Trees, Forests and People 2 (2020) 100028

Contents lists available at ScienceDirect

Trees, Forests and People

journal homepage: www.elsevier.com/locate/tfp

The tree height estimated by non-power models on volumetric models provides reliable predictions of wood volume: The Amazon species height modelling issue

Rodrigo Geroni Mendes Nascimento[a,*], Jerome Klaas Vanclay[b], Afonso Figueiredo Filho[c], Sebastião do Amaral Machado[d], Ademir Roberto Ruschel[e], Nelson Akira Hiramatsu[f], Lucas José Mazzei de Freitas[e]

[a] *Amazon Federal Rural University (UFRA) – Belém Campus, Institute of Agricultural Sciences, Graduate Program in Forest Sciences, Brazil*
[b] *Southern Cross University (SCU) – Lismore Campus – School of Environment, Science and Engineering, Graduate Program in Forest Science and Management, Australia*
[c] *State University of Midwest Paraná (UNICENTRO) – Riozinho Campus, Department of Forest Engineering, Postgraduate Program in Forest Sciences, Brazil*
[d] *Federal University of Paraná (UFPR) – Jardim Botânico Campus, Department of Forest Engineering, Graduate Program in Forest Engineering, Brazil*
[e] *Brazilian Agricultural Research Corporation (EMBRAPA) – Eastern Amazon Head Office, Laboratory of Ecology and Forest Management, Brazil*
[f] *Eucatex - Industry and Commerce S.A. – Forest Unit, Planning and Forest Inventory Coordination, Brazil*

ARTICLE INFO

Keywords:
Allometry
Forest inventory
Tropical rain forest
Wood estimates

ABSTRACT

Allometries that include height as independent variable usually provide greater accuracy on estimates of volume, biomass or individual carbon than other prediction strategies that rely only diameter at breast height as independent variable. However, when these models are applied in Amazon Forest Inventories, it is common to use estimated heights rather than measured heights to prepare volume, biomass or carbon estimates. This practice is common, but rarely discussed and the effect on predictions and precision is usually overlooked. The aim of this study was to examine hypsometric models and evaluate the effect of estimated height on merchantable volume prediction in Eastern Amazonian forests. The study area was a 3,786 ha Forest Management Unit owned by Jari Florestal S.A., in the Jari Valley Region of the State of Pará, Brazil. The data includes 16,099 trees of 25 species, measured and harvested in 2006. Ten percent of the data were reserved for validation of the hypsometric and volumetric estimates. Five hypsometric models and two modelling techniques (linear regression and mixed-effects model) were examined. The choice of best model was based on graphical analyses of residuals, distribution of residuals, heteroscedasticity of error and presence of outliers as assessed by h-values, DFFITS and Cook's distance. The hypsometric relationship and volumetric estimates using DBH and DBH with estimated height were validated with Graybill's test, Theil's error decomposition, Efficiency, Equivalence test and Tukey's test for species estimates level. Heights estimated using a semi-logarithmic mixed-effects model can improve predictions from volume equations. The results show that exploratory data analysis and validation process helped to provide estimates with greater efficiency and should be adopted in related studies. The prediction of height associated with volumetric models for six different species provided volumetric estimates with an error below 5% for the global average volume. The estimated height by the mixed-effect non-power law model should be included in double input models previously developed for volume prediction.

1. Introduction

Since the beginning of the twentieth century mathematical models have been used to describe and predict the height-diameter relationship and volume of trees. These topics have been widely studied in forest biometrics (Trorey 1932; Huxley and Teissier, 1936;Prodan, 1944; Henricksen, 1950; Stoffels and van Soest 1953; Curtis, 1967) and remain central to forest management. Height diameter models are useful because they permit inferences about the forest productivity (Trorey, 1932; Cole and Ewel, 2006), dynamics of population structure (Hunter et al., 2013), health conditions and growth rates (Curtis, 1967), ecological characteristics (Shama and Parton, 2007), effects of silvicultural and cultural treatments (Willmott et al., 2006) and assist in the discrimination between forests and species (Fang and Bailey, 1998; Cole and Ewel, 2006; Vibrans et al., 2015). Because of their importance in the management of both natural forests and forest plantations, it is appropriate to examine closely the attributes of popular hypsometric relationships (Feldpausch et al., 2011; Fang and Bailey, 1998). In addition, the strong relationship between the dominant height and site

---

* Corresponding author.
E-mail address: rodrigo.geroni@ufra.edu.br (R.G.M. Nascimento).

https://doi.org/10.1016/j.tfp.2020.100028
Received 1 June 2020; Received in revised form 19 August 2020; Accepted 24 August 2020

yield potential helps to diagnose the maximum stock capacity in volume, biomass or carbon per area unit (Lima et al., 2012; Rutishauser et al., 2013). The use of hypsometric equations may assist in the identification of preferred sites (Feldpausch et al., 2011; Hunter et al., 2013), or those with the greatest reestablishment capacity after natural or human disturbance (Machado et al., 2008). In short, describe the relationship between height and diameter is a useful indicator to support sustainable forest management.

Recent scientific papers have presented models to estimate stem an total height, merchantable and total volume and above ground biomass in global, continental and regional scales, using large numbers of observations (Feldpausch et al., 2011; Hunter et al., 2013), but these works have not examined independent data for validation (Sileshi, 2014; Ward, 2015), and many such publications reveal conflicting results at different scales, sites and species (Foyolle et al., 2013; Hunter et al., 2013; Kearsley et al., 2013).

Models that include height (h) as independent variable usually provide greater accuracy on estimates of volume, biomass or individual carbon (Marshall et al., 2012; Rutishauser et al., 2013) than models that rely only diameter at breast height (DBH) as independent variable. However, when these models are applied in practice, it is common to use estimated heights rather than measured heights to prepare volume, biomass or carbon estimates. This practice is common, but rarely discussed (Rutishauser et al., 2013), and the effect on predictions and precision is usually overlooked.

A database of forest inventory measurements from the Amazon rainforest was used to examine the following questions: (1) what model best describe the height-diameter relationship of 25 commercial species in the Brazilian Amazon? (2) Does the estimated tree stem height improve prediction of stem wood volume? (3) What exploratory data analysis, model fitting and validation techniques can inform the model selection process? The main emphasis of this paper is to test procedures and analysis techniques for model fitting and validation, and to investigate hypsometric and volumetric models to guide the use of dendrometric relationships for tropical forests.

2. Material and methods

2.1. Study area and data used

This study draws on data from tropical forest in the Eastern Amazon, owned by the Jari Group and administered since 2003 by Jari Florestal S.A.. The area lies within the company's forest reserve of 545,025 hectares of tropical forest, in central-northern portion of the Pará State, on northern channel of the Amazon River, and adjoining the border of the State of Amapá, Brazil (Fig. 1). The climate is classified Ami according to the Köppen classification, and has an annual rainfall of 2234 mm, falling mainly between December and May. The average annual temperature is 25.8 °C with a seasonal variation of ± 2 °C. The vegetation is predominantly evergreen tropical rainforest, on a heavy clayey dystrophic latosol with some red-yellow oxisols (Hiramatsu, 2008).

The data were collected between July 2005 and December 2006 by the forest inventory staff of Jari Florestal S.A. in preparation for harvesting operations. The database contains measurements of 16,099 trees of 25 species harvested from Forest Management Unit (FMU) number 3, a native forest of 3786 ha, one of thirty FMUs used by the company. All harvested trees are merchantable, but six species are particularly important and comprise the value almost 50% of all harvested merchantable volume (Table 1). The company provided data from trees exceeding 140 cm in circumference at breast height (CBH; i.e., 44.6 cm DBH), 1.30 m above the ground or just above buttresses (Table 1). All trees were identified, felled, sectioned and scaled using the Huber method using different size logs as described by Hiramatsu (2008). Merchantable height was measured after the tree was felled, and indicates the total stem length free of branches, defined by the distance from the base of the tree, at ground level, to its crown opening point, including buttresses when present. Ten percent of these data were set aside for validation of the hypsometric equations, as well as for testing of volumetric equations previously developed by Hiramatsu (2008).

The database for validation was selected by constrained random sampling using the combination of DBH size-class and species groups, maintaining a constant 10% sample across each stratum. In the total, 14,489 trees were used to fit the hypsometric models while the remaining 1610 trees were reserved to validate the fitted models. The validation database was also used for testing of volumetric models previously developed, with estimated heights included as the test factor.

2.2. Hypsometric models fitted

We fitted five hypsometric models commonly found in the literature, chosen for their simple mathematical features, extensively documented properties, simple relationships between their coefficients and the structural attributes of the stand, as well as applicability, reproduction and comparison with other studies (Table 2).

