

Nota Técnica

Utilização de variáveis qualitativas de localização em equações volumétricas e a seleção de modelos

Use of qualitative variables of location in volumetric equations and the selection of models

Vinicius Pizzo Ferreira^I 
João Luis Ferreira Batista^{II} 

^IPesquisador Autônomo, Ribeirão Preto, SP, Brasil

^{II}Universidade de São Paulo, Piracicaba, SP, Brasil

RESUMO

A predição de volume da floresta é essencial para diversas finalidades logísticas como o transporte, colheita entre outros fins. Portanto, a medição de tal variável é indispensável para um empreendimento florestal. Para tal medição, são utilizados modelos matemáticos para a quantificação da madeira em campo. O trabalho visa avaliar o impacto da incorporação da informação de localidade, na forma de variáveis indicadoras sobre a acurácia preditiva dos modelos de equações volumétricas, verificando se essa incorporação influencia a escolha do modelo mais apropriado. Ao todo, foram utilizados 11 modelos volumétricos divididos entre modelos de dupla entrada e locais. Para as variáveis qualitativas, foram utilizadas as informações de município, fazenda, estrato e talhão. Foram observados o coeficiente de determinação, erro padrão da estimativa e vários índices relativos aos resíduos, como o desvio absoluto médio, distância interquartil, amplitude de variação, entre outros. Além dos índices, foram analisados os gráficos de dispersão, quantil-quantil e boxplot. Para avaliar a heteroscedasticidade dos modelos, foi calculado o coeficiente de correlação de Spearman e os valores ajustados dos modelos. As variáveis qualitativas ajustadas com um intercepto diferente para cada variável dentro do modelo e em um segundo momento os modelos foram interagidos com as variáveis qualitativas.

Palavras-chave: Inventário florestal; Eucalipto; Interação

ABSTRACT

Forest volume prediction is essential for many logistical purposes, such as transportation, harvesting and others. Therefore, the measurement of such a variable is indispensable for a forest enterprise. For such measurement, mathematical models are used to quantify the wood in the field. The aim is to evaluate the impact of incorporating locality information as indicator variables on the predictive accuracy of volumetric equation models, and to verify whether this incorporation influences the choice of the most appropriate model. A total of 11 volumetric models were used, divided into double entry and local models. For the qualitative variables the information of city, farm, stratum and field was used. It was observed the coefficient of determination, standard error of the estimate and various indexes related to the residuals, such as the mean absolute deviation, interquartile distance, range of variation, among others. In addition to the indices, the scatter plots, quantile-quantile and boxplot were analyzed. To evaluate the heteroscedasticity of the models, the Spearman correlation coefficient and the adjusted values of the models were calculated. The qualitative variables adjusted with a different intercept for each variable within the model and in a second moment the models were interacted with the qualitative variables.

Keywords: Forest Inventory; *Eucalyptus*; Interaction

1 INTRODUÇÃO

A escolha do modelo adequado para a predição de volume do inventário florestal é importante para se obter resultados condizentes com a realidade. Portanto, é necessário o estudo do melhor modelo para a área do trabalho. Na literatura brasileira, os trabalhos de comparação entre os modelos em diferentes regiões e cenários são comuns, já que não existe um modelo superior em todas as situações (MCTAGUE; BATISTA; STEINER, 1989; MACHADO; CONCEIÇÃO; FIGUEIREDO, 2002; Andrade, 2017).

Verifica-se que os modelos mais utilizados para as florestas brasileiras são os modelos de Spurr, Schumacher-Hall, Meyer, Honer, Stoate, Naslund, Ogaya, Takata, Baden-Württemberg (BATISTA; COUTO; SILVA FILHO, 2014; Andrade, 2017), sendo que 41% dos estudos que comparam o uso de equações volumétricas indicam como superior o modelo de Schumacher-Hall, 14%, de Spurr, 17%, de Meyer e 28%, dos demais (Andrade, 2017). Esses resultados indicam que o modelo mais apropriado para a predição volumétrica varia entre as espécies arbóreas ou clones, entre as áreas e entre as situações de aplicação da equação volumétrica.

Uma forma de incorporar nos modelos empíricos a variabilidade causada por situações qualitativamente distintas, como diferentes espécies ou áreas, é a introdução das variáveis qualitativas na forma de variáveis indicadoras (*dummy variables*), o que pode aumentar a acurácia da predição volumétrica. Contudo, as pesquisas realizadas no Brasil têm focado mais na comparação entre os modelos (REZENDE *et al.*, 2006; THOMAS *et al.*, 2006; MIGUEL *et al.*, 2010; SOARES *et al.*, 2011; SALES *et al.*, 2015; ANDRADE, 2017), poucas levam em conta a utilização de variáveis indicadoras para aprimorar o desempenho preditivo dos modelos (MCTAGUE; BATISTA; STEINER, 1989). Sabe-se que a utilização de variáveis indicadoras aumenta a acurácia das predições de equações empíricas, levando em conta diferentes regiões, e também auxilia na verificação de diferenças de crescimento nas diferentes fazendas (SCHNEIDER; TONINI, 2003).

Dessa maneira, o objetivo deste trabalho foi avaliar o impacto da incorporação da informação de localidade (município, fazenda, estrato, talhão), na forma de variáveis indicadoras sobre a acurácia preditiva dos modelos de equações volumétricas, verificando se essa incorporação influencia a escolha do modelo mais apropriado.

2 MATERIAL E MÉTODOS

Os dados utilizados foram obtidos de plantios de eucalipto na região central do estado de São Paulo, nos municípios de Bofete, Botucatu, Itatinga e Salto. Os municípios de Itatinga e Botucatu estão localizados em regiões mais altas, respectivamente 800 e 806 m, o que auxilia nas temperaturas médias serem menores, com precipitações de 1308 e 1324 mm, respectivamente. Já os municípios de Bofete e Salto estão em elevações menores, 520 e 560 m, respectivamente, com médias térmicas mais elevadas e precipitações de 1236 e 1262 mm, respectivamente (CLIMATE-DATA.ORG, 2018). Essas diferenças de elevação influenciam no clima, que podem levar a mudanças nas formações dos solos. Em Itatinga (GONÇALVES *et al.*, 2012), os solos apresentam uma concentração de argila mais elevada que os solos em Bofete, onde as principais

rochas formadoras são os arenitos, com pequenas porções de basalto (INSTITUTO DE PESQUISAS TECNOLÓGICAS, 2009).

O conjunto de dados utilizado compreende uma amostra destrutiva de 1,482 fustes de plantios de *Eucalyptus grandis* em primeira rotação, nos quais foram determinados: a idade, o diâmetro a altura do peito (DAP), a altura total e os volumes total e comercial. Os volumes foram determinados por cubagem rigorosa dos fustes. No conjunto de dados, a localização de cada fuste é definida por quatro variáveis nominais que compõem uma estrutura hierárquica para a localização dos fustes: município, fazenda, estrato e talhão. O “município” identifica as unidades político-administrativas do estado de São Paulo, enquanto que a “fazenda” identifica as propriedades rurais dos plantios dentro dos municípios. Já o “estrato” subdivide os plantios dentro de cada “fazenda” em subáreas de maior homogeneidade em termos de data de plantio e sistema de implantação e manejo, enquanto que a variável “talhão” identifica a “unidade de produção” dentro dos “estratos”, isto é, a unidade que delimita a menor área da floresta para fins de inventário e gerenciamento florestal. Essa estrutura hierárquica é comum nos plantios florestais brasileiros, que podem ter mais ou menos níveis hierárquicos que os quatro níveis que esse conjunto de dados apresenta. A Tabela 1 apresenta o número de categorias para cada um dos níveis hierárquicos.

