

**RAIZA MONIZ FARIA**

**MODELAGEM DO AMARELECIMENTO FATAL NA PALMA DE ÓLEO**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Solos e Nutrição de Plantas, para obtenção do título de *Doctor Scientiae*.

Orientador: Elpídio Inácio Fernandes Filho

Coorientadores: Júlio Cesar Lima Neves  
Márcio Rocha Francelino

**VIÇOSA - MINAS GERAIS  
2022**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade  
Federal de Viçosa - Campus Viçosa**

T

F224m  
2022 Faria, Raiza Moniz, 1991-  
Modelagem do amarelecimento fatal na palma de óleo /  
Raiza Moniz Faria. – Viçosa, MG, 2022.  
1 tese eletrônica (57 f.): il. (algumas color.).

Inclui anexos.

Orientador: Elpídio Inácio Fernandes Filho.

Tese (doutorado) - Universidade Federal de Viçosa,  
Departamento de Solos, 2022.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2022.799>

Modo de acesso: World Wide Web.

1. Agricultura de precisão - Sensoriamento remoto.  
2. Dendzeiro - Doenças e pragas - Controle. 3. Aprendizado do  
computador. I. Fernandes Filho, Elpídio Inácio, 1963-.  
II. Universidade Federal de Viçosa. Departamento de Solos.  
Programa de Pós-Graduação em Solos e Nutrição de Plantas.  
III. Título.

CDD 22. ed. 631.3

**RAIZA MONIZ FARIA**

**MODELAGEM DO AMARELECIMENTO FATAL NA PALMA DE ÓLEO**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Solos e Nutrição de Plantas, para obtenção do título de *Doctor Scientiae*.

APROVADA: 04 de outubro de 2022.

Assentimento:

Raiza Moniz Faria  
Autora

Elpídio Inácio Fernandes Filho  
Orientador

A Deus e a Nossa Senhora Aparecida

A minha família

Eu dedico

## AGRADECIMENTOS

04 de outubro de 2022. Data que ficará registrada em minha vida: o encerramento de mais um ciclo. Um ciclo de 11 anos como estudante da Universidade Federal de Viçosa. Sem dúvidas, o primeiro agradecimento é a Deus. Todos os dias em que entrei pelo saguão do Departamento de Solos, pedi proteção a Nossa Senhora Aparecida, pedia que guiasse meus passos, meus pensamentos, minhas ações em mais um dia de estudos. Obrigada Nossa Senhora.

A UFV foi minha casa. Agradeço a todos os meus colegas de graduação, da pós-graduação e a todos os professores que contribuíram com minha formação. A todos os funcionários desta universidade que sempre mantiveram a minha casa arrumada.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Ao longo destes anos fiz tantos amigos, conheci tantas pessoas. Tive a oportunidade de contribuir de diversas maneiras na formação de muitos estudantes. Mal sabia eu que, a principal beneficiária era eu mesma. Ao final deste ciclo, nestes agradecimentos, não posso me esquecer de mencionar o Laboratório de Geoprocessamento, coordenado pelos professores Elpídio Fernandes e Márcio Francelino. Professor Elpídio, foi e é muito mais que um orientador, nele tenho um grande amigo. Nossa amizade perdurará por longos anos. Professor Márcio, agradeço não só por ter ido ao campo comigo para a coleta dos dados, mais também por ter permitido que o laço da confiança e da amizade fosse criado. Agradeço também a todos os meus colegas de laboratório ao longo dos 10 anos de Labgeo, pelos bons momentos. Muitos saíram do convívio diário, mas não saíram da minha vida. Grandes amizades foram formadas.

Ao professor Júlio César Lima Neves, agradeço pela coorientação, ensinamentos e auxílio em todas as fases deste trabalho. Agradeço também por me apresentado a cultura da palma de óleo.

Ao professor João Ker por ter ido ao campo comigo, por todo auxílio na parte física do solo desta pesquisa.

Agradeço toda a atenção e zelo que Carol e Nayan tiveram com as burocracias existentes até a defesa da tese.

Aos membros da banca examinadora pelas valiosas contribuições.

Saber que este trabalho não ficará engavetado serviu como motivação para todos àqueles dias em que o desânimo “batia”. Esta tese contou com o apoio financeiro da empresa Agropalma, agradeço pela parceria e fomento, oferecendo todo o suporte no desenvolvimento desta pesquisa. Agradeço aos funcionários André Borba, Carlos Garcia, Filipe, Gilson, Hugo, Jailson, Josivaldo, Leonan, Raqueline, Raryssa, Valdinei, e àqueles que após eu retornar a Viçosa me auxiliaram. Aos colegas de empresa que fiz durante esses meses na Agropalma.

Agradeço ao Bernardo, meu braço direito nesta pesquisa.

Aos colegas de pós-graduação do Departamento de Solos. Agradeço a meus grandes amigos Angélica Ladeira, Elton Costa, Isabela Filardi, Fábio Maia, Joelmo, Patrícia Matias, Patrick Souza, Rafael Cachú, Kátia Maciel, Mayara Daher e Yago Pena por estarem comigo nesta caminhada, tornando-a mais leve e cheia de risadas.

Agradeço e dedico parte desta conquista a minha família. A melhor parte de mim, encontro em vocês. Meu esposo Pedro, meu grande amor, que muito me ajudou. A minha mãe Roseni e meu pai José Ivo que me forneceram uma base consolidada para que os estudos chegassem até aqui. Ao Orlando, que entrou a pouco em nossa família e já é tão querido. A minha irmã Maola pelas energias positivas e apoio. Obrigada por terem me incentivado a estudar.

Enfim, agradeço a todos aqueles que sabem da sua importância nesta e nas demais conquistas. E me desculpa se esqueci de mencionar algum nome, é que graças a Deus tenho muitas pessoas boas ao meu redor.

*“Se a única oração que você fizer na vida for obrigada, já será o suficiente”*

*(EckhartJolle)*

## BIOGRAFIA

Raiza Moniz Faria, filha de Roseni Aparecida de Souza e de José Ivo de Faria, natural de Guiricema, estado de Minas Gerais. Nasceu no dia 04 de novembro de 1991.

Em março de 2010 iniciou o curso de Geografia na Universidade Federal de Viçosa, onde graduou-se em licenciatura em janeiro de 2015 e no bacharelado em julho de 2015. Durante a graduação, fez estágio voluntário e foi bolsista de iniciação científica nos laboratórios de Soluções em Inteligência Geográfica para o Meio Ambiente (SIGMA) e Laboratório de Geoprocessamento (Labgeo).

Em agosto de 2015 ingressou no Programa de Pós-Graduação em Solos e Nutrição de Plantas, pela Universidade Federal de Viçosa, defendendo sua dissertação em julho de 2017, com título “*Mapeamento de Áreas Cafeeiras em Relevo Montanhoso no estado de Minas Gerais*”.

Em março de 2018 iniciou seu doutorado no Programa de Pós-Graduação em Solos e Nutrição de Plantas, pela Universidade Federal de Viçosa, submetendo a defesa de sua tese em outubro de 2022.

## RESUMO

FARIA, Raiza Moniz Faria, D.Sc., Universidade Federal de Viçosa, outubro de 2022. **Modelagem do amarelecimento fatal na palma de óleo.** Orientador: Elpídio Inácio Fernandes Filho. Coorientadores: Julio Cesar Lima Neves e Márcio Rocha Francelino.

A cultura da Palma de Óleo é de grande importância no cenário mundial. O óleo extraído de seus frutos é um óleo versátil utilizado em indústrias de cosméticos e de alimentos. No Brasil, os plantios concentram-se nos estados do Pará, Roraima e Bahia. A demanda por óleo de palma cresceu nos últimos anos em concomitância com o aumento das doenças, pragas e distúrbios na cultura. No Brasil, o distúrbio do amarelecimento fatal (AF) vem sendo um entrave para esta cultura, dizimando grandes áreas na região nordeste do estado do Pará. O sensoriamento remoto apresenta uma série de dados que podem servir de subsídio a estudos do distúrbio, com o uso de geotecnologias é possível estabelecer as variáveis com maior correlação no desenvolvimento do AF na cultura da palma de óleo. O objetivo da tese foi desenvolver uma metodologia adequada para as áreas com AF em plantios comerciais na região nordeste do estado do Pará, com a utilização de imagens de média e alta resolução e variáveis bióticas, abióticas e espectrais. A utilização de imagens de altíssima resolução na faixa espectral do visível foi avaliada para a identificação de plantas com AF. Os resultados indicam que a integração do sensoriamento remoto com *machine learning* é promissor para o desenvolvimento de pesquisas de distúrbios como o AF na palma de óleo. A classificação orientada a objetos apresentou acurácia de 94% e índice kappa de 96% com 25 variáveis preditoras. Destas, os índices de vegetação GCI, TVI, MCARI e ERGBVE foram os mais importantes e promissores para serem utilizados na avaliação do status nutricional da cultura e na identificação de doenças e distúrbios. A utilização destes índices e das demais variáveis associadas ao modelo construído, possibilita maior controle fitossanitário dos plantios. As imagens do satélite Sentinel disponíveis de forma gratuita e com periodicidade de 5 dias podem ser utilizadas para o mapeamento do AF. A metodologia desenvolvida e aplicada foi satisfatória na separação entre plantas com distúrbio provocado por AF e plantas saudáveis.

Palavras-chave: Modelagem de dados. Óleo de palma. Agricultura de precisão.

## ABSTRACT

FARIA, Raiza Moniz Faria, D.Sc., Universidade Federal de Viçosa, October, 2022. **Modeling bud rot in oil palm.** Adviser: Elpídio Inácio Fernandes Filho. Co-advisers: Julio Cesar Lima Neves and Márcio Rocha Francelino.

The culture of Palm Oil is of great importance on the world stage. The oil extracted from its fruits is a versatile oil used in the cosmetic and food industries. In Brazil, plantations are concentrated in the states of Pará, Roraima and Bahia. The demand for palm oil has grown in recent years in line with the increase in diseases, pests and disturbances in the crop. In Brazil, the fatal yellowing disorder has been an obstacle for this crop, decimating large areas in the northeastern region of the state of Pará. Remote sensing presents a series of data that can serve as a basis for studies of the disturbance, with the use of geotechnologies it is possible to establish the variables with the greatest correlation in the development of fatal yellowing in the oil palm crop. The main objective of the thesis was to develop an adequate methodology for areas with fatal yellowing in commercial plantations in the northeast region of the state of Pará, using medium and high resolution images and biotic, abiotic and spectral variables. The performance of vegetation indices specific to precision agriculture was also evaluated. The use of very high resolution images in the visible spectral range was evaluated for the identification of plants with fatal yellowing. The results indicate that the integration of remote sensing with machine learning is promising for the development of research on disorders such as fatal yellowing in oil palm. The object-oriented classification showed an accuracy of 94% and a kappa index of 96% with 25 predictor variables. Of these, the vegetation indices GCI, TVI, MCARI and ERGBVE were the most important and promising to be used in the evaluation of the nutritional status of the crop and in the identification of diseases and disturbances. The use of these indices and the other variables associated with the built model, enables greater phytosanitary control of the plantations. Sentinel satellite images available free of charge and every 5 days can be used for AF mapping. The methodology developed and applied was satisfactory in separating plants with disturbance caused by AF and healthy plants.

Keywords: Machine learning. Palm oil. Precision agriculture.

## SUMÁRIO

Introdução Geral .....	11
Referências Bibliográficas.....	12
<b>O USO DO SENSORIAMENTO REMOTO E DE TÉCNICAS DE APRENDIZADO DE MÁQUINA PARA O MAPEAMENTO DO AMARELECIMENTO FATAL NA PALMA DE ÓLEO .....</b>	<b>13</b>
1.0 Introdução.....	13
2.0 Área de estudo .....	16
2.1 Base de dados .....	16
2.2.1 Metodologia para o mapeamento do AF .....	17
2.2.2 Acurácia da classificação .....	20
3.0 Resultados e Discussão.....	21
3.1 Seleção de variáveis e desempenho da classificação.....	21
3.2 Áreas com Amarelecimento Fatal: resultados do mapeamento.....	25
4.0 Conclusões.....	27
Referências Bibliográficas.....	28
Anexo 1: Índices de vegetação utilizados no estudo .....	31
Anexo 2: Covariáveis utilizadas externas aos índices de vegetação. ....	35
Anexo 3: Variáveis utilizadas do banco de dados de fitossanidade. ....	39
<b>MAPEAMENTO DO AMARELECIMENTO FATAL NA PALMA DE ÓLEO UTILIZANDO IMAGENS DE ALTÍSSIMA RESOLUÇÃO .....</b>	<b>40</b>
1.0 Introdução.....	40
2.0 Área de estudo .....	42
3.0 Materiais .....	43
4.0 Metodologia.....	44
5.0 Resultados e Discussão.....	48
6.0 Conclusões.....	52
Referências Bibliográficas.....	53
Conclusão Geral .....	57

## **Introdução Geral**

A agricultura brasileira passa por um aperfeiçoamento de técnicas, de manejo e principalmente no controle de pragas e doenças a partir do estabelecimento da agricultura de precisão, que surge como suporte à agricultura tradicional, aliando máxima produtividade à conservação ambiental. A agricultura de precisão enfoca as características e particularidades de todo o processo produtivo, Tschiedel, (2002) a define como sendo uma ferramenta de gestão ou um sistema de gerenciamento da produção que leva em consideração a variabilidade espacial da produtividade e dos fatores de produção, atrelado ao sistema de navegação por satélite (VIERA-TORRES; SINDE-GONZ; GIL-DOCAMPO, 2020).

O sensoriamento remoto é uma ciência ampla, que apresenta ferramentas importantes no mapeamento do uso e cobertura da terra. A definição mais clássica para sensoriamento remoto o descreve como a ciência que trabalha com a obtenção de imagens da superfície terrestre por meio da detecção e medição quantitativa das respostas das interações da energia eletromagnética com os materiais terrestres (DLAMINI et al., 2019; PONZONI, 2001). O sensoriamento remoto evoluiu rapidamente, sendo possível a aquisição de imagens com resoluções espaciais cada vez maiores (DLAMINI et al., 2019). O uso de imagens para soluções na agricultura é cada vez maior, assim como o de pesquisas que servem de suporte para tais ações.

Diversos trabalhos utilizam produtos do sensoriamento remoto para estudos com palma de óleo, principalmente nas investigações quanto a idade e crescimento de biomassa. Trabalhos que exploram o mapeamento de doenças e distúrbios não são numerosos, principalmente trabalhos que correlacionam a ocorrência do Amarelecimento Fatal (AF), com fatores bióticos e abióticos interligados. O AF é considerado como o problema fitossanitário mais sério do dendezeiro no Brasil (BOARI, 2010), apresentando causa e tratamentos desconhecidos. Este trabalho busca assim, aprofundar os estudos sobre o AF com o uso de produtos do sensoriamento remoto, estabelecendo os fatores que apresentam maior correlação com a ocorrência deste distúrbio.

