Pertanika J. Trop. Agri. Sci. 46 (4): 1127 - 1150 (2023)

TROPICAL AGRICULTURAL SCIENCE

Journal homepage: http://www.pertanika.upm.edu.my/

Assessment of a Monthly Data Structure for Growth and Yield Projections from Early to Harvest Age in Hybrid Eucalypt Stands

Gianmarco Goycochea Casas[1]*, Carlos Pedro Boechat Soares[1], Márcio Leles Romarco de Oliveira[2], Daniel Henrique Breda Binoti[3], Leonardo Pereira Fardin[1], Mathaus Messias Coimbra Limeira[1], Zool Hilmi Ismail[4], Antonilmar Araújo Lopes da Silva[5] and Hélio Garcia Leite[1]

[1]*Department of Forest Engineering, Universidade Federal de Viçosa, 36570-900 Viçosa, Minas Gerais, Brazil*
[2]*Department of Forestry, Universidade Federal dos Vales do Jequitinhonha e Mucuri, 39100-000 Diamantina, Minas Gerais, Brazil*
[3]*Dap Florestal, Rua Papa João XXIII, n 9, Lourdes, 36572-006 Viçosa, Minas Gerais, Brazil*
[4]*Center for Artificial Intelligence and Robotics, Universiti Teknologi Malaysia, 54100 Kuala Lumpur, Malaysia*
[5]*Rua Visconde de Mauá, n 755, Nobre City, 35162-391 Ipatinga, Minas Gerais, Brazil*

ABSTRACT

Whole-stand Models (WSM) have always been fitted with permanent plot data organised in a sequential age-matched database, i.e., *i* and *i*+1, where *i* = 1, 2, ... *N* plot measurements. The objectives of this study were (1) to evaluate the statistical efficiency of a monthly distributed data structure by fitting the models of Clutter (1963), Buckman (1962) in the version modified by A. L. da Silva et al. (2006), and deep learning, and (2) to evaluate the possibility of gaining accuracy in yield projections made from an early age to harvest age of eucalypt stands. Three alternatives for organizing the data were analyzed. The first is with data paired in sequential measurement ages, i.e., *i* and *i*+1, where *i* = 1, 2, ... *N* plot measurements. In the second, all possible measurement intervals for each plot were considered, i.e., *i*, *i*+1; i, *i*+2; ...; *i*, *N*; *i*+1, *i*+2; ..., *N*-1, *N*. The third has data paired by month (*j*), always with an interval of one month, i.e., *j*, *j*+1; *j*+1, *j*+2; *j*+*M*-1, *M*, where *M* is the stand age of the plot measurement in months. This study shows that the accuracy and consistency of the projections depend on the organization of the monthly distributed data, except for the Clutter model. A better alternative to

ARTICLE INFO

Article history:
Received: 29 March 2023
Accepted: 02 May 2023
Published: 22 September 2023

DOI: https://doi.org/10.47836/pjtas.46.4.04

E-mail addresses:
gianmarco.casas@ufv.br (Gianmarco Goycochea Casas)
csoares.ufv@gmail.com (Carlos Pedro Boechat Soares)
marcioromarco@gmail.com (Márcio Leles Romarco de Oliveira)
danielhbbinoti@gmail.com (Daniel Henrique Breda Binoti)
leopfard@gmail.com (Leonardo Pereira Fardin)
mathaus.limeira@ufv.br (Mathaus Messias Coimbra Limeira)
zool@utm.my (Zool Hilmi Ismail)
antonilmar@gmail.com (Antonilmar Araújo Lopes da Silva)
hgleite@ufv.br (Hélio Garcia Leite)
*Corresponding author

ISSN: 1511-3701
e-ISSN: 2231-8542

increasing the statistical assumptions of the forecast from early to harvest age is based on a monthly distributed data structure using a deep learning method.

Keywords: Buckman, clutter, deep learning, forest management, volumetric projection

INTRODUCTION

The growth and yield modeling of even-aged stands is essential to forest management (Campos & Leite, 2017; Davis & Johnson, 1987). The models estimate harvest stocks for different prescriptions and thus support management plans, especially in hierarchical planning. Models are also important to figure out the average productivity at different sites used as inputs in studies related to the sustainability of production.

Growth and yield models can be used for prediction or projection. In the first case, growth and future yield are independent of current yield. In the second, the model's functional relationships use the current yield (in basal area and/or volume) as a predictor variable (i.e., an independent variable). A prediction model can be transformed and applied to differentiate future production according to current production, however, with the same growth trend (Campos & Leite, 2017). Estimating wood stocks for defined ages in a strategic planning model is known as prognosis; this term refers to any procedure that generates an expectation of future production. According to Burkhart and Tomé (2012) and Castro et al. (2013), the prediction describes the change in the size of the individual population over time.

One of the components or elements in forest management is prognosis, which is usually done using growth and yield models. Despite significant advances in studies of growth and yield modeling of forest stands with regression, problems persist. One of these problems is the low accuracy of projections made from early ages, for example, from two or three years of age to harvest age, which in Brazil is generally six or seven years for eucalypt stands to produce cellulose pulp or charcoal. This difficulty is related to the heteroscedastic nature of the relations between dominant height (Hd) and volume (V) with age (A), with less variance at early ages. In most cases, the dispersion of Hd and V at ages 1.5 to 2.5 years, in eucalypt stands, is relatively small. As the years go by, the variance of Hd and V increases, indicating that heteroscedasticity occurs naturally.

The main models used in Brazil are Whole-stand Models (WSM) for the use of wood, most often to produce cellulose pulp or charcoal. These aspects are discussed by Campos and Leite (2017) with reference to exponential and sigmoid models, in addition to the Buckman (1962) and Clutter (1963) models. In their discussions, they report that using simpler functional relationships requires more intense data stratification for modeling.

Buckman model (1962) had not been used on a commercial scale in Brazil despite several researchers intending to demonstrate its feasibility in its use (A. L. da Silva et al., 2006; Burkhart & Sprinz, 1984; Guera et al., 2019; Trevizol Jr., 1985). On the other hand, the Clutter model has always been used in different regions of Brazil since the 1985s (Miguel et al. 2016; Penido et al. 2020; Salles et al., 2012; Soares et al., 2004; Trevizol Jr., 1985).

Authors have shown that for ranking productive capacity (site index) using mathematical regression models, the evaluation of estimates and statistical parameters depends on both the model and the data structure used for modeling (Cao, 1993; Cieszewski, 2002; Hirigoyen et al., 2018; Strub & Cieszewski, 2006).

Another example influenced by the organization of data is also carried out when modeling with metaheuristic techniques such as Artificial Neural Networks (ANNs) for WSM, such is the case of irregular intervals (da Silva Binoti et al., 2015), i.e., i and i+1, where i = 1, 2, ... N plot measurements, for the prediction of the performance of different initial ages and intervals of prediction are necessary for hierarchical planning. Nonetheless, works related to the Buckman (1962) and Clutter (1963) models were not found.

ANNs have been increasingly used for various purposes in forest engineering (Araújo Júnior et al., 2019; de Freitas et al., 2020; Gavilán-Acuña et al., 2021; Lopes et al., 2020; S. Silva et al., 2020). In general, the accuracy of yield estimates at the stand level has been higher when ANN is employed compared to growth and yield models (Casas, Fardin, et al., 2022; de Oliveira Neto et al., 2022). The higher accuracy is partially explained by including categorical variables in the ANN. Despite the gain in accuracy with the use of ANN, the problem of inefficient projection for about 2 years (in the case of eucalypt stands) persists.

Deep learning is part of the broader field of machine learning and artificial intelligence that uses artificial neural networks. The difference with a conventional network is that deep networks allow computational models of multiple processing layers to learn data representations with various levels of abstraction (Aggarwal, 2018; LeCun et al., 2015). Understanding the irrational effectiveness of deep learning is very complex and is still being studied through advanced mathematics and neuroscience, as it is an inspirational source in the brain's architecture (Sejnowski, 2020). Deep learning has been increasingly used in research involving complex structures in different forestry and environmental areas, where studies have been reported on the hypsometric relationship between tree diameter and height (Casas, Gonzáles, et al., 2022), individual tree detection and species classification of Amazonian palms (Ferreira et al., 2020), forest damage assessment (Hamdi et al., 2019), analysis of drone-acquired forest images (Kentsch et al., 2020), automatic identification of charcoal origin (de Oliveira Neto et al., 2021), plant identification in the natural environment (Sun et al., 2017) and wood filtering, and tree species classification from terrestrial laser scanning (Xi et al., 2020).

Most studies using machine learning for prognosis comparisons are made using regression modeling. The following question arises: Is there any gain or loss in statistical efficiency when fitting the Clutter model, Buckman model, and deep learning when using a monthly distributed data organization from continuous inventories with permanent plots? The objectives of this study were to (1) evaluate the statistical efficiency of a monthly distributed data structure by fitting the models of Buckman (1962) and Clutter (1963) in the version modified by A. L. da Silva et al. (2006) and deep learning, and (2) evaluate the possibility of gaining accuracy in yield projections in eucalypt stands made from initial ages of approximately two years to the age of harvest with ages of six and seven years.

