Pertanika J. Sci. & Technol. 30 (2): 1257 - 1272 (2022)

SCIENCE & TECHNOLOGY

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

Improving Yield Projections from Early Ages in Eucalypt Plantations with the Clutter Model and Artificial Neural Networks

Gianmarco Goycochea Casas[1]*, Leonardo Pereira Fardin[1], Simone Silva[1], Ricardo Rodrigues de Oliveira Neto[1], Daniel Henrique Breda Binoti[2], Rodrigo Vieira Leite[1], Carlos Alberto Ramos Domiciano[1], Lucas Sérgio de Sousa Lopes[1], Jovane Pereira da Cruz[3], Thaynara Lopes dos Reis[1] and Hélio Garcia Leite[1]

[1]*Department of Forest Engineering, Federal University of Viçosa. Av. Purdue, s/n, University Campus, 36570-900 Viçosa, MG, Brazil*
[2]*Dap Florestal. Rua Papa João XXIII, n 9, Lourdes - 36572-006 - Viçosa, MG, Brazil*
[3]*Bracell Florestal. Rua Treze de Junho, Centro – 48030-715 – Alagoinhas, BA, Brazil*

ABSTRACT

A common issue in forest management is related to yield projection for stands at young ages. This study aimed to evaluate the Clutter model and artificial neural networks for projecting eucalypt stands production from early ages, using different data arrangements. In order to do this, the changes in the number of measurement intervals used as input in the Clutter model and artificial neural networks (ANNs) are tested. The Clutter model was fitted considering two sets of data: usual, with inventory measurements (I) paired at intervals each year (I_1–I_2, I_2–I_3, …, I_n–I_{n+1}); and modified, with measurements paired at all possible age intervals (I_1–I_2, I_1–I_3, …, I_2–I_3, I_2–I_4, …, I_n–I_{n+1}). The ANN was trained with the modified dataset plus soil type and geographic coordinates as input variables. The yield projections were made up to the final ages of 6 and 7 years from all possible initial ages (2, 3, 4, 5, or 6 years). The methods are evaluated using the relative error (RE%), bias, correlation coefficient ($r_{y\hat{y}}$), and relative root mean square error (RMSE%). The ANN was accurate in all

ARTICLE INFO

Article history:
Received: 2 September 2021
Accepted: 15 December 2021
Published: 11 March 2022

DOI: https://doi.org/10.47836/pjst.30.2.22

E-mail addresses:
gianmarco.casas@ufv.br (Gianmarco Goycochea Casas)
leopfard@gmail.com (Leonardo Pereira Fardin)
simone.silva.ufv@gmail.com (Simone Silva)
rick.neto@gmail.com (Ricardo Rodrigues de Oliveira Neto)
danielhbbinoti@gmail.com (Daniel Henrique Breda Binoti)
rvleite.efl@gmail.com (Rodrigo Vieira Leite)
carlosramosdomiciano@gmail.com (Carlos Alberto Ramos Domiciano)
lucas.s.sergio@ufv.br (Lucas Sérgio de Sousa Lopes)
jovanecruz@gmail.com (Jovane Pereira da Cruz)
thaynaralopes29@gmail.com (Thaynara Lopes dos Reis)
hgleite@ufv.br (Hélio Garcia Leite)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

cases, with RMSE% from 8.07 to 14.29%, while the Clutter model with the modified dataset had values from 7.95 to 23.61%. Furthermore, with ANN, the errors were evenly distributed over the initial projection ages. This study found that ANN had the best performance for stand volume projection surpassing the Clutter model regardless of the initial or final age of projection.

Keywords: Artificial intelligence, data structure, forest growth and yield, forest management, regression

INTRODUCTION

Eucalypt is one of the most planted forest species in the world. There are over 9.6 million hectares of even-aged stands only in Brazil supplying pulp and paper, wood panels, charcoal, and steel industries (IBGE, 2021). Managing these stands demands knowledge of the elements involved in the system, such as the forest growth and yield that can be obtained by fitting predictive models. These models can provide quantitative information on the dynamics and development of commercial forest stands (Scolforo et al., 2019a). They can be classified into three types: individual tree, diameter distribution, and whole-stand models (Castro et al., 2016; Sharma et al., 2019). The latter is the most common type used by forest managers due to its simplicity and effectiveness (de Azevedo et al., 2016; Campos & Leite, 2017).

The model proposed by Clutter (1963) features prominently among the whole-stand models. It is a system of two interdependent models fitted with a two-stage least square regression analysis (Gujarati & Porter, 2011) characterized by density variables (de Abreu Demolinari et al., 2007; Burkhart & Tomé, 2012) which allow, for example, thinning simulations by basal area. Using the Clutter model is widespread in Brazilian forestry companies, mainly because of its characteristics of compatibility and consistency that make feasible production estimates by year or irregular intervals (Vescovi et al., 2020; Penido et al., 2020).

The Clutter model is traditionally fitted using paired data from two consecutive measurements, which for eucalypt stands occurs annually (Campos & Leite, 2017). That might yield greater errors when projections are done for intervals longer than a year (Salles et al., 2012). This study hypothesized that using more measurement intervals to fit the Clutter model could generate more consistent projections considering all age variations. The use of pairs of consecutive measurements to model eucalypt stands volume is an adequate alternative to represent the growth of even-aged stands over time, as it allows greater coverage of the states of each plot (Stankova, 2016; de Alcântra et al., 2018).

