

DESEMPENHO DE ALGORITMOS DE MACHINE LEARNING NA ESTIMATIVA DE NITROGÊNIO DO FEIJÃO COMUM A PARTIR DA LEITURA INDIRETA DE CLOROFILA

MARCOS SILVA TAVARES ¹, JAMILE RAQUEL REGAZZO ², THIAGO LIMA DA SILVA ³, MURILO DE LIMA LEÃO ⁴, MURILO MESQUITA BAESSO ⁵

¹ Doutorando em Engenharia de Sistemas Agrícolas, Depto.de Engenharia de Biossistemas, Escola Superior de Agricultura “Luiz de Queiroz”, USP, Piracicaba – SP, ms.tavares@usp.br.

² Doutoranda em Engenharia de Sistemas Agrícolas, Depto.de Engenharia de Biossistemas, ESALQ/USP, Piracicaba – SP.

³ Doutorando em Engenharia de Sistemas Agrícolas, Depto.de Engenharia de Biossistemas, ESALQ/USP, Piracicaba – SP.

⁴ Graduando em Engenharia de Biossistemas, Depto.de Engenharia de Biossistemas, FZEA/USP, Pirassununga – SP.

⁵ Professor associado, Depto.de Engenharia de Biossistemas, FZEA/USP, Pirassununga – SP.

Apresentado no
LIII Congresso Brasileiro de Engenharia Agrícola - CONBEA 2024
6 a 8 de agosto de 2024 – Natal – RN, Brasil

RESUMO: A cultura do feijão é a fonte proteica mais consumida pelos humanos no mundo. Graças à crescente demanda global por alimentos, a adubação nitrogenada é crucial para obtenção de altas produtividades. A quantificação do nitrogênio foliar é fundamental para a otimização da adubação. No entanto, métodos tradicionais são demorados e caros, o que eleva a necessidade de formas rápidas e não destrutivas para estimar N foliar. O objetivo deste trabalho foi avaliar o desempenho de 2 algoritmos de aprendizado de máquina na predição foliar de N com uso de medidas indiretas de clorofila. A cultivar BRS FC104 foi submetida a 4 níveis de adubação nitrogenada (0, 40, 80 e 120 kg ha⁻¹) em delineamento inteiramente casualizado. As leituras de clorofila foram realizadas 32 DAS com clorofilômetro Falker CFL 1030. Os algoritmos MultiLayerPerceptron e Random Forest foram implementados no software Weka 3.8.6. O nitrogênio influenciou (P<0,01) as variáveis CA, CT e teor de nitrogênio foliar. O Random Forest mostrou-se mais apropriado para predição de N foliar a partir de medidas indiretas de clorofila, com erro de 1,97 g kg⁻¹ e R² = 0,749, superando a rede neural artificial (R² = 0,538 e erro de 2,88 g kg⁻¹).

PALAVRAS-CHAVE: modelos, predição, *phaseolus vulgaris*

PERFORMANCE OF MACHINE LEARNING ALGORITHMS IN ESTIMATING COMMON BEAN NITROGEN FROM INDIRECT CHLOROPHYLL READING

ABSTRACT: Bean crops are the most consumed protein source by humans in the world. Thanks to the growing global demand for food, nitrogen fertilizer is crucial for obtaining high productivity. Quantifying leaf nitrogen is essential for optimizing fertilization. However, traditional methods are time-consuming and expensive, which raises the need for rapid and non-destructive ways to estimate foliar N. The objective of this work was to evaluate the performance of 2 machine learning algorithms in foliar N prediction using indirect chlorophyll measurements. The cultivar BRS FC104 was subjected to 4 levels of nitrogen fertilization (0, 40, 80 and 120 kg ha⁻¹) in a completely randomized design. Chlorophyll readings were performed 32 DAS with a Falker CFL 1030 chlorophyllometer. The MultiLayerPerceptron and Random Forest algorithms were implemented in the Weka 3.8.6 software. Nitrogen influenced (P<0.01) the variables CA, CT and leaf nitrogen content.

Random Forest proved to be more appropriate for predicting leaf N from indirect chlorophyll measurements, with an error of 1.97 g kg⁻¹ and R2 = 0.749, surpassing the artificial neural network (R2 = 0.538 and error of 2.88 g kg⁻¹).

