



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO - UFRPE
UNIDADE ACADÊMICA DE SERRA TALHADA – UAST
BACHARELADO EM AGRONOMIA

Gabriel Cavalcanti de Moura

**Estimativa de Biomassa e índice de Área Foliar de BRS
Capiáu com aplicação de machine learning e sensor RGB em
veículo aéreo não tripulado (VANT)**

Serra Talhada-PE

2025

Gabriel Cavalcanti de Moura

**Estimativa de Biomassa e índice de Área Foliar de BRS
Capiáu com aplicação de machine learning e sensor RGB em
veículo aéreo não tripulado (VANT)**

Trabalho de Conclusão de Curso apresentado
na Unidade Acadêmica de Serra Talhada da
Universidade Federal Rural de Pernambuco,
como requisito parcial para conclusão do
Curso de Bacharelado em Agronomia.

Orientador (a): Dr. Alan Cezar Bezerra

Serra Talhada – PE

2025

Dados Internacionais de Catalogação na Publicação
Sistema Integrado de Bibliotecas da UFRPE
Bibliotecário(a): Suely Manzi – CRB-4 809

M929e Moura, Gabriel Cavalcanti de.

Estimativa de biomassa e índice de área foliar de BRS capiaçu com aplicação de *machine learning* e sensor RGB em veículo aéreo não tripulado (VANT) / Gabriel Cavalcanti de Moura. – Serra Talhada, 2025.

33 f.; il.

Orientador(a): Alan Cezar Bezerra.

Trabalho de Conclusão de Curso (Graduação) – Universidade Federal Rural de Pernambuco, Unidade Acadêmica Serra Talhada - UAST, Bacharelado em Agronomia, Serra Talhada, BR-PE, 2025.

Inclui referências.

1. Sensoriamento remoto. 2. Aprendizado do computador. 3. Planta forrageira. 4. Biomassa 5. Agronomia - Estudo e ensino. I. Bezerra, Alan Cezar, orient. II. Título

CDD 630

Gabriel Cavalcanti de Moura

**Estimativa de Biomassa e índice de Área Foliar de BRS
Capiáu com aplicação de machine learning e sensor RGB em
veículo aéreo não tripulado (VANT)**

**Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado
em Agronomia da Universidade Federal Rural de Pernambuco, Unidade
Acadêmica de Serra Talhada, como requisito parcial para obtenção do
título de Bacharel em Agronomia.**

Aprovado em: 08/08/2025

BANCA EXAMINADORA

Prof. Dr. Alan Cezar Bezerra – UAST/UFRPE
Orientador

Dr. Wagner Martins dos Santos – PGPV/UFRPE
(Examinador Interno)

Dr. José Raliuson Inácio Silva – PGPV/UFRPE
(Examinador Interno)

Agradecimentos

Primeiramente externo minha gratidão a Deus, pois desde o princípio de minha vida esteve comigo e tem me guiado até o presente momento. Acredito que tudo provém dele e que foi ele que me trouxe a UAST e me sustentou durante toda minha graduação.

Aos meus pais, dedico todo o fruto de meu esforço. Antônia Lacerda Cavalcanti Moura e Aluisio Pereira de Moura, vocês são os pilares da minha vida, o motivo e a motivação para chegar até aqui. Sou quem sou hoje, por que vocês me ensinaram a ser. Mãe, obrigado pelos incentivos constantes e por me ensinar a ser forte. Pai, obrigado por sempre ter uma palavra positiva. Nunca me esquecerei de vocês. Os amo.

A minha segunda mãe, pessoa que se importa e cuida de mim desde que eu era criança, sem que eu fosse seu filho. Sempre me incentivou, e principalmente sempre acreditou em mim. Me ajudou de diversas formas, me ajudou inclusive durante a graduação. Sou muito grato por ter a senhora em minha vida. Amo te.

Ao meu orientador, estimado Alan Cezar Bezerra, lembro como hoje o dia que lhe procurei e você me encheu de ideias, instigando minha vontade e curiosidade para diversas possibilidades de pesquisa científica. Agradeço por isso, por acreditar em mim e pela boa orientação. Me ajudou muito no crescimento acadêmico

A minha amiga Edilene Daniel, pessoa que falei abertamente, que sempre me incentivou e sempre teve uma palavra de confiança. Agradeço a Deus por te conhecer, você é uma pessoa espetacular. Sem dúvida nenhuma, uma voz de maturidade. Obrigado pelo acolhimento quando estive disperso, pelos incentivos, pela sua ajuda e amizade.

Às queridas amigas Raphaela, Ana Clara e Jucelia. Pelos bons momentos, pelo acolhimento e pelas boas gargalhadas. Foi muito bom conhecer e conviver com vocês.

A Universidade Federal Rural de Pernambuco, Unidade acadêmica de Serra Talhada, o meu muito obrigado, pela oportunidade de desenvolver aprendizados na Agronomia e no convívio com colegas e professores na unidade.

Externo minha gratidão a todos os professores que fizeram parte da minha formação e crescimento pessoal. Pelos bons conselhos, as experiências transmitidas e o conhecimento passado.

Resumo

A criação de ruminantes é fundamental para a agricultura brasileira, tanto no agronegócio quanto na agricultura familiar. A pecuária brasileira é baseada em pastagens, fontes de nutrientes essenciais. O capim-elefante, em especial a cultivar BRS Capiáçu, se destaca pela alta produtividade e valor nutricional. Essa cultivar é amplamente utilizada para silagem e capineira, requerendo manejo eficiente. VANTs permitem o monitoramento remoto da forragem, utilizando imagens para avaliar características das plantas. A fotogrametria auxilia na extração de índices de vegetação, ajudando a estimar biomassa e qualidade do pasto. Além disso, técnicas de Machine Learning permitem criar modelos preditivos para otimizar esse processo. Sendo assim, esse trabalho objetiva estimar através de imagens de sensor RGB acoplado a veículo aéreo não tripulado (VANT), Biomassa e Índice de Área Foliar da BRS Capiáçu com algoritmos de aprendizagem de máquina (machine learning). De maneira que foram realizadas as modelagens com os algoritmos *Random Forest* (RF) e o *Support Vector Machines* (SVM). Sobre o banco composto pelos dados espectrais obtidos a partir de sensor RGB, bandas (*Red, Green, Blue*) e 15 Índices de Vegetação (IVs). Como também pelos dados de campo aferidos na área de BRS Capiáçu, Biomassa e Índice Área Foliar (IAF). Com o intuito de aprimorar a performance dos algoritmos, procedeu-se com uma seleção de variáveis pós-modelagem. As correlações não apresentaram forte relação com as variáveis de interesse. A performance dos algoritmos foi semelhante para as duas variáveis de interesse, com ligeiras diferenças entre si. A seleção de variáveis pós-modelagem também não melhorou a eficiência da maioria dos modelos, pelo contrário, piorou as métricas desses. Com pequeno incremento apenas para Biomassa com SVM. Os modelos de RF e SVM não produziram ajuste satisfatório para estimar a Biomassa e o IAF. Provavelmente o banco de dados foi inconclusivo para o aprendizado dos modelos. Pelo *underfitting* que ficou nítido para a Biomassa com SVM, e pela dificuldade do sensor RGB em captar a assinatura espectral da cultura. O uso de sensores multiespectrais alinhada ao aumento do volume de dados biofísicos, podem contribuir com esse tipo de modelagem na BRS Capiáçu futuramente.

Palavras-chave: Sensoriamento Remoto, Aprendizado do computador e Planta forrageira.

Abstract

Ruminant livestock production is essential to Brazilian agriculture, both in agribusiness and family farming. Brazilian livestock farming is predominantly based on pasture, sources of essential nutrients. Elephant grass, particularly the BRS Capiáu cultivar, stands out for its high productivity and nutritional value. This cultivar is widely used for silage and cut-and-carry systems, requiring efficient management. Unmanned Aerial Vehicles (UAVs) enable remote monitoring of forage by using imagery to assess plant characteristics. Photogrammetry helps in extracting vegetation indices, which aids in estimating biomass and pasture quality. Furthermore, Machine Learning techniques allow the development of predictive models to optimize this process. Therefore, this study aims to estimate the biomass and leaf area index (LAI) of BRS Capiáu using RGB sensor imagery collected by UAVs and machine learning algorithms. Modeling was performed using the *Random Forest* (RF) and *Support Vector Machines* (SVM) algorithms on a dataset composed of spectral data from the RGB sensor bands (Red, Green, Blue) and 15 vegetation indices (VIs), as well as field data collected from the BRS Capiáu plots (biomass and LAI). To enhance algorithm performance, a post-modeling variable selection was conducted. Correlation analyses did not reveal strong relationships with the target variables. The performance of both algorithms was similar for the two variables of interest, with only slight differences. Post-modeling variable selection did not improve the efficiency of most models; on the contrary, it worsened their metrics, with only a slight improvement observed for biomass estimation with SVM. RF and SVM models did not yield satisfactory fits for estimating biomass and LAI, likely due to the dataset being inconclusive for model training. Evidence of underfitting was observed for biomass estimation using SVM, and the RGB sensor exhibited limitations in capturing the crop's spectral signature. The use of multispectral sensors combined with increased biophysical data volume may contribute to improved modeling for BRS Capiáu in future studies.