All models were adjusted as shown in Table 2. Moreover, all models were calibrated to accommodate species identity, assuming that each species may display a unique hypsometric relationship. This hierarchy of data was modelled using linear mixed-effects models with random effects included for the prediction of height per species. Thus, hypsometric estimates are specific to each group of species. In this case, the general representation of the mixed-effects model is:

$$y_{ij} = \beta_0 + b_{0i} + \left(\beta_1 + b_{1i}\right).x_{ij} + \varepsilon_{ij} \quad (1)$$

Where y_{ij} is the dependant variable j under the effect of the group i; β_0 and β_1 are the fixed factors of analysis, respectively the intercept and slope coefficients of the global relationship; b_{0i} and b_{1i} are random factors that express the effect related to the sample trees within the grouping species; x_{ij} is the independent variable j under the effect of group i, and; ε_{ij} is the model error in estimating the dependant variable. To adjust the model II and III are included in (1) fixed factor β_2, as well as the random factor b_{2i} in the composition of the mixed-effects model. This procedure finds estimators simultaneously for all species and for each group, by incorporating in the established models estimators for random effects that enable hierarchical prediction for the entire forest. These models will be assessed for stand and species level estimates, chosen according to their overall performance.

2.3. Volumetric models tested

The inclusion of the estimated height was tested in seven volumetric models previously fitted by Hiramatsu (2008) with the same data set used in this study (Table 3).

However, in the work of Hiramatsu (2008), individual volumes were modelled only for six species with the highest level of use and commercial value, according to the criteria of Jari Florestal S.A.. The data were divided into seven groups, namely all species, DIEX, DIOD, HYSO, HYSE, MAHU and ROMO, named by their acronyms previously presented in Table 1.

The effect of the estimated height in volume estimates was analysed from the different possible combinations of inclusion the variable in all models presented in Table 3. The models were tested using the database for validation with the following combinations: global volumetric and hypsometric equations (VEH); global volumetric equation and height estimates per species (VHS), volumetric equations for the seven groups of species and global hypsometric equation (VGE), and; volumetric equations for the seven groups of species and hypsometric equation by species (VGS). The validation of the models utilizing observed height as well as the single input models for all species and for the six groups are presented in Hiramatsu (2008), and therefore are not addressed in this work.

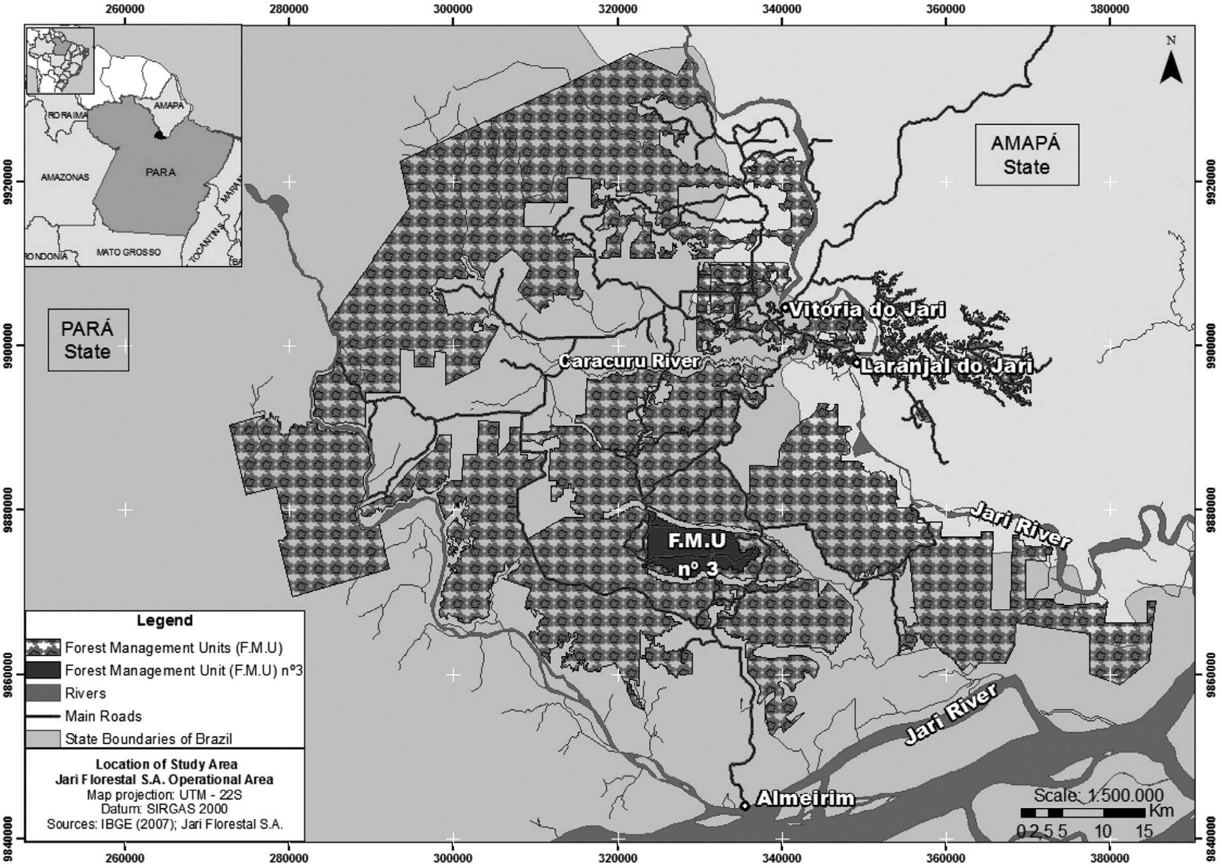

Fig. 1. Location of the study area owned by Jari Florestal S.A., showing the Forest Management Unit (FMU) number 3, in the Eastern Amazon, Pará and Amapá States, Brazil.

Table 1
Distribution of used data (TOT), separated into two databases for model fitting (FIT) and validation (VAL) of the hypsometric relationship and volume estimates of 25 merchantable species in the Eastern Amazon, in Jari Florestal SA, Pará State, Brazil.

SCIENTIFIC NAMES	MERCHANTABLE NAME	GROUP	BASE		
			FIT	VAL	TOT
Bagassa guianensis Aubl.	Tatajuba	BAGU	50	6	56
Bowdichia nitida Spruce ex Benth.	Sucupira-amarela	BONI	173	19	192
Buchenavia parvifolia Ducke	Tanibuca	BUPA	353	39	392
Caryocar glabrum (Aubl.) Pers.	Pequiarana	CAGL	291	32	323
Caryocar villosum (Aubl.) Pers.	Pequiá	CAVI	554	62	616
Dinizia excelsa Ducke	Angelim-vermelho	DIEX*	1196	133	1329
Dipteryx odorata (Aubl.) Willd.	Cumaru	DIOD*	483	53	536
Erisma spp	Cedrinho	ERSP	100	11	111
Eschweilera coriacea (DC.) S.A.Mori	Matamatá	ESCO	60	7	67
Goupia glabra Aubl.	Cupiúba	GOGL	1564	174	1738
Handroanthus serratifolius (Vahl) S.Grose	Ipê	HASE	73	8	81
Hymenaea courbaril L.	Jatobá	HYCO*	226	25	251
Hymenolobium sericeum Ducke	Angelim-pedra	HYSE*	211	24	235
Manilkara huberi (Ducke) A.Chev.	Maçaranduba	MAHU*	4192	465	4657
Mezilaurus lindaviana Schwacke and Mez	Itaúba	MELI	71	8	79
Pouteria elegans (A.DC.) Baehni	Guajará	POEL	50	6	56
Pseudopiptadenia psilostachya (DC.) G.P.Lewis and M.P.Lima	Timborana	PSPS	311	34	345
Qualea paraensis Ducke	Mandioqueira-escamosa	QUPA	1774	197	1971
Roupala Montana Aubl.	Louro-faia	ROMO*	656	73	729
Ruizterania albiflora (Warm.) Marc.-Berti	Mandioqueira-lisa	RUAL	325	36	361
Staminodianthus racemosus (Hoehne) D.B.O.S.Cardoso and H.C.Lima	Sucupira-preta	STRA	51	6	57
Tachigali melanocarpa (Ducke) van der Werff	Taxi-vermelho	TAME	42	5	47
Tachigali spp	Taxi	TASP	163	18	181
Vouacapoua americana Aubl.	Acapu	VOAM	1471	163	1634
Vochysia vismiifolia Spruce ex Warm.	Quaruba	VOVI	49	6	55
*Legend: * Species with higher degree of utilization and commercial value.* **TOTAL**			**14,489**	**1610**	**16,099**

Table 2
Equations used for modelling the hypsometric relationship of 25 species commercialized by Jari Forest SA in the Eastern Amazon, State of Pará, Brazil.

N°	NATURE AND BEHAVIOR	MODEL	ADJUSTMENT FORMAT
I	Arithmetic / Curvilinear	$h = b_0 + b_1 Ln(DBH)$	
II	Arithmetic / Parabolic	$h = b_0 + b_1 DBH + b_2 DBH^2$	
III	Arithmetic / Parabolic	$h = \frac{DBH^2}{b_0+b_1 DBH+b_2 DBH^2}$	$\frac{DBH^2}{h} = b_0 + b_1 DBH + b_2 DBH^2$
IV	Exponential / Curvilinear	$h = e^{[b_0+b_1(\frac{1}{DBH})]}$	$Ln(h) = b_0 + b_1(\frac{1}{DBH})$
V	Exponential / Curvilinear	$h = b_0 DBH^{b_1}$	$Ln(h) = Ln(b_0) + b_1 Ln(DBH)$

Table 3
Volumetric models with one (ALL_D) and two (ALL_DH) variables adjusted by Hiramatsu (2008) for all species and six specific species (DIEX, DIOD, HYCO, HYSE, MAHU and ROMO) with higher degree of utilization and commercial value exploited by Jari Florestal S.A., State of Pará, Brazil.