MATERIALS AND METHODS

Study Area Description

The data used in this study were measured from 1,243 permanent 400 m^2 rectangular plots distributed in 243 project stands of the hybrid *Eucalyptus urophylla* x *Eucalyptus grandis*, with a spacing of 3 x 3 m^2, located in the Midwest region of the Minas Gerais state, Brazil. Stand locations were at a mean altitude of 743 m, a mean annual precipitation of 1,163 mm, and a mean annual temperature of 27°C (Alvares et al., 2013). The database consisted of annual dasometric measurements from 2006 to 2015 for each permanent plot. The variables used in this study were dominant height (m), basal area (m^2/ha), and volume (m^3/ha). The statistical description of the dasometric variables for each region can be found in Table 1.

Structuring of Processing Data

Regression and deep learning methods were used. In the regression method, two WSM models were used: The Buckman

Table 1
Statistical description of dasometric variables of hybrid eucalypt stands located in the Midwest Region in the Minas Gerais state, Brazil

Subregion	Variable	Valid N	Mean	Minimum	Maximum	Variance	Standard deviation	Standard error
A	Age (months)	3,442	46.78	18.00	95.00	373.70	19.33	0.33
	Dominant Height (m)	3,442	21.59	7.63	39.24	35.23	5.94	0.10
	Basal Area (m^2/ha)	3,442	15.98	0.22	40.31	41.10	6.41	0.11
	Volume (m^3/ha)	3,442	153.84	1.22	558.11	9,243.80	96.14	1.64
B	Age (months)	2,576	39.47	20.00	82.00	209.68	14.48	0.29
	Dominant Height (m)	2,576	19.47	7.23	35.87	22.63	4.76	0.09

Table 1 *(Continue)*

Subregion	Variable	Valid N	Mean	Minimum	Maximum	Variance	Standard deviation	Standard error
	Basal Area (m^2/ha)	2,576	14.41	1.53	34.48	26.53	5.15	0.10
	Volume (m^3/ha)	2,576	138.82	5.03	444.22	6,001.68	77.47	1.53
C	Age (months)	1,976	37.10	19.00	73.00	158.39	12.59	0.28
	Dominant Height (m)	1,976	19.87	8.67	33.57	24.56	4.96	0.11
	Basal Area (m^2/ha)	1,976	14.45	3.11	28.91	26.56	5.15	0.12
	Volume (m^3/ha)	1,976	132.17	11.15	397.83	5,953.52	77.16	1.74

model (1962) and the Clutter model (1963), modified by A. L. da Silva et al. (2006). Three different data structures were fit for growth and yield modeling:

I. The two variable density models are equipped with the following data structure: paired data considering ascending intervals without overlap ([A_1-A_2], [A_2-A_3], . . . [A_n-A_{n+1}]). This structure has historically been used for fitting variable-density growth and yield models and other projection models.

II. ANN techniques are used with the following data structure: considering An as age at measurement n = 1, 2, 3, …. N in a measurement plot, the database was organized considering all possible age intervals for each plot, i.e. paired data considering all possible ascending age intervals ([A_1-A_2], [A_1-A_3], ... [A_2-A_3], [A_2-A_4], . . . [A_n-A_{n+1}], [A_n-A_{n+2}]). It was necessary so that the networks could be trained to generalize to different early ages and projection intervals and usually be used for these cases.

III. In this study, the following data structure was proposed: paired data is considered with ascending intervals without overlap by month (j), always with an interval of one month, i.e., j, j+1; j+1, j+2; j+M-1, M, where M is the stand age of the plot measurement in months. Linear interpolations were made between each plot's values to set up this structure.

Monthly structured data must be converted to Data Structure II to apply the deep learning method. In the case of any regression method, it is not necessary to perform the conversion for Data Structure II. By performing this conversion, we obtain the Data Structures II-A and II-B (Figure 1). Training the networks to generalize to different early ages and projection intervals was necessary.

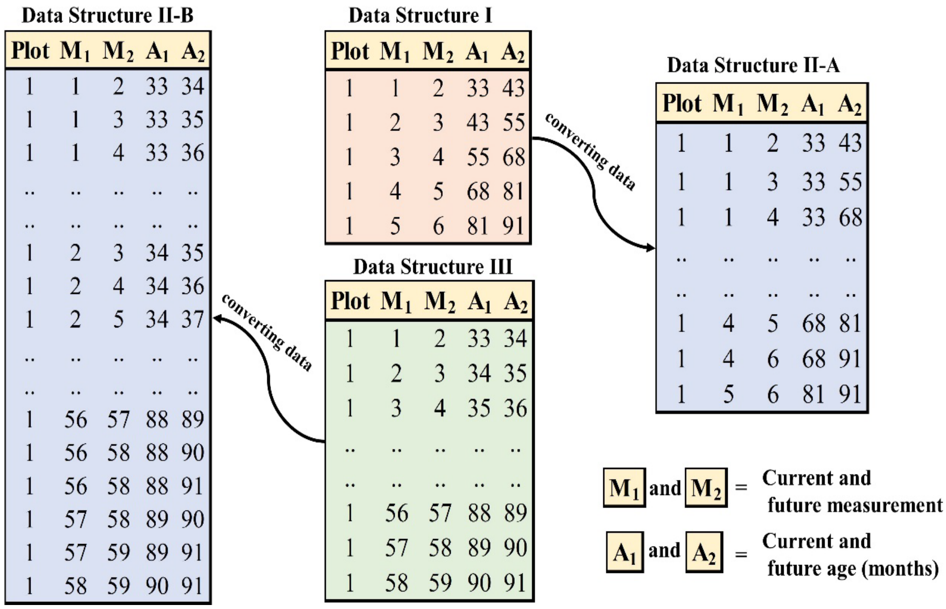

Figure 1. Example of a one-plot database structure used to fit Buckman model (1962) and Clutter model (1963) modified by A. L. da Silva et al. (2006) (Data Structures I and III) and deep learning (Data Structures II-A and II-B) models and the process of converting data from one structure to another structure according to the method to be used

Fitting the Data with the Variable Density Models

The growth and yield models for this study were the Clutter (Equations 1 and 2) and Buckman, modified by A. L. da Silva et al. (2006) (Equations 3 and 4). The Clutter model was fitted by the two-stage least squares method using applied econometrics with an R (AER) package (Kleiber & Zeileis, 2008) and the Buckman model by ordinary least squares in R. The equations are shown below:

$$\left.\begin{aligned} LnB_2 &= LnB_1\left(\frac{A_1}{A_2}\right)+\alpha_0\left(1-\frac{A_1}{A_2}\right)+\alpha_1\left(1-\frac{A_1}{A_2}\right)S+\varepsilon_1 \\ LnV_2 &= \beta_0+\beta_1\left(\frac{1}{A_2}\right)+\beta_2 S+\beta_3 LnB_2+\varepsilon_2 \end{aligned}\right\} \quad \text{(Equations 1 and 2)}$$

$$\left.\begin{aligned} LndB &= \alpha_0+\alpha_1 A_2+\alpha_2 S+\alpha_3 A_2^{-1}+\alpha_4 LnB_{1_1}+\varepsilon_1 \\ LnV_2 &= \beta_0+\beta_1 A_2^{-1}+\beta_2 S+\beta_3 LnB_2+\varepsilon_2 \end{aligned}\right\} \quad \text{(Equations 3 and 4)}$$

where V_2 = future yield (m^3/ha); A_1 and A_2 = current and future age (months); S = site index in the current age (m); B_1 and B_2 = current and future basal areas (m^2/ha); dB = increase in basal area between A_1 and A_2; Ln = neperian logarithm; α and β = model parameters; ε_1 and $\varepsilon_{2;}$ $\varepsilon_k \sim NID(0;\ \sigma^2)$ and $B_2 = B_1 + \sum_{j=A_1}^{j=A_2} dB_j$.

The guide curve method figured out the site indexes for each measurement (Clutter et al., 1983). In our study to figure out the site index, the Gompertz (1825) model has been selected since this model is the most adequate to estimate the dominant height as a function of age for eucalypt stands (Reis et al., 2022). The index age (Ai) was 72 months. This index age is appropriate for the eucalypt species in Brazil.

Thus, considering the classical transformation of the guide curve method, in which:

$Hd = \beta_0 e^{-e^{\beta_1 - \beta_2 A}} + \varepsilon$ with observe Hd at the age A

$S = \beta_0 e^{-e^{\beta_1 - \beta_2 Ai}} + \varepsilon$ at the index age Ai of 72 months

Therefore, the site index equation (Equation 5) was established by taking the differences between the above equations and expressing them explicitly in terms of S:

$$S = Hd\, e^{-e^{\beta_1 - 72\beta_2}} / e^{-e^{\beta_1 - \beta_2 A}}$$

(Equation 5)

where Hd = dominant height (m) observed at age A, S = site index (m), and β_1 and β_2 = parameter estimates of the Gompertz model.

