

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE GEOGRAFIA
CURSO DE ENGENHARIA DE AGRIMENSURA E CARTOGRÁFICA

MÁRCIO JOSÉ FERREIRA NEVES

**ESTIMATIVA DA PRODUTIVIDADE DO MILHO A PARTIR DE IMAGENS DE
ALTA RESOLUÇÃO ESPACIAL**

MONTE CARMELO

2022

MÁRCIO JOSÉ FERREIRA NEVES

ESTIMATIVA DA PRODUTIVIDADE DO MILHO A PARTIR DE IMAGENS DE ALTA
RESOLUÇÃO ESPACIAL

Trabalho apresentado como requisito para obtenção de aprovação na disciplina de Trabalho de Conclusão de Curso II do curso de Engenharia de Agrimensura e Cartográfica da Universidade Federal de Uberlândia campus Monte Carmelo.

Orientador: Prof. Dr. George Deroco Martins.

MONTE CARMELO

2022

AGRADECIMENTOS

Agradeço primeiramente a Deus por me dar fé e ter algo em que acreditar e me apoiar nos momentos de dúvidas e fraqueza, nunca deixando eu desistir.

Aos meus pais, Márcio e Rosilene, pois sem eles eu não seria ninguém e sou eternamente grato a eles por conseguir estar aqui nesse momento. Sempre me apoiaram incondicionalmente e passaram por todos os obstáculos junto comigo.

Ao professor George Deroco por aceitar ser meu orientador e, agradecer por todos os conhecimentos a mim passados, além de todo apoio e conversas.

A minha namorada Flávia, que esteve comigo nesses últimos dois anos, sempre me ajudando e apoiando.

Meus inúmeros amigos e colegas, pelas ajudas e momentos de descontração e diversão nesta dura caminhada.

A coordenação e secretaria deste curso, que sempre atenderam com prontidão as necessidades, e assim, a todo o corpo docente pelos conhecimentos por todos eles transmitidos a mim.

RESUMO

Atualmente, o sensoriamento remoto está se tornando cada vez mais importante devido à sua relação custo/benefício e é cada vez mais usado para monitorar lavouras e plantações. Uma das tecnologias que vem promovendo vigorosamente o sensoriamento remoto na agricultura de precisão são as imagens orbitais de alta resolução, que podem monitorar plantações agrícolas e auxiliar nos processos de tomada de decisão nos principais ambientes agronômicos. A velocidade desta tecnologia a tornou uma das tecnologias mais importantes em agricultura de precisão hoje. Um dos objetivos da agricultura de precisão é ajudar a determinar parâmetros agronômicos que promovam as atividades de campo e aumentem a produtividade das lavouras. Para estimar a produtividade das lavouras, é necessário entender o comportamento de cada cultura analisada para encontrar as características dos possíveis padrões de produção de cada plantação. Portanto, o principal desafio vem de culturas com comportamentos mais específicos, como a cultura branca. Para isso, será usado índices de vegetação combinados com bandas espectrais para estimar produtividade de culturas brancas em uma determinada área, assim utilizando algoritmos de aprendizado de máquinas, pretende-se apresentar a aplicação dos modelos de estimativa a partir de mapas da distribuição espacial da produtividade. Através dos estudos, concluiu-se que o melhor estágio vegetativo para se estimar a produtividade do milho é o R3, através do modelo *Multilayer Perceptron* que apresentou uma acurácia de aproximadamente 78% e um índice de correlação de 0,57.

Palavras-chave: Agricultura de Precisão. Estimativa de Produtividade do Milho. Aprendizado de Máquinas. Sensoriamento Remoto.