Another issue of using the Clutter model is that regression analysis requires statistical assumptions, such as statistical independence, homoscedasticity, normal distribution of errors, and no multicollinearity and autocorrelation (Gujarati & Porter, 2011; Bayat et at., 2019). Some databases might violate one or all these assumptions, particularly when dealing with biological assets. Thus, it is interesting to implement non-parametric methods, such as machine learning tools, which emerge as an alternative for bypassing such limitations (Mongus et al., 2018; Vieira et al., 2018). Several machine learning algorithms have been applied to problems in the forest area with relatively good results, especially using random forest algorithms (Pereira et al., 2021; de Oliveira et al., 2021), neuro-fuzzy (Silva et al., 2021), support vector machines (Nieto et al., 2012; Liu et al., 2020). However, the most used algorithm has been artificial neural networks.

Artificial neural networks (ANN) algorithm has been used in forestry to solve diverse problems such as estimating eucalypt tree heights (Campos et al., 2017), volume (da Silva Binoti et al., 2014; da Silva Tavares Júnior et al., 2019), stem taper (da Cunha Neto et al., 2019), survival (da Rocha et al., 2018), and yield (da Silva Binoti et al., 2015). This method often replaces regression models with more accurate estimates without increasing the number of samples used (da Rocha et al., 2018; da Silva Tavares Júnior et al., 2019). ANN are computational models inspired by the nervous system of living beings. They have the ability to acquire and maintain knowledge based on the information and can be defined as a set of processing units, which are interconnected by a large number of interconnections (Silva et al., 2016).

ANN is superior to regression methods for its ability to learn, generalize, and model categorical (i.e., qualitative) and continuous (i.e., quantitative) variables. This method can be used as an alternative for solving complex problems when data do not meet the assumptions for regression analysis.

This study evaluates using the Clutter model artificial neural networks for projecting eucalypt stands production from early ages, using different data interval structures by analyzing i) whether the accuracy of the Clutter model is independent of the measurement intervals used as input and the range of projection; and ii) whether ANN is more efficient at projecting forest growth and yield.

MATERIAL AND METHODS

Study Area and Database

This study used data from a Continuous Forest Inventory (CFI) of three eucalypt stands (*Eucalyptus urophylla* x *Eucalyptus grandis*) planted in 3 x 3 m tree spacings located in the northeastern region of the State of Bahia (BA), Brazil. The sampling consisted of 375 permanent plots with an average area of 400 m² measured in different years with ages (I) ranging from 2 to 8 years. For each plot, the variable diameter at breast height (Dbh), basal area (B), outside bark volume (V), total height (HT), and dominant height (Hd) were obtained. This study shows their statistical description for each project in Figure 1.

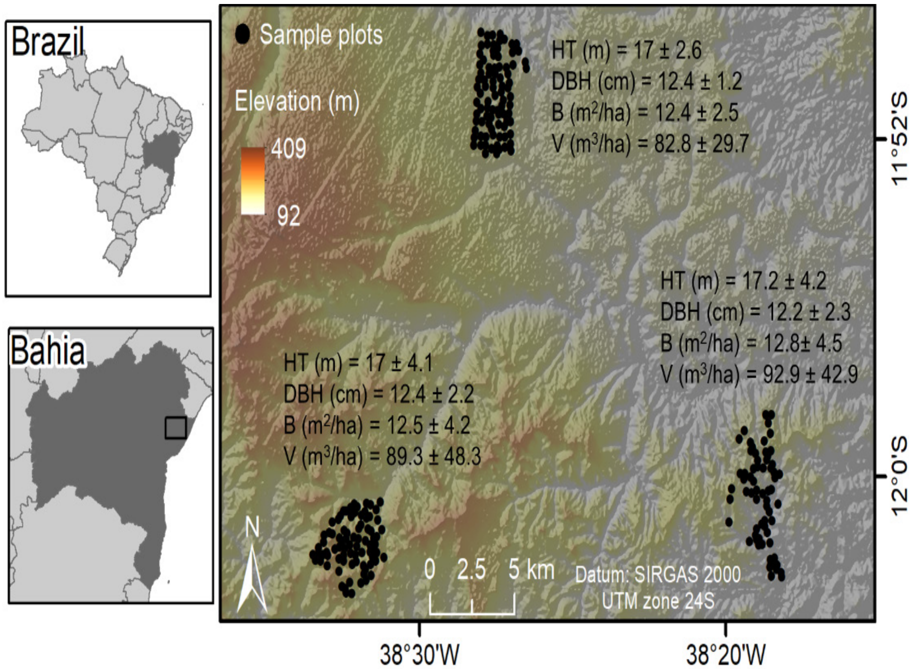

Figure 1. Location of the study area, statistical description, and distribution of the sample plots of the continuous inventory used for the adjustment of the models and the performance of the volume of the project stand

The stands are in the municipalities of Alagoinhas, Aramari and Inhambupe (BA, Brazil) in the Mata Atlântica and Caatinga biomes. The soils in the study area belong to the large PV group (Argisols) (IBGE, 2018). The predominant soil type is Yellow Argisol, characterized by the accumulation of clay on the B horizon. The other soils are oxisol and entisols (quartzipsamments) (dos Santos et al., 2018). The climate of the region is tropical rainforest (Af), according to the Köppen-Geiger classification. The annual rainfall is over 1,500 mm, and in the coldest month, the average minimum and maximum temperatures are 18 and 22°C, respectively.