Fitting the Data with Deep Learning

Input and Output Variables. In the output layer, the variable was future volume (V_2). From this variable in the output layer, we set the variables of the input layer using the following function (Equation 6):

$$V_2 = f\,(\text{Project stand}, A_1, A_2, B_1, H_d, V_1)$$

(Equation 6)

where V_2 = future volume (m^3/ha) in the output layer; A_1 = current age (month); A_2 = future age (month); B_1 = current basal area (m^2/ha); Hd = dominant height (m); and V_1 = current volume (m^3/ha) in the input layer. The project stand had categorical variables in the input layer.

Hyperparameter Tuning with Grid Search for Deep Learning. A Cartesian grid search was performed. It has performed the procedure for implementing the deep learning method and trained models for every combination of the hyperparameter values. For each function, 162 models were trained for the deep learning method. The data were processed in R (R Core Team, 2020) using the H2O package (Fryda et al., 2020). Their architectures had one input layer, two, three, and four hidden layers, and one output layer. The numbers of neurons in the hidden layer were 50:25, 50:25:5, 100:50, 100:50:25, 100:50:25:5, and 200:100:50. The functions of activations in the hidden layer were Tanh (Equation 7), Rectified Linear (Equation 8), and Maxout (Equation 9):

Tanh $\left[f(\alpha) = \frac{e^{\alpha} - e^{-\alpha}}{e^{\alpha} + e^{-\alpha}}\right]; f(x) \in [-1,1]$

(Equation 7)

Rectified Linear

$[f(\alpha) = \max \quad 0, \alpha)]; f(.) \in \mathbb{R}_{+}$ (Equation 8)

Maxout

$[f(.) = \max(w_i x_i + b)]; f(.) \in [-\infty, 1]$;

rescale if max $f(.) \geq 1$ (Equation 9)

In the output layer, it has configured Linear activation function $\left[f(\alpha)=\alpha\right]$ for all trained models, where *f* is the function that represents the non-linear activation used in the entire neural network; *b* is the bias for the neuron activation threshold; x_i and w_i denote the input values of the unit or neuron and their weights; *α* denotes the weighted combination: $\alpha = \sum_{j=1}^{n} w_i x_i + b$

Gaussian distribution function was configured as an equivalent to weighted mean squared error (wMSE) (Equation 10):

$$f(.) = \omega\left(y - f\right)^2 \quad \text{(Equation 10)}$$

where *y* is a true response, *f* is a predicted response and ω is weighted.

The loss function chosen was quadratic (Equation 11):

$$L(W, B \mid j) = \frac{1}{2}\left\| t^{(j)} - o^{(j)} \right\|_2^2$$

(Equation 11)

where $t^{(j)}$ and $o^{(j)}$ are the predicted output and actual output; *j* and *W* are the collection $\{W_i\}$ 1:*N*-1, which W_i denotes the weight matrix connecting layers *i* and *i* + 1 for a network of *N* layers; *B* is the collection $\{b_i\}_{1:N-1}$, which b_i denotes the column vector of biases for layer *i* + 1.

The mini-batch was of size 1, the number of epochs was 1,000, and the type of regularization was the early stopping system, with 5 stop rounds, a stop tolerance of 0.001, and the MSE stop metric.

The adaptive learning rate algorithm (ADADELTA) (Equation 12) (Zeiler, 2012) was used:

$$1.\ \textit{Accumulate gradient}:\ E[g^2]_t = \rho * E[g^2]_{t-1} + (1-\rho) * g_t^2$$

$$2.\ \textit{Compute update}:\ \Delta x_t = -RMS[\Delta x]_{t-1} / RMS[g]_t * g_t$$

$$3.\ \textit{Accumulate updates}:\ E[\Delta x^2]_t = \rho * E[\Delta x^2]_{t-1} + (1-\rho) * (\Delta x_t)^2 \quad \text{(Equation 12)}$$

$$4.\ \textit{Apply update}:\ x_{t+1} = x_t + \Delta x_t$$

where g_t is the gradient at time step *t*; $E\left[g^2\right]_t$ is the running average of the squared gradient at time step *t*; Δx_t is the update at time step *t*; $RMS[\Delta x]_{t-1} = \sqrt{E\left[\Delta x^2\right]_{t-1} + \in}$ is the running average of the squared updates at time step *t*; $RMS[g]_t = \sqrt{E\left[g^2\right]_t + \in}$ is the root mean square of updates at time step *t*; $RMS[g]_t = \sqrt{E\left[g^2\right]_t + \in}$ is the root mean square of gradients at time step *t*; ρ (Rho) is the decay rate; ∈ (Epsilon) is a small constant for numerical stability.

This algorithm works with two parameters (Rho and Epsilon), which were configured with 0.9, 0.95, and 0.999 for Rho and 1e^{-06}, 1e^{-08}, and 1e^{-10} for Epsilon while g_t denoting the parameters at the *t*-th iteration, g_t as the computed gradient, *t* is the time, and RMS is the root mean squared error.

Model Performance

The databases were split into training (70%) and validation (30%). The following statistics were used in the validation data: the linear correlation coefficient between observed and corresponding projected yields ($r_{y\hat{y}}$) (Equation 13), mean absolute deviations (MAD) (Equation 14), percent root mean square error (RMSE%) (Equation 15), Bias (Equation 16), percent Bias% (Equation 17) and percent relative error (RE%) (Equation 18). In addition, the distribution graph of RE% *vs*. fitting was interpreted. The estimators used were:

$$r_{y\hat{y}} = \frac{n^{-1}\sum_{i=1}^{n}(Y_{pi}-\hat{Y}_m)(Y_i-\bar{Y})}{\sqrt{n^{-1}\sum_{i=1}^{n}(Y_{pi}-\hat{Y}_m)^2 \, n^{-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2}} ; Y_m = n^{-1}\sum_{i=1}^{n} Y_{pi} \quad \text{(Equation 13)}$$

$$MAD = \left(n^{-1}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right|\right); \quad \text{(Equation 14)}$$

$$RMSE\% = 100\bar{Y}_i^{-1}\sqrt{n^{-1}\sum_{i=1}^{n}(\hat{Y}_i - Y_I)^2}; \quad \text{(Equation 15)}$$

$$Bias = \sum_{i=1}^{n}\frac{\left(\hat{Y}_i - Y_i\right)}{n}; \quad \text{(Equation 16)}$$

$$Bias\% = 100\bar{Y}_i^{-1}\sum_{i=1}^{n}\frac{\left(\hat{Y}_i - Y_i\right)}{n}; \quad \text{(Equation 17)}$$

$$RE\% = 100\left(\frac{\hat{Y}_i - Y_i}{Y_i}\right) \quad \text{(Equation 18)}$$

where n = number of observations; $\bar{Y}$ = projected yield; $\bar{Y}$ = observed yield; and $\bar{Y}$ = mean observed yield.

RESULTS

Assessment of Growth and Yield Projection

Assessment of Production Capacity. The parameter estimates of the Gompertz model are shown in Table 2, where all coefficients were significant ($p < 0.01$). The same table shows the precision and accuracy statistics obtained from the fit when applied to the validation data with r = 0.9112, MAD = 1.7104, RMSE% = 11.8216, bias = 0.1200, and bias% = 0.0066.

Table 2
Statistical indicators and evaluation parameters obtained using the Gompertz model in hybrid eucalypt stands

Gompertz model								
Parameter	Coefficient	Standard error	*t*-statistic	r	MAD	RMSE (%)	Bias	Bias (%)
β_1	29.3806	0.3676	79.94					
β_2	0.9578	0.0387	24.73	0.9112	1.7104	11.8216	0.12	0.0066
β_3	0.0511	0.002	25.85					

Note. r = Correlation between observed and estimated dominant height (*Hd*); MAD = Mean absolute deviation; RMSE% = Percent mean square error; Bias% = Percent bias

The limit tables of dominant heights were established from the obtained coefficients and applied to the site index equation (Equation 5), being able to construct the site index curves at an index age of 72 months to stand for the productive capacity (Figure 2) illustratively. The site index has been set from 22 to 32 m with an amplitude of 2 m.

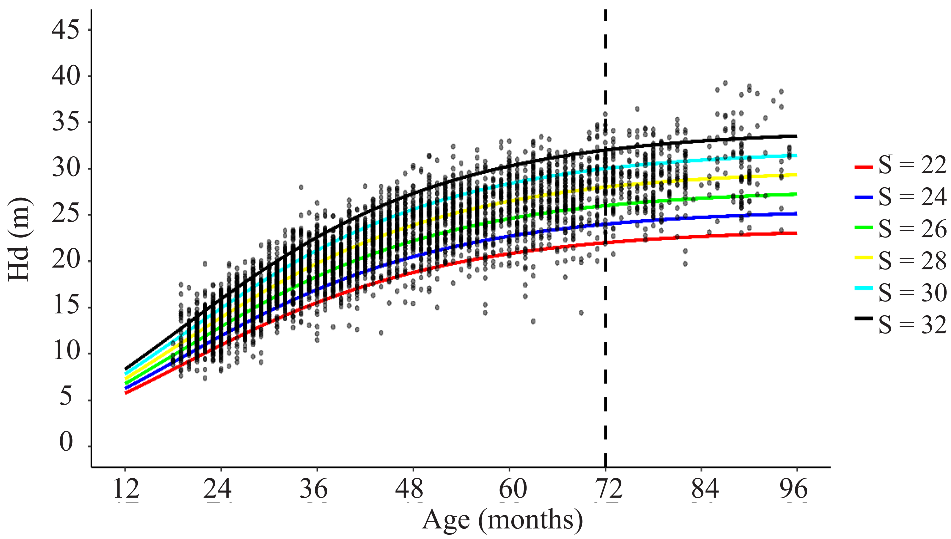

Figure 2. Site index (S) curves at the index age of 72 months in hybrid eucalypt stands

Evaluation of Variable Density Models. The parameter estimates of the Clutter and Buckman model modified by A. L. da Silva et al. (2006) for Data Structures I and III are shown in Table 3, where all coefficients were significant ($p < 0.01$). The results of the statistical parameters evaluated for both models and the data structures evaluated are also shown.

