Australian Journal of Crop Science AJCS
AJCS 18(3):153-159 (2024) ISSN:1835-2707
https://doi.org/10.21475/ajcs.24.18.02.PNE4041

Received: 04/08/2023 | Revised: 31/08/2023 | Accepted: 21/02/2024

Categorical variable modeling for estimating pulp biomass and recovered residues from *Eucalyptus* harvesting

Carla Krulikowski Rodrigues*, Allan Libanio Pelissari, Eduardo da Silva Lopes

Federal University of Paraná, Prefeito Lothário Meissner, 632, 80210-170, Curitiba, Paraná, Brazil
Federal University of Paraná, Prefeito Lothário Meissner, 632, 80210-170, Curitiba, Paraná, Brazil
State University of the Central West, Km 7, PR 153, 84500-000, Irati, Paraná, Brazil

***Corresponding author: carlakr@gmail.com**

Abstract

The definition of stem diameter limits between pulp production and residues recovered from harvesting for energy can influence the dry biomass content of forest products. In this context, the aim of this study was to build linear regression models employing categorical variables to accurately estimate the biomass of pulp and residues from *Eucalyptus saligna* and *E. urophylla* × *E. grandis* clonal stands. Forest inventory and dry biomass data from thirty trees distributed in diameter classes of each stand were used as a database. The measurements of pulp dry mass and residues were treated as dependent variables. Concurrently, variables such as diameter at 1.3 m above the ground, total height, and transformations and combinations of these variables were considered as independent variables. The stem diameter limits of 8, 10, 12, and 14 cm along the stems were considered as categorical variables. To predict pulp and energy biomasses based on stem diameter limits, the categorical variable assumes a value of 1 (indicating an estimate within a particular k-factor), while the other k-factors were set to 0 (zero). The fitted models allowed for the estimation of dry biomass within the forest stands, resulting in percentage errors of estimation per hectare ranging between 4.5% and 11.2%. Moreover, these models incorporating categorical variables facilitated the simplification of estimates for both pulp biomass and recovered residues within commercial *Eucalyptus* stands.

Keywords: *Dummy* variables, Forest biomass, Stem diameter limits, Wood by-products.

Abbreviations: *d*_diameter at 1.3 m above ground, *h*_total height, w_c_dry biomass for pulp, w_e_biomass for energy, $\hat{y}$_estimate of dependent variable, x_independent variable, A_k_categorical variable for a given factor k, $\hat{\beta}_k$_estimated regression coefficient, $\hat{e}$_estimate of the random error, SW_Shapiro-Wilk normality test, BP_Breusch-Pagan homoscedasticity test, DW_Durbin-Watson test, $R^2_{adj.}$_adjusted coefficient of determination, SEE_standard error of the estimate in percent, $\hat{\bar{W}}$_Sample means, $cv\%$_coefficients of variation, E_a_absolute error, $E\%$_percentage error, CI_confidence intervals for the mean, P_probability level.

Introduction

Eucalyptus plantations are on the rise in Brazil, while the market for their products is considered promising. The country has approximately 9.93 million hectares of these plantations established in various states to supply the diverse segments within the forest sector. In addition, Brazil is known for its great forestry vocation for presenting suitable edaphoclimatic conditions, positioning it as a global leader in forest productivity, with an annual average of 38.9 m^3 ha^{-1} (Iba, 2022).

Forest plantations offer an alternative approach to nature conservation, reducing the pressure on native forests (Zhang et al., 2015). Furthermore, these plantations have environmental benefits, such as soil conservation, biomass production, and CO_2 sequestration (Du et al., 2015; Liang et al., 2016). In fact, there is no more space in the market for production segments that do not align economic activity with environmental preservation (Hunt et al., 2016). In this context, eucalyptus species are presented as sustainable forest plantations capable of satisfying environmental premises.

Forest plantations offer a wide diversity of products destined for the various segments, such as sawmills, veneer, pulp, and energy (Santos et al., 2016; Kazmierczak et al., 2017, Carvalho et al., 2019). It is noteworthy that pulp production in 2020 reached its highest historical volume with 22.5 million tons (Iba, 2022), in which 70% was exported. On the other hand, Hunt et al. (2016) note that the pulp and paper companies are reaching their market limits with decreasing demand due to the increasing use of digital media. In addition, there is a worldwide need to replace fossil fuels with sustainably managed products that mitigate CO_2 emissions in the context of new climate change policies.

Energy production derived from forest biomass stands as an alternative to pulp production (Hora, 2017). In recent years, the production and commercialization of renewable energy from planted forests has advanced to 3.6 million tons, increasing the importance of biomass for Brazil's energy security (Iba, 2022). However, the energy forest option changes the management of forest stands by establishing denser plantations and earlier harvests. In addition, multi-products can be obtained from the trees, through assortments of pulp and biomass for energy.

