

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL  
CÂMPUS DE CHAPADÃO DO SUL  
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA

ALESSANDRA RODRIGUES DOS SANTOS

**SENSOR HIPERESPECTRAL PARA A PREDIÇÃO DE NITROGÊNIO  
FOLIAR, PIGMENTOS E FOTOSSÍNTESE LÍQUIDA NA CULTURA  
DO MILHO**

CHAPADÃO DO SUL – MS

2024

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL  
CÂMPUS DE CHAPADÃO DO SUL  
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA

ALESSANDRA RODRIGUES DOS SANTOS

**SENSOR HIPERESPECTRAL PARA A PREDIÇÃO DE NITROGÊNIO  
FOLIAR, PIGMENTOS E FOTOSSÍNTESE LÍQUIDA NA CULTURA  
DO MILHO**

Orientador: Prof. Dr. Cid Naudi Silva Campos

Dissertação apresentada à Universidade Federal de Mato Grosso do Sul, como requisito para obtenção do título de Mestre em Agronomia, área de concentração: Produção Vegetal.

CHAPADÃO DO SUL – MS

2024



Serviço Público Federal  
Ministério da Educação  
Fundação Universidade Federal de Mato Grosso do Sul



## PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA

### CERTIFICADO DE APROVAÇÃO

**DISCENTE:** Alessandra Rodrigues dos Santos

**ORIENTADOR:** Dr. Cid Naudi Silva Campos

**TÍTULO:** Sensor hiperespectral para a predição de nitrogênio foliar, pigmentos e fotossíntese líquida na cultura do milho

#### **AVALIADORES:**

Prof. Dr. Cid Naudi Silva Campos

Pro. Dra. Rita de Cassia Felix Alvarez

Prof. Dra. Dthenifer Cordeiro Santana

Chapadão do Sul, 08 de julho de 2024.

NOTA  
MÁXIMA  
NO MEC

UFMS  
É 10!!!



Documento assinado eletronicamente por **Cid Naudi Silva Campos, Professor do Magisterio Superior**, em 08/07/2024, às 15:51, conforme horário oficial de Mato Grosso do Sul, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

NOTA  
MÁXIMA  
NO MEC

UFMS  
É 10!!!



Documento assinado eletronicamente por **Rita de Cassia Felix Alvarez, Professora do Magistério Superior**, em 08/07/2024, às 15:51, conforme horário oficial de Mato Grosso do Sul, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

NOTA  
MÁXIMA  
NO MEC

UFMS  
É 10!!!



Documento assinado eletronicamente por **Dthenifer Cordeiro Santana, Usuário Externo**, em 08/07/2024, às 21:12, conforme horário oficial de Mato Grosso do Sul, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufms.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufms.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4912698** e o código CRC **9B7FDC31**.

---

### **COLEGIADO DE PÓS-GRADUAÇÃO EM AGRONOMIA**

Rod MS 306, Km 105, Caixa Postal 112

Fone: (67)3562-6351

CEP 79560-000 - Chapadão do Sul - MS

---

**Referência:** Processo nº 23455.000267/2024-30

SEI nº 4912698

## DEDICATÓRIA

*“Dedico este trabalho primeiramente a Deus, o maior orientador da minha vida, autor de meu destino, meu guia, socorro presente na hora da angústia. Ele nunca me abandonou nos momentos de necessidade.”*

*“Pelo amor, carinho, dedicação, cuidado e incentivo que meus pais e meus irmãos me deram durante toda a minha existência, dedico esta monografia a eles. Com muita gratidão.”*

*“Dedico esta dissertação ao meu marido Weslen que foi capaz de suportar todos os meus momentos de estresse durante o processo. Com muita gratidão no coração por fazer parte da minha vida, meu amor.”*

*“A verdadeira motivação vem de realização, desenvolvimento pessoal, satisfação no trabalho e reconhecimento” (Frederick Herzberg).*

## **AGRADECIMENTOS**

Em primeiro lugar, agradeço a Deus, que fez com que meus objetivos fossem alcançados, durante todos os meus anos de estudos e, por ter permitido que eu tivesse saúde e determinação para não desanimar durante a realização deste trabalho.

Sou grata aos meus pais Ademir dos Santos Lima e Rosana Conceição Rodrigues, pelo apoio e incentivo que serviram de alicerce para as minhas realizações; pelo amor incondicional na minha vida sempre, por acreditarem que eu seria capaz de superar os obstáculos que a vida me apresentou e, por todo o esforço investido na minha educação. Esta monografia é a prova de que tais esforços não foram em vão e valeram a pena. Amo vocês.

Aos meus irmãos Amanda Rodrigues de Lima, Alice Fernanda Rodrigues Lima e Ryan Vinícius Rodrigues Lima, pela amizade e atenção dedicadas quando sempre precisei, por estarem ao meu lado e me fazer ter confiança nas minhas decisões. Amo vocês.

Ao meu marido Weslen Fernando Ferreira da Cunha que acima de tudo é um grande amigo, além de cuidar da manutenção do lar enquanto eu permanecia ocupada com este projeto, foi capaz de me incentivar todos os dias. Grata por me ajudar a realizar este sonho, esta é uma das muitas conquistas ao seu lado. Te amo.

À minha família e amigos que me apoiaram; ao Pastor Jonas Ricardo de Lima e aos meus irmãos da fé que oraram e acreditaram em mim sempre.

Ao professor Cid Naudi Silva Campos, por ter sido meu orientador e, desempenhado tal função com dedicação e amizade; pelos conhecimentos imprescindíveis durante minha trajetória no grupo GECENP (Grupo de Estudos do Cerrado em Nutrição de Plantas), ao qual tive a honra de ser integrante.

À minha amiga mestre Marcia Leticia Monteiro Gomes, ótima companheira de trabalho, agradeço por partilhar seus conhecimentos comigo e, sempre me apoiar e ajudar na minha pesquisa de dissertação, por me distrair um pouquinho (com refeições deliciosas) quando eu só ficava estudando.

À instituição de ensino Universidade Federal de Mato Grosso do Sul juntamente com o corpo docente, que foram essenciais no meu processo de formação profissional, pela dedicação, e por tudo o que aprendi ao longo dos anos do curso.

Enfim, a todos aqueles que contribuíram, de alguma forma, para a realização deste trabalho.

## EPÍGRAFE

*“Seja forte, mas lembre-se que ser forte não é suportar tudo, o tempo todo, sem ajuda de ninguém. Ser forte é, apesar de todas as suas vulnerabilidades, seguir tentando. Ser forte não é não sentir medo diante do desconhecido, mas saber que é necessário buscar o caminho do aprendizado, da superação e que, ao se cercar das pessoas certas, você chega muito mais longe. ”*

*Kálita Aquino*

## LISTA DE FIGURAS

Figura 1. Condições meteorológicas no decorrer da safra 2ª de 2023, entre o período de 10/02 a 07/06.....	14
Figura 2. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson (r) (a) e erro médio absoluto (MAE) (b) para os modelos de aprendizagem de máquina (ML) na predição de nitrogênio foliar (NF).....	21
Figura 3. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson (r) (a) e erro médio absoluto (MAE) (b) para os modelos de aprendizagem de máquina (ML), e médias de r para os inputs (entradas) testados (c) na predição de clorofila <i>a</i> (Chl <i>a</i> ).....	22
Figura 4. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson (r) (a) e erro médio absoluto (MAE) (c) para os modelos de aprendizagem de máquina (ML), e médias de r para os inputs (entradas) testados (b) na predição de clorofila <i>b</i> (Chl <i>b</i> ).....	23
Figura 5. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson (r) (c) e erro médio absoluto (MAE) (b) para os modelos de aprendizagem de máquina (ML), e médias de r para os inputs (entradas) testados (a) na predição de clorofila total (Chl <i>a</i> + <i>b</i> ).....	24
Figura 6. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson (r) (c) e erro médio absoluto (MAE) (b) para os modelos de aprendizagem de máquina (ML), e médias de r para os inputs (entradas) testados (a) na predição de carotenoides (Cx + c).....	25
Figura 7. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson (r) (a) para os modelos de aprendizagem de máquina (ML), e erro médio absoluto (MAE) (b) para o desdobramento da interação significativa entre ML e os inputs (entradas) testados na predição de fotossíntese líquida (A).....	26

## LISTA DE TABELAS

Tabela 1. Intervalos de faixa espectral (nm) e comprimento de onda médio (nm) usados no estabelecimento das 28 bandas espectrais.....	17
Tabela 2. Análise de variância para coeficientes de correlação de Pearson (r) e erro médio absoluto (MAE) por diferentes entradas e modelos de aprendizagem de máquina (ML) testados.....	20

## RESUMO

A predição de teor de nitrogênio foliar (NF), pigmentos e, fotossíntese líquida (A) em folhas de milho, por meio de dados hiperespectrais e modelos de aprendizagem de máquina (ML) pode se tornar ferramenta essencial no diagnóstico rápido dos efeitos das adubações nitrogenadas. Objetivou-se avaliar a acurácia dos algoritmos de ML na predição do teor de N, pigmentos e taxa fotossintética em plantas de milho, utilizando como base os dados hiperespectrais. A condução do experimento foi realizada na Universidade Federal de Mato Grosso do Sul, Câmpus Chapadão do Sul na 2ª safra de 2023. O delineamento experimental foi em blocos casualizados, em que os tratamentos consistiram de quatro doses de N: N1 (0%), N2 (30% - 54 kg ha<sup>-1</sup> de N), N3 (60% - 108 kg ha<sup>-1</sup> de N) e N4 (120% - 216 kg ha<sup>-1</sup> de N). No estágio V6 foi realizada a análise espectral de seis amostras foliares de cada parcela utilizando espectrorradiômetro disponibilizando as faixas de 350 a 2500 nm. Obtido os comprimentos de onda, eles foram agrupados em médias de intervalos representativos em bandas. Além da determinação do conteúdo de N e pigmentos, foi realizada a análise de fotossíntese (A). Os dados foram submetidos aos modelos de ML: REPTree (DT), Árvore de Decisão M5P (M5P), Floresta Aleatória (RF), Redes Neurais Artificiais (RNA), Máquina de Vetor Suporte (SVM) e ZeroR (ZR) (controle). Foram testadas duas configurações de entrada: usando apenas os comprimentos de onda na predição das variáveis (ALL) e utilizando as bandas espectrais (SB). Como variáveis de saída foram utilizados: teor de NF, pigmentos (clorofila e carotenoides) e a variável fisiológica A. Como avaliação da acurácia dos algoritmos foram utilizados coeficiente de correlação de Pearson (r) e erro absoluto médio (MAE). O algoritmo SVM obteve o melhor desempenho com maior correlação (r) e menor erro absoluto médio (MAE) entre valores previstos e observados para todas as variáveis exceto carotenoides, com o valor do coeficiente de correlação (r) acima de 0,6, e o erro abaixo de 3,0. A árvore de decisão M5P prediz com eficiência todas as variáveis, exceto clorofila *a*, alcançando um valor de r em torno de 0,6 e o erro abaixo de 0,3. O algoritmo RF foi capaz de prever o conteúdo de N, carotenoides e A, com valor de r em torno de 0,6. Esses três modelos teve uma melhor performance na predição das variáveis usando dados de banda hiperespectral. Dessa forma os algoritmos SVM, árvore de decisão M5P e RF apresentam-se eficazes na predição de teor de NF, pigmentos e A em milho, especialmente quando utilizou-se SB como dados de entrada.