The comparison of fits between Data Structures I and III reveals that, in the case of the Clutter model (1963), Data Structure I provides superior statistical estimates, with r = 0.9262, MAD = 23.6587, RMSE% = 14.6257, bias = -4.3151, and bias% = -0.0187.

In the case of the Buckman model modified by A. L. da Silva et al. (2006), a large difference is observed in their statistical estimations of the future yield. The Structure III database was clearly better than the structure I database with r = 0.8658, MAD = 33.3512, RMSE% = 20.4896, bias = -164,653, and bias% = -0.0712.

The Clutter model fitted using Data Structure I (Figure 4A) showed a better constant variance than Data Structure III (Figure 4B). An improvement in its residual distribution was also observed when fitting the Buckman model modified by A. L. da Silva et al. (2006) using Data Structure III (Figure 4D) rather than using Data Structure I (Figure 4C).

Table 3
Parameter estimates and their respective precision and accuracy statistics of Clutter and Buckman's models fitted with Database Structures I and III in hybrid eucalypt stands

Clutter model (1963)									
Data Structure I									
	Coefficient	Std. error	*t*-statistic	Variable	r	MAD	RMSE (%)	Bias	Bias (%)
α_0	3.932558	0.038727	101.5454						
α_1	-0.008756	0.001392	-6288644	B_2	0.9236	1.3949	9.2528	-0.1118	-0.0053
b_0	0.877906	0.044012	19.94678						
b_1	-10.24837	0.481462	-21.28594						
b_2	0.018141	0.000727	24.94387	V_2	0.9262	23.6587	14.6257	-4.3151	-0.0187
b_3	1.38008	0.014437	95.59182						

Clutter model (1963)									
Data Structure III									
	Coefficient	Std. error	*t*-statistic	Variable	r	MAD	RMSE (%)	Bias	Bias (%)
α_0	3.981974	0.033332	119.4652						
α_1	-0.0128	0.001217	-1051407	B_2	0.9219	1.4789	9.9031	-0.6373	-0.0302
b_0	1.603696	0.007266	220.7228						
b_1	-22.6567	0.100077	-226.3926						
b_2	0.035006	0.000139	251.6177	V_2	0.9143	26.0424	16.1049	-9.8251	-0.0425
b_3	1.06014	0.002439	434.6499						

Buckman model modified by A. L. da Silva et al. (2006)									
Data Structure I									
	Coefficient	Std. error	*t*-statistic	Variable	r	MAD	RMSE (%)	Bias	Bias (%)
α_0	2.735686	0.200096	13.672						
α_1	0.009577	0.001975	4.85						
α_2	0.050674	0.003933	12.883	B_2	0.4655	4.0692	27.3475	-0.1448	-0.0069
α_3	-3.344789	4.096659	-0.816						
α_4	-1.194655	0.040544	-29.466						
b_0	1.008	0.04	24.82						
b_1	0.0193	0	27.19	V_2	0.5766	60.0655	38.0952	-4.495	-0.0194
b_2	-11.42	0.46	-25.04						
b_3	1.333	0.01	102						

Table 3 *(Continue)*

Buckman model modified by A. L. da Silva et al. (2006)									
Data Structure III									
Coefficient	Std. error	*t*-statistic	Variable	r	MAD	RMSE (%)	Bias	Bias (%)
-2.7700256	0.0610517	-45.37						
-0.0173024	0.0004176	-41.43						
-0.0239686	0.000876	-27.36	B_2	0.823	2.3147	14.9104	-1.1643	-0.0552
42.9365767	1.077054	39.87						
0.721791	0.0150107	48.09						
1.598209	0.007283	219.4						
0.034926	0.000139	251.3	V_2	0.8658	33.3512	20.4896	-16.4653	-0.0712
-22.581012	0.100189	-225.4						
1.062248	0.002439	435.6						

Note. r = Correlation between observed and projected future yield (V_2); MAD = Mean absolute deviation; RMSE% = Percent mean square error; Bias% = Percent bias

Evaluation of Deep Learning Models. In the beginning, 167 deep learning models were trained with different hyperparameter configurations for each II-A and II-B data structure. However, in Data Structure II-A, only 159 models found an optimal solution, and in Data Structure II-B, only 157 models found an optimal solution. The best-trained model for Structure II-A presented the following hyperparameter configuration: activation = Maxout, epsilon = 1.00e^{-10}, hidden layer = [200,100,50], rho = 0.9, and epoch = 66.23. The best-trained model for structure II-B presented the following hyperparameter configuration: activation = Rectifier, epsilon = 1.00e^{-08}, hidden layer = [100,50], rho = 0.999, and epoch = 4.59. Initially, 1,000 epochs were configured, but these were not necessary to complete since the early stopping configuration allows the training to stop early during the creation and scoring of the model, thus avoiding overfitting. The Status of neuron layers of the best models can be seen in Table S1. It has also evaluated the learning curves of each best model of the II-A (Figure 3A) and II-B (Figure 3B) Data Structures, both for the training data set and the validation data set in which it has used the RMSE over the number of epochs. Both results were adequately representative; however, using a monthly Data Structure (II-B Data Structure), a greater number of epochs were not needed compared to the II-A Structure.

Statistical indicators of both Data Structures were analyzed (Table 4). The results showed that using the II-B Data Structure, there was a statistical gain in its parameters with R = 0.9758, MAD = 14.0096, RMSE% = 8.5838, bias = -4.5534, and bias% = -0.0197.

Illustratively, regular distribution of the residuals is observed when trained with an II-B Data Structure (Figure 4F), i.e., the residuals were randomly distributed around the zero value and, in fact, with a statistical gain compared to the II-A Structure (Figure 4E).

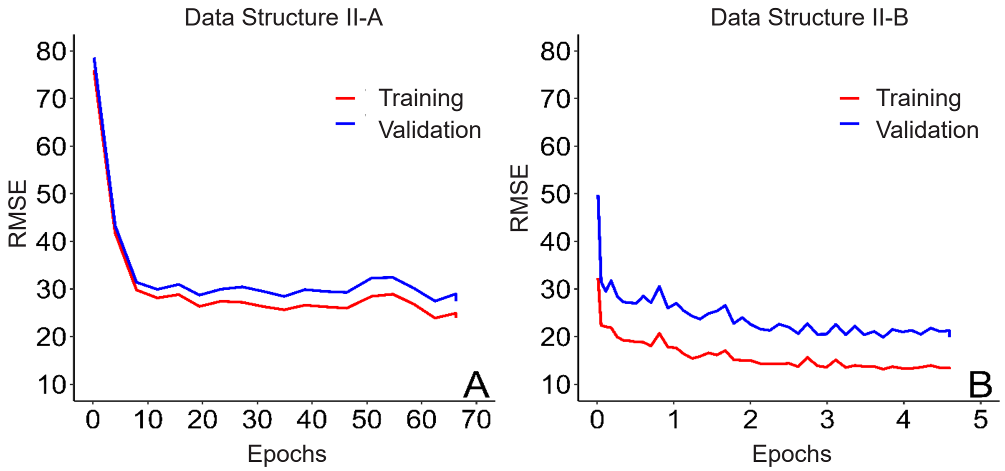

Figure 3. Model learning performance over the number of epochs trained in hybrid eucalypt stands
Note. A = Learning performance using Data Structure II-A; B = Learning performance using Data Structure II-B; RMSE = Root mean square error

Table 4
Precision and accuracy statistics of deep learning model fitted with database structures II-A and II-B in hybrid eucalypt stands

Model	Data Structure	Variable	r	MAD	RMSE (%)	Bias	Bias (%)
Deep learning	II-A	V_2	0.9506	19.4537	11.8697	-1.3161	-0.0057
	II-B	V_2	0.9758	14.0096	8.5838	-4.5534	-0.0197

Note. r = Correlation between observed and projected future yield (V_2); MAD = Mean absolute deviation; RMSE% = Percent mean square error; Bias% = Percent bias