Site Index Curves

This study classifies the productivity capacity (S) of each project (i.e., group of plots) using dominant height and a base age of 6 years. The Gompertz (1825) equation was applied to projects 1, 2, and 3, respectively (Equations 1-3). In order to determine the productive capacity (S) of each project, the guide curve method (Clutter, 1963) was used, establishing site index equations (Equations 4-6). The site index curves are shown in Figure 2.

$$Hd = 29.5092e^{-e^{0.3735-0.0258I}}; \quad R = 0.7498 \tag{1}$$

$$Hd = 30.3402e^{-e^{0.4299-0.0269I}}; \quad R = 0.\ 7977 \tag{2}$$

$$Hd = 20.7956e^{-e^{-0.1745-0.0317I}}; \quad R = 0.6093 \tag{3}$$

The site index (S) of each stand were estimated by:

$$S = Hd e^{-e^{0.3735-0.0258(72)}} e^{e^{0.3735-0.0258I}} \tag{4}$$

$$S = Hd e^{-e^{0.4299-0.0269(72)}} e^{e^{0.4299-0.0269I}} \tag{5}$$

$$S = Hd e^{-e^{0.1745-0.0317(72)}} e^{e^{0.1745-0.0317I}} \tag{6}$$

where: Hd = dominant height (meters); I = stand age (months); S = productivity capacity.

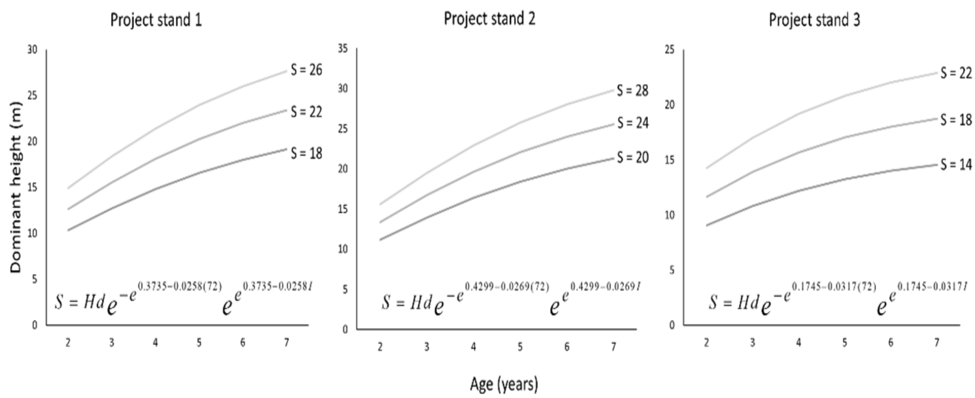

Figure 2. Site index curves for *Eucalyptus urophylla X Eucalytus grandis* clone stands located in the northeastern region of Bahia (BA), Brazil, for an index age of 72 months

Growth and Yield Modeling

The Clutter model is a system of equations (Equations 7 & 8). This study used the two-stage ordinary least squares method to fit the model in its reduced form (without considering the interactions between the variables).

$$LnB_2 = LnB_1\left(\frac{I_1}{I_2}\right) + \alpha_0\left(1-\frac{I_1}{I_2}\right) + \alpha_1\left(1-\frac{I_1}{I_2}\right)S_1 + \varepsilon \tag{7}$$

$$LnV_2 = \beta_0 + \beta_1\left(\frac{1}{I_2}\right) + \beta_2 S + \beta_3 LnB_2 + \varepsilon \tag{8}$$

where: V_2= future volume (m^3 ha^{-1}); I_1 and I_2= current and future age (years); S= site index (m); B_1 and B_2= current and future basal area (m^2 ha^{-1}); L_n= Naperian logarithm, β_n = model parameters; ε = random error.

This study set the database to fit the Clutter model in two ways, denoted as usual and modified datasets. The first is the most used structure, with field measurements paired only in consecutive ages (I_1–I_2, I_2–I_3, …, I_n–I_{n+1}). The second uses all possible combinations of age pairs (I_1–I_2, I_1–I_3, …, I_2–I_3, I_2–I_4, …, I_n–I_{n+1}) (Table 1).

Table 1

Demonstration of datasets used to fit the Clutter model for a plot

Database structure	Survey 1	Survey 2	I_1	I_2	B_1	B_2	V_1	V_2
Usual	1	2	2	3	3.1445	8.4812	9.5583	47.2634
	2	3	3	4	8.4812	11.6959	47.2634	78.0712
	3	4	4	5	11.6959	14.1544	78.0712	121.4490
	4	5	5	6	14.1544	14.8039	121.449	140.0660
Modified	1	2	2	3	3.1445	8.4812	9.55831	47.2634
	1	3	2	4	3.1445	11.6959	9.55831	78.0712
	1	4	2	5	3.1445	14.1544	9.55831	121.4490
	1	5	2	6	3.1445	14.8039	9.55831	140.0660
	2	3	3	4	8.4812	11.6959	47.2634	78.0712
	2	4	3	5	8.4812	14.1544	78.0712	121.4490
	2	5	3	6	8.4812	14.8039	121.449	140.0660
	3	4	4	5	11.6959	14.1544	78.0712	121.4490
	3	5	4	6	11.6959	14.8039	78.0712	140.0660
	4	5	5	6	14.1544	14.8039	121.4490	140.0660

Note. I_1 and I_2 = current and future age (years); B_1 and B_2 = current and future basal area; V_1 and V_2 = current and future volume (m^3 ha^{-1})

The modified data set was also used to train the ANN. This study used Multilayer Perceptron neural networks (MLP) with three layers (input, hidden, and output). The variables in the input layer are chosen based on the recommendations of Martins et al. (2015) and Freitas et al. (2020). The variables were current age (I1), future age (I2), current basal area (B1), Dbh, current volume (V1), Hd, type of soil, and geographic area coordinates. The output layer was the future volume (V2).

