

**UNIVERSIDADE ESTADUAL PAULISTA – UNESP
CÂMPUS DE JABOTICABAL**

**IMAGENS ORBITAIS DE ALTA RESOLUÇÃO E REDES NEURAIAS
ARTIFICIAIS PARA ESTIMATIVA DO TEOR DE SACAROSE EM CANA-
DE-AÇÚCAR**

**Paulo Henrique Cardoso
Engenheiro Agrônomo**

2024

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**IMAGENS ORBITAIS DE ALTA RESOLUÇÃO E REDES NEURAIS
ARTIFICIAIS PARA ESTIMATIVA DO TEOR DE SACAROSE EM CANA-
DE-AÇÚCAR**

Discente: Paulo Henrique Cardoso
Orientador: Prof. Dr. Rouverson Pereira da Silva
Coorientadoras: Prof. Dra. Tatiana Fernanda Canata
Dra. Samira Luns Hatum de Almeida

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

C268i

Cardoso, Paulo Henrique

Imagens orbitais de alta resolução e redes neurais artificiais para estimativa do teor de sacarose em cana-de-açúcar / Paulo Henrique Cardoso. -- Jaboticabal, 2024

68 p. : il., tabs., fotos, mapas

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientador: Rouverson Pereira da Silva

Coorientador: Tatiana Fernanda Canata

1. Agricultura. 2. Ciência e tecnologia. 3. Cana-de-açúcar. 4. Sacarose. 5. Inteligência artificial. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Impacto potencial desta pesquisa

Social: A integração de sensoriamento remoto, aprendizado de máquina e dados multiespectrais no monitoramento da cana-de-açúcar está diretamente alinhada aos Objetivos de Desenvolvimento Sustentável (ODS). Esses avanços permitem a diferenciação de cultivares, a estimativa da qualidade da produção e o mapeamento de matérias-primas de forma não invasiva. Tais práticas promovem a agricultura sustentável, reduzem a intensidade do trabalho, fomentam a educação e o desenvolvimento de habilidades, e aprimoram o desenvolvimento rural, em conformidade com o ODS 1: Erradicação da pobreza e o ODS 4: Educação de qualidade. A troca de conhecimentos e a construção de comunidades são apoiadas, contribuindo para o ODS 10: Redução das desigualdades.

Ambiental: A eficiência no uso de recursos é melhorada, permitindo decisões precisas que reduzem o desperdício e minimizam a pegada ambiental das práticas agrícolas, conforme o ODS 12: Consumo e produção responsáveis e o ODS 13: Ação contra a mudança global do clima. A identificação de tempos ótimos de colheita e aplicação de recursos leva a menores emissões, conservação de energia e promoção da sustentabilidade agrícola. Estas práticas apoiam a saúde do solo, previnem a superexploração de recursos e auxiliam na preservação da biodiversidade, em alinhamento com o ODS 15: Vida terrestre.

Econômico: A otimização do gerenciamento de recursos e das decisões de colheita aumenta a produtividade da cana-de-açúcar, a renda e a lucratividade, contribuindo para o ODS 8: Trabalho decente e crescimento econômico. A melhoria da qualidade da produção aumenta a competitividade no mercado, reduz custos e reforça a sustentabilidade econômica, de acordo com o ODS 9: Indústria, inovação e infraestrutura. A adoção de tecnologias avançadas atrai investimentos, incentiva a inovação e impulsiona o desenvolvimento de infraestrutura, beneficiando economias rurais e urbanas. O fortalecimento da cadeia de valor da cana-de-açúcar, a criação de empregos e a intensificação das exportações promovem o crescimento econômico, assegurando um futuro estável e próspero para as regiões produtoras.

Potential Impact of This Research

Social: The integration of remote sensing, machine learning, and multispectral data in sugarcane monitoring is directly aligned with the Sustainable Development Goals (SDGs). These advancements allow for cultivar differentiation, production quality prediction, and non-invasive raw material mapping. Such practices promote sustainable agriculture, reduce labor intensity, foster education and skill development, and enhance rural development, in line with SDG 1: No Poverty and SDG 4: Quality Education. Knowledge exchange and community building are supported, contributing to SDG 10: Reduced Inequalities.

Environmental: Resource use efficiency is improved, enabling precise decisions that reduce waste and minimize the environmental footprint of agricultural practices, in accordance with SDG 12: Responsible Consumption and Production and SDG 13: Climate Action. Identifying optimal harvest times and resource application leads to lower emissions, energy conservation, and the promotion of agricultural sustainability. These practices support soil health, prevent resource overexploitation, and aid in biodiversity preservation, aligning with SDG 15: Life on Land.

Economic: Optimizing resource management and harvest decisions increases sugarcane productivity, income, and profitability, contributing to SDG 8: Decent Work and Economic Growth. Improving production quality enhances market competitiveness, reduces costs, and reinforces economic sustainability, in line with SDG 9: Industry, Innovation, and Infrastructure. The adoption of advanced technologies attracts investments, encourages innovation, and drives infrastructure development, benefiting rural and urban economies. Strengthening the sugarcane value chain, job creation, and intensified exports promote economic growth, ensuring a stable and prosperous future for producing regions.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: IMAGENS ORBITAIS DE ALTA RESOLUÇÃO E REDES NEURAIS ARTIFICIAIS PARA ESTIMATIVA DO TEOR DE SACAROSE EM CANA-DE-AÇÚCAR

AUTOR: PAULO HENRIQUE CARDOSO

ORIENTADOR: ROVERSON PEREIRA DA SILVA

COORIENTADORA: TATIANA FERNANDA CANATA

COORIENTADORA: SAMIRA LUNS HATUM DE ALMEIDA

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

Prof. Dr. ROVERSON PEREIRA DA SILVA (Participação Virtual)
Departamento de Engenharia e Ciências Exatas / FCAV UNESP Jaboticabal



Documento assinado digitalmente

ROVERSON PEREIRA DA SILVA

Data: 30/07/2024 17:06:22-0300

Verifique em <https://validar.iti.gov.br>

Prof. Dr. MURILO APARECIDO VOLTARELLI (Participação Virtual)
Universidade Federal de São Carlos (UFSCar) / Sorocaba/SP



Documento assinado digitalmente

MURILO APARECIDO VOLTARELLI

Data: 18/08/2024 22:40:00-0300

Verifique em <https://validar.iti.gov.br>

Profa. Dra. FRANCIELE MORLIN CARNEIRO (Participação Virtual)
Universidade Tecnológica Federal do Paraná (UTFPR) / Santa Helena/PR



Documento assinado digitalmente

FRANCIELE MORLIN CARNEIRO

Data: 01/08/2024 09:35:22-0300

Verifique em <https://validar.iti.gov.br>

Jaboticabal, 01 de março de 2024

DADOS CURRILULARES DO AUTOR

PAULO HENRIQUE CARDOSO – nascido em Morro Agudo, São Paulo, no dia 20 de agosto de 1995, filho de Paulo Roberto Cardoso e Maria Aparecida Ramos da Silva Cardoso. cursou o ensino fundamental e médio no Colégio A a Z, finalizando essa etapa de estudos em 2012. Ingressou no ensino superior em 2013, no curso de Engenharia Agrônômica pela Faculdade Dr. Francisco *MAE*da – FAFRAM, no município de Ituverava, São Paulo, obtendo o título de engenheiro agrônomo em 2017. Durante a graduação, foi bolsista da Fundação de Apoio à Pesquisa Agrícola – FUNDAG, sob orientação do Prof. Dr. Hamilton Humberto Ramos, no Instituto Agrônômico – IAC, em Jundiaí, São Paulo. Atuou como instrutor em tecnologia de aplicação de defensivos agrícolas em programa público-privado, fornecendo treinamentos e capacitação para operadores agrícolas, além de colaborar no desenvolvimento e validação de normas nacionais (ABNT) e internacionais (ISO) para inspeção de pulverizadores terrestres em uso no Brasil. Em 2019, ingressou no curso de pós-graduação lato sensu pela Universidade Federal de Viçosa – UFV, onde adquiriu o título de especialista em Proteção de Plantas em 2020. Conduziu atuação profissional em diversas empresas no segmento de tecnologia de aplicação, agricultura de precisão e agricultura digital, participando ativamente na implementação de novas tecnologias e auxiliando produtores rurais no gerenciamento de propriedades agrícolas. Em agosto de 2022, iniciou o curso de mestrado acadêmico em Agronomia (Produção Vegetal), com concentração nas áreas de agricultura digital e máquinas agrícolas, pela Universidade Estadual Paulista – UNESP, na Faculdade de Ciências Agrárias e Veterinárias, campus de Jaboticabal, São Paulo, no Laboratório de Máquinas e Mecanização Agrícola – LAMMA, localizado no Departamento de Engenharia e Ciências Exatas – DECEX, sob orientação do Prof. Dr. Rouverson Pereira da Silva. Durante o curso, realizou pesquisas com ênfase em agricultura digital, utilizando ferramentas de aprendizado de máquina, inteligência artificial e sensoriamento remoto para caracterização do período ideal de colheita para a cultura da cana-de-açúcar. Atualmente, é fundador da empresa Agrocision – Soluções Agrícolas, onde exerce o cargo de diretor e consultor nas áreas de agricultura digital, agricultura de precisão e tecnologia de aplicação de defensivos agrícolas. Em fevereiro de 2024, submeteu-se à banca examinadora para obtenção do título de mestre em Agronomia no programa de Produção Vegetal.

Escolhi a botina porque minha vaidade está abaixo da fome das pessoas. Porque o trabalho árduo não me assusta. Escolhi estar no campo para garantir o conforto dos que moram nos grandes centros. Escolhi aumentar a produtividade, em prol da natureza e extinção da fome no mundo. Escolhi acima de tudo, a simplicidade, a sabedoria e a resignação do produtor rural.

(Autor desconhecido)

Aos meus pais Paulo Roberto Cardoso, Maria Aparecida Ramos da Silva Cardoso e meu irmão Pedro Henrique Cardoso.

Que em cada desafio estiveram ao meu lado, oferecendo amor, encorajamento e a certeza de que eu poderia alcançar qualquer objetivo.

Esta conquista é fruto deste nosso amor compartilhado.

A vocês, **DEDICO!**

AGRADECIMENTOS

- A Deus, por me transformar em um ser humano melhor a cada dia, por me dar força e perseverança na construção deste trabalho, além de me presentear com a oportunidade de ter meus pais e meu irmão presentes para que pudessem presenciar esta conquista;
- Aos meus pais, Paulo Roberto Cardoso e Maria Aparecida Ramos da Silva Cardoso, que nunca deixaram de acreditar no meu potencial. Pelos sacrifícios feitos por mim e meu irmão para que pudéssemos ter uma excelente educação, ética e nos ensinar os verdadeiros valores da vida. Nos momentos de dúvidas e desafios, foi o exemplo de força e dedicação de vocês que me guiou. O apoio emocional e a crença inabalável que depositaram em mim foram as luzes que iluminaram os caminhos mais escuros desta jornada;
- Ao meu irmão, Pedro Henrique Cardoso, quero expressar minha profunda gratidão e apreço por tudo que você significa para mim. Ao longo da jornada da vida, você desempenha o papel não só de irmão, mas de grande amigo, confidente e fiel escudeiro;
- À Faculdade Dr. Francisco Maeda – FAFRAM, por todo embasamento técnico e científico e pelo auxílio em todos os momentos, permitindo-me obter o título de Engenheiro Agrônomo, profissão que exerço com muito amor e dedicação em prol da Agricultura;
- À Universidade Federal de Viçosa - UFV, Campus de Viçosa – MG, em especial ao Programa de Pós-Graduação em Agronomia (Proteção de Plantas), pela oportunidade de obtenção do título de Especialista;
- À Universidade Estadual Paulista “Júlio de Mesquita Filho” - UNESP, Faculdade de Ciências Agrárias e Veterinárias, Campus de Jaboticabal - SP, em especial ao Programa de Pós-Graduação em Agronomia (Produção Vegetal), pela oportunidade de obtenção do título de Mestre;

- À Faculdade de Tecnologia Nilo de Stéfani, por conceder seu laboratório e os equipamentos necessários para que as análises laboratoriais apresentadas neste trabalho fossem possíveis;
- Ao meu orientador, Professor Dr. Rouverson Pereira da Silva, a quem tenho grande admiração e respeito, gostaria de expressar minha mais sincera gratidão. Suas valiosas sugestões, conselhos e encorajamento foram fundamentais para que eu alcançasse os objetivos que estabeleci para mim. Sua influência como mentor e educador terá um impacto permanente em minha vida;
- As minhas coorientadoras, Professora Dra. Tatiana Fernanda Canata e Dra. Samira Luns Hatum de Almeida. A dedicação, conhecimento e paixão que vocês demonstram não apenas enriqueceram minha experiência acadêmica, mas também inspiraram e motivaram minha jornada e desenvolvimento no programa de Produção Vegetal;
- Aos amigos do Laboratório de Máquinas e Mecanização Agrícola - LAMMA, que desempenharam um papel significativo no meu crescimento na área acadêmica. Deixo aqui meu reconhecimento especial a Jarlyson Brunno Costa Souza, Vinicius dos Santos Carreira, Armando Lopes de Brito Filho, Thiago Caio Moura Oliveira, Igor Cristian de Oliveira Vieira e Breno dos Santos Silva por estarem sempre lá, oferecendo um ombro amigo, um ouvido atento e momentos de alegria e descontração. Vocês enriqueceram minha vida com risadas, conversas significativas e lembranças inesquecíveis.

A vocês, AGRADEÇO!

SUMÁRIO

RESUMO	16
ABSTRACT	17
1. INTRODUÇÃO	18
2. REVISÃO DE LITERATURA.....	19
2.1. Cana-de-Açúcar: Panorama Global e Nacional	19
2.2. Agricultura 4.0: A Revolução Digital no Campo	22
2.2.1. O Sensoriamento Remoto e Suas Aplicações na Agricultura.....	24
2.2.2. O Impacto das Redes Neurais Artificiais na Agricultura	27
2.2.3. Adoção de Tecnologias Digitais nas Propriedades Rurais	29
3. MATERIAL E MÉTODOS.....	30
3.1. Caracterização das Áreas Experimentais	30
3.2. Metodologia de Extração do Caldo e Análise do Brix (°)	31
3.3. Metodologia de Análise da Pureza (%).....	32
3.4. Cálculo e Aquisição de Graus-Dias Acumulados (aGDD)	34
3.5. Processo de Aquisição de Dados Multiespectrais	35
3.6. Cálculo de Índices de Vegetação e Extração dos Dados Multiespectrais	36
3.7. Análise Descritiva dos Dados Coletados	38
3.8. Estratégias de Seleção e Estruturação de Variáveis para Estimativa	38
3.9. Aplicação de Redes Neurais Artificiais	39
3.9.1. Modelo de Função de Base Radial (RBF)	39
3.9.2. Modelo Perceptron de Multicamadas (MLP).....	41
3.10. Avaliação do Desempenho dos Modelos de Redes Neurais Artificiais	43
3.10.1. Fase de Treinamento e Teste	43
3.10.2. Transferência e Generalização dos Modelos.....	45
4. RESULTADOS.....	46
4.1. Análise estatística	46

4.2.	Seleção das Variáveis de Entrada.....	47
4.3.	Desempenho dos Modelos Durante os Testes	50
4.3.1.	Comparação entre modelos RBF e MLP na estimativa de Brix (°)	50
4.3.2.	Comparação entre modelos RBF e MLP na estimativa de Pureza (%)	53
4.4.	Capacidade de Transferência e Generalização dos Modelos.....	55
5.	DISCUSSÃO.....	57
5.1.	Dinâmica Temporal de Brix (°) e Pureza (%) na Colheita Inteligente.....	57
5.2.	Relação entre Características Espectrais e Rendimento Qualitativo.....	58
5.3.	Redes Neurais Artificiais na Estimativa de Sacarose: Brix (°) e Pureza (%).....	59
5.4.	Complexidade de Modelos e Aplicabilidade Prática: Uma visão Holística.....	60
6.	CONCLUSÕES.....	61
7.	REFERÊNCIAS	62

LISTA DE FIGURAS

Figura 1 - Área colhida e produção mundial de cana-de-açúcar.....	19
Figura 2 – Principais países produtores de cana-de-açúcar no mundo	20
Figura 3 – Distribuição da produção de cana-de-açúcar por continente	21
Figura 4 - Commodities agrícolas mais produzidas globalmente	22
Figura 5 - Espectro eletromagnético enfatizando os comprimentos de onda da luz visível	25
Figura 6 - Comportamento espectral típico da vegetação agrícola	26
Figura 7 – Localização e pontos amostrais da Área 1	30
Figura 8 – Localização e pontos amostrais da Área 2	31
Figura 9 - Refratômetro digital utilizado para análise do Brix (°).....	32
Figura 10 - Polarímetro digital utilizado para determinar a pol do caldo	33
Figura 11 – Estrutura da Rede Neural RBF (3:3-20-1:1) para estimar Brix (°) utilizando dados multiespectrais.	40
Figura 12 – Estrutura da Rede Neural MLP (3:3-20-12-1:1) para estimar Brix (°) utilizando dados multiespectrais.	41
Figura 13 - Variação temporal de Brix (°) e Pureza (%) em duas áreas de cultivo, medidos em Dias Após o Plantio (DAP).	46
Figura 14 – Aplicação do RFE na estimativa de Brix (°) e Pureza (%), destacando as variáveis selecionadas em diferentes categorias: Bandas Multiespectrais, Índices de Vegetação, e suas combinações.	48
Figura 15 - Desempenho do teste na estimativa de Brix (°) utilizando as redes neurais artificiais RBF e MLP.....	52
Figura 16 - Desempenho do teste na estimativa de Pureza (%) utilizando as redes neurais artificiais RBF e MLP.....	54
Figura 17 – Transferência e generalização dos modelos RBF e MLP na estimativa de Brix (°) e Pureza (%) em cana-de-açúcar através do Erro Médio Absoluto (<i>MAE</i>).....	55

LISTA DE TABELAS

Tabela 1 - Características técnicas do sensor Dove-R PlanetScope	35
Tabela 2 - Apresentação dos índices de vegetação e suas respectivas fórmulas	37
Tabela 3 - Variáveis de importância na estimativa de Brix (°) e Pureza (%) utilizando dados de imagens orbitais.	50

IMAGENS ORBITAIS DE ALTA RESOLUÇÃO E REDES NEURAS ARTIFICIAIS PARA ESTIMATIVA DO TEOR DE SACAROSE EM CANA-DE-AÇÚCAR

RESUMO

A integração de redes neurais artificiais (RNA's) junto ao sensoriamento remoto (SR) na avaliação qualitativa da cana-de-açúcar pode oferecer indicadores confiáveis sobre a qualidade da matéria-prima. Métodos convencionais para análise da sacarose demandam intensa mão de obra e altos custos laboratoriais, especialmente em larga escala. Este estudo propõe estimar parâmetros biométricos da cana-de-açúcar, Brix (°) e Pureza (%), de forma não destrutiva por meio de redes neurais artificiais aplicadas a dados de SR. Para isso, foram adquiridos dados biométricos *in situ* e dados multiespectrais de sensores orbitais. As coletas foram feitas em duas áreas comerciais de cana-de-açúcar: a primeira com 240 pontos amostrais para treinamento e teste, e a segunda com 90 pontos amostrais para a transferência e a generalização dos modelos. As imagens multiespectrais da plataforma PlanetScope CubeSat, com oito bandas espectrais, foram usadas para calcular 16 índices de vegetação (IV's). Dois modelos de RNA's foram empregados para estimar Brix (°) e Pureza (%), Radial Basis Function (RBF) e Multilayer Perceptron (MLP), utilizando como estradas, além das bandas e dos IV's, os graus-dias acumulados (aGDD). Após a seleção de variáveis de entrada, cinco bandas e cinco IV's mostraram-se promissores na estimativa do teor de sacarose por meio do Brix (°) e da Pureza (%). As redes RBF e MLP demonstraram acurácia de 90% e 91% para Brix (°) e 89% e 93% para Pureza (%), respectivamente. Este estudo contribui para o desenvolvimento de métodos de colheita inteligente, promovendo a otimização da qualidade da matéria-prima, proporcionando aos produtores rurais e à indústria subsídios técnicos para a tomada de decisões informadas para definição do período ideal de colheita da cana-de-açúcar.

Palavras-chaves: Saccharum spp; agricultura digital, colheita inteligente; sensoriamento remoto.

