UNIVERSIDADE FEDERAL DE VIÇOSA

FENOTIPAGEM DE ALTO RENDIMENTO E APRENDIZADO DE MÁQUINA NO MELHORAMENTO DA SOJA

Marco Renan Félix
Doctor Scientiae

VIÇOSA - MINAS GERAIS
2024

MARCO RENAN FÉLIX

FENOTIPAGEM DE ALTO RENDIMENTO E APRENDIZADO DE MÁQUINA NO MELHORAMENTO DA SOJA

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

Orientador: Aluizio Borem de Oliveira

Coorientadores: Tavvs Micael Alves
Leonardo Volpato
Odilon L. de M. Filho
Roberto Fritsche Neto

VIÇOSA - MINAS GERAIS
2024

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T
F316f
2024

Félix, Marco Renan, 1991-
Fenotipagem de alto rendimento e aprendizado de máquina no melhoramento da soja / Marco Renan Félix. – Viçosa, MG, 2024.
1 tese eletrônica (65 f.): il. (algumas color.).

Texto em português e inglês.
Inclui apêndices.
Orientador: Aluízio Borém de Oliveira.
Tese (doutorado) - Universidade Federal de Viçosa, Departamento de Agronomia, 2024.
Inclui bibliografia.
DOI: https://doi.org/10.47328/ufvbbt.2025.730
Modo de acesso: World Wide Web.

1. Soja - Melhoramento genético. 2. Seleção fenotípica. 3. Aprendizado do computador. 4. Sensoriamento remoto. I. Oliveira, Aluízio Borém de, 1959-. II. Universidade Federal de Viçosa. Departamento de Agronomia. Programa de Pós-Graduação em Genética e Melhoramento. III. Título.

CDD 22. ed. 633.342

Bibliotecário(a) responsável: Euzébio Luiz Pinto CRB-6/3317

MARCO RENAN FÉLIX

FENOTIPAGEM DE ALTO RENDIMENTO E APRENDIZADO DE MÁQUINA NO MELHORAMENTO DA SOJA

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Doctor Scientiae*.

APROVADA: 3 de maio de 2024.

Assentimento:

Marco Renan Félix
Autor

Aluizio Borem de Oliveira
Orientador

A Deus, à minha família, Gilmar e Nelma,
e à minha avó Maria Eleuza.

Dedico.

AGRADECIMENTOS

Agradeço, primeiramente, a Deus pela vida, saúde, proteção e força para superar os desafios. Aos meus pais, Gilmar e Nelma, minha base e inspiração, pelo apoio incondicional, pelos valores transmitidos e por acreditarem em cada passo desta jornada. À minha avó materna, Maria Eleuza, pelo afeto, cuidado e presença constante.

À Universidade Federal de Viçosa (UFV) e ao Programa de Pós-graduação em Genética e Melhoramento, pela oportunidade de construir minha trajetória acadêmica em uma das melhores instituições públicas do país e de aprender com professores de excelência na produção científica. Aos professores do programa de Pós-graduação em Genética e Melhoramento da UFV pelos ensinamentos transmitidos.

Ao meu orientador, Professor Dr. Aluízio Borém, pelo apoio constante, pelas orientações valiosas e pela confiança ao longo desta jornada de desafios do doutorado, a quem sou grato.

Aos meus coorientadores, Dr. Odilon Lemos de Mello Filho, pesquisador da Embrapa, profissional exemplar que me proporcionou as condições necessárias para conduzir este trabalho; Professor Dr. Roberto Fritsche Neto, pela oportunidade de realizar um período de intercâmbio nos Estados Unidos, no estado da Louisiana, na Louisiana State University Agricultural Center (LSU AgCenter) e integrar sua equipe, bem como pelos valiosos ensinamentos, conselhos e suporte nas análises desta tese; e Dr. Leonardo Volpato, pelo apoio, pelo conhecimento compartilhado e pelas revisões ao longo do doutorado. Aos membros da banca de qualificação e defesa de tese, Dr. Lucas de Paula Corrêdo e Dr. Tiago Olivoto, pelos conselhos e contribuições para melhoria da minha pesquisa.

Este trabalho foi realizado com o apoio das seguintes agências de pesquisa brasileiras: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Aos amigos que fiz durante minha estadia em Viçosa/MG e jornada na UFV, pela troca de conhecimento e pela ajuda, em especial a Caique, Pedro Henrique e Alexandre, meus agradecimentos. Aos amigos da Unesp – FCAV, Jaboticabal/SP, com agradecimento especial aos amigos da República Xicreti. A todos que, de forma direta ou indireta, contribuíram para a concretização deste trabalho. E a todos os amigos e familiares que sempre me apoiaram e fazem parte da minha história.

Muito Obrigado!!

“The important thing is not to stop questioning. Curiosity has its own reason for existing.”
(Albert Einstein)

RESUMO

FÉLIX, Marco Renan, D.Sc., Universidade Federal de Viçosa, maio de 2024. **FENOTIPAGEM DE ALTO RENDIMENTO E APRENDIZADO DE MÁQUINA NO MELHORAMENTO DA SOJA**. Orientador: Aluizio Borem de Oliveira. Coorientadores: Tavvs Micael Alves, Leonardo Volpato, Odilon Lemos de Mello Filho e Roberto Fritsche Neto.

A soja (*Glycine max* (L.) Merrill) é uma cultura agrícola global de grande importância e tem impulsionado o desenvolvimento de tecnologias de fenotipagem de alto rendimento para otimizar a seleção de genótipos promissores. A integração da fenotipagem de alto rendimento com técnicas avançadas de aprendizado de máquina representa um avanço significativo na eficiência dos programas de melhoramento genético. Essa inovação utiliza índices de vegetação (IVs) para predizer a produtividade de grãos e o ciclo da cultura, uma característica relacionada à maturidade, geneticamente controlada e influenciada pelo fotoperíodo. Essa característica é essencial para classificar genótipos de soja. Assim, novas tecnologias são necessárias para otimizar o trabalho dos melhoristas na etapa de seleção, como o aumento da acurácia, da precisão e do rendimento operacional nas avaliações. Este estudo teve como objetivo avaliar o uso de um sistema aéreo não tripulado (VANT), equipado com um sensor multiespectral e RGB, para analisar características agronômicas de maturidade e produtividade de grãos em genótipos de soja, utilizando modelos de aprendizado de máquina e IVs derivados de bandas espectrais. Os IVs NDVI (Normalized Difference Vegetation Index), GNDVI (Green Normalized Difference Vegetation Index), NDRE (Normalized Difference Red Edge Index), e NGRDI (Normalized Green Red Difference Index) foram gerados a partir de mapas de refletância no software Pix4D®. As características avaliadas incluíram o número de dias da semeadura até o florescimento e à maturidade (estágio R8), a produtividade de grãos (kg ha^{-1}) e os IVs correlacionados. Foram utilizadas linhagens da EMBRAPA para este estudo. A análise, realizada com modelos mistos e um modelo espacial de linhas e colunas no pacote SpATS, revelou correlações significativas entre os IVs e as características avaliadas. Além disso, os resultados aprimoraram as estratégias de seleção indireta, como a escolha do melhor dia de voo considerando o estádio fenológico da cultura. Os resultados sugerem a viabilidade dos IVs, destacando a eficácia do índice RGB NGRDI, que se mostrou comparável aos sensores multiespectrais NDVI e GNDVI. O índice RGB demonstrou alta precisão na predição da maturidade ($R^2 > 0,9$) e viabilidade econômica para a predição de genótipos de soja. A aplicação de modelos de aprendizado de máquina, como XGBoost e Random Forest, obteve bons resultados nos

problemas de classificação dos grupos de maturidade e predição do ciclo da cultura, com predições mais precisas para a maturidade aos 90, 95 e 100 DAS, correspondendo aos estádios fenológicos R5, R6 e R7. A inclusão do aprendizado profundo com o modelo de redes neurais convolucionais (CNN) explorou a classificação das classes de maturidade, evidenciando o potencial do aprendizado profundo para futuras otimizações. Portanto, este estudo contribui para o avanço do melhoramento genético, oferecendo novas perspectivas para a seleção otimizada de genótipos de soja e ressaltando a importância de abordagens multidisciplinares para maximizar o ganho genético.

Palavras-chave: índices de vegetação; imagens de drone; maturidade de plantas; fenômica; *Glycine max* ; aprendizado de máquina

ABSTRACT

FÉLIX, Marco Renan, D.Sc., Universidade Federal de Viçosa, May, 2024. **HIGH-THROUGHPUT PHENOTYPING AND MACHINE LEARNING IN SOYBEAN BREEDING**. Adviser: Aluizio Borem de Oliveira. Co-advisers: Tavvs Micael Alves, Leonardo Volpato, Odilon Lemos de Mello Filho and Roberto Fritsche Neto.

Soybean (*Glycine max* (L.) Merrill) is a major global crop and has driven the development of high-throughput phenotyping technologies to optimize the selection of promising genotypes. The integration of high-throughput phenotyping with advanced machine learning techniques represents a significant advance in the efficiency of breeding programs. This innovation uses vegetation indices (VIs) to predict grain yield and the crop cycle, a maturity-related trait that is genetically controlled and influenced by photoperiod. This trait is essential for classifying soybean genotypes. Thus, new technologies are needed to optimize breeders' work in the selection stage, including improvements in accuracy, precision, and operational performance in evaluations. This study aimed to evaluate the use of an unmanned aerial system (UAS), equipped with a multispectral and RGB sensor, to analyze agronomic traits such as maturity and grain yield in soybean genotypes, using machine learning models and vegetation indices (VIs) derived from spectral bands. The vegetation indices: NDVI (Normalized Difference Vegetation Index), GNDVI (Green Normalized Difference Vegetation Index), NDRE (Normalized Difference Red Edge Index), and NGRDI (Normalized Green Red Difference Index) were generated from reflectance maps using Pix4D® software. The evaluated traits included the number of days from sowing to flowering and maturity (R8 stage), grain yield (kg ha^{-1}), and their correlated VIs. EMBRAPA breeding lines were used in this study. The analysis, conducted using mixed models and a row-column spatial model in the SpATS package, revealed significant correlations between the VIs and the evaluated traits. Additionally, the results improved indirect selection strategies, such as identifying the optimal flight day based on the crop's phenological stage. The findings support the viability of the VIs, highlighting the effectiveness of the RGB-derived NGRDI index, which performed comparably to multispectral indices like NDVI and GNDVI. The RGB index showed high accuracy for predicting maturity ($R^2 > 0.9$) and proved economically viable for predicting soybean genotypes. Machine learning models such as XGBoost and Random Forest yielded strong results in classifying maturity classes and predicting the crop cycle, with more accurate predictions for maturity at 90, 95, and 100 DAS, corresponding to phenological stages R5, R6, and R7. Incorporating deep learning via a convolutional neural network

(CNN) model further explored the classification of maturity classes, demonstrating the potential of deep learning for future optimization efforts. Therefore, this study contributes to the advancement of plant breeding, offering new perspectives for the optimized selection of soybean genotypes and reinforcing the importance of multidisciplinary approaches to maximize genetic gain.

Keywords: vegetation indices; drone imagery; plant maturity; phenomics; *Glycine max;* machine learning

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1. INTRODUÇÃO GERAL

A soja (*Glycine max* (L.) Merril) tornou-se uma das principais *commodities* mundiais no agronegócio, especialmente pelo seu elevado teor de proteína, importante para a alimentação animal, e pelo seu teor de óleo para a indústria. No Brasil, atualmente o cultivo de soja na primavera-verão com a sucessão do milho na segunda safra é, notadamente, uma das melhores opções para o sistema de produção de culturas anuais. Atualmente a produção estimada está em 160,17 milhões de toneladas, com produtividade em 3,535 kg ha^{-1}, em área de 45,309 milhões de hectares (CONAB, 2023). A projeção para a produção global de soja na safra 2023/24 é estimada para 398,88 milhões de toneladas (USDA, 2023).

O melhoramento genético da soja é um processo contínuo, voltado ao desenvolvimento de cultivares adaptadas às diferentes condições edafoclimáticas, sendo a seleção de genótipos superiores uma das etapas essenciais. Nesse contexto, a produtividade é o principal foco, devido ao seu impacto direto na viabilidade econômica da cultura. Além disso, a seleção de genótipos precoces é crucial para viabilizar a sucessão com o milho na safrinha. É igualmente importante considerar os fatores abióticos, como temperatura, luz, água e ph do solo, que influenciam a precocidade e a adaptação das cultivares, assegurando que o aumento da produtividade ocorra sem comprometer a sustentabilidade ambiental (Silva *et al.,* 2017).

A seleção de genótipos precoces também desempenha um papel fundamental, não apenas facilitando a implementação da safrinha de culturas como o milho, sob condições ideais de umidade, mas também contribuindo para a rotação de culturas. Essa prática é essencial para manter a saúde do solo, controlar pragas e doenças e reduzir o uso de insumos químicos, fortalecendo a sustentabilidade e resiliência do sistema agrícola.

Recentemente, os veículos aéreos não tripulados (VANTs), equipados com sistemas de sensoriamento remoto, incluindo câmeras e sensores, tornaram-se uma ferramenta indispensável para o levantamento de dados com alta resolução espaço-temporal (Zhang *et al.,* 2019). Esses dispositivos são um dos principais componentes da Agricultura Inteligente, também denominada de "Agricultura 4.0", que pode ser entendida como um conjunto de tecnologias de comunicação, informação e análise à disposição dos produtores rurais. Entre essas tecnologias, destacam-se a conectividade, que permite a integração e o monitoramento em tempo real das operações agrícolas, além de plataformas digitais, softwares, sistemas globais de posicionamento por satélite, sensoriamento remoto e sensores nos campos, que são amplamente utilizados para otimizar a produção (Bolfe *et al.*, 2020, Queiroz *et al.,* 2022).

Nesse contexto, a Agricultura 4.0 surge como uma importante ferramenta para suprir os desafios da produção agrícola em termos de produtividade, impacto ambiental, segurança alimentar e sustentabilidade, pois permite entender a complexidade e a imprevisibilidade dos sistemas agrícolas, monitorando e analisando continuamente os fenômenos durante o desenvolvimento do cultivo (Kamilaris e Prenafeta-Boldú, 2018). Além disso, o sensoriamento remoto e o uso de veículos aéreos não tripulados (VANTs) para fenotipagem de alto rendimento têm se consolidado como técnicas fundamentais para o monitoramento em larga escala com precisão das culturas ao longo de seu desenvolvimento.

Sendo assim, observações em larga escala podem ser realizadas por meio do sensoriamento remoto, com utilização de satélites, veículos aéreos ou equipamentos proximais, com a vantagem de não ser um método destrutivo. Com isso, esse método se tornou uma importante área de pesquisa agrícola, pois utiliza de técnicas inteligentes de análise de dados para interpretação em diversos cenários, como classificação de características de estresse biótico e abiótico (Singh *et al.,* 2016).

A fenotipagem de alto rendimento aplicada em programas de melhoramento de soja com o uso de plataformas e processamento de imagens está progressivamente sendo utilizada. Entre os métodos já aplicados para coleta de dados estão a utilização de Satélites e VANTs, que conseguem capturar imagens de resolução adequada em áreas extensas, permitindo a extração de IVs associados a correções e algoritmos, podendo se correlacionar com várias características morfológicas e fisiológicas das plantas (Maimaitijiang *et al.,* 2019).

Grande parte dos pesquisadores está adaptando e desenvolvendo métodos fenômicos para apoiar programas de melhoramento. Para atingir tal objetivo já se verificou que é necessário conciliar conhecimentos das ciências biológicas, informática, matemática e engenharia (Li, Zhang e Huang, 2014), principalmente para minimizar o viés individual obtido pelas mensurações manuais e garantir acurácia e precisão nas medições (Cobb *et al.,* 2013; Perich *et al.,* 2020).

A aplicação da fenotipagem em estudos de maturidade de soja sugere que há flexibilidade na otimização de um *pipeline* de imagens baseado em VANT e pode auxiliar na seleção de genótipos (Volpato *et al.,* 2021). Em relação à produtividade, estudos recentes, como o de Santana *et al.,* (2022), demonstraram a viabilidade de selecionar genótipos dentro dos programas de melhoramento genético com base em IVs e comprimentos de onda específicos, sendo valores pontuais de reflexão ou absorção de luz em faixas espectrais específicas (vermelho, azul, infravermelho etc.), sem combinar essas bandas em índices. Essa abordagem

é particularmente promissora para identificar genótipos que combinem precocidade e alta produtividade, alinhando-se aos objetivos de otimizar o rendimento das culturas.

Dada a importância da predição da maturidade relativa e produtividade de genótipos da soja para os programas de melhoramento, a utilização de técnicas de análise de imagens e aprendizado de máquina têm se destacado como uma metodologia eficiente a favor da pesquisa agrícola. As técnicas de inteligência artificial possuem a capacidade de aprendizado e generalização, sendo aplicáveis a problemas não linearmente separáveis e não requerendo quaisquer pressuposições dos dados (Cruz e Nascimento, 2018). No entanto, uma implantação robusta, com avaliações em campo das características de desenvolvimento da soja em curtos períodos e utilização de dados de sensoriamento remoto, aliados ao aprendizado de máquina, pode ser um método chave que possibilitaria a visualização detalhada do desenvolvimento da cultura.

Desta forma, este trabalho pretendeu contribuir com os programas de melhoramento genético de soja por meio da aplicação de modelos de aprendizado de máquina na avaliação da fenotipagem de alto rendimento, visando à caracterização de ciclo e seleção de genótipos mais produtivos em larga escala. A contribuição se dá pela capacidade desses modelos de processar grandes volumes de dados fenotípicos coletados em campo por sensores de alta precisão, permitindo identificar padrões e aprimorar a tomada de decisões nos programas de melhoramento. Assim, espera-se que os resultados obtidos possam auxiliar no avanço das pesquisas em melhoramento genético, proporcionando novas ferramentas para otimizar a seleção de genótipos.

REFERÊNCIAS

BOLFE, E. L. Precision and Digital Agriculture: Adoption of Technologies and Perception of Brazilian Farmers. **Agriculture**, [*S. l.*], v. 10, n. 12, article 653, 2020.

