



## CROP SCIENCE

# Site classification for *Eucalyptus* sp. in a tropical region of Brazil

ATAÍDES M. FILHO, SYLVIO P. NETTO, SEBASTIÃO A. MACHADO, ANA P.D. CORTE & ALEXANDRE BEHLING

**Abstract:** The aim was to determine the productive capacity of a forest site by applying different methods of fitting, combined with geostatistical techniques, to stands of *Eucalyptus* sp. in a tropical region of Brazil. Data were collected from 845 plots of a continuous forest inventory over four years. The classification of local production capacity was performed using growth curves obtained by the guide curve (GC) method, algebraic difference approach (ADA) and generalized algebraic difference approach (GADA) methods and ordinary kriging through the spherical, exponential, and Gaussian models to determine the spatial dependence of the variables site, geographical boundaries of site index classes, and their respective areas for each hectare in an annual production unit (APU). The modified Chapman–Richards model, fitted by the generalized difference approach method (GADA), provided the best statistical results, an improvement of 12.23% and 39.80% on the ADA and GC methods, respectively. The exponential model selected to express the spatial distribution of dominant height showed a high degree of spatial dependence.

**Key words:** Generalized algebraic approach, geostatistical techniques, guide curve, continuous forest inventory.

## INTRODUCTION

Growth evaluation in planted forests is essential, due to the requirements of strategic nature: the good performance of silvicultural practices, genetic improvement, and maximization of production gains. The ability of a given species or genetic material to grow in each site encompasses climatic, edaphic, and biological factors (Scolforo 1998).

The site index, which is the average height of dominant trees at a certain reference age (Clutter et al. 1983, Burkhart & Tomé 2012) is strongly correlated to growth factors (soil and agrometeorological variables) and volume yield. It is also strongly influenced by density and treatments like thinning, which provide a reliable source of environmental information.

According to Andrade (2004), the site index is currently the more practical and widespread method to classify forest productivity, since it uses dominant height as a variable, which is the response to inter-related environmental factors, and is highly correlated with the volume yield. Furthermore, dominant height is relatively independent of the stand density.

In accordance with Cieszewski & Strub (2008), modeling dominant height prediction has been evolving since the publication of the Schumacher (1939) models, citing the following authors: Bailey & Clutter (1974), Curtis et al. (1974), Bailey (1980), Newnham (1988), Huang (1994), Amaro et al. (1998), Cieszewski & Bailey (2000) and Krumland & Eng (2005) as authors who furthered the field.

The most commonly used methods for constructing production curves through dominant height, with data from permanent and temporary sample plots and stem analysis are: (1) Guide curve (GC), (2) Algebraic difference approach (ADA), (3) Parameter prediction (PP) and (4) Generalized algebraic difference approach (GADA).

To decide which method will be used, some concepts and desirable peculiarities should be considered for the construction of the site curves (Bailey & Clutter 1974, Cieszewski & Bailey 2000, Castillo Lopez et al. 2013). Standard sigmoid growth with an inflection point and the ability to reach a horizontal asymptote at more advanced ages provides a logical response (for example, dominant height must be equal to zero at zero age and the curve should always keep growing), as does invariance in relation to the simulation path (path invariance), and age of reference (base-age invariance). The condition of invariance, related to the method of performing the simulation, implies that when the dominant height  $H_1$  at the age  $t_1$  is estimated with the curves, the value of dominant height  $H_3$  at age  $t_3$  should result in the same value when estimating firstly the dominant height  $H_2$  at age  $t_2$  and then, with this value, the dominant height at age  $t_3$ .

The condition of invariance also implies that the shape of the curves should not vary, regardless of the reference age, which is used to define the site index. However, the age of reference to be used should be close to the expected rotation age, as height, at this age, presents high correlation with the volumetric yield (Pece de Rios 1993).

Basically, two types of site index curves can be constructed: harmonic curves or anamorphic, and natural curves or polymorphic. The anamorphic curves are constructed by fitting the guide curve as a function of age. Subsequently,

a series of curves with the same shape are built above and below of the guide curve, differing only in magnitude by a fixed percentage.

The anamorphic curves basically present two sources of errors (Batista & Couto 1986, Andrade 2004): (1) Anamorphic curves are only accurate when the dominant height sampling is properly carried out, so that the variation in the site index be equally represented across all ages; (2) Anamorphic curves consider the influence of the site variation in dominant height growth to be uniform in all ages, so that the shape of the height growth curves is the same for all sites.

Polymorphic curves are constructed using data from permanent plots or stem analysis. They assume that the shape of the growth curves of the trees varies from site to site, i.e., they do not have proportionality to the height growth of dominant trees between different site classes. There seems to be a consensus among most researchers (Scolforo 1993, Cieszewski & Bailey 2000, Weiskittel et al. 2011, Burkhart & Tomé 2012) that the polymorphic curves are superior to anamorphic in terms of the accuracy of the site index estimate.

The guide curve method enables only the construction of anamorphic curves and, therefore, is the simpler method and is less indicated among other more elaborated methods.

The methods of algebraic differences and prediction of parameters, however, have been presented as an effective alternative to the construction of the polymorphic site curves, assuming the assumptions of different growth rates between different sites and invariance in the path of the simulation and reference age. Bailey & Clutter (1974) formalized the invariance property in site index age models and presented a technique to derive equations for a dynamic method, known as the algebraic difference approach (ADA), which essentially involves

replacing a parameter of the basic method to express it as a function of the site (in this case, a combination of dominant height and age). The main limitation of the ADA methodology is that most of the derived models are anamorphic or have an asymptote (Bailey & Clutter 1974, Cieszewski & Bailey 2000).

Cieszewski & Bailey (2000) introduced a generalization of the ADA methodology, the generalized algebraic difference approach (GADA). The main advantage of this method is that the basic equation can be expanded according to several growth theories, as the rate of growth and asymptote, allowing more than one parameter of each model to be a function of the site quality, and the growths curves obtained are more flexible (Cieszewski & Bailey 2000, Cieszewski 2001, 2002, 2003). With this generalization, we can get growths curves that are polymorphic and have multiple asymptotes (Cieszewski 2002).

The classification of a traditional site, regardless of the method, as here exposed, considers the average of a specific sample point evaluated, not taking into consideration possible correlations between units surveyed in nearby regions (Bognola et al. 2009, Mello et al. 2006).

Therefore, the use of geostatistical techniques provides significant functional space, since potential existing space dependence between sample units can be explained in an integrated manner. Yet, as a bigger advantage, locations that are not sampled can receive unbiased qualitative or quantitative allocation and are correlated and interpolated to the sampled points through the application of multivariate linear regression techniques.

This tool brings advantages for modeling the site and the sampling procedures on different forms, allowing the definition of geographical spatialization of selected variables,

area quantification, stratification of a more specific population, improvement of statistical estimators and a reduction in sampling error, generating tools for the physical delimitation of production units and for precise silvicultural practices, as demonstrated by Péllico Netto et al. (2014) and Pelissari (2015).

