

# Generic and specific stem volume models for three subtropical forest types in southern Brazil

Alexander C. Vibrans<sup>1</sup> · Paolo Moser<sup>1</sup> · Laio Z. Oliveira<sup>1</sup> · João P. de Maçaneiro<sup>1</sup>

Received: 11 March 2014 / Accepted: 25 March 2015  
© INRA and Springer-Verlag France 2015

## Abstract

• **Key message** We adjusted generic models for species-rich forest types and specific models for 15 species. Regression assumptions, lack of fitness and goodness of fit and comparison between models were assessed analytically. Generic models produced estimates not less reliable than species-specific models. Logarithmic models presented the best results of adjustment and evenness of residual variance.

• **Context** Assessment of dendrometric variables is important to obtain accurate estimates of stand attributes as biomass and carbon stock estimates. Some of them, as tree height and stem volume, are difficult and expensive to measure; volume models, calibrated on large datasets in tropical and subtropical forests, are rare.

• **Aims** This study aimed to construct stem volume models for native tree species in three forest types in southern Brazil, to select models with best fitness, to assess agreement between measured and predicted datasets and to compare species-specific and generic models.

• **Methods** Data from 418 sample plots were used to adjust generic models for forest types and specific models for 15 species. Regression assumptions, modelling efficiency, lack of fitness, goodness of fit and comparison between species-specific and generic models were assessed by analytical methods.

• **Results** Logarithmic models presented the best results of adjustment and evenness of residual variance. Lack of fit  $F$  test showed acceptable adjust quality for nearly all species-specific and generic models;  $R^2_{adj}$ \* and modelling efficiency measure presented values close to 1 for all fitted models; model identity  $F$  test showed differences between specific and generic models in some cases.

• **Conclusion** Since regression assumptions were satisfied and because of their quality of fit, the fitted models compose useful tools for predicting total stem volume (with bark) for forest remnants in southern Brazil. Stratification of datasets by forest type for model fitting showed to be necessary, but, commonly, generic models for forest types produced estimates not less reliable than species-specific models.

**Keywords** Volumetric models · Forest inventory · Regression assumptions · Goodness of fit · Modelling efficiency

**Handling Editor:** Laurent Saint-Andre

**Contribution of the co-authors** The first author is coordinator of this long-term project and also responsible for the sample design and data collection; the second author supervised data analysis and writing of the manuscript; the first three authors performed writing of the manuscript, interpretation and discussion of the results; and the fourth author ran data analysis and revision of literature.

✉ Alexander C. Vibrans  
acv@furb.br

<sup>1</sup> Departamento de Engenharia Florestal, Universidade Regional de Blumenau, Rua São Paulo, 3250, 89030-000 Blumenau-Santa Catarina, Brazil

## 1 Introduction

Precise assessments of dendrometric variables such as tree diameter, height, stem form and volume are important to obtain accurate estimates of attributes of forest stands, tree populations or vegetation communities. Some of these variables, including tree height and stem volume, are difficult and expensive to measure, especially in tropical forests (Picard et al. 2012). However, statistical modelling using observations for relatively small number of trees sampled from the entire population allows prediction of these variables.

In Brazil, few studies at either local or broader geographic scales investigated volume models for natural forests. The first forest surveys that aimed to collect information about timber stock in southern Brazil were conducted in the early 1980 (Brasil 1983; Netto 1984) and adjusted a generic volume model for broadleaved species and a specific model for the Brazilian-pine (*Araucaria angustifolia* Bertol O. Kuntze). The research of Scolforo et al. (2008) in the Minas Gerais (Brazil) Forest Inventory reported construction of generic and species-specific volume models using data of 2060 trees distributed among several forest types. Another important study, despite its locally restricted database, was realized by Machado et al. (2008) who investigated the behaviour of volume models for *Mimosa scabrella* Benth. in mixed ombrophilous forest (Paraná, Brazil). In the National Forest Tapirapé-Aquiri (Pará, Brazil), located in the Amazonian biome, Rolim et al. (2006) constructed eight generic volume models with two predictor variables and four with one predictor variable. Colpini et al. (2009) investigated ten volume models for open ombrophilous forest in the Amazonian biome. Silva et al. (1993) constructed models to estimate the volume of tree species from mangroves in north-eastern Brazil using nonlinear transformations (Box and Cox 1964) to avoid heteroscedasticity in predictor variables. Borsoi (2004) adjusted several generic volume models for three groups of species regarding their commercial value in a forest management study conducted in a secondary mixed ombrophilous forest (Rio Grande do Sul, Brazil). Also, in the Rio Grande do Sul state, Spathelf et al. (2001) adjusted specific models for species in deciduous and dense ombrophilous forest types.

In the USA, in a large-scale study, Cormier et al. (1992) adjusted volumetric and taper models using a dataset consisting of observations for 5350 loblolly-pine (*Pinus taeda* L.) trees collected by the US Forest Service. The results of this study showed that the quality of predictions obtained from weighted regressions (based on residual variances) and non-weighted regressions was similar. Volume models constructed for European tree species were intensively reviewed and discussed by Zianis et al. (2005) who further constructed a relevant database for volume estimation purposes.

The rarity of volume models for tropical and sub-tropical tree species adversely affects estimation of tree volume for both research- and operational-scale applications (Azevedo et al. 2008; Braz et al. 2012). The first measurement cycle of the Santa Catarina Forest and Floristic Inventory (IFFSC), Southern Brazil, was concluded in 2010 and provided a valuable database that supports investigations of quantitative and qualitative variables of the state's forests (Vibrans et al. 2012a). In this context, and motivated by the lack of tree volume models for native forest species in Brazil, the objectives of this paper are twofold: (1) construct generic and species-specific stem volume models for native tree species for three forest types in Santa Catarina and (2) test the

hypothesis that the adjusted models for species-specific datasets separately produce statistically significant better results than the adjusted model for the pooled dataset (or generic model).

## 2 Materials and methods

### 2.1 Study location

The Brazilian state of Santa Catarina has a total area of 95,985 km<sup>2</sup> and is located between latitudes 25° 57' and 29° 21' S and longitudes 48° 62' and 53° 50' W (Fig. 1). Its main three forest types were covered in this study: seasonal deciduous forest (DEC), mixed ombrophilous forest (MIX) and dense ombrophilous forest (DEN) (Klein 1978; IBGE 2012). Altitude in the state varies between sea level and 1560 m a.s.l. (Vibrans et al. 2010).

