



FRANKLIN DANIEL INÁCIO

Estimativa da produtividade de amendoim usando machine learning e Google Earth Engine

LAVRAS-MG

2022

FRANKLIN DANIEL INÁCIO

Estimativa da produtividade de amendoim usando machine learning e Google Earth Engine

Monografia apresentada ao Colegiado do Curso de Agronomia para a obtenção do título de Bacharel.

Prof. Dr. Adão Felipe dos Santos

Orientador

Msc. Jarlyson Brunno C. Souza

Coorientador

LAVRAS - MG

2022

RESUMO

Formas de estimar a produtividade de uma cultura é de fundamental importância, para se traçar melhores estratégias mercadológicas e logísticas. Na cultura do amendoim, mapas de produtividade são ferramentas chaves para implementação de Agricultura de Precisão, pois trazem informações relevantes quanto a variabilidade espacial das lavouras. Entretanto, apesar da grande relevância, essas informações não são geradas por meio de monitores de colheita nas máquinas no momento da colheita. Dessa forma, o objetivo deste trabalho foi desenvolver e implementar de forma automatizada um método para estimar a produtividade de amendoim utilizando series temporais de imagens de satélite provenientes da plataforma Sentinel através da plataforma Google Earth Engine com uso de redes neurais. Para determinação da produtividade real coletamos amostras em campo. Para estimativa, utilizamos imagens multiespectrais do Sentinel, iniciou-se aos 90 dias após a semeadura, coincidindo com as datas de coletas em campo. Os modelos usando os índices tiveram acurácia de 1% de erro. Em geral os modelos MLP se saíram melhores com menos dados para treinamento e validação, enquanto os modelos RBF se saíram melhor quando tiveram uma base de dados maior. Em cada data os modelos obtiveram valores de erro próximo, menores que 300 kg ha⁻¹ na maioria dos casos. Conclui se que podemos usar índices de vegetação para previsão de produtividade com valores bem acurados usando redes neurais.

Palavras-chave: Sensoriamento Remoto. Agricultura de Precisão. Índices de Vegetação, Estimativa de produtividade, Redes Neurais.

Sumário

1. INTRODUÇÃO	5
2. REFERENCIAL TEÓRICO	6
2.1 Cultura do amendoim	6
2.2 Agricultura de precisão	7
2.3 Sensoriamento remoto	9
2.4 Índices de vegetação.....	10
2.5 Estimativa de produtividade.....	12
2.6 Inteligência artificial.....	13
3. MATERIAL E MÉTODOS	14
3.1 Áreas de estudo	14
3.2 Amostras de campo	14
3.3 Aquisição e processamento das Imagens de satélite	15
3.4 Análise dos dados.....	16
3.5 Banco de dados para treinamento e validação.....	18
3.6 Avaliação do desempenho dos modelos.....	19
4. RESULTADOS E DISCUSSÃO	19
5. CONCLUSÃO	26
6. REFERÊNCIAS	27

1. INTRODUÇÃO

Modelos baseados em variáveis climáticas são um dos vários métodos que existem para tentar estimar o rendimento da cultura. Esses métodos são em geral uma abordagem estatística na qual se usa regressões lineares com o intuito de relacionar a produtividade com vários parâmetros climáticos adquiridos em vários locais e datas do ano (JOSHI et al., 2020).

Tecnologias que auxiliem no planejamento, produção e estimativa são ferramentas indispensáveis para o ciclo de qualquer cultura. Os métodos de sensoriamento remoto são técnicas que fornecem informações sem que haja a necessidade de contato direto com os objetos e que analisadas de forma adequada podem dar direcionamento mais assertivo para o planejamento e manejo das operações agrícolas (SOUZA, 2020).

Atualmente temos disponíveis produtos de sensoriamento remoto de alta resolução espaço-temporal. Apesar destes produtos fornecerem boa resposta sobre a vegetação e sanidade das lavouras, são pesados e requerem elevada demanda por processos computacionais. Para contornar este tipo de problemas, ferramentas que desempenham funções computacionais em nuvem, como o Google Earth Engine (GEE), é apresentado como alternativa viável para implementação de series temporais para o monitoramento de diversas culturas. O GEE contém dados geoespaciais (MODIS, Landsat, Sentinel, entre outros), de todo o planeta e que podem ser realizadas múltiplas tarefas em escala global, as aplicações do GEE podem ser ligadas a dinâmica da vegetação, mudanças na cobertura do solo, trabalhos ligados a recursos hídricos, geomorfologia e etc. (PALUDO et al, 2020).

Pensando no monitoramento de series temporais da vegetação por meio de índices de vegetação (IV) e associar esses dados com a produtividade gerando informações para a tomada de decisão e planejamento estratégico, é fundamental que se utilize técnicas modernas de análise de dados, tais como Redes Neurais Artificiais (RNA's). As RNA's são sistemas computacionais que trabalham de forma paralelas, compostas por unidades de processamento simples, denominados neurônios artificiais, conectadas de maneiras específicas para desempenhar uma determinada tarefa (Binoti et al., 2013). Esse tipo de análise de dados consegue processar um grande volume de dados e transformar em informação confiável, principalmente dados que não possuem

um modelo de distribuição linear, ou seja, que é extremamente afetado pela variabilidade espaço-temporal.

Dessa forma, acredita-se que a utilização de computação em nuvem (GEE) para aquisição de imagens de satélites provenientes de plataformas livres associado com inteligência artificial (RNA), seja uma alternativa para o desenvolvimento de modelos de predição da produtividade de amendoim. Assim, objetivou-se desenvolver e testar modelos de predição da produtividade de amendoim usando RNA do tipo Radial Base Function (RBF) e Multi Layer Perceptor (MLP).

2. REFERENCIAL TEÓRICO

2.1 Cultura do amendoim

O amendoim (*arachis hypogaea*) é uma leguminosa assim como a soja (*Glicine max*), e dentre as leguminosas é considerada a segunda cultura de maior importância no mundo. O centro de origem do amendoim é na América do Sul, especificamente na região do sudeste da Bolívia e noroeste da Argentina (FÁVERO et al., 2006). Pertencente ao gênero *Arachis*, compondo mais de 80 espécies descritas existindo espécies selvagens que são espécies em maioria diploides ($2n = 2x = 20$) e tetraploides que são as espécies cultivadas ($2n = 4x = 20$).

Botanicamente o amendoim é uma Fabaceae da subfamília Papilinoideae sendo ela uma dicotiledônea herbácea, anual de ciclo indeterminado (GREGORY et al., 1980). Seu sistema reprodutivo é formado de flores hermafrodita e Cleistogâmica, isso permite que ocorra a autofecundação com baixa taxa de polinização cruzada (NIGAM et al., 1990).

Seu cultivo tem importância ligada a fonte de proteína vegetal e de óleo. No que diz respeito a produção, a China produz 43% de todo o amendoim do mundo tornando a principal produtora do grão (BARBOSA; HOMEM; TARSITANO, 2014). No Brasil o estado de São Paulo é o maior produtor do grão, sendo responsável pela produção de 80 a 90%.

Além da questão alimentar, a cultura do amendoim pode ganhar mais espaço na economia Brasileira devido a crescente busca por tecnologias de alternativas de

combustíveis. Uma vez que óleos vegetais são uma alternativa aos combustíveis provindos de alternativas fósseis. (Ferrari et al., 2005).

Pensando na produção de biodiesel a partir do amendoim, vemos que essa cultura apresenta elevado potencial. O amendoim apresenta um teor de óleo de 40-60% quando comparado ao teor de óleo de outras matérias-primas como a mamona (39,6-59,8 %), o girassol (40-47%), a soja (18-21%), a colza (34-40%) e o caroço de algodão (18-20%). As sementes de amendoim são cultivadas em climas quentes, e podem ser utilizadas para produção de óleo, manteiga de amendoim, entre outros (Nunes, 2021).

No cenário mundial, a China, Índia e Estados Unidos são os maiores exportadores e consumidores do grão, o Brasil fica como o segundo maior produtor e exportador da América Latina, perdendo a liderança apenas para a Argentina, sendo o estado de São Paulo é responsável por 90% da produção nacional do amendoim (EMBRAPA, 2019). Na região nordeste, o amendoim tem sido cultivado em condições de sequeiro, o que expõe a produção a riscos climáticos. Entretanto, cultivares adaptados ao déficit hídrico tem mostrado boa adaptabilidade a estas condições (Araújo & Ferreira, 1997).

O amendoim é a quarta oleaginosa mais plantada no mundo e pode ser usada tanto na alimentação quanto na extração de óleos, sendo o amendoim uma alternativa para a obtenção de biodiesel (PIGHINELLI, 2007). Seu uso de forma alternativa tem mais de 100 anos e foi o criador do motor a Diesel Rudolph Diesel que usou o óleo como combustível alternativo em relação aos combustíveis fósseis (RODRIGUES et al., 2016).

2.2 Agricultura de precisão

A agricultura de precisão é uma forma de gerenciamento agrícola que usa informações exatas para a gestão de uma propriedade de forma específica para cada área levando em consideração as diferenças em cada área específica de uma propriedade. (ROZA, 2000). As tecnologias utilizadas em Agricultura de Precisão (AP) dão suporte a gestão das lavouras para monitorar a eficiência dos insumos e recursos, reduzindo o uso de produtos químicos para evitar danos ambientais e produzir produtos de alta qualidade para satisfazer a crescente demanda por alimentos (ONYANGO et al., 2021). Isso é possível devido a estratégia que reúne, processa e analisa dados temporais,

espaciais e individuais e os combina com outras informações para apoiar as decisões de gestão de acordo com a variabilidade estimada (ISPA, 2019). As principais tecnologias que existem em AP e são utilizadas pelos produtores, incluem, desde o sistema de posicionamento global (GNSS), sistemas de informações geográficas (GIS) e uma infinidade de sensores diferentes para avaliar a variabilidade local e da cultura, fornecendo informações para auxiliar os produtores no gerenciamento mais preciso do sistema agrícola (PATHAK et al., 2019).

Sem dúvidas a adoção de tecnologias de AP é uma decisão que, apesar dos seus benefícios, ainda gera discussão entre a maioria dos produtores, uma vez que esses sabem dos benefícios da adoção, mas criam resistência quanto ao custo de adoção (KERNECKER et al., 2020; LOWENBERG-DEBOER et al., 2019). Além disso, a complexidade para o uso na prática das tecnologia é um fator que diminui a taxa de adoção dos agricultores (PATHAK et al., 2019; LEE et al., 2021), desafio que pode ser superado com a substituição de jovens no comando da maioria dos campos de produção e com o aumento de treinamento específico para utilização de sistemas e sensores na AP (MISAKI et al. 2018).

