**Florestas aleatórias: predição da produtividade de um *feller-buncher***

**Rafaele Almeida Munis¹, Fernando Reis Andrade1, Danilo Simões2**

1 Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agronômicas, Botucatu, São Paulo; 2 Universidade Estadual Paulista (Unesp), Câmpus Experimental de Itapeva, Itapeva, São Paulo. (rafaele.munis@gmail.com).

**RESUMO:** A análise dos dados gerados a partir das operações florestais, como exemplo, na colheita mecanizada de madeira, assegura aos gestores florestais tomadas de decisões mais assertivas. Neste sentido, foi verificado se o algoritmo de aprendizagem de máquinas por florestas aleatórias era capaz de predizer a produtividade de um *feller-buncher* operando em florestas plantadas com *Eucalyptus*. Para isto foram eleitas como variáveis preditoras as componentes derivadas de volume de madeira colhida, além das que dispendiam de tempo. A performance do modelo nos conjuntos de dados teste, resultou em um coeficiente de determinação, ajustado ao risco, de 57,52%. Ademais, as variáveis com maior impacto na modelagem da produtividade do *feller-buncher* são volume de madeira colhida em metros cúbicos, tempo despendido com o abastecimento e lubrificação, e horas de máquina disponíveis.

**Palavras-chave:** aprendizagem de máquinas, colheita de madeira, telemetria

1. **INTRODUÇÃO**

A utilização de algoritmos preditores faz parte da composição de tecnologias aplicadas na informatização de dados pela indústria 4.0. Destarte, o termo Industria 4.0, foi originalmente destinada na promoção da informatização da manufatura (TANG; VEELENTURF, 2019). Neste contexto, houve ampla informatização dos dados nos setores de manufatura, mais especificamente, nas cadeias produtivas dentro das indústrias (LU et al, 2017).

Exemplo disto, são as indústrias de base florestal, nas quais pode-se observar uma ampla gama de informatização dos processos industriais para o beneficiamento da madeira como matéria prima (MÜLLER et al., 2019). Entretanto, de acordo com uma pesquisa realizada pela Deloitte (2018), ainda se observa um amplo nicho de aplicação destas tecnologias nos demais setores que compõe o processo produtivo, com diversificados fins.

Assim, processos externos às indústrias de base florestal, como a colheita de madeira, também são potenciais a serem desenvolvidos por meio da informatização dos dados gerados à gestores florestais. Portanto, a predição de dados inerentes a colheita de madeira pode ser alcançada com a aplicação de algoritmos de aprendizado de máquina como florestas aleatórias (BREIMAN, 2001). Este algoritmo, consiste em um método de classificação ou regressão baseado na concatenação de um grande número de árvores de decisão, geradas aleatoriamente (WOLFF et al., 2020).

Logo, o suporte que este algoritmo preditor fornece em processos decisórios, justifica-se por promover aos gestores florestais diretrizes na adequação das componentes de sistemas de colheita de madeira em resposta as produtividades que devem ser alcançadas para suprir a demanda industrial. Neste sentido, foi verificado se o algoritmo de aprendizagem de máquinas por florestas aleatórias era capaz de predizer a produtividade de um *feller-buncher* operando em florestas plantadas com *Eucalyptus*.

1. **MATERIAL E MÉTODOS**

*Caracterização da área e da máquina florestal autopropelida*

Foram ponderados coeficientes técnicos de florestas plantadas com *Eucalyptus* em espaçamento de 3,30 m x 1,80 m, em relevo suave ondulado e latossolo. A máquina florestal autopropelida observada foi um *feller-buncher* da marca *Tigercat* e modelo L870C com 11.000 horas de uso acumulada, material rodante de esteiras, com um cabeçote da marca *Tigercat*, modelo ST5702, com 24 polegadas de diâmetro, o qual realizava a derrubada das árvores de *Eucalyptus*.

*Variáveis preditoras e aplicação do algoritmo de florestas aleatórias*

A fim de modelar a produtividade do *feller-buncher* (PR), foram ponderas como variáveis de entrada: horas de máquina disponíveis (HM) e volume de madeira colhida (VM) em metros cúbicos.

Além disto, oito variáveis que ocasionaram diminuição da produtividade, por interrupções operacionais (IO), expressas em horas, foram ponderadas, isto é: abastecimento e lubrificação (IO1); aguardando mecânico interno (IO2); aguardando peças (IO3); manutenção do implemento (IO4); manutenção da máquina base (IO5); manutenção preventiva (IO6); análise de falhas e ajustes (IO7) e lavagem da máquina (IO8). Assim, para a aplicação do algoritmo de florestas aleatórias, foi utilizado o *software* R (R CORE TEAM, 2019), pacote *randomForest*, conseguinte, função *randomForest*, sendo que o número de árvores de regressão foi ajustado em 1.000.