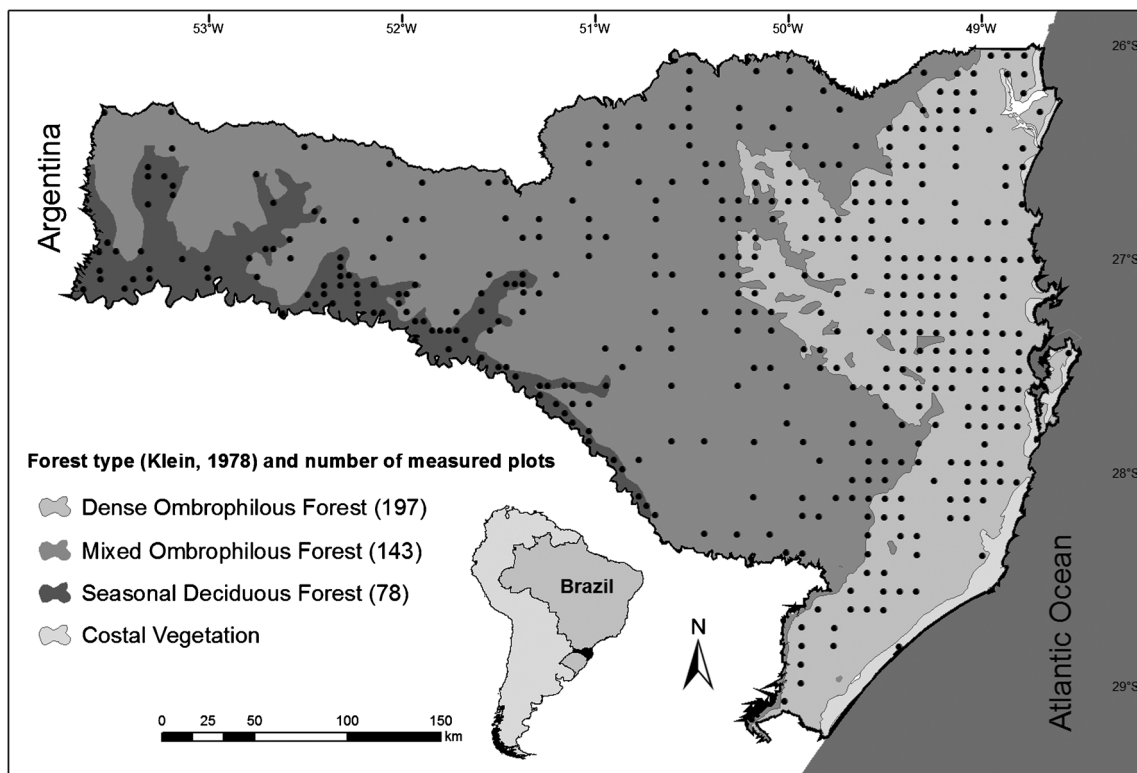
According to Köppen's climatic classification (Kottek et al. 2006), Santa Catarina has two climatic types: Cfa—fully humid temperate climate with warm summer—and Cfb—fully humid temperate climate with cool summer. The climate classes are defined primarily by temperature differences resulting from altitudinal variation. The annual mean temperature varies between 10 and 22 °C, annual mean precipitation varies between 1100 and 2900 mm and is uniformly distributed throughout the year, and annual mean relative humidity varies between 74 and 88 % (EPAGRI 2002).

### 2.2 Data collection

Data for this study were collected by the IFFSC between 2007 and 2011. IFFSC adopted the Brazilian National Forest Inventory methodology (Freitas et al. 2010) featuring a systematic distribution of sample points located at the intersections of a 10 km×10 km grid which covered the entire state (Vibrans et al. 2010). From the 1074 sample plots, 418 satisfied the IFFSC definition of forest land: area with continuous tree vegetation, canopy height of at least 10 m and basal area of at least 10 m<sup>2</sup> ha<sup>-1</sup>. For the DEC forest type, a 5 km×5 km sampling grid was used due to its highly fragmented status. Among the 418 sample plots, 78 were located in the DEC, 143 sample plots were located in the MIX, and 197 sample plots were located in the DEN (Fig. 1).

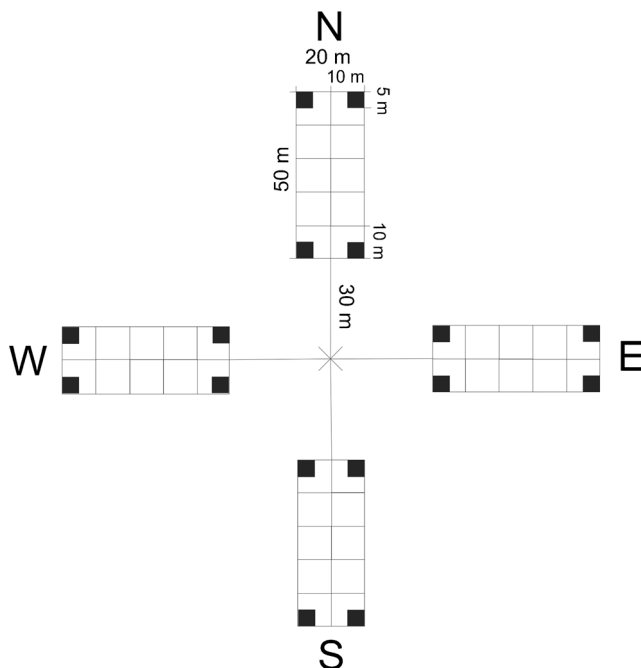
The IFFSC's sample plots consist of a cluster of four cross-wise 1000-m<sup>2</sup> subplots (20 m×50 m), each located at a distance of 30 m from the plot centre; every subplot consists of ten basic subunits with nominal areas of 100 m<sup>2</sup> (10 m×10 m) (Fig. 2).

For as many as two standing trees on each of the four subplots, field crews (including climbers) registered the species and measured tree diameter at breast height ( $d$ , 1.30 m) and circumferences using a tape measure at heights of 0.3, 1.0,



**Fig. 1** Location of the 418 IFFSC's sample plots in Santa Catarina state

1.3, 2.0 and 3.0 m and every full meter. The total stem height ( $h_s$ ) was assigned as the last measurement at the crown base. The aforementioned measures were taken from the stem base (at 0.3 m from the ground) until the crown base; then, the respective tree stem volumes ( $V_s$ ) with bark were calculated,



**Fig. 2** IFFSC's sample plot configuration

using the Smalian method (Avery and Burkhart 2002) due to operational reasons. The stem definition adopted for this study is equivalent to the 'trunk' compartment in Picard et al. (2012, p. 55). Different species and diameter ranges were contemplated for each forest type. The total numbers of measured trees before and after deletion of outliers by species (with at least 30 observations) and forest type are shown in Table 1. The outlier detection was based on two indices:  $h_s/d$  and  $d/h_s$ , which have the property for identifying very tall trees with smaller  $d$  and trees with large  $d$  and smaller  $h_s$ , respectively. Since the standardized scores (mean 0 and standard deviation 1) of these indices presented normal distribution (Shapiro-Wilk test,  $p > 0.05$ ), the outlier status was assigned to trees whose standardized scores were greater than  $\pm 3$  standard deviations of the mean (as 99 % of a normal distributed variable are within the  $\pm 3$  standard deviation interval). *A. angustifolia* trees were not measured, once considering the existence of specific models developed by Netto (1984). Thus, the generic model adjusted in the present study for MIX does not comprise this species.

