



**UNIVERSIDADE FEDERAL DE UBERLÂNDIA
CAMPUS MONTE CARMELO**



**PROGRAMA DE PÓS-GRADUAÇÃO EM AGRICULTURA E
INFORMAÇÕES GEOESPACIAIS – PPGAIG**

ANDRÉA LUCÍLIA PEREIRA DA SILVA

**Análise do ganho de produtividade de soja inoculada com microrganismos
solubilizadores de fósforo por meio de informações geoespaciais**

Monte Carmelo – MG

2025



**UNIVERSIDADE FEDERAL DE UBERLÂNDIA
CAMPUS MONTE CARMELO**



**PROGRAMA DE PÓS-GRADUAÇÃO EM AGRICULTURA E
INFORMAÇÕES GEOESPACIAIS – PPGAIG**

ANDRÉA LUCÍLIA PEREIRA DA SILVA

**Análise do ganho de produtividade de soja inoculada com microrganismos
solubilizadores de fósforo por meio de informações geoespaciais**

Dissertação apresentada ao Programa de Pós-Graduação em Agricultura e Informações Geoespaciais - PPGAIG da Universidade Federal de Uberlândia - UFU, campus Monte Carmelo, como requisito para a obtenção do título de Mestre.

Área de concentração: Informações geoespaciais e tecnologias aplicadas à produção agrícola

Orientador: Prof. Dr. Gilberto de Oliveira Mendes
Coorientador: Prof. Dr. George Deroco Martins.

Monte Carmelo – MG

2025

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

S586 2026	Silva, Andréa Lucília Pereira da, 1975- Análise do ganho de produtividade de soja inoculada com microrganismos solubilizadores de fósforo por meio de informações geoespaciais [recurso eletrônico] / Andréa Lucília Pereira da Silva. - 2026. Orientador: Gilberto de Oliveira Mendes. Coorientador: George Deroco Martins. Dissertação (Mestrado) - Universidade Federal de Uberlândia, Pós-graduação em Agricultura e Informações Geoespaciais. Modo de acesso: Internet. DOI http://doi.org/10.14393/ufu.di.2026.6 Inclui bibliografia. 1. Agronomia. I. Mendes, Gilberto de Oliveira, 1983-, (Orient.). II. Martins, George Deroco, 1987-, (Coorient.). III. Universidade Federal de Uberlândia. Pós-graduação em Agricultura e Informações Geoespaciais. IV. Título. CDU: 631
--------------	---

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:
Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
 Coordenação do Programa de Pós-Graduação em Agricultura e Informações
 Geoespaciais
 Rodovia LMG 746, Km 01, s/nº, Bloco 1AMC, Sala 1A202, Monte Carmelo-MG, CEP 38.500-000
 Telefone: (34) 3810-1033 - ppgaig@iciag.ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Agricultura e Informações Geoespaciais				
Defesa de:	Dissertação de Mestrado Acadêmico				
Data:	28/11/2025	Hora de início:	8:00h	Hora de encerramento:	10:00h
Matrícula do Discente:	32312AIG001				
Nome do Discente:	Andréa Lucilia Pereira da Silva				
Título do Trabalho:	Análise do ganho de produtividade de soja inoculada com microrganismos solubilizadores de fósforo por meio de informações geoespaciais				
Área de concentração:	Informações geoespaciais e tecnologias aplicadas à produção agrícola				
Linha de pesquisa:	Sistemas Integrados de Produção Vegetal				
Projeto de Pesquisa de vinculação:	Processos e produtos microbianos aplicados à produção vegetal				

Reuniu-se em sala virtual a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Agricultura e Informações Geoespaciais, composta pelo prof. Dr. Gilberto de Oliveira Mendes (Universidade Federal de Uberlândia-UFU), orientador da candidata, prof. Dr. Vinicius Francisco Rofatto (Universidade Federal de Uberlândia-UFU), e prof. Dr. Eder Matsuo (Universidade Federal de Viçosa - UFV).

Iniciando os trabalhos o presidente da mesa, Dr. Gilberto de Oliveira Mendes, apresentou a Comissão Examinadora e a candidata, agradeceu a presença do público, e concedeu à discente a palavra para a exposição do seu trabalho. A duração da apresentação da discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir a candidata. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a banca, em sessão secreta, atribuiu o resultado final, considerando a candidata:

Aprovada.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por Gilberto de Oliveira Mendes, Professor(a) do Magistério Superior, em 28/11/2025, às 10:08, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por Vinicius Francisco Rofatto, Professor(a) do Magistério Superior, em 28/11/2025, às 10:20, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por Eder Matsuo, Usuário Externo, em 05/12/2025, às 08:22, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site

https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador 6897863 e o código CRC 01AF8463.



AGRADECIMENTOS

Agradeço primeiramente a Deus por me conceder saúde, sabedoria e força para concluir esta etapa tão importante da minha vida, e a todas as forças do universo que me permitiram perseverar.

À minha família, meu filho Arthur Freitas Martins Giani e ao meu esposo André Luiz de Souza, minha eterna gratidão pelo amor incondicional, apoio e incentivo em todos os momentos. Vocês são meu porto seguro e minha maior inspiração.

Ao meu orientador Prof. Dr. Gilberto de Oliveira Mendes expressei minha mais profunda gratidão pela dedicação, paciência e conhecimento compartilhados.

Ao meu coorientador, Prof. Dr. George Deroco Martins, seu direcionamento e orientações foram fundamentais para o sucesso deste trabalho.

Aos colegas do mestrado, em especial ao Filipe Vieira da Silva que contribuiu com todo o processamento de dados requerido, com muita dedicação, profissionalismo e entrega.

À Universidade Federal de Uberlândia, agradeço pela oportunidade de realizar este trabalho e pelo ensino de excelência que recebi. A instituição me proporcionou um ambiente propício para o desenvolvimento intelectual e pessoal.

Agradeço também a todos que, direta ou indiretamente, contribuíram para a realização deste trabalho. A cada um de vocês, minha sincera gratidão.

RESUMO

Um dos grandes desafios do cultivo de soja em terras tropicais é a baixa fertilidade dos solos, limitando a disponibilidade de fósforo, o que ocasiona o uso intensivo de fertilizantes químicos sintéticos, elevando os custos de produção, gerando impacto ambiental, além de promover a dependência da importação de insumos. Na agricultura moderna, os microrganismos benéficos têm se destacado como alternativas para a melhoria do uso eficiente dos nutrientes, especialmente aqueles capazes de solubilizar fosfatos inorgânicos de baixa solubilidade. Trata-se de uma abordagem promissora, ambientalmente sustentável e economicamente viável, que contribui para o aumento da disponibilidade desse nutriente para as plantas. A interação entre os inoculantes e a planta, além de complexa, pode manifestar-se de forma sutil. Nesses casos, métodos estatísticos clássicos são, em geral, incapazes de evidenciar diferenças de produtividade entre áreas tratadas, especialmente quando se considera toda a dinâmica envolvida nas diversas fases de desenvolvimento da cultura. Neste contexto, ferramentas geoestatísticas têm sido utilizadas para estimar relações espaciais entre a produção agrícola e os diferentes parâmetros de manejo das culturas, onde o objetivo deste estudo é avaliar o ganho de produtividade da soja associado à inoculação com microrganismos solubilizadores de fósforo por meio da análise estatística e geoestatística de dados de produtividade georreferenciados e, adicionalmente, verificar se variáveis espectrais derivadas de imagens multiespectrais explicam e predizem essa produtividade com base em algoritmos de aprendizado de máquina. Os resultados mostraram que o algoritmo Random Forest apresentou as melhores estimativas de produtividade da soja com um menor erro, aliado às imagens multiespectrais que agregaram informações das dinâmicas fenológicas da cultura, gerando estimativas assertivas e com um menor esforço físico. Com uma grande quantidade de dados de treinamento, a técnica de krigagem permitiu gerar uma superfície de ganho médio de 3,93% em produtividade, validando a eficácia do solubilizador de P.

Palavras-chave: geoestatística; *Glycine max*; microrganismo; produtividade.

ABSTRACT

One of the major challenges of soybean cultivation in tropical regions is the low soil fertility, which limits phosphorus availability and leads to the intensive use of synthetic chemical fertilizers. This increases production costs, creates environmental impacts, and promotes dependence on imported inputs. In modern agriculture, beneficial microorganisms have emerged as alternatives for improving nutrient use efficiency, especially those capable of solubilizing low-solubility inorganic phosphates. This represents a promising, environmentally sustainable, and economically viable approach that contributes to increasing the availability of this nutrient to plants. The interaction between inoculants and plants, in addition to being complex, can manifest subtly. In such cases, classical statistical methods are generally unable to reveal productivity differences between treated areas, particularly when considering the full dynamics involved in the various developmental stages of the crop. In this context, geostatistical tools have been used to estimate spatial relationships between agricultural yield and different crop management parameters. The aim of this study is to evaluate soybean yield gains associated with inoculation using phosphorus-solubilizing microorganisms through statistical and geostatistical analysis of georeferenced yield data and, additionally, to determine whether spectral variables derived from multispectral images can explain and predict this yield using machine-learning algorithms. The results showed that the Random Forest algorithm provided the best soybean yield estimates with the lowest error, supported by multispectral images that incorporated information about the crop's phenological dynamics, generating accurate estimates with reduced physical effort. With a large amount of training data, the kriging technique made it possible to generate a surface showing an average yield increase of 3.93%, validating the effectiveness of the P-solubilizing agent.

Keywords: geostatistics; *Glycine max*; microorganisms; productivity.

