

# HEIGHT-DIAMETER MODELS FOR THREE SUBTROPICAL FOREST TYPES IN SOUTHERN BRAZIL

## Modelos hipsométricos para três tipologias florestais do Sul do Brasil

Alexander Christian Vibrans<sup>1</sup>, Paolo Moser<sup>2</sup>, Laio Zimmermann Oliveira<sup>2</sup>, João Paulo de Maçaneiro<sup>2</sup>

### ABSTRACT

Total tree height ( $h$ ) is often difficult to measure in natural forests. Regression models based on easily accessed variables like DBH ( $d$ ) can be an alternative, since their assumptions are validated. The aims of this study are to: (i) calibrate specific and generic  $h$ - $d$  models for three forest types (Seasonal Deciduous Forest, DEC; Mixed Ombrophilous Forest, MIX; and Dense Rainforest, DEN) in Santa Catarina state testing the regression assumptions and evaluating model quality; (ii) verify different  $h$ - $d$  relationship between forest types. The dataset (1,766 measured tree  $h$  and 3,150 estimated  $h$ ) was collected by Santa Catarina Forest and Floristic Inventory (IFFSC) in 418 systematically located sample plots. Models were calibrated for two datasets, one containing hypsometer measurements, the other  $h$  estimations made by field crews. Specific models were calibrated for species with at least 30 sampled trees. Residual normality, randomness and heteroskedasticity were evaluated by analytical methods. Confidence bands were generated by the Working-Hotelling method;  $z$  test for means was applied to compare models based on the two databases. The statistical parameters such as corrected Akaike Information Criterion provided evidences that logarithmic models were better adjusted to the data. Both datasets were statistically different for DEN and MIX. Differences in  $h$ - $d$  relationships were found between forest types. The use of calibrated  $h$ - $d$  models is an alternative for studying the relationships between these variables and to assess vertical structure patterns of forest communities, when  $h$  measurements are not feasible, although, for situations that more accurate  $h$  values are needed, they will not always provide reliable predictions.

**Index terms:** Tree height prediction;  $h$ - $d$  relationship; forest inventory; regression assumptions; model calibration.

### RESUMO

Em florestas naturais, a altura total das árvores ( $h$ ) é, muitas vezes, difícil de medir. Nesses casos, modelos de regressão baseados em variáveis de fácil medição, como o DAP ( $d$ ), são uma alternativa, desde que seus pressupostos sejam atendidos. Neste estudo, objetivou-se i) o ajuste de modelos hipsométricos para três tipologias florestais de Santa Catarina (Floresta Estacional Decidua, Floresta Ombrófila Mista, Floresta Ombrófila Densa), testando os pressupostos e avaliando a qualidade dos ajustes; ii) verificação de diferenças das relações hipsométricas entre as tipologias florestais e entre os modelos ajustados baseados em alturas medidas e estimadas. Os dados de 1.766 alturas medidas e 3.150 alturas estimadas em campo foram coletados pelo Inventário Florístico Florestal de Santa Catarina (IFFSC), em 418 unidades amostrais. Os modelos foram ajustados para ambos os conjuntos de dados. Modelos específicos foram ajustados para espécies com  $\geq 30$  árvores medidas; normalidade, aleatoriedade e homoscedasticidade dos resíduos foram avaliadas analiticamente. Bandas com intervalos de confiança foram construídas pelo método Working-Hotelling; teste  $z$  para médias foi aplicado para comparar os modelos baseados nos dois conjuntos de dados. Parâmetros estatísticos como o AIC evidenciaram melhor ajuste dos modelos logarítmicos. Foram encontradas diferenças nas relações  $h/d$  entre as tipologias florestais (DEN e MIX) e entre os conjuntos de dados. O uso de modelos hipsométricos é uma alternativa para estudar a relação  $h/d$  e para determinar padrões da estrutura vertical em comunidades florestais, quando a medição de  $h$  não é viável, embora em casos da necessidade de maior acurácia das alturas estimadas, os modelos nem sempre possam fornecê-las.

**Termos para indexação:** Relação altura-diâmetro; inventário florestal; pressupostos de regressão; ajuste de modelos.

### INTRODUCTION

Some of the dendrometric variables usually assessed in forest inventories are dependent on each other and the nature of their dependence (e.g. linear, quadratic, logarithmic or exponential) allows estimates through mathematical models that express the relationship between these variables. The use of mathematical modeling turns out to be interesting when the measurement of dependent

variables is costly and the independent variables are easily obtainable. Calibrating regression models is the mathematical technique often used to construct estimates of an independent variable.

Total tree height ( $h$ ) is one of the variables difficult to measure, especially in tropical forests where several layers and a discontinuous and dense canopy occur (Hu, 1992; Vanclay, 1994). This measurement is usually taken with an instrument (hypsometer) that records the extent

<sup>1</sup>Universidade Regional de Blumenau/FURB – Departamento de Engenharia Florestal – Rua São Paulo – 3250 – 89030-000 – Blumenau – Santa Catarina – Brasil – acv@furb.br

<sup>2</sup>Universidade Regional de Blumenau/FURB – Departamento de Engenharia Florestal – Blumenau – SC – Brasil  
Received in october 19, 2014 and approved in january 20, 2015

of the angle and distance from a tree to estimate its total height using trigonometric relations (Crecente-Campo et al., 2010).

This procedure is subject to systematic and random errors due to factors such as distance between hypsometer operators and trees, device type, training level of operators, as well as crown form and total tree height, which can affect the measurement accuracy (Silva et al., 2012). Knowing that the height of a tree is correlated with its diameter at breast height ( $d$ ), an alternative solution to the mentioned problem is the use of  $h-d$  models, which relate these two variables. In fact, the  $h-d$  relationship is one of the most important elements of forest structure (Tewari; Gadow, 1999).