GROUP	EQUATION	R^2	$Syx_{\%}$
ALL_D	$v = -0,367921 + 0,0013446 * DBH^2$	0,721	38,58
ALL_DH	$v = -1,12194 + 0,0327452 * DBH + 0,0000494 * (DBH^2)h$	0,807	32,08
DIEX	$v = -1,02799 + 0,00000303694 * DBH^3 + 0,000105626 * (DBH^2)h - 0,00000043332 * (DBH^3)h$	0,725	30,47
DIOD	$v = -4,7945 + 0,1413018 * DBH - 0,000596 * DBH^2 + 0,0000394 * (DBH^2)h$	0,782	24,01
HYCO	$v = 1,353158 + 0,0002609 * DBH^2 + 0,0000388 * (DBH^2)h$	0,828	18,48
HYSE	$v = -9,70551 + 0,2495734 * DBH - 0,001283 * DBH^2 + 0,0000606 * (DBH^2)h$	0,771	30,98
MAHU	$v = -0,18228 + 0,00000287539 * DBH^3 + 0,0000780196 * (DBH^2)h - 0,000000370194 * (DBH^3)h$	0,749	23,50
ROMO	$v = -6,0988 + 0,175602 * DBH - 0,000963 * DBH^2 + 0,0000477 * (DBH^2)h$	0,804	18,83

2.4. *Verification and validation of all models*

The statistics used for the linear regression of the hypsometric relationship refer to the combined prediction performance of each model adjusted for the 25 species under analysis. Initial comparisons of hypsometric models were based on common statistical criteria: coefficient of determination (R^2) and standard error of estimate in percentage ($Syx_{\%}$), both recalculated for the variable of interest (Machado et al., 2008); correction factor for the logarithmic discrepancy (CF) applied on the logarithmic models estimates, and; the Akaike Information Criterion (AIC). Further analyses were based on graphical comparisons of residual dispersion, normality of error distribution, homoscedasticity of the error distribution, as well as analysis of outlier observations. Outliers in the hypsometric relationship were determined using the following fit statistics: h-values = leverage of each individual observation; DFFITS = the relevance of each observation on the model adjustment process, and Cook's Distance = impact of each observation on the estimated regression coefficients (Hoaglin and Welsch, 1978; Li and Valliant, 2011).

Outlying observations were discarded when they exceeded all the exclusion criteria (Li and Valliant, 2011). In addition, we used the Bayesian Information Criterion (BIC) to compare and select the best fit amongst the adjusted models (Aho et al., 2014). Validation of hypsometric models and combined estimated height and volumetric models relied on graphical analysis of the relationship between the observed and estimated (Piñeiro et al., 2008); Theil's Error Decomposition (Smith and Rose, 1995); Graybill's test (Leite and Oliveira, 2002); Efficiency (Ward, 2015); Equivalence test (Robinson and Froese, 2004), and; Tukey test to compare the average volume estimated by species per each combination of hypsometric and volumetric models (Mayer and Butler, 1993).

Comparisons between observed (OBS) and estimated values (EST) by graphical analysis, Theil's and Graybill's test, were carried out to identify how close the relationship between OBS and EST is to the reference line which indicates perfect fit between these two variables (OBS = EST). The Graybill's test evaluates simultaneously $\beta' = [b0, b1]$ is equal to [0,1] using the F test (Leite and Oliveira, 2002). Moreover, Theil's test assesses whether each component of the sum of squared residuals (SSR), are significant by F test, identifying: model lack-of-fit; bias ($b_0 = 0$); consistency ($b_1 = 1$), and; presence of non-linear deviations not incorporated in the regression (Smith and Rose, 1995). The Equivalence Test uses the dissimilarity as the null hypothesis in the two-tailed t-test, in which the average of the differences between the estimated and observed values is different from zero (Robinson and Froese, 2004). The test's innovation is based on the choice of an indifference region, a percentage of the standard deviation of the residuals, where the average residual can be neglected when it is less than the established prior (Robinson and Froese, 2004). The difference between the Equivalence Test and the t-test is the definition of equivalence region where the predicted values will be considered equal to the observed when the average difference is less than a pre-established error limit (ε). In this study, for all models tested, we used an alpha (α) equal to 1% for t distribution, and epsilon (ε) of 0.25, as suggested by Robinson and Froese (2004).

The Tukey test was applied to compare the volumetric average estimated per species by: VEH; VHS; VGE; VGS; estimated volume only by DBH (VES), and; volumes observed per species (VOB). This analysis enabled the identification of combinations which are statistically similar to the average volume per species. As suggested by Hans and Oderwald (1993), the choice of the best model was performed after the comparison of the model behaviour in different statistical criteria used in each step of adjustment and completing validation procedures. These analyses took into account the behaviour of the model outside and inside the fit interval (analysis only for the tested hypsometric models) as well as the utility of the model and the ability to integrate the model into a system of volumetric estimates. All graphical analysis, verification and validation statistics were performed using the free computer system R (Core Team, 2015), where the scripts and how to approach technics were adapted and replicated from Robinson and Hamann (2011).

3. Results and discussion

3.1. *Adjustment of hypsometric models*

In Fig. 2, it is possible to identify the difference in the quality of the hypsometric models after undergoing the analysis of outliers and modelling per species using mixed-effects model. The adoption of models for prediction at different levels provides a significant gain in accuracy of all the models. With the hierarchical prediction approach, the average error of the estimates of the models decreased by 12.71%, giving an average increase of the 527.75% in the R^2 for all models, and a decrease of the 4.62% on the AIC criteria.

The exclusion of outliers provided a varied percentage of gain amongst the models, as well as being positive for all forms of fit with

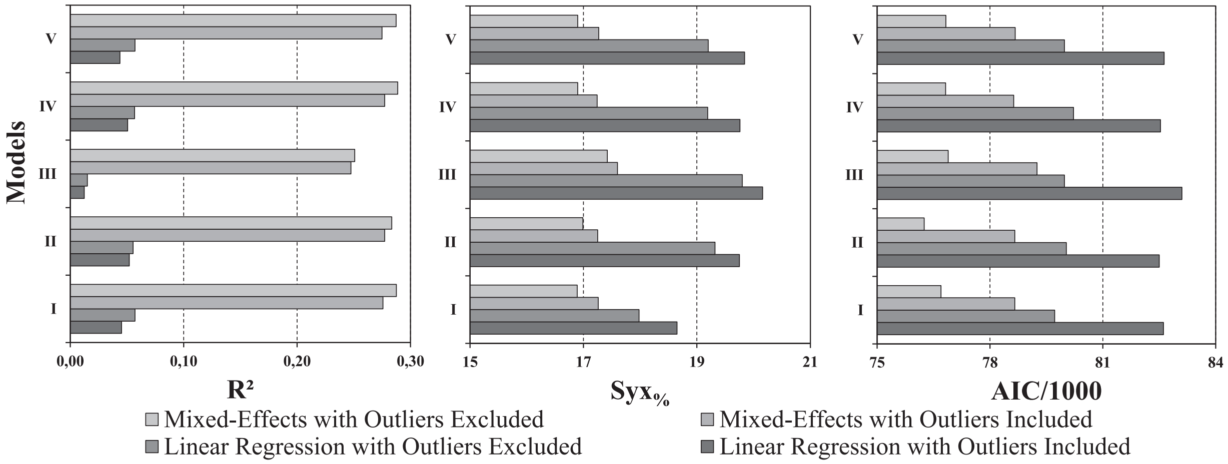

Fig. 2. Statistics of means fit and accuracy of all simple linear regression and mixed-effects model tested for description of the hypsometric relationship of 25 Amazon species, presented for all combinations of model type and outliers' presence. The statistics presented for the linear regression refer to the combined prediction performance of each model for the 25 species under analysis.

respect to all the statistical criteria. Excluding atypical observations provides a significant improvement in the estimate of R^2, contributing an average gain of the 9.52%. Moreover, for the other statistics, Syx% and AIC, these gains were 2.31% and 3.03% respectively. This expressive gain in R^2 becomes more evident when the results are analysed separately. The average percentage gain was 15.86% for all linear regression models as well as 3.19% for the mixed-effects model. These results are expressive because show that the effect of the exploratory data analysis benefits R^2 not affecting at the same proportion the variability measures or the model quality.

The variable performance amongst models is due to their different mathematical nature. Models II and III, respectively known by Trorey (1932) and Prodan (1944) works, present a problematic behaviour eventually. Both models can generate parabolic curves that do not describe the monotonic increase expected in biological relationships between height and diameter. Depending on the data, the curve estimated with these models may include a maximum, with the result that any increase in the independent variable produces a decrease in the dependant variable. This aspect is rarely observed in natural biological relationships, whether in plantations or mixed forests, consequently something undesirable in the description of the hypsometric relationship.

