



**OCTÁVIO PEREIRA DA COSTA**

**IMPACTO DA CORREÇÃO ATMOSFÉRICA EM IMAGENS  
DE SATÉLITE NA PREVISÃO DE PRODUTIVIDADE  
UTILIZANDO APRENDIZADO DE MÁQUINA**

**LAVRAS – MG  
2024**

**OCTÁVIO PEREIRA DA COSTA**

**IMPACTO DA CORREÇÃO ATMOSFÉRICA EM IMAGENS DE SATÉLITE NA  
PREVISÃO DE PRODUTIVIDADE UTILIZANDO APRENDIZADO DE MÁQUINA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agronomia/Fitotecnia, área de concentração Produção Vegetal, para a obtenção do título de Mestre.

Prof. Dr. Adão Felipe dos Santos  
Orientadora

Prof (a). Dra. Lorena Lacerda Nunes  
Coorientadora

**LAVRAS – MG  
2024**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca  
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Costa, Octávio Pereira da.

Impacto da correção atmosférica em imagens de satélite na previsão  
de produtividade utilizando aprendizado de máquina / Octávio Pereira da  
Costa. - 2024.

66 p. : il.

Orientador(a): Adão Felipe dos Santos.

Coorientador(a): Lorena Lacerda Nunes.

Dissertação (mestrado acadêmico) - Universidade Federal de Lavras,  
2024.

Bibliografia.

1. Agricultura de precisão. 2. Sensoriamento remoto. 3.  
Monitoramento agrícola. I. Santos, Adão Felipe dos. II. Nunes, Lorena  
Lacerda. III. Título.

**OCTÁVIO PEREIRA DA COSTA**

**IMPACTO DA CORREÇÃO ATMOSFÉRICA EM IMAGENS DE SATÉLITE NA  
PREVISÃO DE PRODUTIVIDADE UTILIZANDO APRENDIZADO DE MÁQUINA**

**IMPACT OF ATMOSPHERIC CORRECTION IN SATELLITE IMAGERY ON  
PRODUCTIVITY PREDICTION USING MACHINE LEARNING**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agronomia/Fitotecnia, área de concentração em Produção Vegetal, para a obtenção do título de Mestre.

APROVADA em 16 de agosto de 2024.

Dr. Gabriel Araujo e Silva Ferraz UFLA

Dr. Luan Pereira de Oliveira UGA

Prof. Dr. Adão Felipe dos Santos  
Orientadora

Prof (a). Dra. Lorena Lacerda Nunes  
Coorientadora

**LAVRAS – MG  
2024**

## AGRADECIMENTOS

A conclusão desta dissertação só foi possível graças ao apoio de muitas pessoas que estiveram ao meu lado ao longo dessa jornada.

Primeiramente, agradeço à minha família, por todo amor, compreensão e incentivo em cada etapa do meu percurso acadêmico. Especialmente ao meu pai, João Roberto, pois sem o seu suporte incondicional, essa conquista não teria sido possível.

À minha namorada, Anna Carolina, que esteve ao meu lado nos momentos mais difíceis, oferecendo carinho, paciência e apoio. Sua presença foi fundamental para que eu pudesse chegar até aqui.

Expresso minha profunda gratidão ao meu orientador, Prof. Dr. Adão Felipe dos Santos, por sua orientação, paciência e por compartilhar seu conhecimento comigo. Seus conselhos foram essenciais para o desenvolvimento deste trabalho.

Aos membros do Grupo de Pesquisa e Extensão em Agricultura Digital, agradeço pela troca de conhecimentos, experiências e pelo suporte contínuo. O trabalho em equipe e as discussões produtivas contribuíram muito para o enriquecimento desta pesquisa.

Agradeço aos moradores da República Mete o Pé, pelo convívio diário, pelas risadas e pelo apoio nos momentos de descontração e nas horas de tensão. Vocês tornaram essa jornada mais leve e prazerosa.

Agradeço à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio e financiamento deste trabalho.

## RESUMO GERAL

O cultivo do milho (*Zea mays* L.) no Brasil é crucial para a segurança alimentar, e a agricultura de precisão, que usa sensoriamento remoto, pode otimizar o manejo agrícola. Porém a utilização de imagens orbitais apresenta diversos desafios, entre eles a disponibilidade de imagens com ausência de nuvens e a interferência da atmosfera na radiação captada pelos sensores. Métodos de correção atmosférica foram desenvolvidas por diversos autores, entretanto o impactos dos métodos nos monitoramentos contínuo e o impacto em técnicas avançadas de processamento de dados ainda é pouco estudado. Este estudo avalia três métodos de correção atmosférica, Dark Object Subtraction (DOS), Image Correction for Atmospheric Effects (iCOR), e Sentinel-2 Correction (Sen2Cor), e também imagens com processamento LIC que são dados referentes a radiação no Topo da Atmosfera (TOA) para melhorar a qualidade dos dados de imagens orbitais gratuitas do satélite Sentinel-2 na caracterização do ciclo do milho e predição de produtividade utilizando índices de vegetação. A análise demonstrou que a correção iCOR é o método mais eficaz na caracterização do ciclo do milho quando utilizado o NDVI e GNDVI. Já para a predição de produtividade, as correções DOS e Sen2Cor foram as mais eficazes em todos os algoritmos testados, Random Forest (RF), k-Nearest Neighbors (kNN), e Support Vector Machine (SVM), com o RF apresentando o melhor desempenho, especialmente na safrinha, com um  $R^2$  médio de 0,82. A escolha adequada do método de correção atmosférica e do índice de vegetação é essencial para aumentar a precisão no monitoramento e previsão de produtividade, contribuindo para uma gestão agrícola mais eficiente e sustentável.

**Palavras-chave:** Agricultura de precisão; inteligência artificial; sensoriamento remoto; monitoramento agrícola.

## GENERAL ABSTRACT

The cultivation of corn (*Zea mays* L.) in Brazil is crucial for food security, and precision agriculture, which uses remote sensing, can optimize agricultural management. However, the use of orbital images presents several challenges, including the availability of images with the absence of clouds and the interference of the atmosphere in the radiation captured by the sensors. Atmospheric correction methods have been developed by several authors, however the impact of the methods on continuous monitoring and the impact on advanced data processing techniques is still poorly studied. This study evaluates three atmospheric correction methods, Dark Object Subtraction (DOS), Image Correction for Atmospheric Effects (iCOR), and Sentinel-2 Correction (Sen2Cor), as well as images with L1C processing that are data referring to radiation at the Top of the Atmosphere (TOA) to improve the quality of free orbital imagery data from the Sentinel-2 satellite in the characterization of the corn cycle and yield prediction using vegetation indices. The analysis showed that the iCOR correction is the most effective method in characterizing the corn cycle when using NDVI and GNDVI. For yield prediction, the DOS and Sen2Cor corrections were the most effective in all the algorithms tested, Random Forest (RF), k-Nearest Neighbors (kNN), and Support Vector Machine (SVM), with RF presenting the best performance, especially in the off-season, with an average  $R^2$  of 0.82. The proper choice of atmospheric correction method and vegetation index is essential to increase accuracy in monitoring and forecasting productivity, contributing to more efficient and sustainable agricultural management.

**Keywords:** Precision agriculture; artificial intelligence; remote sensing; agricultural monitoring.

## **INDICADORES DE IMPACTO**

Este trabalho investiga o impacto das correções atmosféricas em imagens de satélite no monitoramento e previsão de produtividade na cultura do milho, visando avanços tecnológicos significativos no campo da agricultura de precisão. A correção atmosférica é essencial para aumentar a acurácia das imagens orbitais usadas na análise do desenvolvimento do milho, particularmente em relação a índices de vegetação que servem como indicadores do estado fisiológico da planta. Este estudo potencializa a tomada de decisão para produtores e gestores agrícolas ao fornecer dados confiáveis sobre produtividade, permitindo intervenções mais precisas e sustentáveis. Assim, os resultados deste trabalho alinham-se ao Objetivo de Desenvolvimento Sustentável (ODS) 2, “Fome Zero e Agricultura Sustentável,” promovendo práticas agrícolas mais sustentáveis e eficientes. A pesquisa também contribui para o ODS 12, “Consumo e Produção Responsáveis,” ao otimizar o uso de recursos naturais e insumos agrícolas. Em termos de impacto tecnológico, a metodologia pode ser implementada em diversas regiões agrícolas brasileiras, beneficiando diretamente produtores ao reduzir incertezas nas previsões de produtividade e permitindo um manejo mais preciso e econômico da cultura do milho. A aplicação da tecnologia também atende à área temática de "Meio Ambiente" da extensão universitária, promovendo práticas agrícolas que minimizam impactos ambientais. Além disso, o trabalho envolve a participação da comunidade acadêmica e de produtores locais, reforçando o caráter extensionista e o impacto direto junto a esses grupos, enquanto promove a difusão de práticas tecnológicas avançadas para o setor agrícola.

## **IMPACT INDICATORS**

This study investigates the impact of atmospheric corrections on satellite images for monitoring and predicting maize crop productivity, aiming for significant technological advancements in precision agriculture. Atmospheric correction is essential to increase the accuracy of satellite images used in analyzing maize development, particularly regarding vegetation indices that serve as indicators of the plant's physiological state. This research enhances decision-making for farmers and agricultural managers by providing reliable productivity data, enabling more precise and sustainable interventions. Thus, the results of this study align with Sustainable Development Goal (SDG) 2, “Zero Hunger and Sustainable Agriculture,” promoting more sustainable and efficient agricultural practices. The research also contributes to SDG 12, “Responsible Consumption and Production,” by optimizing the use of natural resources and agricultural inputs. In terms of technological impact, the methodology can be implemented in various Brazilian agricultural regions, directly benefiting producers by reducing uncertainties in productivity predictions and enabling more precise and economical maize crop management. The application of this technology also meets the "Environment" thematic area of university extension, promoting agricultural practices that minimize environmental impacts. Additionally, the study involves the participation of the academic community and local producers, reinforcing its extensionist nature and direct impact on these groups, while promoting the dissemination of advanced technological practices within the agricultural sector.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO GERAL</b>	<b>9</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>11</b>
<b>2.1</b>	<b>A cultura do milho</b>	<b>11</b>
<b>2.2</b>	<b>Sensoriamento remoto</b>	<b>12</b>
<b>2.3</b>	<b>Correção atmosférica de imagens de satélite</b>	<b>14</b>
<b>2.4</b>	<b>Índices de vegetação</b>	<b>15</b>
<b>2.5</b>	<b>Inteligência artificial</b>	<b>17</b>
	<b>REFERÊNCIAS</b>	<b>21</b>
	<b>ARTIGO 1</b>	<b>26</b>
	<b>Impactos da correção de imagens de satélite no desempenho dos índices de reflectância e vegetação sobre a cultura do milho</b>	<b>26</b>
<b>1</b>	<b>INTRODUÇÃO</b>	<b>28</b>
<b>2</b>	<b>MATERIAL E MÉTODOS</b>	<b>29</b>
<b>3</b>	<b>RESULTADOS</b>	<b>32</b>
<b>4</b>	<b>DISCUSSÃO</b>	<b>38</b>
<b>5</b>	<b>CONCLUSÃO</b>	<b>40</b>
	<b>REFERÊNCIAS</b>	<b>41</b>
	<b>ARTIGO 2</b>	<b>45</b>
	<b>Impacto de correções atmosféricas em imagens de satélite na previsão de produtividade de milho utilizando aprendizado de máquina</b>	<b>45</b>
<b>1</b>	<b>INTRODUÇÃO</b>	<b>47</b>
<b>2</b>	<b>MATERIAL E MÉTODOS</b>	<b>48</b>
<b>3</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>53</b>
<b>4</b>	<b>CONCLUSÕES</b>	<b>60</b>
	<b>REFERÊNCIAS</b>	<b>61</b>
	<b>CONSIDERAÇÕES FINAIS</b>	<b>65</b>

## 1 INTRODUÇÃO GERAL

Na era da agricultura de precisão, o uso de tecnologias avançadas, como sensoriamento remoto e aprendizado de máquina, está transformando a forma como monitoramos e gerenciamos as culturas agrícolas. Índices de vegetação, como o Índice de Vegetação por Diferença Normalizada (NDVI), o Índice de Vegetação por Diferença Normalizada Verde (GNDVI), o Índice de Vegetação Melhorado (EVI) e o Índice de Vegetação de Borda Vermelha por Diferença Normalizada (NDRE), são ferramentas cruciais para avaliar a saúde das plantas e prever a produtividade, oferecendo informações detalhadas que facilitam a tomada de decisões mais informadas (Paul *et al.*, 2022; Brambilla *et al.*, 2021; Paustian; Theuvsen, 2017). No entanto, a precisão desses índices pode ser comprometida por distorções atmosféricas, que introduzem ruídos nos dados capturados por sensores de satélite (Hadjimitsis *et al.*, 2010).

Métodos de correção atmosférica, como a Subtração de Objeto Escuro (DOS), Sen2Cor e iCOR, são essenciais para mitigar esses efeitos e melhorar a acurácia dos índices de vegetação (Lantzanakis; Mitraka; Chrysoulakis, 2017; Keukelaere *et al.*, 2018). Essas correções são particularmente importantes em regiões com alta densidade de vegetação, onde o fenômeno da saturação pode limitar a capacidade dos índices de distinguir diferentes níveis de biomassa (Huete *et al.*, 1997; Miller; Shober; Taraila, 2024). Além disso, a qualidade dos dados utilizados é fundamental para o desempenho de modelos de aprendizado de máquina, que têm se mostrado eficazes em tarefas como classificação de culturas e previsão de rendimento (Baburao *et al.*, 2023; Thilakarathne *et al.*, 2022).

A integração de técnicas de aprendizado de máquina com dados de sensoriamento remoto amplia a capacidade de análise em grande escala, permitindo a identificação de padrões e tendências que não são visíveis por métodos tradicionais (Diaz-Gonzalez *et al.*, 2022). No entanto, garantir que os dados sejam corretamente corrigidos e processados é crucial para desenvolver modelos que possam prever rendimentos de culturas com precisão e auxiliar na gestão eficiente dos recursos agrícolas (Sessions; Valtorta, 2006; Gómez, 2020; Guimarães, 2019).

Este trabalho visa explorar como diferentes métodos de correção atmosférica impactam a precisão dos índices de vegetação e a capacidade dos algoritmos de aprendizado de máquina, como Random Forest (RF), Support Vector Machines (SVM) e k-Nearest Neighbors (kNN), em prever a produtividade do milho em duas temporadas de cultivo distintas. Ao analisar essas correções e suas influências sobre os modelos de machine learning, buscamos fornecer insights que possam aprimorar as técnicas de monitoramento agrícola e contribuir para uma gestão mais

sustentável e eficaz da agricultura. Compreender essas dinâmicas é essencial para melhorar a precisão das previsões e otimizar a aplicação de recursos agrícolas, ajudando agricultores a tomar decisões mais informadas e eficazes.

## 2 REFERENCIAL TEÓRICO

### 2.1 A cultura do milho

O milho (*Zea mays* L.) é uma planta da família Poaceae, originária da América Central, com domesticação iniciada há cerca de 5.000 anos (Teixeira; Costa, 2010). Atualmente, o Brasil é o terceiro maior produtor mundial de milho, atrás apenas dos Estados Unidos e da China (Lopes *et al.*, 2021). A cultura tem grande importância econômica, sendo utilizada na alimentação humana e animal, produção de biocombustíveis e outros produtos industriais (Silva *et al.*, 2020). O milho apresenta ampla variabilidade genética devido à sua dispersão pelas Américas e adaptações locais (Teixeira; Costa, 2010). No Brasil, seu cultivo remonta aos povos pré-colombianos, tendo sido descrito por colonizadores no século XVI (Teixeira; Costa, 2010). A adoção de práticas agrícolas modernas, como o uso de fertilizantes sintéticos, herbicidas, inseticidas e mecanização, substituiu métodos tradicionais e impulsionou a produção (Galvão *et al.*, 2014). Além disso, estratégias de manejo, como correção do solo e agricultura de precisão, têm sido fundamentais para o aumento da produtividade (Custodio *et al.*, 2016).

A planta de milho (*Zea mays* L.) apresenta um ciclo de desenvolvimento bem definido, com fases vegetativa e reprodutiva cruciais para o rendimento de grãos (Durães *et al.*, 1995). Fatores ambientais influenciam significativamente os processos fisiológicos, acúmulo de matéria seca e produtividade da cultura (Medeiros, 2009). Um dos pontos cruciais é a eficiência no uso da radiação e da água, sendo fundamental para o desenvolvimento da cultura, com variações observadas em diferentes épocas de semeadura (Medeiros, 2009). O ciclo completo do milho pode variar entre 90 a 150 dias, dependendo da variedade e das condições ambientais (Silva *et al.*, 2020).

O cultivo do milho é altamente dependente das condições climáticas e do solo. As temperaturas ótimas para o crescimento variam de 18 °C a 27 °C, sendo que temperaturas extremas afetam negativamente o desenvolvimento da planta (Maldaner *et al.*, 2014; Silva *et al.*, 2006). O milho requer precipitação anual entre 500-800 mm, distribuída adequadamente ao longo do ciclo de crescimento, especialmente durante o final do estágio vegetativo, pendoamento e enchimento de grãos (Sans *et al.*, 2001). Fatores como disponibilidade de nutrientes, textura do solo e capacidade de retenção de água também influenciam o crescimento e a produtividade da cultura (Silva *et al.*, 2006). O zoneamento de risco climático é crucial para identificar regiões e épocas de plantio adequadas, considerando fatores como coeficientes de cultura, evapotranspiração, precipitação e disponibilidade de água no solo (Maldaner *et al.*, 2014; Sans *et al.*, 2001). Projeções de mudanças climáticas sugerem aumento de temperaturas

em regiões semiáridas, o que pode reduzir os ciclos das culturas e a evapotranspiração total, impactando a produção de milho no futuro (Junior *et al.*, 2018).

O conhecimento das práticas adequadas de manejo para as mais diferentes situações encontradas no território brasileiro é de fundamental importância, desta forma é necessário o produtor estar atento aos estágios de desenvolvimento da planta de milho. O conhecimento dos estágios fenológicos da planta de milho pode otimizar o manejo, e retornar maiores ganhos aos produtores. O fator que está intimamente ligado a produção é a época de semeadura, a precipitação e a temperatura impactam significativamente os rendimentos do milho, com maiores índices de precipitação e temperaturas durante a primeira colheita sendo fatores importantes para a produção (Maldaner *et al.*, 2014; Lopes *et al.*, 2019).

Porém, o cultivo de milho durante a primeira safra no Brasil enfrenta diversos desafios, especialmente em relação a pragas e doenças que podem causar perdas de rendimento superiores a 40%, tornando crucial entender suas condições e hábitos favoráveis para um manejo eficaz (Maia *et al.*, 2024). Entretanto, durante a safrinha, que ocorre após a colheita da soja, o cultivo do milho enfrenta condições climáticas diferentes, com temperaturas mais baixas e menor disponibilidade de água (Campos *et al.*, 2024; Shioga; Gerage, 2010). A adaptação às condições específicas de cada safra e o monitoramento contínuo são essenciais para garantir a produtividade e reduzir perdas na cultura do milho (Brito, 2020; Passos *et al.*, 2019).

Tendo se em vista a quantidade de fatores que interferem na produtividade da cultura do milho se faz necessário lançar mão de métodos para estimativa de produtividade. Uma vez que o produtor precisa realizar o planejamento de diversos processos, como o transporte, armazenamento e avaliação dos prováveis ganhos com a comercialização. Porém, mesmo com a importância a nível mundial o milho não possui técnicas precisas de estimativa de produtividade, os métodos mais utilizados atualmente podem superestimar a produção ou necessitam de adaptações aos diferentes cultivares presentes no mercado (Rodrigues *et al.*, 2005).

## **2.2 Sensoriamento remoto**

Sensoriamento remoto é uma tecnologia que permite obter informações sobre a superfície terrestre à distância, utilizando sensores que detectam radiação eletromagnética (Guedes; Silva, 2018). Essa técnica é amplamente aplicada no estudo de recursos agronômicos e florestais, contribuindo para o mapeamento e monitoramento desses recursos (Shimabukuro; Maeda; Formaggio, 2009). Na agricultura de precisão, o sensoriamento remoto tem sido utilizado tanto em nível terrestre, para caracterização espectral de solos e plantas, quanto orbital,

para estimativas de variações espaciais e temporais de áreas cultivadas e produtividade (Silva *et al.*, 2014). O desenvolvimento de novos sensores com melhores características espaciais, espectrais e temporais tem permitido um monitoramento cada vez mais detalhado dos recursos naturais (Shimabukuro; Maeda; Formaggio, 2009 ).