Do ponto de vista prático, o manejo com aplicações localizadas é um dos principais impulsionadores da AP, de modo que as propriedades do solo, vegetação e mapas de produtividades são combinadas para delinear zonas de manejo homogêneas (ZM) e então aplicar o manejo específico localizado (SHADDAD et al., 2016). A utilização de sensores na AP é importante, e são esses que fazem com que informações sejam obtidas de forma rápida e não destrutiva otimizando a gestão das ZM. Dentre os principais tipos de sensores, destaca-se os sensores proximais, os quais são capazes de coletar dados a uma distância inferior a 2 m do alvo, que pode ser o solo ou dossel das plantas (VISCARRA E ADAMCHUK, 2011).

Os sensores proximais tem sido usado largamente para coleta de dados espacializados e entendimento da variabilidade, principalmente dados de condutividade elétrica do solo, que pode ser utilizado para o manejo da irrigação em taxa variada (SERRANO et al 2020), no estabelecimento de relações entre a produtividade e a quantidade de nitrogênio disponível no solo (CICORE et al., 2019). Ainda no contexto de aplicação em taxa variada, diversas são as aplicações para distribuição de fertilizantes, principalmente as que utilizam sensoriamento on-the-go para aplicação de

nitrogênio (HEIß et al., 2022), o qual faz a aplicação da quantidade de fertilizante de acordo com a leitura do sensor proximal de dossel.

Entretanto, um dos grandes entraves na Agricultura de Precisão é dada pela dificuldade de compreender as causas que provocam variabilidade dos atributos de solo e das plantas, falta ainda pesquisas relacionadas a recomendações agronômicas (curvas de resposta) específicas para aplicação em taxa variável, existe também o alto custo do mapeamento do solo, necessidade de aprimoramento das técnicas ligadas a sensoriamento remoto, falta de procedimentos padronizados e principalmente a falta de mão de obra especializada (BASSOI et al., 2020).

2.3 Sensoriamento remoto

O sensoriamento remoto pode ser definido como uma forma de obter informações sobre um determinado objeto sem que ocorra interação física com o mesmo (ROSA, 2007). O uso de técnicas de sensoriamento remoto na caracterização da natureza como recurso tem ficado bastante sofisticado de tal forma que permite a extração de informações quantitativas a depender da caracterização espectral de um objeto (GAIDA et al., 2020). Em 1970 teve início o programa “landsat” com o objetivo de mapear os recursos naturais terrestres, sendo esse um dos programas mais importantes desenvolvidos pela NASA (National Aeronautics and Space Administration) (GUEDES; SILVA, 2018). No Brasil, o Instituto nacional de pesquisas espaciais (INPE) é responsável pelo desenvolvimento de geotecnologias e faz do país um destaque no hemisfério sul. Em território brasileiro o uso de geotecnologias foram se intensificando com a necessidade de aumento de produtividade na produção de alimentos (SANTOS, 2019).

O Sensoriamento remoto tem como principal característica a capacidade de adquirir informações de um alvo, por meio de sensores, sem que ocorra contato direto. Essas informações são obtidas a partir da reflexão e emissão de energia eletromagnética de cada alvo (CREPANI, 1993; SANTOS, 2019). Guedes e Silva (2018) classificam os sensores como ativos e passivos. Essa classificação é dada em função da fonte de radiação necessária para seu funcionamento. Os sensores que dependem de uma fonte de radiação externa para operar são os chamados sensores passivos, esses sensores usam

via de regra a radiação solar para poder operar. Já os sensores ativos operam com sua própria fonte de energia eletromagnética.

O uso da reflectância espectral de uma cultura pode fornecer de maneira não destrutiva e rápida informações sobre características agronômicas, como acúmulo de biomassa e rendimento da cultura submetidas a estresses abióticos (KOVAR et al., 2019; SUN et al., 2018). Assim, uma redução na eficiência do uso da radiação por parte da planta implica que menos energia é absorvida para ser convertido em compostos orgânicos (SUN et al., 2018).

A luz que é absorvida e refletida carrega amplas informações sobre uma planta. Informações essas que dizem respeito a sua composição bioquímica e condição fisiológica da planta (SOLOVCHENKO et al., 2021). Assim o elo que une os objetos da superfície terrestre e os sensores remotos é a radiação eletromagnética. Essa radiação interage de forma distinta com diferentes matérias dessa forma, rochas, solos, vegetação e água tem comportamentos espectrais distintos (GUEDES; SILVA, 2018).

Nas plantas o perfil de luz absorvidas e refletidas pelas folhas se difere em comprimento de ondas em função da composição das folhas, a clorofila, carotenoides e antocianinas absorvem a luz do visível entre 400 e 700 nm já a água responde nas proximidades do infravermelho próximo que é entre 700 e 1100 nm (MEIRELES et al., 2020). Dessa forma, as características espectrais da radiação que é refletida pela folha podem fornecer respostas fisiológicas nas plantas com algum tipo de estresse (MARÍN-ORTIZ et al., 2020). Ainda Marín-Ortiz et al. (2020) cita que mudanças na concentração foliar de clorofila total são indicadores de estresse foliar e é ligado diretamente a redução de taxas de metabolismo central.

2.4 Índices de vegetação

Índices de Vegetação são combinações de dados espectrais de duas ou mais bandas com o objetivo de sintetizar e realçar de dados da vegetação e parâmetros biofísicos das plantas (PONZONI, 2001). Os índices de vegetação provindos de dados de reflectância espectral são largamente utilizados no sensoriamento remoto, pois conseguem fazer uma representação bastante aproximada da estrutura do dossel, conteúdo de pigmentos das folhas e potencial fotossintético da cultura (CAMPS-VALLS et al., 2021). Camps-Valls et al. (2021) explica que os índices de vegetação

tratam de transformações paramétricas de bandas espectrais projetadas para demonstrar as variações em fenômenos biofísicos específicos bem como a maximização de fenômenos mais sensíveis e reduzir a sensibilidade a fatores como propriedade do solo, iluminação solar e fenômenos atmosféricos.

O conhecimento sobre os índices de vegetação tem se destacado em virtude de sua relevância, uma vez que trabalham como indicadores para análises da cobertura da vegetação em períodos diferentes utilizando técnicas de sensoriamento remoto. O uso desses índices permite um monitoramento da dinâmica espaço-temporal de uma determinada atividade, além da análise de crescimento da vegetação e análise das mudanças da cobertura do solo causada por ações antrópicas (BARROS; FARIAS; MARINHO, 2020). Dessa forma os índices de vegetação são usados como ferramentas para monitorar e gerenciar as dinâmicas vegetativa, uma vez que podem detectar alterações na cobertura vegetal de forma relativamente barata e rápida (LEITE; SANTOS e SANTOS, 2017).

Existem uma variedade de índices de vegetação tal como o Índice de Vegetação por Diferença Normalizada (NDVI) usado para medir a condição da vegetação; Índice de Vegetação ajustada aos Solo (SAVI) indicado em áreas com grande exposição do solo e usado em conjunto com o NDVI; Índice de Diferença Normalizada da Água (NDWI); Índice de Vegetação Melhorado (EVI) semelhante ao NDVI porém é mais usado em vegetações mais altas em função do uso da banda do azul para descontar influencias atmosféricas, e entre outros e cada um com uma aplicabilidade distinta. Janssen et al. (2018) defendem que o Índice normalizado de vegetação (NDVI) é um dos mais usados nos mapeamentos da cobertura vegetal em função da sua alta sensibilidade em relação a vegetação mais rala e a mais densa.

Através de um método de normalização os valores de NDVI ficam entre 0 e 1 (XUE; SU, 2017). Valores abaixo de 0 no índice de vegetação indicam a resposta da água, valores positivos que atingem até 0,2 indicam terras nuas e áreas edificadas, quando os valores estão acima de 0,2 já é indicio de vegetação e áreas com valores mais alto representa culturas com intenso vigor vegetativo (GUHA; GOVIL, 2020). Dessa forma o NDVI é o índice de vegetação mais usado em função gerar uma forte correlação entre fotossíntese e metabolismo central da vegetação (WILSON; NORMAN, 2018).

2.5 Estimativa de produtividade

Estimativas de produtividade podem ser caracterizadas por meio de modelos matemáticos de monitoramento de condições climáticas. Esses modelos levam em consideração que os elementos climáticos exercem certa contribuição na produtividade de uma cultura (SILVA et al., 2011). Medir com precisão a produtividade é um desafio para os sistemas agrícola uma vez que as bases para medição são principalmente auto-relatos de agricultores ou cortes com o objetivo de colheitas para medir o rendimento da lavoura (LOBELL et al., 2019). Ainda Lobell et al. (2019) menciona que quando a estimativa é auto relatada ocorrem erros provindos de usos incorretos de unidades de medida ou a extrapolação de rendimentos com razão da produção por área relatada além da falta de padronização da colheita. A outra forma de estimativa de produtividade é a coleta e corte das culturas no final da estação de crescimento, os grãos são pesados e quantificados sendo esse o padrão-ouro para estimar produtividade (PALIWAL; JAIN, 2020).

Modelos baseados em variáveis climáticas são um dos vários métodos que existem para tentar estimar o rendimento da cultura. Esses métodos são em geral uma abordagem estatística na qual se usa regressões lineares com o intuito de relacionar a produtividade com vários parâmetros climáticos adquiridos em vários locais e datas do ano (JOSHI et al., 2020). Joshi et al. (2020) menciona que foram desenvolvidas várias tentativas com o intuito de gerar modelos de estimativas de produtividade, modelos usando variáveis climáticas e classificação do solo, porém variações nas condições climáticas e a indisponibilidade de dados climáticos tornam essa tarefa complicada.

O sensoriamento remoto pode ser usado para avaliar as condições de crescimento da lavoura e gerar mapas de distribuição e rendimento da cultura, sendo possível realizar um monitoramento contínuo durante o seu período de crescimento e desenvolvimento (ZHOU et al., 2020). O uso de imagens de satélites levou nos últimos anos um rápido desenvolvimento de estimativas de produtividade usando índices de vegetação uma vez que esses índices são sensíveis a cultura, índices como NDVI e EVI descrevem o status de crescimento da cultura e estimar a produtividade da lavoura (HAN et al., 2021).