HIGH-RESOLUTION ORBITAL IMAGES AND ARTIFICIAL NEURAL NETWORKS FOR PREDICTING SUCROSE CONTENT IN SUGARCANE

ABSTRACT

The integration of artificial neural networks (ANNs) with remote sensing (RS) in the qualitative assessment of sugarcane can offer reliable indicators of raw material quality. Conventional methods for sucrose analysis require intensive labor and high laboratory costs, especially on a large scale. This study proposes to estimate biometric parameters of sugarcane, Brix (°) and Purity (%), non-destructively using artificial neural networks applied to RS data. To achieve this, in situ biometric data and multispectral data from orbital sensors were acquired. Data collection was conducted in two commercial sugarcane areas: the first with 240 sampling points for training and testing, and the second with 90 sampling points for model transfer and generalization. Multispectral images from the PlanetScope CubeSat platform, with eight spectral bands, were used to calculate 16 vegetation indices (VIs). Two ANN models were employed to estimate Brix (°) and Purity (%): Radial Basis Function (RBF) and Multilayer Perceptron (MLP), using, in addition to the bands and VIs, accumulated Growing Degree Days (aGDD) as inputs. After selecting input variables, five bands and five VIs proved to be promising for estimating sucrose content through Brix (°) and Purity (%). The RBF and MLP networks demonstrated accuracies of 90% and 91% for Brix (°) and 89% and 93% for Purity (%), respectively. This study contributes to the development of intelligent harvesting methods, promoting the optimization of raw material quality, providing rural producers and the industry with technical support for informed decision-making regarding the optimal sugarcane harvest period.

Keywords: *Saccharum* spp; digital agriculture; smart harvesting; remote sensing.

1. INTRODUÇÃO

A cana-de-açúcar, cultura de significativa importância econômica global, enfrenta o desafio crítico de otimizar sua colheita. O acúmulo de açúcares, essencial para a resistência da planta a condições adversas e para a produtividade industrial, ocorre ao longo do desenvolvimento do cultivo, culminando na fase de maturação. Esta fase, influenciada pela interação entre o genótipo da planta e o ambiente, é crucial para a eficiência econômica da colheita e a qualidade da matéria prima. A determinação do período ideal de colheita é complexa, pois a ultrapassagem do ponto de maturação ideal ou a entrada na fase de floração impacta negativamente a fotossíntese e os níveis de sacarose (Khan et al., 2022; Misra et al., 2023).

Historicamente, a avaliação da maturidade e concentração de sacarose da cana-de-açúcar tem sido feita por meio de métodos convencionais, como a coleta de amostras a campo e análises da concentração de sólidos solúveis (Brix °) e a proporção de sacarose em relação aos sólidos solúveis totais (Pureza %) em laboratório. Embora eficazes, estes métodos apresentam desvantagens, como custos elevados, tempo de análise, além de ser um método destrutivo. Paralelamente, tecnologias emergentes, como o sensoriamento remoto, têm se mostrado promissoras para estas avaliações. A reflectância do dossel, fornece informações valiosas sobre a conversão da energia radiante em sacarose durante a maturação (Chea et al., 2020a).

A aplicação de algoritmos de aprendizado de máquina (ML) e o uso de imagens de aeronaves remotamente pilotadas (ARPs) mostram-se eficientes para a estimativa de parâmetros como o teor de Brix (°) e Pureza (%), contudo, ainda enfrentam limitações, principalmente em termos de escalabilidade global (Chea et al., 2020b; de Oliveira et al., 2022; Barbosa et al., 2023). Estes métodos dependem de equipamentos e pessoas especializadas, além de processos complexos de correção e análise das imagens, e, geralmente não apresentam uma etapa de transferência e generalização destes modelos em outras áreas. Em contraste, estudos recentes apontam para a eficácia do uso de imagens orbitais, que, combinadas com índices biométricos e séries temporais, mostram-se mais precisas na estimativa da biomassa e do teor de sacarose (Bégué et al., 2010).

Este trabalho parte do pressuposto de que o uso de sensoriamento remoto aliado a inteligência artificial pode permitir avanços na colheita mecanizada de cana-de-açúcar, por meio da determinação do período ideal de colheita. Desta forma, este trabalho visa desenvolver um modelo para a estimação de parâmetros biométricos da cana-de-açúcar, integrando imagens orbitais de alta resolução espacial e redes neurais artificiais. Os

objetivos específicos deste trabalho foram: (I) criar um modelo de redes neurais eficaz para estimar os parâmetros de Brix (°) e Pureza (%) da cana-de-açúcar, e (II) transferir e generalizar os modelos encontrados para determinar o período ideal de colheita, a fim de promover a aplicabilidade da colheita inteligente em cana-de-açúcar.

2. REVISÃO DE LITERATURA

2.1. Cana-de-Açúcar: Panorama Global e Nacional

A cana-de-açúcar (*Saccharum spp.*), uma planta de grande porte e longo ciclo, é cultivada em aproximadamente 107 nações situadas nas regiões tropicais e subtropicais do globo terrestre. Esta cultura é responsável por fornecer mundialmente cerca de 80% do açúcar e 35% do etanol (FAOSTAT, 2022). Caracterizada pelo seu rápido desenvolvimento e alta produção de biomassa, sob cultivo semi-perene, a cana-de-açúcar desempenha um papel fundamental no setor agrícola e na bioeconomia (Thorburn et al., 2017).

A produção global de cana-de-açúcar é de aproximadamente 1,92 bilhões de toneladas anuais, com Brasil e Índia respondendo por 68.47% desta produção (FAOSTAT, 2022). Com clima tropical e vasta extensão de terras adequadas para o cultivo, estes países possuem tecnologia avançada na produção de açúcar e etanol, o que contribui para o alto desempenho. Conforme ilustrado na Figura 1, a evolução da área colhida e da produção mundial de cana-de-açúcar de 1995 a 2022 mostra um crescimento contínuo, refletindo a crescente demanda por esta commodity.

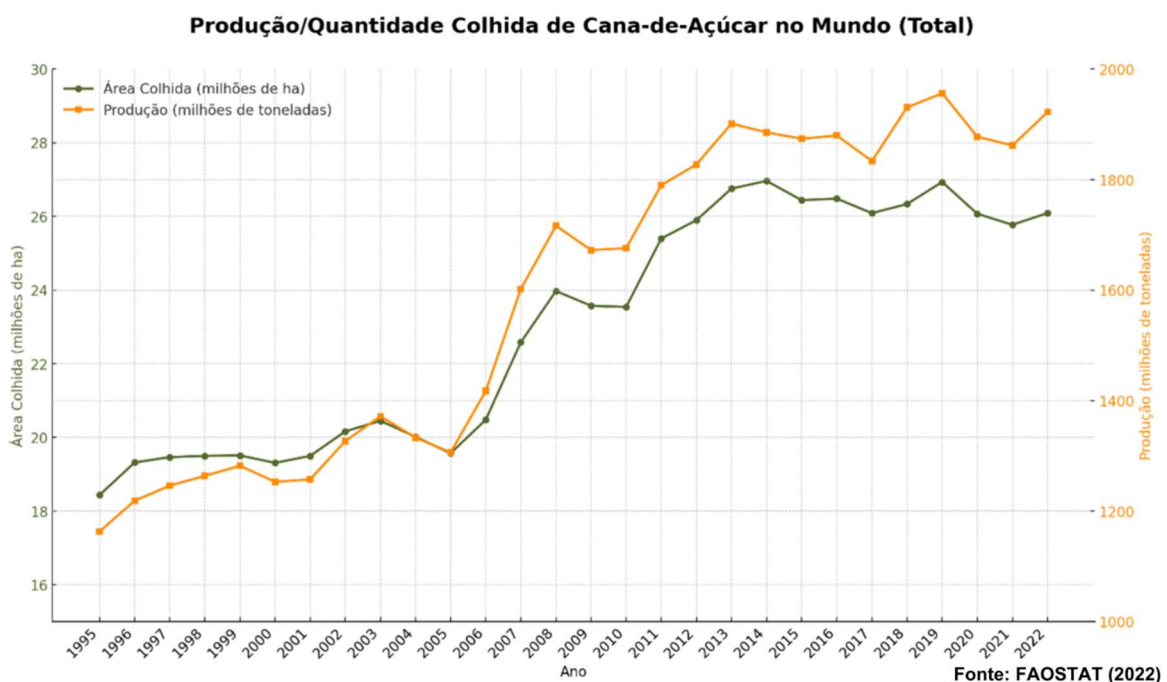


Figura 1 - Área colhida e produção mundial de cana-de-açúcar

A importância econômica da cana-de-açúcar para o Brasil é incontestável, seja na produção de açúcar ou exportação de biocombustível renovável. Durante a safra 2021/2022, a produção foi de 585.179,40 milhões de toneladas em uma área total cultivada de 8,30 milhões de hectares (CONAB, 2023). Nas regiões Norte e Nordeste, a produção ocorre de setembro a março, enquanto na região Centro-Sul, a produção acontece de abril a novembro (Epifanio, 2023).

A produção de cana-de-açúcar no Brasil enfrentou desafios nas últimas safras devido a adversidades climáticas. No entanto, a produtividade atual mostra sinais de recuperação, com estimativas de crescimento para a safra 2023/2024, prevendo uma produção de 652,90 milhões de toneladas (aumento de 11,58% em relação à safra 2021/2022). A área destinada para a colheita será de 8.288,30 milhões de hectares, com produtividade média de 78.779,00 kg/ha (CONAB, 2023). A Figura 2 apresenta a produção dos 10 maiores produtores mundiais de cana-de-açúcar de 1995 a 2022, destacando a dominância do Brasil.

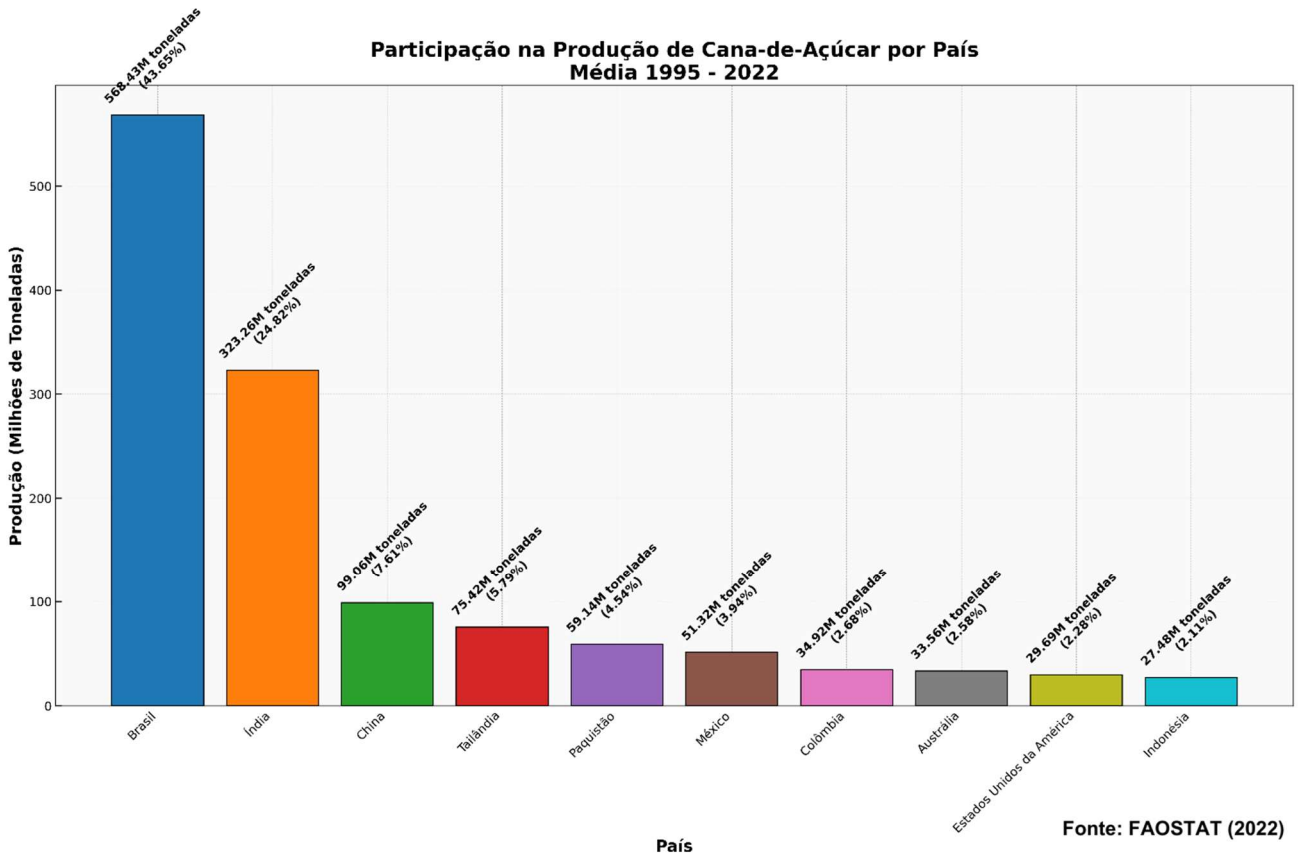


Figura 2 – Principais países produtores de cana-de-açúcar no mundo

Em média, campos de cana-de-açúcar produzem cerca de 70,36 toneladas por hectare, com uma composição aproximada de 10-20% de sacarose e 10-15% de fibra,

variando conforme a tecnologia e o manejo utilizados (Hithamani et al., 2018; Sreedevi, Jayachandran and Rao, 2018; Yang, Ching and Chuah, 2019).

O processo de desenvolvimento da cana-de-açúcar abrange várias etapas, começando pela brotação, passando pelo perfilhamento, crescimento e, finalmente, a maturação. As plantas se caracterizam por caules longos e suculentos, ricos em sacarose. O sistema radicular inclui raízes adventícias aéreas e raízes permanentes. A variação morfológica do caule entre diferentes genótipos é considerável e crucial para a identificação das cultivares da planta (Martin, Abbott and Hughes, 1961).

A distribuição da produção de cana-de-açúcar é predominantemente concentrada nas Américas e na Ásia. A Figura 3 apresenta a participação das diferentes regiões na produção global entre 1995 e 2022.

**Participação na Produção de Cana-de-Açúcar por Continente
Média 1995 - 2022**

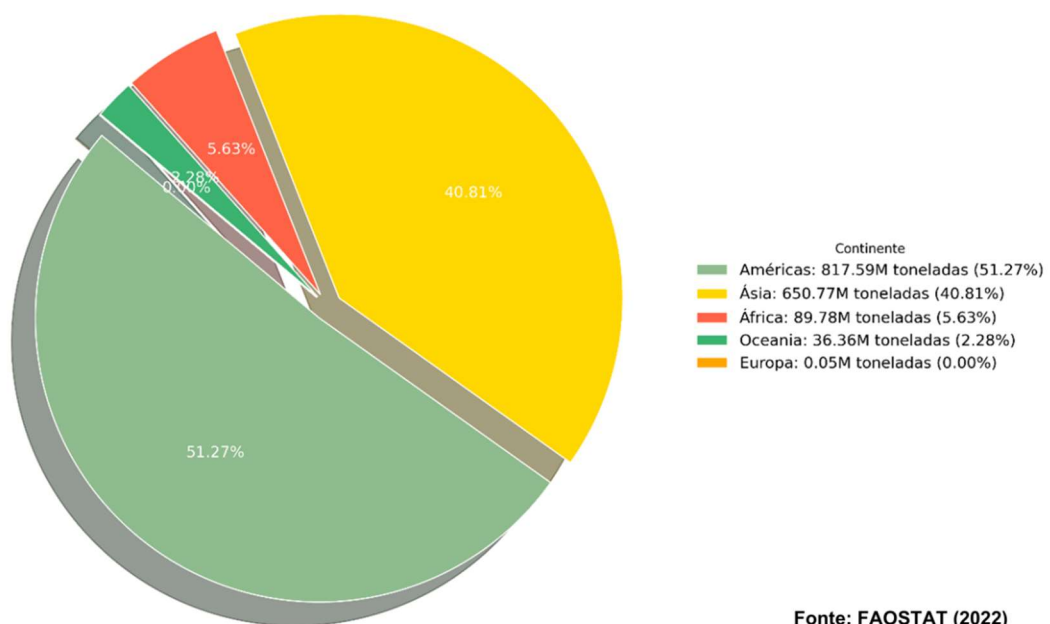


Figura 3 – Distribuição da produção de cana-de-açúcar por continente

A cana-de-açúcar destaca-se como a commodities agrícola mais produzida no mundo, superando outras culturas importantes como milho, arroz e trigo. A Figura 4 ilustra a produção das principais commodities agrícolas entre 1995 e 2022.

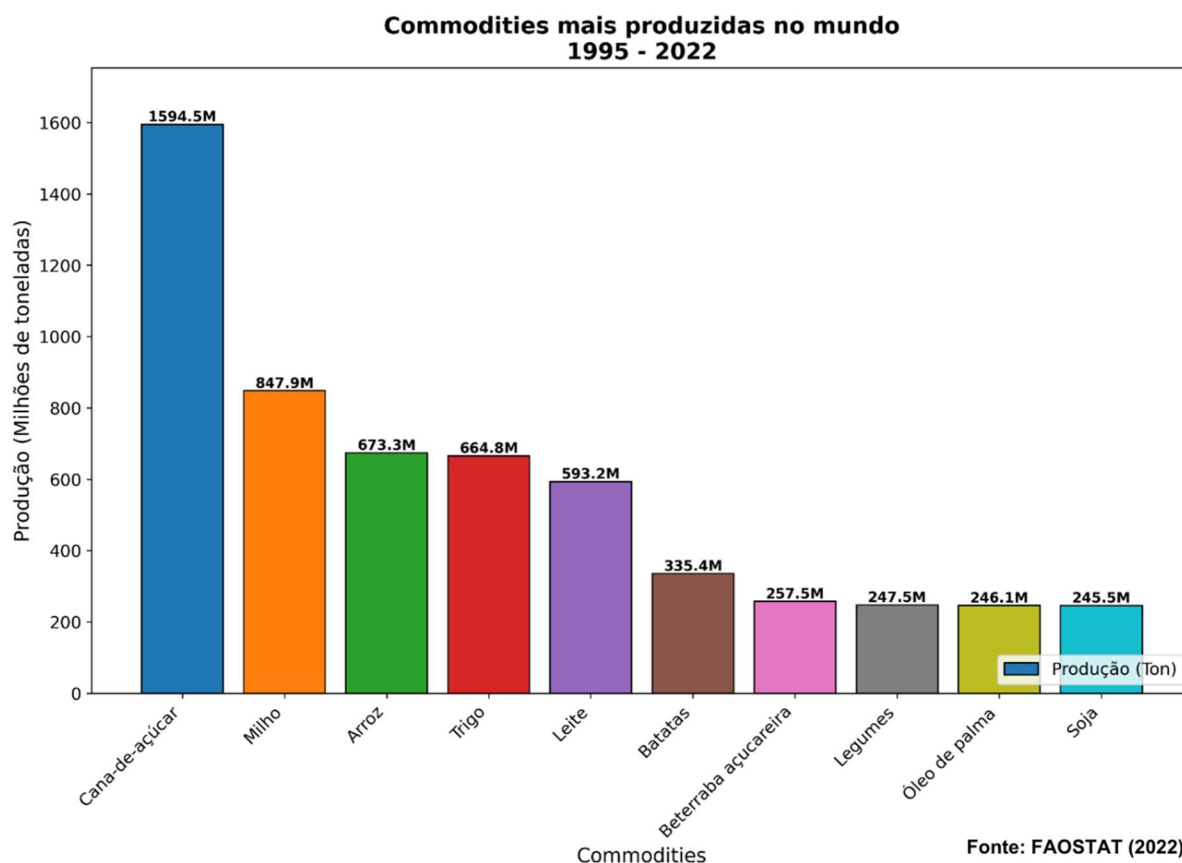


Figura 4 - Commodities agrícolas mais produzidas globalmente

A análise dos dados de produção de cana-de-açúcar revela a importância dessa cultura agrícola no cenário global, com o Brasil desempenhando um papel de liderança. A concentração da produção nas Américas e na Ásia evidencia as condições favoráveis e a relevância econômica da cana-de-açúcar nessas regiões. O crescimento contínuo da área colhida e da produção reflete a demanda crescente por esta commodity, que é fundamental tanto para a indústria alimentícia quanto para a produção de biocombustíveis.

2.2. Agricultura 4.0: A Revolução Digital no Campo

Na evolução da agricultura, a transição para a Agricultura 4.0 representa um marco significativo, caracterizado por um avanço sem precedentes na integração da tecnologia digital na cadeia produtiva agropecuária. Este panorama, frequentemente descrito por meio de termos como "Agricultura Digital" e "Agricultura Inteligente", reflete uma expansão direta dos princípios da Agricultura de Precisão, incorporando o uso intensivo de dados e tecnologias avançadas para otimizar as práticas agrícolas (Hermonn, 2008; Sponchioni et al., 2020).