Most of the  $h-d$  relationship studies refer to even-aged stands, probably due to the aforementioned difficulties of data assessment in native forests. There are few studies concerning tropical forests (Fang; Bailey, 1998; Batista; Couto; Marquesini, 2001; Feldpausch et al., 2011; Scaranello et al., 2012). Batista; Couto and Marquesini (2001) analyzed the  $h-d$  relationship in an equatorial rainforest in Maranhão (Brazil) and in the Atlantic rain forest with *Tabebuia cassinoides* (Lam.) DC. in São Paulo (Brazil). Nonlinear models showed better biological foundation and demonstrated superior performance than linear models.

Machado et al. (2008) tested 13  $h-d$  models to estimate total height of *Araucaria angustifolia* (Bert.) O. Kuntze in Paraná state, Brazil. The authors achieved the best estimates with the logarithmic model proposed by Curtis (1967). Caetano et al. (2014) found that a second order polynomial model was the best for *Bertholletia excelsa* Bonpl. intercropped with *Hevea brasiliensis* ((Willd ex A. Juss.) Muell Arg.) in Minas Gerais state, Brazil.

Scaranello et al. (2012) studied the  $h-d$  relationship on an altitudinal gradient in the Atlantic Forest (Serra do Mar State Park, Brazil); the authors applied 11  $h-d$  models and obtained the best estimates from the Weibull and Chapman-Richards nonlinear models and concluded that models calibrated to specific altitude ranges produced better results than generic models. Feldpausch et al. (2011) concluded that different  $h-d$  relationships occur in different tropical forest types and that a  $h-d$  relationship is influenced by climatic, topographic and forest structural variables. Fang and Bailey (1998) investigated the  $h-d$  relationship of fallen trees on Hainan Island (China) and, among 33 mathematical models, the best estimates were generated by an exponential model for height curves published by Meyer (1940). Tewari and Gadow

(1999) used median regressions and percentile curves of 5 and 95% of the bivariate  $S_{BB}$  distribution to study the  $h-d$  relationship of a pure *Acacia tortilis* Hayne stand in India and in a mixed forest in Germany containing *Fraxinus excelsior* L., *Fagus sylvatica* L. and *Acer* sp. The percentile curves of the  $S_{BB}$  distribution indicated that the variation in  $h$  for a given  $d$  is less pronounced for larger trees. Huang, Price and Titus (2000) calibrated 27  $h-d$  models for *Picea glauca* (Moench) Voss in the Alberta Boreal Forest region and found that logistic models produced more satisfactory predictions. The authors found that the models present different behaviors within different ecoregions and the application of models adjusted for one ecoregion in another can generate different estimates: 29.5% overestimated or 21.92% underestimated.

Dendrometric data collected by the first cycle of Santa Catarina Forest and Floristic Inventory (IFFSC) in southern Brazil give a unique and unprecedented opportunity to explore the quantitative and qualitative aspects of the state's forest cover. IFFSC applied a systematic sample design, measured between 2007 and 2010 at equal probability sample points (Vibrans et al., 2010). Thus, the objectives of this study were to (i) calibrate specific and generic  $h-d$  models for three forest types in Santa Catarina state and to test the regression assumptions; (ii) choose the best model based on statistical indicators; (iii) compare the calibrated models for the measured and estimated datasets; (iv) compare the  $h-d$  relationships of each forest type.

## MATERIAL AND METHODS

### Study location

Santa Catarina state is located between latitudes 26° and 29° S and longitudes 48° and 53° W with an area of 95,346 km<sup>2</sup>. Three forest types, manifestations of important variations of geomorphological, climatic and edaphic site conditions, were covered in this study (Klein, 1978; Vibrans et al., 2010): Seasonal Deciduous Forest (DEC), Mixed Ombrophilous Forest with *Araucaria* (MIX) and Dense Rainforests (DEN). According to Köppen-Geiger climatic classification (Kottek et al., 2006), Santa Catarina has two climatic types: Cfa – fully humid temperate climate with warm summer and Cfb - fully humid temperate climate with cool summer. The climates are defined primarily by the temperature difference due to altitudinal variation. Long-term (> 30 years) averages of mean annual temperature are 18.35; 16.36; 18.90 (°C) for DEC, MIX and DEN,

respectively, while annual mean precipitation is 1,646; 1,632; 1,574mm (EPAGRI, 2002). Altitude range (m a.s.l.) for sample plots is 503-898m for DEC, 514-1,560m for MIX and 2-1,195m for DEN.

### Data collection and analysis

Data used in this study were obtained by IFFSC between 2007 and 2011 (Vibrans et al., 2010), using the Brazilian National Forest Inventory methodology (Freitas et al., 2010). The sampling design features a systematic distribution of sample points located at the intersections of a 10-km x 10-km grid that covers the entire state with the exception of DEC for which a 5km x 5km grid was used due to its highly fragmented status.

Of the 418 sample plots, 78 were located in forest remnants of DEC, 143 in MIX and 197 in DEN. The IFFSC sample plot consists of a cluster of four crosswise 1,000 m<sup>2</sup> subplots (20 m x 50 m), each one located at a distance of 30m from the plot center. Inside the sample plot limits, the field crews measured the  $d$  (diameter at breast height, 1.30 m) of every tree with  $d \geq 10$  cm; total tree heights were visually estimated by the field crews and total heights of up to eight trees per sample plot, spanning all diameter classes, were measured with a hypsometer (Vibrans et al., 2010). In total, 1,766 measured tree heights and 3,150 estimated heights were used to calibrate specific and generic  $h-d$  models as well as to examine the reliability of visual estimates made by the field crews.

The  $h-d$  models (Table 1) were calibrated in a generic way for two datasets: (i)  $h$  measured with hypsometers and (ii)  $h$  estimated by field crews. For further reference, these data are treated as the *measured dataset* and the *estimated dataset*.