Alguns estudos sobre produtividade mostram que existe relação entre as curvas de séries temporais de índices de vegetação como o NDVI e o crescimento da cultura,

encontrando resultados satisfatórios que correlacionam índices de vegetação com a produtividade de uma dada cultura (ZHANG et al., 2022).

A estimativa de produtividade depende de vários atributos diferentes como clima, solo, topografia e entre outros, porém o uso de técnicas de sensoriamento remoto fornece informações espectrais e espaço-temporais que podem auxiliar na avaliação do rendimento da lavoura (CHLINGARYAN; SUKKARIEH; WHELAN, 2018). Dessa forma o grande desafio encontrado na estimativa de produtividade e gerar modelos que faça produtividade em escala de campo e que seja desenvolvido de forma que atenda uma localidade e possa ser calibrado para outras regiões (DONOHUE et al., 2018).

2.6 Inteligência artificial

A inteligência artificial é uma tecnologia computacional que simula o processo humano de aprendizagem, sendo ela uma alternativa a modelagem estatística uma vez que lidam de forma eficiente com problemas complexos e não lineares (KIM et al., 2019). Kim et al. (2019) ainda comenta que os rendimentos das culturas são estimados por processos não lineares e dessa forma o uso de inteligência artificial pode ser uma abordagem mais adequada para esses problemas.

O aprendizado de máquina é um campo da inteligência artificial focada na abordagem prática preditiva, com base em vários outros dados (VAN KLOMPENBURG; KASSAHUN; CATAL, (2020). Van Klompenburg; Kassahun; Catal, (2020) ainda cita que o aprendizado de máquina pode determinar padrões e correlações de um dado conjunto de dados por meio de modelos devidamente treinados e testados.

O aprendizado de máquina é usado na agricultura como suporte para tomada de decisão nas etapas de crescimento e desenvolvimento da cultura, além de existirem algoritmos que trabalham no suporte para previsão de safra (VAN KLOMPENBURG; KASSAHUN; CATAL, 2020). Abordagens utilizando inteligência artificial no campo da agricultura tem sido usada com as mais distintas finalidades, como previsão de safra, definição de propriedades do solo, investigação de salinidade, produtividade algodoeira, umidade do solo, potencial hídrico entre outros (KOUADIO et al., 2018).

As tecnologias usando Inteligência Artificial otimizam vários campos do conhecimento e também ajudam a resolver problemas dos mais variados. Com a

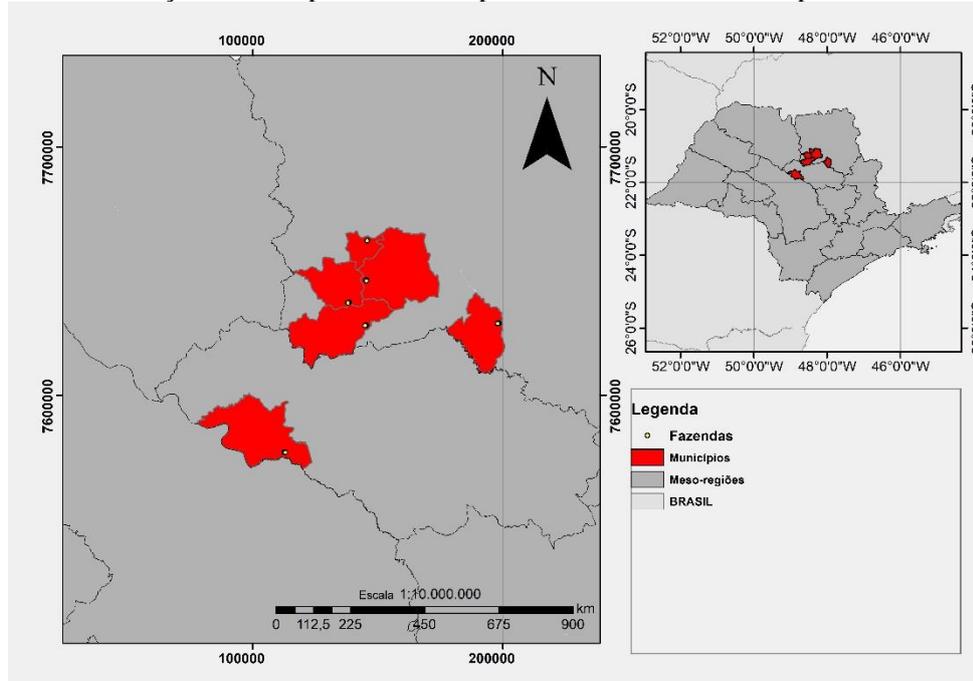
população em crescente aumento será necessário uso de tais ferramentas para otimizar as produções com menos recursos. (TALAVIYA et al., 2020).

3. MATERIAL E MÉTODOS

3.1 Áreas de estudo

O trabalho foi conduzido em seis áreas de produção no interior do estado de São Paulo, Brasil, localizados nos municípios de Jaboticabal, Taquaritinga, Ibitinga, Monte Alto, Taiúva, Guataporá (Figura 1). Em todas as áreas foi semeado, após a colheita da cana-de-açúcar a cultivar IAC OL3 (tipo runner), de ciclo longo (125-130 dias). O espaçamento de semeadura adotado foi de 0,90 m entre linhas para todas as áreas.

Figura 1 - Localização dos campos utilizados para coleta de amostras de produtividade



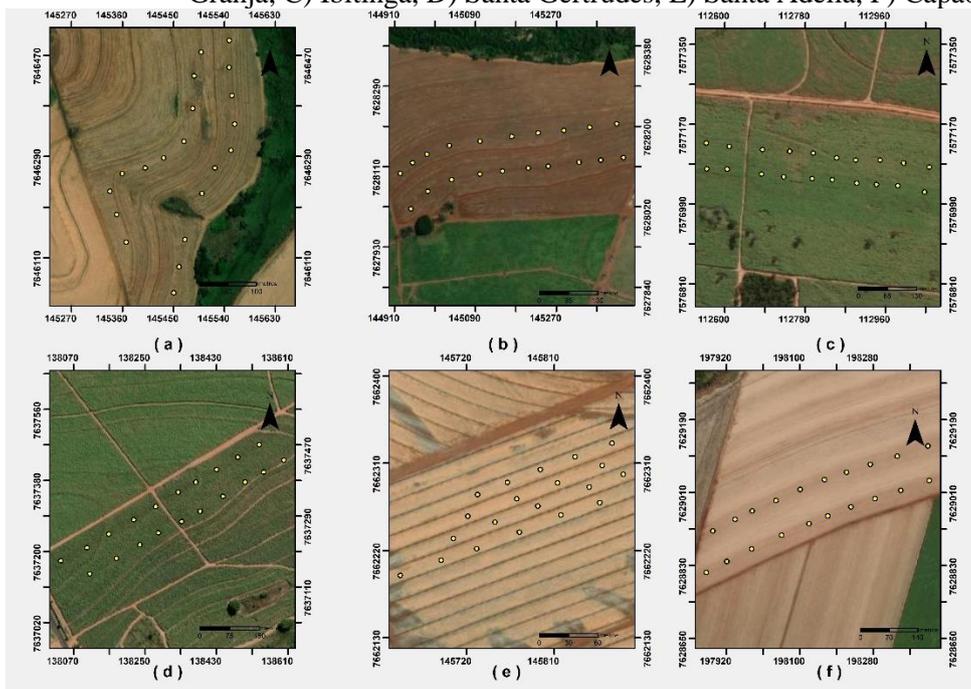
3.2 Amostras de campo

Em cada área foi georreferenciados 20 pontos amostrais espaçados entre si a cada 50 metros (Figura 2). No monitoramento utilizou-se imagens de satélite e iniciou após os 90 DAS e foram realizadas semanalmente até a colheita (arranquio) das áreas experimentais em média 120 DAS.

A determinação da produtividade foi realizada pelas plantas presentes em área amostral de 2 m² quadrados de cada ponto. Todas as plantas dentro da área da armação

foram arrancadas, ensacadas e identificadas para então calcular a produtividade bruta em cada ponto amostral. Após separação das vagens das plantas, essas foram pesadas e posteriormente colocadas em estufa de circulação de ar forçada a 65°C por 72h. Após esse período, procedeu-se com a correção do valor de umidade para 8% e extrapolando os valores para kg ha⁻¹.

Figura 2 - Distribuição dos pontos amostrais de produtividade em A) Frutal, B) Granja, C) Ibitinga, D) Santa Gertrudes, E) Santa Adélia, F) Capão.



Fonte: Do autor (2022)

3.3 Aquisição e processamento das Imagens de satélite

Para aquisição de dados espectrais da cultura de amendoim, imagens de satélite da plataforma Sentinel-2A foram baixadas coincidindo com as datas para o monitoramento. O Sentinel-2 é um satélite europeu multiespectral, que adquire imagens em 13 bandas espectrais, como visível (banda 2–4), Red Edge (RE, banda 5–7), Near Infrared (NIR, banda 8) e Bandas de infravermelho de ondas curtas (SWIR, banda 11–12) com resoluções espaciais de 10 m, 20 m, 10 m e 20 m, respectivamente e tempo de revisita de aproximadamente 5 dias. Para este trabalho foram utilizadas imagens livres de nuvens no período de 27 de janeiro de 2021 ao período de 9 de março de 2021, período que compreende as épocas de monitoramento semanalmente após os 95 DAS em solo, totalizando 6 imagens.

As imagens Sentinel-2 disponíveis no GEE (ID da coleção de imagens: COPERNICUS/S2) foram usadas neste estudo e todas as operações de correção e extração de valores de reflectância foram utilizando a plataforma GEE. As imagens foram convertidas para valores de Refletância pro topo da atmosfera (TOA) escalados em 10.000 por meio de calibração radiométrica (Sentinel-2User, 2022).

Após o pré-processamento das imagens, índices de vegetação (Tabela 1) foram calculados. Para cada área, um arquivo no formato “shape” foi importado no GEE para que os valores médios de índices de vegetação fossem extraídos em cada uma das datas.