Despite the increase in R^2 after excluding outliers, the model III does not behave coherently as height - diameter relationship theory. The transformation on the dependant variable of the model III does not provide a gain in accuracy and results in a suboptimal fit (Fig. 2). These results were similar to the model II performance. Despite good statistics of fit and accuracy, the model II does not describe the relationship h-d correctly in the graphical analysis, either within or outside of the fit interval for all species (Fig. 4). Models I, IV and V have similar descriptive characteristics. Due to the similar mathematical behaviour and nature, the statistics of fit and precision for them presented similar values, regardless of the form of data approach (Fig. 2). These equations were known in forest science as Henricksen (1950), Curtis (1967) and Stoffels and Van Soest (1953) models, due to ease fit characteristics, mathematical simplicity and high degree of correlation of its coefficients with the stand characteristics (Fang and Bailey, 1998; Willmott et al., 2006; Machado et al., 2008; Marshall et al., 2012; Vibrans et al., 2015).

These models are widely used in the description of the hypsometric relationship of natural forests, particularly the model V, which in recent decades has been used by several works aimed at the description of tropical forests (Feldpausch et al., 2011; Kearsley et al., 2013; Rutishauser et al., 2013; Sileshi, 2014; Hunter et al., 2013). Despite being a mathematically advantageous model, model V should not be preferred over others, regardless of the relationship under analyses. Corroborating Picard et al. (2015), it is clear that the model V does not always have the best performance and therefore should not be the only option for modelling the hypsometric relationship. The model I had the best performance amongst the other, either with or without the inclusion of the random effects for the hierarchical modelling. Model I presents the best fit and accuracy statistics of R^2, Syx% and AIC in comparison to the model recommended as ideal by Sileshi (2014), the model V. Hunter et al. (2013) discarded the model I due to an undesired behaviour for description of the rainforest in northern Brazil, but nevertheless this result were not achieved for the forest under analysis, as well as the model I presented the best statistics of fit and accuracy for description of h-d relationship.

After excluding outliers, remaining observations (n) for fitting were not the same for all tested models, due to the mathematical difference in behaviour of the evaluated models. These differences directly affect all commonly used statistics for comparison the performance of fitted models as well as the leverage of each observation on the model fitting process. According to Aho et al. (2014), the BIC usage is preferable to AIC when the number of observations and the number of coefficients of the tested models are not the same at the moment of fit comparison. Model I had the best performance amongst the other when it includes the random effects and without the presence of outliers for describe the h-d relationship (BIC = 76,601.32). This result confirms the general trend of the values found in the other set of statistics in Fig. 2. However, when the atypical values are not excluded, the mixed model II had the best performance amongst the other models (BIC = 78,661.96). This result shows that the influence of outliers on the database can favour models that do not necessarily describe the relationship under analysis as well as the identification and exclusion of outliers provides greater support on models selection procedures (Hoaglin and Welsch, 1978; Li and Valliant, 2011).

In Fig. 3, it is possible to see the difference of model I performance between the two contrasting settings of data approach: simple linear regression, adjusted using raw data, and hierarchical prediction model by mixed-effects approach, including the influence of the species and the exclusion of outliers in the database used for fit process. Fig. 3 illustrates how the performance of the model changes with the approach used to fit the model. The range of values estimated by the model increased substantially after including species identity in the model, while heteroscedasticity of the fit decreased, as can be seen in Fig. 3. The effect of excluding atypical observations and adopting a more complex model is barely noticeable in the standard error statistics, but it is noticeable in Fig. 3 as an improvement in the leverage distributions and greater

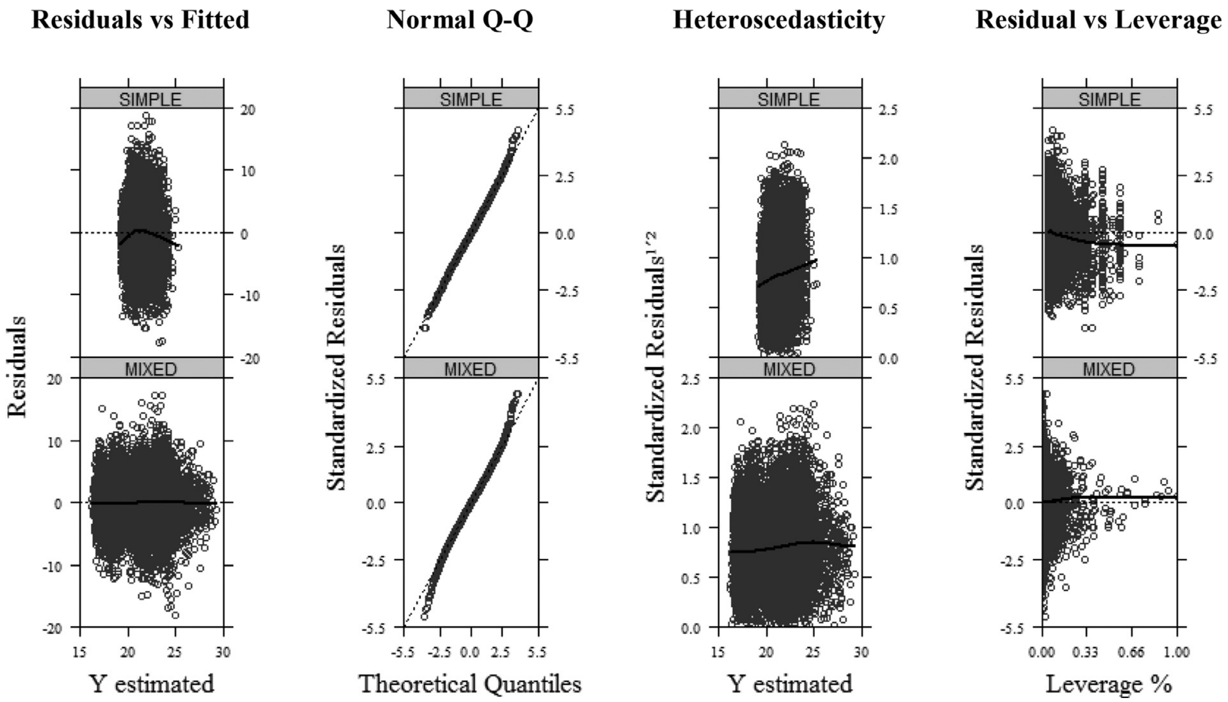

Fig. 3. Residuals distribution, Standardized Residuals Distribution, Heteroscedasticity test and Leverage relationship for two different approaches to fit the hypsometric relationship for 25 species in Eastern Amazon, Pará State, Brazil. Top row shows simple general linear regression; Bottom row shows mixed-effects approach excluding outliers.

Table 4

Statistics of fit and accuracy of the mixed-effects model I adjusted to predict height of 25 species from Amazon Forest situated at Jari Florestal S.A., Pará State, Brazil. Significant terms are bold ($p < 0,05$).

FIXED EFFECTS	STAND LEVEL					
	Estimate	Standard Error	t-value	p-value	ESTIMATES	
					Lower	Upper
β_0	**3,5668**	1,7810	2,0030	0,0075	–0,2293	7,1553
β_1	**4,2033**	0,4326	9,7150	0,0000	3,3506	5,1309

RANDOM EFFECTS	GROUP LEVEL		
	Lower	Estimate	Upper
Standard Desviation (b_{0i})	3,2844	6,3920	10,6409
Standard Desviation (b_{1i})	0,9168	1,5980	2,5470
Correlation (b_{0i},b_{1i})	–0,9783	–0,9386	–0,7729

COEFFICIENTS FOR ALL SUB LEVELS

GROUP	$\beta_0 + b_{0i}$	$\beta_1 + b_{1i}$	GROUP	$\beta_0 + b_{0i}$	$\beta_1 + b_{1i}$
BAGU	3,0471	**4,4693**	**MAHU**	-0,3359	**5,3248**
BONI	4,5774	**4,6317**	**MELI**	0,0160	**4,4992**
BUPA	3,2220	**4,4236**	**POEL**	4,4533	**3,7662**
CAGL	9,1555	1,8443	**PSPS**	-3,2458	**5,9782**
CAVI	11,7314	1,1764	**QUPA**	8,3207	**2,9184**
DIEX	6,6772	**3,5288**	**ROMO**	6,6884	**4,1285**
DIOD	7,8286	2,5566	**RUAL**	6,1348	**3,5367**
ERSP	2,3432	**4,5030**	**STRA**	2,7567	**4,6541**
ESCO	-1,5715	**5,8813**	**TAME**	-5,5852	**5,9428**
GOGL	-2,2857	**4,9778**	**TASP**	7,4620	**3,2048**
HASE	7,7206	**4,3341**	**VOAM**	-3,2084	**5,3514**
HYCO	8,1698	**4,2277**	**VOVI**	4,7899	**4,2709**
HYSE	0,3076	**4,9506**			

symmetry of the spread of observed leverage in the mixed-effects model approach.

