

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

**PREDIÇÃO DO TEOR FOLIAR DE NITROGÊNIO POR
DADOS HIPERESPECTRAIS PARA CULTURA DO MILHO
CULTIVADO EM CAMPO**

Bianca Cavalcante da Silva

Engenheira agrônoma

UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

**PREDIÇÃO DO TEOR FOLIAR DE NITROGÊNIO POR
DADOS HIPERESPECTRAIS PARA CULTURA DO MILHO
CULTIVADO EM CAMPO**

Bianca Cavalcante da Silva

Orientador: Prof. Dr. Renato de Mello Prado

Coorientador: Prof. Dr. Cid Naudi Silva Campos

Coorientador: Prof. Dr. Fabio Henrique Rojo Baio

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutora em Agronomia (Produção Vegetal).

S586p

Silva, Bianca Cavalcante da

Predição do teor foliar de nitrogênio por dados hiperespectrais para cultura do milho cultivado em campo. / Bianca Cavalcante da Silva. -- Jaboticabal, 2025

70 p. : tabs., fotos

Tese (doutorado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientador: Renato de Mello Prado

Coorientador: Cid Naudi Silva Campos

1. Sensoriamento remoto. 2. Aprendizado de máquina. 3. Modelos de predição. 4. Fertilização. 5. Validação de níveis nutricionais. I. Título.

IMPACTO POTENCIAL DESTA PESQUISA

Promover o acesso a tecnologias digitais para produtores, independentemente do seu porte, pode reduzir os impactos ambientais e aumentar a produtividade, fortalecendo os esforços tecnológicos para erradicar a fome, incentivar uma agricultura sustentável e estimular inovações alinhadas aos Objetivos de Desenvolvimento Sustentável (ODS 2 e 9).

POTENTIAL IMPACT OF THIS RESEARCH

Promoting access to digital technologies for producers, regardless of their size, can reduce environmental impacts and increase productivity, strengthening technological efforts to eradicate hunger, encourage sustainable agriculture, and stimulate innovations aligned with the Sustainable Development Goals (SDGs 2 and 9).

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: PREDIÇÃO DO TEOR FOLIAR DE NITROGÊNIO POR DADOS HIPERESPECTRAIS
PARA CULTURA DO MILHO CULTIVADO EM CAMPO

AUTORA: BIANCA CAVALCANTE DA SILVA

ORIENTADOR: RENATO DE MELLO PRADO

COORIENTADOR: CID NAUDI SILVA CAMPOS

COORIENTADOR: FABIO HENRIQUE ROJO BAIO

Aprovada como parte das exigências para obtenção do Título de Doutora em Agronomia (Produção Vegetal), pela Comissão Examinadora:

Documento assinado digitalmente
 **RENATO DE MELLO PRADO**
Data: 21/02/2025 13:44:37-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. RENATO DE MELLO PRADO (Participação Virtual)
Departamento de Ciências da Produção Agrícola / FCAV UNESP Jaboticabal

Documento assinado digitalmente
 **EDILAINE ISTEFAINI FRANKLIN TRASPADINI**
Data: 21/02/2025 19:38:46-0300
Verifique em <https://validar.iti.gov.br>

Pós-Doutoranda EDILAINE ISTEFAINI FRANKLIN TRASPADINI (Participação Virtual)
Departamento de Ciência do Solo / FCAV UNESP Jaboticabal

Documento assinado digitalmente
 **LARISSA PEREIRA RIBEIRO TEODORO**
Data: 23/02/2025 17:11:19-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. LARISSA PEREIRA RIBEIRO TEODORO (Participação Virtual)
Universidade Federal de Mato Grosso do Sul (UFMS) / Chapadão do Sul/MS

Documento assinado digitalmente
 **GABRIEL BARBOSA DA SILVA JÚNIOR**
Data: 27/02/2025 07:43:55-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. GABRIEL BARBOSA DA SILVA JÚNIOR (Participação Virtual)
Departamento de Fitotecnia / Universidade Federal do Piauí (UFPI) - Teresina/PI

DADOS CURRICULARES DA AUTORA

BIANCA CAVALCANTE DA SILVA, nascida em Capanema-PA, em 29 de outubro de 1993, mas foi criada em Peixe-Boi-PA, filha de Isis de Nazaré Cavalcante da Silva e Francisco Ferdinando Silva. Formou-se em Engenharia Agrônômica pela Universidade Federal Rural da Amazônia – Câmpus Capanema, em maio de 2018. Durante a graduação foi bolsista de iniciação científica do CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) por três anos (2014 a 2017) e desenvolveu experimentos em feijão caupi e cupuaçuzeiro focado em nutrição de plantas sob orientação do Professor Dr. Ismael de Jesus Matos Viégas. Em julho de 2020, obteve o título de Mestre em Agronomia (Ciência do Solo), pela Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP) – Câmpus Jaboticabal pelo Programa de Pós-Graduação em Agronomia (Produção Vegetal), sob orientação do Professor Dr. Jairo Oswaldo Cazetta como bolsista CNPq a dissertação intitulada Crescimento inicial e nutrição mineral de plantas de *Moringa oleífera* adubadas com nitrogênio e potássio. Em fevereiro de 2025 obteve o título de Doutora em Agronomia (Produção Vegetal) pela Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), Câmpus Jaboticabal, sob orientação do Professor Dr. Renato de Mello Prado. Durante o período de doutoramento (2021 a 2025), desenvolveu pesquisas sobre o uso da agricultura digital (sensor hiperespectral) e modelos de aprendizado de máquina para predição do status do nitrogênio e pigmentos em milho. É membro do Grupo de Estudos em Nutrição de Plantas da Unesp (Genplant) e neste período realizou pesquisas com o pimentão, avaliando o uso do silício na mitigação de deficiências nutricionais durante o estágio de docência. A autora foi avaliadora de trabalhos em eventos, publicou artigos, participou do projeto sobre predição dos teores nutricionais de nitrogênio em milho por dados multiespectrais e correlação com a fotossíntese, projeto financiado com recursos da Fundect, Mato Grosso do Sul.

“Eu te amarei, ó SENHOR, fortaleza minha. O Senhor é o meu rochedo, e o meu lugar forte, e o meu libertador; o meu Deus, a minha fortaleza, em quem confio; o meu escudo, a força da minha salvação, e o meu alto refúgio. Na angústia invoquei ao Senhor, e clamei ao meu Deus; desde o seu templo ouviu a minha voz, aos seus ouvidos chegou o meu clamor perante a sua face.” (Salmos, 18, 1-2, 6).

A Deus, que me fortalece e direciona meus passos todos os dias da minha vida. Ao meu esposo, Bruno Maciel Reis da Silva, cujo apoio tem sido incansável; aos meus filhos, Maria Alice Cavalcante Aguiar e Benjamin Cavalcante Reis, que são presentes divinos e me motivam a dedicar esforços em prol de suas vidas. À minha mãe, Isis de Nazaré Cavalcante da Silva, e à minha bisavó, Maria da Conceição Silva, que, mesmo diante de adversidades, me ensinaram sobre o amor, valores e caráter, fundamentais para trilhar caminhos éticos e justos. Agradeço aos meus professores durante toda minha trajetória acadêmica desde a educação infantil até a educação superior em especial ao meu ex-orientador, Ismael de Jesus Matos Viégas e ao meu atual orientador, Renato de Mello Prado, que, ao longo da minha trajetória acadêmica, me transmitiram valiosos ensinamentos, não apenas no campo científico, mas também na importância de ser humano e compassivo.

Dedico.

AGRADECIMENTOS

A Deus, meu refúgio nos momentos de angústia, que foram muitos. Com Sua graça e misericórdia, Ele sempre segurou minhas mãos e me guiou por caminhos que jamais imaginei percorrer. Cheguei a Jaboticabal com minha filha de apenas três anos, vinda de Peixe-boi, no Pará, com uma bolsa de mestrado e um curso a ser concluído. Hoje, Maria Alice está com 9 anos, e Deus me abençoou com o nascimento do meu querido e saudável filho, Benjamin. Sinto Sua presença em cada passo dessa jornada e, todos os dias, louvarei ao Senhor em agradecimento por tudo que Ele fez e ainda fará.

Ao meu marido Bruno Maciel, Deus me presenteou com a sua vida. Durante o curso do meu doutorado, nos casamos. Você é um auxiliador, incentivador, amigo, marido e um excelente pai e profissional. Obrigada pelo seu amor, apoio e carinho.

Aos meus filhos, Maria Alice e Benjamin, sou muito grata. Deus me concedeu a dádiva de gerar, amar e cuidar de vocês. É por meio de vocês que me levanto todos os dias para ser não apenas estudante e pesquisadora, mas também mãe, dona de casa e esposa. O amor de vocês é um combustível diário, sempre me motivando, independentemente do dia que se apresenta.

Agradeço profundamente à minha família, especialmente à minha mãe solo, que se dedicou a me educar, amar, proteger e incentivar nos estudos. Mãe, muito obrigada! Minha bisá, que hoje conta com 91 anos, não apenas criou seus filhos, mas também seus netos e bisnetos. Seu coração gigante é uma bênção em minha vida, e sou imensamente grata por sua dedicação em cuidar e instruir tão bem a sua família. Agradeço também à minha avó, Raimunda Benedita, por seus ensinamentos, carinho e atenção.

Ao meu orientador, Prof. Dr. Renato de Mello Prado, que me acolheu, ensinou e orientou com firmeza, acreditando em mim mais do que eu mesma. O resultado foi a publicação dos artigos que antes me pareciam inatingíveis. Obrigada por ser um líder excepcional e por todas as boas lembranças do meu doutorado. Desejo a você muita saúde para continuar sua trajetória como pai, professor e líder do produtivo Genplant (Grupo de Estudos em Nutrição de Plantas da Unesp).

Aos meus coorientadores, Prof. Dr. Cid Naudi Silva Campos, que com paciência me orientou e apoiou em minha pesquisa no Mato Grosso do Sul, e Prof. Dr. Fabio

Henrique Rojo Baio, cuja experiência em agricultura de precisão foi fundamental durante meu doutorado, muito obrigada pelo suporte.

Agradeço aos amigos que estiveram ao meu lado durante essa jornada. Cada troca foi importante para meu amadurecimento pessoal e profissional. Em especial, à Fishua Dango, que me apoiou de diversas formas e me ensinou a ser uma pessoa melhor em todos os aspectos da minha vida. A Ligia, por suas orações e conselhos, ao pastor Auceu e sua família pelo acolhimento, e à irmã Fabiana e ao pastor Claudionor, que iluminaram meu caminho. Meu profundo agradecimento à irmã Lúcia, que continua a orar por mim em momentos difíceis. Que Deus retribua a todos!

Sou grata aos excelentes professores que tive desde a educação infantil até o ensino superior. A educação realmente transforma vidas. Agradeço também aos amigos do grupo de estudos GENPLANT, cujas reuniões foram enriquecedoras para minha formação.

Agradeço à Universidade Estadual Paulista “Júlio de Mesquita Filho” (FCAV–Unesp), especialmente ao departamento de Ciência do Solo, em nome da técnica de laboratório Claudinha, por todo o auxílio e aprendizado, que foram fundamentais para meu desenvolvimento. Agradeço igualmente ao Programa de Pós-graduação em Agronomia (Produção Vegetal) da Unesp, no qual realizei meu doutorado.

Por fim, agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelas bolsas que recebi ao longo de minha trajetória acadêmica, desde a graduação até o doutorado. Sem esse apoio, não teríamos conseguido realizar pesquisas e alcançar êxito profissional. Bolsa CNPq 140525/2021-1.

Obrigada a todos que me auxiliaram, direta ou indiretamente!

Bianca Cavalcante da Silva

SUMÁRIO

RESUMO	iii
ABSTRACT	iv
Capítulo 1 - Considerações gerais.....	1
1. INTRODUÇÃO	1
Capítulo 2. Applicability of Technological Tools for Digital Agriculture with a Focus on Estimating the Nutritional Status of Plants	6
1. Introduction	7
1.1. Precision Agriculture	7
2. Contribution of Artificial Intelligence to Precision Agriculture	9
3. Digital Agriculture and Precision Agriculture:	9
3.1. What Is Digital Agriculture?	9
3.2. How Does It Differ from Precision Agriculture?	10
3.3. Why can digital agriculture	10
4. Machine Learning.....	15
5. Challenges of Digital Agriculture.....	17
6. From Leaf to Harvest: How Interactions Between Spectral Readings, Leaf Nutrient Content, and Yield Affect the Robustness of Predictive Models	18
7. Final Considerations.....	20
8. References.....	21
Capítulo 3. New approach for predicting nitrogen and pigments in maize fromhyperspectral data and machine learning models.....	34
A B S T R A C T	34
1. Introduction	35
2. Materials and methods	37
2.1 Location and characteristics of the experimental area.....	37
2.1.1 Experimental design	38
2.1.2 Conduction of the experiment and application of treatments	38

3.	Analyses.....	39
3.1	Pigments (chlorophyll a + b and carotenoids)	39
3.2	Nitrogen content	40
3.3	Spectral assessments	40
3.4	Machine learning analysis	41
4.	Results	42
4.1	Variability of the reflectance factor spectrum under different N treatments	42
4.2	Evaluation of machine learning models.....	43
5.	Discussion	45
5.1	Variability of the reflectance factor spectrum under different N treatments	45
6.	Conclusions	47
7.	References	48
	Capítulo 4. Considerações finais e perspectivas futuras	55
	Apêndice	56

PREDIÇÃO DO TEOR FOLIAR DE NITROGÊNIO POR DADOS HIPERESPECTRAIS PARA CULTURA DO MILHO CULTIVADO EM CAMPO

RESUMO

Os avanços em sensoriamento remoto, incluindo o uso de sensores hiperespectrais, permitem a coleta de dados detalhados sobre as condições das plantações em tempo real. Essas tecnologias oferecem uma visão mais precisa da saúde das culturas, possibilitando intervenções mais direcionadas e eficazes, reduzindo desperdícios e aumentando a produtividade de forma sustentável. A integração dessas ferramentas com sistemas de gestão agrícola pode garantir uma tomada de decisão mais informada e estratégica. Os modelos de aprendizado de máquina, ao processarem dados hiperespectrais, podem fornecer diagnósticos mais rápidos e precisos, identificando áreas que necessitam de mais ou menos fertilizante com base no conteúdo de nitrogênio e pigmentos. Essa abordagem permite que os agricultores ajustem a aplicação de fertilizantes em tempo real, evitando excessos ou deficiências nutricionais nas plantas, o que pode resultar em uma produção mais sustentável e eficiente. Nosso estudo buscou entender como sensores com dados hiperespectrais e modelos de aprendizado de máquina podem ser integrados para otimizar a gestão nutricional de forma mais eficaz, com práticas agrícolas sustentáveis, que minimizem impactos ambientais e maximizem a produtividade das culturas. Esta pesquisa foi conduzida em campo com o objetivo de avaliar a precisão de cinco algoritmos de Machine Learning (ML) na estimativa dos teores de nitrogênio, clorofila e carotenoides em folhas de milho, em diferentes estádios fenológicos. Para isso, utilizaram-se dados de bandas hiperespectrais, de modo a analisar a assinatura espectral ao longo do desenvolvimento da cultura, verificando a relação dessas bandas com variáveis nutricionais (doses) e sua aplicação como variáveis de entrada nos modelos de predição. Os tratamentos foram organizados em um esquema fatorial com quatro doses de nitrogênio (0, 54, 108 e 216 kg ha⁻¹) e cinco estágios fenológicos de coleta de folhas (V6 a V14). Os modelos de ML testados foram redes neurais artificiais – ANN, árvore de decisão adaptada para problemas de predição – M5P, árvore de decisão REPTree, floresta aleatória - RF, máquina de vetores de suporte polinomial – PSVM e ZeroR - ZR (controle). As bandas espectrais de 530–560 nm e 690–750 nm foram eficazes na estimativa de nutrientes. O modelo RF apresentou o melhor desempenho, com maior correlação e menor erro absoluto médio (MAE) para todas as variáveis, enquanto a SVM e o RF foram eficazes na previsão do conteúdo de nitrogênio. No entanto, para que essas tecnologias digitais sejam amplamente adotadas dentro da nutrição de plantas, é essencial que estudos validem se os níveis identificados correspondem com precisão às condições nutricionais "normais", "deficientes" ou "excessivos", e se essas variações têm, de fato, um impacto significativo no crescimento e no rendimento das culturas.

Palavras chave: Sensoriamento remoto; Aprendizado de máquina (ML); Modelos de predição; Fertilização; Validação de níveis nutricionais.

PREDICTION OF FOLIAR NITROGEN CONTENT USING HYPERSPECTRAL DATA FOR FIELD-GROWN CORN CULTIVATION

ABSTRACT

Advances in remote sensing, including the use of hyperspectral sensors, enable the detailed collection of data on crop conditions in real-time. These technologies provide a more accurate view of crop health, allowing for more targeted and effective interventions, reducing waste and sustainably increasing productivity. The integration of these tools with agricultural management systems can ensure more informed and strategic decision-making. Machine learning models, by processing hyperspectral data, can provide faster and more accurate diagnostics, identifying areas that require more or less fertilizer based on nitrogen and pigment content. This approach allows farmers to adjust fertilizer application in real-time, avoiding excesses or nutritional deficiencies in plants, which can lead to more sustainable and efficient production. Our study aimed to understand how sensors with hyperspectral data and machine learning models can be integrated to optimize nutritional management more effectively, using sustainable agricultural practices that minimize environmental impacts and maximize crop productivity. This field research study evaluated the accuracy of five machine learning (ML) algorithms in estimating nitrogen, chlorophyll, and carotenoid contents in corn leaves at different phenological stages using hyperspectral band data. Hyperspectral measurements were employed to analyze spectral signatures throughout crop development, assess their relationship with nutritional variables (N application rates), and evaluate their effectiveness as input variables in predictive models. The treatments were organized in a factorial scheme with four nitrogen doses (0, 54, 108, and 216 kg ha⁻¹) and five phenological stages of leaf collection (from V6 to V14). The tested ML models included artificial neural networks (ANN), a decision tree adapted for prediction problems (M5P), the REPTree decision tree, random forest (RF), polynomial support vector machine (PSVM), and ZeroR (ZR) as the control. The spectral bands of 530–560 nm and 690–750 nm were effective in estimating nutrient content. The RF model performed the best, with the highest correlation and the lowest mean absolute error (MAE) for all variables, while the SVM and RF were effective in predicting nitrogen content. However, for these digital technologies to be widely adopted in plant nutrition, it is essential for studies to validate whether the identified levels accurately correspond to "normal," "deficient," or "excessive" nutritional conditions, and whether these variations significantly impact crop growth and yield.

Keywords: Remote sensing; Machine learning (ML); Prediction models; Fertilization; Nutritional level validation.

Capítulo 1 - Considerações gerais

1. INTRODUÇÃO

A agricultura de precisão (AP) tem se consolidado como uma ferramenta essencial para a modernização da agricultura, especialmente diante da crescente necessidade de aumentar a produtividade de forma sustentável. Essa abordagem utiliza tecnologias avançadas, como sensoriamento remoto, drones, e sistemas de informação geográfica (SIG), para coletar e analisar dados detalhados sobre as condições do solo, do clima e das culturas (Sharma et al., 2024). Com essas informações, é possível realizar ajustes específicos em tempo real, otimizando o uso de recursos como água, nutrientes e defensivos agrícolas.

Diferentemente do manejo convencional, que considera a área de cultivo como um bloco homogêneo, a AP reconhece as variações existentes dentro de um mesmo talhão, incluindo diferenças de solo, microclima e estado nutricional das plantas. Esse enfoque possibilita que práticas agrícolas, como aplicação de insumos, irrigação e controle de pragas, sejam ajustadas conforme as necessidades específicas de cada área da propriedade

Diferentemente do manejo convencional, que considera a área de cultivo como um bloco homogêneo, a AP reconhece as variações existentes dentro de um mesmo talhão, incluindo diferenças de solo, microclima e estado nutricional das plantas. Essa abordagem permite que as práticas agrícolas, como aplicação de insumos (fertilizantes por exemplo), irrigação, controle de pragas e etc, sejam ajustadas conforme as necessidades específicas de cada ponto da propriedade.

Além disso, a AP contribui para a sustentabilidade, ajudando a minimizar o desperdício de insumos, reduzir impactos ambientais e melhorar a eficiência da produção (Himesh et al., 2018). Através da personalização das práticas agrícolas para diferentes áreas de uma propriedade, a agricultura de precisão permite que os produtores tomem decisões mais informadas, aumentando a rentabilidade e garantindo a segurança alimentar a longo prazo.

Sensores hiperespectrais, divididos em modelos que geram imagens ou fornecem valores espectrais, estão se tornando mais acessíveis, especialmente quando acoplados os veículos aéreos não tripulados (UAVs) (Bhakthan & Loganathan,

2024). Esses sensores capturam uma vasta gama de comprimentos de onda, oferecendo alta resolução espacial e espectral, o que permite identificar e monitorar características das plantas, solo e outros elementos. Equipados com sensores de imagem de alta resolução, os UAVs superam satélites na detecção de estresses nas culturas e na análise do solo, além de não dependerem da ausência de nuvens (Comba et al., 2020).

Sensores hiperespectrais, que capturam uma amostragem contínua do espectro de 350 a 1100 nm, são mais eficientes que sensores multiespectrais, que medem apenas em algumas bandas específicas (Xie et al., 2021; Teodoro et al., 2023). O uso do sensor também pode ser realizado em laboratório, pois algumas pesquisas exigem a redução de ruídos causados por fatores como nuvens, umidade e temperatura em campo, que podem gerar erros e demandar frequentes calibrações do sensor, dificultando seu manuseio. Dados testados em laboratório para minimizar esses erros podem ser posteriormente ajustados, permitindo que a máquina seja alimentada com novos dados coletados em campo.

Quando as plantas apresentam estresse, suas folhas refletem condições específicas, devido à presença de substâncias como a clorofila, que indicam o estado nutricional e fenológico da planta (Teodoro et al., 2023). As regiões visíveis e infravermelhas do espectro eletromagnético apresentam curvas de refletância características (Teodoro et al., 2023). A clorofila absorve luz nas faixas vermelha (620 - 750 nm) e azul (450 - 495 nm), refletindo o verde (495 - 570 nm), enquanto carotenoides refletem luz amarela (570 - 590 nm) e laranja (590 - 620 nm), e a água afeta a reflexão no infravermelho próximo (700 - 2500 nm) (Gitelson et al., 2003).

Como exemplo, o estudo de Silva et al. (2024) identificou atenuações espectrais em comprimentos de onda específicos, particularmente nas regiões verde (aproximadamente 550 nm) e borda vermelha (cerca de 750 nm). Esses resultados sugerem que folhas saudáveis apresentam um padrão característico de absorção e reflexão de luz, com a clorofila absorvendo nas faixas vermelha e azul e refletindo nas faixas verde e infravermelho próximo.

Por sua vez, o estudo de Peng et al. (2020) apontou faixas espectrais sensíveis ao conteúdo de nitrogênio nas folhas, com destaque para as regiões vermelha e infravermelho próximo. Eles sugerem que a análise espectral ao longo de toda a faixa

espectral, e não em bandas específicas, é mais eficaz para identificar deficiências nutricionais. A relação entre a condição fenológica da planta, a clorofila e a leitura espectral são claras durante o crescimento vegetativo, quando a produção de clorofila é alta, resultando em maior absorção de luz na faixa vermelha e menor refletância (Peng et al., 2020). Já na senescência, a redução da clorofila provoca um aumento na refletância, o que pode ser monitorado de forma não destrutiva por sensores espectrais (Peng et al., 2020).

Segundo Wen et al. (2018), a reflectância na luz visível (436–698 nm) foi característica de folhas com deficiência de N, enquanto a reflectância no infravermelho próximo (NIR, 740–1350 nm) foi mais elevada em plantas que receberam fertilização nitrogenada completa. Contudo, a maioria das pesquisas sobre esse assunto concentra-se na análise de índices de vegetação tradicionais para a predição de N, o que limita a precisão e a aplicabilidade dos resultados. Isso ocorre porque essas abordagens analisam apenas a relação entre algumas bandas espectrais, sem explorar o potencial de todas as bandas.

Além disso, o problema da saturação em áreas com alta biomassa verde e índice de área foliar elevado pode levar à subestimação dos valores de biomassa ou gerar resultados pouco variáveis (Gu et al., 2013). Dessa forma, são poucos os estudos que combinam a leitura de diversas bandas espectrais com diferentes métodos de aprendizado de máquina para avaliar o conteúdo de N e os pigmentos fotossintéticos nas folhas de milho utilizando sensores hiperespectrais.