COBB, J. N.; DECLERCK, G.; GREENBERG, A.; CLARK, R.; MCCOUCH, S. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement. **Theoretical and Applied Genetics**, [*S. l.*], v. 126, n. 4, p. 867-887, 2013.

CONAB. Companhia Nacional de Abastecimento. Boletim da Safra de Grãos. **Acompanhamento da Safra Brasileira de Grãos**, Brasília, v. 11, n. 3, 3° levantamento, safra 2023/24, p. 1-137, dez, 2023. Disponível em: https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos. Acesso em: 15 dez. 2023.

CRUZ, C. D.; NASCIMENTO, M. **Inteligência computacional aplicada ao melhoramento genético.** Viçosa-MG: Editora UFV, 2018. 414 p.

KAMILARIS, A.; PRENAFETA-BOLDÚ, F. X. Deep learning in agriculture: A survey. **Computers and Electronics in Agriculture**, [*S. l.*], v. 147, p. 70-90, 2018.

LI, L.; ZHANG, Q.; HUANG, D. A Review of Imaging Techniques for Plant Phenotyping. **Sensors [Basel, Switzerland]**, v. 14, n. 11, p. 20078-20111, 2014.

MAIMAITIJIANG, M. SAGAN, V.; SIDIKE, P.; MAIMAITIYIMING, M.; HARTLING, S.; PETERSON, K. T.; MAW, M. J. W.; SHAKOOR, N.; MOCKLER, T.; FRITSCHI, F. B. Vegetation Index Weighted Canopy Volume Model (CVMVI) for soybean biomass estimation from Unmanned Aerial System-based RGB imagery. **ISPRS Journal of Photogrammetry and Remote Sensing**, [*S. l.*], v. 151, p. 27-41, 2019.

PERICH, G.; HUND, A.; ANDEREGG, J.; ROTH, L.; BOER, M. P.; WALTER, A.; LIEBISCH, F. Assessment of Multi-Image Unmanned Aerial Vehicle Based High-Throughput Field Phenotyping of Canopy Temperature. **Frontiers in Plant Science**, [*S. l.*], v. 11, p. 1-17, 2020.

QUEIROZ, D. M.; VALENTE, D. S. M.; PINTO, F. A. C.; BORÉM, A. **Agricultura Digital.** 2. ed. São Paulo: Oficina de Textos, 2022, 224 p.

SANTANA, D.C.; CUNHA, M.P.; SANTOS, R.G. High-throughput phenotyping allows the selection of soybean genotypes for earliness and high grain yield. **Plant Methods,** [*S. l.*], v. 18, n. 13, p. 1-11, 2022.

SILVA, F. C. S.; SEDIYAMA, T.; OLIVEIRA, R. C. T.; BORÉM, A.; SILVA, F. L.; BEZERRA, A.R.G.; SILVA, A.F. Importância econômica e evolução do melhoramento. In: SILVA, F.L.; BORÉM, A.; SEDIAYAMA, T.; LUDKE, W. (Eds). **Melhoramento da Soja**. Viçosa-MG: Editora da UFV, 2017.

SINGH, A.; GANAPATHYSUBRAMANIAN, B.; SINGH, A.K.; SARKAR, S. Machine Learning for High-Throughput Stress Phenotyping in Plantas. **Trends in Plant Science**, [*S. l.*], v. 21, n. 2, p. 110-124, 2016.

USDA - **United States Department of Agriculture**. Foreign Agriculture Service, (2023). Crop Explorer. Accessed in: https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?cropid=2222000 . Acesso em: 2 Dez. 2023.

VOLPATO, L.; DOBBELS, A.; BORÉM, A.; LORENZ, A. Optimization of temporal UAS-based imagery analysis to estimateplant maturity date for soybean breeding. **The Plant Phenome Journal**, [*S. l.*], v. 4, n. 1, p. e20018, 2021.

ZHANG, L., ZHANG H., NIU Y., HAN W. Mapping maize water stress based on UAV multispectral remote sensing. **Remote Sensing,** [*S. l.*], v. 11, n. 6, p. 605-629, 2019.

2. ARTIGO

HARNESSING HIGH-THROUGHPUT PHENOTYPING AND MACHINE LEARNING TO PREDICT MATURITY AND OPTIMIZE GRAIN YIELD IN SOYBEAN BREEDING

2.1 Abstract

In plant breeding, high-throughput phenotyping (HTP) is a key process for automating the large-scale evaluation of phenotypes using non-destructive methods with high temporal and spatial resolution. As a strategy to support soybean breeding programs, this study aimed to evaluate the performance of machine learning (ML) approaches, such as XGBoost and Random Forest, and deep learning (DL) using convolutional neural networks (CNN), applied to classification and regression tasks for predicting maturity and optimizing grain yield in soybean genotypes. We used an unmanned aerial system (UAS) equipped with multispectral and RGB sensors to collect spectral data and generate four vegetation indices (VIs): NDVI (Normalized Difference Vegetation Index), GNDVI (Green Normalized Difference Vegetation Index), NDRE (Normalized Difference Red Edge Index), and NGRDI (Normalized Green Red Difference Index). The dataset was collected over seven flights at 35, 60, 70, 90, 95, 100, and 110 days after sowing (DAS), covering the main crop phenological stages, from the vegetative phase to maturity. The experimental design included a row-and-column spatial model incorporating different relative maturity groups in soybean breeding trials. Agronomic traits, including days to flowering, days to maturity, and grain yield, were collected as ground-truth data. Maturity showed its strongest correlation ($r = 0.95$) with the NDVI and NGRDI indices at 95 DAS, confirming the indirect selection efficiency (ISE) for this trait at stages R6 and R7. The highest prediction efficiency for grain yield was observed between the R4 and R5 stages (60 and 70 DAS). The ML models were highly accurate in classifying maturity classes, achieving accuracies of 80% (NDVI) and 75% (NGRDI) at 95 DAS, and an R^2 value greater than 0.9 in predicting maturity. The choice of the optimal flight day should consider the crop's phenological stage, the trait of interest, and the VIs evaluated. These results suggest that more cost-effective RGB sensors (suitable for the NGRDI index) can achieve comparable performance to multispectral sensors (suitable for the NDVI index), highlighting the importance of optimized phenotyping strategies for precision agriculture.

Keywords: Vegetation indices. Drone imagery. Plant maturity. Phenomics. *Glycine max*.

2.2 Introduction

Modern agriculture faces the challenge of meeting the growing global demand for food sustainably, requiring innovations that maximize agricultural efficiency and productivity. In the context of soybean (Glycine max L.), a key crop for global food security and the biofuel industry, the selection of superior genotypes is fundamental to improving yield and adaptability to climate change (Dilawari et al., 2022; Silva et al., 2022). Thus, soybean breeding programs are focusing on developing cultivars that combine high grain yield and early maturity for Brazil's growing regions, a continental-scale country and global leader in soybean production.

Conventional breeding methods rely heavily on phenotypic evaluations, which are labor-intensive, time-consuming, destructive, and influenced by the subjective experience of breeders (Araus et al., 2018). Therefore, high-throughput phenotyping (HTP has emerged as a strategy to enhance the accuracy and precision of phenotypic estimates, thereby optimizing yield at various biological levels. The goal is to reduce costs and manual effort through automation, remote sensing, effective data integration, and experimental planning in the application of phenomics (Fritsche-Neto; Borém, 2015). In this context, image-based analysis emerges as a timely opportunity for pattern recognition and classification. For instance, images collected by unmanned aerial vehicles (UAVs) offer high spatial resolution, enabling in-depth interpretation of the collected information (Amaral et al., 2022).

Recently, HTP platforms, such as those equipped with multispectral, hyperspectral and thermal cameras, as well as LiDAR sensors and RGB imaging systems, coupled with machine learning (ML) methods, have emerged as innovative technologies in agriculture, providing non-destructive and efficient methods for the accurate and rapid assessment of phenotypic traits on a large scale (Gill et al., 2022; Herrero-Huerta et al., 2020). This approach facilitates the analysis and interpretation of large datasets, offering an efficient alternative for extracting the necessary information to support decision-making. Additionally, technological advances in more efficient and non-destructive phenotyping tools are improving crop monitoring. Consequently, making it possible to predict complex traits with reduced cost and high efficiency from the earliest stages of growth, where limitations in phenotyping efficiency are increasingly recognized as a major constraint to genetic progress in breeding programs (Araus et al., 2018).

Machine learning enables the analysis of HTP's extensive datasets, facilitating the discovery of intricate patterns, predicting phenotypic traits from spectral image data, and exploring alternative data collection methods (Teodoro et al., 2021). Recent studies have demonstrated the applicability of ML and HTP technologies in agriculture and plant breeding,

with soybean gaining prominence gaining prominence in applications such as grain maturity prediction (Narayanan et al., 2019; Trevisan et al., 2020; Volpato et al., 2021), grain yield (Santana et al., 2022), plant stress (Singh et al., 2018), and fresh biomass (Ranđelović et al., 2023). Approaches such as spectral indices (Tayade et al., 2022) and multitemporal phenology mapping (Silva Junior et al., 2020; Zhong et al., 2016) reinforce the use of these techniques in soybean. Applications in other crops, such as sugarcane in a phenotyping study (Barbosa Júnior et al., 2022), Plant disease resistance in sorghum (Galli et al., 2020), and coffee for yield estimation (Abreu Júnior et al., 2022), also demonstrate the versatility of these technologies. These examples illustrate the broad potential of ML techniques in agronomic contexts, contributing to advances in soybean genetic breeding.

Moreover, advances in deep learning algorithm (DL) technologies enable computational models composed of multiple processing layers, such as convolutional neural networks (CNN), recurrent neural networks (RNN), and "you only look once" (YOLO), to revolutionize the ability to analyze and interpret complex data in various areas, including agriculture (Attri et al., 2023). Specifically, the application of DL in plant breeding has great potential to optimize genotype selection, early identification of plant stress (Singh et al., 2018), estimation of agronomic traits such as grain yield (Lu et al., 2022), physiological maturity (Moeinizade et al., 2022), and identification of flower patterns (Zhu, et al., 2022). These innovative approaches enhance the accuracy and efficiency of breeding programs, supporting data-driven decision-making to optimize agricultural management practices.

Although traditional phenotypic evaluation methods have been valuable, they present limitations in efficiency and scalability, especially in breeding programs. The growing demand for more productive and sustainable agriculture highlights the need for faster, more accurate, and non-destructive tools. In this context, HTP combined with ML has emerged as a promising solution to overcome these challenges. Based on this, we hypothesized that these technologies could optimize genotype evaluation and facilitate the efficient selection of cultivars with desirable agronomic traits.

In this scenario, HTP combined with ML and DL techniques represents a promising solution for predicting flowering stage, maturity, and grain yield—an unprecedented approach in soybeans. Therefore, in this study, we evaluated whether HTP, using both multispectral and RGB imagery, is effective in predicting these traits. Specifically, we investigated the potential of vegetation indices (VIs) derived from spectral data and applied ML and DL algorithms to develop predictive models. We also evaluated the feasibility of using these spectral traits for indirect selection, highlighting their relevance to soybean breeding programs.

2.3 Material and methods

2.3.1 Experimental design and field management

The study was conducted at the experimental field of the Brazilian Agricultural Research Corporation (EMBRAPA) Rice and Beans, located in the municipality of Santo Antônio de Goiás - GO, Brazil (16º 29' 42.14" S, 49º 17' 58.36" W), with data collected by manual and drone-based remote evaluation throughout the soybean crop cycle.

The experimental plots consisting of 4 rows of 5 meters, spaced 0.54 meters apart, at a planting density of 13 seeds per meter sown in November 2021, and with only the useful area of the plot being evaluated, considered as the two central rows. The experimental design was a randomized block design (RBD), analyzed through mixed models in a row and column spatial model. Evaluations of agronomic and spectral traits (remote sensing) were carried out throughout the soybean development cycle. The evaluated experiments were part of Value for Cultivation and Use (VCU) trials within EMBRAPA's soybean breeding program.

The trials were structured to evaluate different relative maturity groups, involving breeding lines carrying RR® (glyphosate-resistant) and Bt® (Bacillus thuringiensis) biotechnologies. Four experimental blocks were used, divided into two third-year Preliminary Evaluation (AP3) blocks and two Final Competition (FC) blocks. In total, the experiments involved 1,004 experimental plots, organized according to the type of biotechnology and the number of replications.

In the AP3 experimental blocks, four experiments were conducted per biotechnology type (Bt and RR), each with three replications. In the FC experimental blocks, two experiments were conducted per biotechnology type (Bt and RR), each with four replications. Thus, the trials were defined by combining each experimental block with its respective experiments, resulting in a total of four experimental blocks and twelve experiments, which were concatenated and treated as a single 'trial' in the statistical model. Treatments were pre-established based on EMBRAPA's classification, including commercial cultivars used as controls or reference genotypes (Table 1).

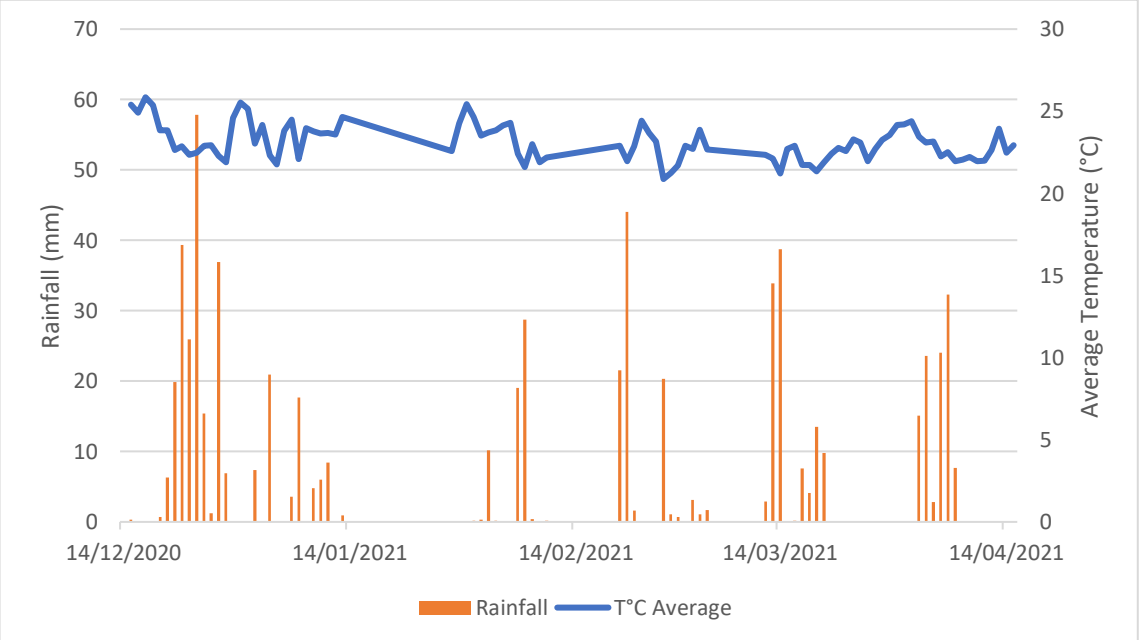
Table 1 - Experimental block design

Biotechnologies	Experimental blocks	Replications	Controls Genotypes	Genotypes	Number of plots	Number of flights
AP3RR	4	3	5	112	372	7
AP3Bt	4	3	3	112	372	7
CFRR	2	4	8	20	116	7
CFBt	2	4	8	28	144	7

Source: Authors (2024).
Experimental area: EMBRAPA Rice and Beans, Santo Antônio de Goiás - GO, Brazil.

Sowing was performed mechanically using a plot seeder. Crop cultivation followed the standard technical recommendations for management. The experiment was conducted under rainfed conditions, with supplementary irrigation provided by a central pivot system. Climatic conditions observed during the experiment are presented in Figure 1.

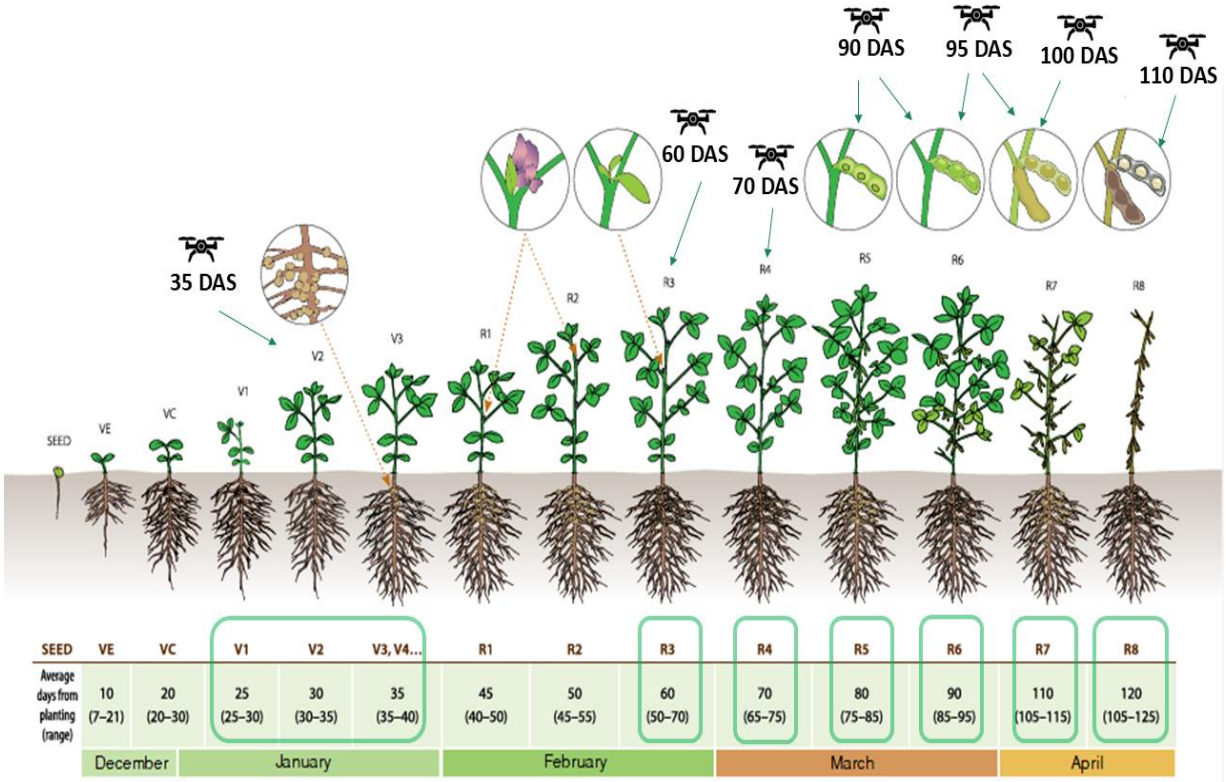
Figure 1 - Rainfall and average temperature during the experiment period



Source: Author (2024).