The model I fit statistics, in a mixed-effects form of prediction, are presented in Table 4. The results presented corroborate with Fig. 2, and underline the quality of the model when incorporated the species effect. The model presented β_0 and β_1 significant for the fixed effects (global level coefficients) as well as estimates curves which are different for each species group, presenting varied slopes and intercepts due to flexibile nature of the model.

It is expected that the b_1 coefficient of this model has an important descriptive power, since the modification of the association range between the height and diameter for a semi-logarithmic scale is consistent with the nature of the hypsometric relationship of mature forests or multiple species and ages (Machado et al., 2008). It is important that in the proposition of models to describe biological relationships, the coefficients of the model present interpretable values (Fang and Bailey, 1998; Feldpausch et al., 2011). Model I has important features such as ease fit, mathematical flexibility as well as simplicity and descriptive power of its monomials. However, its use is deprecated against model V (Hunter et al., 2013), although it has a similar descriptive behaviour and avoids the logarithmic discrepancy in predictions.

In mature plantations and natural forests of mixed species and diverse ages, the hypsometric relationship commonly underperforms with respect to statistics of fit and precision, especially for R^2 and p-value of the β_1 coefficient (Machado et al., 2008). This is due to the greater variation in diameter compared to height in these site conditions, which in turn results in a low coefficient of covariation between height and diameter, hence a low R^2. Under these conditions of stands, the p-value for b_1 is usually not significant due to the parallelism of the curve estimated in relation to the x-axis, thus making $\beta_0 \cong \bar{h}$ e $\beta_1 \cong 0$.

This aspect is not observed for the global hypsometric relationship (Table 4). However, the effect appears only for some sub-levels of the model, specifically for CAGL, CAVI and DIOD species. The current condition of the stand structure can be the cause of this discrepancy in the height-diameter relationship for these three species, since the model was efficient when compared to other tests (Fig. 2) as well as for the other species under analysis (Fig. 3 and Table 4). This result is due to the structural and ecological characteristics of these species within the forest stand, as a result of the greater variability in ages and canopy conditions that they have developed. Since these are long-lifetime pioneer species (Pinheiro et al., 2007), they tend to exhibit great variation in spatial and size distribution in primary forests, with a weak hypsometric relationship under these conditions.

The mixed-effects model fitting methods are based on the Maximum Likelihood (ML) or Residual Maximum Likelihood (REML), unlike the linear regression based on the least squares method. This aspect enables fitting of models in unbalanced data in multiple groups, which have total or partial intersections of random factors over the subgroups. These features are significant for analysis of large databases, with extensive structural variation in sub-levels, such as forest inventory data (Robinson and Hamann, 2011). Under these conditions p-value test for the sub coefficients may indicate an effect in a sublevel that was not stratified, this indicating that there is a component of variation not explained by the proposed hierarchy and possibly connected to different aspects influencing the hypsometric relationship. This result does not invalidate the model, but indicates that improvements are possible, both in the model approach as well as in the database used (Nuzzo, 2014).

Tree height tends to reach its asymptotic maximum value before the biological maximum diameter for all species (Machado et al., 2008). The total height is little affected by the stand density (Machado et al., 2008; Vibrans et al., 2015) and therefore its variation depends directly of the site quality (Trorey, 1932; Willmott et al., 2006; Marshall et al., 2012), silvicultural treatments (Willmott et al., 2006; Sharma and Parton, 2007; Rutishauser et al., 2013), species (Fang and Bailey, 1998; Cole and Ewel, 2006; Fayolle et al., 2013), and age (Trorey, 1932; Curtis, 1967). Conversely, the diameter undergoes significant effect of the stand density and other mentioned factors, presenting higher variation on density distribution compared with height variable in the same age.

In terms of merchantable height, i.e. length of free of branches stem defined by the height to the base of the crown, the covariation between diameter and this height has a larger discrepancy, causing a significant loss of quality for this dendrometric relationship. This is due to the strong relationship between the crown opening height and density of the population, as well as the effect of ecological characteristics of the species in the forest canopy. These aspects are widely discerned in the database used mainly for CAGL, CAVI and DIOD species.

3.2. *Validation of the hypsometric models*

The model II and III presented the worst performance of the h-d relationship, whether in or out of data interval used for fit and validation (Fig. 4). In Fig. 4 are present: the behaviour of all tested models for the estimated fixed effects coefficients (global relationship) on the validation data; relationship between the observed (OBS) and estimated height (EST) for model I using a simple linear relationship, and; relationship between OBS vs EST incorporating the mixed-effects in the model I. The negative highlight is the model II which describes a parable that does not reflect the biological expectation of hypsometric relationship. The models I, IV and V showed similar behaviour for all estimate's curves. These models are always featured in hypsometric relationship description for mature forests or oldest stands, due to the mathematical behaviour which describes well the relationship between height and diameter for these specific site conditions (Willmott et al., 2006; Machado et al., 2008; Lima et al., 2012; Vibrans et al., 2015).

Even presenting non-significant values for Graybill's tests, Theil's decomposition, as well as statistical similarity on the Equivalence Test (Table 5), there is an expressive difference on the predictive behaviour between model I as a simple linear regression (HEL) and fitted as a mixed-effects model (HGR). In Fig. 4, the mean line between the observed and estimated height compared to 1:1 line clearly presents the superiority of the model fitted as mixed-effects model. In the Table 5 the values related to estimative efficiency of adjusted models on the validation database corroborates the results presented in Fig. 4. The expressive difference between HEL and HGR models is verified by the increase of the efficiency in 514,63% on the height estimates by HGR model. The Theil's decomposition of error referring to the components lack-of-fit, consistency and bias, shows a highly non-significant result for both models. However, for the HEL, the calculated F-test at 5% level denotes a significant regression lack-of-fit for this model. This result indicates the existence of a nonlinear component which is not described by HEL in its fullness, aspect widely contemplated by HGR model.

The Equivalence Test presented a positive result for the hypsometric models. The results indicate that both models are valid for describing the relationship h-d, providing estimated volumes within the rejection region ($\varepsilon = 0.25$; $\alpha = 1\%$). The motivation of this test differs from the others, since their use is not only to evaluate the statistical similarity between the mean observed and estimated, but in fact, evaluate the null hypothesis that the model does not reach the required accuracy standards (Robinson and Froese, 2004).

It is possible to see in the Fig. 5, that model I estimated curves on the scatter plot of each species which fits well on the validation data throughout its length. Regardless of the species under analysis, the behaviour of the estimated curve out of the fit interval proved consistent with the hypsometric relationship for even aged and mixed species forests. Moreover, all hypsometric curves are less steep as expected for stands with multiple ages and species, or growing in average or low quality sites, as is the case of these forests in Eastern Amazon (Machado et al., 2008; Lima et al., 2012; Vibrans et al., 2015).

The behaviour of the model out of its fit interval is an important feature for validation of models aimed to description of biological relationships (Mayer and Butler, 1993; Oderwald and Hans, 1993). The HGR model shows consistent behaviour inside and outside the fit interval of the data. This aspect indicates a model that can eventually be applied as an auxiliary tool in the decision-making process out of its domain of action. However, its use in these situations should be based on practical knowledge of the user and carried out only in the absence of specific equations for particular situations.

3.3. *Validation of the volumetric models*

The inclusion of the estimated height in the pre-developed volumetric models, provided a varied behaviour in all validation statistics tested (Table 5). The Theil' test indicates: lack-of-fit of all models on the ob-

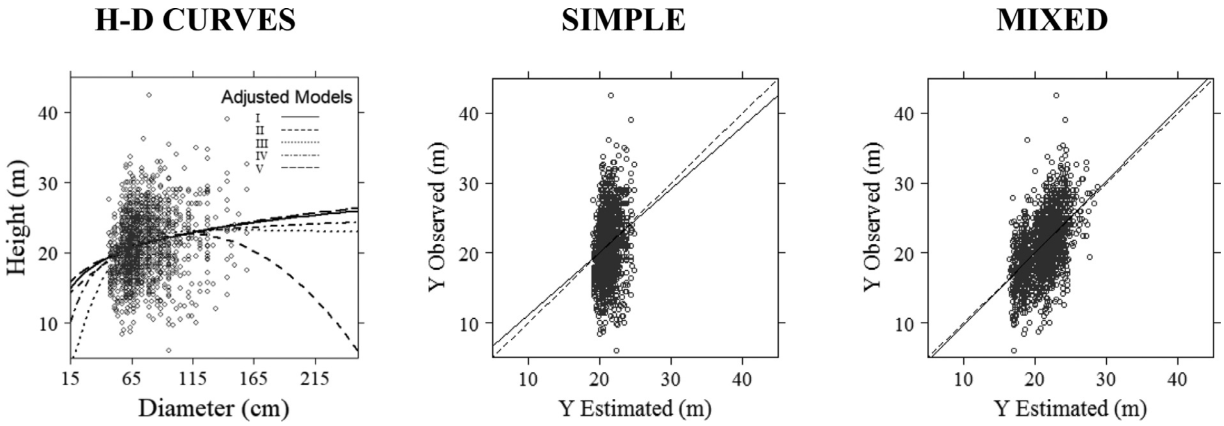

Fig. 4. All five adjusted curves for description of the hypsometric relationship of 25 Amazon trees species situated in Para State, Brazil, as well as the relationship between observed and estimated height for validation the model I, using a linear regression (SIMPLE) and a mixed-effects model (MIXED).

