

Uso da Análise Multivariada para agrupamento de dados no ajuste da relação diâmetro– altura em plantios florestais misto na Amazônia Mato – Grossense

Larissa Mustasso^{1†}, Lucas H. O. Santos², Lila M. G. R. Diaz³, João Paulo S. Madi⁴, Pedro Henrique K. Millikan⁵, Samuel P. C. Carvalho⁶, Carlos A. Silva⁷, Mariana P. L. C. Carvalho⁸

¹Universidade Federal de Mato Grosso/Cuiabá-MT.

²Programa de Pós Graduação em Ciências Florestais e Ambientais/Universidade Federal de Mato Grosso/Cuiabá- MT. E-mail: lucashenderson@ufmt.br.

³Programa de Pós Graduação em Ciências Florestais e Ambientais/Universidade Federal de Mato Grosso/Cuiabá-MT. E-mail: lila.gamarra@gmail.com.

⁴Engenheiro Florestal: joaosardomadi@gmail.com.

⁵Universidade Federal de Mato Grosso/Cuiabá-MT. E-mail: millikanp@gmail.com.

⁶Universidade Federal de Mato Grosso/Cuiabá-MT. E-mail: spccarvalho@ufmt.br.

⁷Universidade Federal de Mato Grosso/Cuiabá-MT. E-mail: carlos_engflorestal@outlook.com.

⁸Universidade Federal de Mato Grosso/Cuiabá-MT. E-mail: marianaperes@ufmt.br.

Resumo: O estudo foi desenvolvido com dados de um reflorestamento no município de Cotriguaçu – MT, pertencente a ONF - Brasil, com os dados de 10 famílias nativas da Amazônia brasileira. Primeiramente, o modelo de regressão foi ajustado de uma forma global para todo o povoamento, posteriormente por família, e então, por grupos de famílias que foram agrupados utilizando análise Multivariada de Dados. Os critérios estatísticos utilizados para a análise dos resultados foram: Raiz Quadrada do Erro Médio (rmse), Raiz Quadrada o Erro Médio Relativa (rmse%), Coeficiente de Correlação de Pearson (r); e o gráfico de dispersão das alturas estimadas versus alturas observadas. A análise de resultados realizada através gráfico de dispersão das alturas permitiu verificar que a utilização da variável família e grupos de família geraram parâmetros específicos para cada ajuste quando comparado ao modelo ajustado globalmente. Constatou - se ainda, com os resultados estatísticos, que o fator família é uma variável que interfere na altura, não tanto quanto o DAP, mas pode ser utilizada para minimizar o erro das estimativas.

Palavras-chave: Relação hipsométrica; Inteligência Artificial; Plantios Mistos; Amazônia Mato-Grossense.

Abstract: This study was conducted using data from a reforestation located in the municipality of Cotriguaçu - MT, owned by ONF - Brazil, containing data from 10 native families of the Brazilian Amazon. Initially, the regression model was fitted for the entire population, then by family, and posteriorly by groups of families that were grouped using Multivariate Data Analysis. The statistical criteria used to analyze the results were: Root Mean Squared Error (rmse), Relative Root Mean Squared Error (rmse%), Pearson Correlation Coefficient (r); and the scatter plot of estimated versus observed heights. The graphical analysis of height dispersion allowed us to verify that the use of the variable family and family groups generated specific parameters for each fit when compared to the globally fitted model. It was also verified, through statistical results, that family is a variable that affects height, not as much as DBH, but can be used to minimize the error of estimates.

Keywords: Amazonia Mato-Grossense; Artificial intelligence; Hypsometric relationship; Mixed Plantings; Amazonia Mato-Grossense.

Introdução

A altura é uma importante variável dendrométrica pois possui alta relação com o volume total das árvores, cujo valor influi diretamente nos planos de manejo florestal, porém, é uma medida onerosa, especialmente em florestas nativas devido às dificuldades de visualização dos indivíduos.

†Autora correspondente: lalimustasso@hotmail.com.