SUMÁRIO

1 INTRODUÇÃO.....	12
2 METODOLOGIA.....	14
2.1 Área de estudo	15
2.2. Aquisição dos dados	16
2.3. Pré-processamento dos dados	17
2.4. Processamento dos dados	21
2.5. Validação dos dados	21
2.6. Superfície estimada por krigagem	21
2.7 Estimativa do ganho em produtividade (%)	22
2.8 Análise estatística clássica.....	22
3 RESULTADOS E DISCUSSÃO.....	23
4 CONCLUSÃO.....	31
5 REFERÊNCIAS.....	32

1 INTRODUÇÃO

A soja (*Glycine max* (L.) Merrill) é uma cultura de grande importância econômica no Brasil, sendo o país o maior produtor e exportador agrícola mundial. Os sistemas de produção agrícola brasileiros são continuamente afetados pela baixa fertilidade dos solos, levando ao uso intensivo de fertilizantes químicos sintéticos, que elevam os custos de produção, geram impacto ambiental, além de promoverem a dependência da importação de insumos. Para suprir a demanda nacional do agronegócio por nutrientes, o País importa, em média, 70% dos fertilizantes nitrogenados e fosfatados e acima de 95% dos fertilizantes potássicos, sendo o quarto maior consumidor global de fertilizantes, atrás apenas da China, da Índia e dos Estados Unidos (Globalfert, 2025).

A disponibilidade de fósforo (P) no solo limita a produtividade de terras agrícolas tropicais. Fertilizantes ajudam, mas são caros e dependem de rocha fosfática, um recurso finito. Enquanto isso, o manejo com alto aporte de P em solos pobres vem aumentando no Brasil, onde os agricultores aplicam cerca de duas vezes mais do que o que colhem para manter altos rendimentos (Roy *et al.*, 2017).

Experimentos de longo prazo em solos brasileiros demonstraram que uma grande proporção (> 70%) do excedente de P adicionado via fertilizantes permanece no solo, principalmente em formas não prontamente disponíveis para as culturas (Pavinato *et al.*, 2020). Dessa forma, torna-se fundamental implementar alternativas ambientalmente sustentáveis e economicamente viáveis para aumentar a disponibilidade desse nutriente para as plantas, como o uso de inoculantes microbianos (Oliveira *et al.*, 2022).

Na agricultura moderna, os microrganismos benéficos são alternativas que possibilitam a melhoria no uso eficiente dos nutrientes, destacando-se os que são capazes de solubilizar fosfatos inorgânicos de baixa solubilidade. Dentre os microrganismos solubilizadores de fosfato (MSF) se destacam as bactérias dos gêneros *Acinetobacter*, *Bacillus*, *Burkholderia*, *Bradyrhizobium*, *Enterobacter*, *Mesorhizobium*, *Paenibacillus*, *Pantoea*, *Pseudomonas*, *Rhizobium* e *Serratia*, e fungos do gênero *Penicillium* e *Aspergillus* (Halder *et al.*, 1990).

A solubilização de fósforo por microrganismos é um fenômeno complexo, que depende de muitos fatores, como as condições nutricionais, fisiológicas e de crescimento da cultura. Dentre os mecanismos utilizados pelos microrganismos na transformação dos fosfatos pouco solúveis em solúveis, estão a ocorrência de processos como o desencadeamento de metabólitos de enzimas extracelulares e ácidos orgânicos, processos de acidificação, reações de trocas e quelatção (Abreu *et al.*, 2002). No entanto, ainda persistem lacunas quanto ao correto

entendimento do efeito da inoculação, especialmente no que se refere ao ganho em produtividade.

Na relação simbiótica entre a planta e os MSF's, o aumento na produtividade das culturas após a aplicação de produtos biológicos é, por vezes, tão sutil que modelos clássicos, como a análise de variância (ANOVA), são incapazes de demonstrar diferenças de produtividade entre os grupos de tratamento e controle em diversas situações. Isso ocorre porque tais modelos não consideram fatores que podem afetar a produtividade, como variáveis fisiológicas e a dependência espacial da produtividade entre a população de plantas em uma determinada parcela (Thay *et al.*, 2020).

Como alternativa aos métodos clássicos a predição, que é uma ferramenta de modelagem para encontrar padrões preditivos generalizáveis através do Método de Machine Learning - ML (Aprendizagem de Máquina), são utilizados nos estudos de experimentos agrícolas, entretanto, os resultados de pesquisas à campo não são claros a respeito da comparação entre análises estatísticas tradicionais e modelos de predição na comparação do ganho em produtividade entre uma área tratada e não tratada. Neste contexto, ferramentas geoestatísticas têm sido utilizadas para estimar relações espaciais entre a produção agrícola e os diferentes parâmetros de manejo das culturas. Em uma área comercial de milho tratada com bioestimulante foi possível evidenciar diferenças no ganho de produtividade utilizando interpoladores como Krigagem e modelos de predição espectral, com base em algoritmos de aprendizado de máquina. Em condições de colheita mecanizada, com grande quantidade de informações georreferenciadas de produtividade, dada a alta acurácia e precisão, recomenda-se o uso da Krigagem para gerar as superfícies de ganho de produtividade (Martins *et al.*, 2023).

Com base no exposto, o objetivo deste estudo é avaliar o ganho de produtividade da soja associado à inoculação com microrganismos solubilizadores de fósforo por meio da análise estatística e geoestatística de dados de produtividade georreferenciados e, adicionalmente, verificar se variáveis espectrais derivadas de imagens multiespectrais explicam e predizem essa produtividade com base em algoritmos de aprendizado de máquina.

2 METODOLOGIA

Os procedimentos metodológicos deste estudo contemplaram as seguintes etapas: aquisição dos dados, pré-processamento, processamento, validação, superfície estimada por krigagem e estimativa do percentual (%) do ganho em produtividade (Fig. 1).

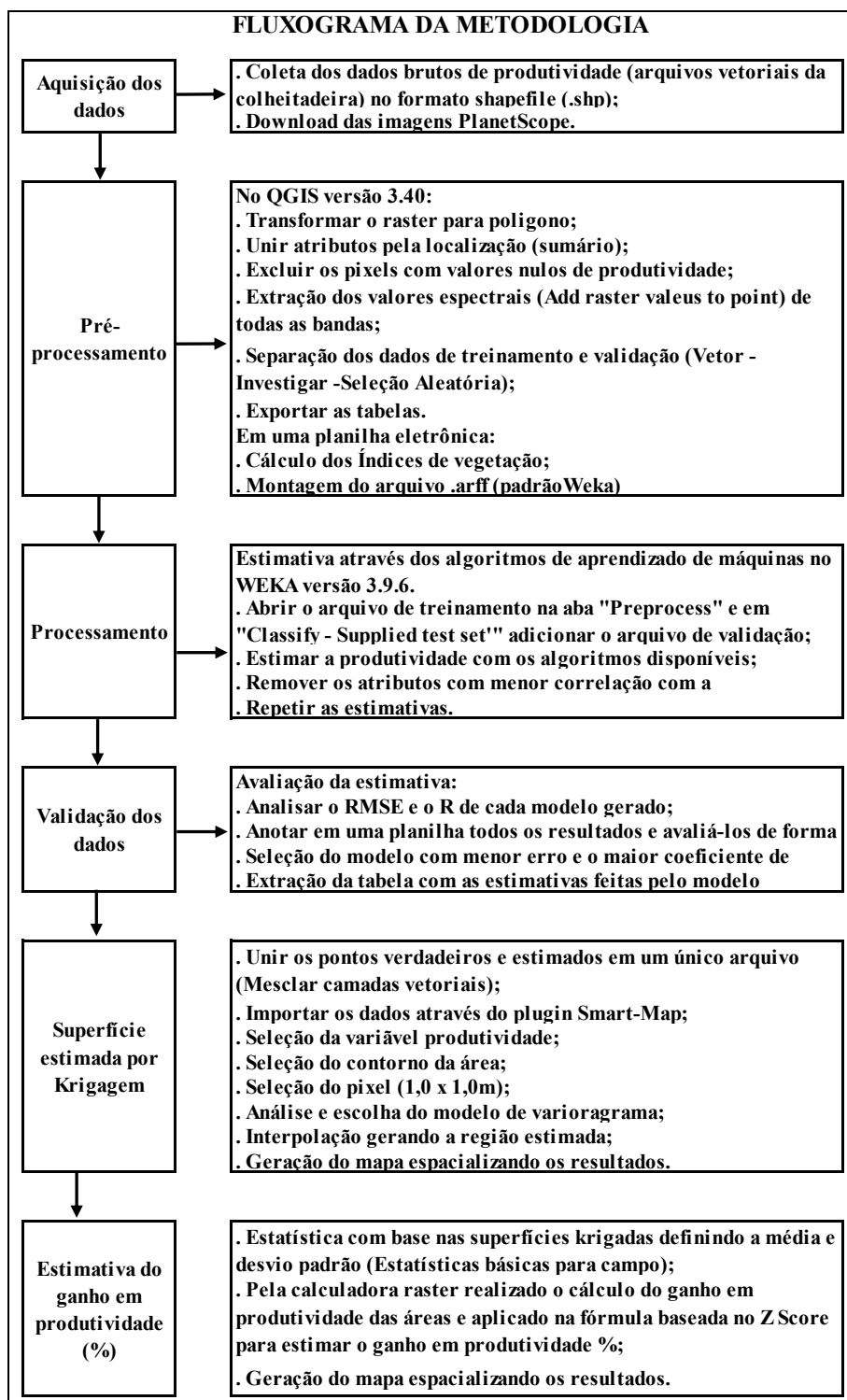


FIGURA 1 – Etapas dos procedimentos metodológicos aplicados no estudo. Fonte: A autora.