Diante disso, torna-se necessário avançar em estudos que validem a acurácia dos níveis de nutrientes identificados por sensores no diagnóstico do estado nutricional das plantas, seja em condições de deficiência, suficiência ou excesso de nutrientes, especialmente em pesquisas de campo que relacionem esses parâmetros com a produtividade das culturas.

Esta pesquisa teve como objetivo integrar ferramentas de agricultura digital para aprimorar a nutrição das plantas, abordando as principais lacunas existentes na área. Além disso, buscou-se aplicar o uso de sensores hiperespectrais e modelos de aprendizado de máquina para a predição do conteúdo de nitrogênio e pigmentos nas plantas, com o intuito de otimizar a gestão nutricional e promover práticas agrícolas eficientes e sustentáveis.

REFERÊNCIAS

- Bhakthan, S. M., & Loganathan, A. (2024). A hyperspectral unmixing model using convolutional vision transformer. *Earth Science Informatics*, 17(3), 2255–2273. <https://doi.org/10.1007/s12145-024-01273-y>
- Comba, L., Biglia, A., Ricauda Aimonino, D., Tortia, C., Mania, E., Guidoni, S., & Gay, P. (2020). Leaf Area Index evaluation in vineyards using 3D point clouds from UAV imagery. *Precision Agriculture*, 21(4), 881–896. <https://doi.org/10.1007/s11119-019-09699-x>
- Gitelson, A. A., Gritz †, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, 160(3), 271–282. <https://doi.org/10.1078/0176-1617-00887>
- Gu, Y., Wylie, B. K., Howard, D. M., Phuyal, K. P., & Ji, L. (2013). NDVI saturation adjustment: A new approach for improving cropland performance estimates in the Greater Platte River Basin, USA. *Ecological Indicators*, 30, 1–6. <https://doi.org/10.1016/j.ecolind.2013.01.041>
- Himesh, S., Rao, E. V. S. P., Gouda, K. C., Ramesh, K. V., Rakesh, V., Mohapatra, G. N., Rao, B. K., Sahoo, S. K., & Ajilesh, P. (2018). Digital revolution and Big Data: a new revolution in agriculture. *CABI Reviews*, 1–7. <https://doi.org/10.1079/PAVSNNR201813021>
- Peng, Y., Zhang, M., Xu, Z., Yang, T., Su, Y., Zhou, T., Wang, H., Wang, Y., & Lin, Y. (2020). Estimation of leaf nutrition status in degraded vegetation based on field survey and hyperspectral data. *Scientific Reports*, 10(1), 4361. <https://doi.org/10.1038/s41598-020-61294-7>
- Sharma, A., Prakash, A., Bhambota, S., & Kumar, S. (2024). Investigations of precision agriculture technologies with application to developing countries. *Environment, Development and Sustainability*. <https://doi.org/10.1007/s10668-024-04572-y>

- Silva, B. C. da, Prado, R. de M., Baio, F. H. R., Campos, C. N. S., Teodoro, L. P. R., Teodoro, P. E., Santana, D. C., Fernandes, T. F. S., Silva Junior, C. A. da, & Loureiro, E. de S. (2024). New approach for predicting nitrogen and pigments in maize from hyperspectral data and machine learning models. *Remote Sensing Applications: Society and Environment*, 33, 101110. <https://doi.org/10.1016/j.rsase.2023.101110>
- Teodoro, L. P. R., Estevão, R., Santana, D. C., Oliveira, I. C. de, Lopes, M. T. G., Azevedo, G. B. de, Rojo Baio, F. H., da Silva Junior, C. A., & Teodoro, P. E. (2023). Eucalyptus Species Discrimination Using Hyperspectral Sensor Data and Machine Learning. *Forests*, 15(1), 39. <https://doi.org/10.3390/f15010039>
- Wen, P.-F., Wang, R., Shi, Z.-J., Ning, F., Wang, S.-L., Zhang, Y.-J., Zhang, Y.-H., Wang, Q., & Li, J. (2018). 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*, 152, 166–181. <https://doi.org/10.1016/j.compag.2018.07.009>
- Xie, Q., Lai, K., Wang, J., Lopez-Sanchez, J. M., Shang, J., Liao, C., Zhu, J., Fu, H., & Peng, X. (2021). Crop Monitoring and Classification Using Polarimetric RADARSAT-2 Time-Series Data Across Growing Season: A Case Study in Southwestern Ontario, Canada. *Remote Sensing*, 13(7), 1394. <https://doi.org/10.3390/rs13071394>



Review

Capítulo 2. Applicability of Technological Tools for Digital Agriculture with a Focus on Estimating the Nutritional Status of Plants

Bianca Cavalcante da Silva ^{1,*}, Renato de Mello Prado ¹, Cid Naudi Silva Campos ², Fábio Henrique Rojo Baio ², Larissa Pereira Ribeiro Teodoro ², Paulo Eduardo Teodoro ² and Dthenifer Cordeiro Santana ³

¹ Department of Soil Science, São Paulo State University “Júlio de Mesquita Filho” UNESP/FCAV, Jaboticabal 14884-900, Brazil; rm.prado@unesp.br

² Department of Agronomy, Federal University of Mato Grosso Do Sul (UFMS), Chapadão Do Sul 79560-000, Brazil; cid.campos@ufms.br (C.N.S.C.); fabio.baio@ufms.br (F.H.R.B.); larissa_ribeiro@ufms.br (L.P.R.T.); paulo.teodoro@ufms.br (P.E.T.)

³ Department of Agronomy, State University of São Paulo (UNESP), Ilha Solteira 15385-000, Brazil; dthenifer.santana@unesp.br

* Correspondence: bianca.cavalcante@unesp.br

Abstract: The global transition to a digital era is crucial for society, as most daily activities are driven by digital technologies aimed at enhancing productivity and efficiency in the production of food, fibers, and bioenergy. However, the segregation of digital techniques and equipment in both rural and urban areas poses significant obstacles to technological efforts aimed at combating hunger, ensuring sustainable agriculture, and fostering innovations aligned with the United Nations Sustainable Development Goals (SDGs 02 and 09). Rural regions, which are often less connected to technological advancements, require digital transformation to shift from subsistence farming to market-integrated production. Recent efforts to expand digitalization in these areas have shown promising results. Digital agriculture encompasses terms such as artificial intelligence (AI), the Internet of Things (IoT), big data, and precision agriculture integrating information and communication with geospatial and satellite technologies to manage and visualize natural resources and agricultural production. This digitalization involves both internal and external property management through data analysis related to location, climate, phytosanitary status, and consumption. By utilizing sensors integrated into unmanned aerial vehicles (UAVs) and connected to mobile devices and machinery, farmers can monitor animals, soil, water, and plants, facilitating informed decision-making. An important limitation in studies on nutritional diagnostics is the lack of accuracy validation based on plant responses, particularly in terms of yield. This issue is observed even in conventional leaf tissue analysis methods. The absence of such validation raises concerns about the reliability of digital tools under real field conditions. To ensure the effectiveness of spectral reflectance-based diagnostics, it is essential to conduct additional studies in commercial fields across different regions. These studies are crucial to confirm the accuracy of these methods and to strengthen the development of digital and precision agriculture.

Keywords: digital agriculture; hyperspectral sensor; sustainable agriculture; smart agriculture; prediction of plant nutritional deficiency

Academic Editors: Qingting Liu, Tao Wu and Zhigang Zhang

Received: 24 March 2025

Revised: 12 May 2025

Accepted: 15 May 2025

Published: 19 May 2025

Citation: Silva, B.C.d.; Prado, R.d.M.; Campos, C.N.S.; Baio, F.H.R.; Teodoro, L.P.R.; Teodoro, P.E.; Santana, D.C. Applicability of Technological Tools for Digital Agriculture with a Focus on Estimating the Nutritional Status of Plants. *AgriEngineering* **2025**, *7*, 161. <https://doi.org/10.3390/agriengineering7050161>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Precision Agriculture

Precision agriculture (PA), also known as site-specific management, offers a sustainable solution to the growing demand for food. By integrating agricultural sciences with information technology, this approach analyzes spatial and temporal variations within fields [1]. PA techniques have been developed to conserve natural resources such as water and soil, and optimize the use of fertilizers, irrigation, pesticides, and weed identification, among others, while reducing production costs [2–4]. The modernization of agriculture, driven by embedded electronics and remote sensing, has revolutionized traditional methods and fostered industry growth [1], adapting agricultural practices to the specific needs of different field areas.

The contribution of digital agriculture to precision agriculture lies in the application of digital technologies to collect, interpret, and communicate data that support decision-making on farms and throughout the production chain, characterizing a data-driven agricultural approach. In this context, precision agriculture has advanced through the use of smart machinery, sensors, drones, and satellites, which generate high-resolution data analyzed by machine learning algorithms. This analysis enables the accurate detection of plant diseases and nutrient deficiencies. Practical examples, such as automated cattle feeding based on weight, highlight the importance of data collection and processing. Thus, digital agriculture enhances the potential of precision agriculture by promoting greater efficiency and improving decision-making in the field.

Remote sensing plays a crucial role in precision agriculture, serving as a tool for acquiring data and information over large cultivated areas in a short period. This technology enables the analysis of target attributes based on their interaction with the electromagnetic spectrum [5]. Remote sensors are equipment designed to measure a wide range of data, including the reflectance of the electromagnetic spectrum coming from an object or plant.

There are two main types of hyperspectral sensors: imaging sensors and those that provide spectral values for each wavelength. With decreasing equipment costs, imaging-based sensors have become increasingly utilized, particularly when mounted on unmanned aerial vehicles (UAVs), expanding their application in aerial mapping and precision agriculture. Hyperspectral sensors capture a broad range of wavelengths, enabling high spatial (visual detail) and spectral (wavelength differentiation) resolution. This capability allows for the precise identification and monitoring of specific plant, soil, and surface characteristics [6], such as crop physiological monitoring, water stress detection, and nutrient assessment.

UAVs equipped with high-resolution imaging sensors are particularly effective in data collection, surpassing satellite imagery in detecting crop stress and analyzing soil attributes [7]. UAVs provide higher spatial resolution images and are not affected by cloud cover, ensuring continuous operation. These data are essential for developing decision-support systems, facilitating intelligent pest control, fertilization, and irrigation management [8].

Plant leaves under stress reflect specific conditions, making them efficient for spectral characterization, as they contain substances such as chlorophyll, which indicate the plant's phenological status [5]. The visible and infrared regions of the electromagnetic spectrum exhibit characteristic reflectance curves [9]. Hyperspectral sensors allow for continuous spectral sampling, covering wavelengths mostly ranging from 350 to 1100 nm, providing greater efficiency compared to multispectral sensors, which only capture electromagnetic reflectance in specific bands [5,10,11].

Chlorophyll primarily absorbs light in the red (600–700 nm) and blue (400–500 nm) spectral bands while reflecting green light (500–600 nm). Additionally, carotenoid

pigments reflect more light in the yellow-orange region (500–600 nm), whereas water molecules influence reflectance in the near-infrared region (700–1300 nm). When plants experience stress, such as water or nutrient deficiency, significant changes occur in reflectance, particularly in the near-infrared and red regions [12].

In a recent study by [11], spectral attenuations were observed at specific wavelengths, particularly in the green spectrum (around 550 nm), forming a cluster between 530 and 560 nm, and in the red-edge spectrum (around 750 nm), forming another cluster between 690 and 820 nm. These results indicate that healthy leaves exhibit a typical pattern of light absorption and reflection, with chlorophyll playing a crucial role in absorbing red and blue wavelengths while reflecting green and near-infrared light [12].

In the study by [13], the spectral bands most sensitive to leaf nitrogen content in degraded temperate vegetation were identified in the red and near-infrared (NIR) regions, specifically at wavelengths of 468, 623, 624, 633, 652, 657, 668, 818, 821, 842, 937, and 938 nm. The nitrogen-sensitive bands were primarily found in the green, green-yellow edge, mid-red, and NIR regions of the spectrum. Similarly, [11,13] also suggested that the identification of nitrogen-sensitive wavelengths should be conducted across the entire spectral range rather than being limited to specific bands. This approach would enhance the accuracy of nitrogen content estimation in leaves, as it captures a broader range of spectral variations, leading to greater precision in detecting nutrient deficiencies.

The relationship between phenological stage, chlorophyll content, and spectral readings is well illustrated during the vegetative growth of plants. At this stage, chlorophyll production is high, resulting in greater light absorption in the red region (600–700 nm) and, consequently, lower reflectance. As the plant progresses to senescence, chlorophyll levels decline, leading to increased reflectance in the same spectral range. This phenomenon can be detected by spectral sensors, which are capable of monitoring these changes non-destructively [14].

The fundamental premise is to move beyond managing fields based on average conditions and instead consider the spatial variability of factors influencing crop productivity. Precision agriculture techniques must necessarily incorporate the concept of “spatial variability” [15]. Therefore, this review aims to discuss research, terminology, and challenges in digital agriculture, with the goal of optimizing plant nutrition effectively, precisely, at a low cost for farmers, and contributing to sustainable agriculture.

In this article, we first address the contribution of artificial intelligence to precision agriculture (Section 2), exploring its role in advancing modern farming practices. Next, we discuss digital agriculture and precision agriculture, with a focus on the related concepts, differences, and contributions to plant nutritional assessment (Section 3), which is subdivided into three parts: Section 3.1, ‘What Is Digital Agriculture?’; Section 3.2, ‘How Does It Differ from Precision Agriculture?’; and Section 3.3, ‘Why Can Digital Agriculture Make Significant Contributions to Assessing Plant Nutritional Status?’

In Section 4, we present key machine learning algorithm models applied to plant nutritional diagnostics in agricultural systems. Following this, Section 5 discusses the main challenges of digital agriculture, highlighting current limitations and opportunities for further development. Section 6, entitled ‘From Leaf to Harvest: How Interactions Between Spectral Readings, Leaf Nutrient Content, and Yield Affect the Robustness of Predictive Models’, explores the importance of validating plant nutritional diagnostics through crop productivity analysis. Finally, we present our final considerations (Section 7), where the main points discussed throughout the article are summarized, and future research directions are suggested. The references used throughout the article are listed at the end.

2. Contribution of Artificial Intelligence to Precision Agriculture

The integration of artificial intelligence (AI) into precision agriculture enhances farming practices by applying data-driven insights for more effective decision-making [16]. AI technologies, such as machine learning and computer vision, process vast amounts of data from various sources, including sensors, UAVs, and satellite imagery [17]. This enables detailed monitoring of crop health, soil conditions, and multiple environmental factors [1].

In recent years, artificial intelligence has gained prominence, largely due to the evolution of data formats, which have shifted from traditional structured formats to unstructured, semi-structured, and heterogeneous architectures, each with distinct characteristics [17]. Advances in large language models (LLMs), machine learning (ML), and neural networks (NNs) have continued to progress [18].

Artificial intelligence does not replicate human cognitive processes but rather simulates aspects of thinking and decision-making through data processing and algorithms [19,20]. Human intelligence is defined as the ability to think, plan, and solve complex problems [21]. Similarly to human cognition, AI has multiple interpretations and is often defined as “the ability of machines to perform tasks that typically require human intelligence” [19].

Remote sensing, combined with artificial intelligence, enables real-time data collection and analysis, optimizing and improving decision-making processes [22,23]. “Smart agriculture” emphasizes the application of AI and data science [24]. Machine learning (ML) algorithms facilitate the processing of sensor-derived data, enabling fast and automated analyses [25]. These algorithms can handle datasets of varying sizes and complexities, depending on model efficiency [26,27].

Artificial intelligence (AI) plays a fundamental role in precision agriculture by simulating human intelligence through machine learning. However, AI does not replace human intelligence, as proper management is required. Moreover, the interaction between artificial intelligence and human expertise is essential for optimizing results. While AI can rapidly process and analyze large volumes of data, the experience and judgment of agricultural professionals and farmers remain indispensable for making informed decisions based on the generated analyses. This collaboration between machines and humans enhances agricultural practices, enabling a more precise and effective approach to crop management.

3. Digital Agriculture and Precision Agriculture:

Concepts, Differences, and Contributions to Plant Nutritional Assessment

3.1. *What Is Digital Agriculture?*

The global transition to the digital era is crucial for society, as most daily activities are driven by digital technologies aimed at increasing productivity and efficiency in food, fiber, and bioenergy production [28–30]. However, the unequal access to digital techniques and equipment in both rural and urban areas represents a significant barrier, threatening technological efforts to combat hunger, ensure sustainable agriculture, and promote innovations aligned with the Sustainable Development Goals (ODS, 02 e 09) [31].

Remote regions, particularly in rural areas of less developed countries, are often disconnected from technological advancements. Digital transformation in these areas is essential for shifting from subsistence farming to production integrated into the global market [28]. By adopting digital technologies, small-scale farmers can efficiently manage agricultural systems [32]. The application of inputs in the field becomes more precise, utilizing collected data to determine the appropriate amount, timing, and location for nutrient,

water, and agrochemical applications, contributing to sustainable and efficient management [33,34].

One of the most common devices in modern agriculture is the mobile phone, which, through applications, facilitates information sharing. However, it is important to emphasize that connectivity is not solely dependent on mobile devices but rather on the availability of mobile network signals. These signals enable communication and telemetry, which are fundamental to digital agriculture. Furthermore, agricultural machinery has already been optimized with specialized software [35,36], allowing for even greater integration between technology and farming practices.

The concept of digital agriculture integrates various technologies, such as automation, robotics, the Internet of Things (IoT), remote sensing, big data, artificial intelligence, and data analytics into agricultural practices, ensuring sustainability [8,37,38]. The authors of [39] pointed out, to avoid confusion, that in this context, digital agriculture refers to the implementation of digital technologies in cropping and livestock systems to collect, interpret, and communicate data that support decision-making on farms and throughout the supply chain. In summary, it is a data-driven approach to agriculture [40].

This approach aims to maximize productivity through integrated sensors in UAVs, connected to mobile devices and agricultural machinery, enabling real-time monitoring of livestock, soil, water, and crops, as well as human interactions [30,41]. A notable example of digital agriculture is cattle monitoring in pastures: as an animal passes through a gate, it is identified and weighed, prompting an automated system to dispense a precise amount of feed. Another example is agricultural fleet management, where a tractor's engine temperature can be remotely monitored from a control center, independent of maps or production variables. The collection and interpretation of these data enable informed decision-making, assisting both in historical analysis and future predictions.

3.2. *How Does It Differ from Precision Agriculture?*

In digital agriculture, precision agriculture plays a key role, particularly in addressing spatial variability, such as fluctuations in soil nutrient levels across different locations. Additionally, precision agriculture accounts for temporal variability, referring to changes in agricultural conditions over time. These include climate fluctuations, crop growth variations, yield productivity shifts, and resource availability changes, such as water and nutrients.

Precision agriculture relies on advanced equipment, including machinery, sensors, UAVs, and satellites, to collect high-resolution data. Artificial intelligence (AI) processes this data through machine learning algorithms, generating predictive models. Once trained on these datasets, agricultural machinery can accurately detect plant diseases and nutrient deficiencies in the field. The earlier example of automated cattle feeding based on weight highlights how data collection and processing are fundamental in this context. Thus, digital agriculture integrates precision agriculture technologies with AI-driven analytics, optimizing agricultural productivity and providing farmers with a comprehensive overview of crop and farm conditions, thereby supporting informed decision-making.

3.3. *Why can digital agriculture make significant contributions to assessing plant nutritional status?*

In recent years, efforts to expand digitalization in agriculture—such as nutrient status sensors (whether mounted on UAVs or stationary) combined with predictive modeling—have yielded positive results in rural areas [42]. Notably, these advancements have been applied to plant nutritional assessment, where hyperspectral electromagnetic spectra have been utilized to predict nitrogen, chlorophyll, and carotenoid content (Table 1). Among predictive models, Random Forest (RF) has emerged as the most effective, owing

to its ability to process large datasets and detect complex spectral patterns, leading to highly accurate nutrient estimations [11].

Table 1. Applications of Remote Sensing and Machine Learning for Plant Nutrient and Stress Analysis.

Nutrient Status/Plant Stress	Sensor and Spectral Bands	Data Type/Visualization	Machine Learning Algorithm/Model	Plant Species	Reference
Prediction of N, chlorophyll a + b, and carotenoid content	Hyperspectral sensor STS-VIS (Ocean Insight, USA). Spectral bands from 450 to 850 nm	Graphs generated from electromagnetic spectrum readings	Artificial Neural Networks (ANN); M5P decision tree; REPTree (REPT); Random Forest (RF); Polynomial Support Vector Machine (SVMP); ZeroR (ZR)	Maize (<i>Zea mays</i> L.)	[11]
N, P, and K contents	Reflectance values obtained in the following spectral bands (SBs): red (660 nm), green (550 nm), near-infrared (NIR, 735 nm), and red edge (790 nm)	Spectral images, spectral bands	J48 decision tree and REPTree; Random Forest (RF); Artificial Neural Network (ANN); Support Vector Machine (SVM); Logistic Regression (LR, used as control)	Soybean genotypes (<i>Glycine max</i> L.)	[43]
Boron deficiency or toxicity	MSI sensor (Sentinel-2)	Orbital images, Vegetation Indices (NDVI, NDRE, EVI, CI, PRSI, CCCI, MTCl, HMSSI)	Spectral Angle Mapping (SAM) algorithm	Ten-year-old clones of <i>Eucalyptus</i> MA-2000	[44]
P, K, Mg, Ca, S, B, Zn, Mn, Fe, Cu	5-band multispectral camera (Micasense Altum, Micasense, USA): blue (465–485 nm), green (550–570 nm), red (663–673 nm), red edge (712–722 nm), and near-infrared (NIR, 820–1000 nm)	Multispectral images from Unmanned Aerial Vehicles (UAVs)	Multiple regression algorithms, such as ElasticNet, Lasso regression, Linear SVM, PLSR, Random Forest (RF), and Ridge regression	Sweet orange seedlings ‘Hamlin’ or ‘Valencia’ grafted onto more than 30 different rootstocks	[45]
Elevated Nitrogen levels	The Sensefly Sequoia multispectral sensor evaluates vegetation indices in the following spectral bands: green (550 nm), red (660 nm), red edge (735 nm), and near-infrared (NIR, 790 nm)	Image (based on reflectance data), followed by processing of vegetation index models (NDVI, NDRE)	Statistical Model: $Y_{ijkl} = \mu + B_k + C_l + V_i + N_j + CV_{il} + CN_{jl} + VN_{ij} + CVN_{ijl} + e_{ijkl}$	Maize genotypes (<i>Zea mays</i> L.)	[46]

N, P, K	Vis-NIR HSI sensor (Visible-Near-Infrared Hyperspectral Imaging) (400–1000 nm)	Visible-Near-Infrared Hyperspectral Image (Vis-NIR HSI)	Partial Least Squares Regression (PLSR) models	Cocoa (<i>Theobroma cacao</i> L.)	[47]
Predicting	Multispectral sensor MicaSense Red-Edge (2017/2018 season) and multispectral sensor Sensefly Parrot Sequoia (2018/2019 season). Bands (G, R, RE, NIR) used for	Images and spectral	REPTree (REPT); RF; K-Nearest Neighbor (K = 1, K = 5, K = 10) (1NN, 5NN, 10NN); SVM-RBF (SVMR);	11 Maize	
leaf nitrogen concentration	vegetation index calculations, with slight variations between sensors/years (e.g., G: 560/550 nm, R: 668/660 nm, RE: 717/735 nm, NIR: 842/790 nm)	bands	Polynomial Support Vector Machine (SVMP); Linear Regression (LR); Radial Basis Function Regression (RBF)	cultivars	[48]
Predicting	Multispectral sensor Parrot Sequoia. Spectral bands captured: green (530–570		RF; ANN; SVM; Decision	33,600 Valencia orange trees (<i>Citrus</i>	
nitrogen content	nm), red (640–680 nm), red edge (730–740 nm), and near-infrared (NIR, 770–810 nm)	Multispectral images	Trees (DT); RNA (Artificial Neural Network—same as ANN)	<i>sinensis</i> ‘Valencia’) planted on Citrumelo oscillating rootstock	[49]
Predicting leaf nitrogen concentration: low (≤ 27 g·kg ⁻¹), medium (> 27 & ≤ 29 g·kg ⁻¹), and high (> 29 g·kg ⁻¹)	Parrot Sequoia sensor. Green (510–590 nm), red (620–700 nm), red edge (725–745 nm), and near-infrared (NIR, 750–830 nm) bands	Multispectral images, spectral bands	Constrained Energy Minimization; Linear Spectral Unmixing; Mixture-Tuned Matched Filtering; Minimum Distance; Orthogonal Subspace Projection; Spectral Angle Mapper; 1 Spectral Information Divergence	Valencia sweet orange (<i>Citrus sinensis</i> ‘Valencia’), grafted onto Citrumelo Swingle rootstock	[50]

Abbreviations used in the table: ANN: Artificial Neural Networks; RF: Random Forest; SVM: Support Vector Machine; PLSR: Partial Least Squares Regression; REPT/REPTree: Reduced Error Pruning Tree; LR: Logistic Regression; MSI: Multispectral Imager; HSI: Hyperspectral Imaging; NIR: Near-Infrared; NDVI: Normalized Difference Vegetation Index; NDRE: Normalized Difference Red Edge Index; UAV: Unmanned Aerial Vehicle.

