

ÁRVORE DE DECISÃO PARA CLASSIFICAR O DESENVOLVIMENTO INICIAL DO EUCALIPTO

ELIZEU DE SOUZA LIMA¹, ZIGOMAR MENEZES DE SOUZA², DIEGO ALEXANDER AGUILERA ESTEBAN³, INGRID NEHMI DE OLIVEIRA⁴, LENON HENRIQUE LOVERA⁵, VAGNER ROBERTO ARIEDI JUNIOR⁶

¹Engenheiro Florestal, Doutorando em Engenharia Agrícola (Água e Solo), Unicamp/Campinas-SP, (19) 3521-1111, elizeu.florestal@gmail.com

²Engenheiro Agrônomo, Professor Doutor, FEAGRI-UNICAMP/Campinas-SP

³Engenheiro Agrícola, Doutorando em Engenharia Agrícola (Água e Solo), FEAGRI-UNICAMP/Campinas-SP

⁴Engenheira Agrícola, Doutoranda em Engenharia Agrícola (Água e Solo), FEAGRI-UNICAMP/Campinas-SP

⁵Engenheiro Agrônomo, Doutorando em Engenharia Agrícola (Água e Solo), FEAGRI-UNICAMP/Campinas-SP

⁶Biólogo, Doutorando em Engenharia Agrícola (Água e Solo), FEAGRI-UNICAMP/Campinas-SP

Apresentado no
XLVIII Congresso Brasileiro de Engenharia Agrícola - CONBEA 2019
17 a 19 de setembro de 2019 - Campinas - SP, Brasil

RESUMO: O cultivo de eucalipto tem se expandido no Brasil, principalmente em regiões que apresentam solos de baixa fertilidade. Para isso, as tomadas de decisões são fundamentais para melhoria da produtividade. O objetivo deste trabalho foi avaliar os atributos de planta, solo e clima no desenvolvimento inicial do *Eucalyptus urograndis* em solo arenoso por meio da árvore de decisão. Para isso foram utilizados 30 atributos, sendo 29 preditivos e um atributo-meta (altura do eucalipto). Foram avaliadas quatro abordagens para seleção de atributos: sem seleção, seleção de atributos baseado em correlação, método do Qui-quadrado e Wrapper por meio da indução de árvore de decisão. A árvore de decisão permitiu o desenvolvimento de um modelo eficiente para classificar o crescimento inicial de eucalipto onde o volume individual de madeira e o diâmetro altura do peito são atributos determinantes para previsão do crescimento do *Eucalyptus urograndis*, além de atributos químicos e físicos do solo como o magnésio, fósforo, alumínio, potássio, acidez potencial e potencial hidrogeniônico, resistência do solo à penetração e relacionado ao clima como a temperatura mínima.

PALAVRAS-CHAVE: *Eucalyptus urograndis*, mineração de dados, volume de madeira.

DECISION TREE TO CLASSIFY THE INITIAL DEVELOPMENT OF EUCALYPTUS

ABSTRACT: Eucalyptus cultivation has expanded in Brazil, especially in regions with low fertility soils. To this end, decision-making is critical to improving productivity. The objective of this work was to evaluate the attributes of plant, soil and climate in the initial development of *Eucalyptus urograndis* in sandy soil through the decision tree. For this, 30 attributes were used, being 29 predictors and one meta attribute (eucalyptus height). Four approaches to attribute selection were evaluated: no selection, selection of attributes based on correlation, Chi-square and Wrapper method through decision tree induction. Decision tree allowed the development of an efficient model to rank the initial growth of eucalyptus where the individual volume of wood and the diameter of the breast are determinant attributes to predict the growth of *Eucalyptus urograndis*, besides soil chemical attributes such as magnesium ,

phosphorus, aluminum, potassium, potential acidity and hydrogen ionic potential, physical attributes such as soil resistance to penetration and climate related as the minimum temperature.

KEYWORDS: *Eucalyptus urograndis*, data mining, volume of wood.