### 2.3 Data analysis

Generic volume models were constructed for each forest type, and an overall generic model was constructed using all available data. Stem volume data ( $V_s$ ) were obtained from stem measurements of 57 species in DEC, 107 species in MIX

**Table 1** Collected data for constructing stem volume models by forest type for Santa Catarina, Brazil—number of measured individuals ( $N$ ), number of individuals after deletion of outliers ( $n$ ) and species' relative frequency across sampled plots considering its respective forest type ( $f\%$ )

Forest type/species	$N$	$n$	$f\%$	$d$ range (cm)
<b>DEC</b>				
<i>Nectandra megapotamica</i>	35	32	93.6	17–42
<i>Ocotea puberula</i>	72	68	83.3	17–56
Generic model	285	275	–	17–60
<b>MIX</b>				
<i>Cedrela fissilis</i>	30	28	46.8	18–50
<i>Clethra scabra</i>	47	44	65.0	20–49
<i>Matayba elaeagnoides</i>	33	31	59.4	20–48
<i>Ocotea puberula</i>	59	57	62.2	17–54
<i>Prunus myrtifolia</i>	34	33	78.3	10–44
Generic model	635	606	–	10–76
<b>DEN</b>				
<i>Alchornea triplinervia</i>	40	37	76.1	13–49
<i>Cedrela fissilis</i>	47	46	64.5	15–51
<i>Hieronyma alchorneoides</i>	38	37	49.7	18–44
<i>Miconia cinnamomifolia</i>	100	98	34.5	16–44
<i>Nectandra oppositifolia</i>	38	37	59.4	14–42
<i>Ocotea puberula</i>	30	29	27.4	15–37
<i>Piptocarpha angustifolia</i>	36	35	24.9	16–43
<i>Tapirira guianensis</i>	32	30	22.8	15–50
<i>Virola bicuhyba</i>	46	45	35.5	17–63
Generic model	1207	1196	–	11–67
Overall generic model	2127	2077	–	10–76

and 207 species in DEN. To explore possible improvements on model quality, the measured trees were stratified by species when their sampling frequency was greater than 30 individuals. Models were constructed for individual species within forest types as follows:

- Seasonal deciduous forest: *Nectandra megapotamica* (Spreng.) Mez and *Ocotea puberula* (Rich.) Nees
- Mixed ombrophilous forest: *Cedrela fissilis* Vell., *Clethra scabra* Pers., *Matayba elaeagnoides* Radlk., *O. puberula* (Rich.) Nees and *Prunus myrtifolia* (L.) Urb.
- Dense ombrophilous forest: *Alchornea triplinervia* (Spreng.) M. Arg., *Cedrela fissilis* Vell., *Hieronyma alchorneoides* Allemão, *Miconia cinnamomifolia* (DC.) Naudin, *Nectandra oppositifolia* Ness, *O. puberula* (Rich.) Nees, *Piptocarpha angustifolia* Dusén ex Malme, *Tapirira guianensis* Aubl. and *Virola bicuhyba* (Schott) Warb.

Based on a literature research, nine stem volume models were constructed for each species or for generic groups using  $h_s$ ,  $d$  and circumference at breast height ( $c$ ) as predictor

variables. The mathematical structure of these models is shown in Table 2.

Depending on the structure of the model, transformations were applied to variables, especially with logarithm scale. For this case, an adjustment term must be added to compensate the bias generated by the scale change. This correction term is expressed by  $s_{\text{res}}^2$  (residual variance) divided by two, which is added to the function that is found in the power of the model (Baskerville 1972). Low values of  $s_{\text{res}}^2$  may permit to neglect the appliance of the correction term due to its low influence upon the final result. Since all  $s_{\text{res}}^2$  were equal or lower than 0.04, predictions will be slightly affected by a maximum multiplicative factor of 1.05, which can be considered negligible; therefore, the correction term was not used.

After data normality verification through the Kolmogorov-Smirnov test ( $\alpha=0.01$ ) and outlier removal, the volume models were constructed using SPSS software. Pearson's product-moment correlation coefficient ( $r$ ), root-mean-square error (RMSE), adjusted coefficient of determination ( $R^2_{\text{adj}}$ ) and Akaike's information criterion (AIC) were calculated for each model. All these measures were computed based on retransformed data (original scale). The  $R^2_{\text{adj}}$  is denoted  $R^2_{\text{adj}}^*$  (pseudo- $R^2_{\text{adj}}$ ) since the assumptions underlying  $R^2$  are not completely satisfied when using nonlinear models (Anderson-Sprecher 1994). The statistical significance of  $r$  values was assessed using the  $t$  tests ( $\alpha=0.05$ ) in function of the number of sampled data. Overall regression significance was also assessed using the  $F$  test ( $\alpha=0.05$ ) of the regression analysis of variance.

After the verification of the significance of the descriptive statistics, a model ranking procedure was used to select the model with the best performance. The model with lower AIC was selected. For the best models, residuals were analyzed to check the basic regression assumptions. Normality was investigated using the Kolmogorov-Smirnov test, run test to randomness and the Brown-Forsythe test (a modification of Levene test) to assess homoscedasticity (Neter et al. 1996). All these hypothesis tests used  $\alpha=0.01$ .

The  $F$  test for lack of fit ( $\alpha=0.05$ ) was conducted on the  $\ln$ -scale to assert that the linear model adjustment presents a good fit for the data (Neter et al. 1996). The modelling efficiency (EF) was calculated on the original scale for the chosen models, following the standards given by Mayer and Butler (1993). This dimensionless statistic presents an overall indication of goodness of fit; a model which gives a negative value is not validated, with preferable values close to 1 indicating a best fitted model, as like  $R^2_{\text{adj}}^*$  (pseudo- $R^2_{\text{adj}}$ ).

