoan, 2013; P.X. Hoan & Ngu, 2003). The large density of small trees will lead to a smaller means in the secondary forest. In the secondary forest, the diameter mean was 14.81 - 15.81 cm; total height was 12.14 - 12.91 m; commercial height was 8.94 - 9.22 m and canopy diameter was 3.46 - 3.62 m. These values in the old forest were 20.57 - 21.53 cm; 16.00 - 17.90 m; 10.04 - 11.24 and 4.33 - 4.41 m, respectively. The results are quite similar to previous studies on the same forest stages in the Central Highlands of Vietnam (Binh, 2014; Ha & Hang, 2010; Sau, 1996).

The density distribution of variable pairs were often skewed to the right. This is due to the tree-size variables being positively related to each other. At the same time, the proportion of small trees was often very large in the stand, making the distribution positively skewed (Binh, 2014; Hung, 2016). The strongest relationship was diameter at breast height vs total height, while the weakest was diameter vs commercial height. Similar findings were obtained by a study for Calabrian pines in Turkey where the strongest relationship was stem diameter vs height (R^2=0.82) and the weakest relationship was between crown diameter and stem diameter (R^2=0.74) (M. D. Avsar, 2004). In contrast, the strongest correlation was between diameter and crown diameter, followed by diameter and height, for oaks in Poland, and Lebanese Cedars in Turkey (D. Avsar & Ayyıldız, 2005; Kazmierczak et al., 2011). Another study in Great Britain on relationships between urban tree variables found that the correlation between diameter and height was weaker than between crown width and diameter (Monteiro et al., 2016). These differences may be due to differences in plant species, location, climate, site fertility, competition among individuals, and other local factors (P. X. Hoan & Ngu, 2003; Monteiro et al., 2016).

4.2 Nonlinear model differences among pairs of variables and forest stages

Nonlinear models were built for pairs of variables in forest stages including height-diameter, crown width-diameter and commercial height-diameter. These models were statistically significant (p-value < 0.001). Therefore, they can be used to predict quantities such as height, crown diameter, and commercial height of forest trees based on the diameter at breast height. It is an easy quantity to measure in forests, especially tropical forests. The ten used models

(Table 1) performed quite well, except for the relationship between commercial height and diameter in the old-growth forest. This relationship is relatively weak (R^2=0.3325).

The ten examined nonlinear functions also showed the same trend, namely that diameter vs height relationship was strongest, followed by diameter vs canopy diameter and diameter vs commercial height. This trend occurred in both stages (secondary forests and old-growth forests) and in both fitting and validation data sets. Based on the analyzed results in both sets, the best function for each pair of variables differed between forest stages (Table 4). This can be explained by different tree species, tree size and site conditions (P.X. Hoan & Ngu, 2003; Monteiro et al., 2016). However, they are all functions that have proven to be suitable for describing the relationship between allometric variables of tropical forest trees. There is considerable range in the models reported in the literature for particular forest ecosystems. The power model best described the relationship between diameter and height of the rainforest in Limpopo Mistbelt Forests and the whole of South Africa (Mensah et al., 2018; Mensah et al., 2017); and of the upper montane forests in Kahuzi Biega National Park, Democratic Republic of Congo (Imani et al., 2017). The relationship between diameter and height of mixed-species forests in Slovakia was modeled best using the Korf function (Petrás et al., 2014). In a study in northeast China, the authors showed that the Ratskowky model had the highest R^2 value (Thanh et al., 2019). The Prodan model was the highest-ranked model for simulating the relationship between diameter and height of Field Maple and Hornbeam species in northwest Iran (Abedi & Abedi, 2020) and old-growth forest trees in Bosnia and Herzegovina (Govedar et al., 2018). Furthermore, the Weibull function was the best fit for the diameter-height relationship of tropical forests in Africa (Imani et al., 2017) and southeastern Brazil (Scaranello et al., 2012).