ABSTRACT

Recently, remote sensing is becoming more important due to the cost/benefit ratio and it's used even more to monitor the tillage and plantations. One of the technologies that has been rigorously promoting the remoting sensing in precision agriculture are the orbitals images of high resolution, that can monitor agricultural plantations and assist in decision-making processes in the main agronomic environments. The speed of this technology has made it one of the most important technologies in precision agriculture nowadays. One of the goals of precision agriculture is to help to determine agronomic parameters that promote the field activities and raise the productivity of the farming. To estimate the productivity of the farming, it is necessary to understand the conduct of each analyzed culture to find the characteristics of possible patterns of each plantation. Therefore, the main challenge comes with cultures with more specific conduct, like the white culture. For this, vegetation indices combined with spectral bands to estimate productivity of white culture in a determined area, so using machine learning algorithms, it's intended to present the application of the estimation models through spatial distribution maps of productivity. The studies conclude that the best vegetative stadia to estimate the corn productivity is the R3, through the Multilayer Perceptron that showed an accuracy about 78% and a correlation index of 0,57.

Keywords: Precision Agriculture. Corn Productivity Estimate. Data Learning. Remote Sensing.

LISTA DE ILUSTRAÇÕES

Figura 1 -	Fases de desenvolvimento da cultura do milho	15
Figura 2 -	Diagrama Redes Neurais Artificiais	21
Figura 3 -	Diagrama <i>Random Forest</i>	21
Figura 4-	Mapa de Localização da Área de Estudo	22
Figura 5-	Pontos coletados em campo	24
Figura 6-	Fluxograma de execução do trabalho	25
Figura 7-	Estádios com imagem adquirida	26
Figura 8 -	Imagens RGB dos estádios fenológicos do milho	29
Figura 8-	Estimativa de Produtividade em R3	36

LISTA DE TABELAS

Tabela 1 -	Bandas e comprimentos de onda do Planet	23
Tabela 2 -	Dados usados para validação	27
Tabela 3 -	Coefficiente de correlação e RMSE em VE	30
Tabela 4 -	Coefficiente de correlação e RMSE em V5	30
Tabela 5 -	Coefficiente de correlação e RMSE em V8	31
Tabela 6 -	Coefficiente de correlação e RMSE em R1	32
Tabela 7 -	Coefficiente de correlação e RMSE em R2	32
Tabela 8 -	Coefficiente de correlação e RMSE em R3	33
Tabela 9 -	Coefficiente de correlação e RMSE em R4	33
Tabela 10 -	Coefficiente de correlação e de RMSE na primeira imagem em R5	33
Tabela 11 -	Coefficiente de correlação e de RMSE na segunda imagem em R5	34
Tabela 12 -	Coefficiente de correlação e de RMSE na terceira imagem em R5	34
Tabela 13 -	Coefficiente de correlação e de RMSE na quarta imagem em R5	34
Tabela 14 -	Coefficiente de correlação e de RMSE na série temporal	35

LISTA DE ABREVIATURAS E SIGLAS

EVI	Enhanced Vegetation Index
EXG	Excess Green Vegetation
GNDVI	Green Normalized Difference Vegetation Index
IV	Índice de Vegetação
NDVI	Normalized Difference Vegetation Index
NIR	Near InfraRed
REM	Radiação EletroMagnética
RGB	Red, Green, Blue
RMSE	Root Mean Square Error
SCCCI	Soil Canopy Chlorophyll Content Index
SPAD	Soil Plant Analysis Development
SR	Simple Ratio
TGI	Triangulation Green Index
VARI	Visible Atmospherically Resistant Index

SUMÁRIO

1 INTRODUÇÃO	12
2 OBJETIVO	13
3 JUSTIFICATIVA	14
4 REVISÃO BIBLIOGRÁFICA	14
4.1 Estádios fenológicos do milho	14
4.2 Sensoriamento Remoto	16
4.3 Índices de Vegetação	16
4.4 Caracterização espectral do milho	17
4.5 Modelos de estimativa de produtividade do milho	18
4.6 Aprendizado de máquinas	19
4.6.1 Redes neurais artificiais	19
4.6.2 Random Forest	21
5 MATERIAL E MÉTODOS	22
5.1 Área de estudo	22
5.2 Material	23
5.3 Métodos	24
5.3.1 Aquisição dos dados de produtividade	25
5.3.2 Aquisição de imagens multiespectrais de alta resolução	26
5.3.3 Cálculo de Índices de Vegetação e extração dos valores de reflectância	26
5.3.4 Geração dos modelos de estimativa	27
6 RESULTADOS E DISCUSSÕES	29
7 CONCLUSÃO	37

