

SITE INDEX IN EUCALYPTUS STANDS APPLYING ORDINARY KRIGING: AN APPROACH WITH DIFFERENT MODELS AND METHODS OF CLASSIFICATION

Pedro Vaz da Rocha^{1*}, Emanuel José Gomes de Araújo², Vinícius Augusto Morais³, Marco Antonio Monte², Danilo Henrique dos Santos Ataíde¹, Laís Candido Silva¹

^{1*}Programa de Pós-Graduação em Ciências Ambientais e Florestais, Universidade Federal Rural do Rio de Janeiro, Seropédica, Rio de Janeiro, Brasil - pedrovaz-522@hotmail.com

¹Programa de Pós-Graduação em Ciências Ambientais e Florestais, Universidade Federal Rural do Rio de Janeiro, Seropédica, Rio de Janeiro, Brasil - daniloataide.florestal@gmail.com; laiscandido30@hotmail.com

²Universidade Federal Rural do Rio de Janeiro, Instituto de Florestas, Departamento de Silvicultura, Seropédica, Rio de Janeiro, Brasil - ejgaraujo@gmail.com; marcomonte.ufrrj@gmail.com

³Universidade do Estado de Mato Grosso, Alta Floresta, Mato Grosso, Brasil – viniciusmorais@unemat.br

Received for publication: 14/07/2020 – Accepted for publication: 20/05/2021

Resumo

Índice de sítio em povoamentos de eucalipto aplicando krigagem ordinária: uma abordagem com diferentes modelos e métodos de classificação. O objetivo deste trabalho foi avaliar a eficiência de modelos e métodos para obtenção do índice de sítio, associados com a krigagem ordinária, na classificação da capacidade produtiva em povoamentos de eucalipto. Assim, a capacidade produtiva foi obtida pela modelagem tradicional em povoamentos clonais (2.119 hectares) localizados em Minas Gerais, Brasil. Foram alocadas aleatoriamente 170 unidades amostrais (*ua*) de 400m², representando uma intensidade amostral de 0,32%. A altura dominante das árvores (Assmann) foi medida aos 24, 36, 48, 60, 72, e 84 meses. O índice de sítio foi estimado pelos métodos da curva-guia e diferença algébrica, utilizando os modelos de Schumacher, Chapman e Richards, e Bailey e Clutter. No ajuste foram utilizadas 136 *ua* e na validação preditiva, 34 *ua*. A dependência espacial do índice de sítio foi avaliada pelo semivariograma experimental e ajuste dos modelos exponencial, esférico e gaussiano. Comprovada a dependência espacial, foi realizada a krigagem ordinária para espacializar o índice de sítio. Na validação preditiva, utilizaram-se os valores de altura dominante aos 72 meses. O método da diferença algébrica proporcionou ótimas estimativas do índice de sítio, que apresentou dependência espacial em todos os ajustes, de moderada a forte. Em geral, o modelo gaussiano foi o mais acurado. Conclui-se que o método da diferença algébrica foi mais eficiente e o índice de sítio apresentou forte dependência espacial em todas as idades, independente do modelo utilizado. Assim, os modelos de regressão para estimativa do índice de sítio podem ser utilizados em combinação com as técnicas de krigagem ordinária.

Palavras-chave: Capacidade produtiva, Dependência espacial, Geoestatística, Modelagem.

Abstract

The objective of this work was to evaluate the efficiency of models and methods to obtain the site index, associated with ordinary kriging, to classify productive capacity in eucalyptus stands. Thus, the site quality was performed considering the traditional modeling in clonal stands (2,119 hectares) located in Minas Gerais state, Brazil. 170 plots of 400m² were randomly allocated, representing a sampling intensity of 0.32%. The dominant height of trees (Assmann) was measured at 24, 36, 48, 60, 72, and 84 months. The site index (S) was estimated by the guide curve and algebraic difference methods, using the models of Schumacher, Chapman and Richards, and Bailey and Clutter. 136 plots were used in the fit and 34 plots in the predictive validation. The spatial dependence of site index was evaluated by experimental semivariogram and adjustment of exponential, spherical, and gaussian models. After confirming the spatial dependence, ordinary kriging was performed to spatialize the site index. For the predictive validation, the dominant height values at 72 months were used. The algebraic difference method provided excellent estimates of site index, which showed spatial dependence in all adjustments, from moderate to strong. In most cases, the gaussian model was the most accurate. It is concluded that the algebraic difference method was more efficient and the site index showed strong spatial dependence at all ages, regardless of the model used. Thus, regression models for site index estimation can be used in combination with ordinary kriging techniques.

Keywords: Productive capacity, Spatial dependence, Geostatistics, Modeling.

INTRODUCTION

The productive capacity refers to the productive potential of wood of a certain area and species, and it can be performed using indirect or direct classification methods (CAMPOS; LEITE, 2017). Direct methods make it possible to represent the productive potential through a quantitative value, called the site index. The site index is a robust indicator of site quality, used to help apply management techniques and assess the economic viability of forest plantations, in addition to being a fundamental variable in growth and production models (LEITE *et al.*, 2011; CAMPOS; LEITE, 2017).

The application of direct methods to classify productive capacity is considered efficient (LEITE *et al.*, 2011), based on the use of regression models that relate dendrometric variables with the age of the population. Among the variables that can be used as an indicator of the site quality, the volume, basal area, total height, and dominant height of eucalyptus stand out (SCOLFORO, 2006). However, the dominant height is considered the variable that provides the most consistent results once it is low influenced by the density of the population and presents a significant correlation with the volume of wood.

The methods generally used to classify the productive capacity do not consider the spatial dependence of the dominant height variable, not even the spatialization of the site index. Some studies have confirmed the strong spatial dependence of the dominant height in commercial stands of *Eucalyptus* sp. (MELLO *et al.*, 2005; GUEDES *et al.*, 2015; LUNDGREN *et al.*, 2017; ATAÍDE *et al.*, 2020). In addition, the regression models provide the average estimate of growth in dominant height and allow site index estimates only in places where there was a measured sample unit. In this case, for a more detailed mapping of the productive capacity, it would be necessary to increase the sampling intensity to represent different locations, increasing sampling costs.