Considering that a height for a given diameter can vary significantly between species (King, 1996) and regions (Huang; Price; Titus, 2000; Feldpausch et al., 2011),  $h-d$  models were also calibrated specifically for some species and for each forest type, in order to verify possible improvements on model quality or characteristic  $h-d$  relations. The specific model was calibrated for species having a sample size of at least 30 trees with exception of *Ocotea puberula* of MIX (22 trees) and *Clethra scabra* of DEN (29 trees) (Table 2).

A random sampling of 3,250 trees amid the entire estimated dataset was performed to facilitate their manipulation. For the measured dataset, this procedure was unnecessary, given the smaller number of measured trees.

**Table 1:** Mathematical structure of  $h-d$  models calibrated for total tree height prediction in Santa Catarina (southern Brazil).

$h-d$ model	Model number
$h = \beta_0 + \beta_1 d + \varepsilon$	1
$h = \beta_0 + \beta_1 d + \beta_2 d^2 + \varepsilon$	2
$h = \beta_0 + \beta_1 d + \beta_2 d^2 + \beta_3 d^3 + \varepsilon$	3
$h = \beta_0 + \beta_1 \frac{1}{d^2} + \varepsilon$	4
$\log h = \beta_0 + \beta_1 \frac{1}{d} + \varepsilon$	5
$\log h = \beta_0 + \beta_1 \ln d + \varepsilon$	6
$\frac{1}{\sqrt{h-1.3}} = \beta_0 + \beta_1 \frac{1}{d} + \varepsilon$	7
$\log h - 1.3 = \beta_0 + \beta_1 \frac{1}{d} + \varepsilon$	8
$\ln h = \beta_0 + \beta_1 \frac{1}{d} + \beta_2 \frac{1}{d^2} + \varepsilon$	9
$\frac{1}{\sqrt[3]{h-1.3}} = \beta_0 + \beta_1 \frac{1}{d} + \varepsilon$	10

$\beta_n$  (regression coefficients);  $\varepsilon$  (random error);  $h$  (total tree height (m));  $d$  (diameter at breast height (cm)), according to Loetsch; Zohrer and Haller (1973) and Finger (1992).

The datasets' normality were evaluated using the Kolmogorov-Smirnov test with  $\alpha = 0.05$ . Detection of outlier data was made by the standardized scores (mean 0 and standard deviation 1) of the indexes  $h/d$  and  $d/h$  which showed a normal distribution according to Shapiro-Wilk's test with  $\alpha = 0.05$ ; values whose absolute score were greater than 3 were considered outliers (for normal data, 99% standardized scores are within an interval of  $\pm 3$  standard deviations). Excluding outliers, the tree species with a sufficient number of measured individuals ( $n \geq 30$ ) for each forest type in Santa Catarina are presented in Table 2.

The  $h-d$  models were linearized for calibration using minimum least square method. The root mean square error (RMSE), pseudo-coefficient of determination ( $R^{2*}$ ), (Næsset, 2011) and corrected Akaike information criterion (AICc) were calculated for each model, all based on back transformed data (original scale). The  $R^{2*}$  was used since the assumptions underlying  $R^2$  are not completely satisfied when using nonlinear models (Anderson-Sprecher, 1994).

**Table 2:** Number of measured individuals (N) and those used to fit the models (n) for the measured and estimated datasets for each forest type in Santa Catarina (southern Brazil).

Forest type / Species	Measured <i>h</i> data			Estimated <i>h</i> data		
	N	n	<i>d</i> range (cm)	N	n	<i>d</i> range (cm)
DEC						
<i>Nectandra megapotamica</i>	46	46	10-61	958	282	10-60
<i>Ocotea puberula</i>	89	89	10-60	1,151	296	10-68
Generic model	341	271	10-62	14,714	966	10-68
MIX						
<i>Ocotea puberula</i>	22	21	11-70	662	250	10-62
Generic model	355	199	10-104	35,623	599	10-78
DEN						
<i>Cedrela fissilis</i>	31	31	12-43	479	215	10-82
<i>Clethra scabra</i>	29	27	10-48	860	273	10-48
<i>Hieronyma alchorneoides</i>	47	47	12-50	1,231	300	10-86
<i>Miconia cinnamomifolia</i>	61	58	10-69	825	264	10-69
<i>Nectandra oppositifolia</i>	30	30	11-60	567	235	10-141
<i>Ocotea puberula</i>	40	39	10-82	447	210	10-82
<i>Pera glabrata</i>	30	27	10-61	673	248	10-132
<i>Tapirira guianensis</i>	37	37	10-57	819	268	10-73
Generic model	1,446	1,317	10-114	5,1691	2,413	10-171
Overall generic model	2,142	1,787	10-114	102,028	3,978	10-171

DEC (Seasonal Deciduous Forest); MIX (Mixed Ombrophilous Forest); DEN (Ombrophilous Forest).

Overall regression significance was assessed using the *F* test ( $\alpha=0.05$ ) of the regression's analysis of variance. A model ranking procedure was used to select the model with the best performance: the significant model (for both overall and parameters significance) with lower AICc was selected. For the best models, residuals were analyzed to check the basic regression assumptions – normality was investigated using the Kolmogorov-Smirnov test, Runs test to randomness and Brown-Forsythe test to heteroskedasticity (Neter et al., 1996). The *F* test for lack of fit (Zar, 2010) was conducted on the linearized scale to assert the statistical significance of the linear relationship between the variables. All the aforementioned hypothesis tests used  $\alpha = 0.05$ .

Confidence bands were generated by the Working-Hotelling method (Neter et al., 1996) with  $\alpha = 0.05$ . The interpretation of these bands is analogous to the

interpretation of confidence intervals for each regression estimate. However, the confidence bands are wider because they relate to the regression line as a whole and not to specific estimates.

Next, a comparison of each pair of models calibrated for the measured and estimated datasets was made. A simulation of 80 random values of *d* (10-80 cm range) was conducted and *h* was predicted by each pair of the models to be compared. The mean values of those 80 predicted *h* were compared through a *z* test for independent means (Triola, 1999) with  $\alpha = 0.05$ . This procedure was conducted 100 times for all the best performance models (generics and specifics) aiming to prevent precipitate inferences based on a single simulation of *d* values.