Tabela 1 – Número de municípios, fazendas, estratos, talhões e árvores que compõem o conjunto de dados

Municípios	Fazendas	Estratos	Talhões	Árvores
Bofete	4	26	50	593
Botucatu	3	11	33	240
Itatinga	5	7	10	165
Salto	5	19	31	484
4	17	63	124	1482

Fonte: Autores (2019)

O estudo utilizou equações volumétricas de dupla entrada e equações locais. As de dupla entrada foram Spurr, Schumacher-Hall, Baden, Meyer e Stoate, já as locais foram Husch, Baden Local, Área Transversal, Parabólico, Schumacher Local e Brenac, que podem ser conferidas na Tabela 2. Os modelos foram escolhidos a partir de uma revisão bibliográfica e identificados os modelos mais utilizados para a predição de volume nas florestas brasileiras, sendo observados em Rolim *et al.* (2006), Thomas *et al.* (2006), Azevedo *et al.* (2011), Miguel e Leal (2012), Batista, Couto e Silva Filho (2014) e Sales *et al.* (2015).

Tabela 2 – Modelos volumétricos de dupla entrada e local, respectivamente, utilizados na estimação do volume dos fustes

Modelo	Fórmula
Spurr	$v_i = \beta_0 + \beta_1(d_i^2 * h_i) + \varepsilon_i$
Schumacher-Hall	$\ln(v_i) = \beta_0 + \beta_1 * \ln(d_i) + \beta_2 * \ln(h_i) + \varepsilon_i$
Baden	$\ln(v_i) = \beta_0 + \beta_1 * \ln(d_i) + \beta_2 * \ln^2(d_i) + \beta_3 * \ln(h_i) + \beta_4 * \ln^2(h_i) + \varepsilon_i$
Meyer	$v_i = \beta_0 + \beta_1 * d_i + \beta_2 * d_i^2 + \beta_3(d_i * h_i) + \beta_4(d_i^2 * h_i) + \beta_5 * h_i + \varepsilon_i$
Stoate	$v_i = \beta_0 + \beta_1 * d_i^2 + \beta_2(d_i^2 * h_i) + \beta_3 * h_i + \varepsilon_i$
Husch	$\ln(v_i) = \beta_0 + \beta_1 * \ln(d_i) + \varepsilon_i$
Baden Local	$\ln(v_i) = \beta_0 + \beta_1 * \ln(d_i) + \beta_2 * \ln^2(d_i) + \varepsilon_i$
Área Transversal	$v_i = \beta_0 + \beta_2 * d_i^2 + \varepsilon_i$
Parabólico	$v_i = \beta_0 + \beta_1 * d_i + \beta_2 * d_i^2 + \varepsilon_i$
Schumacher Local	$\ln(v_i) = \beta_0 + \beta_1(1/d_i) + \varepsilon_i$
Brenac	$\ln(v_i) = \beta_0 + \beta_1 * \ln(d_i) + \beta_2(1/d_i) + \varepsilon_i$

Fonte: Autores (2019)

A incorporação da informação de localização dos fustes nas equações volumétricas foi implementada introduzindo-se variáveis indicadoras (variáveis *dummy*) nos modelos estudados. Cada um dos quatro níveis da estrutura hierárquica de localização do conjunto de dados foi utilizado como uma “*variável de localização*” a ser testada. Logo,

foram testadas quatro variáveis de localização que formam uma sequência crescente de especificidade de localização: município, fazenda, estrato e talhão.

Testou-se duas formas de introdução de cada variável de localização nos modelos de equação volumétricas: "*intercepto*" e "*interação*". Na forma de intercepto, somente o intercepto do modelo é alterado em função da variável de localização. Por exemplo, se a variável de localização sendo testada na forma de intercepto é o "município", as equações volumétricas terão um valor de intercepto para cada um dos municípios (quatro), mas caso a variável de localização seja o "talhão", tem-se um intercepto para cada um dos talhões (124). Na forma de interação, todos os coeficientes de regressão dos modelos são específicos para cada categoria da variável de localização, inclusive o intercepto. Ou seja, têm-se um modelo completo específico para cada categoria da variável de localização. Por exemplo, se a variável é o município, então têm-se quatro modelos, um para cada município, cada qual com seus coeficientes, mas se a variável é o talhão, então são 124 modelos com diferentes coeficientes específicos para cada um dos talhões.

Para a comparação das equações, foram utilizados o coeficiente de determinação ajustado (R^2 Ajustado), erro padrão da estimativa (PSE) e vários índices relativos aos resíduos: o desvio absoluto médio (MAD), a distância interquartil (IQD), a amplitude de variação (Range), o índice de assimetria de Galton (Gskew), o índice de assimetria de Pearson (PSkew) e os gráficos de dispersão, quantil-quantil (Q-Q) normal e boxplot. Para avaliar a heterocedasticidade dos modelos, foi calculado o coeficiente de correlação de Spearman entre a raiz quadrada do valor absoluto dos resíduos padronizados e os valores ajustados (fitted values) dos modelos. Todos os índices foram calculados na escala original da variável resposta, isto é, do volume das árvores.

O R^2 representa a proporção da variabilidade do volume das árvores que é explicada pelo modelo. Já o erro padrão da estimativa expressa a precisão do ajuste dos modelos utilizados, apresentando os erros padrões, quando estão em mesma unidade

de medida. O desvio absoluto médio calcula a distância entre o volume observado e volume predito, assim representando a variabilidade dos resíduos, esta análise é utilizada quando não se deseja a influência dos valores extremos (TIBCO SOFTWARE, 2013). A distância interquartil é a distância entre o primeiro e o terceiro quartil, sendo assim os valores dentro desta diferença somam 50% dos resíduos observados, como medida de variabilidade, ela é utilizada no cálculo de outros índices. A diferença entre o maior e menor valor é conhecida como amplitude de variação, assim mostrando, no nosso caso, a distância entre os valores extremos de resíduo.

O primeiro índice de assimetria tratado é o de Galton, que é dado pelo desvio dos quartis em relação à mediana, sendo construído para comparar conjuntos de dados em diferentes escalas, valores negativos indicam assimetria à esquerda, enquanto valores positivos indicam assimetria à direita. Já o índice de assimetria de Pearson é dado pela diferença entre a média e a mediana pelo desvio padrão, comportando-se da mesma forma que o índice de Galton.