Ordinary kriging is the option commonly used in the spatialization of dendrometric variables in non-sampled locations. It can be cited studies carried out with diameter at 1.30 m from the ground, basal area, the total and dominant height, volume, and average annual increment, being applied for *Eucalyptus* sp. stands, as well as to other species, for example, *Pinus* sp. and *Tectona grandis* (MELLO *et al.*, 2005; PELISSARI *et al.*, 2014; GUEDES *et al.*, 2015; PELISSARI *et al.*, 2015; LUNDGREN *et al.*, 2017; ZECH *et al.*, 2018; ATAÍDE *et al.*, 2020). Even with numerous studies evaluating the spatial structuring of dendrometric variables, few use the site index variable, mainly when obtained by mathematical models. Thus, the results this variable obtained in different adjustment methods and models are still unknown in the spatial estimates of the variable.

Considering that the dominant height has proven spatial dependence, it is possible to assume that the site index will show strong spatial dependence, as the site index is the dominant height at the reference age. If accepted, this hypothesis would make it possible to spatialize the classes of productive potential in forest stands, generating relevant spatial information for management. As there are different procedures to estimate the site index and different geostatistical models to obtain spatial estimates of a variable, it is important to evaluate which ones provide satisfactory results in estimating the site index and spatialization of the productive potential classes.

This work hypothesizes that the choice of the regression model and the method for constructing the site curves influence the spatial continuity of the site index and the spatial estimates through ordinary kriging. Therefore, the objective of this work was to evaluate the effect of different models and methods on the spatial dependence structure of the site index and the application of ordinary kriging to spatialize the productive capacity in eucalyptus stands.

MATERIALS AND METHODS

Study area

The study was carried out in 2,119 hectares with clonal *Eucalyptus* sp. stands, located at coordinates 17°17'26.9" S, 43°42'33.7" W, in the municipality of Bocaiúva, Minas Gerais state. The local climate is of the Aw type, according to the Köppen climate classification, characterized as being tropical wet savanna, dry winters, and rainy summers (ALVARES *et al.*, 2013). The region has an average altitude of 820 m, with rainfall and an average annual temperature of 1,246 mm and 24°C, respectively. The predominant soils are Dystrophic Dark Red Latosol or Dystrophic Red Yellow Latosol.

Data collection and description

Information was obtained from plots allocated in 62 blocks of the stands, with the same genetic material and initial spacing. 170 square plots were selected with a fixed area of 400m², representing a sampling intensity of 0.32% of the total area (one plot for every 12.46 hectares). The distances between the plot ranged from approximately 200 to 44,354 m. In all plots, the dominant heights of trees (Hd) were obtained following the concept of Assmann (1970), at ages 24, 36, 48, 60, 72, and 84 months. The plots were georeferenced based on their central geographic coordinates. The plots were divided into two groups: 136 (1 plot for every 15.58 hectares) were used to adjust the models and evaluate the productive capacity classification procedures; 34 (1 plot for every 62.32 hectares) were used for the predictive validation of the estimates of the best selected procedures. Validation plots were selected throughout the study area, considering at least one validation plot per stand, when possible.

Classification of the productivity capacity

The site index was estimated considering the guide curve and algebraic difference methods, with a reference age of 72 months. In each method, three regression models were adjusted to estimate the dominant height as a function of age (Table 1). The adjustments were realized using the nlstools package (BATY *et al.*, 2015) in

the R environment (R CORE TEAM, 2015). Six equations were obtained to estimate the site index (Table 1) after rearranging the three fitted models, using the guide curve and algebraic difference methods.

The adjusted equations were evaluated using the adjusted coefficient of determination ($R^2_{adj.}$), standard error of the estimate in percentage ($S_{yx}\%$), root mean square error (RMSE), the mean error of prediction in percentage (bias%), and graphical analysis of the normalized residuals. The significance of the coefficients was also verified by the t test and the normality of the residuals using the Shapiro-Wilk test, both with 5% probability ($\alpha=0.05$).

Table 1. Models used to estimate site index through guide curve and algebraic difference methods.

Tabela 1. Modelos utilizados para estimar índice de sítio pelos métodos da curva-guia e diferença algébrica.

Original model	Author	Model by the Guide Curve method	Model by the Algebraic Difference method
$Ln(Hd_i) = \beta_0 + \beta_1 \left(\frac{1}{I}\right) + \varepsilon_i$	Schumacher (1939)	$S = Hd + e \left[\hat{\beta}_1 \left(\frac{1}{I_{ref}} - \frac{1}{I} \right) \right]$	$S = Hd e^{\left[\hat{\beta}_1 \left(\frac{1}{I} - \frac{1}{I_{ref}} \right) \right]}$
$Hd_i = \beta_1 (1 - e^{-\beta_2 I})^{\left(\frac{1}{1-\beta_3} \right)} + \varepsilon_i$	Chapman and Richards (1959)	$S = Hd \left[\frac{1 - e^{(-\hat{\beta}_2 I_{ref})}}{1 - e^{(-\hat{\beta}_2 I)}} \right]^{\left(\frac{1}{1-\hat{\beta}_3} \right)}$	$S = Hd \left[\frac{1 - e^{(-\hat{\beta}_1 I_{ref})}}{1 - e^{(-\hat{\beta}_1 I)}} \right]^{\hat{\beta}_2}$
$Ln(Hd_i) = \beta_0 + \beta_1 \left(\frac{1}{I}\right)^{\beta_2} + \varepsilon_i$	Bailey and Clutter (1974)	$S = Hd + e^{\left\{ \hat{\beta}_1 \left[\left(\frac{1}{I_{ref}} \right)^{\hat{\beta}_2} - \left(\frac{1}{I} \right)^{\hat{\beta}_2} \right] \right\}}$	$S = \frac{Hd}{e^{\hat{\beta}_1 (I - \hat{\beta}_2 I_{ref} - \hat{\beta}_2 I_{ref}^2)}}$

Subtitle: Ln = neperian logarithm; Hd_i = dominant height of trees (m); β_0 , β_1 , β_2 and β_3 = model parameters; I = age of population (months); S the site index (m); I_{ref} = the reference age (months); e = exponential; ε_i = residue.