4.3 Significance of mixed-effect models

Nonlinear mixed-effect models help researchers examine the effects of both fixed and random factors (Chenge, 2021; Pinheiro & Bates, 2006). This study examined the random effects of Section and Plot. The results showed that in both forest stages, the nonlinear mixed model was better and significant in reducing AIC values. The AIC values of the mixed model were all lower than those of the traditional nonlinear models, except for the relationship between crown width and diameter in the old-growth forest. These findings support a number of previous studies. For example, mixed-effect models were better than nonlinear models in predicting total height from diameter based on AIC (Chenge, 2021; Ercanli, 2015; Özçelik et al., 2018; Zang et al., 2016). However, previous studies on mixed-effect models for the height-diameter relationship have had limited applications globally, except for Nigeria

(Chenge, 2021), Turkey (Ercanli, 2015; Tunçkol et al., 2020), Spain (Ogana et al., 2020), and China (Xu et al., 2022). So far, nonlinear mixed-effect models have not been applied to commercial height-diameter and crown width-diameter relationships. Especially these models have not been tested for natural forests in Vietnam. In this study, random effects of factors such as species, tree size, tree density, and tree quality were not considered due to data limitations. Therefore, in future studies, these factors should be included in the mixed model to evaluate more pronounced effects of random factors.

5. Conclusions

The study used measurement data of diameter at breast height, total height, commercial height and crown diameter of forest trees in 2 stages: secondary forests and old-growth forests in Gia Lai, Vietnam. The analysis results showed that the data between plots of each state were relatively homogeneous. The correlation between diameter and height was the strongest. In the secondary forest, the Power, Korf and Ratskowky models were the best for height vs diameter, commercial height vs diameter and crown width vs diameter, respectively. In the old-growth forest, Prodan, Weibull and Power models were the best fit to describe the above relationships. In general, in both forest stages, the mixed nonlinear models outperformed the traditional nonlinear models. All models (fixed-effect and mixed-effect) developed in this study can be applied to determine height, commercial height, canopy diameter based on diameter in the study area and other regions with similar conditions. Future studies can add more random factors such as tree size, tree quality, tree density to the model to test or develop specific models for each family or tree species in the study area. The models will be useful for estimating and auditing biomass and carbon stock accumulating in forests in Vietnam.

Acknowledgments

We would like to thank the Government of Vietnam and Rufford, UK for providing financial support to carry out the field survey in this study. We are also extremely grateful for the support of the board of directors, along with all rangers and staff of Kon Ka Kinh National Park, who provided accommodation and supported field work. We would also like to thank scientists from Vietnam National University of Forestry and TU-Dresen, Germany for their advice and suggestions during the implementation of this study.

References

Abedi R. & Abedi T., 2020, Some non-linear height-diameter models performance for mixed stand in forests in Northwest Iran. Journal of Mountain Science 17(5): 1084–1095. https://doi.org/10.1007/s11629-019-5870-4

Anacioco K.P., Gorio J.A.L., Padsico M.R.S., Lumbres R.I.C., Doyog N.D. & Lee Y.J., 2018, Fitting and evaluation of height-diameter models for Alnus japonica in La Trinidad, Benguet, Philippines. Journal of Mountain Science 15(11): 2422–2432. https://doi.org/10.1007/s11629-018-4866-9

Anderson M.J., 2011, A new method for non-parametric multivariate analysis of variance. Austral Ecology (26): 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Avsar D. & Ayyıldız V.A., 2005, The relationships between diameter at breast height, tree height and crown diameter in Lebanon cedars (*Cedrus libani* A. Rich.) of the Yavsan Mountain, Kahramanmaras, Turkey. Pakistan Journal of Biological Sciences 8: 1228–1232. https://doi.org/10.3923/pjbs.2005.1228.1232

Avsar M.D., 2004, The relationships between diameter at breast height, tree heightand crown diameter in Calabrian pines (*Pinus brutia* Ten.) of baskonus mountain, Kahramanmaras, Turkey. Journal of Biological Sciences (Pakistan) 4(4): 437–440. https://doi.org/10.3923/jbs.2004.437.440

Binh N.T., 2014, Research on forest structure characteristics and biodiversity of closed evergreen mixed broad and needle leaf forest type in Bidoup - Nui Ba National Park. Vietnam Journal of Forest Science 2(2014): 3255–3263.