O coeficiente de correlação de Spearman avalia a relação entre duas variáveis, suas mudanças conjuntas, porém não necessariamente à mesma taxa, variando de -1 a 1, com a aproximação desses valores a correlação é mais acentuada, valores negativos indicam uma correlação inversa.

Os gráficos de dispersão dos resíduos foram analisados para a observação de tendências nas equações de regressão utilizadas que indicam a possibilidade de predição viciada para árvores de diferentes tamanhos. O gráfico Q-Q foi utilizado para uma averiguação visual da proximidade da distribuição dos resíduos para a distribuição normal, porém não se pode afirmar apenas com base no gráfico a natureza da distribuição (FORD, 2015). Já o boxplot foi utilizado para a observação das distribuições e valores atípicos presentes dos resíduos (VALLADARES NETO *et al.*, 2017), além de facilitar a comparação entre os talhões e fazendas em termos da amplitude e distribuição dos resíduos dos modelos.

3 RESULTADOS E DISCUSSÃO

Como apresentado na Tabela 1, o número de fustes por município, fazenda, estrato e talhão foi bastante variável. A distribuição dos fustes da amostra como um todo, segundo as classes de DAP e altura total, é apresentada na Tabela 3. Nota-se que os mínimos e máximos para altura são 5 e 35 m, já para o DAP são 3 e 27 cm. Observando a Tabela percebe-se que a maioria das plantas se encontram entre 9 e 23 m de altura e DAP de 5 a 17 cm. Nota-se que a amostra é mais esparsa entre as grandes árvores, isto é, nas últimas classes de DAP e altura.

Tabela 3 – Número de fustes por classe de DAP (cm) e altura total (m) presentes na amostra destrutiva que constitui o conjunto de dados estudado

	Altura (m)													Total		
	(5,7]	(7,9]	(9,11]	(11,13]	(13,15]	(15,17]	(17,19]	(19,21]	(21,23]	(23,25]	(25,27]	(27,29]	(29,31]		(31,33]	(33,35]
(3,5]	2	16	17	1												36
(5,7]	1	6	64	62	24	1										158
(7,9]			25	70	71	53	8	2								229
(9,11]			7	43	88	103	69	31	4							345
(11,13]			1	10	48	51	125	58	34	5						332
DAP (13,15]					7	15	49	75	53	18	7	1				225
(cm) (15,17]						1	2	15	18	15	13	8				72
(17,19]								1	10	7	11	8	2			39
(19,21]								1	2	2	7	5	4		1	22
(21,23]										2	2	4	3	3	1	15
(23,25]											1	3	1	1		6
(25,27]											1			2		3
Total	3	22	114	186	238	224	253	183	121	49	42	29	10	6	2	1482

Fonte: Autores (2019)

3.1 Modelos de dupla entrada

Os resultados obtidos por meio dos índices selecionados podem ser observados nas Tabelas 4 e 5, as equações de dupla entrada foram organizadas pelas variáveis qualitativas utilizadas para o ajuste das mesmas. Os modelos gerais (sem

variável qualitativa), Tabelas 4 e 5, obtiveram bons resultados, com coeficientes de determinação altos e erros padrão da estimativa relativamente baixos. O modelo de Meyer é o que apresentou melhores resultados dentre os modelos gerais, e o que apresentou os resultados menos expressivos foi o modelo de Schumacher-Hall. As equações apresentaram resultados de coeficiente de determinação, erro padrão, desvio absoluto médio e amplitude relativamente próximos entre si, com pequenas variações entre cada um. Já quando observados os índices de assimetria, os modelos apresentaram uma variação nos resultados, mas sem apresentar altas assimetrias.

Tabela 4 – Tabela de índices utilizados na comparação das equações de dupla entrada em seu modelo geral e na presença de variáveis qualitativas, com interceptos variados

Modelo	N	P	R2 Ajustado	PSE	MAD	Range	GSkew	PSkew	PCor	
Spurr	1482	2	0,994	7,17	3,75	129,88	16	20,3	56,1	
Schumacher-Hall	1482	3	0,993	7,43	3,72	126,63	0,5	22,5	62,3	
Baden	1482	5	0,994	7,28	3,64	134,07	2,8	1,3	61,4	Geral
Meyer	1482	6	0,994	6,93	3,72	119,99	4,2	1,1	58,1	
Stoate	1482	4	0,994	7,01	3,7	123,4	-9,4	-3,7	57	
Spurr	1482	5	0,994	7,12	3,68	129,79	15,1	22	55,1	
Schumacher-Hall	1482	6	0,994	7,3	3,58	118,5	3,6	20,6	63,2	
Baden	1482	8	0,993	7,36	3,42	136,7	4,7	-0,8	62	Município
Meyer	1482	9	0,994	6,88	3,74	117,87	3,1	-1,6	55,8	
Stoate	1482	7	0,994	6,95	3,67	123,08	-5,7	-5,1	56,6	
Spurr	1482	18	0,994	6,92	3,98	128,97	4,6	11,7	47,4	
Schumacher-Hall	1482	19	0,994	6,8	3,33	117,94	-2,9	14,5	60	
Baden	1482	21	0,994	7,1	3,25	133,47	-0,7	-6	60,3	Fazenda
Meyer	1482	22	0,995	6,62	3,79	113,03	-1	-4,7	51,1	
Stoate	1482	20	0,995	6,7	3,82	121,33	-7,4	-5,2	50,6	
Spurr	1482	64	0,996	6,5	3,7	127,25	-9,1	-7,9	44,4	
Schumacher-Hall	1482	65	0,996	6,28	2,9	117,33	-3	12,3	57,7	
Baden	1482	66	0,995	6,26	2,92	125,6	-2,9	-4	58,4	Estrato
Meyer	1482	67	0,995	6,33	3,62	114,61	-3,3	-7,4	44,5	
Stoate	1482	66	0,995	6,37	3,59	121,14	-8,8	-8,6	42,1	

Continua ...

Tabela 4 – Conclusão

Modelo	N	P	R2 Ajustado	PSE	MAD	Range	GSkew	PSkew	PCor	
Spurr	1482	125	0,996	6,34	3,3	128,25	-7,1	-7,8	40,6	
Schumacher-Hall	1482	126	0,996	6,27	2,65	123,59	-5,9	11	55,8	
Baden	1482	128	0,996	6,15	2,71	130,17	-2,1	-3,9	55,7	Talhão
Meyer	1482	129	0,996	6,15	3,34	113,99	-6,3	-8,1	39,2	
Stoate	1482	127	0,996	6,22	3,4	122,21	-7,2	-8,5	36,7	

Fonte: Autores (2019)

Em que: N (número de indivíduos), P (número de parâmetros), R2 Ajustado (coeficiente de determinação), PSE (erro padrão da estimativa), MAD (desvio absoluto médio), Range (amplitude de variação), GSkew (índice de assimetria de Galton), PSkew (índice de assimetria de Pearson), PCor (coeficiente de correlação de Spearman).