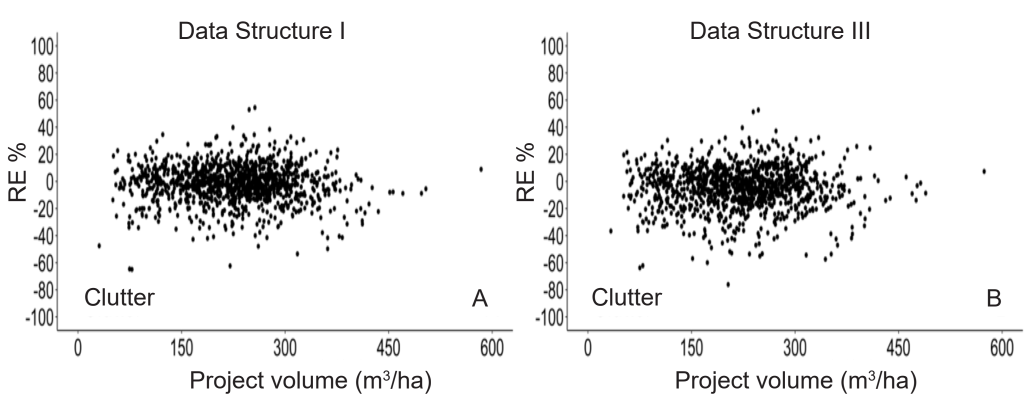

Figure 4. Dispersions of percent relative errors (RE%) as a function of estimated performance for the data structures and models in hybrid eucalypt stands

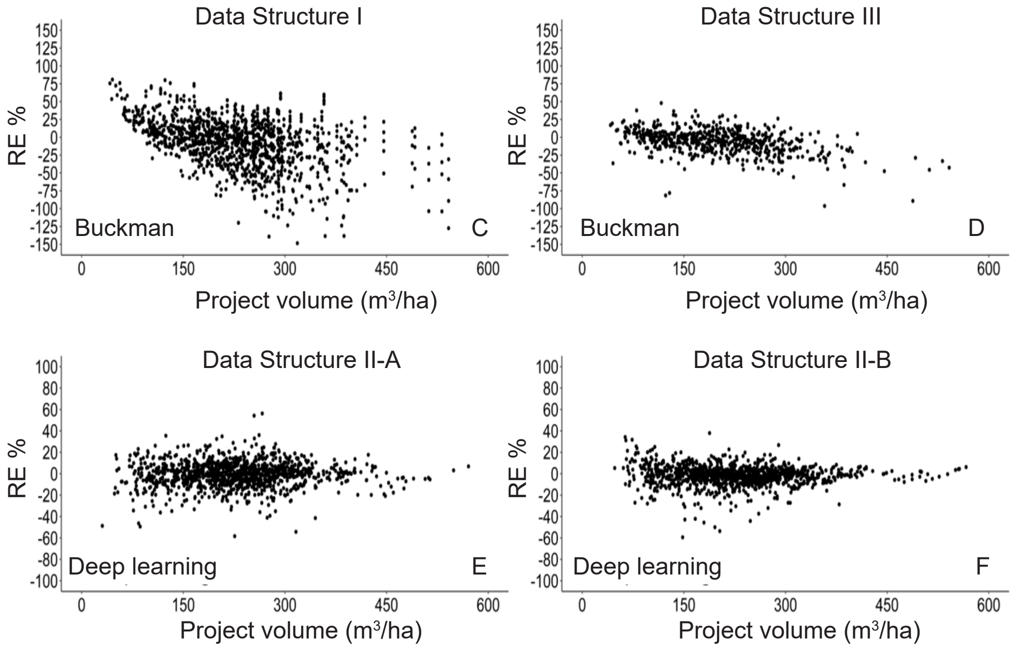

Figure 4. (Continue)
Note. A = Fitting the Clutter model using Data Structure I; B = Fitting the Clutter model using Data Structure III; C = Fitting the modified Buckman model using Data Structure I; D = Fitting the modified Buckman model using Data Structure III; E = Fitting the deep learning model using the Data Structure II-A; F = Fitting the deep learning model using the Data Structure II-B

Evaluation of Growth and Yield Projections from Early Age to Harvest Age

A statistical comparison was made with the methods and data structure evaluated to evaluate the accuracy of the volumetric projections from the early ages (two years) to the harvest age (six and seven years) (Table 5). The RMSE% plot was performed for each projection at the harvest age of 6 years (Figure 5A) and 7 years (Figure 5B) throughout the early ages to understand the behavior of the projections. The deep learning with Data Structure II-B presented greater precision in its statistics, which resulted in a lower variation of values between ages compared to the other methods and data structure. The accuracy generally increases as the early age increases except with the Buckman model with Structure III when the projection is for six years, but it has a resounding effect when the projection is for two years for both projection ages.

The methods used with a monthly data structure in this study present good estimates for volume projections from an early age, but when using artificial intelligence as a method of deep artificial neural networks, there is a tendency to decrease the RMSE%, an important indicator in volume projections, which can be considered in decision-making and influence the management plan. In the case of the Clutter model, despite not having statistical gain with a Structure III, it presents good statistical estimates; however, it is always recommended to fit the model as is typically done, employing Structure I.

Table 5
Statistical results of volumetric projections from early to harvest age by Clutter, Buckman, and deep learning models using a study data structure in hybrid eucalypt stands

Model	Data Structure	Early age	Harvest age	r	MAD	RMSE (%)	Bias	Bias (%)
Clutter	I	2	6	0.6045	33.0581	15.4798	-13.0763	-0.0487
		3		0.7425	25.9512	12.8617	5.2079	0.0201
		4		0.8488	21.0546	10.1621	0.4931	0.0018
		5		0.9332	11.7644	5.6019	-0.5465	-0.002
		2	7	0.4788	32.3557	18.2733	-3.6332	-0.012
		3		0.7516	49.6076	18.9629	-20.3878	-0.0617
		4		0.9058	41.4839	14.9614	-30.8371	-0.0879
		5		0.9105	7.8335	5.3841	-4.6331	-0.0138
	III	2	6	0.5292	40.6503	18.8847	-25.016	-0.0931
		3		0.6717	29.6257	14.33	-0.9692	-0.0037
		4		0.8396	22.4827	10.4329	-3.1825	-0.0118
		5		0.9373	11.5962	5.6741	-5.1036	-0.019
		2	7	0.4031	37.1509	19.9158	-9.5745	-0.0318
		3		0.6969	53.9502	20.3385	-26.7242	-0.0809
		4		0.8759	47.0845	17.0584	-37.7946	-0.1077
		5		0.927	8.8185	5.827	-7.1168	-0.0211
Buckman modified by A. L. da Silva et al. (2006)	I	2	6	0.0045	93.5816	39.5078	-86.1337	-0.3205
		3		0.2467	53.8019	25.3614	-14.2639	-0.055
		4		0.2547	56.9629	25.9555	17.8658	0.0659
		5		0.2706	48.8036	24.7695	26.8213	0.1001
		2	7	-0.1793	61.3515	32.1259	-44.8372	-0.1487
		3		0.025	102.6234	37.2318	-70.2897	-0.2128
		4		0.3843	64.9168	23.739	-20.1564	-0.0574
		5		0.786	11.6204	8.5094	9.1789	0.0272
	III	2	6	0.6481	36.9178	16.8792	-23.2161	-0.0864
		3		0.6345	33.2872	16.3512	-13.6835	-0.0528
		4		0.5674	38.6764	17.7144	-17.3354	-0.0639
		5		0.5943	33.5235	15.9619	-22.4244	-0.0837
		2	7	0.5108	34.3482	18.7104	-9.8646	-0.0327
		3		0.6816	59.2349	21.8192	-37.5963	-0.1138
		4		0.7783	60.4282	20.4033	-48.3494	-0.1378
		5		0.7383	13.5624	8.3983	-5.7556	-0.0171

Table 5 *(Continue)*

Model	Data Structure	Early age	Harvest age	r	MAD	RMSE (%)	Bias	Bias (%)
Deep learning	II-A	2	6	0.7348	26.7587	12.7355	-9.2113	-0.0343
		3		0.7818	23.9604	11.964	1.7124	0.0066
		4		0.8934	17.9795	8.5486	1.4199	0.0052
		5		0.9413	10.3057	5.2914	0.7683	0.0029
		2	7	0.7403	24.1687	14.1892	-0.1785	-0.0006
		3		0.8565	32.7332	13.6318	0.9818	0.003
		4		0.939	23.7861	8.5571	-7.8113	-0.0223
		5		0.9537	5.127	3.3905	-0.1354	-0.0004
	II-B	2	6	0.8992	20.6304	10.0871	-15.4392	-0.0574
		3		0.9353	15.0248	7.3165	-7.045	-0.0272
		4		0.9634	12.2011	5.3415	-3.6078	-0.0133
		5		0.9395	12.9134	5.9563	-6.9718	-0.026
		2	7	0.9066	16.847	9.3742	-7.648	-0.0254
		3		0.9284	21.8407	10.01	-10.6502	-0.0322
		4		0.9711	14.9505	5.7711	-5.8769	-0.0167
		5		0.9777	4.0758	2.8731	-3.0824	-0.0091