To study the differences in *h-d* relationship among the three forest types, it is necessary to conduct an inside analysis of the *h-d* models in a common

range of the predictor variable ( $d$ ). Graybill and Iyer (1994) proposed that a statistical test should not be used by itself but in combination with the construction of simultaneous confidence intervals using dummy variables based on a Student's  $t$  distribution; these intervals present the differences between the regressions instead of a single and static conclusion, which a hypothesis test yields. These intervals must be constructed simultaneously to avoid the increment of the significance level. To achieve this premise, Bonferroni corrections are a usable approach, resulting in conservative confidence intervals. The regression parameters confidence intervals ( $\alpha = 0.05$ ) were constructed for minimum  $d$  of 10 cm and a maximum of 80cm, which is the most sampled diameter range. The models considered in this analysis were the generic ones calibrated for the measured dataset.

## RESULTS AND DISCUSSION

The results of the calibration process are presented in Table 3 for the two datasets. The measured data are graphed over their calibrated models to demonstrate the dispersion around the regression line, circumventing the need for residual graphs. Since the residual analysis was made based on analytical procedures, visual analysis was not necessary. The best-fitted models for each dataset, respectively, were also graphed on the same cartesian plane. The aforementioned graphs for DEC, MIX and for all data (Overall generic) are shown in Figure 1 and 2.

The model residuals were subjected to Kolmogorov-Smirnov (normality), Runs (randomness) and Brown-Forsythe (heteroskedasticity) tests. All models presented an expected residual behavior; the tests showed no evidence to reject the null hypothesis ( $p \geq 0.05$ ). Therefore, the residuals may be regarded as satisfying the aforementioned assumptions. The  $F$  tests for overall regression significance rejected the null hypothesis of absence of regression for every calibrated model. All the  $F$  tests for lack of fit presented non-significant  $p$ -values ( $p \geq 0.05$ ), except the generic model of DEN (model n°9) and Overall generic model (model n°9).

The species-specific and generic models calibrated for measured dataset presented substantially lower AICc than for the estimated dataset. This suggests that data variance is higher on the heights estimated by field crews. This issue reflects on the results of the  $z$  tests for means which rejected the null hypothesis for DEN, MIX and

Overall generic models (all  $p$ -values  $< 0.05$ , according with the 100 runs), meaning that the models calibrated for each dataset did not estimated the same mean height. For DEC the  $z$  test failed to reject the null hypothesis (all  $p$ -values  $\geq 0.05$ ). For all species-specific models the  $z$  tests were not significant, except for *Ocotea puberula* in DEC.

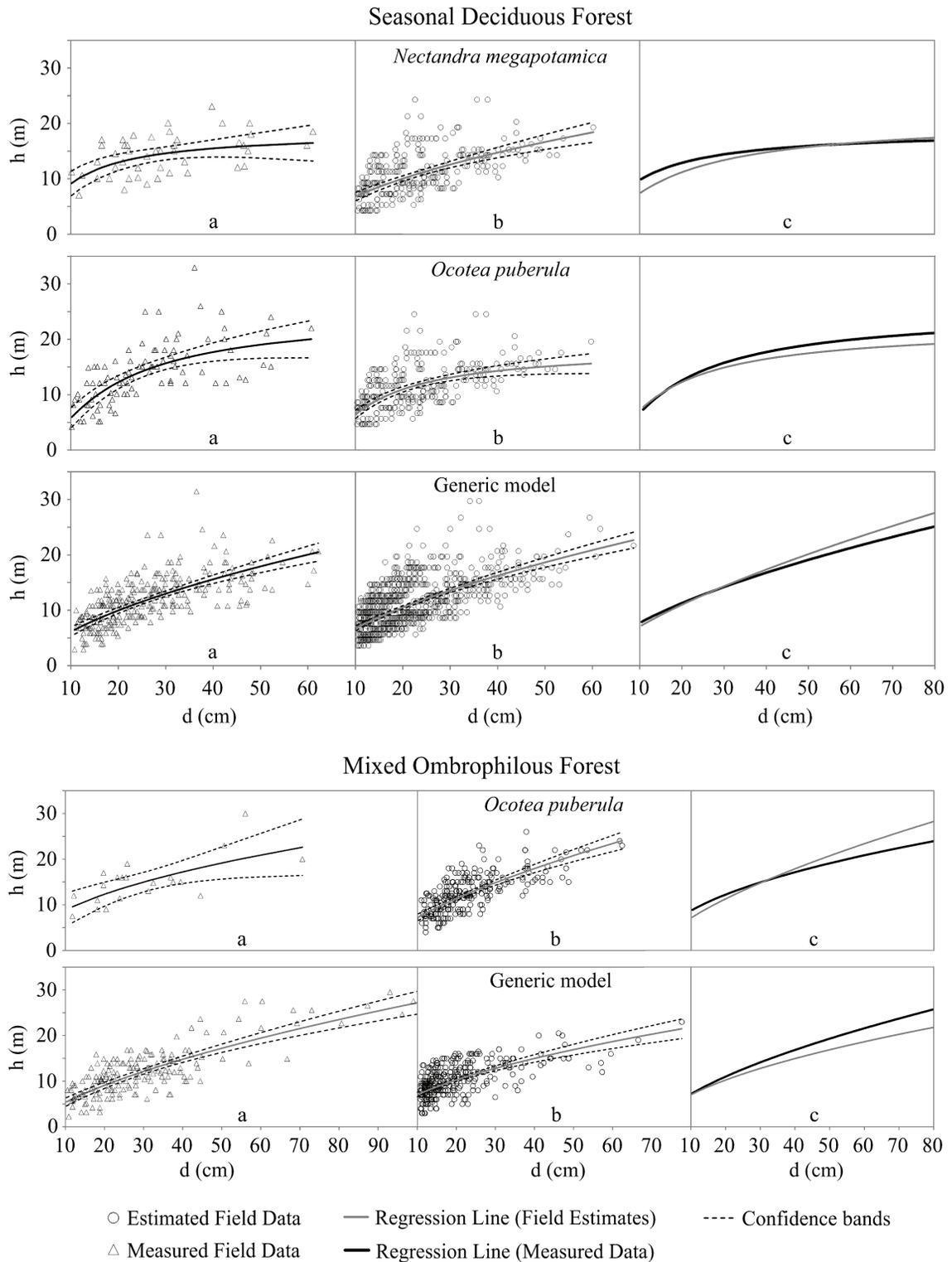
Different  $h$ - $d$  relationships between the forest types were observed, as found by Feldpausch et al. (2011) for different forest types distributed throughout the tropics. These differences can be verified by examining the difference between the adjusted parameters for the regressions (Feldpausch et al., 2011). Table 4 shows the simultaneously generated confidence intervals for these parameters.