**Palavras-chave:** Aprendizagem de Máquina. Trocas Gasosas. Máquina de Vetor Suporte. Sensoriamento Remoto. *Zea mays* L.

## ABSTRACT

The prediction of leaf nitrogen (NF) content, pigments and net photosynthesis (*A*) in maize leaves using hyperspectral data and machine learning (ML) models could become an essential tool in the rapid diagnosis of the behavior of these pigments in the face of nitrogen fertilization. The aim was to evaluate the accuracy of ML algorithms in predicting N content, pigments and photosynthetic rate in maize plants, using hyperspectral data as a basis. The experiment was carried out at the Federal University of Mato Grosso do Sul, Câmpus Chapadão do Sul in the second harvest of 2023. The experimental design was randomized blocks, in which the treatments consisted of four doses of N: N1 (0%), N2 (30% - 54 kg ha<sup>-1</sup> de N), N3 (60% - 108 kg ha<sup>-1</sup> de N) and N4 (120% - 216 kg ha<sup>-1</sup> de N). At the V6 stage, spectral analysis was carried out on six leaf samples from each plot using a spectroradiometer= providing the 350 to 2500 nm bands. Once the wavelengths had been obtained, they were grouped into averages of representative band intervals. In addition to determining N and pigment content, photosynthesis analysis was carried out. The data was submitted to ML models: REPTree (DT), M5P Decision Tree (M5P), Random Forest (RF), Artificial Neural Networks (ANN), Support Vector Machine (SVM) and ZeroR (ZR) (control). Two input configurations were tested: using only the wavelengths to predict the variables (ALL) and using the spectral bands (SB). The following were used as output variables: NF content, pigments (chlorophyll and carotenoids) and the physiological variable *A*. Pearson's correlation coefficient (*r*) and mean absolute error (MAE) were used to assess the accuracy of the algorithms. The SVM algorithm obtained the best performance with the highest correlation (*r*) and lowest mean absolute error (MAE) between predicted and observed values for all variables except carotenoids, with the value of the correlation coefficient (*r*) above 0.6, and the error below 3.0. The M5P decision tree efficiently predicts all variables except chlorophyll *a*, achieving an *r* value of around 0.6 and an error below 0.3. The RF algorithm was able to predict the content of N, carotenoids and *A*, with an *r*-value of around 0.6. These three models performed best in predicting the variables using hyperspectral band data. Thus, the SVM, M5P decision tree and RF algorithms were effective in predicting NF, pigment and *A* content in maize, especially when SB was used as input data.

**Keywords:** Machine Learning. Gas Exchange. Support Vector Machine. Remote Sensing.  
*Zea mays* L.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	12
<b>2</b>	<b>MATERIAL E MÉTODOS</b> .....	13
2.1	LOCALIZAÇÃO E CARACTERIZAÇÃO DA ÁREA EXPERIMENTAL.....	13
2.2	DELINEAMENTO EXPERIMENTAL .....	15
2.3	INSTALAÇÃO DO EXPERIMENTO E APLICAÇÃO DOS TRATAMENTOS...	15
2.4	MANEJO FITOSSANITÁRIO .....	16
2.5	AVALIAÇÕES NUTRICIONAIS E FISIOLÓGICAS .....	16
2.5.1	Trocas gasosas .....	16
2.5.2	Obtenção dos dados espectrais.....	17
2.5.3	Pigmentos (clorofilas <i>a</i> e <i>b</i> e carotenoides).....	18
2.5.4	Teor de Nitrogênio (N).....	18
2.6	ANÁLISES ESTATÍSTICAS .....	19
<b>3</b>	<b>RESULTADOS</b> .....	20
<b>4</b>	<b>DISCUSSÃO</b> .....	27
<b>5</b>	<b>CONCLUSÃO</b> .....	29
<b>6</b>	<b>AGRADECIMENTOS</b> .....	30

## 1 INTRODUÇÃO

O milho (*Zea mays* L.) é destaque nacional principalmente na segunda safra, que representa o maior volume colhido na região Centro-Oeste do país (SANTANA et al., 2021). Um dos macronutrientes mais exigidos pela cultura do milho é o nitrogênio (N), pois está diretamente ligado ao crescimento, produção (desenvolvimento de frutos e sementes) e qualidade da cultura (maior teor de proteínas, resistência a doenças, cor e aparência das folhas) (BERGER et al., 2020). Dessa forma, a aplicação de N garante uma maior produtividade das culturas agrícolas por ser um dos principais componentes da clorofila (Chl) (PRADO, 2021).

A deficiência de N impacta diretamente o teor de Chl, resultando em folhas amareladas (clorose), fechamento dos estômatos foliares (TEWARI et al., 2007), degradação de enzimas fotossintéticas e tilacóides (MU; CHEN, 2021) e, obstrução dos processos de captura de energia luminosa, transferência de elétrons e taxa de carboxilação nas folhas (XIONG et al., 2015). Isso compromete a capacidade fotossintética da planta, e conseqüentemente, seu crescimento e rendimento. Em condições de excesso de N, eleva o risco de doenças e pragas na cultura, além de gerar impactos ambientais negativos (SWARBRECK et al., 2019).

Assim, antes da utilização de fertilizantes nitrogenados nas lavouras, é necessário o conhecimento do estado nutricional que a planta se encontra (PRADO, 2021). Diante disso, a análise química foliar é o método convencional mais utilizado para determinação dos teores de macronutrientes e micronutrientes e, o estado nutricional da planta, identificando a deficiência ou excesso (COMINO et al., 2018). No entanto, apresenta etapas longas e destrutivas, resultando em uma análise de custo elevado, com trabalho intensivo e demorado.

Deste modo, o desenvolvimento de técnicas que permitem monitorar de forma correta, no tempo adequado das culturas e prever a diagnose precoce, é fundamental para que a gestão das culturas obtenha rendimentos elevados (HAN et al., 2023). O autor ressalta ainda que, uma das técnicas que vem sendo explorada dentro do sensoriamento remoto é o uso de sensores hiperespectrais, com o objetivo de monitorar o estado nutricional da planta. Como observado no trabalho de Silva et al. (2024) que, através de comprimentos de luz visíveis, folhas deficientes em N, apresentou menor teor de Chl que resultou em clorose.

Porém, os sensores hiperespectrais podem ter algumas limitações por terem alto custo, gerar grandes volumes de dados, exigindo análise computacional complexa e dispositivos robustos de armazenamento e processamento; tempo de aquisição é lento, calibração e manutenção frequentes, além de exigências de conhecimento especializado (POUR et al., 2024).

A técnica de aprendizagem de máquina (ML) surgiu da inteligência artificial (IA) como uma tecnologia com alto desempenho para desenvolver soluções multidisciplinar tecnológicas na agricultura (LIAKOS et al., 2018), e sua utilização permite obter diferentes respostas sobre as plantas utilizando variáveis espectrais (SANTANA et al., 2022), como por exemplo, teor de N foliar (NF) (OSCO et al., 2020). A ML é uma das formas mais eficazes de processar e analisar as grandes quantidades de dados obtidos pelas técnicas atuais de sensoriamento remoto (ZHENG et al., 2021) e, capaz de prever variáveis através de algoritmos (SONG et al., 2021).

Esta técnica de ML concentra-se principalmente no desenvolvimento de algoritmos preditivos (SONG et al., 2021) que é basicamente oferecer um conjunto de dados (como dados hiperespectrais) a um computador ou máquina e, a partir desses dados, começam a desenvolver e melhorar algoritmos para resolver um certo problema. Alguns algoritmos como, floresta aleatória (RF) e máquina de vetor suporte (SVM) são utilizados para previsão usando variáveis espectrais como entrada em conjuntos de dados (SU et al., 2022).

Esses algoritmos foram desenvolvidos para aprender padrões baseados em dados onde, através de treinamento o computador aprende as propriedades/características do conjunto de dados específico para um certo problema, e logo interpreta essas informações para gerar explicações mais compreensíveis sobre os conjuntos de dados disponíveis (SONG et al., 2021). Chlingaryan et al. (2018) verificaram o teor de N através de previsões de algoritmos de ML.

Nesse contexto, surgem as hipóteses que: doses de N afetam a fisiologia, desenvolvimento e o crescimento do milho, além de, interferir nos teores de Chl, carotenoides e fotossíntese líquida (A) em plantas de milho, e tais efeitos podem ser preditos a partir de comprimento de onda variando entre 350 e 2500 nm. Objetivou-se avaliar a acurácia dos algoritmos de ML na previsão do teor de N, pigmentos e taxa fotossintética em plantas de milho, utilizando como base os dados hiperespectrais.