Brack C., 1999, Forest Measurement and Modelling. Available from: http://fennerschool-associated.anu.edu.au/mensuration/author.htm#copyright.

Chai Z., Tan W., Li Y., Yan L., Yuan H. & Li Z., 2018, Generalized nonlinear height–diameter models for a *Cryptomeria fortunei* plantation in the Pingba region of Guizhou Province, China. Web Ecology 18(1): 29–35. https://doi.org/10.5194/we-18-29-2018

Chenge I.B., 2021, Height–diameter relationship of trees in Omo strict nature forest reserve, Nigeria. Trees, Forests and People, 3, 100051. https://doi.org/10.1016/j.tfp.2020.100051

Ercanli I., 2015, Nonlinear mixed effect models for predicting relationships between total height and diameter of Oriental beech trees in Kestel, Turkey. Revista Chapingo Serie Ciencias Forestales y del Ambiente 21(2): 185–202. https://doi.org/10.5154/r.rchscfa.2015.02.006

Fang Z. & Bailey R., 1998, Height–diameter models for tropical forests on Hainan Island in southern China. Forest Ecology and Management 110(1-3): 315–327. https://doi.org/10.1016/S0378-1127(98)00297-7

Faraway J.J., 2016, Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models. CRC Press, Taylor & Francis Group, New York.

Gałecki A. & Burzykowski T., 2013, Linear Mixed-Effects Models Using R: A Step-by-Step Approach. Springer New York, USA.

Govedar Z., Krstić M., Keren S., Babić V., Zlokapa B. & Kanjevac B., 2018, Actual and balanced stand structure: Examples from beech-fir-spruce old-growth forests in the area of the Dinarides in Bosnia and Herzegovina. Sustainability 10(2), 540. https://doi.org/10.3390/su10020540

Ha D.T. & Hang B.T., 2010, Research on structure of evergreen broadleaved natural forest (IIIa) promoted solution to sustainalbe forest business in Kon Ray, Kon Tum. Vietnam Journal of Forest Science 1(2010): 2215–2226.

Hien C.T.T., Hung B.M., Cuong N.D. & Bich N.V., 2019, Several forest structure characteristics and diversity profile of overstory trees of evergreen broadleaf forest in Ba Be national park. Journal of Forestry Science and Technology 3/2019: 35–45 (in Vietnamese, summary in English).

Hoan N.V., 2013, Report on sustainable forest planning. Kon Ka Kinh National Park, Mang Yang, Gia Lai Province, Vietnam (in Vietnamese).

Hoan P.X. & Ngu H.K., 2003, Silviculture. The Agriculture Publisher, Hanoi, Vietnam (in Vietnamese).

Hung B.M., 2016, Structure and restoration of natural secondary forests in the Central Highlands, Vietnam. Doctoral dissertation. Institute of Silviculture and Forest Protection, Faculty of Environmental Sciences, Dresden University of Technology, Dresden, Germany.

Hung B.M., 2022, Forestry data analysis, using Stata. Scientific and Technical Publishing House, Hanoi, Vietnam.

Huy D.V. & Hung B.M., 2018, Dynamics of cunminghamia lanceolata plantation structure and quality on ages in Si Ma Cai district, Lao Cai Province. Journal of Forestry Science and Technology 4/2018: 22–31 (in Vietnamese, summary in English).

Imani G., Boyemba F., Lewis S., Nabahungu N.L., Calders K., Zapfack L., Riera B., Balegamire C. & Cuni-Sanchez A., 2017, Height-diameter allometry and above ground biomass in tropical montane forests: Insights from the Albertine Rift in Africa. PloS one 12(6): 1–20. https://doi.org/10.1371/journal.pone.0179653

Kazmierczak K., Pazdrowski W., Jedraszak A., Szymanski M. & Nawrot M., 2011, Crown width of a tree and its relationships with age, height and diameter at breast height based on common oak (*Quercus robur* L.). Colloquium Biometricum 41: 221–228.