O sensoriamento remoto tornou-se uma ferramenta essencial na agricultura de precisão, oferecendo diversos métodos de monitoramento de culturas. Imagens de satélite e veículos aéreos não tripulados (VANTs) estão entre as técnicas de coleta de dados mais comuns (Shiratsuchi *et al.*, 2014). Esses métodos podem ser categorizados com base na plataforma e no tipo de sensor utilizado. O sensoriamento orbital é realizado por constelações de satélites equipados com diferentes sensores, enquanto o sensoriamento aéreo ou suborbital envolve VANTs com câmeras que capturam reflectância nos espectros visível e infravermelho (Silva *et al.*, 2014). O sensoriamento proximal utiliza sensores portáteis ou acoplados a veículos terrestres para monitoramento e coleta de dados no campo (Shiratsuchi *et al.*, 2014). VANTs, inicialmente usados para fins militares, foram adaptados para aplicações agrícolas, equipados com Sistemas Globais de Navegação por Satélite (GNSS) e sensores miniaturizados (Souza *et al.*, 2020). A integração dessas tecnologias levou ao desenvolvimento de plataformas agrícolas inteligentes, que processam e gerenciam dados de sensores e imagens de VANTs, melhorando a produtividade e os métodos de produção na agricultura de precisão (Silva; Hachisuca, 2019).

A interação entre a radiação solar e as características da superfície terrestre é influenciada pela sua composição físico-química, textura, densidade e posição relativa ao ângulo de incidência solar e à geometria de imagem (Cruz *et al.*, 2021). Essa interação resulta em assinaturas espectrais únicas para diferentes alvos, que são capturadas por sensores de satélite (Cruz *et al.*, 2021). A tecnologia de sensoriamento remoto avançou significativamente, fornecendo vastas quantidades de dados sobre processos e fenômenos de superfície (Guedes; Silva, 2018). A radiação eletromagnética (REM) serve como a ligação entre os objetos de superfície e os sensores remotos, com sistemas de imagem e não-imagem coletando dados em várias formas (Guedes; Silva, 2018).

Apesar dos avanços significativos no sensoriamento remoto, existem várias limitações no uso de imagens orbitais na agricultura de precisão. Isso ocorre principalmente devido aos altos custos de direitos de uso de imagens de alta resolução, restrições na aquisição de dados devido a fenômenos meteorológicos, interferência da atmosfera na interação com a radiação eletromagnética, falta de mão de obra qualificada para análises e deficiências em soluções tecnológicas para automatizar as análises. Esses fatores podem afetar a adoção em larga escala do sensoriamento remoto nas propriedades agrícolas (Katsigiannis *et al.*, 2016).

### 2.3 Correção atmosférica de imagens de satélite

Imagens de satélite frequentemente sofrem distorções radiométricas e geométricas devido a vários fatores, incluindo problemas de calibração do sensor, efeitos atmosféricos e condições de iluminação (Welter *et al.*, 2020; Leonardi; Souza; Fonseca, 2003). Essas distorções podem impactar significativamente a interpretação e análise dos dados, especialmente em estudos multitemporais (Leonardi; Souza; Fonseca, 2003). Para resolver essas questões, vários métodos de normalização e correção radiométrica foram desenvolvidos e comparados (Leonardi; Souza; Fonseca, 2003; Lima; Ribeiro, 2014). Essas técnicas visam reduzir os efeitos de sombreamento topográfico, interferência atmosférica e outras distorções, melhorando a precisão na extração de informações das imagens de satélite (Lima; Ribeiro, 2014).

Um dos principais desafios no processamento de imagens de satélite está relacionado às características das ondas eletromagnéticas e à interferência causada pela atmosfera. A radiação eletromagnética emitida pela luz solar passa por várias alterações devido às interações com partículas e moléculas de gás presentes na atmosfera, sendo as principais a dispersão e absorção (Hilker, 2018). Portanto, é de extrema importância realizar correções atmosféricas durante o processamento de imagens adquiridas por satélites, a fim de minimizar o impacto dos elementos atmosféricos e, conseqüentemente, aprimorar a precisão das informações contidas nas imagens orbitais (Batista; Dias, 2005; Mognane *et al.*, 2019). Entretanto, a escolha do método de correção e as especificações técnicas do sensor são fatores de extrema importância, pois impactam diretamente na eficácia do processo de correção (Okin; Gu, 2015).

Os métodos de correção atmosférica podem ser divididos em três categorias principais, classificadas de acordo com a abordagem utilizada: métodos físicos, métodos empíricos e métodos híbridos (Gaida *et al.*, 2020). Os métodos físicos são baseados na Teoria de Transferência Radiativa de Chandrasekhar (1960), conhecida como correção absoluta, que utiliza estimativas físicas de parâmetros atmosféricos. Os métodos empíricos são fundamentados em estimativas dos efeitos atmosféricos com base em informações de alvos específicos presentes na própria imagem, sendo geralmente mais simples, mas potencialmente menos precisos. Os métodos híbridos representam uma combinação das abordagens anteriores (Lantzanakis; Mitrika; Chrysoulakis, 2017).

O método “Dark Object Subtraction” (DOS) é uma abordagem que utiliza as respostas espectrais de alvos específicos presentes na imagem (Chavez, 1988). Este método parte do pressuposto de que a imagem contenha alvos escuros, nos quais a reflectância detectada pelo

sensor é assumida como sendo igual a zero em todos os comprimentos de onda, a fim de minimizar os efeitos do espalhamento atmosférico nas medidas de reflectância (Gaida *et al.*, 2020)

O método “Image correction for atmospheric effects” (iCOR) é considerado um método híbrido (Vito, 2017). É um algoritmo desenvolvido para a correção de imagens obtidas por sensores orbitais MultiSpectral Instrument (MSI), que está presente no satélite Sentinel-2. O iCOR permite a correção de efeitos de adjacência através da recuperação de informações contidas na própria imagem, utilizando o algoritmo “Similarity Environmental Correction” (SIMEC), que se baseia em tabelas complexas derivadas do código de transferência radiativa “Moderate resolution atmospheric Transmission” (MODTRAN) (Keukelaere *et al.*, 2018).

O método “Sentinel-2 correction” (Sem2Cor) foi desenvolvido com objetivos semelhantes aos do iCOR, porém o Sem2Cor é exclusivamente aplicado à correção de imagens do sensor MSI presente no satélite Sentinel-2 (Gascon *et al.*, 2017). Em resumo, este método inclui um algoritmo que realiza o processamento das imagens, convertendo a reflectância do topo da atmosfera para a reflectância de superfície e transformando o nível de processamento de 1C para 2ª (Main-Knorn *et al.*, 2017). O modelo leva em consideração estimativas de aerossóis, gases atmosféricos e concentração de ozônio, que variam de acordo com a localização geográfica e as condições climáticas no momento da captura do sensor. Além disso, o Sem2Cor reduz os efeitos causados por nuvens, topografia e adjacência espectral (Main-Knorn *et al.*, 2017).

## 2.4 Índices de vegetação

Os índices de vegetação (IVs) são parâmetros quantitativos derivados de dados de sensoriamento remoto, combinando diferentes bandas espectrais para fornecer informações valiosas sobre as características da vegetação (Xue; Su, 2017). Esses índices são amplamente utilizados para avaliar a cobertura vegetal, o vigor e a dinâmica de crescimento em várias plataformas, incluindo satélites e veículos aéreos não tripulados (Xue; Su, 2017). A reflectância da luz pelas plantas é influenciada por fatores como tipo de planta, teor de água nos tecidos e outras características intrínsecas (Liu *et al.*, 2016). As características de emissão da vegetação (reflectância espectral) é determinada por características químicas e morfológicas da superfície de órgãos ou folhas (Zhang; Kovacs, 2012). Essas métricas são essenciais para avaliar dinâmicas vegetativas, entender processos biológicos e otimizar práticas agrícolas, sendo uma ferramenta crucial para pesquisadores e profissionais envolvidos em sensoriamento remoto e estudos ambientais (Xue; Su, 2017).

Os comprimentos de onda do infravermelho próximo e médio tem sido bastante estudado, e dizem que poder ser atribuídos a quantificação do crescimento e vigor das plantas, relacionado ao teor de água, pigmentos, teor de açúcar e carboidratos, teor de proteínas e aromáticos e outras diversas características (Batten, 1998).

O Índice de Vegetação por Diferença Normalizada (NDVI) representa uma métrica essencial na avaliação da cobertura vegetal, desempenhando um papel significativo na esfera agrícola. Este indicador, calculado a partir de dados obtidos por sensores remotos, fornece uma medida quantitativa da atividade fotossintética das plantas, sendo particularmente valioso na detecção de alterações na saúde das culturas. A capacidade do NDVI em diferenciar entre áreas com vegetação saudável e aquelas sob estresse ambiental, como seca ou deficiência nutricional, oferece aos agricultores insights cruciais para o manejo preciso de suas plantações. Além disso, a aplicação do NDVI não se limita apenas à identificação de problemas; sua utilização abrange a otimização de práticas agrícolas, permitindo uma abordagem mais sustentável e eficiente no uso de recursos (Rouse Jr *et al.*, 1974).

O Índice de Vegetação por Diferença Normalizada Verde (GNDVI) representa uma extensão do NDVI, focando especificamente na região espectral verde do espectro eletromagnético. Este índice é calculado a partir da diferença entre as bandas de reflectância no verde e no infravermelho próximo, proporcionando uma avaliação mais sensível à densidade e estrutura da vegetação. Na agricultura, o GNDVI oferece uma perspectiva aprimorada sobre a saúde das plantas, destacando nuances que podem não ser tão evidentes em índices mais amplos. Sua aplicação abrange a detecção de estresse hídrico, deficiências nutricionais e outros fatores que afetam diretamente a qualidade e o rendimento das culturas (Louhaichi; Borman; Johnson, 2001).

O Índice de Vegetação por Diferença Normalizada de Vermelho-Edge (NDRE) representa uma evolução na avaliação da saúde das plantas, incorporando uma banda de espectro conhecida como “red-edge”. Ao calcular a diferença normalizada entre as reflectâncias nesta banda e no infravermelho próximo, o NDRE oferece uma sensibilidade específica a alterações na clorofila e na estrutura celular das plantas. Essa característica o torna particularmente valioso na detecção precoce de estresses fisiológicos e na avaliação precisa da vitalidade das culturas. Na agricultura de precisão, o NDRE tem sido empregado para mapear variações intra-campo, identificar áreas suscetíveis a doenças ou deficiências nutricionais, e otimizar a aplicação de insumos agrícolas (Tucker, 1979).

O Índice de Vegetação Melhorado (EVI), desenvolvido como uma extensão do NDVI (Índice de Vegetação por Diferença Normalizada), visa superar algumas limitações percebidas

deste último. Ao incorporar correções para influências atmosféricas e características do solo, o EVI oferece uma avaliação mais precisa da vegetação. Sua formulação inclui a banda azul do espectro eletromagnético, permitindo uma maior capacidade de discriminação entre a cobertura vegetal e o solo desnudo. Na agricultura, o EVI tem demonstrado ser valioso para monitorar o vigor das plantas, especialmente em áreas propensas a distorções atmosféricas. Sua aplicação abrange desde a detecção de estresses ambientais até a quantificação precisa da resposta das culturas a práticas agrícolas específicas (Huete *et al.*, 2002).

## 2.5 Inteligência artificial

A grande quantidade de dados produzidos pelas tecnologias digitais, normalmente referida como "Big Data," requer grandes capacidades de armazenamento, além de edição, análise e interpretação. Isso tem um potencial considerável para agregar valor à sociedade, ao meio ambiente e aos tomadores de decisão (Sonka, 2016). As técnicas convencionais de processamento de dados são incapazes de atender às demandas em constante crescimento na nova era da agricultura inteligente, o que representa um obstáculo importante para extrair informações valiosas dos dados de campo. Nesse sentido, a Aprendizagem de Máquina (ML) surgiu, sendo uma subcategoria da inteligência artificial, aproveitando o crescimento exponencial da capacidade computacional (Helm *et al.*, 2020).

O termo IA (Inteligência Artificial) foi cunhado pela primeira vez na conferência de Dartmouth em 1956 por John McCarthy, que a definiu como a ciência e engenharia de criar máquinas inteligentes ou, mais especificamente, programas de computador inteligentes. A tecnologia de IA fornece inteligência computacional às máquinas para que elas possam aprender, entender e reagir de acordo com a situação. A IA abrange subcampos como ML (Aprendizado de Máquina), DL (Aprendizado Profundo), Processamento de Linguagem Natural (NLP), Inteligência de Enxame (SI), sistemas especialistas, lógica difusa e visão computacional (Sharma *et al.*, 2020).

A aprendizagem de máquina (ML) é o subcampo da ciência da computação que confere aos computadores a capacidade de aprender sem serem programados explicitamente. Samuel, em 1959, definiu ML como "a capacidade de aprender sem ser programado explicitamente." Turing, em 2009, propôs o conceito de máquinas de aprendizado e escreveu um artigo de pesquisa intitulado "O Teste de Turing para Inteligência Máquina." Ele realizou um teste e examinou a capacidade da máquina de demonstrar comportamento inteligente semelhante ao dos seres humanos. Uma máquina ou programa de computador inteligente aprende e extrai conhecimento dos dados, constrói um framework para fazer previsões ou decisões inteligentes.

Portanto, o processo de ML é dividido em três partes-chave: entrada de dados, construção de modelo e generalização (Turing, 2009).

Os algoritmos de ML são categorizados como aprendizado supervisionado, aprendizado não supervisionado e aprendizado por reforço. Aprendizado supervisionado, como o nome sugere, é o aprendizado com um supervisor ou professor. Este conjunto de algoritmos trabalha com um conjunto de dados rotulado, o que significa que, para cada entrada, existem saídas correspondentes. O algoritmo constrói uma relação entre entrada e saída com base nesses dados rotulados e, posteriormente, generaliza ou prevê saídas para entradas não vistas. Algoritmos de aprendizado supervisionado usados para prever valores categóricos são conhecidos como algoritmos de classificação, e os algoritmos usados para prever valores numéricos são conhecidos como algoritmos de regressão (Sharma *et al.*, 2020).

A agricultura de precisão habilitada por IA (inteligência artificial) eliminar a subjetividade e ajudar os agricultores da nova geração a otimizar cada etapa do processo agrícola fornecendo soluções em tempo real (Borba *et al.*, 2022). Isso inclui o uso de tecnologia para monitorar e gerenciar o solo, o clima, o uso de água e pesticidas, além de rastrear o crescimento das culturas. Isso permite que os agricultores tomem decisões mais informadas e eficazes, reduzindo os riscos associados à agricultura (Sharma *et al.*, 2020).

A previsão de rendimento de colheitas é uma informação significativa para qualquer agricultor, pois permite antecipar a quantidade de produção e como essa produção pode ser aumentada. Vários parâmetros desempenham um papel importante na previsão do rendimento das colheitas, tais como o valor de pH, tipo e qualidade do solo, padrões climáticos (temperatura, precipitação, umidade, horas de sol), uso de fertilizantes e calendários de colheita (Bhanumathi; Vineeth; Rohit, 2019).

A agricultura manual pode ser vista cientificamente como um sistema de controle de feedback, em que ações corretivas são tomadas assim que um problema na colheita é observado. O rendimento das colheitas depende da eficiência da utilização ótima dos recursos mencionados acima. Se alguma anomalia não for detectada nas fases iniciais, isso pode prejudicar significativamente o rendimento das colheitas (Sharma *et al.*, 2020).

Para ilustrar a importância da previsão precisa, Singh *et al.* (2017) avaliaram o impacto de tempestades de granizo na produção de trigo na Índia e observaram que, somente em fevereiro e março de 2015, eventos de granizo causaram uma queda de 8,4% na produção nacional de trigo. Especialmente para agricultores financeiramente vulneráveis em regiões onde o armazenamento intermitente de safras colhidas é escasso, previsões meteorológicas precisas podem ser fundamentais.

Os modelos de aprendizado de máquina (ML) quando aplicados sistematicamente atuam como um controle de avanço. Eles permitem antecipar os fatores que afetarão o rendimento das colheitas, permitindo assim que ações corretivas sejam tomadas antes que ocorra qualquer anomalia na produção. A previsão de rendimento de colheitas desempenha um papel crítico na agricultura moderna, e os avanços na aprendizagem de máquina e outras técnicas de modelagem estão permitindo que os agricultores otimizem seus métodos de cultivo e tomem medidas corretivas antes que qualquer anomalia possa afetar adversamente a produção (Sharma *et al.*, 2020).

O algoritmo Random Forest, uma técnica de aprendizado de máquina baseada em ensemble, tem se destacado como uma ferramenta poderosa na análise de dados agrícolas e na previsão de produtividade. Essa abordagem utiliza uma coleção de árvores de decisão, cada uma treinada em subconjuntos aleatórios dos dados de entrada, e posteriormente combina suas previsões para obter uma estimativa mais robusta e precisa. No contexto agrícola, o Random Forest pode ser aplicado para prever a produtividade das culturas, levando em consideração uma variedade de variáveis, como dados meteorológicos, índices de vegetação, histórico de cultivo e práticas agronômicas. A capacidade do algoritmo de lidar com grandes conjuntos de dados e de capturar relações não lineares entre variáveis o torna particularmente adequado para modelar a complexidade inerente aos sistemas agrícolas. Ao explorar a literatura sobre a aplicação do Random Forest na agricultura, busca-se compreender as nuances dessa abordagem, identificando suas vantagens e desafios específicos, bem como delineando estratégias para otimizar sua implementação em cenários agrícolas diversos (Gyamerah; Ngare; Ikpe, 2020).

O Support Vector Regression (SVR) é um algoritmo de aprendizado de máquina que se destaca na modelagem de relações não lineares entre variáveis, sendo particularmente útil na previsão de produtividade agrícola. Baseado no conceito de Support Vector Machines, o SVR busca encontrar um hiperplano de melhor ajuste em um espaço multidimensional, minimizando a diferença entre as previsões do modelo e os dados reais. Na agricultura, o SVR pode ser aplicado para prever a produção de culturas, incorporando informações como condições climáticas, práticas agronômicas, e dados de sensoriamento remoto. A capacidade do SVR em lidar com dados complexos e não lineares o torna uma escolha promissora para capturar relações intrincadas presentes nos sistemas agrícolas. Ao revisar a literatura sobre a aplicação do SVR na agricultura, busca-se compreender suas características distintivas, avaliar sua eficácia em comparação com outros algoritmos e explorar estratégias para otimizar sua implementação

em contextos agrícolas diversos, visando uma previsão precisa e adaptável da produtividade das culturas (Deiss *et al.*, 2020).

O algoritmo k-Nearest Neighbors (K-NN) é uma abordagem de aprendizado de máquina que se destaca pela sua simplicidade e eficácia na previsão de padrões em dados agrícolas. Este método classifica ou faz previsões com base na proximidade de um ponto a outros pontos no espaço de características. No contexto agrícola, o K-NN pode ser aplicado para prever a produtividade das culturas, considerando variáveis como histórico de cultivo, práticas agronômicas, dados meteorológicos e índices de vegetação. A flexibilidade do K-NN em lidar com diferentes tipos de dados e a capacidade de se adaptar a padrões sazonais e regionais fazem dele uma escolha atrativa para a modelagem de sistemas agrícolas. Ao revisar a literatura sobre a aplicação do K-NN na agricultura, o objetivo é compreender a sensibilidade do algoritmo a diferentes parâmetros, explorar estratégias para otimizar sua eficácia em cenários agrícolas diversos e avaliar como ele se posiciona em relação a outros métodos de aprendizado de máquina na previsão de produtividade de culturas (Khanal *et al.*, 2018).

A aplicação de algoritmos de aprendizado de máquina na agricultura tem se revelado instrumental, proporcionando insights valiosos para otimizar o manejo de cultivos e prever a produtividade. O Random Forest, com sua capacidade de lidar com grandes conjuntos de dados e capturar relações não lineares, destaca-se pela robustez e flexibilidade. Sua aplicação na agricultura permite a modelagem eficaz de sistemas complexos, abrangendo variáveis diversas. Por outro lado, o Support Vector Regression (SVR) se destaca ao lidar com relações não lineares, adaptando-se a complexidades inerentes aos sistemas agrícolas. Sua eficácia é particularmente notável em contextos em que padrões intrincados demandam modelagem precisa. Já o k-Nearest Neighbors (K-NN), com sua simplicidade e capacidade de adaptação, destaca-se em cenários em que a proximidade espacial de pontos é crucial para previsões. Cada algoritmo apresenta vantagens distintas, mas é importante considerar suas desvantagens, como possíveis sobreajustes ou sensibilidade a parâmetros específicos (Sharma *et al.*, 2020).