Tabela1 - Índices de vegetação utilizados

Índice	Equação	Referencia
Normalized Difference Vegetation Index (NDVI)	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$	Rouse et al. (1974)
Normalized Difference Water Index (NDWI)	$(\text{NIR} - \text{SWIR}) / (\text{NIR} + \text{SWIR})$	(Gao, 1996)
Green Normalized Difference Vegetation Index (GNDVI)	$(\text{NIR} - \text{Green}) / (\text{NIR} + \text{Green})$	Gitelson et al. (1996)
Simple Ratio (SR)	NIR / Red	Jordan (1969)
Non-Linear Index (NLI)	$(\text{NIR}^2 - \text{Red}) / (\text{NIR}^2 + \text{Red})$	Goel and Qin (1994)

Fonte: Santos (2019)

3. 4 Análise dos dados

Como método de estimativa de produtividade, foram utilizadas Redes Neurais Artificiais (RNA) do tipo perceptron de multicamadas (MLP). RNA's são sistemas que consistem de neurônios artificiais, os quais desempenham funções similares aos neurônios humanos (Jiang et al. 2004), que sem pressuposições sobre a distribuição dos dados podem aprender relações entre as variáveis apresentadas como exemplo (Zhang et al. 2008). Uma vantagem da utilização de RNA's é que este método leva em consideração não linearidades na relação entre as informações de entrada e saída (Savegnago et al. 2011). A estrutura básica de rede neural do tipo multicamadas é: uma camada de entrada, duas camadas ocultas e uma camada de saída. O número de camadas ocultas pode variar em função do objetivo do problema que se pretende resolver e dos recursos computacionais disponíveis.

Em uma análise de RNA, a escolha de variáveis, arquitetura de rede neural, algoritmo de treinamento e parâmetros associados são operações críticas e demandam tempo (StatSoft 2013). A abordagem de tentativa e erro comumente adotada consome muito tempo e é impraticável para se fazer de forma manual; para isso a ferramenta

Intelligent Problem Solver (IPS), uma ferramenta embutida no software Statistica 7 que envolve a combinação estratégias heurísticas e sofisticadas de otimização (técnica de minima-becketing e algoritmo simulado de annealing) (Miao et al. 2006). Estes procedimentos foram usados para automaticamente testar mil modelos de diferentes combinações de preditores, arquitetura de redes e parâmetros associados.

A camada de entrada da rede foi constituída pelos índices de vegetação (Tabela 1) e a saída pelo Produtividade em cada ponto. Foi variado o número de neurônios (1 a 20) para as redes MLP e de 1 a 45 para redes com Função de Ativação de Base Radial (RBF), a quantidade de camadas ocultas (1 - 3) para MLP e as entradas das redes. Para cada combinação de entrada foram treinados mil modelos. As redes do tipo multicamadas foram treinadas utilizando o algoritmo de aprendizado supervisionado de retro propagação dos erros (Bryson et al. 1979). A rede neural de retropropagação é treinada com as entradas ajustadas às variáveis de saída em duas fases (Zhang et al. 2008).

Rede MLP

As redes MLP foram interconectadas por forças de conexão representadas por valores que são chamados de pesos sinápticos, os quais são responsáveis pelo armazenamento do conhecimento adquirido. Os valores usados nas camadas de entrada foram normalizados de acordo com a equação 2.

$$y_i = \frac{x_i + x_{max}}{x_{max} - x_{min}} \quad (2)$$

Em que, y_i é o valor do vetor de entrada (p. ex.: índices de vegetação modificados), x_{min} e x_{max} são o mínimo e o máximo valores observados, respectivamente.

O valor de saída de cada neurônio na camada k é expresso por $z_k = g(a_k)$ em que g é a função de ativação de a_k e a_k é a função sináptica, a qual é uma combinação linear dos valores de entrada normalizada e os pesos sinápticos (Equação 3).

$$a_k = \sum_j y_j w_{kj} \quad (3)$$

em que, w_{kj} são os pesos sinápticos ligando os y_j valores de entrada com cada k neurônio.

A transferência ou função de ativação nos neurônios de cada camada oculta foi a função hiperbólica, sendo 'e' o número neperiano (Equação 4).

$$g(a_k) = \frac{e^{ak} - e^{-ak}}{e^{ak} + e^{-ak}} \quad (4)$$

Rede RBF

As redes RBF possuem apenas uma camada oculta e cada neurônio contém uma função de ativação de base radial.

Em cada neurônio a função Gaussiana ou normal foi usada como função de base radial (Bishop 1995), e o valores de distância (desvio) dessa função aumentam ou diminuem a relação para o ponto central (Haykin 1998). A função gaussiana é apresentada pela equação 5:

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

Em que, $v = \|x - \mu\|$ é a distância Euclidiana entre o vetor de entrada e o centro μ da função Gaussiana e σ é a largura. A distância Euclidiana do vetor de entrada para o centro μ é a entrada para função Gaussiana, a qual fornece o valor de ativação da unidade radial.

Assim como na rede neural MLP, os valores de entrada foram normalizados pela equação 2, e os valores na camada de saída forneciam o a produtividade em cada ponto.

A redes neurais RBF foram treinadas pelo algoritmo k-means (Bishop, 1995). Esse algoritmo tenta selecionar o conjunto ótimo de pontos que são colocados nos centroides nos padrões dos dados de treinamento.

3.5 Banco de dados para treinamento e validação

Para o treinamento e validação dos modelos, o banco de dados foi dividido em 80% para o treinamento e 20% para validação. Devido às informações terem sido coletadas em diferentes estádios fenológicos, esta percentagem de partição do banco de dados de treinamento e validação foi aplicada a cada estágio de coleta das informações.

Os procedimentos de treinamento e validação dos modelos neurais foram implementados no pacote Neural Networks do software de análise de dados Statistica (Statistica 7.0, Statsoft Inc., Tulsa, OK).

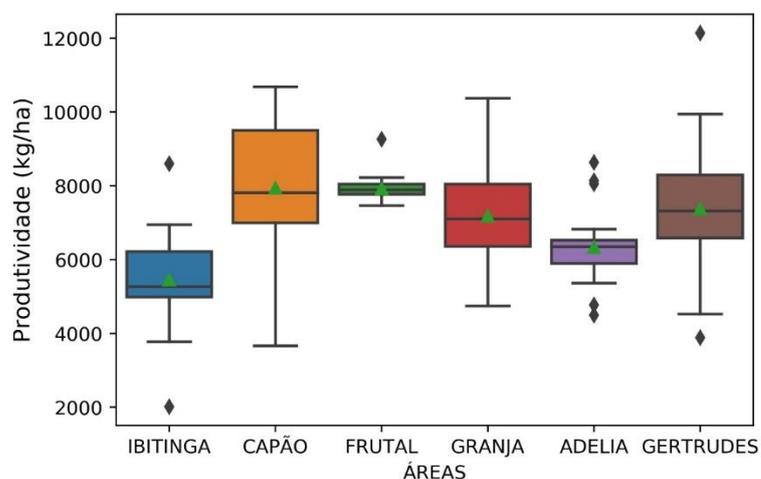
3.6 Avaliação do desempenho dos modelos

As RNA's foram avaliadas em termos de acurácia onde foi avaliado os menores valores de erros para cada modelo. Após a padronização e o pré-processamento dos dados foi utilizado o software (Statistica 7.0, Statsoft Inc., Tulsa, OK) para gerar os modelos. Para melhor representação dos resultados foram definidas 3 classes de acurácias. Até 100 kg ha⁻¹ Alta acurácia; entre 100 e 150 kg ha⁻¹ Media acurácia e maior que 150 kg ha⁻¹ Baixa acurácia. Foram monitoradas 6 datas para cada índice e foi gerado modelos RBF e MLP para cada data. Foram selecionados os modelos com mais alta acurácia e apresentado nos gráficos. Em cada índice foram apresentados 2 modelos, um RBF e um MLP para cada data totalizando 12 modelos para cada índice.

4. RESULTADOS E DISCUSSÃO

Os resultados foram divididos em 2 grupos: um modelo geral para cada índice, dividido nas duas redes (MLP e RBF) e outros modelos separado por datas também divididos em duas redes. O primeiro gráfico é referente a produtividade das 5 áreas do trabalho. É interessante notar a variabilidade da produtividade em cada área como pode ser visto na figura 4.

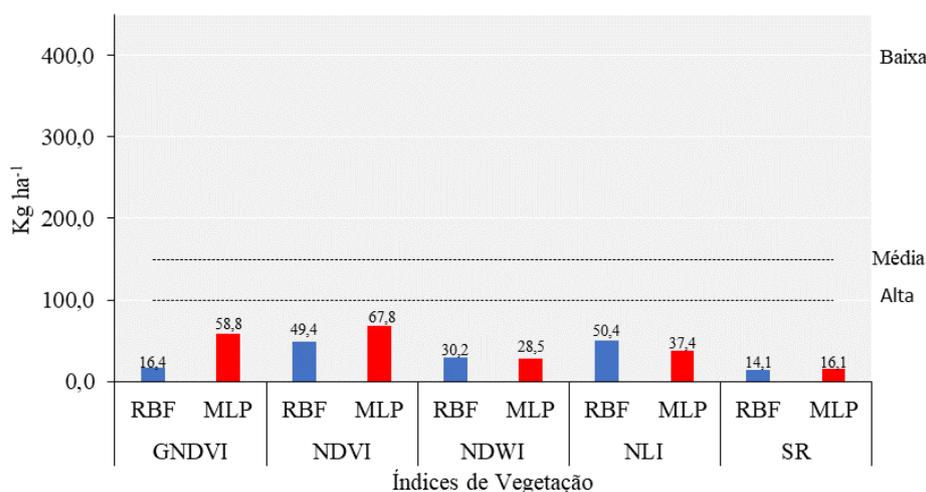
Figura 4 - Gráfico de Boxplot das áreas trabalhadas



Fonte: Do autor (2022)

E visto que as linhas no meio de cada caixa representam a mediana estatística dos dados. Dessa forma a linha divide em 50% dos dados para cima e para baixo. Desconsiderando os outliers, Frutal, Adelia e Ibitinga tiveram as menores variações enquanto Capão e Gertrudes tiveram respectivamente 3800 a 11000 kg ha⁻¹ e 4500 e 10000 kg ha⁻¹ de variação. Essa variabilidade de produtividade pode ser a causa dos resultados encontrados nos modelos de redes neurais uma vez que os dados são muito variáveis. Os resultados dos modelos gerais para cada índice podem ser vistos na figura 5.