1 INTRODUÇÃO

O milho (*Zea mays*) é considerado a maior cultura do mundo e pode ser cultivado em regiões tropicais, subtropicais ou temperadas em altitudes desde o nível do mar até mais de 3600 m acima do nível do mar (CONTINI et al., 2019). Outro fato importante sobre o milho é que ele pode ser usado tanto na alimentação humana quanto animal (GARCIA; MATTOSO; DUARTE, 2006). Dada a importância do milho no cenário mundial, é necessário estimar a produtividade para atender a demanda global pelo cereal (KLEINSCHMITT, 2016). As estimativas de produtividade permitem que os agricultores analisem as lavouras por área, verificando áreas de cultivo que necessitam de tratamento (MELO et al., 2008). Com isso, os agricultores analisam detalhadamente áreas de baixa, média ou alta produtividade com maior segurança na tomada de decisão para o manejo adequado dos insumos (LEAL et al., 2015).

A estimativa de produtividade é um importante parâmetro agrônômico que auxilia os mercados doméstico e internacional em termos de demanda de produtos agrícolas, capacidade de transporte e armazenamento (RODRIGUES et al., 2004). Ao prever a produtividade das culturas, além de prever os preços, levando em consideração a quantidade de produto que será produzida, é possível comercializar antes da colheita (BERTOLIN et al., 2017).

Um método de estimativa de produtividade é combinar a coleta de parâmetros biofísicos amostrados em campo (FIGUEIRAS, VENANCIO, MANTOVANI, 2017), com informações obtidas por meio de técnicas de sensoriamento remoto (SHANAHAN, 2001). Essas técnicas incluem a análise da resposta espectral de culturas e alvos presentes em superfícies adquiridas por sensores multiespectrais acoplados a satélites, aeronaves pilotadas remotamente ou máquinas de campo (MENESES, ALMEIDA, 2012). Esses sensores têm a capacidade de registrar a radiação eletromagnética (REM), possibilitando pesquisas em diferentes tipos de alvos, como a agricultura (JESSEN, 2009).

Entretanto, estimar a produtividade de uma cultura agrícola por imagens vai além da aplicação de técnicas, sendo também necessários estudos que impulsionam e otimizam cada vez mais os métodos a serem aplicados em cada cultura (BERTOLIN et al., 2017). Na agricultura, cada lavoura apresenta suas particularidades, que variam desde o tipo de manejo aplicado em campo, à mudanças estruturais inerentes a cada tipo de cultura e que caracterizam as mesmas, como por exemplo, o tempo de cultivo que cada cultura apresenta (PÁDUA et al., 2018).

As culturas podem ser divididas em anuais e perenes. As culturas anuais podem ser caracterizadas, dentre outras coisas, pelo desenvolvimento em períodos mais curtos, e por

serem removidas do campo após o período da colheita (AMARAL, 2011). Levando em conta a variação das características de cada estado fenológico, surge a necessidade de análise sobre a variação da resposta espectral sobre os mesmos, a fim de se obter o período de estimativa ideal para obtenção das imagens multiespectrais (FORMAGGIO; SANCHES, 2017).

É importante obter o período ideal para estimar a produtividade do milho para que haja um padrão entre os tempos em que os dados multiespectrais são obtidos em cada trabalho. Porque, atualmente, não há consenso sobre o período em que as imagens utilizadas para estimar a produtividade da cultura do milho foram obtidas. Para estimar a produtividade e o vigor da cultura do milho, Wahab, Hall e Jirström (2018) realizaram 3 medições de campo em 4 a 5, 9 e 13 semanas após o plantio, usando imagens multiespectrais de fotogrametria aérea. Em seu trabalho, os autores descobriram a eficiência das estratégias utilizadas medindo estimativas de valores SPAD e pontuações visuais, que incluíam dados das primeiras 4 a 5 semanas.