2.1 Área de estudo

O estudo foi realizado em 20 hectares (ha) de uma área comercial de soja, localizada no município de Araguari, mesorregião do Triângulo Mineiro, Minas Gerais, Brasil, coordenadas 18°50'18.87"Sul e 47°57'56.69"Oeste (Fig. 2). A área experimental está a uma altitude média de 928 m e possui solo classificado como Latossolo Vermelho distrófico. O clima da região é tropical com inverno seco, classificado como Aw (Köppen), com precipitação média anual de 1.370 milímetros e temperatura média anual de 24°C (INMET, 2023).

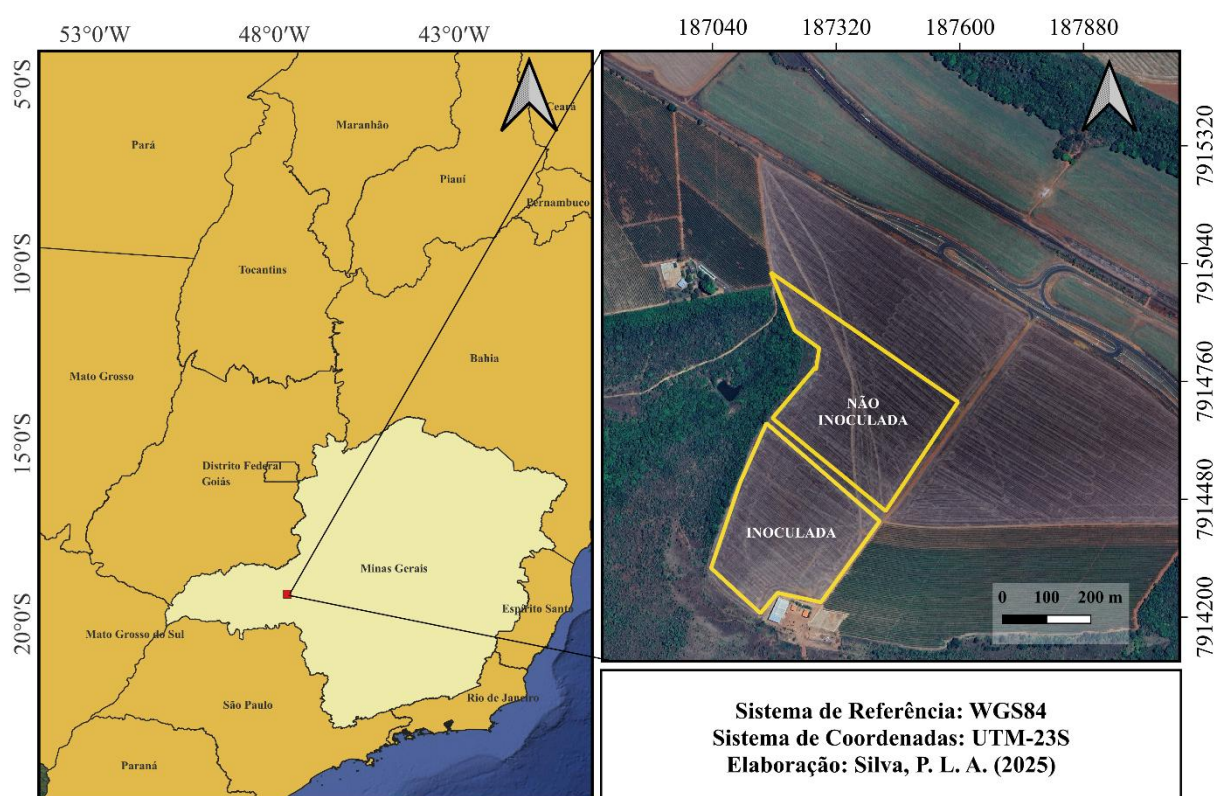


FIGURA 2 - Localização dos experimentos em Araguari, MG. Área Inoculada com solubilizador de P (I) e Área não inoculada (NI).

Na área o produtor plantou a variedade de soja CZ37B43 IPRO no dia 17/10/2024, com espaçamento entre linhas de 0,5 m e distribuição de 15 plantas por metro, com densidade aproximada de 300.000 plantas por ha.

Durante a implantação da cultura, toda a área foi tratada com uma mistura de produtos aplicados via sulco de plantio, na vazão de 50 L/ha, utilizando um pulverizador de jato dirigido tipo micron. A mistura foi composta por 0,15 L/ha de *Trichoderma harzianum* Rifai, ESALQ-

1306 2,0 x 10⁹ conídios viáveis/ml; 0,5 L/ha de *Bradyrhizobium japonicum* - Semia 5079 e 5080 5,0 x 10⁹ UFC/ml; 0,1 L/ha de *Azospirillum brasilense* - AbV5 e AbV6 2,0 x 10⁸ UFC/ml; 0,5 L/ha do ativador de rizosfera à base de NPK e Carbono orgânico e 0,3 L/ha de uma fonte de boro. Foi realizada adubação de plantio em ambas as áreas com 200 Kg/ha de uma fonte de fósforo (fórmula 1% N, 28% P, 5% S e 17% Ca), e 200 Kg/ha a lanço de cloreto de potássio (60% de K₂O). Como forma de não interferir no manejo da área comercial, apenas acrescentamos o produto com a mistura de MSF's (objeto deste estudo) em uma área separada.

A área experimental foi subdividida em duas partes de 10 ha, sendo uma área controle não inoculada (NI) e a outra inoculada (I) com 0,25 L/ha de produto à base de microrganismos solubilizadores de fósforo, que foi incorporado à calda aplicada no sulco de plantio. O produto é composto de *Streptomyces beta-vulgaris* B11 (I-3639) 1 x 10⁹ UFC/ml, *Burkholderia sp.* AGN02 (I-4360) 1 x 10⁹ UFC/ml, *Bacillus megaterium* AGN01 (I-4466) 1 x 10⁹ UFC/ml, *Rhizophagus irregularis* DAOM 197198 1%.

2.2. Aquisição dos dados

Para realização do estudo foram coletados no dia 17/02/2025 os dados brutos de produtividade georreferenciados das áreas I e NI, estimados em toneladas por ha a cada metro colhido, diretamente dos arquivos vetoriais da colheitadeira John Deere, modelo S760. O maquinário possuía o sistema Active Yield, com sensores automatizados no graneleiro que calibram a máquina medindo o grão durante o abastecimento do tanque. As medições foram associadas a um receptor GNSS, sistemas de deslocamento da máquina e largura de corte.

Para gerar o modelo espectral para a predição da produtividade foram utilizadas imagens da constelação de satélites PlanetScope em três fases distintas do desenvolvimento da cultura, vegetativo – V4 (26/11/2024), reprodutivo – R1 (09/12/2024) e pré-colheita – R8 (07/02/2025). As imagens se demonstraram adequadas para a finalidade do estudo pois se destacam por sua alta resolução espacial e frequência de imageamento diária, permitindo o monitoramento quase em tempo real de qualquer local na Terra. O sensor, Bayer Mask CCD tipo PS2, captura os comprimentos de onda do vermelho (590-670 nm), verde (500-590 nm), azul (455-515 nm) e infravermelho próximo (780-860 nm) e resolução radiométrica de 12 bits (Oliveira e Oliveira, 2016).

A caracterização dos estádios de desenvolvimento da planta de soja é essencial para a descrição dos vários períodos que a lavoura atravessa durante o ciclo da cultura, o sistema de classificação mais utilizado e detalhado foi proposto por Ferhr e Caviness (1977), onde foi

sugerido que o ciclo da soja fosse dividido em duas partes, a vegetativa que é representada para letra V e a fase reprodutiva com a letra R, com exceção dos estádios VE (emergência) e VC (cotilédone), as letras V e R são seguidas de índices numéricos que identificam estádios específicos, nessas duas fases de desenvolvimento da planta (fig. 3).

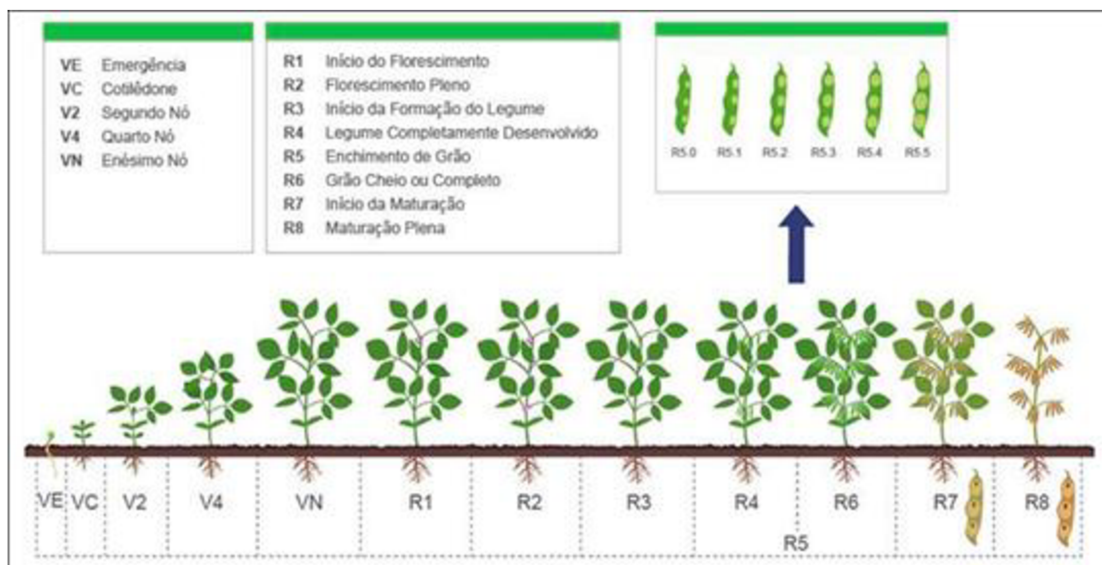


FIGURA 3 - Escala fenológica da soja. Fonte FEHR, CAVINESS (1977).