Table 1. Attributes of pulp biomass production and recovered residues for energy at tree level in *Eucalyptus* stands.

d_i (cm)	*Eucalyptus saligna*				*Eucalyptus urophylla* × *Eucalyptus grandis*			
	$\bar{w}_c \pm s_x$ (kg)	cv_{w_c}%	$\bar{w}_e \pm s_x$ (kg)	cv_{w_e}%	$\bar{w}_c \pm s_x$ (kg)	cv_{w_c}%	$\bar{w}_e \pm s_x$ (kg)	cv_{w_e}%
08	134.6 ± 78.5	58.3	16.3 ± 2.7	16.5	183.1 ± 85.1	46.5	15.9 ± 4.2	26.7
10	122.3 ± 76.3	62.4	28.4 ± 5.1	18.1	171.4 ± 85.1	49.6	29.9 ± 6.4	21.4
12	113.7 ± 78.4	68.9	46.0 ± 7.6	16.5	156.6 ± 85.7	54.7	51.1 ± 11.8	23.1
14	116.4 ± 76.3	65.6	75.6 ± 13.9	18.4	158.9 ± 73.7	46.4	84.9 ± 17.5	20.6

d_i is the stem diameter limit, $\bar{w}_c$ is the sample mean of pulp biomass production, $\bar{w}_c$ is the sample mean of biomass energy production, s_x is the standard deviation of the sample mean, cv_{w_c}% is the coefficient of variation for pulp biomass and cv_{w_e}% is the coefficient of variation for residues recovered for energy.

Some countries, such as the United States, use full trees for energy production originally conducted for pulp production, since their industries require quality chips, with a higher portion of wood and less bark (Baker et al., 2012).

Besides the market and forest business issues, there are operational difficulties in feeding power boilers for the cogeneration of thermal energy in pulp and paper companies. These boilers are supplied by residues recovered from wood harvesting operations, composed of the final stem section usually with a diameter of less than 8 cm; in addition to barks, branches, and leaves that present low quality in physical and chemical characteristics, due to the high content of fine dust, ash, and moisture. According to Dai et al. (2012) and Rackl and Günthner (2016), these characteristics are the main causes of interlocks in screw-type boiler feeders when using residue biomass, which can cause fires and work accidents.

Due to the shortage of quality forest raw material for energy production, low fluidity point oil is utilized. This leads to an increase of up to five times in energy costs compared to utilizing the residues recovered from wood harvesting. Nonetheless, biomass should possess the necessary physical characteristics to ensure the uninterrupted flow of the boiler feed, resulting in neutral carbon emissions and cost savings in the supply. This objective can be attained by selecting the critical stem diameter between the two products or by using fully mature trees initially intended for pulping.

The estimation of forest biomass for pulp and energy purposes within forest stands enables the planning of their utilization. These estimations are typically conducted using methodologies based on forest inventory data, employing biomass factors or equations where diameter, height, and volume provide as the independent variables (Somogyi et al., 2007). To achieve this, destructive sampling is commonly employed, involving the separation of tree components (Silveira et al., 2008).

While studies concerning biomass quantification in *Eucalyptus* stands have been published (Assis et al., 2015; Salvador et al., 2016), their focus was on determining the dry biomass stock and nutrient content present in tree components, along with potential CO_2 sequestration. As a result, biomass quantification becomes essential to guide the determination of economically favorable stem diameter limits between pulpwood and residues recovered from wood harvesting for energy production. This aims to ascertain the economically advantageous ratio between these products."

Regression models with categorical variables can be an alternative for estimating the dry biomass of products considering different stem diameter limits. These models enable the estimation quantitative variables in a single model, considering *dummy* indicator variables which assume values of 0 (zero) or 1 (one) to denote the absence or presence, respectively, of specific attributes or categories (Pal and Bharati, 2019; Fernandes et al., 2022). By means of statistical models with categorical variables, it becomes possible to provide accurate and easily interpreted estimates of pulp and energy forest production in different scenarios.

The aim of this study was to build linear regression models with categorical variables to estimate the biomass of pulp and residues recovered for energy, considering the stem diameter limits. By means of this approach, it is possible to provide information to managers for decision making under different forest management scenarios, as well as to support harvest operations, forest transport plannings, stockyard management, and the determination of consumer unit supply.

Results

The significance of estimated regression coefficients for the first factor ($\hat{\beta}_{A1}$) indicates statistical difference of zero, except for pulp biomass (w_c) in *E. urophylla × E. grandis* stand (Table 2). For this one, the biomass production considering the stem diameter limit of 8 cm can be estimated only as a function of quantitative variable d^2h. This result indicates the average pulp biomass at the height of 8-cm stem diameter is similar to the production considering the full stem.