The ANNs were trained and validated in the Neuro software (version 4.0) (Binoti, 2012). The maximum-minimum normalization was applied to the data. The Resilient

Propagation (RPROP +) training algorithm uses the sigmoid activation function. The stopping criteria achieved a mean error of 0.0001 or 3000 cycles. The number of neurons in the hidden layer was 8, calculated based on variables (Campos & Leite, 2017) (Equation 9).

$$N_{hidden} = \left(\frac{\sum_{i=1}^{n} V_{continuous} + \sum_{i=1}^{n} V_{categorical}}{2} \right) \quad (9)$$

where N_{hidden} is number of neurons in the hidden layer, $V_{continuous}$ is the number of continuous input variables, and $V_{categorical}$ is the number of categorical input variables.

The projections were made for 6 and 7 years, as they are the usual ages for cutting eucalypt plantations in Brazil. The initial ages of the projection were all the possible previous ages before the final age, from two years of age onwards. For both methods (i.e., Clutter and ANN), the data was split into 70% for model fitting or training and 30% for validation. The model bias (Equation 10) is used to evaluate the methods, relative root means square error (RSME%) (Equation 11), the correlation coefficient between estimated and observed volume ($r_{y\hat{y}}$) (Equation 12), relative error ($RE_{\%}$) (Equation 13), and mean absolute deviations (MAD) (Equation 14). The error frequency histograms and the relative error charts by age and volume are also analyzed. The Bartlett test was performed with a significance level of 1% to observe the existence of heterogeneity.

$$Bias = \sum_{i=1}^{n} \frac{\left(\hat{Y}_i - Y_i\right)}{n} \quad (10)$$

$$RMSE\% = \frac{\left(\sqrt{\sum_{i=1}^{n} \frac{\hat{Y}_i - Y_i}{n}}\right)}{\left(\sum_{i=1}^{n} \frac{Y_i}{n}\right)} x100 \quad (11)$$

$$r_{y\hat{y}} = \frac{\left[n^{-1} \sum_{i=1}^{n} (\hat{Y}_i - Y_m)(Y_i - \bar{Y}) \right]}{\left[\sqrt{n^{-1} \sum_{i=1}^{n} (\hat{Y}_i - \hat{Y}_m)^2 n^{-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2} \right]} \quad (12)$$

$$RE_{\%} = \left(\frac{\hat{Y}_i - Y_i}{Y_i} \right) x\ 100 \quad (13)$$

$$MAD = \left(n^{-1} \sum_{i=1}^{n} \left| Y_i - \hat{Y}_i \right| \right) \quad (14)$$

where: n = number of observations, = estimated values, $\hat{Y}_i$ = observed values, $\bar{Y}$ = observed mean values, Y_m = arithmetic means of estimated values.

RESULTS AND DISCUSSION

The parameters of the Clutter model were significant ($p < 0.01$) for the usual (Equations 15 & 16) and modified (Equations 17 & 18) models. When adjusting this model, attention should be paid to the signs of the coefficients, which must be negative and positive. These conditions were also met for both datasets.

$$LnB_2 = LnB_1\left(\frac{I_1}{I_2}\right) + 2.5144\left(1-\frac{I_1}{I_2}\right) + 0.0254\left(1-\frac{I_1}{I_2}\right)S_1 \quad (15)$$

$$LnV_2 = 0.6833 - 10.6582\left(\frac{1}{I_2}\right) + 0.0164S_1 + 1.4261LnB_2 \quad (16)$$

$$LnB_2 = LnB_1\left(\frac{I_1}{I_2}\right) + 2.3239\left(1-\frac{I_1}{I_2}\right) + 0.0331\left(1-\frac{I_1}{I_2}\right)S_1 \quad (17)$$

$$LnV_2 = 1.0793 - 17.1921\left(\frac{1}{I_2}\right) + 0.0295S_1 + 1.2070LnB_2 \quad (18)$$

In addition, heterogeneity ($p < 0.01$) in the variances for basal area per hectare and volume per hectare over the ages of each project tested are observed. Only project 1 did not show heterogeneity ($p > 0.01$) of the variances for basal area per hectare. It was observed using the Barlett statistical test at a significance level of 1%. The performance of the Clutter and Clutter-modified models were similar, with bias ranging from -11.3 to 17.33 $m^3\ ha^{-1}$, RMSE% from 8.93 to 23.61%, MAD from 5.31 to 29.85 $m^3\ ha^{-1}$, and $r_{y\hat{y}}$ from 0.64 to 0.97 (Table 2).

Table 2

Volumetric projection statistics for the proposed methodologies using 6 and 7 years as final projection ages

Final age	Initial age	Clutter				Clutter–modified			
		bias	RMSE%	MAD	$r_{y\hat{y}}$	bias	RMSE%	MAD	$r_{y\hat{y}}$
6	2	12.27	23.05	26.48	0.72	10.99	23.61	27.03	0.70
	3	-1.56	15.42	14.94	0.92	0.21	14.90	14.20	0.91
	4	0.50	10.95	11.47	0.95	0.32	11.28	11.19	0.94
	5	1.58	9.49	9.08	0.97	1.89	9.98	9.26	0.97
7	2	17.33	21.88	29.24	0.65	12.77	21.21	29.85	0.64
	3	-11.30	18.29	15.66	0.85	-5.04	14.51	11.60	0.87
	4	2.33	14.03	12.68	0.90	2.06	13.56	11.83	0.89
	5	-2.46	11.55	9.93	0.97	-2.06	11.99	9.55	0.95
	6	-2.03	8.93	6.66	0.97	-0.58	7.95	5.31	0.97

Table 2 *(Continue)*

Final age	Initial age	ANN			
		bias	RMSE%	MAD	$r_{y\hat{y}}$
6	2	-3.40	14.29	15.27	0.90
	3	2.82	13.08	11.81	0.94
	4	2.44	11.18	10.55	0.95
	5	0.001	9.72	8.13	0.97
7	2	4.70	13.26	15.86	0.91
	3	-0.68	10.08	7.52	0.92
	4	2.95	12.45	9.79	0.92
	5	-2.91	10.92	9.04	0.96
	6	-0.99	8.07	6.62	0.97

These models had a relatively good fit to the data (Figures 3g & 3h—Plot 1), and more pairs of measures did not significantly improve the model performance. In addition, for both Clutter models, errors had a wider range (-48 to 100%) when projections were made from earlier ages (Figures 3a & 3b—Plot 1) and tended to overestimate stands with lower yields (Figures 3d & 3e—Plot 1).