Li Y.-q., Deng X.-w., Huang Z.-h., Xiang W.-h., Yan W.-d., Lei P.-f., Zhou X.-l. & Peng C.-h., 2015, Development and evaluation of models for the relationship between tree height and diameter at breast height for Chinese-fir plantations in subtropical China. PloS one 10(4), e0125118. https://doi.org/10.1371/journal.pone.0125118

Liu M., Feng Z., Zhang Z., Ma C., Wang M., Lian B.-l., Sun R. & Zhang L., 2017, Development and evaluation of height diameter at breast models for native Chinese Metasequoia. PloS one 12(8), e0182170. https://doi.org/10.1371/journal.pone.0182170

Maua J.O., MugatsiaTsingalia H., Cheboiwo J. & Odee D., 2020, Population structure and regeneration status of woody species in a remnant tropical forest: A case study of South Nandi forest, Kenya. Global Ecology and Conservation 21, e00820. https://doi.org/10.1016/j.gecco.2019.e00820

Mensah S., Pienaar O.L., Kunneke A., du Toit B., Seydack A., Uhl E., Pretzsch H. & Seifert T., 2018, Height–Diameter allometry in South Africa’s indigenous high forests: Assessing generic models performance and function forms. Forest Ecology and Management 410: 1–11. https://doi.org/10.1016/j.foreco.2017.12.030

Mensah S., Veldtman R. & Seifert T., 2017, Allometric models for height and aboveground biomass of dominant tree species in South African Mistbelt forests. Southern Forests: a Journal of Forest Science, 79(1): 19–30. https://doi.org/10.2989/20702620.2016.1225187

Miah M.D., Islam K.N., Kabir M.H. & Koike M., 2020, Allometric models for estimating aboveground biomass of selected homestead tree species in the plain land Narsingdi district of Bangladesh. Trees, Forests and People 2, 100035. https://doi.org/10.1016/j.tfp.2020.100035

Monteiro, M. V., Doick, K. J., & Handley, P. (2016): Allometric relationships for urban trees in Great Britain. Urban Forestry & Urban Greening 19: 223–236. https://doi.org/10.1016/j.ufug.2016.07.009

Ogana F.N., Corral-Rivas S. & Gorgoso-Varela J.J., 2020, Nonlinear mixed-effect height-diameter model for *Pinus pinaster* Ait. and *Pinus radiata* D. Don. Cerne 26: 150–161. https://doi.org/10.1590/01047760202026012695

Özçelik R., Cao Q.V., Trincado G. & Göçer N., 2018, Predicting tree height from tree diameter and dominant height using mixed-effects and quantile regression models for two species in Turkey. Forest Ecology and Management 419: 240–248. https://doi.org/10.1016/j.foreco.2018.03.051

Petrás R., Bosela M., Mecko J., Oszlányi J. & Popa I., 2014, Height-diameter models for mixed-species forests consisting of spruce, fir, and beech. Folia Forestalia Polonica, Series A, Forestry: 56(2). https://doi.org/10.2478/ffp-2014-0009

Philip M.S., 1998, Measuring Trees and Forests. CABI Publishing, CAB International, Wallingford, Oxon, OX10, 8DE, UK.

Pinheiro J. & Bates D., 2006, Mixed-effects models in S and S-PLUS. Springer Science & Business Media.

Pinheiro J., Bates D., DebRoy S., Sarkar D. & R Core Team, 2016, nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1–126. http://CRAN.R-project.org/package=nlme

Sau L., 1996, Research on forest structure characteristics and propose some economic technical criteria for selective harvest and sustainable forest use in Kon Ha Nung - Central Highlands. Vietnam National University of Forestry, Xuan Mai, Hanoi, Vietnam.