On the other hand, to estimate nitrogen (N), phosphorus (P), and potassium (K) levels, the study by the authors in [43] used predefined spectral band imagery to calculate various vegetation indices, including NDVI (Normalized Difference Vegetation Index), NDRE (Normalized Difference Red Edge Index), GNDVI (Green Normalized Difference Vegetation Index), SAVI (Soil-Adjusted Vegetation Index), MSAVI (Modified Soil-Adjusted Vegetation Index), MCARI (Modified Chlorophyll Absorption in Reflectance Index), EVI (Enhanced Vegetation Index), and SCCC (Simplified Canopy Chlorophyll Content Index) (Table 1).

In the study mentioned above, in addition to testing different algorithms, three distinct data input methods (spectral bands, vegetation indices, and their combinations) were analyzed to determine which type of spectral data would yield the highest classification accuracy (Table 1). The input configuration that showed the best performance for machine learning algorithms was the use of spectral bands, with the J48 (decision tree) and SVM (support vector machine) algorithms standing out in predicting N, P, and K levels [43], confirming the findings of [11] mentioned in this review (Table 1).

Another significant study investigates the prediction of boron deficiency or toxicity in eucalyptus. Several vegetation indices were analyzed, including CCCI (Canopy Chlorophyll Content Index), CIGreen (Green Chlorophyll Index), EVI (Enhanced Vegetation Index), HMSSSI (Heavy Metal Stress Sensitive Index), MTCI (MERIS Terrestrial Chlorophyll Index), NDRE (Normalized Difference Red Edge Index), and NDVI (Normalized Difference Vegetation Index). Among these, NDVI proved to be the most effective in distinguishing adequate boron levels in *Eucalyptus* spp., both in young and mature plants. Additionally, the CCCI and PRSI (Photochemical Reflectance Stress Index) indices were used to identify deficiencies [44].

Spectral band analysis showed that the red region contributes the most to characterizing boron-deficient levels, with red edge bands at 740 nm, 782 nm, and 864 nm, along with NIR at 835 nm and the water vapor band at 945 nm, which are essential for discriminating adequate boron levels [44].

Some nutritional states were identified through analyses of different parameters, such as wavelength evaluation and algorithm application [45] (Table 1), using predefined wavelengths: blue (465–485 nm), green (550–570 nm), red (663–673 nm), red edge (712–722 nm), and near-infrared (820–1000 nm). For this analysis, multiple regression algorithms were applied, including ElasticNet, Lasso regression, Linear SVM, PLSR, Random Forest, and Ridge regression. The gradient boosting regression algorithm proved particularly effective for small datasets, outperforming the other tested methods [45]. However, one of the main challenges of this research was determining the best model and wavelength for each evaluated nutrient, a question that remains unresolved in the study.

Reflected radiation provides valuable information about plant nutrition,

and vegetation indices are used to quantify this reflectance. Examples include the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Red Edge Index (NDRE), as utilized by [46]. Studies show that high nitrogen doses (180 kg ha^{-1}) are positively related to these vegetation indices, resulting in a better spectral response in maize. The equation presented in Table 1 represents the observation in the k -th block, assessed in the i -th variety, j -th nitrogen fertilization, and l -th harvest; μ is the overall mean; Bk is the block effect, considered fixed; Cl is the harvest effect, considered random; Vi is the variety effect, considered fixed; NNN is the nitrogen level effect, considered fixed; $CVil$ is the harvest by variety interaction, considered random; $CNNMl$ is the harvest by nitrogen level interaction, considered random; $VNij$ is the variety by nitrogen level interaction, considered fixed; $CVNijl$ is the interaction between planting season, variety, and nitrogen, considered random; and ϵ_{ijkl} is the random error associated with observation Y_{ijk} [46].

An interesting study by [47] used a Vis-NIR hyperspectral sensor operating in the 400 to 1000 nm spectral range in a laboratory setting to determine the levels of N, P, K, and Ca. The analysis was carried out using the partial least squares regression (PLSR) model, which can process large datasets generated by hyperspectral images. Wavelengths in the 400 to 450 nm, 530 to 540 nm, 630 to 780 nm, and 990 to 1000 nm ranges were crucial for predicting foliar nitrogen levels. The model also predicted foliar phosphorus levels in the 700 to 1000 nm range. For potassium prediction, the 730 to 1000 nm range presented a higher number of informative wavelengths compared to other nutrients. Meanwhile, wavelengths from 470 to 520 nm, 570 to 650 nm, 710 to 730 nm, 770 to 820 nm, and 860 to 1000 nm were used to estimate calcium levels [47].

To predict nitrogen levels, research was conducted using multispectral images [48– 50]. The input data included spectral bands and vegetation indices, such as the Normalized Difference Vegetation Index (NDVI), the Normalized Difference Red Edge Index (NDRE), the Green Normalized Difference Vegetation Index (GNDVI), and the Soil-Adjusted Vegetation Index (SAVI). These data were analyzed by various models, including REPTree (REPT), Random Forest (RF), K-nearest neighbor ($K = 1, 5, \text{ and } 10$), SVM RBF, Polynomial Support Vector Machine (SVMP), Linear Regression (LR), and RBF Regression. Among these, the Random Forest algorithm showed the best performance, with the vegetation indices contributing more to the model's effectiveness than the spectral bands alone [49].

Among the 33 vegetation indices analyzed by [49], 12 exhibited regression coefficients greater than 0.5. These indices include the Canopy Chlorophyll Content Index (CCCI), Green Chlorophyll Index (CG), Red Edge Chlorophyll Index (Clrededge), Chlorophyll Vegetation Index (IVC), Green Soil-Adjusted Vegetation Index (GSAVI), Modified Soil-Adjusted Vegetation Index (MSAVI), Renormalized Difference Vegetation Index (RDVI), Soil-Adjusted Vegetation Index (SAVI), Simple Ratio 750/550 (SR750/550), Simple Ratio 800/550 (SR800/550), Triangular Vegetation Index (TriVI), and Wide Dynamic Range Vegetation Index (WDRVI). Only five of these indices showed correlation coefficients greater than 0.7 in linear and exponential regressions: MSAVI, SR750/550, SR800/550, TriVI, and WDRVI. The Random Forest (RF) model

outperformed most other machine learning algorithms, achieving a coefficient of determination (R^2) of 0.90, with the XGBoost model also providing comparable metrics [49]. This means that the model explains 90% of the variation in the predictive data, indicating high accuracy in the model's predictions.

To investigate the relationship between nitrogen content in orange trees and surface reflectance obtained from imagery, the study in [48] applied spectral analysis algorithms to high-spatial-resolution images using multiple spectral bands. The analyzed spectral ranges included the green (510–590 nm), red (620–700 nm), red edge (725–745 nm), and near-infrared (750–830 nm) bands. Nitrogen content was classified into three categories: low (less than 27 g kg⁻¹), medium (27 to 29 g kg⁻¹), and high (more than 29 g kg⁻¹). Each nitrogen class (low, medium, and high) had its spectral signature grouped accordingly.

These thresholds were determined based on research correlating foliar nitrogen content (g kg⁻¹) with productivity increases (%) in field experiments [51]. A nitrogen content of 27 g kg⁻¹ was associated with an approximate 100% increase in productivity, whereas above 29 g kg⁻¹, the rate of increase began to decline. Nitrogen classification was performed using the Spectral Angle Mapper (SAM) algorithm, which used spectral curves as training data for the classifier. This algorithm outperformed others tested, achieving an overall accuracy of 87.5% and a kappa index of 0.75 [48].

4. Machine Learning

Algorithm Models

In this section, we will discuss several machine learning algorithm models, including Artificial Neural Networks (ANNs), the M5P tree, REPTree, Random Forest, the Polynomial Kernel Support Vector Machine, and ZeroR. Artificial Neural Networks (ANNs) are computational models inspired by the structure and functioning of the human brain. They consist of interconnected units called neurons, organized into layers. These networks are widely applied in solving complex problems, including pattern recognition, data classification, and predictions across various fields, such as artificial intelligence, image processing, and data analysis [52,53].

The M5P tree is a more advanced model proposed by [54] to solve both linear and nonlinear regression problems [55]. This model uses unidimensional functions at the final nodes of the tree to perform statistical predictions, applying linear regression models at the terminal nodes [54,56].

REPTree (Reduced Error Pruning Tree) is a machine learning algorithm that constructs decision or regression trees based on entropy calculations, adjusting the model to minimize errors caused by data variance. It uses the reduce-error pruning process to prevent overfitting, enhancing the tree's generalization capability [56].

Random Forest (RF) models are a powerful and innovative machine learning approach capable of mapping nonlinear relationships between input variables and their respective responses [57,58]. Among the most notable features of this technique are its relatively small number of advanced hyperparameters,

automatic estimation of generalization errors, ability to handle missing values, and robustness against overfitting [57,59].

The Polynomial Kernel Support Vector Machine (SVMP) is an extension of SVM designed to solve complex problems by mapping data into higher dimensions using a polynomial kernel [60,61]. The polynomial kernel transforms non-linearly separable data into a higher-dimensional space, allowing SVM to identify more complex decision boundaries and better adapt to nonlinear patterns [60].

ZeroR is a machine learning model that serves as a basic benchmark for comparison with other models. It is a simple classifier that does not use any information about data features, relying solely on the most frequent class to make predictions [62]. ZeroR is a useful tool for initial data analysis, but its simplicity means it should be complemented with more sophisticated models to achieve more accurate and meaningful predictions.

From this perspective, these models have been applied in several agricultural projects seeking to solve complex tasks. There is a wide use of machine learning (ML) in agriculture with spectral data due to a combination of factors involving data complexity, prediction potential, and gains in efficiency and accuracy. For example, [63] proposed that using machine learning models with spectral data offers an efficient and low-cost longterm solution for determining caffeine in coffee beans, successfully achieving significant values using SVM combined with spectral information from ground and roasted beans to identify clones, with correlation coefficient values close to 0.7.

Another relevant application of the use of ML models was their use in the prediction of leaf nitrogen and the height of corn plants using multispectral data. The authors concluded that the use of RF is suitable for predicting both agronomic variables in corn, making it easier for farmers to monitor their plants, improving their production rates [64]. The use of a multispectral UAV sensor with a thermal sensor associated with the Random Forest algorithm proved to be efficient when applied to the RF model to predict corn productivity, providing the producer with important information for end-of-harvest policies with the crop still in the middle of the cycle [65]. The use of machine learning algorithms can accurately estimate the diameter (DBH) and height (H) of trees in mixed forest stands using hyperspectral data, especially when using DT, M5P, and SVM, facilitating forest inventories, for example, making the work more efficient [66].

In addition to predictions, the models have classification tasks, for example, monitoring the severity of Asian rust or target spot in soybeans is a task that demands intensive time and work from qualified professionals to identify the disease and its severity so that there is adequate phytosanitary management. Therefore, the application of spectral data to the SVM model yielded values close to 90% for distinguishing the severities of Asian rust and target spot in soybeans, making it a promising model for identifying the severity of the disease in the field [67,68]. This is useful not only for applications in relation to diseases but also for the identification of pest attacks, such as Red gum lerp psyllid (*Glycaspis brimblecombei* More, 1964) in eucalyptus, a pest that has been a concern in the cultivation of the crop. It is possible to identify the severity of the attack using spectral information applied to the SVM model, assisting with control management [69]. For example, [70] used eight algorithms to detect attacks by *S.*

frugiperda and *D. melacanthus*, with Random Forest (RF). The classification of superior soybean genetic materials for industrial grain quality was also possible thanks to the use of ML data with the application of hyperspectral data [71].

Another example of the application of these models is in the prediction of soil nutrients using soil reflectance. In particular, remarkably high levels of accuracy were achieved by the M5P and RF algorithms in predicting soil nutrients, particularly the S content. The authors state that this technology contributes to a reduction in the work involved in the analysis of samples and reduces the use of reagents in laboratory analyses, representing a promising approach for the prediction of soil macronutrients [72]. The use of machine learning has also proven efficient in classifying different soybean genetic materials in the improvement process, specifically in the classification of their primary macronutrient [73,74], secondary macronutrient [75] and micronutrient [76] contents using multispectral reflectance obtained by UAV. This enables the selection of materials with better performance in nutrient absorption.

The application of machine learning algorithms using data from multispectral and hyperspectral sensors represents a promising strategy for monitoring and managing agricultural systems, especially thanks to their ability to detect subtle variations in the physiological state of plants, nutritional and water conditions, soil composition, and other relevant agronomic variables. However, to enable the effective integration of these technologies into real agricultural scenarios, some considerations should be taken into account. First, standardizing data collection and processing protocols is essential, since the heterogeneity between sensors and methodologies limits the comparability and reproducibility of studies. In addition, the construction of robust databases that contemplate different crops, soil and climate conditions, and temporal scales is essential for training models with greater generalization capacity. Additionally, it is essential that the developed models are validated in real field conditions, considering aspects such as spatial variability, spectral resolution, and scalability. Reducing the costs associated with sensors and computational processing, as well as developing more affordable solutions, especially for small and medium-sized producers, should also be prioritized.

5. Challenges of Digital Agriculture

Digital agriculture transforms data collected by sensors into actionable information tailored to farmers' specific needs. The integration of data from multispectral or hyperspectral sensors, satellite and drone imagery, combined with artificial intelligence, enables the development of systems that mimic human intelligence, training machines to understand and execute specific tasks. Precision agriculture, in turn, utilizes autonomous vehicles to apply fertilizers in precise amounts, reducing waste and optimizing time.

However, digital agriculture still faces several challenges, particularly regarding the accuracy of crop nutritional assessments. One key issue is data variability, which includes the organization of structured and unstructured information. Structured data [77] refers to information organized in specific

formats, such as sensor readings, while unstructured data include satellite and drone images that lack a fixed format. It is important to note that although large volumes of data can be difficult to interpret, they provide valuable and robust insights [78].

One way to address this challenge is by generating more comprehensive, detailed, and precise insights. How can this be achieved? The combination of multiple data sources is essential. For instance, in a crop field, agricultural sensor data indicating nutrient deficiencies, such as nitrogen, can be analyzed alongside soil moisture and temperature information. This integrated approach provides a more accurate assessment of field conditions.

This strategy can optimize crop management and nutrient application by adjusting doses and locations based on satellite imagery. Such integration allows for canopy biomass visualization, the identification of plant water stress, and the analysis of crop coverage variability. Plants with lower pigment levels may indicate lower nitrogen content, whereas those with higher pigment levels exhibit increased chlorophyll concentration, resulting in a more intense green coloration and, consequently, higher nitrogen levels. However, it is important to highlight that other nutrient (Mg, Mn, Fe, etc.) are also involved in chlorophyll synthesis, emphasizing the need for accuracy studies to ensure reliable diagnostics.

Based on these data, a predictive mathematical model can be developed to generate fertilizer application maps, optimizing distribution according to the specific needs of different field areas. This approach not only enhances productivity but also reduces fertilizer waste and minimizes groundwater contamination.

6. From Leaf to Harvest: How Interactions Between Spectral Readings, Leaf Nutrient Content, and Yield Affect the Robustness of Predictive Models

Validating plant nutritional diagnostics generated by digital tools, such as spectral analysis, based on actual plant responses in the field is essential to ensuring the advancement of these new techniques on solid foundations. Measuring the accuracy of nutritional diagnostics produced by these technologies through simple field tests is crucial. For instance, if a new method indicates a nutrient deficiency in a specific field, then, theoretically, applying that nutrient should result in a yield increase—provided the diagnosis is correct. To validate this, the nutrient is applied to one part of the field while another portion remains untreated, and the yield of both areas is compared. If the treated area exhibits a yield increase compared to the untreated area, it confirms that the deficiency diagnosis was accurate. Conversely, if no yield improvement is observed, it indicates that the field did not actually suffer from a nutrient deficiency. Field tests to assess the accuracy of new crop nutritional evaluation methods, which rely on leaf tissue analysis to determine nutrient concentrations, have been conducted to verify precision and compare different interpretation methods, such as sufficiency range (conventional method), DRIS/CND, and others. The accuracy of nutritional diagnostics using the conventional method based on leaf tissue analysis has been

evaluated in crops such as soybean [79–81], banana [82], eucalyptus seedlings [83], and sugarcane [84,85].

However, using crop yield as an indicator of response to validate the accuracy of nutritional status assessments remains limited, even for the conventional methods based on leaf tissue analysis. This is a significant concern. With advancements in digital agriculture, new methods for diagnosing plant nutritional status using spectral techniques have been proposed. Unfortunately, these approaches often fail to incorporate yield data to validate accuracy, which needs to be more widely recognized by researchers in this field.

Accuracy testing through yield analysis is crucial, as it confirms whether an identified deficiency is real by demonstrating that nutrient supplementation results in yield gains. In the study by [44], for example, while the researchers successfully discriminated adequate boron levels in eucalyptus using specific near-infrared bands, the digital tool lacked validation to verify the accuracy of diagnoses in commercial fields using yield data, as previously discussed. Similarly, the authors of [12] estimated the chlorophyll content in various species without characterizing deficient areas or assessing the accuracy of these diagnoses in their study.

Some studies emphasize the correlation between spectral readings and foliar nutrient concentrations as an initial step; however, this does not indicate the accuracy of the generated diagnostics. While nutrient concentration data serve as important reference standards, without accuracy validation tests, they cannot be assumed to be highly reliable. If the conventional method using nutrient concentrations exhibits low diagnostic accuracy, and a new spectral reading method has a high correlation with this conventional approach, it does not necessarily mean the spectral method is highly accurate. This only suggests that the technique can detect the nutrient in plant tissues, which is valuable, but it does not confirm its predictive ability for generating reliable diagnoses.

The authors [47,49,50] studied nutrient application and reported spectral correlations with nutrient concentrations. However, their analysis did not establish the correlation between spectral readings and foliar nutrient content. Even in the study by [46], which presented yield data, there was no conclusive evidence indicating nitrogen deficiency in the studied areas. This highlights the need for assessing “accurate” foliar nutrient levels to confirm such deficiencies.

A critical methodology to address this gap would be the targeted application of nutrients in deficient areas, followed by monitoring the yield response. This approach would allow digital tools to detect nutrient deficiencies that directly impact crop productivity, as well as potential nutrient excesses. Incorporating this methodology would enhance the reliability of spectral analysis-based nutritional diagnostics, making them more applicable to precision agriculture.

A literature search in the Web of Science database using the keywords “deficiency, sensor, fertilization, chlorophyll” retrieved ten scientific articles [84,86,87,88,89,90,91,92]. A similar search in the Scopus database yielded a larger number of relevant studies [93,94,95,96,97,98,99,100,101,102,103,104,105,106,107,108,109,110,111,112,113,114,1

15,116,117,118,119,120,121,122,123,124,125,126]. These findings reinforce that research in this area is advancing, yet the number of high-quality studies remains limited.

In general, studies using spectral reflectance in digital agriculture to generate nutritional diagnostics have demonstrated potential for estimating variations in foliar nutrient concentrations, particularly in controlled-condition experiments. However, the applicability of these findings to field conditions remains debatable. Many researchers in this field mistakenly assume that a strong spectral discrimination associated with increased nutrient levels in the soil or plant is sufficient to ensure that the new tool can generate accurate nutritional diagnostics at the field level. This assumption is problematic because field grown crops are influenced by numerous environmental and physiological factors that affect foliar nutrient concentrations and, consequently, the accuracy of nutritional diagnostics.

Another critical issue is that most studies focus primarily on nitrogen (N), highlighting the need for research on other essential nutrients. Nitrogen deficiency, for example, may be linked to the insufficient availability of sulfur (S), iron (Fe), or molybdenum (Mo), which play crucial roles in nitrogen metabolism [127].

Additionally, a major limitation of these studies is the lack of accuracy validation using plant response data, specifically yield, to assess the precision of the generated diagnostics. This issue is observed even in conventional foliar analysis methods. The absence of such validation raises concerns about the practical applicability of digital tools in realworld agricultural settings. Without proper field-level validation, these tools may not provide reliable nutritional status assessments for guiding precise fertilization decisions and promoting sustainable agriculture.

To ensure the real-world applicability of spectral reflectance-based diagnostics, complementary studies must be conducted in commercial fields across different regions. This is essential to verify the performance of indirect plant nutritional assessment methods and confirm their reliability in providing accurate diagnostics, ultimately strengthening digital and precision agriculture. Further studies in varied field conditions, using yield data for validation, are considered crucial to confirm the accuracy and reliability of these methods.

7. Final Considerations

Digital agriculture is emerging as a transformative force in crop management, enabling more efficient and sustainable practices. By integrating data from sensors, satellite imagery, and drones with artificial intelligence, precision agriculture becomes more accessible and effective, allowing informed decision-making based on detailed analyses of field conditions. Despite challenges related to data variability and organization, combining multiple information sources provides valuable insights that can optimize crop management and input application.

The recent digital approach, integrating remote sensing technologies, such

as spectral sensors on UAVs and satellites, and artificial intelligence, specifically machine learning algorithms like Random Forest, shows potential for improving the precision of predicting leaf nutrient concentrations.

This approach not only enhances productivity but also promotes sustainability by reducing waste and mitigating environmental impacts, such as groundwater contamination. Thus, digital agriculture not only modernizes farming practices but also paves the way for a more responsible and conscious future in food production, emphasizing the importance of a data-driven and integrated approach tailored to the specific needs of each crop.

It is crucial not only to correlate nutrient concentrations with spectral variables but also to investigate how these diagnostics perform in crops under normal, deficient, and excessive nutrient conditions. By gaining a deeper understanding of how these variables impact crop productivity, it will be possible to develop more effective tools with high practical utility in the field—supporting farmers in optimizing plant nutrition and improving yields sustainably.

Author Contributions: B.C.d.S. and R.d.M.P.: Conceptualization, Data Curation, Investigation, Project Administration, Resources, Writing—Original Draft, Visualization, Supervision, and Funding Acquisition, Writing—Review and Editing. F.H.R.B.: Visualization, Supervision, Formal Analysis, Conceptualization, Writing—Original Draft, Writing—Review and Editing. C.N.S.C., L.P.R.T., P.E.T. and D.C.S.: Writing—Review and Editing, Writing—Original Draft Preparation, Visualization, Conceptualization. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Council for Scientific and Technological Development (CNPq)—Grant number 140525/2021-1.

Data Availability Statement: The data used in this study were obtained from scientific articles and other published sources, which are properly cited throughout the manuscript. No new dataset was generated or analyzed for this study.

Acknowledgments: The authors acknowledge the support from UNESP (Universidade Estadual Paulista) on behalf of the Postgraduate Agronomy Program (Crop Production), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), UNESP's Plant Nutrition Study Group (GENPLANT), UFMS (Universidade Federal de Mato Grosso do Sul)—UFMS, and the Cerrado Plant Nutrition Study Group (GECENP).

Conflicts of Interest: The authors declare no competing financial interests or known personal relationships that could have influenced the work reported in this article.