The spatial continuity structure of the site index, estimated in all combinations of methods and models, was evaluated using the experimental semivariogram. The anisotropy was evaluated by constructing directional semivariograms (0, 45, 90, and 135°), which confirmed the isotropic behavior of the variable. The exponential, spherical, and gaussian models were fitted to the omnidirectional experimental semivariogram by the Maximum Likelihood Method, using the geoR package (RIBEIRO JÚNIOR; DIGGLE, 2001) in the R software (R CORE TEAM, 2015).

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad \text{Experimental semivariogram}$$

$$\hat{\gamma}(h) = C_0 + C \left[1 - e^{\left(-\frac{h}{a} \right)} \right] \quad \text{Exponential model}$$

$$\hat{\gamma}(h) = C_0 + C \left[1.5 \frac{h}{a} - 0.5 \left(\frac{h}{a} \right)^3 \right] \quad \text{if: } h < a \quad \text{Spherical model}$$

$$\hat{\gamma}(h) = C_0 + C \quad \text{if: } h \geq a$$

$$\hat{\gamma}(h) = C_0 + C \left[1 - e^{\left(-\frac{h^2}{a^2} \right)} \right] \quad \text{Gaussian model}$$

Where: $\hat{\gamma}(h)$ the semivariance, $Z(x_i)$ the value of the regionalized variable at point x , $Z(x_i + h)$ the value of the variable at point, $x + h$, $N(h)$ the number of pairs separated by the distance h , C_0 the nugget effect, C the contribution, a the reach.

The criteria used to select the best semivariance model were: Akaike's information criterion (AIC), reduced mean error (RME) and standard deviation of the reduced error (SDRE) obtained by cross-validation (ARAÚJO *et al.*, 2019). The spatial dependence index (SDI) was obtained for each of the theoretical models adjusted to semivariograms, classified as weak ($SDI \leq 0.25$), moderate ($0.25 < SDI \leq 0.75$), and strong ($IDE > 0.75$) spatial dependence (ZIMBACK, 2003). Once the spatial dependence of the site index was confirmed, ordinary kriging was applied to obtain the spatial estimates of the site index in the non-sampled locations.

Ordinary kriging was used to predict the spatial estimates of site index obtained in each regression model adjusted by the guide curve and algebraic difference methods. The spatial estimates of the ordinary kriging in the

136 plots used in the fit were compared with the values estimated by the regression models. For this, was used the values of the RMSE and the bias%, in addition to the values of Pearson's correlation, Willmott's agreement index, and performance index.

$$SDI = \frac{c}{(C_0 + c)} \quad \text{Spatial dependence index}$$

$$c = dr \quad \text{Performance index}$$

Where: SDI the spatial dependence index, C the contribution, C_0 the nugget effect, c the performance index, r the Pearson correlation coefficient and d the Willmott agreement index.

Predictive validation

The 34 plots not used in the fit of the regression models were used to perform the predictive validation of the site index estimated from the regression models and ordinary kriging. The site index estimates by regression and ordinary kriging were compared with the observed values of the mean height of the dominant trees in the validation plots, at the reference age of 72 months. In ordinary kriging, the maps were converted to raster format with pixel dimensions of 20 m x 20 m, presenting an area equal to the size of the plot. In the maps, the 34 plots separated for validation were inserted and, later, the estimated values of the site index were extracted. The extracted values were compared with the dominant height of the trees at the reference age of 72 months, through the difference between the dominant height observed in the plot and the limit of the respective class where it was positioned.

RESULTS

Adjustment of models for classification of productive capacity

The coefficients of the regression models were significant by the t-test ($p < 0.05$) for all adjustments made (Table 2). The adjusted coefficient of determination ($R^2_{adj.}$) was higher than 0.80 for the two methods of evaluation of the productive capacity. However, the other statistical criteria showed different results between the methods and models, indicating the superiority of the Bailey and Clutter model in the algebraic difference method. The standard error of the estimate in percentage ($S_{yx}\%$) was less than 10% in all adjustments, but with values less than 7% only when the algebraic difference method was used. A similar result was obtained for the RMSE estimates, in which the algebraic difference method provided improvements of about 25% concerning the guide curve, with values around 1.5 m. Regarding the bias% there was a distinct trend: while the guide curve method tended to overestimate, the algebraic difference provided underestimates. In the graphical analysis of the normalized residues (Figure 1), it is observed that the models fitted in the algebraic difference method showed greater homogeneity in the dispersion of residues.

Table 2. Parameters and statistics of adjustment of the models to estimate a dominant height in function of the age. Tabela 2. Parâmetros e estatísticas de ajuste dos modelos para estimar a altura dominante em função da idade.

Method	Model	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$R^2_{adj.}$	$S_{yx}\%$	RMSE (m)	bias%
Guide-curve	Schumacher	3.56*	-24.28*			0.857	9.715	2.076	-0.674
	Chapman and Richards		29.39*	0.03*	0.25*	0.859	9.623	2.055	-1.200
	Bailey and Clutter	3.82*	-11.74*	0.70*		0.857	9.710	2.074	-0.659
Algebraic difference	Schumacher		23.03*			0.840	6.440	1.498	0.657
	Chapman and Richards		0.036*	1.19*		0.818	6.858	1.594	2.913
	Bailey and Clutter		-14.86*	0.81*		0.843	6.363	1.479	0.158

Subtitle: $\beta_0, \beta_1, \beta_2$ e β_3 the coefficients of the adjusted models; $R^2_{adj.}$ the adjusted coefficient of determination; $S_{yx}\%$ the standard error of estimate in percentage; RMSE the root mean square error; bias% the mean error of prediction in percentage; * significant value with 5% probability by t-test.

