MIXED MODELS FOR NUTRIENTS PREDICTION IN SPECIES OF THE BRAZILIAN CAATINGA BIOME

Jadson Coelho de Abreu[2*], José Antônio Aleixo da Silva[3], Rinaldo Luiz Caraciolo Ferreira[3], Samuel José Silva Soares da Rocha[4], Ivaldo da Silva Tavares Júnior[5], Aline Araújo Farias[5], Paulo Henrique Villanova[5], Aguida Beatriz Travaglia Viana[5], Bruno Leão Said Schettini[6], Lucas Arthur de Almeida Telles[5] and Arthur Araújo Silva[6]

[1] Received on 06.05.2022 accepted for publication on 05.06.2023.
[2] Universidade do Estado do Amapá, Departamento de Engenharia Florestal, Macapá, AP - Brasil. E-mail: <jadsoncoelhoabreu@hotmail.com>.
[3] Universidade Federal Rural de Pernambuco, Departamento de Ciência Florestal, Recife, PE - Brasil. E-mail: <jaaleixo@uol.com.br> and <rinaldo.ferreira@ufrpe.br>.
[4] Universidade Federal de Lavras, Departamento de Ciências Florestais, Lavras, MG - Brasil. E-mail: <samuel.rocha@ufla.br>.
[5] Universidade Federal de Viçosa, Programa de Pós-Graduação em Ciência Florestal, Viçosa, MG - Brasil. E-mail: <ivaldo.junior@ufv.br>, <aline.a.araujo@ufv.br>, <paulo.villanova@ufv.br>, <aguida.vianaufv@gmail.com> and <lucasarthur36@gmail.com>.
[6] Universidade Federal de Viçosa, Departamento de Engenharia Florestal, Viçosa, MG - Brasil. E-mail: <blsschettini@gmail.com> and <arthur.araujo@ufv.br>.
*Corresponding author.

ABSTRACT – Nutrient prediction models applied to tree species from Brazilian Caatinga can be a crucial tool in understanding this biome. The study aimed to fit a mixed model to predict nitrogen (N), phosphorus (P), and potassium (K) content in tree species native to the Caatinga biome located in Floresta municipality, Pernambuco State – PE, Brazil. The following species were considered the area's most important and evaluated in the present study: *Poincianella bracteosa* (Tul.) L.P.Queiroz, *Mimosa ophtalmocentra* Mart. ex Benth, *Aspidosperma pyrifolium* Mart, *Cnidoscolus quercifolius* (Mull. Arg.) Pax. & Hoffm, and *Anadenanthera colubrina* var. cebil (Griseb.) Altschul. Four trees, representing the average circumference in each diameter class, were harvested for NPK quantification. The Spurr model was evaluated for NPK prediction, and species inclusion as a random effect was significant ($p < 0.05$) in all models. The Spurr model with fixed and random effects presented better statistics than fixed-effect models in all parameters for all nutrients. Generated NPK predicting equations can be a handy tool to understand the impact of wood extraction over Caatinga's biogeochemical cycles and guide forest management strategies in semi-arid regions of the world.

Keywords: Caatinga Biome; NPK; Fixed and Random Effects.

MODELOS MISTOS PARA PREDIÇÃO DE NUTRIENTES EM ESPÉCIES DO BIOMA CAATINGA, BRASIL

RESUMO – Modelos de predição de nutrientes aplicados a espécies arbóreas da Caatinga brasileira podem ser uma ferramenta crucial para a compreensão do bioma. O estudo teve como objetivo ajustar um modelo misto para prever os teores de nitrogênio (N), fósforo (P) e potássio (K) em espécies arbóreas nativas do bioma Caatinga localizadas no município de Floresta, Pernambuco – PE, Brasil. As seguintes espécies foram as mais importantes da área e avaliadas no presente estudo: ***Poincianella bracteosa*** *(Tul.) L.P.Queiroz,* ***Mimosa ophtalmocentra*** *Mart. ex Benth,* ***Aspidosperma pyrifolium*** *Mart,* ***Cnidoscolus quercifolius*** *(Mull. Arg.) Pax. & Hoffm e* ***Anadenanthera colubrina*** *var. cebil (Griseb.) Altschul. Quatro árvores, representando a circunferência média em cada classe de diâmetro, foram colhidas para quantificação de NPK. O modelo Spurr foi avaliado para predição de NPK e a inclusão de espécies como efeito aleatório foi significativa ($p < 0,05$) em todos os modelos. O modelo de Spurr com efeitos fixos e aleatórios apresentou estatísticas melhores que os modelos de efeito fixo em todos os parâmetros para todos os nutrientes. As equações de previsão de NPK geradas podem*

http://dx.doi.org/10.1590/1806-908820230000012

ser uma ferramenta útil para entender o impacto da extração de madeira sobre os ciclos biogeoquímicos da Caatinga e orientar estratégias de manejo florestal em regiões semiáridas do mundo.

Palavras-Chave: Bioma Caatinga; NPK; Efeitos Fixos e Aleatórios.

1. INTRODUCTION

Brazilian forest conservation is a priority due to its diversity (Soares-Filho et al., 2014), with the remaining 60% of forests covering the country and harboring much of the world forest species in different biomes (Oliveira et al., 2018). Brazilian Caatinga vegetation is one of the largest tropical dry forests remaining areas in the world (Miles et al., 2006) and a complex ecosystem characterized by high environmental variability (Moura et al., 2016).