For testing the hypothesis that fitted models for species-specific datasets separately generate equal parameters as a model fitted to the pooled dataset, the  $F$  test for model identity ( $\alpha=0.05$ ) based on the extra sum of square principle provided by Ratkowsky (1983) was used. The test was applied for the following situations: (a) forest-type generic model versus



**Table 2** Mathematical structure of stem volume models used to predict  $V_s$  (m<sup>3</sup>) (with bark) for native tree species of three forest types in Santa Catarina, Brazil

No.	Volumetric model	Author
1	$V_s = \beta_0 + \beta_1(d^2 h_s) + \varepsilon$	Spurr (1952)
2	$V_s = \beta_0 + \beta_1 d^2 + \beta_2 (d^2 h_s) + \beta_3 h_s + \varepsilon$	Stoate (1945)
3	$\ln V_s = \beta_0 + \beta_1 \ln(d^2 h_s) + \varepsilon$	Spurr (1952)
4	$\ln V_s = \beta_0 + \beta_1 \ln d + \beta_2 \ln h_s + \varepsilon$	Schumacher (1942)
5	$V_s = \beta_0 + \beta_1 d^2 + \varepsilon$	Kopecky and Gehrhardt with Finger (1992)
6	$V_s = \beta_0 + \beta_1 d + \beta_2 d^2 + \varepsilon$	Hohenadl and Krenn with Finger (1992)
7	$\ln V_s = \beta_0 + \beta_1 \log d + \beta_2 \frac{1}{d} + \varepsilon$	Brenac with Finger (1992)
8	$\ln \frac{V_s}{1000} = \beta_0 + \beta_1 \ln c^2 + \beta_2 \ln h_s + \varepsilon$	Adapted from Schumacher and Hall (1933)
9	$\ln \frac{V_s}{1000} = \beta_0 + \beta_1 \ln c^2 h_s + \varepsilon$	Adapted from Schumacher and Hall (1933)

Where  $h_s$  = stem height (m),  $d$  = diameter at breast height (cm),  $c$  = circumference at breast height (cm) and  $\varepsilon$  = random error

species-specific models for each, respectively, forest type and (b) forest-type generic models versus overall generic model. For this means, the procedure requires the sum of squared errors of species-specific models ( $SSE_{sp}$ ) with  $p$  parameters and the sum of squared errors of the model for the pooled data ( $SSE_{pool}$ ) adjusted for the same  $m$  species with  $n$  observations; the  $F$  statistics is computed as follows:

$$F(m \cdot p - p, n - m \cdot p) = \frac{\frac{SSE_{pool} - SSE_{sp}}{m \cdot p - p}}{\frac{SSE_{sp}}{n - m \cdot p}} \quad (1)$$

The species-specific datasets stand for the most frequent species and cannot include the great amount of rare and not sufficiently sampled and measured species.

### 3 Results

Among the volumetric models shown in Table 1, model no. 8 (logarithmic model by Schumacher and Hall 1933) presented the best fitting (lowest AIC) for all situations, for both generic and species-specific models. This model produced the best predictions with  $R^2_{adj}*$  and EF ranging from 0.85 to 0.98, RMSE ranging from 0.029 to 0.098 and AIC ranging from -123.96 to -10,279.87. Table 3 shows the adjusted stem volume models with best overall performance by species and forest type. All selected models had highly significant correlation coefficients ( $p \leq 0.01$ ) and excellent residual behaviour for normality, randomness and homoscedasticity. The tests did not reject any of the null hypotheses ( $p > 0.01$ ) regarding satisfaction of the regression assumptions. This result supports the decision to use regression methods for constructing models and estimating their parameters. None of the adjusted models presented lack of fit according to the  $F$  test, except the model (no. 8) for *O. puberula* in MIX and the overall generic

model (no. 8). Also, for MIX, data concerning *Prunus myrtifolia* did not have enough replications for the  $F$  test, so the test was not able to be conducted.

Figure 3 shows graphic representations of the predictions of  $V_s$  together with the predictor variables  $c$  and  $h_s$  for generic models. The graphs show the agreement between the point clouds of measured and predicted values, showing that estimation on ln-ln scale and backtransforming did not induce significant bias.

The generic models adjusted by forest type presented considerably lower AIC compared to the specific ones; however, the overall generic model presented lower AIC than the forest-type generic models. The RMSE, EF and  $R^2_{adj}*$  values indicate an even residual behaviour among the models, for both generics and specifics. For DEC and MIX,  $F$  tests for comparisons between generic and species-specific models presented no evidences for rejection of the null hypothesis ( $F=0.20$ ;  $p=0.89$  and  $F=0.76$ ;  $p=0.68$ , respectively). However, this was not the case at DEN, where  $F$  test showed evidences for rejection of the null hypothesis ( $F=9.49$ ;  $p<0.01$ ). The final comparison between forest-type generic and overall generic models also showed evidences for rejection of the null hypothesis ( $F=16.70$ ;  $p<0.01$ ).

The graphic representation of model predictions versus the two predictor variables is a three-dimensional surface. This surface is shown in Fig. 4 for DEC's generic model with the respective measured data graphed together (Fig. 4a). The graph (Fig. 4b) shows the plane  $c$  versus  $V_s$  and depicts the relationship between these two variables. The rate of change in  $V_s$  with respect to  $c$  is substantially greater for large trees than for small trees; this conclusion is corroborated by the analysis of the derivatives along the curve generated by the superior border of the surface. The relationship between  $h_s$  and  $V_s$  is shown in Fig. 4c and is the inverse of the aforementioned relationship; the rate of change in  $V_s$  (with respect to  $h_s$ ) is larger for small trees than for large trees; as previously, the conclusion is supported by the analysis of the derivatives.

**Table 3** Best fitted stem volume models, constructed using  $V_s$  measurements in each forest type of Santa Catarina, Brazil