In Brazil, the exhaustion of available areas for forest cultivation in places traditionally used by producers triggered the process of expansion of the production frontier to sites with distinct edaphoclimatic characteristics, where modeling methods of growth and production are still little known.

Allied to these factors, the concept of precision management using geostatistical methods allows a more effective and accurate management of forest stands, even reducing production costs.

However, the site is an important and essential variable used in the improvement of statistical sampling estimators, when it is used as a basis for stratification of the population (Péllico Netto et al. 2014). The purpose of this research was to generate local productivity classes, through three methods (GC, ADA, and GADA) to build up the site curves for *Eucalyptus* sp. in a tropical region of Brazil, by means of permanent sample plots annually measured between 2011–2014, and the application of geostatistics to classify the cartographic bases of the plantations.

The main objective of this study was to evaluate the productive capacity of forest sites based on site index by means of traditional methods and geostatistical techniques for clonal stands of *Eucalyptus* sp. in a tropical region of Brazil. For the more specific objectives of fitting models for the classification of forest sites, the methods of guide curve, algebraic difference approach, and generalized algebraic difference approach were applied. Additionally, ordinary

kriging was applied to the dominant heights at sites to test the spatial dependence of this variable and, if so, plot the spatial boundaries of the site classes in the cartographic base of the forest stands, according to their site index at a reference age.

## MATERIALS AND METHODS

To conduct this research, 2,452 pairs of dominant height ( $h_{dom}$ ) and age from 845 permanent sample plots, systematically distributed in a sampling intensity of 1:2.5 ha, in stands of *Eucalyptus* sp. in a tropical region of Brazil were used, as presented in the characterization of the study area.

Assmann's principle (1970) was used to obtain the average dominant height for each sample plot, based on the 100 trees with the largest diameter per hectare, healthy and free of defects. The classification of productive capacity was tested using three methods for fitting the site index – SI: GC, ADA, and GADA.

The data (Figure 1) feature a range of ages (26–93 months), and heights (11.50–32.8 m), distributed proportionally to the areas of three annual production units (APUs), planted in 2007, 2008, and 2009, respectively. The APUs in 2007, 2008 and 2009 had total areas of 1,085, 840, and 187.5 ha, respectively.

The minimum, maximum, means values and age, as well as the standard deviations (SD) of diameter ( $d_{1,30m}$ ), total height ( $h$ ) of the data are shown in Table I.

The data collection was performed in permanent sample plots, rectangular shape, with variable area (PPAV), allocated through a systematic sampling procedure using a square grid of 158.1 m per side.

Five non-linear models were fitted, which are well known among forestry professionals, because their flexibility to describe a standard sigmoid with an inflection point and ending up in a horizontal asymptote, as follows:

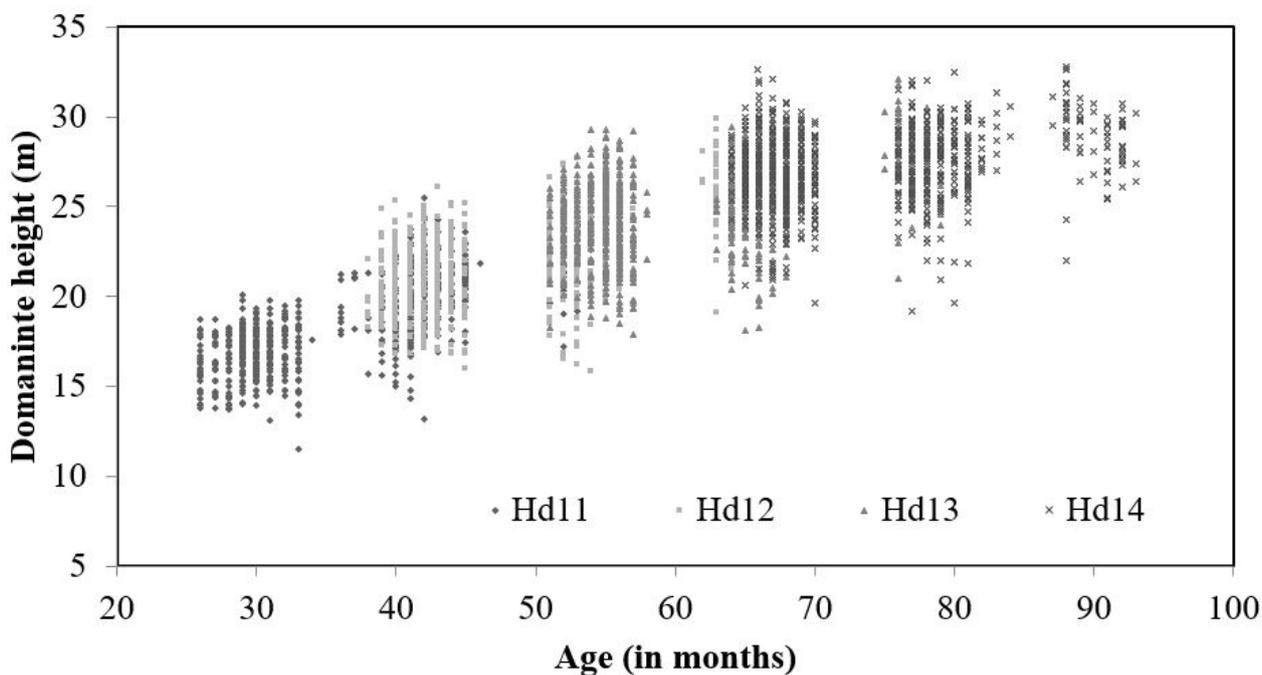


Figure 1. Distribution of the data to *Eucalyptus* sp. in a tropical region of Brazil.

Model 1: Chapman-Richards

Basic equation:  $H = \theta_0 \cdot (1 - e^{-\theta_1 \cdot t})^{\theta_2} + \varepsilon$  (1)

$$SI = \theta_0 \cdot (1 - e^{-\theta_1 \cdot t_0})^{\theta_2} + \varepsilon$$

Guide-model:  $H_1 = SI \left[ \frac{1 - e^{(-\theta_1 \cdot t_1)}}{1 - e^{(-\theta_1 \cdot t_0)}} \right]^{\theta_2} + \varepsilon$  (2)

Model 2: Bailey 4 Parameters

Basic equation:  $H = \theta_0 \cdot (1 - e^{-\theta_1 \cdot t^{\theta_2}})^{\theta_3} + \varepsilon$  (3)

$$SI = \theta_0 \cdot (1 - e^{-\theta_1 \cdot t_0^{\theta_2}})^{\theta_3} + \varepsilon$$

Guide-model:  $H_1 = SI \left( \frac{1 - e^{-\theta_1 \cdot t_1^{\theta_2}}}{1 - e^{-\theta_1 \cdot t_0^{\theta_2}}} \right)^{\theta_3} + \varepsilon$  (4)