Contudo, a altura pode ser estimada, através de uma relação da altura em função do diâmetro que possibilita o ajuste de modelos de regressão. Com a obtenção do DAP (diâmetro a altura do peito) de todas as árvores, e a amostragem de parte das alturas, é possível estimar o valor das demais tornando o inventário florestal mais econômico, quando bem ajustada, apresenta precisão estatisticamente satisfatória.

Com isso, a busca em tornar as relações de diâmetro e altura mais próximas da realidade é de grande interesse para os pesquisadores da área florestal, uma vez que minimiza a quantidade de medidas coletadas em campo. Segundo Prodan et al., (1997), uma única equação ajustada para todo o povoamento englobaria muitas relações diferentes, resultando em alta variabilidade em torno da linha de regressão, ocasionando incertezas na geração de parâmetros dendrométricos, podendo assim, estar subestimando ou superestimando os valores.

Em razão desse problema, uma estratégia utilizada em inventários florestais visando o aumento da precisão e melhorando assim os resultados dos modelos ajustados, é a estratificação dos dados em grupos com características semelhantes para o ajuste das equações (CURTO et al., 2014).

Diante do exposto, este trabalho teve como objetivo avaliar a precisão de equações hipsométricas em dados agrupados por análise multivariada em um plantio florestal misto na Amazônia Mato-grossense.

Material e Métodos

O estudo foi realizado no município de Cotriguaçu, noroeste do estado de Mato Grosso, na Fazenda São Nicolau sob gestão da ONF – Brasil. O clima regional pela classificação de Koppen é do tipo Aw, temperatura média de 25°C e o solo da classe dos Alissolos (ALVARES et al., 2013).

A área total da propriedade é de 10 mil hectares, nos quais quase 2 mil hectares que eram de pastagens foram seccionados em talhões para reflorestamento com diferentes arranjos de espécies e espaçamentos. O reflorestamento da área teve início em 1999 e foi finalizado em 2004.

O inventário do plantio foi realizado anualmente no mês de agosto, e o presente trabalho foi realizado com os inventários dos anos de 2012, 2013, 2014, 2015, 2016 e 2017. Para realizar os ajustes das equações diâmetro - altura, foram selecionadas 10 famílias, nativas da Amazônia, que possuíam mais de 100 indivíduos por ano de medição, sendo elas: Anacardiaceae, Bignoniaceae, Bixaceae, Boraginaceae, Fabaceae, Malvaceae, Meliaceae, Moraceae, Myrtaceae e Simaroubaceae.

Os modelos foram ajustados através do software estatístico “R”. O modelo de regressão escolhido foi o modelo linear simples da altura em função do diâmetro representado na Eq.1, sendo este preliminarmente ajustado de uma forma global para todos os dados e, posteriormente, de forma empírica, foi realizado o ajuste por família. Com o intuito de verificar a possibilidade de formar grupos de família, foi realizado uma Análise Multivariada da Variância (MANOVA), e com resultado da análise significativo, foi prosseguido para a formação dos grupos utilizando a técnica de análise de correlação canônica, em seguida foi realizado o ajuste da Eq.1 utilizando estes grupos como uma variável da relação.

$$H = \beta_0 + \beta_1 . DAP + \varepsilon \quad (1)$$

Com o objetivo de analisar os ajustes de relação diâmetro – altura, utilizou-se o coeficiente de correlação de Pearson (Eq.2) e a raiz quadrada do erro médio (Eq.3). Por fim, foram gerados os gráficos de dispersão das alturas estimadas X alturas observadas para uma análise exploratória dos resultados.

Sigmae, Alfenas, v.8, n.2, p. 156-161, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18º Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).

$$r = \frac{\sum_{i=1}^n (x - \bar{x})(y - \bar{y})}{\sqrt{[\sum_{i=1}^n (x - \bar{x})^2][\sum_{i=1}^n (y - \bar{y})^2]}} \quad (2)$$

$$RMSE = \frac{\sum_{i=1}^n (y - \hat{y})^2}{n - p - 1} \quad (3)$$

Em que: x = valores observados para x ; \bar{x} = média dos valores de x ; \bar{y} = média dos valores de y ; n = número total de dados; p = número de variáveis independentes do modelo; y = valores observados para y e \hat{y} = valores estimados de y pela regressão.