A Agricultura 4.0 distingue-se pela sua capacidade de coletar, integrar e analisar massivas quantidades de dados agrícolas de uma forma automatizada e em tempo real. Esta integração de dados é facilitada pelo emprego de uma ampla gama de tecnologias, incluindo sensores em campo, telemetria, sistemas globais de navegação por satélite (GNSS), e geotecnologias de sensoriamento remoto. Além disso, a utilização de plataformas de computação em nuvem, big data, inteligência artificial (IA), aprendizado de máquina (ML), e a Internet das Coisas (IoT) permite uma análise sofisticada, além de proporcionar tomadas de decisões baseadas em evidências para aprimorar a produtividade e a sustentabilidade nos campos agrícolas (Hermonn, 2008).

Essencialmente, a Agricultura 4.0 visa não apenas melhorar a eficiência e a produtividade, mas também promover práticas agrícolas sustentáveis que minimizem o impacto ambiental, por meio da otimização do uso de recursos naturais e da redução da pegada ecológica. A interoperabilidade entre diferentes sistemas tecnológicos é crucial para assegurar uma análise de dados abrangentes e acurados, facilitando uma gestão agrícola holística que considera todas as variáveis do ecossistema produtivo (CEMA, 2017).

Entretanto, a implementação efetiva da Agricultura 4.0 enfrenta desafios significativos. A superação de obstáculos socioeconômicos é fundamental para garantir a inclusão e beneficiar um espectro mais amplo de comunidades agrícolas. Adicionalmente, questões relacionadas à segurança, proteção das informações coletadas, e à privacidade dos dados agrícolas necessitam de mais atenção, assim como a importância de desenvolver uma infraestrutura de conectividade robusta a fim de oferecer programas de capacitação e educação adaptados às necessidades dos agricultores e trabalhadores rurais (Shang et al., 2021a; Cherubin et al., 2022).

O estudo desta nova forma de fazer agricultura abre novas perspectivas para a adoção de práticas de colheita inteligente e linhas de produção altamente eficientes, promovendo não apenas a otimização da qualidade da matéria prima para a indústria, mas também melhorias significativas no rendimento econômico para produtores rurais. A implementação bem sucedida desta nova era agrícola requer uma abordagem integrada e colaborativa, envolvendo o reforço das infraestruturas, a promoção do desenvolvimento rural, e um compromisso contínuo com a inovação e a educação (Hermonn, 2008; Bolfe et al., 2020).

2.2.1. O Sensoriamento Remoto e Suas Aplicações na Agricultura

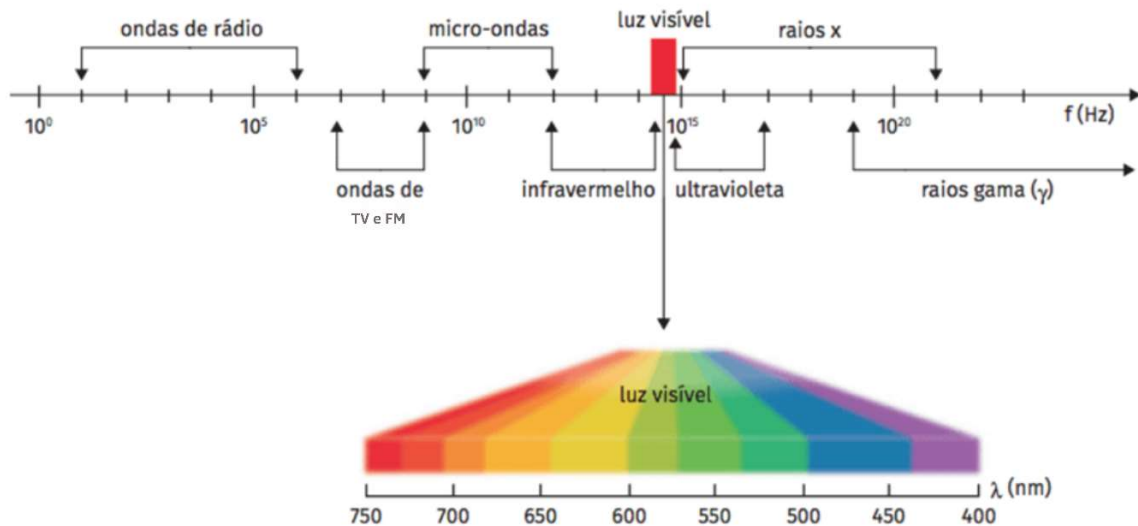
O sensoriamento remoto (SR) é uma tecnologia que utiliza a radiação eletromagnética (REM) para coletar informações sobre objetos ou áreas à distância. Baseia-se nos princípios físicos que regem a interação da REM com diversos materiais e como essa interação é capturada pelos sensores. Esses sensores podem ser imageadores, que geram imagens visuais, ou não imageadores, que coletam dados sem formar imagens visuais. Uma compreensão detalhada deste processo é suportada pela espectrorradiometria, um método que permite quantificar a energia da radiação eletromagnética em diferentes pontos do espectro, essencial para interpretar as imagens produzidas pelo sensoriamento remoto (Meneses, Almeida e Baptista, 2019).

O SR é composto por uma fonte de REM, uma plataforma equipada com sensores e os objetos a serem monitorados. A premissa central do sensoriamento remoto reside na habilidade de deduzir informações a partir da análise dos dados refletidos ou emitidos pelos objetos, os quais exibem propriedades de reflectância distintas em várias faixas do espectro eletromagnético (Formaggio, Formaggio e Sanches, 2017). Diferentes materiais interagem com a radiação eletromagnética de maneiras específicas que são intrínsecas às suas propriedades, podendo ser refletida, transmitida ou absorvida (Meneses, Almeida e Baptista, 2019).

Tomemos, por exemplo, o fenômeno pelo qual folhas de plantas manifestam-se verdes ao olho humano: isso ocorre porque a clorofila, um pigmento presente nas folhas, absorve predominantemente radiação nas faixas do azul e vermelho na região visível do espectro, enquanto reflete mais intensamente a radiação na faixa verde (Ferreira, 2008).

A REM que se encontra com maior disponibilidade na superfície terrestre é a proveniente do sol, que ilumina a superfície da Terra. O Sol emite luz não apenas no comprimento de onda visível, mas também em comprimentos de onda menores, como o ultravioleta, e maiores, como o infravermelho próximo e de ondas curtas. A Terra, sendo também uma fonte de radiação eletromagnética, emite radiação infravermelha termal, com um comprimento de onda significativamente maior do que a luz solar. Além disso, fontes artificiais construídas pelo ser humano podem emitir radiação em comprimentos de onda ainda maiores, como as micro-ondas (Meneses, Almeida and Baptista, 2019).

Essas denominações, entre outras, foram dadas em decorrência de uma proposição arbitrária de divisão do espectro eletromagnético, o qual é apresentado na Figura 5.



Fonte: Adaptado de Meneses, Almeida e Baptista (2019)

Figura 5 - Espectro eletromagnético enfatizando os comprimentos de onda da luz visível

Os sensores utilizados no SR podem ser classificados em quatro tipos com base nas suas características espectrais, espaciais, temporais e radiométricas:

- Sensores espectrais: Capturam dados em múltiplas bandas do espectro eletromagnético, permitindo a análise da composição dos objetos observados. Fornecendo uma visão detalhada das propriedades espectrais dos materiais;

- Sensores espaciais: Referem-se à resolução espacial, que determina o tamanho do menor objeto que pode ser detectado. Sensores de alta resolução espacial são essenciais para mapear detalhes finos nas áreas de cultivo;

- Sensores temporais: Relacionam-se à frequência com que os dados são coletados para a mesma área, conhecida como resolução temporal. Sensores com alta resolução temporal são essenciais para monitorar mudanças rápidas em fenômenos ambientais e agrícolas;

- Sensores radiométricos: Medem a quantidade de energia refletida ou emitida pelos objetos. A resolução radiométrica determina a sensibilidade do sensor em detectar variações sutis na energia refletida ou emitida, sendo crucial para análises precisas de fenômenos naturais.

Os gráficos de reflectância espectral, representações visuais das propriedades de reflexão da luz por parte das plantas em diferentes comprimentos de onda do espectro eletromagnético, são amplamente utilizados para monitorar a saúde de cultivos agrícolas além de estimar vários parâmetros biofísicos da vegetação. O comportamento espectral da

vegetação pode ser caracterizado por regiões distintas no espectro eletromagnético, conforme ilustrado na Figura 6.

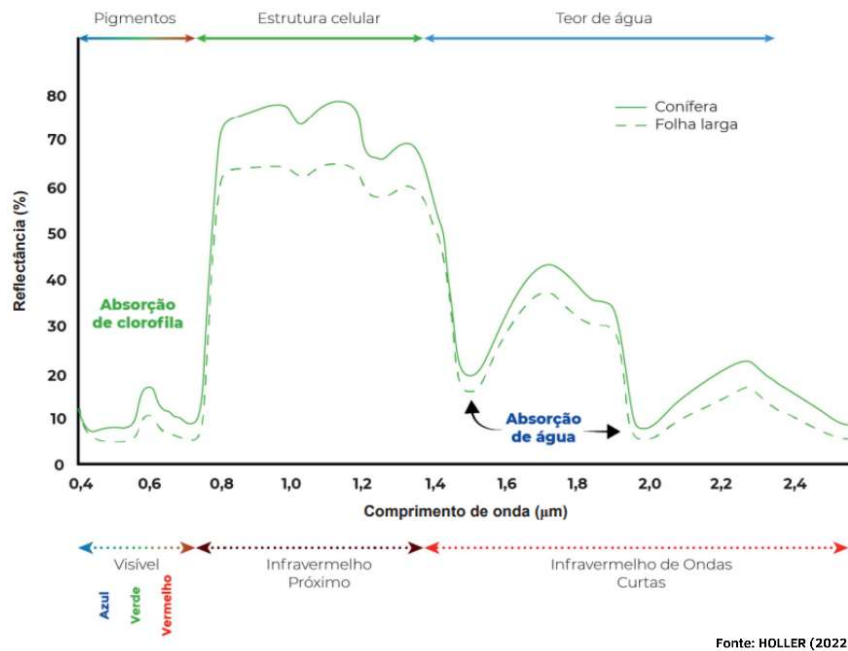


Figura 6 - Comportamento espectral típico da vegetação agrícola

Na região do Visível (400-700 nm), a clorofila absorve fortemente nas bandas azul (cerca de 450-500 nm) e vermelha (cerca de 650-700 nm) para o processo de fotossíntese, resultando em baixa reflectância nessas faixas. A faixa verde (cerca de 500-600 nm) é menos absorvida e mais refletida, conferindo à vegetação saudável sua cor verde característica. No Infravermelho Próximo (NIR, 700-1300 nm), a reflectância aumenta drasticamente devido à estrutura celular das folhas, que dispersa a luz dentro do tecido foliar; a vegetação saudável reflete grande parte da energia incidente no NIR, resultando em um pico acentuado de reflectância, um indicador chave da vitalidade da vegetação. No Infravermelho Médio (1300-2500 nm), a reflectância diminui novamente devido à absorção de água e à presença de outros compostos orgânicos que absorvem radiação nesta faixa (Ferreira, 2008).

O sensoriamento remoto tornou-se uma tecnologia importante com uma variedade de aplicações voltadas ao monitoramento de culturas. Essas aplicações incluem a estimativa de produtividade (Tedesco et al., 2021), detecção de doenças (Abd El-Ghany, Abd El-Aziz and Marei, 2020), bem como monitoramento do crescimento das culturas (Bégué et al., 2010) e a avaliação de aspectos qualitativos relacionados ao teor de sacarose em cana-de-açúcar (Barbosa et al., 2023).

Os padrões distintos de reflectância e absorção ajudam os pesquisadores a distinguir entre diferentes tipos de cobertura vegetal, a avaliar as condições da vegetação e a monitorar mudanças ao longo do tempo, o que é vital para a gestão agrícola, a modelagem do carbono e o estudo das mudanças climáticas.

A análise das fontes revela o sucesso potencial desta geotecnologia em monitoramento de cultivos, detecção de doenças, bem como estimativa de produtividade e índices qualitativos. No entanto, desafios técnicos permanecem, como a calibração de dados de diferentes fontes e a necessidade de resoluções espaciais mais finas. Futuras inovações, como o uso de inteligência artificial para análise de dados, podem superar essas limitações e transformar a agricultura de precisão.

O sensoriamento remoto desempenha um papel crucial na agricultura moderna, melhorando a eficiência, produtividade e sustentabilidade das práticas agrícolas. Ao utilizar as oportunidades oferecidas pela detecção remota, a agricultura poderá tornar-se mais produtiva e sustentável na sua resposta aos desafios futuros.

2.2.2. O Impacto das Redes Neurais Artificiais na Agricultura

No domínio da agricultura, o advento das Redes Neurais Artificiais (RNA's) inaugurou uma nova era de avanço tecnológico, inspiradas na funcionalidade do cérebro humano, emulando funções complexas como geração de padrões, cognição, aprendizado e tomada de decisão (Liakos et al., 2018). Estas redes, por meio de sua capacidade de aprender com dados, estão impactando profundamente as práticas agrícolas, oferecendo soluções inovadoras e eficientes.

As RNA's transformam as metodologias agrícolas ao oferecerem ferramentas analíticas sofisticadas que otimizam os processos de decisão. No campo da agricultura, é amplamente utilizada na resolução de várias tarefas de classificação, estimativa e previsão, podendo substituir os métodos clássicos de modelagem, no qual o espectro da sua aplicabilidade é muito vasto (Kujawa and Niedbała, 2021).

Sua aplicação abrange diversas áreas da agricultura, desde a agricultura de precisão até o manejo de doenças, demonstrando sua versatilidade e eficácia que serão apresentadas abaixo:

- Detecção de doenças e pragas: Uma das aplicações mais críticas das RNA's na agricultura está na detecção e classificação precoce de doenças e pragas em plantas. Susanti et al. (2023) demonstraram que RNA's profundas exibem notável precisão na identificação de doenças e pragas em folhas. Essa capacidade não apenas salva as culturas

de possíveis perdas, mas também reduz a dependência de pesticidas químicos, caminhando para práticas agrícolas mais sustentáveis.

- Estimativa de rendimento de culturas: Prever o rendimento das culturas com alta precisão é vital para o planejamento e garantia da segurança alimentar. Benos et al. (2021) destacaram a eficiência das RNA's no manejo de culturas, especialmente na estimativa de rendimento, sublinhando seu potencial para analisar um vasto banco de dados oriundos de imagens orbitais para prever com precisão a produtividade agrícola. Essa aplicação das RNA's auxilia os agricultores a tomar decisões informadas sobre o manejo do cultivo, resultando na otimização da produtividade agrícola.

- Controle de qualidade e classificação: A indústria agrícola também se beneficia das RNA's em processos de controle de qualidade e classificação. Naranjo-Torres et al. (2020) revisaram a aplicação de Redes Neurais Convolucionais (CNNs), uma classe de aprendizado profundo, no processamento de imagens de frutas para classificação, controle de qualidade e tarefas de detecção. Essas tecnologias possibilitam sistemas de classificação automatizados que reduzem significativamente os custos de mão de obra e melhoram a eficiência e precisão da avaliação de qualidade nos produtos agrícolas.

- Gestão de solo e água: As RNA's contribuem para a gestão sustentável do solo e da água, prevendo as necessidades de irrigação e otimizando o uso da água, através da análise de dados históricos e em tempo real, como condições meteorológicas, tipo de solo e níveis de umidade, permitindo uma programação mais eficiente e precisa da irrigação. Esta aplicação é crucial em regiões enfrentando escassez áridas e semiáridas, garantindo aplicação hídrica adequada conforme as necessidades do cultivo (Maier and Dandy, 2000).

Embora as RNA's ofereçam potencial transformador para a agricultura, desafios na disponibilidade de dados, interpretação de modelos e a necessidade de soluções personalizadas persistem. Pesquisas futuras estão sendo feitas para superar estes obstáculos, desenvolvendo modelos que possam lidar com ambientes agrícolas diversos e simplificando a implantação de soluções baseadas em RNA's nas fazendas.

Esta tecnologia estabelece-se como ferramenta fundamental na agricultura moderna, impulsionando a inovação em diversos setores. Seu impacto vai além das práticas agrícolas imediatas, como plantio, irrigação, manejo de cultivos, controle de pragas, entre outros processos diretos relacionados à produção agrícola, abrangendo objetivos mais amplos de sustentabilidade e segurança alimentar. À medida que a pesquisa progride, a integração das RNA's na agricultura tende a se aprofundar, anunciando um futuro em que a agricultura não só é mais inteligente, mas também mais responsiva aos desafios do século XXI.

2.2.3. Adoção de Tecnologias Digitais nas Propriedades Rurais

A adoção de tecnologias digitais na agricultura, impulsionada por avanços tecnológicos, como a Internet das Coisas (IoT) e a Computação em Nuvem, promove uma revolução na gestão de cultivos. O Big Data desempenha um papel central nesta mudança, caracterizado pelo manejo de grandes volumes de dados diversos, possibilitando análises preditivas, tomadas de decisão em tempo real e redefinição de modelos de negócios para alcançar maior eficiência e sustentabilidade (Wolfert et al., 2017).

A implementação de Sistemas de Informação de Gestão (SIG) no nível da propriedade agrícola enfatiza a importância de compreender tanto os fatores motivadores quanto as barreiras para a adoção dessas tecnologias. Tais sistemas fornecem aos gestores informações necessárias para a tomada de decisões informadas (Laudon KC and Laudon JP, 2018). Aplicadas, frequentemente, para explicar essas dinâmicas, as teorias do quadro de Difusão de Inovações (DOI) destacam o papel das características da propriedade e do agricultor, juntamente com os atributos tecnológicos (Giua, Materia and Camanzi, 2021).

No entanto, a jornada em direção à Agricultura Digital não está isenta de seus desafios sociais e éticos. A abordagem de pesquisa e inovação responsável (RRI) mostra a necessidade de abordar esses desafios de maneira abrangente, focando na inclusão, sustentabilidade e considerações éticas no desenvolvimento e implantação destas tecnologias (Eastwood et al., 2019).

Em regiões como o Oriente Médio e Norte da África, a agricultura digital mostra grande potencial para aumentar a sustentabilidade dos sistemas de cultivo. Apesar de seu potencial, a adoção de tecnologias digitais nessas regiões ainda está no começo, impulsionada principalmente pelos benefícios econômicos na agricultura de alto valor. Há uma necessidade de políticas que promovam não apenas a adoção da tecnologia, mas também garantam acesso e considerem a sustentabilidade social e ambiental (Bahn, Yehya and Zurayk, 2021).

A disparidade digital, termo que se refere à lacuna entre aqueles que têm acesso à moderna tecnologia de informação e comunicação e aqueles que não têm, é uma preocupação crítica no contexto da digitalização agrícola. Tecnologias avançadas como sensores, inteligência artificial e robótica prometem eficiência e sustentabilidade, mas também apresentam riscos de exacerbar desigualdades e injustiças sociais, afetando especialmente os trabalhadores agrícolas e as comunidades rurais (Rotz et al., 2019).

Para fomentar uma compreensão abrangente e uma adoção da agricultura digital, é crucial integrar evidências empíricas do nível da fazenda com perspectivas sistêmicas sobre

a difusão da tecnologia, ajudando a oferecer uma compreensão mais completa e a orientar a implementação eficaz da transformação digital na agricultura (Shang et al., 2021b).

3. MATERIAL E MÉTODOS

3.1. Caracterização das Áreas Experimentais

O estudo foi elaborado no Laboratório de Máquinas e Mecanização Agrícola (LAMMA), pertencente ao Departamento de Engenharia e Ciências Exatas da Faculdade de Ciências Agrárias e Veterinárias (FCAV) da Universidade Estadual Paulista (UNESP), localizado em Jaboticabal, SP. Posteriormente, o projeto foi implementado e as etapas subsequentes foram desenvolvidas em duas áreas comerciais dedicadas ao cultivo de cana-de-açúcar.

A primeira área (Figura 7) apresenta um canal localizado no município de Jaboticabal, São Paulo, Brasil ($48^{\circ}18'39.1''$ W e $21^{\circ}10'14.3''$ S), cultivada em LATOSSOLO Vermelho distrófico com relevo plano a suave ondulado (0 – 8%) e clima com estação seca no verão. O estudo foi conduzido com a cultivar CTC 20 no oitavo ciclo de corte, com três amostragens entre 18 de maio e 31 de julho de 2023, com intervalos médios de 30 dias durante o período de maturação da cultura. Em cada avaliação, os dados foram coletados em 90 pontos de amostragem regularmente distribuídos (grade de 5 x 5 m), totalizando 270 pontos amostrais (90 amostras x 3 datas) em uma área total de 17.86 ha. Quatro colmos foram coletados aleatoriamente em cada ponto amostral nos dias das avaliações, para análise de sacarose através dos parâmetros de Brix ($^{\circ}$) e Pureza (%).