The significance for the other regression coefficients shows the existence of a statistical difference in biomass production between the stem diameter limits (Table 2). The coefficients with negative signs express the reduction in pulp biomass (w_c) with increasing stem diameter limit. On the other hand, the positive coefficients indicate the increase in residue production for energy (w_e) with an increase in stem diameter limit.

The lack of statistical significance by the Shapiro-Wilk (SW) test indicates the residuals are normally distributed, while non-significance by the Breusch-Pagan (BP) test shows homogeneity of variances, both at the 1% level (Table 2). The lack of residual autocorrelation in regression models was proved by the Durbin-Watson test (DW), at the 1% significance level (Table 2). These results corroborate the quality of the fits to satisfy the fundamental linear regression assumptions.

The product of the square of diameter at 1.3 m above ground (d^2) with the total height (h) resulted in a quantitative variable chosen to estimate biomass production (Table 3), presenting adjusted determination coefficients ($R^2_{adj.}$) greater than 0.85. Furthermore, the standard error of estimate (SEE) was satisfactory for the energy biomass modeling (w_e). On the other hand, the error measures for

energy biomass models (w_e) were close to 20%, which expresses the higher difficulty to fit this component and justifies the use of log-transformation (Table 3).
The studentized residuals of models fitted for pulp biomass and residues for energy showed a constant distribution, where the most of values are present in the range of ±2 standard deviations of the t-distribution, indicating the lack of extreme data (Figure 1). However, brief overestimation in *E. saligna* was observed for the estimation of highest pulp biomass production, in which the dispersion of these residuals remained in the desired range values.
When biomass productions were estimated in the sample units by means of categorical variables (Table 4), a reduction in the sample mean ($\hat{\bar{W}}$) of pulp biomass was observed with increasing stem diameter limit (d_i). In opposition, there was an increase in the estimated sample mean for energy biomass. Furthermore, the energy biomass production was higher in *E. saligna* stand, while pulp biomass was greater in *E. urophylla* × *E. grandis* stand.
In general, greater variability in estimates (cv%) was observed in *E. urophylla* × *E. grandis* stand (Table 4). However, the measures of sampling errors (E%) were less than or close to the 10% error limit, resulting in confidence intervals (CI) suitable for predicting the production of pulp biomass and residues for energy at the 95% probability (P) level.

Discussion

In this study, models with categorical variables were fitted to estimate the biomass of pulp and residues recovered for energy at the tree level, using stem diameter limits. The statistical significance of the regression coefficients (Table 2) supports the inclusion of stem diameter limits as factors within the models. These models satisfied the linear regression assumptions, including residual normality, homoscedasticity, and absence of residual autocorrelation (Table 2), allowing appropriate statistical inferences regarding biomass productions.
The combined variable d^2h was employed as a quantitative predictor in biomass models (Table 3). This variable is commonly used in estimating the volume of (Raptis et al., 2020; Behling et al., 2021) and biomass of forest species (Bi et al., 2004; Vargas-Larreta et al., 2017), where the linearization of the relationship with the response variable is its main advantage. For this study, the use of d^2h can be justified, since the aboveground biomass is proportional to the volume of a cylinder of diameter d and height h (Dutcă et al., 2019). Moreover, these variables can be easily acquired through pre-cutting forest inventories, a standard practice in forest companies.
The fitted models can provide information for decision-making under different management scenarios, aiding in the planning of harvesting and transportation operations, yard management, and determining consumer unit supply. To exemplify the use of these fitted models in wood harvesting operations: biomass estimates for pulp and energy support the sizing of harvesting machines, work teams, and transport vehicles, as well as the biomass-to-energy conversion system (Wolfsmayr and Rauch, 2014; Nunes et al., 2020).
In general, harvesting productivity control is based only on pulp biomass. Thus, changing the usual stem diameter limit from 8 cm to larger values could potentially decrease machine productivity, resulting in reduced pulp biomass per unit area and requiring greater machine displacement for a same pulp production. Additionally, it is highlighted that harvesting machines do not quantify recovered residues, which are subsequently processed by the chipping system in the field (Rodrigues et al., 2019). Therefore, quantifying residues becomes essential for sizing chippers and transportation vehicles (Sahoo et al., 2019).
Another consideration involves to storing recovered residues for energy over an extended period to reduce wood humidity, ranging from 30 to 180 days (Oro et al., 2018). With an 8 cm stem diameter limit, the biomass for energy stored in the field is lower. On the other hand, a larger stem diameter leads to higher biomass production (Robinson et al., 2004), which requires more area as a storage yard and more time for drying residues with larger diameters, which should be considered into replacement planting decisions.
Modifying the stem diameter limit from the standard 8 cm between products can enhance the quality of energy chips, potentially mitigating interlocking in screw-type boiler feeders when using biomass from wood harvesting residues (Dai et al., 2012; Rackl and Günthner, 2016). However, utilizing biomass for energy requires more than simple decision-making, in which demands studies on the physical-chemical characteristics of the biomass (Mead and Pimentel, 2006). The company in this study reported occasional utilization of stored pulpwood as a raw material for energy, aiming to improve the biomass characteristics.
In addition to the mentioned examples, it is worth noting the models fitted with categorical variables considered the stem diameter limits of 8, 10, 12, and 14 cm for *E. saligna* and *E. urophylla* × *E. grandis* stands. Therefore, there are sixteen production scenarios if they were considered as independent fits. In this way, the incorporation of categorical variables reduced the estimation process to four models, with two models for each *Eucalyptus* stand, simplifying the procedure for predicting biomass production.