Emergence was recorded when 50% of the plants in each row displayed cotyledons above the soil surface. Field evaluations were conducted throughout the crop cycle, with grain maturity determined at reproductive stage R8, as described by Fehr and Caviness (1977). Phenological stages and corresponding image acquisition dates are presented in Figure 2, identified by a drone icon and matched to the respective developmental stages.

Figure 2 - Soybean growth stages from emergence (VE) to physiological maturity (R8), according to Fehr and Caviness (1977). Drone imagery was acquired at 35, 60, 70, 90, 95, 100, and 110 days after sowing (DAS) and aligned with the corresponding phenological stages



Source: Adapted from Manitoba Pulse e Soybean Growers (2018).
Caption: DAS: Days After Sowing.

Soil samples were collected at a depth of 0–20 cm before the experiment for physical and chemical characterization (Table S1).

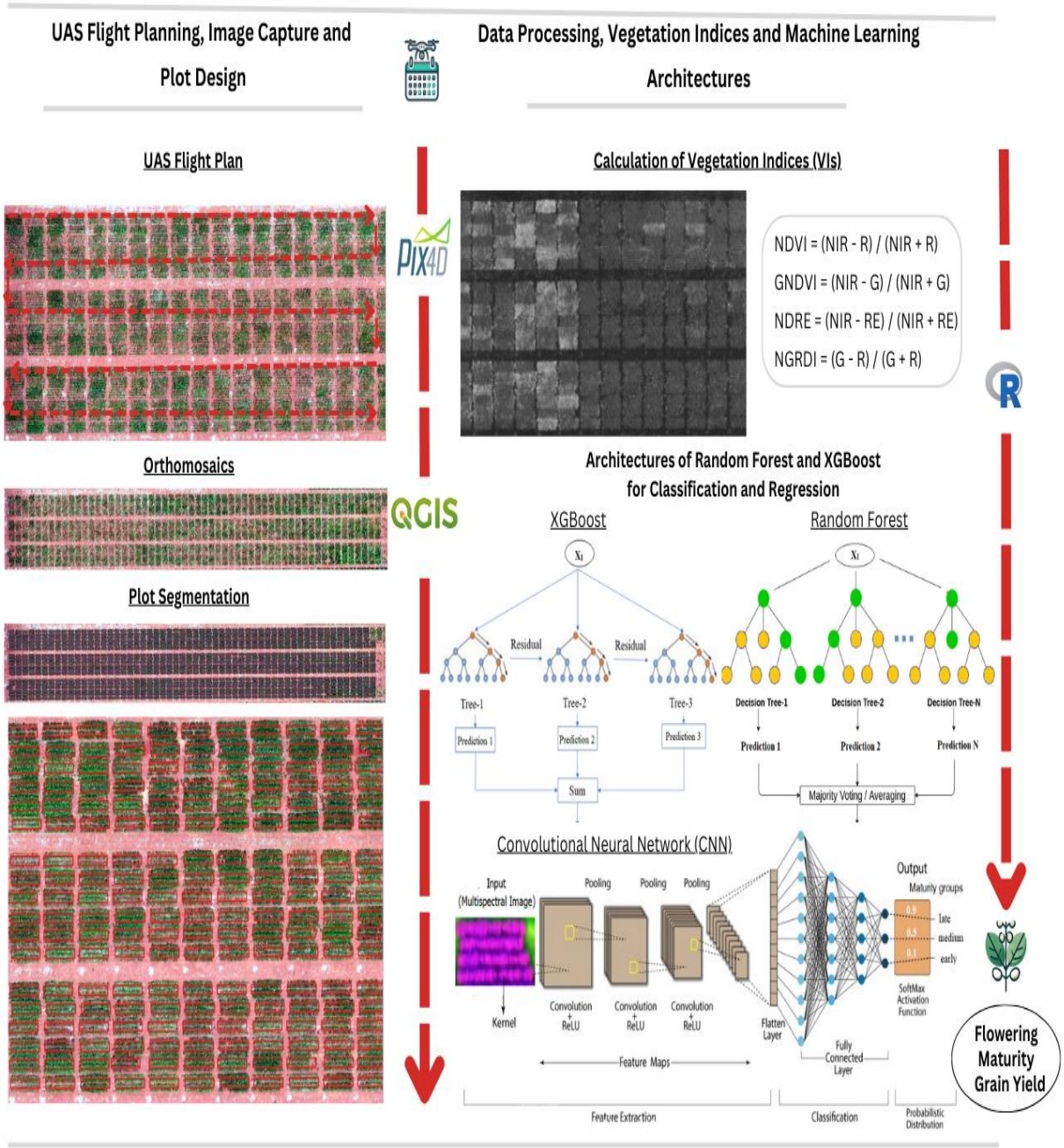
2.3.2 Spectral data collection

Spectral data was acquired using a multirotor UAV (Parrot Bluegrass Fields, Paris, France) equipped with a multispectral passive sensor (Parrot Sequoia, Paris, France). The system consists of a downward-looking multispectral camera and an upward-facing irradiance sensor, which records incoming solar radiation in the same spectral range, enabling radiometric self-calibrations of the camera. The flight plan was standardized at a height of 50 meters above ground, taking into account the terrain's slope, with a ground sample distance (GSD) of 5.6 cm per pixel. During the flights, a 70% overlap was configured for both frontal and lateral views, ensuring adequate coverage and enabling the collection of high-quality data for analysis. The positional accuracy of the orthomosaics was checked using four ground control points (GCPs).

Flights were performed at 35, 60, 75, 90, 95, 100, and 110 DAS, constantly between 11 am and 1 pm, to minimize atmospheric effects and correct possible issues with sensor light sensitivity during spectral image processing (Iqbal et al., 2018). Radiometric calibration was performed using a calibration panel before each flight, adjusting the camera sensor's sensitivity to ensure uniform and reliable responses. This procedure ensured that reflectance values were comparable across flights performed under different illumination conditions. The multispectral sensor captured reflectance data in four bands, namely Green (550 ± 40 nm), Red (660 ± 40 nm), RedEdge (735 ± 10 nm), and NIR (790 nm ± 40 nm).

The collected images were used to generate a reflectance orthomosaic using structure-from-motion software (Pix4D, Prilly, Switzerland). Afterward, VIs were calculated using open-source software (QGIS Association, Zurich, Switzerland), based on the median of each plot, to ensure greater precision in representing the vegetation characteristics and minimize within-plot variation. For spectral data extraction, we used the "FIELDimageR" (Matias, Caraza-Harter, Endelman, 2020) and "Pliman" (Olivoto, 2022) packages. The plots' shapefiles were built using plugins and R scripts in QGIS (code available at: https://github.com/diegojgris/draw-plots-qgis). The aerial images were taken throughout the crop's production cycle, at different phenological stages, up to full maturity (R8) of all the experimental plots. The pipeline for integrating UAS-based spectral data into the analysis is illustrated in Figure 3, which details the key steps from flight planning to predictive modelling.

Figure 3. Pipeline for integrating UAS-based spectral data to predict flowering, maturity, and grain yield in soybean. The left panel illustrates UAS flight planning, orthomosaic generation, and plot segmentation using Pix4D and QGIS. The right panel presents the calculation of vegetation indices (NDVI, GNDVI, NDRE, and NGRDI), followed by the application of machine learning and deep learning models (Random Forest, XGBoost, and CNN) for predictive modeling, supported by the R programming language



Source: Author (2024).

2.3.3 Traits evaluated

During the experiments, the following traits were evaluated: (i) flowering stage, (ii) number of days to maturity, and (iii) grain yield (kg ha^{-1}, standardized to 13% moisture). Maturity was determined when 90% of the plants reached stage R8, identified through periodic field observations conducted throughout the crop cycle. Regular field visits were made to monitor the phenological development of each plot. Subsequently, the genotypes were categorized into three maturity classes: early, medium, and late. To define these classes for soybean genotypes in the Center-West region of Brazil, relative maturity groups were first determined according to the method proposed by Alliprandini et al. (2009). A regression equation, adjusted for the Center-North region, was applied as a function of crop cycle, allowing the classification of genotypes into relative maturity groups by comparing their cycles with those of reference cultivars (controls) used in the experiments.

In the image analyses for each experimental plot, the VIs calculated from the images were used as independent variables in the regression model. The VIs represented relationships between the radiation reflected from two or more bands, being effective algorithms for quantitative and qualitative evaluations of vegetation cover, vigor, and growth dynamics of plant species.

The VIs evaluated included NDVI (Normalized Difference Vegetation Index), NDRE (Normalized Difference Red Edge Index), GNDVI (Green Normalized Difference Vegetation Index), and NGRDI (Normalized Green Red Difference Index) (Table 2). These indices were calculated and analyzed via QGIS software, using the reflectance maps obtained from Pix4D Mapper software.

Table 2 - Vegetation indices evaluated in the experiments

DESCRIPTION	INDEX	FORMULAS	REFERENCES
Normalized Difference Vegetation Index	NDVI	(NIR-R)/(NIR+R)	Rouse et al. (1974)
Green Normalized Difference Vegetation Index	GNDVI	(NIR-G)/(NIR+G)	Gitelson et al. (1996)
Normalized Difference Red Edge Index	NDRE	(NIR-RE)/(NIR+RE)	Gitelson and Merzlyak (1994)
Normalized Green Red Difference Index	NGRDI	(G-R)/(G+R)	Tucker (1979)

Source: Author (2024)

Caption: NIR: reflectance in the near-infrared range; G: reflectance in the green range; R: reflectance in the red range; RE: reflectance in the red-edge transition range.

2.3.4 Statistical analysis

The ground-truth data for flowering, maturity, and grain yield were analyzed using the R Core Team software (2017) via the mixed model approach (Bernardo, 2010), employing the "SpATS" package for spatial analysis (Rodríguez-Álvarez et al., 2018). Initially, statistical analyses of the mixed models were carried out considering the addition of the spatial component, according to the model shown below:

$$y_{ijkmn} = \mu + T_k + R_j + G_i + L_m + C_n + f(col.n, row.n) + \varepsilon_{ijkmn}$$

where:

y_{ijkmn}: represents the phenotypic value observed for genotype *i* in the *j*-th repetition, in the *k*-th trial located in row *m* and column *n*;

μ: is the overall average of the experiment;

T_k: is the fixed effect of the *k*-th trial, which corresponds to the combination of a specific trial and the corresponding experiment;

R_j: is the fixed effect of the *j*-th repetition within each trial;

G_i: is the random effect of the i-th genotype, assuming that Gi follows a normal distribution with zero mean and variance σ_c^2($Gi \sim N(0, \sigma_g^2)$);

L_m: is the random effect of the m-th line, assuming that Lm follows a normal distribution with zero mean and variance σ_l^2($L_m \sim N(0, \sigma_l^2)$);

C_n: is the random effect of the m-th line, assuming that Cn follows a normal distribution with zero mean and variance σ_c^2($C_n \sim N(0, \sigma_c^2)$);

$f(col.n, row.n)$: is the spatial correction function, modeled by a penalized tensor product of two-dimensional B-splines, defined by the specific segments of the columns and rows;

ε_{ijkmn}: is the residual error associated with observation y_{ijkmn} assuming that ε_{ijkmn} follows a normal distribution with zero mean and variance σ_ε^2($\varepsilon_{ijkmn} \sim N(0, \sigma_\varepsilon^2)$);.

The variance components of the agronomic traits evaluated in the field were estimated with mixed models, using the Restricted Maximum Likelihood (REML) method to adjust the parameters. Predictions of random effects were obtained via Best Linear Unbiased Prediction (BLUP). In this context of spatial analysis, heritability was estimated considering generalized heritability, as proposed by Cullis et al. (2006), which is suitable for data with spatial correction.

The efficiency of indirect selection, as described by Falconer and Mackay (1996), based on genetic correlations and the heritability of characters, can be expressed by the formula:

$$ISE = \frac{r_{G,index} * \sqrt{h^2{}_{index}}}{\sqrt{h^2{}_{trait}}}$$

Where: $r_{G,index}$ is the genetic correlation between the vegetative index and the trait of interest; $h^2{}_{index}$ is the heritability of the vegetative indices; and $h^2{}_{trait}$ is the interest trait heritability.

2.3.5 Machine learning methods

In this study, we used two ML algorithms, namely Extreme Gradient Boosting (XGBoost) and Random Forest (RF), for maturity classification and grain yield prediction. Both methods, originating from the ensemble category, aim to enhance overall performance by combining predictions from multiple models into a final model that yields better results. Ensemble methods enhance classification accuracy and robustness by combining the predictions of various classifiers trained independently, leveraging the diversity between models to minimize errors (Dietterich, 2000). Models were performed in the R programming language. In a study by Yoosefzadeh-Najafabadi et al. (2021a), the authors concluded that the use of ensemble methods in a pipeline allows the effective use of reflectance data and machine learning algorithms in soybean selection at pre-harvest stages and analysis of high-throughput data.

As such, these ML methods differ significantly in their approach and training mechanism. RF is a parallelization method that reduces variance employing independently trained decision trees, using different data samples and subsets of variables for each tree to avoid overfitting (Belgiu & Drăguţ, 2016; Breiman, 2001), while XGBoost is a method that applies a boosting strategy, serializing the training process by building decision trees sequentially. Each new tree is trained to correct the errors of the previous trees, continuously adjusting to the residues left by them. This sequential process minimizes the model's bias and variance, thereby optimizing its overall performance by enhancing the accuracy of predictions at each step (Chen & Guestrin, 2016). These differences directly influence error reduction and performance optimization strategies.

2.3.6 Cross-validation for machine learning methods

In this study, the k-fold cross-validation technique with k = 10 (10-fold) employed here is widely recognized for its effectiveness in assessing the generalization capacity of models in different subsets of data (Nakatsu, 2023). This method divides the total data set into ten equal parts, using each one sequentially as a test set and the others for training.

Additionally, we implemented a k-fold cross-validation strategy, in which each genotype was assigned exclusively to either the training or test set within each fold. This genotype-wise partitioning ensured that the model was consistently evaluated on unseen genotypes across iterations, resulting in a more robust and realistic estimation of its generalization performance. As emphasized by Szeghalmy and Fazekas (2023), stratified and distribution-aware cross-validation approaches are essential for reliable evaluation in imbalanced learning scenarios. Furthermore, all observations are used for both training and testing across folds, allowing for a comprehensive and objective comparison of model performance.

2.3.7 Classification method for machine learning

In classification analysis, a model's performance is evaluated by its ability to correctly categorize instances into predefined classes or categories, which in this study correspond to maturity classes - early (0), medium (1), and late (2). This analysis is essential for evaluating the model's effectiveness in correctly assigning labels to data instances during training and testing. Each model tree was trained using 80% of the columns and instances, which were randomly sampled, thereby promoting diversity and robustness in learning. Implementing the stratified 10-fold cross-validation and employing an “early stopping” after 10 rounds without improvement ensures efficiency and helps prevent overfitting of the training set.

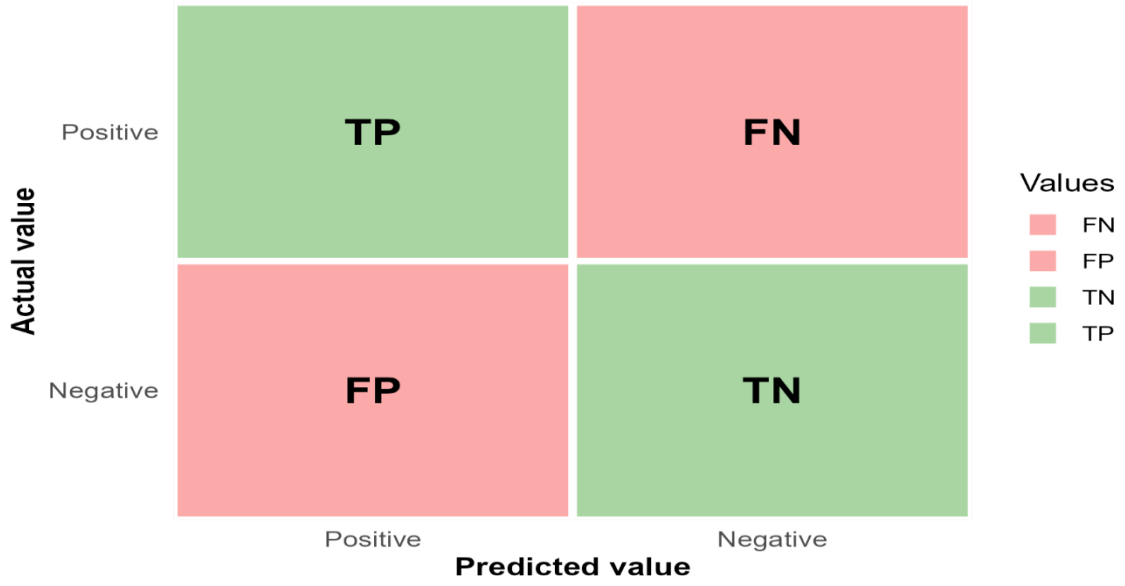
When training the RF model, we set the number of trees to 100 (ntree = 100) and limited the maximum tree depth to 5 (max depth = 5). During training, we adjusted the mtry parameter, which defines the number of variables to be considered in each division. After initial tests with values from 1 to 3, this parameter was optimized to 3, the number allowing maximum accuracy. This configuration searched to balance accuracy and simplicity, optimizing the model's performance.

