



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS AGRÁRIAS
CURSO DE GRADUAÇÃO EM AGRONOMIA

Luciano dos Santos Cardoso Junior

**Machine Learning aplicado ao Monitoramento Agrícola: Estudo de
Caso no Cultivo de Milho em Luís Eduardo Magalhães (BA)**

Florianópolis

2025

Luciano dos Santos Cardoso Junior

**Machine Learning aplicado ao Monitoramento Agrícola: Estudo de Caso no
Cultivo de Milho em Luís Eduardo Magalhães (BA)**

Trabalho de Conclusão de Curso submetido ao curso de Graduação em Agronomia do Centro de Ciências Agrárias da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Engenheiro Agrônomo.

Orientador(a): Prof. Tiago Olivoto,

Florianópolis

2025

AGRADECIMENTOS

Expresso minha profunda gratidão à minha família, que com amor e paciência, esteve presente em toda esta trajetória acadêmica, tornando possível a realização desta graduação. Aos amigos, meus sinceros agradecimentos pela companhia e apoio ao longo destes anos. À minha namorada, meu muito obrigado por estar ao meu lado desde a segunda fase. Você foi, junto da minha família, um dos pilares essenciais para esta conquista.

“O fracasso é uma possibilidade. Se as coisas não estão falhando, você não está inovando o suficiente.” - Elon Musk

RESUMO

A expansão agrícola e a diversificação de culturas evidenciam que o sensoriamento remoto aplicado à agricultura necessita de otimizações nos processos de mapeamento das áreas cultivadas. Este trabalho teve como objetivo avaliar a capacidade preditiva de nove algoritmos de classificação supervisionada para a estimativa da área de milho cultivada no município de Luís Eduardo Magalhães (BA), durante a primeira safra do ano agrícola 2024/2025. Foram utilizadas imagens do satélite Sentinel-2 em duas composições: uma composição simples contendo as 3 bandas: infravermelho próximo (B8), infravermelho de ondas curtas (B11) e vermelho (B4); e outra composta por um stack raster formado por 9 bandas: infravermelho próximo (B8), infravermelho de ondas curtas (B11), vermelho (B4), aerossol (B1), azul (B2), verde (B3) e Red Edge (B5, B6 e B7). Os algoritmos foram aplicados a imagens correspondentes a diferentes períodos fenológicos da cultura: fase de desenvolvimento pleno e final do ciclo (senescência), permitindo verificar a capacidade dos modelos em distinguir a cultura do milho a partir de uma única imagem, cenário comum em regiões com limitações de cobertura temporal. Três métricas foram utilizadas para avaliação do desempenho: o índice Kappa, a variação das áreas estimadas em relação ao mapeamento manual e uma avaliação qualitativa e participativa realizada por pesquisadores e profissionais da área. Os modelos testados apresentaram diferentes valores de Kappa, níveis de preenchimento, ruído e ocorrência de falsos positivos, cada um com suas particularidades. No entanto, o modelo que apresentou o melhor desempenho em relação às métricas propostas foi o Gaussian Mixture Model (GMM), alcançando um índice Kappa de 0,859 na composição simples e 0,928 no stack raster. Além disso, apresentou baixa variação em relação à área do mapeamento manual e obteve as melhores avaliações na análise participativa, indicando que, na percepção dos intérpretes, o GMM pode ser utilizado como um assistente de pré-mapeamento (nas etapas de interpretação e vetorização inicial), contribuindo para a otimização do tempo de trabalho dos analistas humanos. Este estudo trouxe insights relevantes para o setor agrícola, evidenciando que a combinação de diferentes modelos de machine learning constitui uma abordagem promissora, uma vez que alguns algoritmos se destacam na delimitação e preenchimento das áreas, enquanto outros são mais eficientes na interpretação e refinamento dos resultados.

Palavras-chave: Classificação Supervisionada, Mapeamento de Culturas e Sensoriamento Remoto

ABSTRACT

The agricultural expansion and diversification of crops highlight the need for optimized approaches in remote sensing applied to agriculture, particularly in mapping cultivated areas. This study aimed to evaluate the predictive capacity of nine supervised classification algorithms for estimating maize crop area in the municipality of Luís Eduardo Magalhães (BA) during the first harvest of the 2024/2025 agricultural year. Sentinel-2 satellite images were used in two compositions: a simple composition containing 3 bands: near-infrared (B8), shortwave infrared (B11), and red (B4); and a second composition consisting of a raster stack with 9 bands : near-infrared (B8), shortwave infrared (B11), red (B4), aerosol (B1), blue (B2), green (B3), and Red Edge (B5, B6, and B7). The algorithms were applied to images corresponding to different phenological stages of the crop, full development and the end of the cycle (senescence), enabling the assessment of each model's ability to distinguish maize using a single image, a common scenario in regions with limited temporal coverage. Three metrics were used to evaluate performance: the Kappa index, the variation of estimated areas compared to manual mapping, and a qualitative and participatory assessment conducted by researchers and professionals in the field. The tested models presented different Kappa values, levels of area completeness, noise, and false positives, each with its own characteristics. However, the model that showed the best performance according to the proposed metrics was the Gaussian Mixture Model (GMM), achieving a Kappa index of 0.859 in the simple composition and 0.928 in the raster stack. Additionally, it exhibited low variation relative to the manually mapped area and received the highest scores in the participatory evaluation, indicating that, from the interpreters' perspective, GMM can be used as a pre-mapping assistant (in interpretation and initial vectorization stages), helping optimize analysts' working time. This study provides relevant insights for the agricultural sector, demonstrating that combining different machine learning models is a promising approach. While some algorithms excel in delineating and filling crop areas, others are more effective in interpretation and refinement of results, reinforcing the value of hybrid and complementary modeling strategies.

Keywords: Supervised Classification, Crop Mapping, and Remote Sensing

LISTA DE FIGURAS

Figura 1 - Localização do município de Luís Eduardo Magalhães.	14
Figura 2 - Código e interface para baixar imagens do GEE.	16
Figura 3 - Trabalho realizado em outro estado onde estudou a variação do NDVI em imagens dos sensores MODIS e OLI em função do estágio fenológico do milho. Significância a 0,1% de probabilidade pelo teste F.	17
Figura 4 - Resposta do milho em fase inicial do período fenológico tom de marrom e em maturação fisiológica tom de Verde Bandeira. Essa composição se refere a RGB formada pelas bandas 8,11,4 do sentinel 2 com as datas respectivamente de 07/02/2025 e 03/04/2025.	17
Figura 5 - Gráfico Temporal do EVI, destacando em vermelho a característica do milho.	18
Figura 6 - Exemplo das amostras de Milho com contorno Vermelho e as amostras do tipo "0" com contorno e hachuras em cor preta.	21
Figura 7 - Mapa referência do Google Forms	23
Figura 8 - Área de milho cultivada em Luís Eduardo Magalhães (Polígonos em Vermelho) no ano da safra 24/25, mapeada de forma manual. Erro! Indicador não definido.	4
Figura 9. Gráfico de Comparação das notas e índice kappa.	29
Figura 10. Distribuição das notas entre modelos e metodologias.	29

LISTA DE TABELAS

Tabela 1 - Bandas do Sentinel 2.	15
Tabela 2 - Imagens de satélite do Sentinel 2 para aplicação dos modelos de machine learning e mapeamento manual.	16
Tabela 3 - Chave de identificação das classificações do tipo "0" para os algoritmos de machine learning, considerando as bandas B08, B11 e B04 do satélite Sentinel-2.	19
Tabela 4 - Tabela de referência de desempenho do índice kappa.	22
Tabela 5 - Resultados dos Modelos de Machine Learning para composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4).	25
Tabela 6 - Resultados dos Modelos de Machine Learning para o stack infravermelho próximo (8) , infravermelho de ondas curtas (11), vermelho (4), Aerossol (1), Azul (2), Verde (3), Red Edge (5,6,7).	27

LISTA DE ABREVIATURAS E SIGLAS

ANP – Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
CONAB – Companhia Nacional de Abastecimento
DDG – Distillers Dried Grains
DDGS – Distillers Dried Grains with Solubles
ET – Extra Trees
EVI – Enhanced Vegetation Index
GBC – Gradient Boosting Classifier
GMM – Gaussian Mixture Model
IBGE – Instituto Brasileiro de Geografia e Estatística
INPE – Instituto Nacional de Pesquisas Espaciais
KNN – K-Nearest Neighbors
LACIE – Large Area Crop Inventory Experiment
LGB – Light Gradient Boosting Machine (LightGBM)
LR – Logistic Regression
MLP – Multi-Layer Perceptron
NASA – National Aeronautics and Space Administration
NB – Naive Bayes
NDVI – Normalized Difference Vegetation Index
NOAA – National Oceanic and Atmospheric Administration
OP – Órbita/Ponto
REM – Radiação Eletromagnética
RF – Random Forest
SVM – Support Vector Machine
UNEM – União Nacional do Etanol de Milho
USDA – United States Department of Agriculture
WDG – Wet Distillers Grains
XGB – Extreme Gradient Boosting (XGBoost)

SUMÁRIO

1 INTRODUÇÃO	11
2 DESENVOLVIMENTO	14
2.1 Local e Imagens	14
2.2 Software e mapeamento manual	17
2.3 Classificação Supervisionada	18
3 RESULTADOS E DISCUSSÕES	24
3.1 Mapeamento manual	24
3.2 Mapeamento automatizado composição 8-11-4	24
3.3 Mapeamento automatizado composição Stack	26
3.4 Metodologia Qualitativa e Participativa	29
4 CONCLUSÕES	31

1 INTRODUÇÃO

O sensoriamento remoto se define pela utilização de uma gama de sensores e equipamentos de transmissão de dados/processamento que estão alocados em aeronaves, espaçonaves, satélites ou outras plataformas. O principal objetivo do sensoriamento remoto é estudar as interações entre a radiação eletromagnética (REM) e os componentes do objeto (Novo, 2015). Toda matéria existente absorve, reflete, transmite ou emite a radiação eletromagnética (REM) os sensores remotos geram produtos com as características únicas do comportamento de reflectância da área de interesse ou alvo em diferentes regiões do espectro eletromagnético (Formaggio; Sanches, 2017). O sensoriamento agrícola surgiu da aplicação da divisão e exploração do espectro eletromagnético obtido no sensoriamento remoto. Neste sentido, o sensoriamento remoto agrícola pode ser utilizado para quantificar a área plantada e a produtividade de culturas específicas em níveis locais, estaduais e nacionais (Formaggio; Sanches, 2017). O primeiro país que utilizou satélites para o sensoriamento remoto agrícola foi os Estados Unidos com o projeto LACIE (Large Area Crop Inventory Experiment) onde envolveu agência espacial norte-americana (NASA), National Oceanic and Atmospheric Administration (NOAA) e o Departamento de Agricultura Norte-Americana (USDA) com o objetivo de quantificar a área de trigo globalmente (Formaggio; Sanches, 2017). No Brasil, o sensoriamento remoto agrícola iniciou no começo da década de 1970 pelo Instituto Nacional de Pesquisas Espaciais (INPE) e desde então desempenha um importante papel em escala nacional e consolidando o Brasil como um dos mais desenvolvidos nesse campo no hemisfério sul (Formaggio; Sanches, 2017).

A América do Sul possui uma área total de 17.840.000 km², e o Brasil se destaca com 8.509.379,576 km², o que representa aproximadamente 47,7% de toda a extensão territorial do continente. Conforme o levantamento da CONAB de setembro de 2025, os dez estados brasileiros com maior área plantada são: Mato Grosso, Rio Grande do Sul, Paraná, Goiás, Mato Grosso do Sul, Minas Gerais, Bahia, São Paulo, Tocantins e Maranhão.

A Bahia possui uma área territorial total de 564.760,429 km² (6,6% da área do Brasil), bem como Luís Eduardo Magalhães possui uma área territorial 4.033,040 km² (0,7% da área da Bahia) (IBGE, 2025). Para a economia brasileira as culturas de verão consistem por 90% da produção anual de grãos (Formaggio; Sanches, 2017). O Brasil está acostumado com a primeira e a segunda safra, porém alguns estados já

garantem a terceira safra no sequeiro. O manejo da terceira safra consiste em alguns fatores limitantes que são: plantio direto, melhoramento genético e irrigação (em pivôs é comum observar três safras ao ano por conta do manejo hídrico ser mais viável) (Agroreceita, 2023).

A Bahia apresenta grande diversidade agrícola, o que pode ser observado nos números disponibilizados pela CONAB em setembro de 2025. O algodão registrou um acréscimo de 19,4% em relação à safra de 2023/24, alcançando 413,1 mil ha. O amendoim segunda safra manteve-se estável, com 2,6 mil ha de área plantada. As culturas de feijão (cores e caupi), distribuídas em três possíveis safras, registraram uma redução de 11,1%, totalizando 377,3 mil ha. A mamona apresentou aumento de 19,8%, atingindo 69,5 mil ha. O milho, também distribuído em três safras, apresentou decréscimo de 3,2%, reduzindo-se para 741,7 mil ha. Já a soja teve crescimento de 7,9%, somando 2.135,6 mil ha, enquanto o sorgo registrou aumento de 5,1%, chegando a 206 mil ha. Por outro lado, o trigo apresentou queda expressiva de 37,5%, reduzindo sua área para 5 mil ha. No total, a Bahia apresentou um crescimento geral de 4,5% na área plantada em relação à safra anterior, alcançando 3.950,8 mil ha.