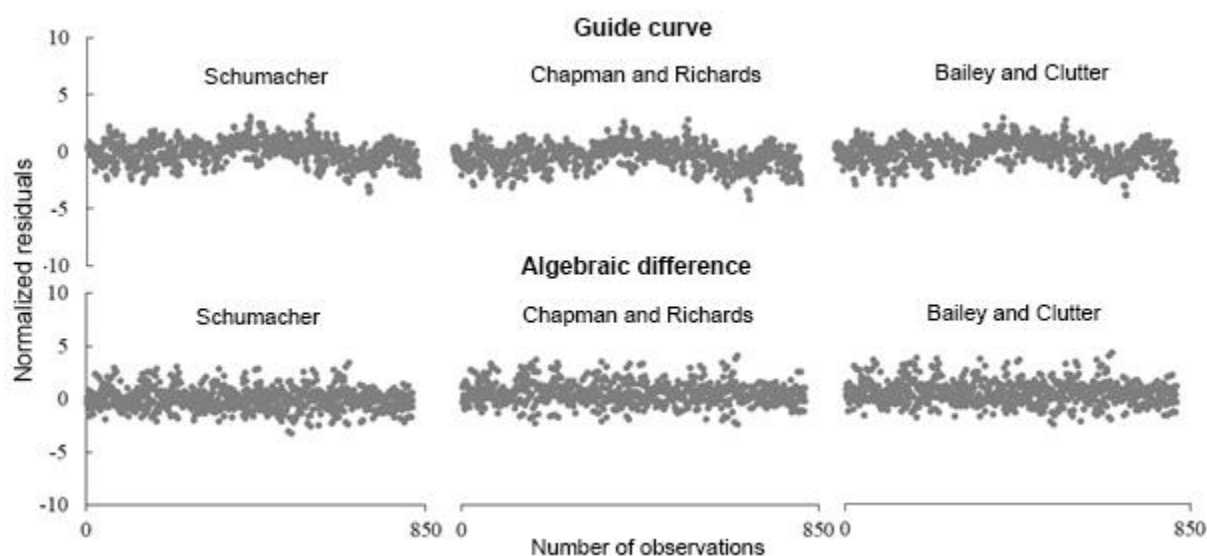


Figure 1. Normalized residuals of the adjusted equations of the Schumacher, Chapman and Richards, and Bailey and Clutter models in the guide curve and algebraic difference methods.

Figura 1. Resíduos normalizados das equações ajustadas dos modelos de Schumacher, Chapman e Richards, e Bailey e Clutter nos métodos da curva-guia e diferença algébrica.

Spatial analysis of the site index

The site index presented spatial dependence ranging from moderate ($0.25 < SDI \leq 0.75$) to strong ($SDI > 0.75$) in all regression models used (Schumacher, Chapman and Richards, Bailey and Clutter) in the two methods of construction of site curves (guide curve and algebraic difference) (Table 3). The SDI ranged from 63.42% to 90.75%, indicating a predominantly strong spatial dependence of the site index variable. The spherical, exponential, and gaussian models presented good fits to the experimental semivariogram, in all the productive capacity classification procedures, with RME values close to 0 and SDRE close to 1.

Table 3. Parameters, adjustment statistics and spatial dependency index of the selected models.

Tabela 3. Parâmetros, estatísticas de ajuste e índice de dependência espacial dos modelos selecionados.

Model	Age	Theoretical model	C ₀	C	A	RME	SDRE	AIC	SDI%
Guide Curve									
Schumacher	24	Spherical	1.71	16.81	4,888	0.0109	1.032	678.8	90.75
	36	Gaussian	1.89	5.63	1,038	0.0050	1.021	615.9	74.86
	48	Exponential	1.80	3.28	2,604	0.0018	1.042	565.5	64.61
	60	Gaussian	1.14	3.27	1,108	0.0002	1.036	540.9	74.23
	72	Gaussian	1.29	4.35	1,471	-0.0021	1.039	553.0	77.11
	84	Gaussian	1.41	3.77	1,758	-0.0026	1.025	540.0	72.81
Chapman and Richards	24	Gaussian	3.35	13.39	2,258	0.0071	1.034	668.1	79.98
	36	Gaussian	2.01	5.96	1,055	0.0049	1.021	622.9	74.76
	48	Exponential	1.87	3.24	2,762	0.0027	1.037	566.6	63.42
	60	Gaussian	1.14	3.23	1,102	0.0000	1.037	540.7	73.86
	72	Gaussian	1.29	4.36	1,471	-0.0022	1.039	553.1	77.15
	84	Gaussian	1.41	3.84	1,749	-0.0025	1.025	541.4	73.16
Bailey and Clutter	24	Gaussian	3.35	13.56	2,257	0.0072	1.034	668.8	80.17
	36	Gaussian	2.00	5.96	1,033	0.0048	1.021	623.8	74.81
	48	Exponential	1.91	3.33	2,733	0.0027	1.038	570.1	63.59
	60	Gaussian	1.16	3.21	1,097	0.0000	1.037	541.9	73.45
	72	Gaussian	1.30	4.15	1,477	-0.0025	1.040	550.5	76.18
	84	Gaussian	1.38	3.54	1,781	-0.0027	1.026	534.8	71.92

Algebraic Difference									
Schumacher	24	Gaussian	3.24	13.62	2,242	0.0073	1.031	666.5	80.77
	36	Gaussian	1.80	5.43	1,002	0.0043	1.016	607.9	75.05
	48	Exponential	1.75	3.28	2,530	0.0019	1.042	564.4	65.24
	60	Gaussian	1.13	3.30	1,112	0.0003	1.036	540.8	74.49
	72	Gaussian	1.29	4.39	1,470	-0.0022	1.039	553.6	77.31
	84	Gaussian	1.41	3.81	1,754	-0.0026	1.025	541.0	72.97
Chapman and Richards	24	Gaussian	2.44	9.31	2,247	0.0072	1.038	623.0	79.25
	36	Gaussian	1.65	5.03	968	0.0048	1.016	598.7	75.24
	48	Exponential	1.66	3.25	2,428	0.0021	1.042	560.7	66.20
	60	Gaussian	1.11	3.34	1,121	0.0000	1.036	539.9	75.13
	72	Gaussian	1.28	4.60	1,464	-0.0020	1.039	556.1	78.19
	84	Gaussian	1.44	4.09	1,729	-0.0024	1.024	546.7	73.96
Bailey and Clutter	24	Gaussian	3.27	13.39	2,255	0.0072	1.033	666.1	80.37
	36	Gaussian	1.91	5.73	1,010	0.0044	1.016	615.0	74.97
	48	Exponential	1.84	3.31	2,643	0.0017	1.042	567.6	64.28
	60	Gaussian	1.15	3.25	1,103	0.0000	1.037	541.4	73.89
	72	Gaussian	1.29	4.25	1,474	-0.0021	1.040	551.8	76.67
	84	Gaussian	1.39	3.65	1,770	-0.0026	1.025	537.5	72.38

Subtitle: C₀ the nugget effect; C the contribution; A the reach; RME the reduced mean error; SDRE the standard deviation of the reduced error; AIC Akaike's information criteria; SDI% the spatial dependence index in percentage.