Materials and Methods

Study area and database

The study was carried out in the Southern region of Brazil at the coordinates 24°26'42" S and 50°45'39" W. The predominant climate is classified as Cfb (Köppen), with an average annual temperature between 18 and 20 ºC and average annual precipitation between 1,600 and 1,900 mm (Alvares et al., 2013). The evaluated forest stands were composed of clones of *Eucalyptus saligna* Smith and interspecific hybrids of *Eucalyptus urophylla* S. T. Blake × *Eucalyptus grandis* W. Hill ex Maiden, which are the most planted genetic materials due to their productivity and characteristics that attend the consumer unit. These stands were 7 years-old, had an average density of 1,111 trees per hectare, spacing of 3.75 m × 2.40 m, and the same site.
A forest inventory was carried out before the wood harvesting operation, aiming to measure the diameter at 1.3 m above ground (d) and total height (h) of the trees. The sampling intensity was determined for finite populations (Kershaw et al., 2017), considering the error limit of 10% at the 5% probability level. Thus, 29 sampling units (s.u.) of

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 A_2 + \hat{\beta}_3 A_3 + \cdots + \hat{\beta}_k A_{k-1} + \hat{e} \quad \text{(Eq. 1)}$$

37.1 m × 23.3 m were randomly allocated in the stands, with 10 s.u. in *E. saligna* and 19 s.u. in *E. urophylla* × *E. grandis*. In addition, five diameter classes with 4 cm intervals were established in each stand, whose distributions were used to define the number of sampled trees for green biomass measurement.

Table 2. Regression coefficients of the categorical variable and statistical tests for estimating pulp biomass and recovered residues for energy in *Eucalyptus* stands.

id.	$\hat{y}$	$\hat{\beta}_{A1}$	$\hat{\beta}_{A2}$	$\hat{\beta}_{A3}$	$\hat{\beta}_{A4}$	SW	BP	DW
	Eucalyptus saligna							
(1)	w_c	-15.6553*	-12.2480*	-28.5073*	-52.1467*	0.984^{ns}	0.158^{ns}	1.729^{ns}
	(p-valor)	(<0.05)	(<0.05)	(<0.05)	(<0.05)	(0.349)	(0.691)	(0.050)
(2)	$\ln(w_e)$	2.8266*	0.5532*	1.0445*	1.5407*	0.985^{ns}	0.534^{ns}	2.249^{ns}
	(p-valor)	(<0.05)	(<0.05)	(<0.05)	(<0.05)	(0.301)	(0.465)	(0.829)
	Eucalyptus urophylla × *Eucalyptus grandis*							
(3)	w_c	-0.7379^{ns}	-11.7000*	-27.4802*	-53.3994*	0.983^{ns}	4.662^{ns}	1.942^{ns}
	(p-valor)	(<0.05)	(<0.05)	(<0.05)	(<0.05)	(0.239)	(0.031)	(0.274)
(4)	$\ln(w_e)$	2.8245*	0.6522*	1.1913*	1.7045*	0.995^{ns}	2.205^{ns}	2.082^{ns}
	(p-valor	(<0.05)	(<0.05)	(<0.05)	(<0.05)	(0.963)	(0.138)	(0.563)

$\hat{y}$ is the estimated dependent variable, w_c is the pulp biomass, w_e is the recovered residues from wood harvesting for energy production, $A1$, $A2$, $A3$ and $A4$ are the stem diameter limits of 8, 10, 12, and 14 cm, SW is the Shapiro-Wilk normality test, BP is the Breusch-Pagan homoscedasticity test, DW is the Durbin-Watson autocorrelation test, * significance at 5% level, and ^{ns} non-significance at 1% level.