In recent years, the population density increase has put pressure on the biome's natural resources and caused changes in land cover, mainly native vegetation; accurate information on land-use change in Caatinga is limited, but in 2009, the biome had 53.4% of the original vegetation cover remaining (Beuchle et al., 2015), being one of the most threatened ecosystems in the country (Arnan et al., 2018). Firewood extraction, pasture, and agricultural field settlements are the main human activities that affect its vegetation (Aguiar et al., 2014; Althoff et al., 2018).

Forest biomass is one of the main energy sources in the region, with 10 million m^3 of wood harvested in the year (Gariglio et al., 2010). In order to supply this energy demand, wood extraction intensifies impacts on the carbon cycle and nutrients (Moura et al., 2016; Althoff et al., 2018). Large nutrient amounts removal can lead to soil depletion and severe adverse effects over long-term productivity (Aquino et al., 2017; Gómez-García et al., 2016; Macedo et al., 2023; Yan et al., 2017). Understanding better the nutrient dynamics in these ecosystems, mainly nitrogen, phosphorus, and potassium, can help in wood harvesting management and provide greenhouse gas emissions and removals better estimates in the region (Althoff et al., 2018).

Nutrients predicting models are a crucial tool in understanding wood extraction impact over biogeochemical cycles in the Caatinga Biome, in addition to forest management strategies guiding (He et al., 2018). Studies with traditional models were developed in Brazil (Barbeiro et al., 2009; Abreu et al., 2016; Oliveira et al., 2018). However, the majority of the datasets utilized for biomass and nutrient modeling in tropical forests have heterogeneous structures, meaning samples in different sites with high species diversity (Miguel et al., 2013; Grau et al., 2017). These factors make traditional regression models present high error of estimates due to the forests' heterogeneity.

Mixed models can be a promising alternative to modeling heterogeneous environments. These models are often utilized to analyze data across a broad area spectrum (Groom et al., 2012; Hu et al., 2018; Poudel et al., 2018; Özkale and Kuran, 2018). Thus, this study aimed to fit a mixed model to predict nitrogen (N), phosphorous (P), and potassium (K) in native species from the Caatinga Biome.

2. MATERIAL AND METHODS

2.1. Study Area

The study was carried out in a 50 ha area (8°30´37" S and 37°59´07" W) with Caatinga vegetation, which is part of the 6,000 ha Itapemirim Farm, located in São Francisco, a mesoregion of Pernambuco State, Brazil.

The Floresta municipality is part of the Pajeú River watershed. According to the Köppen classification, the region's climate is classified as BSh (Hot semi-arid (steppe) climate). The average rainfall for the site is 503 mm, a rainy period from January to April, with an average annual temperature of 26.1 ºC. The municipality area is 3,643.97 km², and the altitude average is 323 m (Araújo Filho et al., 2001).

2.2. Dataset

Forest inventory was carried out by sampling, with 40 plots of 20 × 20 m (400 m^2) spaced 80 m apart, with 50 m of the border and a 6 cm circumference inclusion level at 1.30 m (CBH).

The following five species were selected as the most important ones, according to the Importance Value Index (IVI), based on information from prior forest inventory (Alves et al. 2017): *Poincianella bracteosa* (Tul.) L.P.Queiroz, *Mimosa ophtalmocentra* Mart. ex

Benth, *Aspidosperma pyrifolium* Mart, *Cnidoscolus quercifolius* (Mull. Arg.) Pax. & Hoffm, and *Anadenanthera colubrina* var. cebil (Griseb.) Altschul. Ten individuals per species were sampled for analysis.

2.3. Nutrient Quantification

Nutrient quantification analysis (NPK) in the aerial part was based on the diametric structure found in a new forest inventory. The five most important species were divided into five circumference classes with 3 cm amplitude, starting from a circumference at breast height (CBH) of 6 cm. Four trees representative of the average circumference at each class were harvested for aerial part nutrients analysis. Thus, 10 individuals per species were harvested, totaling 50 trees.

In order to cover diameter classes, individuals were chosen randomly, avoiding, though, partially harvested, burned, or fallen trees. The next step was to measure the chosen trees' CBH. Then, each CBH was converted in diameter at breast height (DBH). Then, total (Ht) and commercial (Hc) trunk heights were measured. Subsequent to dendrometric variable measurements, trunk, branches, and leaves were separated, and their samples were sent to laboratory analysis.

Total weight and wet weight samples obtained in the field were used to calculate dry biomass for each aerial component of the 50 sampled trees, using the expression below.

$$Bs = \frac{Pu(c) * Ps(a)}{Pu(a)} \qquad \text{Eq.1}$$

Where:

Bs = total dry biomass (Kg);

Pu(c) = total wet weight in the field (Kg);

Ps(a) = dry sample weight (Kg);

Pu(a) = wet sample weight (Kg).

The dry matter extracts for P and K analyses were obtained through wet digestion using HNO_3: HCl in proportion (2:1), while N was obtained through sulfuric digestion. Phosphorus (P) levels were analyzed by colorimetry with visible ultraviolet at 420 nm. Potassium (K) was determined by flame emission photometry technique.