De acordo com a União Nacional do Etanol de Milho (UNEM), em menos de uma década, a indústria de etanol de milho transformou o cenário dos biocombustíveis no Brasil e impulsionou o setor de proteína animal, ao mesmo tempo em que valorizou a produção de milho. Hoje, o país ocupa a terceira posição no ranking mundial de produtores, atrás apenas dos Estados Unidos e da China. Nos últimos dez anos, a produção nacional cresceu cerca de 40%. A cadeia do etanol de milho surgiu justamente para dar destino estratégico a esse crescimento, aproveitando excedentes exportáveis como matéria-prima para gerar biocombustíveis e co-produtos de alto valor agregado, como DDG/DDGS e óleo de milho.

Atualmente, o Brasil conta com 24 biorrefinarias em operação, 16 projetos já autorizados pela ANP e outras 16 unidades em fase de planejamento. O setor é reconhecido internacionalmente pela sustentabilidade de seu modelo produtivo. Isso porque utiliza biomassa, uma fonte renovável e de baixo impacto ambiental para geração de energia elétrica e vapor, garantindo autossuficiência energética e ainda fornecendo excedentes para a rede de distribuição. Além do etanol, as biorrefinarias disponibilizam produtos estratégicos como DDG/DDGS/WDG, grãos de destilaria ricos em proteína e energia, fundamentais para a alimentação de bovinos, suínos, aves, peixes, equinos e até pets, óleo de milho, insumo para nutrição animal e também para

indústrias químicas e de alimentos, bioeletricidade, gerada a partir da biomassa, com excedente fornecido ao sistema elétrico e o etanol de milho, tradicionalmente utilizado na mobilidade urbana em veículos leves, ganha cada vez mais espaço em novas soluções energéticas, como ônibus, embarcações e até aviação, contribuindo diretamente para a redução das emissões de gases de efeito estufa.

Entretanto, esse crescimento contínuo de área plantada e a diversificação de culturas (incluindo até três safras anuais) impõem um desafio logístico e operacional ao mapeamento agrícola tradicional. O método manual, que exige a verificação talhão por talhão, torna-se insustentável pela demanda de mão de obra. Embora métodos como a Classificação Supervisionada surjam como assistentes, eles enfrentam limitações, especialmente a perda de precisão causada pela interferência de nuvens (Cristiane et al., 2019). Além disso, o mapeamento exato é dependente da cultura e da região, como no caso do sorgo e milheto na Bahia, que exigem uma separação pós-classificação. Dada a limitação da mão de obra e os desafios de precisão impostos pelos métodos supervisionados no contexto de múltiplas safras e alta diversidade do agronegócio baiano, torna-se imprescindível a pesquisa por novas metodologias. Com isso, o presente trabalho propõe a avaliação de modelos de Machine Learning e Deep Learning voltados ao mapeamento da área cultivada de milho primeira safra no município de Luís Eduardo Magalhães - BA, buscando oferecer uma solução mais rápida, precisa e escalável para a determinação de estatísticas agrícolas (área de culturas agrícolas, diversidade regional, produtividade de culturas agrícolas etc.).

2 DESENVOLVIMENTO

2.1 LOCAL E IMAGENS

O estudo foi conduzido no município de Luís Eduardo Magalhães, localizado na região oeste da Bahia, durante a primeira safra do ano agrícola 2024/2025. Para a obtenção dos dados de sensoriamento remoto, utilizou-se o satélite Sentinel-2. Visto que a área do município se encontra sobreposta a duas Órbitas Pontos (OP), a OP com maior porcentagem de área plantada de interesse, 23LLG (localizada à esquerda), foi selecionada para o mapeamento.

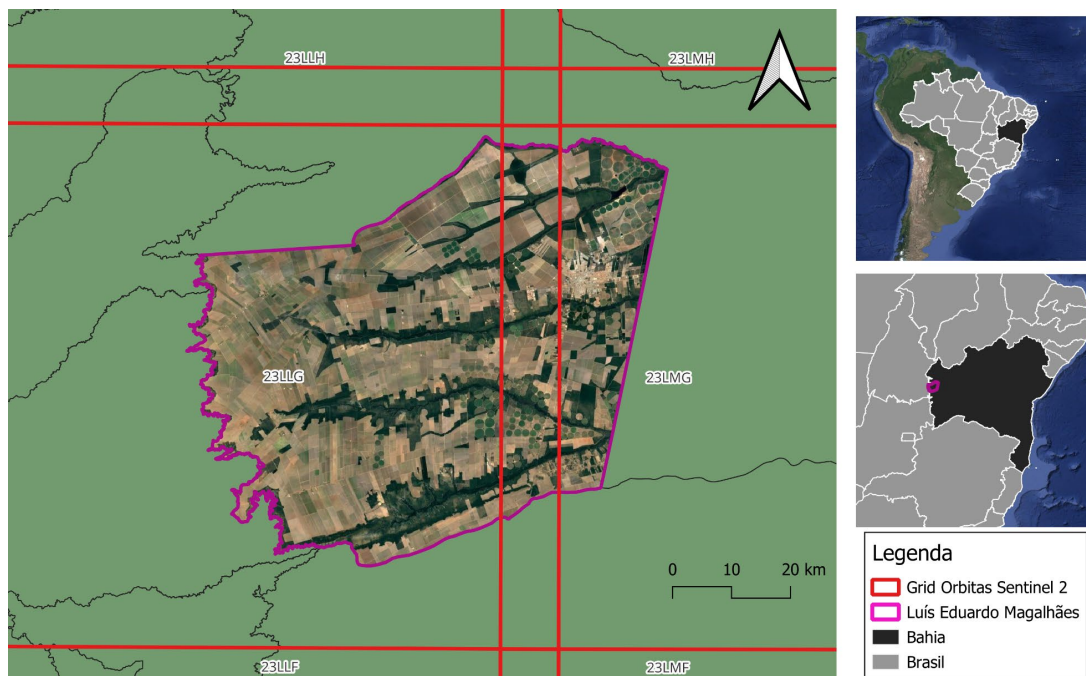


Figura 1. Localização do município de Luís Eduardo Magalhães.

A Bahia inicia seu calendário de plantio de milho da primeira safra entre outubro e janeiro e, a partir dessas datas, a colheita pode ocorrer de fevereiro a agosto. Dessa forma, esse intervalo de plantio foi adotado para a seleção das imagens de satélite. Neste trabalho, optou-se pelo Sentinel-2 devido ao seu intervalo de revisita ao mesmo ponto, que é de aproximadamente 3 dias.

Para a aquisição de imagens foi utilizado o Google Earth Engine (GEE), onde foi criado um código para buscar as melhores imagens nas duas quinzenas de cada mês, com o objetivo de obter duas composições com a menor cobertura de nuvens. De acordo com Formaggio e Sanches (2017) a composição onde as culturas agrícolas

se destacam são as bandas infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4), visando isto foi utilizado B08, B11 e B04 e um outro código para baixar as mesmas imagens sendo um raster stack com as bandas do satélite (8-11-4-1-2-3-5-6-7) (Tabela 1).

Tabela 1. Bandas do sentinel 2.

Resolução	Número da banda	Nome da banda	Comprimento de onda central (nanômetro)
10 m	B02	Azul	490
	B03	Verde	560
	B04	Vermelho	665
	B08	Infravermelho próximo	842
20 m	B05	Red edge 1	705
	B06	Red edge 2	740
	B07	Red edge 3	783
	B08A	Red edge 4	865
	B11	SWIR 1	1610
	B12	SWIR 2	2190
60 m	B01	Aerossol	443
	B09	Vapor d' Água	940
	B10	Cirrus	1375

O algoritmo já realiza o processamento deixando somente o recorte do município de interesse como demonstrado na Figura 2.

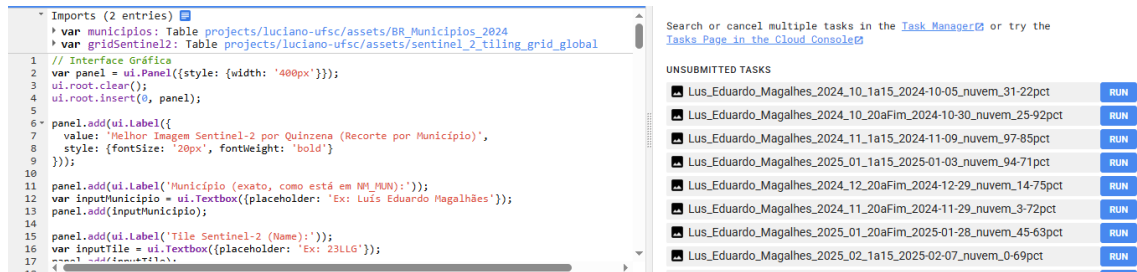


Figura 2. Código e interface para baixar imagens do GEE.

As imagens escolhidas para o presente trabalho foram uma de início de ciclo do dia 07/02/2025 onde a cor característica do milho é marrom/vermelho e outra imagem do fim do ciclo do dia 03/04/2025 onde o milho fica verde (Figura 4). Essa estratégia de segmentação de datas permite verificar se o modelo apresenta maiores diferenças no período que antecede a senescência ou após esse estágio. Por exemplo, caso a região em que está sendo realizado o mapeamento disponha de apenas uma imagem com a menor cobertura de nuvens possível para a execução dos modelos automáticos, é possível avaliar se esses modelos tendem a apresentar melhor desempenho no início do ciclo ou em seu final

Tabela 2. Imagens de satélite do Sentinel 2 para aplicação dos modelos de machine

Órbita Ponto	Bandas espectrais	Data das imagens	Cobertura nuvens (%)
23LLG	Composição 8-11-4	07/02/2025	1,23
		03/04/2025	22,54
	Stack 8-11-4 + 1-2-3-5-6-7	07/02/2025	1,23
		03/04/2025	22,54

learning e mapeamento manual.

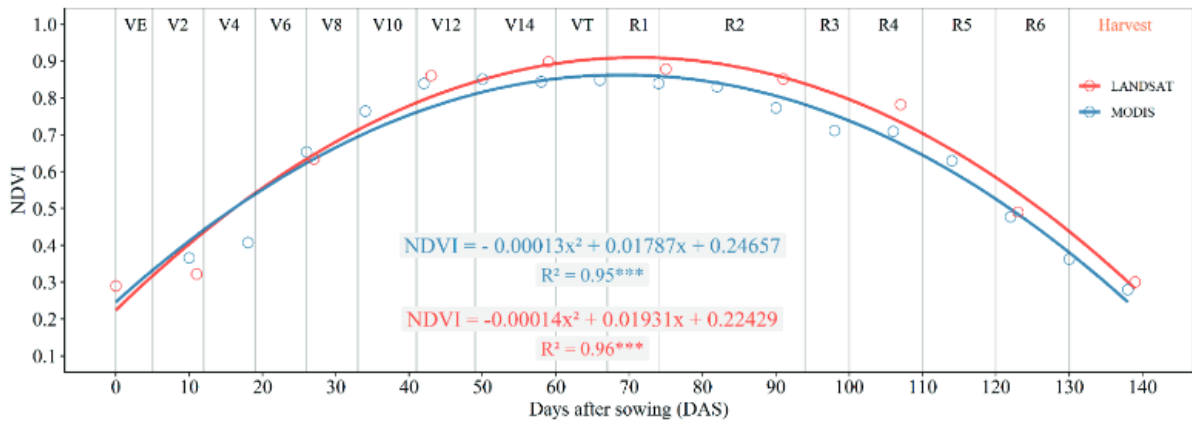


Figura 3. Trabalho realizado em outro estado onde estudou a variação do NDVI em imagens dos sensores MODIS e OLI em função do estágio fenológico do milho. Significância a 0,1% de probabilidade pelo teste F.

Fonte: VENANCIO et al., 2020.

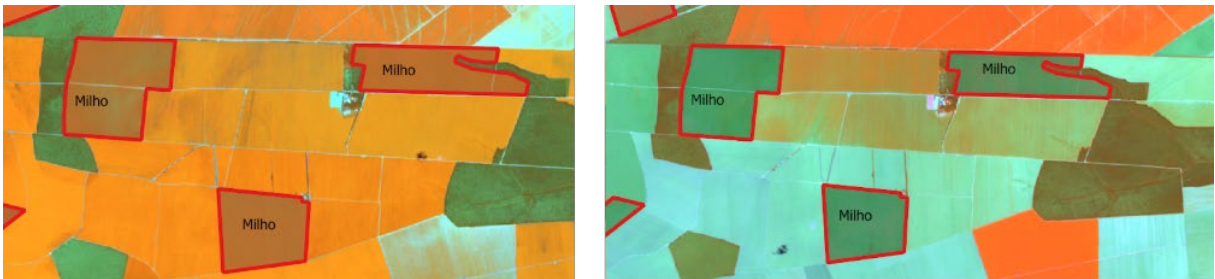


Figura 4. Resposta do milho em fase inicial do período fenológico tom de marrom e em maturação fisiológica tom de verde bandeira. Essa composição refere-se a RGB formada pelas bandas 8,11,4 do Sentinel 2 com as datas respectivamente de 07/02/2025 e 03/04/2025.

2.2 SOFTWARE E MAPEAMENTO MANUAL

Para o mapeamento e processamento dos dados, utilizou-se o sistema de informação geográfica livre e aberto Qgis versão 3.40 (QGIS Development Team, 2025). Manteve-se o projeto no sistema de coordenadas Sirgas 2000 (EPSG 4674), necessitando reprojeter as imagens e arquivos.

O mapeamento manual consistiu na vetorização de polígonos onde se encontram os respectivos talhões de milho, considerando a composição 8-11-4 (Tabela 2). A confirmação dos talhões de milho foi por identificação visual pelo padrão de cor do milho como demonstra a figura 4. utilizou-se um intervalo de imagens do mês de outubro de 2024 até agosto de 2025.