Model 3: Modified Chapman-Richards model, Clutter (1998)

Basic equation:  $H = H_0 \left[ \frac{1 - e^{(-t/t_1)}}{1 - e^{(-t_0/t_1)}} \right]^2 + \varepsilon$  (5)

Where:  $H_0$  is the dominant height at the initial age  $t_0$

Guide-model:  $H_1 = SI \left[ \frac{1 - e^{(-t_1/t_1)}}{1 - e^{(-t_0/t_1)}} \right]^2 + \varepsilon$  (6)

Model 4: Modified Chapman-Richards by Amaro et al. (1998)

Basic equation:  $H = \theta_0 \left\{ 1 - \left[ 1 - \left( \frac{H_0}{\theta_0} \right)^{\frac{1}{\theta_1}} \right]^{\frac{t_0}{t}} \right\}^{\theta_1} + \varepsilon$  (7)

Guide-model:  $H_1 = \theta_0 \left\{ 1 - \left[ 1 - \left( \frac{SI}{\theta_0} \right)^{\frac{1}{\theta_1}} \right]^{\frac{t_0}{t_1}} \right\}^{\theta_1} + \varepsilon$  (8)

Model 5: Chapman-Richards

For this model we can use the basic equation (1)

Parameters of site:  $\theta_0 = e^{(X)}$ ,  $\theta_1 = \beta_0$  and  $\theta_2 = \beta_1 + \beta_2 / X$  (9)

**Table I. Characteristics of the variables for adjusting models to *Eucalyptus* sp. in a tropical region of Brazil.**

| Statistics | $d_{1.30m}$<br>(cm) | $h$<br>(m) | Age<br>(months) |
|------------|---------------------|------------|-----------------|
| Minimum    | 8.6                 | 11.5       | 26              |
| Maximum    | 22.8                | 32.8       | 93              |
| Mean       | 14.2                | 23.3       | 54.8            |
| CV (%)     | 15.7                | 17.4       | -               |

The solution for X with initial values ( $t_0, X_0$ )

$$X_0 = \frac{1}{2} \left[ (\ln H_0 - \beta_1 L_0) + \sqrt{(\ln H_0 - \beta_1 L_0)^2 - 4\beta_2 L_0} \right] \quad (10)$$

$$L_0 = \ln \left[ 1 - e^{-\beta_0 t_0} \right] \quad (11)$$

Dynamic equations:  $H_1 = H_0 \left[ \frac{1 - e^{(-\beta_0 t_1)}}{1 - e^{(-\beta_0 t_0)}} \right]^{\beta_1 + \beta_2 / X_0} + \varepsilon$  (12)

Where:  $H$  is the dominant height of the  $i^{th}$  tree at age  $t$ , in m;  $H_0$  is the dominant height at the initial age  $t_0$ ;  $X$  is the value of the function at age  $t$ , and  $X_0$  is the reference variable defined as the value of the function at age  $t_0$ ,  $t_1$  is age at time 1, in years;  $t_0$  is age at time 0, in years;  $SI$  is the Site-Index, in m;  $\theta_1, \theta_2, \dots, \theta_n$  are parameters of the model;  $\beta_0, \beta_1$  and  $\beta_2$  are the coefficients of the fitted site parameters;  $e$  is the Euler's exponential;  $\varepsilon$  is the random error.

Models 1 and 2 were fitted by the guide curve method, models 3 and 4 by the algebraic differences approach method (ADA), and model 5 by the generalized algebraic differences approach (GADA method), proposed by Cieszewski & Bailey (2000) to obtain the site curves.

The regression coefficients of the fitted models using the GC method and ADA were estimated by Levenberg–Marquardt's method for minimization of the non-linear residual sum of squares (Marquardt, 1963), using Statgraphics Centurion XVI software version 16.1.02. For more details about the GADA approach, review the procedure in (Cieszewski & Bailey 2000,

Krumland & Eng 2005). Other analyses were performed using software R.

### Statistical analysis of the modeling

We evaluated three statistics as a criterion for selecting the best model: lowest standard error of the estimate in percentage ( $S_{xy}\%$ ), higher coefficient of determination ( $R^2_{adj}$ ) and the best residual dispersion in percentage ( $R\%$ ), as presented in (13–18):

$$R^2 = 1 - \left(\frac{SQR}{SQT}\right) \quad (13)$$

$$R^2_{adj} = 1 - \left[\left(R^2\right) \frac{(n-1)}{(n-p)}\right] \quad (14)$$

$$S_{yx} = \sqrt{\frac{SQR}{n-p}} \quad (15)$$

$$S_{yx}\% = \frac{S_{yx}}{\bar{y}} \cdot 100 \quad (16)$$

$$E_i = y_i - \bar{y}_i \quad (17)$$

$$E(\%) = \frac{E_i}{y_i} \cdot 100 \quad (18)$$

Where:  $SQR$  is Residual Sum of Squares and  $SQT$  is the Total Sum of Squares,  $E_i$  are the residuals dispersion and  $E(\%)$  are the percentual distribution of the residuals (which were presented in a plot).

### Application of ordinary kriging

Geostatistics defines a set of mathematical procedures that allows the recognition and description of spatial relationships between sampled points in each domain.

According to Yamamoto & Landim (2013), the kriging uses information from the variogram to find the optimal weights to be associated with samples of known values, so that it is possible to estimate values of unknown points. It is understood to be a series of regression analysis techniques that aims to minimize the estimated

variance from one model, which considers the stochastic dependence between the spatial distributed data.

The main objective of kriging is to define the site limits in each stand of the forest area to compose the stratified estimates.

Ordinary kriging was chosen for the spatial modeling of the site, and is expressed as follows (Pelissari 2015):

$$Z_{krig}(x_0) = \sum_{i=1}^n \lambda_i [Z(x_i)] \quad (19)$$

Where:  $Z_{krig}$  is the estimator for kriging interpolation;  $\lambda_i$  are the weights associated with the  $n$  data; and  $z(x_i)$  are the observed data of dominant height at the reference age.

Data were organized in spreadsheets and the models were fitted using the software R (R Development Core Team 2009, Robinson & Hamann 2011). The geoR package, in which a specific CSV file was defined, containing the coordinates ( $x, y$ ) of the Universal Transverse Mercator (UTM), in the SIRGAS 2000 reference system, was applied for computing the distances between the sampling points. The numerical value of the  $z$  at the reference age of seven years for each sample unit was defined as the dependent variable  $Z$ . ARC GIS 10.2 software was used to generate thematic maps.