Forest type/species	$\beta_0$	$\beta_1$	$\beta_2$	$SE_{\beta_0}$	$SE_{\beta_1}$	$SE_{\beta_2}$	Mn	$R^2_{adj}*$	RMSE	AIC	EF	$p$ value
DEC												
<i>N. megapotamica</i>	-17.72	0.95	0.69	0.271	0.031	0.060	8	0.97	0.037	-205.10	0.97	0.12
<i>O. puberula</i>	-17.52	0.93	0.68	0.192	0.020	0.041	8	0.98	0.029	-476.40	0.98	0.08
Generic model	-17.68	0.95	0.67	0.114	0.013	0.021	8	0.96	0.052	-1621.72	0.96	0.35
MIX												
<i>C. fissilis</i>	-18.57	1.02	0.75	0.550	0.057	0.121	8	0.91	0.098	-123.96	0.92	0.32
<i>C. scabra</i>	-18.31	0.99	0.76	0.485	0.052	0.080	8	0.92	0.040	-278.36	0.92	0.61
<i>M. elaeagnoides</i>	-18.03	0.99	0.61	0.593	0.064	0.088	8	0.94	0.047	-183.23	0.93	0.29
<i>O. puberula</i>	-18.30	1.00	0.74	0.215	0.023	0.047	8	0.97	0.044	-350.44	0.97	0.03
<i>P. myrtifolia</i>	-18.48	1.02	0.74	0.243	0.026	0.059	8	0.97	0.050	-192.07	0.97	<sup>a</sup>
Generic model	-17.96	0.96	0.76	0.102	0.011	0.019	8	0.94	0.057	-3468.65	0.94	0.21
DEN												
<i>A. triplinervia</i>	-18.10	1.05	0.51	0.377	0.042	0.085	8	0.98	0.048	-218.35	0.98	0.99
<i>C. fissilis</i>	-17.80	1.03	0.46	0.261	0.025	0.063	8	0.95	0.088	-218.05	0.94	0.33
<i>H. alchorneoides</i>	-17.74	1.02	0.45	0.291	0.035	0.054	8	0.97	0.046	-221.87	0.97	0.34
<i>M. cinnamomifolia</i>	-17.53	0.97	0.59	0.245	0.025	0.039	8	0.93	0.056	-557.60	0.93	0.99
<i>N. oppositifolia</i>	-17.89	0.99	0.64	0.233	0.028	0.063	8	0.98	0.051	-214.55	0.98	0.30
<i>O. puberula</i>	-17.93	1.01	0.58	0.548	0.056	0.125	8	0.85	0.076	-143.48	0.86	0.20
<i>P. angustifolia</i>	-18.35	1.04	0.59	0.242	0.023	0.045	8	0.96	0.045	-211.67	0.96	0.65
<i>T. guianensis</i>	-17.86	0.95	0.82	0.576	0.054	0.129	8	0.90	0.085	-147.10	0.89	0.99
<i>V. bicuhyba</i>	-17.98	0.98	0.64	0.282	0.034	0.071	8	0.96	0.097	-204.01	0.96	0.05
Generic model	-17.75	0.98	0.57	0.078	0.009	0.018	8	0.92	0.096	-5593.44	0.91	0.43
Overall generic model	-17.84	0.96	0.69	0.057	0.006	0.011	8	0.92	0.084	-10279.87	0.92	0.00

Where  $SE_{\beta_0}$ ,  $SE_{\beta_1}$  and  $SE_{\beta_2}$  are the standard errors of estimated parameters,  $h_s$  = stem height,  $d$  = diameter at breast height,  $c$  = circumference at breast height,  $R^2_{adj}*$  = (pseudo) adjusted coefficient of determination

RMSE root-mean-square error, AIC Akaike's information criterion, EF modelling efficiency,  $p$  value  $p$  value corresponding to lack of fit test

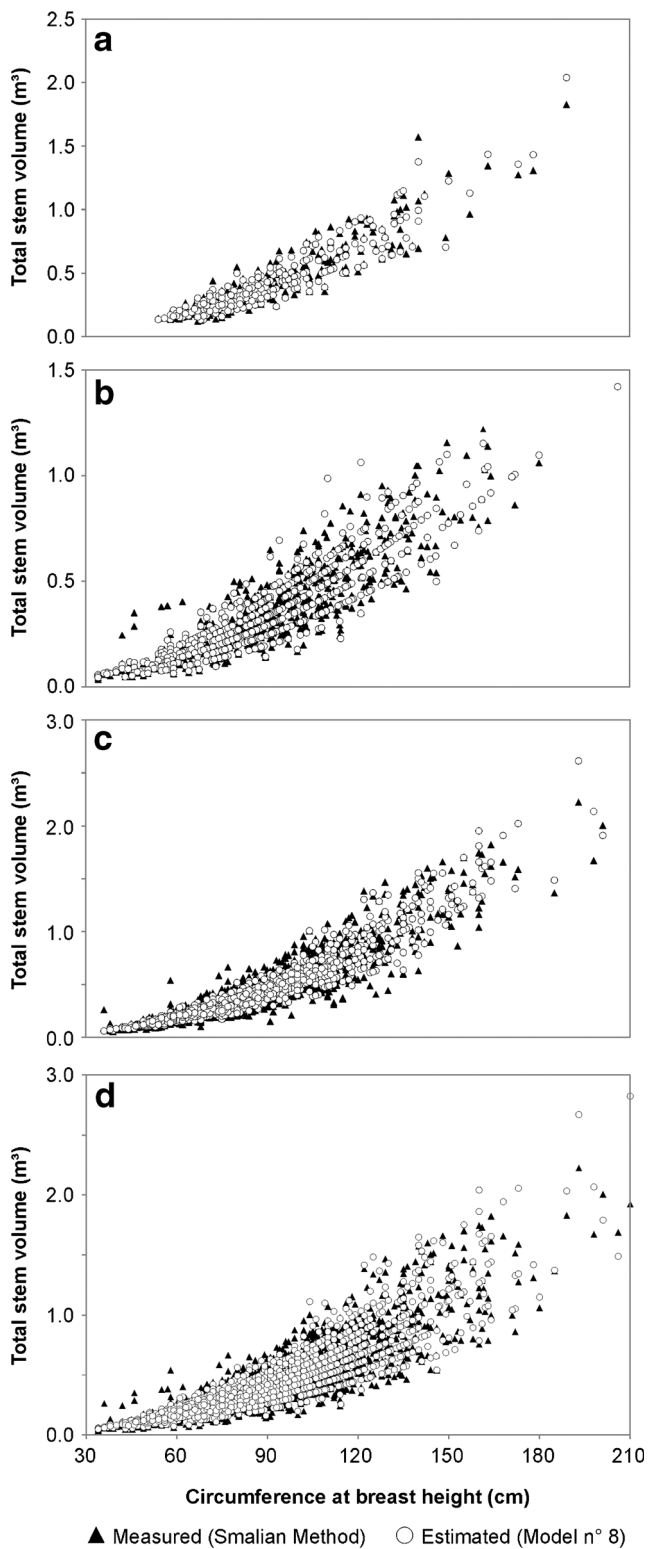
<sup>a</sup> There were no replications in this dataset to allow the execution of the  $F$  test for lack of fit