Figura 5 - erro em kg ha⁻¹ e relativo aos modelos gerais por Índice de Vegetação



Fonte: Do autor (2022)

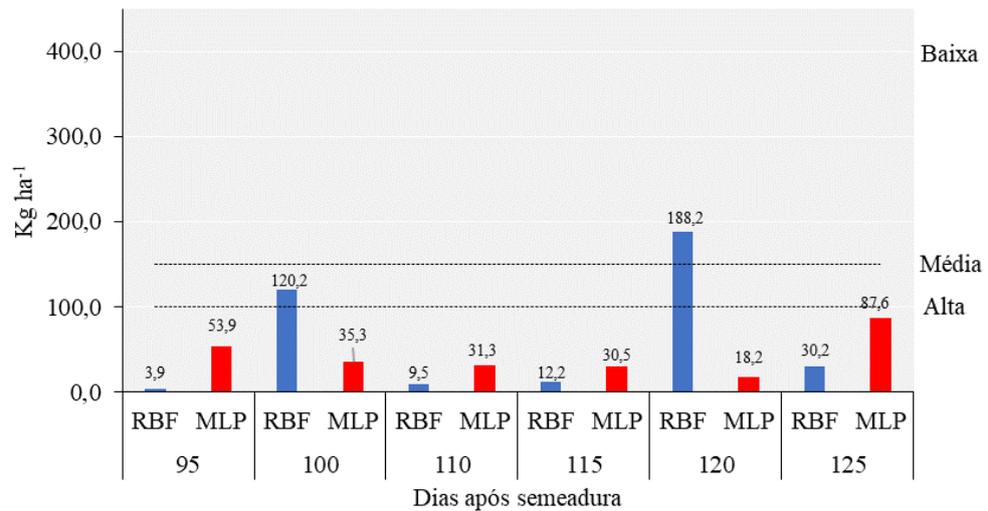
Nos modelos gerais, todos os índices tiveram acurácia ótima, o índice que obteve o menor valor de erro foi SR para ambas as redes. As redes RBF obtiveram um menor erro para os modelos gerais de produtividade. Mesmo levando em consideração o maior erro médio de produtividade que foi a rede MLP para o índice de NDVI o erro foi apenas de 1%. Ainda é necessário fazer testes estatísticos para validar os modelos obtidos, mas esses resultados mostram que o uso de redes neurais pode ser uma ferramenta bastante poderosa para previsão de safra obtendo uma acurácia entre 0,2 e 1,1% do valor médio total produzido.

Os modelos separados por datas tiveram um comportamento semelhante aos modelos gerais, contudo o desempenho dos modelos RBF não seguiu a mesma linha de valores obtidas. Isso pode ser explicado pela própria variabilidade dos dados, uma vez que os modelos gerais foram usados todos os dados para cada índice, sendo esses dados de áreas distintas como mostra as análises do gráfico de caixa (boxplot). O ideal para

ter mais rigor seria agrupar por áreas, mas a ideia desse trabalho é gerar um modelo mais abrangente que pudesse explicar os resultados em áreas diferentes, a fim de gerar modelos que se apliquem em qualquer condição de manejo e características do ambiente.

Os modelos por datas foram divididos de forma temporal levando em consideração 95, 100, 110, 115, 120 e 125 dias após a semeadura. Para cada data foram estimados valores de erros para cada modelo levando em consideração todos os índices de vegetação. Nos gráficos, o tamanho da circunferência representa o valor do erro médio absoluto (MAE) bem como o valor indicado no rótulo da circunferência. O eixo X representa a data após a semeadura e o eixo y representa a média dos erros. A representação dos valores de acurácia segue as classes definidas Alta, Média e Baixa previamente definidas. Na figura 6 temos o primeiro gráfico representando o GNDVI.

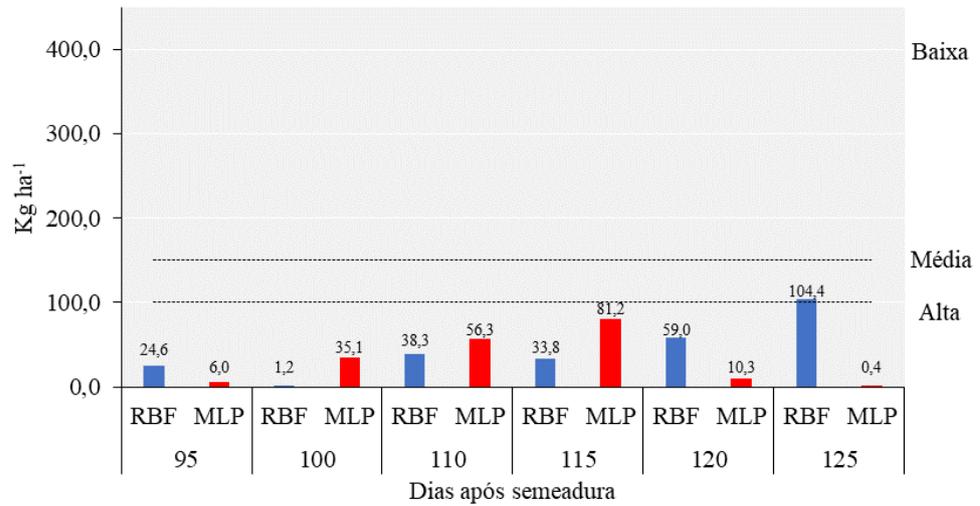
Figura 6 - Gráficos de GNDVI para os modelos MLP e RBF entre 95 e 125 dias após a semeadura.



Fonte: Do autor (2022)

Foi observado que os modelos para o GNDVI tiveram 10 modelos com acurácia Alta, 1 modelos com acurácia Média e 1 com acurácia Baixa. Sendo o melhor modelo usando a rede RBF para o DAS 95 com apenas 3,9 kg ha⁻¹ de erro absoluto e o pior modelo também sendo o RBF para o DAS 120 com um erro absoluto de 188,2 kg ha⁻¹. A figura 7 mostra os gráficos para NDVI.

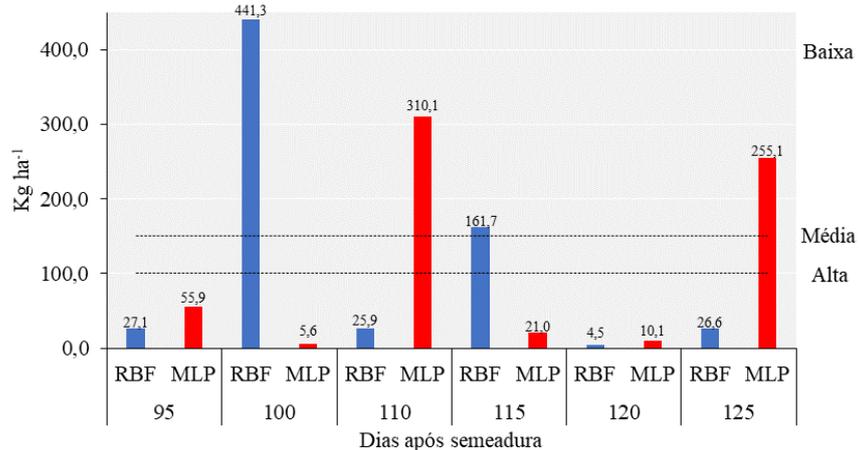
Figura 7. Gráficos de NDVI para os modelos MLP e RBF entre 95 e 125 dias após a semeadura.



Fonte: Do autor (2022)

Os modelos para NDVI por data tiveram melhores resultados com 11 modelos de acurácia Alta e 1 modelos com acurácia Média. Sendo o melhor modelo usando a rede MLP para o 125 DAS com apenas 0,4 kg ha⁻¹ de erro absoluto e o pior modelo sendo o RBF para o 125 DAS com um erro absoluto de 104,4 kg ha⁻¹. A figura 8 mostra os gráficos para NDWI.

Figura 8 - Gráficos de NDWI para os modelos MLP e RBF entre 95 e 125 dias após a semeadura.



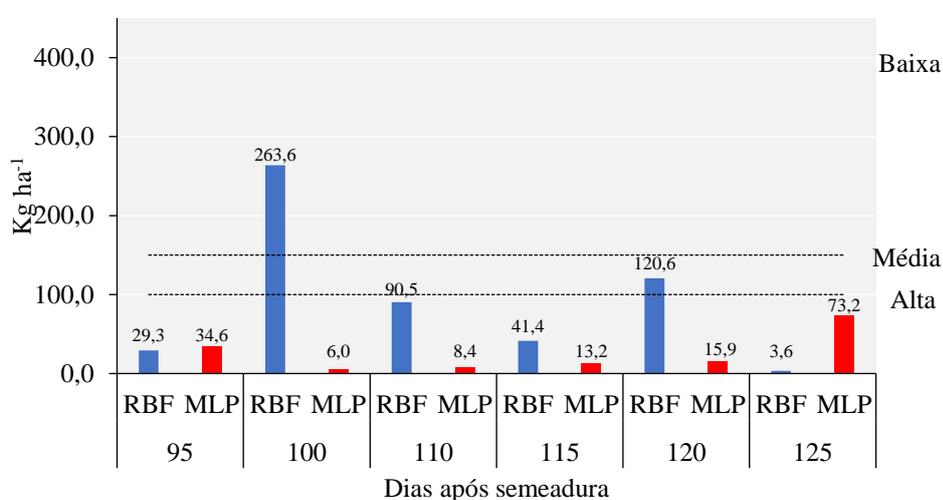
Fonte: Do autor (2022)

Nos modelos para NDWI por data tiveram resultados piores quando comparado com os índices passados. foram 8 modelos de acurácia Alta e 4 modelos com Baixa acurácia. Sendo o melhor modelo usando a rede RBF para o 120 DAS com apenas 4,5

kg ha⁻¹ de erro absoluto e o pior modelo sendo o RBF para o 100 DAS com um erro absoluto de 441,3 kg ha⁻¹.

Nos modelos para NLI por data tiveram resultados mais parecidos quando comparado com os índices GNDI e NDVI. foram 10 modelos de acurácia Alta, 1 modelo com Média e 1 modelo com Baixa acurácia. Sendo o melhor modelo usando a rede RBF para o 125 DAS com apenas 3,6 kg ha⁻¹ de erro absoluto e o pior modelo sendo o MLP para o 100 DAS com um erro absoluto de 263,6 kg ha⁻¹. O gráfico pode ser observado na figura 9 abaixo.