## **2 MATERIAL E MÉTODOS**

### **2.1 LOCALIZAÇÃO E CARACTERIZAÇÃO DA ÁREA EXPERIMENTAL**

O experimento foi conduzido na área experimental da Universidade Federal de Mato Grosso do Sul (UFMS), Câmpus de Chapadão do Sul, com coordenadas geográficas de 18°46'16" S de latitude, 52°37'22" O de longitude e com altitude média de 820 metros, na safra 2023 com a cultura do milho (híbrido X35P618VYHR–Pioneer), em Latossolo Vermelho Distrófico (SANTOS et al., 2018). De acordo com a classificação de Kôppen, o clima predominante na região é do tipo tropical úmido (AW), com estação chuvosa no verão e seca

no inverno, e precipitação média anual de 1850 mm. A temperatura anual oscila entre 13 e 28 °C (medidas diárias).

A Figura 1 demonstra as condições meteorológicas durante a condução do experimento, referentes às médias de temperaturas máximas, mínimas e precipitação pluviométrica durante todo período experimental. Durante o experimento, a temperatura média foi de 21,2 °C. A precipitação mínima foi de 1,0 mm e a máxima de 108 mm com total de 845 mm distribuídos em 37 dias de chuvas (Figura 1), quantidade superior à necessária para o adequado desenvolvimento das plantas de milho, pois a exigência hídrica dessa cultura varia entre 380 e 550 mm (ALBUQUERQUE, 2010).

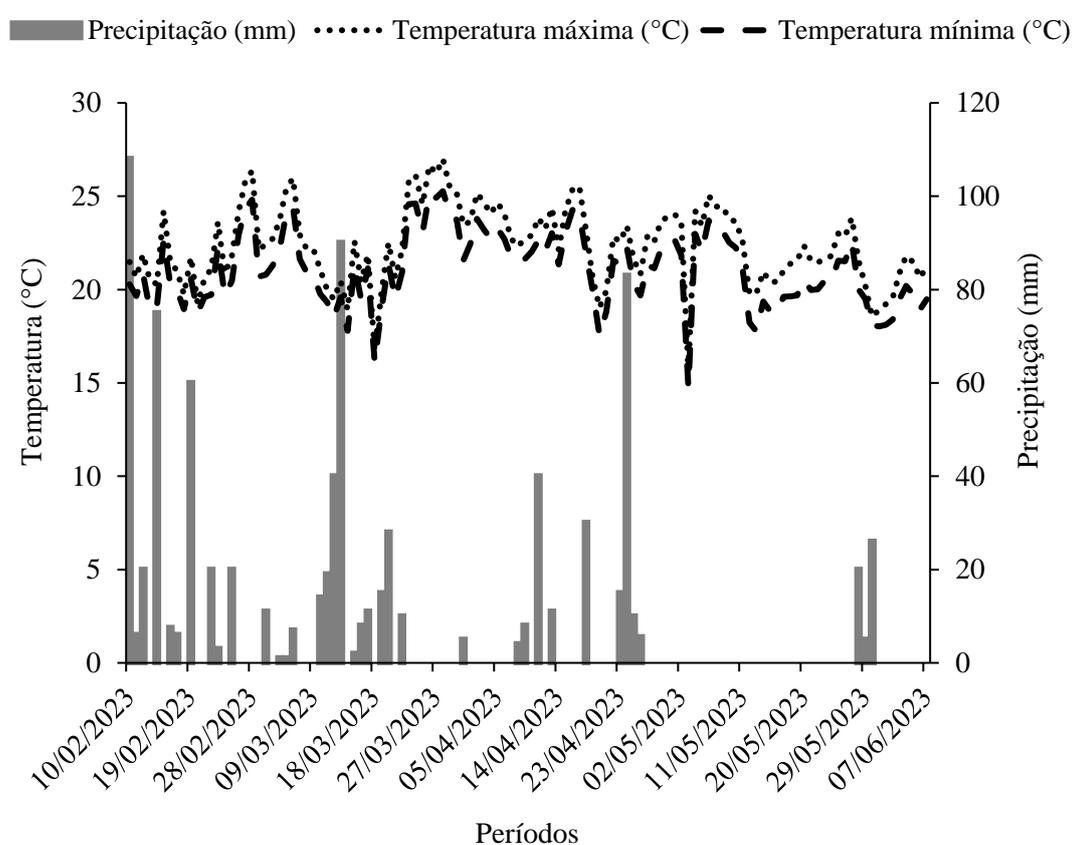


Figura 1. Condições meteorológicas no decorrer da safra 2ª de 2023, entre o período de 10/02 a 07/06.

Antes da instalação do experimento, foi realizada amostragem do solo na camada de 0 a 0,2 m de profundidade, para análise química para fins de fertilidade, de acordo com a metodologia descrita por Raij et al. (2001) e Teixeira et al. (2017). As análises do solo na camada 0 a 0,2 m, apresentaram as seguintes características químicas: pH (CaCl<sub>2</sub>)= 4,6; Ca+Mg= 2,70 cmol<sub>c</sub> dm<sup>-3</sup>; Ca= 2,30 cmol<sub>c</sub> dm<sup>-3</sup>; Mg= 0,40 cmol<sub>c</sub> dm<sup>-3</sup>; Al= 0,16 cmol<sub>c</sub> dm<sup>-3</sup>; H+Al= 5,7 cmol<sub>c</sub> dm<sup>-3</sup>; K= 0,24 cmol<sub>c</sub> dm<sup>-3</sup>; K= 94 mg dm<sup>-3</sup>; P (mel)= 13,3 mg dm<sup>-3</sup>; B= 0,21

mg dm<sup>-3</sup>; Cu= 1,0 mg dm<sup>-3</sup>; Fe= 73 mg dm<sup>-3</sup>; Mn= 10,2 mg dm<sup>-3</sup>; Zn= 4,3 mg dm<sup>-3</sup>; Argila= 485 g dm<sup>-3</sup>; Silte= 50 g dm<sup>-3</sup>; Areia= 465 g dm<sup>-3</sup>; MO= 31,2 g dm<sup>-3</sup>; CTC= 11,7 cmol<sub>c</sub> dm<sup>-3</sup>; V%= 34,0%.

## 2.2 DELINEAMENTO EXPERIMENTAL

O delineamento experimental foi em blocos casualizados (DBC), com quatro repetições, totalizando 16 parcelas compostas por 6 linhas de 5 m de comprimento, espaçadas de 0,50 m entre si, sendo a área útil formada pelas 4 linhas centrais, descartando 1,0 m de cada lado nas bordaduras.

Os tratamentos foram constituídos por quatro doses de nitrogênio (N), correspondendo a N1 (0%), N2 (30% - 54 kg ha<sup>-1</sup> de N), N3 (60% - 108 kg ha<sup>-1</sup> de N) e N4 (120% - 216 kg ha<sup>-1</sup> de N) avaliados em um período no estágio fenológico V6 (sexta folha totalmente desenvolvida).

## 2.3 INSTALAÇÃO DO EXPERIMENTO

No ano anterior (2022), antes da semeadura da soja, foi aplicado calcário na área experimental (33% CaO, 12% MgO e PRNT de 82%) para elevar a saturação por bases (V%) para 60%, conforme indicado por Sousa e Lobato (2004), pois houve essa necessidade pelo fato da V% estar com 34%.

A semeadura foi realizada no dia 10 de fevereiro de 2023, em plantio direto sobre palhada de soja. O híbrido utilizado foi o X35P618VYHR-Pioneer, recomendado para a região Centro-Oeste, no verão e na safrinha. As sementes foram tratadas com os inseticidas Clorantraniliprole – 625 g L<sup>-1</sup> (doses de 48 ml/60.000 sementes) e, Clotianidina – 600 g L<sup>-1</sup> (doses de 70 ml/60.000 sementes).

No plantio todos os tratamentos receberam 400 kg ha<sup>-1</sup> do formulado de NPK 4–14–8, ou seja, 16 kg ha<sup>-1</sup> N, 56 kg ha<sup>-1</sup> P<sub>2</sub>O<sub>5</sub> e 32 kg ha<sup>-1</sup> K<sub>2</sub>O. Para adubação de cobertura, foram aplicados 90 kg ha<sup>-1</sup> de K<sub>2</sub>O em todos os tratamentos em doses divididas, sendo aplicado 45 kg ha<sup>-1</sup> de K<sub>2</sub>O em V4 (quarta folha totalmente desenvolvida) e, 45 kg ha<sup>-1</sup> de K<sub>2</sub>O em V6 (sexta folha totalmente desenvolvida), para tanto utilizou-se a fonte KCl (cloreto de potássio), as adubações foram estabelecidas de acordo com a análise química do solo e recomendação de Sousa e Lobato (2004).

A adubação de cobertura de N foi em função dos tratamentos, onde N1 (0%), N2 (30% - 54 kg ha<sup>-1</sup> de N), N3 (60% - 108 kg ha<sup>-1</sup> de N) e N4 (120% - 216 kg ha<sup>-1</sup> de N) da dose recomendada (180 kg de N ha<sup>-1</sup> – adequada para cultura do milho com produtividade estimada de 10 a 12 t ha<sup>-1</sup>), para isso utilizou-se como fonte de N a ureia (SOUSA; LOBATO, 2004). As

doses foram divididas em duas aplicações, 50% no estágio V4 e 50% no estágio V6. No estágio fenológico V4 – V6 as exigências nutricionais do milho aumentam, então nesse período deve haver uma nutrição adequada para não prejudicar a produtividade da cultura (SILVA et al., 2024).

## 2.4 MANEJO FITOSSANITÁRIO

Durante a condução do experimento, no desenvolvimento da cultura, ocorreu o monitoramento diário, a fim de determinar o momento ideal para efetuar o controle fitossanitário de pragas, doenças e plantas invasoras de acordo com recomendações técnicas. As aplicações foram de forma mecanizada, utilizando pulverizadores tratorizados.