Outro recurso aplicado como auxílio ao mapeamento foi a plataforma online criada pela Embrapa o SatVeg (Embrapa, 2025), onde se pode verificar dados históricos do satélite MODIS obtendo os índices: Vegetação por Diferença Normalizada (NDVI) e Índice de Vegetação Melhorado (EVI) dos talhões mostrados na Figura 4. Assim, foi possível o comportamento espectral já conhecido do milho com o índice EVI em seu ciclo completo (plantio e colheita), bem como consegue-se verificar o tempo da cultura no solo que pode variar para milho de 105 dias até 150.

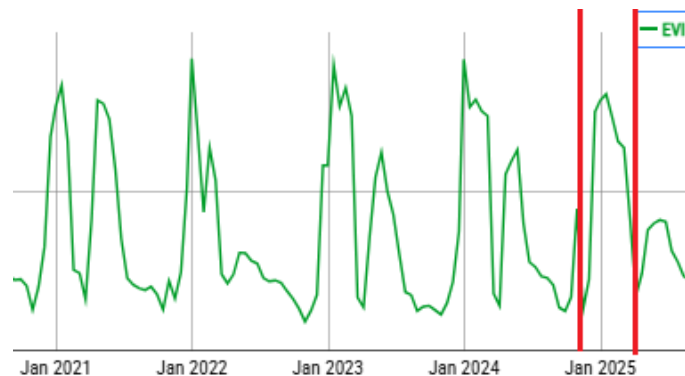


Figura 5. Gráfico Temporal do EVI, destacando em vermelho a característica do milho utilizando o SatVeg.

Fonte: EMBRAPA,2025.


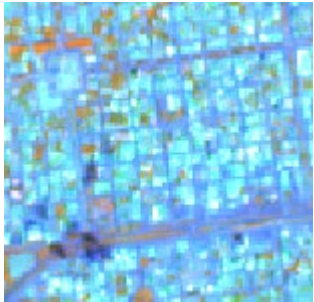
2.3 CLASSIFICAÇÃO SUPERVISIONADA

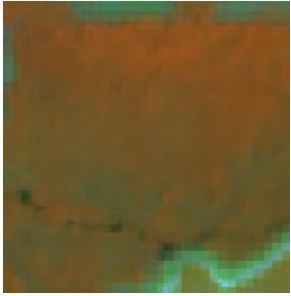

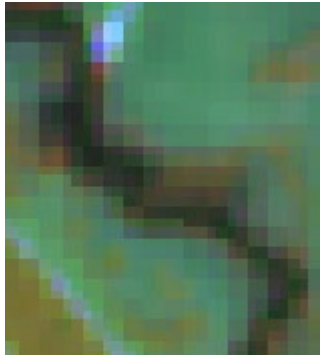
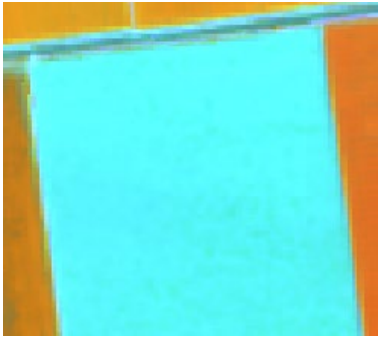
Para realizar a classificação supervisionada de cada imagem do trabalho, utilizou-se o complemento para QGIS denominado Dzetsaka classification tool (Karasiak, 2016). Este Plugin contém algoritmos de classificação baseados em Machine learning que são: Gaussian Mixture Model (GMM), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), XGBoost (XGB), LightGBM (LGB), Extra Trees (ET), Gradient Boosting Classifier (GBC), Logistic Regression (LR), Naive Bayes (NB) e de Deep Learning que é o Multi-layer Perceptron (MLP). Neste trabalho somente não foi utilizado o Support Vector Machine (SVM) e Gradient Boosting Classifier (GBC) por conta de necessitar de um computador mais robusto.

De acordo com o trabalho de Pittman & Hu (2022), para mapear solos os algoritmos de machine learning atingem seu melhor potencial com uma classificação binarizada, logo no presente trabalho tudo que é o milho vira “1” e todo o resto virá

“0”. As classes de uso do tipo “0” são outras culturas agrícolas (soja, algodão e café), áreas urbanas, mata, pastagem, corpos hídricos e solo exposto.

Tabela 3. Chave de identificação das classificações do tipo “0” para os algoritmos de machine learning, considerando as bandas B08, B11 e B04 do satélite Sentinel-2.

Classe	Elementos de interpretação	Recorte da imagem
Soja	Apresenta tom amarelo, textura uniforme e ciclo completo.	
Algodão	Apresenta coloração laranja a vermelha na composição em falsa cor, tonalidade branca em RGB durante a abertura do capulho e ciclo completo.	
Café	Apresenta coloração marrom brilhante, com aspecto de vinil antigo, ausência de solo exposto e ciclo bem definido segundo o SatVeg.	
Áreas Urbanas	Pode-se utilizar o Google Earth, observam-se na composição as ruas totalmente expostas e edificações com forte brilho e alta densidade.	

Mata	Apresenta tom marrom que evolui para avermelhado, com rugosidade perceptível, diferenças visíveis entre extratos vegetais e ciclo fenológico condicionado à sazonalidade da região.	
Pastagem	Apresenta variação cromática típica de vegetação plantada ou nativa, textura rugosa, transições abruptas de cor e ciclo fenológico condicionado à sazonalidade regional, subsidiado pelas métricas do SatVeg.	
Corpos hídricos	Tons azulados ou escuros (dependendo da origem), ocorrendo na mesma área e analisados com apoio do Google Earth.	
Solo exposto	Apresenta tonalidade azul brilhante ou azul-escuro que se mantém estável ao longo de toda a safra, caracterizando área em pousio.	

A quantidade total das amostras de milho para a classificação supervisionada do município foi de 20 amostras com área de 4.127 hectares (Figura 4) e as amostras do tipo “0” onde contém todas as classes não correspondentes ao milho foi de 30 amostras com a área de 48.962 ha (Tabela 3). Ao todo, resultando em 50 amostras com área de 53.089 ha utilizadas no treinamento dos modelos (Figura 6).

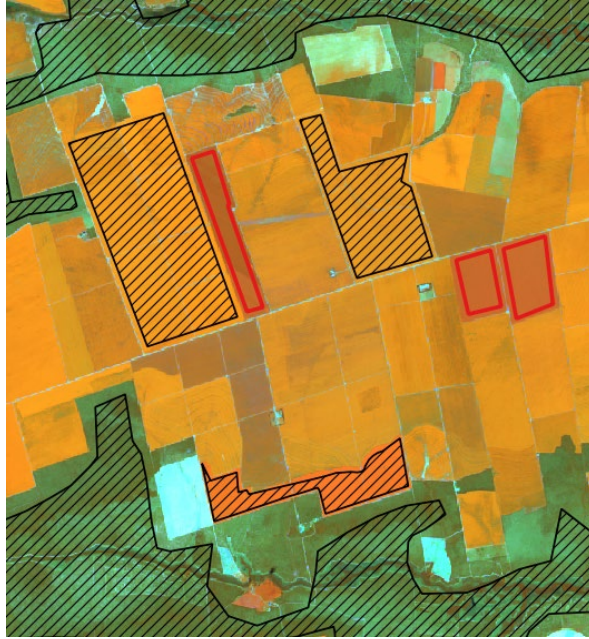


Figura 6. Exemplo das amostras de milho com contorno vermelho e as amostras do tipo “0” com contorno e hachuras em cor preta.

Os modelos de machine learning para a área de mapeamento podem ser de grande valia sendo um assistente de pré-vetorização e interpretação, visto que ocorrem falsos positivos e os ruídos sendo assim necessário revisões de intérpretes. Este trabalho busca fomentar esta ideia de utilizar esses modelos como assistentes, e verifica quais modelos de machine learning diferem com maior precisão o milho de outras culturas, assim observando qual dos modelos acarretaria em menor tempo de retrabalho para os revisores. Sendo assim, cada modelo foi executado para cada imagem, de modo que todos permanecessem no mesmo patamar de comparação, além de não ter sido aplicado nenhum processamento adicional. Foram utilizadas três métricas para avaliar a acurácia e classificação dos modelos. O primeiro foi o índice Kappa que avalia a qualidade de uma classificação, ajustando a acurácia observada pelo efeito do acaso. O Kappa foi calculado considerando a matriz de confusão construída comparando a classificação predita (o valor de classe do raster) com a classificação manual para um conjunto de pixels amostrados.

	Algoritmo	Manual	Total
Algoritmo	a	b	T_A
Manual	c	d	T_b
Total	R_A	R_B	N

Considerando a matriz de confusão, o índice Kappa (k) é dado por $k = \frac{P_o - P_e}{1 - P_e}$ onde $P_o = \frac{a+d}{N}$ é a proporção de concordância observada, ou seja a soma das concordâncias considerando os pixels corretamente classificados pelo algoritmo e manual dividido pelo total de pixels; e $P_e = \frac{(T_A \times R_A) + (T_B \times R_B)}{N^2}$ é a proporção de concordância que seria esperada se as classificações fossem puramente aleatórias ou por mero acaso, dada pela soma dos produtos das probabilidades marginais para cada categoria, dividida pelo quadrado do total de observações. Geralmente, valores de Kappa acima de 0,6 são considerados satisfatórios em estudos de sensoriamento remoto (Tabela 4).

Tabela 4. Tabela de referência de desempenho do índice kappa.

Índice Kappa (k)	Desempenho
< 0	Péssimo
0 a 0,2	Ruim
0,2 a 0,4	Razoável
0,4 a 0,6	Bom
0,6 a 0,8	Muito bom
0,8 a 1	Excelente

Fonte: Landis & Koch (1977, p.165).

A segunda métrica foi a quantidade de hectares gerados pelos modelos e suas variações em relação ao manual. Por fim foi também utilizado uma metodologia de avaliação qualitativa e participativa entre estudantes, pesquisadores da Universidade Federal de Santa Catarina e profissionais da área de Geoprocessamento de empresas conhecidas no mercado que utilizam sensoriamento remoto voltado à agricultura como

a Serasa Experian e Agroconsult. O objetivo desta etapa foi avaliar a acurácia dos modelos em termos de capacidade de preenchimento dos talhões (ou seja, o quão bom foi a identificação das áreas de milho dentro de cada talhão), bem como se o modelo resulta ao produto final poucos ruídos (pixels aleatórios mal classificados). Os participantes receberam o mapa manual (figura 7) como referência de nota 10 (Preenchimento em preto e talhões de milho contornados de vermelho) juntamente com os modelos sem identificação para não enviesar o trabalho. Segue no anexo I no presente trabalho os mapas produzidos pelos modelos de machine learning.

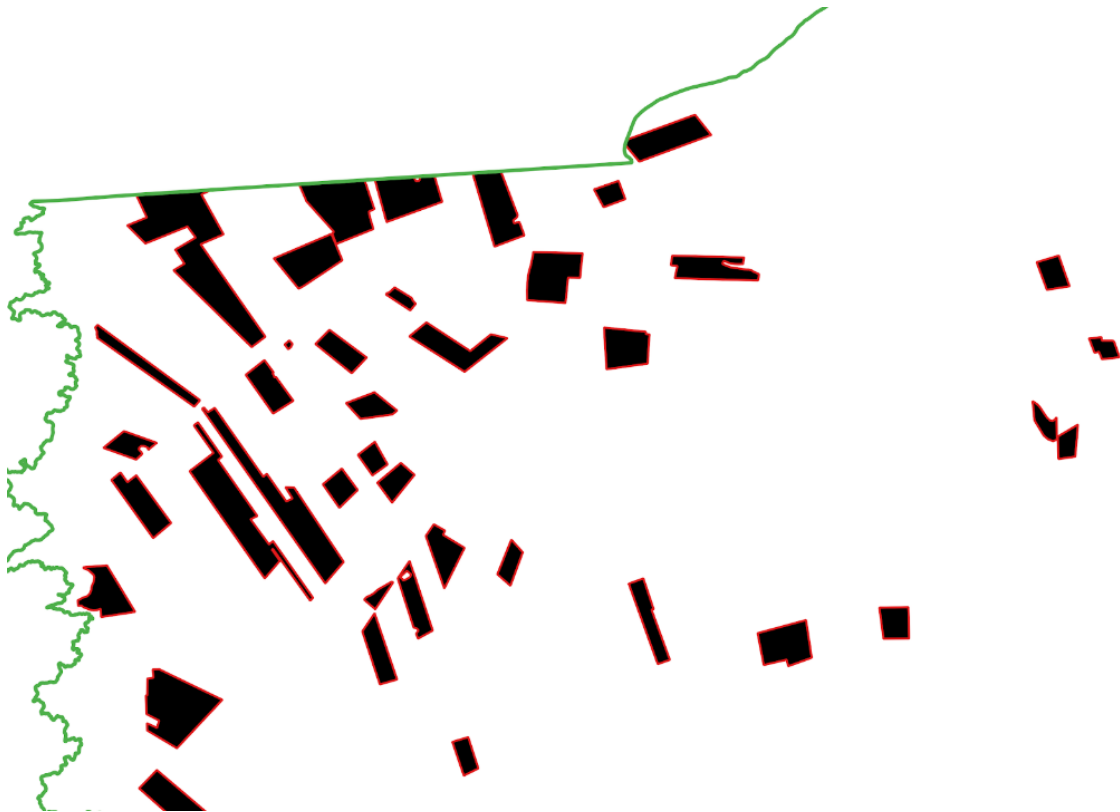


Figura 7. Mapa referência do Google Forms.

3 RESULTADOS E DISCUSSÕES

3.1 MAPEAMENTO MANUAL

Os dados obtidos pelo mapeamento manual indicaram que a área total de milho de primeira safra no município de Luís Eduardo Magalhães (na Órbita Ponto 23LLG), corresponde a 14.562,75 hectares (Figura 8).