Para geração de um modelo de predição de produtividade do milho, Ahmad et al. (2018) utiliza imagens do satélite Landsat-8 com intervalo de 2 a 3 meses após a semeadura. Segundo o autor, o período escolhido para obtenção das imagens corresponde ao pico de crescimento para o milho. No trabalho, os autores utilizam algoritmos de aprendizado de máquinas, e conseguem gerar modelos com R^2 de 0,94. Sendo assim, mais do que estabelecer um padrão, a determinação do período ideal para obtenção de imagens, podem impulsionar os resultados dos estudos que visam mensurar parâmetros agrônômicos da cultura do milho.

Portanto, além de obter imagens multiespectrais de alta qualidade e ter confiança em quando foram obtidas, é necessário processar as informações obtidas por meio de técnicas robustas, capazes de lidar com grandes quantidades de dados e produzir resultados confiáveis (MEKONNEN et al., 2019) . Diante dessa necessidade, o aprendizado de máquina tem se mostrado eficaz na resolução de problemas relacionados ao processamento de dados de parâmetros agrônômicos e é cada vez mais utilizado na agricultura de precisão (TREBOUX; GENOUD, 2018; CHLINGARYAN et al., 2018; MEKONNEN et al., 2019). Segundo Santos (2020), a produtividade das lavouras de cana-de-açúcar pode ser estimada com dois meses de antecedência usando algoritmos de redes neurais artificiais profundas.

2 OBJETIVO

Dada a escassez de trabalhos que analisam a correlação da produtividade com os diferentes estados de desenvolvimento fenológico da cultura do milho, o presente trabalho foi

desenvolvido. Com o objetivo de identificar qual o período ideal para a captura de imagens multiespectrais, para estimar a produtividade na cultura do milho. A complexidade deste estudo parte-se da premissa que a resposta espectral durante o período de colheita, pode ser de alguma forma afetada pelo alto grau de senescência em que a planta se encontra. Desse modo, foi analisada a correlação entre a produtividade da cultura do milho, com os seus estádios fenológicos que antecedem os períodos de colheita, através de modelos oriundos de técnicas de aprendizado de máquinas.

Como objetivo específico, foi verificada a possibilidade de se estimar a produtividade da cultura do milho a partir de modelos espectrais considerando não apenas um período específico, mas sim todo o desenvolvimento fenológico da cultura, ou seja, um modelo de predição baseado em séries temporais de imagens.

3 JUSTIFICATIVA

Tendo em vista a importância atual do mercado de grãos no Brasil, sendo ele o milho um dos pioneiros e mais comercializados mundialmente, entende-se a importância de se prever a estimativa da produtividade para antecipar negociações e assim favorecer o produtor que usa desta tecnologia para negociações antecipadas e obter benefícios frente a seus concorrentes.

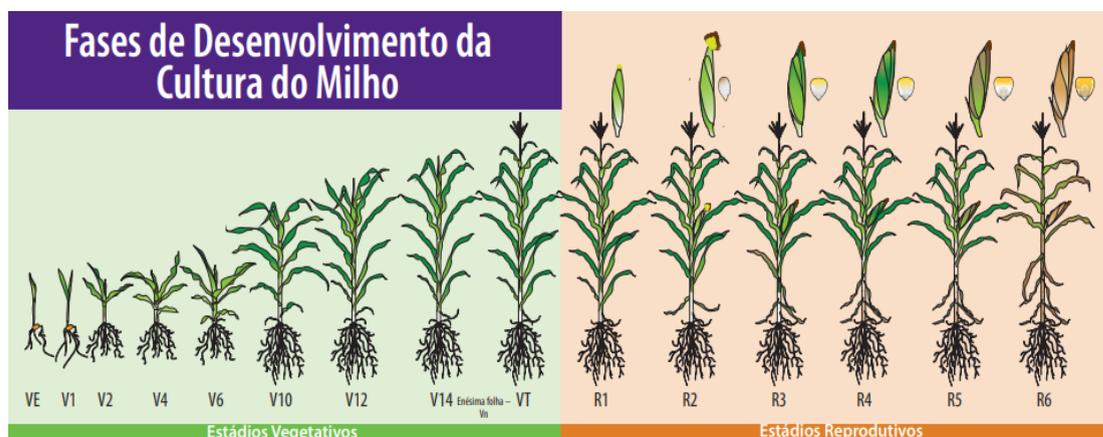
Além disso, com o estudo da estimativa de produtividade é possível a análise da cultura para que se obtenha as áreas com mais e menos produtividade, sendo assim possível a orientação para onde se encaminhe um agrônomo que pode identificar eventuais doenças, escassez hídrica, qualidade do solo, entre outros, naquela área da lavoura.