## REFERÊNCIAS

- BABURAO, Patil Sagar *et al.* Review of Machine Learning Model Applications in Precision Agriculture. *In: International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022)*. Atlantis Press, p. 916-930, 2023.
- BATISTA, G. T.; DIAS, N. W. **Introdução ao sensoriamento remoto e processamento de imagens**. São José dos Campos: INPE, 2005.
- BATTEN, G. D. Plant analysis using near infrared reflectance spectroscopy: the potential and the limitations. **Australian Journal of Experimental Agriculture**, v. 38, n. 7, p. 697-706, 1998.
- BHANUMATHI, S.; VINEETH, M.; ROHIT, N. Crop yield prediction and efficient use of fertilizers. *In: International Conference on Communication and Signal Processing (ICCSP)*. IEEE, p. 0769-0773, 2019.
- BORBA, Marcelo da Costa *et al.* Gestão no meio agrícola com o apoio da Inteligência Artificial: uma análise da digitalização da agricultura. **Revista em Agronegocio e Meio Ambiente**, v. 15, n. 3, p. 1-22, 2022.
- BRAMBILLA, Massimo *et al.* From conventional to precision fertilization: a case study on the transition for a small-medium farm. **AgriEngineering**, v. 3, n. 2, p. 438-446, 2021.
- BRITO, Valéria França. Acompanhamento da produção de milho, segunda safra, em 2019 na fazenda Santo André, município de Caseara - TO. **Monografia (Bacharel em Agronomia) – Universidade Federal do Tocantins, Campus Universitário do Gurupi**, 33 p., 2020.
- CAMPOS, Ghabriel Miranda *et al.* Mudanças climáticas e seus efeitos sobre a água no solo no cultivo de soja e milho em sucessão no cerrado. **Observatório de La Economía Latinoamericana**, v. 22, n. 5, p. e4833-e4833, 2024.
- CHANDRASEKHAR, Subrahmanyam. Radiative transfer. **Courier Corporation**, 1960.
- CHAVEZ, Pat S Jr. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. **Remote sensing of environment**, v. 24, n. 3, p. 459-479, 1988.
- CRUZ, Uilmer Rodrigues Xavier da *et al.* Conceitos, comportamento espectral dos alvos e fotointerpretação de imagens no QGIS. **Revista Científica Multidisciplinar Núcleo do Conhecimento**, 2021.
- CUSTODIO, Cleberti José Silva *et al.* Fatores que contribuíram para o crescimento da produtividade do milho no Brasil. **Revista Eletrônica Interdisciplinar**, v. 1, n. 15, 2016.
- DEISS, Leonardo *et al.* Tuning support vector machines regression models improves prediction accuracy of soil properties in MIR spectroscopy. **Geoderma**, v. 365, p. 114227, 2020.
- DIAZ-GONZALEZ, Freddy A. *et al.* Machine learning and remote sensing techniques applied to estimate soil indicators—review. **Ecological Indicators**, v. 135, p. 108517, 2022.

DURÃES, F. O. M. et al. Índices de vigor de sementes de milho (*Zea mays* L.): associação com emergência em campo, crescimento e rendimento de grãos. **Embrapa**, 1995.

GAIDA, William *et al.* Correção atmosférica em sensoriamento remoto: uma revisão. **Revista Brasileira de Geografia Física**, v. 13, n. 01, p. 229-248, 2020.

GALVÃO, João Carlos Cardoso *et al.* Sete décadas de evolução do sistema produtivo da cultura do milho. **Revista Ceres**, v. 61, p. 819-828, 2014.

GASCON, Ferran *et al.* Copernicus Sentinel-2A calibration and products validation status. **Remote Sensing**, v. 9, n. 6, p. 584, 2017.

GÓMEZ, Carlos Alejandro Ramírez. Aplicación del machine learning en agricultura de precisión. **Revista Cintex**, v. 25, n. 2, p. 14-27, 2020.

GUEDES, Jânio; SILVA, Sebastião Milton Pinheiro da. Sensoriamento remoto no estudo da vegetação: princípios físicos, sensores e métodos. **Acta Geográfica**, v. 12, n. 29, p. 127-144, 2018.

GUIMARÃES, Edson da Silva. Aprendizado de Máquina aplicado à predição da produtividade da cultura da soja utilizando dados de clima e solo. **Doutorado (Tese) Universidade de São Paulo**, 78 p., 2019.

GYAMERAH, Samuel Asante; NGARE, Philip; IKPE, Dennis. Probabilistic forecasting of crop yields via quantile random forest and Epanechnikov Kernel function. **Agricultural and Forest Meteorology**, v. 280, p. 107808, 2020.

HADJIMITSIS, Diofandos G. *et al.* Atmospheric correction for satellite remotely sensed data intended for agricultural applications: impact on vegetation indices. **Natural Hazards and Earth System Sciences**, v. 10, n. 1, p. 89-95, 2010.

HELM, J. Matthew *et al.* Machine learning and artificial intelligence: definitions, applications, and future directions. **Current reviews in musculoskeletal medicine**, v. 13, p. 69-76, 2020.

HILKER, T. **Surface reflectance/bidirectional reflectance distribution function**. 2018.

HUETE, A. R. *et al.* A comparison of vegetation indices over a global set of TM images for EOS-MODIS. **Remote sensing of environment**, v. 59, n. 3, p. 440-451, 1997.

HUETE, Alfredo *et al.* Overview of the radiometric and biophysical performance of the MODIS vegetation indices. **Remote sensing of environment**, v. 83, n. 1-2, p. 195-213, 2002.

JUNIOR, Edmilson Gomes Cavalcante *et al.* Necessidade hídrica da cultura do milho influenciada pelas mudanças climáticas no semiárido nordestino. **Revista Brasileira de Milho e Sorgo**, v. 17, n. 2, p. 251-262, 2018.

KATSIKIANNIS, P. *et al.* Fusion of spatio-temporal UAV and proximal sensing data for an agricultural decision support system. *In: Fourth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2016)*. **SPIE**, p. 564-574, 2016.

KEUKELAERE, Liesbeth de *et al.* Atmospheric correction of Landsat-8/OLI and Sentinel-2/MSI data using iCOR algorithm: validation for coastal and inland waters. **European Journal of Remote Sensing**, v. 51, n. 1, p. 525-542, 2018.

KHANAL, Sami *et al.* Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. **Computers and electronics in agriculture**, v. 153, p. 213-225, 2018.

LANTZANAKIS, Giannis; MITRAKA, Zina; CHRYSOULAKIS, Nektarios. Comparison of physically and image based atmospheric correction methods for Sentinel-2 satellite imagery. *In: Perspectives on atmospheric sciences*. **Springer International Publishing**, p. 255-261. 2017.

LEONARDI, Silvia Shizue; SOUZA, S. M.; FONSECA, Leila Maria Garcia. Comparação do desempenho de algoritmos de normalização radiométrica de imagens de satélite. **Simpósio Brasileiro de Sensoriamento Remoto (SBSR)**, v. 10, p. 2063-2069, 2003.

LIMA, Ricardo Neves de Souza; RIBEIRO, Celso Bandeira de Melo. Comparação de métodos de correção topográfica em imagens LANDSAT sob diferentes condições de iluminação. **Revista Brasileira de Cartografia**, v. 5, n. 66, p. 1097-1116, 2014.

LIU, Chang *et al.* A review of plant spectral reflectance response to water physiological changes. **Chinese Journal of Plant Ecology**, v. 40, n. 1, p. 80, 2016.

LOPES, Leandro H. *et al.* Características agronômicas e rendimento de grãos da cultura do milho (*Zea mays* L.) sob diferentes densidades. *In: SPERS, Eduardo Eugênio (org.). Agrárias: pesquisa e inovação nas ciências que alimentam o mundo*. **Editores Artemis**, v. 7, p. 219-228, 2021.

LOUHAICHI, Mounir; BORMAN, Michael M.; JOHNSON, Douglas E. Spatially located platform and aerial photography for documentation of grazing impacts on wheat. **Geocarto International**, v. 16, n. 1, p. 65-70, 2001.

MAIA, Roberto Gomes *et al.* Abordagem sobre pragas e doenças do milho. **Contribuciones a Las Ciencias Sociales**, v. 17, n. 3, p. e5582-e5582, 2024.

MAIN-KNORN, Magdalena *et al.* Sen2Cor for sentinel-2. *In: Image and Signal Processing for Remote Sensing XXIII*. **SPIE**, p. 37-48. 2017.

MALDANER, Luciano Junior *et al.* Exigência agroclimática da cultura do milho (*Zea mays*). **Revista Brasileira de Energias Renováveis**, v. 3, n. 1, p. 13-23, 2014.

MEDEIROS, Rui Palmeira *et al.* Componentes do balanço de água e de radiação solar no desenvolvimento do milho em quatro épocas de semeadura no agreste de Alagoas. **Dissertação (Mestrado em Agronomia) – Universidade Federal de Alagoas**, 90 p., 2009.

MILLER, Jarrod O.; SHOBER, Amy L.; TARAULA, Jamie. Assessing relationships of cover crop biomass and nitrogen content to multispectral imagery. **Agronomy Journal**, v. 116, n. 3, p. 1417-1427, 2024.

MOGRANE, Mohamed Abdelillah *et al.* Evaluation of five atmospheric correction algorithms over French optically-complex waters for the Sentinel-3A OLCI Ocean Color Sensor. **Remote Sensing**, v. 11, n. 6, p. 668, 2019.

OKIN, Gregory S.; GU, Juan. The impact of atmospheric conditions and instrument noise on atmospheric correction and spectral mixture analysis of multispectral imagery. **Remote Sensing of Environment**, v. 164, p. 130-141, 2015.

PASSOS, Emerson Gabriel Cardoso dos *et al.* Uso de imagens aéreas obtidas por drone no monitoramento do desenvolvimento da cultura do milho. **Anais da Mostra Nacional de Iniciação Científica e Tecnológica Interdisciplinar (MICTI)-e-ISSN 2316-7165**, v. 1, n. 12, 2019.

PAUL, Kenny *et al.* Viable smart sensors and their application in data driven agriculture. **Computers and Electronics in Agriculture**, v. 198, p. 107096, 2022.

PAUSTIAN, Margit; THEUVSEN, Ludwig. Adoption of precision agriculture technologies by German crop farmers. **Precision agriculture**, v. 18, p. 701-716, 2017.

RODRIGUES, Victor do Nascimento *et al.* Comparação entre métodos para estimar a produtividade de grãos de milho. **Ciência e Agrotecnologia**, v. 29, p. 34-42, 2005.

ROUSE JR, John W. *et al.* **Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation**. 1974.

SAMUEL, Arthur L. Some studies in machine learning using the game of checkers. **IBM Journal of research and development**, v. 3, n. 3, p. 210-229, 1959.

SANS, Luiz Marcelo Aguiar *et al.* Zoneamento de riscos climáticos para a cultura de milho na região Centro-Oeste do Brasil e para o estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, v. 9, n. 2, 9 p., 2001.

SESSIONS, Valerie; VALTORTA, Marco. The Effects of Data Quality on Machine Learning Algorithms. **ICIQ**, v. 6, p. 485-498, 2006.

SHARMA, Abhinav *et al.* Machine learning applications for precision agriculture: A comprehensive review. **IEEE Access**, v. 9, p. 4843-4873, 2020.

SHIMABUKURO, Yosio Edemir; MAEDA, Eduardo Eiji; FORMAGGIO, Antônio Roberto. Sensoriamento Remoto e Sistemas de Informações Geográficas aplicados ao estudo dos recursos agronômicos e florestais. **Ceres**, v. 56, n. 4, 2009.

SHIOGA, Pedro Sentaro; GERAGE, Antônio Carlos. Influência da época de plantio no desempenho do milho safrinha no estado do Paraná, Brasil. **Revista Brasileira de Milho e Sorgo**, v. 9, n. 3, p. 236-253, 2010.

SHIRATSUCHI, Luciano Shozo *et al.* Sensoriamento remoto: conceitos básicos e aplicações na agricultura de precisão. In: BERNARDI, Alberto Carlos de Campos *et al.* (editores técnicos). Agricultura de precisão: resultados de um novo olhar. **Embrapa**, 1. ed., p. 58-17, 2014.

- SILVA, C. S. M.; HACHISUCA, A. M. M. Estudo e desenvolvimento de uma plataforma para a agricultura de precisão utilizando o conceito SMART FARM empregando veículos aéreos não tripulados (VANTs). **The Academic Society Journal**, v. 3, n. 1, p. 4-14, 2019.
- SILVA, Luiz Eduardo Bezerra et al. Desenvolvimento da cultura do milho (*Zea mays* L.): revisão de literatura. **Diversitas Journal**, v. 5, n. 3, p. 1636-1657, 2020.
- SILVA, Samuel Ferreira da *et al.* Utilização do sensoriamento remoto na agricultura de precisão: uma análise bibliométrica. **Nucleus**, v. 11, n. 2, p. 459-470, 2014.
- SILVA, Wilson Jesus da *et al.* Exigências climáticas do milho em sistema plantio direto. *In: Cultivo do milho no sistemas de plantio direto. Embrapa Milho e Sorgo, Informe Agropecuário*, v. 27, n. 233, p. 14-25, 2006.
- SINGH, S. K. *et al.* Assessment of hailstorm damage in wheat crop using remote sensing. **Current Science**, p. 2095-2100, 2017.
- SONKA, Steve. Big data: fueling the next evolution of agricultural innovation. **Journal of Innovation Management**, v. 4, n. 1, p. 114-136, 2016.
- SOUZA, Igor Viana *et al.* Veículos aéreos não tripulados e sensores remotos: ferramentas para o gerenciamento dos recursos florestais. **Journal of Biotechnology and Biodiversity**, v. 8, n. 2, p. 148-157, 2020.
- TEIXEIRA, Flavia França; COSTA, Flaviane Malaquias. Caracterização de recursos genéticos de milho. **Embrapa Milho e Sorgo, Comunicado Técnico**, n. 185, 10 p., 2010.
- THILAKARATHNE, Navod Neranjan *et al.* A cloud enabled crop recommendation platform for machine learning-driven precision farming. **Sensors**, v. 22, n. 16, p. 6299, 2022.
- TUCKER, Compton J. Red and photographic infrared linear combinations for monitoring vegetation. **Remote sensing of Environment**, v. 8, n. 2, p. 127-150, 1979.
- TURING, Alan M. Computing machinery and intelligence. **Springer Netherlands**, p. 23-65, 2009.
- VITO. VITO Remote Sensing Unit. Vlaams Instituut voor Technologisch Onderzoek. **iCOR plugin for SNAP toolbox: software user manual**, version 1, 2017.
- WELTER, André *et al.* Distorções causadas pela falta de calibração radiométrica e efeitos atmosféricos na análise de dados de sensoriamento remoto orbital. **Geografia**, v. 45, n. 2, p. 361-383, 2020.
- XUE, Jinru; SU, Baofeg. Significant remote sensing vegetation indices: A review of developments and applications. **Journal of sensors**, v. 2017, 17 p., 2017.
- ZHANG, Chunhua; KOVACS, John M. The application of small unmanned aerial systems for precision agriculture: a review. **Precision agriculture**, v. 13, p. 693-712, 2012.

## ARTIGO 1

### **Impactos da correção de imagens de satélite no desempenho dos índices de reflectância e vegetação sobre a cultura do milho**

#### RESUMO

A utilização de índices de vegetação é de grande importância para a agricultura de precisão, fornecendo dados para monitoramento dos cultivos. Esses índices, derivados de imagens de satélite e em conjunto com tecnologias de análise de dados, oferecem informações valiosas sobre as condições da vegetação, facilitando a gestão eficaz das culturas e a tomada de decisões. No entanto as informações geradas pelos índices de vegetação estão sujeitas a variações devido as condições atmosféricas e ambientais do local de estudo. Uma maneira de minimizar os efeitos atmosféricos e conferir mais confiança aos dados de sensoriamento remoto é a utilização de algoritmos de correção atmosférica. Os métodos de correção atmosférica são divididos em três categorias, empíricas, as quais utilizam dados de alvos conhecidos encontrados na imagem, físicas, que se baseiam na teoria da transferência radiativa, e híbridos, que são a união dos dois métodos anteriores. O objetivo deste trabalho é analisar como diferentes correções atmosféricas impactam os dados de reflectância e índices de vegetação obtidos a partir de imagens orbitais do satélite Sentinel-2 durante o período da safra e safrinha na cultura do milho. Foram utilizados 3 métodos de correção, Subtração de Objeto Escuro (DOS), Correção de Imagem para Efeitos Atmosféricos (iCOR) e Correção Sentinel-2 (Sen2Cor), e as análises foram realizadas para a reflectância das bandas do Azul (B02), Verde (B03), Vermelho (B04), RedEdge (B06) e NIR (B08), também para o Índice de Vegetação por Diferença Normalizada (NDVI), o Índice de Vegetação por Diferença Normalizada Verde (GNDVI), o Índice de Vegetação Melhorado (EVI) e o Índice de Vegetação da Borda Vermelha por Diferença Normalizada (NDRE). Foi realizada a análise da série temporal com imagens obtidas aos 30, 45, 60, 75 e 90 dias após a semeadura, no período da safra e safrinha em campos de produção comercial da região Sul do estado de Minas Gerais. A escolha apropriada do método de correção atmosférica e do índice de vegetação é crucial para melhorar a qualidade dos dados utilizados na agricultura de precisão, contribuindo para uma gestão de culturas mais eficiente e sustentável.

**Palavras-chave:** Agricultura de precisão; correção atmosférica; sensoriamento remoto.

## **Impacts of satellite images correction on reflectance and vegetation index performance for maize in the main and off-season crops**

### **ABSTRACT**

The use of vegetation indices is of great importance for precision agriculture, providing data for crop monitoring. These indices, derived from satellite imagery and in conjunction with data analytics technologies, offer valuable insights into vegetation conditions, facilitating effective crop management and decision-making. However, the information generated by the vegetation indices is subject to variations due to the atmospheric and environmental conditions of the study site. One way to minimize atmospheric effects and give more confidence to remote sensing data is the use of atmospheric correction algorithms. Atmospheric correction methods are divided into three categories: empirical, which use data from known targets found in the image, physical, which are based on the theory of radiative transfer, and hybrid, which are the union of the two previous methods. The objective of this work is to analyze how different atmospheric corrections impact the reflectance data and vegetation indices obtained from orbital images of the Sentinel-2 satellite during the harvest and off-season period in the corn crop. 3 correction methods were used, Dark Object Subtraction (DOS), Image Correction for Atmospheric Effects (iCOR) and Sentinel-2 Correction (Sen2Cor), and the analyses were performed for the reflectance of the Blue (B02), Green (B03), Red (B04), RedEdge (B06) and NIR (B08) bands, also for the Normalized Difference Vegetation Index (NDVI), the Green Normalized Difference Vegetation Index (GNDVI), the Improved Vegetation Index (EVI) and the Normalized Difference Red Edge Vegetation Index (NDRE). The analysis of the time series was carried out with images obtained at 30, 45, 60, 75 and 90 days after sowing, during the harvest and off-season in commercial production fields in the southern region of the state of Minas Gerais. The appropriate choice of atmospheric correction method and vegetation index is crucial to improve the quality of data used in precision agriculture, contributing to more efficient and sustainable crop management.

**Keywords:** Precision agriculture; atmospheric correction; remote sensing.

## 1 INTRODUÇÃO

Índices de vegetação como o Índice de Vegetação por Diferença Normalizada (NDVI), o Índice de Vegetação por Diferença Normalizada Verde (GNDVI), o Índice de Vegetação Melhorado (EVI) e o Índice de Vegetação da Borda Vermelha por Diferença Normalizada (NDRE) são ferramentas essenciais na agricultura de precisão, são utilizados para monitorar o desenvolvimento das culturas, a saúde das plantas e para otimizar o manejo (Paul *et al.*, 2022; Brambilla *et al.*, 2021; Paustian; Theuvsen, 2017). Esses índices, derivados de imagens de satélite e em conjunto com tecnologias de análise de dados, oferecem informações valiosas sobre as condições da vegetação, facilitando a gestão eficaz das culturas e a tomada de decisões (Talaviya *et al.*, 2020). No entanto, a precisão das informações geradas pode ser significativamente afetada por condições atmosféricas, que introduzem ruídos e distorcem os dados de reflectância capturados por sensores orbitais (Hadjimitsis *et al.*, 2010), fazendo necessária a utilização de métodos de correção.

Os métodos de correção atmosférica podem ser divididos em três categorias principais, classificadas de acordo com a abordagem utilizada por Gaida *et al.* 2020: métodos físicos, métodos empíricos e métodos híbridos. Os métodos físicos são baseados na Teoria de Transferência Radiativa de Chandrasekhar (1960), conhecida como correção absoluta, que utiliza estimativas físicas de parâmetros atmosféricos. Os métodos empíricos são fundamentados em estimativas dos efeitos atmosféricos com base em informações de alvos específicos presentes na própria imagem, sendo geralmente mais simples, mas potencialmente menos precisos. Os métodos híbridos representam uma combinação das abordagens anteriores, combinando informações físicas e estatísticas da cena.