The samples were divided among the three laboratories due to limitations in resources and equipment during the research. The nitrogen analyses were performed at the Plant Biochemistry laboratories of Universidade Federal Rural de Pernambuco, while the phosphorus and potassium analyses were conducted at the Laboratory of Organic Chemistry of the Department of Agronomy at Universidade Federal do Piauí in Bom Jesus-PI campus and Universidade Estadual de Londrina, respectively. Nutrient content was determined in g kg^{-1}, while the sampled trees' total nutrient amount was determined by multiplying concentration in g kg^{-1} by the dry biomass total.

2.4. Fitting Equations

The Spurr model (1952), in linear form, was fitted with green biomass, diameter, and total height data:

$$LnNPK = \beta_0 + \beta_1 \, Ln(DBH^2 \times Ht) \pm \varepsilon \qquad \text{Eq.2}$$

Where:

Ln = neperian logarithmic;

NPK = nutrients (nitrogen, phosphorous, and potassium) in kg;

DBH = diameter at breast height, in cm;

Ht = total height, in m;

β_0 and β_1 = model parameters;

$\varepsilon \sim N(0, \sigma 2)$ = random error.

The previous equation was fitted by the Maximum Likelihood Method, using the R programming language (R Core Team, 2014), specifically with the glm2 package. The fit evaluation was done by Akaike Information Criteria (AIC), correlation coefficient ($r_{y\hat{y}}$) between observed and predicted biomass, root mean square error (RMSE%), *bias*, and residual graphical analysis (Binoti et al., 2015).

Equations based on the Spurr model were adjusted considering the structure of mixed linear models, including intercepts and random slope coefficients, with species as a random effect. Mixed models, also known as mixed-effects models or hierarchical models, are a type of statistical model that incorporate both fixed and random effects in the analysis. In these models, fixed effects are used to explain the relationships between independent variables and the dependent variable, while random effects account for variation that is not explained by the fixed effects.

Equations regarding mixed models were fitted by Restricted Maximum Likelihood Method (REML) using the R programming language (R Core Team, 2014), specifically with the nlme package. The same selection criteria used for fixed models was applied to mixed ones. Random effect inclusion result on intercept and slope was verified by maximum likelihood ratio test (Resende et al., 2014), where the significance of differences (D) among deviations [-2log(L)] for models with and without random effect, was done comparing calculated and tabulated values, by χ^2 at 5% significance level. After mixed linear modeling, the resulting mixed model can be complete, partially complete, meaning random effects associated with only some parameters of the original model, or even a fixed-effect model, referring to non-significance of random effects.

Table 1 – Estimates of fixed-effects parameters for the Spurr model to predict native species NPK content regarding trees located in Floresta municipality, Pernambuco State, Brazil.
Tabela 1 *– Estimativas dos parâmetros de efeitos fixos do modelo Spurr para predizer o teor de NPK de espécies nativas, para árvores localizadas no município de Floresta, Pernambuco.*

Model	Nutrient	Model	Effect	β_0	β_1	p-valor
1	Nitrogen	$LnN=\beta_0+\beta_1\times Ln(dbh^2\times Ht)+\varepsilon_i$	Fixed	2.85399	0.35005	<0.0001
2	Phosphorous	$LnP=\beta_0+\beta_1\times Ln(dbh^2\times Ht)+\varepsilon_i$	Fixed	0.88364	0.31707	<0.0001
3	Potassium	$LnK=\beta_0+\beta_1\times Ln(dbh^2\times Ht)+\varepsilon_i$	Fixed	2.08847	0.23923	<0.0001
4	Nitrogen	$LnN=(\beta_0+a_i)+(\beta_1+b_{1i})\times Ln(dbh^2\times Ht)+\varepsilon_i$	Mixed	2.88710	0.34353	<0.0001
5	Nitrogen	$LnN=(\beta_0+a_i)+\beta_1\times Ln(dbh^2\times Ht)+\varepsilon_i$	Mixed	2.87613	0.34589	<0.0001
6	Nitrogen	$LnN=\beta_0+(\beta_1+b_{1i})\times Ln(dbh^2\times Ht)+\varepsilon_i$	Mixed	2.88710	0.34353	<0.0001
7	Phosphorous	$LnP=(\beta_0+a_i)+(\beta_1+b_{1i})\times Ln(dbh^2\times Ht)+\varepsilon_i$	Mixed	1.08482	0.27727	<0.0001
8	Phosphorous	$LnP=(\beta_0+a_i)+\beta_1\times Ln(dbh^2\times Ht)+\varepsilon_i$	Mixed	1.07884	0.28038	<0.0001
9	Phosphorous	$LnP=\beta_0+(\beta_1+b_{1i})\times Ln(dbh^2\times Ht)+\varepsilon_i$	Mixed	1.08482	0.27727	<0.0001
10	Potassium	$LnK=(\beta_0+a_i)+(\beta_1+b_{1i})\times Ln(dbh^2\times Ht)+\varepsilon_i$	Mixed	1.86127	0.28221	<0.0001
11	Potassium	$LnK=(\beta_0+a_i)+\beta_1\times Ln(dbh^2\times Ht)+\varepsilon_i$	Mixed	1.88175	0.27809	<0.0001
12	Potassium	$LnK=\beta_0+(\beta_1+b_{1i})\times Ln(dbh^2\times Ht)+\varepsilon_i$	Mixed	1.89594	0.27718	<0.0001