Table 5
Graybill's test, Theil's Decomposition and Efficiency of the models proposed for the hypsometric relationship and volume prediction for Jari Florestal S.A. forest situated in Pará State, Brazil.

MODEL	GRAYBILL's TEST	THEIL's DECOMPOSITION MODEL LACK OF FIT	THEIL's DECOMPOSITION NO BIAS ($b_0 = 0$)	THEIL's DECOMPOSITION CONSISTENCY ($b_1 = 1$)	THEIL's DECOMPOSITION REGRESSION LACK OF FIT	EF	DISSIMILARITY
HEL	0,5320^{ns}	1,3049^{ns}	0,0074^{ns}	1,0565^{ns}	1,3261*	0,0451	Rejected
HGR	0,3968^{ns}	1,0331^{ns}	0,3255^{ns}	0,4683^{ns}	1,0362^{ns}	0,2772	Rejected
VES	2,3538^{ns}	2,2027**	4,7031*	0,0073^{ns}	2,1950**	0,7038	Rejected
VEH	8,4933**	2,3995**	11,2825**	5,6712*	2,2174**	0,7012	Rejected
VHS	4,0796*	1,9806**	8,1593**	0,0049^{ns}	1,9653**	0,7279	Rejected
VGE	22,8970**	3,1643**	25,6609**	19,8328**	2,9346**	0,7281	Rejected
VGS	12,7157**	1,4560**	21,5051**	3,8876*	1,3943**	0,7568	Rejected

Legend: ns = non-significant difference on F-test; * = significant at 5% level on F-test; ** = significant at 1% level on F-test; All Dissimilarity tests were at $\alpha = 0.01$ and $\varepsilon = 0.25$; EF = Efficiency; HEL = estimated height by linear regression; HGR = estimated height by mixed-effects model; VES = volumetric estimates were made from only DBH; VEH =: global volumetric and hypsometric equations; VHS = global volumetric equation and height estimates per species; VGE = volumetric equations for the seven groups of species and global hypsometric equation, and; VGS = volumetric equations for the seven groups of species and hypsometric equation by species.

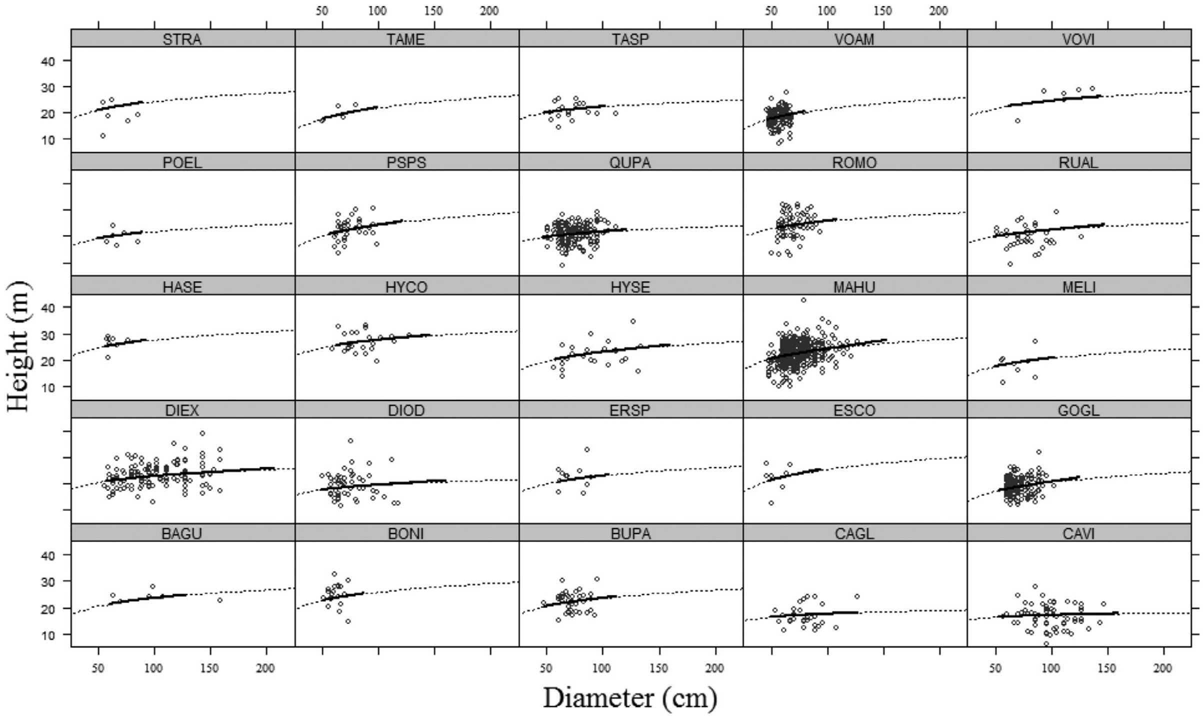

Fig. 5. Height-diameter curves of the 25 commercial species, according to mixed-effects model I, managed by Jari Florestal S.A situated in the Eastern Amazon Forest in Para State, Brazil. The dark line represents fit interval of the model I, and the dotted lines represent the prediction tendency of the model when used outside the fit interval.

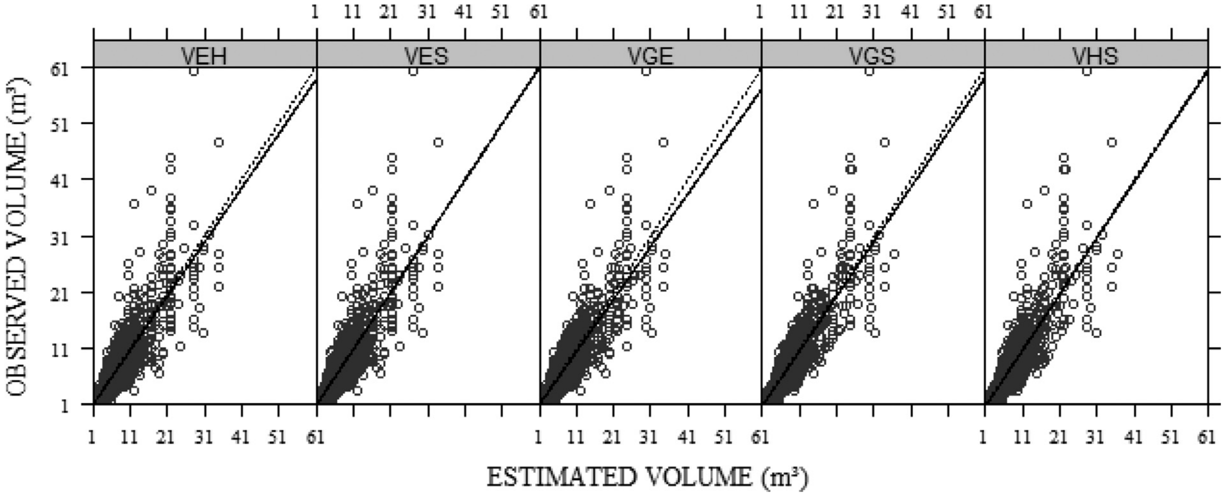

Fig. 6. Relationship between the observed and estimated values for individual variable volume of 25 commercial species from Eastern Amazon present at Jari Florestal S.A., Para State, Brazil. The volumetric estimates were made from only DBH (VES), and different combinations of models under analysis: global volumetric and hypsometric equations (VEH); global volumetric equation and height estimates per species (VHS), volumetric equations for the seven groups of species and global hypsometric equation (VGE), and; volumetric equations for the seven groups of species and hypsometric equation by species (VGS).

served data, the intercept and the slope differ from 0 (for all models) and 1 (except VEH and VHS), respectively, and there are non-linear deviations in the relationship between the volumetric models and the data used for all combinations.

The VHS combination presented significant values at 5% level in Graybill test, and significant result at 1% level for other combinations. The best performances in this criterion were VES and VHS combinations. These two ways of prediction employ a general trend of the predicted mean volume by diameter (VES) as well as a general mean trend of height to predict the volume (VHS). Due to the expressive amount of data used to establish these trends (without random effects for prediction per species), the consistency of the prediction for VES and VHS combinations provided respectively non-significant and a significant result at 5% level in Graybill's test.

Due to the absence of sub-relationships for volumetric prediction per species, the estimated of VES and VHS combinations are relatively less efficient compared to the other prediction's combinations (Table 5). The combinations of higher efficiency are VGE and VGS, both explained respectively 72.81% and 75.68% of observed volume variation of the validation database. Corroborating with the results presented in Table 5, the Fig. 6 indicates that the estimates made by the combination VES, VHS and VGS provided the best graphics results. The line which indicates the average trend between the observed and estimated volumes is slightly distant from the 1:1 line for all models that include the estimated height (Fig. 6). However none of the combinations provided statistical tendency higher than that obtained by VES model (Table 5).