8. References

1. Sharma, A.; Prakash, A.; Bhambota, S.; Kumar, S. Investigations of Precision Agriculture

- Technologies with Application to Developing Countries. *Environ. Dev. Sustain.* **2024**, *26*, 131. <https://doi.org/10.1007/s10668-024-04572-y>.
2. Gebbers, R.; Adamchuk, V.I. Precision Agriculture and Food Security. *Science* **2010**, *327*, 828–831. <https://doi.org/10.1126/sci-ence.1183899>.
 3. Himesh, S.; Rao, E.V.S.P.; Gouda, K.C.; Ramesh, K.V.; Rakesh, V.; Mohapatra, G.N.; Rao, B.K.; Sahoo, S.K.; Ajilesh, P. Digital Revolution and Big Data: A New Revolution in Agriculture. *CABI Rev.* **2018**, 1–7. <https://doi.org/10.1079/PAVSNNR201813021>.
 4. Zhang; Wang, M.; Wang, N. Precision Agriculture—A Worldwide Overview. *Comput. Electron. Agric.* **2002**, *36*, 113–132. [https://doi.org/10.1016/S0168-1699\(02\)00096-0](https://doi.org/10.1016/S0168-1699(02)00096-0).
 5. Teodoro, L.P.R.; Estevão, R.; Santana, D.C.; de Oliveira, I.C.; Lopes, M.T.G.; de Azevedo, G.B.; Rojo Baio, F.H.; da Silva Junior, C.A.; Teodoro, P.E. Eucalyptus Species Discrimination Using Hyperspectral Sensor Data and Machine Learning. *Forests* **2023**, *15*, 39. <https://doi.org/10.3390/f15010039>.
 6. Bhakthan, S.M.; Loganathan, A. A Hyperspectral Unmixing Model Using Convolutional Vision Transformer. *Earth Sci. Inform.* **2024**, *17*, 2255–2273. <https://doi.org/10.1007/s12145-024-01273-y>.
 7. Comba, L.; Biglia, A.; Ricauda Aimonino, D.; Tortia, C.; Mania, E.; Guidoni, S.; Gay, P. Leaf Area Index Evaluation in Vineyards Using 3D Point Clouds from UAV Imagery. *Precis. Agric.* **2020**, *21*, 881–896. <https://doi.org/10.1007/s11119-019-09699-x>.
 8. Abiri, R.; Rizan, N.; Balasundram, S.K.; Shahbazi, A.B.; Abdul-Hamid, H. Application of Digital Technologies for Ensuring Agricultural Productivity. *Heliyon* **2023**, *9*, e22601. <https://doi.org/10.1016/j.heliyon.2023.e22601>.
 9. Teodoro; Rossi, F.S.; Teodoro, L.P.R.; Santana, D.C.; Ratke, R.F.; de Oliveira, I.C.; Della Silva, J.L.; de Oliveira, J.L.G.; da Silva, N.P.; Baio, F.H.R.; Torres, F.E.; et al. Soil CO₂ Emissions under Different Land-Use Managements in Mato Grosso Do Sul, Brazil. *J. Clean. Prod.* **2024**, *434*, 139983. <https://doi.org/10.1016/j.jclepro.2023.139983>.
 10. Xie, K.; Cakmak, I.; Wang, S.; Zhang, F.; Guo, S. Synergistic and Antagonistic Interactions between Potassium and Magnesium in Higher Plants. *Crop J.* **2021**, *9*, 249–256. <https://doi.org/10.1016/j.cj.2020.10.005>.
 11. da Silva, B.C.; Prado, R.d.M.; Baio, F.H.R.; Campos, C.N.S.; Teodoro, L.P.R.; Teodoro, P.E.; Santana, D.C.; Fernandes, T.F.S.; da Silva Junior, C.A.; Loureiro, E.d.S. New Approach for Predicting Nitrogen and Pigments in Maize from Hyperspectral Data and Machine Learning Models. *Remote Sens. Appl.* **2024**, *33*, 101110. <https://doi.org/10.1016/j.rsase.2023.101110>.
 12. Gitelson, A.A.; Grid, Y.; Merzlyak, M.N. Relationships between Leaf Chlorophyll Content and Spectral Reflectance and Algorithms for Non-Destructive Chlorophyll Assessment in Higher Plant Leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. <https://doi.org/10.1078/0176-1617-00887>.

13. Peng, Y.; Zhang, M.; Xu, Z.; Yang, T.; Su, Y.; Zhou, T.; Wang, H.; Wang, Y.; Lin, Y. Estimation of Leaf Nutrition Status in De-graded Vegetation Based on Field Survey and Hyperspectral Data. *Sci. Rep.* **2020**, *10*, 4361. <https://doi.org/10.1038/s41598-020-61294-7>.
14. Song, G.; Wang, Q. Developing Hyperspectral Indices for Assessing Seasonal Variations in the Ratio of Chlorophyll to Carotenoid in Deciduous Forests. *Remote Sens.* **2022**, *14*, 1324. <https://doi.org/10.3390/rs14061324>.
15. Anastasiou, E.; Fountas, S.; Voulgaraki, M.; Psiroukis, V.; Koutsiaras, M.; Kriezi, O.; Lazarou, E.; Vatsanidou, A.; Fu, L.; Bartolo, F.D.; et al. Precision Farming Technologies for Crop Protection: A Meta-Analysis. *Smart Agric. Technol.* **2023**, *5*, 100323. <https://doi.org/10.1016/j.atech.2023.100323>.
16. Wei, H.; Xu, W.; Kang, B.; Eisner, R.; Muleke, A.; Rodriguez, D.; deVoil, P.; Sadras, V.; Monjardino, M.; Harrison, M.T. Irrigation with Artificial Intelligence: Problems, Premises, Promises. *Hum.-Centric Intell. Syst.* **2024**, *4*, 187–205. <https://doi.org/10.1007/s44230-024-00072-4>.
17. Ahmed, Z.; Wan, S.; Zhang, F.; Zhong, W. Artificial Intelligence for Omics Data Analysis. *BMC Methods* **2024**, *1*, 4. <https://doi.org/10.1186/s44330-024-00004-5>.
18. Holl, C. The Content Intelligence: An Argument against the Lethality of Artificial Intelligence. *Discov. Artif. Intell.* **2024**, *4*, 13. <https://doi.org/10.1007/s44163-024-00112-9>.
19. Gignac, G.E.; Szodorai, E.T. Defining Intelligence: Bridging the Gap between Human and Artificial Perspectives. *Intelligence* **2024**, *104*, 101832. <https://doi.org/10.1016/j.intell.2024.101832>.
20. Prasad, G.A.; Kumar, A.V.S.; Sharma, P.; Irawati, I.D.; Chandrashekar, D.V.; Musirin, I.B.; Abdullah, H.M.A.; Rao, L.M. *Artificial Intelligence in Computer Science*; IGI Global: Hershey, PA, USA, 2023; pp. 43–60.
21. Gottfredson, L.S. Mainstream Science on Intelligence. *Wall Street Journal*, 1997.
22. Fabbri, A.G.; Fung, K.B.; Yatabe, S.M. The Usage of Artificial Intelligence in Remote Sensing: A Review of Applications and Current Research. In *Quantitative Analysis of Mineral and Energy Resources*; Springer: Dordrecht, The Netherlands, 1988; pp. 489–512.
23. Akosah, S.; Gratchev, I.; Kim, D.-H.; Ohn, S.-Y. Application of Artificial Intelligence and Remote Sensing for Landslide Detection and Prediction: Systematic Review. *Remote Sens.* **2024**, *16*, 2947. <https://doi.org/10.3390/rs16162947>.
24. Zhang, Q. Opinion Paper: Precision Agriculture, Smart Agriculture, or Digital Agriculture. *Comput. Electron. Agric.* **2023**, *211*, 107982. <https://doi.org/10.1016/j.compag.2023.107982>.
25. Marconi, S.; Weinstein, B.G.; Zou, S.; Bohlman, S.A.; Zare, A.; Singh, A.; Stewart, D.; Harmon, I.; Steinkraus, A.; White, E.P. Continental-Scale Hyperspectral Tree Species Classification in the United States National Ecological Observatory Network. *Remote Sens. Environ.* **2022**, *282*,

113264. <https://doi.org/10.1016/j.rse.2022.113264>.
26. Katharria, A.; Rajwar, K.; Pant, M.; Velásquez, J.D.; Snášel, V.; Deep, K. Information Fusion in Smart Agriculture: Machine Learning Applications and Future Research Directions. *SSRN* **2025**. <https://doi.org/10.2139/ssrn.5187083>.
 27. Padhiary, M.; Saha, D.; Kumar, R.; Sethi, L.N.; Kumar, A. Enhancing precision agriculture: A comprehensive review of machine learning and AI vision applications in all-terrain vehicle for farm automation. *Smart Agric. Technol.* **2024**, *8*, 100483. <https://doi.org/10.1016/j.atech.2024.100483>.
 28. Addison, M.; Bonuedi, I.; Arhin, A.A.; Wadei, B.; Owusu-Addo, E.; Fredua Antoh, E.; Mensah-Odum, N. Exploring the Impact of Agricultural Digitalization on Smallholder Farmers' Livelihoods in Ghana. *Heliyon* **2024**, *10*, e27541. <https://doi.org/10.1016/j.heliyon.2024.e27541>.
 29. Fossen, F.; Sorgner, A. Mapping the Future of Occupations: Transformative and Destructive Effects of New Digital Technologies on Jobs. *Foresight STI Gov.* **2019**, *13*, 10–18. <https://doi.org/10.17323/2500-2597.2019.2.10.18>.
 30. Klerkx, L.; Jakku, E.; Labarthe, P. A Review of Social Science on Digital Agriculture, Smart Farming and Agriculture 4.0: New Contributions and a Future Research Agenda. *NJAS Wagening J. Life Sci.* **2019**, *90*, 1–16. <https://doi.org/10.1016/j.njas.2019.100315>.
 31. ONU. *Objetivos de Desenvolvimento Sustentável: As Nações Unidas No Brasil*; Organização das Nações Unidas: San Francisco, CA, USA, 2021.
 32. MacPherson, J.; Voglhuber-Slavinsky, A.; Olbrisch, M.; Schöbel, P.; Dönid, E.; Mouratiadou, I.; Helming, K. Future Agricultural Systems and the Role of Digitalization for Achieving Sustainability Goals. A Review. *Agron. Sustain. Dev.* **2022**, *42*, 70. <https://doi.org/10.1007/s13593-022-00792-6>.
 33. Balkrishna, A.; Pathak, R.; Kumar, S.; Arya, V.; Singh, S.K. A Comprehensive Analysis of the Advances in Indian Digital Agri- cultural Architecture. *Smart Agric. Technol.* **2023**, *5*, 100318. <https://doi.org/10.1016/j.atech.2023.100318>.
 34. Chandra, R.; Collis, S. Digital Agriculture for Small-Scale Producers. *Commun. ACM* **2021**, *64*, 75–84. <https://doi.org/10.1145/3454008>.
 35. Abioye, E.A.; Hensel, O.; Esau, T.J.; Elijah, O.; Abidin, M.S.Z.; Ayobami, A.S.; Yerima, O.; Nasirahmadi, A. Precision Irrigation Management Using Machine Learning and Digital Farming Solutions. *AgriEngineering* **2022**, *4*, 70–103. <https://doi.org/10.3390/agriengineering4010006>.
 36. Ogubuike, R.; Adib, A.; Orji, R. Masa: AI-Adaptive Mobile App for Sustainable Agriculture. In Proceedings of the 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 27–30 October 2021; IEEE:

Piscataway, NJ, USA, 2021; pp. 1064–1069.

37. Basso, B.; Antle, J. Digital Agriculture to Design Sustainable Agricultural Systems. *Nat. Sustain.* **2020**, *3*, 254–256. <https://doi.org/10.1038/s41893-020-0510-0>.
38. Abbasi, R.; Martinez, P.; Ahmad, R. The Digitization of Agricultural Industry—A Systematic Literature Review on Agriculture
- 4.0. *Smart Agric. Technol.* **2022**, *2*, 100042. <https://doi.org/10.1016/j.atech.2022.100042>.
39. O'Donoghue, T.; Minasny, B.; McBratney, A. Digital Regenerative Agriculture. *npj Sustain. Agric.* **2024**, *2*, 5. <https://doi.org/10.1038/s44264-024-00012-6>.
40. McBratney, A.; Whelan, B.; Ancev, T.; Bouma, J. Future Directions of Precision Agriculture. *Precis. Agric.* **2005**, *6*, 7–23. <https://doi.org/10.1007/s11119-005-0681-8>.
41. Porciello, J.; Coggins, S.; Mabaya, E.; Otunba-Payne, G. Digital Agriculture Services in Low- and Middle-Income Countries: A Systematic Scoping Review. *Glob. Food Secur.* **2022**, *34*, 100640. <https://doi.org/10.1016/j.gfs.2022.100640>.
42. Mentsiev, A.U.; Amirova, E.F.; Afanasev, N. V Digitalization and Mechanization in Agriculture Industry. *IOP Conf. Ser. Earth Environ. Sci.* **2020**, *548*, 032031. <https://doi.org/10.1088/1755-1315/548/3/032031>.
43. Santana, D.C.; Teixeira Filho, M.C.M.; da Silva, M.R.; das Chagas, P.H.M.; de Oliveira, J.L.G.; Baio, F.H.R.; Campos, C.N.S.; Teodoro, L.P.R.; da Silva Junior, C.A.; Teodoro, P.E.; et al. Machine Learning in the Classification of Soybean Genotypes for Primary Macronutrients' Content Using UAV–Multispectral Sensor. *Remote Sens.* **2023**, *15*, 1457. <https://doi.org/10.3390/rs15051457>.
44. Damasceno, A.S.d.S.; Boechat, C.L.; de Souza, H.A.; Capristo-Silva, G.F.; Mendes, W.D.S.; Teodoro, P.E.; Morais, P.G.C.; de Oliveira, R.I.; Della-Silva, J.L.; de Souza, I.A.M.; et al. Nutritional Monitoring of Boron in Eucalyptus Spp. in the Brazilian Cerrado by Multispectral Bands of the MSI Sensor (Sentinel-2). *Remote Sens. Appl.* **2023**, *29*, 100913. <https://doi.org/10.1016/j.rsase.2022.100913>.
45. Costa, L.; Kunwar, S.; Ampadidis, Y.; Albrecht, U. Determining Leaf Nutrient Concentrations in Citrus Trees Using UAV Imagery and Machine Learning. *Precis. Agric.* **2022**, *23*, 854–875. <https://doi.org/10.1007/s11119-021-09864-1>.
46. Santana, D.C.; Cotrim, M.F.; Flores, M.S.; Rojo Baio, F.H.; Shiratsuchi, L.S.; Antonio da Silva Junior, C.; Ribeiro Teodoro, L.P.; Teodoro, P.E. UAV-BASED Multispectral Sensor to Measure Variations in Corn as a Function of Nitrogen Topdressing. *Remote Sens. Appl.* **2021**, *23*, 100534. <https://doi.org/10.1016/j.rsase.2021.100534>.
47. Malmir, M.; Tahmasbian, I.; Xu, Z.; Farrar, M.B.; Bai, S.H. Prediction of Macronutrients in Plant

- Leaves Using Chemometric Analysis and Wavelength Selection. *J. Soils Sediments* **2020**, *20*, 249–259. <https://doi.org/10.1007/s11368-019-02418-z>.
48. Osco, L.P.; Marques Ramos, A.P.; Saito Moriya, É.A.; de Souza, M.; Marcato Junior, J.; Matsubara, E.T.; Imai, N.N.; Creste, J.E. Improvement of Leaf Nitrogen Content Inference in Valencia-Orange Trees Applying Spectral Analysis Algorithms in UAV Mounted-Sensor Images. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *83*, 101907. <https://doi.org/10.1016/j.jag.2019.101907>.
 49. Osco, L.P.; Marques Ramos, A.P.; Roberto Pereira, D.; Akemi Saito Moriya, É.; Nobuhiro Imai, N.; Takashi Matsubara, E.; Estrabis, N.; de Souza, M.; Marcato Junior, J.; Gonçalves, W.N.; et al. Predicting Canopy Nitrogen Content in Citrus-Trees Using Random Forest Algorithm Associated to Spectral Vegetation Indices from UAV-Imagery. *Remote Sens.* **2019**, *11*, 2925. <https://doi.org/10.3390/rs11242925>.
 50. Osco, L.P.; Arruda, M.d.S.; Junior, J.M.; Silva, N.B.d.; Ramos, A.P.M.; Moryia, É.A.S.; Imai, N.N.; Pereira, D.R.; Creste, J.E.; Matsubara, E.T.; Li, J.; et al. A Convolutional Neural Network Approach for Counting and Geolocating Citrus-Trees in UAV Multispectral Imagery. *ISPRS J. Photogramm. Remote Sens.* **2020**, *160*, 97–106. <https://doi.org/10.1016/j.isprsjprs.2019.12.010>.
 51. Quaggio, J.A.; Cantarella, H.; van Raij, B. Phosphorus and Potassium Soil Test and Nitrogen Leaf Analysis as a Base for Citrus Fertilization. *Nutr. Cycl. Agroecosyst.* **1998**, *52*, 67–74. <https://doi.org/10.1023/A:1009763027607>.
 52. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
 53. Haykin, S.S. *Neural Networks and Learning Machines*; Prentice Hall/Pearson: London, UK, 2009; ISBN 9780131471399.
 54. Quinlan, J.R. Learning with Continuous Classes. In Proceedings of the 5th Australian Joint Conference on Artificial Intelligence, Hobart, Australia, 16–18 November 1992; Volume 92, pp. 343–348. [https://doi.org/10.1016/S0020-7373\(87\)80053-6](https://doi.org/10.1016/S0020-7373(87)80053-6).
 55. Pande, C.B.; Al-Ansari, N.; Kushwaha, N.L.; Srivastava, A.; Noor, R.; Kumar, M.; Moharir, K.N.; Elbeltagi, A. Forecasting of SPI and Meteorological Drought Based on the Artificial Neural Network and M5P Model Tree. *Land* **2022**, *11*, 2040. <https://doi.org/10.3390/land11112040>.
 56. Sihag, P.; Mehta, T.; Sammen, S.S.; Pande, C.B.; Puri, D.; Radwan, N. Predictive Modelling of Nitrogen Dioxide Using Soft Computing Techniques in the Agra, Uttar Pradesh, India. *Phys. Chem. Earth Parts A/B/C* **2024**, *134*, 103589. <https://doi.org/10.1016/j.pce.2024.103589>.
 57. Abellán-García, J.; Guzmán-Guzmán, J.S. Random Forest-Based Optimization of UHPFRC under Ductility Requirements for Seismic Retrofitting Applications. *Constr. Build. Mater.* **2021**, *285*, 122869. <https://doi.org/10.1016/j.conbuildmat.2021.122869>.
 58. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32.

59. Aurret, L.; Aldrich, C. Interpretation of Nonlinear Relationships between Process Variables by Use of Random Forests. *Miner. Eng.* **2012**, *35*, 27–42. <https://doi.org/10.1016/j.mineng.2012.05.008>.
60. Guido, R.; Ferrisi, S.; Lofaro, D.; Conforti, D. An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review. *Information* **2024**, *15*, 235. <https://doi.org/10.3390/info15040235>.
61. Pathak, D.K.; Kalita, S.K.; Bhattacharya, D.K. Hyperspectral Image Classification Using Support Vector Machine: A Spectral Spatial Feature Based Approach. *Evol. Intell.* **2022**, *15*, 1809–1823. <https://doi.org/10.1007/s12065-021-00591-0>.
62. Witten, I.H.; Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 2005.
63. Santana, D.C.; Ratke, R.F.; Zanatta, F.L.; Campos, C.N.S.; Seron, A.C.d.S.C.; Teodoro, L.P.R.; Silva, N.P.d.; Oliveira, G.S.; dos Santos, R.G.; Alvarez, R.d.C.F. Caffeine Content Prediction in Coffee Beans Using Hyperspectral Reflectance and Machine Learning. *AgriEngineering* **2024**, *6*, 4480–4492.
64. Osco, L.P.; Junior, J.M.; Ramos, A.P.; Furuya, D.E.; Santana, D.C.; Teodoro, L.P.; Gonçalves, W.N.; Baio, F.H.; Pistori, H.; Junior, C.A.; et al. Leaf Nitrogen Concentration and Plant Height Prediction for Maize Using UAV-Based Multispectral Imagery and Machine Learning Techniques. *Remote Sens.* **2020**, *12*, 3237.
65. Kumar, C.; Dhillon, J.; Huang, Y.; Reddy, K. Explainable machine learning models for corn yield prediction using UAV multi-spectral data. *Comput. Electron. Agric.* **2025**, *231*, 109990. <https://doi.org/10.1016/j.compag.2025.109990>.
66. Manfroio Filho, E.A.; Teodoro, P.E.; Teodoro, L.P.R.; Santana, D.C.; Azevedo, G.T.d.O.S.; da Silva Junior, C.A.; Baio, F.H.R.; de Azevedo, G.B. VNIR-SWIR Spectroscopy and Machine Learning for Measuring Dendrometric Variables in Native Species. *Remote Sens. Appl.* **2025**, *37*, 101522.
67. Santana, D.C.; de Queiroz Otone, J.D.; Baio, F.H.R.; Teodoro, L.P.R.; Alves, M.E.M.; da Silva Junior, C.A.; Teodoro, P.E. Machine Learning in the Classification of Asian Rust Severity in Soybean Using Hyperspectral Sensor. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2024**, *313*, 124113.
68. de Queiroz Otone, J.D.; Theodoro, G.d.F.; Santana, D.C.; Teodoro, L.P.R.; de Oliveira, J.T.; de Oliveira, I.C.; da Silva Junior, C.A.; Teodoro, P.E.; Baio, F.H.R. Hyperspectral Response of the Soybean Crop as a Function of Target Spot (*Corynespora cassiicola*) Using Machine Learning to Classify Severity Levels. *AgriEngineering* **2024**, *6*, 330–343.
69. de Gregori, G.S.; de Souza Loureiro, E.; Amorim Pessoa, L.G.; Azevedo, G.B.d.; Azevedo, G.T.d.O.S.; Santana, D.C.; de Oliveira, I.C.; de Oliveira, J.L.G.; Teodoro, L.P.R.; Baio, F.H.R.

Machine Learning in the Hyperspectral Classification of *Glycaspis Brim-blecombei* (Hemiptera psyllidae) Attack Severity in Eucalyptus. *Remote Sens.* **2023**, *15*, 5657.