For both minimum and maximum  $d$ , the only confidence interval containing zero is for the difference between DEC and MIX, providing no evidence to reject the null hypothesis; so, there is no evidence for assuming differences in the  $h$ - $d$  relationship between the two forest types. The DEN results indicate significant differences from the other two forest types for both minimum and maximum  $d$ . Figure 3 shows the simultaneous confidence intervals for the three models with  $d$  ranging from 10 to 80 cm.

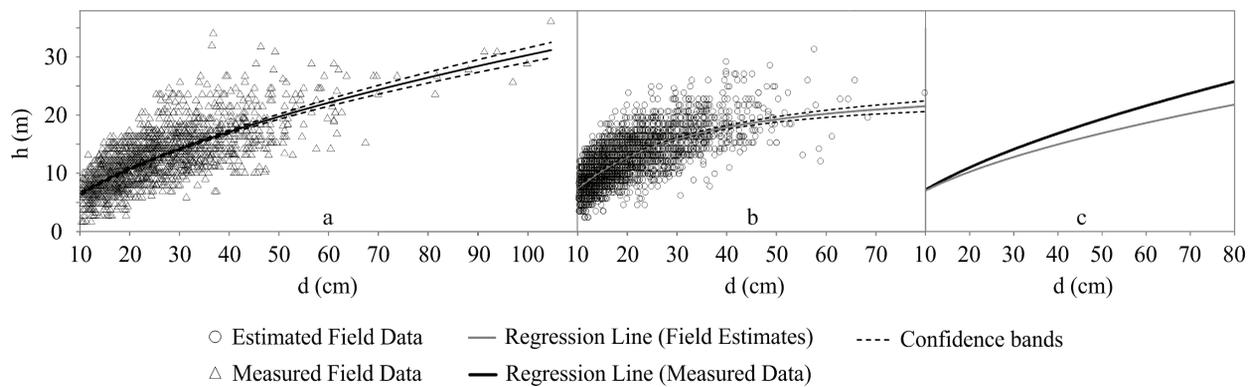
Among the 10 calibrated  $h$ - $d$  models, the logarithmic models of Loetsch; Zohrer and Haller (1973) (models n°5 and 6) proved to be the most effective in 86% of cases. Its advantage is associated, in part, with the fact that the logarithmic scale tends to correct the level of data heteroskedasticity coming from the positive correlation between the mean and variance of the dataset (Sokal; Rohlf, 1987). Log-linear models have good qualities, such as satisfactory precision and easy adjustment and are widely used models for allometric purposes (Brown; Gillespie; Lugo, 1989; Fang; Bailey, 1998; Nogueira et al., 2008; Breidenbach et al., 2014; McRoberts; Westfall, 2014; McRoberts et al., 2015). Although, when model parameters are estimated using data transformed to the logarithmic scale and prediction is desired on the original scale, an adjustment term must be added to the prediction to compensate for bias that accrues due to the transformation (Baskerville, 1972). By the way, this procedure is often ignored by Brazilian forester community. In this study the addition of the correction terms to the predictions were considered negligible because residual variance ( $s^2_{res}$ ) were less than or equal to 0.03; predictions were little affected by a multiplicative factor of 1.02.

**Table 3:** Best-fitted *h-d* models calibrated based on measured and estimated datasets (denoted as M and E on the model number) for each forest type in Santa Catarina (southern Brazil), where:  $\beta_n$  = regression coefficients;  $R^2^*$  = pseudo-coefficient of determination; RMSE = root mean squared error; p = p-value concerning to the lack of fit test; z-t = results of z tests (100 runs) - when most of the p-values were significant, denoted as SIG and when the opposite happens, denoted as NS; DEC = Seasonal Deciduous Forest; MIX = Mixed Ombrophilous Forest; DEN = Ombrophilous Forest.

