

## **PREDIÇÃO DA QUALIDADE DE GRÃOS DE ARROZ BENEFICIADOS UTILIZANDO ESPECTROSCOPIA NO INFRAVERMELHO PRÓXIMO E ALGORITMOS DE APRENDIZADO DE MÁQUINA**

**LETÍCIA DE OLIVEIRA CARNEIRO<sup>1</sup>, PAULO CARTERI CORADI<sup>2</sup>, DÁGILA MELO RODRIGUES<sup>3</sup>, ROSANA SANTOS MORAES<sup>4</sup>, MARISA MENEZES LEAL<sup>5</sup>, RENAN IVAN SHULLER<sup>1</sup>**

<sup>1</sup>Estudante de Graduação em Engenharia Agrícola, Bolsista Iniciação Científica, UFSM-CS, tecnicaleticia@gmail.com, renanschller@gmail.com

<sup>2</sup>Eng<sup>o</sup>. Agrícola, Professor Associado, UFSM-CS, Laboratório de Pós-Colheita (LAPOS), paulo.coradi@ufsm.br

<sup>3</sup>Eng<sup>a</sup>. Agrônoma, Doutoranda em Engenharia Agrícola, UFSM/PPGEA, dagila.rodrigues2012@gmail.com

<sup>4</sup>Eng<sup>a</sup>. Agrícola, Doutoranda em Engenharia Agrícola, UFSM/PPGEA, rosana-moraes95@hotmail.com

<sup>5</sup>Eng<sup>a</sup>. Agrícola, Mestranda em Engenharia Agrícola, UFSM/PPGEA, etecmarisa@gmail.com

Apresentado no  
LII Congresso Brasileiro de Engenharia Agrícola - CONBEA 2023  
18 a 21 de outubro de 2023 – Ribeirão Preto - SP, Brasil

**RESUMO:** O estudo teve como objetivo caracterizar e prever a qualidade de grãos de arroz inteiros beneficiados e quebrados. Coletaram-se amostras de grãos de arroz em pontos superiores e inferiores de quatro silos-secadores com capacidade de 40.000 sacas, com teores de água de 16%, 17%, 18% e 19% submetidos à seca-aeração até 12% (b.u.). Em seguida, as amostras de arroz foram analisadas quanto à qualidade físico-química utilizando tecnologia de espectroscopia no infravermelho próximo. A partir dos dados de entrada, utilizaram-se os modelos de aprendizado de máquina para predição da qualidade. Verificou-se que o modelo de Redes Neurais Artificiais (RNAs) e Árvore Aleatória (RandTree) predisseram satisfatoriamente a qualidade dos grãos inteiros e quebrados, sendo o modelo (RandTree) o mais indicado. Conclui-se que a tecnologia de espectroscopia no infravermelho próximo combinada aos algoritmos de aprendizado de máquina é uma alternativa não destrutiva a classificação física manual, para avaliação da qualidade físico-química do arroz.

**PALAVRAS-CHAVE:** Inteligência artificial, Inovações na pós-colheita, Tecnologia não destrutiva.

### **PREDICTING THE QUALITY OF RICE GRAINS PROCESSED USING NEAR INFRARED SPECTROSCOPY AND MACHINE LEARNING ALGORITHMS**

**ABSTRACT:** The study aimed to characterize and predict the quality of processed whole rice grains and broken. Samples of rice grains were collected at the upper and lower points of four silos-driers with a capacity of 40,000 bags, with water contents of 16%, 17%, 18% and 19% submitted to dry-airing up to 12% (w.b.). Next, the rice samples were analyzed for physicochemical quality, using near-infrared spectroscopy technology. From the input data, Machine Learning models were used to predict quality. It was verified that the model of Artificial Neural Networks (ANNs) and Random Tree (RandTree) satisfactorily predicted the quality of whole and broken grains, being the model (RandTree) the most indicated. It is concluded that near-infrared spectroscopy technology combined with Machine Learning algorithms is a non-destructive alternative to manual physical classification for evaluating the physicochemical quality of rice.