KEYWORDS: models, prediction, *phaseolus vulgaris*

INTRODUÇÃO: O feijão comum (*Phaseolus vulgaris* L.) é uma espécie de leguminosa que contém excelentes características nutricionais, sendo de grande importância para países em desenvolvimento devido ao fato de que é a principal fonte proteica básica e de baixo custo (BALDIN et al., 2021). A produtividade do feijoeiro depende da fertilidade do solo (CARVALHO et al., 2018). O macronutriente mais exigido por essa cultura é o nitrogênio que possui grande dinâmica no sistema solo-planta-atmosfera devido à variabilidade dos processos de volatilização, mineralização, lixiviação e absorção pela cultura (MAIA et al., 2012). Nesse sentido, a quantificação foliar desse elemento é crucial para assegurar uma adubação nitrogenada precisa e eficiente, evitando perdas, favorecimento de pragas e doenças e contaminação do solo e da água. No entanto, análises convencionais são demoradas quando se considera o ciclo fenológico médio de 65 dias, o que motiva a comunidade acadêmica à busca de metodologias mais rápidas. A clorofila pode ser uma medida indireta de nitrogênio foliar por se correlacionar com o N, podendo ser mensurada por meio do clorofilômetro portátil (SAMBORSKI et al., 2009; BASSOI et al., 2022). Avanços tecnológicos profundos em relação à predição de produtividade ocorreram no cenário agrícola, especialmente após o uso de machine learning (ML). Entre vários algoritmos de ML supervisionados, redes neurais, algoritmos de regressão linear, RF e SVM são amplamente utilizados em estudos de estimativa de rendimento de culturas (VAN KLOMPENBURG et al., 2020). Este trabalho teve como objetivo avaliar o desempenho de 2 modelos de machine learning na estimativa de nitrogênio foliar da cultura do feijão por meio de leitura indireta de clorofila.

MATERIAL E MÉTODOS: O experimento foi conduzido em casa de vegetação em recipientes plásticos de 5L, preenchidos com substrato composto por neossolo quartzarênico, palhada de cana triturada e esterco bovino curtido nas proporções 3:2:2. O solo utilizado possuía as seguintes características: pH (CaCl₂): 4.4, P (res): 8 mg dm⁻³, S: 4 mg dm⁻³; K (res), Ca, Mg, Al e H+Al = 0,3, 16, 4, 4, 30 mmolc dm⁻³, respectivamente, e textura arenosa. Foi utilizado o Delineamento Inteiramente casualizado (DIC). A cultivar BRS FC104 foi submetida a 4 níveis de adubação nitrogenada (0 (T1), 40 (T2), 80 (T3) e 120 (T4) kg ha⁻¹ de N, correspondentes a 0, 40, 80 e 120 da dose recomendada). 2 sementes foram colocadas por vaso, espaçadas em 10cm. A irrigação foi feita com uma lâmina diária de 6mm. Foram realizadas 2 leituras de clorofila por tratamento com 3 repetições, totalizando 24 observações. As leituras de clorofila A, clorofila B e clorofila total foram feitas com um clorofilômetro digital, modelo CFL1030 da Falker, aos 32 dias após semeadura. Uma planilha com 5 colunas foi criada contendo dose aplicada, as 3 variáveis de clorofila e a última coluna com os valores de N mensurados em laboratório. Os algoritmos MultiLayerPerceptron (ANN) e Random Forest foram implementados no software Weka 3.8.6; Weka é uma coleção de algoritmos de aprendizado de máquina usado na mineração de dados, possibilitando pré-processamento, classificação, clustering, regressão, associação e visualização de dados (VILLAVICENCIO et al., 2021). O Random Forest e a ANN foram configurados com 100 batchsize e validação cruzada com 10 folds. O coeficiente de determinação (R2) e a raiz quadrática do erro médio (RMSE) foram as métricas usadas para análise do desempenho dos modelos e foram obtidas pelas equações:

$$R2 = 1 - \frac{\sum(Y_i - \hat{P}_i)^2}{\sum(Y_i - \bar{x})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum(Y_i - \hat{P}_i)^2}{n}} \quad (2)$$

em que, \hat{y}_i representa o valor previsto do modelo de regressão, \bar{x} representa o valor médio do valor medido, y_i representa o valor medido e n representa o tamanho da amostra.

RESULTADOS E DISCUSSÃO: O fator nitrogênio influenciou significativamente os valores médios de clorofila A e Clorofila Total, sendo os maiores resultados observados quando as plantas foram submetidas as maiores doses nas duas variáveis (Tabela 1). O conteúdo de nitrogênio foliar variou em função das doses de nitrogênio; o menor valor, 36,6 g kg⁻¹, foi observado na ausência de N (0 kg ha⁻¹) e o maior valor, 46,2 g kg⁻¹, foi obtido quando foi aplicado 120 kg ha⁻¹ de N.