Note. r = Correlation between observed and projected future yield (V_2) from early to harvest age; MAD = Mean absolute deviation; RMSE% = Percent mean square error; Bias% = Percent bias

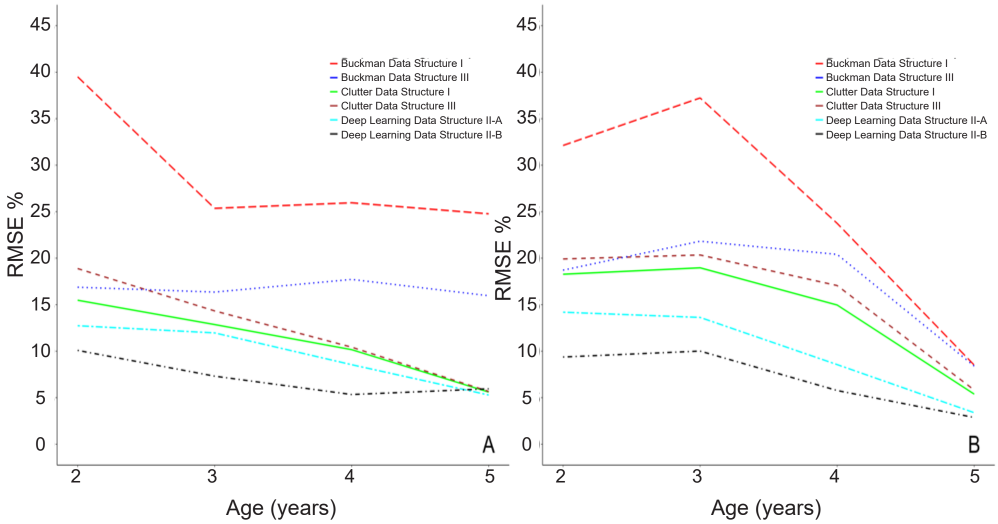

Figure 5. Percent root mean square error (RMSE%) of volumetric projections from early to harvest age by Clutter, Buckman, and deep learning models using a study data structure in hybrid eucalypt stands
Note. A = Harvest age at 6 years; B = Harvest age at 7 years

DISCUSSION

This study analyzed the effect of a monthly data structure on the accuracy and projection of the Clutter, Buckman modified by A. L. da Silva et al. (2006), and deep learning models.

In the case of WSM, the site index was acquired by utilizing the guide curve method and fitting the Gompertz model. This model is very well suited to represent the dominant height growth in eucalypt stands (Reis et al., 2022) since it follows a pattern of biological realism indicating an inflection point at its maximum growth rate, and from there, it indicates its decrease and stagnation of growth (W. S. Silva et al., 2021), a usual asymptotic characteristic in the growth of trees. Although the guide curve method has indeed been criticized for a long time (Socha & Tymińska-Czabańska, 2019), it is the method most used by the forest industries in Brazil, due to its ease of application and without prejudice to growth and yield projections. However, if the purpose is purely to classify the productive capacity, it may present errors, and it is up to the modeler in charge of the study area to be careful.

These two variable density models usually have a database structure with fixed age intervals between two consecutive measurements. Likely, the Clutter model was always fitted using a data structure in which the information was organized by pairing the ages as *i* as *i+1*. For example, if measurements were taken at 2, 3, 4, and 5 years then, to fit the Clutter model, the data were paired in 2–3, 3–4, and 4–5, that is, four records or lines from a permanent plot are transformed into three records.

The idea of changing the database structure with growth intervals to fit growth and production models arose from the principle of implementing non-parametric methods (Mongus et al., 2018; Vieira et al., 2018). Furthermore, statistical assumptions are likely to be violated due to the longitudinal characteristics of the data. Nevertheless, it is essential to consider that the growth intervals can be utilized for obtaining the parameters (Dorado, 2004), and the data can also be organized for the evaluations of volumetric projections.

The Clutter model is widely used in the forestry field and is the most widespread model in Brazil (Campos & Leite, 2017), with many studies conducted and preferred by researchers that over the years, it has served as a reference to express other growth and production (Soares et al., 2004), which gives it a fundamental relevance in forestry measurement works.

Authors highlighted using the modified Buckman model for volumetric estimates (A. L. da Silva et al., 2006; Guera et al., 2019), but the model is not used commercially in Brazil. This study presented results of its volumetric estimates that were highly biased when fitted with the usual data structure (Data Structure I); however, the model had an increase in accuracy when fitted with Data Structure III. The model may not have been used in Brazil due to inconsistencies in the projections because of how the database is organized. The repeated solutions method was used to obtain the basal area, which consists of repeated solutions based on the basal area and involves solving the growth equation for a given site index, age, and population density. The growth is then added to the early density, one month is added to the age, and the equation is solved again (A. L. da Silva et al., 2006). The database structure may rely on an equation for future studies to obtain the monthly data.

The data to fit the Buckman model can, in principle, be arranged in constant one-year (or monthly) intervals for all plots. However, the organization in one-year intervals limits the application of the model in strategic planning because it is often necessary to make production projections for each management unit for a single reference month, for example, December of each year. In this case, monthly projections are needed since the early measurements of the plot are not the same.

The use of artificial neural networks has solved several problems in forest management and, in most cases, with higher accuracy than regression models. In general, many studies have demonstrated the effectiveness of the use of artificial neural networks (da Rocha et al., 2021; da Silva Binoti et al., 2013; de Alcântra et al., 2018; de Freitas et al., 2020; dos Reis Martins et al., 2016), even when comparing artificial neural networks with regression methods, ANNs show superiority in their statistical results (Casas, Fardin, et al., 2022; da Silva Binoti et al., 2014; da Silva Tavares Júnior et al., 2019; Lopes et al., 2020; M. L. M. da Silva et al., 2009; Vendruscolo et al., 2017).

A large amount of data was used in this study, which led to the use of the deep learning method. Deep learning has a hierarchical structure, which makes it particularly suitable for learning knowledge hierarchies (Nielsen, 2015). This evidence can be seen when analyzing their residual distribution graphs, in which deep learning develops with robust and constant variance, and even more so when Data Structure II-B is used (Figure 4F).

The volumetric projections were statistically analyzed for projection ages of 6 and 7 years, and it was seen that the deep learning method presents superiority in its statistical evaluation, especially when using a monthly distributed data structure (Data Structure II-B). Using the deep learning method with an II-B Data Structure, the percent mean square error (RMSE%) decreased more than the other methods and data structure (Figure 5B). However, it still does not achieve statistical performance (when the early age is 2 or 3 years) equal to that achieved when the early age is close to the projection age, i.e., from the early age of five years to a projection age of six or seven years.

Models to predict future yield improve the understanding of tree growth in forest plantations (Lhotka, 2017). Project for close ages is recommended, as yield can be overestimated when projected for lengthy intervals (Weiskittel et al., 2016). This study's findings confirm a consistent pattern observed in growth modeling studies: lower accuracy in projections when conducted from an early age (around two years), irrespective of the chosen model.

This study confirms what has been seen in growth modeling studies, i.e., lower accuracy when the projection is made from an early age (approximately two years), regardless of the model used.

It is highly likely that this could vary if a larger number of categorical variables, such as soil type, were used (Casas, Fardin, et al., 2022); however, not all forest industries have access to such information. This study has made an effort to utilize simple models and readily available variables accessible to the forestry scientific community and the forest industry.

CONCLUSION

A better alternative to increase the statistical assumptions of the forecast from early to harvest age is based on a monthly distributed data structure using the deep learning method. It is set up so that the Clutter model should be fitted in the usual way it has been fitted, i.e., with the Data Structure I. Fitting the Buckman model requires uniform intervals between measurements. In the case of eucalypt stands, it must be a monthly interval to be effectively applied in strategic planning. The Buckman model was most accurate when data were organized month-by-month and without overlap, corresponding to Data Structure III evaluated in this study. Investing more in research is necessary to obtain greater assertiveness and precision in projections from an early age with variables that are easy to get.

FUNDING

This study was financed in part by the "Fundação de Amparo à Pesquisa do Estado de Minas Gerais—FAPEMIG"—Financing Code 507, and the "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil—CAPES"—Financing Code 001, in collaboration with the "Fundação Arthur Bernardes—FUNARBE."

ACKNOWLEDGMENTS

The authors are grateful to the "Fundação de Amparo à Pesquisa do Estado de Minas Gerais—FAPEMIG"—Financing Code 507, "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil—CAPES"—Financing Code 001, "Fundação Arthur Bernardes—FUNARBE" and Federal University of Viçosa, Brazil.