Tabela 5 – Tabela de índices utilizados na comparação das equações de dupla entrada em seu modelo geral e na presença de variáveis qualitativas, com interação destas variáveis

Modelos	N	P	R2	PSE	MAD	Range	GSkew	PSkew	PCor	
Spurr	1482	2	0,99	7,17	3,75	129,88	16	20,3	56,1	
Schumacher-Hall	1482	3	0,99	7,43	3,72	126,63	0,5	22,5	62,3	
Baden	1482	5	0,99	7,28	3,64	134,07	2,8	1,3	61,4	Geral
Meyer	1482	6	0,99	6,93	3,72	119,99	4,2	1,1	58,1	
Stoate	1482	4	0,99	7,01	3,7	123,4	-9,4	-3,7	57	
Spurr	1482	8	0,99	7,11	3,81	128,94	9,7	16,9	53,6	
Schumacher-Hall	1482	12	0,99	7,4	3,61	119,36	3,3	20,3	62,9	
Baden	1482	10	0,99	7,09	3,37	131,23	5,3	2,6	62,2	Município
Meyer	1482	24	1	6,71	3,75	117,49	5,5	-0,5	57,8	
Stoate	1482	16	0,99	6,85	3,76	117,96	-0,3	-0,8	56,2	
Spurr	1482	34	1	6,67	3,44	135,32	6,7	15,5	54,5	
Schumacher-Hall	1482	51	1	6,55	3,22	111,88	2,4	11	60,3	
Baden	1482	85	1	6,87	3,08	152,85	5,9	-2,3	59,5	Fazenda
Meyer	1482	102	1	6,23	3,31	104,08	-1,8	-5,3	54	
Stoate	1482	68	1	6,38	3,42	125,19	-2,6	-5,4	52,5	

Continua ...

Tabela 5 – Conclusão

Modelos	N	P	R2	PSE	MAD	Range	GSkew	PSkew	PCor	
Spurr	1482	126	1	6,16	3,27	123,52	0,3	3,8	48,2	
Schumacher-Hall	1482	189	1	6,27	2,75	128,94	2,9	3,6	58,8	
Baden	1482	315	1	5,68	2,42	80,17	8,8	4,2	53,6	Estrato
Meyer	1482	378	1	5,44	2,62	66,16	1,2	-0,5	45,1	
Stoate	1482	252	1	5,83	2,89	103,65	-2,5	-2,3	46,3	
Spurr	1482	248	1	5,96	2,87	122,58	-5,8	0	43,6	
Schumacher-Hall	1482	372	1	6,67	2,44	155,97	4,6	4	53,3	
Baden	1482	620	1	5,45	1,71	70,71	4,7	4,4	39,5	Talhão
Meyer	1482	744	1	5,06	1,79	52,72	3,5	0	25,3	
Stoate	1482	496	1	5,43	2,34	77,9	-0,2	0	34,4	

Fonte: Autoria própria (2019)

Em que: N (número de indivíduos), P (número de parâmetros), R2 Ajustado (coeficiente de determinação), PSE (erro padrão da estimativa), MAD (desvio absoluto médio), Range (amplitude de variação), GSkew (índice de assimetria de Galton), PSkew (índice de assimetria de Pearson), PCor (coeficiente de correlação de Spearman).

Todos os modelos tiveram significativa melhora quando as variáveis de localização foram introduzidas e o detalhamento aumentou, com a especificação de município para talhão. Tal melhora se mostra mais marcante no erro padrão da estimativa, porém de modo mais sutil nas outras medidas de qualidade de ajuste.

A introdução das variáveis de localização como intercepto alterou pouco o ranqueamento dos modelos, com base no erro padrão da estimativa, porém diminuiu a diferença entre os modelos. O modelo que apresentou melhor resposta no desempenho, erro padrão da estimativa, foi o Schumacher-Hall, seguido do Baden. Já os modelos menos sensíveis foram o Meyer e o Stoate.

Quando se observa a assimetria, todos os modelos mudaram de uma leve assimetria à direita, para uma leve assimetria à esquerda com a introdução das variáveis de localização como intercepto. A exceção foi o modelo de Stoate, que teve sua assimetria à esquerda mantida com a introdução da variável.

Quando utilizadas as variáveis de localização em interação com os modelos, houve uma maior melhora na performance (em termos de erro padrão da estimativa) quando comparada com a introdução como intercepto, porém com a exceção do modelo Schumacher-Hall. O modelo com a maior melhora foi o Meyer, seguido do Baden.

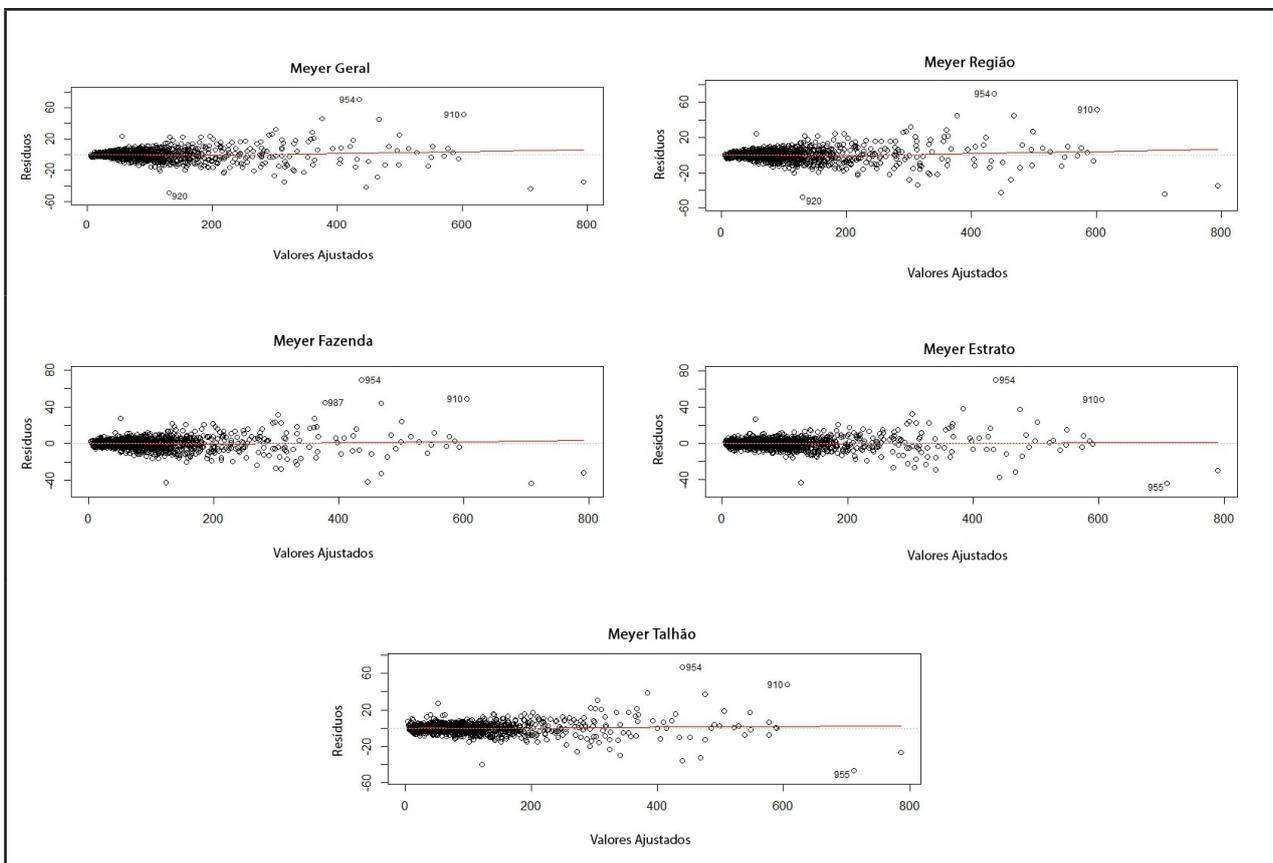
A introdução das variáveis de localização como interação causou uma melhora clara nos outros índices como desvio absoluto médio, amplitude de variação e de modo mais sutil no coeficiente de determinação. Porém, não apresentou mudanças na simetria dos resíduos. Quando observado o índice de assimetria de Galton, todos os modelos permaneceram com leve assimetria à direita, tendo novamente o modelo Stoate como exceção, com sua assimetria à esquerda sendo reduzida a ponto do modelo se tornar praticamente simétrico. Já o índice de assimetria de Pearson indica uma clara redução da assimetria dos modelos Spurr, Schumacher-hall e Stoate e uma leve piora na assimetria à direita do Meyer e Baden.