Resultados e discussão

A Figura 1 se refere ao gráfico do modelo de regressão ajustado de forma global para todos os dados. Pode-se perceber que o modelo global apresenta uma grande quantidade de indivíduos dispersos e distantes da linha de regressão, o que provavelmente se refere a uma baixa relação da altura com o diâmetro, fazendo com que resulte em estimativas pouco precisas e baixos parâmetros estatísticos.

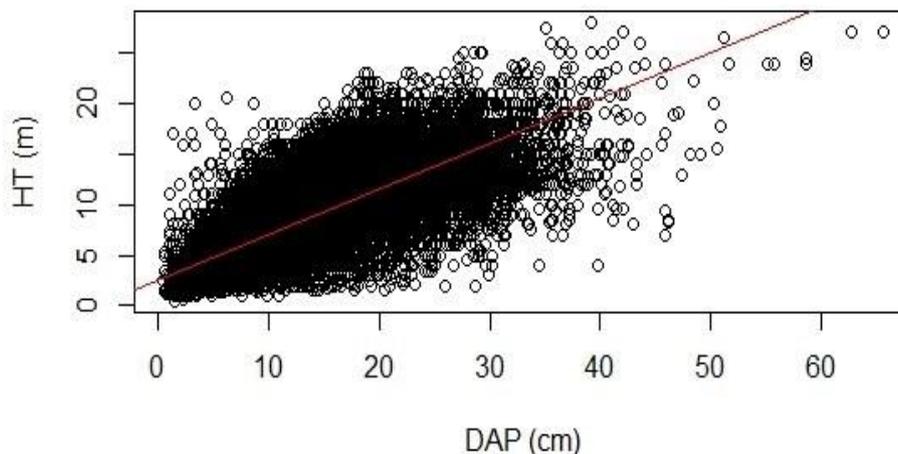


FIGURA 1 – Modelo de regressão ajustado globalmente

A Figura 2 apresenta o gráfico da regressão ajustado por família da espécie, e pode-se observar que os dados se comportaram de forma mais adequada comparando a dispersão do ajuste global.

O resultado da MANOVA gerou um valor significativo a 95% de probabilidade sendo o valor de $F = 2,2^{-16}$, portanto a formação dos grupos foi viável e estão representados no Quadro 1.