Figura 8. Área de milho cultivada em Luís Eduardo Magalhães (Polígonos em Vermelho) no ano da safra 24/25, mapeada de forma manual.

3.2 MAPEAMENTO AUTOMATIZADO COMPOSIÇÃO 8-11-4

Os Mapeamentos automatizados com os modelos de machine learning considerando a composição das bandas infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4) obtiveram resultados satisfatórios (Tabela 5) apresentando, em sua maioria um mapeamento de desempenho excelente para a classificação utilizando imagem do dia 07/02/2025 ($0,699 \leq k \leq 0,859$). Estes resultados sugerem que a classificação mais acurada das áreas de milho é obtida antes à senescência do milho

Tabela 5. Resultados dos Modelos de Machine Learning para composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4).

Composição 8-11-4	Modelos de Machine Learning	Kappa	Área (ha)	Variação (%)**
Mapeamento Manual	-	-	14.562,75	-
07/02/2025	Gaussian Mixture Model	0,859	15.486,78	+6,34
	Naive Bayes	0,843	15.621,47	+7,27
	XGBoost	0,821	18.125,49	+24,26
	LightGBM	0,814	18.592,73	+27,67
	Extra Trees	0,807	18.923,41	+29,94
	Random Forest	0,806	18.987,60	+30,38
	Multi-Layer Perceptron	0,800	18.902,95	+29,8
	K-nearest neighbors	0,771	20.538,12	+41,03
	Logistic Regression	0,699 *(0,696)	290.613,17 *(15.279,05)	+4,91
03/04/2025	Extra Trees	0,662	16.236,26	+11,49
	Random Forest	0,653	16.559,16	+13,7
	Multi-Layer Perceptron	0,64	16.414,27	+12,71
	XGBoost	0,634	12.200,19	-16,22
	LightGBM	0,615	16.821,25	+15,5
	K-nearest neighbors	0,606	18.238,53	+25,24
	Logistic Regression	0	-	-
	Gaussian	0	-	-

	Mixture Model			
--	---------------	--	--	--

* Por conta do logistic Regression tratar as áreas de fora do município como “1” foi necessário recortar o município para retirar esse “erro” do modelo.

** Variação em relação ao mapeamento manual.

Obs: **vermelho** - variação maior que +10 // **Azul** - variação maior que -10

Considerando as métricas de Kappa e a variação dos hectares da metodologia manual o melhor modelo para uma composição básica de B08-B11-B04 foi o Gaussian mixture model com seu kappa de 0,859 e uma área de 15.486,78 (+6,34% acima do mapeamento manual), portanto este modelo pode ser considerado um bom modelo assistente pois necessitará de menor tempo de retrabalho em cima do arquivo vetorizado. Por outro lado, o pior modelo foi do logistic regression, visto que seu índice de desempenho kappa foi de 0,696 neste caso não podendo assim ser considerado um bom modelo assistente por conta dos falsos positivos e dos erros acentuados.

No entanto os modelos rodados para o final de ciclo não obtiveram desempenho satisfatório como assistentes visto que não passam dos 0,7, sendo assim necessário muito retrabalho em cima dos vetores.

3.3 MAPEAMENTO AUTOMATIZADO COMPOSIÇÃO STACK

Os mapeamentos automatizados com os modelos de Machine Learning considerando o raster stack das bandas infravermelho próximo (8), infravermelho de ondas curtas (11), vermelho (4), aerossol (1), azul (2), verde (3), red edge (5,6,7) (Tabela 6) obtiveram resultados melhores que os realizados, de acordo com a métrica Kappa apresentaram em sua maioria um mapeamento de desempenho excelente ($0,746 \leq k \leq 0,928$). Porém neste caso verifica-se valores próximos de 0,93 assim é possível comparar o desempenho dos modelos com a adição dessas bandas evidenciando que alguns apresentaram melhorias enquanto outros tiveram redução de desempenho. Nota-se que agora os modelos conseguem trabalhar com as fenologias de início e fim de ciclo.

Tabela 6. Resultados dos Modelos de Machine Learning para o stack infravermelho próximo (8) , infravermelho de ondas curtas (11), vermelho (4), Aerossol (1), Azul (2), Verde (3), Red Edge (5,6,7).

Stack 8-11-4 +1-2-3-5-6-7	Modelos de Machine Learning	Kappa	Área (ha)	Variação (%)**
Mapeamento Manual	-	-	14.562,75	-
07/02/2025	Gaussian Mixture Model	0,928	14.315,18	-0,017
	Multi-Layer Perceptron	0,917	16.105,97	+10,59
	Extra Trees	0,905	16.169,22	+11,03
	Logistic Regression	0,908 *(0,903)	290.454,06 *(14.941,50)	+2,6
	Random Forest	0,883	16.661,95	+14,41
	LightGBM	0,86	17.524,99	+20,34
	XGBoost	0,852	17.333,90	+19,02
	K-nearest neighbors	0,833	18.904,70	+29,81
	Naive Bayes	0,746	20.139,86	+38,29
03/04/2025	Random Forest	0,818	14.583,57	+0,1
	Extra Trees	0,814	14.896,94	+2,29
	Multi-Layer Perceptron	0,806	16.809,92	+15,43
	Gaussian Mixture Model	0,782	15.693,62	+7,76
	XGBoost	0,782	12.486,14	-14,26
	LightGBM	0,766	17.613,70	+20,95
	K-nearest neighbors	0,734	17.530,10	+20,37

	Logistic Regression	0,461 *(0,460)	300.728,76 *(26.685,76)	-83,24
--	---------------------	-------------------	----------------------------	--------

* Por conta do logistic Regression tratar as áreas do raster como "1" foi necessário recortar o município para retirar esse "erro" do modelo.

** Variação em relação ao mapeamento manual.

Obs: **vermelho** - variação maior que +10 // **Azul** - variação maior que -10

Considerando as métricas de Kappa e a variação dos hectares da metodologia manual o melhor modelo para uma composição stack no início de ciclo foi novamente o Gaussian mixture model com seu Kappa de 0,928 e uma área de 14.315,18 (+0,017% acima do mapeamento manual), portanto este modelo para a condição início de ciclo pode ser considerado um bom modelo assistente. Por outro lado, o pior modelo foi do Naive Bayes, visto que seu índice de desempenho Kappa anterior foi de 0,843 e decaiu para 0,746 sugerindo que o conjunto de dados mais complexo prejudicou a capacidade de classificação do Naive Bayes. Tal redução pode ser explicada devido o Naive Bayes ser um classificador que tem como pressuposto fundamental a independência dos preditores (Arar e Ayan, 2017). Em outras palavras, Naive Bayes assume que cada nova banda adiciona informação nova e independente. Quando bandas altamente correlacionadas (como 5, 6 e 7) são incluídas, o modelo interpreta que está recebendo mais evidências do que de fato possui, atribuindo peso excessivo a informações redundantes. Isso leva a uma superestimação das probabilidades de classe, reduzindo a precisão da predição e, conseqüentemente, o Kappa.

Por outro lado, para final de ciclo, as métricas Kappa melhoraram para o desempenho "muito bom" e "excelente" com o melhor modelo sendo o Random Forest com seu kappa de 0,818 e uma área de 14.583,57 (+0,1 acima do manual) e o pior modelo foi o Naive Bayes com o Kappa de 0,181 demonstrando novamente que este modelo com mais bandas e o desafio de diferenciar milho somente com o final do ciclo não desempenha bem com rasters stack.

O salto de acurácia obtido pelo RF com o stack multiespectral confirma o valor informacional agregado pelas bandas espectrais adicionais (Ex., B5, B6, B7), mesmo que colineares. Enquanto as bandas SWIR são cruciais para a separação de classes com base no teor de umidade da vegetação (Holzman *et al.*, 2021), as bandas Red Edge fornecem informações sobre a estrutura celular e o conteúdo de clorofila (Zhang *et al.*, 2022) o que pode permitir ao RF distinguir culturas em estágios fenológicos

semelhantes ou diferenciar culturas com assinaturas espectrais complexas (ex: separação entre sorgo e milho). O RF, portanto, atuou como um bom classificador para milho em final de ciclo, utilizando a informação espectral rica do Stack sem ser prejudicado pela redundância que afetou o Naive Bayes.

3.4 METODOLOGIA QUALITATIVA E PARTICIPATIVA

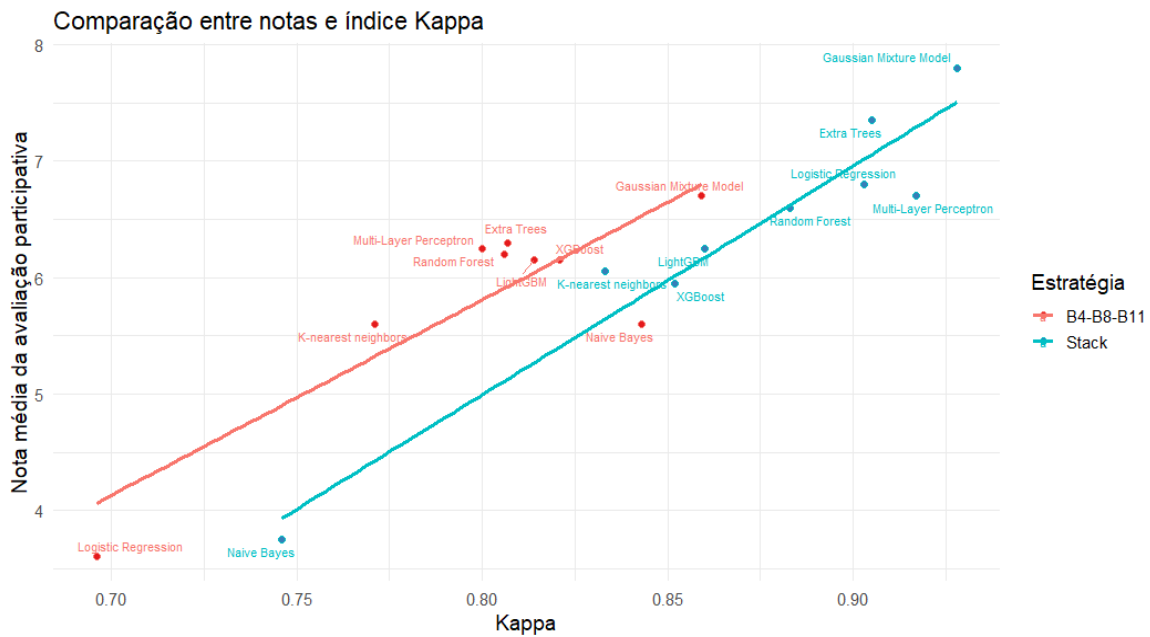


Figura 9. Gráfico de Comparação das notas e índice Kappa.

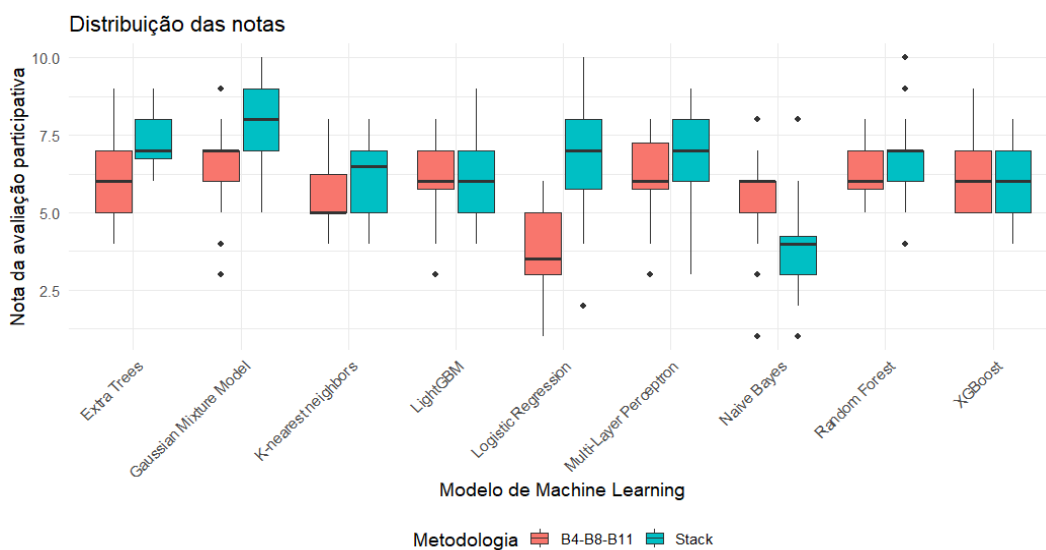


Figura 10. Distribuição das notas entre modelos e metodologias.

Os resultados da composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4), o modelo GMM que apresentou a maior média (6,7), seguido pelos modelos ET (6,3), MLP (6,25), RF (6,2), LGB e XGB (ambos com 6,15), KNN (5,6) e LR com 3,6, indicando desempenho inferior em relação aos demais. As medianas variaram entre 3,5 (LR) até 7 (GMM), enquanto as modas oscilaram principalmente entre 6 e 7, com destaque para os modelos ET, GMM e LGB, que atingiram a moda 7. Em termos de amplitude, os valores mínimos variaram de 1 (LR e LGB) a 5 (RF e XGB) e os máximos ficaram entre 6 (LR) e 9 (ET, GMM e XGB). O desvio-padrão foi relativamente baixo para todos os modelos, variando entre 0,98 (RF) e 1,56 (NB) o que indica consistência nas avaliações, com destaque para o Random Forest, que apresentou menor variabilidade nos resultados. Estes resultados da metodologia de avaliação participativa sugerem que o modelo com os melhores parâmetros de média, mediana, moda e nota máxima da composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4) foi o Gaussian Mixture Model, justamente o modelo com maior índice Kappa, bem como os piores valores foram referentes ao Logistic Regression que foi o pior modelo em relação a índice Kappa nesta composição.