Após atingir o estágio inicial de maturidade (R7) as vagens apresentam a cor cinza, bege ou marrom, como a planta está se aproximando da senescência, o teor de clorofila das folhas diminui e a cor fica amarela/cinza, o que pode resultar em mudanças substanciais na refletância espectral. Portanto, é possível quantificar o nível de maturidade da planta usando as combinações de refletância espectral em bandas de ondas sensíveis às características do dossel da cultura, como teor de clorofila, teor de nitrogênio e estresse hídrico (Zhou *et al.*, 2019).

Diferentes fases de desenvolvimento da cultura permitem um maior o número de imagens, com uma maior resolução temporal, onde a quantidade de bandas nos permite explorar o cálculo de diversos índices correlacionados com a produtividade, agregando mais informações aos modelos para melhorar a estimativa da produtividade pelos métodos de ML.

2.3. Pré-processamento dos dados

No pré-processamento os dados brutos (arquivos vetoriais da colheiteira) foram processados no QGIS versão 3.40. Um total de 6.852 e 6.357 pontos foram amostrados para a

área I e NI, respectivamente. Com isso, foram selecionados somente os pixels que intersectam os pontos de produtividade para garantir que a resposta espectral seja referente ao dossel da lavoura e esteja vinculada com os valores de produtividade coletados.

Para a execução desse trabalho foi utilizado o método Holdout, com os dados sendo separados em um conjunto contendo 80% dos pontos para a geração do modelo e outro acervo com os 20% restante dos pontos para a validação. Os pontos foram divididos em dois arquivos, um para o treinamento e outro para validação do modelo, por meio de uma ferramenta de seleção aleatória dos pontos, disponível no QGIS. Este procedimento foi aplicado para os pontos das áreas NI e I.

Por se tratar de amostras com mais de 3.500 pontos avaliados, temos que um subconjunto com 20% dos pontos pode ser considerado representativo para toda a lavoura, além de ser um método com baixo esforço computacional, tornando a utilização do método uma escolha viável e validada em outros projetos científicos (Allibhai, 2018).

Na sequência, aos dados foram acrescentados os índices de vegetação (IV) calculados em uma planilha eletrônica, foram considerados os que apresentaram grande correlação com os dados de produtividade, Normalized Difference Vegetation Index (NDVI), Green Normalized Difference Vegetation (GNDVI) e o Green Chlorophyl Index (CIgreen ou CIG).

Os IV's agregam informações a respeito do estado nutricional da lavoura, biomassa e pigmentação. Por representarem a resposta espectral do dossel de plantas, possibilitam identificar variações no desenvolvimento das plantas e localizar os fatores causadores. O NDVI indica variações na biomassa de um cultivo vegetal ao longo do seu ciclo de desenvolvimento. Ele é um dos índices mais utilizados para análises de vegetação em diversas áreas do conhecimento, em especial, nos estudos de ciências agrárias. Seu uso permite o monitoramento do nível de absorção de energia e clorofila da planta. Em sua estimativa, considera a relação antagônica da reflectância nas bandas do vermelho (RED) e do infravermelho próximo (NIR – do inglês Near InfraRed) frente ao acúmulo de biomassa. Entre as aplicações mais frequentes estão as inferências sobre parâmetros biofísicos, monitoramento e a modelagem do rendimento de grãos ou mesmo a indicação da sanidade das plantas, como doenças e deficiências nutricionais. Tal índice é calculado a partir da razão normalizada entre bandas do vermelho e infravermelho próximo, pela soma da reflectância das mesmas bandas (Eq. 1).

$$NDVI = \frac{NIR-RED}{NIR+RED} \quad (1)$$

O GNDVI possui a capacidade de avaliar as variações de biomassa verde na escala do dossel considerando o NDVI, NIR e a banda verde (GREEN), é calculado de utilizando-se a Equação 2:

$$GNDVI = \frac{NIR - GREEN}{NDVI + GREEN} \quad (2)$$

O CIG é utilizado de forma complementar ao NDVI, permitindo avaliar os estádios iniciais de desenvolvimento da soja, já que correlaciona o índice de área foliar com a clorofila presente (Hatfield e Prueger, 2010) (Eq. 3).

$$CI_{green} = \frac{NIR}{GREEN-1} \quad (3)$$

Para finalizar o pré-processamento realizou-se a preparação do arquivo de entrada do WEKA (Waikato Environment for Knowledge Analysis), que permitiu o teste dos modelos. Neste arquivo constava as medidas de produtividade, os valores espectrais de cada banda relacionados aos pontos e aos IV's que foram definidos. Os arquivos das áreas NI e I foram formatados no tipo .arff (Attribute – Relation File Format), padrão requerido pelo software em questão.

Vale ressaltar que o Support Vector Machine (SVM), Random Forest (RF), Rede Neural - Neural Net (NN) ou Random Tree (RT) são alguns exemplos de algoritmos que o software disponibiliza para uso na estimativa de uma variável de interesse. Além da estimativa, ele retorna ao usuário o Coeficiente de Correlação (R) e o Root Mean Squared Error (RMSE), para a avaliação do modelo treinado. Também é possível pela ferramenta “Select Attributes” ranquear os atributos de acordo com a correlação de cada um com a produtividade, da maior para menor, possibilitando a exclusão deles, evitando assim uma possível confusão no algoritmo que influencie negativamente a estimativa.

O ML foi utilizado nesta fase porque apresenta-se como um conjunto de técnicas de inteligência artificial que visam reconhecer padrões em base de dados, agregando na qualidade na geração de informação e conhecimento, auxiliando e melhorando os processos de tomadas de decisão. Entre as principais tarefas que se podem ser executadas com a aplicação de técnicas de aprendizagem de máquina está a predição (classificação e estimação), que no caso deste

estudo foi a da produtividade. Ele ser dividido em dois tipos: supervisionado e não-supervisionado. O supervisionado, consiste no modelo (ou algoritmo) ser construído a partir dos dados de entrada, que são apresentados na forma de pares ordenados (entrada — saída desejada). Já no não-supervisionado não existe a informação dos rótulos históricos, ou seja, não temos as das saídas desejadas a serem estimadas e, por este motivo, dizemos que nossos dados são não-rotulados. Assim, o algoritmo não recebe durante o treinamento os resultados esperados, devendo descobrir por si só, por meio da exploração dos dados, os possíveis relacionamentos entre eles (Escovedo, 2020).

Random Forest trata-se de uma combinação de preditores de árvores tais que cada árvore depende dos valores de um vetor aleatório amostrado de forma independente e com a mesma distribuição para todas as árvores da floresta (Breiman, 2001). Além disso, não é exigido que os dados admitam distribuição normal. Por este motivo, no RF o processamento de dados advindos de fontes diferentes é facilitado, como é o caso da pesquisa onde temos dados mensurados em campos e outros derivados de imagens de satélites (Budu, 2023).

A NN é considerada uma máquina ou um processador modelado para trabalhar como o cérebro humano, o qual têm a capacidade de armazenar conhecimento, processar as informações e retornar um resultado baseados nas experiências passadas. Ela passa por um processo repetitivo de interações durante o treinamento do algoritmo com os dados na camada dos neurônios de entrada, adiante eles são processados pelas camadas ocultas e geram os neurônios de saída, que são os resultados. Este algoritmo traz como benefício a não linearidade, adaptabilidade, o mapeamento de entrada-saída, informação contextual e resposta a evidências (Haykin, 2001).

E o weka também retorna com a opção do algoritmo SMOreg, que é a implementação de uma versão de máquina de vetores de suporte (SVM), que faz o uso de métodos de aprendizado de máquina para classificação, detecção de outliers e regressão. Para isso, o algoritmo utiliza-se de vetores suporte (ou registros essenciais) do conjunto de treino para criar hiperplanos que separam os dados em duas classes, além de calcular as margens dos hiperplanos para identificar o melhor dentre todos os criados (Silva *et al.*, 2017).

Visando a aplicação do aprendizado de máquina de forma supervisionada é necessário trabalhar as melhores ferramentas para tratar os dados em alta dimensionalidade (hiperespectrais), pois oferecem um poder discriminante bem mais elevado do que dados tradicionais (em baixa dimensionalidade).

2.4. Processamento dos dados

No WEKA (versão 3.9.6.), foi adicionado o arquivo de treinamento na aba "Preprocess" e na aba "Classify - Supplied test set" o de validação. Em seguida, foram testados os algoritmos e anotado em uma planilha eletrônica os valores do RMSE e R de cada estimativa computada. Tais parâmetros são calculados segundo as equações 4 e 5 respectivamente.

$$R = \frac{\sum (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum (y_i - \bar{y})^2 \sum (\hat{y}_i - \bar{\hat{y}})^2}} \quad (4)$$

$$RMSE (\%) = \frac{\sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}}{\frac{\sum y_i}{n}} * 100 \quad (5)$$

Onde: y_i representa o valor medido em campo de produtividade; $\bar{y}$ expressa a média da produtividade medida em campo; $\hat{y}_i$ indica o valor estimado de produtividade; $\bar{\hat{y}}$ representa a média da produtividade estimada e n o número de amostras.

2.5. Validação dos dados

Para validação do modelo foi realizada a análise do R (Eq. 4) e do RMSE (Eq. 5), onde o que apresentasse o maior coeficiente de correlação e o menor erro seria definido como a melhor performance para os dados e melhor acurácia. Foram considerados como atributos os dados de produtividade, os valores de todas as bandas da imagem coletada e os índices de vegetação.

Na sequência, foi extraída uma tabela dos valores preditos pelo RF (algoritmo que apresentou a melhor performance) e os mensurados em campo das áreas, sendo para NI, arquivo contendo os dados da área não inoculada com a predição da área I, e outro com os dados da área I com a predição da NI.