## 4 Discussion

Besides the fact that nonlinear models are available and provide the possibility of fitting by maximum likelihood any model for the variance (Zuur et al. 2009; Picard et al. 2012), the logarithmic scaling and back transforming to original scale have a long and continuing history in constructing volume and biomass allometric models (Baskerville 1972; Brown et al. 1989; Chave et al. 2005; Segura and Kanninen 2005; Litton and Kauffman 2008; Basuki et al. 2009; Guendehou et al. 2012; Breidenbach et al. 2014; McRoberts and Westfall 2014; McRoberts et al. 2015). In Brazil, there is a clear trend for using this kind of transformation in volume model adjustment for native forests. The logarithmic model by Schumacher and Hall (1933) presented best overall adjustment in this study, similarly as found by Scolforo et al. (2008), who used the same model to calculate stem and total volume estimates for different forest types for the Minas Gerais (Brazil) state forest inventory: cerrado sensu stricto (or forested savanna, Oliveira-Filho 2009), campo cerrado (or parkland savanna, Oliveira-Filho 2009), ombrophilous

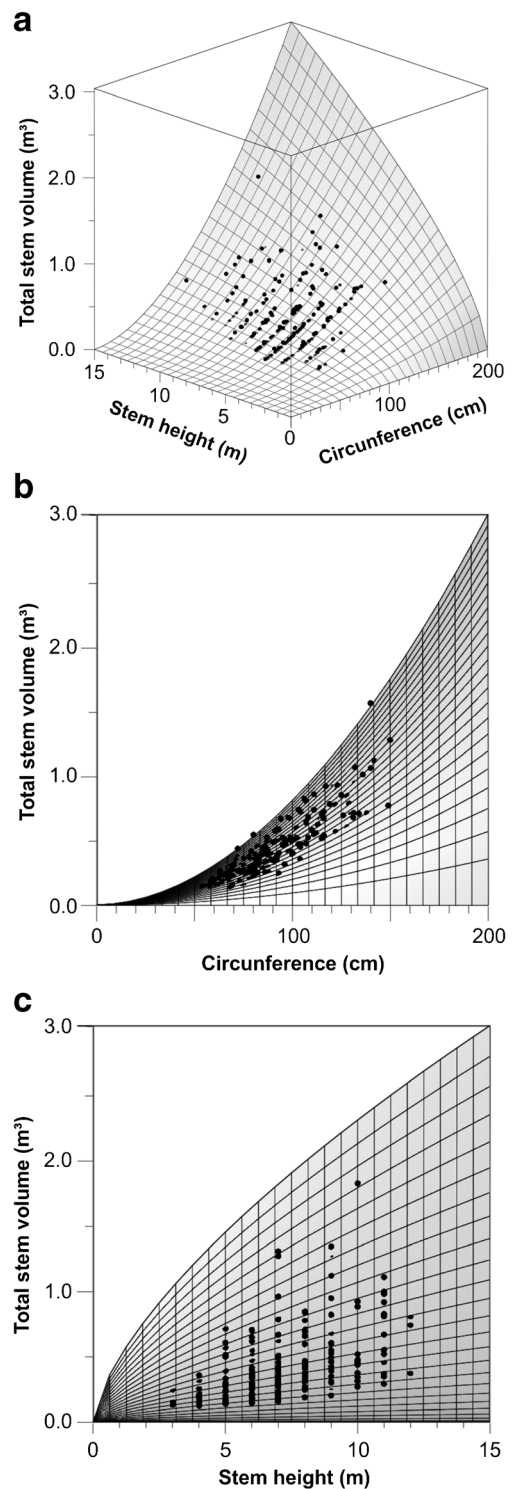
forest, seasonal semideciduous forest and seasonal deciduous forest. Rolim et al. (2006) and Colpini et al. (2009) also used this logarithmic model generically for open ombrophilous forest for the Brazilian state of Mato Grosso and for both open and dense ombrophilous forest for the Brazilian state of Pará.

Logarithmic models constructed by the first forest inventories of southern Brazil (Brasil 1983; Netto 1984) are often used for broadleaved species (Glufke et al. 1994; Vaccaro 2002; Figueiredo-Filho et al. 2006) and for Brazilian pine (Figueiredo-Filho et al. 2006; Vibrans et al. 2012b). Machado et al. (2008) reported that Schumacher and Hall (1933) model produced most accurate predictions for the cultural and ecological important species *M. scabrella* Benth. in the MIX forest type, in southern Brazil. Scolforo et al. (2008) also adjusted specific logarithmic models for two native tannin-producing species (*Anadenanthera colubrina* (Benth.) Brenan and *Stryphnodendron adstringens* (Mart.) Coville), for a cork-producing species (*Kielmeyera coriacea* Mart. & Zucc.) and for an oil-producing species (*Eremanthus erythropappus* (DC.)



**Fig. 3** Comparison between measured (Smalian method) and predicted values of  $V_s$ ; estimates were evaluated through model no. 8, adjusted for generic models of seasonal deciduous forest (a), mixed ombrophilous forest (b), dense ombrophilous forest (c) and the overall generic model (d)

MacLeish). Going out of the ‘logarithmic trend’, Spathelf et al. (2001) adjusted specific volume models for the taxa *Ilex*



**Fig. 4** Predicted  $V_s$  surface (model no. 8) and measured data (points) for the generic model for seasonal deciduous forest (a), binary relationship between  $c$  and  $V_s$  (b) and between  $h_s$  and  $V_s$  (c), where  $h_s$  = stem height,  $d$  = diameter at breast height and  $c$  = circumference at breast height

*brevicuspis* Reissek., *Erythroxylum deciduum* A.St.-Hil., *Cytherexylum montevidense* Spreng. and *Myrsine coriacea* (Sw.) R. Br. in Rio Grande do Sul state and founded that the second-order polynomial model presented the best fitting.

Despite that studies conducted in Brazil performed a model selection procedure, several critics may be useful and relevant for Brazilian forest research community dedicated to construct reliable models and estimates. The major part of cited authors did not test for satisfaction of regression assumptions and models' goodness of fit. The logarithmic models are frequently used in Brazil, although very few studies applied the back transformation from log-log scale to original scale using the correcting factor (Baskerville 1972; Sprugel 1983). Also, the  $R^2$  is often used to describe model's adjustment quality; however, very few studies calculated the  $R^2$  for back-transformed data (or original scale), evidencing a misleading use of  $R^2$ , since it has limitations of use on nonlinear models (Anderson-Sprecher 1994). Other parameters like RMSE are rarely calculated, based on residuals' original scale. Also, robust methods for availing model adjustment quality like AIC or Bayesian information criterion are rarely used and should be incorporated on allometric model adjustment routines.