Quando considerada a passagem gradativa das variáveis de localização de município a fazenda, desta a estrato e deste a talhão, os índices de assimetria tiveram um comportamento errático. Com exceção do modelo Schumacher-Hall, que teve seu melhor desempenho em termos de erro padrão da estimativa na introdução das variáveis de localização como intercepto, todos os demais modelos apresentaram o menor erro padrão da estimativa, desvio absoluto médio e amplitude de variação e maior coeficiente de determinação na introdução da variável de localização "talhão" como interação. Avaliando-se individualmente todos os modelos estudados, verifica-se que a melhora de performance com a introdução da variável "talhão", seja como intercepto, seja como interação, foi maior que a diferença dos modelos entre si sem as variáveis de localização.

Os modelos com as variáveis de localização apresentaram resultados próximos, porém na observação dos gráficos de cada equação se pode constatar que, mesmo com coeficientes de determinação e erro padrão da estimativa próximos, os modelos apresentam diferenças, principalmente na amplitude de dispersão dos resíduos.

Como exemplo, a Figura 1 apresenta os gráficos de dispersão dos resíduos do modelo Meyer do ajuste geral e dos ajustes das variáveis de localidade (na forma de interação). O padrão de redução da amplitude da dispersão dos resíduos observado para esse modelo se repetiu em todos os modelos. Os gráficos de dispersão dos resíduos demonstraram os resultados analisados das Tabelas, sendo que os melhores gráficos eram das equações ajustadas com a variável “talhão”. Os gráficos Q-Q normal também indicaram os modelos com a variável “talhão” como os mais adequados, pois apresentaram menores assimetrias que os modelos com as demais variáveis de localização (município, fazenda e estrato). Entretanto, os gráficos dos modelos com qualquer uma das variáveis de localização sempre se mostraram superiores aos modelos sem variáveis de localização.

Figura 1 – Gráficos de dispersão (Resíduos x Valores Ajustados) do modelo de Meyer geral e Meyer com interação de região, fazenda, estrato e talhão



Fonte: Autores (2019)

3.2 Modelos locais

Os modelos locais apresentaram resultados piores em comparação aos modelos de dupla entrada, porém com a introdução das variáveis de localização houve melhoras marcantes (Tabelas 6 e 7). Os modelos gerais apresentaram resultados de coeficiente de determinação e erro padrão com grandes diferenças entre si, sendo que existem modelos com erro padrão menores que 20 e um deles com erro padrão maior que 50. Tais diferenças continuam presentes nos demais índices analisados, sendo Schumacher local o pior modelo, pois seus índices apresentaram os piores resultados, e o modelo Parabólico foi o modelo com melhor desempenho.

Tabela 6 – Tabela de índices utilizados na comparação das equações locais em seu modelo geral e na presença de variáveis qualitativas, com interceptos variados

Modelos	N	P	R2	PSE	MAD	Range	GSkew	PSkew	PCor	
Husch	1482	2	0,962	17,61	10,31	228,08	0,3	25,7	58,8	Geral
Baden Local	1482	3	0,961	17,95	10,59	273,11	2,9	4,6	59,6	
Área Transversal	1482	2	0,948	20,66	17,74	240,96	0,1	1,7	14,6	
Parabólico	1482	3	0,967	16,51	10,56	223,24	3,1	3	50,7	
Schumacher Local	1482	3	0,686	50,75	13,48	527,12	-7,4	68,9	62,1	
Brenac	1482	3	0,963	17,38	10,48	260,97	4,7	8,5	58,8	
Husch	1482	5	0,962	17,62	10,04	272,63	3,1	25,8	60,1	Município
Baden Local	1482	6	0,96	18,25	10,12	320,34	2,7	4,7	59,6	
Área Transversal	1482	5	0,965	20,27	17,17	241,33	3,1	3,6	12,1	
Parabólico	1482	6	0,967	16,38	10,81	223,21	-6,5	-7,4	47,6	
Schumacher Local	1482	6	0,707	49,07	13,59	511,48	3,8	67,8	62,9	
Brenac	1482	6	0,962	17,71	10,17	307,74	0,8	6,6	59,3	
Husch	1482	18	0,97	15,77	7,72	235,35	5,4	18,1	61,8	Fazenda
Baden Local	1482	19	0,97	15,73	7,63	255,63	9,5	10,4	61,4	
Área Transversal	1482	18	0,954	19,26	16,46	238,8	-1,6	3,1	14,3	
Parabólico	1482	19	0,972	15,29	9,92	214,92	-13,8	-18,8	41,9	
Schumacher Local	1482	19	0,764	43,98	10,68	544,25	-5,9	62,9	62,6	
Brenac	1482	19	0,971	15,65	7,62	248,26	9,7	12,4	61,4	

Continua ...

Tabela 6 – Conclusão

Modelos	N	P	R2	PSE	MAD	Range	GSkew	PSkew	PCor	
Husch	1482	64	0,979	13,25	5,15	268,94	-2,2	13,6	56,2	Estrato
Baden Local	1482	65	0,978	13,3	5,14	269,89	-4,9	13,1	55,9	
Área Transversal	1482	64	0,961	17,77	14,02	243,51	-3,7	6,9	7,6	
Parabólico	1482	65	0,976	14,01	8,84	218,5	-15	-23,6	31,3	
Schumacher Local	1482	65	0,805	39,96	7,21	484,88	0,2	61,7	58,9	
Brenac	1482	65	0,978	13,33	5,15	270,31	-4,6	14,3	56,4	
Husch	1482	125	0,984	11,55	4,53	231,53	1,7	14,3	56,8	Talhão
Baden Local	1482	126	0,984	11,59	4,53	232,27	1,8	15,1	56,5	
Área Transversal	1482	125	0,965	16,97	12,6	221,04	-6,1	0,4	0,5	
Parabólico	1482	126	0,979	13,02	8,14	194,95	-13,3	-27,4	24,6	
Schumacher Local	1482	126	0,819	38,57	6,7	460,2	3,4	59,4	58,5	
Brenac	1482	126	0,984	11,63	4,53	232,98	0,7	16,3	55,8	

Fonte: Autores (2019)

Em que: N (número de indivíduos), P (número de parâmetros), R2 Ajustado (coeficiente de determinação), PSE (erro padrão da estimativa), MAD (desvio absoluto médio), Range (amplitude de variação), GSkew (índice de assimetria de Galton), PSkew (índice de assimetria de Pearson), PCor (coeficiente de correlação de Spearman).