Dentre os métodos de correção mais comuns pode-se destacar, o Dark Object Subtraction (DOS), Sentinel-2 Correction (Sen2Cor) e Image Correction for Atmospheric Effects (iCOR), que são amplamente utilizados. O DOS é um método empírico, enquanto os métodos Sen2Cor e iCOR são híbridos. O Sen2Cor foi desenvolvido para correção das imagens das constelações Sentinel, integrado à Plataforma de Aplicações Sentinel (SNAP), além da correção fornece dados que permitem a caracterização do local de estudo, incluindo mapas de espessura óptica do aerossol e vapor de água (Lantzanakis; Mitraka; Chrysoulakis, 2017). O iCOR, projetado para imagens Sentinel e Landsat, é amplamente utilizado para ajustar os efeitos atmosféricos sob corpos d'água (Keukelaere *et al.*, 2018), com suas aplicações em agricultura pouco exploradas.

Além da interferência atmosférica, um desafio persistente associado aos índices de vegetação é o fenômeno da saturação, onde os índices se tornam menos sensíveis a variações em vegetação de alta densidade. Por exemplo, o NDVI, índice amplamente utilizado em aplicações agrícolas, frequentemente satura em áreas com vegetação densa, levando a um platô nos valores e limitando sua eficácia em distinguir entre diferentes níveis de densidade da vegetação (Huete; Liu; Van Leeuwen, 1997). De forma semelhante, o GNDVI e o NDRE, apesar de serem menos sensíveis à saturação, também enfrentam limitações em situações de alta biomassa (Miller; Shoher; Taraila, 2024). Desta maneira, compreender o impacto dessas correções nas propriedades de saturação dos índices de vegetação é crucial, pois podem influenciar significativamente a análise temporal dos dados de reflectância e, conseqüentemente, a precisão do monitoramento de culturas.

Estudos recentes destacaram a importância da correção atmosférica na redução de problemas de saturação e na melhoria da sensibilidade dos índices de vegetação em áreas de alta biomassa. Muchsin *et al.* (2023) demonstraram que vários algoritmos de correção atmosférica melhoraram a precisão da classificação em campos de arroz. De forma semelhante, Moravec *et al.* (2021) testaram seis métodos de correção atmosférica em diferentes tipos de cobertura do solo, revelando variações substanciais nos valores de NDVI e melhorando a intercomparabilidade entre imagens Landsat 8 e Sentinel-2. Porém, o impacto das correções no monitoramento contínuo e em diferentes períodos de cultivo ainda é pouco estudado.

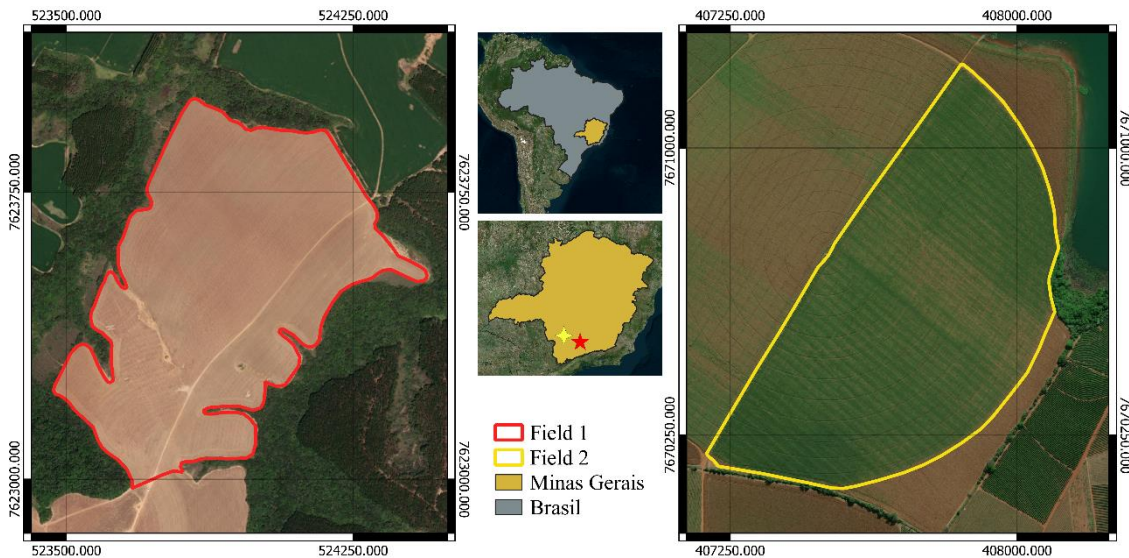
A maioria da literatura atual sobre a aplicação de índices de vegetação na agricultura de precisão desconsidera ou utiliza predominantemente correções atmosféricas simples. No entanto, o impacto dessas correções em diferentes períodos de cultivo de milho permanece inexplorado. Este estudo tem como objetivo aplicar e avaliar o impacto de diferentes métodos de correção atmosférica em imagens Sentinel-2 adquiridas durante todo o período de desenvolvimento da cultura do milho, tanto na safra quanto na safrinha. Além disso, busca identificar e quantificar possíveis limitações e discrepâncias nos índices de vegetação calculados usando cada tipo de correção.

## **2 MATERIAL E MÉTODOS**

Foram analisadas duas lavouras comerciais de milho na região sul de Minas Gerais. A primeira lavoura está no município de Luminárias (latitude 21° 30' 39" S, longitude 44° 54' 10" O, altitude 1.141 m) (Figura 2 – A), os dados são da primeira safra de 2021/2022, semeadura em outubro de 2021 e colheita em março de 2022, foi utilizado o híbrido P2501 da empresa Pionner, a área tem 49,56 hectares e não tem sistema de irrigação. A segunda está localizada

no município de Campo do Meio (latitude 21° 06' 25" S, longitude 45° 49' 48" O, altitude 774 m) (Figura 2 – B), os dados são referentes a safrinha 2020, semeadura em fevereiro e colheita em junho, foi utilizado o híbrido AS 1633 Pro3 da empresa Bayer, a área tem 50,48 hectares e conta com sistema de irrigação por pivô central.

**Figura 1** – Mapa de localização das áreas de cultivo de milho safra (A) e safrinha (B).



Fonte: Do autor (2024).

As imagens foram obtidas dos satélites da missão Sentinel-2 do programa espacial da Agência Espacial Europeia (ESA) na plataforma Copernicus Data Space Ecosystem (<https://dataspace.copernicus.eu/>, acessada em 23 de fevereiro de 24), com nível de processamento L1C, referentes à reflectância no Topo da Atmosfera (TOA). As imagens foram adquiridas da constelação Sentinel-2 S2A, com o sensor MSI, nível de processamento L1C, processadas pela versão N0500, na órbita relativa R138 e na região T23KMS. Segue abaixo a Tabela 1 com a descrição das imagens com data e porcentagem de cobertura de nuvens.

**Tabela 1** – Descrição da growing season, data, days after sowing (DAS) e cloud coverage das imagens utilizadas neste trabalho.

<b>Growing Season</b>	<b>Data</b>	<b>DAS</b>	<b>Cloud Coverage</b>
Season	11/17/2021	10	15.08
Season	12/27/2021	50	34.83
Season	01/21/2022	75	29.94
Season	02/25/2022	110	11.57
Season	03/07/2022	120	0.45
Off Season	03/27/2020	30	9.24
Off Season	04/26/2020	60	0.12
Off Season	05/06/2020	70	32.53
Off Season	05/26/2020	90	0.02
Off Season	06/25/2020	120	81.83

Fonte: Do autor (2024).

O processamento das imagens foi realizado nos softwares QGIS (2024) e SNAP (2024). A correção DOS foi feita usando o plugin Semi-Automatic Classification Plugin (Congedo, 2021) no QGIS utilizando a aba de “Pré-processamento” >> “Conversão de Imagens”, com a opção “Aplicar correção DOS 1” ativada. A correção iCOR foi realizada com o plugin da empresa VITO no software SNAP da ESA, na aba utilizando a banda B01 para detecção de nuvens, a banda B08 para detecção de corpos de água e o tipo de aerossol "Rural". A correção Sen2Cor foi realizada no SNAP da ESA, utilizando o plugin Sen2Cor 2.11.0, com parâmetros de “Classificação” e “Correção de nuvens cirrus” ativados, e tipo de aerossol "Rural".

Após o processamento das imagens, foi realizado o procedimento de remoção da bordadura de 30 m, para minimizar a interferência de fatores externos ao campo de cultivo. Foi realizada a reamostragem da banda do RedEdge (B06) de 20 x 20 m para 10 x 10 m por pixel utilizando a ferramenta “Reprojetar coordenadas” do menu “Raster” do software QGIS. As áreas de estudo foram divididas em grade regular de 100 x 100 m, devido ao formato irregular da área algumas unidades amostrais foram cortadas, foram selecionadas as amostras com pelo menos 70% da área da unidade amostral completa.

Os índices de vegetação Índice de Vegetação da Diferença Normalizada (NDVI), Índice de Vegetação da Borda Vermelha Normalizada (NDRE), Índice de Vegetação da Diferença Normalizada Verde (GNDVI) e Índice de Vegetação Melhorado (EVI) foram calculados para as três diferentes correções atmosféricas e para as imagens com nível de processamento L1C (TOA) utilizando o software R v.4.3.2, utilizando os pacotes "raster" (Hijmans, 2023), "sp" (Pebesma; Bivand, 2005; Bivand; Pebesma; Gomez-Rubio, 2013) e "sf" (Pebesma; Bivand, 2018).

**Tabela 2** - Índices de vegetação.

Nome	Equação	Autor
NDVI	$\frac{NIR - RED}{NIR + RED}$	Rouse <i>et al.</i> (1974)
NDRE	$\frac{NIR - REDEEDGE}{NIR + REDEEDGE}$	Tucker (1979)
GNDVI	$\frac{NIR - GREEN}{NIR + GREEN}$	Louhaichi; Borman; Johnson, 2001
EVI	$G * \frac{(NIR - RED)}{(NIR + C1 * RED - C2 * BLUE + L)}$	Huete <i>et al.</i> (2002)

Legenda: G = 2,5; C1 = 6; C2 = 7,5; L = 1.

Fonte: Do autor (2024).

Após estes procedimentos foram extraídos e tabulados os valores de reflectância das seguintes bandas do Sentinel-2: Azul (Banda 2, 490 nm), Verde (Banda 3, 560 nm), Vermelho (Banda 4, 665 nm), Red Edge (Banda 5, 705 nm) e NIR (Banda 8, 842 nm) e dos IV's NDVI, NDRE, GNDVI e EVI. A extração e tabulação dos valores de reflectância e dos IV's foi realizada no software R v. 4.3.2, utilizando os pacotes "raster" (Hijmans, 2023), "sf" (Pebesma; Bivand, 2018), "exactextractr" (Boston, 2023) e "dplyr" (Wickham *et al.*, 2023).

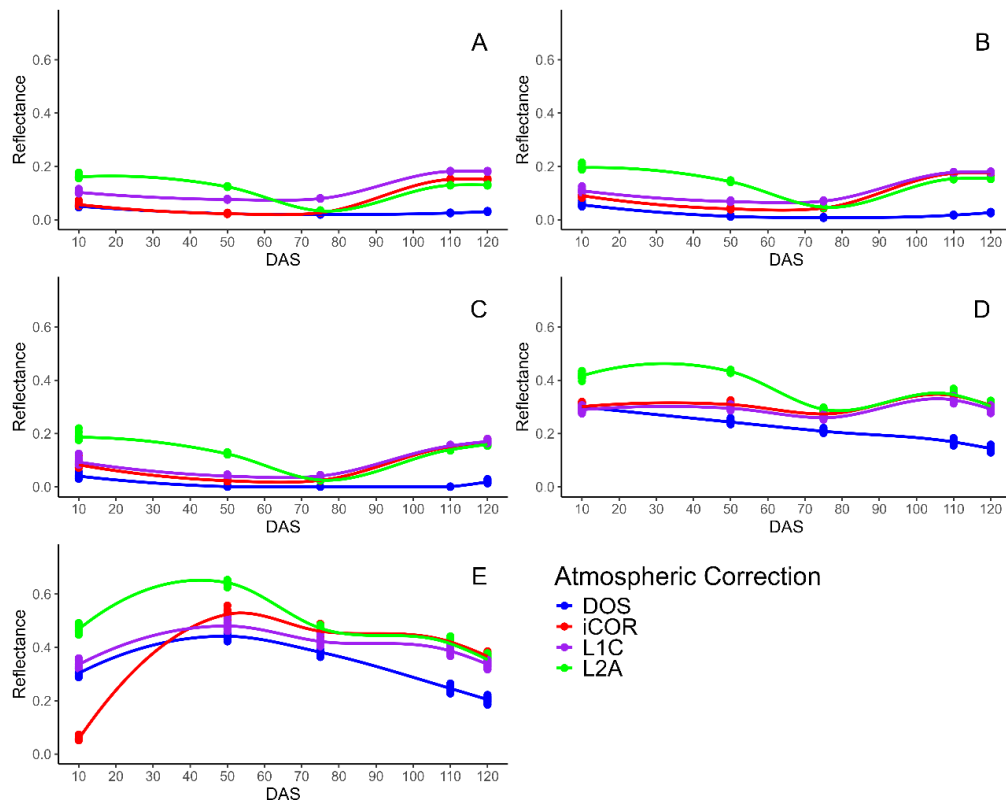
A análise dos dados seguiu uma abordagem estruturada para garantir a precisão e relevância dos resultados. Inicialmente, foi realizada a limpeza dos dados, removendo valores ausentes e outliers, de modo a assegurar a qualidade das informações a serem analisadas. Em seguida, aplicamos o método de Regressão Local (Loess) (Cleveland; Loader, 1996), uma técnica não paramétrica que suaviza curvas e superfícies para estimar padrões temporais. Esses ajustes foram feitos tanto para as bandas espectrais quanto para os índices de vegetação calculados previamente (NDVI, NDRE, GNDVI e EVI). Os resultados permitiram comparar a variação entre diferentes correções atmosféricas e identificar quais bandas e índices de vegetação apresentaram maior estabilidade para o monitoramento da cultura do milho ao longo do tempo.

### 3 RESULTADOS

A análise temporal das bandas provenientes do satélite Sentinel-2 (S-2), na região espectral do visível, para os períodos de safra e safrinha apresentaram comportamentos distintos em relação a reflectância durante o desenvolvimento, porém a característica das correções atmosféricas foi semelhante.

Durante a primeira safra (Figura 2 A - C) e a segunda safra da cultura do milho (Figura 3 A-C) observou-se estabilidade na reflectância nas bandas do espectro visível (azul, verde e vermelho) independente do estágio de desenvolvimento da planta de milho para a correção DOS. Já para a correção atmosférica iCOR durante o período da safra e safrinha, o comportamento foi de aumento após os 70 DAS, atingindo seu ponto máximo próximo ao período de colheita. Em ambas as épocas de cultivo (safra e safrinha), a correção Sen2Cor apresentou valores de reflectância mais altos no início e no final do ciclo de desenvolvimento da cultura do milho. Especialmente para o período da safra, os valores da reflectância para a correção Sen2Cor atingiram seu ponto mínimo próximo ao 70 DAS, próximo ao pico máximo da vegetação das plantas de milho.

**Figura 2** – Valores de reflectância de imagens de satélite (Sentinel-2) após correção atmosférica utilizando algoritmos DOS, iCor, Sen2Cor e L1C para as bandas (A) Azul; (B) Verde; (C) Vermelho; (D) Red Edge; (E) NIR; no período de primeira safra da cultura do milho.



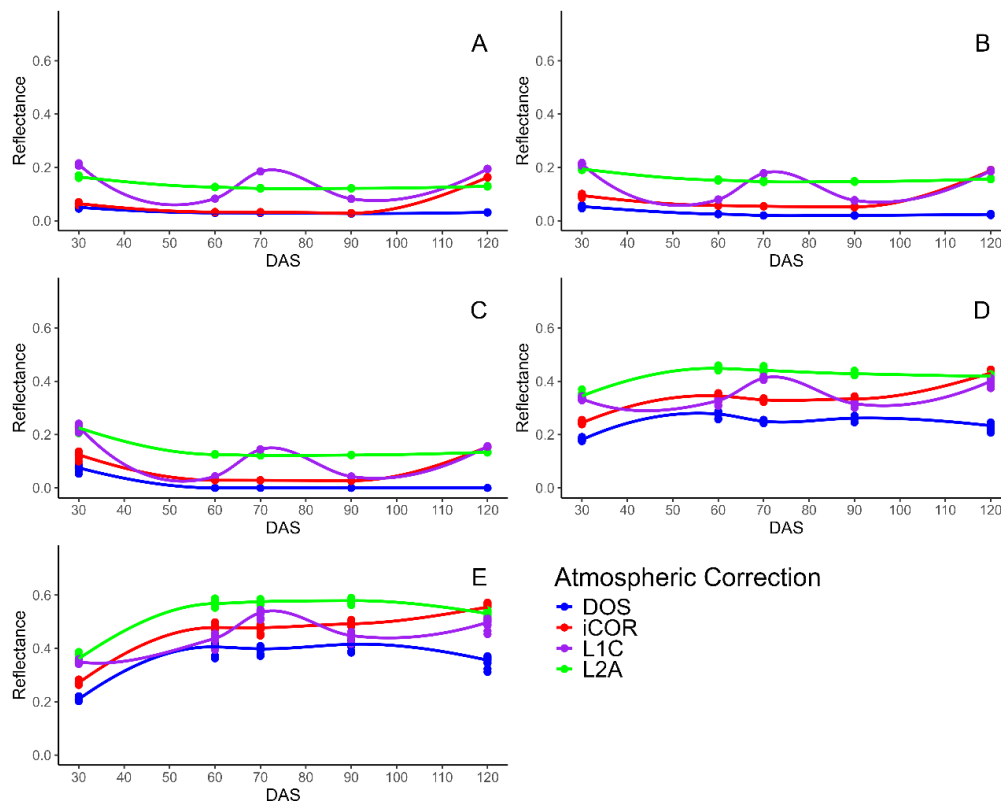
Fonte: Do autor (2024).

Diferente do espectro visível o comportamento das correções atmosféricas nas bandas do Red Edge e NIR foram diferentes entre os períodos de safra (Figura 2 – D e E) e safrinha (Figura 3 – D e E).

Para a banda do Red Edge no período de safra (Figura 3 – D) observou-se os valores máximos de reflectância para todas as correções aos 10 DAS. O Sen2Cor apresentou os maiores valores de reflectância em comparação as outras correções (DOS, iCOR e L1C) até os 75 DAS, após este período os valores foram similares. O método DOS apresentou valores similares ao iCOR em estágios iniciais, porém, com o desenvolvimento da cultura houve uma redução constante nos valores de reflectância para a banda Red Edge, atingindo o ponto mínimo próximo a colheita.

Durante a safrinha o comportamento da banda do Red Edge (Figura 4 – D) para a correção Sen2Cor foi de estabilidade após os 60 DAS. Já para a correção iCOR, tem-se padrão de aumento dos valores de reflectância dos 30 até 60 DAS, e chegando ao seu ponto máximo aos 120 DAS. E assim como observado durante o período de safra, os valores de reflectância para as correções iCOR e Sen2Cor apresentam valores similares em estágios de desenvolvimento mais avançados da cultura do milho e a correção DOS apresentou os valores mais baixos.

**Figura 3** – Valores de reflectância de imagens de satélite (Sentinel-2) após correção atmosférica utilizando algoritmos DOS, iCOR, Sen2Cor e L1C para as bandas (A) Azul; (B) Verde; (C) Vermelho; (D) Borda do Vermelho; (E) NIR; no período de safrinha da cultura do milho (safrinha).



Fonte: Do autor (2024).

O efeito das diferentes correções atmosféricas foi mais evidente na banda do Infravermelho Próximo (NIR) safra (Figura 2 – E), uma vez que o iCOR apresentou valores inferiores as demais correções aos 10 DAS e após este período demonstrou similaridade aos valores do algoritmo Sen2Cor. A correção DOS, após os 50 DAS apresentou os valores mais baixos de reflectância para a banda do NIR, assim como observado para as outras bandas.

A semelhança entre as correções atmosféricas também foi observada durante o período de safrinha para banda do NIR (Figura 3 – E), porém, esta proximidade entre os valores de RS observados para cada correção depende da data de obtenção das imagens utilizadas neste trabalho. É possível observar que aos 30 DAS, a reflectância das correções DOS e iCOR apresentaram pequena variação. Entretanto, com o crescimento das plantas de milho houve alteração na resposta, na qual os valores de reflectância da correção iCOR demonstrou mais proximidade aos do Sen2Cor, especialmente próximo a colheita.