Regarding the obtained results in this study for species-specific models, they did not present an expressive lower AIC compared to generic models for the three forest types, as found by Brandeis et al. (2005), who used Puerto Rican forest inventory data of subtropical moist and wet forests. The significant results from  $F$  tests for comparisons of generic and species-specific models (DEN) and overall generic and forest-type generic models may occur due to the greater sample size and number of considered species. Indeed, this fact seems to introduce a considerable amount of residual variance in these generic models due to the large number of species and its allometric relationships variability. This result was verified by a closer look on the algorithm of  $F$  test, where sample size and, consequently, the squared errors, have substantial influence on the final ratio.

As for DEC and MIX, the  $F$  tests for comparisons between generic and species-specific models presented no evidences for rejection of the null hypothesis; this means that those specific models do not present statistically significantly better result than generic models. However, this was not the case at DEN, where  $F$  test showed evidences for rejection of the null hypothesis, created by the major amount of residual variance of the nested model, much more than considered for DEC and MIX. Thus, besides these different findings, forest-type generic models embrace larger ranges of  $d$  (Table 1) and species (57 in DEC, 107 in MIX and 207 in DEN), allowing broader applications.

The final comparison between forest-type generic and overall generic models also showed evidences for rejection of the null hypothesis ( $F=16.70$ ;  $p<0.01$ ). This result, together with lack of fit of the overall generic model (Table 3), suggests that volume estimates should be computed separately by forest type, due to dissemblance of their species composition and structure. Schorn et al. (2012) and Gasper et al. (2012) found important structural differences and low floristic

similarity between these forest types in Santa Catarina (Sørensen Similarity Index): DEC/DEN (0.42), DEC/MIX (0.58) and MIX/DEN (0.62), and only 165 species are common to all three forest types, amidst a total of 840 tree and shrub species with  $d\geq 10$  cm.

However, McRoberts and Westfall (2014) applied species-specific and generic models for conifer and deciduous species using data from the US National Forest Inventory and found that generic models produced smaller deviations on predictions than species-specific models. Therefore, these results suggest that (for large area and generic volume estimations), the use of specific models is not mandatory. The authors stress that national forest inventories should concentrate their efforts in reducing sample variability by using bigger sample sizes and statistically more confident sample designs. On the other hand, tropical and subtropical forests with high species richness, floristic and structural heterogeneity and higher degree of naturalness compared to 'domesticated' temperate and boreal forests may require the use of specific models, regarding forest types or regions, for example.

## 5 Conclusions and recommendations

The following conclusions may be drawn from this study. First, as the assumptions underlying the use of regression methods for constructing volume models were validated, the results allow estimates with valid confidence intervals. Second, because of their quality of fit, the adjusted logarithmic models compose useful tools for predicting total stem volume (with bark) for forest remnants in southern Brazil. Third, contrary to what was expected, generic models for two forest types produced estimates not less reliable than species-specific ones, concerning descriptive statistics, and can be used for large-scale applications, where specific volume estimates are not required. Fourth, stratification of dataset by forest type for model fitting showed to be necessary, as significant results were found in model identity tests among forest types; these could be induced by ecological drivers such as species composition and structure, competition, environmental conditions, among others. Further studies should elucidate drivers of stem form differentiation between forest types and geographic regions.

**Acknowledgments** The authors are grateful to Ronald E. McRoberts (US Forest Service) for the valuable review and comments on this manuscript.

**Funding** We 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 and to FURB and CAPES for the Master's scholarship granted to the third and fourth authors.