Scaranello M.A. d. S., Alves L.F., Vieira S.A., Camargo P.B. d., Joly C.A. & Martinelli L.A., 2012, Height-diameter relationships of tropical Atlantic moist forest trees in southeastern Brazil. Scientia Agricola 69(1): 26–37. https://doi.org/10.1590/S0103-90162012000100005

Sharma R., Vacek Z. & Vacek S., 2016, Nonlinear mixed effect height-diameter model for mixed species forests in the central part of the Czech Republic. Journal of Forest Science 62(10): 470–484. https://doi.org/10.17221/41/2016-JFS

Sullivan M.J., Lewis S.L., Hubau, W., Qie L., Baker T.R., Banin L.F., Chave J., Cuni-Sanchez A., Feldpausch T.R. & Lopez-Gonzalez G., 2018, Field methods for sampling tree height for tropical forest biomass estimation. Methods in Ecology and Evolution 9(5): 1179–1189. https://doi.org/10.1111/2041-210X.12962

Team R.C., 2021, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Thanh T.N., Tien T.D. & Shen H.L., 2019, Height-diameter relationship for Pinus koraiensis in Mengjiagang Forest Farm of Northeast China using nonlinear regressions and artificial neural network models. Journal of Forest Science 65(4): 134–143. https://doi.org/10.17221/5/2019-JFS

Thormann B., Raupach M.J., Wagner T., Wägele J.W. & Peters M.K., 2011, Testing a short nuclear marker for inferring staphylinid beetle diversity in an African tropical rain forest. PloS one 6(3), e18101. https://doi.org/10.1371/journal.pone.0018101

Trieu N.V., 2017, Research on structural characteristics and biodiversity of natural forests at Tam Dao National Park, Vinh Phuc. Vietnam National University of Forestry, Xuan Mai, Hanoi, Vietnam.

Truax B., Gagnon D., Lambert F. & Fortier J., 2015, Multiple-use zoning model for private forest owners in agricultural landscapes: A case study. Forests 6(10): 3614–3664. https://doi.org/10.3390/f6103614

Tuan H.M., 2017, Research on forest structure and plant diversity of natural forests at Ba Be National Park, Bac Kan. Vietnam National University of Forestry, Xuan Mai, Hanoi, Vietnam.

Tunçkol B., Aksoy N., Çoban S. & Zengin H., 2020, Diversity and ecology of forest communities in Küre Mountains National Park of Turkey. BOSQUE 41(3): 289–305. https://doi.org/10.4067/S0717-92002020000300289

Van P.Q. & Hien C.T.T., 2018, Some structural characteristics and tree species diversity of forest state IIIA in An Lao district, Binh Dinh province. Journal of Forestry Science and Technology 1/2018: 69–78 (in Vietnamese, summary in English).

West B.T., Welch K.B. & Gatecki A.T., 2015, Linear Mixed Models: A practical Guide Using Statistical Software. CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, Florida, USA.

Xu Q., Lei X., Zang H. & Zeng W., 2022, Climate Change Effects on Height–Diameter Allometric Relationship Vary with Tree Species and Size for Larch Plantations in Northern and Northeastern China. Forests 13(3), 468. https://doi.org/10.3390/f13030468

Zang H., Lei X. & Zeng W., 2016, Height–diameter equations for larch plantations in northern and northeastern China: a comparison of the mixed-effects, quantile regression and generalized additive models. Forestry: An International Journal of Forest Research 89(4): 434–445. https://doi.org/10.1093/forestry/cpw022

Zar J.H., 2010, Biostatistical Analysis (5th Edition). Prentice Hall, Upper Saddle River, New Jersey 07458, USA.

Zhang L., 1997, Cross-validation of non-linear growth functions for modelling tree height–diameter relationships. Annals of Botany 79(3): 251–257. https://doi.org/10.1006/anbo.1996.0334

Zuur A.F., Ieno E.N., Walker N.J., Saveliev A.A. & Smith G.M., 2009, Mixed Effects Models and Extensions in Ecology with R. Springer, USA.