Estimates were obtained using the fitted semivariogram, applying the following theoretical models: spherical (29), exponential (30), and Gaussian (31) (Pelissari 2015).

$$\gamma(h) = \begin{cases} C_0 + C \left[ \left(\frac{3}{2}\right) \cdot \left(\frac{h}{A}\right) - \left(\frac{1}{2}\right) \cdot \left(\frac{h}{A}\right)^3 \right] & \text{para } h \leq A \\ C_0 + C & \text{para } h > A \end{cases} \quad (20)$$

$$\gamma(h) = C_0 + C(1 - e^{-h/A}) \quad (21)$$

$$\gamma(h) = C_0 + C(1 - e^{-h^2/A^2}) \quad (22)$$

Where:  $\gamma(h)$  is the semivariance of dominant height at the reference age  $z(x_i)$ ;  $h$  is the distance

expressed in meters;  $C_0$  is the nugget effect;  $C$  is the variance a priori;  $A$  is the range.

The objective of this specific procedure was to evaluate the quality of the fitted model by the coefficient of determination ( $R^2$ ) and the weighed sum of squares of deviations ( $SQDP$ ). The degree of spatial dependence ( $DSD$ ) of the variable site index was estimated in (23).

$$DSD = \frac{C_0}{C_0 + C} \cdot 100 \quad (23)$$

Finally, cross-validation was applied to compare the observed ( $z(x_i)$ ) and estimated ( $Z_{est}(x_i)$ ) values, through adjustment of a straight-line equation (24).

$$Z(x_i) = \theta_1 + \theta_2 \cdot Z_{est}(x_i) \quad (24)$$

## RESULTS AND DISCUSSION

### Site classification

All estimated parameters of the fitted models by the *GC* method, *ADA*, and *GADA* were significantly different from zero at 95% probability level, and presented in Table II. As expected, the *GC* approach presented the worst adjusted coefficient of determination values ( $R^2_{adj} = 0.753\text{--}0.755$ ) and precision ( $S_{yx}\% = 8.62\text{--}8.65\%$ ). The fitted models by *ADA* presented significant improvement on the adjusted coefficient of determination values ( $R^2_{adj} = 0.778\text{--}0.807$ ) and precision ( $S_{yx}\% = 5.60\text{--}6.02\%$ ). Chapman's model, fitted by *GADA* method, provided the best adjusted coefficient of determination values ( $R^2_{adj} = 0.899$ ) and precision ( $S_{yx}\% = 4.92\%$ ).

The graphical residual distributions (Figure 2) indicate greater homogeneity as the stands are close to the technical rotation for all curve construction methods, however, it is noticeable that the *GADA* method provides a more expressive effect to detect the yield capacity, proving the better efficiency of this method.

**Table II. Coefficients and statistics of the models for forest sites classification of *Eucalyptus* sp. in a tropical region of Brazil.**

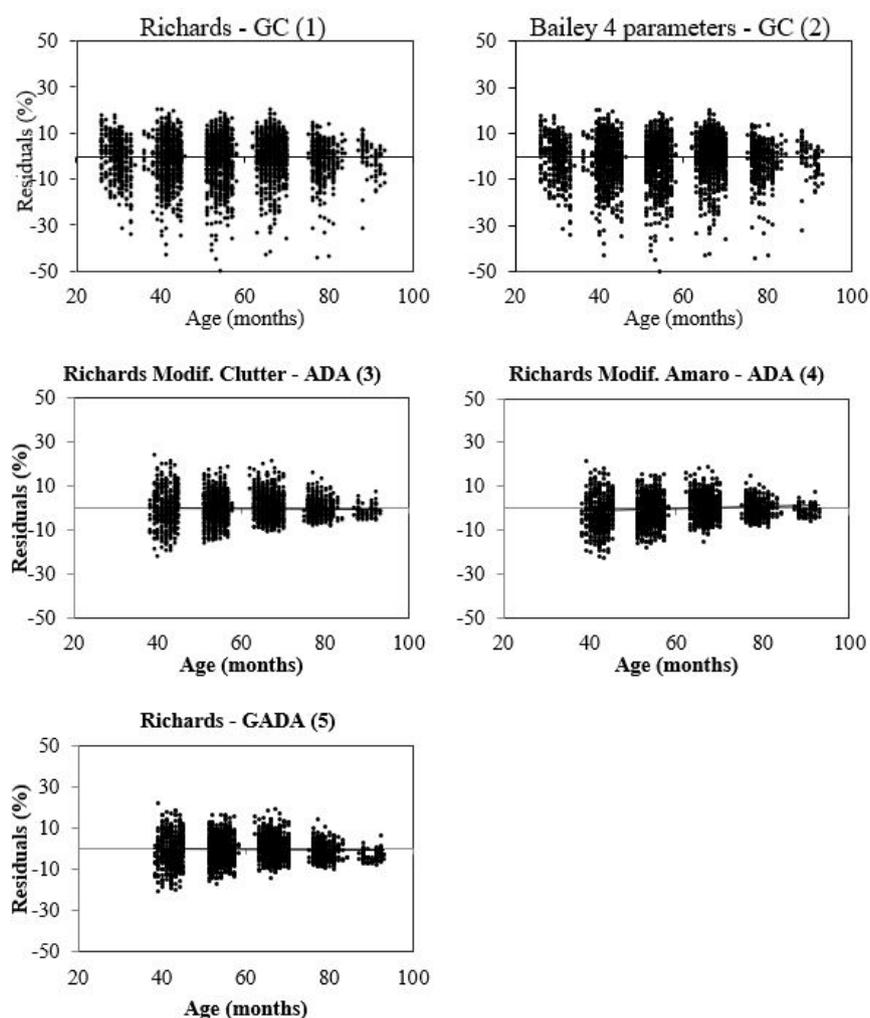
| Method      | Model N° | Functional equation | $S_{yx}$ (m) | $S_{yx}\%$ | $R^2_{Ajust.}$ |
|-------------|----------|---------------------|--------------|------------|----------------|
| <i>GC</i>   | (1)      | ***                 | 2.01         | 8.65       | 0.753          |
| <i>GC</i>   | (2)      |                     | 2.01         | 8.62       | 0.755          |
| <i>ADA</i>  | (3)      |                     | 1.49         | 6.02       | 0.778          |
| <i>ADA</i>  | (4)      |                     | 1.39         | 5.60       | 0.807          |
| <i>GADA</i> | (5)      |                     | 1.22         | 4.92       | 0.899          |

\*\*\* The coefficients of all equations were significant.

In general, the five fitted models resulted good statistics of adjusted coefficient of determination, precision, and residual distribution. However, it is noticeable that the *GC* methodology presented the more expressive residual dispersion, ranging between  $-45\%$  and  $25\%$ , i.e., the worst results. The distributions residual dispersions of the *ADA* and *GADA* were balanced and showed more adherence to the data variation, whose values occurred between  $-22\%$  and  $23\%$  for *ADA*, and between  $-20\%$  and  $20\%$  for *GADA*.

Evaluating the mean curve generated for each methodology (*GC*, *ADA*, and *GADA*), differences in the curvature pattern, scale, and asymptotic points are not observed, except for the fitted Bailey model by the *GC* method, which features a horizontal asymptote at an earlier age than that seen in the other fitted models (Figure 3).