REFERENCES

Aggarwal, C. C. (2018). *Neural networks and deep learning*. Springer. https://doi.org/10.1007/978-3-319-94463-0

Alvares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. D. M., & Sparovek, G. (2013). Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, *22*(6), 711-728. https://doi.org/10.1127/0941-2948/2013/0507

Araújo Júnior, C. A., de Souza, P. D., de Assis, A. L., Cabacinha, C. D., Leite, H. G., Soares, C. P. B., da Silva, A. A. L., & Castro, R. V. O. (2019). Artificial neural networks, quantile regression, and linear regression for site index prediction in the presence of outliers. *Pesquisa Agropecuária Brasileira*, *54*, e00078. https://doi.org/10.1590/S1678-3921.pab2019.v54.00078

Buckman, R. E. (1962). *Growth and yield of red pine in Minnesota*. AgEcon Search: Research in Agricultural and Applied Economics. https://doi.org/10.22004/ag.econ.170997

Burkhart, H. E., & Sprinz, P. T. (1984). Compatible cubic volume and basal area projection equations for thinned old-field loblolly pine plantations. *Forest Science*, *30*(1), 86–93. https://doi.org/10.1093/forestscience/30.1.86

Burkhart, H. E., & Tomé, M. (2012). *Modeling forest trees and stands*. Springer. https://doi.org/10.1007/978-90-481-3170-9

Campos, J. C. C., & Leite, H. G. (2017). *Mensuração florestal: Perguntas e respostas* (5. ed. atual. e ampl.) [Forest measurement: Questions and answers (5th ed.)]. Editora Universidade Federal de Viçosa.

Cao, Q. V. (1993). Estimating coefficients of base-age-invariant site index equations. *Canadian Journal of Forest Research*, *23*(11), 2343–2347. https://doi.org/10.1139/x93-288

Casas, G. G., Fardin, L. P., Silva, S., de Oliveira Neto, R. R., Binoti, D. H. B., Leite, R. V., Domiciano, C. A. R., de Sousa Lopes, L. S., da Cruz, J. P., do Reis, T. L., & Leite, H. G. (2022). Improving yield projections from early ages in eucalypt plantations with the clutter model and artificial neural networks. *Pertanika Journal of Science and Technology*, *30*(2), 1257–1272. https://doi.org/10.47836/pjst.30.2.22

Casas, G. G., Gonzáles, D. G. E., Villanueva, J. R. B., Fardin, L. P., & Leite, H. G. (2022). Configuration of the deep neural network hyperparameters for the hypsometric modeling of the *Guazuma crinita* Mart. in the Peruvian Amazon. *Forests*, *13*(5), 697. https://doi.org/10.3390/f13050697

Castro, R. V. O., Soares, C. P. B., Martins, F. B., & Leite, H. G. (2013). Growth and yield of commercial plantations of eucalyptus estimated by two categories of models. *Pesquisa Agropecuária Brasileira*, *48*(3), 287–295. https://doi.org/10.1590/S0100-204X2013000300007

Cieszewski, C. J. (2002). Comparing fixed-and variable-base-age site equations having single versus multiple asymptotes. *Forest Science*, *48*(1), 7–23. https://doi.org/10.1093/forestscience/48.1.7

Clutter, J. L. (1963). Compatible growth and yield models for loblolly pine. *Forest Science*, *9*(3), 354–371.

Clutter, J. L., Fortson, J. C., Pienaar, L. v., Brister, G. H., & Bailey, R. L. (1983). *Timber management: A quantitative approach*. John Wiley & Sons Inc.

da Rocha, J. E. C., Junior, M. R. N., da Silva Tavares Júnior, I. S., de Souza, J. R. M., de Sousa Lopes, L. S., da Silva, M. L. (2021). Configuration of artificial neural networks for height-diameter relationship of *Eucalyptus* spp. *Scientia Forestalis*, *49*(132), e3706. https://doi.org/10.18671/scifor.v49n132.08

da Silva Binoti, M. L. M., Binoti, D. H. B., & Leite, H. G. (2013). Application of artificial neural networks to estimate the height of equine eucalyptus stands. *Revista Árvore*, *37*(4), 639–645. https://doi.org/10.1590/S0100-67622013000400007

da Silva Binoti, M. L. M., Binoti, D. H. B., Leite, H. G., Garcia, S. L. R., Ferreira, M. Z., Rode, R., & da Silva, A. A. L. (2014). Neural networks for estimating of the volume of trees. *Revista Árvore*, *38*(2), 283–288. https://doi.org/10.1590/S0100-67622014000200008

da Silva Binoti, M. L. M., Leite, H. G., Binoti, D. H. B., & Gleriani, J. M. (2015). Stand-level prognosis of eucalyptus clones using artificial neural networks. *Cerne*, *21*(1), 97-105. https://doi.org/10.1590/01047760201521011153

da Silva Tavares Júnior, I., da Rocha, J. E. C., Ebling, Â. A., de Souza Chaves, A., Zanuncio, J. C., Farias, A. A., & Leite, H. G. (2019). Artificial neural networks and linear regression reduce sample intensity to predict the commercial volume of *Eucalyptus* clones. *Forests*, *10*(3), 268. https://doi.org/10.3390/f10030268

da Silva, A. L., Campos, J. C. C., Leite, H. G., de Souza, A. L., & Lopes, P. F. (2006). Growth and yield prediction using the modified Buckman model. *Revista Árvore*, *30*(5), 787–793. https://doi.org/10.1590/S0100-67622006000500012

da Silva, M. L. M., Binoti, D. H. B., Gleriani, J. M., & Leite, H. G. (2009). Adjustment of the Schumacher and Hall model and application of artificial neural networks to estimate volume of eucalyptus trees. *Revista Árvore, 33*(6), 1133–1139. https://doi.org/10.1590/S0100-67622009000600015

Davis, L. S., & Johnson, K. N. (1987). *Forest management* (3rd ed.). McGrawHill Book Company.

de Alcântra, A. E. M., de Albuquerque Santos, A. C., da Silva, M. L. M., Binoti, D. H. B., Soares, C. P. B., Gleriani, J. M., & Leite, H. G. (2018). Use of artificial neural networks to assess yield projection and average production of eucalyptus stands. *African Journal of Agricultural Research*, *13*(42), 2285–2297. https://doi.org/10.5897/ajar2017.12942

de Freitas, E. C. S., de Paiva, H. N., Neves, J. C. L., Marcatti, G. E., & Leite, H. G. (2020). Modeling of eucalyptus productivity with artificial neural networks. *Industrial Crops and Products*, *146*, 112149. https://doi.org/10.1016/j.indcrop.2020.112149

de Oliveira Neto, R. R., Leite, H. G., Gleriani, J. M., & Strimbu, B. M. (2022). Estimation of eucalyptus productivity using efficient artificial neural network. *European Journal of Forest Research*, *141*, 129–151. https://doi.org/10.1007/s10342-021-01431-7

de Oliveira Neto, R. R., Rodrigues, L. F., Mari, J. F., Naldi, M. C., Milagres, E. G., Vital, B. R., de Cássia Oliveira Carneiro, A., Binoti, D. H. B., Lopes, P. F., & Leite, H. G. (2021). Automatic identification of charcoal origin based on deep learning. *Maderas. Ciencia y tecnología*, *23*(65), 1-12. https://doi.org/10.4067/S0718-221X2021000100465

Dorado, F. C. (2004). *Modelo dinámico de crecimiento para las masas de "*Pinus radiata *D. Don" en Galicia* [Dynamic growth model for stands of "*Pinus radiata* D. Don" in Galicia] [Doctoral dissertation, Universidade de Santiago de Compostela]. Dialnet. https://dialnet.unirioja.es/servlet/tesis?codigo=106294

dos Reis Martins, E., da Silva Binoti, M. L. M., Leite, H. G., Binoti, D. H. B., & Dutra, G. C. (2016). Configuração de redes neurais artificiais para estimação do afilamento do fuste de árvores de eucalipto [Configuration of artificial neural networks to estimate the thinning of the shaft of *Eucalyptus* trees]. *Revista Brasileirade Ciencias Agrarias*, *11*(1), 33–38. https://doi.org/10.5039/agraria.v11i1a5354

Ferreira, M. P., de Almeida, D. R. A., de Almeida Papa, D., Minervino, J. B. S., Veras, H. F. P., Formighieri, A., Santos, C. A. N., Ferreira, M. A. D., Figueiredo, E. O., & Ferreira, E. J. L. (2020). Individual tree detection and species classification of Amazonian palms using UAV images and deep learning. *Forest Ecology and Management*, *475*, 118397. https://doi.org/10.1016/j.foreco.2020.118397

Fryda, T., LeDell, E., Gill, N., Aiello, S., Fu, A., Candel, A., Click, C., Kraljevic, T., Nykodym, T., Aboyoun, P., Kurka, M., Malohlava, M., Poirier, S., Wong, W., Rehak, L., Eckstrand, E., Hill, B., Vidrio, S., Jadhawani, S., … Novotny, M. (2020). *H2O: R interface for the 'H2O' scalable machine learning platform*. https://CRAN.R-project.org/package=h2o

Gavilán-Acuña, G., Olmedo, G. F., Mena-Quijada, P., Guevara, M., Barría-Knopf, B., & Watt, M. S. (2021). Reducing the uncertainty of radiata pine site index maps using an spatial ensemble of machine learning models. *Forests*, *12*(1), 77. https://doi.org/10.3390/f12010077

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c. *Philosophical Transactions of the Royal Society of London*, *115*, 513-583. https://doi.org/10.1098/rstl.1825.0026