These conflicting results, between the estimate efficiency as well as Graybill's and Theil's tests, provided doubts about the quality of all models when the estimated height is included for prediction of individual volume. These tests used allow evaluation of model's accuracy; however, the acceptance process of the tested model is reduced a binary procedure which does not include a criteria that express the modeller required accuracy (Robison and Froese, 2004). The absence of a criteria of accuracy in these tests may be an inappropriate aspect for validation process, since the model's efficiency is closely linked to the cost of producing it as well as the variability of the data which it purport to describe. This last aspect directly affects the precision of the analysis of validation. It was found that all prediction models meet the required standard by the Equivalency Test (Table 5).

All models presented an estimated average volume statistically equivalent to the observed value, i.e. within the critical region of dissimilarity hypothesis. It is possible to analyse in the Fig. 7, the performance of all combinations between volumetric and hypsometric equations for prediction each sublevel via Tukey test. The combination which presents the best performance is the VGS, where the average estimated volume for each species is statistically equal to the observed for all species.

All other combinations present significant difference on volumetric prediction for VOAM, QUPA, RUAL, MAHU, DIEX, GOGL, CAGL and CAVI simultaneously (Fig. 7). The inclusion of the estimated height by HEL, or volume prediction only by DBH, did not provide a good performance of the VEH and VES respectively for the mentioned species. However, VHS combination failed to estimate the average volume for DIEX, presenting an estimated lower than the observed average. The estimate of the mean volume by VGE was above the observed value for the species CAVI, providing a statistical similarity to the VEH and VES performance on volume prediction to this species. The effect of estimated height by species is significant for the prediction of the average volume per species when it is used the volumetric models of the DIEX, DIOD, HYSO, HYSE, MAHU and ROMO species. This outcome, associated with Equivalence Test result as well as the visual analysis present in Fig. 6, indicates a good performance of the VGS combination for global and per species predictions. Conversely, these good results were not obtained in the Graybill's test and Theil's Error Decomposition test. The best performing model is the one which present the highest biological coherence (Fang and Bailey, 1998; Machado et al., 2008), good predictive quality and extrapolation results (Oderwald and Hans, 1993; Smith and Rose, 1995; Robinson and Froese, 2004). Linking hypsometric models with models to estimate the biomass of tropical forests have shown conflicting and unreliable results in several recent reviews (Hunter et al., 2013; Fayolle et al., 2013; Kearsley et al., 2013; Ward, 2015). These publications show errors that exceed 15%, reaching values above 50% for estimates per unit of area. Moreover, these works present a widely varied results and highly dependency of the hypsometric equation used (Hunter et al., 2013; Kearsley et al., 2013). The average volume estimate for the validation database does not exceed 4.45% using VGS combination. However, the most accurate combination in the conducted analyses was the VHS prediction, which presented 2.54% of error in the global average volume forecast.

The use of a part of the database to assess the predictive efficiency of the model is a simple assessment procedure and suitable for the validation of statistical fits (Nuzzo, 2014; Sileshi, 2014; Ward, 2015). Verification of bias and deviations (Leite and Oliveira, 2002; Piñeiro et al., 2008), behaviour out of the fit data interval (Mayer and Butler, 1993; Oderwald and Hans, 1993), lack of fit (Smith and Rose, 1995;

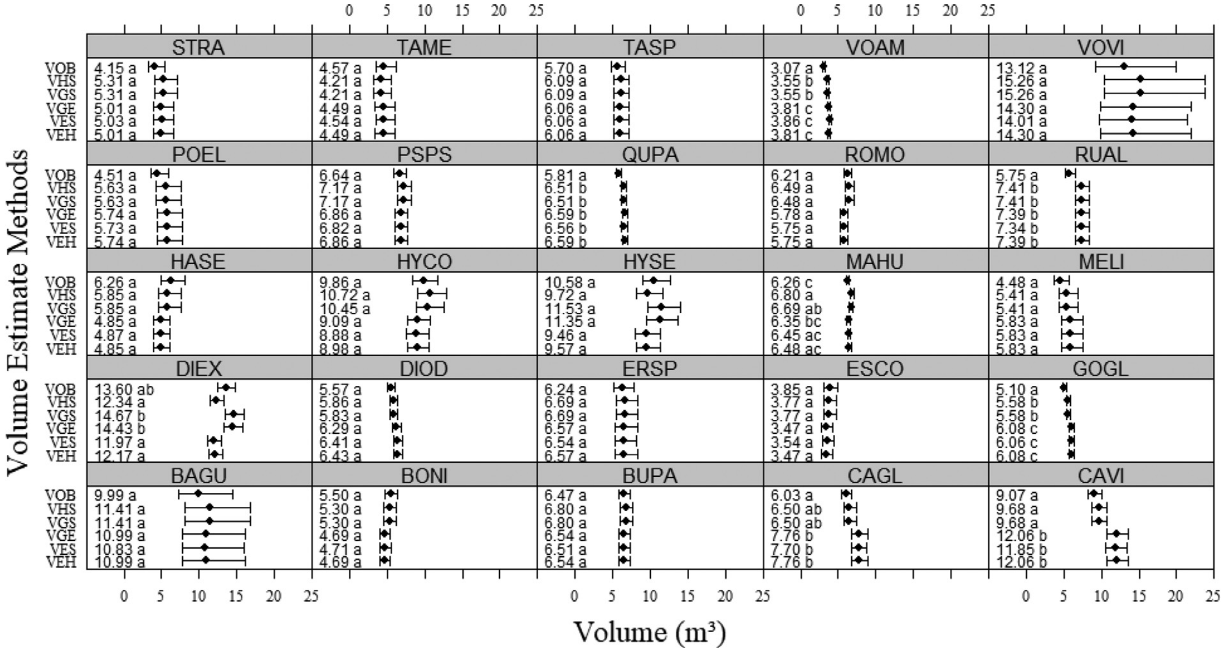

Fig. 7. Comparison of the average observed volume per species (VOB), and estimated only by DBH (VES), VEH, VHS, VGE and VGS combinations for 25 commercial species of Eastern Amazon situated Jari Florestal S.A., Pará State, Brazil. The average values which present the same letter do not differ statistically by Tukey test at 5% level of probability.

Sileshi, 2014), as well as its use in combination with other models (Robinson and Froese, 2004; Ward, 2015), are verification and validation procedures that support the selection of quality models.

Conclusions

The treatment of the database before fit the models (analysis of atypical observations), as well as post-fit the models (validation tests using an extra database), produced more reliable results and greater basement to select a hypsometric model and the best combination of hypsometric and volumetric models. The results show that the exploratory data analysis and the validation of all fits provide final estimates with greater efficiency and predictive utility compared to the behaviour of the observed data.

The most reliable hypsometric relationships for commercial tree species in the Eastern Amazon are those that include species-specific parameters. In general, simple linear regression offered models with inferior performance to those established by hierarchical modelling (mixed-effects model), especially for model I with random effects included. This approach provided the best performance in estimating the merchantable height for all species under analysis.

The estimated height by the mixed-effect model I should be included in models of double input previously developed for volume prediction. The prediction of height associated with volumetric models for the species DIEX, DIOD, HYSO, HYSE, MAHU and ROMO provided volumetric estimates with an error below 5% for the average volume. The validation tests indicate: the VGS combination provides an estimated average volume equal to the observed average for all species in the validation database; statistical equivalence of the VGS model prediction to the observed data in relation to the global average volume, and; the VGS model presents the greater statistical efficiency amongst other combinations of hypsometric and volumetric equations tested.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This contribution is part of a Doctoral dissertation at the Federal University of Paraná (UFPR), which was supported by a fellowship from the Brazilian National Research Council (CNPq). We thank Augusto Praxedes Neto, Luciana Di Paula and Jari Florestal S.A., who provided the data on which this work is based as well as Timni Vieira, Walmes Marques Zeviani and Décio José de Figueiredo for their valuable suggestions in the map design and statistical analysis with R program.

References

Aho, K., Derryberry, D., Peterson, T., 2014. Model selection for ecologists: the worldviews of AIC and BIC. Ecol. Wash. 95 (3), 631–636. doi:10.1890/13-1452.1.

Cole, T.G., Ewel, J.J., 2006. Allometric equations for four valuable tropical tree species. For. Ecol. Manage. Amsterdam 229, 351–360. doi:10.1016/j.foreco.2006.04.017.

Curtis, R.O., 1967. Height-diameter and height-diamater-age equations for second-growth Douglas-Fir. For. Sci. Bethesda v.13 (4), 365–375. doi:10.1093/forestscience/13.4.365.

Fang, Z., Bailey, R.L., 1998. Height-diameter models for tropical forests on Hainan Island in Southern China. For. Ecol. Manag. Amsterdam 110, 315–327. doi:10.1016/S0378-1127(98)00297-7.