To select the best models at each age, the lowest AIC value was used as the main selection criteria. The models are considered statistically different if the difference between the AIC values is greater than 2. Thus, it was found that the gaussian model presented the best fits in most ages and productive capacity classification procedures (combination of 3 models with 2 methods). Exceptions occurred for the adjustments made at the age of 48 months, where the exponential model stood out as the best, and at 24 months using the Schumacher model by the guide curve method, in which the spherical semivariance model presented the best fit.

The nugget effect values ranged from 1.11 to 3.35, while the contribution from 3.21 to 16.81, and the range of spatial dependence was between 968 to 4,888. The highest contribution values in relation to the nugget effect prove the spatial dependence of the variable, making it possible to apply ordinary kriging to spatialize the site index in the non-sampled locations. Figure 2 shows the comparison statistics between the estimated site index values using the adjusted regression equations and applying the ordinary kriging.

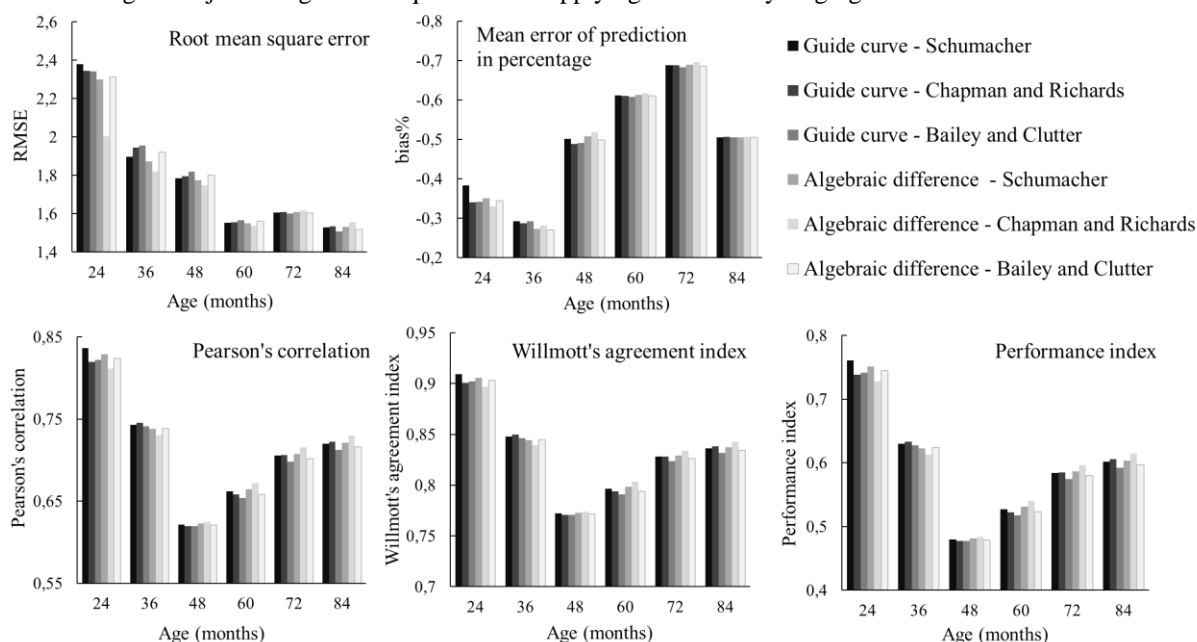


Figure 2. Statistics of the site index estimates by ordinary kriging and the adjusted equations.
 Figura 2. Estatísticas das estimativas de índice de sítio pela krigagem ordinária e as equações ajustadas.

The RMSE decreased exponentially from 24 to 84 months, indicating an increase in the accuracy of ordinary kriging in estimating the site index over age. The bias% showed negative values, demonstrating a mean tendency to overestimate the site index applying ordinary kriging. Pearson's correlation values, Willmott's agreement index, and performance showed a tendency to decrease up to 48 months and, from that age onwards, they started to increase, for the different procedures to classify the productive capacity.

Spatialization and validation of productive capacity

The site index estimated at 72 months, based on the algebraic difference method and ordinary kriging, for the three adjusted regression models were compared with the dominant height observed at 72 months, in the 34 plots of validation (Figure 3). The site index estimated by regression models in the algebraic difference method showed greater precision and accuracy than those estimated via ordinary kriging. Determination and angular coefficients greater than 0.90 demonstrate the predictive power of these models. In the spatial estimates of the site index, applying the ordinary kriging, these coefficients were reduced to values between 0.6 and 0.7, demonstrating that in the spatialization process there is a reduction in the accuracy.