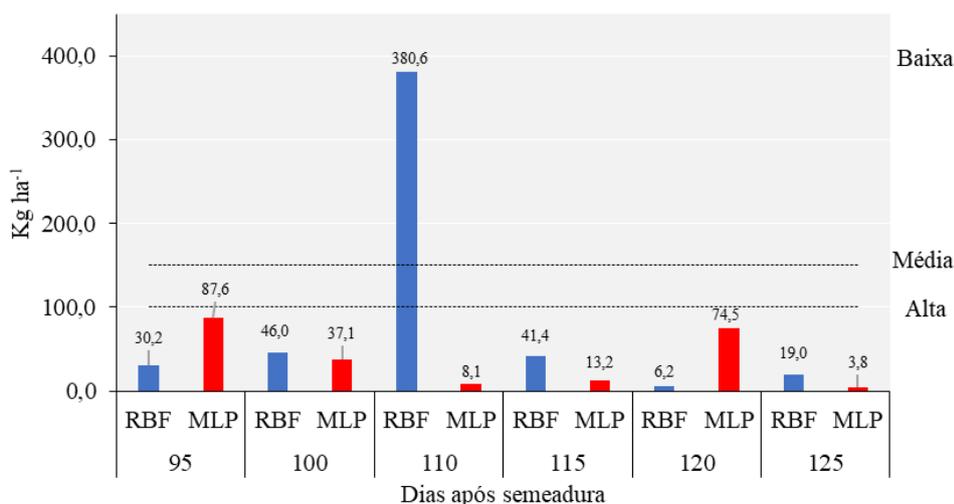
Figura 9 - Gráficos de NLI para os modelos MLP e RBF entre 95 e 125 dias após a semeadura.



Fonte: Do autor (2022)

Nos modelos para SR por data tiveram 11 modelos de acurácia Alta e um modelo com Baixa acurácia. Sendo o melhor modelo usando a rede RBF para o 125 DAS com apenas 3,8 kg ha⁻¹ de erro absoluto e o pior modelo sendo o MLP para o 110 DAS com um erro absoluto de 380,6 kg ha⁻¹ como pode ser visto na figura 10.

Figura 10 - Gráficos de SR para os modelos MLP e RBF entre 95 e 125 dias após a semeadura.



Fonte: Do autor (2022)

No geral todos os índices obtiveram modelos com acurácia Alta, ou seja, com o erro menor que 100 kg ha⁻¹. A tabela 2 mostra as datas de melhor acurácia para cada índice bem como a rede que foi usada para gerar o modelo.

Tabela 2. Melhores acurácias por Índice e data

Índice de Vegetação	DAS	Modelo	Acurácia em kg ha ⁻¹
GNDVI	95	RBF	3,9
NDVI	125	MLP	0,4
NDWI	100	RBF	4,5
NLI	125	MLP	3,6
SR	125	MLP	3,8

Fonte: Do autor (2022)

Dentre os melhores modelos temos o NDVI com o menor erro absoluto, NLI, SR e GNDI foram os melhores erros seguido do modelo de NDWI. Com a exceção do GNDVI, os melhores estimadores foram os modelos MLP e 3 deles estimaram melhores valores no 125 DAS.

Kayad et al. (2019) relata em seu trabalho sobre previsão de produtividade de milho que GNDVI foi o índice com maior correlação de produtividade de campo, em seu trabalho ele comparou vários índices diferentes e entre eles o NDVI. Em termos de comparação o NDVI se saiu com acurácia melhor quando comparado com GNDVI no presente trabalho, porem ambos conseguiram modelos com alta acurácia. Em outro

trabalho Tedesco et al. (2021) usa o GNDVI e o NDVI para prever o rendimento de lavoura de batata doce. Ele comenta sobre o fato de ambos os índices utilizarem a banda NIR do espectro invisível em sua composição. Essa banda está relacionada a estrutura celular da folha, dessa forma a diferença no desempenho das estimativas de produtividade pode ser relacionada a quantidade de luz que a planta reflete no espectro do visível. O NDVI usa a banda do vermelho que é mais absorvida pelas clorofilas. O GNDVI usa a banda do verde que é menos absorvida pelas folhas sendo mais sensíveis a alterações. A mesma lógica pode ser aplicada ao NLI e ao SR uma vez que ambos os índices possuem em sua composição apenas as bandas do vermelho e do NIR.

O pior índice monitorado por data foi o NDWI que teve 4 datas com acurácia acima de 150 kg ha⁻¹ para ambos os modelos. Sendo 3 para os modelos MLP e apenas 1 para o RBF tendo esse o maior erro entre todos os modelos. Isso pode ser explicado pela própria composição do índice. Guo et al. (2020) comenta em seu trabalho para estimativa de produtividade de milho que a banda mais sensível a clorofila é a banda do vermelho. Quando analisamos a composição do NDWI é visto que o vermelho não está presente, esse fato pode explicar por que o NDWI teve as piores acurácias. Liu et al. (2020) explica que os índices de vegetação são indicadores da capacidade fotossintética da cultura e é provinda da biomassa viva sobre o solo. A quantidade de biomassa verde acumulada até o estágio de pico de crescimento determina em grande parte o rendimento final da lavoura. O NDWI é um índice, porem ele é pouco relacionado a vegetação, dessa forma pode se explicar o porquê de o NDWI ter um desempenho pior quando comparado com os outros índices.

No presente trabalho tanto as redes MLP quanto as redes RBF obtiveram resultados interessantes, com acurácia alta. Porem as redes RBF em todos os índices em pelo menos uma data, obtiveram valores de erro acima de 100 kg ha⁻¹. Para a otimização de produtividade de soja, Yoosefzadeh-Najafabadi, Tulpan e Eskandari (2021) obtiveram resultados satisfatório com as redes RBF, porém o seu experimento contou com 250 genótipos de soja de 2000 parcelas e 4 ambientes distintos. Nos resultados dos modelos para cada índice de vegetação, 3 dos 5 índices tiveram resultados melhores usando os Modelos RBF, isso pode ser explicado pela quantidade de dados usados para treinar cada rede, enquanto nos modelos gerais foram usados aproximadamente 470 dados, nos modelos por data foram usados em média 70 dados. Santos *et al.* (2021) constatou em seu trabalho que os modelos RBF superestimaram os valores de produção

em escalas menores enquanto em escalas maiores os valores são subestimados isso reforça os resultados encontrados no presente trabalho. Matematicamente as RBF's são interpoladores usados para resolver equações diferenciais paramétricas. Equações essas que modelam transferência de calor entre outros fenômenos naturais. Ahmad *et al.* (2020) fala em seu trabalho que as redes RBF's usam pontos vizinhos para estimar valores locais em uma matriz, isso pode explicar os resultados diferentes para os modelos RBF's no presente trabalho. Se foi usado valores vizinhos para interpolação dos dados e a rede possui uma base reduzida de dados de entrada, logo uma pequena variabilidade desses dados pode gerar um resultado insatisfatório no resultado final, porém se a base de dados de entrada for grande o suficiente, qualquer anomalia nos dados será compensada por valores dos dados vizinhos. Como existe a limitação da obtenção de dados uma alternativa seria o teste de outros algoritmos para tentar otimizar os resultados.

5. CONCLUSÃO

Os valores de erros de produtividade foram muito baixos, em alguns casos não chegando a 10 kg ha^{-1} de erro para alguns modelos. Nos modelos por índice o SR teve os melhores resultados tanto para os modelos RBF quanto para os MLP. Por data o SR e NDVI tiveram o maior número de modelos com alta acurácia totalizando 11 modelos e o resultado mais acurado foi usando o NDVI aos 125 DAS que obteve um valor de $0,4 \text{ kg ha}^{-1}$.

Os modelos obtiveram comportamentos distintos em função das diferentes quantidades de dados. Enquanto o MLP trabalhou bem com a quantidade reduzida de dados, o RBF gerou alguns modelos com acurácia baixa. Já no caso inverso, o RBF teve melhores resultados absolutos quando alimentado com um banco de dados maior.

Por fim podemos dizer que tanto os modelos RBF quanto MLP conseguem estimar a produtividade de uma lavoura usando dados de sensoriamento remoto com uma acurácia alta e vale ressaltar que esses resultados são interessantes, mas ainda são necessárias mais pesquisas para implementação de modelos mais confiáveis.

6. REFERÊNCIAS

AHMAD, I. et al. **Solution of Multi-Term Time-Fractional PDE Models Arising in Mathematical Biology and Physics by Local Meshless Method**. *Symmetry*, [S.L.], v. 12, n. 7, p. 1195, 19 jul. 2020. MDPI AG. <http://dx.doi.org/10.3390/sym12071195>. Disponível em: <https://www.mdpi.com/2073-8994/12/7/1195/htm>. Acesso em: 19 abr. 2022.

ANDREAS, D. et al. **Real-time control for multi-parametric data fusion and dynamic offset optimization in sensor-based variable rate nitrogen application**, *Computers and Electronics in Agriculture*, Volume 196, 2022,

ARAÚJO, W.F.; FERREIRA, L.G.R. Efeito do déficit hídrico durante diferentes estádios do amendoim. *Pesquisa Agropecuária Brasileira*, Brasília, v.32, n.5, p. 481-484, maio, 1997.

BARBOSA, R. M.; HOMEM, B. Fernandes M.; TARSITANO, M. A. A.; **Custo de produção e lucratividade da cultura do amendoim no município de Jaboticabal**, São Paulo. *Revista Ceres*, [S.L.], v. 61, n. 4, p. 475-481, ago. 2014. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/0034-737x201461040005>.

BARROS, A. S.; FARIAS, L. M.; MARINHO, J. L. A.; **Aplicação do Índice de Vegetação por Diferença Normalizada (NDVI) na Caracterização da Cobertura Vegetativa de Juazeiro Do Norte – CE**. *Revista Brasileira de Geografia Física*, [S.L.], v. 13, n. 6, p. 2885, 26 nov. 2020. *Revista Brasileira de Geografia Física*. <http://dx.doi.org/10.26848/rbgf.v13.6.p2885-2895>.

BASSOI, L. H. *et al.* **Agricultura de precisão e agricultura digital**. *Teccogs: Revista Digital de Tecnologias Cognitivas*, [S.L.], n. 20, p. 17-36, 26 maio 2020. Pontifical Catholic University of Sao Paulo (PUC-SP). <http://dx.doi.org/10.23925/1984-3585.2019i20p17-36>.

BATCHELOR, B.; WHIGHAM, K.; DEWITT, J.; (1997) **Precision agriculture: introduction to precision agriculture**. Iowa Cooperative Extension, 4p.