Esta tese é composta por uma breve parte introdutória com informações do solo e da cultura e de dois artigos. O primeiro intitulado de “O uso do Sensoriamento Remoto para o Mapeamento do Amarelecimento Fatal na Palma de Óleo” e o segundo intitulado de

“Mapeamento do amarelecimento fatal na palma de óleo utilizando imagens de alta resolução e tecnologias geoespaciais”.

### **Referências Bibliográficas**

- BOARI, A. DE J. O desafio das pesquisas com a etiologia do Amarelecimento Fatal (AF) da cultura da palma de óleo. [s.l: s.n.].
- DLAMINI, S. N. et al. Review of remotely sensed data products for disease mapping and epidemiology. *Remote Sensing Applications: Society and Environment*, v. 14, n. January, p. 108–118, 2019.
- PONZONI, F. Comportamento espectral da vegetação. *Sensoriamento remoto: reflectância de alvos naturais*, v. 1, n. Brasília: UNB, p. 157–199, 2001.
- TSCHIEDEL, M.; FERREIRA, M. F. Introdução À Agricultura De Precisão: Conceitos E Vantagens. *Ciência Rural*, v. 32, n. 1, p. 159–163, 2002.
- VIERA-TORRES, M.; SINDE-GONZ, I.; GIL-DOCAMPO, M. Generating the Baseline in the Early Detection of Bud Rot and Red Ring Disease in Oil Palms by Geospatial Technologies. 2020.

## O USO DO SENSORIAMENTO REMOTO E DE TÉCNICAS DE APRENDIZADO DE MÁQUINA PARA O MAPEAMENTO DO AMARELECIMENTO FATAL NA PALMA DE ÓLEO

**Abstract:** Oil palm culture has been affected by the bud rot disorder (BD). BD is a disorder with no known cause and stands out for the gradual yellowing of the leaves, especially of the younger leaves, evolving to necrosis. The crown architecture changes, thus modifying the spectral response of plants affected by BD. In this work, the objective was to map how PA occurrence areas in commercial oil palm plants, using remote sensing and physical environment data through remote sensing and machine learning techniques. The database consisted of 826 variables: Sentinel 2A image, vegetation indices, terrain attributes and plant health dataset. The classification methodology for OBIA and the random forest algorithm. Our results indicate that an integration of remote sensing data with machine learning is promising for the development of disease and disorder mapping research, such as the case of BD. The GCI, TVI and MCARI indices are promising and of great importance without mapping BD in oil palm culture.

Keywords: oil palm, machine learning, remote sensing, bud rot, diseases mapping.

### 1.0 Introdução

A palma de óleo do gênero *Elaeis spp.* de nome científico *Elaeis guineenses* Jacq é a oleaginosa de maior importância mundial, com taxa de extração de óleo maior que a de outras oleaginosas, como a soja, e apresenta produtividade média de 5 ton/ha/ano (RAMALHO FILHO et al., 2010). É originária da África Ocidental e é uma das culturas que mais se expandiram nos trópicos, presente em diversos países como Malásia, Indonésia, Colômbia, Equador, República Democrática do Congo, Brasil, dentre outros. Destes países, Malásia e Indonésia lideram o ranking mundial, sendo responsáveis por 85% de todo o óleo produzido no mundo (MPOC, 2012).

No Brasil, temos plantios comerciais nos estados de Roraima, Bahia e Pará, com aproximadamente 240 mil hectares cultivados. Deste total, 211,2 mil hectares estão concentrados na região nordeste do estado do Pará, nos municípios de Tailândia, Acará e Tomé-Açu. A demanda por óleo de palma nos últimos anos vem aumentando, e além do setor alimentício e indústrias de cosméticos, inicia-se a demanda por óleo para a produção de

biodiesel. O zoneamento agrícola elaborado para a cultura da palma no estado do Pará (RAMALHO FILHO et al., 2010) esboça que existem 2 milhões de hectares de áreas antropizadas e degradadas que poderiam ser destinadas ao plantio de palma. O zoneamento também aponta outro ponto favorável à expansão no estado: a presença de um polo consolidado da cadeia produtiva da palma de óleo.

Apesar deste cenário promissor, as perspectivas da dendeicultura no país não têm sido animadoras, visto que, como qualquer cultura a palma está sujeita a surtos populacionais de pragas e doenças, que se tornam fatores limitantes à produção. Dentre as pragas de maior importância, pode-se citar lagartas desfolhadoras e coleópteras bloqueadoras, e dentre as doenças e distúrbios, o amarelecimento fatal e o anel vermelho são sem dúvidas os maiores problemas fitossanitários da cultura. O amarelecimento fatal ou AF como é conhecido popularmente, apresenta uma série de sintomas característicos, começando com o amarelecimento progressivo das folhas mais jovens (BOARI, 2010; LAING, 2009; ROT; WILT; RING, 2010), seguido de clorose e necrose das mesmas. Sua etiologia é desconhecida e por este fato, dizemos ser um distúrbio fisiológico. Sua primeira ocorrência data de 1929, e caracteriza-se por ser um desafio para muitos pesquisadores. O AF é responsável por acentuadas quedas de produtividade em diversos plantios no estado do Pará, e todos os anos centenas de plantas são dizimadas em razão de sua ocorrência. Em alguns casos, relacionados com incidências menos graves de AF, as plantas afetadas podem se recuperar e retornar aos níveis de produtividade obtidos anteriormente ao início da doença, no entanto, o mais comum é o avanço do distúrbio até a morte da palmeira (LAING, 2012).

As aplicações do sensoriamento remoto em estudos nos cultivos de palma de óleo concentram-se em mapeamentos do uso e cobertura da terra, estudos com imagens de satélite e de radar voltados principalmente a fisiologia. Torres (2009), menciona que a maioria dos trabalhos desenvolvidos até o presente momento dão ênfase a estimativa do crescimento da biomassa e ao mapeamento da idade de maturação das plantas (MARZUKHI; NOOR; SAUTI, 2018), assim como a contagem de plantas (CHONG et al., 2017). Trabalhos que abordam o mapeamento de doenças, como ganoderma e AF (referenciados como Bud rot ou pudridón del congollo) são numerosos (BENÍTEZ; GARCÍA, 2014; SUNDRAM; INTAN-NUR, 2017; VIERA-TORRES; SINDE-GONZ; GIL-DOCAMPO, 2020), no entanto, em sua maioria utilizam imagens de baixa resolução, dados dos satélites landsat por exemplo. Também são encontrados trabalhos que utilizaram imagens hiperespectrais ou imagens de baixa altitude adquiridas com veículos não tripulados (AZMAN; MOHD NOOR, 2002; SANTOS;

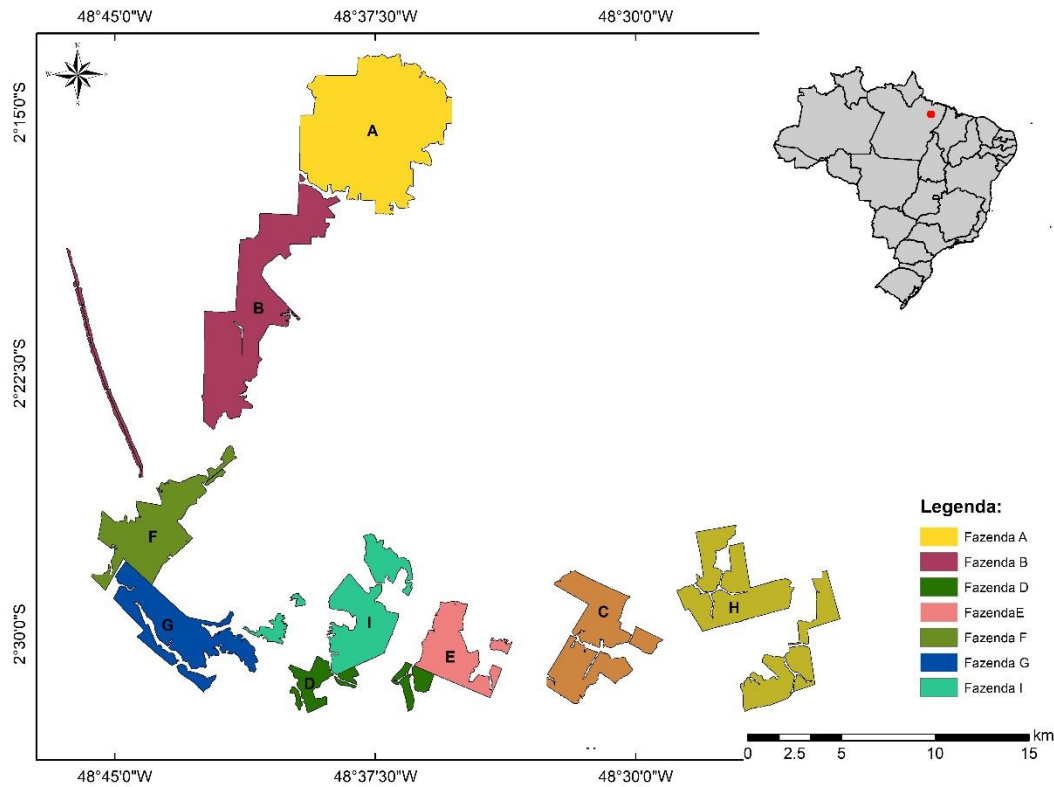
MESSINA, 2008; SHAFRI, 2014; TORRES, 2009; VIERA-TORRES; SINDE-GONZ; GIL-DOCAMPO, 2020). Grandes plantações de palma, requerem metodologias de mapeamento de doenças, capazes de obter resultados precisos em curto espaço de tempo, visto que, quanto antes os sintomas forem identificados, mais rápido de medidas fitossanitárias serem tomadas, para conter o avanço das doenças. Neste sentido, pesquisas que utilizem imagens de média resolução de livre acesso e com baixo tempo de revisita fazem-se importantes. As imagens do satélite Sentinel são imagens gratuitas e que apresentam tempo de revisita de 5 dias na linha do Equador, faixa esta onde estão concentradas grande parte dos cultivos de palma de óleo no mundo. As imagens do sensor MSI Sentinel apresentam resolução espacial de 10 metros, sendo consideradas de média resolução espacial, e recobrem partes importantes do espectro eletromagnético para o mapeamento da vegetação. Song et al. (2017), utilizou as imagens do satélite Sentinel 2 para o mapeamento da podridão radicular em cultivos de algodão, apresentando precisão de mapeamento de 91,2%. No Brasil, Almiro et al. (2020) é um dos primeiros a explorar uso destas imagens no mapeamento do anel vermelho, em cultivos de palma de óleo, aplicando índices de vegetação. Os autores concluíram que o uso de índices de vegetação é satisfatório e faz-se importante para o mapeamento das doenças estudadas nos cultivos, conduzindo à ideia de que as imagens do satélite Sentinel 2A, são promissoras para trabalhos na cultura da palma de óleo.

O comportamento espectral da vegetação no espectro eletromagnético está condicionado a diversos fatores como a estrutura da folha, teor de nutrientes, conteúdo de água, dentre outros. À medida que as folhas envelhecem, diminuem a atividade fotossintética nos comprimentos de onda do visível, principalmente no azul e no vermelho (PONZONI, 2001). O AF caracteriza-se pelo amarelecimento gradual das folhas das plantas, principalmente das mais jovens, que evolui para a necrose. A perda de pigmento, pode ocorrer por três razões centrais: envelhecimento natural das folhas, carência nutricional ou estresse hídrico. Além da modificação do comportamento espectral, ocorre também a alteração da arquitetura de copa das plantas, visto que, a reemissão de novas folhas fica comprometida pela lesão causada no ápice.

Neste trabalho, os autores tem por objetivo, mapear as áreas de ocorrências do AF em plantios comerciais de palma de óleo no noroeste do estado do Pará, utilizando dados de sensoriamento remoto e do meio físico por meio de técnicas de aprendizado de máquina. Adicionalmente pretendemos identificar o comportamento espectral de plantas saudáveis e de plantas com AF e determinar quais variáveis espectrais e do meio físico são mais importantes para o mapeamento deste distúrbio.

## 2.0 Área de estudo

A área de estudo são plantios comerciais localizados nos municípios de Tailândia, Moju e Acará no estado do Pará. A área é de 17.272,45 mil hectares, distribuídos em 10 fazendas (Figura 1).



**Figura 1:** Localização da área de estudo.

### 2.1 Base de dados

As imagens de satélite utilizadas foram as do satélite Sentinel 2, adquiridas no repositório digital do projeto Copernicus (<https://scihub.copernicus>). A imagem é do mês de setembro do ano de 2020, e está livre de nuvens. Todas as bandas da imagem adquiridas pelo sensor MSI – Sentinel foram obtidas com o nível de processamento 1C, ortorretificadas e com valores de reflectância do topo da atmosfera (TOA). Com o uso do software SNAP- Sen2cor, foi possível realizar a correção dos valores de reflectância da atmosfera para reflectância de superfície, melhorando assim o nível do processamento das imagens para 2A. Com este processo, pixels defeituosos com valores espúrios, e interferências atmosféricas são eliminadas. Foram utilizadas as bandas de 10 e de 20 metros de resolução espacial. Após uma vasta revisão de literatura, foram calculados 30 índices de vegetação referenciados como potenciais à agricultura

de precisão, como: NDVI, VARI, GCI, ERGBVE, ARVI, VEG. Os índices e suas respectivas siglas são apresentados no anexo 1.

O modelo digital de elevação NASADEM de 30 metros de resolução espacial e 47 índices morfométricos derivados dele (anexo 2), a partir de processamento realizado no software RSAGA (BRENNING; BANGS; BECKER, 2007) foram inseridos no banco de covariáveis. Posteriormente todos os arquivos gerados foram re-amostrados para 10 metros e alinhados às bandas da imagem Sentinel. Também foram adicionados em formato raster, o mapa de solos da área e o mapa geológico (VASQUEZ; SOUSA; CARVALHO, 2008)

Para compor o banco de dados, duas bases de dados com informações sobre os cultivos foram inseridas ao banco de dados. O primeiro, é um banco de dados com informações gerais, deste foram utilizados os tipos de material genético cultivados nas parcelas, ano de plantio e número de plantas por parcela. O segundo banco de dados é de fitossanidade, com informações sobre anomalias e áreas que sofreram ataques de lagartas desfolhadoras. Os dados utilizados e a origem do banco de dados são apresentados no anexo 3. No total o banco de dados foi composto por 118 rasters.

### **2.2.1 Metodologia para o mapeamento do AF**

A abordagem metodológica de classificação utilizada foi a Orientada a Objetos (OBIA), que é uma técnica de sensoriamento remoto desenvolvida para analisar imagens de alta resolução espacial, que utiliza características espectrais, espaciais e topológicas (LANG, 2008). O OBIA baseia-se em segmentar a imagem em diversos polígonos, que passam a ser denominados de objetos. Caracterizam-se por grupo homogêneos em padrões espectrais e espaciais. Durante a segmentação, são consideradas a heterogeneidade espacial (forma), a cor (espectral) e a diferenciação do entorno, a partir do método de crescimento de regiões. São estabelecidos valores para estes critérios, definidos pelo foto-intérprete. O processo é feito até alcançar o nível de segmentação que atenda às exigências do foto-intérprete.