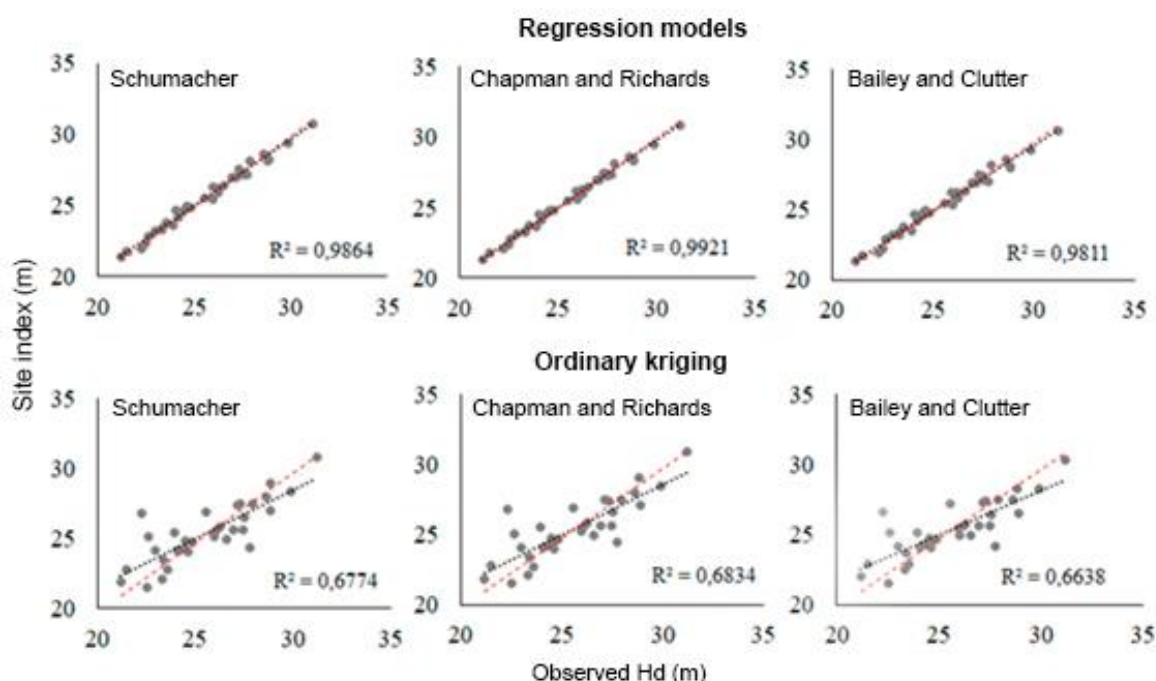


Figure 3. Validation of site index estimates using the Schumacher, Chapman and Richards, and Bailey and Clutter models, using the algebraic difference method, and by ordinary kriging.

Figura 3. Validação das estimativas do índice de sítio utilizando os modelos Schumacher, Chapman e Richards e Bailey e Clutter, pelo método da diferença algébrica, e pela krigagem ordinária.

Figure 4 shows the spatialization through ordinary kriging using 3 and 7 site index classes (amplitude of 5 and 2 m, respectively) obtained by the models of Schumacher, Chapman and Richards, and Bailey and Clutter, in the method of the algebraic difference. The model by Chapman and Richards showed greater accuracy, with 20 plots (58.83%) classified in their respective classes and 14 plots (41.17%) classified in other classes. The Schumacher and Bailey and Clutter models showed lower accuracy, with 15 plots (44.11%) and 16 plots (47.05%), respectively, classified outside the correct site classes. However, the greatest difference found among all models was 3.72 m, equivalent to 14.54% of the observed average. Most of the errors obtained were not greater than 1 m (3.90%), demonstrating that even using smaller amplitude classes, the spatialization of the site index is performed with high accuracy.

By increasing the amplitude of the class and reducing the number of classes to 3, the following classification was established for the classes: good (Class I) ranging from 27 to 32 m, regular (Class II) from 22 to 27 m and poor (Class III) from 18 to 22 m. Under these conditions, the accuracy of the productive capacity mapping increases considerably for all models, as 28 plots of validation (82.36%) fit the correct site class and only 6 plots (17.64%) did not fit, but with errors less than 1 m.

Considering the same class amplitude, the differences between the ordinary kriging maps obtained by the Schumacher, Chapman and Richards, and Bailey and Clutter models are imperceptible, indicating that the site

mapping is little influenced by the regression model used. In this case, the choice of the method of construction of the site curves will have a more significant influence, with superiority for the algebraic difference method.

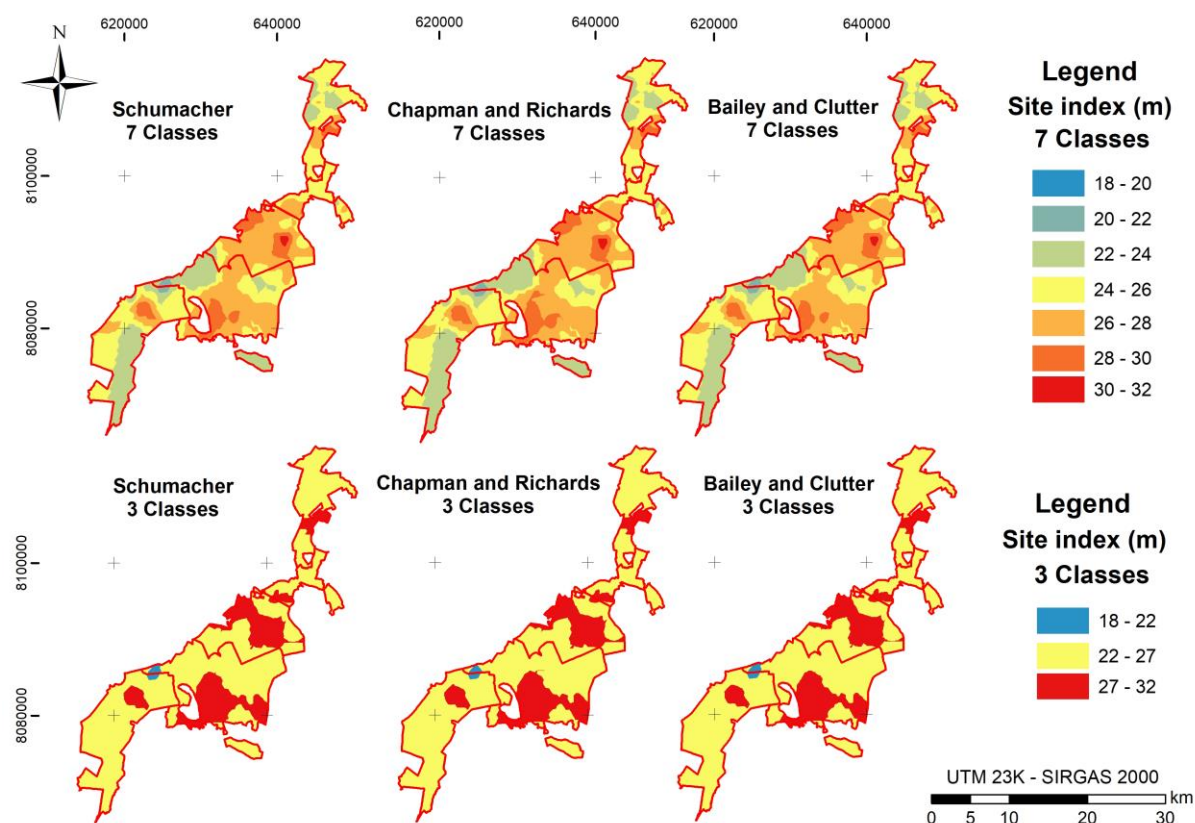


Figure 4. Spatialization of the site index, with 7 and 3 classes, in the method of algebraic difference at 72 months, for the models of Schumacher, Chapman and Richards and Bailey and Clutter.