The three proposed approaches produced similar results, however the *GADA* method generated results with smallest standard error of the estimate, highest coefficient of determination, and best residual dispersion, being  $12.23\%$  higher than the *ADA* and  $39.80\%$  higher than the *GC* method. Considering this evaluation, the Chapman-Richards model fitted



**Figure 2.** Residual distributions of forest site classification in stands of *Eucalyptus* sp. in a tropical region of Brazil.

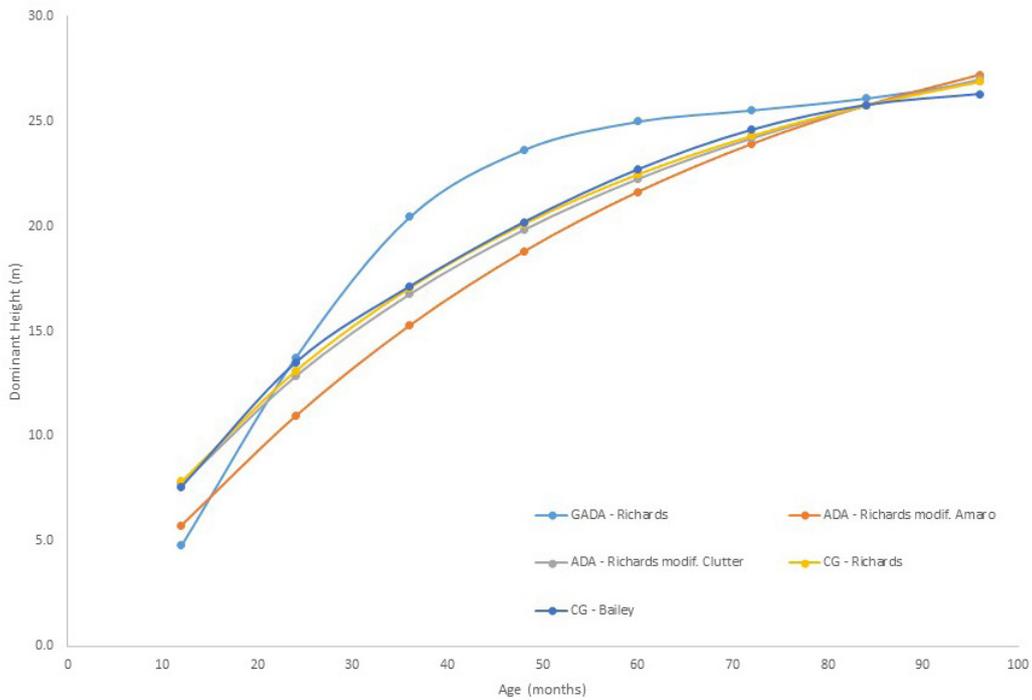
by the GADA method was the best to represent the data in the present study, therefore the selected model to generate the site index.

The complex polymorphic growths curves developed using the Chapman-Richards model (GADA) and the four years' observations of the forest inventory (inventory in 2011 to inventory 2014) can be observed in Figure 4. The stand was classified in two site classes (I and II), due to its homogeneity and the strategy to be adopted for applying stratified sampling.

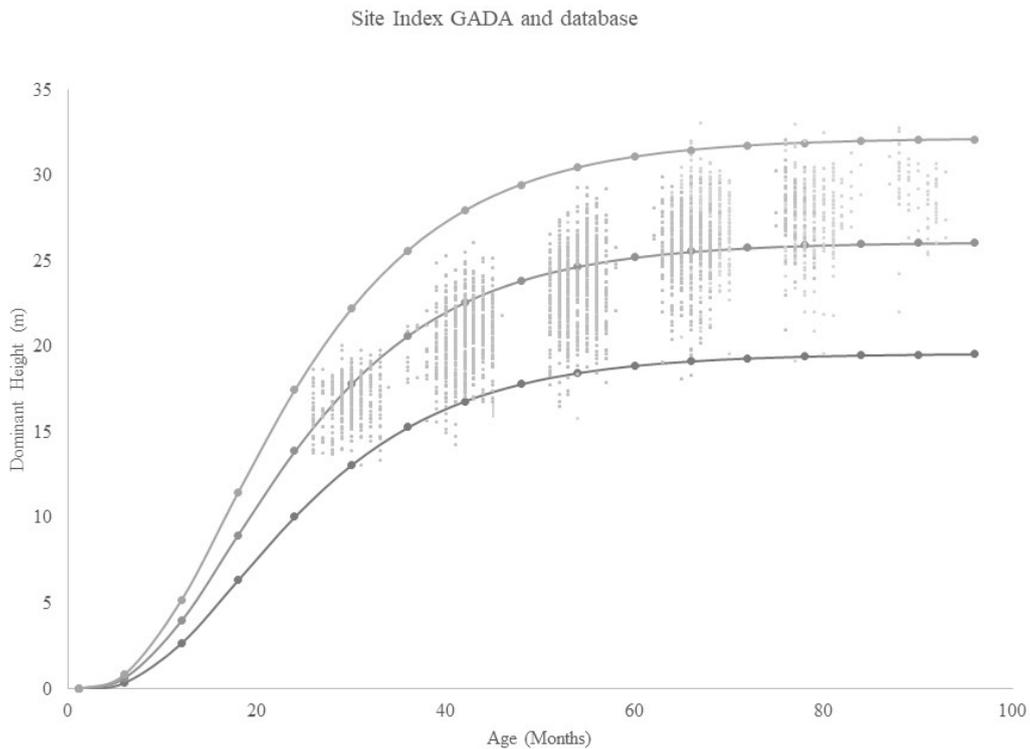
According to Scolforo (1993), the height growth curve tends to present a more pronounced sigmoid form at more productive sites, a proof that the proportionality between

the curves of the site is a failure. Already, in less productive sites, the height growth pattern tends to be smoother, i.e., the inflection point is reached later than in the more productive sites. Therefore, polymorphic models produce growth curves closer to biological reality.