INTRODUÇÃO: O cultivo de eucalipto no Brasil tem aumentado nos últimos anos em razão do seu rápido crescimento, diversificação no uso da madeira e facilidade de adaptação às diferentes condições edafoclimáticas. No entanto, o cultivo de eucalipto vem ocorrendo em regiões do cerrado com o predomínio de solos altamente intemperizados, ácidos e de baixa disponibilidade de nutrientes. Para manter e aumentar a produtividade, é necessário prevalecer-se de métodos de monitoramento e avaliação para determinar quais variáveis mais afetam o desenvolvimento e a produtividade da cultura. Com isso, a utilização de sistemas computacionais na tomada de decisão é uma estratégia eficiente que possibilita avaliar grande quantidade de informações obtendo resultados que poderão auxiliar no aumento da produtividade. Dentre as técnicas computacionais, a mineração de dados tem como finalidade obter resultados com o objetivo de inferir uma variável dependente a partir de um conjunto de atributos relacionados (CRIVELENTI et al., 2009). Entre as técnicas de mineração de dados, a indução por árvore de decisão se destaca por permitir classificar conjuntos de dados alcançando resultados satisfatórios que auxilia no processo de tomada de decisão, considera os atributos mais relevantes e a facilidade a interpretação e compreensão dos resultados. Portanto, objetivou-se avaliar os atributos de planta, solo e clima no desenvolvimento inicial do *Eucalyptus urograndis* em solo arenoso por meio da árvore de decisão (HAN et al., 2011).

MATERIAL E MÉTODOS: O estudo foi desenvolvido em área experimental localizada no município de Três Lagoas, MS, com precipitação média anual de 1.300 mm e temperatura média de 23,7 °C. A área experimental está sob cultivo de *Eucalyptus urograndis* e o solo foi classificado como Neossolo Quartzarênico (SANTOS et al., 2018).

Foram avaliados os atributos dendrométricos: altura individual das árvores de eucalipto (ALT), diâmetro altura do peito (DAP) e volume de madeira (VOL). Os atributos do solo foram resistência do solo à penetração (RP), umidade gravimétrica (UG), umidade volumétrica (UV), densidade do solo (DS), densidade da partícula (DP), porosidade total (PT), areia, silte, argila, fósforo (P), matéria orgânica (MO), potencial hidrogeniônico (pH), potássio (K), cálcio (Ca), magnésio (Mg), acidez potencial (H+Al), alumínio (Al), soma de bases (SB), capacidade de troca catiônica (CTC), saturação por bases (V), relação cálcio com capacidade de troca catiônica (Ca/CTC), relação magnésio com capacidade de troca catiônica (Mg/CTC) e saturação por alumínio (m), coletados nas camadas de 0,00-0,20 m e 0,20-0,40 m. Estes foram determinados de acordo com metodologia proposta por Raij et al. (2001), Stolf et al. (2014) e TEIXEIRA et al., 2017. Os atributos de clima foram precipitação (Prec), temperaturas máximas (Tmx), média (Tmed) e mínima (Tmin) ocorrida durante o ano agrícola (365 dias).

O conjunto de dados original foi composto por 300 observações para cada atributo. O atributo-meta refere-se à altura do *Eucalyptus urograndis* como alvo de classificação. As diferentes alturas do *E. urograndis* foram submetidas a um procedimento de discretização em categorias, ou seja, os dados contínuos foram transformados em dados discretos (intervalos).

Para seleção dos atributos foram avaliadas quatro abordagens: sem seleção de atributos, seleção de atributos baseado em correlação (CFS), método do Qui-quadrado (χ^2) e abordagem Wrapper. Para a classificação dos dados foi utilizado o método de árvore de decisão binária

induzida no software Weka 3.6 (WITTEN et al., 2011). O algoritmo de indução utilizado foi o J48, amplamente conhecido como C4.5.

RESULTADOS E DISCUSSÃO:

O desempenho obtido pelos diferentes métodos de seleção de atributos foi semelhante, com acurácia de 84 a 85,67%, e Kappa de 0,76 a 0,79, sendo a classificação de todos os métodos testados considerada muito boa com o número de regras geradas entre de 12 e 21.

Dentre os métodos de seleção de atributos avaliados, o método sem seleção de atributos foi o que demonstrou melhor eficiência do modelo (acurácia de 85% e Kappa de 0,78), indicando uma classificação muito boa. Este método utilizou todos os atributos e gerou uma árvore de decisão com 18 regras. Além disso, esse método apresentou alta precisão para as classes avaliadas, acertando 89% da classe baixa, 78% da classe média e 89% da classe alta, indicando eficiência do modelo em classificar todas as classes estudadas.