Performance metrics, including precision, recall, F1-score, confusion matrix, and accuracy, were used to evaluate the classification algorithms. These metrics offer a

comprehensive view of the model's performance, encompassing crucial aspects such as correct predictions, recovery of positive instances, balance between precision and recall, precision distribution, and overall accuracy (Ahmed & Yadav, 2023).

Confusion matrices are fundamental tools that categorize model results into four types: True Positive (TP), in which the model correctly predicts the positive class; True Negative (TN), in which the model correctly predicts the negative class; False Positive (FP), cases in which the model incorrectly identifies an instance as being in the positive class; and False Negative (FN), cases in which the model does not correctly identify the positive class. Afterward, we present a basic explanatory model of the confusion matrix that provides a comprehensive view of the model's performance by comparing the predicted classifications with the actual values (Figure 4).

Figure 4 - Confusion matrix for the classification model



Source: Authors (2024).

Defining T as the total sum of all instances or observations:

$$T = (TP + TN + FP + FN)$$

Accuracy represents the proportion of correctly classified instances out of the total number of instances, providing an overall measure of the model's performance and indicating how often the model makes corrects predictions.

$$Acuraccy = (TP + TN)/T$$

Precision is called positive predictive value.

$$Precision = TP/(TP + FP)$$

The Recall is the proportion of real positives that have been correctly identified.

$$Recall = TP/(TP + FN)$$

The F1-score combines the model's precision and recall and defines the harmonic average of the precision and recall models.

$$F1 - Score = 2 * (\frac{Precision * Recall}{Precision + Recall})$$

2.3.8 Regression method for machine learning

In regression analysis, the model's performance was evaluated by its ability to predict continuous values from independent variables. The dependent variables analyzed in this study were the agronomic traits, while the VIs obtained from the spectral bands acted as independent (or explanatory) variables. This analysis is crucial for understanding the relationship between the variables and for applying the model to make estimates and projections.

To evaluate model performance, we employed specific metrics: the coefficient of determination (R^2) to measure the proportion of variance explained by the model, the mean absolute error (MAE) to assess the average magnitude of prediction error, and the root mean square error (RMSE) to emphasize larger deviations. A higher R^2 value, closer to 1, indicates a better model fit and predictive accuracy for continuous values. Conversely, lower MAE and RMSE values indicate higher model accuracy.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \hat{Y}_i|$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$$

Where:

n: is the total number of observations;

Y_i: are the actual values;

$\hat{Y}_i$: are the predictions;

$\bar{Y}$: is the average of the actual values.

2.3.9 Image-based deep learning for classification of soybean maturity

For the DL analysis, we integrated Python into the R interface to implement convolutional neural network (CNN) models for classifying soybean relative maturity into three groups: early, medium, and late. This approach focused on complex patterns, utilizing field information and cropped images of the experimental soybean plots as the basis. A total of 1,004 images were divided into two sets: 70% were used to train the model, and the remaining 30% provided sufficient test samples for validation, ensuring an adequate balance between training and testing. We employed CNN models using the Keras (Chollet, 2015) and TensorFlow (Abadi et al., 2015) packages for multiclass classification. Keras is a robust application programming interface (API) for DL, integrated as the standard modeling interface within TensorFlow. Images were loaded with the imager package, which represents pixel values in the [0,1] range prior to training.

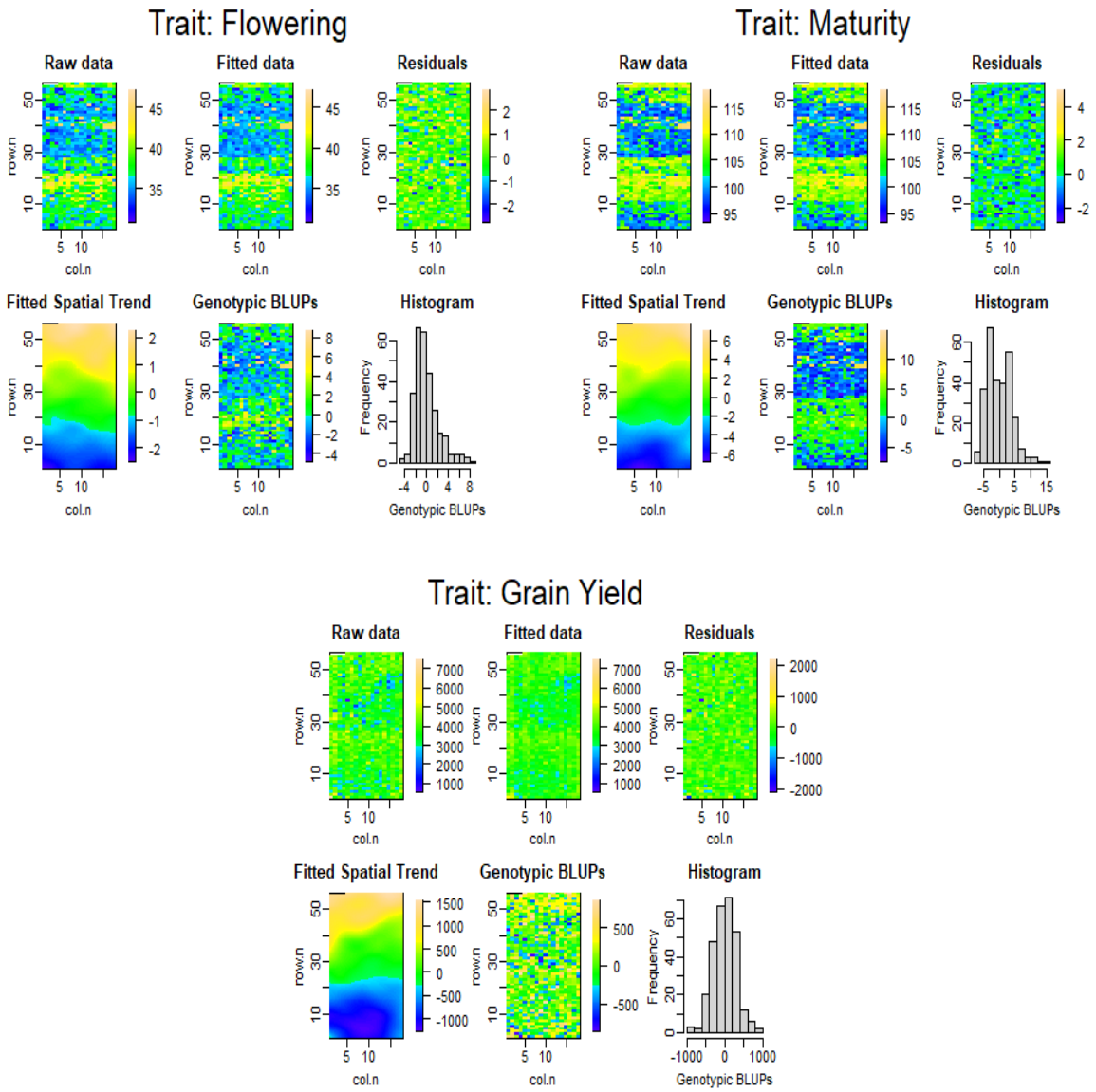
The architecture begins with a convolutional layer for 79 × 67-pixel images in 4 channels, based on the 1,004 plot images. This is followed by successive convolutional layers with 16, 32, 64, 128, and 256 filters, combined with ReLU activation and 2 × 2 pooling layers. Layers with ReLU activation and a dropout rate of 0.5 were followed by a softmax output layer with 15 units. The model was compiled with the Adam optimizer (learning rate = 0.001) and the Sparse Categorical Crossentropy loss function and trained for 20 epochs with a batch size of 32, employing validation to monitor performance.

2.4 Results

Our mixed model analysis was carried out after developing the images from the field data collection, employing the row and column spatial model. The results allowed for the obtention of BLUP values for genotypes, residues, and genetic parameters related to the evaluated trait heritability. Our results demonstrate the effectiveness of using splines to model spatial trends, as seen in the Fitted Spatial Trend graphs for the flowering, maturity, and grain yield variables. This approach allowed for an adjusted analysis of spatial variations in the experimental field (Figure 5).

Figure 5 - Adjusted spatial model, BLUPs based on phenotypic data and residuals for flowering, maturity, and grain yield traits

Source: Authors (2024).

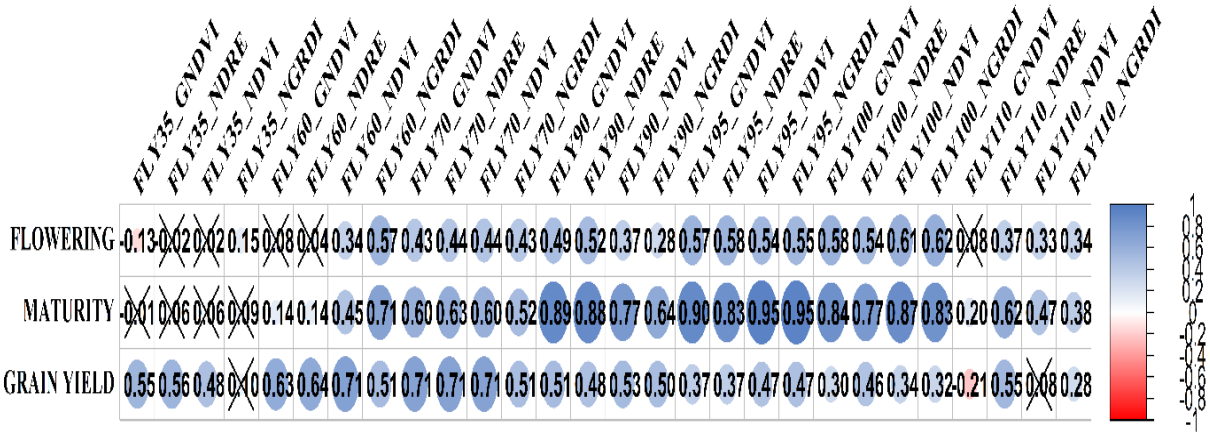


The frequency distribution histogram displays the spatial division of the genotypes and justifies the adopted analysis model. Genotypes were grouped according to the maturity groups, following the experimental design, which resulted in excellent data quality. The analysis revealed high heritability (h^2 = 0.94; 0.96) for flowering and maturity traits, respectively, indicating significant genetic control and considerable potential for selection and genetic breeding. For the grain yield trait, a moderate heritability was observed (h^2 = 0.69), indicating polygenic genetic control and a strong environmental influence. Residual analysis revealed homogeneous and controlled distributions for flowering and maturity, with more pronounced

variability for grain yield, which reinforces the greater sensitivity of this trait to environmental conditions. These findings underscore the importance of future research into genotype-environment interactions that influence grain yield, as well as the development of strategies to maximize grain yield under various conditions. The consistency of the residuals for flowering and maturity also reinforces the feasibility of using VIs for indirect selection in these traits.

We subsequently conducted Pearson's correlations between flowering, maturity, and grain yield regarding the VIs across the different data collection dates. Pearson's correlation is a statistical measure that assesses the strength and direction of the linear relationship between two variables, ranging from -1 to +1. A value of -1 indicates a perfect negative correlation, +1 indicates a perfect positive correlation, and 0 means no linear correlation. The significance of the correlations was determined at the 5% probability level, with non-significant values indicated by an “x” symbol (Figure 6).

Figure 6 - Pearson's correlation matrix between the evaluated traits and the vegetation indices across seven flight days (35, 60, 70, 90, 95, 100, and 110 DAS)



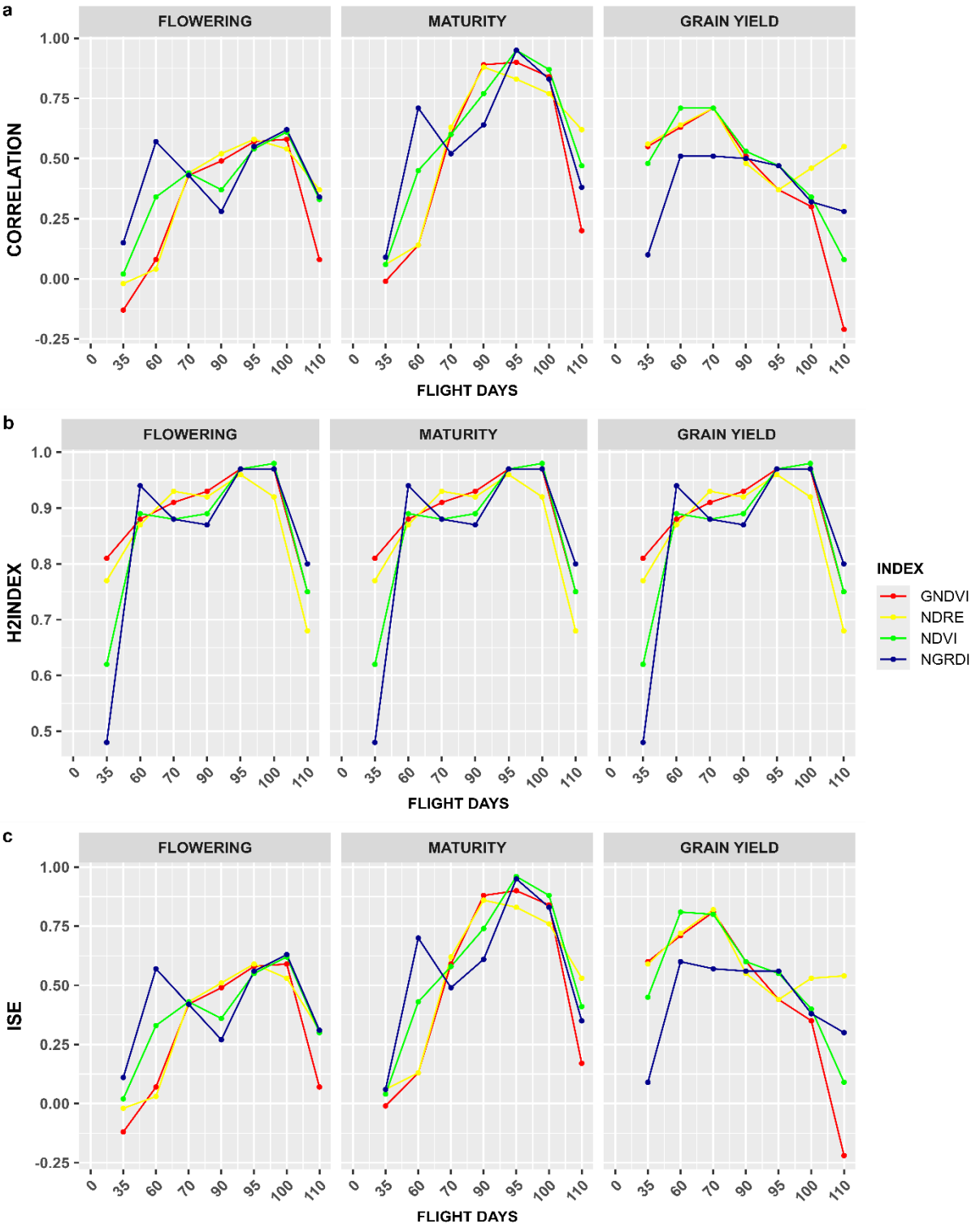
Source: Authors (2024).

A strong correlation, with a value of 0.95, existed between maturity and the NDVI and NGRDI during the 95 DAS flight. This strong correlation was more frequent in the reproductive phenological stages R5, R6, and R7 (90, 95, and 100 DAS), which correspond to the start of grain filling, full filling (green grain), and the start of maturity, respectively, for the different VIs evaluated. These final stages of the cycle reflect the genotype maturity stage, suggesting that the ideal flight days for evaluating maturity are those close to the senescence phase when the plant redirects its resources to the final filling of the grains.

Regarding grain yield, a high correlation of 0.71 was observed with the NDVI at 60 DAS, and similar correlations were found with the GNDVI, NDRE, and NDVI indices at 70 DAS, corresponding to the reproductive phenological stages R3 and R4. These stages represent the plant's peak vegetative biomass, characterized by greater leaf cover and enhanced photosynthetic capacity, which directly influence the potential for grain filling and, consequently, grain yield. Concerning flowering, moderate correlation values of 0.57 were recorded with the NGRDI at 60 DAS, at phenological stage R3.

In the phenotypic statistical analyses, 84 combinations of data were carried out based on three traits (flowering, grain yield, and maturity) and four VIs (NDVI, GNDVI, NDRE, and NGRDI) on seven flight dates (35, 60, 70, 90, 95, 100, and 110) and, therefore, the graphs of the traits and VIs correlations, heritability (h^2) of the vegetation index and ISE (indirect selection efficiency) were analyzed (Figure 7).

Figure 7 - Correlation between traits and vegetation indices: (a) heritability of vegetation indices (h^2); (b) indirect selection efficiency (ISE); and (c) variation across seven flight days based on drone images, considering three field traits: flowering, maturity, and grain yield



Source: Authors (2024).

Caption: Colored lines represent the vegetation indices: GNDVI (red), NDRE (yellow), NDVI (green) and NGRDI (blue).

The correlation patterns presented in Figure 7a highlight a strong association between traits and VIs, particularly for maturity, with the highest values observed at 90, 95, and 100 DAS. These time points correspond to key phenological stages: at 90 DAS, grain filling begins (R5); at 95 DAS, full grain filling occurs (R6); and at 100 DAS, physiological maturity begins (R7), reflecting this phenological transition (Figure 2). Regarding the h^2 index (Figure 7b), the VIs exhibited high heritability, underscoring their potential for genetic studies and breeding programs. Moreover, the identical curves across traits demonstrate the consistency of heritability values among VIs, regardless of the evaluated trait.