TABELA 1. Síntese dos valores de análise de variância e do teste de médias para as variáveis Clorofila A, Clorofila B, Clorofila Total e Nitrogênio.

FATOR	Clorofila A (ICF)	Clorofila B (ICF)	Clorofila Total (ICF)	Nitrogênio (g kg ⁻¹)
Nitrogênio (N)				
0	33,7 b	7,5 a	39,1 b	36,6 c
40	35,2 b	7,9 a	41,7 b	40,1 b
80	39,5 a	7,4 a	46,8 a	41,4 b
120	43,6 a	6,5 a	50,1 a	46,2 a
Valor de F	24,3 **	1,10 NS	25,9 **	23,2 **
C.V. (%)	8,6	19,8	9,4	10,9

NS: não significativo (P>0,05); *: significativo (P<0,05); **: significativo (P<0,01); C.V.: coeficiente de variação. Médias seguidas pela mesma letra não diferem entre si pelo teste de Tukey a 5% de probabilidade.

A clorofila A e a Clorofila Total demonstraram alta correlação positiva em relação ao N, 0,89 e 0,87, respectivamente (Figura 1 - A). O desdobramento da análise de regressão para o teor foliar de N, dentro de cada nível da adubação nitrogenada, expressou tendência linear com alto coeficiente de determinação, R² = 0,95 (Figura 1 - B).

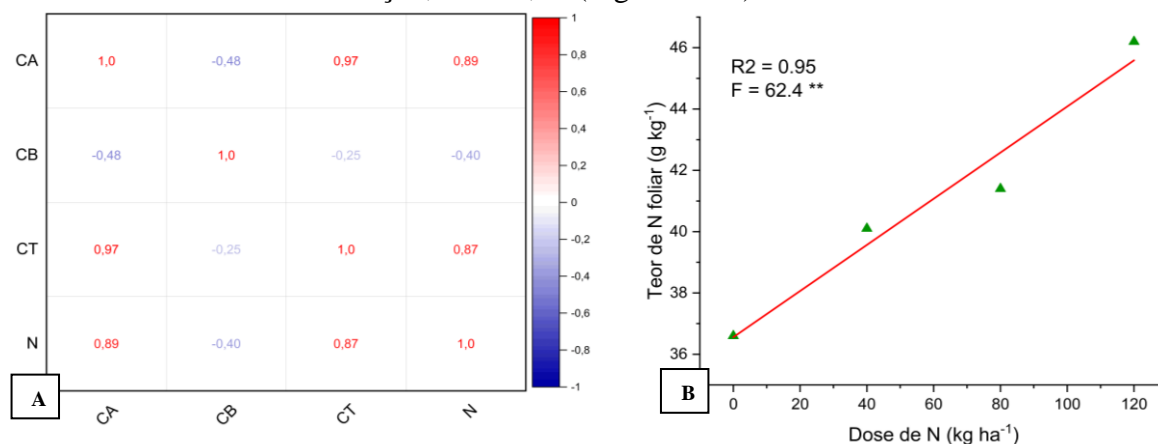


Figura 1 – A: Correlação de Pearson para as variáveis clorofila A (CA), clorofila B (CB), clorofila total (CT) e nitrogênio foliar (N); B: Teor foliar de nitrogênio foliar em função da adubação nitrogenada

Esses dados assemelham-se aos observados por (HAIM et al., 2012) que verificaram correlação de 0,79 entre N foliar e clorofila na cultivar ICA Pijao. O Random Forest expressou melhor resultado preditivo com baixo erro, RMSE = 1,97 (Figura 2 - A). O uso da rede neural artificial elevou o erro, RMSE = 2,88, e teve um desempenho 32% menor quando comparada ao modelo Random Forest (Figura 2 - B).