2.6. Superfície estimada por krigagem

Para a análise geoestatística no QGIS, foi realizada a junção dos pontos verdadeiros e estimados em um único arquivo (mesclar camadas vetoriais) e feita a interpolação por krigagem ordinária. Os dados foram importados através do plugin Smart-Map, selecionado o contorno das áreas, seleção do pixel de 1,0 x 1,0m, logo após o programa retornou com o melhor modelo de semivariograma para os parâmetros definidos. Os semivariogramas foram gerados para cada um dos cenários amostrais avaliados, NI e I.

Após este procedimento foi realizada a interpolação e com o objetivo de visualizar a distribuição espacial das estimativas ao longo da lavoura. Foram gerados os mapas com os

resultados, com a mesclagem dos pontos de treinamento e validação da área NI mais predição da área I, e outro com mesclagem dos pontos de treinamento e validação da área I mais predição.

2.7 Estimativa do ganho em produtividade (%)

Visando compatibilizar a escala da região não inoculada com a inoculada para estimar o percentual do ganho em produtividade, foi realizada a normalização dos dados da seguinte forma: no QGIS foi aplicada a estatística com base nos dados das superfícies krigadas definindo a média e desvio padrão (em estatísticas básicas para campo) de cada uma delas. Pela calculadora raster foi configurada a equação 3 para calcular a área Z_1 (se refere à área NI), o mesmo procedimento foi aplicado para a equação 4 e cálculo da área Z_2 (área I) com isto gerou os mapas das áreas Z_1 e Z_2 , e, pela diferença entre eles foi gerado o mapa espacializando os resultados.

$$Z_1 = \frac{X_1 - \mu_1}{\sigma_1} \quad (3)$$

$$Z_2 = \frac{X_2 - \mu_1}{\sigma_1} \quad (4)$$

$$\% \text{ Ganho} = \frac{100 \times [(Z_2 - Z_1) \cdot \sigma_1]}{\mu_1 + Z_1 \cdot \sigma_1} \quad (5)$$

Onde,

Z_1 = Z score da população 1 (NI), na escala da própria população 1

Z_2 = Z score da população 2 (I), na escala da população 1

X_1 = dados da superfície krigada como referência a área NI

X_2 = dados da superfície krigada como referência a área I

μ_1 = média da população 1

σ_1 = desvio padrão da população 1

2.8 Análise estatística clássica

Também realizamos esta análise para avaliar o comportamento dos dados coletados caso somente fosse esta opção de análise a ser executada. Por estatística clássica entende-se aquela que se utiliza de parâmetros como média e desvio padrão para representar um fenômeno e se baseia na hipótese principal de que as variações de um local para outro são aleatórias. Transformar dados observados em informação condensada que permita inferências rigorosas (Fischer, 1992).

Com os dados brutos de produtividade das áreas I e NI foi realizado o teste não-paramétrico Wilcoxon. Tal teste é frequentemente empregado como alternativa ao teste t de Student

pareado, quando os dados não satisfazem os pressupostos paramétricos, como a distribuição normal, baseando-se nos postos (ranks) dos valores em vez dos valores em si.

3 RESULTADOS E DISCUSSÃO

Nas tabelas 1 e 2 constam os resultados obtidos para os modelos contemplando as três datas de coleta das imagens comparados ao do arquivo contendo todos os dados de todas as datas (TT).

TABELA 1 – Desempenho dos algoritmos para a área inoculada.

Data	R	RMSE (%)	Algoritmo
26/11/2024	0,1815	8,50	REDE NEURAL
26/11/2024	0,2018	8,22	SMOreg
26/11/2024	0,1805	8,43	RANDOM FOREST
09/12/2024	0,1519	8,78	REDE NEURAL
09/12/2024	0,1348	8,30	SMOreg
09/12/2024	0,1486	8,53	RANDOM FOREST
07/02/2025	0,2543	8,31	REDE NEURAL
07/02/2025	0,2794	8,03	SMOreg
07/02/2025	0,2379	8,28	RANDOM FOREST
TT	0,2936	8,47	REDE NEURAL
TT	0,3293	7,90	SMOreg
TT	0,3435	7,88	RANDOM FOREST

Fonte: A autora.

TABELA 2 – Desempenho dos algoritmos para a área não inoculada.

Data	R	RMSE (%)	Algoritmo
26/11/2024	0,3949	10,6	REDE NEURAL
26/11/2024	0,3771	9,16	SMOreg
26/11/2024	0,4615	8,75	RANDOM FOREST
09/12/2024	0,4328	10,75	REDE NEURAL
09/12/2024	0,4133	9,00	SMOreg
09/12/2024	0,4022	9,06	RANDOM FOREST
07/02/2025	0,5154	10,52	REDE NEURAL
07/02/2025	0,5058	8,52	SMOreg
07/02/2025	0,5029	8,52	RANDOM FOREST
TT	0,5559	10,54	REDE NEURAL
TT	0,5407	8,32	SMOreg
TT	0,5811	8,03	RANDOM FOREST

Fonte: A autora.

O algoritmo que apresentou o melhor desempenho tanto para os dados da área I (RMSE 7,88%) quanto para a NI (RMSE 8,03%), foi o RF considerando os dados TT, portanto com

precisão na previsão do modelo em 91,97% e 92,12% respectivamente. Avaliando os valores da correlação entendemos que uma correlação alta não garante precisão absoluta, mas indica que o modelo acerta a tendência geral do fenômeno. Por isso, recomenda-se usar correlação apenas como medida complementar, junto com métricas de erro (Willmott e Matsuura, 2005).

Apesar do algoritmo SMOreg ter demonstrado resultado muito próximo ao RF, este foi escolhido por já ter demonstrado sua robustez na modelagem da produtividade da soja na análise de diferentes fases de desenvolvimento (Santos, 2025). RF consegue minimizar o erro de previsão através de um conjunto de árvores de decisão além da funcionalidade apresentada para situações que envolvem grandes conjuntos de dados e de diferentes tipos, como numéricos, binários e categóricos (Breiman, 2001) (Budu, 2023).

Na escolha do algoritmo, o resultado de melhor desempenho, com os dados de produtividade aliados aos de todas as imagens coletadas em diferentes datas e aos IV's, aprimoraram a estimativa permitindo capturar dinâmicas fenológicas da cultura, além de minimizar as variações ocasionais que podem ocorrer no início do desenvolvimento da cultura ou próximas ao plantio. Estimativas feitas apenas no início da fase vegetativa, são comprometidas, pois o solo se encontra mais exposto, o que influencia nas respostas espectrais das imagens e tende a saturar alguns índices, como NDVI. Já quando são realizadas apenas nos estágios mais avançados, especialmente a partir do enchimento de grãos (R5 e R6), reduz a sensibilidade do modelo devido ao efeito de saturação espectral (Santos, 2025).

Os semivariogramas de cada um dos cenários amostrais avaliados, NI e I, constam na figura 4 com os detalhes dos seus parâmetros, a saber, o efeito pepita (C_0), o patamar ($C_0 + C$) e o alcance (A), a proporção da variabilidade total dos dados (R^2) e o RMSE são apresentados na tabela 3.

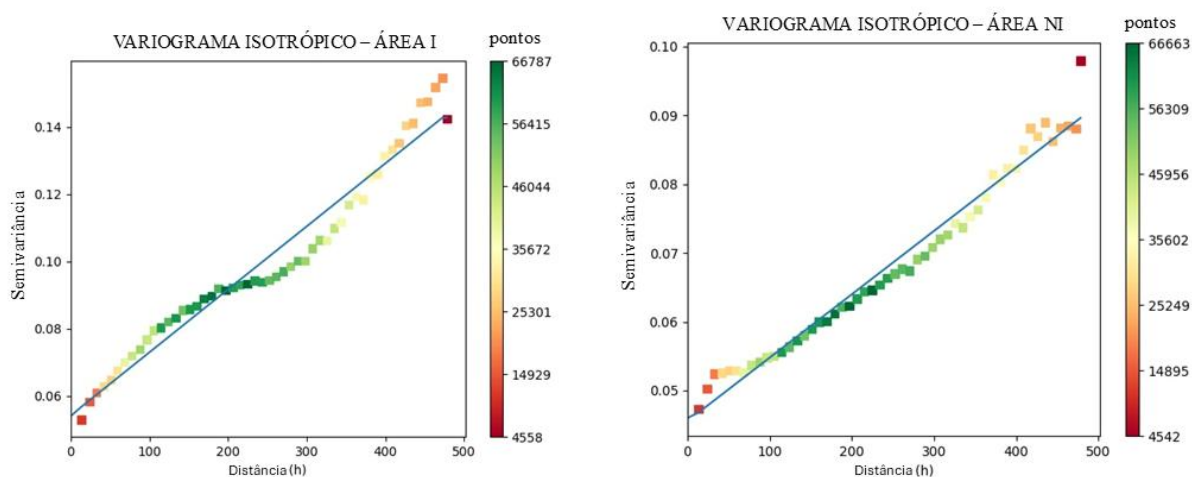


FIGURA 4 - Semivariogramas das áreas krigadas I e NI. Fonte: A autora.

TABELA 3 – Parâmetros dos semivariogramas

Parâmetros	ÁREA I (LINEAR TO SILL)	Área NI (Linear)
Co	0,054	0,046
Co + C	0,143	0,073
A	472,458	297,692
RMSE	0,002	0,000
R ²	0,957	0,974

Fonte: A autora.