BINOTI, D. H. B., BINOTI, M. L. M. S., & LEITE, H.G. (2013). BR Patent No 13410-5. Brasil: **Instituto Nacional da Propriedade Industrial**.

CAMPS-VALLS, G. et al. **A unified vegetation index for quantifying the terrestrial biosphere**. *Science Advances*: s, [S.L.], v. 7, n. 9, p. 1-32, 26 fev. 2021. American Association for the Advancement of Science (AAAS). <http://dx.doi.org/10.1126/sciadv.abc7447>. Disponível em: <https://www.science.org/doi/full/10.1126/sciadv.abc7447>. Acesso em: 06 abr. 2022.

CHLINGARYAN, A.; SUKKARIEH, SALAH.; WHELAN, BRETT. **Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review**. *Computers And Electronics In Agriculture*, [S.L.], v. 151, p. 61-69, ago. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.compag.2018.05.012>. Disponível em:

https://www.sciencedirect.com/science/article/pii/S0168169917314710?casa_token=rrv5KOa3hawAAAAA:d0JyXDLlzCwicurCJkub2QeM4hMFzEO40bk8BxHgH_Om1J6YPEp8v5pvvmHybnHC1A9Vh9ThPyU. Acesso em: 8 abr. 2022.

CICORE, P.L. et al **Relationship between soil apparent electrical conductivity and forage yield in temperate pastures according to nitrogen availability and growing season.** *Crop. Pasture Sci.* 2019, 70, 908–916.

DONOHUE, R. J. et al. **Towards a national, remote-sensing-based model for predicting field-scale crop yield.** *Field Crops Research*, [S.L.], v. 227, p. 79-90, out. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.fcr.2018.08.005>. Disponível em: https://www.sciencedirect.com/science/article/pii/S0378429018302454?casa_token=mnIY7raq9oAAAAA:b4ygJZSDqTrwgnxNV1vMFp6nxoNZ7_-YiNsEhtcbZ-EO2v9yhP2_0q8rAUyLqEbD_pWjMTuA1c8. Acesso em: 8 abr. 2022

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária. **Estudo Mapeia Áreas De Produção De Amendoim No Brasil Para Prevenir A Doença Do Carvão.** Disponível em <https://www.embrapa.br/busca-de-noticias/-/noticia/40697528/estudo-mapeia-areas-de-producao-de-amendoim-do-brasil-para-prevenir-doenca-do-carvao>. Acesso em 22/03/2022.

FÁVERO, A.R. et al. **Study of the evolution of cultivated peanut through crossability studies among *A. ipaensis*, *A. duraanensis*, and *A. hypogaea*.** *Crop Science*. v. 46, p.1546-1555, 2006.

FERRARI, R. A.; OLIVEIRA, V. S.; SCABIO, A. **Biodiesel de soja: Taxa de conversão em ésteres etílicos, caracterização físico-químicas e consumo em geradores de energia.** *Química Nova*, v.18, p.19-23, 2005

GAO, B.C. **NDWI –A normalized difference water index for remote sensing of vegetation liquid water from space** *Remote Sens. Environ.*, 58 (3) (1996), pp. 257-266

GAIDA, W. *et al.* **Correção Atmosférica em Sensoriamento Remoto.** *Revista Brasileira de Geografia Física*, [s. l], v. 13, n. 1, p. 229-248, 1 mar. 2020. Disponível em: http://mtc-m21c.sid.inpe.br/col/sid.inpe.br/mtc-m21c/2020/05.08.12.48/doc/gaida_correcao.pdf. Acesso em: 6 abr. 2022.

GOEL, N.S. QIN, W. **Influences of canopy architecture on relationships between various vegetation indices and LAI and FPAR** *Remote Sens. Rev.*, 10 (4) (1994), pp. 309-347

GUEDES, J.C.F.; SILVA, S.M.P. **SENSORIAMENTO REMOTO NO ESTUDO DA VEGETAÇÃO: PRINCÍPIOS FÍSICOS, SENSORES E MÉTODOS.** *Acta Geografia*, Boa Vista, v. 12, n. 29, p. 127-144, ago. 2018. Disponível em: <https://revista.ufr.br/actageo/article/view/4001>. Acesso em: 06 abr. 2022.

GUHA, S.; GOVIL, H. **An assessment on the relationship between land surface temperature and normalized difference vegetation index.** *Environment, Development And Sustainability*, [S.L.], v. 23, n. 2, p. 1944-1963, 24 fev. 2020.

Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s10668-020-00657-6>. Disponível em: <https://link.springer.com/content/pdf/10.1007/s10668-020-00657-6.pdf>. Acesso em: 6 abr. 2022.

GREGORY, W.C. KRAPOVICKAS, A. GREGORY, M. P. **Structure, variation, evolution and classification in Arachis**. In: BUNTING, S. *Advances in Legume Science*. Kew: London, 1980. p. 469-481.

GUO, Y. et al. **Modified Red Blue Vegetation Index for Chlorophyll Estimation and Yield Prediction of Maize from Visible Images Captured by UAV**. *Sensors*, [S.L.], v. 20, n. 18, p. 5055, 5 set. 2020. MDPI AG. <http://dx.doi.org/10.3390/s20185055>. Disponível em: <https://www.mdpi.com/1424-8220/20/18/5055/htm>. Acesso em: 20 abr. 2022.

HAN, DONG *et al.* **Improving Wheat Yield Estimates by Integrating a Remotely Sensed Drought Monitoring Index Into the Simple Algorithm for Yield Estimate Model**. *Ieee Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, [S.L.], v. 14, p. 10383-10394, 2021. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/jstars.2021.3119398>. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9568740>. Acesso em: 8 abr. 2022.

ISPA (2019) <https://www.ispag.org/about/definicion> Acessado em 04/04/2022

JANSSEN, T. A. J. et al. **Extending the baseline of tropical dry forest loss in Ghana (1984-2015) reveals drivers of major deforestation inside a protected area**. *Biological Conservation* [online] 218. Disponível: <https://doi.org/10.1016/j.biocon.2017.12.004>. Acesso: 23 mar. 2022.

JORDAN, C.F. **Derivation of leaf-area index from quality of light on the forest floor** *Ecology*, 50 (4) (1969), pp. 663-666

JOSHI, V. R. *et al.* **In-season weather data provide reliable yield estimates of maize and soybean in the US central Corn Belt**. *International Journal Of Biometeorology*, [S.L.], v. 65, n. 4, p. 489-502, 21 nov. 2020. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s00484-020-02039-z>. Disponível em: <https://link.springer.com/content/pdf/10.1007/s00484-020-02039-z.pdf>. Acesso em: 8 abr. 2022.

Kernecker, M. et al. **Experience versus expectation: Farmers' perceptions of smart farming technologies for cropping systems across Europe**. *Precis. Agric.* 2020, 21, 34–50. [Google Scholar] [CrossRef]

KIM, N. *et al.* **A Comparison Between Major Artificial Intelligence Models for Crop Yield Prediction: case study of the midwestern united states, 2006-2015**. *Isprs International Journal Of Geo-Information*, [S.L.], v. 8, n. 5, p. 240, 21 maio 2019.

KOUADIO, L. *et al.* **Artificial intelligence approach for the prediction of Robusta coffee yield using soil fertility properties**. *Computers And Electronics In Agriculture*, [S.L.], v. 155, p. 324-338, dez. 2018. Elsevier BV. <http://dx.doi.org/10.1016/j.compag.2018.10.014>. Disponível em: https://www.sciencedirect.com/science/article/pii/S0168169918304733?casa_token=K

rBoRVCEVJwAAAAA:qrHApmDvmpJib5xJkMb-jCBHUr5A9gCgLWB9glYVcna1MRNtDqFZLLywPiQuWmme274S_1XGiac. Acesso em: 8 abr. 2022

LEE, C. L.; STRONG, R.; DOOLEY, K. E. **Analyzing Precision Agriculture Adoption across the Globe: A Systematic Review of Scholarship from 1999–2020.** Sustainability 2021, 13, 10295. <https://doi.org/10.3390/su131810295>

LEITE, A. P., SANTOS, G. R., SANTOS, J. É. O., 2017. **Análise Temporal dos Índices de Vegetação NDVI E SAVI na Estação Experimental de Itatinga Utilizando Imagens Landsat 8.** Revista Brasileira de Energias Renováveis [online] 6. Disponível: <http://dx.doi.org/10.5380/rber.v6i4.45830>. Acesso: 23 mar. 2022.

LIU, J. et al. **Crop Yield Estimation in the Canadian Prairies Using Terra/MODIS-Derived Crop Metrics.** Ieee Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, [S.L.], v. 13, p. 2685-2697, 2020. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/jstars.2020.2984158>. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9103945>. Acesso em: 20 abr. 2022.

LOBELL, D. B. *et al.* **Sight for Sorghums: comparisons of satellite- and ground-based sorghum yield estimates in mali.** Remote Sensing, [S.L.], v. 12, n. 1, p. 100, 27 dez. 2019. MDPI AG. <http://dx.doi.org/10.3390/rs12010100>. Disponível em: <https://www.mdpi.com/2072-4292/12/1/100/htm>. Acesso em: 7 abr. 2022.

LOWENBERG-DEBOER, J.; ERICKSON, B. **Setting the Record Straight on Precision Agriculture Adoption.** Agron. J. 2019, 111, 1552–1569. [Google Scholar] [CrossRef]

MARÍN-ORTIZ, J. C. *et al.* **Linking physiological parameters with visible/near-infrared leaf reflectance in the incubation period of vascular wilt disease.** Saudi Journal Of Biological Sciences, [S.L.], v. 27, n. 1, p. 88-99, jan. 2020. Elsevier BV. <http://dx.doi.org/10.1016/j.sjbs.2019.05.007>.

MEIRELES, José Eduardo *et al.* Leaf reflectance spectra capture the evolutionary history of seed plants. New Phytologist, [S.L.], v. 228, n. 2, p. 485-493, 24 jul. 2020. Wiley. <http://dx.doi.org/10.1111/nph.16771>

MISAKI, E.; APIOLA, M.; GAIANI, S., & TEDRE, M. (2018). **Challenges facing sub-Saharan small-scale farmers in accessing farming information through mobile phones: A systematic literature review.** The Electronic Journal of Information Systems in Developing Countries, 84(4), e12034. <https://doi.org/10.1002/isd2.12034>.

NIGAM, S. N.; RAO, M. J. V.; GIBBONS, R. W. **Artificial hybridization in groundnut. Índia:** ICRISAT. v. 29, p.29-35. 1990.