As aplicações foram feitas no plantio com o herbicida Atrazina ( $500 \text{ g L}^{-1}$ ) e, no estágio fenológico: V4 com os inseticidas Tiametoxam ( $141 \text{ g L}^{-1}$ ), Lambda-Cialotrina ( $106 \text{ g L}^{-1}$ ); V6 com o inseticida Acetamiprido ( $75 \text{ g L}^{-1}$ ), Fenpropatrina ( $112,5 \text{ g L}^{-1}$ ); Bifentrina ( $50 \text{ g L}^{-1}$ ), Carbosulfano ( $150 \text{ g L}^{-1}$ ).

## 2.5 AVALIAÇÕES NUTRICIONAIS E FISIOLÓGICAS

As avaliações foram realizadas nos estádios fenológicos V6 quando a sexta folha estava totalmente aberta, com os bordos de cada folíolo não mais se tocando. Primeiro foi mensurada (em campo) a taxa fotossintética em seis folhas (jovens completamente desenvolvidas) por parcela, descartando a folha cartucho. Logo estas mesmas folhas foram coletadas e armazenadas em uma caixa de isopor, evitando assim a perda de turgidez; posteriormente foram realizadas (em laboratório) as leituras espectrais e, análises de pigmentos e teor de N.

### 2.5.1 Trocas gasosas

Na primeira folha completamente desenvolvida com bainha visível foram realizadas as análises de trocas gasosas através do equipamento portátil de fotossíntese (Infrared Gas Analyzer – IRGA, modelo Li6400XT (LiCor Inc., Lincoln, Nebraska, USA) com fluxo de fótons fotossinteticamente ativos de  $1200 \mu\text{mol m}^{-2} \text{ s}^{-1}$  e concentração de  $\text{CO}_2$  ambiente de  $(372 \pm 10 \text{ mols m}^{-2} \text{ s}^{-1})$ . A variável fisiológica avaliada foi: fotossíntese ( $A$ ,  $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$ ). As medições foram realizadas no período da manhã, entre 8:00 h e 11:00 h, em 6 plantas de cada parcela, em dias sem nuvens, com temperatura de  $26 \text{ }^\circ\text{C}$  e umidade relativa do ar entre 60 e 80%.

## 2.5.2 Obtenção dos dados espectrais

Para análise espectral das folhas foi utilizado o espectrorradiômetro (FieldSpec 3 Jr da ASD - Analytical Spectral Devices (Dispositivos espectrais analíticos), pois esse equipamento obtém dados na faixa de 350 a 2500 nm, registra o sinal na região do VIS-NIR (350–1000 nm) e duas regiões do SWIR (1001–1800 nm e 1801-2500 nm), com intervalo de amostragem espectral de 1,4 e 2 nm, e resolução espectral (Full Width at Half Maximum (largura total na metade do máximo) - FWHM) de 3 nm e 10 nm, respectivamente (ASD, 2010).

Para a leitura das folhas foram utilizados o ASD Plant Probe (sonda de planta), pois ele é projetado para realizar medições espectrais de contato em materiais sólidos. Foi utilizado uma placa branca (feita com sulfato de bário) que reflete 100% da luz, para calibração do equipamento. O sensor foi conectado a um computador para os registros de cada leitura por meio de um software próprio do sensor chamado RS<sup>3</sup>, que registra as leituras feitas pelo equipamento. Posteriormente os arquivos foram importados para outro software, ViewSpectroPro para extração dos dados em extensão .txt, facilitando análises estatísticas posteriores.

Obtido os comprimentos de onda, foram agrupados em médias de intervalos representativos em bandas (Tabela 1). Conforme sugerido por Silva Junior et al. (2018), tais faixas estão relacionadas com processos fisiológicos importantes para a planta, características fisiológicas e suas respostas espectrais também foi descrito por Silva Junior et al. (2018).

Tabela 1. Intervalos de faixa espectral (nm) e comprimento de onda médio (nm) usados no estabelecimento das 28 bandas espectrais.

<b>Banda (nº)</b>	<b>Faixa espectral (nm)</b>	<b>Comprimento de onda médio (nm)</b>
1	350–369	359.5
2	370	370.0
3	371–419	395.0
4	420	420.0
5	421–424	422.5
6	425	425.0
7	426–444	435.0
8	445–475	460.0
9	480	480.0
10	481–500	490.5

Banda (n°)	Faixa espectral (nm)	Comprimento de onda médio (nm)
11	501–530	515.5
12	531–539	535.0
13	540	540.0
14	541–649	595.0
15	650	650.0
16	661–670	665.5
17	675	675.0
18	676–684	680.0
19	685–689	687.0
20	690–700	695.0
21	701–709	705.0
22	710	710.0
23	711–730	720.5
24	960	960.0
25	1100	1100.0
26	1400	1400.0
27	1930	1930.0
28	2200	2200.0

### 2.5.3 Pigmentos (clorofilas *a* e *b* e carotenoides)

Para analisar os teores de pigmentos, coletou-se discos com massa entre 0,025 e 0,030 g de amostras frescas das seis folhas. Em seguida, foram colocadas em tubos Falcon, contendo 1,5 ml de acetona a 80%. As amostras ficaram a 4°C por 48 horas, protegidas da luz. Os extratos foram medidos através de espectrofotômetro (modelo SF 325NM da BEL) em 663 nm (clorofila *a*), 647 nm (clorofila *b*) e 470 nm (carotenoides), e a concentração de cada pigmento obtida conforme o método de Lichtenthaler (1987). Os resultados foram expressos em mg g<sup>-1</sup> de massa de matéria fresca.

### 2.5.4 Teor de Nitrogênio (N)

As folhas coletadas foram lavadas em água corrente, depois colocadas em solução detergente neutro a 0,1%, solução de HCl (ácido clorídrico) a 0,3% e água deionizada (Prado, 2021). O material vegetal foi seco em estufa de circulação forçada de ar (65 ± 5 °C) até peso

constante, sendo obtida a matéria seca das folhas de milho. O material foi moído e pesado (0,1 g) após a secagem, seguindo a técnica de titulação Kjeldahl, o material passou por digestão, destilação e, titulação com ácido sulfúrico ( $H_2SO_4$ ), assim determinou o teor de N em extratos obtidos pela digestão sulfúrica.

## 2.6 ANÁLISES ESTATÍSTICAS

Os dados foram submetidos as análises de aprendizagem de máquina (ML): REPTree (DT), Árvore de Decisão M5P (M5P), Floresta Aleatória (RF), Redes Neurais Artificiais (RNA), Máquina de Vetor Suporte (SVM) e ZeroR (ZR – controle (refere-se a um padrão de comparação para verificar os resultados de algo testado)). Utilizamos estes algoritmos porque foram aplicados com sucesso em estudos de predição usando variáveis espectrais como entrada em conjuntos de dados. O ZR (padrão do Weka) foi usado como uma técnica de controle (referência) para avaliar a eficácia de outros algoritmos, pois ele é capaz de determinar se novos modelos realmente adicionam valor em termos de precisão, aplicando o valor médio da variável alvo e disponibilizando esse valor como uma predição (FAN et al., 2019).

Foram testadas duas configurações de entrada (input): usando apenas os comprimentos de onda na predição da variável (ALL) e utilizando as bandas espectrais (SB). Para a seleção dessas SB foi observado o máximo desvio e absorção de radiação entre as regiões visível e infravermelho (SILVA JUNIOR et al., 2018). Como variáveis de saída (output) foram utilizados: teor de nitrogênio foliar (NF), pigmentos e a variável fisiológica fotossíntese líquida (A). A predição foi realizada por meio de validação cruzada estratificada com k-fold = 10 e dez repetições (100 execuções para cada modelo). Todos os parâmetros dos modelos de ML foram estabelecidos de acordo com a configuração padrão do software Weka 3.8.5, exceto para RNA em que foram utilizados 10 neurônios na primeira camada e, 10 neurônios na segunda camada.

Para avaliar a acurácia dos modelos de predição testados, foram utilizadas as métricas do coeficiente de correlação de Pearson ( $r$ ) e do erro médio absoluto (MAE). Para verificar a significância das entradas testadas, dos algoritmos de ML e da interação entre ambos, foi realizada uma análise de variância (ANOVA). Quando significativos, foram gerados gráficos de caixa com as médias de  $r$  e MAE agrupadas pelo teste de Scott-Knott ao nível de 5% de probabilidade.

### 3 RESULTADOS

A tabela 2 apresenta os valores de  $p$  obtidos por ANOVA para  $r$  e MAE, considerando os diferentes modelos de aprendizagem de máquina (ML) e as duas configurações de entrada: usando apenas os comprimentos de onda na predição da variável (ALL) e utilizando as bandas espectrais (SB). A interação entre os algoritmos de ML e os dados de entrada foi significativa (teste  $F - P < 0,05$ ) para o erro médio absoluto (MAE) para a variável fisiológica fotossíntese líquida ( $A$ ), sendo necessário o desdobramento da interação (tabela 2). Para os modelos de ML testados, tanto o coeficiente de correlação de Pearson ( $r$ ) quanto o MAE foram significativos (teste  $F - P < 0,05$ ) para todas as variáveis. Os dados de entrada, tiveram o  $r$  significativos para algumas variáveis como, clorofila  $a$  (Chl  $a$ ), clorofila  $b$  (Chl  $b$ ), clorofila total (Chl  $a + b$ ) e, carotenoides (Cx + c), (tabela 2).

Tabela 2. Análise de variância para coeficientes de correlação de Pearson ( $r$ ) e erro médio absoluto (MAE) por diferentes entradas e modelos de aprendizagem de máquina (ML) testados.