Forest type / Species	Model umber	$\beta_0$	$\beta_1$	$\beta_2$	$R^2^*$	RMSE	AICc	p	z-t
DEC									
<i>N. megapotamica</i>	5 <sup>M</sup>	1.2662	-3.1068	-	0.22	3.13	109.29	0.83	NS
	5 <sup>E</sup>	1.2805	-4.5634	-	0.36	3.27	671.54	0.68	
<i>O. puberula</i>	5 <sup>M</sup>	1.3390	-6.2690	-	0.45	4.01	251.50	0.67	SIG
	6 <sup>E</sup>	0.2489	0.2709	-	0.35	3.76	787.51	0.99	
Generic model	6 <sup>M</sup>	0.3121	0.2488	-	0.45	3.19	634.36	0.68	NS
	6 <sup>E</sup>	0.1613	0.2919	-	0.40	3.36	2005.97	0.99	
MIX									
<i>O. puberula</i>	6 <sup>M</sup>	0.4648	0.2089	-	0.47	3.59	58.40	0.25	NS
	6 <sup>E</sup>	0.2006	0.2856	-	0.54	2.86	528.57	0.47	
Generic model	6 <sup>M</sup>	0.2483	0.2655	-	0.66	3.29	425.47	0.72	SIG
	5 <sup>E</sup>	1.2670	-4.6446	-	0.42	2.70	668.51	0.81	
DEN									
<i>C. fissilis</i>	5 <sup>M</sup>	1.4091	-5.1854	-	0.46	2.98	72.31	0.35	NS
	5 <sup>E</sup>	1.3896	-5.1591	-	0.61	2.54	405.26	0.61	
<i>C. scabra</i>	6 <sup>M</sup>	0.5670	0.1756	-	0.58	1.53	27.66	0.42	NS
	6 <sup>E</sup>	0.4391	0.2117	-	0.32	2.17	427.63	0.81	
<i>H. alchorneoides</i>	10 <sup>M</sup>	0.3448	1.6154	-	0.43	2.82	101.88	0.68	NS
	10 <sup>E</sup>	0.3421	1.8571	-	0.51	2.65	590.69	0.99	
<i>M. cinnamomifolia</i>	5 <sup>M</sup>	1.4685	-6.6504	-	0.59	2.68	118.43	0.13	NS
	5 <sup>E</sup>	1.4390	-5.8937	-	0.61	2.67	524.25	0.99	
<i>N. oppositifolia</i>	6 <sup>M</sup>	0.4439	0.2240	-	0.47	3.31	76.33	0.29	NS
	6 <sup>E</sup>	0.4314	0.2306	-	0.54	2.72	475.12	0.32	
<i>O. puberula</i>	6 <sup>M</sup>	0.3233	0.2611	-	0.49	2.99	89.81	0.77	NS
	6 <sup>E</sup>	0.2992	0.2647	-	0.64	2.27	348.61	0.95	
<i>P. glabrata</i>	6 <sup>M</sup>	0.4135	0.2410	-	0.62	2.12	45.14	0.50	NS
	6 <sup>E</sup>	0.3980	0.2372	-	0.43	2.17	389.53	0.51	
<i>T. guianensis</i>	6 <sup>M</sup>	0.7078	0.1536	-	0.45	2.52	72.83	0.99	NS
	6 <sup>E</sup>	0.4604	0.2182	-	0.51	2.62	521.64	0.42	
Generic model	6 <sup>M</sup>	0.3799	0.2476	-	0.58	3.02	2260.51	0.71	SIG
	9 <sup>E</sup>	3.2847	-16.5885	46.5339	0.56	2.60	4610.16	<0.001	
Overall generic model	6 <sup>M</sup>	0.3757	0.2435	-	0.54	3.18	4089.74	0.72	SIG
	9 <sup>E</sup>	3.2250	-15.1420	33.9081	0.53	2.71	5490.79	<0.001	



**Figure 1:** Best-fitted  $h$ - $d$  models calibrated to measured (a) and estimated (b) datasets for DEC and MIX (the points are the real field data); (c) shows the overlap of the measured and estimated models.



**Figure 2:** Best-fitted  $h$ - $d$  models calibrated to measured (a) and estimated (b) datasets for the Overall generic model (the points are the real field data); (c) shows the overlap of the measured and estimated models.

**Table 4:** Simultaneous confidence intervals ( $\alpha = 0.05$ ) for the differences among the  $h$ - $d$  relationship for the generic models calibrated for the measured dataset for the three forest types in Santa Catarina (southern Brazil).

$d$ range (cm)	Simultaneous confidence intervals
$d_{\min}$ (10 cm)	$-0.256 \leq \text{DEC} - \text{MIX} \leq 0.628$
	$-1.779 \leq \text{DEC} - \text{DEN} \leq -1.012$
	$-1.863 \leq \text{MIX} - \text{DEN} \leq -1.299$
$d_{\max}$ (80 cm)	$-5.630 \leq \text{DEC} - \text{MIX} \leq 4.255$
	$-8.516 \leq \text{DEC} - \text{DEN} \leq -0.013$
	$-6.771 \leq \text{MIX} - \text{DEN} \leq -0.383$

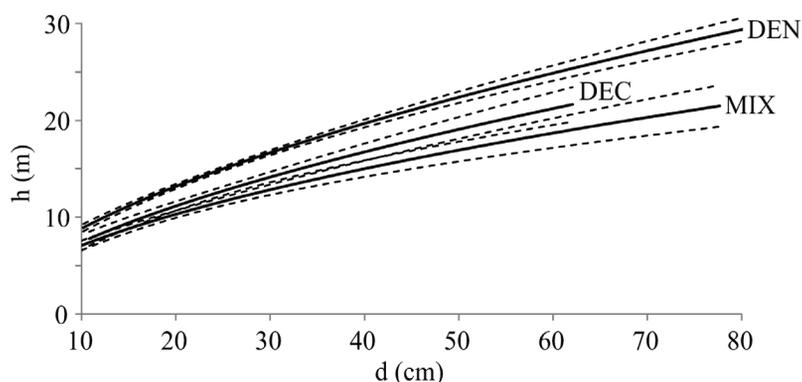
DEC (Seasonal Deciduous Forest); MIX (Mixed Ombrophilous Forest); DEN (Dense Ombrophilous Forest).

The variability within the  $h$ - $d$  data is an obstacle for the calibration of models for native forests. There is a large dispersion of the observed values and of the model-predicted values as found by Huang et al. (2000). The encountered magnitude of the  $R^{2*}$  values was not large, having a maximum value of 0.66. However, the  $F$  tests revealed that all regressions were statistically significant, providing evidence for the existence of correlations between  $h$  and  $d$ , but not necessarily of causal relationship between these variables. Like site conditions in even-aged forest stands, tree height is probably influenced by other random explanatory factors not included in this study.

The use of multivariate models may be an alternative that could increase the prediction and explanation of  $h$  (Neter et al., 1996).