De modo geral a correção DOS resultou em valores mais baixos de reflectância quando comparados as correções iCOR e Sen2Cor em todas as bandas espectrais analisadas. Este comportamento do DOS indica possível limitação ao desconsiderar os efeitos aditivos da atmosfera em sua metodologia de correção, o que dificulta as perdas causadas pela interferência atmosférica. Além disso, os métodos iCOR e Sen2Cor incluem técnicas que buscam minimizar a interferência de aerossóis que estão dispersos na BOA quando comparado com a DOS que é mais simples e não considera aspectos relacionados a filtragem, e são mais acurados para a classificação de imagens (Sola *et al.*, 2018)

Os índices de vegetação são ferramentas de grande importância no contexto da agricultura de precisão (AP), sendo empregados com diversas finalidades, desde o monitoramento das lavouras, estimativa de produtividade e recomendação de fertilizantes (Paul *et al.*, 2022; Brambilla *et al.*, 2021; Paustian; Theuvsen, 2017). Desta forma, é importante verificar a interferência da correção atmosférica no padrão dos índices de vegetação ao longo do ciclo de desenvolvimento da cultura do milho.

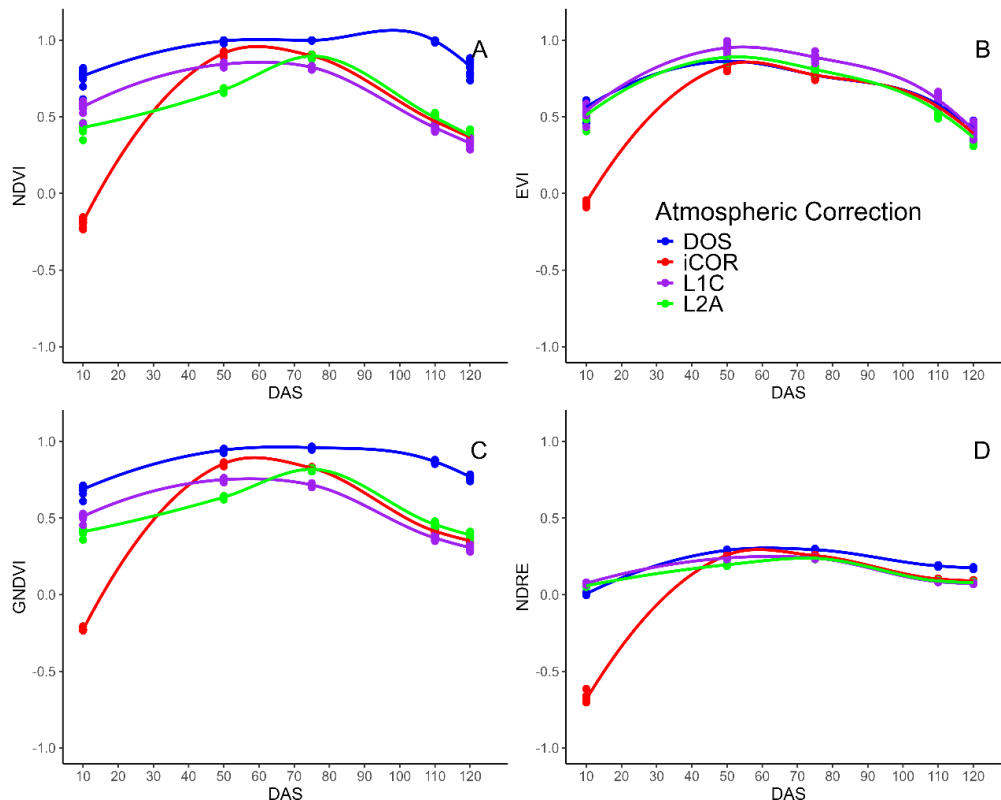
Assim, observa-se que a aplicação do algoritmo de correção atmosférica DOS para calcular o NDVI, obteve-se valores próximos a saturação (Figura 4 - A). Já as correções iCOR e Sen2Cor apresentam aumento do NDVI até os 50 e 60 DAS respectivamente, com redução nos valores com a maior proximidade do período de colheita. Para o índice GNDVI na safra (Figura 4 – C), o comportamento das correções atmosféricas foi similar aos observados para o NDVI, porém com valores mais baixos, demonstrando menor tendência a saturação.

Em contrapartida a análise do comportamento dos índices GNDVI e NDVI (Figura 5 – A - C) apresentaram valores similares para as correções iCOR e Sen2Cor aos 30 e 120 DAS.

Porém entre os 60 e 90 DAS, os valores dos índices para o iCOR foram superiores aos da correção Sen2Cor em cerca de 37%. A correção DOS obteve valores semelhantes ao da safra, o NDVI saturou após os 60 DAS, e com valores superiores as outras correções para o GNDVI, porém não apresentou saturação.

O efeito de saturação de NDVI foi mais pronunciado durante a segunda safra, em que os valores desse IV quando calculado a partir da correção DOS apresentaram valor constante, de 1, para todas as datas analisadas. Conseqüentemente, o GNDVI foi mais adequado para captar a variabilidade ao longo do ciclo de desenvolvimento das plantas de milho cultivadas independente da correção atmosférica utilizada e da safra analisada. Esses resultados reforçam a necessidade de escolha mais adequada da correção atmosférica para se utilizar índices de vegetação dentro do contexto de agricultura de precisão.

**Figura 4** – Valores de índices de vegetação NDVI (A), EVI (B), GNDVI (C) e NDRE (D) calculados a partir de imagens de satélite (Sentinel-2) após correção atmosférica utilizando algoritmos DOS, iCOR, Sen2Cor e L1C no período de primeira safra da cultura do milho.

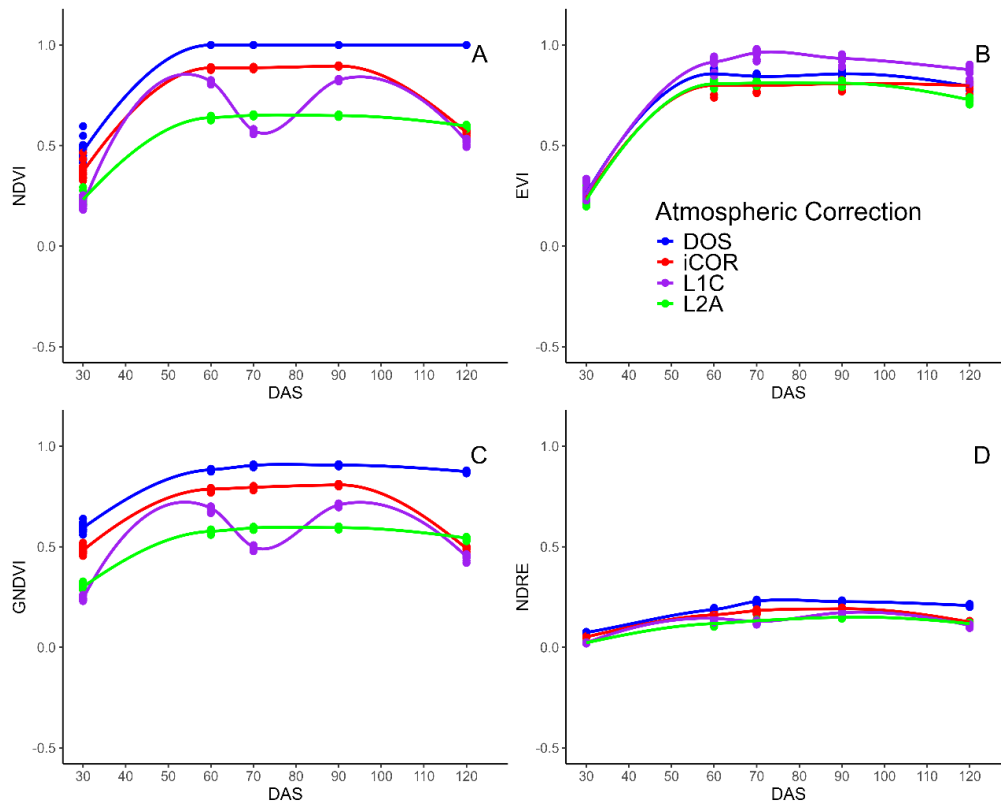


Fonte: Do autor (2024).

Os valores de EVI durante a safra e safrinha foram semelhantes (Figura 5 e 6 – B), sendo que os maiores valores foram encontrados aos 60 DAS. A amplitude média de variação dos

valores de EVI entre as correções (DOS, iCOR, Sen2Cor e L1C) foi baixa, com exceção dos resultados aos 10 DAS para a correção iCOR durante a safra. Os menores valores para o EVI foram observados aos 10 e 120 DAS, próximo aos períodos de semeadura e colheita. Porém a diferença entre as correções atmosféricas foi menor em comparação ao NDVI e GNDVI, demonstrando que o EVI é menos sensível as condições atmosféricas da área de estudo nos dois períodos analisados.

**Figura 5** – Valores de índices de vegetação NDVI (A), EVI (B), GNDVI (C) e NDRE (D) calculados a partir de imagens de satélite (Sentinel-2) após correção atmosférica utilizando algoritmos DOS, iCOR, Sen2Cor e L1C no período de segunda safra da cultura do milho.



Fonte: Do autor (2024).

O NDRE apresentou valores baixos para o período da safra e safrinha (Figura 4 e 5 – D) quando comparados aos outros índices de vegetação analisados (NDVI, GNDVI e EVI), independente da correção atmosférica utilizada (DOS, Sen2Cor e iCOR). Com a correção iCOR apresentando valores negativos para o índice aos 10 DAS. Para ambas as correções e safras analisadas os maiores valores de NDRE foram observados próximos ao 60 DAS. Assim como o EVI, os valores de NDRE não foram impactados pelas correções utilizadas (DOS, iCOR, Sen2COR e L1C) e pela época de cultivo do milho (safra e safrinha)

## 4 DISCUSSÃO

Essa diferença no comportamento da RS entre as duas safras pode ter sido afetada por diversos fatores, tais como, maior exposição de solo, orientação das folhas, condições atmosféricas, biomassa e índice de área foliar (IAF) (Gao *et al.*, 2023). Os principais motivos que explicam a perda de acurácia e precisão em relação o a RS das imagens orbitais é a presença de vapor de água e aerossóis dispersos na atmosfera, onde a depender das condições estes valores podem ser sub ou superestimados (Herman; Browning, 1975). As condições climáticas do local de estudo podem favorecer a formação de nuvens cirrus sub-visuais, que interage de diversas maneiras com a radiação eletromagnética na região do infravermelho (Jensen *et al.*, 1996; Liou, 2002).

Além disso, as variações na reflectância na região do visível pode estar relacionado a efeitos de estresse, onde as plantas tendem a aumentar a reflectância devido à redução da atividade fotossintética (Ponzoni; Disperati, 2001). Isso porque as plantas apresentam máxima absorção da radiação eletromagnética nas bandas do azul (490 nm) e vermelho (665 nm) devido à alta atividade das clorofilas nesses comprimentos de onda e máxima reflectância na banda do verde (560 nm) (Gao *et al.*, 2024).

Já as variações na reflectancia na região do NIR, especialmente a redução desta RS, se relaciona a proximidade da colheita, no qual estas plantas entram em período de senescência, que comumente coincide com a redução nos valores de reflectância na maioria das culturas (Santos *et al.*, 2021). Isso acontece pois existe uma relação direta entre a estrutura celular das folhas com a taxa de reflectância na região do NIR (Slaton; Hunt Jr; Smith, 2001). Desta forma, a perda de tecidos fotossinteticamente ativos faz com que a interação da radiação eletromagnética passa a apresentar maiores taxas de dissipação ao invés de refletir neste comprimento de ondas.

De forma geral, pode-se observar que a correção DOS apresentou valores mais baixos de RS, e valores mais altos de todos os IV's analisados, com exceção do EVI, em comparação a correção Sen2Cor e iCOR. Isso se deve ao fato de que embora a correção DOS possa corrigir os efeitos de dispersão aditiva, ele falha em considerar os efeitos de transmitância multiplicativa, o que pode levar a valores de reflectância imprecisos (Chavez, 1996). No entanto, o DOS continua amplamente utilizado devido à sua simplicidade e aplicabilidade a dados históricos (Lu *et al.*, 2002). Estudos como o de Chang *et al.* (2008), buscam realizar modificações no método de correção DOS, como a Subtração de Objeto Escuro baseada em Solo (GDOS), têm mostrado potencial para melhorar a precisão da reflectância.

Isso indica que métodos de correção atmosférica baseados em modelos físicos superam as técnicas baseadas em imagem para as imagens do Sentinel-2 (Lantzanakis; Mitraka; Chrysoulakis, 2017; Muchsin *et al.*, 2023). O algoritmo Sen2Cor apresentam a maior precisão em tarefas de classificação e introduz maiores correções, especialmente nas bandas do infravermelho (Sola *et al.*, 2018; Muchsin *et al.*, 2023). No entanto, a eficácia desses métodos pode variar dependendo dos tipos de cobertura do solo e das condições atmosféricas (Sola *et al.*, 2018; Lantzanakis; Mitraka; Chrysoulakis, 2017). Apesar de seu desempenho superior, os métodos baseados em modelos físicos são computacionalmente exigentes e requerem informações atmosféricas auxiliares, o que torna sua implementação desafiadora (Lantzanakis *et al.*, 2017). No entanto, todos os algoritmos de correção atmosférica geralmente melhoram a precisão da classificação e são essenciais para aplicações que exigem reflectâncias confiáveis da base da atmosfera (Muchsin *et al.*, 2023; Sola *et al.*, 2018).

Durante a análise dos índices de vegetação GNDVI, NDVI, EVI e NDRE, observamos que o GNDVI, especialmente durante o período de safrinha, com menor presença de vapor d'água na atmosfera, mostrou os melhores resultados para definir o ciclo de desenvolvimento da cultura do milho com a correção atmosférica iCOR. O GNDVI é um índice ainda pouco utilizado, mesmo sendo sensível ao conteúdo de clorofila e menos afetado pelo solo exposto tornando-o adequado para aplicações de agricultura de precisão (Peter *et al.*, 2021; Burns *et al.*, 2022).

Enquanto o NDVI é amplamente utilizado, e pode ser menos sensível a mudanças sutis na saúde da vegetação em comparação com índices baseados em verde (Peter *et al.*, 2021), levando ao efeito de saturação. O NDVI utiliza o contraste entre alta reflectância no infravermelho próximo (NIR) e baixa reflectância no vermelho em vegetação saudável, é sensível à estrutura do dossel e ao conteúdo químico em vegetação esparsa, mas satura em valores mais altos de índice de área foliar (Gamon *et al.*, 1995). Como observado nos resultados, a adoção do método de correção atmosférica DOS estimou valores de reflectância na região do vermelho muito abaixo das correções iCOR e Sen2Cor, próximos a 0, como resultado, o NDVI se aproxima de seu valor máximo teórico de 1.

O EVI apresentou comportamento similar para todas as correções atmosféricas realizadas durante o período de safra e safrinha. Porém Matsushita *et al.* (2007) revelaram que o EVI é mais sensível aos efeitos topográficos do que o NDVI devido ao seu fator de ajuste de solo, enfatizando a necessidade de correção topográfica em terrenos acidentados. Somvanshi e Kumari (2020) compararam vários índices de vegetação em áreas com alta poluição de

partículas, concluindo que índices resistentes à atmosfera, como o EVI e o ARVI, apresentam melhor desempenho nessas condições.

## **5 CONCLUSÃO**

Portanto, a análise para as bandas espectrais demonstrou que milho cultivado nos períodos de safra e safrinha apresentaram comportamento diferentes, especialmente para as bandas do RedEdge e NIR. Porém as diferenças entre as correções atmosféricas foram similares para as duas épocas de cultivo, apresentando estabilidade temporal.

Para a análise das bandas do satélite Sentinel-2 a correção DOS apresentou valores mais baixos quando comparado as correções Sen2Cor e iCOR, especialmente para a banda do vermelho, na qual demonstrou valores próximos a 0 para várias datas. Já correções Sen2Cor e iCOR demonstraram comportamento diferente durante os estádios iniciais, porém em estádios avançados de desenvolvimento da planta de milho apresentaram comportamento similares.

Os índices NDVI e GNDVI apresentaram comportamentos similares em cada época de cultivo analisada, porém o GNDVI sendo menos suscetível a saturação, demonstrando ser uma ferramenta adequada para utilização em lavouras de milho. Já os índices EVI e NDRE não foram impactos pelas correções atmosféricas e nem pela época de cultivo.

## **AGRADECIMENTOS**

Agradecemos à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio e financiamento deste trabalho.

## REFERÊNCIAS

- BARBOZA, Thiago Orlando Costa *et al.* Advanced Farming Strategies Using NASA POWER Data in Peanut-Producing Regions without Surface Meteorological Stations. **AgriEngineering**, v. 6, n. 1, p. 438-454, 2024.
- BIVAND, R.; PEBESMA, E.; GOMEZ-RUBIO, V. Applied spatial data analysis with R. **Springer**, 2. ed., 2013. Disponível em: < <https://asdar-book.org/> >.
- BOSTON, Daniel. Exactextractr: fast extraction from raster datasets using polygons. **R package**, version 0.10.0, 2023. Disponível em: < <https://CRAN.R-project.org/package=exactextractr> >.
- BRAMBILLA, Massimo *et al.* From conventional to precision fertilization: a case study on the transition for a small-medium farm. **AgriEngineering**, v. 3, n. 2, p. 438-446, 2021.
- BURNS, Brayden W. *et al.* Determining nitrogen deficiencies for maize using various remote sensing indices. **Precision Agriculture**, v. 23, n. 3, p. 791-811, 2022.
- CHANDRASEKHAR, Subrahmanyam. Radiative transfer. **Courier Corporation**, 1960.
- CHANG, Jiyul *et al.* Evaluating Modified Atmospheric Correction Methods for Landsat Imagery: Image-Based and Model-Based Calibration Methods. **Communications in soil science and plant analysis**, v. 39, n. 9-10, p. 1532-1545, 2008.
- CHAVEZ, Pat S. Image-based atmospheric corrections-revisited and improved. **Photogrammetric engineering and remote sensing**, v. 62, n. 9, p. 1025-1035, 1996.
- CLEVELAND, William S.; LOADER, Clive. Smoothing by local regression: Principles and methods. *In: Statistical Theory and Computational Aspects of Smoothing: Proceedings of the COMPSTAT'94 Satellite Meeting held in Semmering, Austria, 27-28 August 1994.* **Physica-Verlag HD**, p. 10-49, 1996.
- CONGEDO, Luca. Semi-Automatic Classification Plugin: A Python tool for the download and processing of remote sensing images in QGIS. **Journal of Open Source Software**, v. 6, n. 64, p. 3172, 2021.
- GAIDA, William *et al.* Correção atmosférica em sensoriamento remoto: uma revisão. **Revista Brasileira de Geografia Física**, v. 13, n. 01, p. 229-248, 2020.
- GAMON, John A. *et al.* Relationships between NDVI, canopy structure, and photosynthesis in three Californian vegetation types. **Ecological applications**, v. 5, n. 1, p. 28-41, 1995.
- GAO, Feng *et al.* Near real-time detection of winter cover crop termination using harmonized Landsat and Sentinel-2 (HLS) to support ecosystem assessment. **Science of Remote Sensing**, v. 7, p. 100073, 2023.
- GAO, Weishi *et al.* Remote sensing estimation of sugar beet SPAD based on un-manned aerial vehicle multispectral imagery. **Plos one**, v. 19, n. 6, p. e0300056, 2024.
- GUISAN, Antoine; ZIMMERMANN, Niklaus E. Predictive habitat distribution models in ecology. **Ecological modelling**, v. 135, n. 2-3, p. 147-186, 2000.

HADJIMITSIS, Diofandos G. *et al.* Atmospheric correction for satellite remotely sensed data intended for agricultural applications: impact on vegetation indices. **Natural Hazards and Earth System Sciences**, v. 10, n. 1, p. 89-95, 2010.

HERMAN, Benjamin M.; BROWNING, Samuel R. The effect of aerosols on the earth-atmosphere albedo. **Journal of Atmospheric Sciences**, v. 32, n. 7, p. 1430-1445, 1975.

HIJMANS, R. Raster: geographic data analysis and modeling. **R package**, version 3.6-26, 2023. Disponível em: < <https://CRAN.R-project.org/package=raster> >.

HUETE, Alfredo *et al.* Overview of the radiometric and biophysical performance of the MODIS vegetation indices. **Remote Sensing of Environment**, v. 83, n. 1-2, p. 195-213, 2002.

HUETE, Alfredo R.; LIU, HuiQing; VAN LEEUWEN, Wim JD. The use of vegetation indices in forested regions: issues of linearity and saturation. *In: IGARSS'97. 1997 IEEE international geoscience and remote sensing symposium proceedings. Remote sensing-a scientific vision for sustainable development.* **IEEE**, p. 1966-1968, 1997.

JENSEN, Eric J. *et al.* On the formation and persistence of subvisible cirrus clouds near the tropical tropopause. **Journal of Geophysical Research: Atmospheres**, v. 101, n. D16, p. 21361-21375, 1996.