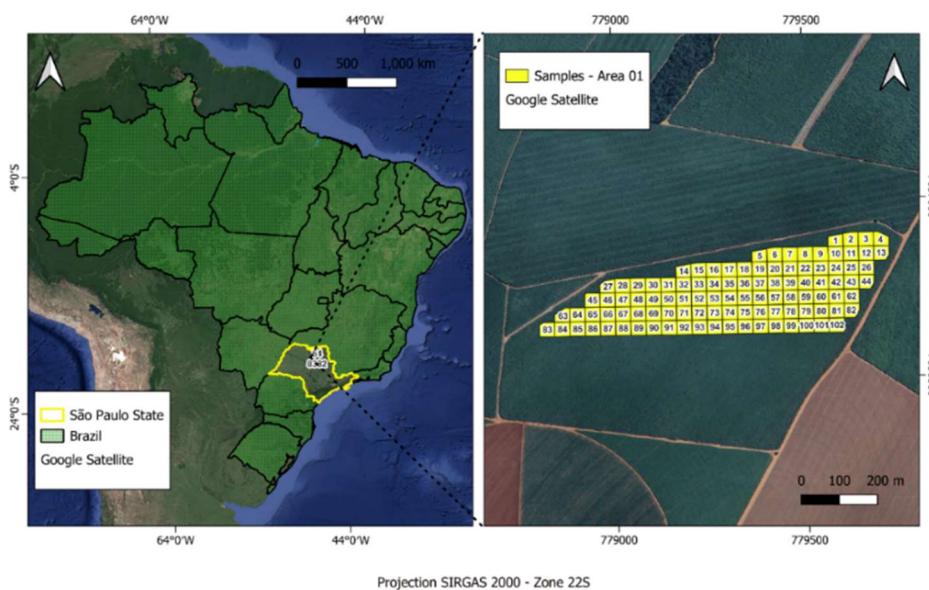


Figura 7 – Localização e pontos amostrais da Área 1

A segunda área de estudo (Figura 8) está localizada próximo ao município de Pradópolis, São Paulo, Brasil (48°09'35.6" W e 21°18'54.4" S), também em solo do tipo LATOSSOLO Vermelho distrófico, com declividade plano-ondulado (0 - 8%) e estação seca no verão. Nesta área, conduzida com a cultivar RB 97-5201 em sexto corte, foram feitas três amostragens ao longo da fase de maturação da cultura entre 28 de fevereiro e 8 de maio de 2022, com intervalos médios de 30 dias entre as amostragens. Em cada avaliação, os dados foram coletados em 30 pontos amostrais regularmente distribuídos (9 x 9 m) espaçados a cada dois metros, totalizando 90 pontos amostrais (30 amostras x 3 datas) em uma área de 0,37 ha. Quatro colmos foram coletados aleatoriamente em cada ponto amostral nos dias das avaliações, para análise de Brix (°) e Pureza (%).

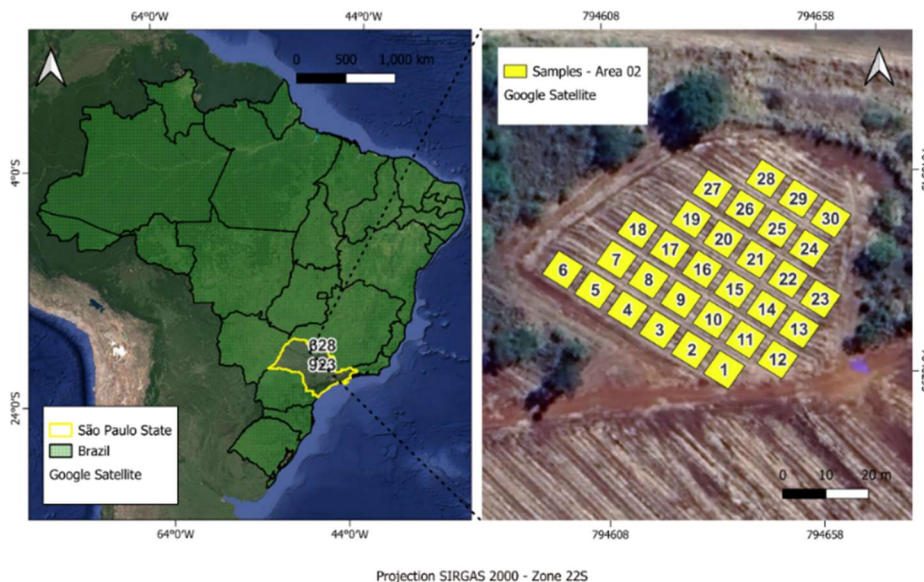


Figura 8 – Localização e pontos amostrais da Área 2

3.2. Metodologia de Extração do Caldo e Análise do Brix (°)

Os quatro colmos inteiros de cana-de-açúcar correspondentes a cada ponto amostral foram cortados manualmente a nível do solo, o ponto de crescimento e as folhas foram retirados antes de serem enviados a uma prensa hidráulica operando sob uma pressão constante de 24,5 MPa durante tempo mínimo de um minuto. Este procedimento assegura uma extração eficiente do caldo, minimizando a variação entre as amostras.

A determinação do Brix ° (teor de sólidos solúveis por cento, em peso, de caldo), foi realizada no laboratório da Faculdade de Tecnologia – FATEC. Utilizou-se um refratômetro digital (modelo ABBE, Atago Pal-1, Tóquio, Japão) com uma resolução máxima de 0,1 °Brix

para realizar as medições, conforme ilustrado na Figura 9. As leituras do Brix ($^{\circ}$) é um indicador crucial da concentração de sacarose, minerais e outras substâncias dissolvidas no caldo da cana-de-açúcar, fornecendo informações valiosas sobre a maturidade e a qualidade da cana.



Figura 9 - Refratômetro digital utilizado para análise do Brix ($^{\circ}$)

Para garantir a consistência e a precisão dos resultados, as leituras do Brix ($^{\circ}$) foram ajustadas para uma temperatura padrão de 20 $^{\circ}$ C, conforme o protocolo convencional proposto pelo Conselho Nacional dos Produtores de Cana-de-Açúcar do Estado de São Paulo (CONSECANA, 2006). Este ajuste de temperatura é essencial para compensar as variações térmicas que podem influenciar a densidade e, conseqüentemente, a leitura do refratômetro. Este procedimento assegura que as medições de Brix ($^{\circ}$) sejam precisas e confiáveis.

3.3. Metodologia de Análise da Pureza (%)

A Pureza % é um indicador vital que reflete a proporção de sacarose em relação aos sólidos solúveis totais presentes no caldo da cana-de-açúcar. Esta análise é crucial para avaliar a qualidade da safra e a eficiência do processamento, além de influenciar diretamente a eficiência na produção de açúcar e etanol. A metodologia adotada, conforme descrita pelo Conselho Nacional dos Produtores de Cana-de-Açúcar do Estado de São Paulo (CONSECANA, 2006), envolve um conjunto de etapas analíticas e equações específicas, detalhadas a seguir.

Inicialmente, 12 g de uma solução clarificante composta por celite 545 (4 partes), cloreto de alumínio (2 partes) e hidróxido de cálcio (1 parte) são diluídos em 200 ml do caldo

extraído. Esta solução é então filtrada para remover impurezas e partículas sólidas, garantindo a precisão das medições subsequentes. A pol do caldo (teor de sacarose aparente por cento, em peso, de caldo) é determinada por meio de um polarímetro digital (Anton Paar, Bremen, Alemanha), conforme ilustrado na Figura 10. Este instrumento mede a rotação da luz polarizada ao passar a solução clarificada.



Figura 10 - Polarímetro digital utilizado para determinar a pol do caldo

A pol do caldo é calculada através da equação (1):

$$\text{Pol} = \text{L Pb} \cdot (0,2605 - 0.0009882 \cdot \text{Brix}) \quad (1)$$

A transformação da leitura sacarimétrica com a mistura clarificante à base de alumínio, equivalente ao subacetato de chumbo é feita pela equação (2) abaixo:

$$\text{L Pb} = 1.00621 \cdot \text{L AI} + 0,05117 \quad (2)$$

Em que:

- LPb = Leitura sacarimétrica equivalente ao subacetato de chumbo;
- LAI = Leitura sacarimétrica obtida com a mistura clarificante a base de alumínio;

A pureza aparente do caldo, é definida como a porcentagem da Pol em relação ao Brix, é calculada pela equação (3):

$$\text{Pureza} = \frac{\text{Pol}}{\text{Brix}} \times 100 \quad (3)$$

Em que:

- LPb = Leitura sacarimétrica equivalente ao subacetato de chumbo;
- LAI = Leitura sacarimétrica obtida com a mistura clarificante a base de alumínio;
- Pol = Teor de sacarose aparente;
- Brix = Teor de sólidos solúveis.

Estas fórmulas são partes críticas dos testes laboratoriais que detalham sobre a composição química e qualidade da cana-de-açúcar fornecida as indústrias.

3.4. Cálculo e Aquisição de Graus-Dias Acumulados (aGDD)

Os graus-dias de crescimento (GDD) são uma medida amplamente utilizada na agronomia para estimar o crescimento e desenvolvimento das plantas, correlacionando-os com as condições ambientais. Este indicador é crucial na modelagem do desenvolvimento fenológico das culturas, como a cana-de-açúcar. O cálculo dos GDD permite prever os estágios de desenvolvimento da planta, como brotação, desenvolvimento vegetativo, maturação e florescimento.

Para cálculo do GDD, emprega-se a equação (4):

$$GDD = \left(\frac{T_{max} + T_{min}}{2} \right) - T_b \quad (4)$$

Para a cana-de-açúcar, a temperatura basal (T_b) é um parâmetro crítico que se altera conforme a variedade e as condições locais de solo e clima. Neste estudo, adotou-se 18 °C. A escolha desta temperatura está alinhada com as características fisiológicas da cultura, que tende a ter crescimento otimizado em condições de temperaturas mais elevadas (Bacchi, Osny OS, 1978).

Os valores diários de GDD são acumulados ao longo do tempo para obter os graus-dias acumulados (aGDD). Esta acumulação permite monitorar o progresso do desenvolvimento da cultura e prever os estágios fenológicos com maior precisão. A equação (5) demonstra o cálculo do GDD:

$$aGDD = \sum_{i=1}^n GDD_i \quad (5)$$

Em que:

- Tmax = temperatura máxima diária.
- Tmin = temperatura mínima diária.
- Tb = temperatura basal ou limite abaixo do qual o desenvolvimento da cultura é mínimo ou não ocorre.
- GDDi = valor dos Graus-Dia de Crescimento calculado para o dia i;
- n = número total de dias considerados no ciclo de crescimento.

Este processo continua ao longo do ciclo da cana-de-açúcar, proporcionando ferramenta valiosa para a gestão agrônômica e a estimativa de estágios de desenvolvimento da cultura.

3.5. Processo de Aquisição de Dados Multiespectrais

Para coletar os dados espectrais, foram utilizadas imagens capturadas pelo sensor Dove-R da plataforma PlanetScope. O produto Surface Reflectance é ortorretificado, composto por 8 bandas com correção geométrica, radiométrica e de superfície.

Os produtos são fornecidos no formato GeoTIFF de 16 bits com calibração radiométrica, e os valores de reflectância são escalonados por 10.000. Durante o processamento dos dados, realiza-se a divisão por 10.000 para recuperar o valor original de reflectância. As características técnicas detalhadas do sensor Dove-R estão apresentadas na Tabela 1.

Tabela 1 - Características técnicas do sensor Dove-R PlanetScope

Parâmetro	Descrição
Tipo de Sensor	(Dove-R) Gerador de imagens de oito bandas com filtro butcher-block
Resolução Espectral	Azul Costeiro: 431 - 452 nm Azul: 465 – 515 nm Verde I: 513 - 549 nm Verde: 547 – 583 nm Amarela: 600 - 620 nm Vermelha: 650 – 680 nm RedEdge: 697 – 713 nm NIR: 845 – 885 nm
Resolução Espacial	3.0 m

Tamanho do Quadro	24 km x 16 km (aproximado)
Máxima Faixa de Imagem	20,000 km ²
Resolução Temporal	Diário
Capacidade de Captura de Imagem	200 milhões de km ² /dia
Componentes do Produto	Arquivo de imagem (formato GeoTIFF) Arquivo de metadados (formato XML) Arquivo de miniatura (formato GeoTIFF) Arquivo UDM de máscara de dados inutilizável (formato GeoTIFF) Arquivo UDM2 de máscara de dados utilizável (formato GeoTIFF)
Projeção	UTM WGS84
Precisão	<10 m RMSE

A combinação dessas bandas espectrais proporciona um conjunto robusto de dados para análises agrícolas, permitindo uma análise detalhada da reflectância da superfície terrestre em diferentes comprimentos de ondas.

A aquisição diária de imagens permite o monitoramento contínuo das condições da vegetação ao longo do tempo, essencial para a gestão agrícola e a tomada de decisões mais informadas. A capacidade de monitorar a evolução das condições da vegetação oferece insights valiosos sobre a saúde das culturas, a necessidade de irrigação, a presença de pragas ou doenças e a eficiência das práticas agrícolas implementadas. Garantindo uma abordagem integrada através do sensoriamento remoto, facilitando uma gestão agrícola mais precisa e eficaz.

3.6. Cálculo de Índices de Vegetação e Extração dos Dados Multiespectrais

Com a obtenção dos dados multiespectrais e laboratoriais, o software QGIS, ferramenta de Sistema de Informação Geográfica (SIG) de código aberto (QGIS Development Team, Open-Source Geospacial Foundation, Chicago, IL, EUA), foi utilizado para extrair os valores das bandas multiespectrais por meio da ferramenta “estatísticas zonais” e posteriormente aplicar os cálculos dos índices de vegetação. A utilização do QGIS para esta tarefa se deve à capacidade em manipular dados especializados, combinada com interface amigável e ampla gama de funcionalidades para o processamento de dados de sensoriamento remoto e análises de geoprocessamento.

A seguir, constam na Tabela 2, os 16 índices de vegetação, que variam nas especificidades e aplicações, sendo selecionados com base na literatura disponível para a estimativa de vigor de plantas, parâmetros qualitativos e produtividade de culturas agrícolas

(Souza et al., 2021; Tedesco et al., 2021; Barbosa Júnior et al., 2022; de Oliveira et al., 2022; Barbosa et al., 2023). Permitindo avaliação abrangente para seleção dos melhores índices preditores de sacarose na cana-de-açúcar.

Tabela 2 - Apresentação dos índices de vegetação e suas respectivas fórmulas

Índice	Fórmula	Citação do Autor
BGI	$\frac{(\rho_{Blue})}{(\rho_{Green})}$	(Zarco-Tejada et al., 2005)
CIRE	$\frac{(\rho_{NIR})}{\rho_{RedEdge}}$	(Gitelson, Gritz and Merzlyak, 2003)
CIVE	$(0.441 \times \rho_{Red} - 0.811 \times \rho_{Green} + 0.385 \times \rho_{Blue} + 18.79)$	(Kataoka et al., 2003)
GLI	$\frac{2 \times (\rho_{Green} - \rho_{Red} - \rho_{Blue})}{2 \times (\rho_{Green} + \rho_{Red} + \rho_{Blue})}$	(Louhaichi, Borman and Johnson, 2001)
GNDVI	$\frac{(\rho_{NIR} - \rho_{Green})}{(\rho_{NIR} + \rho_{Green})}$	(Gitelson, Kaufman and Merzlyak, 1996)
NDVI	$\frac{(\rho_{NIR} - \rho_{Red})}{(\rho_{NIR} + \rho_{Red})}$	(Rouse et al., 1974)
YNDVI	$\frac{(\rho_{NIR} - \rho_{Yellow})}{(\rho_{NIR} + \rho_{Yellow})}$	(J Jasper and S Reush, 2009)
EVI	$2.5 \times \frac{(\rho_{NIR} - \rho_{Red})}{\rho_{NIR} + (6 \times \rho_{Red} - 7.5 \times \rho_{Blue}) + 1}$	(Huete, Justice and Liu, 1994)
EVI2	$2.5 \times \frac{(\rho_{NIR} - \rho_{Red})}{\rho_{NIR} + (2.4 \times \rho_{Red}) + 1}$	(Jiang et al., 2008)
OSAVI	$\frac{1.5 \times \rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red} + 0.16}$	(Rondeaux, Steven and Baret, 1996)
SR	$\frac{\rho_{NIR}}{\rho_{Red}}$	(Cohen, 1991)
MSR	$\frac{\frac{\rho_{NIR}}{\rho_{Red}} - 1}{\sqrt{\frac{\rho_{NIR}}{\rho_{Red}} + 1}}$	(Leblanc, Chen and Cihlar, 1997)
ExG	$2 \times \rho_{Green} - \rho_{Red} - \rho_{Blue}$	(Gitelson, Gritz and Merzlyak, 2003)
TVI	$0.5 \times (120 \times (\rho_{NIR} - \rho_{Green}) - 200 \times (\rho_{Red} - \rho_{Green}))$	(Broge and Leblanc, 2001)
VARI	$\frac{(\rho_{Green} - \rho_{Red})}{(\rho_{Green} + \rho_{Red} - \rho_{Blue})}$	(Gitelson et al., 2002)
SFDVI	$\frac{\rho_{NIR} + \rho_{Green}}{2}$	(Baptista, 2015)

BGI: Blue Green Index; CIRE: Chlorophyll Index Red Edge; CIVE: Chlorophyll Index Visible; GLI: Green Leaf Index; GNDVI: Green Normalized Difference VI; NDVI: Normalized Difference VI; YNDVI: Yellow Normalized Difference VI; EVI: Enhanced Vegetation Index; EVI2: Enhanced Vegetation Index 2; OSAVI: Optimized Soil

Adjusted VI; SR: Simple Ratio; MSR: Modified Simple Ratio; ExG: Excess Green Index; TVI: Transformed Vegetation Index; VARI: Visible Atmospherically Resistant VI; SFDVI: Soil and Atmospherically Resistant VI.

3.7. Análise Descritiva dos Dados Coletados

A análise da resposta das variáveis Brix (°) e Pureza (%) ao longo dos intervalos de coleta foi conduzida utilizando estatística descritiva, com a aplicação de gráficos Box Plot, em função dos dias após o plantio (DAP) para cada uma das áreas de estudo. Foram consideradas três datas de coleta, além de uma quarta análise geral, que examinou a variação dos dados desde a primeira até a última coleta de campo.

Na Área 1, a técnica do intervalo interquartil foi empregada para a remoção de outliers, resultando na exclusão de mais de 7% dos dados originais. Após essas etapas, o conjunto de dados final para a Área 1 foi reduzido para 240 pontos amostrais. Da mesma forma, na Área 2, a aplicação da técnica do intervalo interquartil levou à exclusão de 7% dos dados, resultando em um conjunto final de 84 pontos amostrais, utilizados para fins de transferência e generalização dos resultados.

3.8. Estratégias de Seleção e Estruturação de Variáveis para Estimativa

Para a construção de um banco de dados adequado à modelagem estimativa dos parâmetros de Brix (°) e Pureza (%), foram compiladas 25 variáveis distintas, abrangendo 8 Bandas Multiespectrais (BM), 16 Índices de Vegetação (IVs) e a variável de Graus-Dias Acumulados (aGDD). A metodologia utilizada para a seleção das variáveis mais relevantes para a estimativa baseou-se na técnica de Eliminação Recursiva de Atributos (RFE), aplicada por meio de regressão linear com o uso da biblioteca scikit-learn, no ambiente Python (versão 3.10). O principal objetivo desta etapa do estudo foi identificar, entre as bandas multiespectrais e índices de vegetação disponíveis, aquelas que apresentassem maior poder preditivo em relação aos valores de Brix (°) e Pureza (%).

A técnica de RFE é caracterizada por ser um processo iterativo que, inicialmente, treina um modelo de regressão linear utilizando todas as 25 variáveis disponíveis no conjunto de dados. Este modelo inicial é empregado para avaliar a importância relativa de cada variável. No contexto da regressão linear, essa importância é geralmente medida pelos coeficientes absolutos das variáveis no modelo.

Após o treinamento inicial, as variáveis são ranqueadas com base em sua importância. A variável que apresenta a menor relevância é classificada com o maior valor numérico, indicando menor importância. Consequentemente, a variável menos relevante é

eliminada, e o modelo é re-treinado utilizando as variáveis remanescentes. Este processo de reavaliação da importância das variáveis e subsequente eliminação é repetido de forma iterativa até que se atinja o número especificado de variáveis.