This study formulated the hypothesis that the number of sequential measures used as input to the models could overcome yield projection errors, especially from the early ages. Although they had similar statistics, the Clutter model fitted with the modified database was more accurate than the usual database, as the relative errors were closer to zero. However, the error becomes significantly greater when projected from younger ages. The two forms of adjustment proposed for the Clutter model were biased, overestimating production at younger ages and underestimating at older ages. The difficulty in estimating production from young ages is commonly observed when using the clutter model, as shown by de Abreu Demolinari et al. (2007) and Dias et al. (2005). The ANN also showed this tendency, but the projection errors were more uniform over the ages. Forest stands yield modeling helps forest managers plan forest activities, so the over or underestimation of stand's production can negatively affect the decision-making process (Salles et al., 2012). Therefore, there is a great concern about obtaining reliable and accurate estimates.

The projections with the ANN were accurate from all initial and final ages of projection. The bias ranged from -0.99 to 4.7, RMSE% from 8.07 to 14.29, MAD from 6.62 to 15.27, and $r_{y\hat{y}}$ from 0.91 to 0.97 (Table 2). The model's performance improves as close to the final age of projection as expected. This study observed that the difference was much smaller with the projections with ANN indicating its superiority (Figure 3—Plot 2). The ANN adjusted better to the data, and the error range was relatively narrower in the early ages

(Figure 3c—Plot 2). Although the study can observe a trend to underestimate volume in stands with a lower yield, the bias is less pronounced than with the Clutter models (Figure 3f —Plot 2).

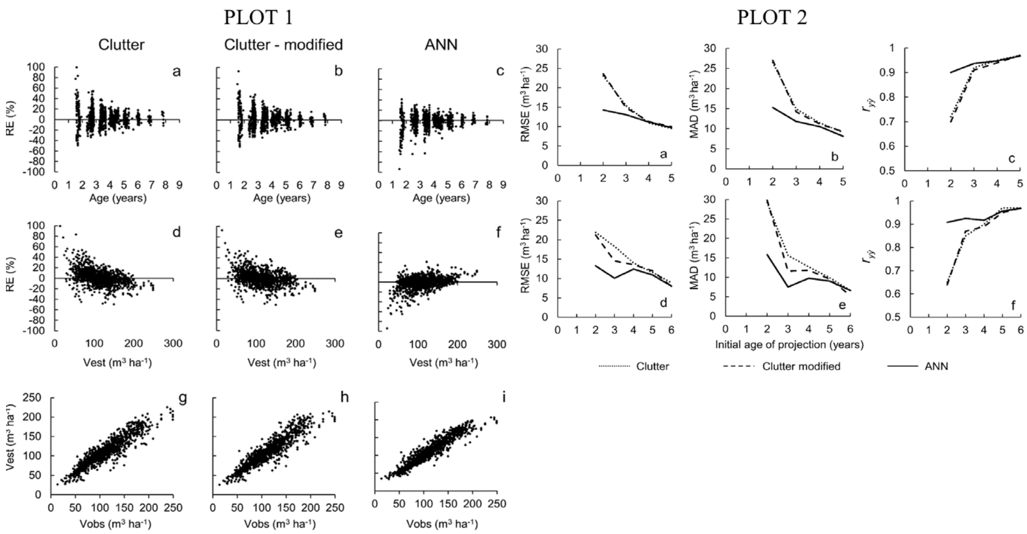

Figure 3. Plot 1 - Relative errors of stand yield estimates by age, stand volume (m^3 ha^{-1}), and the relationship between observed and estimated volume for Clutter (usual) (a, d, g), Clutter (modified) (b, e, h) and ANN (c, f, i). Plot 2—Statistics for stand yield projections with the methods Clutter, Clutter modified and Artificial Neural Networks (ANN) to 6 (upper-charts—a, b and c) and 7 years (lower-charts – d, e and f)

Growth and yield models must be reliable enough to describe the growth dynamics of a forest stand (Burkhart & Tomé, 2012), and the model's accuracy is a factor that can influence future yield estimates (Campos & Leite, 2017). Modifying the input dataset was proposed in this study to bypass projection errors, mainly from early ages. One of the drawbacks of using regression models is that data from forest plantations has characteristics that violate some statistical assumptions, such as homogeneity of variances (García, 1988; Gujarati & Porter, 2011). This study also observed this issue in modeling growth and yielded with the Clutter model. ANN for modeling does not require data to meet these principles (Braga et al., 2007) and is also more effective when dealing with dispersed data (Reis et al., 2018). Another ANN characteristic is tolerating data noise and easily modeling nonlinear problems (Binoti et al., 2013; Chiarello et al., 2019). For these reasons, the ANN may have resulted in better performances over the Clutter model in estimating the volume yield of eucalypt stands. Bayat et al. (2021), when estimating the increase in volume using MLP ANN—the same type of network used in this study—and multiple linear regression, reported the superiority of ANN, especially when working with heterogeneous data and complex nonlinear relationships.