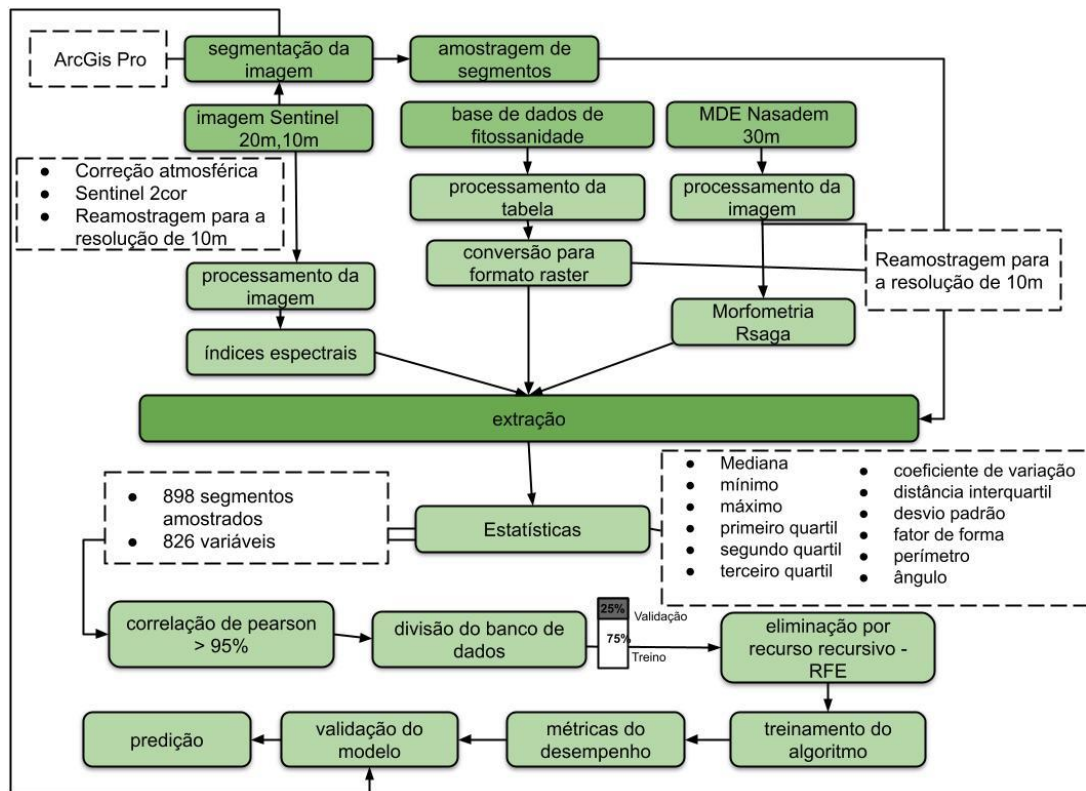
A classificação orientada a objetos (OBIA) foi realizada com as bandas de resolução espacial de 10 metros (B2, B3, B4 e B8) com adição do índice de vegetação NDVI no software ArcGis Pro (ESRI DATA, 2019). Os padrões para a segmentação da imagem foram 15.2 de detalhamento espacial, 16.0 de detalhamento espectral e 400 m<sup>2</sup> como área mínima para a criação de segmentos. As classes de uso e cobertura da terra mapeadas estão descritas na tabela 1.

**Tabela 1:** Classes de uso da terra classificadas.

<b>Código da classe</b>	<b>Classe de uso e cobertura</b>	<b>Descrição</b>
1	Palma saudável	Parcelas homogêneas, e com 0% de AF no banco de dados de fitossanidade.
2	Palma com AF	Parcelas heterogêneas, e com presença de AF pelo banco de dados de fitossanidade.
3	Solo exposto	Estradas
4	Palma sem AF e com outro distúrbio	Palma sem AF e com infestação de lagartas desfolhadoras na data da imagem.

Foi realizado um levantamento de campo, com o uso de um receptor do sinal Global Navigation Satellite System (GNSS), para o georreferenciamento das plantas, nas classes descritas na tabela 1. Além desses pontos, foram realizados 5 voos utilizando um veículo aéreo não tripulado (VANT) em parcelas produtivas das fazendas A e B. As imagens de alta resolução foram utilizadas como auxílio para a coleta de amostras para as áreas com AF (classe 2). Essa etapa foi de grande importância para o entendimento do padrão espectral da doença. Também foram coletadas diretamente na imagem de satélite. No total 898 amostras foram coletadas, de modo a garantir uma distribuição homogênea por toda a área de estudo.

Para cada polígono amostrado foi feita a extração dos valores dos pixels nos 118 rasters que representam as variáveis explicativas. Os valores extraídos foram submetidos a estatísticas básicas para cada polígono. Foram elas: mínimo, mediana, máximo, primeiro e terceiros quartis, coeficiente de variação (CV), distância interquartílica e desvio padrão. Foram ainda calculadas variáveis morfométricas, representando a forma dos polígonos: área, perímetro, fator de forma, raio do menor círculo envolvente e ângulo de orientação. Toda as determinações anteriores foram feitas no software R (R CORE TEAM, 2019). Na figura 2 encontram-se estes parâmetros, assim como os demais procedimentos metodológicos desenvolvidos.



**Figura 2:** Esquema metodológico mostrando a sequência das etapas metodológicas para o mapeamento das áreas com AF.

O processo do cálculo das estatísticas e da forma dos polígonos gerou um banco de dados com 898 linhas representando as amostras coletadas e 826 colunas representando as variáveis explicativas. Foi feito o particionamento dos dados usando 75% dos dados para treinamento e 25% para teste do modelo.

Inicialmente, foi realizada a eliminação das variáveis altamente correlacionadas, utilizando a função findcorrelation do pacote caret (KUHN; JOHNSON, 2013). Esse procedimento visa eliminar variáveis altamente correlacionadas já que podem ser considerados redundantes em termos de informação. O limiar máximo de correlação permitido foi de 95%. O processo de remoção é feito par a par e quando um par de variáveis tem correlação acima do limiar é feito o cálculo da soma da correlação de cada uma delas com todas as demais variáveis existentes no banco de dados. A variável do par que apresentar maior somatório de correlação é considerada de maior redundância, e é eliminada.

A próxima etapa consiste na identificação do conjunto de variáveis mais importantes para o objetivo do estudo, chamadas de covariáveis potenciais por Gomes et al (2019). Para isso o algoritmo Recursive Feature Elimination do pacote Caret (Kuhn and Johnson, 2013) foi utilizado. O RFE é um algoritmo de seleção retroativa que elimina variáveis que apresentam pouca contribuição com o modelo com base na importância da mesma (Kuhn and Johnson, 2013). As variáveis selecionadas pelo RFE serão as de maior importância no mapeamento do AF, e é com este conjunto de variáveis que foi feito o ajuste do modelo, e a predição para toda a área de estudo e cálculo das métricas de desempenho do modelo.

Para a classificação das imagens foi utilizado o algoritmo *Random Forest* (RF) proposto por Breiman (2001). O RF é baseado na ideia de aprendizado em conjunto (ensemble learning – bagging) conhecido por sua capacidade em reduzir o ruído experimental e aumentar a acurácia da predição agregando diferentes predições (GOMES et al., 2019). Foi utilizado o pacote *randomforest* (LIAW; WIENER, 2002).

Em estudos de uso e cobertura da terra, o RF foi considerado estável e eficiente, envolvendo o ajuste de poucos hiperparâmetros para gerar resultados confiáveis (GISLASON; BENEDIKTSSON; SVEINSSON, 2006; LEBOURGEOIS et al., 2017) quando comparados a outros classificadores. Além disso em trabalhos que abordaram a podridão basal do caule na palma de óleo (SANTOSO; TANI; WANG, 2017), o RF foi considerado o melhor algoritmo em termos de predição geral, sendo sua aplicação no mapeamento de doenças e distúrbios, promissor. Amirruddin et al. (2020), considerou o RF como um algoritmo robusto para avaliação dos níveis de suficiência das clorofilas *a* e *b* na cultura da palma de óleo, com acurácia variando de 93% a 96%.

### 2.2.2 Acurácia da classificação

A avaliação da precisão da classificação foi realizada por meio do procedimento de validação externa, com 25% dos dados. Após esta etapa foi calculada a matriz de confusão utilizando a função existente no pacote Caret (KUHN; JOHNSON, 2013), que possibilitou o cálculo do índice kappa (COHEN, 1960) e da exatidão global, índices que refletem o nível de concordância real da classificação. O valor do índice Kappa varia de 0 (concordância nula) e 1 (concordância perfeita).

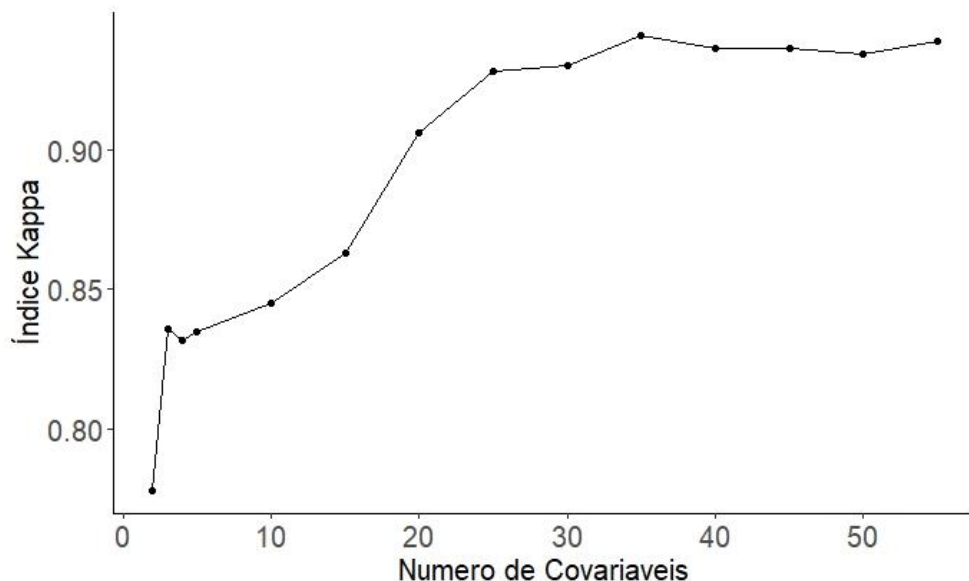
$$Kappa = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e}$$

$p_o$  é a taxa de aceitação e  $p_e$  a taxa hipotética de aceitação (CONGALTON, 1991).

### 3.0 Resultados e Discussão

#### 3.1 Seleção de variáveis e desempenho da classificação

Do conjunto de dados inicial, contendo 826 variáveis, 532 foram eliminadas por apresentarem alta correlação. Neste conjunto, dominam variáveis espectrais e do terreno. Das covariáveis de forma dos segmentos, apenas a área do menor círculo se manteve no conjunto de dados. Dessa forma, 294 variáveis foram submetidas ao RFE. Dessas, 35 foram indicadas como as mais importantes para o mapeamento das áreas com AF. O índice kappa sofre alterações significativas até o número de 25 variáveis, estabilizando quando 35 variáveis são selecionadas (figura 4). Considerando o número inicial de covariáveis isso corresponde a uma redução de 95,7 % permitindo a criação de um modelo mais parcimonioso e com demanda computacional muito menor do que seria exigido se fosse utilizado o conjunto total de variáveis.



**Figura 1:** Gráfico que representa o valor do índice Kappa versus o número de covariáveis selecionadas

A ordem de importância das variáveis utilizadas para a predição do mapa final é apresentada na tabela 2. Nenhuma variável do terreno foi selecionada, isto pode ser em razão da baixa variação da altitude na área de estudo. O relevo da área é plano a suave ondulado com interflúvios longos. Outra razão possível é a resolução espacial do MDE NASADEM, que não apresenta nível de detalhamento apropriado para as mudanças sutis que ocorrem na paisagem.

**Tabela 2:** Variáveis selecionadas pelo algoritmo RFE-RF como as variáveis mais importantes para o mapeamento de AF na cultura da palma de óleo. A frente do nome da variável, é identificada qual a estatística da variável foi mais importante.

Nome da Variável	Ordem de importância	Origem	Nome da Variável	Ordem de importância	Origem
GCI_max	1 <sup>a</sup>	E	Banda7_max	19 <sup>a</sup>	E
TVI_q4	2 <sup>a</sup>	E	Idade2021_Media	20 <sup>a</sup>	P
MCARI1_max	3 <sup>a</sup>	E	Idade2021_Min	21 <sup>a</sup>	P
ERGBVE_Min	4 <sup>a</sup>	E	Banda2_Iqr	22 <sup>a</sup>	E
RGI_Min	5 <sup>a</sup>	E	Idade2021_max	23 <sup>a</sup>	P
Banda8_max	6 <sup>a</sup>	E	Tvi_Iqr	24 <sup>a</sup>	E
RGVBI_max	7 <sup>a</sup>	E	Banda6_max	25 <sup>a</sup>	E
Mcari2_Min	8 <sup>a</sup>	E	Banda9_Min	26 <sup>a</sup>	E
Banda8_Min	9 <sup>a</sup>	E	N_Plant_Min	27 <sup>a</sup>	P
GCI_Min	10 <sup>a</sup>	E	N_Plant_Min	28 <sup>a</sup>	P
Banda3_Min	11 <sup>a</sup>	E	Banda5_Iqr	29 <sup>a</sup>	E
Banda5_Min	12 <sup>a</sup>	E	Banda9_max	30 <sup>a</sup>	E
Banda11_Min	13 <sup>a</sup>	E	Sipi_Iqr	31 <sup>a</sup>	E
Banda8a_Min	14 <sup>a</sup>	E	N_Plant_Q05	32 <sup>a</sup>	P
Banda8a_max	15 <sup>a</sup>	E	ARI1_Min	33 <sup>a</sup>	E
Banda11_max	16 <sup>a</sup>	E	N_Plant_Q25	34 <sup>a</sup>	P
Banda3_Iqr	17 <sup>a</sup>	E	ARI1_Media	35 <sup>a</sup>	P
Banda7_Min	18 <sup>a</sup>	E			

Onde: Min = mínimo, media = média, Iqr = amplitude interquartil, Q05 = valor inferior do limite. E = espectral, P = plantio.

Deste banco de dados, 27 variáveis são espectrais e oito são referentes à número de plantas por parcela e idade. Nove bandas do satélite MSI/Sentinel foram selecionadas, com seus valores mínimos e máximos dominando o ranking de importância. Nove índices de vegetação foram selecionados e da mesma maneira que as bandas do Sentinel, os valores mínimos e máximos dominaram a importância, com a adição da média e da amplitude interquartil. Das bandas selecionadas, duas são do comprimento de onda do visível (bandas 2 e 3– 490 a 560 nm), quatro do Red Edge (bandas 5, 6, 7 e 8A - 705 a 865 nm), duas do infravermelho de ondas

curtas (banda 9 – 940 nm e banda 11 – SWIR 1610 nm). Dos índices de vegetação selecionados, três utilizam apenas bandas do visível (ERGBVE, RGI e RGI), os demais utilizam a banda referente ao comprimento de onda do vermelho (banda 4) em associação com bandas do infravermelho em seus diferentes comprimentos. Chemura et al. (2017) conclui em sua pesquisa de mapeamento da ferrugem em talhões com café que a utilização de índices de vegetação associados a bandas de satélite do Red Edge apresenta melhor desempenho na discriminação de doenças na vegetação e que, imagens Sentinel 2A são potenciais para a avaliação do status nutricional de culturas, melhorando o manejo de terras agrícolas e a gestão do meio ambiente no controle de pragas e doenças.