Keywords: Remote sensing, Machine learning, and Forage crops.

Sumário

1. Introdução.....	8
2. Objetivo.....	10
2.1. Geral.....	10
2.2. Específicos.....	10
3. Fundamentação Teórica.....	10
3.1. BRS Capiaçu no contexto da criação animal.....	10
3.2. Monitoramento de plantas forrageiras com sensoriamento remoto.....	11
3.3. Modelagem com aplicação de Machine Learning (ML).....	12
4. Metodologia.....	14
4.1. Caracterização da área de estudo.....	14
4.2. Dados espectrais.....	16
4.3. Dados de campo.....	19
4.4. Modelagem sobre o banco de dados.....	20
4.5. Avaliação dos modelos.....	20
5. Resultados e Discussões.....	21
5.1. Análise descritiva dos dados.....	21
5.2. Análise da modelagem.....	22
6. Conclusões.....	27
Bibliografia.....	27

1. Introdução

A criação de animais ruminantes desempenha um papel importante na agricultura brasileira, seja na vertente do agronegócio, um dos principais setores da economia nacional, que esboça uma produção ostensiva que visa a comercialização de seus produtos em mercado internacional, principalmente da carne bovina, visto que segundo o IBGE (2025), em 2024, 39,27 milhões de cabeças de bovinos foram abatidas, o que representa um aumento de 15,2% frente a 2023, movimento que segue à tendência de crescimento observada em 2022. Na agricultura familiar, em geral, de menor porte, que visa a subsistência, por meio dos produtos gerados, tanto no consumo direto, como pela comercialização, responsável pelo abastecimento do mercado local.

Plantas forrageiras são cruciais no contexto da agricultura, em especial na atividade pecuária, pelo fato de constituírem em alimento indispensável na dieta de ruminantes. Lisboa (2020) afirma que, a base alimentar da pecuária brasileira é a pasto, e as espécies forrageiras, apresentam um banco de elementos essenciais à nutrição de animais ruminantes, entre eles minerais, fibras e vitaminas. Outra alternativa é a utilização de espécies de grande porte que apresentam alta produtividade e teor de massa seca.

O capim-elefante (*Pennisetum purpureum* Schumach), por exemplo, se destaca entre as gramíneas, por ser altamente produtivo. Diante disso, a Embrapa Gado de Leite desenvolveu a BRS Capiapu, cultivar de porte alto, caracterizada pelo seu valor nutricional, elevada capacidade de produção de biomassa e considerável tolerância ao estresse hídrico (Pereira *et al.*, 2016)

A BRS Capiapu é amplamente utilizada como capineira ou produção de silagem, assim, para o uso da forragem, o manejo eficiente é o caminho a ser seguido objetivando a máxima eficiência, e os melhores resultados da produção animal. Nesse sentido, o conhecimento da massa de forragem a ser colhida, e a aferição de seu potencial produtivo, é uma estratégia promissora.

Neste contexto os Veículos Aéreos não tripulados (VANTs) ou aeronaves remotamente pilotadas (ARP), são ferramentas eficientes aplicadas à coleta de dados, através de imagens de alta qualidade, de forma remota (Magalhães *et al.*, 2013). A fotogrametria com esse equipamento possibilita o

desenvolvimento de metodologias de monitoramento capazes de proporcionar o entendimento de diversas características de cultivos agrícolas. A extração de índices de vegetação a partir das imagens coletadas é uma dessas possibilidades e em plantas forrageiras esses índices se associam às aferições de biomassa, altura do pasto e teores de nitrogênio (XUE e SU, 2017; TONG *et al.*, 2019)

O *Machine Learning* (ML) é um braço da Inteligência Artificial (IA), que é capaz de explorar uma base de dados e criar modelos de predições e prescrições aplicáveis a diversas situações. A partir do aprendizado de uma base de dados (dados de treino), que possui uma série de informações sobre um objeto de estudo. Essas informações são denominadas de *features* ou variáveis (Liakos *et al.*, 2018). Segundo Jung *et al.* (2021), é possível obter um grande número de modelos preditivos e prescritivos, com aplicação dos algoritmos de ML.

Entre os diferentes algoritmos ML, existe o *Random Forest* (RF) que trabalha com árvores de decisão binárias, de entradas aleatórias, buscando a máxima exploração dos dados (Bulagang *et al.*, 2020). E o SVM, que modela buscando um hiperplano capaz de maximizar a distância de separação entre diferentes classes de dados. Apresentando elevada eficiência para aprendizado com grandes densidades de dados e com massiva aplicação em trabalhos de classificação binária e regressão (Godoy e Resende, 2024).

Com isso, o presente trabalho visa aprimorar o caminho para a aferição não destrutiva de IAF e biomassa na cultura, através da aplicação de uma estrutura de modelagem de dados moderna e robusta, que é o aprendizado de máquina. Avaliando seu real potencial para determinação de características biofísicas na BRS Capiáçu. Ao fim do processo, se objetiva a construção de uma metodologia que tenha potencial para auxiliar, com custo acessível e embasamento técnico, o melhor manejo da cultura.

2. Objetivo

2.1. Geral

Estimar através de imagens de sensor *RGB* acoplado a Veículo Aéreo Não Tripulado (VANT), Biomassa e Índice de Área Foliar de BRS Capiapu com algoritmos de *machine learning*.

2.2. Específicos

- Aferir a correlação entre Biomassa e Índice de Área Foliar, e os dados de índices de vegetação e das bandas do RGB.
- Estimar as variáveis Biomassa e Índice de Área Foliar com a aplicação dos algoritmos de machine learning sobre os índices e bandas que apresentarem as melhores correlações.
- Indicar os algoritmos que melhor representam o Índice de Área Foliar e a Biomassa.

3. Fundamentação Teórica

3.1. BRS Capiapu no contexto da criação animal

Segundo Lisboa *et al.* (2020) a base alimentar da pecuária brasileira é a pasto, sendo que as espécies forrageiras apresentam um banco de elementos essenciais à nutrição dos animais ruminantes. Por outro lado, as variações climáticas ao longo do ano, desfavorecendo o pleno desenvolvimento das pastagens, sendo responsáveis por gerar uma diminuição da oferta de forragem.

Neste sentido Monção *et al.* (2020) afirmaram que os produtores têm investido no cultivo de capineiras. Como meio para driblar o problema da diminuição da oferta de forragem. Entre elas, o capim-elefante se mostra como uma forrageira de alto potencial produtivo de suplemento volumoso. Sendo amplamente utilizado, principalmente nas médias e pequenas propriedades rurais no Brasil (Pereira *et al.*, 2016).

Visando o suprimento forrageiro, a Embrapa Gado de Leite através do programa de melhoramento genético do capim-elefante desenvolveu a BRS Capiapu, cultivar destinada à formação de capineira. Que se destaca pelo seu

elevado potencial de produção de massa de forragem por unidade área, apresentando um decente valor nutricional, destinado geralmente para produção de silagem ou servido picado (Pereira *et al.*, 2016).

Entre as características da cultivar que a categorizam em posição de destaque entre as alternativas forrageiras, estão sua resistência ao tombamento e ausência de pelos, o que facilita as atividades manuais com a cultura e a mecanização de sua colheita. Suas touceiras são densas e eretas; sua produtividade que pode chegar a 50 toneladas de massa de forragem por hectare a cada ano, superando culturas como a cana-de-açúcar e o milho. Comparada a outras cultivares de capim-elefante, a mesma apresenta maior teor de carboidratos solúveis e proteína bruta, resultando em uma silagem de elevada qualidade. A BRS Capiáçu se destaca ainda por apresentar tolerância ao estresse hídrico, sendo opção para o cultivo em regiões que apresentam risco elevado de incidência de veranicos (Pereira *et al.*, 2016; Pereira *et al.*, 2021).