Figura 4. Espacialização do índice de sítio, com 7 e 3 classes, no método da diferença algébrica aos 72 meses, para os modelos de Schumacher, Chapman e Richards e Bailey e Clutter.

DISCUSSION

The Schumacher, Chapman and Richards, and Bailey and Clutter models presented satisfactory fits, with superiority of the algebraic difference method compared to the guide curve method. The superiority of the algebraic difference method in the construction of site index curves was also observed by Castro *et al.* (2015), in a eucalyptus stand in Minas Gerais, with the standard error values of the estimate close to 6%. Pissinin and Schneider (2017) observed that the algebraic difference method presented the smallest estimation errors compared to the guide curve, using the Chapman and Richards model.

Kitikidou *et al.* (2011), applying the algebraic difference method to construct site index curves in a *Pinus brutia* stand, observed that the Bailey and Clutter model was superior to Chapman and Richards, corroborating the results obtained in this study. In the present work, the Bailey and Clutter model in the algebraic difference method presented the best fit statistics, with values (in module) of $S_{yx}\%$, RQEM, and bias% of 6.363, 1.479, and 0.158, respectively. The superiority of the algebraic difference method was confirmed from the graphic analysis of the normalized residuals, which presented a more homogeneous distribution compared to the guide curve method (Figure 2). A similar characteristic was found by Pissinin and Schneider (2017) when they adjusted different models in the guide curve and algebraic difference methods in eucalyptus stands, in which the algebraic difference method presented a homogeneous distribution of residuals and smaller errors.

In the spatial structuring of the site index, there was moderate to strong spatial dependence in all analyzed scenarios, with SDI values above 60% (Table 4). This is due to the strong spatial structure of the dominant height in species of the genus *Eucalyptus* sp. (MELLO *et al.*, 2005; GUEDES *et al.*, 2015; ATAÍDE *et al.*, 2020). This characteristic was also observed in stands of other forest species, such as *Tectona grandis* (PELISSARI *et al.*, 2015) and *Pinus* sp. (ZECH *et al.*, 2018). The most accentuated nugget effect values were obtained in the initial growth phase of the stands, between 24 and 48 months, ages in which a greater variability was observed in the

stands. The nugget effect is considered a random error at small distances which cannot be explained in the spatial analysis, related to the proximity between the plot and the high variation of the data in the initial phase of the population. Guedes *et al.* (2015) observed a similar characteristic when evaluating the spatial dependence structure of the variables dominant height, volume, and mean annual increment in a eucalyptus stand. On that occasion, the authors found that the nugget effect was also higher in the first measurement, performed at 2.7 years of age.

The range of spatial dependence showed the highest values at ages up to 48 months, decreasing at ages 60 to 84 months. The range represents the average distance at which the site index has spatial correlation and depends on the size of the study area and the average distance between the plots. For the conditions of this study, the range ranged from 968 m to 4,888 m, with an average of 1,777 m, indicating that plots distanced up to 1,777 m from each other, on average, show greater similarity in terms of site index, compared to the most distant ones. Other studies applied to eucalyptus stands showed range values ranging from 150 to 450 m in the evaluation of the spatial structure of the variables volume, total height, and dominant height (MELLO *et al.*, 2005; GUEDES *et al.*, 2015). However, these studies were carried out in areas smaller than 1,100 hectares, using the systematic sampling process, with greater sampling intensity, justifying the range values that were much lower compared to the present study.

The results obtained demonstrate that it is possible to model the spatial structure of the site index, from the experimental semivariogram, and apply the ordinary kriging in large areas, with lower sampling intensity and using simple casual sampling. Despite the reduction in the levels of precision and accuracy when estimating the site index in the plots, applying ordinary kriging, the information obtained is of great importance for the definition of the method and process of sampling in forest inventories (MELLO *et al.*, 2005; PELISSARI *et al.*, 2014), in addition to enabling the stratification of stands, aiming at reducing measurement costs and maintaining high accuracy in the continuous forest inventory (MELLO *et al.*, 2005; PELISSARI *et al.*, 2014), in addition to enabling the stratification of stands, aiming at cost reduction in measurement and maintaining high precision in the continuous forest inventory.

The spatialization of the site index allowed mapping the productive capacity classes in the stands and, consequently, demonstrated that the selected regression model has little influence on the stratification of the stands. The same spatial pattern was obtained for the Schumacher, Chapman and Richards, and Bailey and Clutter models, using 7 or 3 classes of site index (Figure 4). A characteristic of ordinary kriging is the smoothing effect, which is a tendency to overestimate the smallest values of the variable of interest and to underestimate the largest values, this characteristic being a determining factor in the definition of spatialized strata. Plots that were classified outside their respective site classes, even when using a greater class amplitude, presented site index instability over time or outliers, which differ from the natural trend due to environmental factors.

Importantly, geostatistics was proposed in the mapping of forest sites of *Tectona grandis* at index age (PELISSARI *et al.*, 2015), and in the insertion of the spatial component in growth and production models of *Eucalyptus* sp. by latent variable (PEREIRA *et al.*, 2016), demonstrating its potential in forest management. Associating regression models with geostatistical estimates allows for spatializing the site index in the continuous forest inventory, aiding the prognosis of volume based on growth and production models, such as Clutter's (1963).

CONCLUSION

- Among the evaluated procedures, the one that used the Bailey and Clutter model in the algebraic difference method presented the best site index estimates;
- The site index presents moderate to strong spatial dependence regardless of the regression model and the method used to construct the site curves, making it feasible to spatialize and map the productive capacity in eucalyptus stands by applying any of the procedures presented in this study;
- The choice of the regression model and the method of construction of the site curves do not influence the spatial continuity of the site index. Only the method has an effect on spatialization through ordinary kriging, with superiority to the algebraic difference method in mapping the productive capacity.