Where: Ln = neperian logarithmic; N = nitrogen, in kg; P = phosphorous, in kg; K = potassium, in kg; DBH = diameter at breast height, in cm; Ht = individual total height, in m; β0 to β2 = model fixed parameters; a_i = random intercept for i-th species; b_{1i} = random slope coefficient for i-th specie; $\varepsilon_i \sim N(0, \sigma 2)$ = random error.
Onde: Ln = logaritmo neperiano; N = nitrogênio, em kg; P = fósforo, em kg; K = potássio, em kg; DAP = diâmetro à altura do peito, em cm; Alt = altura total individual, em m; β0 até β2 = parâmetros fixos do modelo; a_i *= intercepto para a i-ésima espécie;* b_{1i} *= coeficiente de inclinação para a i-ésima espécie;* $\varepsilon_i \sim N(0, \sigma 2)$ *= erro aleatório.*

Table 2 – Random effects estimates regarding the Spurr equation to predict NPK content in native species located in Floresta municipality, Pernambuco State, Brazil.
Tabela 2 *– Estimativas de efeitos aleatórios da equação de Spurr para prever o teor de NPK em espécies nativas localizadas no município de Floresta, Pernambuco, Brasil.*

Species	Parameters					
	β0	β1	β0	β1	β0	β1
	Model 4		Model 5		Model 6	
1	1.12×10^{-09}	4.82×10^{-02}	2.35×10^{-02}	-	-	4.83×10^{-02}
2	-7.05×10^{-10}	1.96×10^{-03}	-1.56×10^{-02}	-	-	1.96×10^{-03}
3	-1.19×10^{-10}	-2.05×10^{-02}	-8.52×10^{-02}	-	-	-2.06×10^{-02}
4	-2.86×10^{-10}	-2.00×10^{-02}	-8.89×10^{-02}	-	-	-2.00×10^{-02}
5	3.69×10^{-12}	-9.65×10^{-03}	-4.49×10^{-02}	-	-	-9.65×10^{-03}
	Model 7		Model 8		Model 9	
1	3.35×10^{-02}	9.46×10^{-02}	5.31×10^{-01}	-	-	1.00×10^{-01}
2	-2.59×10^{-02}	3.64×10^{-02}	1.36×10^{-01}	-	-	3.20×10^{-02}
3	8.39×10^{-03}	-9.44×10^{-02}	-4.52×10^{-01}	-	-	-9.29×10^{-02}
4	-2.71×10^{-02}	-5.11×10^{-02}	-3.02×10^{-01}	-	-	-5.57×10^{-02}
5	1.11×10^{-02}	1.44×10^{-02}	8.63×10^{-02}	-	-	1.64×10^{-02}
	Model 10		Model 11		Model 12	
1	1.66×10^{-01}	4.84×10^{-02}	4.20×10^{-01}	-	-	7.62×10^{-02}
2	-3.58×10^{-01}	4.62×10^{-02}	-1.14×10^{-01}	-	-	-1.72×10^{-02}
3	8.54×10^{-02}	-6.15×10^{-02}	-2.17×10^{-01}	-	-	-4.81×10^{-02}
4	3.62×10^{-01}	1.90×10^{-02}	4.57×10^{-01}	-	-	8.36×10^{-02}
5	-2.55×10^{-01}	-5.21×10^{-02}	-5.45×10^{-01}	-	-	-9.45×10^{-02}

Where: 1 = *Poincianella bracteosa* (Tul.) L.P.Queiroz ; 2 =*Mimosa ophtalmocentra* Mart. ex Benth; 3= *Aspidosperma pyrifolium* Mart; 4= *Cnidoscolus quercifolius* (Mull. Arg.) Pax. & Hoffm; 5= *Anadenanthera colubrina* var. cebil (Griseb.) Altschul.
Onde: 1 = Poincianella bracteosa (Tul.) L.P.Queiroz; 2 = ***Mimosa ophtalmocentra*** *Mart. ex Benth; 3 = Aspidosperma pyrifolium Mart; 4 = Cnidoscolus quercifolius (Mull. Arg.) Pax. & Hoffm; 5 = Anadenanthera colubrina var. cebil (Griseb.) Altschul.*