Segundo Pereira *et al.* (2021) como consequência a sua elevada capacidade produtiva, a cultivar acaba por extrair nutrientes do solo em grandes quantidades. Sendo assim é recomendado a aplicação de fertilizantes químicos e orgânicos para a sua plena produção, bem como que o cultivo seja realizado em solos profundos, de boa drenagem e fertilidade natural.

3.2. Monitoramento de plantas forrageiras com sensoriamento remoto

Atualmente o que mais se utiliza para aferições das características biofísicas e de biomassa em pastagens são coletas destrutivas, onde realiza-se amostragens dentro da área. Para determinação de biomassa o material é colhido, passa por uma dessecação, e posteriormente é submetido a análises laboratoriais (Lisboa, 2020). É importante a busca por metodologias indiretas que possuam capacidade de aferir essas características das pastagens. Uma vez que os princípios da amostragem podem não ser atendidos em muitos locais, por possíveis dificuldades no processo. Como em áreas montanhosas de difícil acesso ou áreas com grandes extensões (Lisboa, 2020).

Trabalhos já indicam aplicações do monitoramento remoto com lógicas já bem desenvolvidas, e com potencial para a mensuração de atributos biofísicos em pastagens (Coelho, 2024).

Segundo Hoot *et al.* (2019) a utilização de VANTs possibilita diversas ações importantes no monitoramento de pastagens, uma vez que correlações entre produção de biomassa e os índices de vegetação extraídos podem ser feitas. Sendo possível o acompanhamento do desenvolvimento da forrageira ao longo do tempo, identificação de falhas ao longo da área, aferição do vigor das plantas, identificação de pragas e doenças, e extração do modelo digital do terreno (MDT) que possibilita a estimativa do volume de biomassa e altura de plantas, e ainda a identificação de pragas e doenças pode ser feita.

De forma que o sensoriamento remoto com a utilização de sensores embarcados em VANTs, possibilitam uma melhor e mais rápida coleta de dados de campo. Sendo que os custos em relação ao processo têm se tornado cada vez mais acessíveis (Tsouros *et al.*, 2019).

Um produto importante extraído dos sensores embarcados em VANTs, são os Índices de Vegetação. Os IVs são combinações matemáticas das bandas espectrais captadas em diferentes comprimentos de onda. De forma que essas formulações diminuem a complexidade dos produtos gerados, minimizam o efeito de ruídos e problemas de angulação entre o sensor e a luz solar. Permitindo o entendimento de características inerentes à vegetação monitorada, com base em sua resposta espectral nos diferentes comprimentos de ondas captados pelos sensores. Permitindo por exemplo a estimativa de biomassa, altura de pasto e teores de nitrogênio (Rey-Caramés *et al.*, 2015; XUE e SU, 2017; TONG *et al.*, 2019).

3.3. Modelagem com aplicação de Machine Learning (ML)

Jung *et al.* (2021) afirmam que com os avanços nos processos de aprendizado de máquina, é possível a obtenção de modelos preditivos e prescritivos em grandes volumes, através de algoritmos de aprendizagem.

Conforme Paixão (2023), a metodologia é um braço da inteligência artificial que busca o aprendizado autônomo a partir de técnicas computacionais. Analisando e extraíndo conhecimento de bases de dados pré-existentes (principalmente padrões). Necessitando de extensas bases de dados para adquirirem conhecimento satisfatório e modelagem de forma eficiente.

O ML é capaz de identificar padrões e gerar correlações entre os diversos dados, aprendendo sobre suas informações no processo. De forma que desse banco de dados, uma parte de seu volume é destinado ao treinamento da máquina, operação essencial para a exploração é identificação das características da base em questão. A outra fração dos dados é destinada a o processo de testes, fase posterior do ML, que é responsável por avaliar a eficácia do modelo preditivo (van Klompenburg *et al.*, 2020)

O método de validação cruzada K-fold é amplamente utilizado. Nessa técnica, o conjunto de dados é particionado em K subconjuntos. Em cada ciclo do processo, um desses subconjuntos é reservado para a avaliação do modelo, enquanto os demais são utilizados para o treinamento. Esse procedimento é realizado sucessivamente até que todas as partes tenham sido usadas como conjunto de teste. O desempenho final do modelo corresponde à média aritmética de todos os modelos que foram gerados (Spezia, 2024).

Segundo Liakos *et al.*, (2018) existem duas modalidades de aprendizado de máquinas. O aprendizado supervisionado e o aprendizado não supervisionado. No aprendizado supervisionado o conjunto de dados é dividido em dados de entrada (base de treinamento) e dados de saída (base de testes). Onde a base de dados treino é explorada pela máquina, identificando padrões e os rotulando. Posteriormente, o modelo treinado prevê os rótulos dos conjuntos de testes. O aprendizado não supervisionado consiste em um processo onde o conjunto de dados não é dividido, nem rotulado. A máquina busca simplesmente entender os padrões e ordenamentos da base por completo.

Dentro do aprendizado de máquinas, existem diversos algoritmos responsáveis por executar o processo. Cada um possui suas particularidades de funcionamento. O *Random Forest* (RF) trata-se de um algoritmo que treina várias árvores individualmente, procurando dentro de cada árvore classes de dados frequentes. Criando ao final do processamento modelos de regressão a partir da média dessas árvores, ou modelos de classificação com base nas classes mais frequentes. O algoritmo realiza o ajuste dos modelos através de amostragem bootstrap, que permite maior precisão estatística. De forma que o processo se repete até que ocorra a exploração máxima do banco de dados, ou quando o agrupamento de dados não seja mais divisível. O RF é

massivamente utilizado para atividades de ML que trabalham com classificações e regressões (Bulagang *et al.*, 2020; Breiman, 2001; Niseimon, 2020).

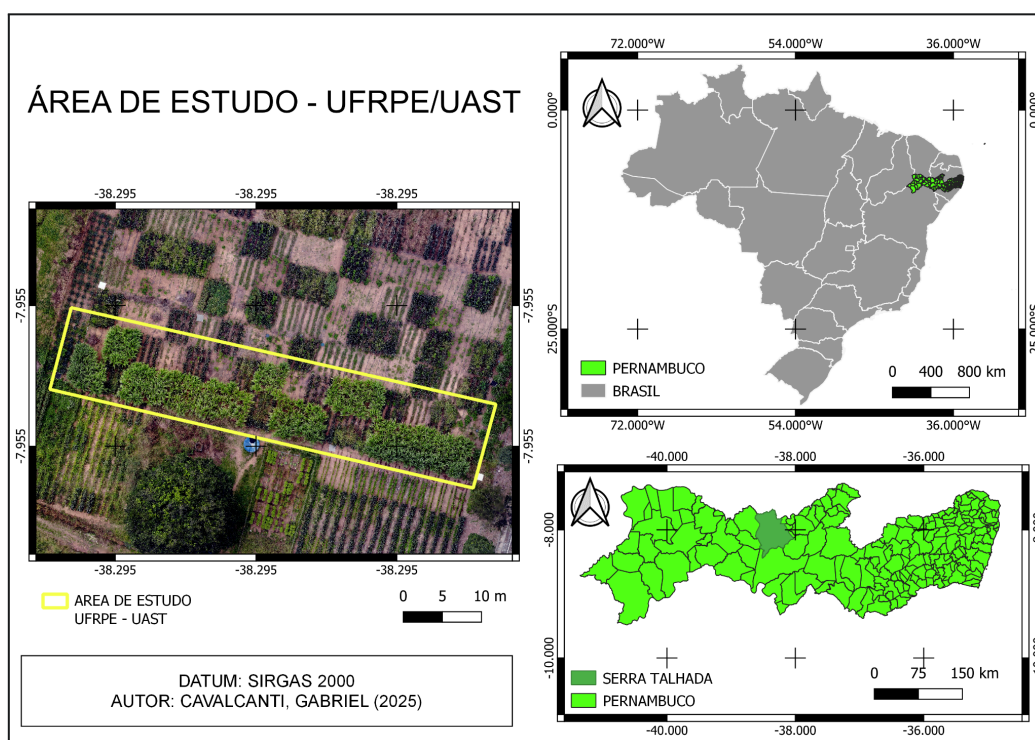
Outro algoritmo que tem demonstrado eficiência na resolução de demandas envolvendo modelagem, é o *Support Vector Machines* (SVM). Baseado no método estatístico Vetores de Suporte (*Support Vector*) proposto e relatado por Cortes e Vapnik (1995). Possui uma série de qualidades que justificam sua grande utilização em atividades de classificação e regressão, entre elas está sua capacidade de modelar de forma generalista e eficácia para trabalhar com grandes volumes de dados. O algoritmo trabalha buscando um hiperplano, que é uma superfície que atua como uma fronteira de separação entre classes de dados distintos. Esse hiperplano é responsável por maximizar a margem de separação entre as classes de dados, localizando o hiperplano nas distâncias mais próximas entre os pontos observados e o hiperplano. Esses pontos de controle são os vetores de suporte (Lorena e Carvalho, 2007; Nielsen, 2015). Segundo Nisieimon (2020), a metodologia se popularizou devido a sua multidisciplinaridade em trabalhar com problemas de regressão e de classificação, como também pela sua habilidade de realizar uma modelagem generalista no aproveitamento dos dados.