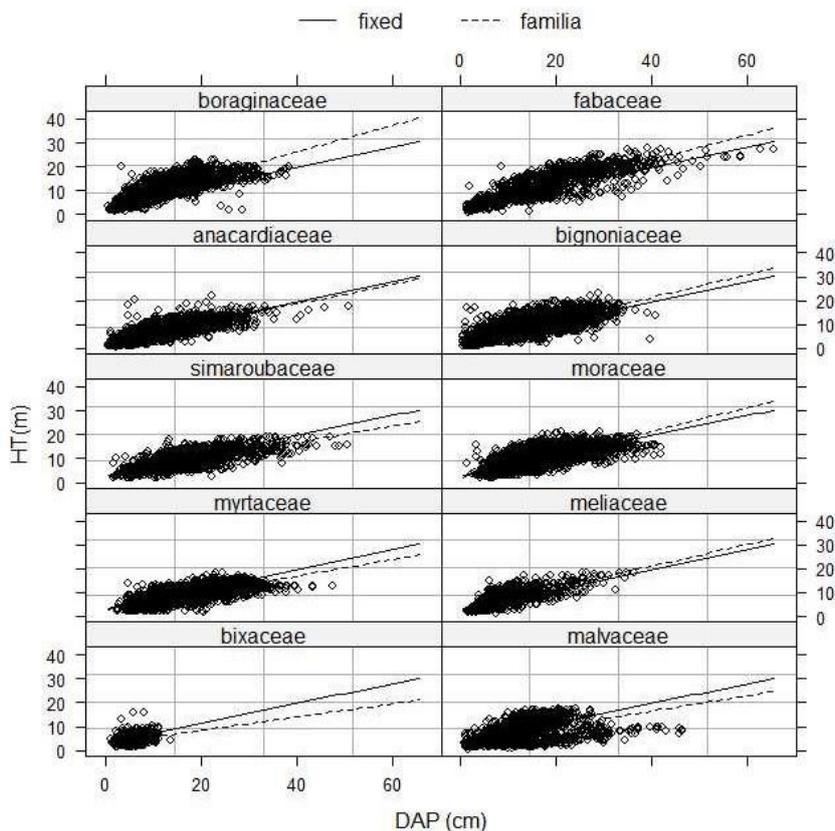


FIGURA 2 – Modelo de regressão ajustado por família

QUADRO 1 – Grupos de famílias mais semelhantes

Grupo 1	Anacardiaceae
	Bignoniaceae
	Bixaceae
	Malvaceae
Grupo 2	Myrtaceae
	Simaroubaceae
Grupo 3	Fabaceae
	Moraceae
Grupo 4	Boraginaceae
	Meliaceae

A Figura 3 corresponde ao gráfico de regressão ajustado por grupos de família, e é notável que a dispersão dos dados é maior quando comparado ao modelo por família, o que já era esperado devido ao meio número que indivíduos por grupo, assim como espécies de famílias diferentes.

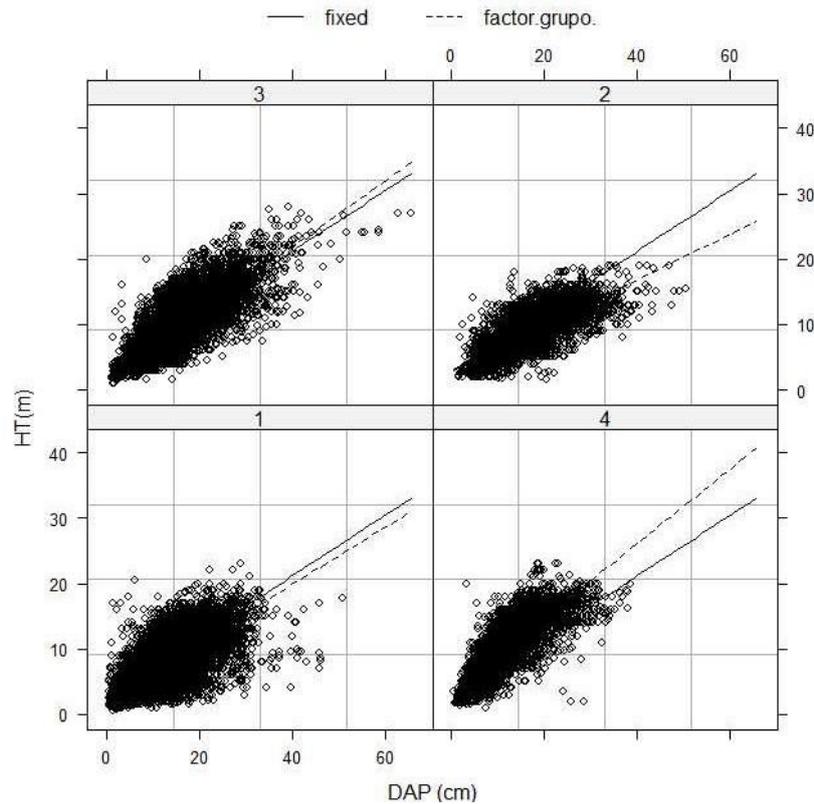


FIGURA 3 – Modelo de regressão ajustado por grupo de família.

Analisando a Tabela 2 a qual apresenta os resultados estatísticos dos ajustes, pode-se observar que o modelo ajustado de uma forma global para todos os dados, apresentou o menor coeficiente de correlação, tal como um erro médio elevado, o que justifica que um modelo de regressão ajustado para todo um povoamento pode resultar em erros significativos na predição da altura, devido a grande quantidade de indivíduos de diferente espécies, famílias e idades.

TABELA 1 – Resultados estatísticos dos modelos ajustados.