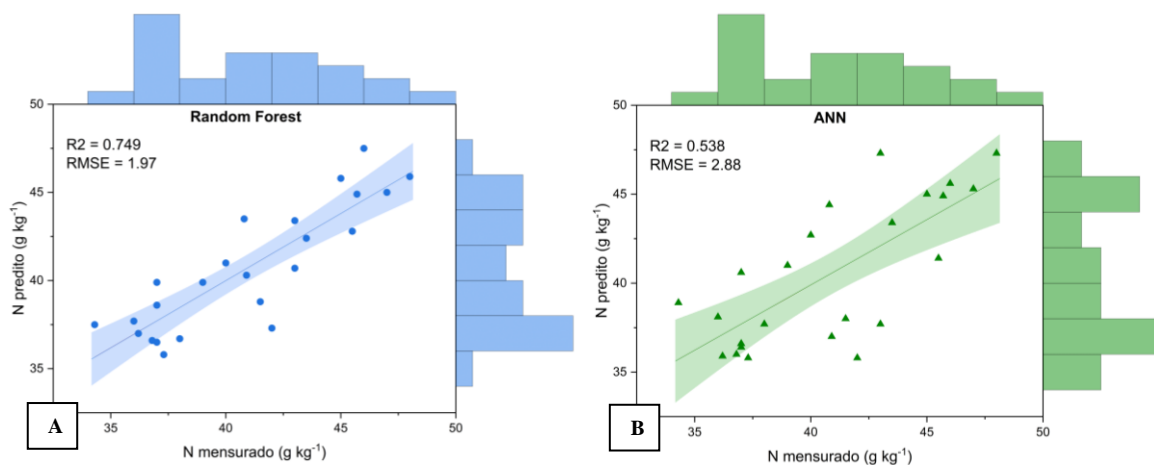


Figura 2 – A: Predição de nitrogênio com Random Forest; B - Predição de nitrogênio com Rede Neural Artificial

CONCLUSÕES: A leitura indireta de clorofila mostrou-se altamente correlacionada com o teor foliar de nitrogênio na cultivar BRS FC 104. O Random Forest ($R^2 = 0,749$, $RMSE = 1,97$) superou a rede neural artificial ($R^2 = 0,538$, $RMSE = 2,88$) na estimativa de N foliar. Novos estudos envolvendo algoritmos de aprendizado de máquina são demandados para consolidação de técnicas de predições de N foliar com uso de CA, CB e CT como variáveis preditoras.

AGRADECIMENTOS: A FEALQ pelo auxílio financeiro para apresentação do trabalho.

REFERÊNCIAS:

BALDIN, R. C., KAVALCO, S. A. F., WOYANN, L. G., RODRIGUES, A. A., GOBATTO, D. R., SILVA, G. R. D., ... & FINATTO, T. (2021). Yield stability of common bean genotypes in the state of Santa Catarina, Brazil. *Pesquisa Agropecuária Brasileira*, 56, e02302.

BASSOI, L. H., da SILVA, T. M. M., COSTA, B. R. S., OLDONI, H., & FOSCHINI, M. (2022). **Calibração de medidor portátil para estimativa de teor foliar de N em função de zonas de vigor vegetativo de videiras fertirrigadas.** In: CONGRESSO BRASILEIRO DE AGRICULTURA DE PRECISÃO-ConBAP, 2022, Campinas.

CARVALHO, M. D., NASCENTE, A. S., FERREIRA, G. B., MUTADIUA, C. A., & DENARDIN, J. E. (2018). Phosphorus and potassium fertilization increase common bean grain yield in Mozambique. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 22, 308-314.

HAIM, P. G., ZOFFOLI, B. C., ZONTA, E., & ARAÚJO, A. P. (2012). Diagnose nutricional de nitrogênio em folhas de feijoeiro pela análise digital de imagens. *Pesquisa Agropecuária Brasileira*, 47, 1546-1549.

MAIA, S. C. M., SORATTO, R. P., NASTARO, B., & FREITAS, L. B. D. (2012). The nitrogen sufficiency index underlying estimates of nitrogen fertilization requirements of common bean. *Revista Brasileira de Ciência do Solo*, 36, 183-192.

SAMBORSKI, S. M., TREMBLAY, N., & FALLON, E. (2009). Strategies to make use of plant sensors-based diagnostic information for nitrogen recommendations. *Agronomy journal*, 101(4), 800-816.

VAN KLOMPENBURG, T., KASSAHUN, A. e CATAL, C. (2020). Previsão do rendimento agrícola usando aprendizado de máquina: uma revisão sistemática da literatura. *Computadores e Eletrônicos na Agricultura*, 177, 105709.

VILLAVICENCIO, C. N., MACROHON, J. J. E., INBARAJ, X. A., JENG, J. H., & HSIEH, J. G. (2021). Covid-19 prediction applying supervised machine learning algorithms with comparative analysis using weka. *Algorithms*, 14(7), 201.