ACKNOWLEDGEMENTS

To the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and to the Post-Graduate Program in Environmental and Forest Sciences at UFRRJ.

REFERENCES

- ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. M.; SPAROVEK, G. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, Berlin, v. 22, n. 6, p. 711 - 728, 2013.
- ARAÚJO, E. J. G.; MORAIS, V. A.; DAVID, H. C. D.; SCOLFORO, J. R. S.; MELLO, J. M.; EBLING, A. A.

Spatialization of tree species diversity in the state of Minas Gerais. **Floresta e Ambiente**, Seropédica, v. 26, n. 1, p. 1 - 13, 2019.

ASSMANN, E. **The principles of forestry yield study**. Oxford: Pergamon Press, 1970, 506 p.

ATAÍDE, D. H. S.; ARAÚJO, E. J. G.; MONTE, M. A.; MENDONÇA, B. A. F.; SILVEIRA, C. E.; ALBUQUERQUE, A. M.; SILVA, L. C. Métodos geoestatísticos e determinísticos na espacialização da altura dominante em povoamento de eucalipto. **Revista de Ciências Agrárias**, Recife, v. 43, n. 1, p. 98-108, 2020.

BATY, F.; RITZ, C.; CHARLES, S.; BRUTSCHE, M.; FLANDROIS, J.; DELIGNETTE-MULLER, M. A Toolbox for nonlinear regression in R: the package nlstools. **Journal of Statistical Software**, Innsbruck, v. 66, n. 5, p. 1 - 21, 2015.

CAMPOS, J. C. C.; LEITE, H. G. **Mensuração florestal: perguntas e respostas**. Viçosa: Editora UFV, 5 ed. 2017. 636 p.

CASTRO, R. V. O.; CUNHA, A. B.; SILVA, L. V.; LEITE, H. G.; SILVA, A. A. L. Modelagem do crescimento e produção para um povoamento de *Eucalyptus* utilizando dois métodos para quantificação do índice de local. **Scientia Forestalis**, Piracicaba, v. 43, n. 105, p. 83 – 90, 2015.

CLUTTER, J. L. Compatible growth and yield models for Loblolly Pine. **Forest Science**, Oxford, v. 9, n. 3, p. 354 - 371, 1963.

GUEDES, I. C. L.; MELLO, J. M.; SILVEIRA, E. M. O.; MELLO, C. R.; REIS, A. A.; GOMIDE, L. R. Continuidade espacial de características dendrométricas em povoamentos clonais de *Eucalyptus* sp. avaliada ao longo do tempo. **Cerne**, Lavras, v. 21, n. 4, p. 527 - 534, 2015.

KITIKIDOU, K.; BOUNTIS, D.; MILIOS, E. Site index models for calabrian pine (*Pinus brutia* Ten.) in Thasos island, Greece. **Ciência Florestal**, Santa Maria, v. 21, n.1, p. 125 - 131, 2011.

LEITE, H.; CASTRO, R.; SILVA, A.; JÚNIOR, C.; BINOTI, D.; CASTRO, A. F. Classificação da capacidade produtiva de povoamentos de eucalipto utilizando diâmetro dominante. **Silva Lusitana**, Lisboa, v. 19, n. 2, p. 181 - 195, 2011.

LUNDGREN, W. J. C.; SILVA, J. A. A.; FERREIRA, R. L. C. Estimação do volume de eucaliptos por krigagem e cokrigagem no Semiárido Pernambucano. **Floresta e Ambiente**, Seropédica, v. 24, 2017.

MELLO, J. M.; BATISTA, J. L. F.; OLIVEIRA, M. S.; RIBEIRO JÚNIOR, P. J. Estudo da dependência espacial de características dendrométricas para *Eucalyptus grandis*. **Cerne**, Lavras, v. 11, n. 2, p. 113 - 126, 2005.

PELISSARI, A. L.; FIGUEIREDO FILHO, A.; CALDEIRA, S. F.; MACHADO, S. A. Geoestatística aplicada ao manejo de povoamentos florestais de teca, em períodos pré-desbaste seletivo, no estado do Mato Grosso. **Revista Brasileira de Biomassa**, São Paulo, v. 32, n. 3, p. 430 - 444, 2014.

PELISSARI, A. L.; CALDEIRA, S. F.; FIGUEIREDO FILHO, A.; MACHADO, S. A. Propostas de mapeamentos da capacidade produtiva de sítios florestais por meio de análises geoestatísticas. **Scientia Forestalis**, Piracicaba, v. 43, n. 107, p. 601 - 608, 2015.

PEREIRA, C. P.; DIAS, P. A. S.; MERGULHÃO, R. C.; THIERSCH, C. R.; FARIA, L. C. Modelo de crescimento e produção de Clutter adicionado de uma variável latente para predição do volume em um plantio de *Eucalyptus urograndis* com variáveis correlacionadas espacialmente. **Scientia Forestalis**, Piracicaba, v. 44, n. 110, p. 393-403, 2016.

PISSININ, L. Z.; SCHNEIDER, P. R. Tendências de crescimento em altura dominante de dois clones de *Eucalyptus saligna* Smith implantados em distintas condições de solo. **Ciência Florestal**, Santa Maria, v. 27, n. 1, p. 263 - 275, 2017.

R CORE TEAM. **R: A language and environment for statistical computing**. Vienna, Austria. R Foundation for Statistical Computing, 2015.

RIBEIRO JÚNIOR, P. J.; DIGGLE, P. J. GeoR: a package for geostatistical analysis. **RNEWS**, Austria, v. 1, n. 2, p. 15 - 18, 2001.

SCOLFORO, J. R. S. **Biometria florestal: modelos de crescimento e produção florestal**. Lavras: UFLA/FAEPE, 2006. 393 p.

ZECH, F. D.; REFOSCO, J. C.; YOSHIZUMI, L. T.; ABREU, R. O. S. Uso de interpoladores espaciais na estratificação estatística de *Pinus taeda*. **Scientia Forestalis**, Piracicaba, v. 46, n. 117, p. 87 - 96, 2018.

ZIMBACK, C. R. L. **Geoestatística**. Botucatu: UNESP, 1 ed., 2003, 25 p.