F.V.	G.L.	NF		Chl $a$		Chl $b$	
		$r$	MAE	$r$	MAE	$r$	MAE
ML	5	1,35*	3,90*	0,77*	0,01*	1,13*	0,00*
Entrada	1	0,00	0,09	0,07*	0,00	0,03*	0,00
ML*Entrada	5	0,00	0,02	0,00	0,00	0,00	0,00
Erro	108	0,00	0,06	0,04	0,01	0,00	0,00

F.V.	G.L.	Chl $a + b$		Cx + c		A	
		$r$	MAE	$r$	MAE	$r$	MAE
ML	5	0,97*	0,01*	1,02*	0,00*	1,09*	14,65*
Entrada	1	0,05*	0,00	0,02*	0,00	0,01	0,47
ML*Entrada	5	0,00	0,00	0,00	0,00	0,00	0,41*
Erro	108	0,00	0,10	0,00	0,00	0,00	0,16

\*: significativo a 5% pelo teste F; F.V.: fonte de variação; G.L.: graus de liberdade; NF: nitrogênio foliar; Chl  $a$ : clorofila  $a$ ; Chl  $b$ : clorofila  $b$ ; Chl  $a + b$ : clorofila total; Cx + c: carotenoides; A: fotossíntese líquida;  $r$ : coeficiente de correlação de Pearson; MAE: erro médio absoluto; ML: aprendizagem de máquina; ML\*Entrada: interação.

Buscando identificar quais modelos de ML tem mais eficiência na predição das variáveis, foi construído um gráfico de caixa com médias de  $r$  e MAE para os diferentes algoritmos de ML. Estes gráficos permitem visualizar a distribuição dos dados de uma maneira simples e objetiva; além de mostrar a mediana, quartis e possíveis outliers dos dados. Isso é fundamental para entender a dispersão e a centralidade dos dados antes de adotar um modelo de ML (WICKHAM, 2016). O autor ressalta ainda que os outliers (pontos fora de uma curva normal –

valores que foge da normalidade) podem ter um efeito significativo nos modelos de ML, dessa forma, os gráficos de caixa ajudam na identificação desses pontos fora do comum, e permite uma tomada de decisão sobre como lidar com eles (por exemplo, remoção ou tratamento).

Nota-se que para predição da variável NF, os melhores modelos foram M5P, RF e SVM, pois ambos tiveram maiores correlações ( $r$ ) acima de 0,6 e menores erros abaixo de 3,0 (Figura 2).

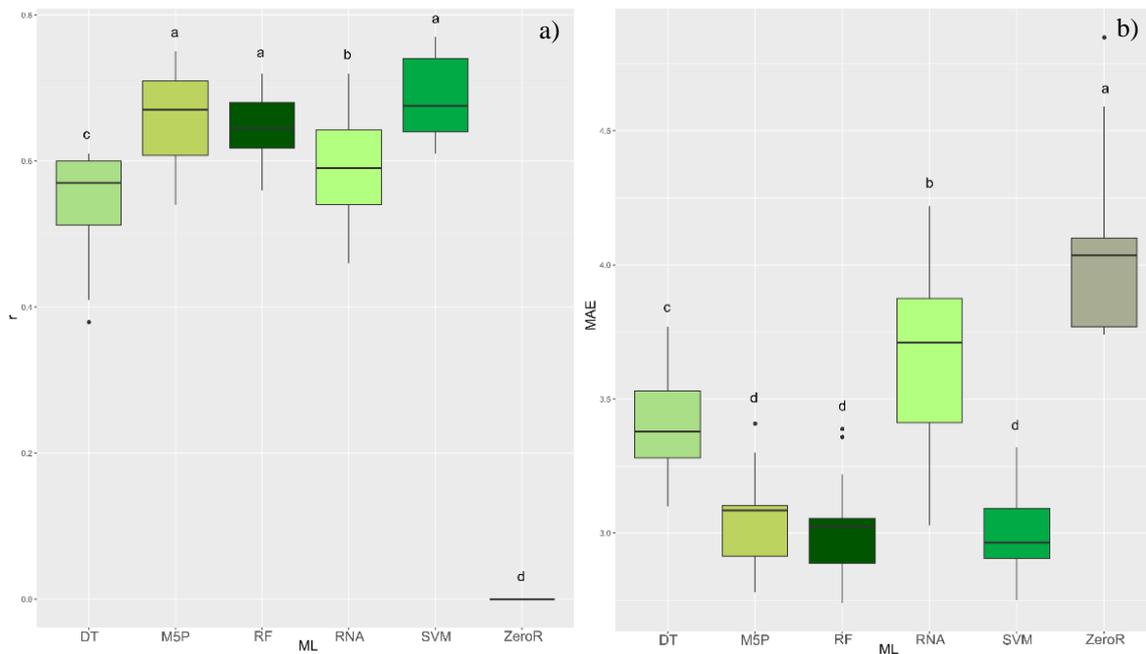


Figura 2. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson ( $r$ ) (a) e erro médio absoluto (MAE) (b) para os modelos de aprendizagem de máquina (ML) na predição de nitrogênio foliar (NF).

As médias seguidas pelas mesmas letras não diferem entre si pelo teste de Scott-Knott a 5% de probabilidade.

REPTree (DT), árvore de decisão M5P (M5P), floresta aleatória (RF), redes neurais artificiais (RNA), máquina de vetor suporte (SVM) e ZeroR.

O modelo SVM apresentou maior precisão, consistindo em uma melhor técnica de ML para prever a variável Chl  $a$ , principalmente quando utilizou como entrada para o algoritmo as bandas espectrais (SB) (Figura 3). Isso porque forneceu os maiores valores de  $r$  e os menores valores de MAE. Além disso, apresentou gráficos de caixa menores, evidenciando menor dispersão dos dados utilizados para obtenção de  $r$  e MAE, indicando resultados mais consistentes, alta confiabilidade e qualidade do modelo, também fácil interpretação e comunicação e, acurácias na predição (JAMES et al., 2013).

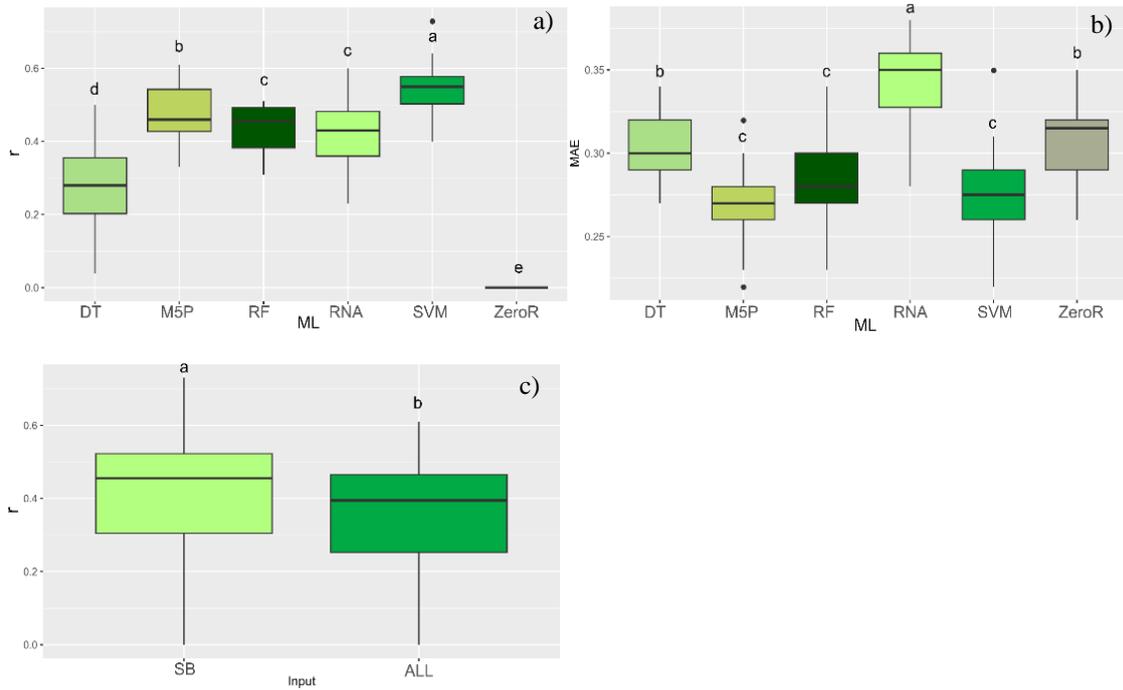


Figura 3. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson ( $r$ ) (a) e erro médio absoluto (MAE) (b) para os modelos de aprendizagem de máquina (ML), e médias de  $r$  para os inputs (entradas) testados (c) na predição de clorofila  $a$  (Chl  $a$ ).

As médias seguidas pelas mesmas letras não diferem entre si pelo teste de Scott-Knott a 5% de probabilidade.

REPTree (DT), árvore de decisão M5P (M5P), floresta aleatória (RF), redes neurais artificiais (RNA), máquina de vetor suporte (SVM) e ZeroR.

Bandas espectrais (SB) e comprimentos de onda na predição da variável (ALL).

Para a predição da variável Chl  $b$ , os melhores algoritmos de ML foram M5P e SVM, pois forneceu os maiores valores de  $r$  em torno de 0,6, indicando uma relação linear positiva moderada, porque os valores de  $r$  variam de -1 a 1 (RATNER, 2009); e menores valores de MAE, abaixo de 0,1 (Figura 4). Como entrada dos algoritmos, SB apresentou melhor acurácia.