Os resultados da composição stack, o modelo GMM apresentou o melhor desempenho médio (7,8), seguido pelo ET (7,35), LR (6,8), MLP (6,7), RF (6,6) e LGB (6,25), KNN (6,05), XGB (5,95) e o NB obteve a menor média (3,75), indicando desempenho inferior. As medianas variaram de 4 (NB) a 8 (GMM), enquanto as modas mostraram comportamento semelhante, com destaque para LR, MLP e GMM, que apresentaram moda 8. Os valores mínimos oscilaram entre 1 (NB) e 6 (ET), e os máximos variaram de 8 (NB, KNN e XGB) a 10 (GMM e RF). Quanto à dispersão dos resultados, o desvio-padrão foi mais baixo para ET (1,01) e XGB (1,02), sugerindo maior estabilidade nesses modelos, enquanto MLP (1,55) e LR (1,91) apresentaram maior variação entre as execuções. De modo geral, obteve o mesmo resultado, onde o modelo Gaussian Mixture Model obteve a maior média, mediana e moda. Novamente o pior modelo respondido pelos intérpretes foi o pior índice kappa, o Naive bayes.

4 CONCLUSÕES

O algoritmo Gaussian Mixture Model (GMM) demonstrou alta eficácia na diferenciação da cultura do milho primeira safra no município de Luís Eduardo Magalhães, Bahia, atingindo um Coeficiente Kappa de 0,928, com área predita que variou apenas -0,017% da área mapeada manualmente. O desempenho robusto do GMM o qualifica como uma ferramenta promissora de pré-mapeamento, podendo automatizar a vetorização e a interpretação preliminar dos talhões, resultando em uma significativa redução do tempo de retrabalho e do custo operacional para o intérprete humano.

Este trabalho valida a aplicação de modelos de Machine Learning para a agricultura e reforça que a escolha do classificador deve considerar a dimensionalidade dos dados. Para trabalhos futuros, procuraremos utilizar abordagens híbridas que combinem o desempenho robusto do Random Forest com o potencial do GMM para otimizar etapas específicas, como a vetorização das classes. Os resultados obtidos demonstram que se ocorrer essa integração podem elevar tanto a acurácia quanto a eficiência do mapeamento agrícola, especialmente quando cada algoritmo é direcionado às tarefas em que apresenta maior vantagem. Pode-se avaliar a inclusão de outros classificadores, técnicas de pós-processamento espacial e a incorporação de índices temporais (NDVI, EVI, SAVI, GCVI etc.), buscando aprimorar a consistência das classes ao longo da safra e aumentar a capacidade de generalização dos modelos em diferentes regiões produtoras.

5 REFERÊNCIAS

AGROLINK. **Estatísticas Agropecuárias.** Disponível em: <<https://www.agrolink.com.br/regional/ba/luis-eduardo-magalhaes/estatistica>>. Acesso em: 23 set. 2025.

AGRORECEITA. **Terceira safra: mais sustentabilidade e produtividade no campo.** Disponível em: <<https://agroreceita.com.br/terceira-safra/>>. Acesso em: 23 set. 2025.

ARAR, Ömer Faruk; AYAN, Kürşat. **A feature dependent Naive Bayes approach and its application to the software defect prediction problem.** Applied Soft Computing, v. 59, p. 197–209, 2017.

COMPANHIA NACIONAL DE ABASTECIMENTO - CONAB. **Acompanhamento da Safra Brasileira.** Disponível em: <<https://www.gov.br/conab/pt-br/atuacao/informacoes-agropecuarias/safras>>. Acesso em: 23 set. 2025.

COMPANHIA NACIONAL DE ABASTECIMENTO - CONAB. **Calendário de Plantio e Colheita de Grãos no Brasil 2022.** Disponível em: <<https://www.gov.br/conab/pt-br/aceso-a-informacao/institucional/publicacoes/arquivos-de-paginas/calendariozplantiozezcolheitazjunz2022.pdf>>. Acesso em: 23 set. 2025.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA (2025). **SATVeg - Sistema de Análise Temporal da Vegetação.** Disponível em: <<http://www.satveg.cnptia.embrapa.br>>. Acesso em: 23 set. 2025.

FORMAGGIO, Antonio Roberto; SANCHES, Ieda Del'arco. **Sensoriamento remoto em agricultura.** São Paulo: Oficina de Textos, 2017. Disponível em: <<https://books.google.com.br/books?id=hk88DwAAQBAJ&printsec=frontcover&hl=ptBR#v=onepage&q&f=false>>. Acesso em: 23 set. 2025.

GORELICK, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). **Google Earth Engine: Análise geoespacial em escala planetária para todos. Sensoriamento Remoto do Meio Ambiente .** Disponível em: <<https://code.earthengine.google.com/>>. Acesso em: 23 set. 2025.

HOLZMAN, Mauro Ezequiel; RIVAS, Raúl Eduardo; BAYALA, Martín Ignacio. **Relationship between TIR and NIR-SWIR as Indicator of Vegetation Water Availability.** Remote Sensing, v. 13, n. 17, p. 3371, 2021.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. **Anuário estatístico dos Estados.** Disponível em: <<https://www.ibge.gov.br/cidades-e-estados>>. Acesso em: 23 set. 2025.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. **Anuário estatístico da Bahia.** Disponível em: <<https://www.ibge.gov.br/cidades-e-estados/ba.html>>. Acesso em: 23 set. 2025.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. **Anuário estatístico de Luís Eduardo Magalhães**. Disponível em: <<https://www.ibge.gov.br/cidades-e-estados/ba/luis-eduardo-magalhaes.html>>. Acesso em: 23 set. 2025.

KARASIAK, Nicolas. (2016) **Dzetsaka Qgis Classification plugin**. Disponível em: <<https://github.com/nkarasiak/dzetsaka>>. DOI: 10.5281/zenodo.2552284. Acesso em: 23 set. 2025.

LANDIS, J.R.; KOCH, G.G (1977). **The measurement of observer agreement for categorical data**, In: Biometrics, v. 33, n. 1, p. 159-174.

NOVO, E. M. L. de M. **Sensoriamento Remoto - Princípios e Aplicações** – Editora Edgar Blucher Ltda., 2015. 4. ed. 387 p.

PITMAN, R; HU, B (2022). **IMPROVING THE BINARY CLASSIFICATION OF PEAT LOCALITIES FROM MULTI-SOURCE REMOTELY-SENSED DATA USING CNN**. Disponível em: <<https://isprs-archives.copernicus.org/articles/XLIII-B3-2022/983/2022/isprs-archives-XLIII-B3-2022-983-2022.pdf>>. Acesso em: 23 set. 2025.

QGIS Development Team. **QGIS Geographic Information System**. Bratislava: Open Source Geospatial Foundation Project, 2025. Disponível em: <<http://qgis.org>>. Acesso em: 23 set. 2025.

SALGADO, C. B. et al. **Análise da interferência de nuvens na classificação de séries temporais MODIS-NDVI na região da Amazônia**, município de Capixaba, Acre. Disponível em: <<https://www.scielo.br/j/sn/a/FG9GJcBVP8jCphkKv74Q5hN/?format=pdf&lang=pt>>. Acesso em: 23 set. 2025.

UNIÃO NACIONAL DO ETANOL DE MILHO - UNEM. Disponível em: <<https://etanoldemilho.com.br/>>. Acesso em: 23 set. 2025.

VENANCIO, Luan Peroni; FILGUEIRAS, Roberto; CUNHA, Fernando França da; SILVA, Francisco Charles dos Santos; SANTOS, Robson Argolo dos; MANTOVANI, Everardo Chartuni. **Mapeamento dos estádios fenológicos da cultura do milho utilizando o NDVI derivado dos sensores OLI e MODIS**. DOI: 10.5433/1679-0359.2020v41n5p1517. Disponível em: <<https://ojs.uel.br/revistas/uel/index.php/semagrarias/article/view/36641/27518>>. Acesso em: 17 nov. 2025.

ZHANG, Hu et al. **A novel red-edge spectral index for retrieving the leaf chlorophyll content**. Methods in Ecology and Evolution, v. 13, n. 12, p. 2771–2787, 2022.

ANEXO I

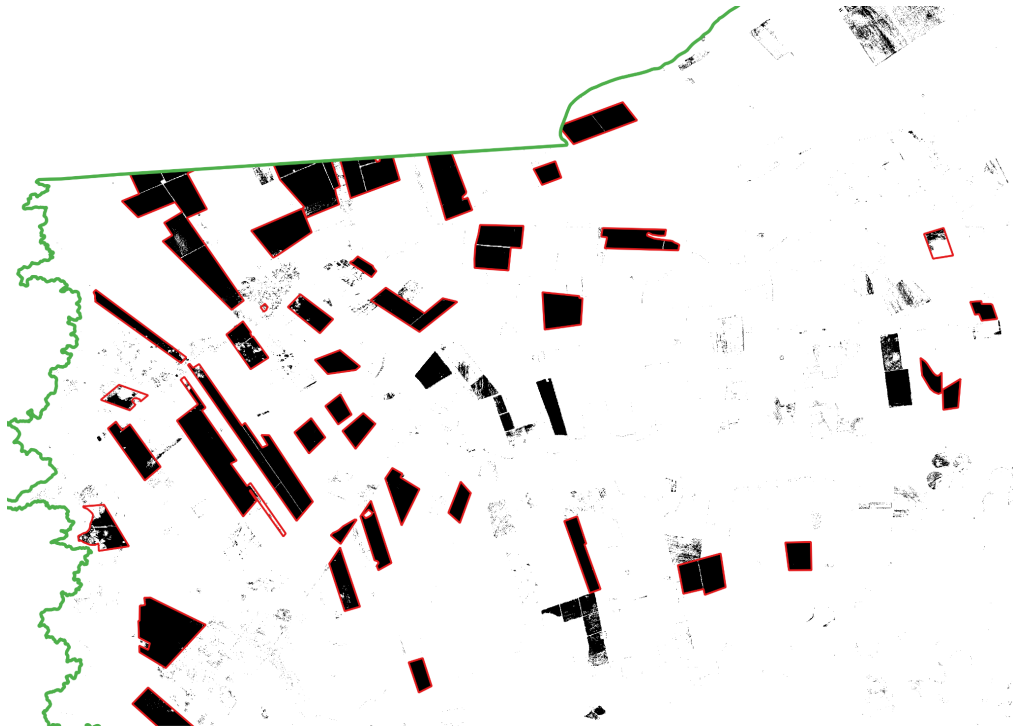


Figura anexo 1. Imagem do preenchimento e ruídos do modelo Extra Trees na composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4).

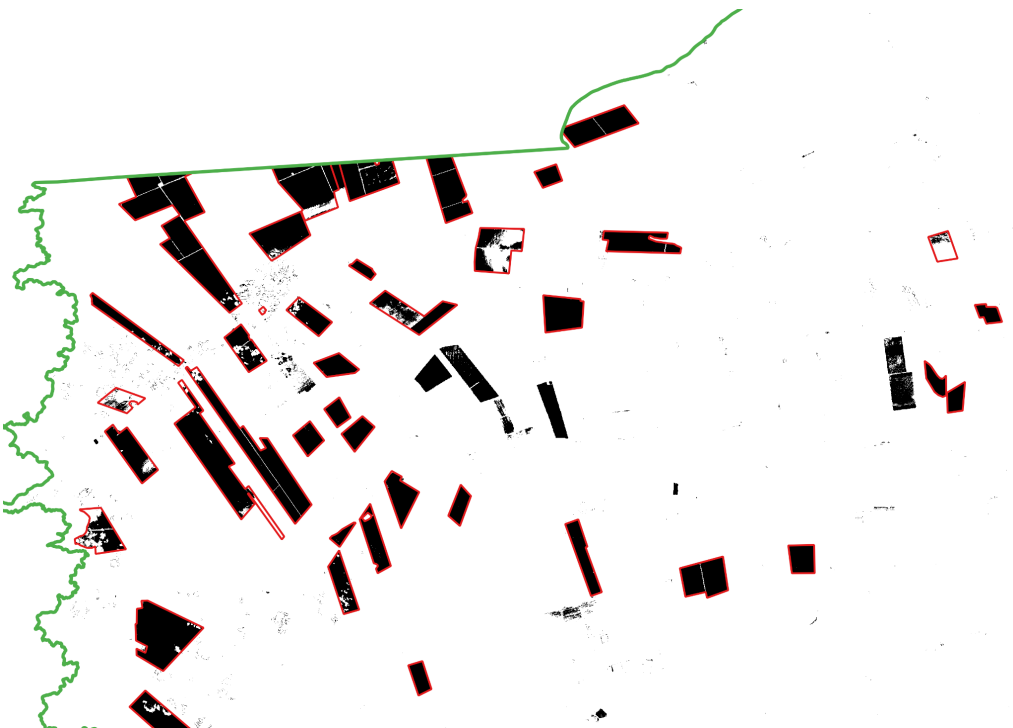


Figura anexo 2. Imagem do preenchimento e ruídos do modelo Gaussian Mixture Model na composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4).