4 REVISÃO BIBLIOGRÁFICA

4.1 Estádios fenológicos do milho

O desenvolvimento do milho é dividido em dois, o Vegetativo (V) e o Reprodutivo (R), ambas são designadas numericamente, porém o vegetativo é designado em V1, V2, V3 até V(n), pois não se tem ao certo quantas folhas serão germinadas na planta; o número de cada folha totalmente visível determina o numeral; terminando em V(t) que é o estágio do pendoamento, o primeiro estágio vegetativo é o VE, que é quando a planta emerge no solo (MAGALHÃES, 2006; DURÃES, 2006).

Figura 1: Fases de desenvolvimento da cultura do milho



Fonte: IPNI.

O estágio reprodutivo começa no R1 a R6, onde R1 é dado pelo embonecamento da espiga, ou seja, quando começam a aparecer os “cabelos” do milho (MAGALHÃES, 2006; DURÃES, 2006).

O R2 é chamado de Grão Bolha D’água, pelas semelhanças dos grãos a uma bolha d’água, sua coloração é branca e seu interior composto basicamente de açúcares e sua umidade é de aproximadamente 85% (MAGALHÃES, 2006; DURÃES, 2006).

R3 também conhecido como grão leitoso, sua cor já está amarelada e no seu interior, os açúcares estão se transformando em amido. É um estágio importante pois pode se definir a densidade dos grãos, aqui a sua umidade é de 80% (MAGALHÃES, 2006; DURÃES, 2006).

O estágio R4, grão pastoso, ainda acumulam amido, porém ao invés do interior dos grãos estarem leitosos, estão numa forma pastosa, é caracterizado pelo ganho de peso do grão, e ao serem apertados, já estão consistentes, aqui a umidade está em cerca de 70% (MAGALHÃES, 2006; DURÃES, 2006).

O R5 é comparado com a formação de dente, pois aparece uma concavidade na parte superior do grão, o estado no seu interior está passando de pastoso para farináceo, aqui percebe-se a divisão dos estádios R4 e R5, o grão vai ficando mais resistente e amarelo, “empurrando” o amido do R4 para sua base, ficando com uma colocação mais branca, aqui a umidade dos grãos está por volta de 55%, é o estágio em que pode-se usar o milho a silagem (MAGALHÃES, 2006; DURÃES, 2006).

Assim chega-se ao último estágio reprodutivo o R6, onde o grão atinge sua maturidade fisiológica, o amido neste estágio já foi completamente para dentro da espiga, uma camada

preta começa a se formar, tendo início na ponta da espiga até sua base, avançando progressivamente, sendo caracteriza pela umidade já estar baixa, entre 30% e 38%, é o ponto máximo de produtividade que se pode obter na cultura sendo assim ideal para colheita, porém, é aconselhável que para armazenamento, o grão esteja entre 13% a 15% (MAGALHÃES, 2006; DURÃES, 2006).