O objetivo final deste procedimento é refinar o conjunto de variáveis, preservando apenas aquelas que contribuem significativamente para a precisão do modelo preditivo. O processo é concluído quando se alcança o conjunto ideal de variáveis, que são as que demonstram maior importância e relevância para os modelos de redes neurais artificiais.

3.9. Aplicação de Redes Neurais Artificiais

A modelagem preditiva dos valores de Brix (°) e Pureza (%) na cana-de-açúcar exige o uso de ferramentas robustas e precisas, capazes de lidar com a complexidade dos dados agrônômicos. Neste estudo, foi utilizado o software STATISTICA, versão 7 (StatSoft, Inc., Tulsa, OK, USA), conhecido por sua eficiência na análise estatística e na modelagem de dados, para implementar, treinar e validar os modelos de Redes Neurais Artificiais (RNAs). A escolha deste software deve-se à sua interface intuitiva e à capacidade de realizar análises profundas e customizadas, essenciais para as necessidades deste trabalho.

Para a estimativa dos parâmetros Brix (°) e Pureza (%), foram empregadas duas arquiteturas de redes neurais artificiais: Multilayer Perceptron (MLP) e Radial Basis Function (RBF). As arquiteturas escolhidas foram ajustadas de acordo com as especificidades dos dados coletados, utilizando diferentes combinações de bandas espectrais e índices de vegetação associados aos graus-dias acumulados (aGDD).

3.9.1. Modelo de Função de Base Radial (RBF)

A Rede de Função de Base Radial (RBF) é um modelo de rede neural artificial amplamente utilizado em tarefas de classificação e regressão devido à sua capacidade de realizar transformações não-lineares complexas. Nesta rede, as camadas são organizadas de forma a permitir que as variáveis de entrada sejam processadas e transformadas em características ocultas que, posteriormente, são combinadas para gerar a saída predita.

A estrutura da rede RBF é composta por três camadas principais: a camada de entrada, onde as variáveis iniciais são recebidas; as camadas intermediárias ou ocultas, onde funções de base radial são aplicadas para mapear os dados em um novo espaço de características; e a camada de saída, que combina as ativações obtidas nas camadas ocultas para produzir a estimativa final.

A Figura 11 ilustra a arquitetura de uma rede RBF aplicada ao Brix (°) como camada de saída, utilizando bandas multiespectrais como entrada demonstrando o fluxo de informação através das camadas (3:3-20-1:1) e destacando a importância das transformações realizadas nas camadas ocultas para alcançar resultados satisfatórios de estimativas.

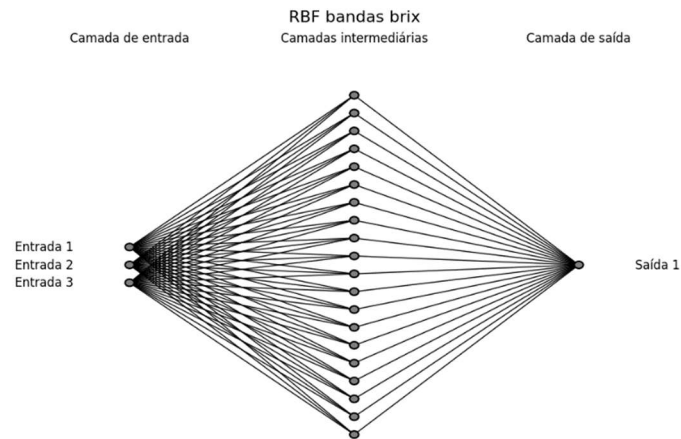


Figura 11 – Estrutura da Rede Neural RBF (3:3-20-1:1) para estimar Brix (°) utilizando dados multiespectrais.

A função de base radial, utilizada na camada oculta, varia conforme a distância a um ponto central. A função gaussiana (5), frequentemente utilizada como função de base radial nas redes RBF, é expressa pela fórmula:

$$\varphi = \exp\left(-\frac{v^2}{2\sigma^2}\right) \quad (5)$$

Em que:

- φ = o valor de ativação da unidade radial (neurônio da camada oculta).
- $v = \|x - \mu\|$ representando a distância euclidiana entre o vetor de entrada x e o centro μ da função gaussiana. Esta distância é a "entrada" para a função gaussiana.
- σ = o parâmetro que determina a largura da função gaussiana. É conhecido como fator de suavização e afeta o quão rápido o valor de ativação φ diminui à medida que a distância v aumenta.

A função de ativação gaussiana tem um pico no centro μ e decresce rapidamente conforme a distância do ponto de entrada aumenta. Isso significa que cada neurônio na

camada oculta "responde" a diferentes regiões do espaço de entrada, dependendo de quão próximos ou distantes estão os pontos de entrada de seus centros μ .

3.9.2. Modelo Perceptron de Multicamadas (MLP)

O Perceptron de Multicamadas (MLP) é uma arquitetura de rede neural amplamente utilizada, caracterizada pelo uso do algoritmo de retropropagação de erros em um processo de treinamento supervisionado. Essa abordagem permite que a rede ajuste seus parâmetros internos, os pesos sinápticos, para minimizar o erro entre as previsões da rede e os valores reais dos dados de treino.

A estrutura básica de uma rede MLP é composta por três tipos de camadas. Primeiramente, as camadas de entrada são responsáveis por receber os dados de entrada, que, no contexto deste estudo, incluem informações espectrais obtidas de sensores, índices de vegetação e graus-dias acumulados, seja em combinação ou isoladamente. Esses dados são então transmitidos para as camadas ocultas, onde ocorre a transformação dos dados. No estudo em questão, o MLP foi configurado com duas camadas ocultas, contendo 20 e 12 neurônios, respectivamente. Por fim, a camada de saída fornece a previsão final da rede.

A Figura 12 ilustra a estrutura da rede MLP (3:3-20-12-1:1) utilizada para a estimativa dos valores de Brix ($^{\circ}$), onde a rede possui três neurônios na camada de entrada, vinte neurônios na primeira camada oculta, doze neurônios na segunda camada oculta, e um neurônio na camada de saída sendo Brix ($^{\circ}$) e posteriormente a Pureza (%).

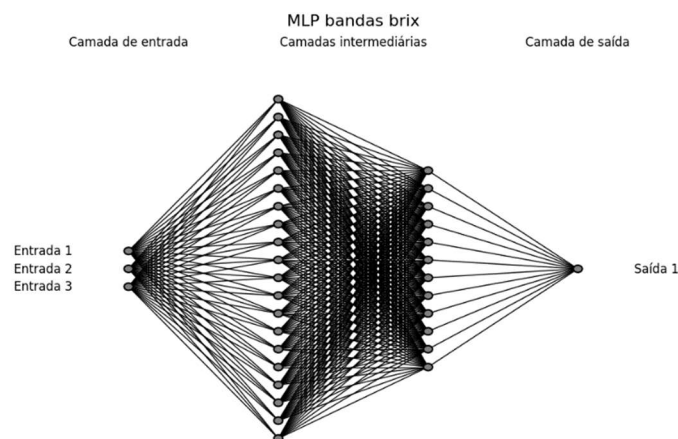


Figura 12 – Estrutura da Rede Neural MLP (3:3-20-12-1:1) para estimar Brix ($^{\circ}$) utilizando dados multiespectrais.

As conexões na rede são feitas através de elementos chamados pesos sinápticos, que armazenam o conhecimento adquirido. Os dados inseridos nas camadas de entrada foram ajustados usando as fórmulas (6), (7) e (8) apresentadas abaixo:

$$y_i = \frac{x_i - x_{\max}}{x_{\max} + x_{\min}} \quad (6)$$

Em que:

- y_i = valor normalizado da i -ésima característica. Este valor é obtido após a aplicação da fórmula de normalização e será usado como entrada na rede neural.
- x_i = valor original da i -ésima característica antes da normalização.
- x_{\max} = valor máximo observado para a característica i no conjunto de dados. Este valor é usado como parte do denominador para ajustar a escala do valor original.
- x_{\min} = valor mínimo observado para a característica i no conjunto de dados. Também faz parte do denominador e é usado para ajustar a escala do valor original.

Esta fórmula é usada para normalizar ou escalar os valores de entrada da rede neural, x_i , para que todos fiquem em uma escala comum entre 0 e 1. A normalização é feita subtraindo o menor valor observado, x_{\min} , do valor atual, x_i , e dividindo pelo intervalo de valores (a diferença entre o maior valor observado, x_{\max} , e o menor valor observado, x_{\min}).

$$a_k = \sum_j y_j w_{kj} \quad (7)$$

Em que:

- a_k = soma ponderada das entradas para o k -ésimo neurônio na camada subsequente. É o resultado da soma de cada entrada multiplicada pelo seu respectivo peso, antes da aplicação de qualquer função de ativação.
- \sum = símbolo de somatória, indicando que a expressão seguinte deve ser somada para todos os valores de j , que representam os índices dos neurônios na camada anterior.
- y_j = valor de saída do j -ésimo neurônio da camada anterior, que pode ser o valor de entrada original ou o resultado de uma função de ativação aplicada a esse neurônio.

- w_{kj} = peso da conexão entre o j-ésimo neurônio da camada anterior e o k-ésimo neurônio da camada atual. Este peso ajusta a importância do sinal de entrada y_j ao ser somado para calcular a_k

Esta fórmula calcula o valor de ativação a_k de um neurônio na rede neural. É uma soma ponderada na qual o valor de entrada normalizado y_j é multiplicado por um peso sináptico w_{kj} correspondente, e então todos esses produtos são somados juntos. Os pesos sinápticos são ajustáveis e são aprendidos pela rede durante o treinamento.

$$g(a_k) = \frac{e^{a_k} - e^{-a_k}}{e^{a_k} + e^{-a_k}} \quad (8)$$

Em que:

- $g(a_k)$ = valor de saída do neurônio após a aplicação da função de ativação tangente hiperbólica ao seu potencial de ativação a_k .

- a_k = potencial de ativação do neurônio, que é a soma ponderada das entradas deste neurônio, como foi descrito na fórmula (7) anterior.

- e^{a_k} = função exponencial elevada ao potencial de ativação a_k .

- e^{-a_k} = função exponencial elevada ao negativo do potencial de ativação a_k .

Após calcular o valor de ativação a_k , aplica-se a função de ativação hiperbólica, que é a função $g(a_k)$. Esta função transforma a_k em nova saída y_k , que pode ser usada como entrada para a próxima camada na rede ou como a saída final se for a última camada. A função hiperbólica é especialmente útil porque pode lidar com valores negativos e positivos e os comprime entre -1 e 1, permitindo que a rede neural lide com não-linearidades complexas nos dados.

3.10. Avaliação do Desempenho dos Modelos de Redes Neurais Artificiais

3.10.1. Fase de Treinamento e Teste

Na primeira etapa do estudo, os modelos de Redes Neurais Artificiais (RNAs) passaram por um processo rigoroso de treinamento e teste. Cada modelo foi treinado ao longo de 2000 iterações, um número cuidadosamente escolhido para garantir que as redes tivessem tempo suficiente para aprender os padrões nos dados. Durante essa fase, foram

utilizadas duas abordagens principais: o Perceptron de Multicamadas (MLP) e as Redes de Função de Base Radial (RBF). A estrutura das redes foi mantida constante entre as duas abordagens, garantindo uniformidade na comparação dos resultados. Para explorar diferentes níveis de complexidade, o número de neurônios nas camadas ocultas variou entre 1 e 20, conforme sugerido por (Souza et al., 2022). Essa variação permitiu testar como a complexidade da rede impactava o desempenho durante o treinamento.

Para avaliar a eficácia dos modelos, gráficos de dispersão foram utilizados para comparar as estimativas geradas pelos modelos com os valores laboratoriais observados, em uma proporção de 1:1. Essa abordagem visual é fundamental para entender quão bem os modelos conseguem replicar os dados reais. Além disso, a precisão e a qualidade do ajuste das estimativas foram quantificadas utilizando duas métricas principais:

Erro Médio Absoluto (*MAE* – Eq. 9), métrica foi utilizada para medir a magnitude média dos erros de estimativa. O *MAE* fornece uma medida direta da diferença média entre as estimativas geradas pelos modelos e os valores observados, expressa na mesma unidade dos dados originais.

$$MAE = \frac{1}{n} \sum_{i=1}^n (P_i - O_i) \quad (9)$$

Onde:

- P_i = valor previsto.
- O_i = valor observado.
- n = número total de observações.

Coefficiente de Determinação Ajustado $R^2_{ajustado}$ – Eq.10), métrica foi utilizada para avaliar a qualidade do ajuste das estimativas do modelo em relação aos valores observados, levando em conta o número de parâmetros no modelo. O $R^2_{ajustado}$ corrige o valor do R^2 tradicional para refletir a complexidade do modelo, penalizando excessos de complexidade que não resultam em ganhos significativos de precisão.

$$R^2_{ajustado} = 1 - \left(\frac{SQR/(n - p - 1)}{SQT/(n - 1)} \right) \quad (10)$$

Onde:

- SQR = soma dos quadrados dos resíduos
- SQT = soma dos quadrados totais
- p = número total de preditores (neurônios) no modelo
- n = número total de observações.

A combinação dessas métricas permitiu uma análise detalhada sobre quais modelos foram mais eficazes na estimativa dos valores de Brix (°) e Pureza (%), considerando tanto a magnitude dos erros quanto a qualidade do ajuste.

3.10.2. Transferência e Generalização dos Modelos

Na segunda etapa do estudo, a mesma estrutura de rede utilizada durante o treinamento e teste foi aplicada para avaliar a capacidade de generalização dos modelos. Esta fase foi essencial para verificar se os modelos, uma vez treinados, conseguiriam manter seu desempenho ao serem aplicados a novos conjuntos de dados, diferentes daqueles utilizados na fase inicial.

Durante essa etapa, a consistência e a acurácia dos modelos foram avaliados utilizando novamente a métrica MAE . A análise focou-se em determinar se os modelos mantinham um desempenho robusto ao serem expostos a diferentes cenários, considerando variações nos dados de entrada, como bandas multiespectrais, índices de vegetação e suas combinações.

O objetivo principal desta fase foi identificar quais modelos demonstraram a melhor capacidade de generalização, ou seja, quais redes foram capazes de replicar com acurácia e precisão os padrões aprendidos durante o treinamento quando confrontadas com novos dados. O modelo que apresentou o menor valor de MAE foi considerado o mais eficiente na generalização das estimativas de Brix (°) e Pureza (%).

Modelos menos complexos, com um menor número de neurônios nas camadas ocultas, apresentaram uma capacidade de generalização superior às redes mais complexas. Destacando a importância de se buscar um equilíbrio adequado entre a complexidade do modelo e sua capacidade de generalizar, ressaltando que redes mais complexas nem sempre proporcionam um desempenho superior.

4. RESULTADOS

4.1. Análise estatística

A estatística descritiva (Figura 13) mostra a variação dos valores de Brix (°) e Pureza (%) em duas condições de cultivo distintas, identificadas como (a) área 1 e (b) área 2. As três primeiras caixas demonstram a variação dos dados em cada etapa de coleta.

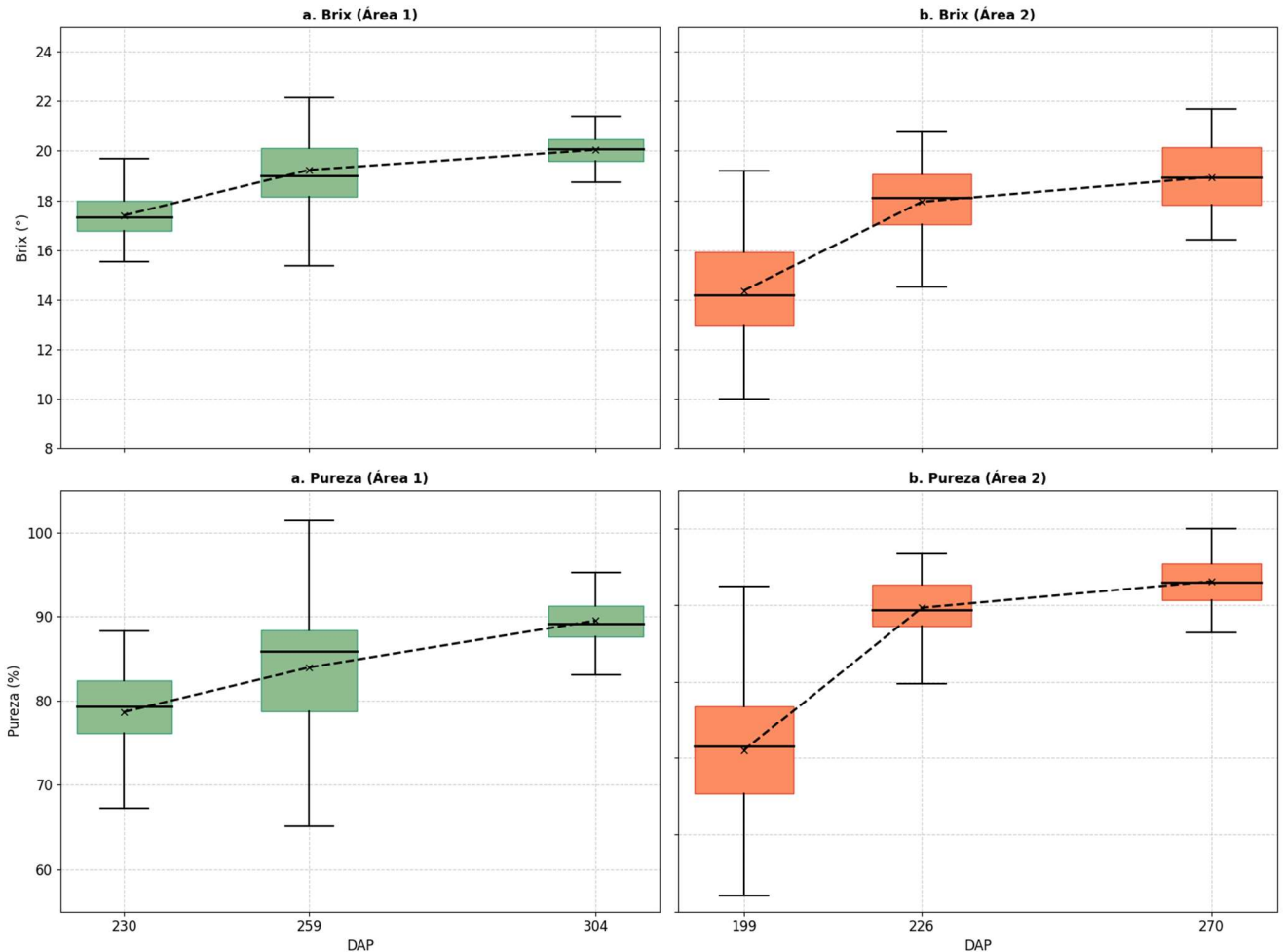


Figura 13 - Variação temporal de Brix (°) e Pureza (%) em duas áreas de cultivo, medidos em Dias Após o Plantio (DAP).

A análise de Brix (°), revelou padrões distintos entre as duas condições de cultivo ao longo do tempo. Na condição (a), observamos um intervalo mais amplo de variação nos índices de Brix em comparação com a condição (b), especialmente nos 230 DAP com 995.5 GDD para 304 DAP com 1023.1 GDD para (a); e 199 DAP (1104.3 GDD) e 270 DAP (1293.52 GDD) para (b).

Em relação à pureza, as áreas apresentaram tendência de aumento da pureza com o avanço do tempo de cultivo. No entanto, a condição (b) demonstrou maior consistência na pureza ao longo dos DAP analisados, indicando uma possível estabilidade no processo de acumulação de sacarose nesta condição de cultivo, consequentemente indicando o ponto ideal de colheita.

Os valores médios de Brix (°) e Pureza foram marcados por "X" nos gráficos de caixa, fornecendo uma representação visual rápida da média em comparação com a distribuição geral dos dados. A condição de cultivo (b) pode oferecer um ambiente mais estável para o desenvolvimento de sacarose, o que é evidenciado pela menor variação nos índices de Brix e pela Pureza mais consistente ao longo do tempo. A condição (a), embora mostre maior variabilidade, pode ser indicativa de outros fatores ambientais ou de manejo influenciando a composição dos sólidos solúveis nas amostras.

A compreensão detalhada dessas variações pode ajudar na otimização das práticas de cultivo para maximizar a qualidade e o rendimento do produto. Futuras investigações poderão se aprofundar nos fatores que contribuem para as diferenças observadas entre as condições de cultivo (a) e (b).