Nota-se que em ambos os modelos o efeito pepita (Co) foi pequeno, próximo de zero, indicando alta continuidade espacial nos pontos próximos, sem ruído/erro significativo nos dados. O melhor modelo de semivariograma que o programa retornou para a área NI foi do tipo “linear”, indicando que a dependência espacial continua a aumentar indefinidamente com a distância, onde todos os pontos estão de alguma forma espacialmente correlacionados, não importando quão distantes estejam. Para a área I foi “linear to sill”, a semivariância aumentou linearmente com a distância h até atingir o patamar (sill) de 0,143 a uma distância de 472,458, e a partir daí apresenta uma tendência a ser constante, demonstrando que a variável estudada, produtividade, apresenta uma correlação espacial clara que se estabiliza a uma certa distância.

Os semivariogramas das áreas krigadas evidenciaram a subida gradual e alcance longo, mostrando que a variável produtividade é contínua e a dependência espacial se estende por distâncias maiores (dependência espacial clara). Na área I ele atinge um patamar, onde o método identificou o alcance para considerar as amostras vizinhas de forma eficiente, indicando que o método aplicado forneceu uma maior homogeneidade às áreas em estudo.

Vale ressaltar que alguns cuidados são requeridos com modelos puramente lineares, por não apresentar um sill teórico finito, por isto foi adotado um alcance conservador (entre 400 e 500 m). A Krigagem favorece a aplicação neste cenário por ser uma ferramenta de interpolação que utiliza a probabilidade para atribuir os pesos às amostras, minimizando o erro (Carvalho e Vieira, 2001).

Nas Figuras 5 e 6 estão representados os mapas com os resultados da mesclagem dos pontos de treinamento e validação com os preditos, neles pode ser constatado a geração de uma

superfície interpolada sem ruídos, permitindo a identificação clara das tendências espaciais, podendo identificar as regiões com menor e maior produtividade.

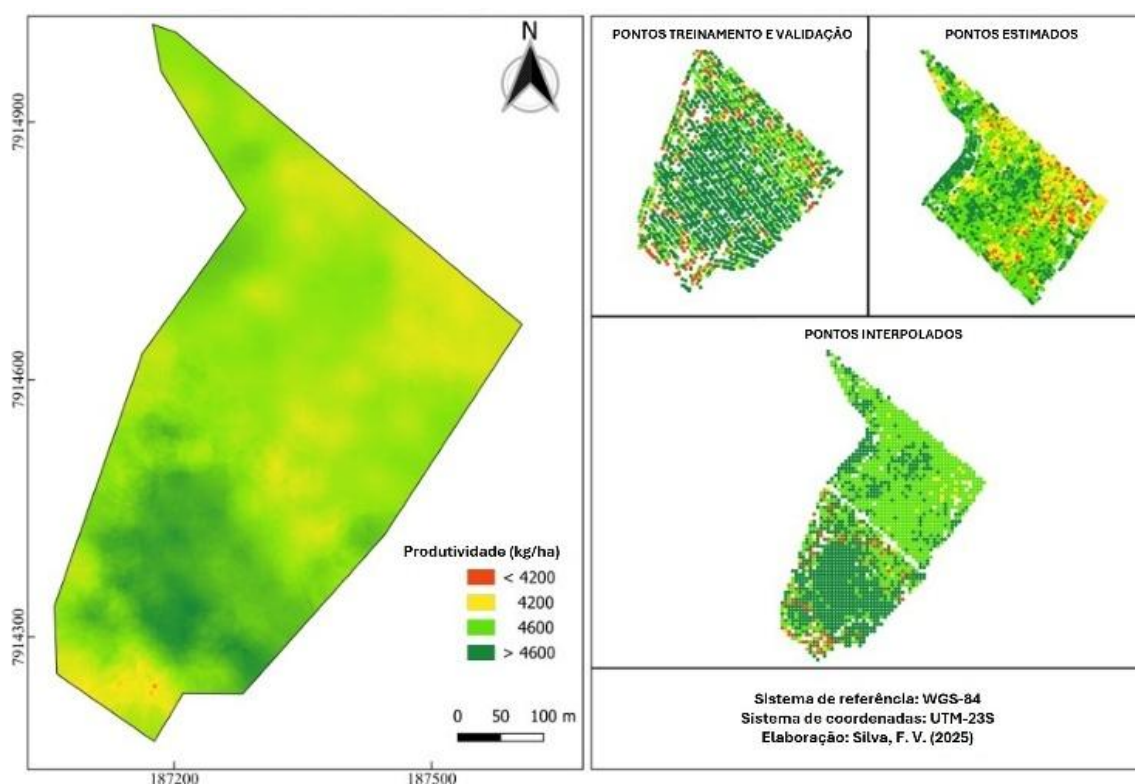


FIGURA 5 - Mapa da predição da produtividade, área NI tendo como referência a área I.

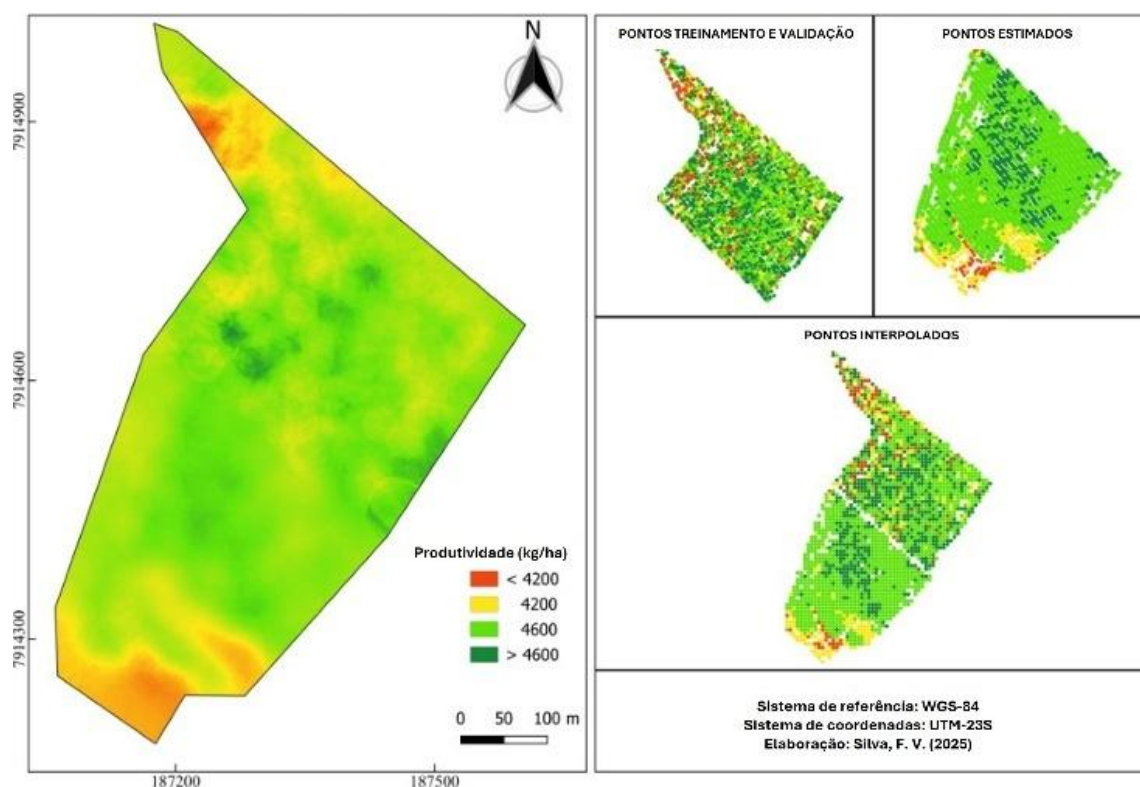


FIGURA 6 - Mapa da predição da produtividade, área I tendo como referência a área NI.

Na figura 7 está representado o mapa da diferença entre as áreas Z_1 e Z_2 , espacializando os resultados da estimativa da predição do ganho percentual em produtividade se fosse aplicado o solubilizador de P em toda a área do estudo.