4.2 Sensoriamento Remoto

A vegetação é muito importante para o funcionamento do ecossistema, o conhecimento da variabilidade de padrões de distribuição, ciclo fenológico e mudanças na fisiologia e morfologia dos vegetais fornece informações importantes de uma área (WEIERS et al., 2004; JACKSON e JENSEN, 2005). Com o avanço da tecnologia vem se buscando desenvolver sensores e algoritmos de processamento digitais para extrair informações biofísicas importantes sobre a vegetação a partir de dados de sensoriamento remoto. Muitas técnicas de sensoriamento remoto são genéricas por natureza e podem ser aplicadas a uma variedade de paisagens de vegetação, uma dessas paisagens vegetativas se trata da agricultura (JENSEN, 2011).

O sensoriamento remoto é a ciência onde se estuda um objeto sem estar em contato físico com ele, e ele pode ser utilizado para monitorar e medir características biofísicas da superfície terrestre. Os sensores remotos emitem ou refletem REM que podem ser usados para interpretar superfícies da Terra como solo, água, vegetação, construções, etc (JENSEN, 2011).

Nos sensores existem os sistemas multiespectrais, onde é possível receber reflectância de REM em múltiplas bandas da área de interesse, sendo esses dados coletados de forma digital (JENSEN, 2011). Esses dados são coletados por sensores passivos, onde é captado pelas câmeras de satélites a reflectância das bandas do visível e do infravermelho, gerada pelo meio (SHIPPERT, 2004).

4.3 Índices de Vegetação

O índice de vegetação é uma medida adimensional de radiação que expressa a abundância relativa e a atividade da vegetação verde, incluindo índice de área foliar, cobertura verde, teor de clorofila, biomassa e radiação fotossinteticamente ativa absorvida. Idealmente, o índice de vegetação ideal deve (RUNNING et al., 1994; HUETE e JUSTICE, 1999):

- Maximizar a sensibilidade aos parâmetros biofísicos da planta, preferencialmente de forma linear, para que essa sensibilidade seja confiável para uma ampla gama de condições de vegetação e facilite a validação e calibração do índice;
- Modelar influências externas como distúrbios atmosféricos para comparações espaciais e temporais;
- Normalizar efeitos internos, como mudanças no substrato abaixo do dossel, incluindo solo, e diferenças na senescência;
- Acoplado a alguns parâmetros biofísicos mensuráveis, como biomassa e índice de área foliar, para validação e controle de qualidade.

Os índices de vegetação podem ser calculados com combinações das bandas do visível e do infravermelho. Existem alguns índices de vegetação utilizados na agricultura, são eles:

NDVI (*Normalized Difference Vegetation Index*): Talvez seja o índice mais famoso e conhecido na agricultura de precisão. Este índice é capaz de medir a vegetação verde e sua saúde vegetal. A combinação de sua formulação de diferença normalizada e o uso das regiões de maior absorção e reflectância da clorofila o tornam robusto em uma ampla gama de condições. Porém, pode tender a ficar mais saturado em condições de vegetação muito densa, ou seja, quando o índice de área foliar é alto (ROUSE, 1973).

GNDVI (*Green Normalized Difference Vegetation Index*): Este índice é semelhante ao NDVI, a diferença é que ele usa a banda do verde (540 a 570 nm) ao invés do vermelho, isso o torna mais sensível à concentração de clorofila do que o NDVI (GITELSON, 1998).

4.4 Caracterização espectral do milho

A maioria das pesquisas de sensoriamento remoto é baseada na modelagem da relação entre a energia eletromagnética refletida ou emitida em bandas específicas e as características químicas, biológicas e físicas do fenômeno sob investigação, como o dossel do milho (JENSEN, 2011).

A resolução espectral é o número e o tamanho da faixa de comprimento de onda das bandas ou canais no espectro eletromagnético aos quais os instrumentos de sensoriamento remoto são sensíveis. Em um sistema multiespectral, a resposta REM é registrada em várias bandas (JENSEN, 2011).

Certas bandas espectrais do eletromagnético são adequadas para obter informações biofísicas. As bandas são geralmente escolhidas para maximizar o contraste entre o objeto de interesse e seu substrato. Selecionado cuidadosamente o número de bandas espectrais pode melhorar a probabilidade de extrair a informação desejada do sensor remoto (JENSEN, 2011).