KEUKELAERE, Liesbeth De *et al.* Atmospheric correction of Landsat-8/OLI and Sentinel-2/MSI data using iCOR algorithm: validation for coastal and inland waters. **European Journal of Remote Sensing**, v. 51, n. 1, p. 525-542, 2018.

LANTZANAKIS, Giannis; MITRAKA, Zina; CHRYSOULAKIS, Nektarios. Comparison of physically and image based atmospheric correction methods for Sentinel-2 satellite imagery. *In: Perspectives on atmospheric sciences.* Springer International Publishing, p. 255-261, 2017.

LIU, Kuo-Nan. An introduction to atmospheric radiation. **Elsevier**, 2002.

LOUHAICHI, Mounir; BORMAN, Michael M.; JOHNSON, Douglas E. Spatially located platform and aerial photography for documentation of grazing impacts on wheat. **Geocarto International**, v. 16, n. 1, p. 65-70, 2001.

LU, D. *et al.* Assessment of atmospheric correction methods for Landsat TM data applicable to Amazon basin LBA research. **International journal of remote sensing**, v. 23, n. 13, p. 2651-2671, 2002.

MATSUSHITA, Bunkei *et al.* Sensitivity of the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) to topographic effects: a case study in high-density cypress forest. **Sensors**, v. 7, n. 11, p. 2636-2651, 2007.

MILLER, Jarrod O.; SHOBER, Amy L.; TARAILA, Jamie. Assessing relationships of cover crop biomass and nitrogen content to multispectral imagery. **Agronomy Journal**, v. 116, n. 3, p. 1417-1427, 2024.

MORAVEC, David *et al.* Effect of atmospheric corrections on NDVI: Intercomparability of Landsat 8, Sentinel-2, and UAV sensors. **Remote Sensing**, v. 13, n. 18, p. 3550, 2021.

MUCHSIN, Fadila *et al.* Effect of atmospheric correction algorithm on landsat-8 and sentinel-2 classification accuracy in paddy field area. **International Journal of Remote Sensing and Earth Sciences (IJReSES)**, v. 20, n. 1, p. 57-65, 2023.

NASA POWER. Prediction of worldwide energy resources. **Nasa Power**, 2024. Disponível em: <<https://power.larc.nasa.gov/>>. Acesso em: 16 mar. 2024.

PAUL, Kenny *et al.* Viable smart sensors and their application in data driven agriculture. **Computers and Electronics in Agriculture**, v. 198, p. 107096, 2022.

PAUSTIAN, Margit; THEUVSEN, Ludwig. Adoption of precision agriculture technologies by German crop farmers. **Precision agriculture**, v. 18, p. 701-716, 2017.

PEBESMA, E.; BIVAND, R. Classes and methods for spatial data in R. **R News**, v. 5, n. 2, p. 9-13, 2005. Disponível em: <<https://CRAN.R-project.org/doc/Rnews/>>.

PEBESMA, E.; BIVAND, R. Spatial ata Science: with applications in R. **The R Journal**, v. 10, n. 1, p. 436-446, 2018.

PETER, Brad G. *et al.* A case for green-based vegetation indices: Plot-scale sUAS imagery related to crop chlorophyll content on smallholder maize farms in Malawi. **Remote Sensing Letters**, v. 12, n. 8, p. 778-787, 2021.

PONZONI, F. J.; DISPERATI, A. A. Comportamento espectral da vegetação. **Sensoriamento remotorefletância dos alvos naturais. Brasília, DF: Universidade de Brasília**, p. 157-99, 2001.

QGIS: spatial without compromise – spatial visualization and decision-making tools for everyone. **QGIS**, 2024. Disponível em: <<https://qgis.org/>>. Acesso em: 15 jan. 2024.

ROUSE JR, John W. *et al.* Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. 1974.

SANTOS, Adão F. *et al.* Using UAV and multispectral images to estimate peanut maturity variability on irrigated and rainfed fields applying linear models and artificial neural networks. **Remote Sensing**, v. 14, n. 1, p. 93, 2021.

SLATON, Michèle R.; HUNT JR, E. R.; SMITH, William K. Estimating near-infrared leaf reflectance from leaf structural characteristics. **American journal of botany**, v. 88, n. 2, p. 278-284, 2001.

SNAP: Sentinel Application Platform. **SNAP**, 2024. Disponível em: <<https://step.esa.int/main/toolboxes/snap/>>. Acesso em: 18 jan. 2024.

SOLA, Ion *et al.* Assessment of atmospheric correction methods for Sentinel-2 images in Mediterranean landscapes. **International journal of applied earth observation and geoinformation**, v. 73, p. 63-76, 2018.

SOMVANSHI, Shivangi S.; KUMARI, Maya. Comparative analysis of different vegetation indices with respect to atmospheric particulate pollution using sentinel data. **Applied Computing and Geosciences**, v. 7, p. 100032, 2020.

TALAVIYA, Tanha *et al.* Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides. **Artificial Intelligence in Agriculture**, v. 4, p. 58-73, 2020.

TAYYEH, Halah Kadhim; MOHAMMED, Ruqayah. Analysis of NASA POWER reanalysis products to predict temperature and precipitation in Euphrates River basin. **Journal of Hydrology**, v. 619, p. 129327, 2023.

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

WICKHAM, H. *et al.* Dplyr: a grammar of data manipulation. **R package**, version 1.1.4, 2023. Disponível em: < <https://CRAN.R-project.org/package=dplyr> >.

## ARTIGO 2

### **Impacto de correções atmosféricas em imagens de satélite na previsão de produtividade de milho utilizando aprendizado de máquina**

#### RESUMO

Técnicas de aprendizado de máquina (AM) são cada vez mais aplicadas na agricultura para melhorar o manejo do solo, detecção de doenças e na previsão de rendimento das culturas. Aliado a estas técnicas modernas de análise de dados, o sensoriamento remoto é uma importante ferramenta, fornecendo dados para monitoramento contínuo de grandes áreas agrícolas, permitindo a identificação de padrões que métodos tradicionais de amostragem não conseguem capturar. Os satélites são a principal fonte de dados devido a sua grande área de cobertura e facilidade de acesso. No entanto, as condições atmosféricas podem impactar a qualidade das imagens coletadas, exigindo correções para melhorar a precisão dos dados. A qualidade dos dados é fundamental para que as técnicas de AM sejam empregadas de forma eficaz, neste sentido é importante investigar quais as melhores técnicas para o processamento digital de imagens (PDI) que otimizem a integração dessas ferramentas na agricultura moderna. Neste contexto, o objetivo deste trabalho foi analisar a influência das técnicas de correção atmosférica Dark Object Subtraction (DOS), Sentinel-2 Correction (Sen2Cor), Image Correction for Atmospheric (iCOR) e dados L1C provenientes da radiação no Topo da Atmosfera (TOA) nos algoritmos de aprendizado de máquina Random Forest (RF), k-Nearest Neighbors (kNN) e Support Vector Machine (SVM) na acurácia e precisão da previsão de produtividade da cultura do milho nos períodos de safra e safrinha, utilizando dados da série temporal do monitoramento e datas únicas ao longo das safras. Os resultados demonstraram que quando utilizados series temporais os diferentes algoritmos não apresentam ganho em precisão e acurácia dos modelos. Já para os modelos de ML analisados o RF apresentou melhores resultados em comparação a SVM e kNN. Para o período de safra os modelos previsão da produtividade se beneficiam da utilizam de séries temporais, enquanto na safrinha a utilização de dados de uma data especifica demonstrou melhores resultados.

**Palavras-chave:** Agricultura de Precisão; sensoriamento remoto; inteligência artificial; processamento digital de imagens.

## **Impact of atmospheric corrections on satellite imagery for maize yield prediction using machine learning**

### **ABSTRACT**

Machine learning (ML) techniques are increasingly applied in agriculture to improve soil management, disease detection, and crop yield prediction. Combined with these modern data analysis techniques, remote sensing is an important tool, providing data for continuous monitoring of large agricultural areas, allowing the identification of patterns that traditional sampling methods cannot capture. Satellites are the main source of data due to their large coverage area and ease of access. However, atmospheric conditions can impact the quality of the images collected, requiring corrections to improve the accuracy of the data. Data quality is essential for ML techniques to be employed effectively, in this sense it is important to investigate the best techniques for digital image processing (DIP) that optimize the integration of these tools in modern agriculture. In this context, the objective of this work was to analyze the influence of the atmospheric correction techniques Dark Object Subtraction (DOS), Sentinel-2 Correction (Sen2Cor), Image Correction for Atmospheric (iCOR) and L1C data from the radiation at the Top of the Atmosphere (TOA) in the machine learning algorithms Random Forest (RF), k-Nearest Neighbors (kNN) and Support Vector Machine (SVM) on the accuracy and precision of corn yield prediction in the harvest and off-season periods. using data from the monitoring time series and unique dates throughout the harvests. The results showed that when time series were used, the different algorithms did not present gains in precision and accuracy of the models. For the ML models analyzed, RF presented better results compared to SVM and kNN. For the harvest period, the yield prediction models benefit from the use of time series, while in the off-season the use of data from a specific date showed better results.

**Keywords:** Precision Agriculture; remote sensing; artificial intelligence; digital image processing.

## 1 INTRODUÇÃO

Técnicas de aprendizado de máquina são cada vez mais aplicadas em múltiplas áreas da agricultura, incluindo manejo do solo, previsão de rendimento, detecção de doenças, gestão de água e monitoramento climático (Veeragandham; Santhi, 2020; Yadav; Alfayeed; Wadhawan, 2020; Sharma *et al.*, 2020). Essas técnicas modernas de análise de dados ajudam a melhorar a produtividade sustentável, a qualidade dos produtos e a eficiência geral do gerenciamento agrícola a medida que o setor agrícola enfrenta desafios como o crescimento populacional e as mudanças climáticas (Veeragandham; Santhi, 2020; Sharma *et al.*, 2020). Modelos de aprendizado de máquina como Random Forest (RF), Support Vector Machines (SVM) e k-Nearest Neighbors (kNN) têm se mostrado eficazes na agricultura de precisão para tarefas como classificação de culturas, previsão de rendimento e recomendação de manejo (Baburao *et al.*, 2023; Thilakarathne *et al.*, 2022).

Aliadas ao sensoriamento remoto, as técnicas de machine learning podem trazer soluções inovadoras, ampliando a capacidade de monitoramento e análise em grande escala de forma não destrutiva e com menor exigência de mão de obra. O sensoriamento remoto, particularmente através de imagens de satélite, possibilita a coleta contínua e não invasiva de dados em grandes áreas agrícolas, essenciais para o monitoramento da saúde das culturas, estimativa de rendimento e detecção de mudanças no uso do solo (Shiratsuchi *et al.*, 2014; Diaz-Gonzalez *et al.*, 2022). Quando integrados com o aprendizado de máquina, os dados de sensoriamento remoto podem revelar padrões e tendências que não são visíveis por métodos tradicionais (Diaz-Gonzalez *et al.*, 2022).

Porém, a utilização de imagens de satélite enfrenta diversos desafios, como distorções atmosféricas, incluindo a interferência por aerossóis e vapor d'água, que podem impactar significativamente a qualidade dos dados e dos índices de vegetação (Vergaz *et al.*, 2000; Welter *et al.*, 2020). Devido a isso, para aplicações em agricultura de precisão a correção atmosférica é essencial para a análise dos índices de vegetação e para a determinação de assinaturas espectrais (Pereira; Silva; Pamboukian, 2016). Métodos como o Dark Object Subtraction (DOS), Sentinel-2 Correction (Sen2Cor) e Image Correction for Atmospheric Effects (iCOR) são amplamente utilizados para corrigir as imagens de satélite, cada um com suas próprias abordagens e vantagens específicas (Ding *et al.*, 2015; Lantzanakis; Mitraka; Chrysoulakis, 2017). Porém as características atmosféricas podem variar significativamente de acordo com a região e a época do ano, tornando ainda mais difícil a escolha do melhor método de correção dos dados orbitais.

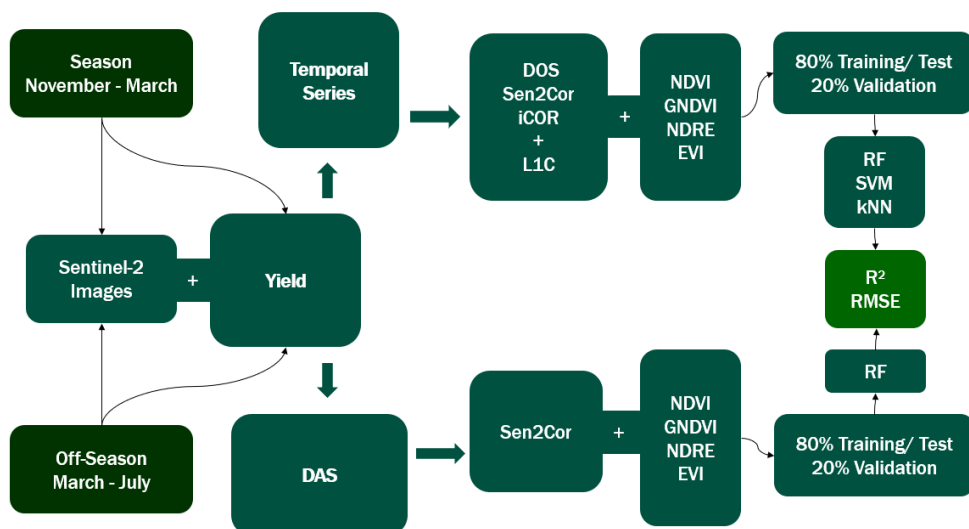
Aliado a isso, para a implementação de algoritmos de aprendizado de máquina a qualidade dos dados utilizados é de fundamental importância, pois dados de baixa qualidade podem levar a previsões imprecisas e tomadas de decisão erradas (Sessions; Valtorta, 2006). Dados ruidosos ou distorcidos podem resultar em modelos que não generalizam bem para novos dados, prejudicando a capacidade de prever rendimentos de culturas ou identificar áreas problemáticas em tempo hábil (Gómez, 2020). Portanto, garantir que os dados sejam corretamente corrigidos e processados é uma etapa crucial no desenvolvimento de modelos eficazes de machine learning para a agricultura.

Assim, este trabalho tem como objetivo avaliar a influência das correções atmosféricas DOS Sen2Cor e iCOR na capacidade dos algoritmos de aprendizado de máquina RF, SVM e kNN de prever a produtividade do milho em duas épocas de cultivo. Compreender essas dinâmicas é essencial para melhorar a precisão das previsões e otimizar a aplicação de recursos agrícolas, ajudando agricultores a tomar decisões mais informadas e eficazes.

## 2 MATERIAL E MÉTODOS

A Figura 1 apresenta o fluxograma que descreve de maneira visual e detalhada o material e métodos adotados neste trabalho. Este diagrama apresenta uma compreensão clara e estruturada dos passos envolvidos, desde a etapa de coleta de dados até a análise e métricas utilizadas na interpretação dos resultados. A seguir serão apresentados mais detalhes dos passos apresentados na Figura 1.

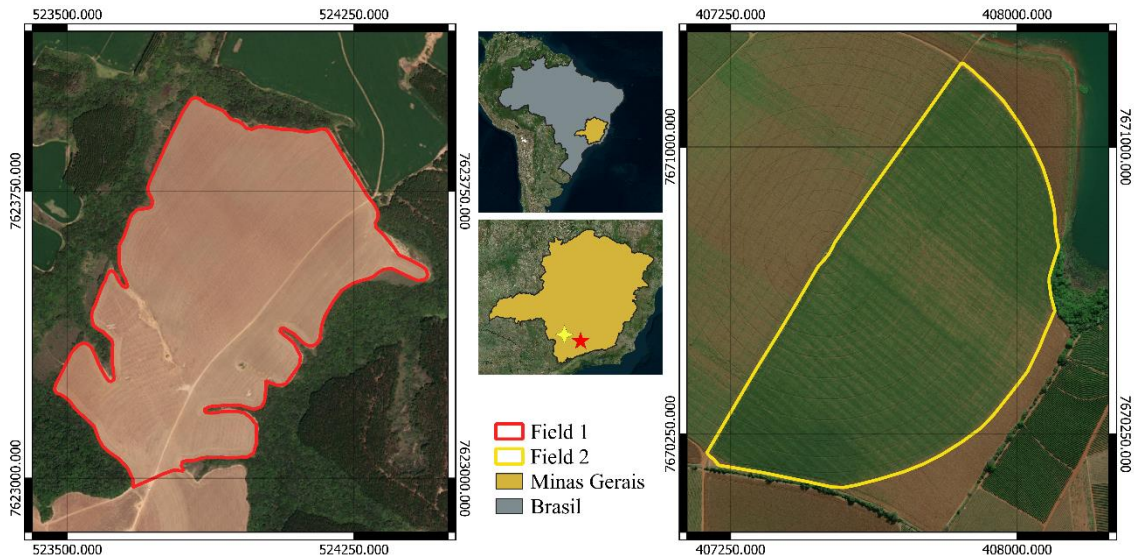
**Figura 1** – Fluxograma do material e métodos para análise do impacto das diferentes técnicas de correção atmosférica quando aliados a modelos de aprendizado de máquina.



Fonte: Do autor (2024).

Foram analisadas duas lavouras comerciais de milho na região sul de Minas Gerais. A primeira lavoura está no município de Luminárias (latitude 21° 30' 39" S, longitude 44° 54' 10" O, altitude 1.141 m) (Figura 2 – A), os dados são da primeira safra de 2021/2022, semeadura em outubro de 2021 e colheita em março de 2022, foi utilizado o híbrido P2501 da empresa Pionner, a área tem 49,56 hectares e não tem sistema de irrigação. A segunda está localizada no município de Campo do Meio (latitude 21° 06' 25" S, longitude 45° 49' 48" O, altitude 774 m) (Figura 2 – B), os dados são referentes a safrinha 2020, semeadura em fevereiro e colheita em junho, foi utilizado o híbrido AS 1633 Pro3 da empresa Bayer, a área tem 50,48 hectares e conta com sistema de irrigação por pivô central.

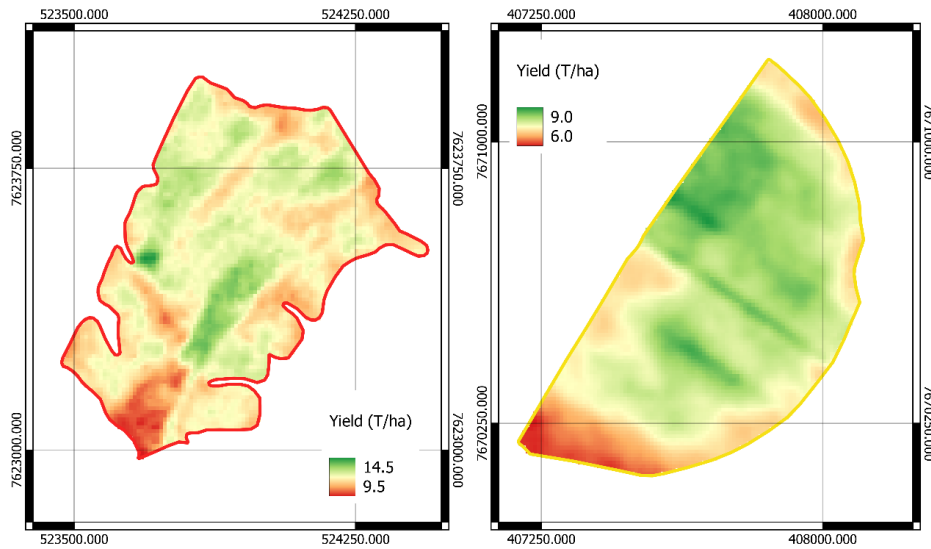
**Figura 2** – Mapa de localização das áreas de cultivo de milho safra (A) e safrinha (B).



Fonte: Do autor (2024).

Os dados de produtividade da Figura 3, foram obtidos diretamente do painel das colhedoras, primeiramente foi realizada a limpeza dos dados brutos pela remoção da bordadura de 30 m. Em seguida foi utilizado o software Map Filter (Maldaner; Molin; Spekken, 2021), utilizando os seguintes parâmetros, “Limite de Variação Geral” de 100%, “Dependência espacial” de 25 m e “Limite de Variação Local de 10%”, para remoção dos outliers presentes nos mapas devido erros no momento da operação de colheita. Após a etapa de limpeza, os dados de produtividade foram interpolados utilizando a técnica de krigagem ordinária. Para o campo 1 (Safra), foram utilizados 12543 pontos e produtividade com reamostragem para 5000, o melhor modelo foi o Exponencial com RMSE de 0,024 e  $R^2$  de 0,93, já para o campo 2 (safrinha) foram utilizados 14367 pontos com reamostragem para 5000, o melhor modelo foi o Linear to Sill com RMSE de 0,050 e  $R^2$  de 0,95.