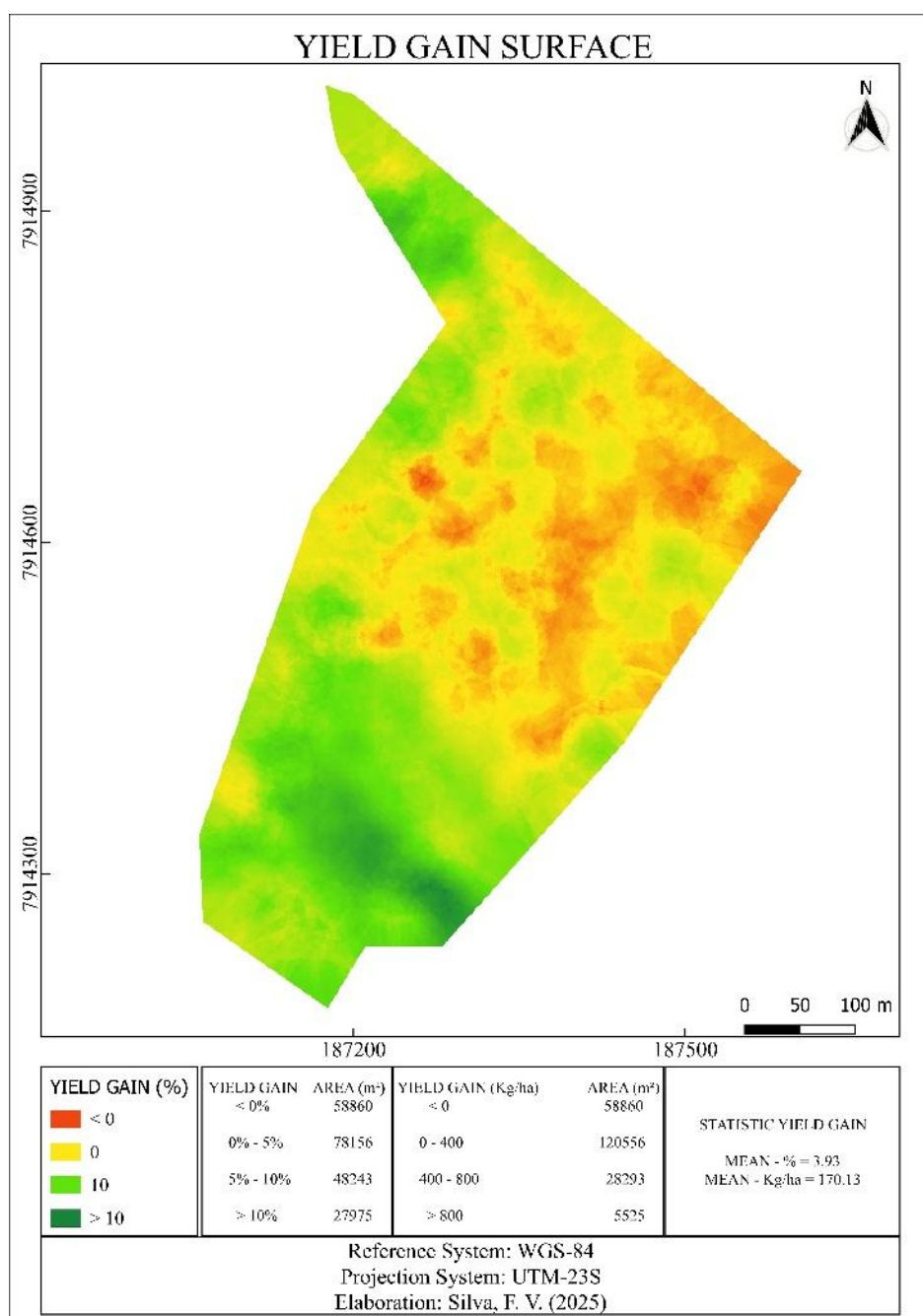


FIGURA 7 - Mapa da predição do percentual de ganho em produtividade

Notamos que pela predição do ganho em produtividade ao aplicar o solubilizador de P em toda a área, aproximadamente 213.234 m² (21,32 ha) teríamos um ganho médio de 3,93% em produtividade, equivalente a 0,170 ton/ha. Podendo alcançar ganhos acima de 5% em 36% da área, ou seja, em 76.218 m² (7,62 ha). A análise dos dados demonstra que o solubilizador de P foi benéfico para a cultura. Esta metodologia também demonstrou eficiência em outro estudo

utilizando inoculantes no tratamento de sementes de soja, por meio de uma superfície estimada de produtividade foi evidenciado o ganho em produtividade (Silva e Martins, 2023).

O teste não-paramétrico Wilcoxon foi escolhido porque os dados não atendiam pressupostos para teste t, onde o resultado está representado na figura 8.

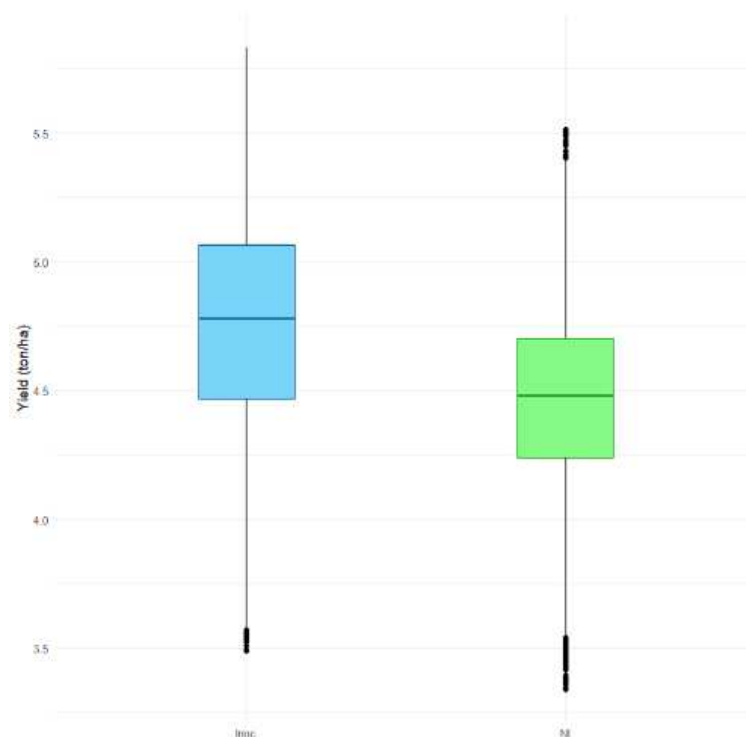


FIGURA 8 - Bloxpot do teste de Wilcoxon, comparando a diferença em produtividade entre a área Inoculada e Não Inoculada.

O teste apontou que a média de produtividade da área inoculada foi de 4,75 ton/ha, sendo superior ($p < 0.05$, Wilcoxon rank sum test) à não inoculada que apresentou 4,46 ton/ha. Com isto foi alcançado o ganho de 0,290 ton/ha (6,6 %) em produtividade.

O bloxpot compara a produtividade (ton/ha) entre dois grupos, área Inoculada e Não Inoculada. Podemos observar que na área I a mediana está entre 4.8–4.9 ton/ha, com uma amplitude interquartilica (IQR) relativamente larga, indicando maior variação interna. A distribuição na parte superior da caixa é maior indicando uma leve assimetria para cima (mais dispersão nos valores altos). Para os valores extremos (outliers), observa-se alguns pontos abaixo do bigode inferior (em torno de ~3.5 ton/ha), sugerindo valores atípicos na parte baixa.

Já na observação dos dados da área NI, comparados ao da área I a mediana está menor, na faixa de 4.5 ton/ha, o IQR apresenta-se um pouco menor também, indicando menos

variabilidade no grupo; a distribuição apresenta uma caixa mais deslocada para baixo, com valores de produtividade menores no geral. Os outliers possuem alguns valores altos isolados (entre 5.5–5.6 ton/ha), indicando possíveis outliers superiores.

As análises apontam que o tratamento com o solubilizador de P apresenta rendimento mediano mais alto, tem maior dispersão (caixa mais alta), sugerindo a resposta em produtividade para a inoculação, com uma assimetria positiva (tendência a valores altos). A área sem o solubilizador apresenta os dados de produtividade mais consistente, mas com a média menor. Para os outliers a área tratada tem outliers baixos, enquanto a NI tem outliers altos (possíveis plantas com desempenhos extremos em produtividade).

Pela análise apresentada fica evidente que há muita variabilidade nos dados e outliers, onde pela clássica (ANOVA, teste t, média, erro padrão etc.) assume a normalidade dos resíduos, homogeneidade de variâncias, ausência de outliers e distribuição simétrica dos dados, essas suposições são violadas, levando a estimativas de média distorcidas, erros-padrão inflados ou subestimados, perda de poder estatístico e conclusões incorretas. Neste caso, assume a independência entre as observações e faz inferências generalizadas sobre a população, desconsiderando toda a dinâmica envolvida nas diversas fases de desenvolvimento da cultura.

Neste cenário torna-se primordial implementar modelos que aprendem padrões diretamente dos dados, sem precisar assumir normalidade ou homogeneidade de variâncias, como as técnicas de Machine Learning, e, ao utilizar a krigagem como ferramenta geoestatística, melhoramos a capacidade preditiva do modelo, assumimos a dependência (autocorrelação espacial) entre as amostras, acrescentamos o elemento fundamental na análise que é a localização geográfica (coordenadas), aliado às informações das bandas das imagens capturadas em diversas fases de desenvolvimentos da cultura e aos IV's que aprimoram as estimativas. Tal método permite modelar fenômenos espacialmente contínuos, entender a variabilidade espacial e prever valores em locais não amostrados. Logo, a espacialização dos dados permite comparações mais diretas com outros fatores como ambientais, agronômicos, e edafoclimáticos, que facilitam a análise e a tomada de decisão para o manejo da cultura. Recomendamos também onde não é possível evidenciar diferenças estatísticas entre área inoculada e sem inoculação.

Há espaço para aprimoramento para buscar maior exatidão e precisão, como utilizar imagens de outros estádios fenológicos, aliados a utilização de outros índices de vegetação que sejam mais relevantes para cada fase estudada. Considerar diferentes variáveis agronômicas como variabilidade da disponibilidade de P no solo em pontos georreferenciados, dados climáticos e

edáficos. Reforçamos também que técnicas de modelagem podem ser melhoradas como aplicar os métodos de ML e após a predição executar a krigagem do resíduo, ou aplicar outros métodos de krigagem como forma comparativa para aperfeiçoar os modelos.

4 CONCLUSÃO

Neste estudo estimamos o ganho percentual em produtividade da soja ao aplicar solubilizador de P via sulco de plantio, utilizando variáveis espectrais derivadas de imagens multiespectrais para predição da produtividade com base em algoritmos de aprendizagem de máquina, aliados a aplicação da técnica de Krigagem Ordinária por meio do plugin Smart-Map em uma área comercial com grande quantidade de dados amostrais georreferenciados. A metodologia aplicada apresentou ganho médio de 3,93%, podendo obter acima de 5% em 36% da área.

A análise proposta demonstrou que o solubilizador de P foi benéfico para a cultura, sendo uma ferramenta válida e promissora, pois ao acrescentar dados espectrais ao modelo, são agregadas informações importantes quanto à saúde da planta, gerando estimativas assertivas e com um menor esforço físico.

Em condições em que há uma grande quantidade de dados de produtividade georreferenciados, com utilização de colheita mecanizada, recomendamos o uso da Krigagem para gerar a superfície de ganho de produtividade.