70. Furuya, D.E.G.; Ma, L.; Pinheiro, M.M.F.; Gomes, F.D.G.; Gonçalves, W.N.; Junior, J.M.; de Castro Rodrigues, D.; Blassioli-Moraes, M.C.; Michereff, M.F.F.; Borges, M. Prediction of Insect-Herbivory-Damage and Insect-Type Attack in Maize Plants Using Hyperspectral Data. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102608.
71. Bernardy, R.; Cañizares, L.d.C.C.; Meza, S.L.R.; Rodrigues, L.A.; Jappe, S.N.; Oliveira, M.d. Identifying soybean genotypes with artificial intelligence and near infrared reflectance. *Cienc. e Agrotecnologia* **2024**, *48*, e005224. <https://doi.org/10.1590/1413-7054202448005224>.
72. Nozari, S.; Pahlavan-Rad, M.R.; Brungard, C.; Heung, B.; Borůvka, L. Digital soil mapping using machine learning-based methods to predict soil organic carbon in two different districts in the Czech Republic. *Soil Water Res.* **2024**, *19*, 32–49. <https://doi.org/10.17221/119/2023-SWR>.
73. Mukherjee, S.; Laskar, S. Vis-NIR-Based Optical Sensor System for Estimation of Primary Nutrients in Soil. *J. Opt.* **2019**, *48*, 87–103. <https://doi.org/10.1007/s12596-019-00517-1>.
74. Zhou, J.; Beche, E.; Vieira, C.C.; Yungbluth, D.; Zhou, J.; Scaboo, A.; Chen, P. Improve Soybean Variety Selection Accuracy Using UAV-Based High-Throughput Phenotyping Technology. *Front. Plant Sci.* **2022**, *12*, 768742. <https://doi.org/10.3389/fpls.2021.768742>.
75. Santana, D.C.; de Oliveira, I.C.; Cavaleiro, S.B.; das Chagas, P.H.M.; Teixeira Filho, M.C.M.; Della-Silva, J.L.; Teodoro, L.P.R.; Campos, C.N.S.; Baio, F.H.R.; da Silva Junior, C.A.; et al. Classification of Soybean Genotypes as to Calcium, Magnesium, and Sulfur Content Using Machine Learning Models and UAV-Multispectral Sensor. *AgriEngineering* **2024**, *6*, 1581–1593. <https://doi.org/10.3390/agriengineering6020090>.
76. Keshavarzi, A.; Kaya, F.; Başayığit, L.; Gyasi-Agyei, Y.; Rodrigo-Comino, J.; Caballero-Calvo, A. Spatial prediction of soil micronutrients using machine learning algorithms integrated with multiple digital covariates. *Nutr. Cycl. Agroecosyst.* **2023**, *127*, 137–153. <https://doi.org/10.1007/s10705-023-10303-y>.
77. Lowenberg-DeBoer, J.; Huang, I.Y.; Grigoriadis, V.; Blackmore, S. Economics of Robots and Automation in Field Crop Production. *Precis. Agric.* **2020**, *21*, 278–299. <https://doi.org/10.1007/s11119-019-09667-5>.
78. Zhang, Q.; Yang, L.T.; Chen, Z.; Li, P. A Survey on Deep Learning for Big Data. *Inf. Fusion* **2018**, *42*, 146–157. <https://doi.org/10.1016/j.inffus.2017.10.006>.
79. Beverly, R.B. DRIS Diagnoses of Soybean Nitrogen, Phosphorus, and Potassium Status Are Unsatisfactory. *J. Plant Nutr.* **1993**, *16*, 1431–1447. <https://doi.org/10.1080/01904169309364625>.
80. Traspadini, E.I.F.; Wadt, P.G.S.; de Mello-Prado, R.; Garcia-Roque, C.; Roberto-Wassolowski,

- C.; Vidal-Perez, D. Efficiency of Critical Level and Compositional Nutrient Diagnosis Methods to Evaluate Boron Nutritional Status in Soybean. *Chil. J. Agric. Res.* **2022**, *82*, 309–319. <https://doi.org/10.4067/S0718-58392022000200309>.
81. Traspadini, E.I.F.; Wadt, P.G.S.; de Mello Prado, R.; Roque, C.G.; Wassolowski, C.R. Prescient Diagnostic Analysis for Boron Nutritional Status in Soy Crops. *Sci. Rep.* **2023**, *13*, 2281. <https://doi.org/10.1038/s41598-022-26263-2>.
 82. Villaseñor-Ortiz, D.; de Mello Prado, R.; Pereira da Silva, G.; Lata-Tenesaca, L.F. Applicability of DRIS in Bananas Based on the Accuracy of Nutritional Diagnoses for Nitrogen and Potassium. *Sci. Rep.* **2022**, *12*, 18125. <https://doi.org/10.1038/s41598-022-22554-w>.
 83. de Moraes, T.C.B.; Prado, R.d.M.; Traspadini, E.I.F.; Wadt, P.G.S.; de Paula, R.C.; Rocha, A.M.S. Efficiency of the CL, DRIS and CND Methods in Assessing the Nutritional Status of *Eucalyptus* Spp. Rooted Cuttings. *Forests* **2019**, *10*, 786. <https://doi.org/10.3390/f10090786>.
 84. Pereira da Silva, G.; Prado, R.d.M.; Wadt, P.G.S.; da Silva, F.C.; Vaz, G.J.; Moda, L.R.; Caione, G. Modeling Formulas of the Comprehensive Diagnosis and Recommendation System (DRIS) for Phosphorus in Sugarcane. *J. Plant Nutr.* **2021**, *44*, 1316–1329. <https://doi.org/10.1080/01904167.2020.1862192>.
 85. da Silva, G.P.; Wadt, P.G.S.; Prado, R.d.M.; Caione, G.; Moda, L.R. Accuracy of Plant Response Potential to Fertilization in Nutritional Diagnoses for Phosphorus in Sugarcane. *J. Plant Nutr.* **2022**, *45*, 1702–1711. <https://doi.org/10.1080/01904167.2021.2014869>.
 86. Barker, D.W.; Sawyer, J.E. Using Active Canopy Sensing to Adjust Nitrogen Application Rate in Corn. *Agron. J.* **2012**, *104*, 926–933. <https://doi.org/10.2134/agronj2012.0030>.
 87. Corp, L.A.; McMurtrey, J.E.; Middleton, E.M.; Mulchi, C.L.; Chappelle, E.W.; Daughtry, C.S.T. Fluorescence Sensing Systems: In Vivo Detection of Biophysical Variations in Field Corn Due to Nitrogen Supply. *Remote Sens. Environ.* **2003**, *86*, 470–479. [https://doi.org/10.1016/S0034-4257\(03\)00125-1](https://doi.org/10.1016/S0034-4257(03)00125-1).
 88. Motomiya, A.V.d.A.; Molin, J.P.; Motomiya, W.R.; Biscaro, G.A. Diagnose Nutricional Com o Uso de Sensor Óptico Ativo Em Algodoeiro. *Rev. Bras. de Eng. Agrícola e Ambient.* **2012**, *16*, 1159–1165. <https://doi.org/10.1590/S1415-43662012001100003>.
 89. Padilla, F.M.; Farneselli, M.; Gianquinto, G.; Tei, F.; Thompson, R.B. Monitoring Nitrogen Status of Vegetable Crops and Soils for Optimal Nitrogen Management. *Agric. Water Manag.* **2020**, *241*, 106356. <https://doi.org/10.1016/j.agwat.2020.106356>.
 90. Pancorbo, J.L.; Camino, C.; Alonso-Ayuso, M.; Raya-Sereno, M.D.; Gonzalez-Fernandez, I.; Gabriel, J.L.; Zarco-Tejada, P.J.; Quemada, M. Simultaneous Assessment of Nitrogen and Water Status in Winter Wheat Using Hyperspectral and Thermal Sensors. *Eur. J. Agron.* **2021**, *127*, 126287. <https://doi.org/10.1016/j.eja.2021.126287>.

91. Quemada, M.; Gabriel, J.; Zarco-Tejada, P. Airborne Hyperspectral Images and Ground-Level Optical Sensors As Assessment Tools for Maize Nitrogen Fertilization. *Remote Sens.* **2014**, *6*, 2940–2962. <https://doi.org/10.3390/rs6042940>.
92. Wang, Y.; Dunn, B.L.; Arnall, D.B. Assessing Nitrogen Status in Potted Geranium through Discriminant Analysis of Ground Based Spectral Reflectance Data. *HortScience* **2012**, *47*, 343–348. <https://doi.org/10.21273/HORTSCI.47.3.343>.
93. Adams, M.L.; Norvell, W.A.; Philpot, W.D.; Peverly, J.H. Spectral Detection of Micronutrient Deficiency in ‘Bragg’ Soybean. *Agron. J.* **2000**, *92*, 261–268. <https://doi.org/10.2134/agronj2000.922261x>.
94. Avgoustaki, D.D.; Avgoustakis, I.; Miralles, C.C.; Sohn, J.; Xydis, G. Autonomous Mobile Robot with Attached Multispectral Camera to Monitor the Development of Crops and Detect Nutrient and Water Deficiencies in Vertical Farms. *Agronomy* **2022**, *12*, 2691. <https://doi.org/10.3390/agronomy12112691>.
95. Ayala-Silva, T.; Beyl, C.A. Changes in Spectral Reflectance of Wheat Leaves in Response to Specific Macronutrient Deficiency. *Adv. Space Res.* **2005**, *35*, 305–317. <https://doi.org/10.1016/j.asr.2004.09.008>.
96. Bhadra, S.; Sagan, V.; Maimaitijiang, M.; Maimaitiyiming, M.; Newcomb, M.; Shakoor, N.; Mockler, T.C. Quantifying Leaf Chlorophyll Concentration of Sorghum from Hyperspectral Data Using Derivative Calculus and Machine Learning. *Remote Sens.* **2020**, *12*, 2082. <https://doi.org/10.3390/rs12132082>.
97. Campbell, P.K.E.; Middleton, E.M.; Corp, L.A.; Kim, M.S. Contribution of Chlorophyll Fluorescence to the Apparent Vegetation Reflectance. *Sci. Total Environ.* **2008**, *404*, 433–439. <https://doi.org/10.1016/j.scitotenv.2007.11.004>.
98. Cartelat, A.; Cerovic, Z.G.; Goulas, Y.; Meyer, S.; Lelarge, C.; Prioul, J.-L.; Barbottin, A.; Jeuffroy, M.-H.; Gate, P.; Agati, G.; et al. Optically Assessed Contents of Leaf Polyphenolics and Chlorophyll as Indicators of Nitrogen Deficiency in Wheat (*Triticum aestivum* L.). *Field Crops Res.* **2005**, *91*, 35–49. <https://doi.org/10.1016/j.fcr.2004.05.002>.
99. Crema, A.; Boschetti, M.; Nutini, F.; Cillis, D.; Casa, R. Influence of Soil Properties on Maize and Wheat Nitrogen Status Assessment from Sentinel-2 Data. *Remote Sens.* **2020**, *12*, 2175. <https://doi.org/10.3390/rs12142175>.
100. Gomes, R.L.; Sousa, M.C.; Campos, F.G.; Boaro, C.S.F.; de Souza Passos, J.R.; Ferreira, G. Near-Infrared Leaf Reflectance Modeling of Annona Emarginata Seedlings for Early Detection of Variations in Nitrogen Concentration. *J. For. Res.* **2023**, *34*, 269–
282. <https://doi.org/10.1007/s11676-022-01557-3>.
101. Hong, S.; Schepers, J.S.; Francis, D.D.; Schlemmer, M.R. Comparison of Ground-Based Remote

Sensors for Evaluation of Corn Biomass Affected by Nitrogen Stress. *Commun. Soil Sci. Plant Anal.* **2007**, *38*, 2209–2226. <https://doi.org/10.1080/00103620701549157>.

102. Hunt, E.R.; Horneck, D.A.; Spinelli, C.B.; Turner, R.W.; Bruce, A.E.; Gadler, D.J.; Brungardt, J.J.; Hamm, P.B. Monitoring Nitrogen Status of Potatoes Using Small Unmanned Aerial Vehicles. *Precis. Agric.* **2018**, *19*, 314–333. <https://doi.org/10.1007/s11119-017-9518-5>.
103. Jang, S.; Han, J.; Cho, J.; Jung, J.; Lee, S.; Lee, D.; Kim, J. Estimation of Apple Leaf Nitrogen Concentration Using Hyperspectral Imaging-Based Wavelength Selection and Machine Learning. *Horticulturae* **2023**, *10*, 35. <https://doi.org/10.3390/horticulturae10010035>.
104. Kruse, J.K.; Christians, N.E.; Chaplin, M.H. Remote Sensing of Nitrogen Stress in Creeping Bentgrass. *Agron. J.* **2006**, *98*, 1640–1645. <https://doi.org/10.2134/agronj2006.0022>.
105. Lorenzo, G.A.; Mascarini, L.; Gonzalez, M.N. Dosis de N Sobre Reflectancia Espectral, Contenido de Clorofila y Nutrientes En Plantas de Gerbera. *Hortic. Bras.* **2017**, *35*, 278–285. <https://doi.org/10.1590/s0102-053620170220>.
106. Maresma, Á.; Lloveras, J.; Martínez-Casasnovas, J. Use of Multispectral Airborne Images to Improve In-Season Nitrogen Management, Predict Grain Yield and Estimate Economic Return of Maize in Irrigated High Yielding Environments. *Remote Sens.* **2018**, *10*, 543. <https://doi.org/10.3390/rs10040543>.
107. Mariotti, M.; Ercoli, L.; Masoni, A. Spectral Properties of Iron-Deficient Corn and Sunflower Leaves. *Remote Sens. Environ.* **1996**, *58*, 282–288. [https://doi.org/10.1016/S0034-4257\(96\)00070-3](https://doi.org/10.1016/S0034-4257(96)00070-3).
108. Masoni, A.; Ercoli, L.; Mariotti, M. Spectral Properties of Leaves Deficient in Iron, Sulfur, Magnesium, and Manganese. *Agron. J.* **1996**, *88*, 937–943. <https://doi.org/10.2134/agronj1996.00021962003600060015x>.
109. Milton, N.M.; Eiswerth, B.A.; Ager, C.M. Effect of Phosphorus Deficiency on Spectral Reflectance and Morphology of Soybean Plants. *Remote Sens. Environ.* **1991**, *36*, 121–127. [https://doi.org/10.1016/0034-4257\(91\)90034-4](https://doi.org/10.1016/0034-4257(91)90034-4).
110. Motomiya, A.V.d.A.; Valente, I.M.Q.; Molin, J.P.; Motomiya, W.R.; Biscaro, G.A.; Jordan, R.A. Índice de Vegetação No Algodoeiro Sob Diferentes Doses de Nitrogênio e Regulador de Crescimento. *Semin. Cienc. Agrar.* **2014**, *35*, 169–178. <https://doi.org/10.5433/1679-0359.2014v35n1p169>.
111. Niaz, N.; Gulzar, S.; Kazmi, J.H.; Aleem, S.A.; Pham, M.-P.P.; Mierzwa-Hersztek, M.M.-H.; Riaz Ahmed, S.; Lahori, A.H.; Mush-taq, Z.N. Assessment of Chlorophyll Content in Leaves of Crops and Orchards Based on SPAD, Multispectral, and Hyperspectral Techniques. *Ecol. Quest.* **2023**, *35*, 161–174. <https://doi.org/10.12775/EQ.2024.025>.
112. Oliveira, R.A.; Näsi, R.; Niemeläinen, O.; Nyholm, L.; Alhonoja, K.; Kaivosoja, J.; Jauhiainen, L.; Viljanen, N.; Nezami, S.; Markelin, L.; et al. Machine Learning Estimators for the Quantity and

Quality of Grass Swards Used for Silage Production Using Drone-Based Imaging Spectrometry and Photogrammetry. *Remote Sens. Environ.* **2020**, *246*, 111830. <https://doi.org/10.1016/j.rse.2020.111830>.

113. Padilla, F.M.; Teresa Peña-Fleitas, M.; Gallardo, M.; Thompson, R.B. Evaluation of Optical Sensor Measurements of Canopy Reflectance and of Leaf Flavonols and Chlorophyll Contents to Assess Crop Nitrogen Status of Muskmelon. *Eur. J. Agron.* **2014**, *58*, 39–52. <https://doi.org/10.1016/j.eja.2014.04.006>.
114. Pandey, P.; Singh, S.; Khan, M.S.; Semwal, M. Non-Invasive Estimation of Foliar Nitrogen Concentration Using Spectral Characteristics of Menthol Mint (*Mentha arvensis* L.). *Front. Plant Sci.* **2022**, *13*, 680282. <https://doi.org/10.3389/fpls.2022.680282>.
115. Padilla, F.M.; Karaca, C.; Peña-Fleitas, M.T.; Gallardo, M.; Rodríguez, A.; Thompson, R.B. Cultivar Effect on Proximal Optical Sensor Measurements and Estimation of Leaf N Content in Muskmelon and Sweet Pepper. *Eur. J. Agron.* **2024**, *159*, 127249. <https://doi.org/10.1016/j.eja.2024.127249>.
116. Putra, B.T.W. New Low-Cost Portable Sensing System Integrated with on-the-Go Fertilizer Application System for Plantation Crops. *Measurement* **2020**, *155*, 107562. <https://doi.org/10.1016/j.measurement.2020.107562>.
117. Rodriguez, D.; Fidgerald, G.J.; Belford, R.; Christensen, L.K. Detection of Nitrogen Deficiency in Wheat from Spectral Reflectance Indices and Basic Crop Eco-Physiological Concepts. *Aust. J. Agric. Res.* **2006**, *57*, 781–789. <https://doi.org/10.1071/AR05361>.
118. Roy Choudhury, M.; Christopher, J.; Das, S.; Apan, A.; Menzies, N.W.; Chapman, S.; Mellor, V.; Dang, Y.P. Detection of Calcium, Magnesium, and Chlorophyll Variations of Wheat Genotypes on Sodic Soils Using Hyperspectral Red Edge Parameters. *Environ. Technol. Innov.* **2022**, *27*, 102469. <https://doi.org/10.1016/j.eti.2022.102469>.
119. Rustioni, L.; Grossi, D.; Brancadoro, L.; Failla, O. Characterization of Iron Deficiency Symptoms in Grapevine (*Vitis* Spp.) Leaves by Reflectance Spectroscopy. *Plant Physiol. Biochem.* **2017**, *118*, 342–347. <https://doi.org/10.1016/j.plaphy.2017.06.031>.
120. Sanaeifar, A.; Yang, C.; Min, A.; Jones, C.R.; Michaels, T.E.; Krueger, Q.J.; Barnes, R.; Velte, T.J. Noninvasive Early Detection of Nutrient Deficiencies in Greenhouse-Grown Industrial Hemp Using Hyperspectral Imaging. *Remote Sens.* **2024**, *16*, 187. <https://doi.org/10.3390/rs16010187>.
121. Sharaf-Eldin, M.A.; Elsayed, S.; Elmetwalli, A.H.; Yaseen, Z.M.; Moghanm, F.S.; Elbagory, M.; El-Nahrawy, S.; Omara, A.E.-D.; Tyler, A.N.; Elsherbiny, O. Using Optimized Three-Band Spectral Indices and a Machine Learning Model to Assess Squash Characteristics under Moisture and Potassium Deficiency Stress. *Horticulturae* **2023**, *9*, 79. <https://doi.org/10.3390/horticulturae9010079>.

122. Tang, Z.; Wang, X.; Xiang, Y.; Liang, J.; Guo, J.; Li, W.; Lu, J.; Du, R.; Li, Z.; Zhang, F. Application of Hyperspectral Technology for Leaf Function Monitoring and Nitrogen Nutrient Diagnosis in Soybean (*Glycine max* L.) Production Systems on the Loess Plateau of China. *Eur. J. Agron.* **2024**, *154*, 127098. <https://doi.org/10.1016/j.eja.2024.127098>.
123. Wang, S.; Guan, K.; Wang, Z.; Ainsworth, E.A.; Zheng, T.; Townsend, P.A.; Liu, N.; Nafziger, E.; Masters, M.D.; Li, K.; et al. Airborne Hyperspectral Imaging of Nitrogen Deficiency on Crop Traits and Yield of Maize by Machine Learning and Radiative Transfer Modeling. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102617. <https://doi.org/10.1016/j.jag.2021.102617>.
124. Shen, Z.; Wang, K.; Wang, R. Feasibility of Field Evaluation of Rice Nitrogen Status from Reflectance Spectra of Canopy. *Pedosphere A Q. J. Soil Sci.* **1998**, *8*, 121–126.
125. Yi, J.; Krusenbaum, L.; Unger, P.; Hüging, H.; Seidel, S.J.; Schaaf, G.; Gall, J. Deep Learning for Non-Invasive Diagnosis of Nutrient Deficiencies in Sugar Beet Using RGB Images. *Sensors* **2020**, *20*, 5893. <https://doi.org/10.3390/s20205893>.
126. Zhao, D.; Raja Reddy, K.; Kakani, V.G.; Read, J.J.; Carter, G.A. Corn (*Zea mays* L.) Growth, Leaf Pigment Concentration, Photosynthesis and Leaf Hyperspectral Reflectance Properties as Affected by Nitrogen Supply. *Plant Soil* **2003**, *257*, 205–218. <https://doi.org/10.1023/A:1026233732507>.
127. de Mello Prado, R. *Mineral Nutrition of Tropical Plants*; Springer International Publishing: Cham, Swiderland, 2021; 339 p. ISBN 978-3-030-71261-7.

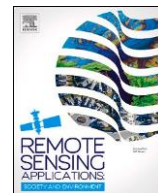
Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

journal homepage: www.elsevier.com/locate/rsase

Capítulo 3. New approach for predicting nitrogen and pigments in maize from hyperspectral data and machine learning models

Bianca Cavalcante da Silva^a, Renato de Mello Prado^a, Fábio Henrique Rojo Baio^b,
Cid Naudi Silva Campos^b, Larissa Pereira Ribeiro Teodoro^b,
Paulo Eduardo Teodoro^b, Dthenifer Cordeiro Santana^c,
Thiago Feliph Silva Fernandes^a, Carlos Antonio da Silva Junior^{d,*},
Elisangela de Souza Loureiro^b

^a Department of Soils and Fertilizers, Paulista State University "Júlio de Mesquita Filho" UNESP/FCAV, Jaboticabal, Brazil

^b Federal University of Mato Grosso Do Sul (UFMS), Chapadão Do Sul, 79560-000, MS, Brazil

^c Department of Agronomy, State University of São Paulo (UNESP), Ilha Solteira, 15385-000, SP, Brazil

^d Department of Geography, State University of Mato Grosso (UNEMAT), Sinop, 78550-000, MT, Brazil

ARTICLE INFO

Keywords:

Zea mays

Precision agriculture

Leaf N content

Nutritional diagnosis

ABSTRACT

Fast diagnostics from hyperspectral data and machine learning (ML) models to predict nitrogen (N) and pigment content in maize crops is challenging to optimize nitrogen fertilization. This research assessed the efficiency of the five ML algorithms, the best phenological stage, and the sensitivity of the 90 spectra to estimate N and pigment content. Therefore, this field research proposes as a novelty to test which of the five ML algorithms accurately estimates nitrogen, chlorophyll, and carotenoid content in maize leaves at different phenological stages using hyperspectral band data. The treatments were arranged in a factorial scheme with four N doses (0, 54, 108, and 216 kg ha⁻¹) combined with five leaf collection seasons at phenological stages V6 to V14. The ML models tested were artificial neural networks – ANN, decision tree adapted for prediction problems – M5P, REPTree decision tree, random forest – RF, polynomial support vector machine – PSVM, and ZeroR – ZR (control). Spectral bands 530–560 nm and 690–750 nm are effective wavelengths because the visible region with lower reflectance (530–560 nm) affects N uptake and chlorophyll and carotenoid content, while the red-edge and near-infrared region affects N and chlorophyll content. The random forest (RF) model performed better with higher correlation (r) and mean absolute error (MAE) between predicted and observed values for all variables, with the correlation coefficient (r) value being around 0.6 and the MAE below 0.5 for the prediction of chlorophyll a+b. For the prediction of flavonoids, the r was around 0.6 and the error was 0.07. Support vector machine (SVM) and RF efficiently predicted nitrogen content, in predicting of NF, the r values for both algorithms were above 0.35 and the error was below 2.75.

* Corresponding author.

E-mail addresses: bianca.cavalcante@unesp.br (B.C.d. Silva), rm.prado@unesp.br (R.d.M. Prado), fabio.baio@ufms.br (F.H.R. Baio), cid.campos@ufms.br (C.N.S. Campos), larissa_ribeiro@ufms.br (L.P.R. Teodoro), paulo.teodoro@ufms.br (P.E. Teodoro), dthenifer.santana@unesp.br (D.C. Santana), thiago.feliph@unesp.br (T.F.S. Fernandes), carlosjr@unemat.br (C.A.d. Silva Junior), elisangela.loureiro@ufms.br (E.d.S. Loureiro).

<https://doi.org/10.1016/j.rsase.2023.101110>

Received 7 September 2023; Received in revised form 21 November 2023; Accepted 22 November 2023

Available online 27 November 2023

2352-9385/© 2023 Elsevier B.V. All rights reserved.

1. Introduction

Nitrogen is the nutrient most absorbed by maize and that which most limits maize yield in different regions of the world. Farmers widely use nitrogen fertilizer in crops, but crop response depends on its nutritional status, requiring foliar diagnosis before planning for fertilization (de Mello Prado, 2021). The method most used for foliar diagnosis is the conventional method that chemically analyzes plants. However it demands time, generates waste, and interferes with the logistics of the task performed, especially in large areas (Li et al., 2019). Furthermore, its cost can be high and its application may include many diagnostic areas.

In order to solve these problems, technologies such as machine learning can be used to obtain and process data. Some examples of authors that used it are Coulibaly et al. (2022), who obtained optimal results, and Rengasamy et al. (2022) who estimated the nutritional status of crops. In this scenario, precision agriculture has emerged as a management technique for farmers seeking greater productivity with less environmental impact, with precision exceeding that of conventional techniques (Coulibaly et al., 2022). This technique can generate spectral responses (such as hyperspectral data) from the collection of many data. Moreover, it is nondestructive and fast and can be used to estimate plant nutritional content (Osco et al., 2019). It thus positively reflects food security, environmental sustainability, and human health (D. Li et al., 2022).

Recent advances in remote sensing lead to different spectral reflectance values due to the varying physical and chemical characteristics of the materials. These characteristics can be captured by spectroscopic instruments and subsequently learned by machine learning with algorithms (Mi et al., 2020; Su et al., 2018). Some algorithms that have been successfully applied in prediction studies using spectral variables as input in datasets are random forest (RF), artificial neural networks (ANNs), decision trees (DTs), and support vector machines (Su et al., 2022; P. K. Goel et al., 2003; Mercier and Lennon 2003; Fujita et al., 2020a,b; Safa et al., 2019; Barzin et al., 2021). The random forest (RF) algorithm builds and merges several independent regression trees to obtain more accurate and stable predictions (Shao et al., 2012), as in maize yield grain prediction using multispectral data (Marques Ramos et al., 2020).

In turn, SVM makes predictions by projecting the original finite dataset into a hyperplane using a kernel function, which can achieve accuracy and generalization ability with samples that are not large enough or when the dimensionality of the features is higher than the number of samples (Shao et al., 2012; Abdel-Rahman et al., 2013). In addition to prediction, it is a good algorithm for classifying soybean genotypes in terms of primary macronutrients, for example (Santana et al., 2023a,b). The use of the SVM algorithm also showed good performance in predicting the SPAD index in corn plants, which can be related to several plant stress factors (Guo et al., 2022). The SVM had high yield prediction accuracies in rice in which phenological variables improved the accuracy of yield predictions (Guo et al., 2021).

Artificial Neural Networks are one of the main machine learning algorithms used in agriculture. This is because ANNs adjust the weights of node connections in hidden layers to obtain a higher correlation between observed and predicted values and have strong fault tolerance and learning capability (Lee et al., 2020). A practical example of its application is the use of a type of neural network in estimating water quality using multi-sensor data, showing promise in predicting imminent outbreaks of water pollution, such as algal blooms (Sagan et al., 2020).