Differences in productivity performance between the site classes and the early asymptotic culmination occur owing to the edaphic factors in the region, once sand soils of quartz of low natural fertility are dominant in these areas, formed by stains of Quartzarenic Neosol, with less than 10% clay content (EMBRAPA 2006). Another important factor is the choice of genetic materials, given that it is a pioneer region,



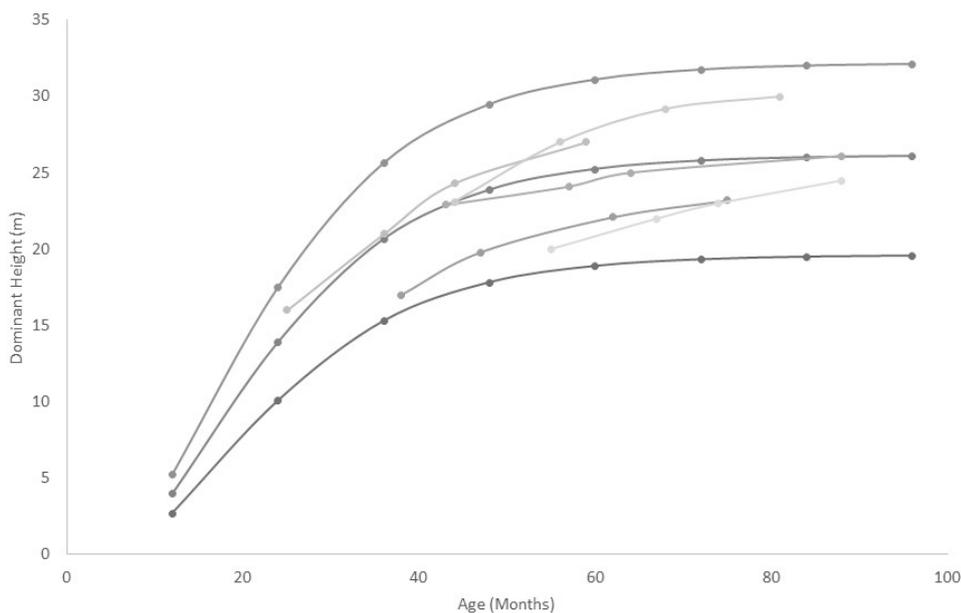
**Figure 3.** Growth curves of the five fitted models for stands of *Eucalyptus* sp. in a tropical region of Brazil.



**Figure 4.** Site index for stands of *Eucalyptus* sp. in a tropical region of Brazil.

**Table III.** Site Index values for *Eucalyptus* sp. in a tropical region of Brazil.

| Age (years) | Site Index  |             |             |             |             |             |
|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|             | I           |             |             | II          |             |             |
|             | UL (m)      | C (m)       | LL (m)      | UL (m)      | C (m)       | LL (m)      |
| 1           | 11.5        | 8.7         | 7.1         | 7.1         | 5.3         | 4.1         |
| 2           | 19.6        | 15.6        | 13.1        | 13.1        | 10.2        | 8.0         |
| 3           | 24.7        | 20.6        | 17.8        | 17.8        | 14.3        | 11.5        |
| 4           | 27.9        | 24.2        | 21.5        | 21.5        | 17.7        | 14.6        |
| 5           | 29.8        | 26.8        | 24.3        | 24.3        | 20.5        | 17.2        |
| 6           | 31.0        | 28.6        | 26.4        | 26.4        | 22.9        | 19.5        |
| <b>7</b>    | <b>31.8</b> | <b>29.9</b> | <b>28.0</b> | <b>28.0</b> | <b>24.8</b> | <b>21.5</b> |
| 8           | 32.3        | 30.8        | 29.2        | 29.2        | 26.3        | 23.2        |
| 9           | 32.5        | 31.4        | 30.1        | 30.1        | 27.5        | 24.6        |



**Figure 5.** Stable behavior of dominant heights in the same site index for *Eucalyptus* sp. in a tropical region of Brazil.

where the first productive cycles are just being completed for the first time.

When dominant height amplitude in the reference age of seven years was observed, two site index curves were defined at the center-of-class values of 29.9 m for site I and 24.8 m for site II, whose upper and lower limits and center-of-class are presented in (Table III).

A desirable characteristic is that the data are kept in the same site class along the measurements (David 2014, Clutter et al. 1983).

The mean dominant heights in their respective ages of measurement are shown in (Figure 5), in which the stability of the plots within the same site class is proven for 2011, 2012, 2013, and 2014 measurements.

**Ordinary kriging**

The fitting of the semivariograms for the dominant height at the age of seven years ( $h_{dom}$ ) for the exponential, spherical, and Gaussian models are presented in Table IV, where a high

**Table IV. Results of the semivariograms adjusted for  $h_{dom}$  at the age of reference for *Eucalyptus* sp. in a tropical region of Brazil.**

| Model       | Nugget effect ( $C_0$ ) | Variance a priori (C) | Range (A) (m) | Degree of Spatial Dependence DSD (%) | $R^2$ | SQDP   |
|-------------|-------------------------|-----------------------|---------------|--------------------------------------|-------|--------|
| Exponential | 0.989                   | 3.291                 | 795.37        | 76.9%                                | 0.94  | 0.0002 |
| Spherical   | 1.101                   | 3.179                 | 823.10        | 74.3%                                | 0.86  | 0.0006 |
| Gaussian    | 1.127                   | 3.153                 | 812.30        | 73.8%                                | 0.90  | 0.0004 |

Where:  $R^2$  is the coefficient of determination; SQDP is the sum of squares of the weighted deviations.

degree of spatial dependence was found (DSD > 0.75). This evidence allows the application of geostatistical techniques for modeling this variable (Pelissari 2015).

The exponential model resulted in a higher degree of spatial dependence (DSD = 76.9%), which was considered strong when equal to or greater than 75% (Pelissari 2015, Cambardella et al. 1994); a lower nugget effect ( $C_0 = 0.989$ ); a higher coefficient of determination ( $R^2 = 0.94$ ); and a smaller sum of squares of the weighed deviations (SQDP = 0.0002). These results confirm the high spatial dependence of the variable in the present study, corroborating positively to appropriate estimates in non-sampled sites, for delimitation and quantification of site classes in the stands. In addition, Figure 6 shows the semivariograms adjusted to the exponential model, in which their appropriate features explicitly represent the data of the stand.

The exponential model was used for the preparation of the ordinary kriging, and after its completion, the observed and predicted values were used for the preparation of cross validation. In the cross-validation, for the exponential model results in a linear coefficient of 2.8753, angular coefficient of 0.8757, number of neighbors of 6,  $R^2$  of 0.64 and  $S_{yx}$ % of 5.25% (Table V). The  $R^2$  value was close to ideal (> 0.70). Similar results were obtained by Pelissari

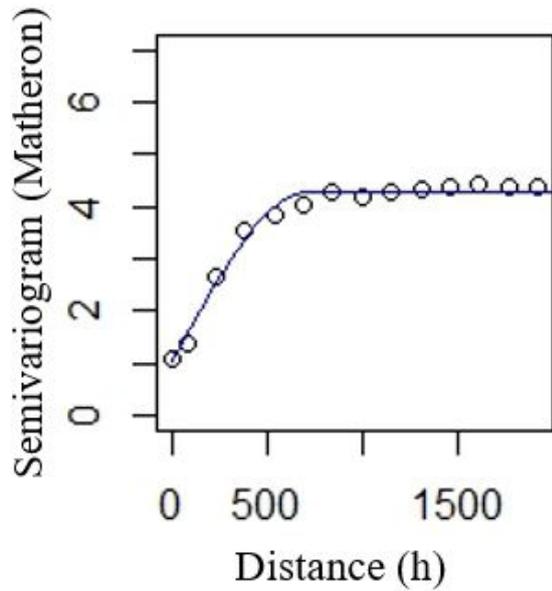
(2015). Despite  $R^2$  being smaller than expected, estimates for the dependent variable resulted in a standard error of estimate ( $S_{yx}$ %) of less than 10%, and was therefore considered adequate to represent the domain under study.