Subsequently, the ISE (Figure 7c) evaluates the efficiency of indirect selection of plant traits via correlated indices, regardless of the direction of the correlation with the desired trait (Falconer; Mackay, 1996). Instead of directly measuring the target trait, such as maturity or grain yield, the decision was made to observe and evaluate associated traits using comparable vegetation indices (VIs), which correlate positively with the desired traits. Furthermore, when analyzing maturity based on ISE, the flight at 95 DAS was particularly informative. Among the indices evaluated (NDVI, GNDVI, and NGRDI), no differences were observed in the ISE analysis, indicating the effectiveness of these VIs in indirect selection for maturity. Particularly, NDVI and the NGRDI achieved high ISE, exceeding the threshold of 0.9, which emphasizes the potential of these indices in accurately predicting maturity. These results highlight the feasibility of using cost-effective sensors, such as RGB, to derive VIs (e.g. NGRDI), offering an affordable alternative for similar applications.

Regarding the flowering trait, the moderate ISE values suggest that manual evaluation in the field is more reliable for this methodology, as it is challenging to identify the initial R1 stage in drone images of the soybean crop, even at 5 cm pixel-size or finer scale. Regarding grain yield, a high ISE value of above 0.8 for the NVDI at 60 DAS was identified. The NDVI, GNDVI, and NDRE also displayed ISE values above 0.8 at 70 DAS, corresponding to phenological stages R3 to R4, characterized by the start of pod formation and full pod formation, respectively.

About the statistical analysis of the correlations between the traits and the VIs over the seven flight collection dates, an ANOVA was carried out to evaluate the interactions and main effects (Table 3).

Table 3 - Analysis of variance for combinations between three traits (flowering, maturity, and grain yield) and four vegetation indices (NDVI, GNDVI, NDRE, and NGRDI) across seven flight days

SOURCES OF VARIATION	DF	SUM S	M SQUARE	F-VALUE	PR (F)
TRAIT	2	0.5655	0.2827	23.365	*
INDEX	3	0.0849	0.0283	2.339	NS
FLIGHT	6	2.2700	0.3783	31.263	*
INDEX × VOO	18	0.6215	0.0345	2.853	*
TRAIT × INDEX	6	0.1000	0.0167	1.377	NS
TRAIT × FLIGHT	12	1.6378	0.1365	11.278	*
RESIDUALS	36	0.4356	0.0121		

Source: Authors (2024).
Caption: *Significant by F test ($p<0.05$). NS - non-significant at $p < 0.05$.

In the ANOVA, a total of 84 combinations were analyzed, resulting from the three traits (flowering, maturity, and grain yield), four vegetation indices (NDVI, GNDVI, NDRE, and NGRDI), and 7 flight dates. Significant interactions were observed between flight days and VIs ($F = 2.853$, $p < 0.05$), as well as between flight days and phenotypic traits ($F = 11.278$, $p < 0.05$), indicating that the relationship between VIs and phenotypic traits varies over time. This suggests that the phenological stage at which the flights were carried out, as well as weather conditions (e.g., cloud cover) and moisture on the plants, may affect the nature of the correlation between the phenotypic traits and the VIs (Figure 6).

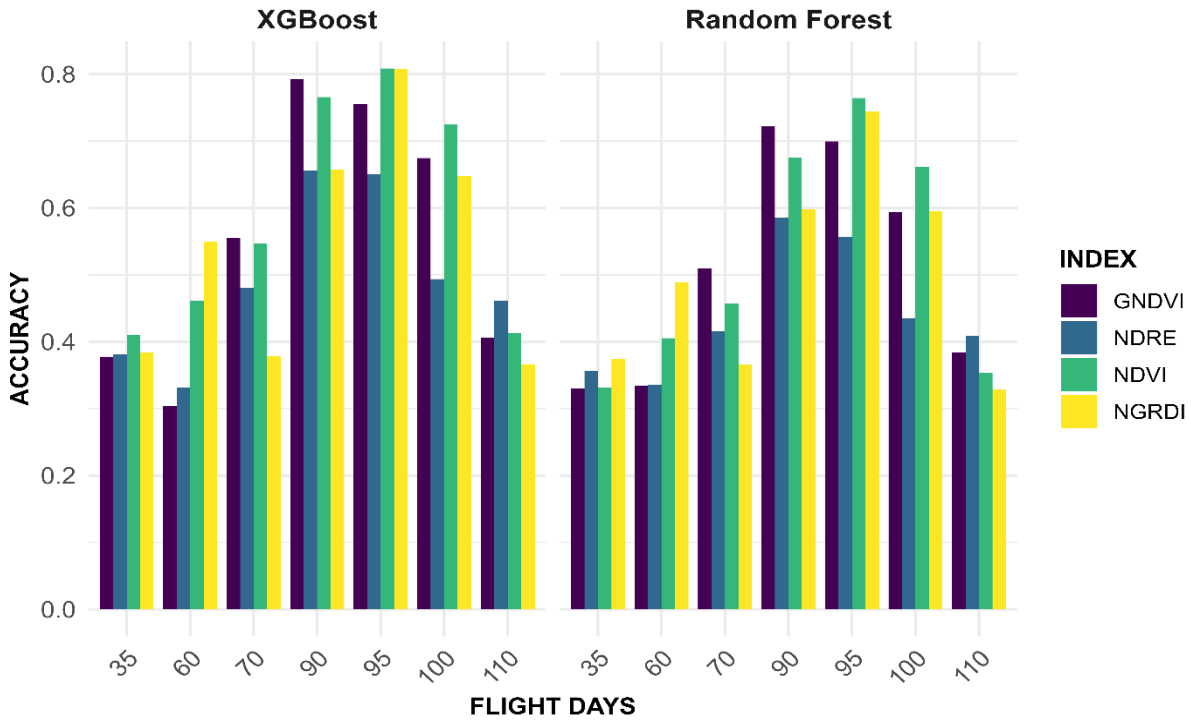
Conversely, the interaction between the phenotypic traits and the VIs was not statistically significant ($F = 1.377$, $p > 0.05$), indicating that, regardless of the vegetative index considered, its relationship with the phenotypic traits remained consistently the same across the different genotypes tested. The Tukey test was used to analyze the sources of variation in the ANOVA (Table S2). Therefore, the ability of the VIs to explain the variation in phenotypic traits was not influenced by the type of index used, corroborating the feasibility of using cost-effective sensors, such as RGB, to derive VIs (e.g., NGRDI) for this type of approach.

2.4.1 Machine learning models for classification

The ML classification models (XGBoost and Random Forest) were developed for the maturity traits, categorized into three maturity groups (early, medium, and late), based on seven flights and four vegetation indices. The results illustrate the model's accuracy in making correct predictions.

In the classification of soybean maturity for XGBoost, the evaluations highlighted high accuracies of 0.79 for GNDVI at 90 DAS and 0.81 for NDVI and NGRDI at 95 DAS. For the RF model, we identified an accuracy of 0.72 for GNDVI at 90 DAS and 0.76 for NDVI and 0.74 for NGRDI at 95 DAS. The best performance was observed at 95 DAS, during the phenological stages between R6 and R7, which corroborates the ISE analyses in both ML methods. Regarding the VIs at 95 DAS, the NDVI and NGRDI indices were found to have high accuracies for both ML methods, with the economic RGB NGRDI index standing out as an excellent alternative for HTP studies (Figure 8).

Figure 8 - Accuracy of the XGBoost and Random Forest machine learning models in evaluating vegetation indices across seven flight days in the soybean crop



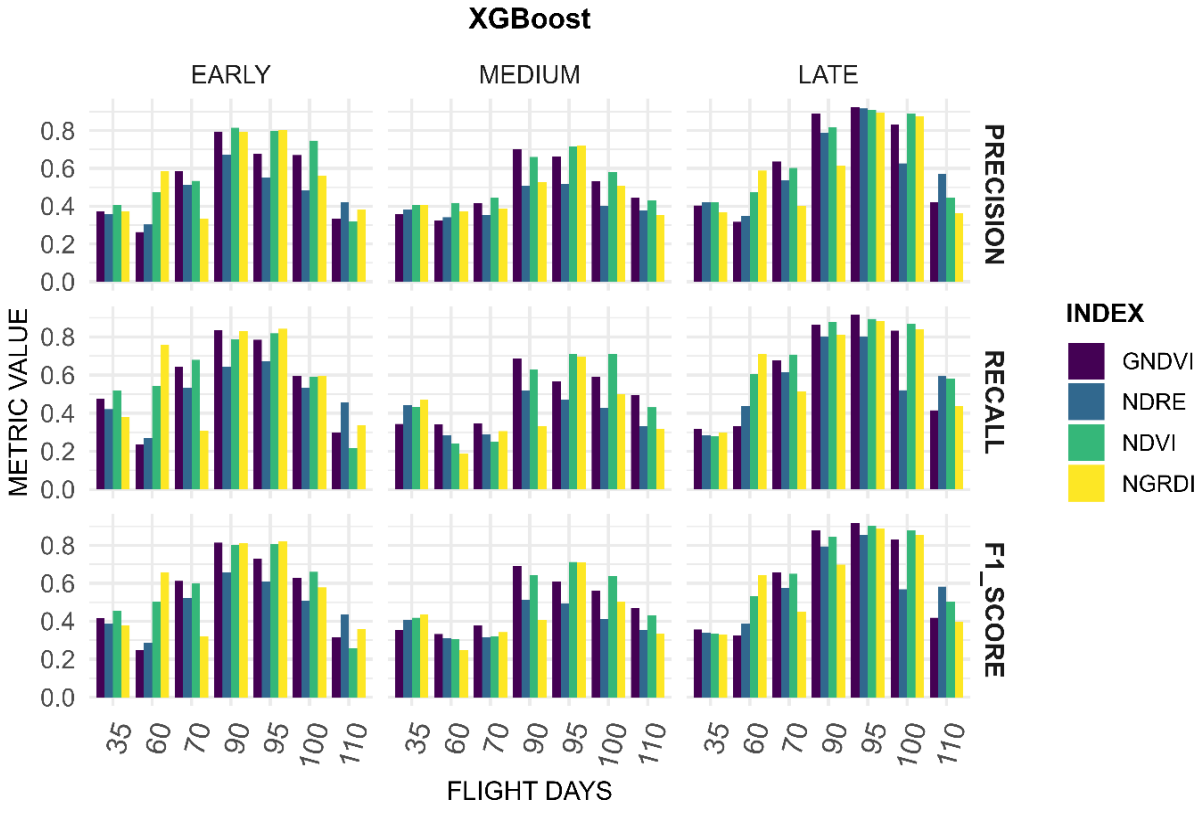
Source: Authors (2024).

In the context of metric evaluations for these results, confusion matrices were applied for each combination of flight day, VIs, and ML method (Table S3). These matrices were categorized into three classes to represent soybean growth stages: early (0), medium (1), and late (2). We observed significant improvements in classification for the maturity classes at 90, 95, and 100 DAS (phenological stages R5, R6, and R7, respectively) in both ML algorithms. For XGBoost, NDVI, and NGRDI proved to be the most effective, accurately classifying the

classes early (0) and late (2) at 95 DAS. Similarly, RF, NDVI, GNDVI, and NGRDI performed well for the same classes and periods. However, NDRE showed the highest error rate in both models. In this way, NDVI was the most effective in identifying the medium class (1) that presents the most confusion, compared to the other VIs for both algorithms.

The XGBoost model produced more accurate predictions, particularly at 95 DAS, for the NDVI, GNDVI, and NGRDI (Figure 9). Regarding the early class, the precision was 0.8 for NDVI and NGRDI, while for the late class, all indices showed a precision of 0.9. For the medium class, the precision was 0.72 for both NDVI and NGRDI. Regarding the recall, the values were 0.82 for NDVI and 0.84 for NGRDI in the early class. In the late class, NDVI reached 0.89, GNDVI 0.92, and NGRDI 0.88, while NDRE was 0.8. For the middle class, the recall was 0.71 for NDVI and 0.70 for NGRDI. The F1-score was 0.81 for NDVI and 0.82 for NGRDI in the early class, and higher than 0.9 for NDVI and GNDVI, 0.89 for NGRDI, and 0.86 for NDRE in the late class. For the middle one, the F1-score was 0.71 for NDVI and NGRDI.

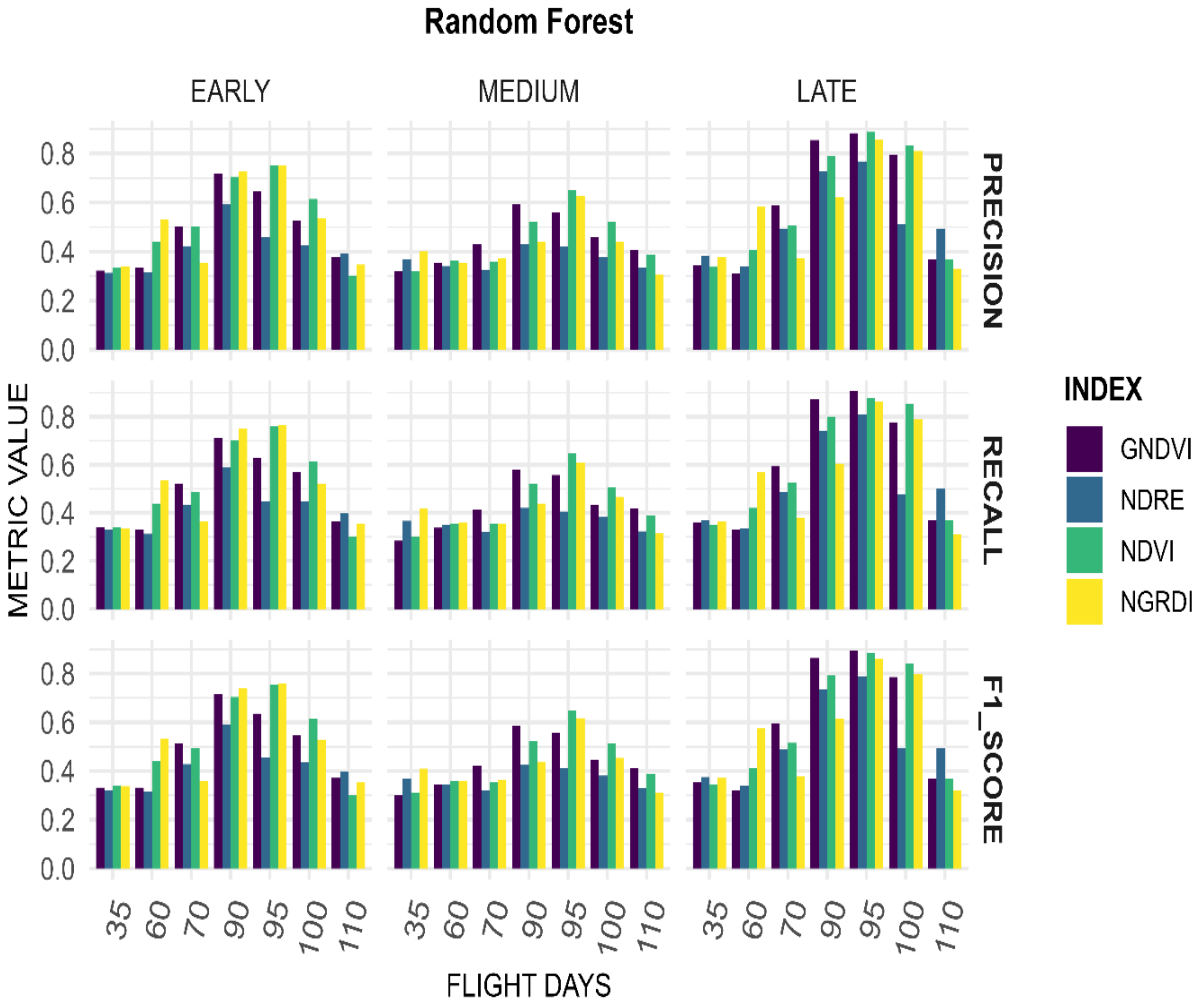
Figure 9 - Evaluation of maturity classes using XGBoost classification metrics applied to vegetation indices (NDVI, GNDVI, NDRE, and NGRDI) across seven flight days. The evaluated metrics include precision, recall, and F1-score



Source: Authors (2024).

The RF model showed better predictions, especially at 95 DAS, for the NDVI, GNDVI, and NGRDI (Figure 10). The results for NDVI showed a precision of 0.75, a recall of 0.76, and an F1-score of 0.76 for the early class; all the metrics were 0.65 for the medium class; and for the late class, a precision of 0.89, a recall of 0.88, and an F1-score of 0.88. For NGRDI, the precision was 0.75, recall 0.77, and F1-Score 0.76 in the early class; for the medium class, precision 0.63, recall 0.61, and F1-Score 0.62; and for the late class, all values were 0.86. The GNDVI yielded the best results in the late class, with values approaching 0.9 for the evaluated metrics.