Dos 29 atributos disponíveis, o método Wrapper selecionou apenas cinco (DAP, VOL, DS, PT e H+A1) com acurácia de 85,67%, e Kappa de 0,79, indicando que a utilização desses cinco atributos é suficiente para se alcançar uma classificação muito boa, no entanto, este método gerou uma árvore de decisão com número regras e atributos selecionados baixo.

A partir da indução da árvore de decisão, sem seleção de atributos, foi gerada uma árvore de decisão com 18 regras, sendo cinco regras pertencentes à classe baixa (27%), nove pertencentes à classe média (50%) e quatro pertencentes à classe alta (22%). A árvore de decisão gerada (Figura 1) demonstra que o volume de madeira (VOL) apresentou maior importância devido ao maior ganho de informação exibido como destaque na raiz da árvore.

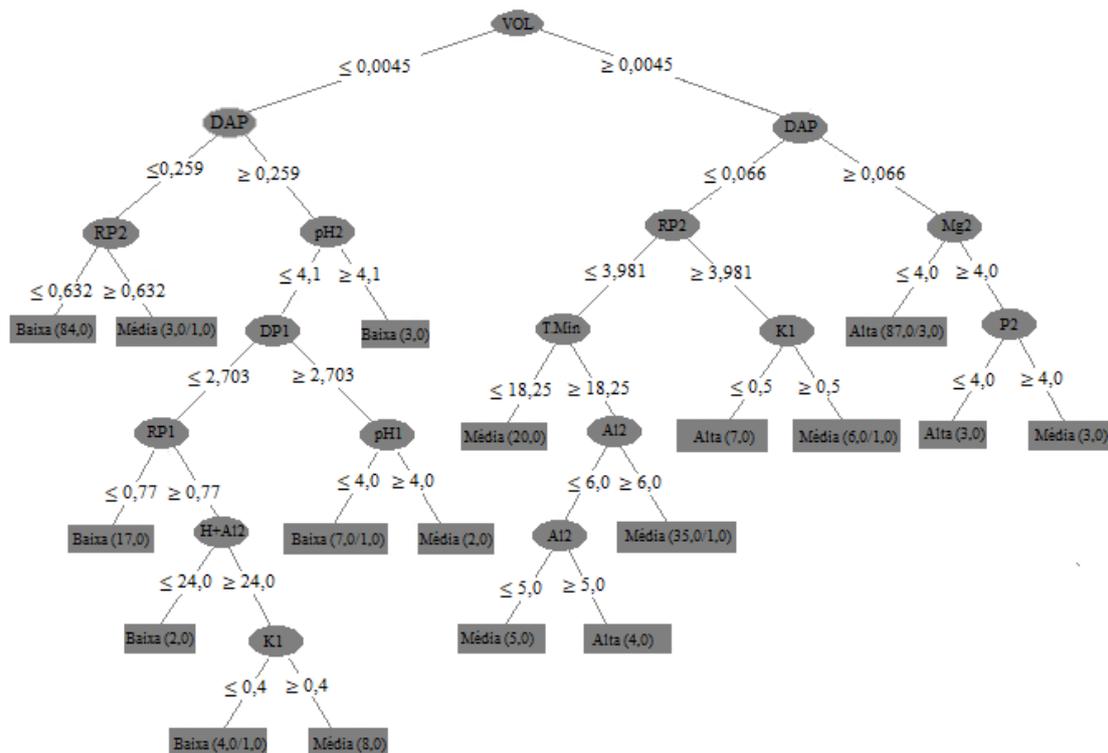


FIGURA 1. Árvore de decisão para prever o crescimento inicial de *Eucalyptus urograndis*.