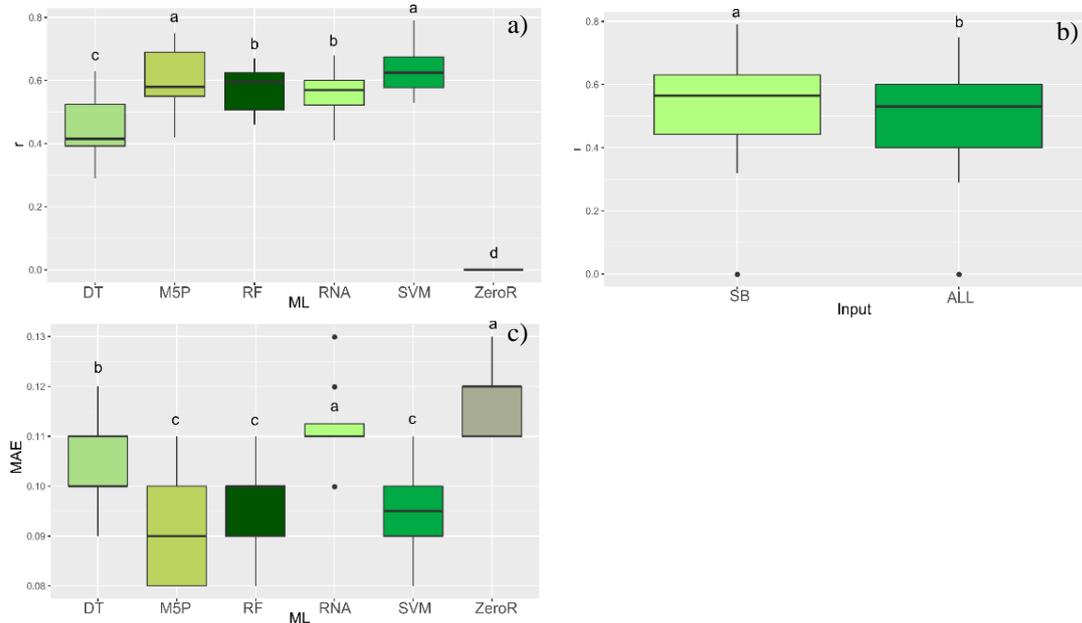


Figura 4. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson ( $r$ ) (a) e erro médio absoluto (MAE) (c) para os modelos de aprendizagem de máquina (ML), e médias de  $r$  para os inputs (entradas) testados (b) na predição de clorofila  $b$  (Chl  $b$ ).

As médias seguidas pelas mesmas letras não diferem entre si pelo teste de Scott-Knott a 5% de probabilidade.

REPTree (DT), árvore de decisão M5P (M5P), floresta aleatória (RF), redes neurais artificiais (RNA), máquina de vetor suporte (SVM) e ZeroR.

Bandas espectrais (SB) e comprimentos de onda na predição da variável (ALL).

Os melhores modelos de ML para prever a variável Chl  $a + b$ , foram M5P e SVM, pois ambos tiveram maiores valores de  $r$  (em torno de 0,6) e menores MAE (abaixo de 0,25) (Figura 5). A melhor configuração de entrada para estes modelos foram as SB.

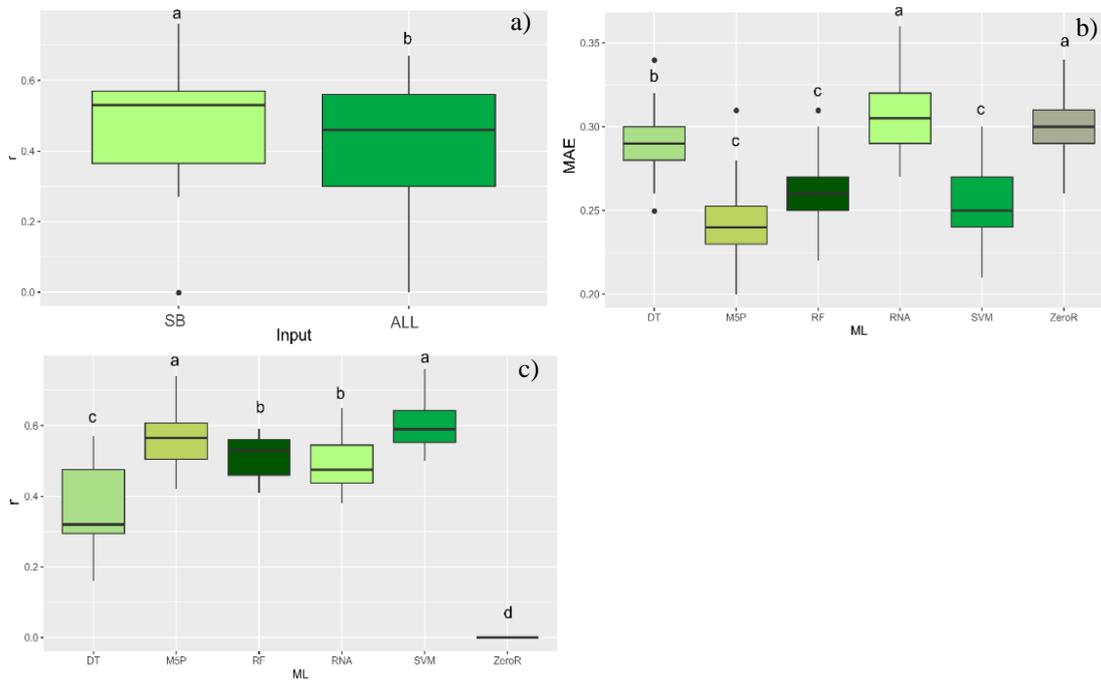


Figura 5. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson ( $r$ ) (c) e erro médio absoluto (MAE) (b) para os modelos de aprendizagem de máquina (ML), e médias de  $r$  para os inputs (entradas) testados (a) na predição de clorofila total (Chl  $a + b$ ).

As médias seguidas pelas mesmas letras não diferem entre si pelo teste de Scott-Knott a 5% de probabilidade.

REPTree (DT), árvore de decisão M5P (M5P), floresta aleatória (RF), redes neurais artificiais (RNA), máquina de vetor suporte (SVM) e ZeroR.

Bandas espectrais (SB) e comprimentos de onda na predição da variável (ALL).

Os algoritmos de ML M5P e RF, tiveram maiores valores de  $r$  (próximo de 0,6) e menores erros (abaixo de 0,06) para a predição da variável carotenoides; utilizando SB proporcionou melhor acurácia aos algoritmos (Figura 6).

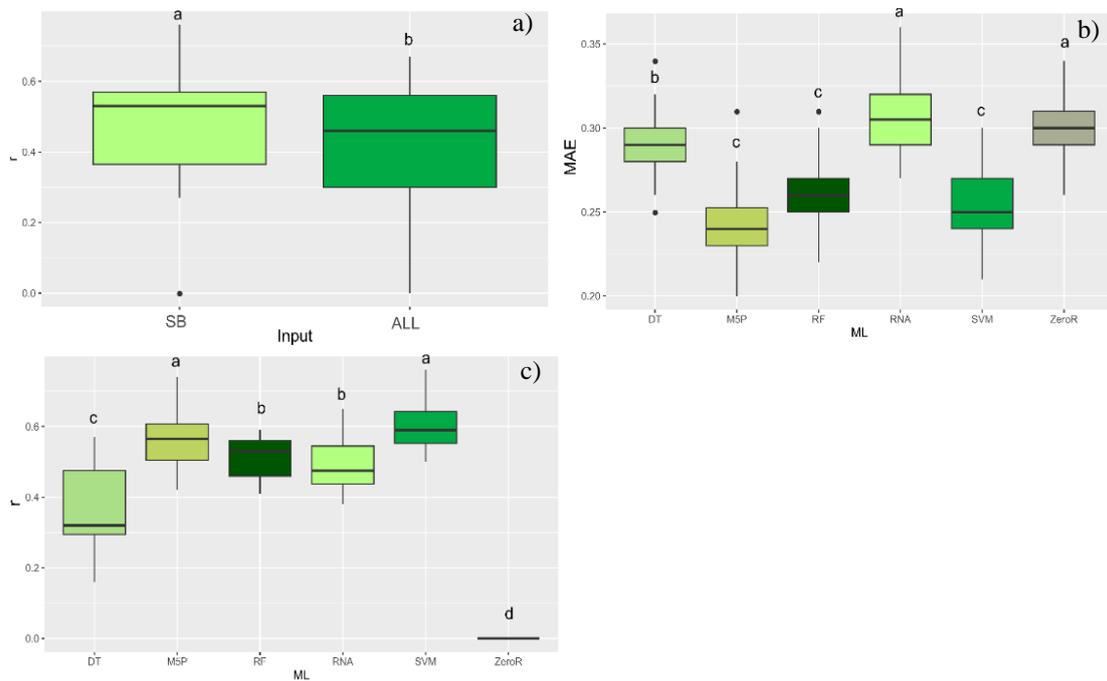


Figura 6. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson ( $r$ ) (c) e erro médio absoluto (MAE) (b) para os modelos de aprendizagem de máquina (ML), e médias de  $r$  para os inputs (entradas) testados (a) na predição de carotenoides (Cx + c).

As médias seguidas pelas mesmas letras não diferem entre si pelo teste de Scott-Knott a 5% de probabilidade.

REPTree (DT), árvore de decisão M5P (M5P), floresta aleatória (RF), redes neurais artificiais (RNA), máquina de vetor suporte (SVM) e ZeroR.

Bandas espectrais (SB) e comprimentos de onda na predição da variável (ALL).

Para a predição da variável A, os melhores algoritmos de ML foram M5P, RF e SVM (Figura 7), com maiores valores de  $r$ , em torno de 0,6. Para os modelos M5P, RF as entradas SB e ALL não foram significativas para ambos, porém para o SVM, utilizando as SB como entrada, este modelo erra menos.