Moreover, the M5P algorithm is a decision tree adapted for prediction problems. It consists of the steps of branching data to grow a complete tree; developing a regression model at each node for prediction, then pruning the tree to avoid the overfitting problem; and monitoring the tree to compensate for the sharp discontinuities caused by splitting (L. Li et al., 2022). Another decision tree algorithm is REPTree, a fast decision tree model that builds a decision/regression tree using information

on its gain/variance by reduced error pruning with backfitting (Wen et al., 2018).

Algorithms improve model and data interpretation in cases with complex, nonlinear data. Moreover, traditional models like linear regression usually result in low prediction accuracy, especially in spectral data analysis. The preferred algorithm in the literature is the random forest (RF) classifier. This is because it performs well in terms of accuracy and robustness with a small number of data and the least overfitting (Su et al., 2022). Besides RF, researchers can use ANNs (Goel et al., 2003) and SVM (Mercier and Lennon 2003) to treat hyperspectral data. In turn, ZeroR-ZR learning algorithms train only the seen class and predict the unseen class, *i.e.*, information that is not in the training data (Fujita et al., 2020a,b).

Some researchers like Safa et al. (2019) used the ANN model to predict grassland N content. Barzin et al. (2021) reported that the SVM model is accurate in estimating leaf N content. The latter authors used red, red-edge, and NIR spectral bands to calculate various vegetation indices (VIs). They found that the VIs SCCCI and ERVI in association with SVM were the most effective VIs to estimate corn leaf N concentration. Different authors also attested the accuracy the SVM model (D. Li et al., 2022). Najafzadeh and Niazmardi (2021), by proposing a modification of the SVM regression algorithm (called the multiple kernel support vector regression algorithm), found an outperformance of this model in estimating biochemical oxygen demand (BOD) and chemical oxygen demand (COD). In addition, research conducted by Shao et al. (2012), which used reflectance spectroscopy in the visible and near-infrared spectral regions, indicated SVM as a promising model to quantify N status in rice.

Similarly, data from Abdel-Rahman et al. (2013) indicated that RF regression applied to hyperspectral data is suitable for predicting N content in sugarcane. Lee et al. (2020) tested the SVM and RF methods by adding more variables to the models. Their study indicated that RF performed better and that RF outputs are easier to interpret than SVM outputs. For nitrogen prediction in maize, the partial least squares (PLS) regression model showed the highest coefficient of determination in the spectral range of 325–1075 nm (L. Li et al., 2022).

It is paramount to improve these alternative methods for nutritional diagnosis to increase the accuracy of the results. An important point in this regard concerns N content estimation, in which near-visible spectrum (VIS) wavelengths and near-infrared (NIR) band reflectance correlate positively with leaf chlorophyll content and N sensitivity (R. Li et al., 2019). Principal component analysis (PCA) is used to extract principal components, or clusters, that represent the different distributions of correlated data (Chang et al., 2022).

This analysis considers the amount of information in each spectral band for the main component as a wavelength selection criterion to ensure that the optimal spectral band combination contains most of the variation in the data (Chen et al., 2019). The PCA method has been used in several agricultural studies, such as for clustering soybean genotypes according to N, P and K contents (Santana et al., 2023a,b), association of spectral bands reflected by canopy and Amazonian species (Della-Silva et al., 2022), clustering of soybean genotypes for industrial traits (Santana et al., 2023a,b), clustering according to storage conditions and initial moisture of corn kernels (Coradi et al., 2022).

The use of the hyperspectral spectroscopy technique in N-deficient leaves showed lower chlorophyll content as well as chlorosis from the visible wavelengths. According to Wen et al. (2018), visible light reflectance (436–698 nm) characterized deficient maize leaves, while NIR reflectance (740–1350 nm) was higher in plants that received full nitrogen fertilization. Notwithstanding, the main studies on this topic focus on analyzing the performance of traditional vegetation indices designed for N prediction, which limits the accuracy and the practical value of the results. This is because these studies do not analyze all bands but rather the relationship of a spectral band with the other. In addition, the authors bring about the problem of saturation in places with high amounts of green biomass and with high leaf area index. This can either underestimate biomass values or bring invariable results (Gu et al., 2013). Therefore, few studies correlate the reading of several spectral bands with different machine learning approaches in assessing maize leaf N content and photosynthetic pigments using a hyperspectral sensor.

The hypotheses of this research were the following: nitrogen doses affect maize physiologically,

and these effects can manifest themselves through spectra and influence the content of chlorophyll and carotenoids, making it possible to use hyperspectral data to predict the nitrogen status in corn crops. The objective of the work was to evaluate the prediction accuracy of machine learning algorithms in predicting maize pigments using hyperspectral data.

This research thus identified which ML algorithms accurately estimate nitrogen, chlorophyll, and carotenoid content in maize leaves at different phenological stages. For that, the authors considered data from 90 hyperspectral bands.

2. Materials and methods

2.1 Location and characteristics of the experimental area

The experiments were carried out in the experimental area of the Universidade Federal de Mato Grosso do Sul, Chapadão do Sul Campus, Mato Grosso do Sul State, Brazil. Geographical coordinates are 18° 46' 16" S latitude, 52° 37' 22" longitude, with an average altitude of 820 m (Fig. 1). The soil of the study area is an Oxisol with previous maize cultivation (2020/2021 crop yr⁻¹). The predominant climate in the region is humid tropical (Aw), with a rainy season in the summer and a dry season in the winter. Average annual rainfall is 1850 mm. Annual temperature ranges from 13 to 28 °C (daily measurements).

The authors obtained weather data from the National Institute of Meteorology station (INMET, 2022) in Chapadão do Sul city (18° 48' S and 52° 36' W, 821 m altitude). In the 2021/2022 maize crop, global radiation in the five experimental months - from sowing to harvest - ranged between 7.6 and 30.5 MJ m² day, and temperatures ranged between 19.7 and 27.81 °C (Fig. 2a). In that crop year, rainfall ranged between 0 and 105.0 mm, totaling 1305.5 mm over the experimental period (Fig. 2b).

The experiment used the early-cycle hybrid maize B2401PWU. Seed treatment included the use of insecticides Chlorantraniliprole (625 g L⁻¹) and Clothianidin (600 g L⁻¹) at doses of 48 mL and 70 mL, respectively, applied on 60 thousand seeds.

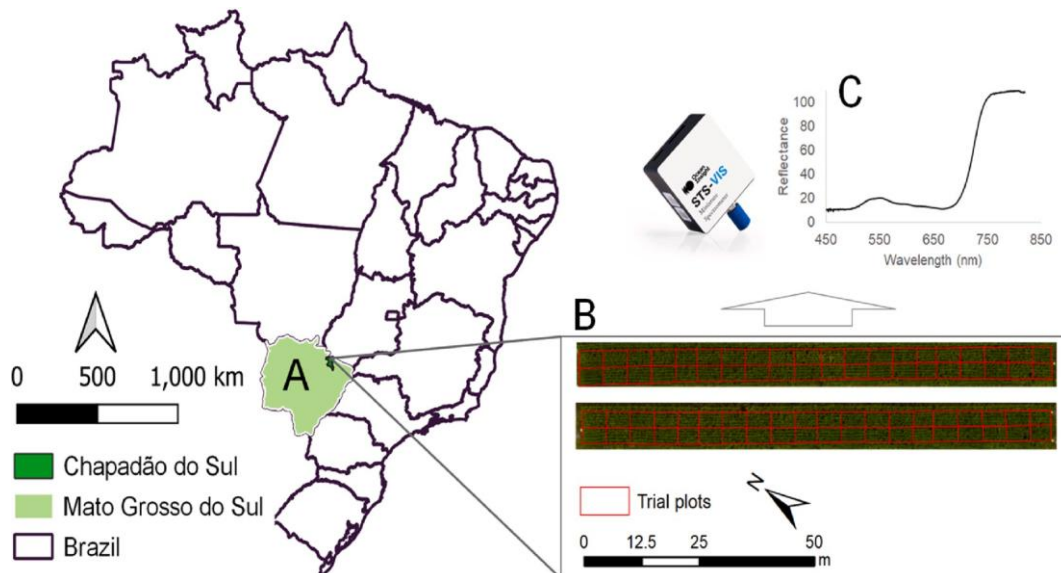


Fig. 1. Location of Chapadão do Sul city, Brazil (A), experimental plot design (B), and sensor used for spectral data acquisition (C).

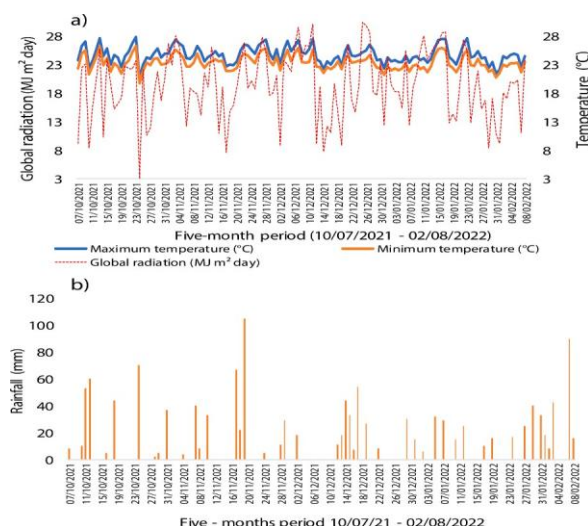


Fig. 2. Global radiation ($\text{MJ m}^{-2} \text{ day}$), minimum and maximum temperature ($^{\circ}\text{C}$) (a), and rainfall (b) during the experimental period from July 10, 2021 (sowing) to August 02, 2022 (harvest).

2.1.1 Experimental design

The experimental design was in randomized blocks, with treatments consisting of four N levels: 0 kg ha^{-1} (N1), 54 kg ha^{-1} (N2), 108 kg ha^{-1} (N3), and 216 kg ha^{-1} (N4) evaluated at five periods (phenological stages V6, V8, V10, V12, and V14). The experiment included the collection of new fully expanded leaves, discarding the leaf cartridge. Phenological stages corresponded to periods in which the leaves were fully developed (V6 - sixth fully opened leaf, with the edges of each leaflet no longer touching; V8 - eighth fully opened leaf; V10 - tenth fully opened leaf; V12 - twelfth fully opened leaf; and V14 - fourteenth fully opened leaf). The experimental unit consisted of six maize rows spaced 0.50 m apart and 5 m long.

2.1.2 Conduction of the experiment and application of treatments

In the previous crop year (2020/2021), liming was performed in the experimental area to raise base saturation ($\text{K} + \text{Ca} + \text{Mg} / \text{K} + \text{Ca} + \text{Mg} + \text{H} + \text{Al}$) to 60%, as indicated by (de Sousa and Lobato 2004). In addition, before soybean sowing, 500 kg ha^{-1} of CaSO_4 was applied and then millet was sown and used as mulch.

The experiment was then installed with soil preparation and installation of the maize crop for the 2021/2022 agricultural year during the period from October 07 (sowing) to February 08 (harvest). Glyphosate herbicide (588 g L^{-1}) was applied at the phenological stage V4. Insecticides Thiamethoxam (141 g L^{-1}) and Lamb-da-Cyhalothrin (106 g L^{-1}) were applied at stages V4 and V8. In addition, 300 mL ha^{-1} of the fungicides Azoxystrobin (200 g L^{-1}) and Cyproconazole (80 g L^{-1}) were applied 40 days after planting, and 500 mL ha^{-1} of the fungicides Azoxystrobin and Flutriafol (125 g L^{-1}) were applied 60 days after planting.

Planting and top dressing fertilizers for each nutrient were established based on initial soil analysis according to the recommendation of soil correction and fertilization for the maize crop, with expected yield between 10 and 12 t ha^{-1} (de Sousa and Lobato 2004). The soil was chemically analyzed for fertility with the following results (Table 1).

Sowing methods included the application of 20 kg ha⁻¹ N, 70 kg ha⁻¹ P₂O₅, and 40 kg ha⁻¹ K₂O. Nitrogen and potassium top dressing fertilizers were applied in split doses at phenological stages V4 and V8. At phenological stage V4 (fourth fully developed maize leaf), 45 kg ha⁻¹ K₂O was applied. The same quantity was applied at phenological stage V8 using potassium chloride. A total of 2 kg ha⁻¹ B in the form of Borax (15% B) was applied at planting by broadcasting in the total area. A solution containing 0.5% zinc sulfate and manganese sulfate was applied to the leaves, with syrup volume of 200 L ha⁻¹. Soon after the first nitrogen fertilization, 100 mL ha⁻¹ of sodium molybdate (15.5% Mo) was applied, with a syrup volume of 200 L ha⁻¹.

Table 1

Chemical analysis of the soil from the experimental area.							
pH (CaCl ₂)	Ca	Mg	Al	H + Al	K	P (mehlich)	S
-----cmolc dm ⁻³ -----					------(mg dm ⁻³)----		
5.30	4.50	1.10	0.05	4.2	0.32	34.0	13.6
B	Cu	Fe	Mn	Zn	Organic matter	Organic carbon	Cation exchange capacity
-----mg dm ⁻³ -----					-----g dm ⁻³ -----	cmolc dm ⁻³	
0.28	1.5	49.0	14.7	4.8	33.2	19.3	10.1
M.O. (g dm ⁻³)	CTC (cmolc)	V%	Sat.Al %				
33.2	10.1	58.5	0.8				

pH CaCl₂, Ca: calcium, Mg: magnesium, Al: aluminum, H + Al: potential acidity, K: potassium, P: phosphorus (mehlich), S: sulfur, B: boron, Cu: copper, Fe: iron, Mn: manganese, Zn: zinc, MO: organic matter, CEC: cation exchange capacity. V%: base saturation, Sat.Al%: aluminum saturation (Van Raij and Quaggio, 2001).

Nitrogen (20 kg ha⁻¹ N) in the form of urea was applied for initial crop development in all plots.

For top dressing fertilization, nitrogen was used at doses 0, 54, 108, and 216 kg ha⁻¹ N, corresponding to 0%, 30%, 60%, and 120% of the recommended dose of 180 kg ha⁻¹ N (de Sousa and Lobato 2004). Half of these doses were applied at stage V4, and the remaining were applied at stage V8. Yield potential analysis considers from the V4 to the V6 stage since the nutritional requirements of maize increase during this period in which nutrition should be adequate so as not to impair crop yield.

Leaves were sprayed when 50% of the plants had the fifth leaf fully developed (stage sV5). Spraying application consisted of a single rate of 250 mL ha⁻¹, using strain BR 11005 (Sp 245) of *Azospirillum brasilense* (1 x 10⁹ viable cells mL⁻¹) (Kappes et al., 2017).

3. Analyses

Analyses were performed at phenological stages V6 (sixth fully developed leaf), V8 (eighth fully expanded leaf), V10 (tenth fully developed leaf), V12 (twelfth fully expanded leaf), and V14 (fourteenth fully developed leaf). For this, samples of 10 new and fully expanded leaves per plot were taken. The collected leaves were washed with 0.1% neutral detergent solution, 0.3% HCl solution, and deionized water (de Mello Prado, 2021). Spectral readings were taken and the same leaves were used for pigment and N content analysis. A principal component analysis was also carried out.

3.1 Pigments (chlorophyll a + b and carotenoids)

The analyses of chlorophyll (chlorophyll *a* + *b*) and carotenoid contents followed the methodology proposed by Lichtenthaler (1987). For that, discs were collected with 0.025–0.030 g of fresh maize leaf samples considering the upper third of the fourth leaf. Then, 1.5 mL of 80% acetone was added to the samples in Eppendorf tubes, with a 72-h waiting period. The samples were subsequently analyzed by an absorption spectrophotometer at 663 nm for chlorophyll *a*, 647 nm for chlorophyll *b*, and 470 nm for carotenoids.

3.2 Nitrogen content

The plant material was dried in an oven with forced air circulation (65 ± 5 °C) until constant weight, providing maize leaf dry matter. The plant material was ground and chemically analyzed by sulfuric digestion. Total N content was determined according to the method described by Bataglia et al. (1983).

3.3 Spectral assessments

Maize spectral bands were measured in the laboratory after the removal of leaves 6, 8, 10, 12, and 14 in the field at stages V6, V8, V10, V12, and V14, respectively. Styrofoam was used to store the leaves, thus avoiding turgidity loss. The STS-VIS hyperspectral sensor (Ocean Insight, USA) was calibrated at each start of the readings for each treatment block. The reflectance range comprised the spectral bands 450.15–820.80 nm, each with 0.45 nm in length, totaling 780 spectral bands that were used as input variables in the machine learning models. A barium sulfate plate was used in this calibration process. Laboratory readings were performed with artificial illumination by two 120 W halogen lamps pointed at the reading point.

The sensor was connected to a computer for recording and displaying the readings. The latter equipment was connected to a program called OceanView, which records readings made in the electromagnetic spectrum range between 450 and 850 nm with a spectral resolution of 1.5 nm (FWHM). This equipment has a signal-to-noise lower than >1500:1, and dynamic range performance of 4600:1. Among the spectrum range, the procedure considered spectral information measured every 0.5 nm, accounted as the average calculated every 5 nm, totaling 90 measurements of reflectance.

Wavelengths were grouped in representative interval means into bands as recommended by da Silva Junior et al. (2018) to relate the ranges within each band to the variables and doses evaluated (Table 2).

Table 2

Wavelength intervals (nm) used in establishing 16 spectral bands for statistical analysis.

Bands (n°)	Spectral Range (nm)	Mean wavelength (nm)
B1	450–475	462.5
B2	480	480
B3	481–500	490.47
B4	501–530	515.34
B5	531–539	534.89
B6	540	540
B7	541–649	594.80
B8	650	650
B9	661–670	665.54
B10	675	675
B11	676–684	680.05
B12	685–689	686.83
B13	690–700	695.07
B14	701–709	705.03

B15	710	710
B16	711–730	720.60

3.4 Machine learning analysis

Hyperspectral data were used as input variables while N, chlorophyll, and carotenoid contents were used as output variables in the prediction models. Supervised machine learning models were: artificial neural networks – ANN; decision tree adapted for prediction problems – M5P; REPTree – REPT Saha et al. (2020), random forest – RF Belgiu and Drăguț (2016), polynomial support vector machine – SVMP (Nalepa and Kawulok 2019), and ZeroR (ZR).

A random cross-validation sampling strategy with k-fold = 10 and 10 repetitions (total of 100 runs) was applied. This strategy was used to evaluate the performance of the six supervised machine learning models. All model parameters were set according to the default setting of the Weka 3.8.5 software. The ANN tested consists of a single hidden layer formed by a number of neurons that is equal to the number of attributes plus the number of classes, all divided by 2. The ANN adopted a learning rate of 0.3 and a momentum rate of 0.2, and used the backpropagation algorithm to learn a multi-layer perceptron to predict the variables (Egmont-Petersen et al., 2002). REPTree model is an adaptation of the C4.5 classifier that can be used in regression problems with an additional pruning step based on an error reduction strategy (with backfitting), and only sorts values for numeric attributes once (Snouy et al., 2011). The REPTree model in Weka default adopts a minimum total weight of the instances in a leaf equal to 2 and a minimum proportion of the variance on all the data that needs to be present at a node in order for splitting to be performed in regression trees equal to 0.001. The M5P model is a reconstruction of Quinlan's M5 algorithm Quinlan Bassar (1992) based on the conventional decision tree with the addition of a linear regression function to the leaf nodes (Blaifi et al., 2018). In M5P algorithm, the minimum number of instances to allow at a leaf node adopted was 4. RF model produces several prediction trees for the same data set and use a voting scheme among all learned trees to predict new values (Belgiu and Drăguț, 2016). The RF was built using a number of trees equal to 100 and adopting the default settings of the Weka software for the remaining hyperparameters. In turn, SVM used the optimizer proposed by Chen et al. (2019) that improves sequential minimal optimization (SMO) for regression with a polynomial kernel. Moreover, ZR was used as a control technique, applying the mean value of the target variable and making that value available as a prediction. The prediction method can thus return a number of merit higher than 1 (since the relative increase can exceed 100%).

The performance of each model and the relationship between the predicted and observed values of the variables were assessed using the Pearson correlation coefficient (r) metrics and the mean absolute error (MAE) between the predicted and observed values by each algorithm considering the 10 folds. For a series of n measurements of the observed ($\mathbf{x}_i$) and predicted ($\mathbf{y}_i$) values, where $i = 1, 2, 3, \dots, n$, the correlation coefficient between the observed and predicted values ($\mathbf{r}_{xy}$) is calculated by the following expression:

$$\mathbf{r}_{xy} = \left(\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum_{i=1}^n (x_i - \bar{x})^2) \sum_{i=1}^n (y_i - \bar{y})^2}} \right), \text{ where } \bar{x} \text{ and } \bar{y} \text{ are the sample means of the the observed and}$$

predicted mean values, respectively. Mean absolute error can be estimated by the following expression: $MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$, where n is the number of samples, y_i is the observed absolute value for each sample, $\hat{y}_i$ is the predicted absolute value by the model for each sample.

The means of r and MAE for each ML model were submitted to analysis of variance and the means were grouped by the Scott-Knott test at 5% probability.

4. Results

4.1 Variability of the reflectance factor spectrum under different N treatments

The distributions of larger reflectance factor spectrum indicate greater contributions of each wavelength to predict N status in maize (Fig. 3). Reflectance factor spectrum were similar for prediction characteristics because total chlorophyll (Chl a+b), carotenoid (Cx + c), and nitrogen contents are closely linked variables, which hinders the differentiation of N doses in the different vegetative stages. In this regard, some wavelengths showed important spectral attenuations, *e.g.* green spectra (about 550 nm) forming a group (530–560 nm) and red-edge spectra (about 750 nm) forming another group (690–820 nm). Yao et al. (2015) describe that the visible region of the green band strongly correlates with increased chlorophyll reflectance; the decrease in reflectance in the red region (690–820 nm) is a sensitive indicator of chlorophyll activity, that is, the lower the reflectance in this range, the greater its absorption, indicating greater chlorophyll activity that can be influenced by nitrogen fertilization.

Fig. 3 illustrates the separation of the inflection intervals in maize leaves, which are the curved and concave portions present in the mean curves of the spectra, as described by Nanni et al. (Nalepa and Kawulok 2019) and da Silva Junior et al. (2018) for soybean crop. The inflection points in the reflectance factor are from features known as phototropism (370 nm), a-carotene (420 nm), b-carotene (425 nm), Chlorophyll *a* (430 nm), a-carotene (440 nm), Xanthophyll (445 nm), Chlorophyll synthesis (445 nm), b-carotene (450 nm), Chlorophyll *b* (453 nm), a-carotene (470 nm), Xanthophyll (475 nm), b-carotene (480 nm), Chlorophyll synthesis (650 nm), Chlorophyll absorption (960 nm) (da Silva Junior et al., 2018).

Principal component analysis (Fig. 4) gathered 96.1% of the accumulated variance in the first two components. The approximation of spectral bands B5 (531–539 nm), B6, (540 nm) and B13 (690–700 nm) with nitrogen doses N3 (108 kg ha⁻¹) and N4 (216 kg ha⁻¹) is remarkable (Fig. 4). Leaf nitrogen content, the evaluated pigments, and bands B4 (501–530 nm), B11 (676–684 nm), and B12 (685–689 nm) were close to the nitrogen dose N3 (108 kg ha⁻¹), while B7 (541–649 nm) and B8 (650 nm) correlated with a lower dose of nitrogen (N1, 0 kg ha⁻¹). In a study using PCA for clustering soybean genotypes for macronutrient contents, Santana et al. (2023a,b) found two groups of genotypes clustered by similarity in N, P, and K contents. Principal component analysis identifies similar groups, differentiating groups of soybean genotypes based on industrial traits used as output variables in machine learning

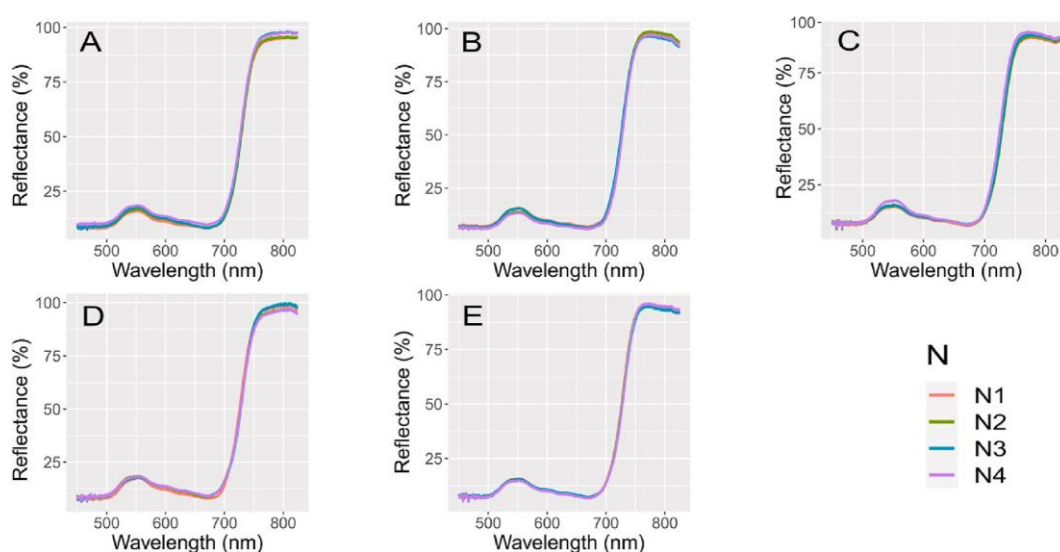


Fig. 3. Reflectance factor spectrum of maize evaluated at stages V6 (A), V8 (B), V10 (C), V12 (D), and V14 (E) in response to nitrogen doses (N1: 0; N2: 54; N3: 108; and N4: 216 kg ha⁻¹).