NUNES, R. F.; PAULA, D. C. G.; MARTINS, M. G. **Produção de biodiesel do óleo de amendoim por transesterificação metílica e aplicação da superfície de resposta.** Scientia Plena, [S.L.], v. 17, n. 4, p. 1-16, 13 maio 2021. Associação Sergipana de Ciência. <http://dx.doi.org/10.14808/sci.plena.2021.044201>.

ONYANGO, C. et al. **Precision Agriculture for Resource Use Efficiency in Smallholder Farming Systems in Sub-Saharan Africa: A Systematic Review.** Sustainability 2021, 13, 1158. [Google Scholar] [CrossRef]

PATHAK, H.S.; BROWN, P.; BEST, T. **A systematic literature review of the factors affecting the precision agriculture adoption process.** Precis. Agric. 2019, 20, 1292–1316. [Google Scholar] [CrossRef]

PALIWAL, A.; JAIN, M. **The Accuracy of Self-Reported Crop Yield Estimates and Their Ability to Train Remote Sensing Algorithms.** Frontiers In Sustainable Food Systems, [S.L.], v. 4, n. 1, p. 1-10, 11 mar. 2020. Frontiers Media SA. <http://dx.doi.org/10.3389/fsufs.2020.00025>. Disponível em: <https://www.frontiersin.org/articles/10.3389/fsufs.2020.00025/full>. Acesso em: 7 abr. 2022.

PALUDO, A. B. et al. **Mapping summer soybean and corn with remote sensing on Google Earth Engine cloud computing in Parana state – Brazil.** International Journal Of Digital Earth, [S.L.], v. 13, n. 12, p. 1624-1636, 5 jun. 2020. Informa UK Limited. <http://dx.doi.org/10.1080/17538947.2020.1772893>.

PIGHINELLI, A.L.M.T. **Extração mecânica de óleos de amendoim e de girassol para a produção de biodiesel via catálise básica.** 2007. 93f. Dissertação (Mestrado em Engenharia Agrícola) – Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas, 2007.

PINHEIRO, R. M. et al. **Inteligência artificial na agricultura com aplicabilidade no setor sementeiro.** Diversitas Journal, [S.L.], v. 6, n. 3, p. 2984-2995, 2021. Universidade Estadual de Alagoas. http://dx.doi.org/10.48017/diversitas_journal-v6i3-1857.

Ponzoni, F. J. **Comportamento espectral da vegetação.** In. Sensoriamento Remoto - Reflectância dos alvos naturais. Brasília: UnB, EMBRAPA, 2001. p.157-199.

RODRIGUES, L. G. S. M. *et al.* **Amendoim (*Arachis sp.*) como fonte na matriz energética brasileira.** Journal Of Bioenergy And Food Science, [S.L.], v. 3, n. 3, p. 178-190, 30 set. 2016. Instituto Federal do Amapá. <http://dx.doi.org/10.18067/jbfs.v3i3.80>.

ROSA, R. **Introdução ao sensoriamento remoto.** 7. ed. Uberlândia: EDUFU, 2009. 262 p ISBN 978-87-7078-219-9.

ROZA, D. **Novidade no campo: Geotecnologias renovam a agricultura.** Revista InfoGEO, n 11 - jan/fev, 2000. Disponível na Internet. http://www.infogeo.com.br/Revista/materia_11.htm em 21 Mai. 2000.

ROUSE, J.W.; HAAS, R.H. J.A.; SCHELL, D.W. **Deering Monitoring Vegetation Systems in the Great Plains** with ERTS NASA special publication (1974), p. 309 - 351(1974)

SANTOS, A. F. **MÉTODO NÃO DESTRUTIVO PARA PREDIÇÃO DA MATURAÇÃO DE AMENDOIM (*Arachis hypogaea* L.) UTILIZANDO SENSORIAMENTO REMOTO**. 2019. 107 f. Tese (Doutorado) - Curso de Agronomia, Universidade Estadual Paulista - Unesp, Jaboticabal, 2019.

SANTOS, V. B. et al. **Machine learning algorithms for soybean yield forecasting in the Brazilian Cerrado**. Journal Of The Science Of Food And Agriculture, [S.L.], p. 1-16, 27 dez. 2021. Wiley. <http://dx.doi.org/10.1002/jsfa.11713>. Disponível em: https://onlinelibrary.wiley.com/doi/full/10.1002/jsfa.11713?casa_token=_Og9WkSsZ2gAAAAA%3AGg2sFxTJxhfiJH7hq-aGHu7JUEMrK2bqrsux8l7dnHe_cZKkeop0BnhBZsJOKqTLCwUjk1XugUE4n7A. Acesso em: 19 abr. 2022.

Sentinel-2User Handbook. Available online: https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook (Acessado em 06/04/2022).

SHADDAD, S.M. et al. **Data fusion techniques for delineation of site-specific management zones in a field in UK**. *Precis. Agric.* 2016, 17, 200–217. [Google Scholar] [CrossRef]

SILVA, S, de A. et al. **Modelo agrometeorológico na estimativa da produtividade de duas variedades de café arábica considerando a variabilidade espacial**. *Irriga*, v.16, n.1, p.1-10, 2011. Disponível em: . Acesso em: 02 maio. 2022.

SOLOVCHENKO, A. *et al.* **Linking Tissue Damage to Hyperspectral Reflectance for Non-Invasive Monitoring of Apple Fruit in Orchards**. *Plants*, [S.L.], v. 10, n. 2, p. 310, 5 fev. 2021. MDPI AG. <http://dx.doi.org/10.3390/plants10020310>

SEPE, P. M.; GOMES, S. **Indicadores ambientais e gestão urbana: desafios para a construção da sustentabilidade na cidade de São Paulo**. São Paulo: Secretaria Municipal do Verde e do Meio Ambiente/Centro de Estudos da Metrópole, 2008.

SERRANO, J. ET AL. **Mapping management zones based on soil apparent electrical conductivity and remote sensing for implementation of variable rate irrigation: Case study of Corn under a center pivot**. *Water* 2020, 12, 3427. [Google Scholar] [CrossRef]

SOUZA, J. B. C. **Qualidade da aquisição do NDVI na cultura do amendoim utilizando sensores terrestres**. *South American Sciences*, [S.L.], v. 1, n. 1, p. 1-9, 2020. Galoa Events Proceedings. <http://dx.doi.org/10.17648/sas.v1i1>.

SUN, C.X. *et al.* **Reflectance and biochemical responses of maize plants to drought and re-watering cycles**. *Annals Of Applied Biology*, [S.L.], v. 172, n. 3, p. 332-345, 22 fev. 2018. Wiley. <http://dx.doi.org/10.1111/aab.12423>.

TEDESCO, D. *et al.* **Use of remote sensing to characterize the phenological development and to predict sweet potato yield in two growing seasons**. *European Journal Of Agronomy*, [S.L.], v. 129, p. 126337, set. 2021. Elsevier BV.

<http://dx.doi.org/10.1016/j.eja.2021.126337>. Disponível em: https://www.sciencedirect.com/science/article/pii/S1161030121001088?casa_token=ZuMdr1nlhyYAAAAA:NhCAcP-kb2iJPn6b2bZUSDWsBAf-eL1WZiLf1u6GjmL8gOmyppoWXwTdY2XcJIgmZ1iz8mam7DY. Acesso em: 20 abr. 2022.

VAN KLOMPENBURG, T.; KASSAHUN, A.; CATAL, C. **Crop yield prediction using machine learning: a systematic literature review**. *Computers And Electronics In Agriculture*, [S.L.], v. 177, p. 105709, out. 2020. Elsevier BV. <http://dx.doi.org/10.1016/j.compag.2020.105709>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0168169920302301>. Acesso em: 8 abr. 2022.

VIEIRA, S. R. **Geoestatística em estudos de variabilidade espacial do solo**. In: NOVAIS, R. F.; ALVAREZ, V. H.; SCHAEFER, G. R. (Ed.). *Tópicos em ciência do solo*. Viçosa: Sociedade Brasileira de Ciência do Solo, 2000. p. 1-54. v. 1.

VISCARRA ROSSEL, R.A.; ADAMCHUK, V.I. Proximal soil sensing. *Vadose Zo. J.* 2011, 10, 1340. [Google Scholar]

WILSON, N. R.; NORMAN, L. M. **Analysis of vegetation recovery surrounding a restored wetland using the normalized difference infrared index (NDII) and normalized difference vegetation index (NDVI)**. *International Journal Of Remote Sensing*, [S.L.], v. 39, n. 10, p. 3243-3274, 12 fev. 2018. Informa UK Limited. <http://dx.doi.org/10.1080/01431161.2018.1437297>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/01431161.2018.1437297?scroll=top&needAccess=true>. Acesso em: 06 abr. 2022.

XUE, J.; SU, B. **Significant Remote Sensing Vegetation Indices: a review of developments and applications**. *Journal Of Sensors*, [S.L.], v. 2017, p. 1-17, 2017. Hindawi Limited. <http://dx.doi.org/10.1155/2017/1353691>. Disponível em: <https://downloads.hindawi.com/journals/js/2017/1353691.pdf>. Acesso em: 06 abr. 2022.

YOOSEFZADEH-NAJAFABADI, M.; TULPAN, D.; ESKANDARI, M. **Application of machine learning and genetic optimization algorithms for modeling and optimizing soybean yield using its component traits**. *Plos One*, [S.L.], v. 16, n. 4, p. 1-9, 30 abr. 2021. Public Library of Science (PLoS). <http://dx.doi.org/10.1371/journal.pone.0250665>. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0250665>. Acesso em: 19 abr. 2022.

ZHANG, J. *et al.* **Improving wheat yield estimates using data augmentation models and remotely sensed biophysical indices within deep neural networks in the Guanzhong Plain, PR China**. *Computers And Electronics In Agriculture*, [S.L.], v. 192, p. 106616, jan. 2022. Elsevier BV. <http://dx.doi.org/10.1016/j.compag.2021.106616>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0168169921006335>. Acesso em: 8 abr. 2022.

ZHOU, X. *et al.* **Reconstruction of time series leaf area index for improving wheat yield estimates at field scales by fusion of Sentinel-2, -3 and MODIS imagery.** *Computers And Electronics In Agriculture*, [S.L.], v. 177, p. 105692, out. 2020. Elsevier BV. <http://dx.doi.org/10.1016/j.compag.2020.105692>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0168169920311248>. Acesso em: 8 abr. 2022.