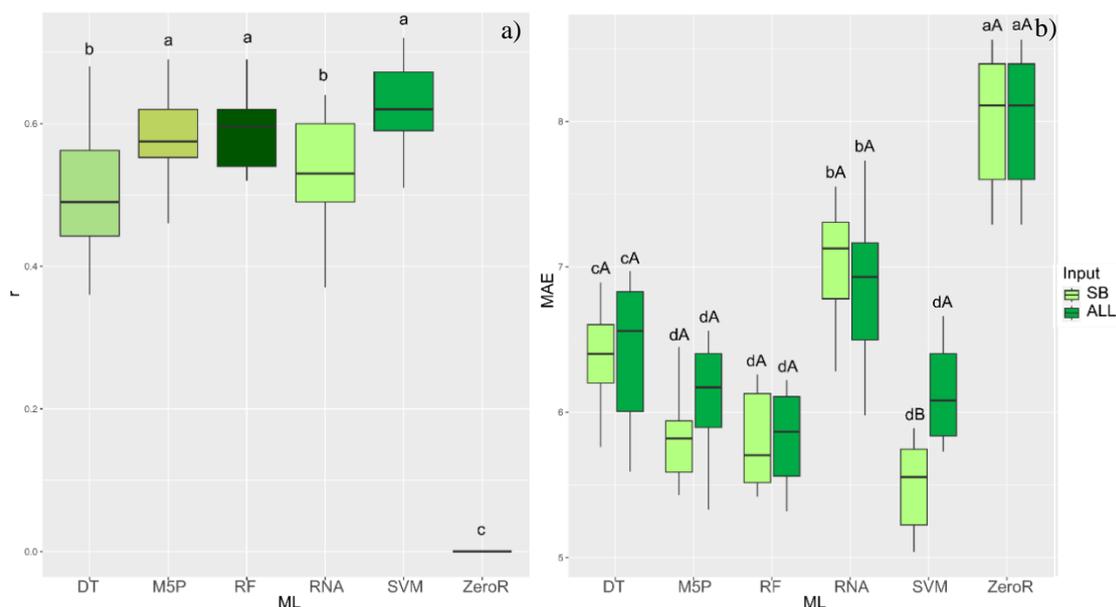


Figura 7. Gráficos de caixa (3 quartis, valor de máximo e mínimo) para coeficiente de correlação de Pearson ( $r$ ) (a) para os modelos de aprendizagem de máquina (ML), e erro médio absoluto (MAE) (b) para o desdobramento da interação significativa entre ML e os inputs (entradas) testados na predição de fotossíntese líquida ( $A$ ).

Médias seguidas pelas mesmas letras maiúsculas para as diferentes entradas e mesmas letras minúsculas para os diferentes algoritmos de ML não diferem pelo teste Scott-Knott a 5% de probabilidade.

REPTree (DT), árvore de decisão M5P (M5P), floresta aleatória (RF), redes neurais artificiais (RNA), máquina de vetor suporte (SVM) e ZeroR.

Bandas espectrais (SB) e comprimentos de onda na predição da variável ( $ALL$ ).

Ao avaliar os modelos de ML testados para predição, através das métricas do  $r$  e do MAE, o algoritmo M5P quando utilizou como entrada as SB, teve maior acurácia para prever as variáveis NF, Chl  $b$ , Chl total, carotenoides e  $A$ . Para o modelo RF, ao usar as SB foi mais eficaz na predição das variáveis NF, carotenoides e  $A$ . O algoritmo SVM teve melhor desempenho usando as SB como configuração de entrada para prever as variáveis NF, Chl  $a$ , Chl  $b$ , Chl total e  $A$ .

Sendo assim, os modelos M5P e SVM usando SB como entrada, foram os que mais se destacaram e tiveram maiores precisões, consistindo em uma melhor técnica de ML para predição na cultura de milho. Além de apresentar diagramas de caixa menores, demonstrando menor dispersão dos dados utilizados para obtenção de  $r$  e MAE.

## 4 DISCUSSÃO

Os sistemas dispositivos espectrais analíticos são usados para determinar a condição fisiológica da planta, incluindo o estado da doença, situação nutricional relacionada à absorção de N e ao equilíbrio de umidade (WANG et al., 2022). Para verificar a reflectância na faixa de 350–2500 nm, Wen et al. (2018) comparam folhas de milho com deficiência e suficiência de N, que resultou na reflectância da luz do visível de 436–698 nm mais alta em plantas com deficiência de N, por outro lado, para as plantas com suficiência de N, a reflectância NIR do milho de 740–1350 nm foi maior.

Os sensores hiperespectrais combinados com algoritmos de aprendizagem de máquina (ML) permitem análises precisas e detalhadas da fisiologia da planta. As bandas espectrais (SB) utilizadas como variáveis de entrada para os algoritmos, são uma fonte confiável para prever nitrogênio foliar (NF), pigmentos e fotossíntese líquida (*A*) em plantas de milho, pelo fato que os sensores hiperespectrais capturam dados em centenas de bandas estreitas ao longo do espectro eletromagnético, permitindo uma análise mais detalhada e específica (MULLA, 2013).

Ao comparar o desempenho de diferentes modelos de predição, os gráficos de caixa são eficazes para mostrar a variabilidade dos algoritmos, sendo muito útil para avaliar a consistência e a robustez dos modelos (JAMES et al., 2013).

O N está relacionado à capacidade fotossintética das culturas, sendo o macronutriente necessário para a síntese de clorofila (Chl), assim, plantas com níveis adequados de N produzem mais Chl *a* e *b*, o que eleva a capacidade de absorção de luz e, conseqüentemente, a taxa de *A* (ZHANG et al., 2020). Na ausência de Chl, os carotenoides não conseguem desempenhar sua função, pois a energia que os carotenoides produzem ao absorver luz precisa ser transferida para as moléculas de Chl, que é a responsável por transformar a energia luminosa em energia química, logo usam essa energia para excitar os elétrons e iniciar a cadeia de transporte de elétrons na fotossíntese (ZULFIQAR et al., 2021).

Os carotenoides são pigmentos acessórios importantes no processo de fotossíntese, pois absorvem luz em diferentes partes do espectro visível (principalmente na faixa azul e verde) que a clorofila não absorve de forma eficiente (TAIZ; ZEIGER, 2010). Isso amplia o espectro de luz utilizável para a fotossíntese. O autor ressalta ainda que, em condições de alta intensidade de luz, os carotenoides ajudam a dissipar o excesso de energia como calor, protegendo as células fotossintéticas do estresse oxidativo e danos que poderiam ser causados por espécies reativas de oxigênio.

Diante disso, a limitação da produtividade das culturas e os impactos ambientais negativos, estão na aplicação excessiva ou insuficiente de fertilizantes nitrogenados (FU et al.,

2021). Com isso, a predição eficaz do rendimento e o manejo melhorado do N, são importantes na agricultura, pois através de dados espectrais sobre o estado nutricional da planta, permite rápidos diagnósticos nas práticas de manejo. Para melhorar a produção e a gestão das culturas, e ao mesmo tempo reduzir os custos operacionais e o impacto ambiental, os sistemas de detecção remota (RS) estão sendo bastante utilizados para a construção de ferramentas de suporte à decisão para os sistemas agrícolas contemporâneos (CHLINGARYAN et al., 2018).

Por ser um elemento estrutural, a deficiência ou excesso de N ocasionam alterações fisiológicas na planta (GOMES et al., 2023). Dessa forma, o algoritmo máquina de vetor suporte (SVM) atrelado aos dados hiperespectrais permitem a detecção precoce de deficiências nutricionais antes que essas manifestem sinais visíveis, permitindo uma análise mais profunda da fisiologia das plantas. Uma vez que, antes dos sintomas visuais na planta, perdas fisiológicas já ocorreram nas folhas (OSCO et al., 2020). Por conseguinte, técnicas capazes de avaliar precocemente eventuais deficiências nutricionais de N, são indispensáveis para minimizar danos nas culturas (SILVA et al., 2024).

Esses dados hiperespectrais fornecem uma vasta quantidade de informações espectrais que, quando analisadas com o modelo de ML árvore de decisão M5P, resultam em predições precisas de Chl e carotenoides. Assim, o algoritmo M5P para análise de dados, torna-se promissor para prever resultados de fácil interpretação e precisão e, sua estrutura é construída por meio da seleção de variáveis preditivas (THAI et al., 2020). O algoritmo M5P é uma árvore de decisão que foi projetada para problemas de predição, onde realiza-se a ramificação de dados para construir uma árvore completa e, em seguida desenvolve-se um modelo de regressão em cada nó terminal para fazer predições; posteriormente, a árvore é podada para evitar o overfitting (o modelo decora ao invés de aprender), e ocorre o monitoramento da árvore (SILVA et al., 2024).

A ML é particularmente eficaz no processamento e análise de grandes volumes de dados, inclusive os hiperespectrais. Algoritmos como SVM, floresta aleatória (RF) são capazes de extrair padrões e relações complexas dos dados, que seriam difíceis de identificar com métodos tradicionais. Os parâmetros fisiológicos das plantas, como o teor de Chl, podem ser preditos com alta precisão através de algoritmos treinados, identificando correlações entre as SB capturadas pelos sensores e os parâmetros fisiológicos de interesse (IM et al., 2022; SUDU et al., 2022).

A partir dos dados hiperespectrais, este estudo demonstrou que em plantas de milho, os algoritmos RF, SVM e M5P foram eficazes para a predição de NF, pigmentos fotossintéticos e

A. Sendo o SVM e M5P de maior desempenho por predizerem um maior número de variáveis analisadas pelas SB.

A combinação de dados hiperespectrais e o modelo SVM é uma abordagem robusta para a predição de Chl *a* em plantas de milho. Esse algoritmo pode lidar eficientemente com a complexidade dos dados, resultando em predições precisas dos níveis de Chl (RUSZCZAK et al., 2023). RF também proporcionou precisão dos resultados para as variáveis NF, fotossíntese e carotenoides, visto que, apresentou menor dispersão dos dados utilizados para obtenção de *r* e MAE.

Nossos resultados demonstram que é possível predizer variáveis de milho com mais precisão, usando SB como entrada, nos modelos de ML M5P, SVM e RF. Isso é devido à capacidade dessas bandas de capturar informações detalhadas – em centenas de bandas estreitas, sobre a nutrição e fisiologia das plantas, que podem refletir pequenas mudanças nos componentes bioquímicos das folhas, como o teor de N e Chl, que são essenciais para o crescimento e desenvolvimento das culturas (SMITH et al., 2020). Estudos demonstram que ao combinar dados de múltiplas SB na análise de dados hiperespectrais, os modelos, por exemplo RF, se tornam mais precisos para predizer variáveis do milho, como NF (SMITH et al., 2021).