Os resultados demonstraram que o único atributo físico capaz de influenciar as altas produções de eucalipto foi a resistência do solo à penetração na camada de 0,20-0,40 m. Para Lima et al. (2010), tal fato dá-se a restrição do sistema radicular do eucalipto que correlaciona

linearmente com a produtividade de madeira. Dos atributos climáticos, apenas a temperatura mínima se faz presente na árvore de decisão, sendo que temperaturas superiores a 18,25 proporcionam maior crescimento do eucalipto. Moraes et al. (2015) ressalta que plantas jovens de eucalipto são mais sensíveis a baixas temperaturas do que as plantas adultas podendo levá-las até a morte devido não estarem completamente estabelecidas no ambiente onde estão plantadas. Atributos relacionados à acidez do solo, acidez potencial na camada de 0,20-0,40 m e pH nas camadas de 0,00-0,20 m e 0,20-0,40 m foram decisivos na geração de regras relacionadas a classe baixa, estando estes atributos presentes em 60% das regras relacionadas a esta classe de crescimento do eucalipto. Lima et al. (2010) observaram que o pH na camada superficial do solo, foi o atributo do solo que mostrou a melhor interação com a produtividade do *Eucalyptus camaldulensis* para as condições do cerrado, podendo ser tomado como o melhor indicador da qualidade do solo avaliado, quando destinado à produtividade de madeira.

CONCLUSÕES:

Os resultados demonstraram que os modelos de classificação gerados para o desenvolvimento inicial de eucalipto foram eficientes para previsão de novos casos em diferentes classes de produção onde, o volume de madeira e o diâmetro altura do peito, os teores de magnésio, fósforo, alumínio, potássio, acidez potencial, pH, resistência do solo à penetração e a temperatura mínima são atributos determinantes para previsão do crescimento do *Eucalyptus urograndis*.

REFERÊNCIAS:

- CRIVELANTI, R. C.; COELHO, R. M.; ADAMI, S. F.; OLIVEIRA, S. R. M. Mineração de dados para inferência de relações solo-paisagem em mapeamentos digitais de solo. **Pesquisa Agropecuária Brasileira**, Brasília, v.44, n.12, p.1707-1715, 2009.
- HAN, J.; PEI, J.; KAMBER, M. **Data mining: concepts and techniques**. 3rd edition. San Francisco: Elsevier. 2011, 673 p.
- LIMA, C. G. R.; CARVALHO, M. P.; NARIMATSU, K. C. P.; SILVA, M. G.; QUEIROZ, H. A. Atributos físico-químicos de um latossolo do cerrado brasileiro e sua relação com características dendrométricas do eucalipto. **Revista Brasileira de Ciência do Solo**, Viçosa, v.34, p.163-173, 2010.
- MORAES, C. B.; CARVALHO, E.V.; ZIMBACK, L.; LUZ, O. S. L.; PIERONI, G. B.; MORI, E. S.; LEAL, T. C. A. B. Variabilidade genética em progênies de meios-irmãos de eucaliptos para tolerância ao frio. **Revista Árvore**, Viçosa, v.39, n.6, p.1047-1054, 2015.
- RAIJ, B.; ANDRADE, J. C.; CANTARELLA, H.; QUAGGIO, J. A. **Análise química para avaliação da fertilidade de solos tropicais**. Campinas: Instituto Agrônomo, 2001. 285 p.
- SANTOS, H.G.; JACOMINE, P.K.T.; ANJOS, L.H.C.; OLIVEIRA, V.A.; LUMBRERAS, J. F.; COELHO, M.R.; ALMEIDA, J.A.; ARAUJO FILHO, J.C.; OLIVEIRA, J.B.; CUNHA, T.J.F. **Sistema Brasileiro de Classificação de Solos**. 5^a Ed. Revista e Ampliada, Brasília: Embrapa, 2018. 353 p.
- STOLF, R.; MURAKAMI, J.H.; BRUGNARO, C.; SILVA, L.G.; SILVA, L.C.F.; MARGARIDO, L.A.C. Penetrômetro de impacto Stolf - programa computacional de dados em EXCEL-VBA. **Revista Brasileira de Ciência do Solo**, v.38, n.3, p.774-782, 2014.
- TEIXEIRA, P.C.; DONAGEMMA, G.K.; FONTANA, A.; TEIXEIRA, W.G. **Manual de Métodos de Análise de Solos**. 3^a edição revista e ampliada, Brasília: Embrapa, 2017. 573p.
- WITTEN, I. H.; FRANK, E.; HALL, M. A. **Data mining: practical machine learning tools and techniques**. São Francisco, CA: The Morgan Kaufmann series in data management systems, 2011. 665 p.