After the validation of the ordinary kriging and the exponential model, the processing routines to create the thematic map were generated (Figure 7) and the computation of areas per APU, stands, and site classes. The summary of areas and sites per APU, with their respective weights ( $W_{sj}$ ), is presented in Table VI.

### Some recommendations

(1) The results obtained confirmed better efficiency of the GADA for the establishment of the productive capacity in relation to the other tested methods. Therefore, it is recommended due to its more flexibility and its invariance for a reference age and simulation path. This allows obtaining polymorphic or anamorphic curves when it is included in the same equation more than one parameter as a function of the site.

(2) The generated curves and/or the coefficients of the models can be applied to the tropical regions of Brazil with similar edaphoclimatic characteristics in a reliable way, given the wide range of sample data, concerning ages of two to eight years and all detected site classes.

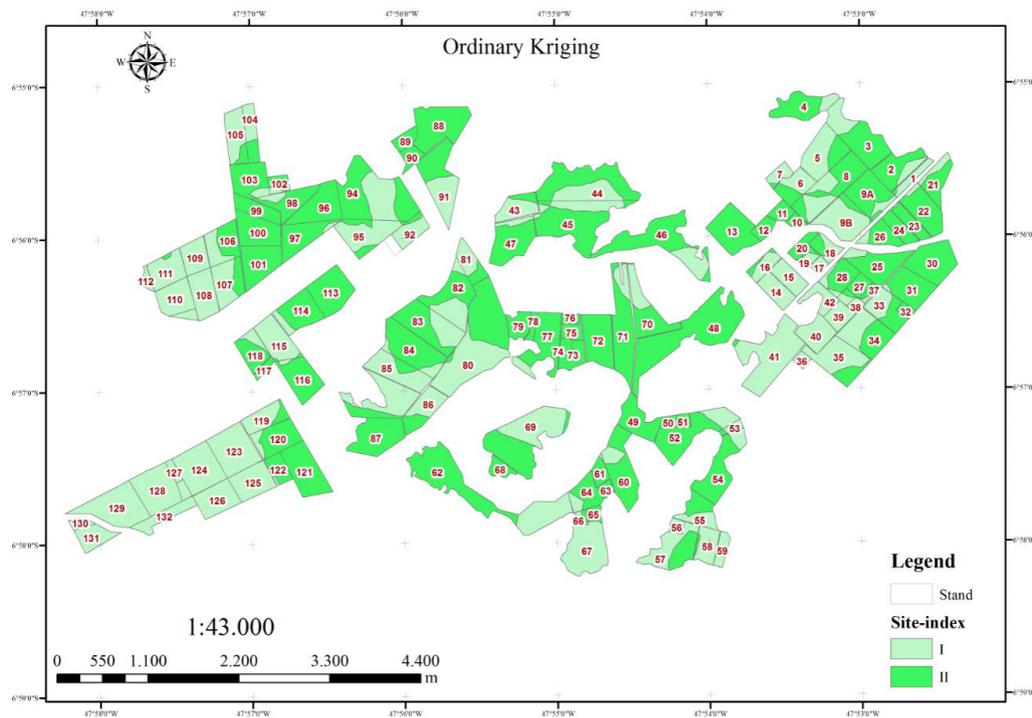


**Figure 6.** Semivariogram of the exponential model for *Eucalyptus* sp. in a tropical region of Brazil.

**Table V.** Cross-validation results for the exponential model for *Eucalyptus* sp. in a tropical region of Brazil.

| Model       | Linear coefficient | Angular coefficient | Number of neighbors | $R^2$ | $S_{yx}$ % |
|-------------|--------------------|---------------------|---------------------|-------|------------|
| Exponential | 2.8753             | 0.8757              | 6                   | 0.64  | 5.25%      |

Where:  $R^2$  is the coefficient of determination;  $S_{yx}$  % is the standard error of the estimate.



**Figure 7.** Thematic map of site classification based on ordinary kriging for *Eucalyptus* sp. in a tropical region of Brazil.

**Table VI. Summary of areas and sites per APU for *Eucalyptus* sp. in the tropical region of Brazil.**

| APU   | Sites | Areas (ha) | Weights WSj |
|-------|-------|------------|-------------|
| 2.5   | SI    | 516.43     | 24.5%       |
|       | SII   | 568.54     | 26.9%       |
| 3.5   | SI    | 507.42     | 24.0%       |
|       | SII   | 330.16     | 15.6%       |
| 4.5   | SI    | 180.09     | 8.5%        |
|       | SII   | 8.04       | 0.4%        |
| Total | -     | 2,110.66   | 100.0%      |

Where: UL is the upper limit; LL is the lower limit; and C is the center of the site index.

(3) The fitting of curves in young forest stands (less than three years) should be used with caution owing to the possible influence of silvicultural treatments and fertilization.

(4) Geostatistical techniques are indispensable tool for forest inventories because they allow generating accurate inferences on dominant height for classification of traditional sites in locations not sampled in the field, with a high degree of spatial dependence. We recommend the use of ordinary kriging to delimit, and calculate areas per site class in stands of *Eucalyptus* sp. In addition, other variables can be tested to differentiate the productive capacity, such as the physicochemical characteristics of the soil, rainfall, and evapotranspiration.

## CONCLUSIONS

Among the tested models of Chapman-Richards and Bailey 4-parameters used in the Guide Curve – GC approach, the modified Chapman-Richards model in two versions used in the ADA approach and the Chapman-Richards model with a set of dynamic equations used in GADA approach, the last one is more flexible and precise to

describe the relationship between dominant height growth at different ages and degrees of productivity and, thus, more appropriate than GC and ADA approaches.

The dominant height has spatial continuity structure, allowing the appropriate interpolation of productivity classes in the forest stand.

The ordinary kriging, applying the exponential model, resulted in the best performance with a high degree of spatial dependence, and was therefore an efficient tool to delimit the site index and allow the calculation of stratified areas for local production level.