No mapeamento do AF, a variável mais importante foi o índice de vegetação Green Chlorophyll Index (GCI) proposto por Gitelson et al. (2003). Os maiores valores de GCI foram indicados como mais importantes para o mapear AF, e os valores mínimos de GCI foram a 10ª variável mais importante. O GCI é um índice sensível ao conteúdo de clorofila na folha. Palmeiras com atividade fotossintética maior, com conteúdo de clorofila maior, distinguem-se das palmeiras que apresentam algum distúrbio, no caso às com presença de AF. Lichtenthaler (1998) afirma que o conteúdo de clorofila foliar e a concentração de nitrogênio são bioindicadores importantes da fisiologia vegetal, atuando ativamente no processo de fotossíntese. A diminuição do conteúdo de clorofila na folha é característica do AF, visto que, as folhas amarelam gradualmente até alcançar a necrose, comprometendo a atividade fotossintética.

Após o GCI, temos os índices Transformed Vegetation Index (TVI), Modified Chlorophyll Absorption in Reflectance Index (MCARI) e Enhanced Red-Green-Blue Vegetation Index (ERGBVE), juntamente com a banda 8, dominando até a décima posição no ranking das variáveis importantes. Estes índices, assim como o GCI, relacionam-se com o conteúdo de clorofila. As bandas utilizadas para o cálculo destes, são bandas no comprimento do visível associadas a bandas do infravermelho médio e infra vermelho próximo, regiões do espectro eletromagnético, capazes de fornecer informações sobre o estresse de nutrientes (RENDANA et al., 2015). Ela afirma que imagens, com resolução espacial igual ou menor a 10 metros são capazes de identificar estresses até mesmo na colheita, assim como inferir sobre a gravidade dos mesmos. Isso em razão das modificações da reflectância que ocorre na cultura, identificadas primeiramente no infravermelho médio e infravermelho próximo. O estresse, afeta a estrutura interna da folha, que podem ser diagnosticadas nestes comprimentos de onda, antes mesmo dos sintomas tornarem-se visíveis. A seleção destas variáveis corrobora com os

trabalhos de Srestasathiern and Rakwatin (2014) e Viera-torres et al. (2020) que utilizaram o VARI e o GCI em seus trabalhos no mapeamento dos sintomas do ganoderma na palma de óleo e para a contagem de plantas em plantios da Malásia, e ambos apresentarem satisfatórios para o mapeamento do ganoderma.

Além das variáveis espectrais, a idade mínima dos plantios (1 ano) e a idade média deles (20 anos) também foram selecionadas dentre as variáveis mais importantes, separando assim, plantios recentes de plantios com desenvolvimento fenológico mais avançado. Existem parcelas na fazenda A em estágio inicial de desenvolvimento, mas também existem parcelas com plantios chegando a 30 anos. Tais características justificam a seleção das variáveis idade mínima e idade média. Além do mais, plantios com idade próxima a 20 anos apresentam maior incidência de AF. A variável do número de plantas por parcela, refere-se à densidade de plantas. Os valores mínimos, primeiro quartil e média foram indicados como importantes, a partir da 27<sup>a</sup> variável. O AF é responsável por dizimar centenas de plantas ano a ano, o que altera a densidade de plantas em cada parcela, o que justifica a seleção destas variáveis.

Após o ajuste do modelo com as variáveis selecionadas pelo RFE, foram calculadas a matriz de confusão e as métricas de desempenho de ajuste do modelo (Tabela 3). Seguindo o procedimento de cálculo da acurácia definido por Congalton (1991), o mapeamento apresenta 96.14% de precisão geral, de acurácia e índice Kappa de 0.94. A matriz de confusão da classificação, mostra que as maiores confusões ocorreram entre as áreas com e sem AF. Esta confusão pode ser em razão dos polígonos coletados de maneira incorreta, ou da baixa alteração fenológica que ocorre nas plantas com os primeiros sintomas do AF, que não altera a arquitetura da copa das plantas, no entanto, já compromete a atividade fotossintética.

A matriz de confusão (tabela 3) da classificação mostra que o classificador conseguiu separar de maneira satisfatória a classe solo exposta, das demais classes. Algumas confusões ocorreram entre as áreas com AF das áreas de palma sem AF e com outro distúrbio. Na cultura da palma além do AF, existem outras anomalias como anel vermelho, podridão basal ou até mesmo a ação de uma descarga elétrica. Além disso, na data que a imagem foi adquirida pelo satélite Sentinel 2, existia um grande ataque de lagartas desfolhadoras em alguns departamentos. Este ataque expôs o solo, justificando assim a confusão entres estas classes.

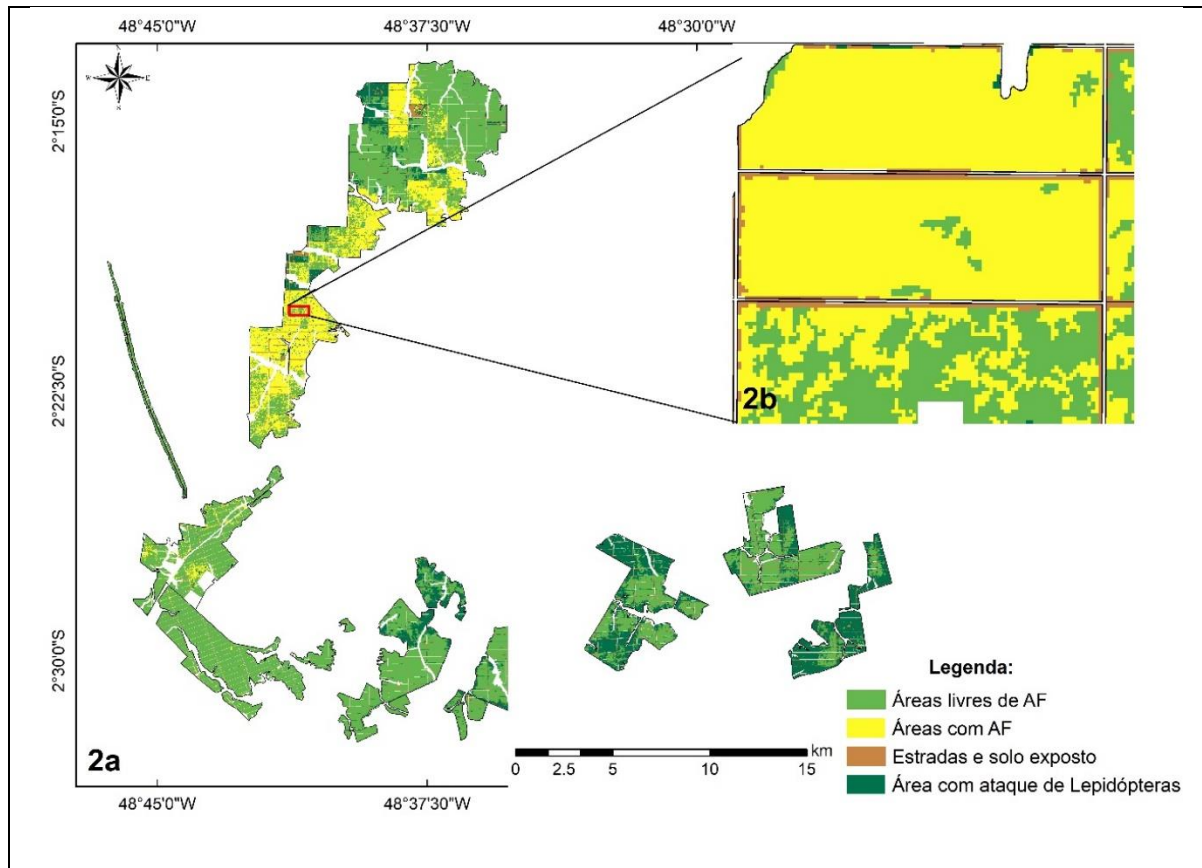
**Tabela 3:** Matriz de confusão.

<b>Matriz de Confusão</b>				
	Áreas livres de AF	Áreas com AF	Solo exposto	Palma sem AF e com outro distúrbio
Áreas livres de AF	<b>295</b>	9	0	1
Áreas com AF	10	<b>137</b>	0	0
Solo exposto	1	0	<b>170</b>	0
Palma sem AF e com outro distúrbio	2	0	0	<b>51</b>
<b>Valor do índice Kappa</b>	0.94			
<b>Valor da Acurácia</b>	96.14%			

### 3.2 Áreas com Amarelecimento Fatal: resultados do mapeamento

A distribuição das plantas com amarelecimento fatal é uma informação importante para o manejo dos plantios no campo, sendo ferramenta para a gestão dos plantios. O conhecimento das áreas mais propensas ao AF, vai de encontro com a tomada de decisões para mitigar o surgimento de novos casos e até mesmo no desenvolvimento de pesquisas nesta linha. A complementação da nutrição dos plantios é um exemplo de ação que pode ser tomada na tentativa de conter o surgimento de novos casos. O conhecimento desta informação, pode servir de base para o planejamento de novas áreas e de replantios, visto que, as parcelas com alto grau de infestação são conhecidas e quantificadas.

Dentro da área de estudo, o amarelecimento fatal não é uniforme. O AF está presente em 15,2% das fazendas, concentrando-se nas fazendas A, B e F. Nas fazendas E, C e H até o momento não foi observado nenhum caso de AF no banco de dados de fitossanidade.



**Figura 2:** Mapeamento das áreas de ocorrência do AF e do uso e cobertura da terra.

Nas fazendas, a média de plantas infectadas com AF é de 428 plantas por parcela. Os casos concentram-se em 3 fazendas. A fazenda B é a de maior número de parcelas com AF. Das 139, 101 estão com AF. Nesta fazenda, a média é de 46% de AF por parcela. Existem parcelas tomadas pelo AF, ao lado de parcelas que ainda não apresentam um número elevado de plantas com o distúrbio (figura 2b). A parcela central na figura 2b apresenta 90,6% de plantas com AF, o que corresponde a 4819 plantas. Nesta parcela, em 2016 a produtividade foi de 32,5% ton/ha/ano, e neste último ano a produtividade foi de 6,4 ton/ha/ano. Esses dados corroboram com as pesquisas de Boari (2010), Hoe and Shukor (2019) e Rot et al. (2010) que afirmam que o AF influencia na produtividade da cultura da palma de óleo.

Os primeiros casos de AF foram identificados na fazenda A, e dizimou muitas parcelas na década passada, levando aos gestores da empresa a realizar o replantio com o híbrido interespecífico, resultado do cruzamento das espécies *Elaeis guinenses* e *Elaeis oleífera*. Boari 2010, aponta que esta variedade é resistente ao AF. Esta variedade apresenta volume de raiz muito maior, quando comparado à Deli x Lamé. O plantio de híbridos é apontado como uma opção para as regiões com alta incidência de AF, comprovados em estudos realizados em

plantios na Colômbia e Equador (LAING, 2014; ROT; WILT; RING, 2010; TIEMANN et al., 2018).

Nas fazendas C e H na data da obtenção da imagem, outubro de 2020, ocorreu um ataque de lepidópteras que chegou a atingir alguns talhões das fazendas E, I e X como mostrado na figura 2. Durante o treinamento do modelo de classificação, essas áreas se confundiram com AF, em razão de ambos os fenômenos alterarem a arquitetura de copa das plantas, um desfolhando as plantas e no outro, com a perda de folhas e amarelecimento das mesmas. A separação entre essas classes, só foi possível após a inserção do banco de dados de fitossanidade que continha o número de mariposas registrado por planta e o número de plantas com presença de mariposas. Durante o processo de classificação é de suma importância que além das informações espectrais, sejam considerados fatores bióticos da cultura ou do uso da terra mapeado, e que esses fatores sejam incluídos no treinamento do modelo na forma de covariáveis. A integração do sensoriamento com machine learning tem como principal vantagem a inserção de arquivos de formatos e fontes dos mais diversos formatos.

As fazendas F e G, segundo o banco de dados de fitossanidade, apresentaram plantas com AF nos últimos cinco anos. São plantios com idade entre 17 e 20 anos, que alcançaram o patamar produtivo, e que em razão do AF tem sua produtividade afetada. Em ambos os departamentos temos Latossolos Amarelos distróficos argissólicos em associação com Plintossolos Pétricos, o que leva a possibilidade da existência da relação entre AF e classe de solos. Os Plintossolos apresentam mosqueados vermelhos em forma de plintitas, constituídos principalmente por ferro e alumínio, presentes na camada de 15 cm ou mais de espessura, que endurecem quando expostos a ciclos de umedecimento e secagem (Resende et al, 2014), podendo constituir como impedimento a percolação da água e à penetração das raízes, o que pode representar como fator inicial dos sintomas do AF nas plantas.

#### **4.0 Conclusões**

Nossos resultados indicam que a integração de dados de sensoriamento remoto com *machine learning* são promissores para o desenvolvimento de pesquisas para o mapeamento do AF, aumentando a janela de ações de manejo e de controle dos plantios.

O sensoriamento remoto apresenta muito potencial para ser utilizado a agricultura de precisão e a integração dos dados de fitossanidade, geoespaciais, machine learning e ciência computacional são promissores ao estudo de distúrbios e doenças.

Os índices de vegetação GCI, TVI e MCARI mostram-se promissores em estudos do mapeamento do AF na cultura da palma.

A metodologia desenvolvida e aplicada foi satisfatória na separação entre plantas com distúrbio provocado por AF e plantas saudáveis.

Imagens do satélite Sentinel disponíveis de forma gratuita e com periodicidade de 5 dias podem ser utilizadas para o mapeamento do AF.