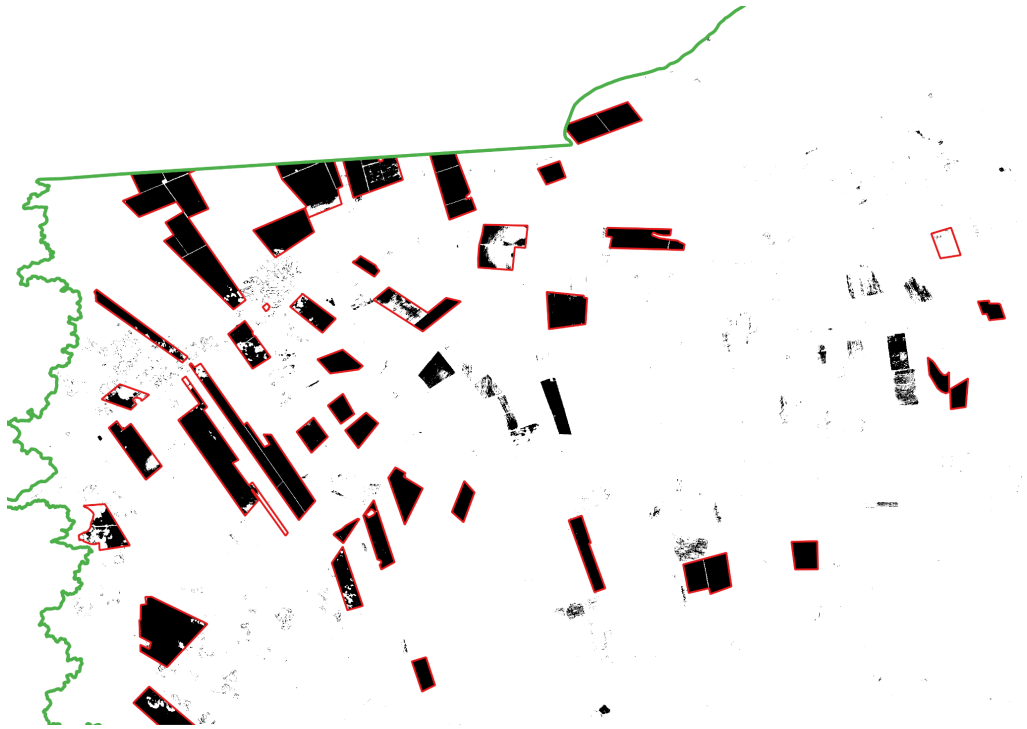


Figura anexo 3. Imagem do preenchimento e ruídos do modelo Naive Bayes na composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4).

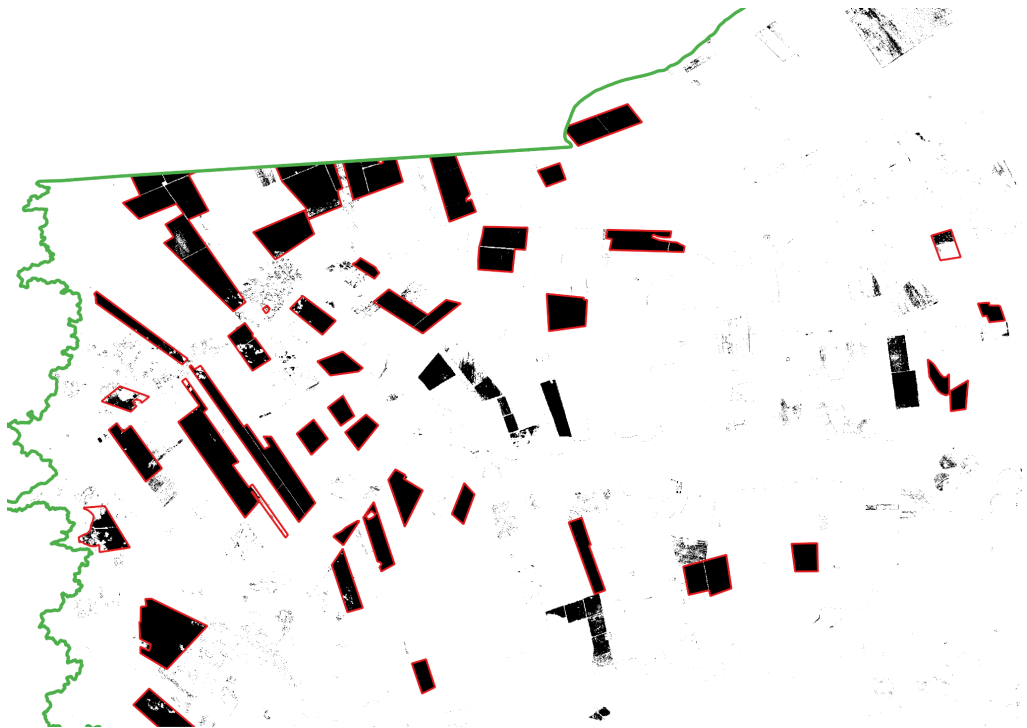


Figura anexo 4. Imagem do preenchimento e ruídos do modelo LightGBM na composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4).

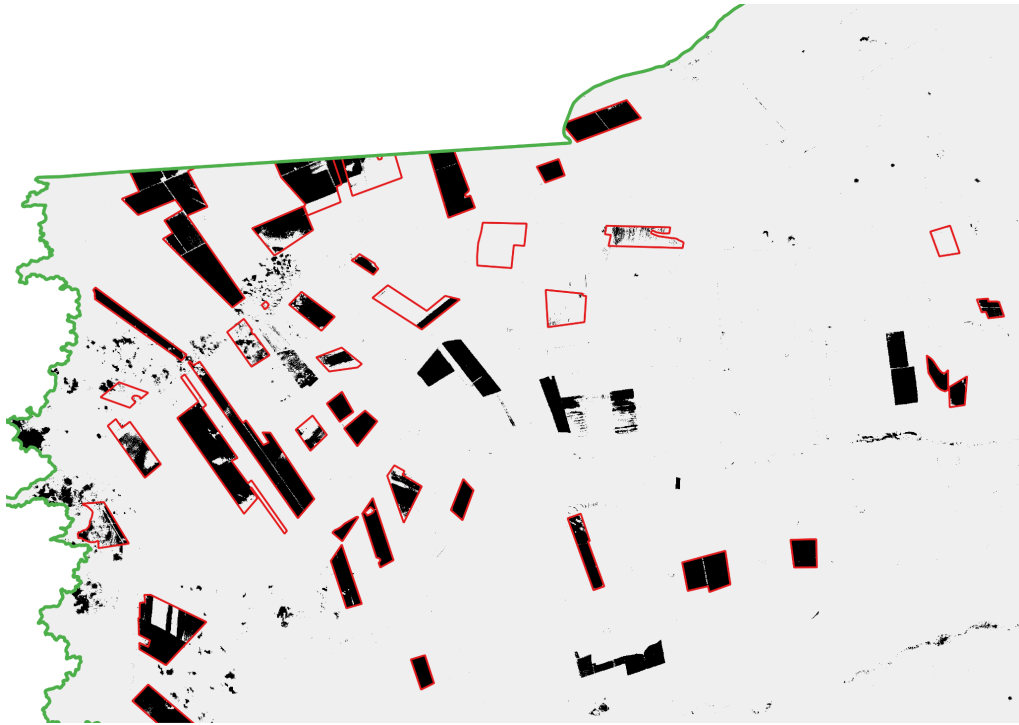


Figura anexo 5. Imagem do preenchimento e ruídos do modelo Logistic Regression na composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4).

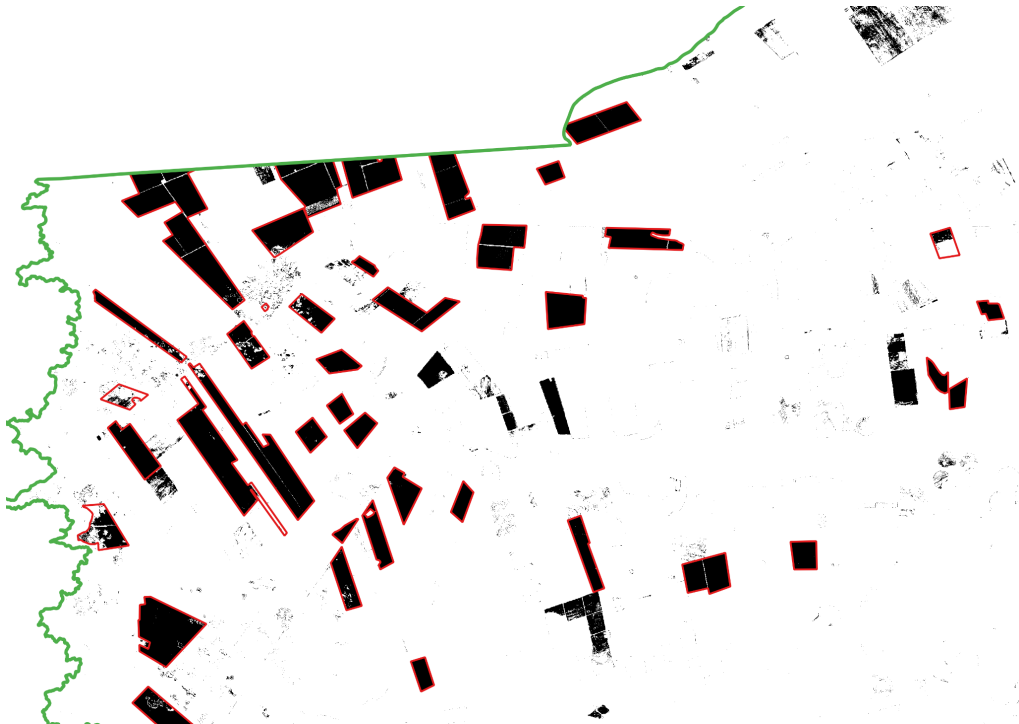


Figura anexo 6. Imagem do preenchimento e ruídos do modelo Multi-layer Perceptron na composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4).

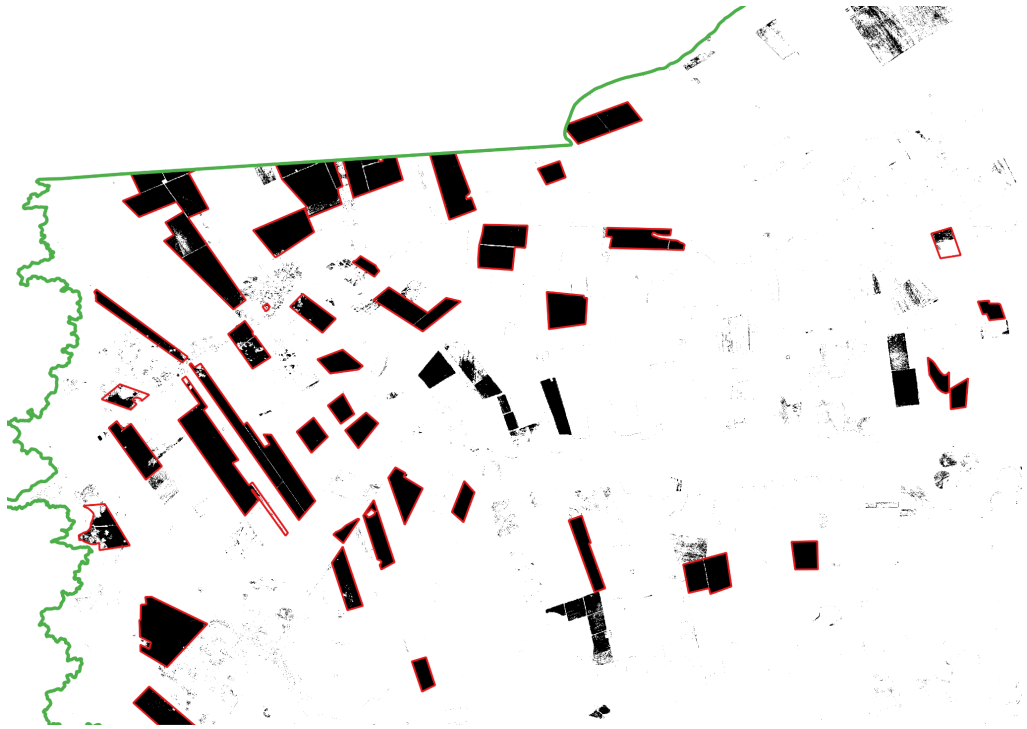


Figura anexo 7. Imagem do preenchimento e ruídos do modelo Random forest na composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4).

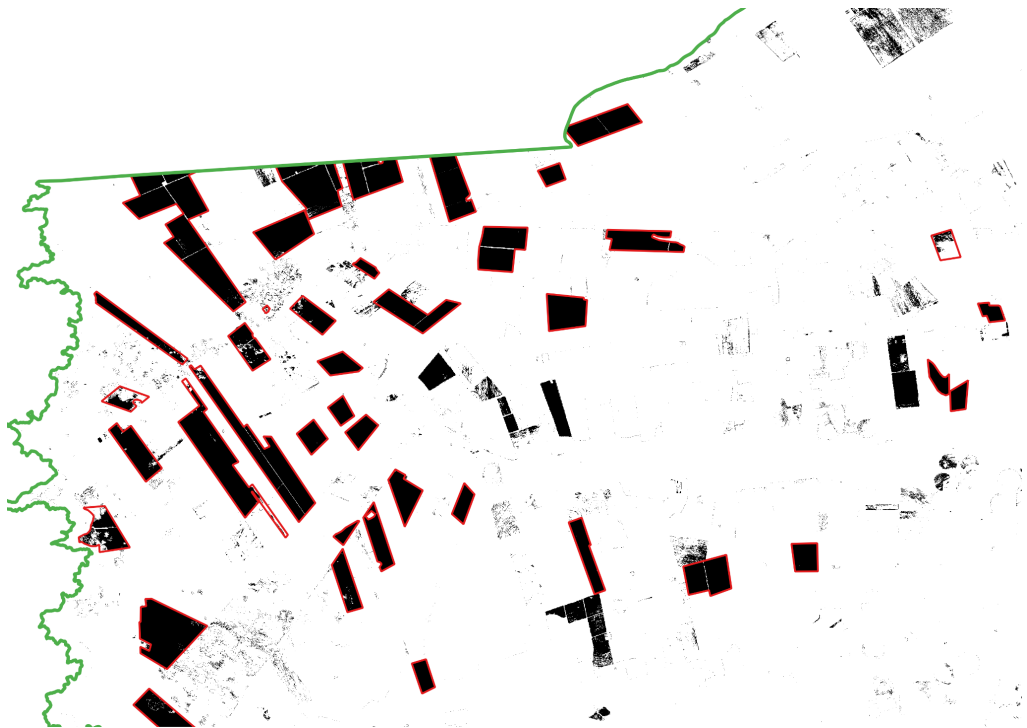


Figura anexo 8. Imagem do preenchimento e ruídos do modelo K-Nearest Neighbors na composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4).