Todo objeto tem uma estrutura molecular e atômica em que os elétrons são distribuídos em torno de núcleos de diferentes níveis de energia que podem absorver diferentes quantidades de Radiação Eletromagnética (REM), e quanto mais o objeto absorve, menos energia ele reflete. O local onde ocorre a absorção de energia é chamado de banda de absorção. Esse fenômeno de um objeto interagindo com o incidente REM é chamado de comportamento espectral (FIGUEIREDO, 2015).

Na faixa do visível o comportamento espectral da vegetação apresenta uma variação de 400 a 750 nm, onde tem seu máximo de reflectância em VT (pendoamento do milho) (SUN et al., 2010), sendo um primeiro pico em 480 nm, relacionado a pigmentação de carotenóides, um segundo pico com 680 nm, relacionado a clorofila e um último com 500 nm que é a coloração verde da folha (RSI, 2013). Nesta etapa é apresentada dois picos por conta de água presente na folha, não são precisos para a cobertura vegetal porém muito importante para o estudo da vegetação (RSI, 2013).

Já na faixa do infravermelho próximo, de 750 a 1000 nm, a reflectância sobe gradualmente de V6 até R1, e depois declina até R2 voltando a subir até R5 (SUN et al., 2010).

4.5 Modelos de estimativa de produtividade do milho

No estudo de Barzin et al., (2020), foram utilizados 5 bandas espectrais e 26 IVs e os modelos de regressão e árvore em diferentes estádios vegetativos do milho sendo eles V3, V4-5, V6-7, V10-11 e VT, onde, a influência das variáveis mudou de acordo com o estágio fenológico, o mais impactante foi SCCCI, um índice baseado no infravermelho próximo. Analisando pelo índice de regressão, o maior R^2 foi em V10 e VT com valores de 0,97 e 0,95, respectivamente. A acurácia desse estudo foi aumentando conforme o milho foi se desenvolvendo.

Já no estudo de Sunoj et al., (2021), com vôos feitos do VE até R5, com câmera capaz de captar as bandas RGB e NIR, ele teve como principais objetivos comparar os IVs NDVI, GNDVI, EVI1, SR, EXG e TGI, para ver a capacidade de estimar o rendimento e o efeito da aplicação de cobertura na estimativa, ambos em R4 e avaliar ver os efeitos das aplicações de

IV no estágio/ tempo de crescimento. O NDVI e EVI2 tiveram as melhores estimativas de produtividade e os dados com menos confiabilidade foram de V6 até R1. Assim, determinaram que para uma estimativa de rendimento razoável, onde $R^2 > 0,80$, o sensoriamento deve ser realizado nos estágios entre R1 a R4.

4.6 Aprendizado de máquinas

A pesquisa em aprendizado de máquina segundo Mitchell et al., (1990), busca desenvolver sistemas computacionais que melhorem automaticamente seu desempenho por meio da experiência. Embora existam hoje formas especializadas de programas de aprendizado, o objetivo final é desenvolver sistemas mais amplamente aplicáveis com recursos de aprendizado mais robustos.

A longo prazo, a tecnologia pode levar a um tipo totalmente novo de software de computador que, ao contrário dos programas atuais, melhora com a experiência. Se bem-sucedida, a pesquisa de aprendizado de máquina pode produzir sistemas de computador, como robôs que aprendem a operar em novos ambientes, sistemas de consultoria baseados em conhecimento que trabalham com especialistas humanos para resolver problemas difíceis e adquirir novas estratégias de resolução de problemas. Uma solução final para um problema, ou um programa de computador que adquire a capacidade de resolver problemas de física ou cálculo lendo capítulos de livros didáticos e resolvendo problemas práticos no final do capítulo (MITCHELL et al., 1990).