4. Metodologia

4.1. Caracterização da área de estudo

O experimento foi conduzido na Universidade Federal Rural de Pernambuco, Unidade Acadêmica de Serra Talhada, localizado no município de Serra Talhada, Pernambuco. A região é caracterizada pelo clima tipo BSwh – semiárido, quente e seco e com temperatura média anual de 32° C (Almeida, 2016; Lundgren *et al.*, 2013). A área experimental se localiza nas seguintes coordenadas geográficas: 7°57'16.95"S e 38°17'42.61"O, a altitude média do local é de 431 m (Figura 1).

Figura 1. Mapa de localização da área de estudo

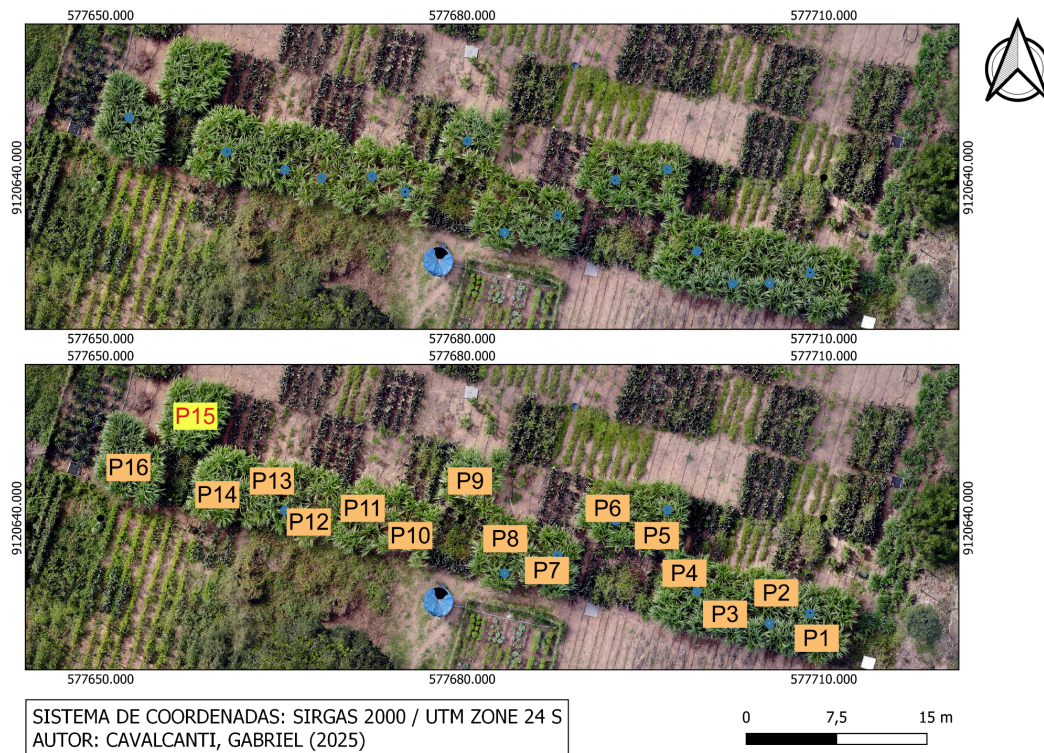


Fonte: Figura do autor

A área é dividida em 4 blocos com 4 parcelas cada, totalizando 16 parcelas ao todo. As plantas são cultivadas em quatro linhas por parcela, sendo que as duas linhas das extremidades são bordaduras. O cultivo possui espaçamento de 0,20 m x 1,0 m. Cada parcela possui 5 metros de comprimento por 4 de largura, totalizando 20 metros quadrados por parcela. Como cada ao todo 16 parcelas, a área total é de 320 m². De forma que as parcelas são submetidas a diferentes lâminas de irrigação, com o objetivo de se avaliar a eficiência hídrica de cada uma delas na cultura. São 50%, 75%, 100% e 125% da capacidade de campo, de maneira que a distribuição das lâminas é diferente em cada parcela.

A parcela 15 não foi georreferenciada, não sendo contabilizados dados para ela na modelagem. A Figura 2 traz o mapa amostral da área de estudo.

Figura 2. Mapa amostral das parcelas. P: parcela.



Fonte: Figura do autor.

4.2. Dados espectrais

O monitoramento da cultura para esse trabalho iniciou em outubro de 2023, se estendendo até fevereiro de 2025, compreendendo quatro ciclos. Nesse período foram feitos no total oito levantamentos espectrais. Sendo três desses em períodos de desenvolvimento inicial ou intermediário, novembro de 2023, janeiro de 2024 e maio de 2024. E coincidindo com a colheita em outubro de 2023, março de 2024, junho de 2024 e fevereiro de 2025. Imagens das sete datas foram utilizadas na modelagem de IAF, porém, para a biomassa foram aproveitadas apenas as três primeiras datas de final de ciclo. Isso porque os dados de campo para biomassa de fevereiro de 2024 apresentaram discrepância em relação às demais datas, o que prejudicaria o modelo gerando mais *outliers*.

Para registro das imagens foi utilizado o veículo aéreo não tripulado (VANT) DJI modelo Mavic 2 Enterprise Dual que apresenta autonomia de voo de aproximadamente 30 minutos, tipo quadricóptero, com GPS, e equipado

com uma câmera RGB e um sensor termal. O plano de voo foi definido com sobreposição das imagens nas faixas e entre faixas de 90%, altura de voo de 30 m, *Ground sample distance* (GSD) de 5,1 cm por pixel. As imagens coletadas no voo, foram processadas no agisoft metashape (versão 1.5.2), para geração dos ortomosaicos das imagens RGB. A partir dos ortomosaicos, no software de sistema de informações geográficos Qgis (versão 3.0), foram extraídas as estatísticas zonais para cálculo dos IVs.

Em cada parcela foi escolhida uma planta para realizar a estimativa dos parâmetros de interesse. Posteriormente com equipamento GNSS RTK, modelo Stonex 850A, essas plantas foram georreferenciadas.

A Tabela 1 apresenta a formulação dos 15 índices de vegetação apresentados por Freire-Silva *et al.*, (2019) e Swaef *et al.*, (2021). Que foram utilizados como variáveis independentes para a modelagem junto com as bandas RGB.

Tabela 1. Índices de vegetação e as respectivas equações de uso.

Índice	Sigla	Fórmula	Referência
<i>Visible Atmospherically Resistant Index</i>	VARI	$\frac{g-r}{g+r-b}$	(Gitelson et al., 2002)
<i>Modified Green Red Vegetation Index</i>	MGRVI	$\frac{g^2-r^2}{g^2+r^2}$	(Bending et al., 2015)
<i>Green Leaf Index</i>	GLI	$\frac{(2g-r-b)}{(2g+r+b)}$	(Louhaichi et al., 2001)
<i>Red Green Blue Vegetation Index</i>	RGBVI	$\frac{(g^2-b*r)}{(g^2+b*r)}$	(Bending et al., 2015)
<i>Green Red Vegetation Index</i>	NGRDI	$\frac{G-R}{G+R}$	Tucker (1979)
<i>Excess Red Vegetation Index</i>	ExR	$1.4r - g$	(Meyer & Neto, 2008)
<i>Excess Blue Vegetation Index</i>	ExB	$1.4b - g$	(Guijarro et al., 2011)
<i>Excess Green Vegetation Index</i>	ExG	$2g - r - b$	(Guijarro et al., 2011)
<i>Excess Green Red Vegetation Index</i>	ExGR	$ExG - ExR$	(Guijarro et al., 2011)
<i>Color Index of Vegetation</i>	CIVE	$0.441r - 0.881g + 0.385b + 18.78745$	(Guijarro et al., 2011)
<i>Vegetative</i>	VEG	$\frac{g}{r^{0.667} * b^{0.333}}$	(Hague, Tillett & Wheeler, 2006)