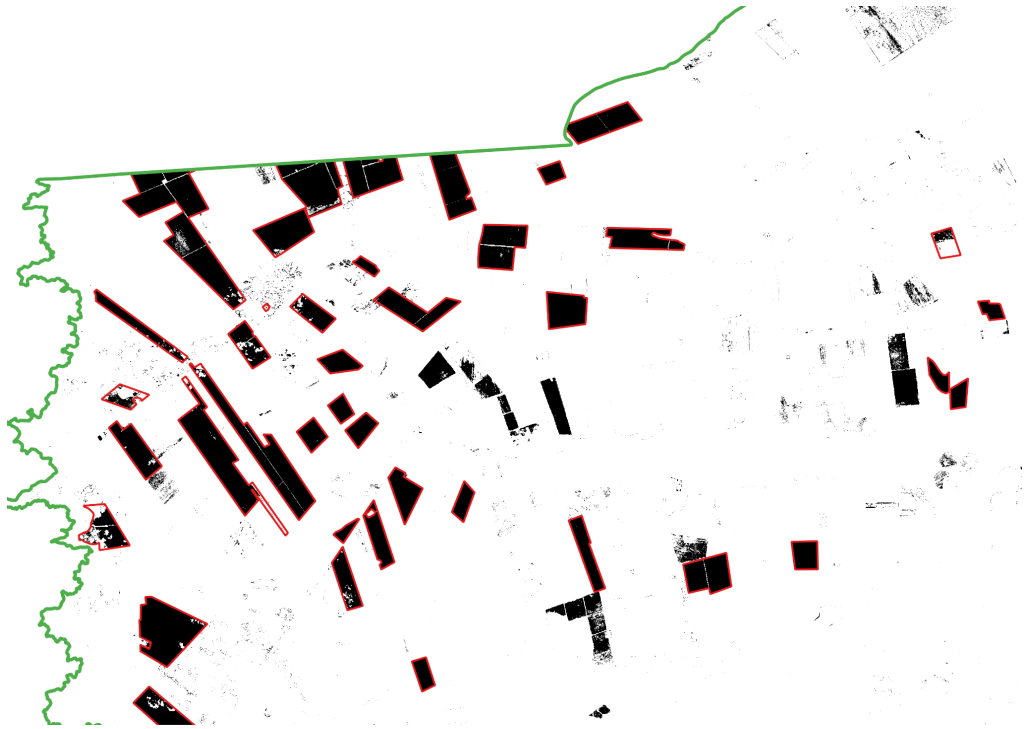


Figura anexo 9. Imagem do preenchimento e ruídos do modelo XGBoost na composição infravermelho próximo (8), infravermelho de ondas curtas (11) e vermelho (4).

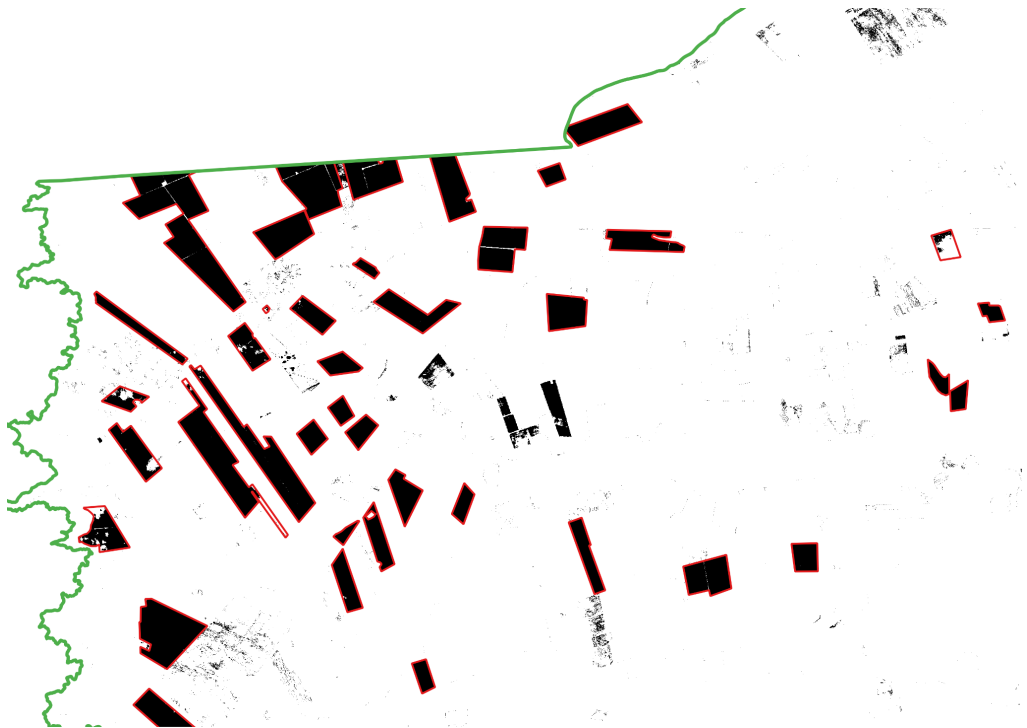


Figura anexo 10. Imagem do preenchimento e ruídos do modelo Extra Trees na composição infravermelho próximo (8), infravermelho de ondas curtas (11), vermelho (4), aerossol (1), azul (2), verde (3), red edge (5,6,7).

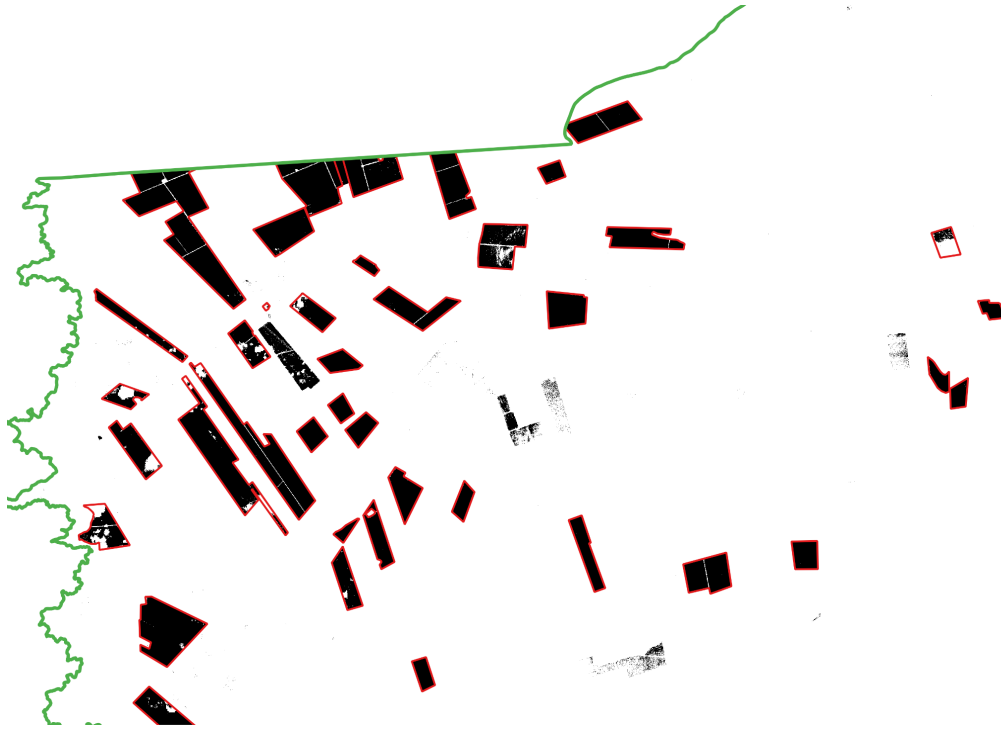


Figura anexo 11. Imagem do preenchimento e ruídos do modelo Gaussian Mixture Model na composição infravermelho próximo (8), infravermelho de ondas curtas (11), vermelho (4), aerossol (1), azul (2), verde (3), red edge (5,6,7).

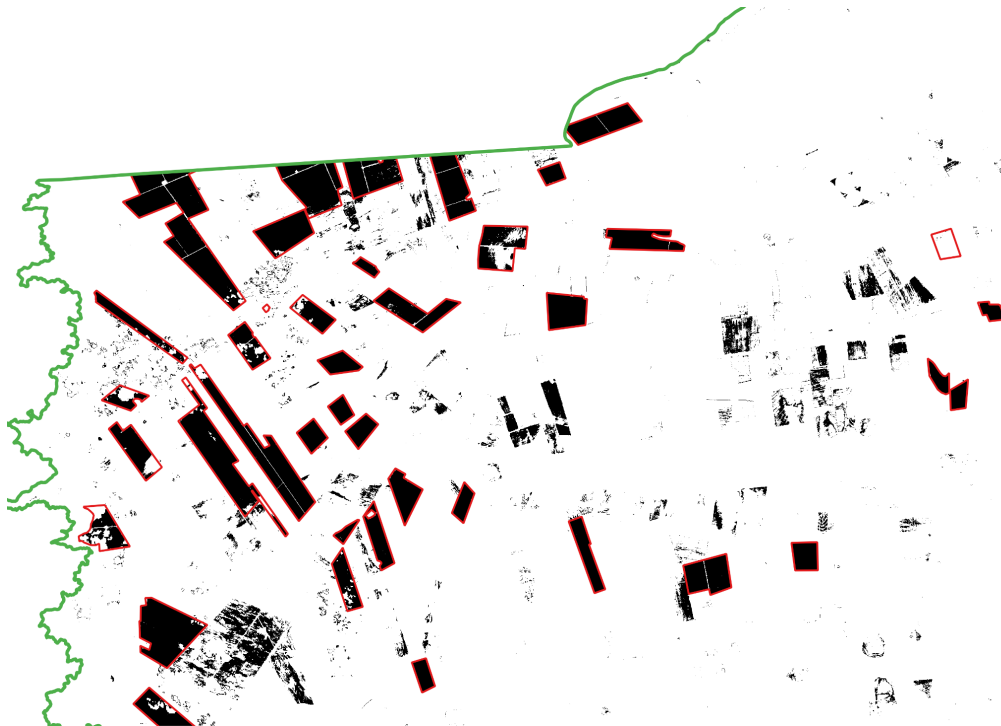


Figura anexo 12. Imagem do preenchimento e ruídos do modelo Naive Bayes na composição infravermelho próximo (8), infravermelho de ondas curtas (11), vermelho (4), aerossol (1), azul (2), verde (3), red edge (5,6,7).

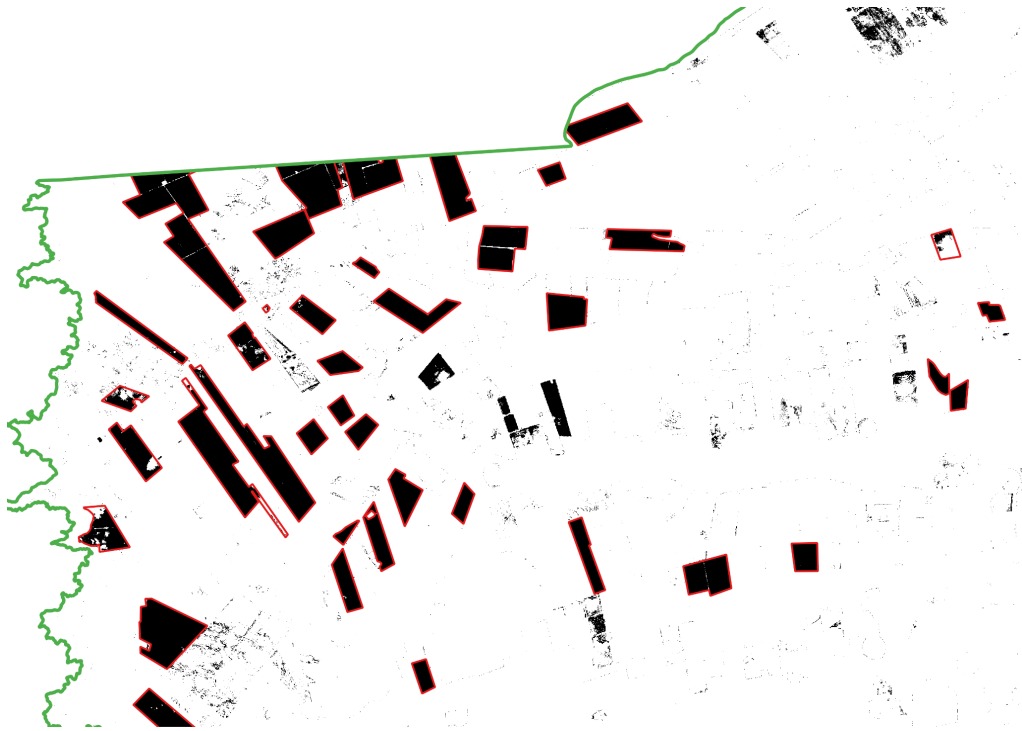


Figura anexo 13. Imagem do preenchimento e ruídos do modelo LightGBM na composição infravermelho próximo (8), infravermelho de ondas curtas (11), vermelho (4), aerossol (1), azul (2), verde (3), red edge (5,6,7).