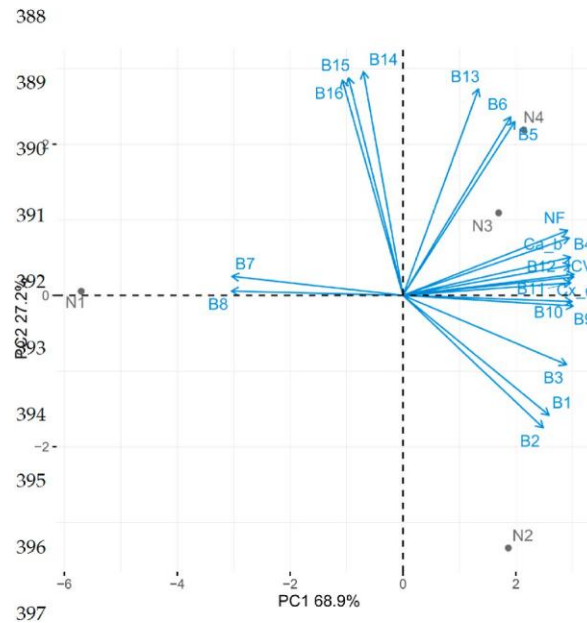


Fig. 4. Principal component analysis for the bands formed with the wavelengths and for the analyzed pigments and leaf nitrogen.

models (Santana et al., 2023a,b). Furthermore, Rossi et al. (2023) verified that it is possible to rank the variables according to the degree of vegetation cover, moisture content, and land cover using PCA.

4.2 Evaluation of machine learning models

Fig. 5, Fig. 6, and Fig. 7 group the box plots for the values of the Pearson correlation coefficient (r) and mean absolute error (MAE) between the predicted and observed values for maize total chlorophyll (Chl a + b), carotenoid (Cx + c), and nitrogen contents obtained with different machine learning models (ANN, M5P, REPT, RF, SVM, and ZR). The random forest (RF) model showed higher accuracy, consisting of a better machine learning technique for prediction of all variables. This is because it provided the highest values of r and the lowest values of MAE. Moreover, it presented smaller box diagrams, evidencing less dispersion of the data used to obtain r and MAE. For the variable N content, the SVM model obtained a significant effect of r , equal to the RF model, and the lowest MAE values, however with greater data variability (Fig. 6). In a study for predicting N input and export, (Zheng et al., 2022) verified that the SVM model was more accurate than the RF model. In a similar study using hyperspectral data (spectral bands and vegetation indices) on wheat leaves, the RF algorithm outperformed the other models for estimating total chlorophyll

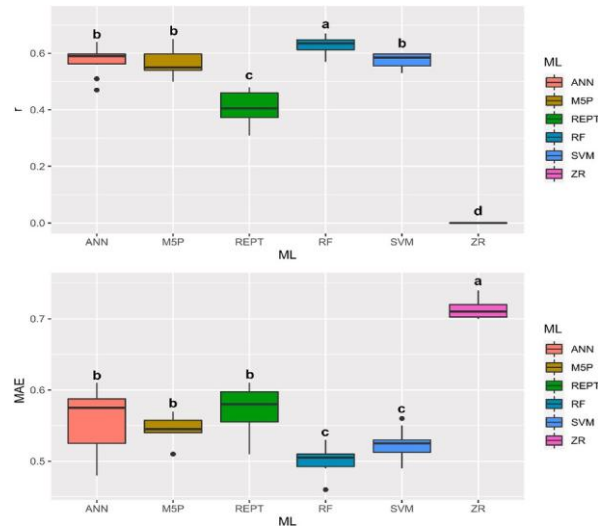


Fig. 5. Mean clustering for the Pearson correlation coefficient (r) and mean absolute error (MAE) between the observed values of chlorophyll $a + b$ content and the values predicted by different machine learning (ML) algorithms using hyperspectral data as input. Means followed by equal letters in the same figure do not differ by the Scott-Knott test at 5% probability.

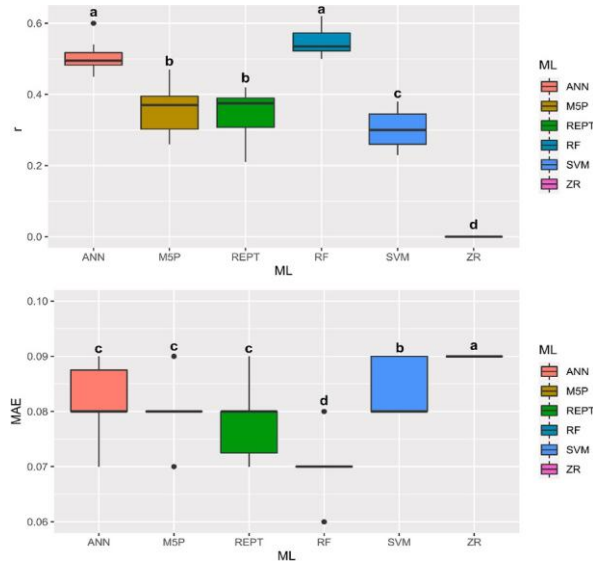


Fig. 6. Mean clustering for the Pearson correlation coefficient (r) and mean absolute error (MAE) between the observed values of carotenoid content ($c + x$) and the values predicted by different machine learning (ML) algorithms using hyperspectral data as input. Means followed by equal letters in the same figure do not differ by the Scott-Knott test at 5% probability.

when using all spectral bands as input variables instead of vegetation indices (Shah et al., 2019). Carotenoid prediction studies, such as those by Oivukkamäki et al. (2023); Wong et al. (2019) and Filella et al. (2009) have used only vegetation indices as variables in the dataset. Our work brings a novelty by also using spectral bands isolated as input data for carotenoid prediction. In a study with a similar data analysis approach, considering only vegetation indices or in combination with spectral bands as input variables, Teodoro et al. (2021) verified the superiority of RF. The RF algorithm associated with vegetation indices best predicted citrus canopies with lower nitrogen content (Osco et al., 2020).

5. Discussion

5.1 Variability of the reflectance factor spectrum under different N treatments

The maize spectral curves evaluated at different phenological stages (Fig. 3) behaved similarly, but there was variability between readings, significantly affecting the performance of the algorithms. This suggests that the technique can early predict N status, even before maize plants present visual symptoms of N deficiency. Before visual symptoms in the plant, physiological losses have already occurred in the leaves (Osco et al., 2020). Therefore, it is important to have foliar diagnosis techniques that can reliably detect eventual N deficiencies in maize plants at young stages of the plant so that decisions can be made quickly and crop yield losses can be minimized.

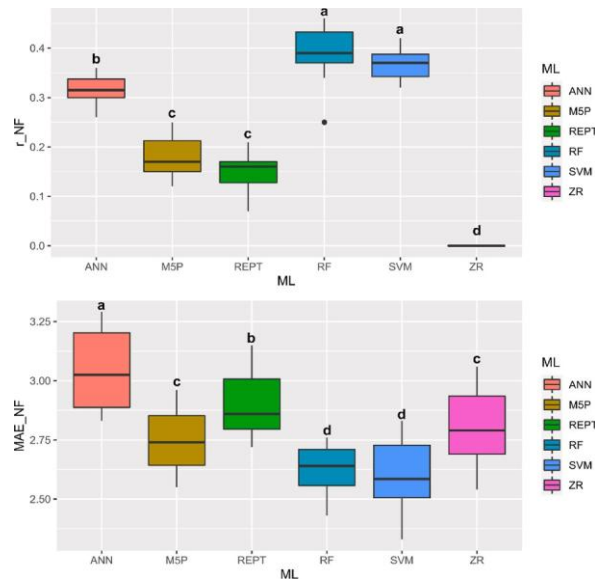


Fig. 7. Mean clustering for Pearson correlation coefficient (r) and mean absolute error (MAE) between the observed values of leaf nitrogen content and the values predicted by different machine learning (ML) algorithms using hyperspectral data as input. Means followed by equal letters in the same figure do not differ by the Scott-Knott test at 5% probability.

This alternative method of diagnosing N is likely to be more advantageous than the conventional method based on chemical analysis in maize. This is because the detection of possible deficiency in the latter only occurs in the reproductive phase, which is the time indicated for leaf collection (de Mello Prado, 2021). Thus, if there is an eventual deficiency of the nutrient it cannot be corrected in time to avoid yield losses, which is a critical point of the conventional method of foliar diagnosis of N in maize.

With increasing N content in maize leaves, spectral reflectance significantly increased in the red-edge region and decreased in the visible region (Fig. 3). When fertility is sufficient, crops are denser and more photosynthetically active, meaning that they have higher absorbance in the visible region and show higher reflectance in the red-edge and near-infrared regions (Liang et al., 2016).

According to Burns et al. (2022), N prediction in maize should consider the growth stage. These authors found that the best stage for detection of maize grain yield was V8, in which the best indices were green CI, GNDVI, and RENDVI. The authors indicate that the relationship between the predictive power of chlorophyll indices and yield increases with the progression of the growth stage. Furthermore, they state that remote sensing data should be acquired as late as possible (*i.e.* late vegetative stages or early reproductive stages) for optimal sensitivity. This is because soil interference may occur at early stages, increasing data variation. At the early stage of plant growth, soil exposition accounts for a large variation in the signal detected by a sensor (Fu et al., 2021).

However, the present study did not use vegetation indices and the soil did not interfere as the spectra were collected in the laboratory. This explains the little variation in the data with similar behavior in the different phenological stages of the reflectance factor spectrum. These curves (hyperspectral) (Fig. 3) indicated higher intensity in the reflectance factor at wavelengths near 550 nm (green band) and 750 nm (red-edge band). The peak of green reflectance was observed at about 550 nm, caused by the strong absorption of some photosynthetic pigments mainly in the blue and red spectrum region and the weak absorption in the green region. The visible spectral region of 400–600 nm had relatively lower reflectance, caused by the strong absorption of photosynthetic pigments (chlorophylls and carotenoids) (Y. Wang et al., 2019).

Fu et al. (2021) reported similar results, demonstrating that the wavelengths that correlate with N status occur mainly in the 400–780 nm region. Furthermore, two clusters of wavebands are close to the chlorophyll absorption peaks in the blue and red regions of the spectrum (490–530 nm and 600–780 nm), pointing to a strong correlation between chlorophyll and N results (Fu et al., 2021).

The absorption peak of chlorophyll in the blue region overlaps with the absorbance for carotenoids (Fu et al., 2021). Carotenoids cannot act in photosynthesis if there is no chlorophyll. The energy that carotenoids produce by absorbing light is transferred to chlorophyll, and it plays its role in photosynthesis. Reflectance increased sharply at approximately 690 nm, peaking at 750 nm. These results corroborate Zhang et al. (2021), who reported higher reflectance from 690 nm, and Sun et al. (2020) who obtained a higher correlation with chlorophyll at the wavelength of 704 nm, reflecting at the red-edge region. Leaf pigments have low reflectance in spectra below 500 nm (Blackburn, 2007). Thus, the lower reflectance at this wavelength in the present research.

Research conducted by Sonobe and Wang (2017) revealed that chlorophyll strongly absorbs light in the blue (400–500 nm) and red (600–700 nm) regions. A study conducted by Che Halin and Azhari (2016) employing the UV-VIs spectra attested that the peak at approximately 615 nm was the spectrum for chlorophyll-b and the peak at 664 nm was the spectrum for chlorophyll *a*. These bands had thus lower reflectance as higher absorbance of light was occurring (Fig. 3).

The present study corroborated these findings. More specifically, our research indicated that these bands were active during Measurement with the spectrometer mainly over the visible (400–700 nm) and red-edge regions (Fig. 3). Other authors, such as Zulfa et al. (2020) found similar findings directly correlating reflectance with chlorophyll content, as the higher the reflectance, the higher the chlorophyll content (Zulfa et al., 2020). This occurred because leaves with high chlorophyll content absorbed less light, increasing reflectance (Feng et al., 2017).

Spectral bands B5, B6 (540 nm) and B13 (690–700 nm) have reflectance ranges that correlate directly with *a*-carotene, *b*-carotene, and chlorophyll synthesis (Jensen, 2009). Such bands were close to the highest doses of N because nitrogen is highly required by the corn crop, playing a fundamental role in the plant's chlorophyll (Sangoi et al. (2015)). In other words, the reflectance of leaves is highly influenced by their nitrogen content. The levels of chlorophyll in plants are extremely sensitive to N supplementation, which significantly interferes with what is reflected by the leaves (Taiz et al., 2015). Another result that compared this interference of N in the reflectance was the proximity of foliar nitrogen content, of the evaluated pigments, and of the B4, B11, and B12 bands (bands also related to the foliar pigments) with dose N3.

Assessing the relationship between nitrogen and pigments and between these and leaf reflectance provides valuable information on the photosynthetic rate of the maize plant using multispectral data (Chou et al., 2020). Being the N essential structural element, deficit or excess alterations cause plant metabolic disturbances. As N is a structural element, the deficiency or excess of N in plants can lead to metabolic disorders that show visible signs and alter reflectance (Gomes et al., 2023). Therefore, adequate nitrogen levels favor photosynthetic pigments, which makes the reflectance in the mentioned bands high.

1.1. Machine learning models

After analyzing the spectra from 350 to 800 nm and considering MAE data, the authors

compared some machine learning algorithms to clarify which models effectively predicted nitrogen, chlorophyll *a+b*, and carotenoid contents. Among these, the random forest algorithm stood out (lower mean MAE and higher mean *r*).

The RF model performed better because its internal structure relies on multiple sets of decision trees. Its algorithm can manipulate many model parameters, reduce estimation bias, and does not present overfitting problems (Jensen, 2009). Recent studies have rated RF as an effective and versatile machine learning method for predicting crop yield as well as chlorophyll and N (Singh et al., 2022; Teodoro et al., 2021; S. Wang et al., 2021). RF has also been shown to be effective and robust in predicting nitrogen concentration in corn in the vegetative growth phase (Yang et al., 2023).

Support vector machine (SVM) is another successful algorithm that has been widely used with a Gaussian kernel function (Biau and Scornet 2016). Previous studies have also applied this model for chlorophyll estimation (Singh et al., 2022; S. Wang et al., 2021). Support vector machine (SVM) is another successful algorithm that has been widely used with a Gaussian kernel function (Costa et al., 2022). The SVM method excels at selecting wavelengths or vegetation indices related to plant physiological characteristics, *i.e.* chlorophyll-related wavelengths that can predict N content. According to L. Li et al. (2020), the SVM regression method showed excellent performance for estimating leaf N content, with $R^2 = 0.823$. (Fig. 6). Such results are similar to those found here regarding the superiority of SVM over other methods; however, our *r* values were lower (*r* = 0.4).

Hyperspectral reflectance can detect physiological signals, and many researchers use it to detect plant stress. Reflectance values at some wavelengths allow assessing the risk of early mortality and nitrogen deficiency or sufficiency even before the plant shows visual symptoms.

Hyperspectral data based on several contiguous wavelengths emerge as a fast, economical, and reliable technique to assess attributes of interest in the maize crop. Previous studies have already demonstrated an adequate prediction of chlorophyll, carotenoid, and nitrogen status with known bands or using vegetation indices by remote sensing techniques (Costa et al., 2022; Osco et al., 2020; Zhou et al., 2019). However, studies that predict nitrogen deficiency and sufficiency with contiguous data from different wavelengths are new and scarce in the literature. Even rare are studies using the six machine learning approaches with N doses at different phenological stages.

Thus, our findings have increased knowledge on the subject, allowing analyses of spectral bands with wavelengths ranging between 350 and 800 nm and using the ratio of reflected radiation in the visible and red-edge bands with machine learning methods to predict N and pigment status. Applying artificial intelligence techniques together with hyperspectral data is an efficient approach for precision agriculture practices, improving maize management and decision-making on nitrogen fertilization.

In this scenario, the optimization of nitrogen fertilization, with a precise and rapid diagnosis, should guide the application or nonapplication of nitrogen, providing the maximum increase in yield and reducing soil nutrient losses. This should increase the sustainability of maize cultivation, which has global implications since this crop grows on all continents. Moreover, N fertilization is also common practice for this crop, given the high plant demand in different regions worldwide.

The study has notable limitations, highlighting the lack of research into near infrared (NIR) and its correlation with the molecular structure of plants. This gap arises from the limitation of the spectral range of the sensor used, restricting the scope of the research to the visible region of the spectrum up to the near infrared. Therefore, future research investigating methods for selecting spectral bands (inputs) can further improve the classification of SVM and ANN, which perform better without correlated input data.

6. Conclusions

Spectral bands 530–560 nm and 690–750 nm are effective wavelengths because the visible region with lower reflectance (530–560 nm) correlates with N uptake, chlorophyll and carotenoid

contents, while the red-edge and near-infrared region correlates with N and chlorophyll contents. The random forest (RF) model outperformed the other algorithms by its higher correlation (r) and lower error (MAE) between predicted and observed values for all variables. Support vector machine (SVM) and RF efficiently predicted nitrogen content. These wavelengths and models can be applied to maize grown in precision agriculture.

The accuracy of this alternative method of N diagnosis is promising and allows this diagnosis to be performed at different vegetative stages. This method is much more flexible than the conventional method that is performed only at the reproductive stage. This has important practical implications for efficient N management early in the crop, which favors maize response. However, the method does not come to replace the traditional one, but to add to it, in order to reduce the time and the effects of contamination on the soil and on the groundwater table. This study will help in further research using these wavelengths sensitive to pigments and N content, in addition to the positive implication of the RF and SVM models. Future perspectives include applying the method in the field using a sensor that can cover wavelengths from the visible to the near-infrared region to verify the spectral behavior upon N limitation.

Ethical

All ethical practices have been followed in relation to the development, writing, and publication of the article entitled “New approach for predicting nitrogen and pigments in maize from hyperspectral data and machine learning models”.

CRedit authorship contribution statement

Bianca Cavalcante da Silva: Conceptualization, Data curation, Investigation, Project administration, Resources, Writing – original draft. **Renato de Mello Prado:** Conceptualization, Project administration, Supervision, Writing – original draft. **Áblio Henrique Rojo Baio:** Data curation, Formal analysis. **Cid Naudi Silva Campos:** Data curation, Formal analysis, Project administration. **Larissa Pereira Ribeiro Teodoro:** Formal analysis, Investigation, Visualization, Writing – review & editing. **Paulo Eduardo Teodoro:** Formal analysis, Visualization, Writing – review & editing. **Dthenifer Cordeiro Santana:** Data curation. **Thiago Feliph Silva Fernandes:** Investigation, Writing – review & editing. **Carlos Antonio da Silva Junior:** Formal analysis, Writing – review & editing. **Elisangela de Souza Loureiro:** Investigation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to thank the Universidade Federal de Mato Grosso do Sul (UFMS), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – Grant numbers 303767/2020-0, and 304979/2022-8, and Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT) TO numbers 88/2021, 07/2022, 318/2022 and 94/ 2023, and SIAFEM numbers 30478, 31333, 32242 and 33111. UNESP (Universidade Estadual Paulista) on behalf of the Postgraduate Agronomy Program (Crop Production); UNESP's Plant Nutrition Study Group (GENPLANT); UFMS (Universidade Federal de Mato Grosso do Sul) – UFMS, and the Cerrado Plant Nutrition Study Group (GECENP). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) – Financial Code 001.

7. References

Abdel-Rahman, E.M., Ahmed, F.B., Ismail, R., 2013. Random forest regression and spectral band selection

for estimating sugarcane leaf nitrogen concentration using EO-1 Hyperion hyperspectral data. *Int. J. Rem. Sens.* 34 (2), 712–728. <https://doi.org/10.1080/01431161.2012.713142>.

Barzin, R., Lotfi, H., Varco, J.J., Bora, G.C., 2021. Machine learning in evaluating multispectral active canopy sensor for prediction of corn leaf nitrogen concentration and yield. *Rem. Sens.* 14 (1), 120. <https://doi.org/10.3390/rs14010120>.

Bataglia, O.C., Furlani, A.M.C., Teixeira, J.P.F., Furlani, P.R., Gallo, J.R., 1983. *Methods of Chemical Analysis of Plants. Boletim Tecnico-Instituto Agronomico. (Brazil). no. 78.*

Belgiu, M., Drăguț, L., 2016. Random forest in remote sensing: a review of applications and future directions. *ISPRS J. Photogrammetry Remote Sens.* 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.

Biau, G., Scornet, E., 2016. A random forest guided tour. *Test* 25 (2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>.

Blackburn, G.A., 2007. Hyperspectral remote sensing of plant pigments. *J. Exp. Bot.* 58 (4), 855–867. <https://doi.org/10.1093/jxb/erl123>.

Blaifi, S., Moulahoum, S., Benkercha, R., Taghezouit, B., Saim, A., 2018. M5P model tree based fast fuzzy maximum power point tracker. *Sol. Energy* 163, 405–424. <https://doi.org/10.1016/j.solener.2018.01.071>.

Burns, B.W., Green, V.S., Hashem, A.A., Massey, J.H., Shew, A.M., Adviento-Borbe, M.A.A., Milad, M., 2022. Determining nitrogen deficiencies for maize using various remote sensing indices. *Precis. Agric.* 23 (3), 791–811. <https://doi.org/10.1007/s11119-021-09861-4>.

Chang, L.-C., Liou, J.-Y., Chang, F.-J., 2022. Spatial-temporal flood inundation nowcasts by fusing machine learning methods and principal component analysis. *J. Hydrol.* 612, 128086. <https://doi.org/10.1016/j.jhydrol.2022.128086>.

Che Halin, D.S., Azhari, A.W., 2016. Synthesis of chlorophyll thin film from noni leaves via dip coating process. *Trans Tech Publ. Mater. Sci. Forum* 857, 142–145. Chen, Shi, Gong, Sun, Chen, Du, et al., 2019. True-color three-dimensional imaging and target classification based on hyperspectral LiDAR. *Rem. Sens.* 11 (13), 1541.

<https://doi.org/10.3390/rs11131541>.

Chou, S., Chen, B., Chen, J., Wang, M., Wang, S., Croft, H., Shi, Q., 2020. Estimation of leaf photosynthetic capacity from the photochemical reflectance index and leaf pigments. *Ecol. Indicat.* 110, 105867. <https://doi.org/10.1016/j.ecolind.2019.105867>.

Coradi, P.C., Lutz, É., dos Santos Bilhalva, N., Jaques, L.B.A., Leal, M.M., Teodoro, L.P.R., 2022. Prototype wireless sensor network and Internet of Things platform for real-time monitoring of intergranular equilibrium moisture content and predict the quality corn stored in silos bags. *Expert Syst. Appl.* 208, 118242. <https://doi.org/10.1016/j.eswa.2022.118242>.

Costa, L., Kunwar, S., Ampatzidis, Y., Albrecht, U., 2022. Determining leaf nutrient concentrations in citrus trees using UAV imagery and machine learning. *Precis.*

Agric. 23 (3), 854–875. <https://doi.org/10.1007/s11119-021-09864-1>.

Coulibaly, S., Kamsu-Foguem, B., Kamissoko, D., Traore, D., 2022. Deep learning for precision agriculture: a bibliometric analysis. *Intelligent Syst. Applications* 16, 200102. <https://doi.org/10.1016/j.iswa.2022.200102>.

da Silva Junior, C.A., Nanni, M.R., Shakir, M., Teodoro, P.E., de Oliveira-Júnior, J.F., Cezar, E., et al., 2018. Soybean varieties discrimination using non-imaging hyperspectral sensor. *Infrared Phys. Technol.* 89, 338–

350. <https://doi.org/10.1016/j.infrared.2018.01.027>.

de Mello Prado, R., 2021. Mineral Nutrition of Tropical Plants. Springer Nature.

de Sousa, D.M.G., Lobato, E., 2004. *Cerrado: correção do solo e adubação*. Brasília, DF: Embrapa Informação Tecnológica. Embrapa, Planaltina, DF.