Tabela 7 – Tabela de índices utilizados na comparação das equações de dupla entrada em seu modelo geral e na presença de variáveis qualitativas, com interação destas variáveis

Modelos	N	P	R2 Ajustado	PSE	MAD	Range	GSkew	PSkew	PCor	
Husch	1482	2	0,962	17,61	10,31	228,08	0,3	25,7	58,8	Geral
Baden Local	1482	3	0,961	17,95	10,59	273,11	2,9	4,6	59,6	
Área Transversal	1482	2	0,948	20,66	17,74	240,96	0,1	1,7	14,6	
Parabólico	1482	3	0,566	16,51	10,56	223,24	3,1	3	50,7	
Schumacher Local	1482	3	0,686	50,75	13,48	527,12	-7,4	68,9	62,1	
Brenac	1482	3	0,963	17,38	10,48	260,97	4,7	8,5	58,8	

Continua ...

Tabela 7 – Conclusão

Modelos	N	P	R2 Ajustado	PSE	MAD	Range	GSkew	PSkew	PCor	
Husch	1482	8	0,966	16,81	9,99	204,9	5	24,8	60	Município
Baden Local	1482	12	0,962	17,69	10,03	280,03	7	6	61,3	
Área Transversal	1482	8	0,955	19,19	15,26	222,53	-0,4	-5,6	21,9	
Parabólico	1482	12	0,969	16,1	10,2	204,02	-1,1	-4,6	53,6	
Schumacher Local	1482	12	0,722	47,75	13,33	521,65	5,7	66,9	61,5	
Brenac	1482	12	0,964	17,34	10,05	267,82	3,9	6,1	61,5	
Husch	1482	34	0,972	14,98	7,41	225,44	7	10,3	61	Fazenda
Baden Local	1482	51	0,974	14,65	7,27	204,05	8,6	5,8	60,3	
Área Transversal	1482	34	0,967	16,39	10,7	194,11	2,1	-8,7	33,2	
Parabólico	1482	51	0,976	13,9	7,79	201,83	6,2	-1,3	53,2	
Schumacher Local	1482	51	0,808	39,44	10,36	530,68	4,9	59	64	
Brenac	1482	51	0,974	14,59	7,33	194,54	8,4	7	60	
Husch	1482	126	0,98	12,62	4,78	201,78	4,4	-6,4	59,4	Estrato
Baden Local	1482	189	0,984	11,66	4,71	188,11	-0,5	3,8	60,2	
Área Transversal	1482	126	0,983	11,81	5,67	193,74	-5,4	-2	42,6	
Parabólico	1482	189	0,985	11,09	5,04	189,96	-1	1,7	52,2	
Schumacher Local	1482	189	0,877	30,98	6,94	427,89	-5	48	60,8	
Brenac	1482	189	0,983	11,88	4,78	190,86	-0,7	2	60,5	
Husch	1482	248	0,988	9,86	3,71	167,96	6,5	-5,8	53,5	Talhão
Baden Local	1482	372	0,991	8,91	3,31	125,42	3,7	4,5	50,2	
Área Transversal	1482	248	0,989	9,22	4,34	125,39	-4,5	-1,3	33,8	
Parabólico	1482	372	0,992	8,17	3,57	108,03	2,8	0	38,9	
Schumacher Local	1482	372	0,889	28,61	5,89	367,71	-3	44,6	53,5	
Brenac	1482	372	0,989	9,2	3,32	137,44	4,4	2,9	50,3	

Fonte: Autores (2019)

Em que: N (número de indivíduos), P (número de parâmetros), R2 Ajustado (coeficiente de determinação), PSE (erro padrão da estimativa), MAD (desvio absoluto médio), Range (amplitude de variação), GSkew (índice de assimetria de Galton), PSkew (índice de assimetria de Pearson), PCor (coeficiente de correlação de Spearman).

A introdução das variáveis de localização como intercepto resultou na melhora de todos os modelos, com maior acentuação como nos de dupla entrada, no erro padrão

da estimativa, e menor nos demais índices. A melhora foi gradativa com o aumento da especificação da localização (da variável “município” para a variável “talhão”), alterando o ranqueamento dos modelos, segundo o erro padrão da estimativa. O modelo que apresentou o melhor desempenho do erro padrão da estimativa foi Husch, seguido de Baden. Entretanto, o modelo de Schumacher local obteve uma melhoria de desempenho muito alta, porém seu desempenho continuou sendo ruim. Os modelos Parabólico, Área Transversal e Schumacher local tiveram menor desenvolvimento que os demais, pois seus índices não apresentaram melhoras significativas com a introdução das variáveis de localização como intercepto.

Considerando o índice de assimetria de Galton, alguns modelos apresentaram diminuição na assimetria, como Baden, Schumacher local e Brenac, enquanto os demais modelos apresentaram pioras em sua simetria. Por outro lado, segundo o índice de assimetria de Pearson, houve uma séria piora nos modelos de Baden, Parabólico e Brenac.

A introdução das variáveis de localização como interação resultou em melhoras significativas, quando comparado com a introdução como intercepto, com o erro padrão da estimativa apresentando as maiores melhoras. O ranqueamento foi pouco afetado com a introdução da localização como interação, tendo o modelo Parabólico com o melhor desempenho e melhora, seguido do modelo de Baden. Novamente, o modelo de Schumacher local apresentou boa melhoria, porém seu desempenho continuou sendo pior que dos demais modelos.

Todos os índices apresentaram claras melhoras com a introdução das variáveis de localização como interação em todos os modelos, como o desvio absoluto médio, amplitude de variação e coeficiente de determinação. Porém, os índices de simetria apresentaram comportamentos diferentes para cada modelo. O índice de assimetria de Galton foi mantido praticamente inalterado nos modelos de Brenac e Parabólico, apresentando piora nos demais modelos. O índice de assimetria de Pearson, por outro lado, apresentou melhora em quase todos os modelos, com exceção dos modelos de

Baden, que não foi alterado, e de Área Transversal que passou de assimétrico à direita para assimétrico à esquerda.

Todos os modelos locais apresentaram melhores desempenhos quando acrescidos das variáveis de localização, sendo que a melhora ocorreu de modo gradativo com o aumento da especificação da localização de “município” até “talhão”. Comparados aos modelos sem informação de localização, os modelos com a variável “talhão” tiveram as menores diferenças entre si, tanto com a introdução da variável como intercepto quanto com a introdução da variável como interação.