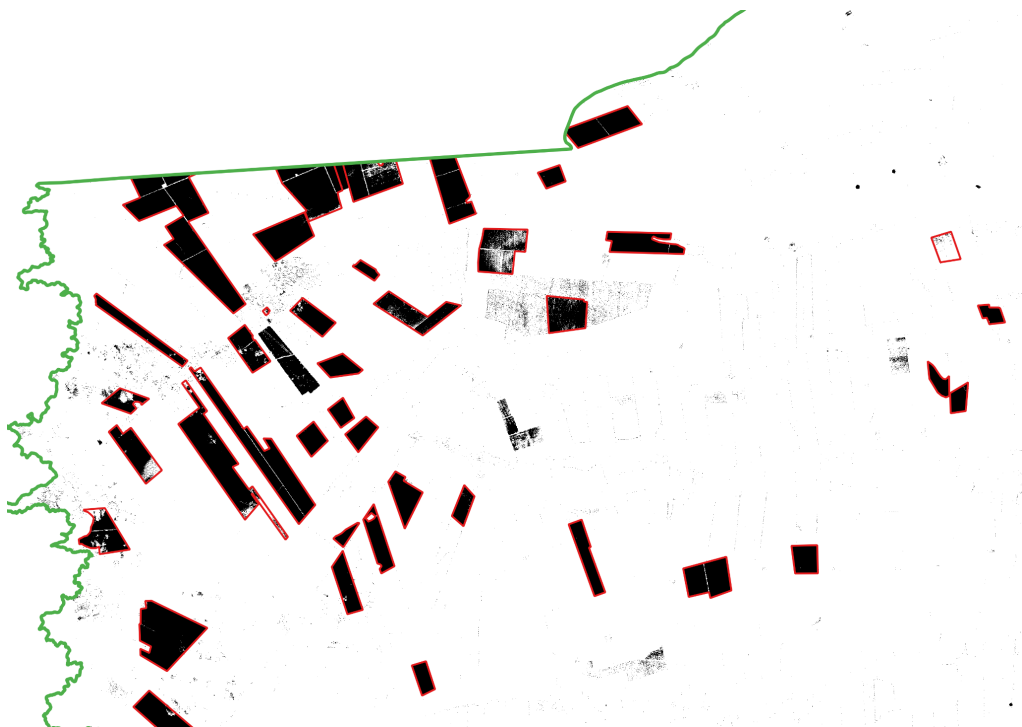


Figura anexo 14. Imagem do preenchimento e ruídos do modelo Logistic Regression na composição infravermelho próximo (8), infravermelho de ondas curtas (11), vermelho (4), aerossol (1), azul (2), verde (3), red edge (5,6,7).

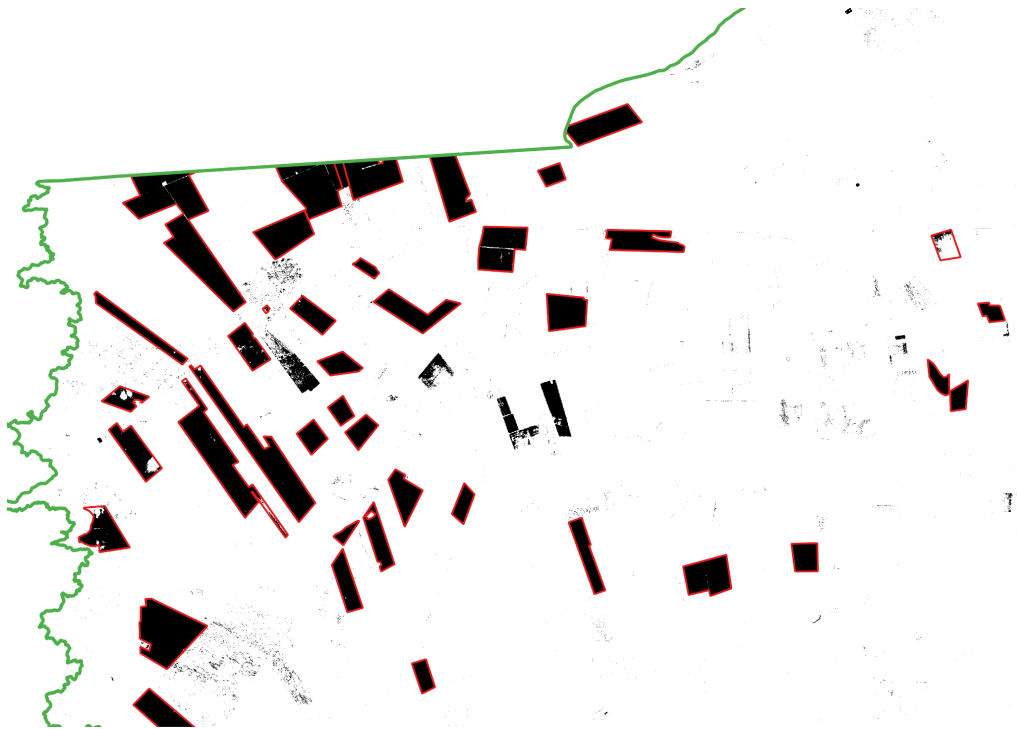


Figura anexo 15. Imagem do preenchimento e ruídos do modelo Multi-layer Perceptron na composição infravermelho próximo (8), infravermelho de ondas curtas (11), vermelho (4), aerossol (1), azul (2), verde (3), red edge (5,6,7).

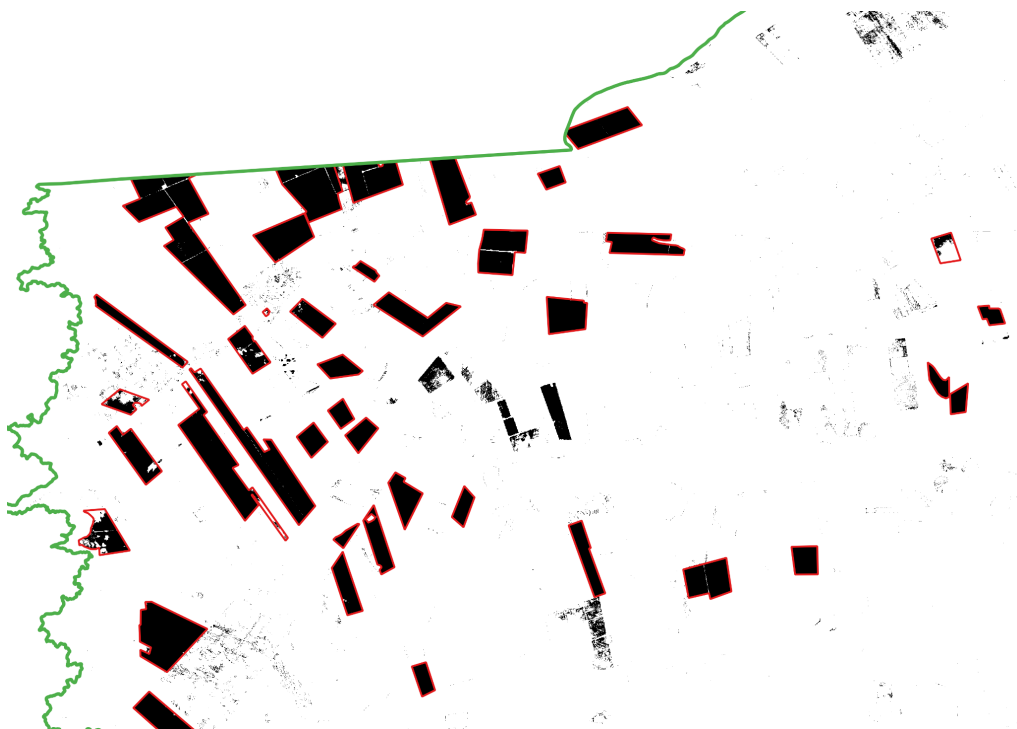


Figura anexo 16. Imagem do preenchimento e ruídos do modelo Random Forest na composição infravermelho próximo (8), infravermelho de ondas curtas (11), vermelho (4), aerossol (1), azul (2), verde (3), red edge (5,6,7).

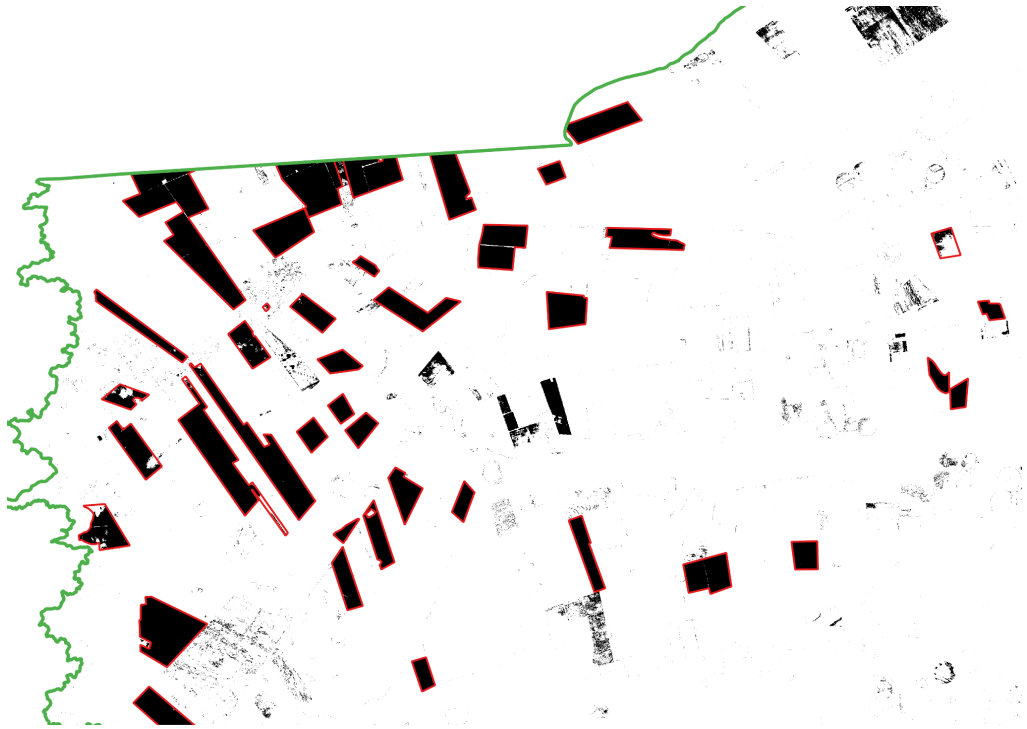


Figura anexo 17. Imagem do preenchimento e ruídos do modelo K-Nearest Neighbors na composição infravermelho próximo (8), infravermelho de ondas curtas (11), vermelho (4), aerossol (1), azul (2), verde (3), red edge (5,6,7).

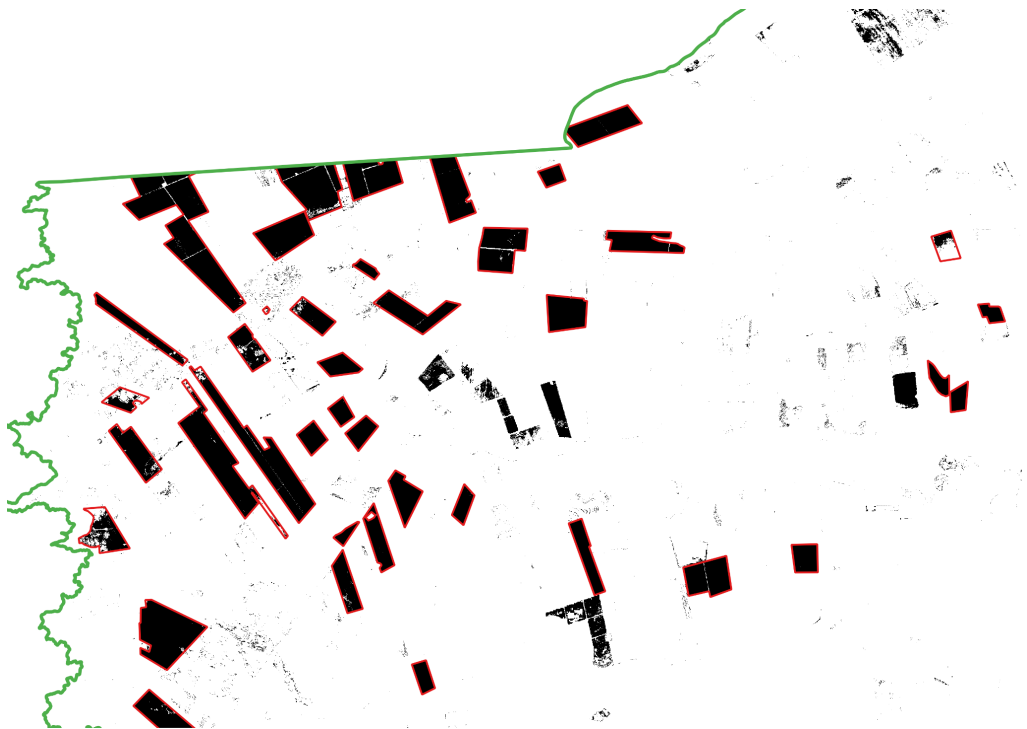


Figura anexo 18. Imagem do preenchimento e ruídos do modelo XGBoost na composição infravermelho próximo (8) , infravermelho de ondas curtas (11), vermelho (4), aerossol (1), azul (2), verde (3), red edge (5,6,7).