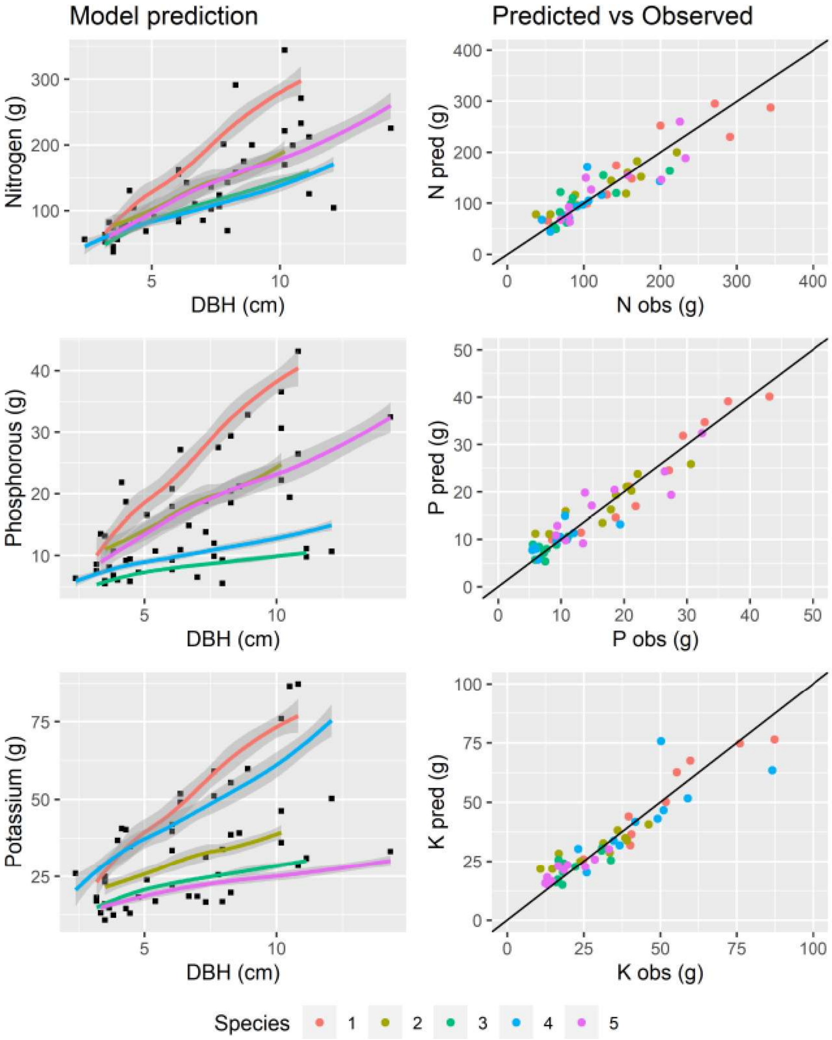

Figure 1 – Observed and predicted values for equations in mixed forms in Floresta municipality, Pernambuco State, Brazil.
Figura 1 – *Valores observados e previstos para equações em formas mistas no Município de Floresta, Pernambuco.*

3. RESULTS

Mixed models allow for the incorporation of both fixed and random effects, which can help to explain the sources of variability in the data and improve the accuracy of the results obtained. In their fixed or mixed forms, the Spurr model equations showed significant estimates for fixed effects parameters (Table 1).

Random coefficients considering the Spurr model structure were generated for each species to predict the NPK content in the region evaluated (Table 2).

Residuals showed adequate distribution along a straight line, with a mean around zero and constant variance. The hypothesis of homogeneity is not rejected concerning equations with random effects on DBH and Ht (Figure 1).

Species inclusion as a random effect was significant ($p < 0.05$) in all models according to the maximum likelihood ratio test. Thus, the final model showed fixed and random effects. (Table 3).

The AIC value for Model 10, which includes a random effect only in the slope of the height variable, was the lowest among all models tested (Table 4). This indicates that Model 10 is the best model for potassium analysis, Model 6 is the best model for

Table 3 – Maximum likelihood ratio test for equations that predict NPK in native species in Floresta municipality, Pernambuco State, Brazil.
Tabela 3 *– Teste de razão de máxima verossimilhança para equações de NPK em espécies nativas no Município de Floresta, Pernambuco.*

Model	Effect	MLE	Test	Ratio MLE
4	Mixed	-4.55232	1 vs 4	7.88577
5	Mixed	-5.81343	1 vs 5	5.36355
6	Mixed	-4.55232	1 vs 6	7.88577
1	Fixed	-8.49521		
7	Mixed	-6.28711	2 vs 7	45.55119
8	Mixed	-9.93831	2 vs 8	38.24879
9	Mixed	-6.28711	2 vs 9	45.55119
2	Fixed	-29.06271		
10	Mixed	-2.20533	3 vs 10	56.29369
11	Mixed	-3.68907	3 vs 11	53.32622
12	Mixed	-3.53774	3 vs 12	53.62888
3	Fixed	-30.35218		

Where: MLE: Maximum-Likelihood Estimation. The Spurr model with fixed and random effects showed the best statistics than fixed models for all parameters and nutrients.
Onde: MLE: Estimador de Máxima Verossimilhança. O modelo Spurr com efeitos fixos e aleatórios apresentou as melhores estatísticas em comparação com os modelos fixos para todos os parâmetros e nutrientes.

Table 4 – Precision statistics of the Spurr model in its fixed and mixed forms in Floresta municipality, Pernambuco State, Brazil.
Tabela 4 *– Estatísticas de precisão do modelo Spurr em suas formas fixa e mista no Município de Floresta, Pernambuco.*

Model	AIC	BIC	$r_{y\hat{y}}$	RSME%
1	22.99	28.72	0.8614	42.64
4	19.1	28.66	0.8982	30.96
5	19.62	27.27	0.8973	33.61
6	17.1	24.75	0.8982	30.96
2	64.12	69.86	0.6688	7.07
7	22.57	32.13	0.9126	3.01
8	27.87	35.52	0.9115	3.29
9	20.57	28.22	0.9126	3.01
3	66.7	72.44	0.5412	16.13
10	14.41	23.97	0.9379	6.81
11	15.37	23.07	0.9311	7.19
12	15.07	22.72	0.9319	7.28

Where: AIC= Akaike information criterion; BIC= Bayesian information criterion; $r_{y\hat{y}}$ = correlation coefficient; RSME% = root mean square error.
Onde: AIC = Critério de Informação de Akaike; BIC = Critério de Informação Bayesiano; ryy = coeficiente de correlação; RSME% = erro quadrático médio percentual.

nitrogen analysis, and Model 9 is the best model for phosphorus analysis.