Table 3. Models fitted with categorical variables for the estimation of pulp biomass and recovered residues for energy in *Eucalyptus* stands.

id.	Model	$R^2_{adj.}$	SEE%
	Eucalyptus saligna		
(1)	$w_c = 0.0169\, d^2h - 15.6553\, d_8 - 27.9033\, d_{10} - 44.1627\, d_{12} - 67.8020\, d_{14}$	0.996	7.58
(2)	$\ln(w_e) = -\,0.000006\, d^2h + 2.8266\, d_8 + 3.3798\, d_{10} + 3.8712\, d_{12} + 4.3674\, d_{14}$	0.879	20.17
	Eucalyptus urophylla × *Eucalyptus grandis*		
(3)	$w_c = 0.0176\, d^2h - 0.7379\, d_8 - 12.4379\, d_{10} - 28.2181\, d_{12} - 54.1373\, d_{14}$	0.990	10.98
(4)	$\ln(w_e) = -\,0.000010\, d^2h + 2.8245\, d_8 + 3.4767\, d_{10} + 4.0158\, d_{12} + 4.5290\, d_{14}$	0.857	23.69

w_{cel} is the pulp biomass (kg), w_e is the recovered residues for energy (kg); d is the diameter at 1.3 m above ground (cm), h is the total height (m), d_i is the stem diameter limits of 8, 10, 12, and 14 cm, $R^2_{aj.}$ is the adjusted coefficient of determination and SEE% is the standard error of the estimate in percentage.

Green biomass was quantified using the destructive method, in which tree components were separated, measured, and sampled. Each tree was segmented in the compartments: 1) stem with bark, for pulp, 2) end stem section, for energy, and 3) branches and leaves, for energy. The stem mass was measured in four end section scenarios, considering stem diameter limits (d_i) of 8, 10, 12, and 14 cm in commercial height between pulp and energy products. In addition, we simulated the sectioning of 7.2 m long logs for pulp and the harvesting of logs longer than 2.4 m due to the distance from the stakes in the log-carrying vehicles.

The sampled green biomass was measured with a digital balance of 0.01 g precision and 500 kg capacity. For the stem, 5 cm thick discs were taken at the relative diameter heights of 0% (base), 25%, 50%, and 75% of the total height of trees. Leaf and branch samples were obtained from the tip, middle and base of the branches located at the mid-third of the crowns. Biomass samples were dried in an oven with air renewal and circulation until constant mass. A 0.01 g precision balance was used to measure mass, in which the dry biomass of trees was determined using the moisture content of the samples in each component.

Fitting regression models with categorical variables

The dry biomass for pulp (w_c) was determined for each tree, as well as the residues recovered from wood harvesting for energy (w_e), such as branches, leaves, whole stem bark, and final stem section, were sampled considering the stem diameter limit scenarios (Table 1). Using these data, linear regression models with categorical variables were fitted at the tree level, in which the diameter limits (d_i) of 8, 10, 12, and 14 cm along the stem were considered as the four k factors of the categorical variable in the models (A_k).

The dry biomass for pulp and residues recovered for energy were considered as dependent variables ($\hat{y}$). The diameter at 1.3 m above ground (d) and the total height (h), as well as the logarithmic, quadratic power, inverse, and their multiplicative and division combinations, were assumed as quantitative independent variables (x) in the regression models (Equation 1).

where $\hat{y}$ is the estimate of dependent variable, x is the quantitative independent variable, A_k is the categorical variable for a given factor k (stem diameter limit), $\hat{\beta}_k$ is the estimated regression coefficient and $\hat{e}$ is the estimate of the random error.