Figure 10 - Evaluation of maturity classes using Random Forest classification metrics, applied to vegetation indices (NDVI, GNDVI, NDRE, and NGRDI) across seven flight days. The evaluated metrics include precision, recall, and F1-score



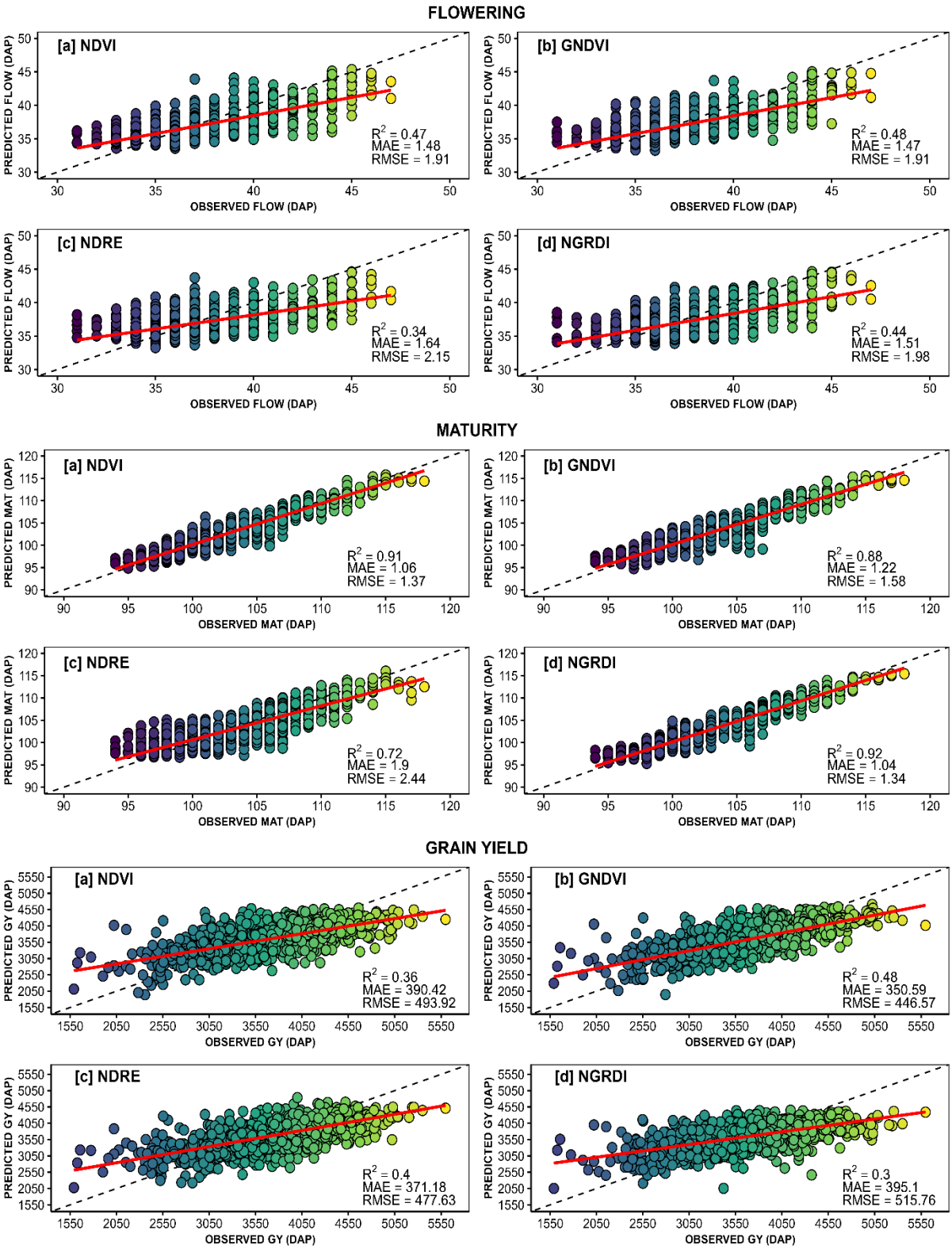
Source: Authors (2024).

Thus, the results of both machine learning models highlight the effectiveness of the NDVI and NGRDI, demonstrating their ability to accurately classify soybean maturity classes, especially the early and late classes. In addition, the results suggest that indices based on the visible spectrum (RGB) can be an efficient economic alternative to multispectral indices.

2.4.2 Machine learning models for regression

Regression analyses for ML were carried out for the flowering, maturity, and grain yield traits. We incorporated data from the four VIs collected over seven days of flights, covering the crop cycle. Thus, the regression was carried out using the complete time series of these VIs, capturing the variation over the analyzed days. When using the XGBoost model, the VIs demonstrated a distinct ability to predict the studied traits (Figure 11). Although the R^2 is commonly used to measure the proportion of variability explained by the model, it is important to complement it with error metrics, such as RMSE and MAE, for a more robust evaluation of the model's performance. In the case of the maturity trait, the predictions were more accurate, with an R^2 of 0.91 for NDVI and 0.92 for NGRDI, indicating an excellent model fit. In contrast, the model's performance was moderate for grain yield and flowering predictions, with an R^2 value of 0.48 for both traits using GNDVI. These variations in accuracy were also reflected in the error metrics, such as RMSE and MAE. For the maturity trait, for example, the RMSE was 1.34 and the MAE was 1.04 for the NGRDI, showing significant accuracy in the predictions.

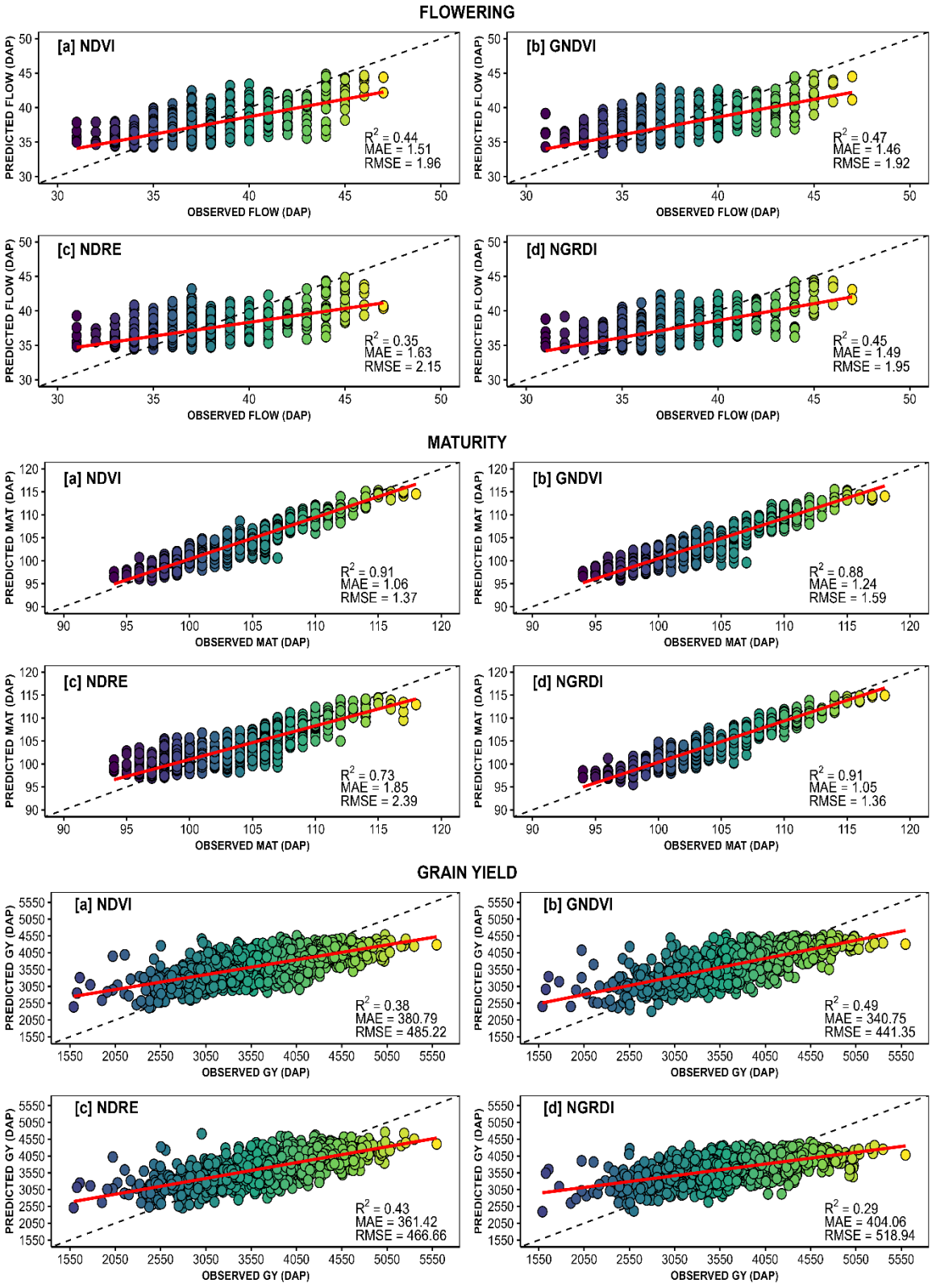
Figure 11 - Graphical performance of regression analyses using XGBoost applied to vegetation indices (NDVI, GNDVI, NDRE, and NGRDI) across seven flight days for the flowering, maturity, and grain yield traits. The evaluated metrics include the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE)



Source: Authors (2024).

For the RF model, both NDVI and NGRDI provided highly accurate predictions for maturity, with identical R^2 values of 0.91. The corresponding MAE and RMSE values were also similar (1.06 and 1.37 for NDVI; 1.05 and 1.37 for NGRDI), demonstrating the consistency and robustness of the model (Figure 12). However, grain yield and flowering showed more moderate performance, with GNDVI reaching an R^2 of 0.49 for grain yield and 0.47 for flowering. The RMSE and MAE values corroborated these results, reinforcing the RF's ability to estimate maturity with greater precision, which contrasted the errors observed in the grain yield and flowering predictions that displayed greater variability.

Figure 12 - Graphical performance of regression analyses using Random Forest applied to vegetation indices (NDVI, GNDVI, NDRE, and NGRDI) across seven flight days for the flowering, maturity, and grain yield traits. The evaluated metrics include the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE)



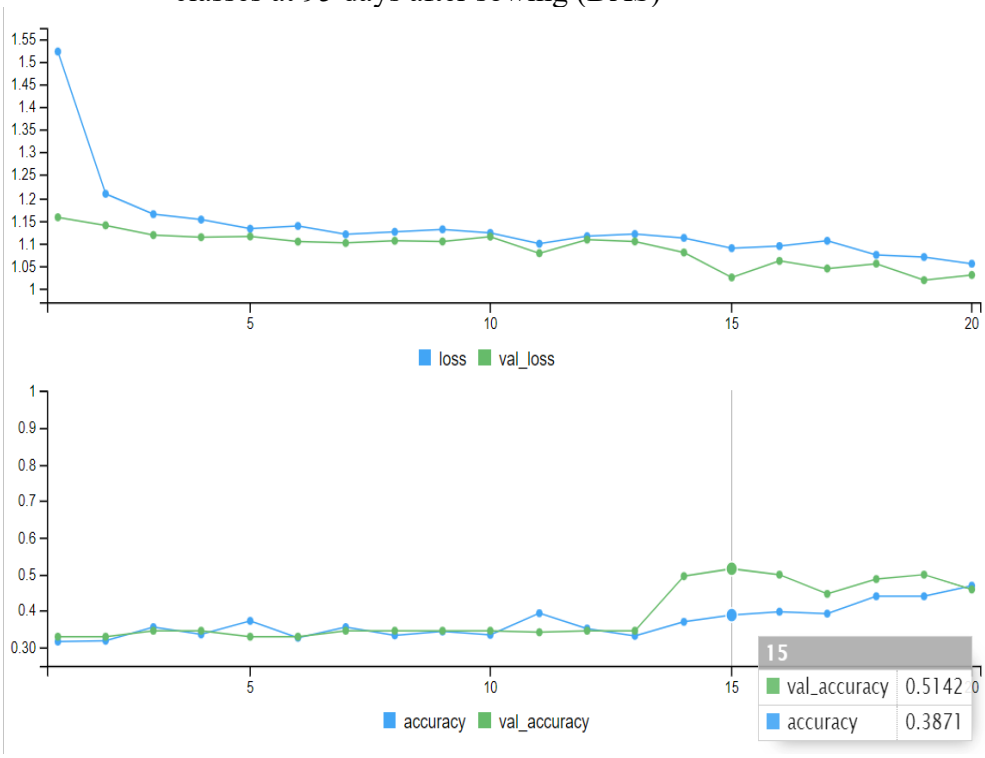
Source: Authors (2024).

2.4.3 Deep learning model for soybean maturity classification

To investigate soybean maturity groups, we categorized field images of soybean plots captured at 95 DAS into three classes: early (0), medium (1), and late (2). A convolutional neural network (CNN) was applied to classify these maturity classes. The model's performance was evaluated using learning curves and a confusion matrix (Figure 13).

The CNN achieved a validation accuracy of 51%, indicating moderate classification performance. The confusion matrix revealed a clearer distinction between early and late classes, while the medium class showed greater overlap, suggesting difficulty in defining boundaries. Although the model demonstrates potential, the results highlight the need for improvements in model architecture and training dataset composition. The final validation loss (1.01) reflects the challenges in minimizing classification errors, indicating room for refinement.

Figure 13 - Learning curves and confusion matrix for the classification of soybean maturity classes at 95 days after sowing (DAS)



	predicted classes	
Actual	0	2
0	85	17
1	43	52
2	29	68

Source: Author (2024).

2.5 Discussion

2.5.1 Exploratory data analysis

Our results demonstrate the potential of using UAS and VIs derived from spectral bands in breeding programs, as well as employing ML models and DL techniques in approaches such as classification and regression for predicting agronomic traits in large-scale high- throughput phenotyping evaluations.

The spatial mixed model adopted in this study effectively captured the growth gradient in the soybean plots, as evidenced by the distribution of the data collected in the field and the adjusted spatial trend. The uniformity of the residue in flowering and maturity traits indicates that most of the spatial variation was adequately modeled, suggesting that the model had a good fit for these highly heritable traits. Conversely, the higher residual variability for yield reflects greater sensitivity to the environment, highlighting the need to refine the model for this trait. The distribution of phenotypic BLUPs reflects the genetic diversity between the genotypes tested and, together with the histogram, suggests a balanced distribution of phenotypic values (Figure 5). These findings highlight the potential of the spatial mixed model in breeding programs, where precision in capturing spatial variability can direct the selection of genotypes resilient to environmental conditions. Our results reinforce the accuracy of the experimental design and the efficiency of phenotyping, highlighting the feasibility of identifying superior performance genotypes for key traits, such as grain yield and crop cycle, which are essential for improving soybean production.

Pearson's correlation coefficient (r) indicated a robust correlation between maturity and VIs. This strong correlation suggests that vegetation indices effectively capture variations associated with maturity. Although maturity is a highly heritable trait controlled by a few genes, its correlations with those indices reflect relevant phenotypic variations influenced by genetic and environmental factors. The high correlation between VIs and maturity observed in the final reproductive stages (R6 and R7) suggests that flight days near the end of the cycle are ideal for capturing the maturity phases of the genotypes. These stages reflect the process of senescence and the redirection of resources toward final grain filling, which reinforces the choice of these periods for crop cycle evaluations, as reported by e Silva et al. (2024). Strong correlation values and high heritability above 0.9 were observed between maturity and RGB VIs in the soybean crop across various environments (Narayanan et al., 2019), corroborating the findings of this work.

Although Pearson's correlations between the VIs and flowering were moderate, they highlight the interaction between genetic and environmental factors in determining the timing of flowering. Although highly heritable, flowering is sensitive to environmental variations such as temperature and photoperiod (Figure 1). This sensitivity, combined with the timing of the flights, which did not adequately capture the transition from the vegetative to the reproductive period (R1), limited the accuracy of the VIs in reflecting the phenotypic changes of flowering. Recent work with the flowering traits indicates the difficulty and challenge of working with more complex DL models and modern sensors to identify phenotypic variations of flowers in the approach to soybean cultivation (Zhu et al., 2022).

Regarding grain yield, the highest correlations with VIs were observed at stages R3 and R4 (60 and 70 days after sowing, DAS), when the plant reaches its peak of vegetative biomass. This period represents the maximum potential for photosynthesis and resource allocation for initial grain formation, highlighting the role of the biomass accumulated at these stages as a critical factor in yield. The relationship between biomass and yield has been observed in studies with other grain crops, such as corn, and has been identified as fundamental for predicting final yield, especially in time intervals extending up to 120 DAS (Macedo et al., 2023; Venancio et al., 2020). Regarding grain yield in soybeans, a variable correlation was observed with the VIs, reflecting the complexity of this trait, which has moderate to low heritability, is quantitative and polygenic, and is highly influenced by environmental factors, making it a crucial character in breeding programs (Singh & Vatsa, 2009).

Heritabilities and correlations between the traits determine the ISE. Indirect selection becomes more efficient than direct selection when the product between the square root of the heritability of the target trait in the indirect selection and the genetic correlation with the direct selection trait exceeds the square root of the direct selection trait heritability (Falconer; Mackay, 1996). Maturity showed a high ISE with the VIs, especially at 95 and 100 DAS, reflecting its high heritability (0.96) and correlations (higher than 0.9). In a study conducted with tropical wheat, Silva et al. (2024) demonstrated that vegetation indices derived from HTP platforms were highly efficient for indirect selection, particularly in terms of yield and cycle traits.

The results indicate a peak of grain yield between 60 and 70 DAS, coinciding with the greatest accumulation of vegetative biomass. This initial accumulation may have provided the necessary structure to sustain higher grain yield. However, physiological maturity, observed between 95 and 100 DAS, represents the final phase of the cycle, when the plant redistributes its resources from vegetative biomass to reproductive organs, completing its development.

These results suggest that utilizing VIs in maturity evaluations is an effective strategy in HTP (Narayanan et al., 2019; Volpato et al., 2021). Identifying the optimal days of flight, along with the phenological stages of the crop for spectral image analysis, yields significant savings in time, computational resources, and operational costs, thereby optimizing the genetic breeding process.

This study highlights that the NGRDI economic RGB index achieved comparable results to the multispectral NDVI in all analyses. Vegetation indices derived from UAS imagery, especially those incorporating the NIR band, have demonstrated strong correlations with agronomic variables (Lima et al., 2022). Similarly, Al Rosyid et al. (2022) highlighted the effectiveness of the RGB index, particularly the NGRDI, which demonstrated comparable performance to the multispectral NDVI across all growth stages of the rice crop. This similarity supports the high correlation between the two indices, emphasizing their applicability in HTP studies for various crops, reinforcing the potential of cost-effective RGB-based indices for phenotyping applications. Recent studies have further highlighted the relationship between vegetation indices and key plant traits, reinforcing their role in high-throughput phenotyping using UAV imagery (Castilho et al., 2024). Regarding the ISE for grain yield, strong associations were observed at 60 DAS for NDVI and at 70 DAS for NDVI, GNDVI, and NDRE, all with values above 0.8. These periods correspond to the R3–R4 phenological stages, characterized by pod formation and full pod formation, respectively, confirming their importance as key moments for indirect selection. Working with soybeans, Santana et al. (2022) demonstrated that early and productive genotypes can be identified based on vegetation indices and specific wavelengths.