5 REFERÊNCIAS

- ABREU, C. S.; FIGUEIREDO, J. E. F.; OLIVEIRA, C. A.; SANTOS, V. L.; GOMES, E. A.; ALVES, L.; MENDONZA, E. A.; SILVA FILHO, G. N. **Microrganismos solubilizadores de fosfatos e o crescimento de pinus e eucalipto**. Revista Brasileira de Ciência do Solo, v. 36, n. 4, p. 939-947, 2002. Disponível em: <https://doi.org/10.1590/S0100-06832002000400011>. Acesso em: 16 ago. 2025.
- ALLIBHAI, E. **Hold-out vs. cross-validation in machine learning**. [S. l.], 2018. Disponível em: <https://medium.com/@ejaz/holdout-vs-cross-validation-in-machine-learning-7637112d3f8f>. Acesso em: 16 ago. 2025.
- BREIMAN, L. **Random Forests**. Statistics Department - University of California, Berkeley, CA 94720, 2001. Disponível em: <https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf>. Acesso em: 18 set. 2025.
- BUDU, E. **Random Forest vs. extremely: randomized trees**. [S. l.], 2023. Disponível em: <https://www.baeldung.com/cs/random-forest-vs-extremely-randomized-trees>. Acesso em 15 set. 2025.
- CARVALHO, J. R. P. e VIEIRA, S. R. **Avaliação e comparação de estimadores de Krigagem para variáveis agronômicas-uma proposta**. Campinas: Embrapa Informática Agropecuária, 2001. (Documento 3). Disponível em: <https://www.infoteca.cnptia.embrapa.br/bitstream/doc/8320/1/DOCUMENTO3int.pdf>. Acesso em: 18 set. 2025.
- ESCOVEDO, T. **Machine Learning: conceitos e modelos — parte I: aprendizado supervisionado**. [S. l.], 2020. Portal: Medium. Disponível em: [https://tatianaesc.medium.com/machine-learning-conceitos-e-modelos-f0373bf4f445#:~:text=Podemos%20dividir%20o%20aprendizado%20em,\(entrada%20%E2%80%94%20sa%C3%ADda%20desejada\)](https://tatianaesc.medium.com/machine-learning-conceitos-e-modelos-f0373bf4f445#:~:text=Podemos%20dividir%20o%20aprendizado%20em,(entrada%20%E2%80%94%20sa%C3%ADda%20desejada)). Acesso em: 23 out. 2025.
- FISHER, R. A. (1992). **Métodos Estatísticos para Pesquisadores**. In: Kotz, S., Johnson, NL (eds) Breakthroughs in Statistics. Springer Series in Statistics. Springer, Nova Iorque, NY. 1992. Disponível em: https://doi.org/10.1007/978-1-4612-4380-9_6. Acesso em: 17 out. 2025.
- FEHR, W.R.; CAVINESS, C.E. 1977. **Stages of soybean development**. Ames: Iowa State University of Science and Technology. 11 p. (Special Report 80). Disponível em: <https://dr.lib.iastate.edu/handle/20.500.12876/90239>. Acesso em: 23 out. 2025.
- GLOBALFERT. **Anuário GlobalFert 2024**. Disponível em: <https://globalfert.com.br/Anuario-OGF-2024/>. Acessado em: 12 ago. 2025.

- HALDER, A. K.; MISHIRA, A. K.; BHATTACHARYYA, P.; CHAKRABARTTY, P. K. **Solubilization of rock phosphate by Rhizobium and Bradyrhizobium**. Journal Genetics Applied Microbiology, v. 36, p. 81-92, 1990. Disponível em: <https://doi.org/10.2323/jgam.36.81>. Acesso em: 10 ago. 2025.
- HATFIELD, J. L.; PRUEGER, J. H. **Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices**. Remote Sensing. v. 2, n. 2, p. 562-578, 2010. Disponível em: <https://doi.org/10.3390/rs2020562>. Acesso em: 10 ago. 2025.
- HAYKIN, S. **Redes neurais: princípios e práticas**. Porto Alegre, RS. ed. Bookman, 2001. cap 1, p 27-74.
- Instituto Nacional de Meteorologia, **Dados históricos anuais**. [s. l.], 2023. Disponível em: <https://portal.inmet.gov.br/dadoshistoricos>. Acesso em: 09 ago. 2025.
- JENSEN, J.R. **Sensoriamento Remoto do Ambiente: uma perspectiva em recursos terrestres**. 2. ed., São José dos Campos: Parenteses, 2009. Disponível em: <https://pdfcoffee.com/sensoriamento-remoto-do-ambiente-john-r-jensen-pdf-free.html>. Acesso em: 18 set. 2025.
- MARTINS, F. B. *et al.* **Classificação climática de Köppen e de Thornthwaite para Minas Gerais: Cenário atual e Projeções futuras**. Revista Brasileira de Climatologia, [s. l.], v. 1, 2018. DOI:<http://dx.doi.org/10.5380/abclima.v1i0.60896>.
- MARTINS, G. D. XAVIER, L. C. M.; OLIVEIRA, G. P.; GALLO, M. L. B.; JÚNIOR, C. A. M. A.; VIEIRA, B. S.; MARQUES, D. J.; SILVA, F. V.. **Utilizando Informações Geoespaciais para Mapear Ganho de Produtividade com o Uso de Azospirillum brasilense em Sulco**. Agronomia 2023, V. 13, N. 13. Disponível em: <https://doi.org/10.3390/agronomy13030808>. Acesso em: 12 mai. 2025.
- OLIVEIRA-PAIVA, C. A.; ALVES, V. M. C.; GOMES, E. A.; SOUSA, S. M.; LANA, U. G. P.; MARRIEL, I. E. **Microrganismos solubilizadores de fósforo e potássio na cultura da soja**. Bioinsumos na cultura da soja. 1ª ed., Embrapa, Brasília, DF, BRASIL, 2022, Cap. 9, pag. 163-179.
- OLIVEIRA, G. B.; OLIVEIRA, M. **Nanossatélites Planet – a Realidade do Monitoramento Contínuo**. Disponível em: <http://mundogeo.com/blog/2016/12/09/artigo-nanossatelites-planet-a-realidade-do-monitoramento-continuo/>. Acesso em: 26 ago. 2025.
- PAVINATO, P. S.; CHERUBIN, M. R.; ROCHA, G. C.; CHADWICK, DAVE, R.; JONES, D. L. **Revealing soil legacy phosphorus to promote sustainable agriculture in Brazil**. Scientific Reports, v. 10, article 15615, 2020. DOI: 10.1038/s41598-020-72302-1. Disponível em: <https://www.nature.com/articles/s41598-020-72302-1>. Acesso em: 12 mai. 2025.
- ROY, E.; WILLING, E.; RICHARDS, P. D.; MARTINELLI, L. A.; VAZQUEZ, F. F.; PEGORINI, L.; SPERA, S. A.; PORDER, S. **Capacidade de sorção de fósforo do solo após três décadas de fertilização intensiva em Mato Grosso, Brasil**. Agric. Ecos.

Environ. 249, 206–214 (2017). DOI: 10.1016/j.agee.2017.08.004. Disponível em: <https://repositorio.usp.br/item/002865907>. Acesso em: 12 mai. 2025.

SANTOS, M. C. L. **Modelos multiespectrais para a estimativa de produtividade da cultura da soja (Glycine max L)**. 2025, 50 f. Trabalho de Conclusão de Curso (Graduação em Engenharia de Agrimensura e Cartográfica) - Universidade Federal de Uberlândia, Monte Carmelo, 2025. Disponível em: <https://repositorio.ufu.br/handle/123456789/44891>. Acesso em: 26 ago. 2025.

SILVA, J. P. da.; ZULLO JÚNIOR, J.; ROMANI, L. A. S. **Active learning e sua aplicação no monitoramento da cana-de-açúcar utilizando o algoritmo SVM**. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 11., 2017, Campinas-SP, Anais [...]. Campinas: Editora da Unicamp, 2017. p. 119-128. Disponível em: <https://www.alice.cnptia.embrapa.br/alice/handle/doc/1083299>. Acesso em: 22 out. 2025.

SILVA, F. V.; MARTINS, G. D. **Estimativa do ganho de produtividade da soja com uso imagens multiespectrais**. Artigo apresentado na I Semana do Engenheiro Agrimensor e Cartógrafo, 20 a 24 de março de 2023. Disponível em: <https://grcmlesydpd.objectstorage.sa-saopaulo-1.oci.customer-oci.com/p/OQwcvnO-c63O08Gc2Kv4OTbJttj5ik60dguiDIyyQ0wuo5SWn-jHOLW9wNbylNqI/n/grcmlesydpd/b/dtysppobjmntbcp01/o/media/doity/submissoes/artigo-e2643f4adacd2a4f53aae7a953103298d04f7f05-arquivo.pdf>. Acesso em: 18 set. 2025.

THAI, T. H.; OMARI, R. A.; BARKUSKY, D.; KIMURA, B. **Análise estatística versus o algoritmo de aprendizado de máquina M5P para analisar a produtividade do trigo de inverno em um experimento de fertilização de longo prazo**. Agronomy 2020, 10, 1779. <https://doi.org/10.3390/agronomy10111779>. Acesso em: 12 mai. 2025.

WILLMOTT, Cort J.; MATSUURA, K. **Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance**. Climate Research, vol. 30, no. 1, 2005, pp. 79–82. JSTOR, Disponível em: <http://www.jstor.org/stable/24869236>. Acesso em: 15 set. 2025.

ZHOU, J.; YUNGBLUTH, D.; VONG, C. N.; SCABOO, A. **Estimativa da data de maturação de linhagens de melhoramento de soja usando imagens multiespectrais obtidas por VANT**. Sens. Remoto 2019, 11, 2075. Disponível em: <https://doi.org/10.3390/rs11182075>. Acesso em: 17 out. 2025.