The second hypothesis of this study was that projection range affects the accuracy of the estimates, which was also verified. The dispersion of errors is greater when projections are made from younger stands, which is aggravated when using regression models (Castro et al., 2016). The findings are confirmed by running an additional test projecting the volume with the modified database and different numbers of ANN configurations (Figure 4). The early age projections had RMSE% and MAD similar to the other ages regardless of the architecture and input variables tested on the ANN. Particularly, the RMSE% decreases slightly in the upper ages of projection (Figures 4a & 4c). However, the best performing ANN showed an RMSE% equal to 6% and MAD equal to 6 m³ ha-1, while Clutter-usual and Clutter-modified had 10% and 9 $m^3\ ha^{-1}$, respectively. In this study, categorical variables are used and found that such variables increased the accuracy of the estimates. ANN benefits from using categorical variables, which can be useful when field measurement data is scarce (de Alcântra et al., 2018). It is noteworthy that the projections in this study were for ages that represent the point of maximum forest production considering the ecological characteristics of the ecosystem within an economic context; that is, in Brazil, it is common to find eucalypt trees with growth curves showing the maximum annual average increment around 6 and 7 years (Rodriguez et al., 1997; Campos & Leite, 2017).

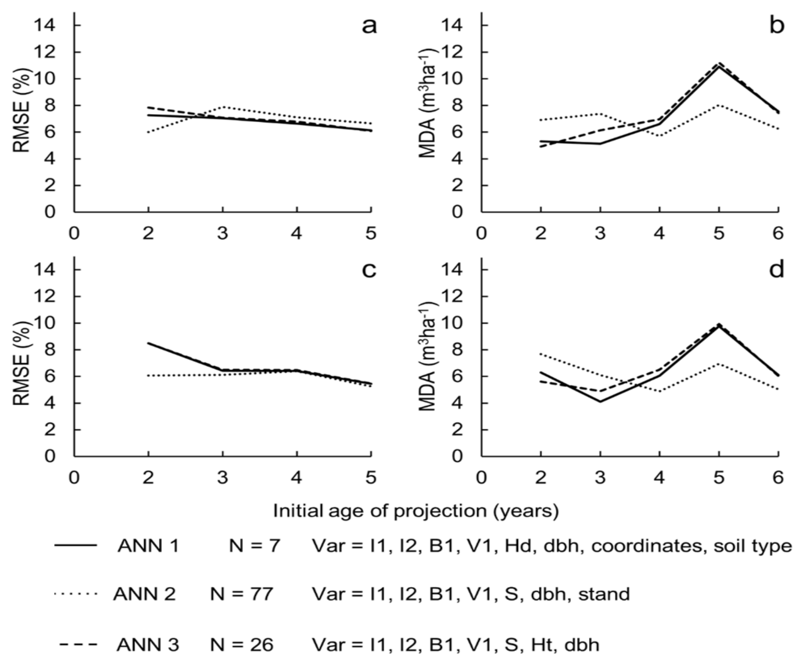

Figure 4. Statistics of yield projection for year 6 (upper plots) and 7 (lower plots), with initial ages of 2, 3, 4, 5 and 6 years (x-axis) for the three ANN that were evaluated (y-axis). A = root mean square error (RMSE); B = mean absolute deviations (MAD). N = number of neurons; Var = input variables in the ANN, namely: I1 = age at time 1; I2 = age at time 2; B1 = basal area at time 1; Hd = dominant height; dbh = diameter at 1.3 meters above ground; and S = Site index

This validation confirms the efficiency of an ANN for yield projection even from early ages, which represents a problem for the Clutter model. Regression models are subject to estimation errors that can significantly affect decision-making by the forest manager (Scolforo et al., 2019b). The ANN can therefore be an important tool for growth and yield modeling for its characteristics can overcome some of the limitations found in regression modeling.

CONCLUSION

Accurate yield projections from early ages are a common issue in eucalypt plantations management. This study found that using a larger number of measurement intervals as input variables in a growth and yield model can improve the projection estimates. The Clutter model with more measurement pairs (i.e., Clutter with the modified dataset) had lower errors than usual data inputs. Despite this, the accuracy became lower when projections were made from young age stands. The Clutter model limitations were solved using artificial neural networks (ANNs). This method was accurate in all cases with similar errors when projecting volume from early or later ages. Future work should investigate ANN structures and the number of observations in training models for forest planning and reduce fieldwork measurement costs.

ACKNOWLEDGEMENTS

The authors are grateful to “Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG” - Identifier Code 507, in collaboration with the “Fundação Arthur Bernardes – FUNARBE” for the financial support. This work was carried out with the support of the “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil—CAPES”—Financing Code 001.

REFERENCES

Bayat, M., Ghorbanpour, M., Zare, R., Jaafari, A., & Pham, B. T. (2019). Application of artificial neural networks for predicting tree survival and mortality in the Hyrcanian forest of Iran. *Computers and Electronics in Agriculture*, *164*, Article 104929. https://doi.org/10.1016/j.compag.2019.104929

Bayat, M., Bettinger, P., Hassani, M., & Heidari, S. (2021). Ten-year estimation of Oriental beech (*Fagus orientalis* Lipsky) volume increment in natural forests: A comparison of an artificial neural networks model, multiple linear regression and actual increment. *Forestry: An International Journal of Forest Research, 94*(4), 598-609. https://doi.org/10.1093/forestry/cpab001

Binoti, D. H. B. (2012). *Emprego de Redes Neurais Artificiais em Mensuração e Manejo Florestal* [Use of Artificial Neural Networks in Measurement and Forest Management] (PhD Thesis). Universidade Federal de Viçosa, Brazil.