*Importância relativa das variáveis preditoras e validação cruzada*

Foram estimadas as importâncias relativas das variáveis preditoras quando ajustadas a PR e por meio da técnica de *k-fold* de validação cruzada, seguiu-se a divisão aleatória das observações do conjunto de dados em 10 *folds* de tamanho aproximadamente igual. Destarte, os dados foram previamente separados por: 70% para treinamento e 30% para teste.

1. **RESULTADOS E DISCUSSÃO**

Com a aplicação do algoritmo de florestas aleatórias, as variáveis preditoras de PR, com maiores importâncias relativas foram: VM, IO1 e HM (Figura 1). Em consonância com Rocha et al. (2009), a interrupção operacional de abastecimento e lubrificação impacta significativamente a produtividade do *feller-buncher*, sendo que para os autores, ocasiona, consequentemente, baixa eficiência operacional.

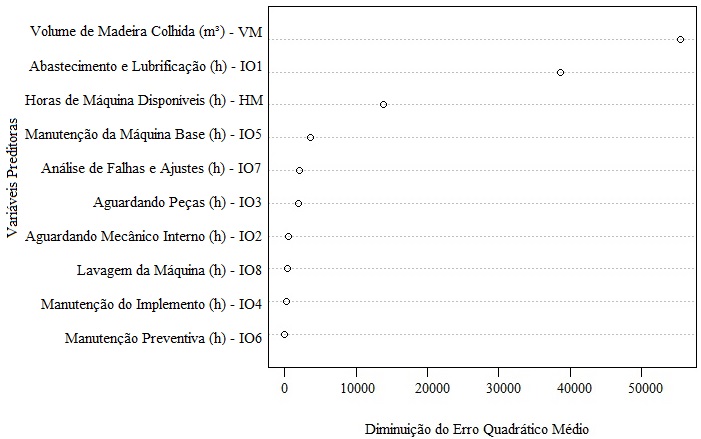


Figura 1. Importância relativa das variáveis preditoras da produtividade do *feller-buncher* em florestas plantadas de *Eucalyptus*.

Ao analisar o poder preditivo do algoritmo de florestas aleatórias para PR, com o modelo eleito na validação cruzada, foi obtido um coeficiente de determinação ajustado ao risco (R²-ajus) de 86,73%, entretanto, R²-ajus de 57,52% na avaliação da performance do modelo com o conjunto de dados teste.

À vista disto, os resultados refletiram o potencial para melhorar o planejamento de processamento de dados de colheita de madeira, sendo que para Müller et al. (2019), apoiados nesta e em outras tecnologias da indústria 4.0, pode-se, porém, processar e ainda testar o planejamento em diversos cenários digitais, antecipadamente.

1. **CONCLUSÕES**

A performance do modelo de florestas aleatórias é preconizada para predizer a produtividade do *feller-buncher* em floresta plantadas com *Eucalyptus*, devido ao coeficiente de determinação ajustado ao risco.

A avaliação da importância relativa das variáveis preditoras da produtividade do *feller-buncher* revela que volume de árvores colhidas, tempo despendido com o abastecimento e lubrificação, além das horas de máquina disponíveis são variáveis com maior impacto na modelagem.

1. **REFERÊNCIAS**

BREIMAN, L. Random Forests. **Machine learning**, v. 45, p. 5–32, 2001.

DELOITTE, I. 2018. **The fourth industrial revolution is here: are you ready**? Disponívelem:<https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/manufacturing/Industry4-0\_Are-you-ready\_Report.pdf> Acesso em: 09/09/2020.

LU, Y.; KAUSHAL, N.; DENIER, N.; WANG, J. S. H. Health of newly arrived immigrants in Canada and the United States: Differential selection on health. **Health & Place**, v. 48, p. 1-10, 2017.

MÜLLER, F.; JAEGER, D.; HANEWINKEL, M. Digitization in wood supply – A review on how Industry 4.0 will change the forest value chain. **Computers and Electronics in Agriculture**, v. 162, p. 206-218, 2019.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2019. Disponível em: <https://www.R-project.org/> Acesso em: 15/09/2020.

ROCHA, E. B.; FIEDLER, N. C.; ALVES, R. T.; LOPES, E. S.; GUIMARÃES, P. P.; PERONI, L. Produtividade e custos de um sistema de colheita de árvores inteiras. **Cerne**, v. 15, n. 3, p. 372-381, 2009.

TANG, C. S.; VEELENTURF, L. P. The strategic role of logistics in the industry 4.0 era. **Transportation Research Part E**, v. 129, p. 1-11, 2019.

WOLFF, S.; O’DONNCHA, F.; CHEN, B. Statistical and machine learning ensemble

modelling to forecast sea surface temperature. **Journal of Marine Systems**, v. 208, 2020.