### **Referências Bibliográficas**

- ALMIRO, J. et al. Detection of Diseases in Oil Palm Plantations in the Brazilian Amazon through Orbital Image. v. 12, n. 4, p. 138–147, 2020.
- AZMAN, I.; MOHD NOOR, M. The Optimal Age of Oil Palm Replanting. Oil Palm Industry Economic Journal, v. 2, n. 1, p. 11–18, 2002.
- BOARI, A. DE J. O desafio das pesquisas com a etiologia do Amarelecimento Fatal (AF) da cultura da palma de óleo.
- BREIMAN, L. Random forests. Machine Learning, v. 45, n. 1, p. 5–32, 2001.
- BRENNING, A.; BANGS, D.; BECKER, M. RSAGA: SAGA geoprocessing and terrain analysis in R. R package version 0.9, 2007.
- CHEMURA, A.; MUTANGA, O.; DUBE, • TIMOTHY. Separability of coffee leaf rust infection levels with machine learning methods at Sentinel-2 MSI spectral resolutions. v. 18, p. 859–881, 2017.
- CHONG, K. L. et al. A review of remote sensing applications for oil palm studies. Geo-Spatial Information Science, v. 20, n. 2, p. 184–200, 2017.
- COHEN, J. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, v. 20, n. 1, p. 37–46, 1960.
- CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, v. 37, n. 1, p. 35–46, 1991.
- DOUGLAS LAING. Informações agronômicas. n. 19, p. 1–17, 2012.
- ESRI DATA. Arcmap Redlands, California, United States, 2019. Disponível em:

- <<https://www.esri.com/pt-br/home>>
- GISLASON, P. O.; BENEDIKTSSON, J. A.; SVEINSSON, J. R. Random forests for land cover classification. *Pattern Recognition Letters*, v. 27, n. 4, p. 294–300, 2006.
- GITELSON, A. A.; † Y. G.; MERZLYAK, M. N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, v. 160, n. 3, p. 271–282, 2003.
- GOMES, L. C. et al. Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma*, v. 340, 2019.
- KUHN, M.; JOHNSON, K. *Applied Predictive Modeling*.
- LAING, D. La Causa de Pudrición de Cogollo (PC) en Palma de Aceite: Hipótesis Abiótica - Edáfica. *Presentación CIAT*, p. 113, 2009.
- LAING, D. Causa de lapudricion de cogollo en la palma de aceite papel del calcio en una hipótesis abiótica-edáfica. *Informaciones Agronómicas*, v. 3, n. June, p. 27, 2014.
- LANG, S. Object-based image analysis for remote sensing applications: modeling reality – dealing with complexity. In: *Object-based image analysis for remote sensing applications*. [s.l: s.n.]. p. 3–27.
- LEBOURGEOIS, V. et al. A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM) *Remote Sensing*, 2017.
- LIAW, A; WIENER, M. Classification and Regression by randomForest. *R news*, v. 2, n. December, p. 18–22, 2002.
- MARZUKHI, F.; NOOR, S.; SAUTI, S. Application of remote sensing in the investigation of maturity age of palm oil trees In Pasir Puteh, Kelantan. *Jurnal Intelek*, v. 13, n. 1, p. 44–57, 2018.
- MPOC. MALASIAN PALM OIL COUNCIL (MPOC). Disponível em: <<http://mpoc.org.my/>>. Acesso em: 11 abr. 2021.
- PONZONI, F. Comportamento espectral da vegetação. *Sensoriamento remoto: reflectância de alvos naturais*, v. 1, n. Brasília: UNB, p. 157–199, 2001.
- RAMALHO FILHO, A. et al. Zoneamento Agroecológico, Produção e Manejo para a Cultura da Palma de Óleo na Amazônia. Rio de Janeiro: [s.n.]. v. 1
- RENDANA, M. et al. A Review of Methods for Detecting Nutrient Stress of Oil Palm in Malaysia. *J. Appl. Environ. Biol. Sci. Journal of Applied Environmental and Biological Sciences*, v. 5, n. 6, p. 60–64, 2015.

- ROT, B.; WILT, S.; RING, R. Pudrición del cogollo, Marchitez sorpresiva, Anillo rojo y Marchitez letal en la palma de aceite en América Bud Rot, Sudden Wilt, Red Ring and Lethal Wilt of Introduction. v. 31, n. 1, p. 43–53, 2010.
- SANTOS, C.; MESSINA, J. P. Multi-Sensor Data Fusion for Modeling African Palm in the Ecuadorian Amazon. n. June, p. 711–723, 2008.
- SANTOSO, H.; TANI, H.; WANG, X. Random Forest classification model of basal stem rot disease caused by *Ganoderma boninense* in oil palm plantations. *International Journal of Remote Sensing*, v. 38, n. 16, p. 4683–4699, 2017.
- SHAFRI, H. Hyperspectral Imagery for Mapping Disease Infection in Oil Palm Plantation Using Vegetation Indices and Red Edge Techniques. n. June 2009, 2014.
- SONG, X. et al. Evaluation of Sentinel-2A Satellite Imagery for Mapping Cotton Root Rot. *Remote Sensing*, v. 9, p. 906, 2017.
- SRESTASATHIERN, P.; RAKWATIN, P. Oil palm tree detection with high resolution multi-spectral satellite imagery. *Remote Sensing*, v. 6, n. 10, p. 9749–9774, 2014.
- TIEMANN, T. T. et al. Feeding the Palm: A Review of Oil Palm Nutrition. *Advances in Agronomy*, v. 152, p. 149–243, 2018.
- TORRES, L. A. Caracterización de firma espectral a partir de sensores remotos para el manejo de sanidad vegetal en el cultivo de palma de aceite Characterization of the Spectral Signature from Remote Sensors for the Management of Plant Health in the Oil Palm Crop Intro. v. 30, n. 3, p. 63–79, 2009.
- VASQUEZ, M. L.; SOUSA, C. S. DE; CARVALHO, J. M. DE A. Mapa Geológico do Estado do Pará. Ministério de Minas e Energia, p. 1, 2008.
- VIERA-TORRES, M.; SINDE-GONZ, I.; GIL-DOCAMPO, M. Generating the Baseline in the Early Detection of Bud Rot and Red Ring Disease in Oil Palms by Geospatial Technologies. 2020.

## Anexo 1: Índices de vegetação utilizados no estudo

Nome	Referência	Descrição
Índice de vegetação ajustado ao solo (SAVI)	Huete (1988)	Desenvolvido como uma modificação do NDVI para corrigir a influência do brilho do solo quando a cobertura vegetal é baixa.
Índice de vegetação de diferença normalizada da borda do vermelha (NDVI705)	GitelsonandMerzlayk (1994)	Usado em agricultura de precisão, monitoramento florestal e detecção de estresse de vegetação.
Índice de Vegetação Resistente Ajustado Atmosfericamente (ARVI)	Kaufman andTanre (1992)	O índice é um aprimoramento do NDVI e usa refletância azul para corrigir a refletância vermelha para espalhamento atmosférico. É mais útil em regiões de alto conteúdo de aerossol atmosférico, incluindo regiões tropicais contaminadas por fuligem da agricultura de corte e queima.
Índice de razão simples (SR)	<a href="http://www.exelisvis.com/docs/BroadbandGreensess.html">http://www.exelisvis.com/docs/BroadbandGreensess.html</a>	Usado para estudar propriedades espectrais de solos e vegetação verde.
Índice de Refletância Carotenóide 1 (CRI1)	<a href="http://www.exelisvis.com">http://www.exelisvis.com</a>	Medida da vegetação estressada como consequência dos efeitos nocivos do excesso de luz.

Índice de Refletância Carotenóide 2 (CRI2)	<a href="http://www.exelisvis.com">http://www.exelisvis.com</a>	Medida modificada da vegetação estressada como consequência dos efeitos nocivos do excesso de luz.
Índice de refletância de antocianina 1 (ARI1)	Gitelson et al. (2001)	Medida da vegetação estressada e indica mudanças na folhagem por meio de novo crescimento ou morte.
Índice de refletância de antocianina 2 (ARI2)	Gitelson et al. (2001)	Detecta concentrações mais altas de antocianinas na vegetação. Mede a vegetação estressada e indica mudanças no dossel na folhagem por meio de um novo crescimento ou morte..
Índice de área foliar (LAI)	Bréda (2003)	Caracteriza os dosséis das plantas e prediz a produção fotossintética primária, evapotranspiração e é uma ferramenta de referência para o crescimento da cultura
Índice de vegetação de diferença verde (GDVI)	Sripada et al (2006)	O índice minimiza os efeitos do solo de fundo enquanto enfatiza a vegetação verde e usa coeficientes globais que pesam os valores de pixel para gerar novas bandas transformadas.
Índice de vegetação de diferença normalizada verde (GNDVI)	Gitelson and Merzyak (1997)	Usado para prever as necessidades de nitrogênio para o milho.
Índice de vegetação de proporção verde (GRVI)	Sripada et al. (2006)	O índice minimiza os efeitos do solo de fundo enquanto enfatiza a vegetação verde e usa coeficientes globais que pesam os valores de pixel para gerar novas bandas transformadas.
Índice de Vegetação Ajustado do Solo Otimizado (OSAVI)	Rondeaux et al. (1996)	O índice fornece maior variação do solo do que o SAVI para cobertura vegetal baixa, ao mesmo

---

Índice Visível de Resistência Atmosférica (VARI)	Gitelson et al. (2002)	tempo que demonstra maior sensibilidade à cobertura vegetal superior a 50% e é melhor usado em áreas com vegetação relativamente esparsa onde o solo é visível através do dossel. Usado para estimar a fração da vegetação em uma cena com baixa sensibilidade aos efeitos atmosféricos.
Índice de clorofila verde (GCI)	Gitelson et al 2003)	O índice é usado para estimar o conteúdo de clorofila na folha em uma ampla gama de espécies de plantas. Ter amplos comprimentos de onda NIR e verde fornece uma melhor previsão do conteúdo de clorofila, permitindo mais sensibilidade e uma relação sinal-t-ruído mais alta.
Índice de Vegetação Ajustado para Solo Otimizado Verde (GOSAVI)	Shipada et al 2005	Este índice foi originalmente desenvolvido com fotografia infravermelha colorida para prever as necessidades de nitrogênio para conr. É semelhante ao OSAVI, mas substitui a faixa vermelha pelo verde.
Índice de área da folha (LAI)	Boegh et al (2002)	Este índice é usado para estimar a cobertura de folhagem e prever o crescimento e a produção da safra. O ENVIE calculou o LAI verde usando a seguinte fórmula empírica de Boegh et al (2002).
Índice de refletância fotoquímica (PRI)	Penuelaas et al (1995)	Este índice é uma medida de refletância que é sensível a mudanças nos pigmentos carotenóides na folhagem viva. Os pigmentos carotenóides são indicativos da eficiência do uso da luz fotossintética ou da taxa de absorção de dióxido de

---

---

carbono pela folhagem por unidade de energia absorvida. Relacionado com a produtividade da vegetação e estresse.

Índice de vegetação por  
diferença normalizada  
(NDVI)

Handisky et al (1983)

---

**Anexo 2: Covariáveis utilizadas externas aos índices de vegetação.**

<b>Co-variável</b>	<b>Descrição</b>	<b>Tipo</b>
10 Bandas espectrais	Bandas 2 a 8a e bandas 11 e 12 da imagem Sentinel 2A	Espectral
Aspect	Orientação das vertentes	
Convergence Index	Índice de Convergência	
Cross SectionalCurvature	Curvatura Transversal	
Curvatureclassification	Classificação da curvatura	
Difference	Diferença do gradiente hidrológico	
Diffuse Solar Radiation (January)	Radiação difusa de janeiro	
Diffuse Solar Radiation (June)	Radiação difusa de junho	
Digital Elevation Model	Modelo digital de elevação	Topografia
Direct Solar Radiation (January)	Radiação direta de janeiro	
Direct Solar Radiation (June)	Radiação direta de junho	
DiurnalAnisotropicHeating	Aquecimento anisotrópico diurno	
FlowLineCurvature	Curvatura da linha de fluxo	
General Curvature	Curvatura geral	
Gradient	Gradiente hidrológico	
Hill	Morros	

<b>Co-variável</b>	<b>Descrição</b>	<b>Tipo</b>
Hill Index	Índice de morros	
Landforms	Formas do relevo	
Longitudinal Curvature	Curvatura longitudinal	
Mass Balance Index	Índice de balanço entre erosão e deposição	
Maximal Curvature	Curvatura máxima	
Mid-Slope Position	Posição de inclinação média	
Minimal Curvature	Curvatura mínima	
Morphometric Protection Index	Índice de proteção de um ponto em relação ao relevo do entorno	
Multiresolution Index of Ridge Top	Índice que identifica áreas planas em altas altitudes	
Normalized Height	Distância vertical entre a base e o cume do declive normalizada	
Planar Curvature		
Profile Curvature	Curvatura no perfil	
Real Surface Area	Cálculo real da área da célula	
Slope	Declividade	

<b>Co-variável</b>	<b>Descrição</b>	<b>Tipo</b>
SlopeHeight	Distância vertical entre a base e o cume do declive	
Slope Index	Índice de declividade	
StandardizedHeight	Inclinação entre a base e o cume do declive normalizado	
Surface Specific Points	Pontos quantitativos da superfície do terreno	
Tangencial Curvature	Curvatura tangencial	Topografia
TerrainRuggedness Index	Índice quantitativo da rugosidade da topografia	
Terrain Surface Convexity	Convexidade superficial do terreno	
Topographic Position Index	Diferença vertical entre a base e o cume do declive padronizada	
Total Curvature	Curvatura total	
Total Solar Radiation (January)	Radiação solar total de janeiro	
Total Solar Radiation (June)	Radiação solar total de junho	
Valley	Vale	
Valley Depth	Profundidade do vale calculada pela distância vertical ao nível da hidrografia	
Valley Index	Índice de elevação dos morros com base em Valley Depth	

<b>Co-variável</b>	<b>Descrição</b>	<b>Tipo</b>
Vector RuggednessMeasure	Medida da variação da rugosidade do terreno da declividade	
Wetness Index	Índice do controle topográfico nos processos hidrológicos	
Mapa de Solos	Mapa de solos da área de estudo, reorganizado para o terceiro nível categórico.	Solos
Geologia	Mapa das unidades lio estratigráficas da área de estudo	Geologia

**Anexo 3: Variáveis utilizadas do banco de dados de fitossanidade.**

<b>Nome da variável</b>	<b>Tipo de banco de dados</b>
Material genético	Geral
Ano de plantio	Geral
Anomalias	Fitossanidade
Tipo de lagarta desfolhadora	Fitossanidade
Total de plantas encontradas com algum tipo de lagarta desfolhadora	Fitossanidade
Total de lagartas acumulado por planta	Fitossanidade

## **MAPEAMENTO DO AMARELECIMENTO FATAL NA PALMA DE ÓLEO UTILIZANDO IMAGENS DE ALTÍSSIMA RESOLUÇÃO**

**Abstract:** The present study aimed to generate a method for identify plants with bud rot using very high resolution images. The image acquisition was performed with a camera equipped with red, green and blue channels, in stands with bud rot. The work was developed in two stages: the first, drone flights for the georeferencing of plants and obtaining the orthomosaic and the second with the processing of the orthomosaic and generation of vegetation indices ERBVE, VARI, GVI and RGVBI. Field validation of plants with bud rot was carried out to compare the results. The plant count performed in phase A, with the use of MDS and MD, showed good results for the georeferencing of plants. The use of vegetation indices and machine learning techniques to map bud rot in oil palm was efficient in identify plants with bud rot and healthy plants.

Palavras-chave: óleo de palma, sensoriamento remoto, agricultura de precisão, detecção de doenças, contagem de plantas, imagens de alta resolução, machine learning

### **1.0 Introdução**

A palma de óleo é uma oleaginosa originada no continente africano e é uma cultura de ciclo produtivo longo. O óleo extraído de seus frutos é comestível e emergiu como um dos produtos agrícolas de maior importância mundial, apresentando a maior taxa de extração de óleo, quando comparado a oleaginosas como a soja, milho, algodão, etc. (MPOC, 2012). O Brasil é o quinto maior produtor do mundo de óleo de palma, com cultivos nos estados do Pará, Bahia e Roraima (MPOC, 2012; RAMALHO FILHO et al., 2010).