4.2. Seleção das Variáveis de Entrada

A seleção criteriosa de variáveis de entrada é um passo fundamental na construção de modelos preditivos robustos e precisos, especialmente no contexto da agronomia, onde a complexidade dos dados pode impactar significativamente os resultados. A escolha correta das variáveis é essencial para garantir que o modelo capte as nuances dos processos biológicos e agrônômicos, permitindo previsões mais acuradas e confiáveis. Para enfrentar essa complexidade e garantir que apenas as variáveis mais informativas sejam utilizadas, optamos por empregar o método de Eliminação Recursiva de Atributos (RFE), amplamente reconhecido por sua capacidade de refinar a seleção de variáveis ao longo de várias iterações.

O RFE é um método iterativo que ajusta repetidamente os modelos, eliminando características menos importantes até que o conjunto ideal de variáveis seja identificado. Esse processo é especialmente valioso em contextos em que a inclusão de dados irrelevantes pode não só introduzir ruído, mas também aumentar a complexidade do modelo sem ganho real em precisão. Através do RFE, conseguimos focar nas variáveis que realmente contribuem para a capacidade preditiva do modelo, assegurando que este seja eficiente e interpretável. Ao eliminar gradualmente as variáveis de menor impacto, o RFE

ajuda a construir modelos que são não apenas mais simples, mas também mais robustos e adaptáveis a diferentes cenários de aplicação.

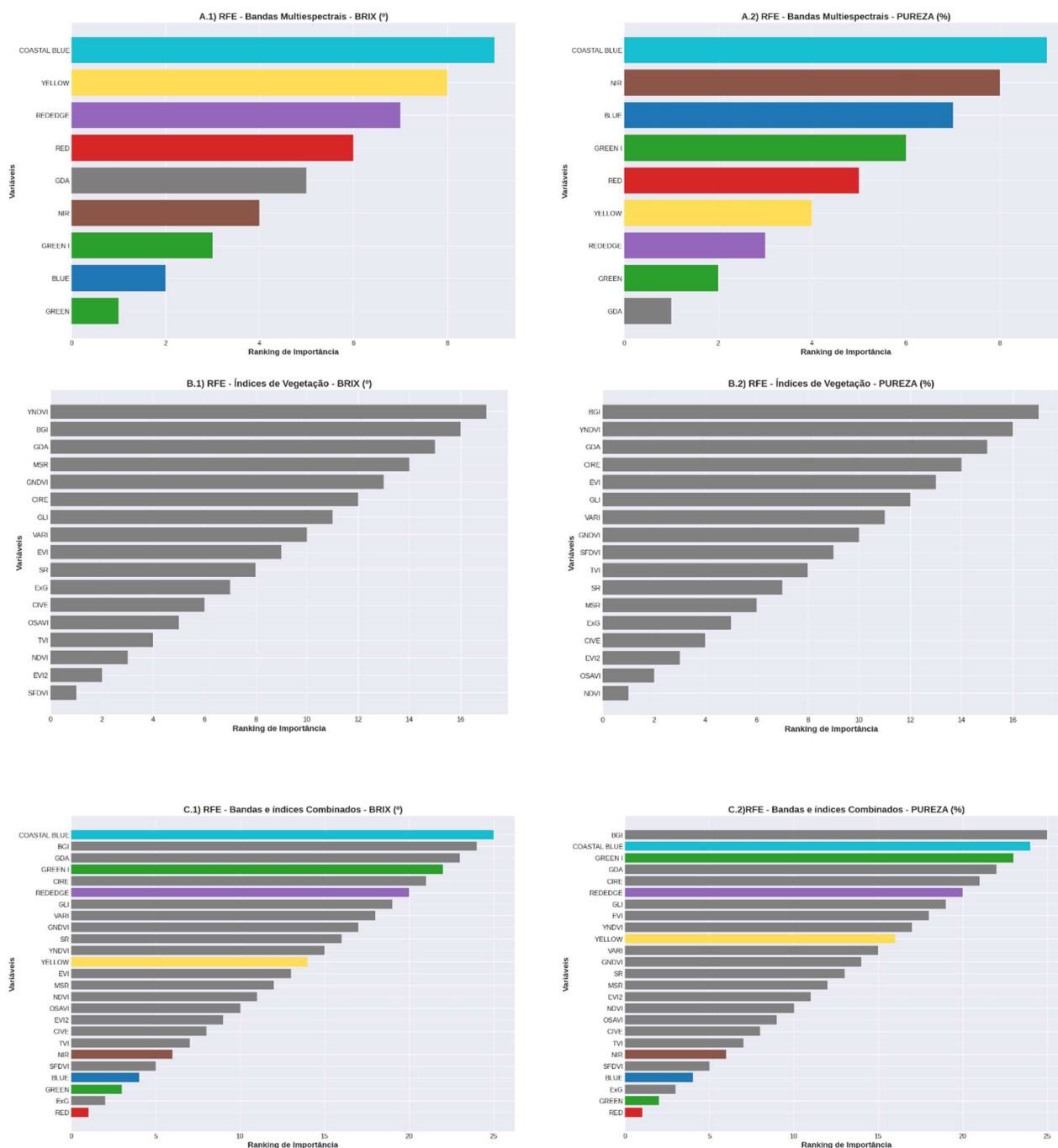


Figura 14 – Aplicação do RFE na estimativa de Brix (°) e Pureza (%), destacando as variáveis selecionadas em diferentes categorias: Bandas Multiespectrais, Índices de Vegetação, e suas combinações.

A análise revelou que, para a estimativa de Brix (°), as bandas Verde, Azul, Verde I e NIR, quando associadas aos aGDD, mostraram-se particularmente influentes. Estas

bandas, que capturam variações críticas na absorção e reflexão da luz pelas plantas, são essenciais para compreender a dinâmica da fotossíntese e da acumulação de açúcares, refletidos nos valores de Brix (°). A incorporação de aGDD complementa essas informações, adicionando uma dimensão temporal que correlaciona o desenvolvimento fenológico da cultura com a variabilidade do Brix (°). Em contraste, para a estimativa de Pureza (%), as bandas Verde, Vermelho, RedEdge e Amarelo, também em conjunto com os aGDD, destacaram-se como as principais contribuintes. Essas bandas são sensíveis a aspectos como a clorofila, a saúde das folhas e o estágio de maturação, todos fatores que influenciam diretamente a concentração de sacarose e, portanto, a Pureza (%).

No que diz respeito aos Índices de Vegetação, o SFDVI, EVI2, NDVI, TVI e OSAVI desempenham um papel fundamental na estimativa de Brix (°). Estes índices são derivados de combinações específicas de bandas multiespectrais e são reconhecidos por sua capacidade de realçar características como a densidade foliar e o teor de clorofila, que estão intimamente relacionados ao acúmulo de açúcares na planta. Para Pureza (%), os índices NDVI, OSAVI, EVI2, CIVE e ExG foram considerados os mais relevantes. Cada um desses índices captura diferentes aspectos da condição das plantas, desde o vigor vegetativo até a composição bioquímica das folhas, o que pode influenciar diretamente a pureza do caldo de cana.

Ao explorar a combinação de Bandas Multiespectrais e Índices de Vegetação, observamos que, para Brix (°), as variáveis Vermelho, ExG, Verde, Azul e SFDVI proporcionaram uma visão mais abrangente. Esta abordagem integrada sugere que a combinação dessas variáveis pode potencializar a robustez do modelo preditivo, pois cada conjunto de dados oferece uma perspectiva única sobre o estado fisiológico da cultura. Essa combinação permite capturar a complexidade do crescimento da planta e a variação nos teores de açúcar, resultando em uma modelagem mais precisa e confiável. Para Pureza (%), as variáveis Vermelho, Verde, ExG, Azul e SFDVI também se mantiveram como as mais significativas, reforçando a importância de uma seleção de variáveis que considere múltiplas fontes de dados. A abordagem multivariada mostrou-se particularmente eficaz na identificação dos fatores que influenciam a pureza, embora a complexidade desse parâmetro ainda represente um desafio na modelagem preditiva.

A Tabela 3, a seguir, exemplifica as variáveis de entrada selecionadas para construção dos modelos para estimativa do Brix (°) e Pureza (%).

Tabela 3 - Variáveis de importância na estimativa de Brix (°) e Pureza (%) utilizando dados de imagens orbitais.

Variáveis de Entrada	Parâmetros Qualitativos	Variáveis de Importância
BM + aGDD	Brix (°)	<i>Verde; Azul; Verde 1; NIR; aGDD</i>
VI's + Agdd	Brix (°)	<i>SFDVI; EVI 2; NDVI; TVI; OSAVI</i>
Combinação + aGDD	Brix (°)	<i>Vermelho; EXG; Verde; Azul; SFDVI</i>
BM + aGDD	Pureza (%)	<i>aGDD; Verde; RedEdge, Amarelo; Vermelho</i>
VI's + aGDD	Pureza (%)	<i>NDVI; OSAVI; EVI 2; CIVE; ExG</i>
Combinação + aGDD	Pureza (%)	<i>Vermelho; Verde; ExG; Azul; SFDVI</i>

BS: Bandas de Satélite; IV's: Índices de Vegetação; CIRE: Chlorophyll Index Red Edge; CIVE: Chlorophyll Index Visible; GLI: Green Leaf Index; NDVI: Normalized Difference VI; EVI: Enhanced Vegetation Index; OSAVI: Optimized Soil Adjusted VI; SR: Simple Ratio; MSR: Modified Simple Ratio e aGDD: Graus-Dias Acumulados.

O método RFE demonstrou-se valioso ao equilibrar a necessidade de precisão com a simplicidade do modelo, facilitando o desenvolvimento de modelos que são ao mesmo tempo precisos e manejáveis. Esse processo, ao focar nas variáveis realmente impactantes, não apenas fortalece a capacidade de estimativa dos modelos atuais, mas também estabelece uma base sólida para investigações futuras. A integração de técnicas avançadas de seleção de variáveis ou a inclusão de novos dados poderá, ainda mais, aprimorar a robustez e a aplicabilidade dos modelos desenvolvidos, tornando-os ferramentas essenciais na gestão agrícola inteligente.

4.3. Desempenho dos Modelos Durante os Testes

4.3.1. Comparação entre modelos RBF e MLP na estimativa de Brix (°)

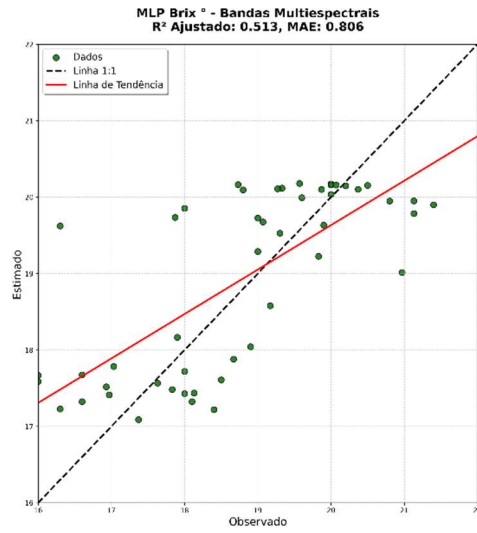
Na comparação entre os modelos RBF e MLP na estimativa de Brix (°), foram avaliados diferentes conjuntos de dados de entrada e métodos de modelagem. Os resultados dessas comparações são ilustrados em gráficos de dispersão que correlacionam os valores observados no eixo X com os valores estimados no eixo Y pelos modelos. Para o parâmetro Brix (°), os gráficos foram organizados em duas linhas: a linha 1, composta por a.1, a.2 e a.3, representa os modelos utilizando MLP, enquanto a linha 2, composta por b.1, b.2 e b.3, corresponde aos modelos baseados em RBF. Esses gráficos revelam que o desempenho dos modelos variou significativamente de acordo com o tipo de dado de entrada e a arquitetura utilizada.

Ao analisar os resultados para Brix ($^{\circ}$), observa-se que o modelo MLP, quando utilizando bandas multiespectrais como variáveis de entrada (a.1), apresentou um R^2 ajustado de 0,513 e MAE de 0,806. Esses resultados indicam uma moderada precisão na estimativa, sugerindo que as bandas multiespectrais capturam informações relevantes para a estimativa de Brix ($^{\circ}$). Quando foram utilizados apenas os índices de vegetação (a.2), o desempenho do modelo MLP melhorou ligeiramente, com o R^2 ajustado aumentando para 0,517 e o MAE diminuindo para 0,786. Esses valores sugerem que os índices de vegetação são variáveis relevantes para Brix ($^{\circ}$), capturando bem as características necessárias para uma boa estimativa. No entanto, ao combinar as bandas multiespectrais com os índices de vegetação (a.3), o desempenho do MLP apresentou uma leve diminuição, com o R^2 ajustado caindo para 0,506 e o MAE aumentando para 0,841. Isso indica que a adição de mais variáveis introduziu uma complexidade maior no modelo, sem resultar em uma melhora significativa na precisão das estimativas.

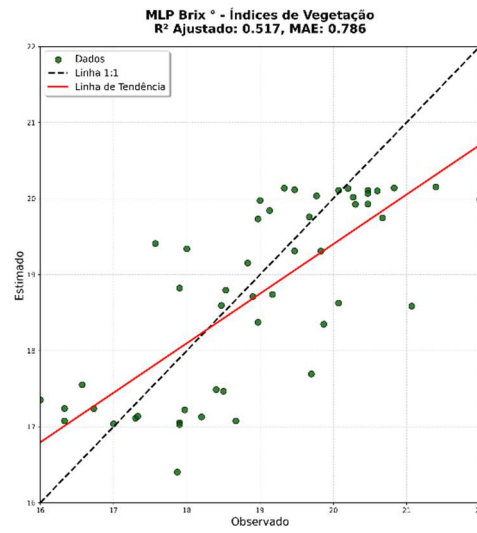
Por outro lado, os resultados obtidos com o RBF foram, em geral, inferiores aos obtidos com o MLP. Quando utilizando bandas multiespectrais (b.1), o RBF alcançou um R^2 ajustado de 0,429 e um MAE de 0,853, sugerindo uma menor eficácia dessa arquitetura em capturar as informações relevantes para a estimativa de Brix ($^{\circ}$). Com o uso exclusivo dos índices de vegetação (b.2), o desempenho do RBF foi ainda mais fraco, com o R^2 ajustado caindo para 0,307 e o MAE aumentando para 0,893. Isso reflete uma performance de estimativa mais fraca em comparação com o MLP e com o uso das bandas multiespectrais. Quando foram combinados os dois conjuntos de variáveis, tanto as bandas quanto os índices (b.3), o RBF apresentou um R^2 ajustado de 0,438 e um MAE de 0,921, evidenciando que a adição de complexidade também não resultou em ganhos na estimativa para essa arquitetura.

Esses resultados indicam que, para a estimativa de Brix ($^{\circ}$), os modelos MLP apresentam uma performance superior em comparação aos modelos RBF, especialmente quando utilizam bandas multiespectrais ou índices de vegetação isoladamente. A combinação de variáveis, embora tenha aumentado a complexidade dos modelos, não necessariamente melhorou a precisão das estimativas, sugerindo que uma seleção cuidadosa das variáveis de entrada é essencial para otimizar a performance na estimativa dos modelos. Visualmente representados na Figura 15.

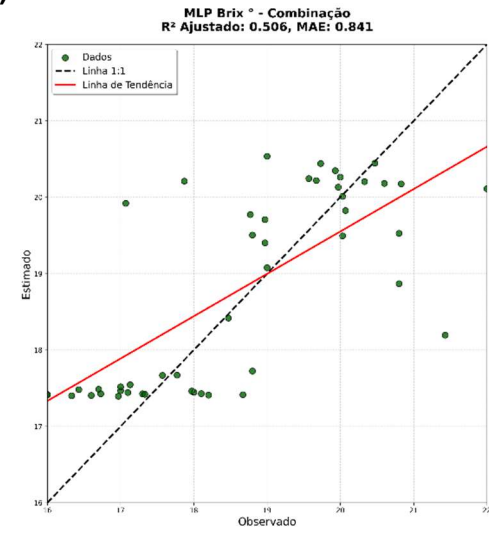
a.1)



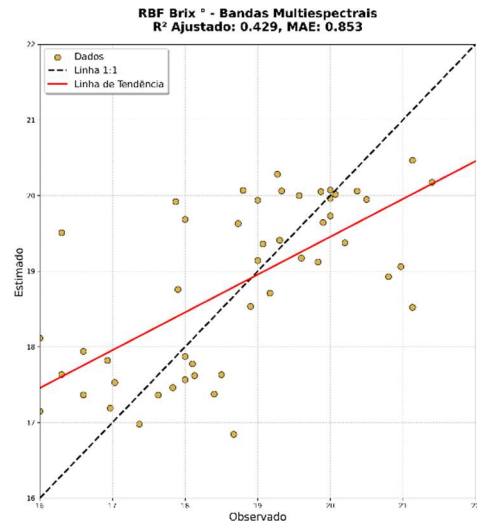
a.2)



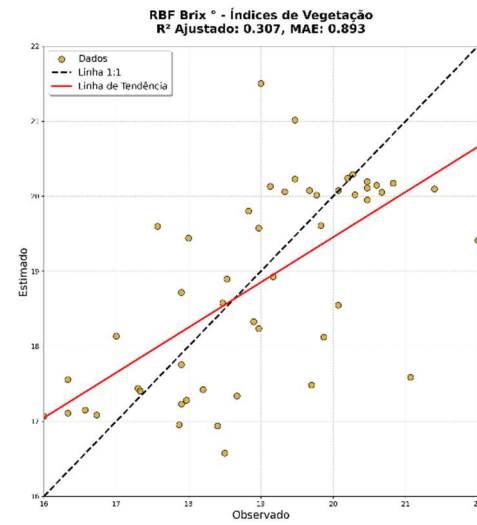
a.3)



b.1)



b.2)



b.3)

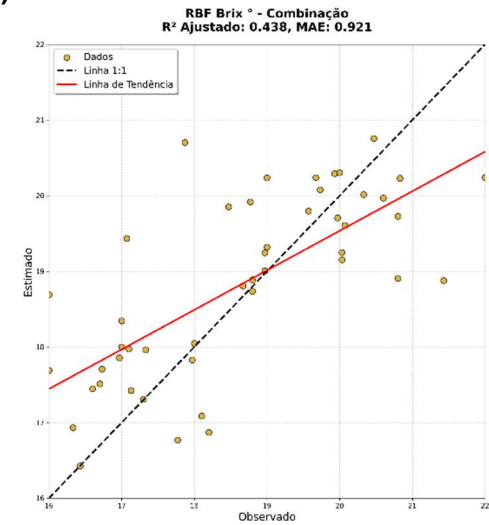


Figura 15 - Desempenho do teste na estimativa de Brix (°) utilizando as redes neurais artificiais RBF e MLP.

4.3.2. Comparação entre modelos RBF e MLP na estimativa de Pureza (%)

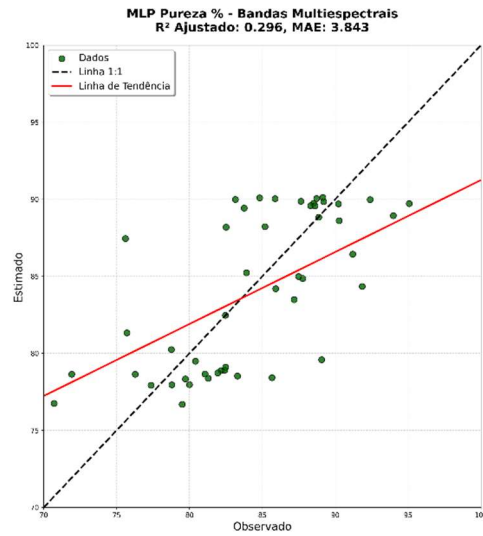
Ao analisar os resultados da estimativa de Pureza (%), utilizou-se uma abordagem semelhante, aplicando redes neurais RBF e MLP aos dados de entrada. Conforme apresentado na Figura 16, as dispersões entre os valores observados e estimados pelos modelos são evidentes, com a performance quantificada por meio de *MAE* e R^2 ajustado. Os resultados indicam que, em geral, os modelos tiveram um desempenho inferior na estimativa de Pureza (%) em comparação a Brix (°), especialmente os baseados em RBF, que mostraram maior dispersão e erros mais elevados.

Especificamente, ao utilizar bandas multiespectrais, o MLP (c.1) apresentou um R^2 ajustado de 0,296 e um *MAE* de 3,843, enquanto o RBF (d.1) obteve um R^2 ajustado de 0,000 e um *MAE* de 4,988, indicando que o RBF não conseguiu estimar Pureza (%) com as variáveis de entrada utilizadas. Quando foram usados apenas os índices de vegetação, o MLP (c.2) alcançou R^2 ajustado de 0,209 e *MAE* de 3,899, mas o RBF (d.2) novamente falhou em estimar Pureza (%), com R^2 ajustado de 0,000 e *MAE* de 5,288. A combinação de bandas e índices também não melhorou o desempenho do RBF (d.3), que manteve R^2 ajustado de 0,000 e *MAE* de 5,585, confirmando a incapacidade do RBF de estimar adequadamente Pureza (%) com os conjuntos de variáveis testados.