Despite the fact that the purpose of mathematical modeling is to provide an average estimate for the desired value of the response variable, the models calibrated for high variant data will not generate fully reliable predictions. The calibration of  $h$ - $d$  models appears to be an alternative for studying different relationships between these variables (Feldpausch et al., 2011; Scaranello et al., 2012) and to assess vertical structure patterns of forest communities, when  $h$  measurements are not feasible; although, for situations that more accurate  $h$  values are needed they will not always provide reliable predictions.

The comparison between the models adjusted for the measured and estimated datasets showed that most of the species-specific models predicted the same mean values for  $h$ . This result demonstrates an overall quality of visual estimates made by IFFSC's field crews. Thus, visual estimation may be a feasible operational technique, since instrumental measurements are often laborious and time-consuming to obtain. However, the generic models did not predict the same mean  $h$  (except for DEC) for the measured and estimated datasets. These facts, together with the higher AICc of generic models, may be evidences that species-specific models are more appropriated when predictions of  $h$  are desired. In addition, species-specific models are also appropriate for analyzing the different  $h$ - $d$  relationships between species (King, 1996) or regions/sites (Huang; Price; Titus, 2000; Scaranello et al., 2012).



**Figure 3:** Representation of the best fitted generic  $h-d$  models estimates with their respective simultaneous confidence bands for DEC, MIX and DEN.

Some factors may explain the results concerning the  $h-d$  relationships of the three forest types. Variables such as basal area are closely related to vegetation structure and can produce variation in  $h-d$  allometry (King et al., 2009; Feldpausch et al., 2011). This observation made by Feldpausch et al. (2011) may, in part, explain the different  $h-d$  relationships observed for DEN in relation to the other forest types. The total basal area ( $\text{m}^2 \cdot \text{ha}^{-1}$ ) of this forest type proved to be statistically different from that of the other forest types (Vibrans et al., 2012). The influence of basal area is explained by the high competition for light where individuals tend to grow first in height to reach the upper forest layers (Feldpausch et al., 2011). King et al. (2009) observed different  $h-d$  relationships in two well-developed forests in Malaysia with different basal areas. These differences were also observed by King (1981) and Hummel (2000), who found that trees growing in more open environments have larger diameters compared to trees growing in closed environments such as tropical rainforests; therefore basal area perhaps should be included in new models as a predictor variable to account for competition.

Geomorphological and climatic factors may also influence the  $h-d$  relationship, as well as altitude (Huang et al., 2000; Lawton, 1982; Feldpausch et al., 2011; Scaranello et al., 2012). While DEN occurs in coastal sedimentary plains and the granulitic complex of the coastal range in altitudes between sea level and 1,200m, the MIX occurs in the sedimentary and basaltic plateaus of Santa Catarina, at altitudes ranging from 500 to 1,500m a.s.l., where the winter season is characterized by the frequent occurrence of ground frost and temperatures that are considerably lower than those of the coastal region. According to Woodward

(1993), the wind speed increases with increasing altitude. The highest altitude of the Santa Catarina plateau exposes the vegetation to stronger winds and may influence the  $h-d$  relationship, as observed by Lawton (1982). Scaranello et al. (2012) noticed that the decrease in air and soil temperature at higher altitudes can also influence  $h-d$  relationships.

## CONCLUSIONS

The regression assumptions for all the presented  $h-d$  models were validated; this allows for estimates with valid confidence intervals. The high variance within  $h-d$  datasets of native forests is a limitation for the calibration of models which are intended to be used as reliable predictors of  $h$ . The calibration of  $h-d$  models showed to be an alternative for studying different relationships between these variables, although, for situations that more accurate  $h$  values are needed they will not always provide reliable predictions. Compared to the generic models the species-specific ones are more indicated when predictions of  $h$  are desired. Most of the species-specific models calibrated for the measured dataset and estimated dataset estimated the same mean  $h$ . This is an evidence that the visual estimation can be used when field crews are appropriately trained; it is necessary to use the  $h$  measurements of some trees in the field to create a reference for further estimations. There are differences in the  $h-d$  relationship of the DEN when compared to the other two forest types. However, no significant differences were found in the  $h-d$  relationships between MIX and DEC, possibly due to the similar climatic and geomorphological conditions as well as the similar structural and floristic closeness of these two forest types.

## ACKNOWLEDGMENTS

The authors are grateful to Fundação de Amparo à Pesquisa e Inovação de Santa Catarina (FAPESC) for supporting the IFFSC, to CNPq for the research grant awarded to the first author, to CAPES and FURB for the Master's scholarship granted to the third and fourth author. We are indebted to Ronald McRoberts for the kind review and the anonymous reviewers for their valuable contribution to the manuscript.