Apesar de apresentar números representativos de produção, quando comparado aos principais produtores, o Brasil tem baixa produtividade. A Colômbia extrai cerca de 900 ton/ha de óleo, enquanto a taxa de extração do Brasil não passa de 350 ton/ha. Qual seria o fator responsável por esta diferença tão significativa? São inúmeros os fatores, perpassando por material genético cultivado, ataque de doenças e pragas, regime hídrico, método de extração etc. Um hectare de palma de óleo plantado pode produzir até dez vezes mais do que outras oleaginosas (SRESTASATHIERN; RAKWATIN, 2014).

Nos plantios do estado do Pará, o distúrbio Amarelecimento Fatal (AF) vem sendo considerado como o maior entrave ao aumento da produtividade, sendo responsável por dizimar grandes áreas de palma de óleo todos anos (BOARI, 2010; RAMALHO FILHO et al., 2010). O AF não tem agente causal identificado, caracteriza-se pelo amarelecimento gradual das folhas jovens com presença de necrose apical que progride em direção a base da planta (BOARI, 2010; LAING, 2014; ROT; WILT; RING, 2010; VIERA-TORRES; SINDE-GONZ; GIL-DOCAMPO, 2020).

O sensoriamento remoto apresenta uma série de dados que podem servir como subsídio para a otimização dos plantios de palma de óleo, considerando a grande extensão que apresentam e a heterogeneidade que neles ocorre. A inserção de tecnologias e de ferramentas que possibilitem o gerenciamento da unidade produtiva, visando a otimização dos recursos, aliado à máxima produtividade é de grande importância.

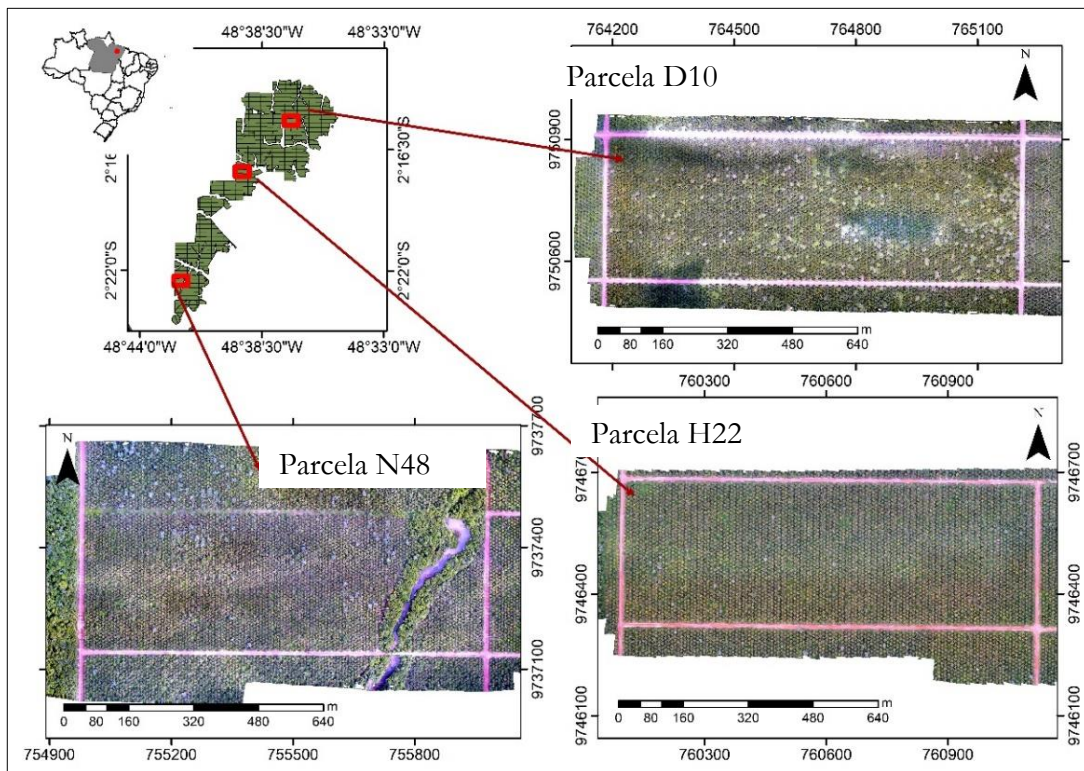
Já é comum o uso de ferramentas da geotecnologia no manejo da cultura da palma. Almiro et al., (2020), utilizou imagens Sentinel 2A e índices de vegetação para o mapeamento do anel vermelho em plantios comerciais no estado do Pará, indicando que o uso de imagens de satélite e de índices de vegetação tem muito a contribuir para a cultura da palma de óleo. As conclusões dos referidos autores deixam claro que estudos que utilizam imagens de alta resolução, que permitam o detalhamento espectral da palma de óleo nos plantios da Amazônia brasileira são de grande importância. No entanto, nenhum estudo para esta região utilizou imagens de alta resolução, obtidas por aeronaves remotamente tripuladas no mapeamento de doenças, mesmo que, na literatura muitos sejam os trabalhos que apresentem a importância do uso destas imagens atreladas ao mapeamento de doenças (BALLESTER et al., 2019; DLAMINI et al., 2019; KERKECH; HAFIANE; CANALS, 2018; SONG et al., 2017; VIERA-TORRES; SINDE-GONZ; GIL-DOCAMPO, 2020).

Outro ponto de relevância para os plantios de palma é a contagem de plantas. Conhecer quantas plantas existem dentro das parcelas e a sanidade das mesmas, facilita a previsão de safra e auxílio no manejo das áreas (SRESTASATHIERN; RAKWATIN, 2014). Os métodos convencionais de contagem de plantas são realizados manualmente em imagem de satélites ou com marcação com o uso de aparelho receptor do sinal GNSS (MUBIN et al., 2019). Assim, o desenvolvimento de uma metodologia que utilize imagens de alta resolução apresenta relevância em função da rapidez e acurácia do resultado.

Assim o presente estudo tem por objetivo desenvolver uma metodologia de mapeamento do amarelecimento fatal (AF) com o uso de imagens de alta resolução obtidos por VANT. Além disso, desenvolver um método de contagem de plantas, visto que a contagem de plantas para o mapeamento de doenças e distúrbios dos plantios é um fator chave para o gerenciamento e monitoramento dos plantios.

## 2.0 Área de estudo

A área de estudo é composta por três parcelas de plantio de palma de óleo localizadas no município de Acará, estado do Pará (figura 1). Estas parcelas pertencem a duas fazendas de plantios comerciais, pertencentes ao Agropalma que atua na região. Cada uma das parcelas tem área aproximada de 35 hectares. São cultivadas palmas das variedades Deli x Lamé com 19 anos (N48), 20 anos (H22) e 32 anos (D10).



**Figura 1:** Mapa de localização das parcelas estudadas.

Em todas estas parcelas ocorre a incidência do distúrbio AF. Os sintomas do AF apresentam-se primeiramente nas folhas mais jovens, com o amarelecimento das mesmas, que evolui gradualmente a necrose. O definhamento move-se lentamente para o núcleo do

meristema, comprometendo a emissão de novas folhas, restringindo o desenvolvimento das plantas. Na figura 2 é mostrado imagens de duas plantas com AF.



**Figura 2:** Imagens de duas plantas apresentando sintomas mais comuns do AF. Fotos registradas em janeiro de 2019 (direita) e outubro de 2020 (direita). Fotos autorais.

No levantamento de campo disponibilizado pela empresa, para o ano de 2019, a parcela H22 apresentava 25% de sua área com plantas com AF. Já a parcela N48 possuía 2,20% de comprometimento com a AF, enquanto que a D10 apresentava 0,85%. Para a parcela H22, a porcentagem de infecção somava 1.216 plantas que não estão atingindo o potencial produtivo na fase adulta. Na parcela N48, o número de plantas com AF registrado é de 105 plantas e para a parcela D10 422 plantas.

### 3.0 Materiais

Os voos de VANT foram realizados para a aquisição de imagens com o equipamento DJI Mavic Pro equipado com uma câmera que recobre os comprimentos de onda do visível. As especificações do voo são listadas na tabela 1.

**Tabela 1:** Especificações da câmera utilizada no presente estudo.

Câmera	Resolução	Resolução espacial	Resolução espectral	Altura de voo
	20 MP	4 cm	400 nm, 550 nm, 700 nm	150 m

Um aparelho receptor do sinal Global Navigation Satellite System (GNSS) foi utilizado para a obtenção das coordenadas de plantas ao longo dos talhões de trabalho. O plano de voo foi realizado no software DJI pilot e o processamento das imagens foi realizado no software AgisoftMetashape Professional 1.5.2. A contagem automática das plantas foi feita com o uso do software Saga e a definição das copas das plantas e o cálculo dos índices de vegetação com o uso do software Arcmap 10.8.1 (ESRI DATA, 2019). A classificação das plantas foi realizada no software R (R CORE TEAM, 2019). O banco de dados com o número de plantas com AF para as parcelas no ano de 2019 disponibilizado pelo Agropalma também foi inserido ao estudo.

#### 4.0 Metodologia

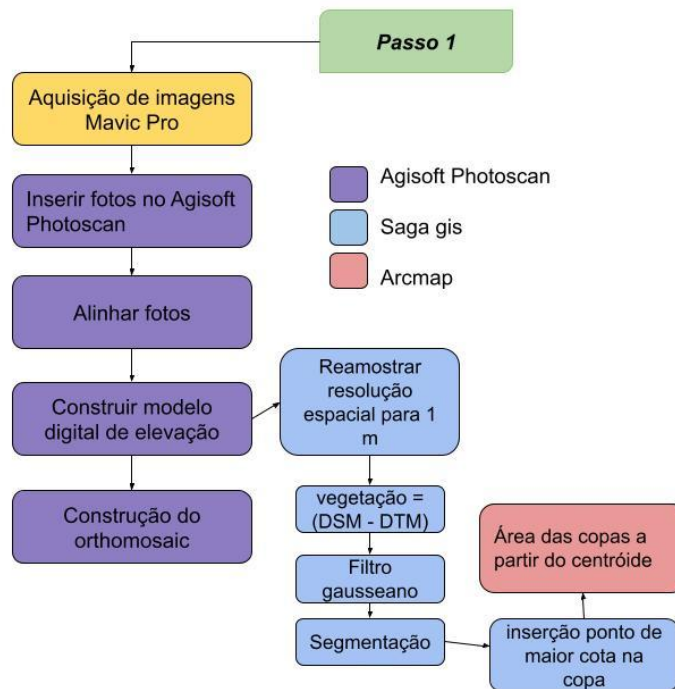
Após os voos, foi realizado o processo de geração dos ortomosaicos e dos Modelos Digitais de Elevação (MDE), ambos georreferenciados. Os números digitais das imagens não foram calibrados e convertidos para valores de reflectância. Após a geração dos ortomosaicos, a metodologia deste trabalho pode ser dividida em duas partes:

##### Fase A

Nesta fase foi realizada a contagem das plantas e a identificação da copa das árvores. A fim de se obter um arquivo vetorial com a indicação central das plantas, o modelo digital do terreno (MDT) e o modelo digital de superfície (MDS) foram gerados com resolução espacial de 1 metro. O MDS foi subtraído no MDT, individualizando a informação da altura da vegetação em cada parcela. Considerando a idade dos plantios, a altura média das plantas é de 8 metros, sendo assim, foi estabelecido o intervalo de 5 a 9 metros como a altura para plantas de palma de óleo nestas parcelas. Toda e qualquer vegetação fora deste intervalo, foi eliminada da análise. Nestas parcelas, é encontrado com certa frequência árvores da espécie *Cecropiapachystachya*, de nome popular Embaúba. Estas árvores caracterizam-se por terem altura entre 4 a 8 metros, e ocorrem muitas vezes sobre a plantas de palma de óleo, levando a altura ser superior a este intervalo. Com o intervalo estabelecido para a vegetação, grande parte

dessas plantas invasoras dos plantios foram eliminadas. Os pontos que foram indicados nas Embaúbas dentro das parcelas, foram eliminados manualmente após inspeção visual.

De posse do arquivo de vegetação, ele foi segmentado e para cada segmento gerado foi atribuído um ponto ou seed como mostrado no fluxograma das etapas (figura 3). Cada ponto foi atribuído a cota máxima da vegetação dentro do segmento. Quando a indicação do ponto foi muito distante da região central da copa da planta, foi realizado o ajuste manual.



**Figura 3:** Sequência das ferramentas e passos utilizados para o desenvolvimento metodológico.

Para a definição do tamanho médio das copas, sobre a imagem de cada parcela e dos seeds foi feita a amostragem de dez plantas com diferentes características: planta com AF e com grau de severidade avançado, planta com AF, mais com copa sem muitas alterações e plantas saudáveis. A partir dessas plantas, foram criados buffer's com raio variando de 0,10 m até 4,50 metros com intervalo de 0,10 cm. Foi selecionado de foram visual o buffer que cobrisse toda a copa das palmeiras com o mínimo de sobreposição com os polígonos vizinhos. Como resultado, foi obtido uma camada vetorial do tipo polígono, onde cada polígono refere-se a uma planta.

Para cada parcela, foi criado um arquivo de amostras de 110 pontos, divididos regularmente entre as classes 'plantas com AF' e 'plantas saudáveis'. Em campo e com o uso

do aparelho receptor do sinal GNSS, foi realizado o georreferenciamento de algumas plantas e anotado a sua sanidade. O caminhamento ao longo das parcelas contou com o auxílio de um profissional de campo do Agropalma. Estes pontos serviram de subsídio para novas amostras nas imagens das parcelas, completando assim o arquivo de amostras.

### Fase B

Nesta fase, com as imagens adquiridas nos voos, quatro índices de vegetação foram calculados. Índice aprimorado de vegetação vermelho-azul-verde (ERGBVE – Enhanced Red-Green-Blue Vegetation Index), índice de vegetação verde-vermelho (Gree-Red Vegetation Index – GVI), índice vermelho-verde (Red Green Index - RGI) e índice visível de resistência a atmosfera (Visible Atmospherically Resistant Index – VARI). Suas fórmulas e referências são apresentadas na tabela 2.

**Tabela 2:** Fórmulas dos índices de vegetação utilizados no estudo.

	Índice de Vegetação	Equação	Referências
ERGBV E	Enhanced Red-Green-Blue Vegetation Index	$\pi * ((\rho_G^2 - (\rho_R * \rho_B))$ $(\rho_G^2 + (\rho_R * \rho_B))$	Themistocleous, 2019
GVI	Gree-RedVegetation Index	$(\rho_G - \rho_R)$ $(\rho_G + \rho_R)$	Kerkech; hafiane; canals, 2018
RGVBI	Red Green Index	$\rho_R$ $\rho_G$	Gamon & Surfus, 1999; Coops et al., 2006
VARI	VisibleAtmosphericallyResista nt Index	$(\rho_G - \rho_R)$ $(\rho_G + \rho_R - \rho_B)$	Gitelson, 2003

Onde:  $\rho_G$  corresponde a banda do verde,  $\rho_B$  corresponde a banda do azul e  $\rho_R$  corresponde a banda do vermelho.