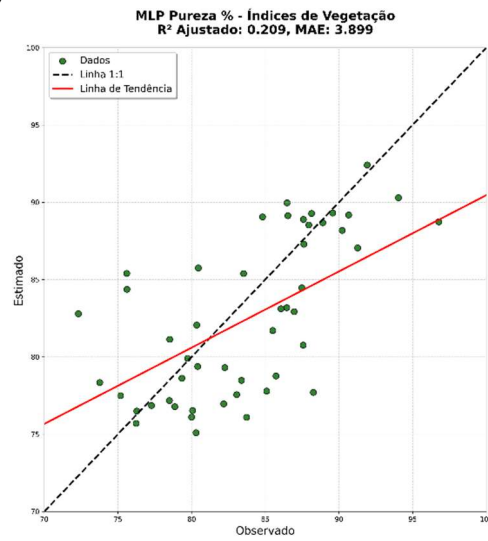
Por outro lado, o MLP mostrou-se mais robusto, embora ainda com limitações. A combinação de bandas e índices no MLP (c.3) resultou em um R^2 ajustado de 0,384 e *MAE* de 4,158, indicando uma superioridade em relação ao RBF, mas ainda evidenciando desafios na modelagem. Embora o MLP siga a direção geral dos dados observados, há variabilidade não capturada, sugerindo que melhorias adicionais são necessárias na modelagem ou na qualidade dos dados de entrada.

Esses resultados enfatizam que, para a estimativa de Pureza (%), os modelos RBF não foram eficazes com as variáveis de entrada utilizadas, independentemente de se tratar de bandas, índices ou uma combinação de ambos. Em contraste, o MLP, embora tenha apresentado um desempenho melhor, ainda enfrenta desafios significativos na estimativa de Pureza (%), destacando a complexidade inerente a essa métrica e a necessidade de refinamento nos métodos aplicados.

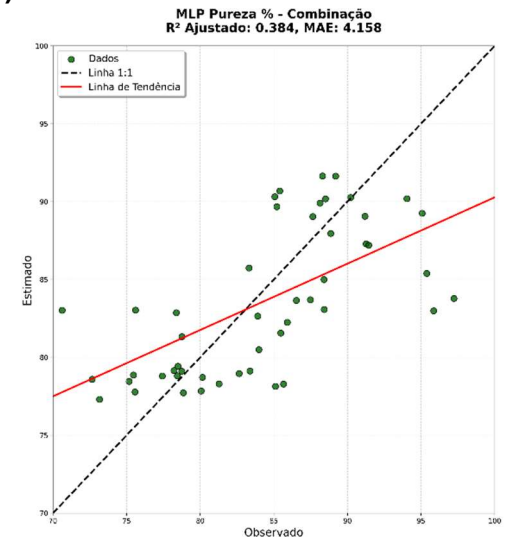
c.1)



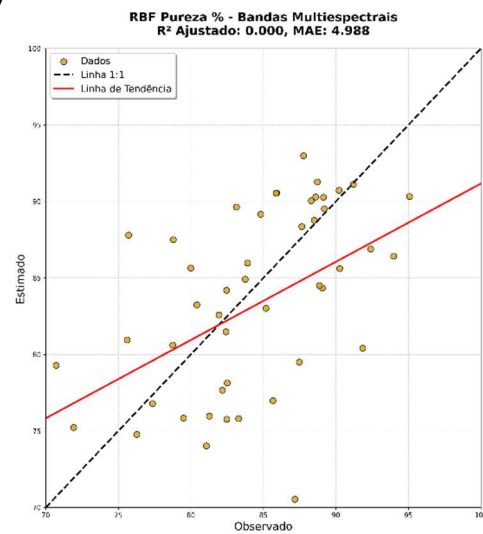
c.2)



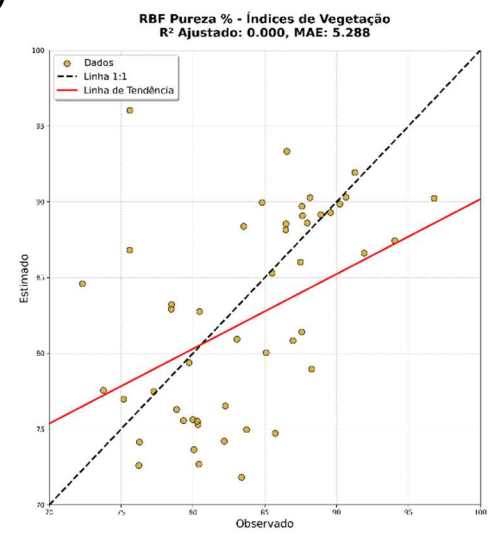
c.3)



d.1)



d.2)



d.3)

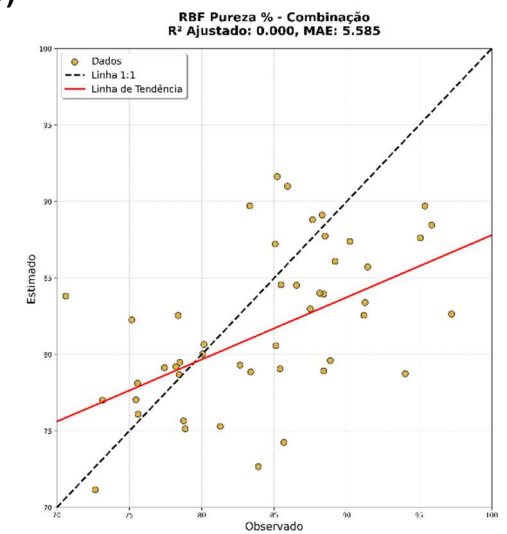


Figura 16 - Desempenho do teste na estimativa de Pureza (%) utilizando as redes neurais artificiais RBF e MLP.

4.4. Capacidade de Transferência e Generalização dos Modelos

Na Figura 17, explora-se a performance comparativa de dois modelos de redes neurais artificiais, Radial Basis Function (RBF) e Multilayer Perceptron (MLP), na estimativa de Brix (°) e Pureza (%) em cana-de-açúcar. A análise foi realizada em dois conjuntos de dados distintos: a Área 1, utilizada para treinamento e teste inicial dos modelos, e a Área 2, que serviu para avaliar a capacidade de transferência e generalização dos modelos para novos dados.

Os gráficos apresentados na Figura 17 comparam o Erro Médio Absoluto (MAE) obtido para cada modelo, utilizando diferentes variáveis de entrada, como bandas multiespectrais (BM), índices de vegetação (IVs) e em combinação, somados aos graus-dias acumulados (aGDD).

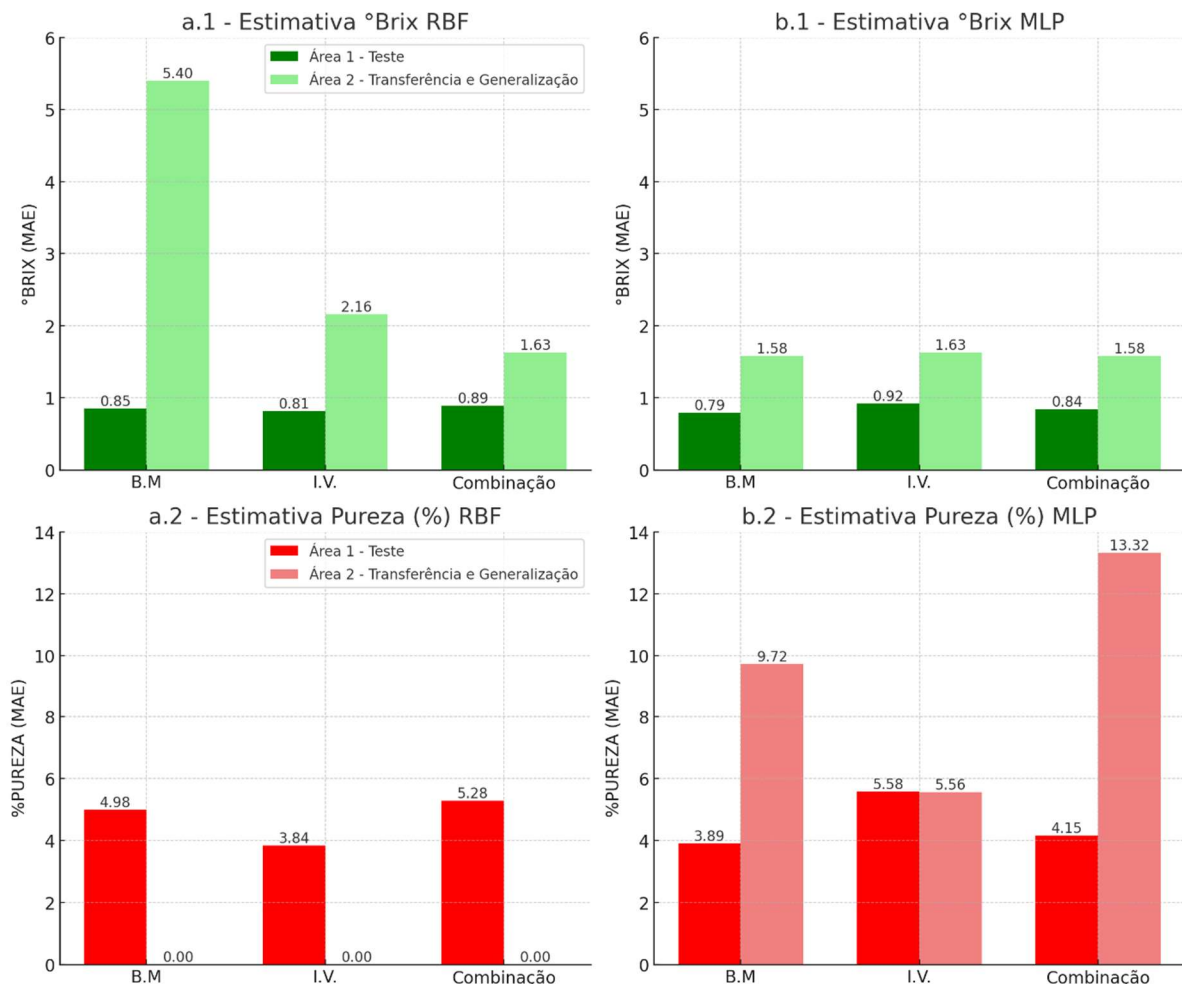


Figura 17 – Transferência e generalização dos modelos RBF e MLP na estimativa de Brix (°) e Pureza (%) em cana-de-açúcar através do Erro Médio Absoluto (MAE).

No gráfico (a.1), referente ao modelo RBF, o *MAE* registrado na Área 1 foi de 0,85 ao se utilizar a combinação de variáveis, o que corresponde a uma assertividade superior a 95% para valores de Brix acima de 18°. Esse resultado reflete a alta acurácia do modelo durante a etapa de teste. Na Área 2, que avaliou a capacidade de transferência e generalização do modelo, o *MAE* aumentou para 5,40 no uso exclusivo de bandas multiespectrais. Apesar desse aumento, a precisão do modelo RBF ainda se manteve acima de 70%, o que indica que, embora haja uma diminuição na performance, os erros permanecem dentro de uma margem aceitável para aplicações práticas no campo, especialmente em situações em que a variabilidade dos dados é maior.

Por outro lado, o modelo MLP, conforme ilustrado no gráfico (b.1), apresentou uma performance notavelmente mais consistente. O *MAE* na estimativa de Brix (°) se manteve baixo em ambos os conjuntos de dados, variando de 0,79 a 1,58, o que se traduz em uma assertividade superior a 90% mesmo na fase de transferência e generalização. Esses resultados evidenciam a robustez do MLP, que, devido à sua capacidade de generalização, oferece um desempenho estável e confiável, essencial para sua aplicação prática em cenários de produção agrícola.

Em relação à pureza, os gráficos (a.2) e (b.2) mostram um cenário interessante. No modelo RBF, inicialmente, o *MAE* foi de 4,98 na Área 1. Porém, ao avaliarmos a Área 2, verificamos que os valores do *MAE* são nulos, indicando que, embora o modelo tenha mostrado um erro baixo nos gráficos de dispersão, o R^2 ajustado foi igual a zero. Esse resultado significa que o modelo não conseguiu capturar a variabilidade dos dados com as variáveis utilizadas, tornando impossível a análise e interpretação dos resultados para estimativas de pureza na Área 2.

O MLP, por sua vez, apresentou um desempenho superior na estimativa de pureza, com o *MAE* variando de 3,89 na Área 1 a 13,32 na Área 2. Mesmo com esse aumento na Área 2, o modelo manteve uma assertividade próxima de 90% em todos os cenários para estimativas de pureza. Isso indica uma alta capacidade de adaptação a novas condições de dados, reafirmando a adequação do MLP para aplicações em práticas de agricultura de precisão, onde a consistência das estimativas é crucial para a tomada de decisões fundamentadas.

Embora ambos os modelos apresentem variações em seus desempenhos, os erros observados, mesmo nos piores cenários, são relativamente baixos e permitem que se obtenha uma alta assertividade nas estimativas de Brix (°) e Pureza (%). A robustez

demonstrada pelo modelo MLP, em particular, sugere que ele pode ser implementado com sucesso em ambientes de produção agrícola, contribuindo para a otimização da colheita e melhoria da qualidade da matéria prima.

Os resultados obtidos reforçam a importância da seleção criteriosa de modelos de redes neurais artificiais para aplicações em agricultura de precisão. O MLP, com sua capacidade de lidar com interações complexas e não-lineares entre variáveis, se destaca como uma ferramenta eficaz para estimativas confiáveis, essencial para a tomada de decisões baseada em dados no campo.

5. DISCUSSÃO

5.1. Dinâmica Temporal de Brix (°) e Pureza (%) na Colheita Inteligente

O ciclo fenológico da cana-de-açúcar envolve uma série de mudanças fisiológicas que afetam diretamente a qualidade da matéria-prima. À medida que a planta envelhece, açúcares se acumulam em diferentes partes, como folhas, caules e raízes, fundamentais para a manutenção de suas funções vitais. Nas fases mais avançadas do ciclo, esses açúcares são redistribuídos das folhas mais velhas para as partes mais novas da planta, garantindo o seu desenvolvimento contínuo. Entretanto, durante a fase de floração, ocorre uma redução no conteúdo total de açúcares, o que pode comprometer a qualidade do caldo extraído (Khan et al., 2022).

O uso de sensoriamento remoto combinado com redes neurais artificiais permite mapear essas variações ao longo do crescimento, fornecendo excelentes estimativas dos parâmetros de qualidade do caldo. A maior concentração de sólidos solúveis foi observada na terceira avaliação, indicando um ponto ótimo para a colheita. No entanto, com o início da floração, esses parâmetros tendem a diminuir, enquanto a planta se torna mais fibrosa e passa por um processo de isoporização. Essas alterações não apenas complicam a colheita, mas também aumentam os custos de produção devido à maior resistência mecânica durante o processamento (Silva, de Moraes and Molin, 2011; Poltroniere et al., 2021)

Na indústria sucroalcooleira, a qualidade da cana é um fator crucial para garantir a eficiência da colheita e do processamento. Indicadores como Brix (°) e Pureza (%) são fundamentais para agricultores e usinas avaliarem a qualidade da matéria-prima e maximizarem o rendimento. Idealmente, níveis elevados desses indicadores sugerem uma maior concentração da sacarose, resultando em um caldo de melhor qualidade. Contudo, a presença excessiva de minerais e açúcares não sacarose, como glicose e frutose, pode comprometer o processo de refino e dificultar o manejo da colheita (Khan et al., 2022).

Para uma colheita mecanizada eficiente, os valores de Brix (°) e Pureza (%) devem ser superiores a 18° e 85%, respectivamente. Embora a cana possa alcançar esses valores ideais antes ou durante a maturidade fisiológica, o crescimento contínuo da planta pode levar a uma diminuição gradual e não linear desses indicadores. Esse declínio ao longo do tempo apresenta um desafio significativo para indústrias e produtores, que precisam identificar o período ideal para a colheita a fim de otimizar a qualidade e a eficiência operacional e o retorno econômico.

Nesse cenário, a agricultura digital se destaca como uma solução eficaz. Por meio de dados digitais obtidos através do sensoriamento remoto e inteligência artificial, produtores rurais e a indústria podem tomar decisões informadas sobre o período ideal de colheita, ajustando suas práticas de manejo para maximizar a qualidade, o rendimento e o retorno econômico. Essa abordagem não apenas melhora a eficiência operacional, mas também contribui para a sustentabilidade do setor, ao reduzir mão de obra, custos excessivos e maximizar o uso dos recursos.

A aplicação de redes neurais artificiais para analisar a dinâmica temporal da qualidade da cana-de-açúcar oferece uma nova dimensão de eficiência e gestão de operações no campo. A combinação ao sensoriamento remoto, essa tecnologia possibilita ajustes imediatos nas estratégias de manejo e colheita, promovendo uma agricultura inteligente. Isso é fundamental para o avanço da agricultura digital e para a competitividade da indústria sucroalcooleira, que busca continuamente melhorar a produção de alimentos e combustíveis de forma sustentável.

5.2. Relação entre Características Espectrais e Rendimento Qualitativo

A pesquisa revela que as bandas multiespectrais, quando corretamente interpretadas, são um indicativo poderoso das condições fisiológicas das plantações de cana-de-açúcar. Por exemplo, variações no índice de Brix (°) podem ser detectadas pela análise de bandas específicas que são sensíveis à composição química das folhas.

Apresentamos que a análise das bandas multiespectrais Verde, Azul, Verde I e NIR são eficazes na detecção de variações nos índices de Brix (°). Isso se deve à sensibilidade dessas bandas à composição química e estrutural da cultura. Estudos realizados por Barbosa et al. (2023) evidenciam que as bandas Verde, Azul e NIR, quando incluídas nos modelos, melhoram a robustez de previsões para Brix (°). Bandas como Verde e Azul podem refletir mudanças na estrutura celular e no conteúdo de clorofila, o que indiretamente afeta a concentração de açúcar.

Quando usamos esses comprimentos de onda específicos, eles funcionam como um tipo de "ativador" para a clorofila, que é uma parte importante das folhas das plantas. Isso faz com que a clorofila reflita luz em uma faixa específica que podemos medir, ou até mesmo produza uma espécie de brilho, conhecido como fluorescência Rodrigues et al. (2020, 2022). Isso é importante porque nos ajuda a entender melhor como as plantas estão absorvendo e usando a luz, o que está relacionado com a quantidade de açúcar que produzem.

As bandas Verde, RedEdge, Amarelo e Vermelho podem melhorar a estimativa de Pureza na cana-de-açúcar, sendo as duas primeiras variáveis também evidenciadas nos modelos gerados por Barbosa et al. (2023). Conforme a cana-de-açúcar cresce, ela realiza mais fotossíntese, que é o processo pelo qual as plantas usam a luz do sol para crescer. Isso causa mudanças químicas nas partes das células das plantas chamadas cloroplastos. Essas mudanças incluem o aumento de açúcar nas folhas que vai para o caule e afetam como a clorofila funciona. Isso muda a aparência da planta, fazendo com que ela pareça mais ou menos verde (Chea et al., 2020a).

Desta forma, estabelece-se uma correlação com os valores qualitativos de rendimento. Essa correlação não apenas proporciona uma visão mais clara da fenologia da cultura, mas também permite identificar períodos de estresse vegetativo que poderiam prejudicar o rendimento qualitativo. A capacidade de capturar essas alterações bioquímicas por meio de imagens de satélite abre caminho para uma gestão agrícola mais eficiente. O monitoramento contínuo e a interpretação dessas características espectrais proporcionam uma base sólida para a tomada de decisões sobre práticas de manejo, irrigação e aplicação de insumos, visando a otimização do rendimento e da qualidade da matéria-prima.

5.3. Redes Neurais Artificiais na Estimativa de Sacarose: Brix (°) e Pureza (%)

Realizou-se a seleção das bandas multiespectrais e dos índices de vegetação disponíveis e gerados através das imagens orbitais, com o objetivo de identificar os atributos que mais contribuíam para a precisão e a acurácia dos modelos. Esse processo de seleção não apenas aprimorou o desempenho dos modelos, mas também destacou os parâmetros mais influentes na determinação dos índices de qualidade da cana-de-açúcar, como Brix (°) e Pureza (%).