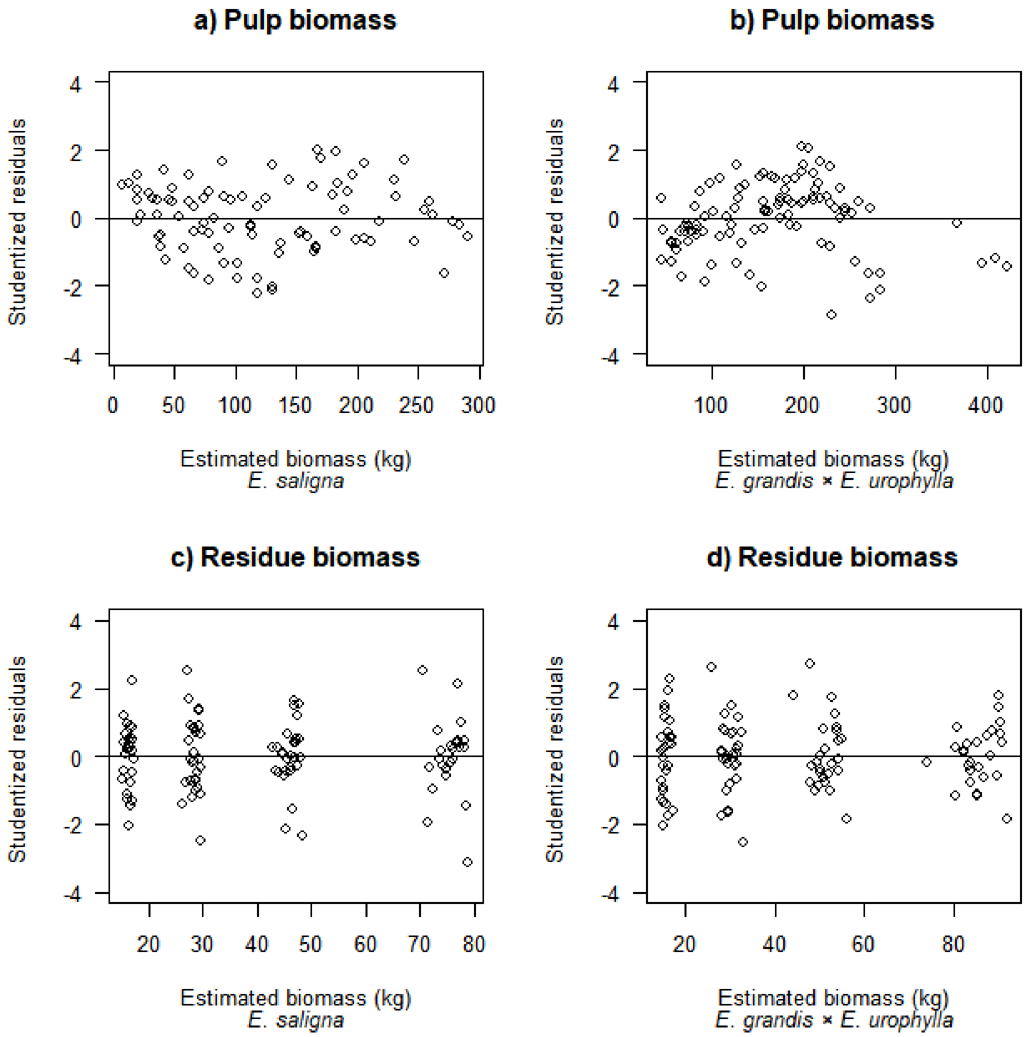

Figure 1. Studentized residuals of the models fitted with categorical variables for estimating pulp biomass and recovered residues for energy in *Eucalyptus* stands.

Table 4. Estimates of pulp biomass production and recovered residues for energy in *Eucalyptus* stands by means of the fitted models with categorical variables.

d_i(cm)	Product	$\hat{\bar{W}}$(Mg ha^{-1})	$cv\%$	E_a(Mg ha^{-1})	$E\%$	$CI[LL(\text{Mg ha}^{-1}) \leq \bar{W} \leq UL(\text{Mg ha}^{-1})] = P$
				Eucalyptus saligna		
08	Pulp	203.09	11.8	±16.76	±8.3	186.33 – 219.85
	Energy	18.01	7.1	±0.89	±4.5	17.01 – 18.90
10	Pulp	189.11	12.2.	±16.23	±8.6	172.87 – 205.34
	Energy	31.31	7.1	±1.56	±5.0	29.75 – 32.86
12	Pulp	170.54	13.0	±15.56	±9.1	154.98 – 186.09
	Energy	51.17	7.1	±2.55	±5.0	48.63 – 53.72
14	Pulp	125.67	15.9	±14.05	±11.2	104.78 – 158.66
	Energy	84.05	7.1	±4.18	±5.0	79.87 – 88.24
Eucalyptus urophylla × *Eucalyptus grandis*						
08	Pulp	209.47	14.2	±14.02	±6.7	195.45 – 223.50
	Energy	15.95	18.5	±1.39	±8.7	14.56 – 17.34
10	Pulp	197.04	14.1	±13.08	±6.6	183.96 – 210.11
	Energy	30.62	18.5	±2.67	±8.7	27.95 – 33.29
12	Pulp	180.27	13.9	±11.84	±6.6	168.44 – 192.11
	Energy	52.49	18.5	±4.58	±8.7	47.91 – 57.08
14	Pulp	152.73	13.8	±9.89	±6.5	142.84 – 162.63
	Energy	87.68	18.5	±7.65	±8.7	80.03 – 95.34

d_i is the stem diameter limit, $\hat{\bar{W}}$ is the estimate of the sample mean, $cv\%$ is the coefficient of variation, E_a is the absolute error of the estimate, $E\%$ is the percent error of the estimate, CI is the confidence interval for the mean, LL is the estimated lower limit, $\bar{W}$ is the mean population value, UL is the estimated upper limit and P is the 95% probability level.