A metodologia desenvolvida para a classificação das plantas com AF, foi a classificação orientada a objetos, OBIA. O processo de classificação baseia-se em atribuir classes a grupos homogêneos em padrões espectrais e espaciais. Na OBIA a análise de parâmetros vão além da resposta espectral de cada pixel, agregando ao método informações de forma, textura, e detalhamento espacial (THENKABAIL, 2015). As copas das plantas identificadas na fase A, tornaram-se objetos da OBIA.

O polígono que representa a área das copas realizada na fase A, foi utilizado para o recorte dos índices de vegetação, o que permitiu a visualização das informações apenas para as plantas, excluindo os elementos do solo que circundam as mesmas. Para cada planta foi feita a extração dos valores dos índices de vegetação e submetido ao cálculo das seguintes estatísticas descritivas: mínimo, média, mediana, amplitude interquartil, coeficiente de variação, os centis 5, 25, 75 e 95 e o valor máximo. Este processo no arquivo de amostras, gerou um banco de dados com 312 linhas e 48 colunas. As linhas referem-se as amostras, e as colunas as variáveis explicativas.

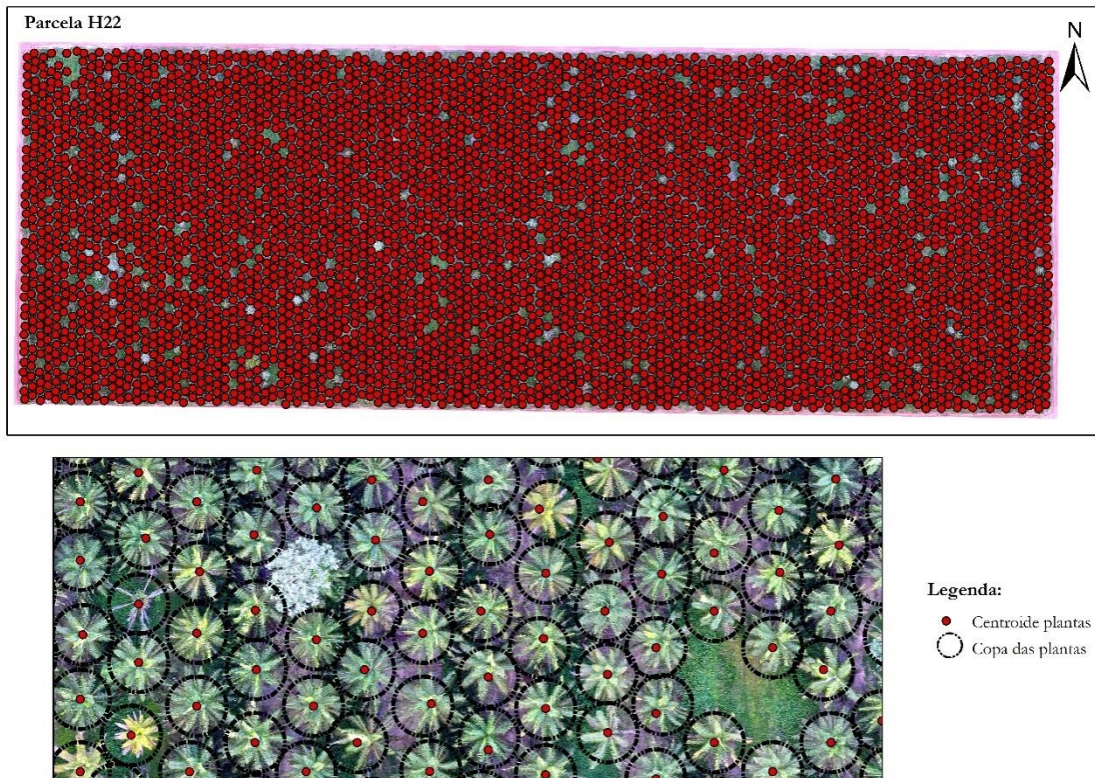
O banco de dados para o treinamento do modelo foi particionado em 75% dos dados para treino e 25% para teste do modelo. Com o uso da função *find correlacion* (KUHN; JOHNSON, 2013), as covariáveis com correlação maior que 95% foram eliminados. Após esta etapa o conjunto de dados foi submetido a seleção de variáveis por importância pelo método Recursive Feature Elimination (RFE) do pacote caret (KUHN; JOHNSON, 2013). O RFE é um método de seleção de variáveis do tipo backward, que inicialmente ajusta o modelo com todas as covariáveis e paulatinamente elimina as variáveis com menor importância, e ao final seleciona o melhor conjunto de variáveis explicativas para a variável resposta (HENRIQUE BRANT DIAS et al., 2021)

Para a classificação foi utilizado o algoritmo Random Forest (BREIMAN, 2001) que é baseado na ideia do aprendizado em conjunto. É um modelo conhecido por sua eficiência em reduzir o ruído experimental e aumentar a acurácia da predição (GOMES et al., 2019). A acurácia da classificação foi calculada feita com o conjunto de teste com 25% dos dados e a métrica de exatidão usada foi o índice Kappa, proposto por Cohen (1960). O índice Kappa reflete o nível de exatidão da classificação e varia de 0 a 1, sendo que, quanto mais próximo a 1, maior é a concordância dos dados preditos pelo modelo com os dados observados em campo. A importância das variáveis também foi calculada.

O processo de treinamento do modelo foi repetido 100 vezes. Em cada repetição, o particionamento dos dados em treino e teste era repetido de forma aleatória. Nesse procedimento cria conjunto de treino e teste diferentes a cada execução levando conseqüentemente aos ajustes de modelos diferentes, gerando variabilidade a partir de um único conjunto de dados iniciais. A métrica de exatidão (Kappa) foi calculada a partir da média dos 100 valores de kappa, foi calculado também o desvio padrão do índice kappa

## 5.0 Resultados e Discussão

A figura 4 traz o resultado da etapa de detecção de plantas de palma de óleo para a parcela H22 (fase A). A figura 4a é a visualização geral para toda a parcela, e a figura 4b um recorte para uma pequena área da parcela. O ponto de maior altura para as plantas nem sempre foi indicado na região central delas, em razão da alteração da arquitetura de copa que o AF ocasiona. Com a emissão de novas folhas comprometida, ocorre o afunilamento da copa, ou inúmeras flechas são acumuladas na região central das plantas. Ambas as situações influenciam para que o ponto indicativo não seja na região central. O valor de *buffer* que satisfaz as condições pré-estabelecidas foi de 4,10 metros (figura 4b).



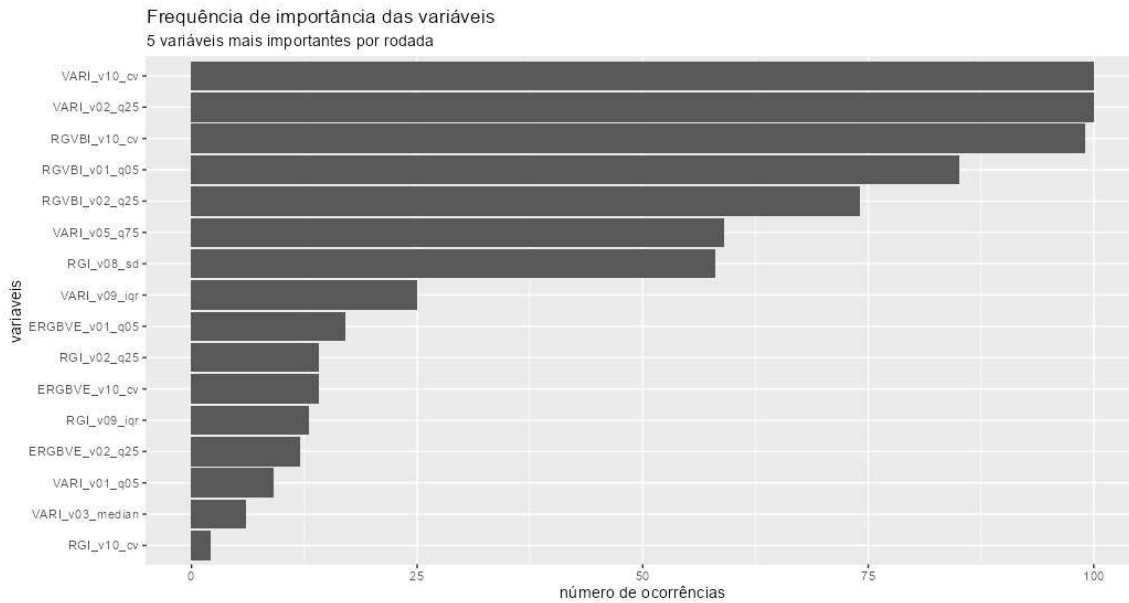
**Figura 3:** Resultado da detecção automática de plantas baseada na altura das plantas.

O método para detectar plantas de palma de óleo de maneira individual foi muito adequado com a indicação de apenas um polígono por planta, sem muitas interferências de plantas invasoras ou do substrato (figura 4). O intervalo de 5 a 9 metros de altura das plantas também foi satisfatório. Srestasathiern and Rakwatin (2014) desenvolveram para plantios de palma de óleo na Tailândia, um método de contagem de plantas baseado em imagens multiespectrais de alta resolução. Estes autores definiram que o centro da planta é o ponto de maior reflectância quando comparado a áreas de seu entorno, e a esta área, foi atribuída o ponto

de identificação da copa das plantas. Tal método apresenta grandes avanços, no entanto, não é aplicado ao mapeamento de doenças como o AF. Os sintomas do AF iniciam-se pelas folhas mais jovens, na região central, que apresentará uma lesão, sendo esta região com menor valor de reflectância, se comparado ao seu entorno, ainda livre de AF. A metodologia que utiliza a diferença entre o MDS e MDT, como a altura das plantas, torna-se muito mais relevante, podendo ser aplicado ao inventário dos plantios e ao monitoramento dos plantios e de outras doenças na palma de óleo. Outra abordagem com o uso desta metodologia de detecção de plantas, é o cálculo do número de plantas que morreram e o cálculo do extrato herbáceo e de plantas invasoras contidas nos talhões, que pode auxiliar no direcionamento de ações de poda e capina dessas áreas.

Quanto a importância das variáveis (Figura 5), nas 100 vezes que o modelo foram executadas variáveis VARI\_CV e VARI\_q\_25 estavam presentes (figura 5). O índice de vegetação VARI baseia-se na variabilidade do vigor e do estresse das plantas (MEIVEL; MAHESWARI, 2022), a indicação do coeficiente de variação do VARI como a variável mais importante faz-se relação com a diferença de coloração e de forma das folhas das plantas saudáveis (VIERA-TORRES; SINDE-GONZ; GIL-DOCAMPO, 2020) e com plantas com AF. Plantas com AF terão folhas amareladas e mais curtas se comparadas às plantas saudáveis.

O valor do coeficiente de variação do VARI para as plantas com AF apresenta alta dispersão, em razão dos diferentes graus de severidade do AF. Baixos valores de VARI, representados pelo primeiro quartil (q25), foi a segunda variável de maior importância e completa a ideia dos diferentes graus de severidade do AF e as alterações que ocorrem na copa das plantas de palma de óleo. Plantas saudáveis, não estarão sob o mesmo nível de estresse que plantas com AF. O estresse nutricional é uma interferência na saúde da cultura causada pela falta de elementos nutritivos para apoiar a exigência nutricional da cultura (LICHTENTHALER, 1998; RENDANA et al., 2015).



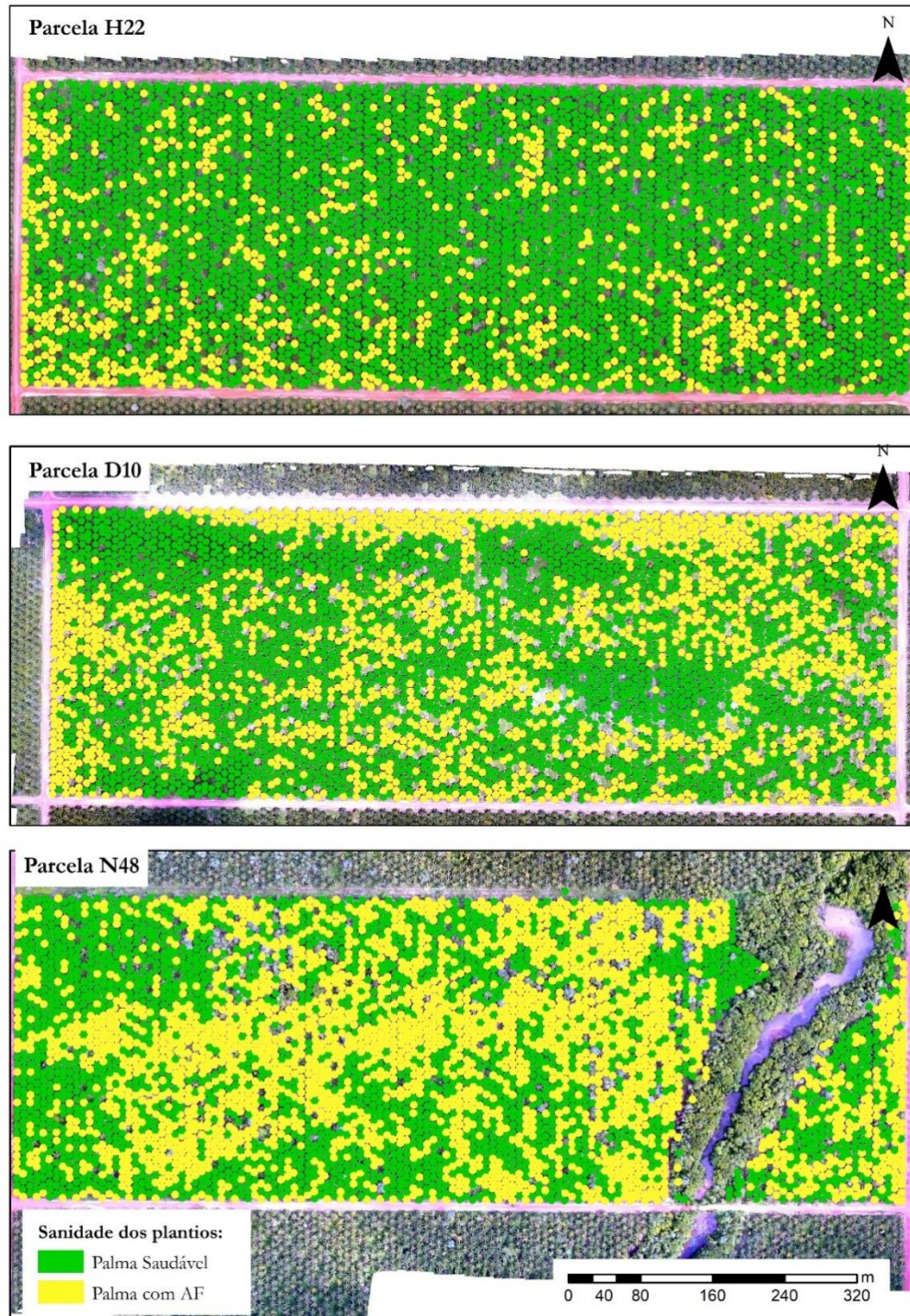
**Figura 5:** Gráfico com a frequência de importância das variáveis nas 100 repetições. Cv = coeficiente de variação.