**KEYWORDS:** Artificial intelligence, Post-harvest innovations, Non-destructive technology.

**INTRODUÇÃO:** A classificação física é uma operação aplicada nas unidades de recebimento, armazenamento e beneficiamento de grãos para caracterizar fisicamente a qualidade do produto. Esse método convencional de avaliação demanda tempo e mão-de-obra qualificada, onde os resultados muitas vezes são subjetivos, uma vez que a avaliação é visual. Diante da necessidade de tornar a operação mais rápida, assertiva e menos dependente, tecnologias não destrutivas e inteligência computacional pode ser alternativa para caracterização a qualidade dos grãos. Atualmente, dentre as tecnologias disponíveis, a espectroscopia no infravermelho próximo (NIR) poderia ser aplicada para avaliação físico-química dos grãos de arroz beneficiados (MULLER et al., 2021). Ainda, a aplicação de algoritmos de aprendizado de máquina (ML) também tem apresentado resultados satisfatórios para predição da qualidade de sementes (ANDRÉ et al., 2022) e grãos (LUTZ & CORADI, 2022). Diante disto, a medição da qualidade físico-química de grãos de arroz por meio de tecnologias não destrutivas torna possível realizar a predição da qualidade dos grãos aplicando modelos de aprendizado de máquina. Assim, este estudo teve como objetivo caracterizar e prever a qualidade de grãos de arroz inteiros beneficiados e quebrados aplicando tecnologia de espectroscopia no infravermelho próximo e modelos de aprendizado de máquina.

**MATERIAL E MÉTODOS:** Os grãos de arroz em casca da variedade IRGA 424, foram submetidos à secagem em quatro unidades de silo-secadores em escala real, modelo SFP-18314 (Pagé industry, Araranguá, Santa Catarina, Brasil) até alcançar 12% (b.u.). Coletaram-se amostras em 11 pontos diferentes nos quatro silo-secadores (16%, 17%, 18% e 19% de teores de água iniciais), sendo seis pontos superiores e cinco pontos inferiores do lote. As amostras foram beneficiadas no provador de arroz, modelo Paz-1/DTA (Zaccaria company, Limeira, São Paulo, Brasil). Na sequência, os grãos inteiros e quebrados foram separados. A qualidade físico-química dos grãos inteiros e quebrados foi analisada pela tecnologia espectroscopia no infravermelho próximo (NIR). Os dados espectrais foram registrados no modo de refletância na faixa espectral de 400–2500 nm, determinando o teor de amido (A), gordura (GD), cinzas (CZ) e fibra bruta (FB). As variáveis grãos inteiros, quebrados e teores de água foram utilizadas como dados de entrada nos modelos de Rede Neural Artificial (RNAs), Árvore de Decisão (REPTree), Árvore Aleatória (RandTree), Algoritmo M5 de Quinlan (M5P), Floresta Aleatória (RF) e Regressão Linear Múltipla (MLR) para predição da qualidade de grãos, utilizando o software Weka versão 3.9.5. Todos os modelos foram comparados com a regressão linear múltipla. Foi estabelecida uma análise de variância, com agrupamento das médias pelo teste de Scott-knott a 5% de probabilidade e teste de normalidade dos dados (Shapiro-Wilk).

**RESULTADOS E DISCUSSÃO:** O percentual de amido está correlacionado com os percentuais de grãos inteiros. Os teores de cinzas foram menores nos grãos com teores de água iniciais de 18 e 19% (b.u.). Da mesma maneira para a variável fibra bruta, quando os maiores percentuais foram observados em grãos com menores teores de água iniciais (NUNES et al., 2022). Na avaliação de amido, o coeficiente de correlação foi semelhante entre os modelos RNAs e RLM (0,8251 e 0,8169, respectivamente), com melhor ajuste para RNAs, devido ao MAE ser inferior ao modelo de RLM (0,6235 e 0,7657, respectivamente) (Fig.1A). Para a análise de cinza, os modelos de RNAs, ReepTree e RandTree apresentaram correlações semelhantes em r e MAE (0,5125, 0,5160, 0,5160 e 0,0621, 0,0620, 0,0620, respectivamente) (Fig.1B).