Binoti, D. H. B., Binoti, M. L. M. S., Leite, H. G., & Silva, A. (2013). Redução dos custos em inventário de povoamentos equiâneos [Reduction in inventory costs in even-aged stand]. *Revista Brasileira de Ciências Agrárias - Brazilian Journal of Agricultural Sciences*, *8*(1), 125-129. https://doi.org/10.5039/agraria.v8i1a2209

Braga, A. P., Carvalho, A. P. L. F., & Ludemir, T. B. (2007). *Redes neurais artificiais: Teoria e aplicações* [Artificial neural networks: Theory and applications] (2nd Ed). LTC Editora.

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

Campos, B. P. F., da Silva, G. F., Binoti, D. H. B., de Mendonça, A. R., & Leite, H. G. (2017). Predição da altura total de árvores em plantios de diferentes espécies por meio de redes neurais artificiais [Estimation of total tree height in plantations of different species through artificial neural networks]. *Pesquisa Florestal Brasileira*, *36*(88), 375-385. https://doi.org/10.4336/2016.pfb.36.88.1166

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)]. UFV.

Castro, R. V. O., Araújo, R. A. A., Leite, H. G., Castro, A. F. N. M., Silva, A., Pereira, R. S., & Leal, F. A. (2016). Modelagem do crescimento e da produção de povoamentos de eucalyptus em nível de distribuição diamétrica utilizando índice de local [Modeling of growth and yield of eucalyptus stands in level of diameter distribution using site index]. *Revista Árvore*, *40*(1), 107-116. https://doi.org/10.1590/0100-67622016000100012

Chiarello, F., Steiner, M. T. A., Oliveira, E. B. D., Arce, J. E., & Ferreira, J. C. (2019). Artificial neural networks applied in forest biometrics and modeling: State of the art (january/2007 to july/2018). *CERNE*, *25*(2), 140-155. https://doi.org/10.1590/01047760201925022626

Clutter, J. L. (1963). Compatible growth and yield models for loblolly pine. *Forest Science*, *9*(3), 354-371. https://doi.org/10.1093/forestscience/9.3.354

da Cunha Neto, E. M., Bezerra, J. C. F., de Miranda, L. C., do Mar, A. L., Vaz, M. M., da Silva Melo, M. R., & da Castro Rocha, J. E. (2019). Kozak model and artificial neural networks in eucalyptus fuser sharing estimate. *Revista de Engenharia e Tecnologia*, *11*(3), 150-158.

da Rocha, S. J. S. S., Torres, C. M. M. E., Jacovine, L. A. G., Leite, H. G., Gelcer, E. M., Neves, K. M., Schettini, B. L. S., Villanova, P. H., da Silva, L. F., Reis, L. P., & Zanuncio, J. C. (2018). Artificial neural networks: Modeling tree survival and mortality in the Atlantic Forest biome in Brazil. *Science of The Total Environment*, *645*, 655-661. https://doi.org/10.1016/j.scitotenv.2018.07.123

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). Redes neurais artificiais para estimação do volume de árvores [Neural networks for estimating of the volume of tree]. *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). Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais [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), Article 268. https://doi.org/10.3390/f10030268

de Abreu Demolinari, R, Soares, C. P. B., Leite, H. G., & de Souza, A. L. (2007). Crescimento de plantios clonais de eucalipto não desbastados na região de Monte Dourado (PA) [Growth of unthinned clonal eucalyptus plantations in the region of Monte Dourado (PA)]. *Revista Árvore*, *31*(3), 503-512. https://doi.org/10.1590/S0100-67622007000300016

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 Azevedo, G. B., de Oliveira, E. K. B., Azevedo, G. D. O., Buchmann, H. M., Miguel, E. P., & Rezende, A. V. (2016). Modeling production by stand and diameter distribution in eucalyptus plantations. *Scientia Forestalis*, *44*(110), 383-392.

de Oliveira, B. R., da Silva, A. A. P., Teodoro, L. P. R., de Azevedo, G. B., de Oliveira Sousa Azevedo, G. T., Baio, F. H. R., Sobrinho, R. L., da Silva Junior, C. A., & Teodoro, P. E. (2021). Eucalyptus growth recognition using machine learning methods and spectral variables. *Forest Ecology and Management 497*, Article 119496. https://doi.org/10.1016/j.foreco.2021.119496

Dias, A. N., Leite, H. G., Campos, J. C. C., Couto, L., & Carvalho, A. F. (2005). Emprego de um modelo de crescimento e produção em povoamentos desbastados de eucalipto [The use of a growth and yield model in thinned eucalypt stands]. *Revista Árvore*, *29*(5), 731-739. https://doi.org/10.1590/S0100-67622005000500008

dos Santos, H. G., Jacomine, P. K. T., dos Anjos, L. H. C., de Oliveira, V. A., Lumbreras, J. F., Coelho, M. R., de Almeida, J. A., de Araujo Filho, J. C., de Oliveira, J. B., & Cunha, T. J. F. (2018). *Sistema brasileiro de classificação de solos* (5a edição revista e ampliada) [Brazilian soil classification system (5th edition revised and expanded)]. Embrapa.

Freitas, E. F. S., 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*, Article 112149. https://doi.org/10.1016/j.indcrop.2020.112149

García, O. (1988). Growth modelling - A review. *New Zealand Forestry*, *33*(3), 14-17.

Gompertz, B. (1825). XXIV. 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/rspl.1815.0271

Gujarati, D. N., & Porter, D. C. (2011). *Basic econometrics* (5th Ed). AMGH Editora.