O índice de vegetação RGVBI também apresentou alta frequência na importância das variáveis. Dentre as estatísticas calculadas para este índice o coeficiente de variação e os valores mínimos e o primeiro quartil, foram as mais importantes. Este índice possibilitou a separação das plantas de palma de óleo que estavam com número de folhas reduzido em razão do AF, sendo um bom indicador de estresse na planta.

O RBVBI se correlaciona com a fenologia e com o teor de água das folhas, os resultados obtidos por Ballester et al. (2019) no mapeamento do estresse hídrico na cultura do algodão, compararam o desempenho do RGVBI aos do índice de vegetação por diferença normalizada (NDVI) que utiliza de bandas do infravermelho para o seu cálculo. Em termos de custos, o índice RBVBI é muito inferior, visto que não é necessário o investimento em câmeras multiespectral como é necessário para o NDVI. Em palma de óleo, o RGVBI foi aplicado no mapeamento da podridão basal e foi indicado pelos autores como um dos melhores índices de vegetação para diferenciar plantas saudáveis de plantas infectadas pela podridão basal (KHAIRUNNIZA-BEJO et al., 2015).

O índice Kappa médio da classificação foi de 0,78, considerado bom por Landis and Koch (1977). A classificação final para cada uma das parcelas é apresentada na figura 8. Comparando o percentual de plantas com AF com o banco de dados de campo do Grupo Agropalma para a parcela H22, o percentual registrado por eles é de 24,01% e o percentual

encontrado pela classificação realizada foi de 25,95%, que mostra uma alta concordância entre o valor estimado pelo modelo e o valor obtido no campo (tabela 3).



**Figura 7:** Mapas com a sanidade dos plantios estudados.

**Tabela 3:** Informações descritivas das parcelas analisadas com comparação dos dados mapeados e dos dados levantamentos pela Agropalma.

Parcela	Mapeamento realizado				Inventário de campo		
	Plantas sadias	Plantas com AF	% de plantas com AF	Total de plantas	Total de plantas	Plantas com AF	% de plantas com AF
<b>D10</b>	2696	1807	40,1%	4503	4850	-	-
<b>N48</b>	2029	2401	54,2%	4430	4793	1995	41,62%
<b>H22</b>	3600	1076	24,01	4463	4864	-	-

Para as demais parcelas a porcentagem de casos de AF mapeado em campo no banco de dados da Agropalma, com a porcentagem mapeada nesta classificação apresenta grande discrepância (tabela 3). Tal fato pode ser atribuído a não calibração dos números digitais das imagens em valores de reflectância. A calibração radiométrica corrige os valores em termos de reflectância, e minimiza os efeitos da inclinação solar nos diferentes dias de tomada das imagens aéreas. A calibração radiométrica visa não apenas converter os valores digitais para valores de reflectância, mais também atenuar os efeitos da atmosfera e da modificação do ângulo de iluminação solar (GUO et al., 2019). Dos Santos (et al. 2021) afirma que a calibração radiométrica das imagens de Vant é um pré-requisito necessário para as imagens aéreas. Nos voos foram realizados no mesmo dia, mais em horários diferentes, o que também pode alterar os padrões espectrais entre as parcelas.

## 6.0 Conclusões

A contagem automática de plantas com a utilização do MDS e do MDT com a obtenção da altura da vegetação, mostrou-se satisfatória e eficiente para os plantios. A metodologia que utiliza a altura das plantas foi satisfatória e promissora.

Os resultados mostraram que com a utilização de índices de vegetação e de o uso de técnicas de aprendizagem de máquina é possível mapear o amarelecimento fatal na palma de óleo, bem como identificar as variáveis de maior importância para o modelo utilizado.

O uso dos índices de vegetação obtidos com imagens na região espectral do visível mostraram-se eficientes no mapeamento do distúrbio amarelecimento fatal. A aplicação do

índice de vegetação VARI é promissor para o mapeamento do distúrbio amarelecimento fatal na palma de óleo, sendo potencial para a utilização em doenças na palma de óleo.

### Referências Bibliográficas

- ALMIRO, J. et al. Detection of Diseases in Oil Palm Plantations in the Brazilian Amazon through Orbital Image. v. 12, n. 4, p. 138–147, 2020.
- AMIRRUDDIN, A. D. et al. Hyperspectral remote sensing for assessment of chlorophyll sufficiency levels in mature oil palm (*Elaeis guineensis*) based on frond numbers: Analysis of decision tree and random forest. *Computers and Electronics in Agriculture*, v. 169, n. December 2019, p. 105221, 2020.
- AZMAN, I.; MOHD NOOR, M. The Optimal Age of Oil Palm Replanting. *Oil Palm Industry Economic Journal*, v. 2, n. 1, p. 11–18, 2002.
- BALLESTER, C. et al. Monitoring the Effects of Water Stress in Cotton Using the Green Red Vegetation Index and Red Edge Ratio. 2019.
- BENÍTEZ, É.; GARCÍA, C. Historia de las investigaciones en pudrición del cogollo de la palma de aceite (*Elaeis guineensis* Jacq.) en Colombia. *Agronomía Colombiana*, v. 32, n. 3, p. 390–398, 2014.
- BOARI, A. DE J. O desafio das pesquisas com a etiologia do Amarelecimento Fatal (AF) da cultura da palma de óleo. [s.l.: s.n.].
- BREIMAN, L. Random forests. *Machine Learning*, v. 45, n. 1, p. 5–32, 2001.
- BRENNING, A.; BANGS, D.; BECKER, M. RSAGA: SAGA geoprocessing and terrain analysis in R. R package version 0.9, 2007.
- CHEMURA, A.; MUTANGA, O.; DUBE, • TIMOTHY. Separability of coffee leaf rust infection levels with machine learning methods at Sentinel-2 MSI spectral resolutions. v. 18, p. 859–881, 2017.
- CHONG, K. L. et al. A review of remote sensing applications for oil palm studies. *Geo-Spatial Information Science*, v. 20, n. 2, p. 184–200, 2017.
- COHEN, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, v. 20, n. 1, p. 37–46, 1960.
- CONGALTON, R. G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, v. 37, n. 1, p. 35–46, 1991.
- DLAMINI, S. N. et al. Review of remotely sensed data products for disease mapping and epidemiology. *Remote Sensing Applications: Society and Environment*, v. 14, n. January,

- p. 108–118, 2019.
- DOS SANTOS, R. A. et al. Surface reflectance calculation and predictive models of biophysical parameters of maize crop from RG-NIR sensor on board a UAV. *Precision Agriculture*, n. February, 2021.
- DOUGLAS LAING. *Informações agronômicas*. n. 19, p. 1–17, 2012.
- ESRI DATA. ArcmapRedlands, California, United States, 2019. Disponível em: <<https://www.esri.com/pt-br/home>>
- GISLASON, P. O.; BENEDIKTSSON, J. A.; SVEINSSON, J. R. Random forests for land cover classification. *Pattern Recognition Letters*, v. 27, n. 4, p. 294–300, 2006.
- GITELSON, A. A.; † Y. G.; MERZLYAK, M. N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, v. 160, n. 3, p. 271–282, 2003.
- GITELSON, A. A.; GRITZ †, Y.; MERZLYAK, M. N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, v. 160, n. 3, p. 271–282, 2003.
- GOMES, L. C. et al. Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma*, v. 340, 2019.
- GUO, Y. et al. Radiometric Calibration for Multispectral Camera of Different Imaging Conditions Mounted on a UAV Platform. v. 11, p. 978, 2019.
- HENRIQUE BRANT DIAS, S. et al. Reference evapotranspiration of Brazil modeled with machine learning techniques and remote sensing. 2021.
- HOE, J. H.; SHUKOR, S. A. A. Monitoring Oil Palm Tree Health-A Review. *IOP Conference Series: Materials Science and Engineering*, v. 705, n. 1, 2019.
- KERKECH, M.; HAFIANE, A.; CANALS, R. Deep learning approach with colorimetric spaces and vegetation indices for vine diseases detection in UAV images. *Computers and Electronics in Agriculture*, v. 155, n. July, p. 237–243, 2018.
- KHAIRUNNIZA-BEJO, S. et al. Identification of Healthy and BSR-Infected Oil Palm Trees Using Color Indices. *International Journal of Biological, Biomolecular, Agricultural, Food and Biotechnological Engineering*, n. January, 2015.
- KUHN, M.; JOHNSON, K. *Applied Predictive Modeling*. [s.l: s.n.].
- LAING, D. La Causa de Pudrición de Cogollo ( PC ) en Palma de Aceite : Hipótesis Abiótica - Edáfica. *Presentación CIAT*, p. 113, 2009.
- LAING, D. Causa de la pudricion de cogollo en la palma de aceite papel del calcio en una

- hipotesis abiotica-edafica. *Informaciones Agron[omicas]*, v. 3, n. June, p. 27, 2014.
- LANDIS, J. R.; KOCH, G. G. The Measurement of Observer Agreement for Categorical Data  
Data for Categorical of Observer Agreement The Measurement. v. 33, n. 1, p. 159–174,  
1977.
- LANG, S. Object-based image analysis for remote sensing applications : modeling reality –  
dealing with complexity. In: *Object-based image analysis for remote sensing applications*.  
[s.l: s.n.]. p. 3–27.
- LEBOURGEOIS, V. et al. A combined random forest and OBIA classification scheme for  
mapping smallholder agriculture at different nomenclature levels using multisource data  
(simulated Sentinel-2 time series, VHRS and DEM) *Remote Sensing*, 2017.
- LIAW, A; WIENER, M. Classification and Regression by randomForest. *R news*, v. 2, n.  
December, p. 18–22, 2002.
- LICHTENTHALER, H. K. The Stress Concept in Plants: An Introduction. *Annals of the New  
York Academy of Sciences*, v. 851, n. 1, p. 187–198, 1998.
- MARZUKHI, F.; NOOR, S.; SAUTI, S. Application of remote sensing in the investigation of  
maturity age of palm oil trees In Pasir Puteh, Kelantan. *Jurnal Intelek*, v. 13, n. 1, p. 44–  
57, 2018.
- MEIVEL, S.; MAHESWARI, S. Monitoring of potato crops based on multispectral image  
feature extraction with vegetation indices *Multidimensional Systems and Signal  
Processing*, 2022.
- MPOC. MALASIAN PALM OIL COUNCIL (MPOC). Disponível em: <<http://mpoc.org.my/>>.  
Acesso em: 11 abr. 2021.
- MUBIN, N. A. et al. Young and mature oil palm tree detection and counting using  
convolutional neural network deep learning method. *International Journal of Remote  
Sensing*, v. 40, n. 19, p. 7500–7515, 2019.
- PONZONI, F. Comportamento espectral da vegetação. *Sensoriamento remoto: reflectância de  
alvos naturais*, v. 1, n. Brasília: UNB, p. 157–199, 2001.
- R CORE TEAM. No Title Vienna, Austria, 2019. Disponível em: <<http://www.r-project.org/index.html>>
- RAMALHO FILHO, A. et al. Zoneamento Agroecológico, Produção e Manejo para a Cultura  
da Palma de Óleo na Amazônia. Rio de Janeiro: [s.n.]. v. 1
- RENDANA, M. et al. A Review of Methods for Detecting Nutrient Stress of Oil Palm in  
Malaysia. *J. Appl. Environ. Biol. Sci. Journal of Applied Environmental and Biological*

- Sciences, v. 5, n. 6, p. 60–64, 2015.
- ROT, B.; WILT, S.; RING, R. Pudrición del cogollo , Marchitez sorpresiva , Anillo rojo y Marchitez letal en la palma de aceite en América Bud Rot , Sudden Wilt , Red Ring and Lethal Wilt of Introducción. v. 31, n. 1, p. 43–53, 2010.
- SANTOS, C.; MESSINA, J. P. Multi-Sensor Data Fusion for Modeling African Palm in the Ecuadorian Amazon. n. June, p. 711–723, 2008.
- SANTOSO, H.; TANI, H.; WANG, X. Random Forest classification model of basal stem rot disease caused by *Ganoderma boninense* in oil palm plantations. *International Journal of Remote Sensing*, v. 38, n. 16, p. 4683–4699, 2017.
- SHAFRI, H. Hyperspectral Imagery for Mapping Disease Infection in Oil Palm Plantation Using Vegetation Indices and Red Edge Techniques. n. June 2009, 2014.
- SONG, X. et al. Evaluation of Sentinel-2A Satellite Imagery for Mapping Cotton Root Rot. *Remote Sensing*, v. 9, p. 906, 2017.
- SRESTASATHIERN, P.; RAKWATIN, P. Oil palm tree detection with high resolution multi-spectral satellite imagery. *Remote Sensing*, v. 6, n. 10, p. 9749–9774, 2014.
- SUNDRAM, S.; INTAN-NUR, A. M. A. South American Bud rot: A biosecurity threat to South East Asian oil palm. *Crop Protection*, v. 101, p. 58–67, 2017.
- THENKABAIL, P. S. Remotely Sensed Data Characterization, Classification, and Accuracies. [s.l: s.n.].
- TIEMANN, T. T. et al. Feeding the Palm: A Review of Oil Palm Nutrition. *Advances in Agronomy*, v. 152, p. 149–243, 2018.
- TORRES, L. A. Caracterización de firma espectral a partir de sensores remotos para el manejo de sanidad vegetal en el cultivo de palma de aceite Characterization of the Spectral Signature from Remote Sensors for the Management of Plant Health in the Oil Palm Crop Intro. v. 30, n. 3, p. 63–79, 2009.
- TSCHIEDEL, M.; FERREIRA, M. F. Introdução À Agricultura De Precisão: Conceitos E Vantagens. *Ciência Rural*, v. 32, n. 1, p. 159–163, 2002.
- VASQUEZ, M. L.; SOUSA, C. S. DE; CARVALHO, J. M. DE A. Mapa Geológico do Estado do Pará. Ministério de Minas e Energia, p. 1, 2008.
- VIERA-TORRES, M.; SINDE-GONZ, I.; GIL-DOCAMPO, M. Generating the Baseline in the Early Detection of Bud Rot and Red Ring Disease in Oil Palms by Geospatial Technologies. 2020.

## **Conclusão Geral**

O sensoriamento remoto apresenta grande potencial para ser utilizado na agricultura de precisão. A integração dos dados de fitossanidade, geoespaciais, machine learning e inteligência artificial sendo promissores ao estudo de distúrbios e doenças.

A contagem automática de plantas com a utilização do MDS e do MDT com a obtenção da altura da vegetação, mostrou-se satisfatória e eficiente para os plantios. A metodologia que utiliza a altura das plantas foi satisfatória e promissora.

Os resultados mostraram que com a utilização de índices de vegetação e de o uso de técnicas de aprendizagem de máquina é possível mapear o amarelecimento fatal na palma de óleo, bem como identificar as variáveis de maior importância para o modelo utilizado.