For the prediction of pulp and energy biomasses by stem diameter limit, the categorical variable assumes on a value equal to 1 (one) for the estimate at a given k-factor. Furthermore, the other k-factors assume value of 0 (zero) in the same regression model. Thus, considering the incorporation of stem diameter limits as an additive effect categorical variable to estimate biomass per tree, different intercepts were estimated, resulting in one level per k-factor (Equation 2), in which the final models for estimating biomass production ($\hat{w}$) can be rewrite by means of (Equation 3).

$$\hat{w} = \begin{cases} \hat{\beta}_0 + \hat{\beta}_1 x & \text{to factor } A_1\text{: stem diameter limits } (d_i) \text{ of 8 cm} \\ (\hat{\beta}_0 + \hat{\beta}_2) + \hat{\beta}_1 x & \text{to factor } A_2\text{: stem diameter limits } (d_i) \text{ of 10 cm} \\ (\hat{\beta}_0 + \hat{\beta}_3) + \hat{\beta}_1 x & \text{to factor } A_3\text{: stem diameter limits } (d_i) \text{ of 12 cm} \\ (\hat{\beta}_0 + \hat{\beta}_4) + \hat{\beta}_1 x & \text{to factor } A_4\text{: stem diameter limits } (d_i) \text{ of 14 cm} \end{cases} \quad \text{(Eq. 2)}$$

$$\hat{w} = \hat{\beta}_1 x + \hat{\beta}_{A1} A_1 + \hat{\beta}_{A2} A_2 + \hat{\beta}_{A3} A_3 + \hat{\beta}_{A4} A_4 + \hat{e} \quad \text{(Eq. 3)}$$

where $\hat{w}$ is the estimate of biomass production for pulp and energy and A_1, A_2, A_3 and A_4 are the stem diameter limits (d_i) of 8, 10, 12 and 14 cm, respectively.

The estimated regression coefficients ($\hat{\beta}_k$) per k-factor were compared at the 5% significance level by the t-test, making it possible to indicate the influence of each factor on the biomass estimation. Furthermore, Shapiro-Wilk normality test (SW) and Breusch-Pagan homoscedasticity test (BP) were applied to verify the linear regression assumptions, considering non-significance at the 1% probability level. Also, Durbin-Watson test (DW) was performed to verify the lack of residual autocorrelation in regression models at the 1% probability level. The adjusted coefficient of determination ($R^2_{adj.}$) and standard error of the estimate in percent (SEE) were applied to evaluate the quality of estimates. Graphical analysis of the studentized residuals was also used to evaluate biases in the estimates and outliers.

Estimation of pulp and residues biomass production

By means of fitted models with categorical variables, we estimated the biomass of pulp and residues for energy in the inventory samples, considering the stem diameter limits and the estimators of simple random sampling (Kershaw et al., 2017). Sample means ($\hat{\overline{W}}$) and coefficients of variation ($cv\%$) were used as descriptive measures, while absolute (E_a) and percentage ($E\%$) errors made it possible to evaluate the quality of the estimates. Finally, confidence intervals for the mean (CI) were obtained at the 95% probability level (P).

Conclusion

The biomass models fitted with categorical variables, considering the stem diameter limits of 8, 10, 12, and 14 cm, made it possible to simplify the estimates of pulp biomass and recovered residues production in commercial *Eucalyptus* stands. These models allow decision-making in different management scenarios, aiming to support the machines, vehicles, and equipment sizing, while also ensuring a consistent supply of high-quality raw materials for pulp and energy production.

References

Alvares CA, Stape JL, Sentelhas PC, Gonçalves JLM, Sparovek G (2013) Köppen's climate classification map for Brazil. Meteorol Z. 22(1):711-728.

Assis MR, Trugilho PF, Rosado SCS, Protasio TP, Goulart SL (2015) Modelagem da biomassa e do estoque de carbono em plantas jovens de *Eucalyptus.* Sci For. 43(105):01-13.

Baker SA, Greene WD, Wilson A (2012) Fuels characteristics of woods-run whole tree southern pine chips. Biomass Bioenerg. 37:67-72.

Behling M, Koehler HS, Behling A (2021) Compatibility between the Spurr function volume and the Kozak's taper function and quintic polynomial volumes for black wattle trees. Floresta. 51(2):521-530.

Bi H, Turner J, Lambert MJ (2004) Additive biomass equations for native eucalypt forest trees of temperate Australia. Trees. 18(4).

Carvalho DE, Rocha MP, Timofeiczyk Júnior R, Klitzke RJ (2019) Rendimento e variedade de produtos no desdobro de toras de *Eucalyptus* spp. Tecno-Lógica. 23(1):08-13.

Dai J, Cui H, Grace JR (2012) Biomass feeding for thermochemical reactors. Prog Energ Combust. 38(5):716-736.

Du H, Zeng F, Peng W, Wang K, Zhang H, Liu L, Song T (2015) Carbon storage in a eucalyptus plantation chrono sequence in Southern China. Forests. 6:1763-1778.