IBGE. (2018). *Macrocaracterização - Tipos de Solos* [Macrocharacterization - Soil types]. Instituto Brasileiro de Geografia e Estatística. https://portaldemapas.ibge.gov.br/portal.php#homepage

IBGE. (2021). *Produção da extração vegetal e da silvicultura – PEVS, 2020* [Vegetal extraction and forestry production - PEVS, 2020]. Instituto Brasileiro de Geografia e Estatística. https://www.ibge.gov.br/

estatisticas/economicas/agricultura-e-pecuaria/9105-producao-da-extracao-vegetal-e-da-silvicultura.html?edicao=29153&t=resultados

Liu, L., Lim, S., Shen, X., & Yebra, M. (2020). Assessment of generalized allometric models for aboveground biomass estimation: A case study in Australia. *Computers and Electronics in Agriculture, 175*, Article 105610. https://doi.org/10.1016/j.compag.2020.105610

Martins, E. R., Binoti, M. L. M. S., Leite, H. G., Binoti, D. H. B., & Dutra, G. C. (2015). Configuração de redes neurais artificiais para prognose da produção de povoamentos clonais de eucalipto [Configuration of artificial neural network for prognosis the production of eucalyptus clonal stands]. *Revista Brasileira de Ciências Agrárias (Agrária), 10*(4), 532-537. https://doi.org/10.5039/agraria.v10i4a5350

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

Nieto, P. J. G., Torres, J. M., Fernández, M. A., & Galán, O. C. (2012). Support vector machines and neural networks used to evaluate paper manufactured using *Eucalyptus globulus*. *Applied Mathematical Modelling, 36*(12), 6137-6145. https://doi.org/10.1016/j.apm.2012.02.016

Penido, T. M. A., Lafetá, B. O., Nogueira, G. S., Alves, P. H., Gorgens, E. B., & Oliveira, M. L. R. (2020). Modelos de crescimento e produção para a estimativa volumétrica em povoamentos comerciais de eucalipto [Growth and production models for volumetric estimates in commercial eucalypt stands]. *Scientia Forestalis, 48*(128), Article e3340. https://doi.org/10.18671/scifor.v48n128.06

Pereira, K. D., Carneiro, A. P. S., Santos, G. R., Carneiro, A. C. O., Leite, H. G., & Borges, F. P. (2021). Study of the influence of wood properties on the charcoal production: Applying the random forest algorithm. *Revista Árvore 45*, Article e4502. http://dx.doi.org/10.1590/1806-908820210000002

Reis, L. P., de Souza, A. L., dos Reis, P. C. M., Mazzei, L., Soares, C. P. B., Torres, C. M. M. E., da Silva, L. F., Ruschel, A. R., Rêgo, L. J. S., & Leite, H. G. (2018). Estimation of mortality and survival of individual trees after harvesting wood using artificial neural networks in the amazon rain forest. *Ecological Engineering*, *112*, 140-147. https://doi.org/10.1016/j.ecoleng.2017.12.014

Rodriguez, L. C. E., Bueno, A. R. S., & Rodrigues, F. (1997). Rotações de eucaliptos mais longas: análise volumétrica e econômica [Longer eucalypt rotations: volumetric and economic analysis]. *Scientia Forestalis*, *51*(1), 15-28.

Salles, T. T., Leite, H. G., de Oliveira Neto, S. N., Soares, C. P. B., de Paiva, H. N., & dos Santos, F. L. (2012). Modelo de Clutter na modelagem de crescimento e produção de eucalipto em sistemas de integração lavoura-pecuária-floresta [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

Scolforo, H. F., McTague, J. P., Burkhart, H., Roise, J., Campoe, O., & Stape, J. L. (2019a). Eucalyptus growth and yield system: Linking individual-tree and stand-level growth models in clonal Eucalypt plantations in Brazil. *Forest Ecology and Management*, *432*, 1-16. https://doi.org/10.1016/j.foreco.2018.08.045

Scolforo, H. F., McTague, J. P., Burkhart, H., Roise, J., McCarter, J., Alvares, C. A., & Stape, J. L. (2019b). Stand-level growth and yield model system for clonal eucalypt plantations in Brazil that accounts for water availability. *Forest Ecology and Management*, *448*, 22-33. https://doi.org/10.1016/j.foreco.2019.06.006

Sharma, R. P., Vacek, Z., Vacek, S., & Kučera, M. (2019). Modelling individual tree height–diameter relationships for multi-layered and multi-species forests in central Europe. *Trees*, *33*(1), 103-119. https://doi.org/10.1007/s00468-018-1762-4

Silva, I. N., Spatti, D. H., & Flauzino, R. A. (2016). *Redes Neurais Artificiais para engenharia e ciências aplicadas curso prático.* (2ª edição revisada e ampliada) [Artificial Neural Networks for engineering and applied sciences: practical course. (2nd edition revised and expanded)]. Artliber.

Silva, J. P. M., da Silva, M. L. M., de Mendonça, A. R., da Silva, G. F., de Barros Junior, A. A., da Silva, E. F., Aguiar, M. O., Santos, J. S., & Rodrigues, N. M. M. (2021). Prognosis of forest production using machine learning techniques. *Information Processing in Agriculture*, 1-14. https://doi.org/10.1016/j.inpa.2021.09.004

Stankova, T. V. (2016). A dynamic whole-stand growth model, derived from allometric relationships. *Silva Fennica, 50*(1), 1406. http://dx.doi.org/10.14214/sf.1406

Vescovi, L. B., Leite, H. G., Soares, C. P. B., de Oliveira, M. L. R., Binoti, D. H. B., Fardin, L. P., Silva, G. C. C., de Sousa Lopes, L. S., Leite, R. V., de Oliveira Neto, R. R., Silva, S. (2020). Effect of growth and yield modelling on forest regulation and earnings. *African Journal of Agricultural Research, 16*(7), 1050-1060. https://doi.org/10.5897/AJAR2020.14755

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*, 1473-1481. https://doi.org/10.1016/j.scitotenv.2017.11.138