## REFERENCES

- AMARO A, REED D, TOMÉ M & THEMIDO I. 1998. Modeling dominant height growth: *Eucalyptus* plantations in Portugal. *For Sci* 44(1): 37-46.
- ANDRADE LO. 2004. *Biometria Florestal Aplicada*. Três Barras: Ed. pelo autor, 189 p.
- ASSMANN E. 1970. *The Principles of Forest Yield Study*. Oxford: Pergamon Press Ltd, 506 p.
- BAILEY RL. 1980. The potential of Weibull-type functions as flexible growth-curves: discussion. *Can J For Res* 10: 117-118.
- BAILEY RL & CLUTTER JL. 1974. Base-age invariant polymorphic site curves. *For Sci* 20: 55-159.
- BATISTA JL & COUTO HTZ. 1986. Escolha de modelos matemáticos para a construção de curvas de índice de sítio para florestas implantadas de *Eucalyptus* sp. no estado de São Paulo. *IPEF* 32: 33-41.
- BOGNOLA IA, HIGA AR, STOLLE L, OLIVEIRA EB & FRANCISCON L. 2009. Modelagem da variabilidade espacial do rendimento produtivo de *Pinus taeda* L. com uso da geoestatística. In: *Anais XIV Brazilian Symposium on Remote Sensing*, Natal, Brazil, p. 3591-3596.
- BURKHART HE & TOMÉ M. 2012. *Modeling forest trees and stands*. Dordrecht: Springer, 457 p.
- CAMBARDELLA CA, MOORMAN TB, NOVAK JM, PARKIN TB, KARLEN DL, TURCO RF & KONOPKA AE. 1994. Field-scale variability of soil properties in central Iowa soils. *Soil Sci Soc Am J* 58(5): 1501-1511.
- CASTILLO LÓPEZ A, VARGAS-LARRETA B, RIVAS JJC, NÁJERA LUNA JA, CRUZ COBOS F & HERNÁNDEZ FJ. 2013. Modelo compatible

- altura-índice de sitio para cuatro especies de pino en Santiago Papasquiari, Durango. *Rev Mex Cienc Forestales* 4(18): 86-103.
- CIESZEWSKI CJ. 2001. Three methods of deriving advanced dynamic site equations demonstrated on inland Douglas-fir site curves. *Can J For Res* 31:1 65-173.
- CIESZEWSKI CJ. 2002. Comparing fixed and variable-base-age site equations having single versus multiple asymptotes. *For Sci* 48: 7-23.
- CIESZEWSKI CJ. 2003. Developing a well-behaved dynamic site equation using a modified Hossfeld IV function  $Y_3 = (axm)/(c + xm-1)$ , a simplified mixed-model and scant subalpine fir data. *For Sci* 49: 539-554.
- CIESZEWSKI CJ & BAILEY RL. 2000. Generalized algebraic difference approach: Theory based derivation of dynamic site equations with polymorphism and variable asymptotes. *For Sci* 46: 116-126.
- CIESZEWSKI CJ & STRUB M. 2008. Generalized algebraic difference approach derivation of dynamic site equations with polymorphism and variable asymptotes from exponential and logarithmic functions. *For Sci* 54: 303-315.
- CLUTTER JL. 1983. Timber management: a quantitative approach. New York: Wiley & Sons, 333 p.
- CURTIS RO, DEMARS DJ & HERMAN FR. 1974. Which dependent variable in site-index -height-age regressions. *For Sci* 20: 74-87.
- DAVID HC. 2014. Avaliação de sítio, relação dendrométrica e otimização de regimes de manejo de *Pinus taeda* L. nos estados do Paraná e Santa Catarina. Dissertação de Mestrado, Engenharia Florestal, Universidade Federal do Paraná, Curitiba, Brasil, p. 152.
- HUANG S. 1994. Ecologically based reference-age invariant polymorphic height growth and site-index curves for White spruce in Alberta. Land and Forest Services, Alberta Environmental Protection, Technical Report, no T305, Edmonton, AB, Canada, p. 128.
- KRUMLAND B & ENG H. 2000. Site-index systems for major Young-growth forest and woodland species in northern California. Forest Report, no. 4. California Department of Forestry and Fire Protection, Sacramento, CA, p. 219.
- MELLO JM, OLIVEIRA MS, BATISTA JLF, RIBEIRO JÚNIOR PJ & KANEGAE JUNIOR H. 2006. Uso do estimador geoestatístico para predição volumétrica por talhão. *Floresta* 36(2): 251-260.
- NEWNHAM RM. 1988. A modification of the Ek-Payandeh nonlinear regression model for site-index curves. *Can J For Res* 18: 115-120.
- PECE DE RIOS MGDV. 1993. Um modelo de crescimento e produção aplicado a plantações de *Eucalyptus pellita*. Dissertação de Mestrado, Ciências Florestais, Universidade Federal de Viçosa, Brasil, p. 89.
- PELLISSARI AL. 2015. Geoestatística aplicada ao manejo de povoamentos de *Tectona grandis* L. Tese de Doutorado, Engenharia Florestal, Universidade Federal do Paraná, p. 119.
- PÉLICO NETTO S, STEFANELLO FR, PELLISSARI AL & DAVID HC. 2014. Mapping of sites in forest stands. *An Acad Bras Cienc* 86: 2025-2037.
- R DEVELOPMENT CORE TEAM R: a language and environment for statistical computing. 2009. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- ROBINSON AP & HAMANN JD. 2011. Forest analytics with R: An Introduction. New York: Springer, 339 p.
- SCHUMACHER FX. 1939. A new growth curve and its application to timber yield studies. *J Forest* 37: 819-820.
- SCOLFORO JSR. 1993. Mensuração florestal: Avaliação de produtividade florestal através da classificação do sítio. Lavras: ESAL/FAEPE, 451 p.
- SCOLFORO JSR. 1998. Modelagem do crescimento e da produção de florestas plantadas e nativas. Lavras: UFLA/FAEPE, 451 p.

WEISKITTEL AR, HANN DW, KERSHAW JR JA & VANCLAY JK. 2011. Forest Growth and Yield Modeling. USA: J Wiley & Sons, 344 p.

#### How to cite

FILHO AM, NETTO SP, MACHADO SA, CORTE APD & BEHLING A. 2023. Site classification for *Eucalyptus* sp. in a tropical region of Brazil. *An Acad Bras Cienc* 94: e20200038. DOI 10.1590/0001-3765202320200038.

*Manuscript received on January 1, 2020;  
accepted for publication on October 20, 2020*

#### ATAÍDES M. FILHO

<https://orcid.org/0000-0001-8237-3430>

#### SYLVIO P. NETTO

<https://orcid.org/0000-0002-3178-1810>

#### SEBASTIÃO A. MACHADO

<https://orcid.org/0000-0003-1010-4623>

#### ANA P.D. CORTE

<https://orcid.org/0000-0001-8529-5554>

#### ALEXANDRE BEHLING

<https://orcid.org/0000-0002-7032-2721>

Universidade Federal do Paraná, Departamento de Ciências Florestais, Av. Pref. Lothário Meissner, 900, Campus III, Jardim Botânico, 80210-170, Curitiba, PR, Brazil

Correspondence to: **Sylvio P. Netto**  
E-mail: [sylviopelliconetto@gmail.com](mailto:sylviopelliconetto@gmail.com)

#### Author contributions

Ataídes Marinheski Filho: data collection, modeling the data and writhed the manuscript. Sylvio Péllico Netto: performed the original idea, review the manuscript and translated it to English. Sebastião do Amaral Machado: he made important suggestions for site index evaluation and review the manuscript. Ana Paula Dalla Corte: review the manuscript. Alexandre Behling: review the manuscript.