A complexidade inerente aos modelos de aprendizado de máquina, como RBF e MLP, frequentemente levanta questões sobre a interpretação e a análise dos resultados. No contexto agrônomo, é fundamental que as estimativas geradas por esses modelos sejam compreensíveis e aplicáveis, para que os profissionais do setor possam adotar e

implementar essas tecnologias de forma eficaz. Embora esses modelos tenham se mostrado ferramentas poderosas para a estimativa dos teores de sacarose, refletidos nos valores de Brix (°) e Pureza (%), eles também apresentam desafios significativos. O ajuste fino dos parâmetros e a definição da arquitetura das redes neurais exigem uma compreensão profunda tanto da teoria do aprendizado de máquina quanto das características específicas da cultura da cana-de-açúcar.

As implicações práticas do uso desses modelos são vastas, oferecendo aos agricultores ferramentas avançadas para otimizar a colheita e o processamento da cana-de-açúcar. Além disso, este estudo abre novas oportunidades para futuras investigações, que possam explorar a integração de outras fontes de dados, como informações climáticas, características do solo, relevo e disponibilidade hídrica. A inclusão dessas variáveis adicionais pode enriquecer significativamente a capacidade preditiva dos modelos, permitindo uma análise mais holística e precisa das condições que influenciam os teores de Brix (°) e Pureza (%) na cana-de-açúcar.

Essa abordagem integrada não apenas aumentaria a acurácia das estimativas, mas também proporcionaria insights mais profundos sobre os fatores ambientais que afetam a qualidade da cana-de-açúcar. Assim, poderíamos potencializar ainda mais as práticas de agricultura de precisão, promovendo uma gestão agrícola mais informada e eficiente, que contribua para a sustentabilidade e a qualidade na produção de cana-de-açúcar.

5.4. Complexidade de Modelos e Aplicabilidade Prática: Uma visão Holística

Uma das questões centrais deste estudo foi navegar pela complexa interação entre a sofisticação dos modelos de redes neurais artificiais e sua aplicabilidade prática no campo. A agricultura de precisão, que se caracteriza por uma abundância de dados e pela necessidade de decisões em tempo real, apresenta um ambiente ideal para o emprego de tais modelos. Contudo, a implementação efetiva desses modelos requer um equilíbrio delicado entre a precisão das estimativas e a facilidade de interpretação.

Modelos altamente complexos podem oferecer uma precisão extraordinária, mas sua operação e interpretação podem se tornar desafiadoras, especialmente para profissionais que não possuem formação especializada em ciência de dados. Agricultores e agrônomos necessitam de justificativas claras e compreensíveis para as estimativas geradas, de modo que possam confiar e agir com base nessas informações. Portanto, este estudo enfatizou a necessidade de desenvolver modelos que não apenas capturem a complexidade dos dados agrícolas, mas que também sejam suficientemente transparentes para serem interpretados

e utilizados por não especialistas, como profissionais da indústria sucroalcooleira e empresas prestadoras de serviços.

A precisão dos modelos é, sem dúvida, crucial. No entanto, um modelo que seja altamente preciso, mas inaplicável em condições reais de campo, tem pouca utilidade prática. Assim, este trabalho investigou como a simplificação dos modelos pode ser realizada sem comprometer significativamente a precisão e a acurácia das estimativas, permitindo sua aplicação prática na estimativa de Brix (°) e Pureza (%) para a cana-de-açúcar. Esse compromisso entre complexidade e aplicabilidade prática é vital para a adoção da agricultura digital em larga escala nos campos agrícolas.

Os modelos desenvolvidos e testados neste estudo demonstram um potencial significativo para otimizar o uso de recursos, reduzir o desperdício e aumentar a eficiência da produção. Esses objetivos estão em consonância com os princípios de sustentabilidade, contribuindo para uma cadeia de suprimentos de alimentos mais resiliente e eficiente. Além disso, o estudo delinea um roteiro promissor para pesquisas futuras na interseção entre agronomia, ciência de dados e tecnologia de sensoriamento remoto.

A complexidade dos desafios enfrentados na produção agrícola moderna exige uma abordagem interdisciplinar. A colaboração entre agrônomos, cientistas de dados e especialistas em tecnologia é fundamental para acelerar o desenvolvimento de soluções que sejam tecnicamente viáveis e prontamente adotáveis pelos produtores rurais. A integração dessas disciplinas pode não apenas melhorar a eficácia das soluções desenvolvidas, mas também facilitar sua implementação prática, garantindo que as inovações tecnológicas beneficiem efetivamente a agricultura em um cenário global cada vez mais desafiador.

6. CONCLUSÕES

Este estudo confirma a relevância de selecionar adequadamente variáveis espectrais e índices de vegetação na modelagem de Brix (°) e Pureza (%) da cana-de-açúcar, ambos fundamentais para a qualidade da matéria-prima para a indústria. As bandas *Verde; Azul; Verde e NIR*, em conjunto com o aGDD, mostraram-se eficazes na estimativa de Brix (°), enquanto a Pureza (%) foi melhor estimada por meio do aGDD e das bandas *Verde; RedEdge; Amarelo e Vermelho*.

A análise também revelou que, ao combinar dados espectrais com índices de vegetação, as variáveis-chave para Brix (°) e Pureza (%) foram consistentes, destacando-

se as bandas *Vermelho; Verde; ExG; Azul e SFDVI*. Isso sugere que uma abordagem integrada pode ser vantajosa para monitorar esses parâmetros simultaneamente.

A implementação dessas variáveis em modelos de inteligência artificial, especialmente nas redes neurais MLP, demonstrou uma capacidade robusta de generalização e eficiência no monitoramento da qualidade da cana-de-açúcar, além de oferecer potencial para otimização das operações agrícolas. No entanto, reconhece-se que há margem para melhorias, como a inclusão de variáveis adicionais relacionadas ao solo, clima e disponibilidade hídrica, o que pode aumentar a precisão dos modelos.

Em conclusão, esta pesquisa contribui para o avanço da agricultura de precisão e agricultura digital oferecendo soluções inovadoras que podem ser amplamente aplicadas e adaptadas para enfrentar os desafios da produção agrícola moderna. A integração de novos dados e o desenvolvimento contínuo de modelos de inteligência artificial proporcionarão uma base sólida para tomadas de decisões baseada em dados, assegurando que a qualidade da cana-de-açúcar seja aprimorada, beneficiando toda a cadeia de produção, desde o campo até a indústria.

7. REFERÊNCIAS

Abd El-Ghany NM, Abd El-Aziz SE and Marei SS (2020) A review: application of remote sensing as a promising strategy for insect pests and diseases management. **Environmental Science and Pollution Research** 27:33503–33515.

Bacchi, Osny OS and JAGCS (1978) Minimum threshold temperature for sugarcane growth. **INTERNATIONAL SOCIETY SUGAR CANE TECHNOLOGISTS** 16 1733–1741.

Bahn RA, Yehya AAK and Zurayk R (2021) Digitalization for Sustainable Agri-Food Systems: Potential, Status, and Risks for the MENA Region. **Sustainability** 13:3223.

Baptista GMM (2015) Aplicação do Índice de Vegetação por Profundidade de Feição Espectral (SFDVI - Spectral Feature Depth Vegetation Index) em dados RapidEye. in **Anais XVII Simpósio Brasileiro de Sensoriamento Remoto - SBSR** 2277–2284.

Barbosa Júnior MR, Moreira BR de A, de Brito Filho AL, Tedesco D, Shiratsuchi LS and da Silva RP (2022) UAVs to Monitor and Manage Sugarcane: Integrative Review. **Agronomy**. MDPI.

Barbosa MR, Moreira BR de A, de Oliveira RP, Shiratsuchi LS and da Silva RP (2023) UAV imagery data and machine learning: A driving merger for predictive analysis of qualitative

yield in sugarcane. **Frontiers in Plant Science** 14.

Bégué A, Lebourgeois V, Bappel E, Todoroff P, Pellegrino A, Baillarin F and Siegmund B (2010) Spatio-temporal variability of sugarcane fields and recommendations for yield forecast using NDVI. **International Journal of Remote Sensing** 31:5391–5407.

Benos L, Tagarakis AC, Dolias G, Berruto R, Kateris D and Bochtis D (2021) Machine Learning in Agriculture: A Comprehensive Updated Review. **Sensors** 21:3758.

Bolfe ÉL, Jorge LA de C, Sanches ID, Júnior AL, Costa CC da, Victoria D de C, Inamasu RY, Grego CR, Ferreira VR and Ramirez AR (2020) Precision and Digital Agriculture: Adoption of Technologies and Perception of Brazilian Farmers. **Agriculture** 2020, Vol. 10, Page 653 10:653.

Broge NH and Leblanc E (2001) Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. **Remote Sensing of Environment** 76:156–172.

CEMA (2017) *Digital Farming: what does it really mean?* **European Agricultural Machinery**. Available at: <https://www.cema-agri.org/publication/position-papers/254-digital-farming-what-does-it-really-mean> (Accessed: 6 December 2023).

Chea C, Saengprachatanarug K, Posom J, Wongphati M and Taira E (2020a) Sugar Yield Parameters and Fiber Prediction in Sugarcane Fields Using a Multispectral Camera Mounted on a Small Unmanned Aerial System (UAS). **Sugar Tech** 22:605–621.

Chea C, Saengprachatanarug K, Posom J, Wongphati M and Taira E (2020b) Sugar Yield Parameters and Fiber Prediction in Sugarcane Fields Using a Multispectral Camera Mounted on a Small Unmanned Aerial System (UAS). **Sugar Tech** 22:605–621.

Cherubin MR, Damian JM, Tavares TR, Trevisan RG, Colaço AF, Eitelwein MT, Martello M, Inamasu RY, Pias OH de C and Molin JP (2022) Precision Agriculture in Brazil: The Trajectory of 25 Years of Scientific Research. **Agriculture (Switzerland)** 12.

Cohen WB (1991) Response of vegetation indices to changes in three measures of leaf water stress. **Photogrammetric Engineering & Remote Sensing** 57:195–202. Available at: [https://www.researchgate.net/profile/Kamel-Nagaz/post/How-to-calculate-Leaf-Water-Content-Index-](https://www.researchgate.net/profile/Kamel-Nagaz/post/How-to-calculate-Leaf-Water-Content-Index-LWCI/attachment/5e7205c1cfe4a7d888e33716/AS%3A870399755702272%401584530881141/download/1991_feb_195-202.pdf)

[LWCI/attachment/5e7205c1cfe4a7d888e33716/AS%3A870399755702272%401584530881141/download/1991_feb_195-202.pdf](https://www.researchgate.net/profile/Kamel-Nagaz/post/How-to-calculate-Leaf-Water-Content-Index-LWCI/attachment/5e7205c1cfe4a7d888e33716/AS%3A870399755702272%401584530881141/download/1991_feb_195-202.pdf) (Accessed: 29 December 2023).

CONAB (2024) *Produção de cana-de-açúcar e subprodutos*. **Companhia Nacional de Abastecimento**. Available at: <https://www.conab.gov.br/info-agro/safras/cana/boletim-da-safradecana-de-acucar> (Accessed: 8 September 2023).

CONSECANA (2006) *Manual CONSECANA-SP. Conselho dos Produtores de Cana-de-Açúcar, Açúcar e Álcool do Estado de São Paulo*. Available at: [https://edisciplinas.usp.br/pluginfile.php/3116035/mod_resource/content/1/Manual do Consecana.pdf](https://edisciplinas.usp.br/pluginfile.php/3116035/mod_resource/content/1/Manual%20do%20Consecana.pdf) (Accessed: 28 October 2023).

Eastwood C, Klerkx L, Ayre M and Dela Rue B (2019) Managing Socio-Ethical Challenges in the Development of Smart Farming: From a Fragmented to a Comprehensive Approach for Responsible Research and Innovation. **Journal of Agricultural and Environmental Ethics** 32:741–768.

Epifanio HR (2023) A importância da produção de cana-de-açúcar no Brasil: Revisão sobre os aspectos sócio-econômico-ambientais do processo produtivo do açúcar no Brasil.

FAOSTAT (2022) *Production quantities of Sugar cane by country*. **Crops and livestock products**. Available at: <https://www.fao.org/faostat/en/#data/QCL/visualize>.

Ferreira LG (2008) Sensoriamento remoto da vegetação. **Acta Scientiarum. Biological Sciences**.

Formaggio AR, Formaggio RA and Sanches ID (2017) *Sensoriamento Remoto em Agricultura*.

Gitelson AA, Kaufman YJ, Stark R and Rundquist D (2002) Novel algorithms for remote estimation of vegetation fraction. **Remote Sensing of Environment** 80:76–87.

Gitelson AA, Gritz Y and Merzlyak MN (2003) Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. **Journal of plant physiology** 160:271–282.

Gitelson AA, Kaufman YJ and Merzlyak MN (1996) Use of a green channel in remote sensing of global vegetation from EOS-MODIS. **Remote Sensing of Environment** 58:289–298.

Giua C, Materia VC and Camanzi L (2021) Management information system adoption at the farm level: evidence from the literature. **British Food Journal** 123:884–909.

Hermonn W (2008) Precision farming. **GIS-Business** 37–39.

Hithamani G, Medappa H, Chakkaravarthi A, Ramalakshmi K and Raghavarao KSMS (2018) Effect of adsorbent and acidulants on enzymatic browning of sugarcane juice. **Journal of**

Food Science and Technology 55:4356–4362.

Huete A, Justice C and Liu H (1994) Development of vegetation and soil indices for MODIS-EOS. **Remote Sensing of Environment** 49:224–234.

J Jasper and S Reush (2009) Active sensing of the N status of wheat using optimized wavelength combination: impact of seed rate, variety and growth stage. **Precision agriculture '09** [Preprint].

Jiang Z, Huete AR, Didan K and Miura T (2008) Development of a two-band enhanced vegetation index without a blue band. **Remote Sensing of Environment** 112:3833–3845.

Kataoka T, Kaneko T, Okamoto H and Hata S (2003) Crop growth estimation system using machine vision. in **IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM**. Institute of Electrical and Electronics Engineers Inc. 1079–1083.

Khan Q, Qin Y, Guo DJ, Zeng XP, Chen JY, Huang YY, Ta QK, Yang LT, Liang Q, Song XP, Xing YX and Li YR (2022) Morphological, agronomical, physiological and molecular characterization of a high sugar mutant of sugarcane in comparison to mother variety. **Plos One** 17:e0264990.

Kujawa S and Niedbała G (2021) Artificial neural networks in agriculture. **Agriculture (Switzerland)**. Multidisciplinary Digital Publishing Institute 497.

Laudon KC and Laudon JP (2018) *Management Information Systems: Managing the Digital Firm*.

Leblanc SC, Chen JM and Cihlar J (1997) Ndvi directionality in boreal forests: A model interpretation of measurements. **Canadian Journal of Remote Sensing** 23:369–380.

Liakos KG, Busato P, Moshou D, Pearson S and Bochtis D (2018) Machine learning in agriculture: A review. **Sensors (Switzerland)**. Multidisciplinary Digital Publishing Institute 2674.

Louhaichi M, Borman MM and Johnson DE (2001) Spatially located platform and aerial photography for documentation of grazing impacts on wheat. **Geocarto International** 16:65–70.

Maier HR and Dandy GC (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. **Environmental Modelling & Software** 15:101–124.

Martin JP, Abbott E V. and Hughes CG (1961) Sugar-cane diseases of the world. Vol. 1.

Sugar-cane diseases of the world. Vol. 1. [Preprint].

Meneses PR, Almeida T de and Baptista GM de M (2019) Reflectância dos materiais terrestres. in.

Misra V, Mall AK, Ansari SA and Ansari MI (2023) Sugar Transporters, Sugar-Metabolizing Enzymes, and Their Interaction with Phytohormones in Sugarcane. **Journal of Plant Growth Regulation**. Springer 4975–4988.

Naranjo-Torres J, Mora M, Hernández-García R, Barrientos RJ, Fredes C and Valenzuela A (2020) A Review of Convolutional Neural Network Applied to Fruit Image Processing. **Applied Sciences** 10:3443.

de Oliveira RP, Barbosa Júnior MR, Pinto AA, Oliveira JLP, Zerbato C and Furlani CEA (2022) Predicting Sugarcane Biometric Parameters by UAV Multispectral Images and Machine Learning. **Agronomy** 12.

Poltroniere SC, Aliano Filho A, Caversan AS, Balbo AR and Florentino H de O (2021) Integrated planning for planting and harvesting sugarcane and energy-cane for the production of sucrose and energy. **Computers and Electronics in Agriculture** 184:105956.

Rodrigues M, Nanni MR, Cezar E, dos Santos GLAA, Reis AS, de Oliveira KM and de Oliveira RB (2020) Vis–NIR spectroscopy: from leaf dry mass production estimate to the prediction of macro- and micronutrients in soybean crops. **Journal of Applied Remote Sensing** 14.

Rodrigues M, Cezar E, Abrantes dos Santos GLA, Reis AS, Furlanetto RH, de Oliveira RB, D'Àvila RC and Nanni MR (2022) Estimating technological parameters and stem productivity of sugarcane treated with rock powder using a proximal spectroradiometer Vis-NIR-SWIR. **Industrial Crops and Products** 186:115278.

Rondeaux G, Steven M and Baret F (1996) Optimization of soil-adjusted vegetation indices. **Remote Sensing of Environment** 55:95–107.

Rotz S, Gravely E, Mosby I, Duncan E, Finnis E, Horgan M, LeBlanc J, Martin R, Neufeld HT, Nixon A, Pant L, Shalla V and Fraser E (2019) Automated pastures and the digital divide: How agricultural technologies are shaping labour and rural communities. **Journal of Rural Studies** 68:112–122.

Rouse JW, Haas RH, Schell JA, Deering DW and Harlan JC (1974) Monitoring the vernal advancement of natural vegetation. **Final report, NASA/GCSFC, Greenbelt, MD** 6.

Shang L, Heckelei T, Gerullis MK, Börner J and Rasch S (2021a) Adoption and diffusion of

digital farming technologies - integrating farm-level evidence and system interaction. **Agricultural Systems** 190:103074.

Shang L, Heckelei T, Gerullis MK, Börner J and Rasch S (2021b) Adoption and diffusion of digital farming technologies - integrating farm-level evidence and system interaction. **Agricultural Systems** 190:103074.

Silva CB, de Moraes MAFD and Molin JP (2011) Adoption and use of precision agriculture technologies in the sugarcane industry of São Paulo state, Brazil. **Precision Agriculture** 12:67–81.

Souza JBC, De Almeida SLH, De Brito de Filho AL, Monteles MS, Silva LB and Da Silva RP (2021) Uso de redes neurais para predição da maturação do amendoim com imagens de UAV. **South American Sciences** 2:e21129.

Souza JBC, de Almeida SLH, Freire de Oliveira M, Santos AF dos, Filho AL de B, Meneses MD and Silva RP da (2022) Integrating Satellite and UAV Data to Predict Peanut Maturity upon Artificial Neural Networks. **Agronomy** 12:1512.

Sponchioni G, Vezzoni M, Bacchetti A, Pavesi M and Renga FM (2020) *The 4.0 revolution in agriculture: a multi-perspective definition*. ME -.

Sreedevi P, Jayachandran LE and Rao PS (2018) Browning and bioactive composition of sugarcane juice (*Saccharum officinarum*) as affected by high hydrostatic pressure processing. **Journal of Food Measurement and Characterization** 12:1962–1971.

Susanti R, Nofendra R, Zaini Z, Bin Suhaimi MSA and Rusydi MI (2023) The Use of Artificial Neural Networks in Agricultural Plants. **Andalas Journal of Electrical and Electronic Engineering Technology** 2:62–68.

Tedesco D, Almeida Moreira BR de, Barbosa Júnior MR, Papa JP and Silva RP da (2021) Predicting on multi-target regression for the yield of sweet potato by the market class of its roots upon vegetation indices. **Computers and Electronics in Agriculture** 191.

Thorburn PJ, Biggs JS, Palmer J, Meier EA, Verburg K and Skocaj DM (2017) Prioritizing crop management to increase nitrogen use efficiency in australian sugarcane crops. **Frontiers in Plant Science** 8:272232.

Wolfert S, Ge L, Verdouw C and Bogaardt M-J (2017) Big Data in Smart Farming – A review. **Agricultural Systems** 153:69–80.

Yang J, Ching YC and Chuah CH (2019) Applications of Lignocellulosic Fibers and Lignin in

Bioplastics: A Review. **Polymers** 2019, Vol. 11, Page 751 11:751.

Zarco-Tejada PJ, Berjón A, López-Lozano R, Miller JR, Martín P, Cachorro V, González MR and De Frutos A (2005) Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. **Remote Sensing of Environment** 99:271–287.