**Figura 3** – Mapa de produtividade das áreas de cultivo de milho em durante a primeira safra (A) e segunda safra (B).



Fonte: Do autor (2024).

As imagens foram obtidas dos satélites da missão Sentinel-2 do programa espacial da Agência Espacial Europeia (ESA) na plataforma Copernicus Data Space Ecosystem (<https://dataspace.copernicus.eu/>), acessada em 23 de fevereiro de 24), com nível de processamento L1C, referentes à reflectância no Topo da Atmosfera (TOA). As imagens foram adquiridas da constelação Sentinel-2 S2A, com o sensor MSI, nível de processamento L1C, processadas pela versão N0500, na órbita relativa R138 e na região T23KMS. Segue abaixo a tabela 1 com a descrição das imagens com data e porcentagem de cobertura de nuvens.

**Tabela 1** – Descrição da growing season, data, days after sowing (DAS) e cloud coverage das imagens utilizadas neste trabalho.

Growing Season	Data	DAS	Cloud Coverage
Season	11/17/2021	10	15.08
Season	12/27/2021	50	34.83
Season	01/21/2022	75	29.94
Season	02/25/2022	110	11.57
Season	03/07/2022	120	0.45
Off Season	03/27/2020	30	9.24
Off Season	04/26/2020	60	0.12
Off Season	05/06/2020	70	32.53
Off Season	05/26/2020	90	0.02
Off Season	06/25/2020	120	81.83

Fonte: Do autor (2024).

O processamento das imagens foi realizado nos softwares QGIS (2024) e SNAP (2024). A correção DOS foi feita usando o plugin Semi-Automatic Classification Plugin (Congedo, 2021) no QGIS utilizando a aba de “Pré-processamento” >> “Conversão de Imagens”, com a opção “Aplicar correção DOS 1” ativada. A correção iCOR foi realizada com o plugin da empresa VITO no software SNAP da ESA, na aba utilizando a banda B01 para detecção de nuvens, a banda B08 para detecção de corpos de água e o tipo de aerossol "Rural". A correção Sen2Cor foi realizada no SNAP da ESA, utilizando o plugin Sen2Cor 2.11.0, com parâmetros de “Classificação” e “Correção de nuvens cirrus” ativados, e tipo de aerossol "Rural".

Após o processamento das imagens, foi realizado o procedimento de remoção da bordadura de 30 m, para minimizar a interferência de fatores externos ao campo de cultivo. Foi realizada a reamostragem da banda do RedEdge (B06) de 20 x 20 m para 10 x 10 m por pixel utilizando a ferrameta “Reprojetar coordenadas” do menu “Raster” do software QGIS. As áreas de estudo foram divididas em grade regular de 10 x 10 m, devido ao formato irregular da área algumas unidades amostrais foram cortadas, foram selecionadas as amostras com pelo menos 70% da área da unidade amostral completa.

Os índices de vegetação Índice de Vegetação da Diferença Normalizada (NDVI), Índice de Vegetação da Borda Vermelha Normalizada (NDRE), Índice de Vegetação da Diferença Normalizada Verde (GNDVI) e Índice de Vegetação Melhorado (EVI) (Tabela 2) foram calculados para as imagens com três diferentes correções atmosféricas e para as imagens com nível de processamento L1C (TOA) utilizando o software R v.4.3.2, utilizando os pacotes "raster" (Hijmans, 2023), "sp" (Pebesma; Bivand, 2005; Bivand; Pebesma; Gomez-Rubio, 2013) e "sf" (Pebesma; Bivand, 2018).

**Tabela 2** – Equação dos índices de vegetação utilizados para compor o modelo de previsão de produtividade.

Nome	Equação	Autor
NDVI	$\frac{NIR - RED}{NIR + RED}$	Rouse <i>et al.</i> (1974)
NDRE	$\frac{NIR - REDEGE}{NIR + REDEGE}$	Tucker (1979)
GNDVI	$\frac{NIR - GREEN}{NIR + GREEN}$	Louhaichi; Borman; Johnson, 2001
EVI	$G * \frac{(NIR - RED)}{(NIR + C1 * RED - C2 * BLUE + L)}$	Huete <i>et al.</i> (2002)

Legenda: G = 2,5; C1 = 6; C2 = 7,5; L = 1.

Fonte: Do autor (2024).

Após estes procedimentos foram extraídos e tabulados os valores de reflectância das seguintes bandas do Sentinel-2: Azul (Banda 2, 490 nm), Verde (Banda 3, 560 nm), Vermelho (Banda 4, 665 nm), Red Edge (Banda 5, 705 nm) e NIR (Banda 8, 842 nm) e dos IV's NDVI, NDRE, GNDVI e EVI. A extração e tabulação dos valores de reflectância e dos IV's foi realizada no software R v. 4.3.2, utilizando os pacotes "raster" (Hijmans, 2023), "sf" (Pebesma; Bivand, 2018), "exactextractr" (Boston, 2023) e "dplyr" (Wickham *et al.*, 2023).

Após a etapa de extração e tabulação, os dados foram divididos em três conjuntos: treino, teste e validação, visando treinar e avaliar os modelos de aprendizado de máquina de forma robusta. A divisão foi realizada de forma estratificada para garantir a representação proporcional da variabilidade da produtividade em cada conjunto. O conjunto de treino, que compreende 60% dos dados totais, foi utilizado para ajustar os modelos. O conjunto de teste, representando 20% dos dados, foi empregado para avaliar o desempenho dos modelos durante o desenvolvimento e auxiliar na seleção dos melhores hiperparâmetros. O conjunto de validação, também com 20% dos dados, foi utilizado para a avaliação final dos modelos, fornecendo uma medida imparcial do desempenho preditivo.

A validação cruzada foi aplicada aos conjuntos de treino e teste para otimizar os modelos, minimizando o risco sobre ajuste e garantindo uma boa generalização dos modelos para dados não vistos. A utilização de um conjunto de validação separado permitiu uma avaliação final mais precisa, fornecendo uma estimativa confiável de como o modelo se comportaria em cenários reais.

Para a previsão da produtividade, na primeira etapa foram utilizados três algoritmos de aprendizado de máquina: Random Forest (RF), Support Vector Machine (SVM) e k-Nearest Neighbors (KNN). As análises foram realizadas utilizando o software R versão 4.2.3 e o pacote caret para assegurar uma implementação padronizada e comparável entre os modelos. O modelo RF foi ajustado utilizando a função train do pacote caret, com otimização dos hiperparâmetros de número de árvores (ntree) e o número de variáveis selecionadas em cada divisão (mtry) por meio de uma busca em grade (grid search) com validação cruzada de 10-fold. Para o modelo SVM, foi utilizada a função train com o método svmRadial para treinar o modelo com kernel radial, ajustando os hiperparâmetros de regularização (C) e do kernel (sigma) também através de busca em grade com validação cruzada. Já para o modelo KNN, o número de vizinhos (k) foi determinado por validação cruzada utilizando a função train do pacote caret.

Para a segunda etapa de análises, foi selecionado o algoritmo RF em conjunto com a correção atmosférica Sen2Cor. Os dados foram divididos por DAS para treinamento dos

modelos de previsão da produtividade. Seguindo o mesmo método de divisão e ajuste de parâmetros por Cross Validation da primeira etapa de análises.

Todos os modelos foram avaliados utilizando as métricas de Coeficiente de Determinação ( $R^2$ ) e Raiz do Erro Quadrático Médio (RMSE) para determinar o desempenho preditivo.

Coeficiente de Determinação ( $R^2$ ):

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

Raiz do Erro Quadrático Médio (RMSE):

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

### 3 RESULTADOS E DISCUSSÃO

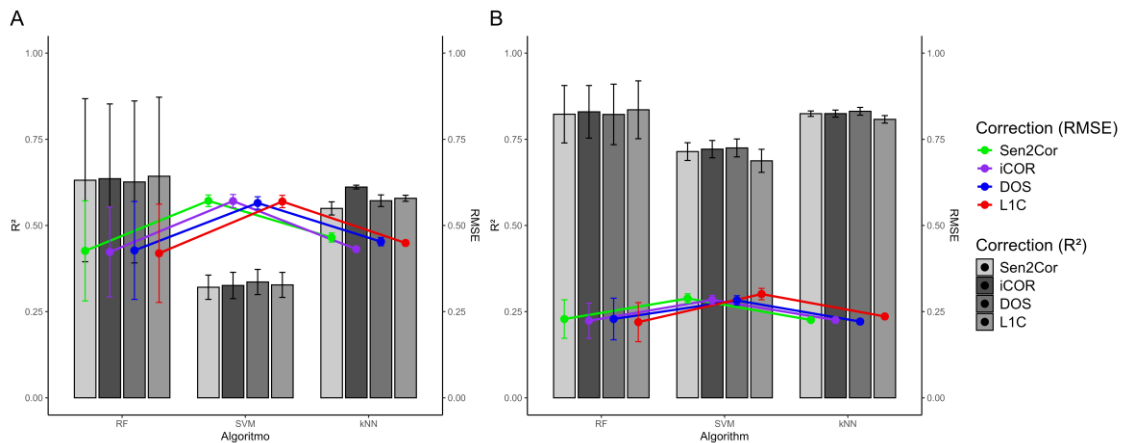
Os modelos de aprendizado de máquina (AM) para previsão de produtividade gerados utilizando dados de índices de vegetação (IV) durante o período da safra e safrinha, não apresentaram diferença entre as correções atmosféricas analisadas durante o treinamento (Figura 4). Os intervalos de confiança calculados para as correções atmosféricas dos modelos apresentaram sobreposição, isso indica que, dentro das margens de incerteza analisadas, os índices de vegetação utilizados são robustos o suficiente para minimizar variações introduzidas por diferentes métodos de correção atmosférica. O que comprova, que a elaboração das equações dos IV's visa minimizar efeitos da luminosidade do solo, ângulos de visualização e condições atmosféricas (Huete; Justice; Liu, 1994).

Em relação aos algoritmos de AM analisados durante o período de safra e safrinha (Figura 4) o Random Forest (RF) se destaca com maiores valores de  $R^2$ , em comparação a SVM e kNN. Porém, diferentemente dos modelos SVM e kNN, que apresentaram baixa amplitude do intervalo de confiança para valores de  $R^2$  durante o treinamento por CV, o RF apresentou maior amplitude entre os limites superior e inferior do intervalo de confiança durante a safra (LS = 0.55 e LI = 0.75) e safrinha (LS = 0.78 e LI = 0.86).

O intervalo de confiança é um parâmetro que auxiliar na escolha do modelo, menores valores de na amplitude no intervalo de confiança indicam previsões mais confiáveis e

transparentes, auxiliando na tomada de decisão e reduzindo a incerteza do modelo. Durante a safra o modelo RF apresentou maior variação em relação ao ajuste, indicando maior dificuldade de generalização da previsão de produtividade em comparação ao período de safrinha. Por isso, destaca-se a maior importância no ajuste dos parâmetros de número de árvores (ntree) e número de variáveis selecionadas (mtry) para o modelo RF, quando comparados aos ajustes dos modelos SVM e kNN, que são menos susceptíveis a variação nos parâmetros, sendo assim mais impactados pela característica dos dados de entrada.

**Figura 4** – Resultados do treinamento de modelos de aprendizado de máquina utilizando diferentes correções atmosféricas (Sen2Cor, iCOR, DOS e L1C) para o período de safra (A) e entressafra (B).



Fonte: Do autor (2024).

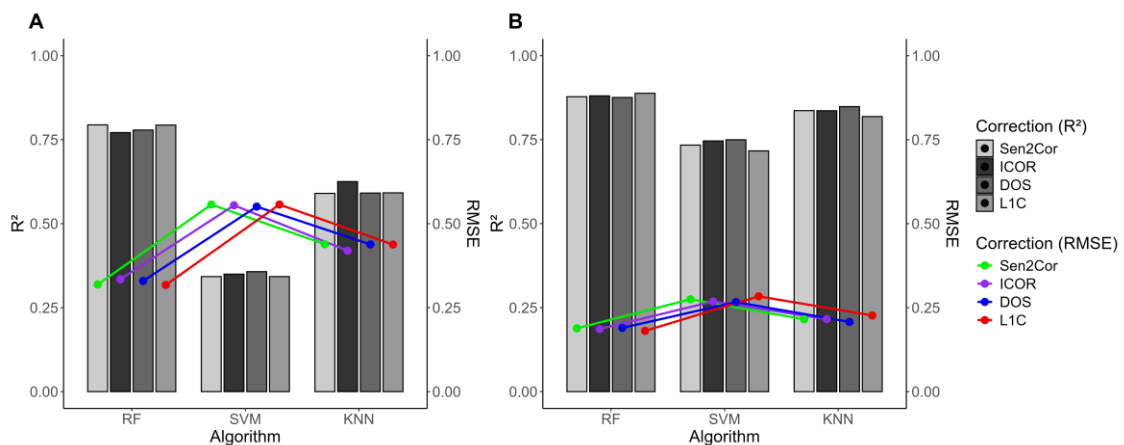
Na etapa de validação dos modelos de AM, assim como na etapa de treinamento, as correções atmosféricas não tiveram impacto na precisão da previsão da produtividade da cultura do milho durante a safra e na safrinha (Figura XX). Os índices de vegetação utilizados como entrada dos modelos (GNDVI, NDVI e NDRE) são baseados na razão entre diferentes bandas, em especial a banda do NIR, assim apresentam a característica de serem menos influenciados por impactos atmosféricos leves, que são aqueles onde a interferência de nuvens, vapor de água e aerossóis é pequena (Vergaz *et al.*, 2000). No entanto, a correção atmosférica continua sendo crucial para cálculos precisos dos IVs quando se objetiva comparar diferentes coberturas vegetais (Francisco *et al.*, 2020).

Os modelos de RF foram os que apresentaram os maiores valores de precisão ( $R^2 = 0,79$ ) e menores valores de erro ( $RMSE = 0,31 \text{ T ha}^{-1}$ ) durante o período da safra, quando comparado com os demais modelos (SVM e kNN). Os valores de  $R^2$  para os modelos SVM foram menores que 0,36 com erros de até de  $0,50 \text{ T ha}^{-1}$ . Da mesma forma, na safrinha, a performance do RF foi superior ao demais (Figura 5B), indicando que este modelo apresenta maior precisão e

acurácia, independente da época de cultivo. O algoritmo de aprendizado de máquina Random Forest tem a capacidade de trabalhar com estruturas de dados complexas, alta dimensionalidade e tamanho amostrais variados (Qi, 2012), consegue assim compensar pequenas variações nos dados de entrada.

Dentre as correções atmosféricas analisadas, a Sen2Cor se destaca, pois foi desenvolvido especialmente para processamento de dados da missão Sentinel, disponibilizando dados de profundidade óptica do aerossol (OAT) vapor de água (WV) e classificação da cena (SCL) (Main-Knor *et al.*, 2017). Como observado, as diferenças entre os dados corrigidos por vários algoritmos podem ser reduzidas em cenários com efeitos atmosféricos leves (Valdivieso-Ros; Alonso-Sarria; Gomariz-Castillo, 2021). Porém, em outras situações, o Sen2Cor se mostrou eficiente na correção dos efeitos atmosféricos (Marujo *et al.*, 2021), no aumento da precisão em algoritmos de ML que utilizam IV para classificação (Sola *et al.*, 2018) e na validação dos valores de reflectância da superfície observados em diferentes alvos (Sola *et al.*, 2018).

**Figura 5** – Resultados da validação de modelos de aprendizado de máquina utilizando diferentes correções atmosféricas (Sen2Cor, iCOR, DOS e L1C) para o período de safra (A) e entressafra (B).



Fonte: Do autor (2024).

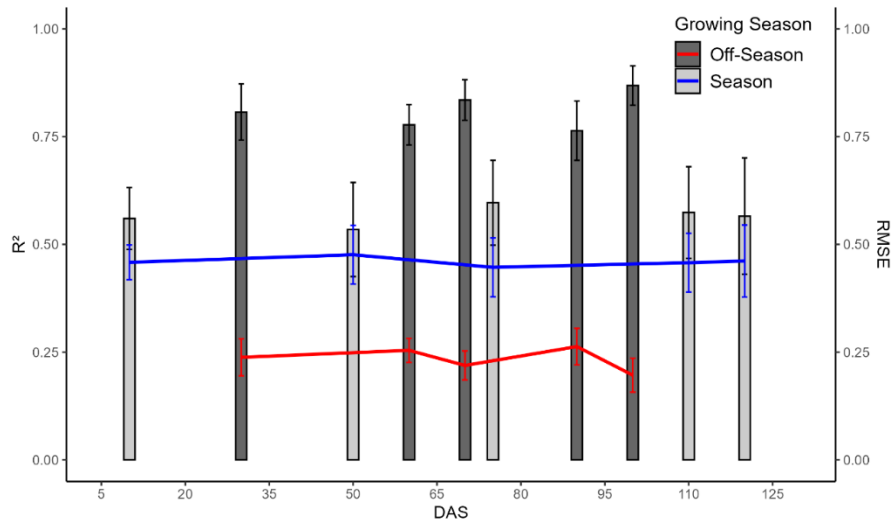
Após a etapa inicial de análise, na qual foi identificado que o modelo Random Forest (RF) apresentou o melhor desempenho entre os métodos de aprendizado de máquina testados (RF, SVM e KNN) e que as diferentes correções atmosféricas (iCOR, Sen2Cor, DOS e L1C) não resultaram em variações na acurácia dos modelos, optou-se por utilizar o RF e a correção Sen2Cor nas análises subsequentes. O foco dessas novas análises foi examinar o impacto de diferentes datas de coleta de imagens na previsão da produtividade do milho. Ao considerar

múltiplos períodos ao longo do ciclo de desenvolvimento da cultura, buscou-se avaliar se a data de aquisição das imagens influenciava a precisão do modelo na predição da produtividade, permitindo uma compreensão mais abrangente da relação entre o estágio de desenvolvimento da cultura e a performance do modelo preditivo.

A análise do modelo RF com dados da correção atmosférica Sen2Cor apresentou resultados diferentes para safra e safrinha (Figura 6). Durante a safra pode-se observar maior amplitude dos intervalos de confiança para valores de  $R^2$  e RMSE, em comparação a safrinha, indicando maior dificuldade no ajuste dos parâmetros do modelo. Além disso a capacidade de previsão dos modelos da safra foi inferior para a safra em comparação aos modelos da safrinha. Isso indica que as condições atmosféricas, variabilidade genética dos cultivares e ciclo de desenvolvimento do milho afetam a capacidade de previsão da produtividade dos modelos utilizados.

Durante a safra, a maior variabilidade climática, com chuvas intensas, alta umidade e nebulosidade, interagem com fatores fisiológicos das plantas, gerando maior complexidade aos fatores que determinam a produtividade de várias culturas e assim dificultando a previsão (Picini *et al.*, 1999; Ely; Almeida; Neto, 2003; Galon *et al.*, 2010) e assim dificultando a previsão. Devido a isso, estudos anteriores já comprovaram que a inclusão de dados climáticos aos modelos de previsão de produtividade, para as culturas do milho e da cana-de-açúcar, fornece um ganho na precisão e acurácia, sendo que a importância das variáveis climáticas pode até preceder os IV's em relação a importância para a previsão (Silva-Fuzzo; Praela-Pantano; Camargo, 2015). Além dos fatores abióticos, a variabilidade genética do milho, decorrente do uso de diferentes cultivares entre safra e safrinha, também influencia os resultados dos modelos de previsão. A relação entre produtividade e índices de vegetação é complexa, pois os diferentes índices de vegetação (IVs) respondem de maneira distinta às plantas de milho devido a uma série de fatores, incluindo características genéticas, estado nutricional das plantas e estágio de desenvolvimento da cultura.

**Figura 6** – Resultados do treinamento dos modelos de aprendizado de máquina do algoritmo Random Forest, utilizando correção Sen2Cor em diferentes datas, para colheita e entressafra.

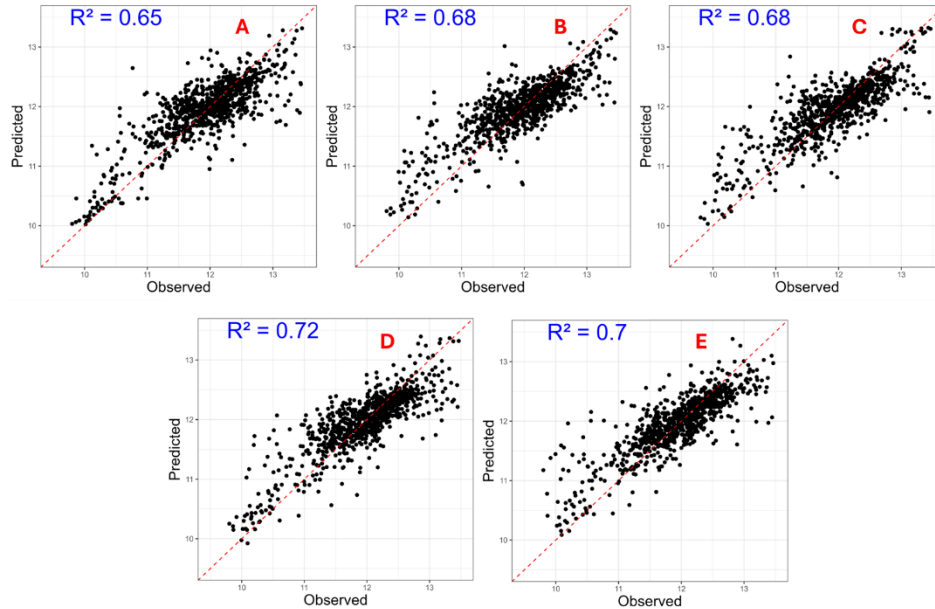


Fonte: Do autor (2024).