Seguindo os resultados dos modelos de dupla entrada, os gráficos de dispersão e Q-Q normal dos resíduos dos modelos locais apresentaram os melhores resultados quando ajustados com a variável de localização “talhão”, contudo, tais resultados ainda se mostraram inferiores aos resultados dos modelos de dupla entrada.

3.3 Discussão

A Tabela 8 resume os resultados encontrados em termos de PSE, porém o modelo Schumacher local foi excluído devido a seu péssimo desempenho.

A inclusão da variável qualitativa “talhão” resulta uma melhora de todos os modelos. Essa melhora é maior que a diferença entre os modelos tanto no ajuste geral quanto no ajuste com a interferência de talhão. Portanto, nas equações volumétricas de dupla entrada, a diferença de desempenho entre os modelos é, via de regra, maior quando os modelos são ajustados com a informação de “talhão”. Porém, nas equações volumétricas locais, a diferença de desempenho entre os modelos é via de regra maior no ajuste geral, isto é, sem informação de talhão.

A introdução de variáveis qualitativas de localização nos modelos volumétricos tem um forte impacto sobre o desempenho dos modelos, tanto os modelos de dupla-entrada quanto nos locais. Isso já era indicado na literatura (MCTAGUE; BATISTA; STEINER, 1989; ROLIM *et al.*, 2006; MIGUEL *et al.*, 2010; MIGUEL; LEAL, 2012), mas, neste trabalho, observou-se uma melhoria crescente na acuidade de predição dos modelos

à medida que a variável qualitativa indicou localizações progressivamente mais específicas, ou seja, na sequência: ajuste geral, ajuste por município, por propriedade, por estrato e por talhão. O fato de que a melhoria na acuidade da predição entre o ajuste geral e o ajuste por talhão (intercepto ou interação) seja maior em todos os modelos individualmente que as diferenças entre os modelos numa dada localização (geral ou talhão) indica que a escolha da especificidade de localização mais apropriada para o ajuste das equações volumétricas é tanto ou mais importante que a escolha do modelo a ser ajustado.

Tabela 8 – 1ª tabela são os modelos de dupla entrada e 2ª tabela os modelos locais

Modelos de Dupla-Entrada						
	TALHÃO	Spurr	Schumacher-Hall	Baden	Meyer	Stoate
GERAL	---	7,17	7,43	7,28	6,93	7,01
Spurr	5,96	1,21	0,26	0,11	0,24	0,16
Schumacher-Hall	6,67	0,71	0,76	0,15	0,50	0,42
Baden	5,45	0,51	1,22	1,83	0,35	0,27
Meyer	5,05	0,91	1,62	0,40	1,88	0,08
Stoate	5,43	0,53	1,24	0,02	0,38	1,58

Modelos Locais							
	TALHÃO	Husch	Baden Local	Área Transversal	Parabólico	Brenac	Schumacher-Local
GERAL	---	17,61	17,95	20,66	16,51	17,38	50,75
Husch	9,86	7,75	0,34	3,05	1,10	0,23	33,14
Baden Local	8,91	2,95	9,04	2,71	1,44	0,57	32,80
Área Transv.	9,22	0,64	0,31	11,44	4,15	3,28	30,09
Parabólico	8,17	1,69	0,74	1,05	8,34	0,87	34,24
Brenac	9,20	0,66	0,29	0,02	1,03	8,18	33,37
Schumacher-Local	38,57	28,71	29,66	29,35	30,40	29,37	12,18

Fonte: Autores (2019)

Em que: Valores de PSE e de diferença de PSE entre os modelos. 1ª linha (GERAL): PSE dos modelos no ajuste geral. 1ª coluna (TALHAO): PSE dos modelos no ajuste com informações de talhão (interação). Diagonal principal: diferença de PSE de cada modelo entre ajuste geral e com talhão. Matrix triangular superior: diferença de PSE entre os modelos no ajuste geral. Matrix triangular inferior: diferença de PSE entre os modelos no ajuste com talhão.

O ganho em acuidade de precisão obtido na seleção do melhor modelo pode ser relativamente pequeno face ao ganho possível na decisão sobre a escala em que o modelo será ajustado. Contudo, o fato de os modelos terem apresentado melhoria diferenciada em desempenho à medida que as variáveis de localização foram introduzidas, mostra que a seleção do melhor modelo e a escolha da localização mais apropriada para ajuste não podem ser realizadas de modo independente. É necessário comparar modelos e localizações conjuntamente.

Como esperado, os modelos de equação de dupla-entrada se mostraram superiores aos modelos de equação local em todas as localizações. Contudo, as equações locais ajustadas para talhões individualmente são comparáveis às equações de dupla-entrada com ajuste geral. Esse resultado ilustra o tradicional dilema da escolha entre modelos para localizações específicas que requerem um menor número de medidas (equações locais). A solução para esse dilema deve ser realizada caso-a-caso sendo dependente do bom senso do biometrista florestal.

Além da acuidade da predição, medida pelo PSE, outros índices da qualidade do desempenho dos modelos podem ser frequentemente utilizados na seleção de equações. Os resultados obtidos indicam que outros índices calculados diretamente dos resíduos do ajuste, como MAD e a amplitude de variação, são semelhantes ao PSE. Os índices que indicam a simetria da distribuição dos resíduos, contudo, mostram grande variabilidade tanto entre os modelos quanto em função da especificidade de localização do ajuste. Via de regra, a simetria tende a aumentar em todos os modelos à medida que se aumenta a especificidade de localização do ajuste. Embora raramente avaliada, a simetria dos erros de predição é um atributo intuitivamente esperado de todo bom modelo de predição, o que reforça a importância de uma cuidadosa consideração da especificidade de localização de ajuste na construção das equações volumétricas.

Na avaliação da localização apropriada para a construção de equações volumétricas, é importante considerar que o número de árvores disponíveis para cada localização se torna menor à medida que essas se tornam progressivamente mais específicas. No presente trabalho, os 1.482 fustes do conjunto de dados representam

370 fustes para cada município, mas apenas 12 fustes por talhão. Esse aspecto enfatiza a importância de que a informação de localização seja introduzida nos modelos volumétricos como variáveis qualitativas (variáveis indicadoras) dos modelos e jamais como modelos ajustados individualmente para as localizações específicas. Embora modelos com alta especificidade de localização tenham muitos parâmetros, sendo praticamente impossível apresentá-los de forma explícita, isso não representa um impedimento para a utilização prática profissional. Por exemplo, o modelo de Meyer ajustado com a variável qualitativa “talhões” possui 744 parâmetros ajustados (coeficientes de regressão). A impossibilidade da apresentação explícita desse modelo não impede, contudo, que ele seja facilmente utilizado em softwares e ambientes computacionais profissionais, como o ambiente R, por exemplo.