<i>Combination</i>	COM	$0.25ExG + 0.3ExGR + 0.33CIVE + 0.12V$	(Guijarro et al., 2011)
<i>Woebbecke Index</i>	WI	$\frac{g-b}{r-b}$	(Woebbecke et al., 1995)
<i>Kawashima Index</i>	IKAW	$\frac{r-b}{r+b}$	(Shigeto & Makoto, 1998)
<i>Color Intensity Index</i>	INT	$\frac{R+G+B}{3}$	(Shigeto & Makoto, 1998)

Fonte: Elaboração própria

Foram criados buffers nos ortomosaicos ao redor de cada planta georreferenciada, a fim de se extrair os valores médios dos pixels com estatísticas zonais, no QGIS, as dimensões desses buffers foram definidas para cada uma das duas variáveis (IAF e Biomassa). Para o IAF, aplicou-se um buffer com raio de 25 cm ao redor de cada planta georreferenciada, isso se justifica devido ao fato de que os dados biométricos das folhas, que abastecem o cálculo da variável serem coletados individualmente, por planta, no caso em duas de cada parcela. Também pelo fato de dados de imageamentos de períodos de menor desenvolvimento vegetativo da cultura, serem utilizados para sua estimativa. Sendo assim, 25 cm apresentou uma melhor cobertura da área efetiva da parte aérea das plantas individualmente (Figura 3), sobretudo em estádios iniciais de desenvolvimento, evitando problemas de saturação de índices pela interferência do solo.

Figura 3. Buffer de 25 cm em plantas de BRS Capiáçu em estágio inicial de desenvolvimento.



Fonte: Figura do autor.

A dimensão que melhor se aplicou as estatísticas zonais da estimativa de biomassa, foi de 1 metro (Figura 4). Uma vez que o peso de oito plantas é aferido no final do ciclo da cultura, onde a planta apresenta sua maior área de

parte aérea. Essa maior extensão do buffer cobre com mais efetividade a área da planta, em estágio final de desenvolvimento, fornecendo dados mais consistentes para a modelagem.

Figura 4. Buffer de 1 m em plantas de BRS Capiacu em estágio final de desenvolvimento.



Fonte: Figura do autor.

4.3. Dados de campo

A obtenção dos dados biométricos e de biomassa da cultura foram obtidas com biometrias periódicas de características foliares durante o desenvolvimento da cultura e contabilização da biomassa no final do ciclo.

As biometrias de caracterização e desenvolvimento foliar, foram realizadas com uma periodicidade de 14 dias. Sendo contabilizadas as medidas de comprimento e largura da terceira folha com no mínimo 30% de sua área verde, bem como o número de folhas vivas e número de folhas mortas.

Para o cálculo do índice de Área Foliar (IAF), primeiro foi calculada a Área Foliar (AF), através dos dados biométricos e depois dividido o valor pela área da parcela. O modelo utilizado para AF foi baseado no proposto por Hermann e Câmara (1999):

$$AF = C * L * 0,75 * (N+2) \quad (1)$$

Onde, C é o comprimento da folha, L a largura, 0,75 o fator de forma e o N+2 representa o número de folhas totalmente abertas e com no mínimo 20% de sua área verde. Esse modelo é uma atualização do proposto por Francis *et al.* (1969), para a estimativa da AF na cultura do milho.

Para esse trabalho o modelo foi ajustado para:

$$AF = C * L * 0,75 * (N+3) \quad (2)$$

Ajustando para que o número de folhas totalmente abertas seja de no mínimo 30% de sua área verde. Que corrobora com a metodologia de coleta dos dados biométricos fornecidos para esse trabalho.

A biomassa foi obtida através de uma métrica do peso de oito plantas que é realizada sempre no final do ciclo e utilizada para estimar a produtividade no experimento. Sendo dividido o valor do peso de oito plantas por 8, obtendo-se o peso de uma planta que multiplicado por 100, valor que quantifica o número de plantas da parcela. Dividindo o valor da biomassa total da parcela, por sua área em metros quadrados, obteve-se a biomassa por unidade de área, expressa em quilogramas por metro quadrado (kg/m²).

4.4. Modelagem sobre o banco de dados

O banco composto pelos dados espectrais (bandas RGB e IVs) e pelos dados de campo (IAF e biomassa), foram submetidas a testes de correlações pelo método de *Spearman*, para a aferição das relações entre as duas modalidades de dados. Uma matriz de correlações foi gerada, para a demonstrando da relação das variáveis dependentes com as 18 variáveis preditoras.

Posteriormente realizou-se com a modelagem de aprendizado de máquinas com os algoritmos RF e SVM. Particionando os dados em proporção de 70% para subconjunto de treino e 30% para teste. Configurando e ajustando seus hiperparâmetros de modo a extrair a melhor performance dos mesmos na predição das variáveis de interesse. No final, extraíndo o modelo, seus parâmetros estatísticos, o gráfico de dispersão e a importância das variáveis preditoras para a modelagem.

Com o intuito de aprimorar a performance dos algoritmos, procedeu-se com uma seleção de variáveis pós-modelagem, excluindo as variáveis de menor importância, e preservando as de maior contribuição para a construção do modelo.

4.5. Avaliação dos modelos

Segundo Oliveira (2025), as métricas adequadas permitem aferir a qualidade e eficácia dos modelos, uma vez que mensuram seu nível de

eficiência em estimar corretamente a partir dos dados fornecidos. Etapa extremamente importante para a modelagem.

Nesse sentido, na avaliação da capacidade preditiva dos modelos, aplicou-se validação cruzada (k fold=10), onde foram estimados 10 modelos, e ao final escolhido o melhor. A validação dos modelos foi feita pela mensuração do erro quadrático médio (RMSE), o erro absoluto médio (MAE), e o coeficiente de compacidade (R^2).

Toda a execução da modelagem foi realizada em ambiente R, utilizando a interface do Rstudio (R CORE TEAM, 2024).

5. Resultados e Discussões

5.1. Análise descritiva dos dados

Na Tabela 1 são demonstrados os valores de mínimo, máximo, média, mediana, desvio padrão e coeficiente de variação, para as variáveis mensuradas em campo (Biomassa e IAF). Gomes (2000), afirma que o coeficiente de variação é uma medida válida para a indicação da variabilidade de resultados experimentais. O autor inclusive definiu faixas de classificação para o indicado. São elas: baixo (0 a 10%); médio (10 a 20%); alto (20 a 30%); muito alto (>30%).

A biomassa apresentou um CV de 0,246 ou 24,60%, que segundo a classificação é considerado um valor alto. Mesmo sendo mensurada sempre em final de ciclo, onde a planta está em estágio final de desenvolvimento e expressa seu máximo potencial produtivo. Houve variação considerável nos valores de biomassa entre os ciclos avaliados. O IAF, por sua vez, demonstrou um valor ainda maior (0,49 ou 49,0%), que é considerado muito alto pela classificação. Esse comportamento ocorre pelo fato de que o banco de dados para o IAF, possui resultados de biometrias de diferentes estágios de desenvolvimento da cultura em cada ciclo estudado.

Tabela 2. Análise descritiva dos dados de Biomassa e de Índice de Área Foliar

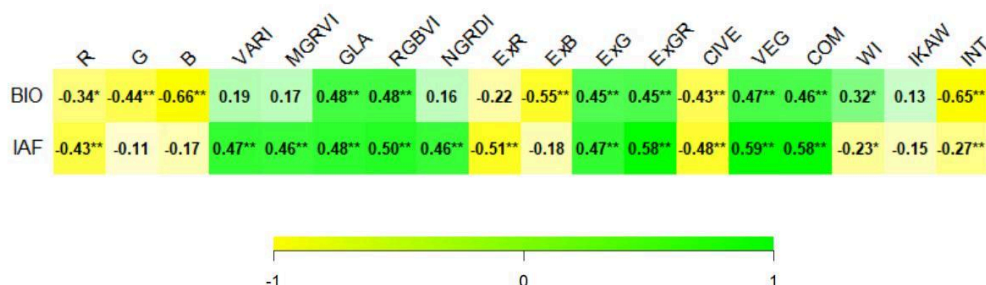
Parâmetros Estatísticas	Biomassa (Kg/m ²)	Índice de Área Foliar
Mínimo	1,140	0,095
Máximo	3,530	3,420
Média	2,200	1,585
Mediana	2,140	1,664
Desvio Padrão	0,540	0,776
CV	0,246	0,490

CV: coeficiente de variação.