Dutcă I, McRoberts RE, Næsset E, Blujdea VNB (2019) A practical measure for determining if diameter (D) and height (H) should be combined into D^2H in allometric biomass models. Forestry. 92:627-634.

Fernandes AAR, Solimun S, Nurjannah N (2022) Computational Statistics with Dummy Variables. In: López-Ruiz R (ed) Computational Statistics and Applications. IntechOpen, London.

Hora AB (2017) Panoramas setoriais 2030: papel e celulose. In: Panoramas setoriais 2030: desafios e oportunidades para o Brasil. BNDES, Rio de Janeiro.

Hunt JD, Guillot V, Freitas MAV, Solari RSE (2016) Energy crop storage: An alternative to resolve the problem of unpredictable hydropower generation in Brazil. Energy. 101:91-99.

Iba. Brazilian Tree Industry (2022) Relatório anual IBÁ 2022. Fundação Getúlio Vargas, São Paulo.

Kazmierczak S, Hillig E, Iwakiri S (2017) Painéis compensados fabricados com lâminas de três espécies de eucaliptos. FLORAM. 24:e00111614.

Kershaw JA, Ducey MJ, Beers TW, Husch B (2017) Forest mensuration. Wiley, New York. 5.

Liang J, Reynolds T, Wassie A, Collins C, Wubalem A (2016) Effects of exotic *Eucalyptus* spp. plantations on soil properties in and around sacred natural sites in the northern Ethiopian Highlands. AIMS Agric Food. 1(2):175-193.

Mead DJ, Pimentel D (2006) Use of energy analyses in silvicultural decision-making. Biomass Bioenergy. 30(4):357-362.

Nunes LJR, Causer TP, Ciolkosz D (2020) Biomass for energy: A review on supply chain management models. Renew Sustain Energy Rev. 120:109658.

Oro D, Lopes ES, Silva DA, Hillig E, Pelz SK (2018) Biomass energetic potential from timber harvesting at different times of storage. Floresta. 48(1):09-18.

Pal M, Bharati P (2019) The regression models with dummy explanatory variables. In: Applications of regression techniques. Springer, Singapore.

Rackl M, Günthner WA (2016) Experimental investigation on the influence of different grades of wood chips on screw feeding performances. Biomass Bioenerg. 88:106-115.

Raptis DI, Kazana V, Kazaklis A, Stamatiou C (2020)

Development and testing of volume models for *Pinus nigra* Arn., *Fagus sylvatica* L., and *Quercus pubescens* Willd. South Forests. 82(4):331-341.

Robinson KM, Karp A, Taylor G (2004) Defining leaf traits linked to yield in short-rotation coppice Salix. Biomass Bioenergy. 26(5):417-431.

Rodrigues CK, Lopes ES, Figueiredo A, Pelissari AL, Silva MK (2019) Modeling residual biomass from mechanized wood harvesting with data measured by forest harvester. An Acad Bras Cienc. 91(4):e20190194.

Salvador SM, Schumacher MV, Viera M, Stahl J, Consensa CB (2016) Biomassa e estoque de nutrientes em plantios clonais de *Eucalyptus saligna* Smith. em diferentes idades. Sci For. 44(110):311-321.

Sahoo K, Bilek E, Bergman R, Kizha AR, Mani S (2019) Economic analysis of forest residues supply chain options to produce enhanced-quality feedstocks. Biofuel Bioprod Biorefin. 13(3):514-534.

Santos R, Mello Júnior JA, Caraschi JC, Ventorim G, Pereia FA (2016) Polpação kraft e kraft/aq da madeira pré-hidrolisada de híbrido de *Eucalyptus urophylla* × *grandis*. Cienc Florest. 26(4):1281-1290.

Silveira P, Koehler HS, Sanquetta CR, Arce JE (2008) O estado da arte na estimativa de biomassa e carbono em formações florestais. Floresta. 38(1):185-206.

Somogyi Z, Cienciala E, Mäkipää R, Muukkonen P, Lehtonen A, Weiss P (2007) Indirect methods of large forest biomass estimation. Eur J For Res. 126(2):197-207.

Vargas-Larreta B, López-Sánchez CA, Corral-Rivas JJ, López-Martínez JO, Aguirre-Calderón CG, Álvarez-González JG (2017) Allometric equations for estimating biomass and carbon stocks in the temperate forests of north-western Mexico. Forests. 8(8):269.

Wolfsmayr UJ, Rauch P (2014) The primary forest fuel supply chain: A literature review. Biomass Bioenergy. 60:203-221.

Zhang D, Stenger A, Haroud PA (2015) Policy instruments for developing planted forests: Theory and practices in China, the U.S., Brazil, and France. J Econ. 21:223-237.