4. DISCUSSION

Mixed-effects models offer a flexible and powerful tool for analyzing pooled data while estimating both fixed and random model parameters. The fixed effects are average values of the population similar to parameters obtained by ordinary least squares regression. Random effects can be estimated for each hierarchical level in a data set and various parameters in a model (Ou et al., 2016). These models are essential tools used in the forestry sector as they provide an adequate framework for assessing the growth and forests condition. Mixed models allow calibrations for a given location or tree and can provide individual and species-specific predictions (Miguel et al. 2013; Huff et al. 2018).

The all fixed effects parameters significance confirms the DBH and Ht inserting importance as model-predictive variables (Calegario et al., 2015). In a mixed model, if response variable information is available for a new species, random coefficients are estimated considering each species-specific response instead the population mean response. In the average population, the random coefficients vector of for a new individual has expected value equal to zero (Burkhart and Tomé, 2012).

The species-included significance as a random effect in all models indicates that this variable can be inserted as another tree NPK predictor (Garber and

Maguire 2003; Huff et al. 2018) in order to improve estimates precision. Statistics from mixed effect models were superior to fixed-effect models when predicting NPK in native species, which highlights the improvement due to random effect inclusion (Adame et al., 2008; Crecente-Campo et al., 2010; Ruslandi et al., 2017).

The residual distribution was considered adequate. Data outside the range were insignificant, since it is a small amount in relation to the sample size, not actively interfering with the model estimates (Gouveia et al., 2015). When a sample is available to estimate random effects, the performance of a mixed model is better than a fixed model (Temesgen et al., 2008). This statement is proved by residues of all equations with random effects that have a smaller amplitude than the equation in its fixed form.

These results are important because they suggest that the models used in the study are reliable and provide accurate estimates of the effects of the variables being analyzed. In particular, the fact that the residuals follow a straight line with a mean value close to zero suggests that the models are unbiased and that the random effects included in the equations effectively account for the variability in the data. Furthermore, the constant variance observed in the residuals indicates that the models are valid across the range of values of the predictor variables, suggesting that the relationships between the variables being studied are consistent throughout the dataset. This is important because it indicates that the results obtained from the models are likely to be robust and applicable to other similar datasets (Bates et al., 2015).

In the present study, improvement in NPK predictions due to random effect inclusion corroborates with the affirmation of Huff et al. (2018), in which the authors stated that species included as a random effect improve the estimates of mixed models compared to fixed ones. It is worth mentioning that other variables can be inserted as a random effect, such as forest type, region or site quality classes, precipitation, soil, elevation, among other geographical characteristics (Meng et al., 2007; Boubeta et al., 2015; Ou et al., 2016; Özçelik et al. 2018).

Morphological changes that occur between species, together with intraspecific differences caused by climatic and other environmental factors, require that individual equations are used to predict biomass in varied regions (Huff et al., 2018). Thus, the mixed model approach for species macronutrients modeling in the Caatinga biome is an alternative to obtain accurate predictions.

It is worth mentioning that new studies with environmental variables can be carried out and can improve the estimates. Mixed linear models provide a more flexible approach to analyze non-normal data when random effects are present. Finally, generated equations can support decision-making and guide politics towards better conservation practices in the Caatinga Biome.

5. CONCLUSION

Species inclusion as a random effect promoted an RMSE reduction of at least 4% in mixed models compared to fixed models. Thus, the proposed equations capture each species' effect and can be applied to better estimate NPK in trees from the Caatinga Biome.

The generated equations can be a handy tool to understand the impact of wood extraction over biogeochemical cycles of the Caatinga Biome and support forest management strategies in semi-arid regions of the world.

Overall, the use of mixed models in the study of tree nutrition in the Caatinga ecosystems can help provide a more comprehensive understanding of the complex relationships between nutrient availability, tree physiology, and ecosystem dynamics, ultimately contributing to the development of more effective and sustainable management strategies for these valuable and threatened ecosystems.

AUTHOR CONTRIBUTIONS

Conceptualization: Abreu JC, Silva JAA, Ferreira RLC, Rocha SJSS, Tavares Júnior IS; Experimental testing: Abreu JC, Rocha SJSS, Tavares Júnior IS, Farias AA, Villanova PH, Viana ABT, Schettini BLS, Telles LA; Statistical analysis: Abreu JC, Silva JAA, Ferreira RLC, Rocha SJSS; Results analysis: Abreu JC, Silva JAA, Ferreira RLC, Rocha SJS, Tavares Júnior IS, Farias AA, Villanova PH, Viana ABT, Schettini BLS, Telles LA; Writing, review, and editing: Abreu JC, Ferreira RLC, Rocha SJSS, Tavares Júnior

IS, Farias AA, Villanova PH, Viana ABT, Schettini BLS, Telles LA, Silva AA; Research supervision and coordination: Abreu JC, Silva JAA, Ferreira RLC.