## REFERENCES

- ANDERSON-SPRECHER, R. Model comparisons and  $R^2$ . **The American Statistician**. 48(2):113-117, 1994.
- BASKERVILLE, G.L. Use of logarithmic regression in the estimation of plant biomass. **Canadian Journal of Forest Research**. 2(1):49-53, 1972.
- BATISTA, J.L.F.; COUTO, H.T.Z.; MARQUESINI, M. Desempenho de modelos de relações hipsométricas: estudo em três tipos de floresta. **Scientia Forestalis**. (60):149-163, 2001.
- BREIDENBACH, J. et al. Quantifying the contribution of biomass model errors to the uncertainty of biomass stock and change estimates in Norway. **Forest Science**. 60: 25-33, 2014.
- BROWN, S., GILLESPIE, A.J.R., LUGO, A.E. Biomass estimation methods for tropical forests with application to forest inventory data. **Forest Science**. 35(4):881-902, 1989.
- CAETANO, A.C. et al. Height-diameter equations for Brazil nut intercropped with rubber tree in the south of Minas Gerais. **Floresta**. 44(3):497-504, 2014.
- CRECENTE-CAMPO, F. et al. Generalized nonlinear mixed-effects height-diameter model for *Eucalyptus globulus* L. in northwestern Spain. **Forest Ecology and Management**. 259(5):943-952, 2010.
- CURTIS, R.O. Height-diameter and height-diameter-age equations for second-growth Douglas Fir. **Forest Science**. 13:365-375, 1967.
- EPAGRI. Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina. **Atlas Climatológico do Estado de Santa Catarina**. Florianópolis (Brazil): EPAGRI, 2002. CD-ROM
- FANG, Z.; BAILEY, R.L. Height-diameter models for tropical forests on Hainan Island in southern China. **Forest Ecology and Management**. 110(1-3):315-327, 1998.
- FELDPAUSCH, T.R. et al. Height-diameter allometry of tropical forest trees. **Biogeosciences Discussions**. 7(5):7727-7793, 2011.
- FINGER, C.A.G. **Fundamentos de biometria florestal**. Santa Maria/Brazil: UFSM/CEPEF/FATEC, 1992. 269p.
- FREITAS, J.V. et al. National Forest Inventory Reports: Brazil. In: TOMPPA, E. et al. **National Forest Inventories: Pathways for Common Reporting**. London: Springer, 2010, p.89-96.
- GRAYBILL, F.A.; IYER, H.K. **Regression analysis: concepts and applications**. Belmont, CA: Duxbury Press. 1994. 648p.
- HU, Y.J. **The tropical forest of Hainan Island**. Guang Dong Province: The College Education Press, 1992. 453p.
- HUANG, S.; PRICE, D.; TITUS, S.J. Development of ecoregion-based height-diameter models for white spruce in boreal forests. **Forest Ecology and Management**. 129(1-3):125-140, 2000.
- HUMMEL, S. Height, diameter and crown dimensions of *Cordia alliodora* associated with tree density. **Forest Ecology and Management**. 127(1-3):31-40, 2000.
- KING, D. Tree dimensions: maximizing the rate of height growth in dense stands. **Oecologia**. 51(3):351-356, 1981.
- KING, D.A. Allometry and life history of tropical trees. **Journal of Tropical Ecology**. 12(1):25-44, 1996.
- KING, D.A. et al. Trees approach gravitational limits to height in tall lowland forests of Malaysia. **Functional Ecology**. 23(2):284-291, 2009.
- KLEIN, R.M. Mapa fitogeográfico do estado de Santa Catarina. **Flora Ilustrada Catarinense**. parte5:1-24, 1978.

- KOTTEK, M. et al. World map of the Köppen-Geiger climate classification updated. **Meteorologische Zeitschrift**. 15(3):259-263, 2006.
- LAWTON, R.O. Wind stress and elfin stature in a montane rain forest tree: an adaptive explanation. **American Journal of Botany**. 69:1224-1230, 1982.
- LOETSCH, P., ZÖHRER, P., HALLER, K.E. **Forest inventory**. Munich: BLV, 1973. 469p.
- MACHADO, S.A. et al. Comportamento da relação hipsométrica de *Araucaria angustifolia* no capão da Engenharia Florestal da UFPR. **Pesquisa Florestal Brasileira**. 56:5-16, 2008.
- MCROBERTS, R.E., MOSER, P., OLIVEIRA, L.Z., VIBRANS, A.C. A general method for assessing the effects of uncertainty in individual tree volume model predictions on large area volume estimates with a subtropical forest illustration. **Canadian Journal of Forest Research**. 45(1):44-51, 2015.
- MCROBERTS, R.E., WESTFALL, J. Effects of uncertainty in model predictions of individual tree volume on large area volume estimates. **Forest Science**. 60(1):34-42, 2014.
- MEYER, H.A. A mathematical expression for height curves. **Journal of Forestry**. 38(5):415-420, 1940.
- NÆSET, E. Estimating above-ground biomass in young forests with airborne laser scanning. **International Journal of Remote Sensing**. 32(2):473-501, 2011.
- Neter, J. et al. **Applied Linear Statistical Models**, 4. ed. Boston: McGraw-Hill, 1996. 439p.
- NOGUEIRA, E.M. et al. Estimates of forest biomass in the Brazilian Amazon: new allometric equations and adjustments to biomass from wood-volume inventories. **Forest Ecology and Management**. 256(11):1853-1867, 2008.
- SCARANELLO, M.A.S. et al. Height-diameter relationships of tropical Atlantic moist forest trees in southeastern Brazil. **Scientia Agricola**. 69(1):26-37, 2012.
- SILVA, G.F. et al. Influência de diferentes fontes de erro sobre as medições de alturas de árvores. **Cerne**. 18(3):397-405, 2012.
- SOKAL, R.R.; ROHLF, F.J. **Introduction to Biostatistics**. New York: Freeman and Co, 1987. 363p.
- TEWARI, V.P.; GADOW, K.V. Modelling the relationship between tree diameters and heights using  $S_{BB}$  distribution. **Forest Ecology and Management**. 119(1-3):171-176, 1999.
- TRIOLA, M.F. **Introdução à estatística**. Rio de Janeiro: Editora LTC, 1999. 410p.
- VANCLAY, J.K. **Modeling forest growth and yield: Applications to mixed tropical forests**. Wallingford: CABI, 1994. 312p.
- VIBRANS A.C. et al. Inventário florístico florestal de Santa Catarina (IFFSC): aspectos metodológicos e operacionais. **Pesquisa Florestal Brasileira**. 30(64):291-302, 2010.
- VIBRANS, A.C. et al. Análise estatística do IFFSC e estimativas dendrométricas. In: VIBRANS, A.C. et al. **Diversidade e conservação dos remanescentes florestais**. Inventário Florístico Florestal de Santa Catarina. Blumenau: Edifurb, 2012, vol.1, p.79-95.
- WOODWARD, F.I. The lowland-to-upland transition-modelling plant responses to environmental change. **Ecological Applications**. 3(3):404-408, 1993.
- ZAR, J.H. **Biostatistical Analysis**. 5. ed. New York: Pearson Prentice-Hall, 2010. 944p.