4 CONCLUSÃO

A introdução da informação de localização das árvores e fuste, na forma de variáveis qualitativas (variáveis indicadoras), é um importante aspecto na construção de equações volumétricas. A escolha da especificidade dessa localização, ou, dito de outra forma, da escala espacial apropriada para a construção das equações volumétricas, é tão relevante quanto a seleção do melhor modelo. Mais ainda, a escolha da escala de ajuste e a seleção dos modelos são aspectos inseparáveis que não podem ser realizados de modo independente. Embora maior especificidade de localização no ajuste, isto é, maior resolução de escala espacial, tenda a resultar, via de regra, em melhor desempenho em todos os modelos de equação volumétrica, a escala mais apropriada à aplicação da equação volumétrica não é necessariamente a mais específica. Outros aspectos de ordem prática devem ser considerados como a infraestrutura computacional que será utilizada na aplicação da equação volumétrica construída e a introdução da informação de localização no modelo volumétrico construído. A introdução da informação de localização nas equações volumétricas deve ser realizada, preferencialmente, na forma de variáveis qualitativas (variáveis indicadoras) introduzidas no procedimento de regressão linear.

REFERÊNCIAS

ANDRADE, V. C. L. Modelos volumétricos de dupla entrada para aplicar em povoamentos florestais brasileiros. **Floresta Ambiente**, Seropédica, v. 24, e00135415, 2017.

AZEVEDO, G. B. *et al.* Estimativas volumétricas em povoamentos de eucalipto sob regime de alto fuste e talhadia no sudoeste da Bahia. **Pesquisa Florestal Brasileira**, Colombo, v. 31, n. 68, p. 309-318, 2011.

BATISTA, J. L. F.; COUTO, H. T. Z.; SILVA FILHO, D. F. **Quantificação de recursos florestais**. 1. ed. São Paulo: Oficina de Textos, 2014. 384 p.

CLIMATE-DATA.ORG. **Dados climáticos para cidades mundiais**. [S. l.], 2018. Disponível em: <https://pt.climate-data.org/>. Acesso em: 10 abr. 2018.

FORD, C. **Understanding Q-Q Plots**. Charlottesville: University of Virginia Library, 2015. Disponível em: <http://data.library.virginia.edu/understanding-q-q-plots/>. Acesso em: 12 abr. 2018.

GONÇALVES, J. L. M. *et al.* Mapeamento de solos e da produtividade de plantações de *Eucalyptus grandis*, com uso de sistema de informação geográfica. **Scientia Forestalis**, Piracicaba, v. 40, n. 94, p. 187-201, 2012.

INSTITUTO DE PESQUISAS TECNOLÓGICAS (SP). **Ordenamento territorial geomineiro do município de Bofete**. São Paulo: Secretaria de Desenvolvimento do Estado de São Paulo; Prefeitura Municipal de Bofete, 2009. Relatório Técnico N° 17501-301.

MACHADO, S. A.; CONCEIÇÃO, M. B.; FIGUEIREDO, D. J. Modelagem do volume individual para diferentes idades e regimes de desbaste em plantações de *Pinus oocarpa*. **Ciências Exatas e Naturais**, Guarapuava, v. 4, n. 2, p. 185-197, 2002.

MCTAGUE, J. P.; BATISTA, J. L. F.; STEINER, L. H. Equações de volume total, volume comercial e forma do tronco para plantações de *Eucalyptus* nos estados de São Paulo e Rio de Janeiro. **IPEF**, Piracicaba, n. 41/42, p. 56-63, 1989.

MIGUEL, E. P. *et al.* Ajuste de modelo volumétrico e desenvolvimento de fator de forma para plantios de *Eucalyptus grandis* localizados no município de Rio Verde – GO. **Enciclopédia Biosfera**, Goiânia, v. 6, n. 11, p. 1-13, 2010.

MIGUEL, E. P.; LEAL, F. A. Seleção de equações volumétricas para a predição do volume total de *Eucalyptus urophylla* S. T. Blake na região Norte do Estado de Goiás. **Enciclopédia Biosfera**, Goiânia, v. 8, n. 14, p. 1372, 2012.

REZENDE, A. V. *et al.* Comparação de modelos matemáticos para estimativa do volume, biomassa e estoque de carbono da vegetação lenhosa de um cerrado *sensu stricto* em Brasília, DF. **Scientia Forestalis**, Piracicaba, n. 71, p. 65-76, ago. 2006.

ROLIM, S. G. *et al.* Modelos volumétricos para a Floresta Nacional do Tapirapé-Aquirí, Serra dos Carajás (PA). **Acta Amazônica**, Manaus, v. 36, n. 1, p. 107-114. 2006.

SALES, F. C. V. *et al.* Ajuste de modelos volumétricos para clone *Eucalyptus grandis* x *E. urophylla* cultivados no agreste de Pernambuco. **Floresta**, Curitiba, v. 45, n. 4, p. 663-670, 2015.

SCHNEIDER, P. R.; TONINI, H. Utilização de variáveis dummy em equações de volume para *Acacia mearnsii* De Wild. **Ciência Florestal**, Santa Maria, v. 13, n. 2, p. 121-129, 2003.

SOARES, C. P. B. *et al.* Equações hipsométrica, volumétricas e de *Taper* para onze espécies nativas. **Revista Árvore**, Viçosa, MG, v. 35, n. 5, p. 1039-1051, 2011.

THOMAS, C. *et al.* Comparação de equações volumétricas ajustadas com dados de cubagem e análise de tronco. **Ciência Florestal**, Santa Maria, v. 16, n. 3, p. 319-327, 2006.

TIBCO SOFTWARE. **Visão geral de medidas**: desvio médio absoluto. Palo Alto, 2013. Disponível em: https://docs.tibco.com/pub/spotfire_web_player/6.0.0-november-2013/pt-BR/WebHelp/GUID-3BBFA6DA-7780-4914-B187-06E3B3E35104.html. Acesso em: 26 maio 2018.

VALLADARES NETO, J. *et al.* Boxplot: um recurso gráfico para a análise e interpretação de dados quantitativos. **Revista Odontológica do Brasil Central**, Goiânia, v. 26, n. 76, p. 1-6, 2017.

Contribuição de Autoria

1 – Vinicius Pizzo Ferreira

Engenheiro Florestal, Pesquisador Autônomo

<https://orcid.org/0000-0003-2112-2034> • viniciusferreira12@gmail.com

Contribuição: Curadoria de dados, Análise Formal, Metodologia, Software, Visualização de dados, Escrita – primeira redação, Escrita – revisão e edição

2 – João Luis Ferreira Batista

Engenheiro Florestal, Dr., Professor

<https://orcid.org/0000-0002-4730-6897> • batista.jlf@usp.br

Contribuição: Conceituação, Metodologia, Administração do projeto, Software, Supervisão, Validação

Como citar este artigo

Ferreira, V. P.; Batista, J. L. F. Utilização de variáveis qualitativas de localização em equações volumétricas e a seleção de modelos. **Ciência Florestal**, Santa Maria, v. 31, n. 4, p. 1968-1990, 2021. DOI 10.5902/1980509840926. Disponível em: <https://doi.org/10.5902/1980509840926>.