Guera, O. G. M., da Silva, J. A. A., Ferreira, R. L. C., Lazo, D. A. Á., & Medel, H. B. (2019). Alternatives to growth and yield prognosis for *Pinus caribaea* var. *caribaea* Barrett & Golfari. *Floresta e Ambiente*, *26*(4), e20170381. https://doi.org/10.1590/2179-8087.038117

Hamdi, Z. M., Brandmeier, M., & Straub, C. (2019). Forest damage assessment using deep learning on high resolution remote sensing data. *Remote Sensing*, *11*(17), 1976. https://doi.org/10.3390/rs11171976

Hirigoyen, A., Franco, J., & Diéguez, U. (2018). Dynamic stand model for *Eucalyptus globulus* (L.) in Uruguay. *Agrociencia Uruguay*, *22*(1), 63–80. https://doi.org/10.31285/AGRO.22.1.7

Kentsch, S., Caceres, M. L. L., Serrano, D., Roure, F., & Diez, Y. (2020). Computer vision and deep learning techniques for the analysis of drone-acquired forest images, a transfer learning study. *Remote Sensing*, *12*(8), 1287.https://doi.org/10.3390/rs12081287

Kleiber, C., & Zeileis, A. (2008). *Applied econometrics with R*. Springer. https://doi.org/10.1007/978-0-387-77318-6

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*, 436–444. https://doi.org/10.1038/nature14539

Lhotka, J. M. (2017). Examining growth relationships in *Quercus* stands: An application of individual-tree models developed from long-term thinning experiments. *Forest Ecology and Management*, *385*, 65–77. https://doi.org/10.1016/j.foreco.2016.11.029

Lopes, L. S. S., Rode, R., Pauletto, D., Baloneque, D. D., dos Santos, F. G., Silva, A. R., Binoti, D. H. B., & Leite, H. G. (2020). Uso de regressão e redes neurais artificiais na estimativa do volume de *Khaya ivorensis* [Use of regression and artificial neural networks in the estimation of *Khaya ivorensis* volume]. *Ciência da Madeira*, *11*(2), 74-84. https://doi.org/10.12953/2177-6830/rcm.v11n2p74-84

Miguel, E. P., de Oliveira, C. S., Marthins, T. O., Matias, R. A. M., Rezende, A. V., Angelo, H., & Martins, I. S. (2016). Growth and yield models by total stand (MPT) in *Eucalyptus urophylla* (s.t. Blake) plantations. *Australian Journal of Basic and Applied Sciences*, *10*(13), 79–85.

Mongus, D., Vilhar, U., Skudnik, M., Žalik, B., & Jesenko, D. (2018). Predictive analytics of tree growth based on complex networks of tree competition. *Forest Ecology and Management*, *425*, 164–176. https://doi.org/10.1016/j.foreco.2018.05.039

Nielsen, M. A. (2015). *Neural networks and deep learning*. Determination Press.

Penido, T. M. A., Lafetá, B. O., Nogueira, G. S., Alves, P. H., Gorgens, E. B., & de Oliveira, M. L. R. (2020). Growth and production models for volumetric estimates in commercial eucalypt stands. *Scientia Forestalis*, *48*(128), e3340. https://doi.org/10.18671/scifor.v48n128.06

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org

Reis, T., Silva, S., Junior, I. T., Leite, R., Taquetti, M., Atanazio, K., Fardin, L., Cruz, J., & Leite, H. (2022). Assessing alternatives for growth and yield modelling in eucalypt stands. *Southern Forests*, *84*(1), 34–43. https://doi.org/10.2989/20702620.2021.2011468

Salles, T. T., Leite, H. G., de Oliveira Neto, S. N., Soares, C. P. B., de Paiva, H. N., & dos Santos, F. L. (2012). Clutter model in modeling growth and yield of eucalyptus in crop-livestock-forest integration systems. *Pesquisa Agropecuária Brasileira*, *47*(2), 253–260. https://doi.org/10.1590/S0100-204X2012000200014

Sejnowski, T. J. (2020). The unreasonable effectiveness of deep learning in artificial intelligence. *Proceedings of the National Academy of Sciences*, *117*(48), 30033–30038. https://doi.org/10.1073/pnas.1907373117

Silva, S., de Oliveira Neto, S. N., Leite, H. G., de Alcântara, A. E. M., de Oliveira Neto, R. R., & de Souza, G. S. A. (2020). Productivity estimate using regression and artificial neural networks in small familiar areas with agrosilvopastoral systems. *Agroforestry Systems*, *94*, 2081–2097. https://doi.org/10.1007/s10457-020-00526-1

Silva, W. S., Fernandes, F. A., Muniz, F. R., Muniz, J. A., & Fernandes, T. J. (2021). *Eucalyptus grandis* x *Eucalyptus urophylla* growth curve in different site classifications, considering residual autocorrelation. *Revista Brasileira de Biometria*, *39*(1), 122–138. https://doi.org/10.28951/rbb.v39i1.511

Soares, C. P. B., Leite, H. G., de Oliveira, M. L. R., & Carvalho, A. (2004). Specifying a model of forest growth and yield. *Revista Árvore*, *28*(6), 831–837. https://doi.org/10.1590/S0100-67622004000600007

Socha, J., & Tymińska-Czabańska, L. (2019). A method for the development of dynamic site index models using height-age data from temporal sample plots. *Forests*, *10*(7), 542. https://doi.org/10.3390/f10070542

Strub, M., & Cieszewski, C. J. (2006). Base–age invariance properties of two techniques for estimating the parameters of site index models. *Forest Science*, *52*(2), 182-186.

Sun, Y., Liu, Y., Wang, G., & Zhang, H. (2017). Deep learning for plant identification in natural environment. *Computational Intelligence and Neuroscience*, *2017*, 7361042. https://doi.org/10.1155/2017/7361042

Trevizol Jr., T. L. (1985). *Análise de um modelo compatível de crescimento e produção em plantações de* Eucalyptus grandis *(W. Hill ex Maiden)* [Analysis of a compatible model of growth and production in plantations of *Eucalyptus grandis* (W. Hill ex Maiden)] [Unpublished Master's dissertation]. Universidade Federal de Viçosa.

Vendruscolo, D. G. S., Chaves, A. G. S., Medeiros, R. A., da Silva, R. S., Souza, H. S., Drescher, R., & Leite, H. G. (2017). Height estimation of *Tectona grandis* L.f. trees using regression and artificial neural networks. *Nativa: Pesquisas Agrárias e Ambientais*, *5*(1), 52-58. https://doi.org/10.5935/2318-7670.v05n01a09

Vieira, G. C., de Mendonça, A. R., da Silva, G. F., Zanetti, S. S., da Silva, M. M., & dos Santos, A. R. (2018). Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. *Science of The Total Environment*, *619–620*, 1473–1481. https://doi.org/10.1016/j.scitotenv.2017.11.138

Weiskittel, A., Kuehne, C., McTague, J. P., & Oppenheimer, M. (2016). Development and evaluation of an individual tree growth and yield model for the mixed species forest of the Adirondacks Region of New York, USA. *Forest Ecosystems*, *3*, 26. https://doi.org/10.1186/s40663-016-0086-3

Xi, Z., Hopkinson, C., Rood, S. B., & Peddle, D. R. (2020). See the forest and the trees: Effective machine and deep learning algorithms for wood filtering and tree species classification from terrestrial laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *168*, 1–16. https://doi.org/10.1016/j.isprsjprs.2020.08.001

Zeiler, M. D. (2012). ADADELTA: An adaptive learning rate method. *arXiv*. https://doi.org/10.48550/arXiv.1212.5701

SUPPLEMENTARY

Table S1

Status of neuron layers of the best models trained with the deep learning method for each Data Structure II-A and II-B in hybrid eucalypt stands

Data Structure II-A								
Status of neuron layers: projection V_2, regression, Gaussian distribution, quadratic loss, 96,751 weights/biases, 183,078 training samples, mini-batch size 1								
Layer	Units	Type	mean_rate	rate_rms	mean_weight	weight_rms	mean_bias	bias_rms
2	200	Maxout	0.1394	0.3614	0.0002	0.0803	0.4891	0.0516
3	100	Maxout	0.1435	0.3666	-0.0018	0.0826	0.9977	0.0056
4	50	Maxout	0.3001	0.4548	-0.0015	0.1142	0.9996	0.0019
5	1	Linear	0	0	-0.0486	0.1955	0.0022	0
Data Structure II-B								
Status of neuron layers: projection V_2, regression, Gaussian distribution, quadratic loss, 16,701 weights/biases, 3,104,698 training samples, mini-batch size 1								
Layer	Units	Type	mean_rate	rate_rms	mean_weight	weight_rms	mean_bias	bias_rms
2	100	Rectifier	0.1125	0.1714	-0.0114	0.3045	-0.4842	0.6457
3	50	Rectifier	0.1088	0.1225	-0.1333	0.4122	0.5021	0.5974
4	1	Linear	0.0049	0.0058	0.0006	0.2037	0.3875	0