As diferentes datas de coleta das imagens não apresentaram diferenças em relação a precisão e acurácia para a safra e safrinha na etapa de treinamento dos modelos, no qual as métricas de  $R^2$  e RMSE apresentaram sobreposição dos intervalos de confiança. Isso se deve ao fato de que, nas fases iniciais de desenvolvimento da cultura do milho, o potencial das sementes já está relacionado a produtividade final, no qual originadas de sementes de alta qualidade demonstram desempenho superior em área foliar, acúmulo de matéria seca e taxas de crescimento em comparação com sementes de baixa qualidade (Ludwig *et al.*, 2010), o que são fatores decisivos para a determinação da produtividade final.

No entanto, é importante destacar que, embora as previsões feitas nas fases iniciais do desenvolvimento possam ser úteis, elas apresentam riscos associados a variações que podem ocorrer ao longo do restante do ciclo de crescimento. Mudanças nas condições ambientais, pragas e doenças, podem impactar o desenvolvimento das plantas e, conseqüentemente, afetar a produtividade final. Portanto, a coleta de imagens em estágios iniciais do ciclo de desenvolvimento do milho fornece dados importantes para composição do modelo de previsão de produtividade, porém, se utilizadas de forma isolada, a precisão da previsão pode ser influenciada por eventos imprevistos que ocorrem à medida que a cultura se desenvolve.

**Figura 7** – Resultado da validação dos modelos de Random Forest utilizando correção Sen2Cor em 10 DAS (A), 50 DAS (B), 75 DAS (C), 110 DAS (D) e 120 DAS (E), para o período de safra.



Fonte: Do autor (2024).

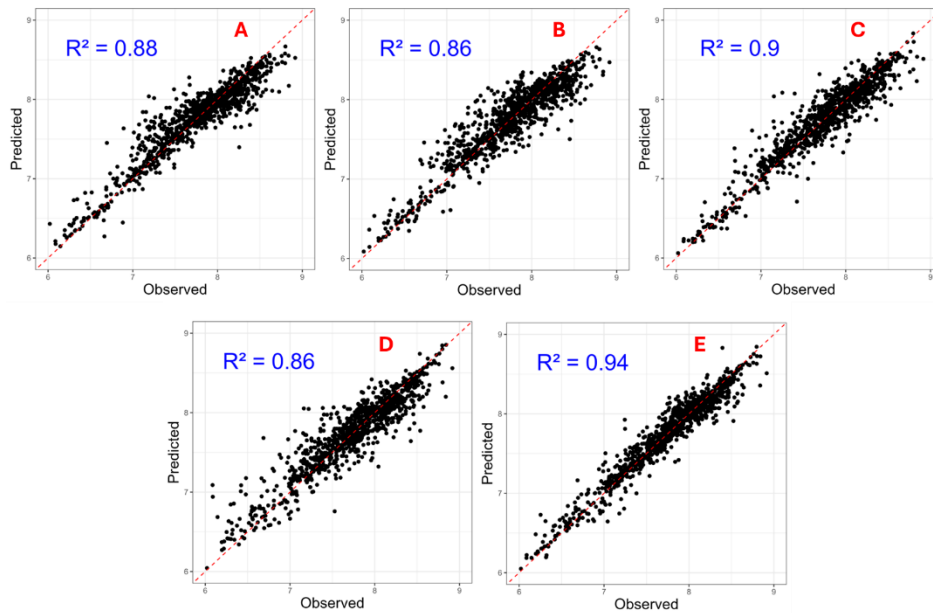
Os resultados da fase de validação dos modelos RF para o período de safra apresentaram os maiores valores aos 110 DAS ( $R^2 = 0,72$ ) (Figura X – D) e os menores valores aos 10 DAS ( $R^2 = 0,65$ ) (Figura X – A). Para o período de safrinha, os valores mais altos aos 120 DAS ( $R^2 = 0,94$ ) (Figura X – E) e valores mais baixos aos 60 DAS ( $R^2 = 0,86$ ) (Figura X – B). Ambos os períodos analisados (Safra e Safrinha) apresentaram uma variação de 10% nos valores de  $R^2$  entre o melhor e o pior resultado, demonstrando um ganho de precisão e acurácia relacionado a escolha do melhor momento para obtenção dos dados para compor o modelo.

A cultura do milho apresenta uma dinâmica de crescimento e desenvolvimento em que os componentes de rendimento (como o número de grãos por espiga, o peso dos grãos e o número de espigas por planta) são consolidados em fases mais avançadas, principalmente após o florescimento (estágio R1) e durante o enchimento de grãos (R2-R6).

Nessas fases, os processos fisiológicos que determinam o rendimento final estão praticamente concluídos, como a alocação de nutrientes e assimilados para os grãos, o que confere uma maior estabilidade à produtividade esperada. Além disso, nessa etapa, o índice de área foliar (LAI), a interceptação da radiação solar e a eficiência de uso da água e nutrientes atingem seus picos, o que reflete diretamente nos índices de vegetação e na performance dos modelos de previsão.

Por isso, os modelos que utilizam dados próximos ao período de colheita têm mais informações consolidadas sobre o estado final da cultura, permitindo uma maior precisão nas estimativas de produtividade.

**Figura 8** – Resultado da validação dos modelos de Random Forest utilizando correção Sen2Cor em 30 DAS (A), 60 DAS (B), 70 DAS (C), 90 DAS (D) e 120 DAS (E), para o período de entressafra.



Fonte: Do autor (2024).

Em contraste, o modelo de Random Forest (RF) que utilizou dados de toda a série temporal (Figura 4) apresentou um desempenho superior, alcançando um  $R^2$  de 0,79, o que representa um aumento de 10% em relação ao melhor valor obtido na segunda etapa das análises (Figura 6). No entanto, para o período da safrinha, o mesmo modelo que considerou toda a série temporal (Figura 5A) obteve resultados 10% inferiores quando comparado ao melhor modelo que utilizou apenas uma data específica (Figura 6). As diferenças entre a safra e a safrinha ficam claras pela forma como os índices refletem os componentes de produção das plantas de milho ao longo do ciclo.

Na safra, análise de séries temporais de IVs pode fornecer insights sobre os padrões de crescimento das culturas e ajudar a detectar anomalias no desenvolvimento das plantas (Silva *et al.*, 2021). A integração de dados de séries temporais de IV com modelos globais dinâmicos de vegetação pode melhorar a avaliação e a previsão das respostas da vegetação às mudanças ambientais (Maignan *et al.*, 2011). Como o desenvolvimento na safra não é significativamente

impactado por estresses, os IVs fornecem uma visão clara do crescimento, resultando em modelos mais precisos.

No caso da safrinha, a variabilidade imposta pelo déficit hídrico torna os IVs menos estáveis ao longo do ciclo. Índices de vegetação (IVs), como o NDVI, são amplamente utilizados para monitorar o desenvolvimento das culturas, no entanto, a sensibilidade dos IVs varia de acordo com o tipo de vegetação, estágio fenológico e ao estresse hídrico (Araújo; Galvão; Dalagnol, 2023; Jackson; Slater; Pinter, 1983). O estresse hídrico durante os estágios de desenvolvimento, leva a quedas abruptas nos valores dos índices de vegetação, mesmo que a planta ainda tenha potencial para produzir bem em momentos posteriores, em que nos estágios críticos, como o pendramento e enchimento de grãos é feito o uso do sistema de irrigação. Assim, o uso de uma série temporal completa para o período de safrinha incorpora períodos de estresse severo que não necessariamente refletem o potencial produtivo final da cultura, resultando em uma menor precisão nos modelos.

#### **4 CONCLUSÕES**

Para as condições atmosféricas das lavouras analisadas neste trabalho, não houve diferença entre os métodos de correção analisados (DOS, Sen2Cor, iCOR e L1C). Porém, dentre os três algoritmos de ML (RF, SVM e kNN) o RF se destacou com melhor poder de previsão, apresentando os melhores resultados em termos de precisão e acurácia.

Quando analisada a melhor data para a obtenção das imagens do satélite Sentinel-2 para compor os modelos de ML, no período da safra foi aos 110 DAS e na safrinha aos 120 DAS. Isso demonstra que estágios finais do desenvolvimento, quando os componentes de rendimento já estão estabelecidos, fornece uma estimativa mais acurada e precisa.

Porém, os resultados de modelos de séries temporais e modelos com data única apresentaram resultados distintos para safra e safrinha. Durante a safra, com menos variação no estado hídrico das plantas de milho, o uso de séries temporais apresentou resultados superiores, já no período de safrinha o uso de uma data específica obteve melhores resultados.

#### **AGRADECIMENTOS**

Agradecemos à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio e financiamento deste trabalho.

## REFERÊNCIAS

- ARAÚJO, Juliana de Abreu; GALVÃO, Lênio Soares; DALAGNOL, Ricardo. Sensitivity of hyperspectral vegetation indices to rainfall seasonality in the Brazilian savannahs: an analysis using PRISMA data. **Remote Sensing Letters**, v. 14, n. 3, p. 277-287, 2023.
- BABURAO, Patil Sagar *et al.* Review of Machine Learning Model Applications in Precision Agriculture. *In: International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022)*. **Atlantis Press**, p. 916-930, 2023.
- BIVAND, R.; PEBESMA, E.; GOMEZ-RUBIO, V. Applied spatial data analysis with R. **Springer**, 2. ed., 2013. Disponível em: < <https://asdar-book.org/> >.
- BOSTON, Daniel. Exactextractr: fast extraction from raster datasets using polygons. **R package**, version 0.10.0, 2023. Disponível em: < <https://CRAN.R-project.org/package=exactextractr> >.
- CONGEDO, Luca. Semi-Automatic Classification Plugin: A Python tool for the download and processing of remote sensing images in QGIS. **Journal of Open Source Software**, v. 6, n. 64, p. 3172, 2021.
- DIAZ-GONZALEZ, Freddy A. *et al.* Machine learning and remote sensing techniques applied to estimate soil indicators—review. **Ecological Indicators**, v. 135, p. 108517, 2022.
- DING, Huimei *et al.* An improved dark-object subtraction technique for atmospheric correction of Landsat 8. *In: MIPPR 2015: Remote Sensing Image Processing, Geographic Information Systems, and Other Applications*. **Spie**, p. 128-135, 2015.
- ELY, Deise Fabiana; ALMEIDA, Ivan Rodrigues de; NETO, João Lima Sant'Anna. Variabilidade climática e o rendimento da cultura do milho no estado do Paraná: algumas implicações políticas e econômicas. **GEOGRAFIA (Londrina)**, v. 12, n. 1, p. 495-508, 2003.
- FRANCISCO, Cristiane Nunes *et al.* Análise do impacto da correção atmosférica no cálculo do Índice de Vegetação por Diferença Normalizada a partir de Imagem Landsat 8/OLI. **Revista Brasileira de Geografia Física**, v. 13, n. 1, p. 076-086, 2020.
- GALON, Leandro *et al.* Influência dos fatores abióticos na produtividade da cultura do milho. **Revista Tropic—Ciências Agrárias e Biológicas**, v. 4, n. 3, p. 18, 2010.
- GÓMEZ, Carlos Alejandro Ramírez. Aplicación del machine learning en agricultura de precisión. **Revista Cintex**, v. 25, n. 2, p. 14-27, 2020.
- HIJMANS, R. Raster: geographic data analysis and modeling. **R package**, version 3.6-26, 2023. Disponível em: < <https://CRAN.R-project.org/package=raster> >.
- HUETE, Alfredo *et al.* Overview of the radiometric and biophysical performance of the MODIS vegetation indices. **Remote Sensing of Environment**, v. 83, n. 1-2, p. 195-213, 2002.
- HUETE, Alfredo; JUSTICE, Christopher; LIU, Huiqing. Development of vegetation and soil indices for MODIS-EOS. **Remote Sensing of environment**, v. 49, n. 3, p. 224-234, 1994.

JACKSON, R. D.; SLATER, P. N.; PINTER JR, P. J. Discrimination of growth and water stress in wheat by various vegetation indices through clear and turbid atmospheres. **Remote sensing of environment**, v. 13, n. 3, p. 187-208, 1983.

LANTZANAKIS, Giannis; MITRAKA, Zina; CHRYSOULAKIS, Nektarios. Comparison of physically and image based atmospheric correction methods for Sentinel-2 satellite imagery. *In: Perspectives on atmospheric sciences*. **Springer International Publishing**, p. 255-261, 2017.

LOUHAICHI, Mounir; BORMAN, Michael M.; JOHNSON, Douglas E. Spatially located platform and aerial photography for documentation of grazing impacts on wheat. **Geocarto International**, v. 16, n. 1, p. 65-70, 2001.

LUDWIG, MARCOS PAULO *et al.* Desempenho de sementes e plantas de milho híbrido originadas de lotes de sementes com alta e baixa qualidade fisiológica. **Revista Brasileira de Milho e Sorgo**, v. 8, n. 01, 2009.

MAIGNAN, F. *et al.* Evaluation of a global vegetation model using time series of satellite vegetation indices. **Geoscientific Model Development**, v. 4, n. 4, p. 1103-1114, 2011.

MAIN-KNORN, Magdalena *et al.* Sen2Cor for sentinel-2. *In: Image and Signal Processing for Remote Sensing XXIII*. **SPIE**, p. 37-48. 2017.

MALDANER, Leonardo Felipe; MOLIN, José Paulo; SPEKKEN, Mark. Methodology to filter out outliers in high spatial density data to improve maps reliability. **Scientia Agricola**, v. 79, n. 1, p. e20200178, 2021.

MARUJO, Rennan F. B. *et al.* Evaluating the impact of LaSRC and Sen2cor atmospheric correction algorithms on Landsat-8/OLI and Sentinel-2/MSI data over AERONET stations in Brazilian territory. **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 3, p. 271-277, 2021.

PEBESMA, E.; BIVAND, R. Classes and methods for spatial data in R. **R News**, v. 5, n. 2, p. 9-13, 2005. Disponível em: <<https://CRAN.R-project.org/doc/Rnews/>>.

PEBESMA, E.; BIVAND, R. Spatial Data Science: with applications in R. **The R Journal**, v. 10, n. 1, p. 436-446, 2018.

PEREIRA, Ludmily da Silva; SILVA, Débora de Oliveira; PAMBOUKIAN, Sergio Vicente Denser. Sensoriamento remoto aplicado à agricultura de precisão no cultivo de bambu. **Revista Mackenzie de Engenharia e Computação**, v. 16, n. 1, 2016.

PICINI, Angélica Giarolla *et al.* Desenvolvimento e teste de modelos agrometeorológicos para a estimativa de produtividade do cafeeiro. **Bragantia**, v. 58, p. 157-170, 1999.

QGIS: spatial without compromise – spatial visualization and decision-making tools for everyone. **QGIS**, 2024. Disponível em: <<https://qgis.org/>>. Acesso em: 15 jan. 2024.

QI, Yanjun. Random forest for bioinformatics. **Ensemble machine learning: Methods and applications**, p. 307-323, 2012.

ROUSE JR, John W. *et al.* Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. 1974.

SESSIONS, Valerie; VALTORTA, Marco. The Effects of Data Quality on Machine Learning Algorithms. **ICIQ**, v. 6, p. 485-498, 2006.

SHARMA, Abhinav *et al.* Machine learning applications for precision agriculture: A comprehensive review. **IEEE Access**, v. 9, p. 4843-4873, 2020.

SHIRATSUCHI, Luciano Shozo *et al.* Sensoriamento remoto: conceitos básicos e aplicações na agricultura de precisão. *In*: BERNARDI, Alberto Carlos de Campos *et al.* Agricultura de precisão: resultados de um novo olhar. **Embrapa**, 1. ed., p. 58-73, 2014.

SILVA, Pedro A. de Azevedo *et al.* Time series analysis of the enhanced vegetation index to detect coffee crop development under different irrigation systems. **Journal of Applied Remote Sensing**, v. 15, n. 1, p. 014511-014511, 2021.

SILVA, Vinícius Bueno da; SHIMADA, Belmiro Saburo; CANDIDO, Gabriel. Os componentes de produção e os aspectos fisiológicos da planta. **Revista Multidisciplinar de Educação e Meio Ambiente**, v. 2, n. 3, p. 26-26, 2021.

SILVA-FUZZO, Daniela Fernanda da; PRELA-PANTANO, Angélica; CAMARGO, Marcelo Bento Paes. Modelagem agrometeorológica para estimativa de produtividade de soja para o Vale do Médio Paranapanema-SP. **Irriga**, v. 20, n. 3, p. 490-501, 2015.

SNAP: Sentinel Application Platform. **SNAP**, 2024. Disponível em: <<https://step.esa.int/main/toolboxes/snap/>>. Acesso em: 18 jan. 2024.

SOLA, Ion *et al.* Assessment of atmospheric correction methods for Sentinel-2 images in Mediterranean landscapes. **International journal of applied earth observation and geoinformation**, v. 73, p. 63-76, 2018.

THILAKARATHNE, Navod Neranjan *et al.* A cloud enabled crop recommendation platform for machine learning-driven precision farming. **Sensors**, v. 22, n. 16, p. 6299, 2022.

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

VALDIVIESO-ROS, Carmen; ALONSO-SARRIA, Francisco; GOMARIZ-CASTILLO, Francisco. Effect of different atmospheric correction algorithms on sentinel-2 imagery classification accuracy in a semiarid mediterranean area. **Remote Sensing**, v. 13, n. 9, p. 1770, 2021.

VEERAGANDHAM, Syamasudha; SANTHI, H. A review on the role of machine learning in agriculture. **Scalable Computing: Practice and Experience**, v. 21, n. 4, p. 583-589, 2020.

VERGAZ, R. *et al.* Estudio de la influencia de los aerosoles sobre la reflectancia de los canales 1 y 2 del sensor AVHRR-NOAA y el NDVI. **Revista de teledetección: Revista de la Asociación Española de Teledetección**, n. 13, p. 2, 2000.

WELTER, André *et al.* Distorções causadas pela falta de calibração radiométrica e efeitos atmosféricos na análise de dados de sensoriamento remoto orbital. **Geografia**, v. 45, n. 2, p. 361-383, 2020.

WICKHAM, H. *et al.* Dplyr: a grammar of data manipulation. **R package**, version 1.1.4, 2023. Disponível em: < <https://CRAN.R-project.org/package=dplyr> >.

YADAV, Neha; ALFAYEED, S. M.; WADHAWAN, A. Machine learning in agriculture: techniques and applications. **International Journal of Engineering Applied Sciences and Technology**, v. 5, n. 7, p. 118-122, 2020.

## CONSIDERAÇÕES FINAIS

Este estudo destaca a importância de escolher cuidadosamente as correções atmosféricas e as técnicas de machine learning (ML) para aprimorar o monitoramento e a previsão de produtividade do milho. Embora este estudo tenha focado em três métodos de correção atmosférica Dark Object Subtraction (DOS), Image Correction for Atmospheric Effects (iCOR), e Sentinel-2 Correction (Sen2Cor) e em três algoritmos de ML-Random Forest, k-Nearest Neighbors, e Support Vector Machine é importante reconhecer que existem outras correções atmosféricas e técnicas de ML disponíveis que podem oferecer resultados ainda mais precisos.

Apesar das limitações, os resultados obtidos já indicam uma direção clara: as correções Image Correction for Atmospheric Effects (iCOR) e Sentinel-2 Correction (Sen2Cor) demonstraram ser eficazes na caracterização do ciclo do milho e na previsão de produtividade, especialmente quando aplicadas junto ao algoritmo Random Forest, que se destacou em termos de desempenho. A continuidade deste trabalho pode explorar outras correções atmosféricas, como Second Simulation of the Satellite Signal in the Solar Spectrum (6S) e Aerosol Optical Depth correction (AOD), além de outras técnicas de ML, como redes neurais profundas (Deep Learning) e Gradient Boosting Machines, que podem melhorar ainda mais a precisão das previsões. Portanto, o estudo já demonstra resultados promissores, mas futuras investigações serão essenciais para otimizar o uso de dados de sensoriamento remoto na agricultura de precisão.