In a study on HTP using UAV, Ren et al. (2023) indicated that the flowering and grain-filling phases are ideal for yield estimation, using a combination of variables including VIs and maturity. Identifying ideal stages for image acquisition helps to decide flight days and which VIs to use to optimize the evaluation of traits such as grain yield and maturity. Additionally, Yoosefzadeh-Najafabadi et al. (2021b) investigated 35 hyperspectral indices (HVI) collected at R5, finding that indices associated with the 670 nm (red) and 800 nm (near infrared) regions were the most effective at predicting grain yield, which corroborates the results of this study.

2.5.2 Performance of machine learning models for classification

The quality of training data is fundamental to the performance of ML models, particularly in maturity classification tasks. Therefore, the classification of the maturity classes

showed high explanatory power, with an accuracy of 0.81 for NDVI and NGRDI at 95 DAS using the XGBoost model. This result suggests that the model effectively captured the variability present in the data, particularly between the most distinct classes, early (0) and late (2), as detailed in the classification metrics (Figures 9 and 10) and observed in the confusion matrices (Table S3). The RF model also achieved consistent results, although with slightly lower accuracies overall. The magnitude of precision provides confidence in the model's ability to predict soybean maturity classes, making it particularly useful for practical applications in precision agriculture, where reliable data can enhance decision-making in breeding programs.

The choice of the ideal performance metric depends on the specific objective for the prediction. For example, recall proved particularly effective, reaching 0.92 for GNDVI in the late class and showing similarly high values for NDVI and NGRDI in XGBoost (Figure 9). The RF model also showed robust recall values in the late class (up to 0.88), reinforcing the relevance of this metric across algorithms (Figure 10). Therefore, this metric is preferable when the focus is on minimizing false negatives (FN), which is crucial to ensure the correct identification of instances across classes, as opposed to the accuracy metric, which combines true positives (TP) and true negatives (TN) and may not accurately reflect the model's ability to classify all classes correctly (Naik et al., 2017). In contrast, precision relates to false positives (FP), while the F1-score balances both precision and recall. However, this balance may dilute the emphasis on minimizing FN in critical scenarios, such as maturity classification.

In this study, the most distinct soybean maturity classes, early (0) and late (2), exhibited clear differences in phenotypic and spectral patterns, facilitating both visual and automatic classification. In contrast, the medium class (1), representing an intermediate stage of maturity, was more prone to misclassification with either early or late classes. This may be due to the methodology used to define maturity classes, planting location, photoperiod, and abiotic and biotic factors. Consequently, the precision for accurately classifying the medium maturity class is low, reflecting the complexity of distinguishing intermediate maturity stages based on the available training data (Hassanijalilian et al., 2020). In these cases, it is more common to use sophisticated ML models, such as boosting and ensemble models (e.g. XGBoost and Random Forest), as well as DL approaches, due to their superior ability to deal with complexity and overlaps between classes (Araújo et al., 2023), as in the differentiation between early, medium and late soybean maturity classes. However, it is essential to compare their performance with that of other models developed for the same classification task and incorporate additional data into the model, such as genotypes with known maturity classes and genotype-by-environment interactions, to enhance its accuracy and applicability.

2.5.3 Performance of machine learning models for regression

The regression analyses highlight both the potential and limitations of the ML XGBoost (Figure 11) and RF (Figure 12) models in predicting important agronomic traits over a time series of VIs. Both methods demonstrated comparable performance across different traits, reinforcing their suitability as predictive tools. The high accuracy of these models in predicting maturity, especially when using NDVI and NGRDI, highlights the ability of these indices to capture essential information related to the final stages of plant development, with accuracies above 0.9 for both indices.

The accuracy observed can be attributed to the NDVI sensitivity to the vegetative activity of the plant quantified by the leaf area index (LAI) (Bajocco et al., 2022) and the ability of NGRDI to reflect photosynthetic efficiency through its correlation with chlorophyll (Pazhanivelan, 2023), both of which are fundamental for evaluating the maturity stage. In a study with soybeans, Zhou et al. (2019) demonstrated that models based on multispectral images obtained by UAV, using partial least squares regression (PLSR) with 10-fold cross-validation, are promising and practical for estimating maturity dates. The study selected proven image characteristics for prediction and evaluated the rate of change between data collection days, achieving high accuracy in estimating maturity.

Concerning flowering, the inefficient capture of this trait can be attributed to the flight schedules, which occurred predominantly at the beginning of the crop cycle, at intervals of 15 to 20 days. Considering the duration of flowering and the variety of genotypes with different cycles, this periodicity was insufficient to adequately record the rapid changes associated with the onset of the reproductive stage (R1) of soybean development. Additionally, the analysis of soybean flowers requires high-resolution images. However, only low-resolution images were obtained due to flight height and flower covering which compromised their detection and limited their monitoring.

The moderate success of the regression models in predicting grain yield and flowering indicates the complexity of these traits, which are affected by multiple biotic and abiotic factors. The lower R^2 values for grain yield suggest that other crucial factors, such as environmental ones, were not captured by the indices and need to be considered to improve the accuracy of predictions. These findings underscore the need for ongoing research into developing more sophisticated indices and models to capture the greater complexity of crop trait variability.

2.5.4 Performance of deep learning model for soybean maturity classification

The deep learning model's ability to accurately differentiate between soybean genotypes in early (0), medium (1), and late (2) maturity classes was limited, with a moderate accuracy of 51%. This performance was influenced by the 79x67-pixel image resolution obtained by drones flying at a height of 50 meters, resulting in a granularity of approximately 5 cm per pixel. This barrier may be overcome by reducing flight height to capture higher-resolution images and investing in better equipment, such as more advanced drones and cameras, although this implies additional costs. In addition, the low-resolution images compromise the supply of detailed data to the model and the precision of classification. Therefore, feeding the model with a higher-quality data set to capture the subtle variations between maturity stages is essential.

Future directions include integrating multispectral data and utilizing pre-processing techniques to enhance the sharpness and relevance of the captured information. Trevisan et al. (2020) demonstrated in a study on soybean maturity that the RMSE displayed the most significant increase when models were used with the lowest resolution (750 mm/px). This result underscores the importance of high-resolution images in achieving greater accuracy in estimating maturity, particularly in HTP contexts. In addition, applying advanced DL methods, such as an improved CNN, can intensify the identification of maturity class traits. Expanding the data set with more representative examples of each class, including witnesses, can also increase the classification's accuracy, considering the genetic diversity of the genotypes. However, these approaches require greater computing power and present challenges in compiling and storing data, which may affect their viability for adoption (Albahar, 2023).

Ultimately, these results offer valuable insights into soybean breeding strategies, highlighting the need for more integrative approaches that combine spectral data with genotypic, phenotypic, and environmental information. Resende et al. (2024) emphasize that integrating statistics, environment, and remote sensing is crucial for understanding the complex interactions among genetics, environment, and management (G × E × M). This innovative approach opens up new perspectives for advances in HTP with VIs obtained from UAS platforms, enhancing the indirect selection of soybean genotypes and increasing the precision and efficiency of breeding programs.

2.6 Conclusion

This study confirmed the feasibility of phenotyping with UAS equipped with RGB and multispectral sensors, utilizing vegetation indices (VIs) to evaluate agronomic traits, such as maturity and grain yield, in soybean genotypes. Additionally, the NGRDI index (based on RGB bands) demonstrated performance comparable to the NDVI, supporting the use of more cost-effective equipment for phenotyping applications.

Indirect selection using VIs showed high efficiency for maturity (ISE values ≥ 0.9 for NDVI, GNDVI, and NGRDI at 95 DAS, corresponding to stages R6–R7) and grain yield (ISE values ≥ 0.8 for NDVI at 60 DAS and for NDVI, GNDVI, and NDRE at 70 DAS, corresponding to stages R3–R4). These results highlight the importance of VIs in capturing critical phenological stages and enhancing the selection and development of cultivars in soybean breeding programs.

The ML XGBoost and RF models classified maturity classes with 80% and 75% accuracy for NDVI and NGRDI at 95 DAS, respectively. For regression, both models showed R^2 values greater than 0.9 for maturity, confirming the effectiveness of VIs in predicting maturity and classifying soybean genotypes. Deep learning with CNNs achieved moderate accuracy in maturity class classification, suggesting that improvements in image resolution and network architecture could enhance performance.

2.7 Acknowledgements

The authors thank the Brazilian Agricultural Research Corporation (EMBRAPA) for providing the experiment location and plant material, as well as the Remote Sensing and Geoprocessing Lab (LAPIG)/Pro-VANT Nucleus for their support in this research, including the use of UAS and sensors. We also extend our gratitude to the LSU AgCenter and the members of the Roberto Fritsche Neto laboratory for their assistance with the data analyses. Additionally, we acknowledge the Federal University of Viçosa (UFV) and its Graduate Program in Genetics and Breeding for their contributions to human resource training. Finally, we are grateful to the National Council for Scientific and Technological Development (CNPq) and the Coordination for the Improvement of Higher Education Personnel (CAPES) for financial support and scholarships.

2.8 References

ABADI, M.; AGARWAL, A.; BARHAM, P.; BREVDO, E.; CHEN, Z.; CITRO, C.; CORRADO, G. S.; DAVIS, A.; DEAN, J.; DEVIN, M.; GHEMAWAT, S.; GOODFELLOW, I.; HARP, A.; IRVING, G.; ISARD, M.; JIA, Y.; JOZEFOWICZ, R.; KAISER, L.; KUDLUR, M.; ZHENG, XI. **TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.** 2015. [Software]. Available at: https://www.tensorflow.org/. Accessed on: May 25, 2023.

ABREU JÚNIOR, C.A.M.D.; MARTINS, G.D.; XAVIER, L.C.M.; VIEIRA, B.S.; GALLIS, R.B.D.A.; FRAGA JUNIOR, E.F.; MARTINS, R.S.; PAES, A.P.B.; MENDONÇA, R.C.P.; LIMA, J.V.D.N. Estimating Coffee Plant Yield Based on Multispectral Images and Machine Learning Models. **Agronomy**, [*S. l.*], v. 12, n. 12, p. 3195-3220, 2022.

AHMED, I.; YADAV, P. K. Plant disease detection using machine learning approaches**. Expert Systems,** [*S. l.*], v. 40, n. 5, p. e13136, 2023.

ALBAHAR, M. A. Survey on Deep Learning and Its Impact on Agriculture: Challenges and Opportunities. **Agriculture,** [*S. l.*], v. 13, n. 3, p. 540, 2023,

ALLIPRANDINI, L. F.; ABATTI, C.; BERTAGNOLLI, P. F.; CAVASSIM, J. E.; GABE, H. L.; KUREK, A.; MATSUMOTO, M. N.; OLIVEIRA, M. A. R.; PITOL, C.; PRADO, L. C.; STECKLING, C. Understanding soybean maturity groups in Brazil: Environment, cultivar classification, and stability. **Crop Science**, [*S. l.*], v. 49, p. 801-808, 2009.

AL ROSYID, A. N. I.; ENDIVIANA, O. A.; ASTIKA, I. W.; SETIAWAN, Y.; MUTTAQIN, K. H.; IMPRON; IMANTHO, H.; SUGIARTO, S. W.; YULIAWAN, T. Visible Band Optimization of Unmanned Aerial Vehicle for Estimating Synthetic NDVI on Rice Vegetation. **Jurnal Keteknikan Pertanian,** [*S. l.*], v. 10, n. 3, p. 281-290, 2022.

AMARAL, L. R.; FREITAS, R. G. de; BARBOSA JÚNIOR, M. R.; SIMÕES, I. O. P. da S. Application of drones in agriculture. In: QUEIROZ, D. M. de; VALENTE, D. S. M.; PINTO, F. A. de A. de C.; BORÉM, A.; SCHUELLER, J. K. (Eds.). Digital Agriculture. Cham: Springer, 2022.

ARAÚJO, S. O.; PERES, R.S.; RAMALHO, J.C.; LIDON, F.; BARATA, J. Machine Learning Applications in Agriculture: Current Trends, Challenges, and Future Perspectives. **Agronomy,** [*S. l.*], v. 13, n. 12, p. 2976, 2023.

ARAUS, J.L.; KEFAUVER, S.C.; ZAMAN-ALLAH, M.; OLSEN, M.S.; CAIRNS, J.E. Translating high-throughput phenotyping into genetic gain. **Trends Plant Science,** [*S. l.*], v. 23, p. 451-466, 2018.

ATTRI, I.; AWASTHI, L. K.; SHARMA, T. P.; RATHEE, P. A review of deep learning techniques used in agriculture. **Ecological Informatics**, *[S. l.]*, v. 77, p. 102217, 2023.

BAJOCCO, S.; GINALDI, F.; SAVIAN, F.; MORELLI, D.; SCAGLIONE, M.; FANCHINI, D.; RAPARELLI, E.; BREGAGLIO, S. U. M. On the use of NDVI to estimate LAI in field crops: Implementing a conversion equation library. **Remote Sensing,** [*S. l.*], v. 14, n. 15, p. 3554, 2022.

BARBOSA JÚNIOR, M. R.; TEDESCO, D.; CARREIRA, V. S.; PINTO, A. A.; MOREIRA, B. R. A.; SHIRATSUCHI, L. S.; ZERBATO, C.; SILVA, R. P. The time of day is key to discriminate cultivars of sugarcane upon imagery data from unmanned aerial vehicle. Drones, [S. l.], v. 6, n. 5, p. 112, 2022.

BELGIU, M.; DRĂGUŢ, L. Random Forest in Remote Sensing: A Review of Applications and Future Directions. **ISPRS Journal of Photogrammetry and Remote Sensing**, [*S. l.*], v. 114, p. 24-31, 2016.

BERNARDO, R. **Breeding for quantitative características in plants**. 2. ed. Woodbury: Stemma, 2010. 400 p.

BREIMAN, L. Random Forests. **Machine Learning,** [*S. l.*], v. 45, n. 1, p. 5-32, 2001. ok

CASTILHO, D.; TEDESCO, D.; HERNANDEZ, C.; MADARI, B. E.; CIAMPITTI, I. A global dataset for assessing nitrogen-related plant traits using drone imagery in major field crop species. **Scientific Data,** v. 11, p. 585, 2024.

CHEN, T.; GUESTRIN, C. XGBoost: A Scalable Tree Boosting System. **Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**, p. 785-794, 2016.

CHOLLET, F. Keras. 2015. Available at: https://keras.io. Accessed on 12/16/2024. Accessed on: May 25, 2023.

CULLIS, B. R.; SMITH, A. B.; COOMBES, N. E. On the design of early generation variety trials with correlated data. **Journal of agricultural, biological, and environmental statistics**, [*S. l.*], v. 11, p. 381-393, 2006.

DIETTERICH, T. G. Ensemble Methods in Machine Learning. **Proceedings of the First International Workshop on Multiple Classifier Systems - Springer**, p. 1-15, 2000.

DILAWARI, R.; KAUR, N.; PRIYADARSHI, N.; PRAKASH, I.; PATRA, A.; MEHTA, S.; SINGH, B.; JAIN, P.; ISLAM, M. A. Soybean: A Key Player for Global Food Security. In: WANI, S. H.; SOFI, N. U. R.; BHAT, M. A.; LIN, F. (Eds.). **Soybean Improvement: Physiological, Molecular and Genetic Perspectives.** Cham: Springer International Publishing, 2022. p. 1–46.

FALCONER, D. S.; MACKAY, T. F. C. **Introdução à Genética Quantitativa.** Reino Unido: Grupo Longman, Essex, 1996, 464 p.

FEHR, W. R.; CAVINESS, C. E. Stages of soybean development. Ames, Yowa: Yowa State University of Science and Technology, **Cooperative Extension Service**, 1977. 11 p. (Special Report, n. 80).

FRITSCHE-NETO R.; BORÉM A. **Phenomics How Next-Generation Phenotyping is Revolutionizing Plant Breeding.** Switzerland: Springer Press, 2015.

GALLI, G.; HORNE, D. W.; FRITSCHE-NETO, R.; AND ROONEY, W. L. Optimization of UAS-based high-throughput phenotyping to estimate plant health and grain yield in sorghum. **Plant Phenom.** J. [*S. l.*], v. 3, p. 1-14, 2020.

GILL, T.; GILL, S. K.; SAINI, D. K.; CHOPRA, Y.; DE KOFF, J. P.; SANDHU, K. S. A comprehensive review of high throughput phenotyping and machine learning for plant stress phenotyping. **Phenomics**, [*S. l.*], v. 2, n. 3, p. 156-183, 2022.

GITELSON, A. A.; MERZLYAK, M. N. Quantitative estimation of chlorophyll-a using reflectance spectra: experiments with autumn chestnut and maple leaves. **Journal of Photochemistry and Photobiology B: Biology**, [S. l.], v. 22, n. 3, p. 247-252, 1994.

GITELSON, A. A.; KAUFMAN, Y. J.; MERZLYAK, M. N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. **Remote Sensing of Environment**, [S. l.], v. 58, n. 3, p. 289-298, 1996.

HASSANIJALILIAN, O.; IGATHINATHANE, C.; BAJWA, S.; NOWATZKI, J. Rating iron deficiency in soybean using image processing and decision-tree based models. **Remote Sensing,** [*S. l.*], v. 12, p. 4143, 2020.

HERRERO-HUERTA, M.; RODRIGUEZ-GONZALVEZ, P.; RAINEY, K.M. Yield prediction by machine learning from UAS-based multi-sensor data fusion in soybean. **Plant Methods**, [*S. l.*], v. 16, n. 78, p. 1-16, 2020.

IQBAL, F., LUCIEER, A.; BARRY, K. Simplified radiometric calibration for UAS-mounted multispectral sensor. **Eur. J. Remote Sens.** [*S. l.*], v. 51, p. 301-313, 2018.

LIMA, G. S. A.; FERREIRA, M. E.; MADARI, B. E.; CARVALHO, M. T. M. Carbon estimation in an integrated crop-livestock system with imaging sensors aboard unmanned aerial platforms. **Remote Sensing Applications: Society and Environment**, v. 28, p. 100867, 2022.

LU, W.; DU, R.; NIU, P.; XING, G.; LUO, H.; DENG, Y.; SHU, L. Soybean yield preharvest prediction based on bean pods and leaves image recognition using deep learning neural network combined with GRNN**. Frontiers in Plant Science**, v. 12, p. 791256, 2022.