6. REFERENCES

Abreu JC de, da Silva JAA, Ferreira RLC, Junior FTA. Ajuste de modelos matemáticos lineares e não lineares para estimativa de biomassa e nutrientes de *Anadenanthera colubrina* var. cebil no semiárido pernambucano. Scientia Forestalis. 2016; 44:739–750. https://doi.org/dx.doi.org/10.18671/scifor.v44n111.20

Adame P, del Río M, Cañellas I. A mixed nonlinear height-diameter model for pyrenean oak (*Quercus pyrenaica* Willd.). Forest Ecology and Management. 2008; 256:88–98. https://doi.org/10.1016/j.foreco.2008.04.006

Aguiar MI, Fialho JS, Campanha MM, Oliveira TS. Carbon sequestration and nutrient reserves under differen t land use systems. Revista Árvore. fevereiro de 2014;38:81–93. https://doi.org/10.1590/S0100-67622014000100008

Althoff TD, Menezes RSC, Pinto A de S, Pareyn FGC, Carvalho AL de, Martins JCR, Carvalho EX de, Silva ASA da, Dutra ED, Sampaio EV de SB. Adaptation of the century model to simulate C and N dynamics of Caatinga dry forest before and after deforestation. Agriculture, Ecosystems and Environment. 2018; 254:26–34. https://doi.org/10.1016/j.agee.2017.11.016

Alves AR, Ferreira RLC, Silva JAA da, et al. Conteúdo de nutrientes na biomassa e eficiência nutricional em espécies da Caatinga. Ciênc Florest. 2017; 27:377–390. https://doi.org/10.5902/1980509827686

Aquino DDN, Andrade EMD, Palácio HADQ, Pereira LR. Nutrient cycling and CO2 emissions in areas of preserved and thinned Caatinga. Revista Árvore. 2017;41. https://doi.org/10.1590/1806-90882017000300008

Araújo Filho JC de, Silva AB da, Silva F.B.R., Leite AP. Diagnóstico Ambiental do Município de Floresta, Pernambuco. 2001; https://www.embrapa.br/busca-de-publicacoes/-/publicacao/338510/diagnostico-ambiental-do-municipio-de-floresta-pernambuco [accessed 10.05.2021]

Arnan X, Arcoverde GB, Pie MR, Ribeiro-Neto JD, Leal IR. Increased anthropogenic disturbance and aridity reduce phylogenetic and functional diversity of ant communities in Caatinga dry forest. Science of the Total Environment. 2018; 631–632:429–438. https://doi.org/10.1016/j.scitotenv.2018.03.037

Barbeiro L da SS, Vieira G, Sanquetta CR. Equações para estimativa da biomassa individual de *Nectandra grandiflora* Ness (canela-amarela). FLORESTA. 2009; https://doi.org/10.5380/rf.v39i4.16318

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of statistical software, 67(1), 1-48.

Beuchle R, Grecchi RC, Shimabukuro YE, Seliger R, Eva HD, Sano E, Achard F. Land cover changes in the Brazilian Cerrado and Caatinga biomes from 1990 to 2010 based on a systematic remote sensing sampling approach. Applied Geography. 2015; 58:116–127. https://doi.org/10.1016/j.apgeog.2015.01.017

Binoti MLM da S, Leite HG, Binoti DHB, Gleriani JM. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. CERNE. 2015; 21:97–105. https://doi.org/10.1590/01047760201521011153

Boubeta M, Lombardía MJ, Marey-Pérez MF, Morales D. Prediction of forest fires occurrences with area-level Poisson mixed models. Journal of Environmental Management. 2015; 154:151–158. https://doi.org/10.1016/j.jenvman.2015.02.009

Burkhart HE, Tomé M. Modeling forest trees and stands, Springer Dordrecht Heidelberg, New York, London. 2012.

Calegario N, Daniels RF, Maestri R, Neiva R. Modeling dominant height growth based on nonlinear mixed-effects model: A clonal Eucalyptus plantation case study. Forest Ecology and Management. 2015; 204:11–20. https://doi.org/10.1016/j.foreco.2004.07.051

Crecente-Campo F, Tomé M, Soares P, Diéguez-Aranda U. A generalized nonlinear mixed-effects height-diameter model for Eucalyptus globulus L. in northwestern Spain. Forest Ecology and Management. 2010; 259:943–952. https://doi.org/10.1016/j.foreco.2009.11.036

Garber SM, Maguire DA. Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures. Forest Ecology and Management. 2003; 179:507–522. https://doi.org/10.1016/S0378-1127(02)00528-5

Gariglio MA, Sampaio EV de SB, Cestaro LA, Kageyama PY. Uso Sustentável e Conservação dos Recursos Florestais da Caatinga, Serviço Florestal Brasileiro, Brasília. 2010.

Gómez-García E, Diéguez-Aranda U, Cunha M, Rodríguez-Soalleiro R. Comparison of harvest-related removal of aboveground biomass, carbon and nutrients in pedunculate oak stands and in fast-growing tree stands in NW Spain. Forest Ecology and Management. 2016; 365:119–127. https://doi.org/10.1016/j.foreco.2016.01.021

Gouveia JF, Silva JAA da, Ferreira RLC, Gadelha FHL, Lima Filho LM de A. Modelos volumétricos mistos em clones de *Eucalyptus* no polo gesseiro do Araripe, Pernambuco. FLORESTA. 2015; 45:587–598. https://doi.org/10.5380/rf.v45i3.36844

Grau O, Peñuelas J, Ferry B, Freycon V, Blanc L, Desprez M, Baraloto C, Chave J, Descroix L, Dourdain A, Guitet S, Janssens IA, Sardans J, Hérault B. Nutrient-cycling mechanisms other than the direct absorption from soil may control forest structure and dynamics in poor Amazonian soils. Scientific Reports. 2017; 7:1–11. https://doi.org/10.1038/srep45017