Feldpausch, T.R., Banin, L., Phillips, O.L., Baker, T.R., Lewis, S.L., Quesada, C.A., Affum-Baffoe, K., Arets, E.J.M.M., Berry, N.J., Bird, M., Brondizio, E.S., Camargo, P.de, Chave, J., Djagbletey, G., Domingues, T.F., Drescher, M., Fearnside, P.M., França, M.B., Fyllas, N.M., Lopez-Gonzalez, G., Hladik, A., Higuchi, N., Hunter, M.O., Lida, Y., Salim, K.A., Kassim, A.R., Keller, M., Kemp, J., King, D.A., Lovett, J.C., Marimon, B.S., Marimom-Junior, B.H., Lenza, E., Marshall, A.R., Metcalfe, D.J., Mitchard, E.T.A., Moran, E.F., Nelson, B.S., Nilus, R., Nogueira, E.M., Palace, M., Patiño, S., Peh, K.S.H., Raventos, M.T., Reitsma, J.M., Saiz, G., Schrodt, F., Sonké, B., Taedoumg, H.E., Tan, S., White, L., Woll, H., Lloyd, J., 2011. Height-diameter allometry of tropical forest trees. Biogeosci. Göttingen 8 (5), 1081–1106. doi:10.5194/bgd-7-7727-2010.

Foyolle, A., Doucet, J.L., Gillet, J.F., Bourland, N., Lejeune, P., 2013. Tree allometry in Central Africa: testing the validity of pantropical multi-species allometric equations for estimating biomass and carbon stocks. For. Ecol. Manag., Amsterdam 305, 29–37. doi:10.1016/j.foreco.2013.05.036.

Henricksen, H.A., 1950. Height diameter curve with logarithmic diameter: brief report on a more reliable method of height determination from height curves, introduced by the State Forest Research Branch. Dan. Skovforeningerns Tidsskr. Frederiksberg 35 (4), 193–202.

Hiramatsu, N.A., 2008. Equação Do Volume Comercial Para Espécies Nativas Na Região do Vale do Jari, Amazônia Oriental. Universidade Federal do Paraná, Curitiba 107 f. Dissertação (Mestrado em Engenharia Florestal).

Hoaglin, D.C., Welsch, R.E., 1978. The hat matrix in regression and ANOVA. Am. Stat. Assoc. Alexandria 32 (1), 17–22. doi:10.1080/00031305.1978.10479237.

Hunter, M.O., Keller, M., Victoria, D., Morton, D.C., 2013. Tree height and tropical forest biomass estimation. Biogeosci. Göttingen 10, 8385–8399. doi:10.5194/bg-10-8385-2013.

Huxley, J.S., Teissier, G., 1936. Terminology of relative growth. Nat. London 137, 780–781. doi:10.1038/137780b0.

Kearsley, E., Haulleville, T.de, Hufkens, K., Kidimbu, A., Toirambe, B., Baert, G., Huygens, D., Kebede, Y., Defourny, P., Bogaert, J., Beeckman, H., Steppe, K., Boeckx, P., Verbeeck, H., 2013. Conventional tree height-diameter relationships significantly overestimate aboveground carbon stocks in the Central Congo Basin. Nat. Commun. London 4, 1–8. doi:10.1038/ncomms3269.

Leite, H.G., Oliveira, F.H.T.de;, 2002. Statistical procedure to test identity between analytical methods. Commun. Soil Sci. Plant Anal. London 33, 7–8. doi:10.1081/CSS-120003875.

Li, J., Valliant, R., 2011. Linear regression influence diagnosis for unclustered survey data. J. Off. Stat. 27 (1), 99–119.

Lima, A.J.N., Suwa, R., Ribeiro, G.H.P.de, Kajimoto, M., dos Santos, T., da Silva, J., de Souza, R.P., de Barros, C.A.S., Noguchi, P.C., Ishizuka, H., Higuchi, M., 2012. Allometric models for estimating above-and below-ground biomass in Amazonian forests at São Gabriel da Cachoeira in the upper Rio Negro. Braz. For. Ecol. Manag. Amsterdam 227, 163–172. doi:10.1016/j.foreco.2012.04.028.

Machado, S.A., Nascimento, R.G.M., Augustynczik, A.L.D., Silva, L.C.R., Figura, M.A., Pereira, E.M., Téo, S.J., 2008. Behavior of the hypsometric relationship of Araucaria angustifolia in the forest copse of the faculty of forest - Federal University of Paraná. Braz. Pesquisa Florestal Bras. Colombo (56) 5–16.

Marshall, A.R., Willcock, S., Platts, P.J., Lovett, J.C., Balmford, A., Burgess, N.D., Latham, J.E., Munishi, P.K.T., Salter, R., Shirina, D.D., Lewis, S.L., 2012. Measuring and modelling above-ground carbon and tree allometry along a tropical elevation gradient. Biol. Conserv. Amsterdam 154, 20–33. doi:10.1016/j.biocon.2012.03.017.

Mayer, D.G., Butler, D.G., 1993. Statistical validation. Ecol. Modell. Amsterdam 68, 21–32. doi:10.1016/0304-3800(93)90105-2.

Nuzzo, R., 2014. Statistical errors: p values, the "gold standard" of statistical validity, are not as reliable as many scientists assume. Nat. London 506 (7487), 150–152. doi:10.1038/506150a.

Oderwald, R.G., Hans, R.P., 1993. Corroborating models with model properties. For. Ecol. Manag. Amsterdam 62, 271–283. doi:10.1016/0378-1127(93)90054-Q.

Picard, N., Rutishauser, E., Ploton, P., Ngomanda, A., Henry, M., 2015. Should tree biomass allometry be restricted to power models? For. Ecol. Manag. Amsterdam 353, 156–163. doi:10.1016/j.foreco.2015.05.035.

Piñeiro, G., Perelman, S., Guerschman, J., Paruelo, J.M., 2008. How to evaluate models: observed vs predicted or predicted vs observed? Ecol. Model. Amsterdam 216, 316–322. doi:10.1016/j.ecolmodel.2008.05.006.

Pinheiro, K.A.O., Carvalho, J.C., Quanz, P de, Francez, B., de, L.M., Schwartz, B., 2007. Phytosociology of a permanent preservation area in east of Amazon: indication of species for recovering altered areas. Floresta, Curitiba 37 (2), 175–187. doi:10.5380/rf.v37i2.8648.

Prodan, M, 1944. Zuwachs-und Ertragsuntersuchungen im Plenterwald. Universidade de Freiburg, Freiburg 127 f. Dissertação (Mestrado em Ciências).

R Core Team. R: a language and environment for statistical computing. 2015. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Robinson, A.P., Froese, R.E., 2004. Model validation using equivalence tests. Ecol. Modell. Amsterdam 176, 349–358. doi:10.1016/j.ecolmodel.2004.01.013.

Robinson, A.P., Hamann, J.D., 2011. Forest Analytics with R: an Introduction. Springer, Baltimore, MD, p. 355. doi:10.1007/978-1-4419-7762-5.

Rutishauser, E., Noor'an, F., LaumonieR, Y., Halperin, J., Rufi'ie Hergoualc'h, K., Verchot, L., 2013. Generic allometric models including height best estimate forest biomass and carbon stocks in Indonesia. For. Ecol. Manag. Amsterdam 307, 219–225. doi:10.1016/j.foreco.2013.07.013.

Sharma, M., Parton, J., 2007. Height-diameter equations for boreal tree species in Ontario using a mixed-effects modelling approach. For. Ecol. Manag. Amsterdam 249, 187–198. doi:10.1016/j.foreco.2007.05.006.

Sileshi, G.W., 2014. A critical review of forest biomass estimation models, common mistakes and corrective measures. For. Ecol. Manag. Amsterdam 329, 237–254. doi:10.1016/j.foreco.2014.06.026.

Smith, E.P., Rose, K.A., 1995. Model goodness-of-fit analysis using regression and related techniques. Ecol. Model., Amsterdam 77 (1), 49–64. doi:10.1016/0304-3800(93)E0074-D.

Stoffels, A., van Soest, J., 1953. Principiele vraagstukken bij proefperken 3. Hoogteregressie. Nederlands bosbouwtijdschrift. Groningen 25, 190–199.

Trorey, L.G., 1932. A mathematical method for the construction of diameter height curves based on site. For. Chron. Mattawa 8 (2), 121–132. doi:10.5558/tfc8121-2.

Vibrans, A.C., Moser, P., Oliveira, L.Z., Maçaneiro, J.P., 2015. Height-diameter models for three subtropical forest types in Southern Brazil. Ciênc. Agrotecnol. Lavras 39 (3), 205–215. doi:10.1007/s13595-015-0481-x.

Ward, P.J., 2015. Prediction intervals: placing real bounds on regression-based allometric estimates of biomass. Biom. J. Torino 57 (4), 695–711. doi:10.1002/bimj.201400070.

Willmott, C., Ackleson, S., Davis, R., Feddema, J., Klink, K., Legates, D., O'Donnell, J., Zhao, W., Mason, E.G., Brown, J., 2006. Modelling height-diameter relationships of *Pinus radiata* plantations in Canterbury, New Zealand. N.Z. J. For. Wellingt. 51, 23–27.