Fonte: Elaboração própria

Com base na matriz de correlações (Figura 5) as maiores correlações com a biomassa foram negativas, na figura das variáveis B, INT e ExB (B= -0,66, INT= -0,65, ExB= -0,55). Seguidas por GLA e RGBVI com correlações positivas de 0,48. Para o IAF as maiores correlações foram com os índices VEG, COM, ExGR, ExR e RGBVI (VEG= 0,59, COM= 0,58, ExGR= 0,58, ExR= -0,51, RGBVI= 0,50). Com mais cinco índices apresentando correlações de 0,48 a 0,46.

Sendo assim, observa-se que as correlações não apresentaram forte relação com as variáveis de interesse. Com três variáveis de Biomassa apresentando valores acima de 0,50 mas abaixo de 0,70. E em relação a o IAF entre 0,50 e 0,60.

Figura 5. Matriz de correlações de Spearman entre a Biomassa (BIO) e Índice de área Foliar (IAF) em relação às bandas RGB e os IVs. Significância a 5% representada por: **.**Fonte:** Figura do autor.

5.2. Análise da modelagem

A modelagem de aprendizado de máquinas indicou performance semelhante entre os algoritmos para as duas variáveis de interesse, com ligeiras diferenças entre si (Tabela 3). A seleção de variáveis pós-modelagem também não melhorou a eficiência da maioria dos modelos, pelo contrário, piorou as métricas desses. Apresentando algum incremento, ainda assim baixo, apenas para as métricas da Biomassa com SVM. Dessa forma, optou-se por trazer os gráficos de dispersão e de importância das variáveis apenas para os modelos sem seleção.

Ao se avaliar os modelos para a biomassa (Tabela 3), temos (RMSE= 0,140, MAE= 0,108, R^2 = 0,64) para o RF sem seleção e (RMSE= 0,159, MAE= 0,124, R^2 = 0,46) para RF com seleção; SVM sem seleção (RMSE= 0,177, MAE= 0,144, R^2 = 0,77), SVM com seleção (RMSE= 0,176, MAE= 0,143, R^2 = 0,78). Com relação ao IAF (RMSE= 0,171, MAE= 0,132, R^2 = 0,45) para RF sem seleção e (RMSE= 0,175, MAE= 0,137, R^2 = 0,43) para RF com seleção; SVM sem seleção (RMSE= 0,172, MAE= 0,127, R^2 = 0,41), SVM com seleção (RMSE= 0,189, MAE= 0,145, R^2 = 0,39).

Tabela 3. Parâmetros estatísticos de estimativa das variáveis com os algoritmos de aprendizagem de máquina.

Variável de Interesse	Algoritmo	Sem seleção			Com seleção		
		RMSE	MAE	R^2	RMSE	MAE	R^2
Biomassa (Kg/m ²)	RF	0,140	0,108	0,64	0,159	0,124	0,46
	SVM	0,177	0,144	0,77	0,176	0,143	0,78
IAF	RF	0,171	0,132	0,45	0,175	0,137	0,43
	SVM	0,172	0,127	0,41	0,189	0,145	0,39

SVM: máquinas de vetores de suporte, RF: floresta aleatória, RMSE: erro quadrático médio, MSE: erro absoluto médio, R^2 : coeficiente de capacidade.

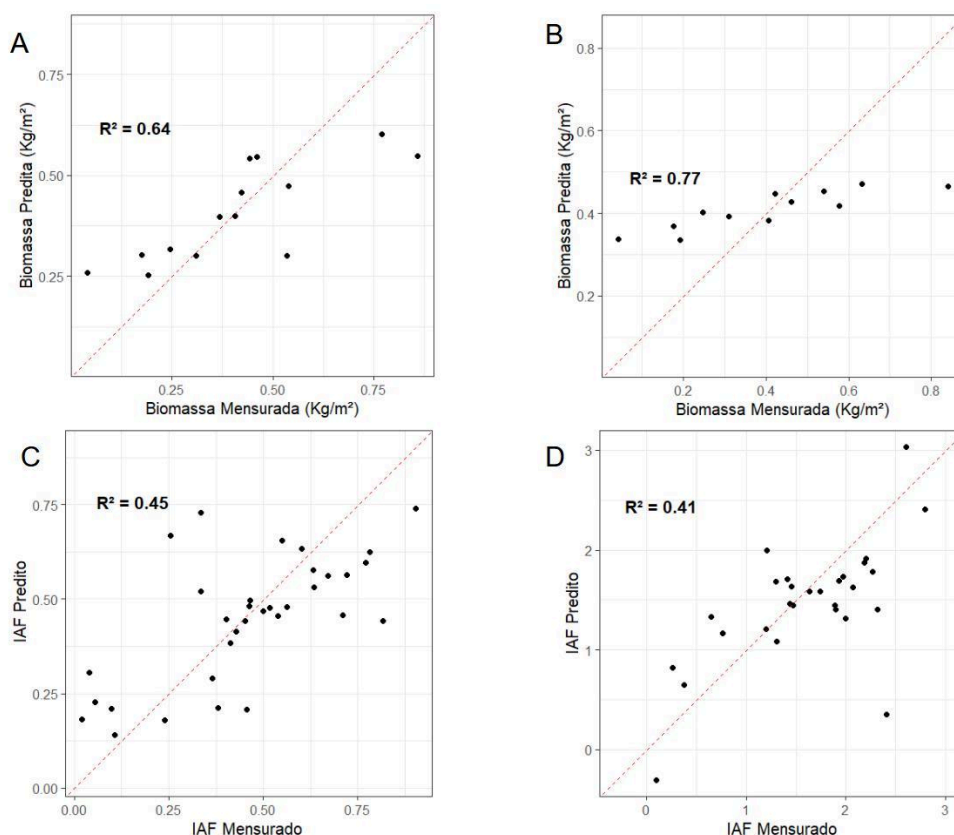
Fonte: Elaboração própria.

Observando os gráficos da dispersão das medidas mensuradas em campo e os valores estimados (Figura 6), se evidencia que os modelos não conseguiram reproduzir com eficiência predições assertivas para as variáveis de interesse. A distribuição dos pontos nos planos cartesianos é pouco precisa,

apresentando um ordenamento aleatório para os dois algoritmos nas duas variáveis.

No que se refere ao R^2 , Poley e Mcdermid (2020) em sua revisão, analisaram trabalhos sobre mensuração de biomassa de várias culturas através de imagens espectrais obtidas por VANT's e modelagem preditiva. Observaram que o R^2 é o parâmetro mais utilizado para a validação dos modelos, e criaram uma escala para classificar os modelos em relação a o indicador. Onde, é considerado pobre = 0 a 0,25; moderado= 0,25 a 0,50; boa= 0,50 a 0,75; Excelente= 0,75 a 1. A partir dessa escala, e considerando todas as modelagens, a performance de Biomassa seria considerada boa e a modelagem para o IAF considerada de performance moderada. Porém, os gráficos de dispersão demonstram dificuldades por parte dos modelos no ajuste das estimativas.

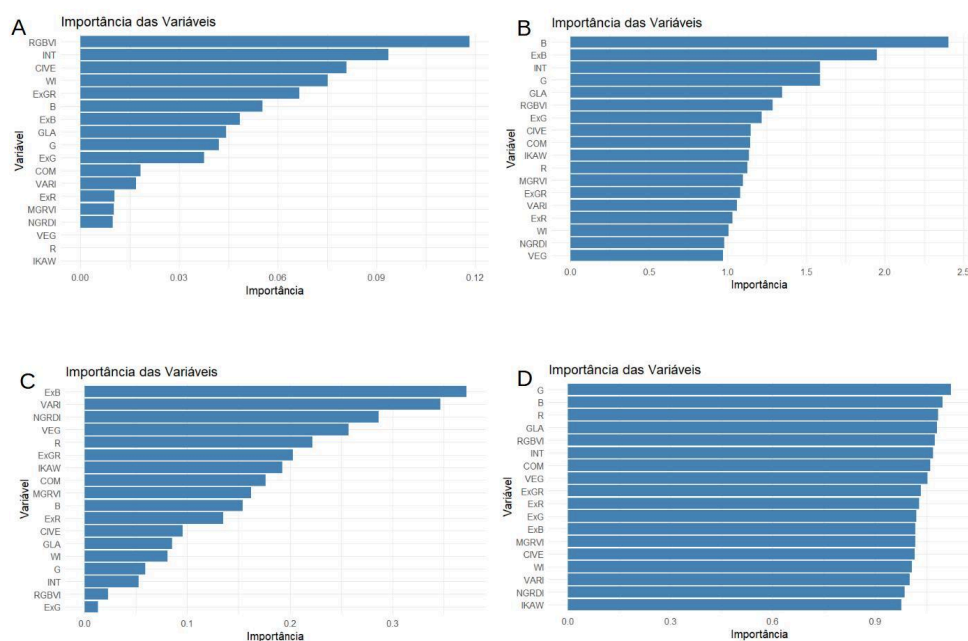
Figura 6. Ajuste dos modelos de biomassa e índice de área foliar - A) Biomassa com RF, B) Biomassa com SVM, C) IAF com RF e D) IAF com SVM.