Della-Silva, J.L., Silva Junior, C. A. da, Lima, M., Ribeiro, R. da S., Shiratsuchi, L.S., Rossi, F.S., et al., 2022. Amazonian species evaluation using leaf-based spectroscopy data and dimensionality reduction approaches. *Remote Sens. Appl. Soc. Environ.* 26, 100742. <https://doi.org/10.1016/j.rsase.2022.100742>.

Egmont-Petersen, M., de Ridder, D., Handels, H., 2002. Image processing with neural networks—a review. *Pattern Recogn.* 35 (10), 2279–2301. [https://doi.org/10.1016/S0031-3203\(01\)00178-9](https://doi.org/10.1016/S0031-3203(01)00178-9).

Feng, H., Chen, G., Xiong, L., Liu, Q., Yang, W., 2017. Accurate digitization of the chlorophyll distribution of individual rice leaves using hyperspectral imaging and an integrated image analysis pipeline. *Front. Plant Sci.* 8, 1238.

Filella, I., Porcar-Castell, A., Munné-Bosch, S., Bäck, J., Garbulsky, M.F., Peñuelas, J., 2009. PRI assessment of long-term changes in carotenoids/chlorophyll ratio and short-term changes in de-epoxidation state of the xanthophyll cycle. *Int. J. Rem. Sens.* 30 (17), 4443–4455. <https://doi.org/10.1080/01431160802575661>.

Fu, Y., Yang, G., Pu, R., Li, Z., Li, H., Xu, X., et al., 2021. An overview of crop nitrogen status assessment using hyperspectral remote sensing: current status and perspectives. *Eur. J. Agron.* 124, 126241. <https://doi.org/10.1016/j.eja.2021.126241>.

2020a. Fujita, H., Fournier-Viger, P., Ali, M., Sasaki, J. (Eds.), 2020a, Trends in Artificial Intelligence Theory and Applications. Artificial Intelligence Practices, vol.

12144, 2020a. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-030-55789-8>.

Fujita, H., Fournier-Viger, P., Ali, M., Sasaki, J., 2020b. Trends in artificial intelligence theory and applications. In: Artificial Intelligence Practices: 33rd International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2020, Kitakyushu, Japan, September 22–25, 2020, Proceedings, vol. 12144. Springer Nature.

Goel, P.K., Prasher, S.O., Landry, J., Patel, R.M., Viau, A.A., Miller, J.R., 2003. Estimation of crop biophysical parameters through airborne and field hyperspectral remote sensing. *Trans. ASAE* 46 (4), 1235.

Gomes, R.L., Sousa, M.C., Campos, F.G., Boaro, C.S.F., de Souza Passos, J.R., Ferreira, G., 2023. Near-infrared leaf reflectance modeling of *Annona emarginata* seedlings for early detection of variations in nitrogen concentration. *J. For. Res.* 34 (1), 269–282.

Gu, Y., Wylie, B.K., Howard, D.M., Phuyal, K.P., Ji, L., 2013. NDVI saturation adjustment: a new approach for improving cropland performance estimates in the Greater Platte River Basin, USA. *Ecol. Indic.* 30, 1–6. <https://doi.org/10.1016/j.ecolind.2013.01.041>.

Guo, Y., Fu, Y., Hao, F., Zhang, X., Wu, W., Jin, X., et al., 2021. Integrated phenology and climate in rice yields prediction using machine learning methods. *Ecol.*

Indicat. 120. <https://doi.org/10.1016/j.ecolind.2020.106935>.

Guo, Y., Chen, S., Li, X., Cunha, M., Jayavelu, S., Cammarano, D., Fu, Y.H., 2022. Machine learning-based approaches for predicting SPAD values of maize using multi-spectral images. *Rem. Sens.* 14 (6). <https://doi.org/10.3390/rs14061337>.

INMET Information. Instituto Nacional de Meteorologia (INMET); 2022. Available online: <http://portal.inmet.gov.br/> (accessed on 28 August 2023). Jensen, J.R., 2009. Remote Sensing of the

Environment: an Earth Resource Perspective 2/e. Pearson Education India.

Kappes, C., da Silva, R.G., Ferreira, V.E.N., 2017. Foliar application of *Azospirillum brasilense* and nitrogen on maize off-season. *Sci. Agraria Paranaensis* 16 (3), 366–373.

Lee, H., Wang, J., Leblon, B., 2020. Using linear regression, random forests, and support vector machine with unmanned aerial vehicle multispectral images to predict canopy nitrogen weight in corn. *Rem. Sens.* 12 (13), 2071. <https://doi.org/10.3390/rs12132071>.

Li, R., Wang, R., Zhang, J., Xie, C., Liu, L., Wang, F., et al., 2019. An effective data augmentation strategy for CNN-based pest localization and recognition in the field.

IEEE Access 7, 160274–160283. <https://doi.org/10.1109/ACCESS.2019.2949852>.

Li, L., Lin, D., Wang, J., Yang, L., Wang, Y., 2020. Multivariate analysis models based on full spectra range and effective wavelengths using different transformation techniques for rapid estimation of leaf nitrogen concentration in winter wheat. *Front. Plant Sci.* 11, 755.

Li, D., Miao, Y., Ransom, C.J., Bean, G. Mac, Kitchen, N.R., Fernández, F.G., et al., 2022a. Corn nitrogen nutrition index prediction improved by integrating genetic, environmental, and management factors with active canopy sensing using machine learning. *Rem. Sens.* 14 (2), 394.

Li, L., Chang, L., Ji, Y., Qin, D., Fu, S., Fan, X., et al., 2022b. Quantification and dynamic monitoring of nitrogen utilization efficiency in summer maize with hyperspectral technique considering a non-uniform vertical distribution at whole growth stage. *Field Crops Res.* 281, 108490. <https://doi.org/10.1016/j.fcr.2022.108490>.

Liang, L., Qin, Z., Zhao, S., Di, L., Zhang, C., Deng, M., et al., 2016. Estimating crop chlorophyll content with hyperspectral vegetation indices and the hybrid inversion method. *Int. J. Rem. Sens.* 37 (13), 2923–2949.

Lichtenthaler, H.K., 1987. [34] Chlorophylls and carotenoids: pigments of photosynthetic biomembranes. Elsevier. *Methods Enzymol.* 148, 350–382.

Marques Ramos, A.P., Prado Osco, L., Elis Garcia Furuya, D., Nunes Gonçalves, W., Cordeiro Santana, D., Pereira Ribeiro Teodoro, L., et al., 2020. A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices. *Comput. Electron. Agric.* 178, 105791. <https://doi.org/10.1016/j.compag.2020.105791>.

Mercier, G., Lennon, M., 2003. Support vector machines for hyperspectral image classification with spectral-based kernels. In: *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No. 03CH37477)*, vol. 1. IEEE, pp. 288–290.

Mi, Z., Zhang, X., Su, J., Han, D., Su, B., 2020. Wheat stripe rust grading by deep learning with attention mechanism and images from mobile devices. *Front. Plant Sci.*

11, 558126.

Najafzadeh, M., Niazmardi, S., 2021. A novel multiple-kernel support vector regression algorithm for estimation of water quality parameters. *Nat. Resour. Res.* 30 (5), 3761–3775. <https://doi.org/10.1007/s11053-021-09895-5>.

Nalepa, J., Kawulok, M., 2019. Selecting training sets for support vector machines: a review. *Artif. Intell. Rev.* 52 (2), 857–900. <https://doi.org/10.1007/s10462-017-9611-1>.

Oivukkamäki, J., Atherton, J., Xu, S., Riikonen, A., Zhang, C., Hakala, T., et al., 2023. Investigating foliar macro- and micronutrient variation with chlorophyll fluorescence and reflectance measurements at the leaf and canopy scales in potato. *Rem. Sens.* 15 (10), 2498. <https://doi.org/10.3390/rs15102498>.

- Osco, L.P., Marques Ramos, A.P., Saito Moriya, É.A., de Souza, M., Marcato Junior, J., Matsubara, E.T., et al., 2019. Improvement of leaf nitrogen content inference in Valencia-orange trees applying spectral analysis algorithms in UAV mounted-sensor images. *Int. J. Appl. Earth Obs. Geoinf.* 83, 101907. <https://doi.org/10.1016/j.jag.2019.101907>.
- Osco, L.P., Junior, J.M., Ramos, A.P.M., Furuya, D.E.G., Santana, D.C., Teodoro, L.P.R., et al., 2020. Leaf nitrogen concentration and plant height prediction for maize using UAV-based multispectral imagery and machine learning techniques. *Rem. Sens.* 12 (19), 3237. <https://doi.org/10.3390/rs12193237>.
- Quinlan Bassar, J.R., 1992. Learning With Continuous Classes, 343–348 (Accessed 23 August 2023).
- Rengasamy, D., Mase, J.M., Kumar, A., Rothwell, B., Torres, M.T., Alexander, M.R., et al., 2022. Feature importance in machine learning models: a fuzzy information fusion approach. *Neurocomputing* 511, 163–174. <https://doi.org/10.1016/j.neucom.2022.09.053>.
- Rossi, F.S., La Scala, N., Capristo-Silva, G.F., Della-Silva, J.L., Teodoro, L.P.R., Almeida, G., et al., 2023. Implications of CO₂ emissions on the main land and forest uses in the Brazilian Amazon. *Environ. Res.* 227, 115729. <https://doi.org/10.1016/j.envres.2023.115729>.
- Safa, M., Martin, K.E., Kc, B., Khadka, R., Maxwell, T.M.R., 2019. Modelling nitrogen content of pasture herbage using thermal images and artificial neural networks. *Therm. Sci. Eng. Prog.* 11, 283–288. <https://doi.org/10.1016/j.tsep.2019.04.005>.
- Sagan, V., Peterson, K.T., Maimaitijiang, M., Sidike, P., Sloan, J., Greeling, B.A., et al., 2020. Monitoring Inland Water Quality Using Remote Sensing: Potential and Limitations of Spectral Indices, Bio-Optical Simulations, Machine Learning, and Cloud Computing.
- Saha, S., Saha, M., Mukherjee, K., Arabameri, A., Ngo, P.T.T., Paul, G.C., 2020. Predicting the deforestation probability using the binary logistic regression, random forest, ensemble rotational forest, REPTree: a case study at the Gumani River Basin, India. *Sci. Total Environ.* 730, 139197. <https://doi.org/10.1016/j.scitotenv.2020.139197>.
- Sangoi, L., Silva, L. M. M. da, Mota, M.R., Panison, F., Schmitt, A., Souza, N. M. de, et al., 2015. DESEMPENHO agrônômico do milho em razão do TRATAMENTO de SEMENTES com azospirillum SP. E da aplicação de doses de nitrogênio mineral. *Rev. Bras. Ciência do Solo* 39, 1141–1150.
- Santana, D.C., Teixeira Filho, M.C.M., da Silva, M.R., Chagas, P. H. M. das, de Oliveira, J.L.G., Baio, F.H.R., et al., 2023a. Machine learning in the classification of soybean genotypes for primary macronutrients' content using UAV–multispectral sensor. *Rem. Sens.* 15 (5), 1457.
- Santana, D.C., Teodoro, L.P.R., Baio, F.H.R., Santos, R. G. dos, Coradi, P.C., Biduski, B., et al., 2023b. Classification of soybean genotypes for industrial traits using UAV multispectral imagery and machine learning. *Remote Sens. Appl.* 29, 100919. <https://doi.org/10.1016/j.rsase.2023.100919>.
- Shah, S.H., Angel, Y., Houborg, R., Ali, S., McCabe, M.F., 2019. A random forest machine learning approach for the retrieval of leaf chlorophyll content in wheat. *Rem. Sens.* 11 (8), 920. <https://doi.org/10.3390/rs11080920>.
- Shao, Y., Zhao, C., Bao, Y., He, Y., 2012. Quantification of nitrogen status in rice by least squares support vector machines and reflectance spectroscopy. *Food Bioprocess Technol.* 5 (1), 100–107. <https://doi.org/10.1007/s11947-009-0267-y>.
- Singh, H., Roy, A., Setia, R.K., Pateriya, B., 2022. Estimation of nitrogen content in wheat from proximal hyperspectral data using machine learning and explainable artificial intelligence (XAI) approach. *Model. Earth Syst. Environ.* 8 (2), 2505–2511. <https://doi.org/10.1007/s40808-021-01243-z>.
- Snousy, M. B. Al, El-Deeb, H.M., Badran, K., Khilil, I. A. Al, 2011. Suite of decision tree-based classification

algorithms on cancer gene expression data. *Egypt. Inform. J.*

12 (2), 73–82. <https://doi.org/10.1016/j.eij.2011.04.003>.

Sonobe, R., Wang, Q., 2017. Towards a universal hyperspectral index to assess chlorophyll content in deciduous forests. *Rem. Sens.* 9 (3), 191.

Su, J., Liu, C., Coombes, M., Hu, X., Wang, C., Xu, X., et al., 2018. Wheat yellow rust monitoring by learning from multispectral UAV aerial imagery. *Comput. Electron. Agric.* 155, 157–166. <https://doi.org/10.1016/j.compag.2018.10.017>.

Su, J., Yi, D., Coombes, M., Liu, C., Zhai, X., McDonald-Maier, K., Chen, W.-H., 2022. Spectral analysis and mapping of blackgrass weed by leveraging machine learning and UAV multispectral imagery. *Comput. Electron. Agric.* 192, 106621. <https://doi.org/10.1016/j.compag.2021.106621>.

Sun, Y., Qin, Q., Ren, H., Zhang, T., Chen, S., 2020. Red-edge band vegetation indices for leaf area index estimation from sentinel-2/MSI imagery. *IEEE Trans. Geosci. Rem. Sens.* 58 (2), 826–840. <https://doi.org/10.1109/TGRS.2019.2940826>.

Taiz, L., Zeiger, E., Møller, I.M., Murphy, A., 2015. *Plant Physiology and Development*. Sinauer Associates Incorporated.

Teodoro, P.E., Teodoro, L.P.R., Baio, F.H.R., da Silva Junior, C.A., dos Santos, R.G., Ramos, A.P.M., et al., 2021. Predicting days to maturity, plant height, and grain yield in soybean: a machine and deep learning approach using multispectral data. *Rem. Sens.* 13 (22), 4632. <https://doi.org/10.3390/rs13224632>.

Van Raij, B., Quaggio, J.A., 2001. Determinação de fósforo, cálcio, magnésio e potássio extraídos com resina trocadora de íons. B. van Raij, JC, andrade, H. Cantarella,

J.A. Quaggio (Eds.), *Análise química para avaliação da fertilidade de solos tropicais*, 189–199.

Wang, Y., Hu, X., Jin, G., Hou, Z., Ning, J., Zhang, Z., 2019. Rapid prediction of chlorophylls and carotenoids content in tea leaves under different levels of nitrogen application based on hyperspectral imaging. *J. Sci. Food Agric.* 99 (4), 1997–2004.

Wang, S., Guan, K., Wang, Z., Ainsworth, E.A., Zheng, T., Townsend, P.A., et al., 2021. Airborne hyperspectral imaging of nitrogen deficiency on crop traits and yield of maize by machine learning and radiative transfer modeling. *Int. J. Appl. Earth Obs. Geoinf.* 105, 102617. <https://doi.org/10.1016/j.jag.2021.102617>.

Wen, P.-F., Wang, R., Shi, Z.-J., Ning, F., Wang, S.-L., Zhang, Y.-J., et al., 2018. Effects of N application rate on N remobilization and accumulation in maize (*Zea mays* L.) and estimating of vegetative N remobilization using hyperspectral measurements. *Comput. Electron. Agric.* 152, 166–181. <https://doi.org/10.1016/j.compag.2018.07.009>.

Wong, C.Y.S., D’Odorico, P., Bhathena, Y., Arain, M.A., Ensminger, I., 2019. Carotenoid based vegetation indices for accurate monitoring of the phenology of photosynthesis at the leaf-scale in deciduous and evergreen trees. *Remote Sens. Environ.* 233, 111407. <https://doi.org/10.1016/j.rse.2019.111407>.

Yang, H., Yin, H., Li, F., Hu, Y., Yu, K., 2023. Machine learning models fed with optimized spectral indices to advance crop nitrogen monitoring. *Field Crops Res.* 293. <https://doi.org/10.1016/j.fcr.2023.108844>.

Yao, X., Huang, Y., Shang, G., Zhou, C., Cheng, T., Tian, Y., et al., 2015. Evaluation of six algorithms to monitor wheat leaf nitrogen concentration. *Rem. Sens.* 7 (11), 14939–14966. <https://doi.org/10.3390/rs71114939>.

Zhang, Y., Hui, J., Qin, Q., Sun, Y., Zhang, T., Sun, H., Li, M., 2021. Transfer-learning-based approach for leaf chlorophyll content estimation of winter wheat from hyperspectral data. *Remote Sens. Environ.* 267, 112724. <https://doi.org/10.1016/j.rse.2021.112724>.

Zheng, Y., Wang, Q., Zhang, X., Yu, J., Li, C., Chen, L., Liu, Y., 2022. Nitrogen and phosphorus retention risk assessment in a drinking water source area under anthropogenic activities. *Rem. Sens.* 14 (9), 2070. <https://doi.org/10.3390/rs14092070>.

Zhou, X., Huang, W., Zhang, J., Kong, W., Casa, R., Huang, Y., 2019. A novel combined spectral index for estimating the ratio of carotenoid to chlorophyll content to monitor crop physiological and phenological status. *Int. J. Appl. Earth Obs. Geoinf.* 76, 128–142. <https://doi.org/10.1016/j.jag.2018.10.012>.

Zulfa, A.W., Norizah, K., Hamdan, O., Zulkifly, S., Faridah-Hanum, I., Rhyma, P.P., 2020. Discriminating trees species from the relationship between spectral reflectance and chlorophyll contents of mangrove forest in Malaysia. *Ecol. Indicat.* 111, 106024. <https://doi.org/10.1016/j.ecolind.2019.106024>.

Capítulo 4 - Considerações finais e perspectivas futuras

A aplicação de sensores hiperespectrais em campo oferece uma abordagem não invasiva para monitoramento nutricional das culturas. Esses sensores capturam dados de diferentes faixas espectrais, que podem ser correlacionados com o conteúdo de nutrientes nas plantas, como o nitrogênio, a clorofila e os carotenoides. Ao integrar esses dados com os valores de referência de nutrientes para as plantas, é possível verificar se os teores estão dentro dos padrões ideais, ou se há deficiências ou excessos, permitindo ajustes na aplicação de fertilizantes. Essa prática pode reduzir o uso excessivo de insumos, promover o manejo mais sustentável e contribuir para a agricultura de precisão.

Ao correlacionar os espectros hiperespectrais com os teores de nutrientes, é possível identificar padrões específicos de absorção de luz que estão associados à quantidade de nutrientes nas plantas. Isso possibilita a criação de modelos de predição que podem ser utilizados em tempo real para estimar com precisão o conteúdo nutricional das plantas, fornecendo informações cruciais para a gestão da fertilização. Essas correlações podem ser usadas para detectar variações nas plantas, o que pode indicar deficiências ou excessos, permitindo um manejo mais adequado e preciso.

Além da estimativa nutricional, é fundamental integrar os dados de produtividade com as medições dos nutrientes obtidas pelos sensores. A produtividade das culturas é fortemente influenciada pelos níveis nutricionais das plantas, portanto, correlacionar os dados de teores de nutrientes com os dados de produtividade pode fornecer informações valiosas sobre a eficiência da fertilização e o impacto das condições nutricionais no crescimento e no rendimento da cultura. Essa integração permitirá ajustes finos na aplicação de fertilizantes e na gestão da cultura, visando não apenas a maximização da produtividade, mas também a sustentabilidade e o uso responsável dos recursos naturais.

Em suma, a integração de sensores hiperespectrais em campo, correlacionado à análise dos teores de nutrientes e aos dados de produtividade, oferece um caminho promissor para otimizar a nutrição das plantas, garantir o uso eficiente de fertilizantes e aumentar a produtividade. Com isso, a precisão e a robustez do modelo aumentam, permitindo uma gestão mais eficiente e sustentável da fertilização e produtividade das culturas.

Ainda faltam trabalhos que aprofundem a aplicação prática de sensores hiperespectrais em campo, especialmente com a correlação entre os teores de nutrientes e os dados de produtividade. Essa lacuna na pesquisa limita o pleno potencial da tecnologia, mas também cria uma oportunidade para novos estudos que podem trazer avanços significativos.

APÊNDICE

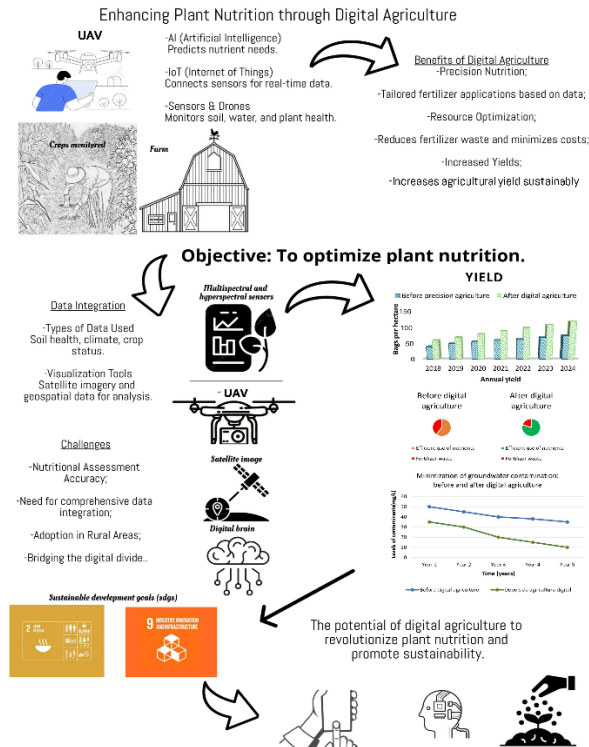


Figura 1. Graphical Abstract (Artigo 1)

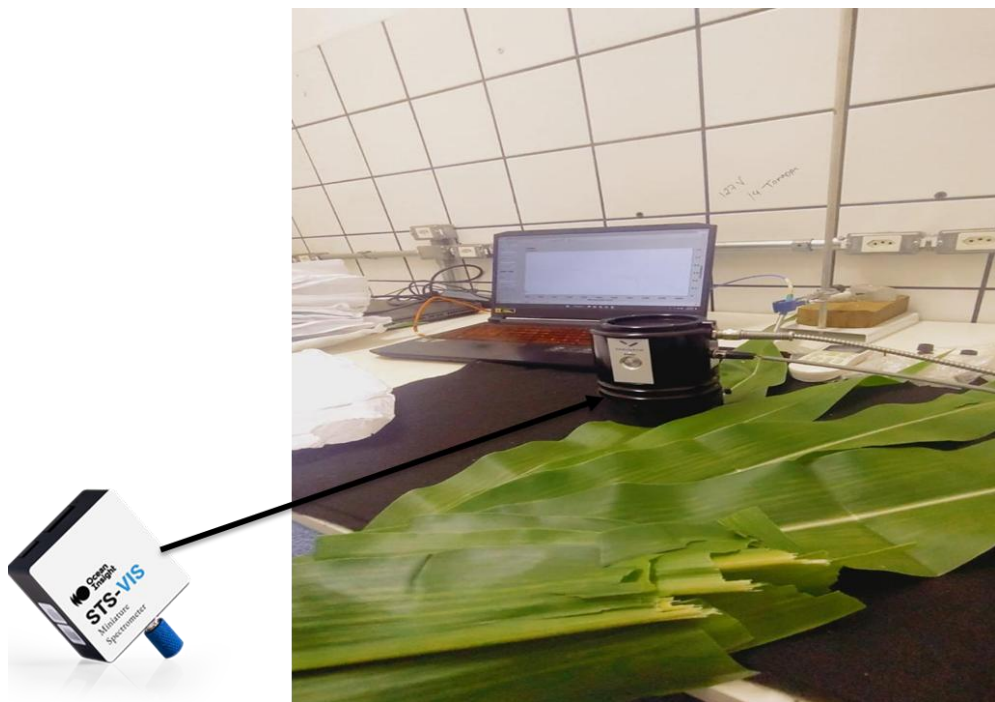


Figura 2. Análise espectral de folhas de milho em laboratório, utilizando o sensor STS-VIS, sob diferentes doses de nitrogênio e em distintos estádios fenológicos.