Ajuste	rmse (m)	rmse (%)	r (%)
Global	2,56	31,53	59,08
Família	2,31	28,31	67,01
Grupos de Família	2,65	29,07	65,20

Os modelos de regressão utilizando o fator família como uma variável, e o fator grupo de família o qual foi utilizado para agrupamento a Análise de Correlação Canônica, tiveram resultados próximos tanto de erro médio como de coeficiente de correlação, o modelo por família teve um resultado um pouco mais significativo que o de grupos de família, o que já era esperado devido ao maior número de indivíduos no ajuste assim como indivíduos de famílias diferente. No entanto, ambos os ajustes apresentaram resultados melhores que o modelo global.

Sanquetta et al. (2013), em um estudo de ajuste de relação diâmetro-altura para espécies lenhosas em um fragmento de floresta Ombrófila Mista no Sul do Paraná, concluiu que os modelos testados que englobaram todas as espécies em conjuntos não se comportaram adequadamente,

Sigmae, Alfenas, v.8, n.2, p. 156-161, 2019.

64^a Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18^o Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).

gerando baixo grau de ajuste ($R^2 = 41\%$) e conseqüentemente, maior erro de estimativa. Ainda segundo os autores, buscando uma alternativa para melhorar o desempenho das equações, realizou-se a estratificação dos dados, dividindo-os em famílias botânicas. Essa técnica possibilitou um melhor ajuste para algumas famílias, portanto outras não resultaram em coeficientes de determinação significativos. Souza et al. (2003) em um estudo que buscava a estratificação vertical em floresta inequidiana, concluiu que as técnicas de análise de agrupamento e análise discriminante para realizar tal atividade são viáveis, possibilitando um melhor conhecimento da estrutura das florestas, com isso gerando subgrupos mais homogêneos.

Conclusão

A utilização de Análise Multivariada de Dados pode ser uma alternativa para diminuir a quantidade de medições de altura no inventário florestal. Foi constatado que o fator família tem influência na altura das árvores, e o agrupamento destas não resulta em estatísticas inferiores ao ajuste realizado para cada uma das famílias.

Trabalhos futuros serão realizados avaliando a inclusão de outras variáveis que tenham alguma importância biológica para o povoamento, e que possam ser uma alternativa para o aprimoramento de formações de grupos e conseqüentemente melhores resultados estatísticos e melhores estimativas de altura.

Agradecimentos

CNPq, UFMT, PPGCFA e a CAPES.

Referências bibliográficas

- ALVARES, C.A.; STAPE, J.L.; SENTELHAS, P.C.; GONÇALVES, J.L.M.; SPAROVEK, G. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, v. 22, n. 6, p. 711–728, 2013.
- CURTO, R. A.; LOUREIRO, G. H.; MÔRA, R.; MIRANDA, R. O. V.; PÉLLICO NETTO, S. SILVA, G. F. Relações hipsométricas em Floresta Estacional Semidecidual. *Revista de Ciências Agrárias*, v. 57, n. 1, p. 57-66, 2014.
- PRODAN, M.; PETERS, R.; COX, F.; REAL, P. Mensura forestal. San José, Costa Rica: IICA, 1997. 586 p.
- SANQUETTA, C. R.; CORTE, A. P. D.; ROGLIN, A.; PIMENTEL, A. Relações diâmetro-altura para espécies lenhosas em um fragmento de floresta Ombrófila Mista no Sul do Paraná. *Revista Iheringa, Série Botânica*. v-68, p. 103 – 114. 2013.
- SOUZA, D. R.; SOUZA, A. L.; GAMA, J. R. V.; LEITE, H. G. Emprego de análise multivariada para estratificação vertical de florestas inequidiana. R. *Árvores*. V.27, p. 59-63, 2003.
- TRUGILHO, P. F.; LIMA, J. T.; MORI, F. A. Correlação canônica das características químicas e físicas da madeira de clones de *Eucalyptus grandis* e *Eucalyptus saligna*. *Cerne*, v.9, p. 66 – 80, 2003.

Sigmae, Alfenas, v.8, n.2, p. 156-161, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18º Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).