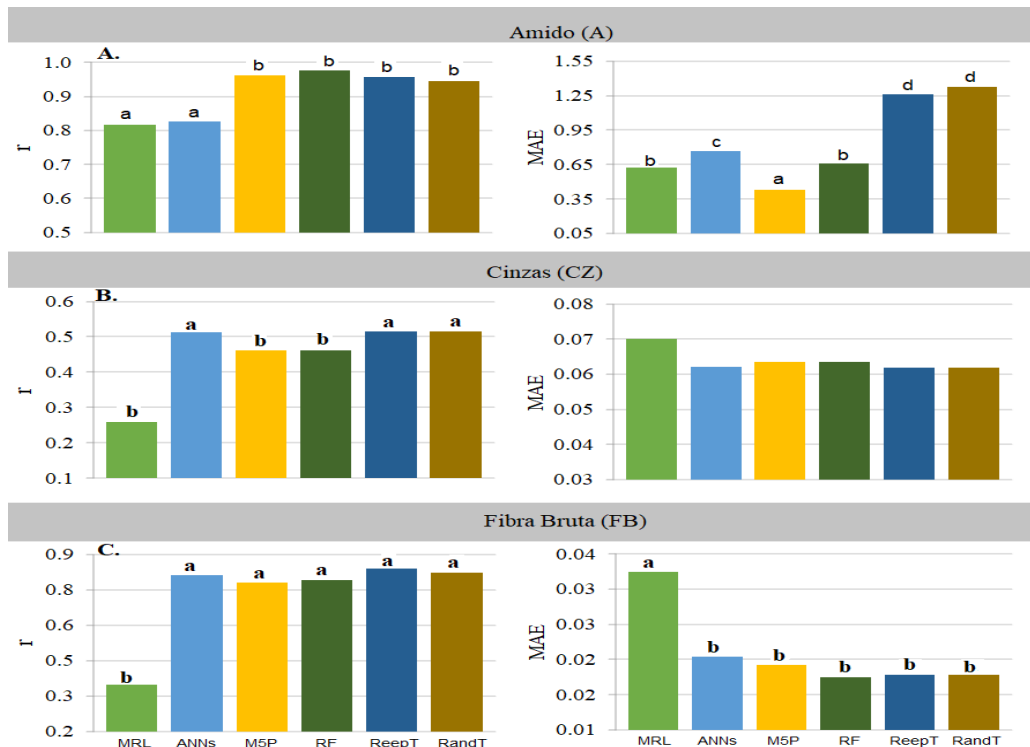


FIGURA 1. Ajustes obtidos pelo coeficiente de correlação ( $r$ ) entre os valores observados e preditos por cada modelo de Machine Learning e o erro médio absoluto (MAE) dos valores preditos em relação aos valores observados na predição de amido-A (A.), teores de cinzas-CZ (B.) e fibra bruta-FB (C.) em grãos inteiros. Rede Neural Artificial (RNAs), Árvore de Decisão (REPTree), Árvore Aleatória (RandTree), Algoritmo M5 de Quinlan (M5P), Floresta Aleatória (RF) e Regressão Linear Múltipla (MLR).