Groom JD, Hann DW, Temesgen H. Evaluation of mixed-effects models for predicting Douglas-fir mortality. Forest Ecology and Management. 2012; 276:139–145. https://doi.org/10.1016/j.foreco.2012.03.029

He H, Zhang C, Zhao X, Fousseni F, Wang J, Dai H, Yang S, Zuo Q. Allometric biomass equations for 12 tree species in coniferous and broadleaved mixed forests, Northeastern China. PLoS ONE. 2018; 13:1–16. https://doi.org/10.1371/journal.pone.0186226

Hu M, Peñuelas J, Sardans J. Stoichiometry patterns of plant organ N and P in coastal herbaceous wetlands along the East China Sea : implications for biogeochemical niche. 2018; 273–288. https://doi.org/10.1007/s11104-018-3759-6

Huff S, Poudel KP, Ritchie M, Temesgen H. Quantifying aboveground biomass for common shrubs in northeastern California using nonlinear mixed effect models. Forest Ecology and Management. 2018; 424:154–163. https://doi.org/10.1016/j.foreco.2018.04.043

Macedo RS, Moro L, Lambais ÉO, Lambais GR, Bakker AP de. Effects of degradation on soil attributes under Caatinga in the Brazilian semi-arid. Rev Árvore. 23 de janeiro de 2023;47:e4702.

Meng Q, Cieszewski CJ, Madden M, Borders B. A linear mixed-effects model of biomass and volume of trees using Landsat ETM+ images. Forest Ecology and Management. 2007; 244:93–101. https://doi.org/10.1016/j.foreco.2007.03.056

Miguel S, Guzmán G, Pukkala T. A comparison of fixed- and mixed-effects modeling in tree growth and yield prediction of an indigenous neotropical species (Centrolobium tomentosum) in a plantation system. Forest Ecology and Management. 2013; 291:249–258. https://doi.org/10.1016/j.foreco.2012.11.026

Miles L, Newton AC, DeFries RS, Ravilious C, May I, Blyth S, Kapos V, Gordon JE. A global overview of the conservation status of tropical dry forests. In: Journal of Biogeography. 2006; pp 491–505

Moura PM, Althoff TD, Oliveira RA, Souto JS, Souto PC, Menezes RSC, Sampaio EVSB. Carbon and nutrient fluxes through litterfall at four succession stages of Caatinga dry forest in Northeastern Brazil. Nutrient Cycling in Agroecosystems. 2016 105:25–38. https://doi.org/10.1007/s10705-016-9771-4

Oliveira EV da S, Prata AP do N, Pinto A de S. Caracterização e atributos da vegetação herbácea em um fragmento de Caatinga no Estado de Sergipe, Brasil. Hoehnea. 2018; 45:159–172. https://doi.org/10.1590/2236-8906-70/2017

Ou G, Wang J, Xu H, Chen K, Zheng H, Zhang B, Sun X, Xu T, Xiao Y. Incorporating topographic factors in nonlinear mixed-effects models for aboveground biomass of natural Simao pine in Yunnan, China. Journal of Forestry Research. 2016; 27:119–131. https://doi.org/10.1007/s11676-015-0143-8

Özçelik R, Cao Q V, Trincado G, Göçer N. Predicting tree height from tree diameter and dominant height using mixed-effects and quantile regression models for two species in Turkey. Forest Ecology and Management. 2018; 419–420:240–248. https://doi.org/10.1016/j.foreco.2018.03.051

Özkale MR, Kuran Ö. Principal components regression and r-k class predictions in linear mixed models. Linear Algebra and Its Applications. 2018; 543:173–204. https://doi.org/10.1016/j.laa.2018.01.001

Poudel KP, Temesgen H, Gray AN. Estimating upper stem diameters and volume of Douglas-fir and Western hemlock trees in the Pacific northwest. Forest Ecosystems. 2018; 5:1–12. https://doi.org/10.1186/s40663-018-0134-2

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2014; http://www.R-project.org/

Resende MDV de, Silva FF, Azevedo CF. Estatística matemática, biométrica e computacional: Modelos mistos, multivariados, categóricos e generalizados (REML/BLUP), inferência bayesiana, regressão aleatória, seleção genômica, QTL-GWAS, estatística espacial e temporal, competição, sobrevivência, Suprema gráfica e Editora Ltda, Viçosa. 2014.

Ruslandi, Cropper WP, Putz FE. Tree diameter increments following silvicultural treatments in a dipterocarp forest in Kalimantan, Indonesia: A mixed-effects modelling approach. Forest Ecology and Management. 2017; 396:195–206. https://doi.org/10.1016/j.foreco.2017.04.025

Soares-Filho B, Rajao R, Macedo M, Carneiro A, Costa W, Coe M, Rodrigues H, Alencar A (2014) Cracking Brazil's Forest Code. Science. 2014; 344:363–364. https://doi.org/10.1126/science.1246663

Spurr, S.H. 1952. Forestry inventory. Ronald Press, New York. 476p.

Temesgen H, Monleon VJ, Hann DW. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. Canadian Journal of Forest Research. 2008; 38:553–565. https://doi.org/10.1139/X07-104

Yan T, Zhu J, Yang K, Yu L, Zhang J. Nutrient removal under different harvesting scenarios for larch plantations in northeast China: Implications for nutrient conservation and management. Forest Ecology and Management. 2017; 400:150–158. https://doi.org/10.1016/j.foreco.2017.06.004