MACEDO, F. L.; NÓBREGA, H.; FREITAS, J. G. R.; RAGONEZI, C.; PINTO, L.; ROSA, J.; PINHEIRO DE CARVALHO, M. A. A. Estimation of productivity and above-ground biomass for corn (*Zea mays*) via vegetation indices in Madeira Island. **Agriculture**, [*S. l.*], v. 13, n. 6, p. 1115, 2023.

MANITOBA PULSE; SOYBEAN GROWERS. **Soybean Plant Development**. 2018. Available in: https://www.manitobapulse.ca/2016/08/soybean-plant-development/. Adapted. Accessed on: October 30, 2023.

MATIAS, F. I.; CARAZA-HARTER, M. V.; ENDELMAN, J. B. FIELDimageR: An R package to analyze orthomosaic images from agricultural field trials. **The Plant Phenome Journal**, [*S. l.*], v. 3, n. 1, p. e20005, 2020.

MOEINIZADE, S.; PHAM, H.; HAN, Y.; DOBBELS, A.; HU, G. An Applied Deep Learning Approach for Estimating Soybean Relative Maturity from UAV Imagery to Aid Plant Breeding Decisions. **Machine Learning with Applications**, [*S. l.*], v. 7, p. e10023, 2022.

NAIK, H.S.; ZHANG, J.; LOFQUIST, A.; ASSEFA, T.; SARKAR, S.; ACKERMAN, D.; SINGH, A.; SINGH, A.K.; GANAPATHYSUBRAMANIAN, B. A real-time phenotyping framework using machine learning for plant stress severity rating in soybean. **Plant Methods.** [*S. l.*], v. 13, n. 23, p. 1-12, 2017.

NAKATSU, R. T. Validation of machine learning ridge regression models using Monte Carlo, bootstrap, and variations in cross-validation. **Journal of Intelligent Systems**, [*S. l.*], v. 32, n. 1, pp. 20220224, 2023.

NARAYANAN, B.; FLOYD, B.; TU, K.; RIES, L.; HAUSMANN, N. **Improving soybean breeding using UAS measurements of physiological maturity.** Proc. SPIE 11008, Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping IV, 110080U. 2019.

OLIVOTO, T. Lights, camera, pliman! An R package for plant image analysis. **Methods in Ecology and Evolution.** v. 13, n. 4, p. 789-798, 2022.

PAZHANIVELAN, S. Quantification of Biophysical Parameters and Economic Yield in Cotton and Rice Using Drone Technology. **Agriculture,** [*S. l.*], v. 13, n. 9, p. 1668, 2023.

RANĐELOVIĆ, P.; ĐORĐEVIĆ, V.; MILADINOVIĆ, J.; PRODANOVIĆ, S.; ĆERAN, M.; VOLLMANN, J. High-throughput phenotyping for non-destructive estimation of soybean fresh biomass using a machine learning model and temporal UAV data. **Plant Methods,** [*S. l.*], v. 19, n. 89, 2023.

REN, P.; LI, H.; HAN, S.; CHEN, R.; YANG, G.; YANG, H.; FENG, H.; ZHAO, C. Estimation of Soybean Yield by Combining Maturity Group Information and Unmanned Aerial Vehicle Multi-Sensor Data Using Machine Learning. **Remote Sensitive,** [*S. l.*], v. 15, n. 17, p. 4286, 2023.

RESENDE, R. T.; HICKEY, L.; AMARAL, C. H.; PEIXOTO, L. L.; MARCATTI, G. E.; XU, Y. Satellite-enabled Enviromics to Enhance Crop Improvement. **Molecular Plant**., [*S.* l.], v. 17, n. 6, p. 848-866, 2024.

RODRÍGUEZ-ÁLVAREZ, M. X.; BOER, M. P.; EEUWIJK, F. A.; EILERS, P. H. C. Correcting for spatial heterogeneity in plant breeding experiments with P-splines. **Spatial Statistics,** v. 23, p. 52-71, 2018.

ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W. Monitoring vegetation systems in the Great Plains with ERTS. **Third ERTS-1 Symposium**, Washington, DC: NASA SP-351, p. 309-317, 1974.

SANTANA, D. C.; CUNHA, M. P. O.; SANTOS, R. G.; COTRIM, M. F.; TEODORO, L. P. R.; SILVA JUNIOR, C. A.; BAIO, F. H. R.; TEODORO, P. E. High-throughput phenotyping allows the selection of soybean genotypes for earliness and high grain yield. **Plant Methods**, [*S. l.*], v. 18, n. 13, 2022.

SILVA JUNIOR, C. A.; LEONEL-JUNIOR, A. H. S.; ROSSI, F. S.; CORREIA FILHO, W. L. F.; SANTIAGO, D. B.; OLIVEIRA-JÚNIOR, J. F.; TEODORO, P. E.; LIMA, M.; CAPRISTO-SILVA, G. F. Mapping soybean planting area in midwest Brazil with remotely sensed images and phenology-based algorithm using the Google Earth Engine platform. **Comput. Electron. Agric.,** [*S. l.*], v. 169, p. 105194, 2020.

SILVA, F.; BORÉM, A.; SEDIYAMA, T.; CÂMARA, G. **Soja: do plantio a colheita**. 2. ed. São Paulo: Oficina de textos, 2022.

SILVA, C. M. e; MEZZOMO, H. C.; RIBEIRO, J. P. O.; SIGNORINI, V. S.; LIMA, G. W.; VIEIRA, E. F. T., PORTES, M. F., MOROTA, G., CORREDO, L. P., NARDINO, M. Insights on multi-spectral vegetation indices derived from UAV-based high-throughput phenotyping for indirect selection in tropical wheat breeding. **Euphytica,** [*S. l.*], v. 220, n. 35, 2024.

SINGH, A. K.; GANAPATHYSUBRAMANIAN, B.; SARKAR, S.; SINGH, A. Deep learning for plant stress phenotyping: trends and future perspectives. **Trends Plant Sci.** [*S. l.*], v. 23, p. 883-898, 2018.

SINGH M.; VATSA V. K. Genetic analysis of yield and its component in soybean [Glycine max (L.) Merrill]. **Vegetos**, [*S. l.*], v. 22, p. 91-96, 2009.

SZEGHALMY, S.; FAZEKAS, A. A. Comparative Study of the Use of Stratified Cross-Validation and Distribution-Balanced Stratified Cross-Validation in Imbalanced Learning. **Sensors**, [*S. l.*], v. 23, n. 4, p. 2333, 2023.

VENANCIO, L. P.; MANTOVANI, E. C.; AMARAL, C. H.; NEALE, C. M. U.; GONÇALVES, I. Z.; FILGUEIRAS, R.; EUGENIO, F. C. Potential of using spectral vegetation indices for corn green biomass estimation based on their relationship with the photosynthetic vegetation sub-pixel fraction. **Agricultural Water Management**, [*S. l.*], v. 236, p. 106155, 2020.

TAYADE, R.; YOON, J.; LAY, L.; KHAN, A. L.; YOON, Y.; KIM, Y. Utilization of Spectral Indices for High-Throughput Phenotyping. **Plants**, [S. l.], v. 11, n. 13, article 1712, 2022.

TEODORO, P. E.; TEODORO, L. P. R.; BAIO, F. H. R.; SILVA JUNIOR, C. A. DA.; SANTOS, R. G. DOS.; RAMOS, A. P. M.; PINHEIRO, M. M. F.; OSCO, L. P.; GONÇALVES, W. N.; CARNEIRO, A. M.; MARCATO JUNIOR, J.; PISTORI, H.; SHIRATSUCHI, L. S. Predicting Days to Maturity, Plant Height, and Grain Yield in Soybean: A Machine and Deep Learning Approach Using Multispectral Data. **Remote Sensitive**, v. 13, n. 22, p. 4632, 2021.

TREVISAN, R.; PÉREZ, O.; SCHMITZ, N.; DIERS, B.; MARTIN, N. High-throughput phenotyping of soybean maturity using time series UAV imagery and convolutional neural networks. **Remote Sensing**, [*S. l.*], v. 12, n. 21, p. 3617, 2020.

TUCKER, C. J. Red and photographic infrared linear combinations for monitoring vegetation. **Remote Sensing of Environment**, [S. l.], v. 8, n. 2, p. 127-150, 1979.

VOLPATO, L.; DOBBELS, A.; BORÉM, A.; LORENZ, A. Optimization of temporal UAS-based imagery analysis to estimate plant maturity date for soybean breeding. **The Plant Phenome Journal**, [*S. l.*], v. 4, n. 1, p. e20018, 2021.

YOOSEFZADEH NAJAFABADI, M.; EARL, H.; TULPAN, D.; SULIK, J.; ESKANDARI, M. Application of Machine Learning Algorithms in Plant Breeding: Predicting Yield from Hyperspectral Reflectance in Soybean. **Frontiers in Plant Science**, [*S. l.*], v. 11, p. 624273, 2021a.

YOOSEFZADEH NAJAFABADI, M.; TULPAN, D.; ESKANDARI, M. Using hybrid artificial intelligence and evolutionary optimization algorithms for estimating soybean yield and fresh biomass using hyperspectral vegetation indices. **Remote Sensing**, [S. l.], v. 13, n. 13, p. 2555, 2021b.

ZHONG, L.; HU, L.; YU, L.; GONG, P.; BIGING, G. S. Biging. Automated mapping of soybean and corn using phenology. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 119, p. 151-164, 2016.

ZHOU, J.; YUNGBLUTH, D.; VONG, C.N.; SCABOO, A.; ZHOU, J. Estimation of the Maturity Date of Soybean Breeding Lines Using UAV-Based Multispectral Imagery. **Remote Sens**. [*S. l.*], v. 11, n. 18, p. 2075, 2019.

ZHU, R.; WANG, X.; YAN, Z.; QIAO, Y.; TIAN, H.; HU, Z.; ZHANG, Z.; LI, Y.; ZHAO, H.; XIN, D.; CHEN, Q. Exploring Soybean Flower and Pod Variation Patterns During Reproductive Period Based on Fusion Deep Learning. **Front. Plant Sci.** [*S. l.*], v. 13, n. 922030, 2022.

2.9 Supplementary material

Table S1 - Soil chemical and physical properties before experiment installation

Experimental Block	pH in water	K	P	Zn	CTC	SB	MO
		---------- mg dm^{-3} ----------			---------- % ----------		-- d dm^{-3} --
	5.3	60.3	70	6.6	7.65	67.32	29
P3RR	**Ca**	**Mg**	**H+Al**	**Clay**	**Silte**		**Sand**
	---------- $cmol_c$ dm^{-3} ----------			---------- % ----------			
	3.5	1.5	2.5	500	110		390
Experimental Block	**pH in water**	**K**	**P**	**Zn**	**CTC**	**SB**	**MO**
		---------- mg dm^{-3} --------			---------- % ----------		-- d dm^{-3} --
	5.4	62.1	108	5.4	6.56	57.32	33
P3Bt	**Ca**	**Mg**	**H+Al**	**Clay**	**Silte**		**Sand**
	---------- $cmol_c$ dm^{-3} ----------			---------- % ----------			
	2.6	1	2.8	490	100		410
Experimental Block	**pH in water**	**K**	**P**	**Zn**	**CTC**	**SB**	**MO**
		---------- mg dm^{-3} ----------			---------- % ----------		-- d dm^{-3} --
	5.8	56.6	30.8	5.5	7.14	71.99	29
CFRR	**Ca**	**Mg**	**H+Al**	**Clay**	**Silte**		**Sand**
	---------- $cmol_c$ dm^{-3} ----------			---------- % ----------			
	3.5	1.5	2	480	100		420
Experimental Block	**pH in water**	**K**	**P**	**Zn**	**CTC**	**SB**	**MO**
		---------- mg dm^{-3} --------			---------- % ----------		-- d dm^{-3} --
	5.2	82.6	127	5.4	7.11	57.81	31
CFBt	**Ca**	**Mg**	**H+Al**	**Argila**	**Silte**		**Areia**
	---------- $cmol_c$ dm^{-3} ----------			---------- % ----------			
	2.9	1	3	510	120		370

Source: Author (2024).

Caption: SB: Sum of Bases; OM: Organic Matter; CEC: Total Cation Exchange Capacity; Ca: Calcium; Mg: Magnesium; H+Al: Hydrogen and Aluminum (exchangeable acidity); pH in water; K: Potassium; P: Phosphorus; Zn: Zinc. Source: Solocria Agricultural Laboratory.

Table S2 - Tukey HSD test comparing correlation values across seven flight days (35, 60, 70, 90, 95, 100, and 110 DAS), vegetation indices (NDVI, GNDVI, NDRE, and NGRDI), and agronomic traits (flowering, maturity, and grain yield)

Flights	Correlation		VIs	Correlation		Trait	Correlation	
95	0.629459	a	NDRE	0.4839955	a	MAT	0.5632729	a
100	0.5884569	a	NDVI	0.4828097	a	GY	0.4509365	b
90	0.5725732	a	NGRDI	0.4633981	a	FLOW	0.3627818	c
70	0.5598798	a	GNDVI	0.4057849	a			
60	0.4127896	b						
110	0.2902587	bc						
35	0.1595623	c						

Source: Author (2024).
Caption: Groups with different letters are statistically different at a probability level of 0.05. FLOW: Flowering; MAT: Maturity; GY: Grain Yield.

Table S3 - Confusion matrix results for classification performance of ML models

(continue)

Flight Days	Index	Models							
		XGBoost				Random Forest			
FLY_35	NDVI		Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
		0	169	120	128	0	111	109	111
		1	95	146	119	1	105	102	111
		2	61	71	95	2	109	126	120
	GNDVI		Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
		0	154	134	124	0	111	116	117
		1	97	116	109	1	101	96	101
		2	74	87	109	2	113	125	124
	NDRE		Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
		0	137	121	125	0	107	116	118
		1	122	149	120	1	113	124	98
		2	66	67	97	2	105	97	126
	NGRDI		Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
		0	124	93	116	0	109	101	111
		1	110	159	124	1	105	141	106
		2	91	85	102	2	111	95	125
FLY_60	NDVI		Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
		0	176	117	79	0	143	94	88
		1	57	81	56	1	97	119	110
		2	92	139	207	2	85	124	144
	GNDVI		Observed				Observed		

Table S3 - Confusion matrix results for classification performance of ML models

(continuation)

Flight Days	Index	Models							
		XGBoost				Random Forest			
		Predicted	0	1	2	Predicted	0	1	2
		0	77	108	108	0	107	100	113
		1	118	115	121	1	92	114	116
		2	130	114	113	2	126	123	113
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NDRE	0	88	104	99	0	102	106	113
		1	92	96	93	1	113	118	114
		2	145	137	150	2	110	113	115
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NGRDI	0	246	136	40	0	174	113	40
		1	47	63	59	1	114	122	107
		2	32	138	243	2	37	102	195
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NDVI	0	221	139	54	0	159	106	51
		1	58	84	46	1	102	119	111
		2	46	114	242	2	64	112	180
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	GNDVI	0	209	122	27	0	169	108	59
		1	81	117	83	1	104	139	79
FLY_70		2	35	98	232	2	52	90	204
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NDRE	0	173	120	44	0	141	119	73
		1	90	97	87	1	123	108	103
		2	62	120	211	2	61	110	166
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NGRDI	0	100	96	103	0	119	110	106
		1	99	103	63	1	95	119	106
		2	126	138	176	2	111	108	130
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NDVI	0	256	58	0	0	225	89	5
FLY_90		1	68	212	41	1	90	168	63
		2	1	67	301	2	5	66	268
	GNDVI		Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2

Table S3 - Confusion matrix results for classification performance of ML models

(continuation)

Flight Days	Index	Models							
		XGBoost				Random Forest			
		0	271	70	0	0	228	88	2
		1	53	231	46	1	89	188	41
		2	1	36	296	2	3	47	293
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NDRE	0	209	99	3	0	189	114	16
		1	105	175	65	1	110	136	71
		2	11	63	274	2	21	73	249
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NGRDI	0	270	67	4	0	241	74	17
		1	39	112	61	1	64	141	116
		2	16	158	277	2	15	108	203
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NDVI	0	266	67	0	0	247	80	1
		1	59	239	36	1	78	219	40
		2	0	31	306	2	0	38	301
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	GNDVI	0	255	120	1	0	204	110	2
		1	70	191	28	1	119	188	30
FLY_95		2	0	26	313	2	2	39	310
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NDRE	0	219	156	21	0	146	147	25
		1	103	159	46	1	148	137	40
		2	3	22	275	2	31	53	277
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NGRDI	0	274	67	0	0	249	82	0
		1	51	235	40	1	76	205	47
		2	0	35	302	2	0	50	295
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NDVI	0	192	64	1	0	200	116	9
FLY_100		1	130	239	44	1	116	171	41
		2	3	34	297	2	9	50	292
	GNDVI		Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2

Table S3 - Confusion matrix results for classification performance of ML models

(conclusion)

Flight Days	Index	Models							
		XGBoost				Random Forest			
		0	193	90	5	0	185	141	25
		1	122	199	52	1	121	146	51
		2	10	48	285	2	19	50	266
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NDRE	0	174	124	62	0	146	120	78
		1	114	144	102	1	111	129	101
		2	37	69	178	2	68	88	163
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NGRDI	0	193	138	12	0	169	132	14
		1	121	169	43	1	141	157	58
		2	11	30	287	2	15	48	270
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NDVI	0	70	88	62	0	98	112	114
		1	111	145	81	1	104	131	102
		2	144	104	199	2	123	94	126
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	GNDVI	0	97	91	103	0	119	94	102
		1	113	167	97	1	93	141	114
FLY_110		2	115	79	142	2	113	102	126
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NDRE	0	148	136	69	0	130	124	77
		1	114	112	70	1	122	109	94
		2	63	89	203	2	73	104	171
			Observed				Observed		
		Predicted	0	1	2	Predicted	0	1	2
	NGRDI	0	109	86	90	0	116	113	105
		1	95	107	102	1	108	106	130
		2	121	144	150	2	101	118	107

Source: Author (2024).