Fonte: Figura do autor.

Ao se analisar os gráficos de importância das variáveis (Figura 6), percebe-se uma maior disparidade de relevância das variáveis nos modelos RF, em relação a SVM. Existe uma diminuição acentuada, quase linear, do índice de importância entre as variáveis no RF. Comportamento que difere para o algoritmo SVM onde a diminuição do índice é menos acentuada, principalmente para o IAF. Com a biomassa fugindo um pouco desse padrão nas primeiras cinco variáveis (da banda B até o IV GLA). Seguido por uma diminuição da variação do GLA ao VEG.

Figura 7. Importância das variáveis para biomassa e índice de área foliar - A) Biomassa com RF, B) Biomassa com SVM, C) IAF com RF e D) IAF com SVM.



Fonte: Figura do autor.

Em relação à concepção da metodologia RF, Boelter (2021), afirma que consoante a alta capacidade das árvores de decisões em classificar bancos de dados, surgiu a ideia da criação de diversas árvores, cada uma explorando determinado aspecto do conjunto de dados e sendo treinadas a partir dele. E o resultado sendo o somatório de todas as árvores. Isso evidencia a capacidade seletiva iminente do RF, para a classificação das tendências do banco de dados explorado. O que pode justificar a disparidade de importância entre as variáveis utilizadas para criação dos modelos baseados nesta técnica, no presente trabalho.

Nisieimon (2020), relatou em seu trabalho que o SVM necessitou de mais atributos que os algoritmos baseados em árvores de decisão (RF e GBM), para realizar as predições com maior assertividade. E que gerou um modelo passível de *overfitting*. Por sua abordagem generalista no aproveitamento dos dados, o algoritmo estima o modelo com base na aproximação proporcionada pelo hiperplano orientado pelos vetores de suporte. O algoritmo aproveita um número maior de variáveis na modelagem o que é visualizado nos gráficos de importância, a desvantagem é que essa estratégia pode proporcionar que valores dispersos (*outliers*) sejam utilizados na criação do modelo, ocasionando o *overfitting*.

É claro que no trabalho de Nisieimon as técnicas utilizadas são metodologias mais eficientes como Relief e GA (*Genetic Algorithm*), que são diferentes da simples metodologia executada no presente estudo. São realizadas antes da modelagem do algoritmo, mapeando as melhores variáveis e otimizando o custo operacional da máquina na predição já no primeiro momento, apesar de envolver o custo operacional para a seleção. Sendo um problema a ser resolvido com base na situação particular de cada trabalho de modelagem.

Não menos importante, o principal motivo para a dificuldade dos algoritmos no ajuste do modelo é a pequena amostragem de dados para as duas variáveis de interesse, mas principalmente para Biomassa. Shen e Qian (2022), afirmam que os algoritmos de *machine learning* necessitam de dados em volumes consideráveis para conseguirem compreender em essência o objeto de estudo. Nesse sentido, bancos de dados pequenos tendem a não representar na totalidade o que se está investigando. Segundo Li *et al.*, (2018), os algoritmos apresentam baixo desempenho quando a base de dados utilizada é pequena e com distribuição desuniforme. Nessas situações existem lacunas de informações.

Esse comportamento é chamado de *underfitting*, que se caracteriza pelo aprendizado insuficiente do algoritmo, implicando em um ajuste deficitário aos dados de treino. O modelo não é capaz de extrair os padrões de todas as camadas de dados, se constituindo em simples e ineficiente para a predição diante do conjunto de testes. Sua ocorrência está ligada a baixa representatividade dos dados utilizados, a deficiências nos ajustes dos

parâmetros dos algoritmos e a curtos espaços de tempo para o treinamento (Barboza, 2023; Chacon, 2021; Silva, 2023).

Outro fato, é que a base composta pelas bandas e pelos índices obtidos através do sensor RGB, pode não ter sido eficiente na representação espectral da cultura. A maior reflectância neste espectro por partes das plantas, é a banda *green*, que está relacionada à coloração da clorofila. Diferente dos sensores multiespectrais que além do RGB, captam as bandas de maior comprimento de onda, como infravermelho e infravermelho próximo, é onde está a maior resposta espectral das plantas (Araujo, 2022).

6. Conclusões

Os modelos de RF e SVM, a partir dos IVs e das bandas RGB, não produziram ajuste satisfatório para representarem a Biomassa e o IAF. Provavelmente o banco de dados foi inconclusivo para o aprendizado dos modelos. Tanto por possuir um pequeno volume, como pela dificuldade do sensor RGB em captar a assinatura espectral da cultura.

O *underfitting* ficou nítido para o banco de dados de Biomassa, com o gráfico de dispersão plotando as estimativas do modelo praticamente em linha reta para o SVM. Demonstrando um aprendizado insuficiente e incoerente da tendência dos dados.

O uso de sensores multiespectrais pode gerar uma base de dados espectrais de maior qualidade, extração de textura e modelos digitais de terreno e de elevação (MDT e MDE). Como também o aumento do volume de dados biofísicos da cultura, pode fornecer um banco com maior capacidade de contribuir com esse tipo de modelagem na BRS Capiáçu futuramente.

Bibliografia

ALMEIDA, H. A. de. **Climatologia Aplicada à Geografia**. Paraíba: Universidade Estadual da Paraíba, 21 ed. p. 317, 2016.

ARAUJO, B. C. **Drones como ferramenta de gestão florestal: desempenho dos índices IRAV (Índice Resistente à Atmosfera na Região Visível) e IVT (Índice Triangular Verde) na estimativa de volume de Eucalyptus**. Trabalho de Conclusão de Curso (Graduação em Engenharia Florestal) - Unidade Acadêmica Especializada em Ciências Agrárias, Universidade Federal do Rio Grande do Norte. p. 2, 2022.

BARBOSA, J. D. M. **Manutenção Preditiva com Recurso a Inteligência Artificial**. Tese de Doutorado. Universidade de Coimbra. p. 43, 2023

BOELTER, J. S. **Classificação de sinais eletromiográficos utilizando redes neurais artificiais, análise discriminante linear e floresta aleatória**. Trabalho de Conclusão de Curso (Bacharel em Engenharia Eletrônica) - Universidade do Vale do Rio dos Sinos. p. 50, 2021.

BULAGANG, A.F.; WENG, N.G.; MOUNTSTEPHENS, J.; TEO, J. A Review of Recent Approaches for Emotion Classification Using Electrocardiography and Electrodermography Signals. **Inform. Med. Unlocked**, 20, 100363, 2020.

BREIMAN, L. **Random forests**. *Mach Learn*; 45:5–32, 2001

COELHO, R. M. **Caracterização do estabelecimento e previsão da massa de forragem de pastos de capins *Cynodon* spp. por sensoriamento remoto**. 2024. Dissertação (Mestrado em Ciência Animal e Pastagens) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo. doi:10.11606/D.11.2024.tde-08102024-162721, 2024

CORTES, C.; VAPNIK, V. **Support-vector networks**. **Machine Learning**, v. 20, p. 273-297, 1995.

CHACON, A. B. S. **Serviço de predição de dados para IoT utilizando Machine Learning com aprendizado incremental**. Trabalho de Conclusão de Curso (Graduação em Análise e Desenvolvimento de Sistemas) - Unidade Acadêmica Especializada em Ciências Agrárias, Universidade Federal do Rio Grande do Norte. p. 32, 2021.

CHEN, T.; GUESTRIN, C. **XGBoost: A scalable tree boosting system**. In: **Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining**. New York, NY, USA, 13–17. p. 785–794, 2016.

FRANCIS, C. A.; RUTGER, J. N.; PALMER, A. F. E. A rapid method for plant leaf area estimation in maize (*Zea mays* L). **Crop Science**, Madison, v. 9, p. 537-539, 1969.

FREIRE, S. J., LIMA, S. P. P., PEREIRA, J. A., CANDEIAS, A. L. B. Índices de vegetação do sensoriamento remoto para processamento de imagens na faixa do visível (*RGB*). **Journal of Hyperspectral Remote Sensing**, v.9, n.4, p. 228-239, 2019.