Além disso, a associação de características espectrais com dados da imagem desenvolve ainda mais a performance dos modelos preditivos (JOHNSON; WANG, 2021). Proporcionando, assim, o aumento da eficiência na utilização destes sensores, e técnicas de espectroscopia de refletância. As bandas hiperespectrais permite a análise rápida de inúmeras amostras, otimiza o tempo e reduz o custo para análises convencionais do tecido vegetal, e ainda minimiza análises de N que possam ter longos processos e, quantidades de fertilizantes químicos aplicados. Sendo assim, a quantidade de resíduos gerados pelo uso de reagentes em análises químicas é menor, o que a curto e longo prazo são prejudiciais para o pesquisador, técnico e para natureza.

## **5 CONCLUSÃO**

Este estudo provou que algoritmos de aprendizagem de máquina (ML) atrelados aos dados hiperespectrais, podem ser utilizados como método alternativo de diagnóstico de nitrogênio em plantas de milho. A utilização dos modelos de ML máquina de vetor suporte, árvore de decisão M5P e floresta aleatória, apresentam-se promissores na predição de teor de nitrogênio foliar, pigmentos e fotossíntese líquida em milho, especialmente quando usado as bandas espectrais como dados de entrada para estes algoritmos.

## **6 AGRADecIMENTOS**

À Fundação de apoio ao desenvolvimento do ensino, ciência e tecnologia do estado de Mato Grosso do Sul - FUNDECT-Termo de outorga: 318/2022. SIAFEM: 32242.

À Fundação Universidade Federal de Mato Grosso do Sul, ao Grupo de Estudos do Cerrado em Nutrição de Plantas – UFMS (GECENP), e ao Grupo de Estudos em Nutrição de Plantas da UNESP (Genplant) e, a Pioneer pelas sementes fornecidas.

## REFERÊNCIAS

ALBUQUERQUE, P. E. P. Manejo de irrigação na cultura do milho. **Embrapa Milho e Sorgo**. Sistema de Produção, 1 ISSN 1679-012X Versão Eletrônica - 6ª edição Set, 2010.

BERGER, K. et al. Crop nitrogen monitoring: Recent progress and principal developments in the context of imaging spectroscopy missions. **Remote Sensing of Environment**, v. 242, n. December 2019, p. 111758, 2020.

CHLINGARYAN, A. et al. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. **Computers and electronics in agriculture**, v. 151, p. 61-69, 2018.

COMINO, F. et al. Near-infrared spectroscopy and X-ray fluorescence data fusion for olive leaf analysis and crop nutritional status determination. **Talanta**, v. 188, n. June, p. 676-684, 2018.

FAN, L. et al. Hyperspectral-based estimation of leaf nitrogen content in corn using optimal selection of multiple spectral variables. **Sensors**, v. 19, n. 13, p. 2898, 2019.

FU, Y. et al. Uma visão geral da avaliação do status de nitrogênio da cultura usando sensoriamento remoto hiperespectral: Status atual e perspectivas. **European Journal of Agronomy**, v. 124, p. 126241, 2021.

GOMES, R. L. et al. Near-infrared leaf reflectance modeling of *Annona emarginata* seedlings for early detection of variations in nitrogen concentration. **Journal of Forestry Research**, v. 34, n. 1, p. 269-282, 2023.

HAN, P. et al. Dissection of Hyperspectral Reflectance to Estimate Photosynthetic Characteristics in Upland Cotton (*Gossypium hirsutum* L.) under Different Nitrogen Fertilizer Application Based on Machine Learning Algorithms. **Plants**. v.12, p. 455, 2023.

IM, G. M. et al. Estimating chlorophyll-a concentration from hyperspectral data using various machine learning techniques: A case study at Paldang Dam, South Korea. **Water**, v. 14, n. 24, p. 4080, 2022.

JAMES, G. et al. An introduction to statistical learning. **New York: springer**, 2013.

JOHNSON, R.; WANG, M. Fusion of spectral and texture features for improved predictive modeling in precision agriculture. **Sensors**, 21(6), p. 2147, 2021.

LIAKOS, K. G. et al. Machine learning in agriculture: A review. **Sensors**, v. 18, n. 8, p. 2674, 2018.

LICHTENTHALER, H. K. Chlorophylls and carotenoids: pigments of photosynthetic biomembranes. Em: **Methods in enzymology**. [s.l.] Elsevier. v. 148, p. 350-382, 1987.

MU, X.; CHEN Y. The physiological response of photosynthesis to nitrogen deficiency. **Plant Physiology and Biochemistry**, v. 158, p. 76-82, 2021.

MULLA, D. J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. **Biosystems Engineering**, 114(4), p. 358-371, 2013.

OSCO, L. P. et al. Leaf nitrogen concentration and plant height prediction for maize using UAV-based multispectral imagery and machine learning techniques. **Remote Sensing**, v. 12, n. 19, p. 3237, 2020.

PRADO, R. M. Mineral nutrition of tropical plants. Cham, Switzerland: **Springer**, 2021.

POUR, A. B. et al. Hyperspectral Remote Sensing from Spaceborne and Low Altitude Aerial/Drone-Based Platforms—**Differences in Approaches, Data Processing Methods, and Applications**. 2024.

RAIJ, B. V. et al. Análise química para avaliação da fertilidade de solos tropicais. Campinas: **IAC**, 2001.

RATNER, B. The correlation coefficient: Its values range between +1/-1, or do they? **Journal of Targeting**, Measurement and Analysis for Marketing, 17(2), p. 139-142, 2009.

RUSZCZAK, B. et al. Estimating chlorophyll content from hyperspectral data using gradient features. In: International Conference on Computational Science. Cham: **Springer Nature Switzerland**, p. 196-203, 2023.

SANTANA, D. C. et al. High-throughput phenotyping allows the selection of soybean genotypes for earliness and high grain yield. **Plant Methods**, v. 18, n. 1, p. 13, 2022.

SANTANA, D. C. et al. UAV-BASED multispectral sensor to measure variations in corn as a function of nitrogen topdressing. **Remote Sensing Applications: Society and Environment**, v. 23, n. May, p. 100534, 2021.

SANTOS, H. G. et al. Sistema Brasileiro de Classificação de Solos. [s.l.] Brasília, DF: **Embrapa**, 2018.

SCOTT, A. J; KNOTT, M. A Cluster Analysis Method for Grouping Means in the Analysis of Variance. **Biometria**, 30, 507-512, 1974.

SILVA, B. C. et al. New approach for predicting nitrogen and pigments in maize from hyperspectral data and machine learning models. **Remote Sensing Applications: Society and Environment**, v. 33, p. 101110, 2024.

SILVA JUNIOR, C. A. et al. Soybean varieties discrimination using non-imaging hyperspectral sensor. **Infrared Physics & Technology**, v. 89, p. 338-350, 2018.

SMITH, A. B.; JONES, C. D.; TAYLOR, E. F. Hyperspectral imaging for precision agriculture: Monitoring crop health and development. **Journal of Agricultural Science and Technology**, 15(4), p. 210-225, 2020.

SMITH, J.; BROWN, A.; LEE, K. Enhanced prediction of leaf nitrogen content in maize using optimized spectral variables in random forest models. **Remote Sensing**, 13(10), p. 1985, 2021.

SONG, H. et al. Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms. **Construction and Building Materials**, v. 308, p. 125021, 2021.

SOUSA, D. M. G.; LOBATO, E. Cerrado: correção do solo e adubação. [s.l.] Brasília, DF: **Embrapa Informação Tecnológica**; Planaltina, DF: Embrapa ..., 2004.

SUDU, B. et al. Retrieving SPAD values of summer maize using UAV hyperspectral data based on multiple machine learning algorithm. **Remote Sensing**, v. 14, n. 21, p. 5407, 2022.

SU, J. et al. Spectral analysis and mapping of blackgrass weed by leveraging machine learning and UAV multispectral imagery. **Computers and Electronics in Agriculture**, v. 192, p. 106621, 2022.

SWARBRECK, S. M. et al. A roadmap for lowering crop nitrogen requirement. **Trends in plant science**, v. 24, n. 10, p. 892-904, 2019.

TAIZ, L.; ZEIGER, E. Plant Physiology. Sinauer Associates, 2010.

TEIXEIRA, P. C. et al. Manual de métodos de análise de solo. [s.l.] **Embrapa** Brasília, DF, 2017.

TEWARI, R. K. et al. Oxidative stress and antioxidant responses in young leaves of mulberry plants grown under nitrogen, phosphorus or potassium deficiency. **Journal of Integrative Plant Biology**, v. 49, n. 3, p. 313-322, 2007.

THAI, Thi Huyen et al. Statistical analysis versus the M5P machine learning algorithm to analyze the yield of winter wheat in a long-term fertilizer experiment. **Agronomy**, v. 10, n. 11, p. 1779, 2020.

XIONG, D. et al. SPAD-based leaf nitrogen estimation is impacted by environmental factors and crop leaf characteristics. **Scientific reports**, v. 5, n. 1, p. 13389, 2015.

ZHANG, J. Y. et al. Photosynthetic performance and photosynthesis-related gene expression coordinated in a shade-tolerant species *Panax notoginseng* under nitrogen regimes. **BMC Plant Biology**, v. 20, p. 1-19, 2020.

ZHENG, C. et al. Remote sensing and machine learning in crop phenotyping and management, with an emphasis on applications in strawberry farming. **Remote Sensing**, v. 13, n. 3, p. 531, 2021.

ZULFIQAR, S. et al. Role of carotenoids in photosynthesis. **Carotenoids: Structure and function in the human body**, p. 147-187, 2021.

WANG, A. et al. Hyperspectral Sensing of Plant Diseases: Principle and Methods. **Agronomy**, 12(6), p. 1451, 2022.

WEN, P. F. et al. Effects of N application rate on N remobilization and accumulation in maize (*Zea mays* L.) and estimating of vegetative N remobilization using hyperspectral measurements. **Computers and Electronics in Agriculture**, v. 152, p. 166-181, 2018.

WICKAM, H. ggplot2: elegant graphics for data analysis. **Springer-V erlag. Accessed March**, v. 16, p. 2021, 2016.