## References

- Anderson-Sprecher R (1994) Model comparisons and  $R^2$ . *The American Statistician* 48:113–117
- Avery TE, Burkhardt HE (2002) *Forest measurements*, 5th edn. McGraw-Hill, New York
- Azevedo CP, Sanquetta CR, Silva JN, Machado SA, Souza CR, Oliveira MM (2008) Simulação de estratégias de manejo florestal na Amazônia com o uso do modelo SYMFOR. *Acta Amaz* 38:51–69
- Baskerville GL (1972) Use of logarithmic regression in the estimation of plant biomass. *Can J For Res* 2:49–53
- Basuki TM, van Laake PE, Skidmore AK, Hussin YA (2009) Allometric models for estimating above-ground biomass in tropical lowland Dipterocarp forests. *For Ecol Manag* 257:1684–1694
- Borsoi GA (2004) Subsídios para o manejo de uma Floresta Ombrófila Mista em estágio avançado de regeneração natural. Doctoral thesis. Universidade Federal de Santa Maria, Brazil
- Box GEP, Cox DR (1964) An analysis of transformation. *J R Stat Soc* 26: 211–243
- Brandeis TJ, Kuegler O, Knowe SA (2005) Equations for merchantable volume for subtropical moist and wet forests of Puerto Rico. United States Department of Agriculture-Forest Service. Research Paper SRS-39
- Braz EM, Schneider PR, Mattos PP, Thaines F, Selle GL, Oliveira MF, Oliveira TC (2012) Manejo da estrutura diamétrica remanescente de florestas tropicais. *Ciênc Flor* 22:787–794
- Brasil (1983) Ministério da Agricultura. Instituto Brasileiro de Desenvolvimento Florestal. Inventário florestal nacional: Florestas nativas do Rio Grande do Sul. IBDF, Brasília
- Breidenbach J, Antón-Fernández C, Petersson H, Astrup P, McRoberts RE (2014) Quantifying the contribution of biomass model errors to the uncertainty of biomass stock and change estimates in Norway. *For Sci* 60:25–33
- Brown S, Gillespie AJR, Lugo AE (1989) Biomass estimation methods for tropical forests with application to forest inventory data. *For Sci* 35:881–902
- Chave J, Andalo C, Brown S, Cairns MA, Chambers JQ, Eamus D, Fölster H, Fromard F, Higuchi N, Kira T, Lescure J, Nelson BW, Ogawa H, Puig H, Riéra B, Yamakura T (2005) Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecol* 145:87–99
- Colpini C, Travagin DP, Soares TS, Silva VSM (2009) Determinação do volume, do fator de forma e da porcentagem de casca de árvores individuais em uma Floresta ombrófila Aberta na região noroeste de Mato Grosso. *Act Amaz* 39:97–104
- Cormier KL, Reich RM, Czaplewski RL, Bechtold WA (1992) Evaluation of weighted regression and sample size in developing a taper model for loblolly pine. *For Ecol Manag* 53:65–76
- EPAGRI. Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina, (2002) Atlas Climatológico do Estado de Santa Catarina. EPAGRI, Florianópolis. CD-ROM
- Figueiredo-Filho A, Dias AN, Watzlawick LF (2006) Inventário das florestas naturais na Floresta Nacional de Irati, Estado do Paraná. UNICENTRO, Irati
- Finger CAG (1992) Fundamentos de biometria florestal. UFSM/CEPEF/FATEC, Santa Maria
- Freitas JV, Oliveira YM, Rosot MAD, Gomide GLA, Mattos PP (2010) National forest inventory reports: Brazil. In: Tomppo E, Gschwantner T, Lawrence M, McRoberts RE (eds) *National Forest Inventories: pathways for common reporting*. Springer, London, pp 89–96
- Gasper AL, Meyer L, Sevegnani L, Sobral MEG, Bonnet A (2012) Flora vascular de Santa Catarina. In: Vibrans AC, Sevegnani L, Gasper AL, Lingner DV (eds) *Diversidade e conservação dos remanescentes florestais*. Edifurb, Blumenau, pp 99–111
- Glufke C, Mainardi GL, Schneider PR, Alvarez Filho A (1994) Produção de uma floresta natural em Santa Maria, RS. *Ciênc Flor* 4:61–76
- Guendehou GHS, Lehtonen A, Moudachirou M, Mäkipää R, Sinsin B (2012) Stem biomass and volume models of selected tropical tree species in West Africa. *South For* 74:77–88
- IBGE-Instituto Brasileiro de Geografia e Estatística, (2012) *Manuais Técnicos em Geociências. Manual Técnico da Vegetação Brasileira*, 2<sup>th</sup> edn. IBGE, Rio de Janeiro
- Klein RM (1978) Mapa fitogeográfico do estado de Santa Catarina. *Flor Ilust Catar* 5:1–24
- Kottek M, Grieser J, Beck C, Rudolf B, Rubel F (2006) World map of the Köppen-Geiger climate classification updated. *Meteorol Z* 15:259–263
- Litton CM, Kauffman JB (2008) Allometric models for predicting above-ground biomass in two widespread woody plants in Hawaii. *Biotrop* 40:313–320
- Machado SA, Figura MA, Silva LCR, Téo SJ, Stolle L, Urbano E (2008) Modelagem volumétrica para bracinga (*Mimosa scabrella*) em povoações da Região Metropolitana de Curitiba. *Pesq Flor Bras* 56:17–29
- Mayer DG, Butler DG (1993) Statistical validation. *Ecol Model* 68:21–32
- McRoberts RE, Westfall J (2014) Effects of uncertainty in model predictions of individual tree volume on large area volume estimates. *For Sci* 60:34–42
- McRoberts RE, Moser P, Oliveira LZ, Vibrans AC (2015) 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. *Can J For Res* 45:44–51
- Neter J, Kutner M, Wasserman W, Nachtsheim C (1996) *Applied linear statistical models*. McGraw-Hill, Boston
- Netto SP (1984) Instituto Brasileiro de Desenvolvimento Florestal. Inventário florestal nacional: florestas nativas, Paraná, Santa Catarina. IBDF, Brasília
- Oliveira-Filho AT (2009) Classificação das fitofisionomias da América do Sul cisandina tropical e subtropical: proposta de um novo sistema-prático e flexível-ou uma injeção a mais de caos? *Rodriguésia* 60:237–258
- Picard N, Saint-André L, Henry M (2012) *Manual for building tree volume and biomass allometric equations: from field measurement to prediction*. FAO, Rome and CIRAD, Montpellier
- Ratkowsky DA (1983) *Nonlinear regression modeling: a unified practical Approach*. Mekker
- Rolim SG, Couto HTZ, Jesus RM, França JT (2006) Modelos volumétricos para a Floresta Nacional do Tapirapé-Aquiri, Serra dos Carajás (PA). *Act Amaz* 36:107–114
- Schorn LA, Gasper AL, Meyer L, Vibrans AC (2012) Síntese da estrutura dos remanescentes florestais de Santa Catarina. In: Vibrans AC, Sevegnani L, Gasper AL, Lingner DV (eds) *Diversidade e conservação dos remanescentes florestais*. Edifurb, Blumenau, pp 125–140
- Schumacher F, Hall F (1933) Logarithmic expression of timber-tree volume. *J Agric Res* 47:719–734
- Schumacher FX (1942) *Forest mensuration*, 1st edn. McGraw-Hill, New York
- Scolforo JR, Oliveira AD, Acerbi Junior FW (eds) (2008) *Inventário Florestal de Minas Gerais: Equações de volume, Peso de Matéria Seca e Carbono para diferentes fisionomias da flora nativa*. Editora UFLA, Lavras
- Segura M, Kanninen M (2005) Allometric models for tree volume and total aboveground biomass in a tropical humid forest in Costa Rica. *Biotropica* 37:2–8
- Silva JAA, Melo MRCS, Borders BE (1993) A volume equation for mangrove trees in northeast Brazil. *For Ecol Manag* 58:129–136
- Spathelf P, Berger R, Vaccaro S, Tonini H, Borsoi GA (2001) Growth of native species of the Deciduous/Dense Tropical Forest of Rio Grande do Sul. *Ciênc Flor* 11:103–119

- Sprugel DG (1983) Correcting for bias in log-transformed allometric equations. *Ecol* 64:209–210
- Spurr SH (1952) *Forestry inventory*. Ronald Press, New York
- Stoate T (1945) The use of a volume equation in pine stands. *Aust For* 9: 48–52
- Vaccaro S (2002) Crescimento de uma Floresta Estacional Decidual, em três estágios sucessionais, no município de Santa Tereza, RS, Brasil. Doctoral thesis. Universidade Federal de Santa Maria, Brazil
- Vibrans AC, Sevegnani L, Lingner DV, Gasper AL, Sabbagh S (2010) The Floristic and Forest Inventory of Santa Catarina State (IFFSC): methodological and operational aspects. *Braz J For Res* 30:291–302
- Vibrans AC, Moser P, Lingner DV, Maçaneiro JP (2012a) Análise estatística do IFFSC e estimativas dendrométricas. In: Vibrans AC, Sevegnani L, Gasper AL, Lingner DV (eds) *Diversidade e conservação dos remanescentes florestais*. Edifurb, Blumenau, pp 79–95
- Vibrans AC, Moser P, Maçaneiro JP, Lingner DV, Silva LS (2012b) Equações hipsométricas, volumétricas e de peso seco para Floresta Ombrófila Mista em Santa Catarina. In: Vibrans AC, Sevegnani L, Gasper AL, Lingner DV (eds) *Floresta Ombrófila Mista*. Edifurb, Blumenau, pp 95–127
- Zianis D, Muukkonen P, Mäkipää R, Mencuccini M (2005) Biomass and stem volume equations for tree species in Europe. *Silva Fenn. Monographs* 4 p 63 <http://www.metla.fi/silvafennica/sfmono.htm>
- Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) *Mixed effects models and extensions in ecology with R*. Springer, New York