GODOY, F. C.; RESENDE, L. F. T. Sinergia entre aprendizado de máquinas e dinâmicas moleculares para predição de propriedades moleculares: uma revisão da literatura. Anais Workshop do Programa de Mestrado Profissional Tecnologia em Química e Bioquímica da USP. **Blucher**. Vol. 5. p. 9-32, 2024.

GOMES, F. P. **Curso de estatística experimental**. Piracicaba: USP, ESALQ. 14ª ed. p. 477, 2000.

HERMANN, E. R.; CÂMARA, G. M. S. Um método simples para estimar a área foliar de cana-de-açúcar. **Revista da STAB**, [Piracicaba], v. 17, n. 5, p. 32-34, 1999.

HOTT, M. C.; ANDRADE, R. G.; MAGALHAES JUNIOR, W. C. P. de; BENITES, F. R. G. **Uso de veículo aéreo não tripulado (VANT) para estimativa de vigor e de correlações agrônômicas em genótipos de capim Cynodon**. In: SILVA, H. C. da (org.). Engenharia sanitária e ambiental. Ponta Grossa: Atena. p. 235-244, 2019.

Instituto Brasileiro de Geografia e Estatística (IBGE). 2024 registra recorde no abate de bovinos, frangos e suínos. **Agência de notícias - IBGE**, 2025. Disponível em: <https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/42898-2024-registra-recorde-no-abate-de-bovinos-frangos-e-suinos>. Acesso em: 20 de julho de 2025.

IZBICKI, R.; SANTOS, T. M. **Aprendizado de máquina: uma abordagem estatística**. v. 1. p. 272. ISBN: 978-65-00-02410-4, 2020

JUNG, J.; MAEDA, M.; Anjin CHANG, A.; BHANDARI, M.; ASHAPURE, A.; LANDIVAR-BOWLES, J. The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. **Current Opinion in Biotechnology**, v. 70, p. 15–22, 2021.

LIAKOS, K.G.; BUSATO, P.; MOSHOU, D.; PEARSON, S.; BOCHTIS, D. Machine learning in agriculture: a review. **Sensors**. v. 18, p. 2674, 2018.

LI, D. C.; LIN, W, K.; CHEN, C. C.; CHEN, H. Y.; LIN, L. S. Rebuilding sample distributions for small dataset learning. **Decision Support Systems**, [s.l.], v. 105, p 66-76, 2018.

LISBOA, A. M. **Sensoriamento remoto na avaliação de pasto de Brachiaria decumbens**. Dissertação (Mestrado em Zootecnia) - Universidade Federal de Viçosa. p 13-14, 2020.

LUNDGREN, W, J, C.; SILVA, L, F.; ALMEIDA, A, Q. Influência das espécies exóticas arbóreas urbanas na área de cobertura da cidade de Serra Talhada - PE. Soc. Bras. de Arborização Urbana. **REVSBAU**. v.8, n.3, p 96-107, 2013.

SILVA, L. R. M. **Classificação do potencial energético de combustíveis oriundos da bioenergia com o auxílio de aprendizagem de máquina**. Monografia (Curso de Engenharia de Energias) Universidade da Integração Internacional da Lusofonia Afro-Brasileira. p. 32, 2023.

LORENA, A. C.; CARVALHO, A. C. P. L. F. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43-67, 2007.

MAGALHÃES, P. S. G.; SANCHES, G. M; MOLIN, J. P.; PORTZ, G.; AMARAL, L. R.; KOLLN, O. T.; Duft, D. G. **Potential solution to contribute for sugar cane spatial variability management**. Anais... Proceedings of EFITA, 2013. v. 1. p. 1-8, 2013.

MONÇÃO, F.P.; COSTA, M.A.M.S.; RIGUEIRA, J.P.S.; SALES, E.C.J. DE; LEAL, D.B.; SILVA, M.F.P. DA; GOMES, V.M.; CHAMONE, J.M.A.; ALVES, D.D.; CUNHA SIQUEIRA CARVALHO, C. DA; MURTA, J.E.J.; JÚNIOR, V.R.R. Productivity and nutritional value of BRS capiaçu grass (Pennisetum

purpureum) managed at four regrowth ages in a semiarid region. **Tropical Animal Health and Production**. v.52, p.235–241, 2020.

NISIEIMON, V. H. **Estimativa de biomassa de pastagem utilizando técnicas de seleção de atributos e aprendizagem de máquina a partir de imagens do Sentinel-2**. Trabalho de Conclusão de Curso (Bacharel em Ciências Agrícolas) - Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas. p. 8-20, 2020.

NIELSEN, M. A. **Neural Networks and Deep Learning**. San Francisco: Determination Press, 2015.

OLIVEIRA, D. H. B. **Aprimoramento das previsões de demanda: uma abordagem integrada com Machine Learning e tratamento eficiente de dados**. Monografia (Graduação em Engenharia de Controle e Automação) – Escola de Minas, Universidade Federal de Ouro Preto. p. 25, 2025.

PAIXÃO, R. C. F. **Comparação de técnicas de aprendizado de máquina para prever a resistência à compressão do concreto**. Monografia (Graduação em Engenharia de Controle e Automação) - Escola de Minas, Universidade Federal de Ouro Preto. p. 18, 2023.

PEREIRA, A.V.; AUAD, A.M.; SANTOS, A.M.B.; MITTELMANN, A.; GOMIDE, C.A.M.; MARTINS, C.E.; PACIULLO, D.S.C.; LÉDO, F.J.S.; OLIVEIRA, J.S.; LEITE, J.L.B.; MACHADO, J.C.; MATOS, L.L.; MORENZ, M.J.F.; ANDRADE, P.J.M.; BENDER, S.E.; ROCHA, W.S.D. **BRS Capião e BRS Kurumi: cultivo e uso**. Brasília, DF: Embrapa, 2021.

PEREIRA, A. V.; LEDO, F. J. S.; MORENZ, M. J. F.; LEITE, J. L. B.; SANTOS, A. M. B.; MARTINS, C. E.; MACHADO, J.C. BRS Capião: cultivar de capim-elefante de alto rendimento para produção de silagem. Comunicado Técnico, **Embrapa CNPGL**, Outubro, n. 79, p.1-2, 2016.

REY CARAMÉS, C.; DIAGO, M. P.; MARTÍN, M. P.; Lobo, A.; TARDAGUILA, J. Using RPAS Multi-Spectral Imagery to Characterise Vigour, Leaf Development, Yield Components and Berry Composition Variability within a Vineyard. **Remote Sensing**. 7. 14458-14481. 10.3390/rs71114458, 2015.

SPEZIA A. M. **Estudo de estratégias de validação cruzada baseadas em clusters**. Monografia (Graduação em Engenharia da Computação) - Instituto de Informática, Universidade Federal do Rio Grande do Sul, p. 18. 2024.

SHEN, L.; QIAN, Q. A virtual sample generation algorithm supporting machine learning with a small-sample dataset: A case study for rubber materials. **Computational Materials Science**, [s. l.], v. 211, e-article 111475, August 2022.

SWAEF, T.; MAES, W. H.; APER, J.; BAERT, J.; COUGNON, M.; REHEUL, D.; STEPPE, ROLDÁN-RUIZ, I.; LOOTENS, P. Applying RGB-and Thermal-Based Vegetation Indices from UAVs for High-Throughput Field Phenotyping of

Drought Tolerance in Forage Grasses. **Remote Sensing**. 13. 147. 10.3390/rs13010147, 2021.

POLEY, L., G.; MCDERMID, G., J. A Systematic Review of the Factors Influencing the Estimation of Vegetation Aboveground Biomass Using Unmanned Aerial Systems. **Remote Sensing**. v. 12, n. 7, p. 1052, 2020.

VAN KLOMPENBURG, T.; KASSAHUN, A.; CATAL, C. Crop yield prediction using machine learning: A systematic literature review. **Computers Electronics Agriculture**. v. 177, 105709, 2020.

R CORE TEAM. R: **A Language and Environment for Statistical Computing** Vienna, Austria, 2024.

TONG, X.; DUAN, L.; LIU, T.; SINGH, V. P. Uso combinado de índices de vegetação hiperespectral in situ para estimar a biomassa da pastagem no pico do período produtivo para decisão de colheita. **Precision Agriculture**, v.20, p.477–495, 2019.

TSOUROS, D. C.; BIBI, S.; SARIGIANNIDIS, P. G. A review on UAV-based applications for precision agriculture. **Information (Switzerland)**, v. 10, n. 11, 2019.

XUE, J., SU, B. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. **Journal of Sensors**, v.17, 2017.