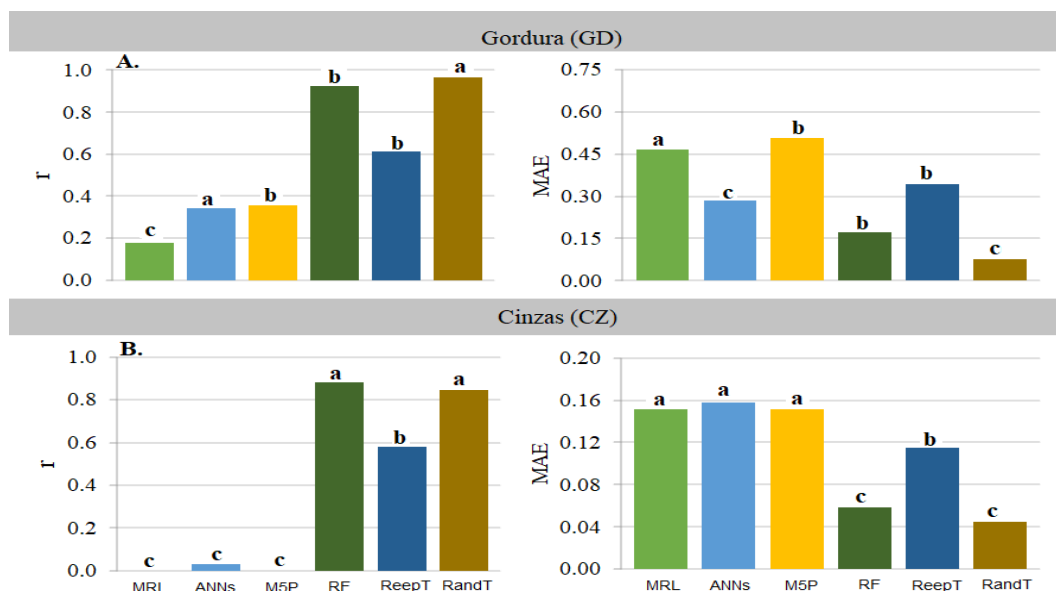


FIGURA 2. Ajustes obtidos pelo coeficiente de correlação ( $r$ ) entre os valores observados e preditos por cada modelo de aprendizado de máquina e o erro médio absoluto (MAE) dos valores preditos em relação aos valores observados na predição de gordura-GD (A) e teores de cinzas-CZ (B.) em grãos quebrados. Rede Neural Artificial (RNAs), Árvore de Decisão (ReepTree), Árvore Aleatória (RandTree), Algoritmo M5 de Quinlan (M5P), Floresta Aleatória (RF) e Regressão Linear Múltipla (MLR).

Na avaliação de fibra bruta, verificou-se que os modelos RNAs, FA e RandTree tiveram melhores ajustes (Fig.1C). Os grãos quebrados tiveram correlação positiva entre as análises de gordura e cinzas. Foi verificado que quanto maior o teor de água dos grãos de arroz, maiores foram os percentuais de quebrados no beneficiamento, e consequentemente menores foram os percentuais de gordura e cinzas. Na predição de gordura no grão de arroz quebrado, verificou-se que o modelo de Árvore de Decisão Aleatória (RandTree) e as Redes Neurais Artificiais (RNAs), apresentaram as maiores correlações, com valores de r e MAE (0,7325, 0,6490 e 0,0415, 0,0614, respectivamente) (Fig. 2A). Enquanto que, para avaliação de cinzas foi verificado que os modelos de Floresta Aleatória e Árvore de Decisão Aleatória (RandTree) tiveram os melhores resultados.

**CONCLUSÕES:** A aplicação da espectroscopia no infravermelho próximo com os modelos de aprendizado de máquina caracterizou satisfatoriamente a composição físico-química de grãos de arroz inteiros e quebrados, sendo uma alternativa ao método convencional de classificação física. O modelo de árvore aleatória (RandTree) foi o modelo indicado para prever a qualidade físico-química nos grãos de arroz inteiros e quebrados.

**AGRADECIMENTOS:** Os autores agradecem a Universidade Federal de Santa Maria, Programa de Pós-Graduação em Engenharia Agrícola, ao Research Group at Postharvest Innovation: Technology, Quality & Sustainability (UFSM), Laboratório de Pós-Colheita (LAPOS-UFSM), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) pela concessão de bolsas de estudos, recursos financeiros e espaços físicos para desenvolvimento dos experimentos.

## REFERÊNCIAS

- ANDRÉ, G. S., CORADI, P. C., TEODORO, L. P. R., & TEODORO, P. E. Predicting the quality of soybean seeds stored in different environments and packaging using machine learning. **Scientific Reports**, v. 12, n. 1, p. 1-13, 2022.
- LUTZ, É., & CORADI, P. C. Applications of new technologies for monitoring and predicting grains quality stored: sensors, internet of things, and artificial intelligence. **Measurement**, v. 188, n. 110609, 2022.
- MÜLLER, A., CORADI, P. C., NUNES, M. T., GROHS, M., BRESSIANI, J., TEODORO, P. E., ANSCHAU, K. F., & FLORES, E. M. M. Effects of cultivars and fertilization levels on the quality of rice milling: a diagnosis using near-infrared spectroscopy, x-ray diffraction, and scanning electron microscopy. **Food Research International**, v. 147, n. 110524, 2021.
- NUNES, M. T., CORADI, P. C., MULLER, A., CARNEIRO, L. O., STEINHAUS, J. I., ANSCHAU, K. F., SOUZA, G. C., MÜLLER, E. I., TEODORO, P. E., & DUTRA, A. P. Stationary rice drying: influence of initial moisture contents and impurities in the mass grains on the physicochemical and morphological rice quality. **Journal of Food Processing and Preservation**, v. 46, n. 5, p. 1-10, 2022.