O objetivo da pesquisa de aprendizado de máquina é gerar tecnologias de habilitação agnósticas de domínio para uma ampla gama de aplicativos de computador. Avanços no aprendizado de máquina podem ter grandes implicações para uma variedade de aplicativos de computador, como robótica, design auxiliado por computador, bancos de dados inteligentes e sistemas de consultoria baseados em conhecimento. Muitos aplicativos de computador são cada vez mais baseados em conhecimento, ou seja, contam com um grande número de fatos específicos sobre o domínio da tarefa. O aprendizado de máquina oferece o potencial de remover os gargalos de aquisição de conhecimento que limitam o desempenho e aumentam o custo de desenvolvimento desses sistemas (MITCHELL et al., 1990).

4.6.1 Redes neurais artificiais

A estrutura de uma Rede Neural Artificial (RNA) apresenta duas fases de estruturação, sendo elas a arquitetura e o algoritmo de aprendizagem, sendo ela construída assim pela forma em que a rede é treinada (RAUBER, 2014).

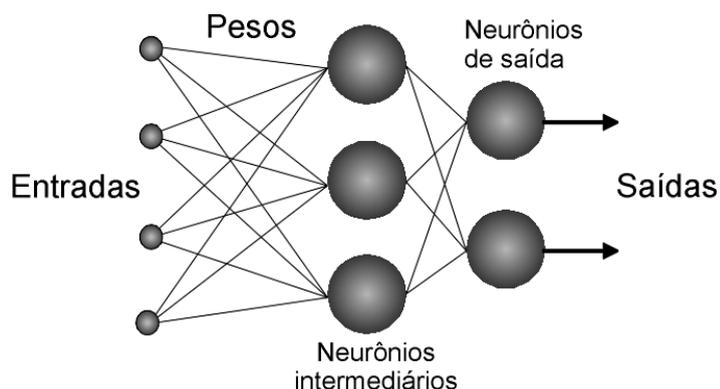
Um ponto marcante na história das redes neurais artificiais é a apresentação de modelos de neurônios artificiais (MCCULLOCH; PITTS, 1943). A atividade nessa área de pesquisa acabou levando ao conceito de perceptrons e um modelo semelhante chamado de Adaline. O perceptron é capaz de classificar entre classes linearmente separáveis, um exemplo, pode ser usado para reconhecer caracteres. O aplicativo gerou grande entusiasmo, considerado exagerado em relação à imaginação das capacidades dos futuros robôs inteligentes. Uma característica importante do *perceptron* é que ele propõe um algoritmo de aprendizado capaz de ajustar os pesos internos dos neurônios de forma a resolver problemas de classificação linear, no caso de classes linearmente separáveis (RAUBER, 2014).

Ao contrário de computadores convencionais que são programados, a RNA é treinada por entrada de dados. O algoritmo generaliza os dados informados na estrada e memoriza-os dentro de parâmetros, atribuindo a eles pesos, sendo assim, o sistema a ser criado baseado em RNA apresenta dois graus de liberdade, um define qual será a rede, que é composta por neurônios, que irá resolver o problema dos dados de entrada e o outro sendo o algoritmo que vai treinar essa rede baseado nos pesos a eles impostos (RAUBER, 2014). Os algoritmos podem ser classificados como aprendizagem supervisionada e não supervisionada.

No aprendizado supervisionado é fornecido dados de entrada com classes conhecidas ao algoritmo, ou seja, cada classe é representada por um vetor com valores de atributos, rotulados por classe associada a cada um, assim o algoritmo treina para construir uma classificação para classe de atributos ainda não rotulados, a partir dos dados de entrada. Quando os rótulos são discretos, são caracterizados como classificação, quando os valores são contínuos, intitula-se como regressão.

O aprendizado não-supervisionado, o algoritmo analisa o produto por inteiro, e determina quais atributos (ou pixel) podem ser agrupados de uma forma em que possam formar classes, formando assim agrupamentos (CHEESEMAN; STUTZ, 1990). Após o agrupamento é necessária uma análise para se determinar qual rótulo pode ser aplicado a cada classe (MONARD; BARANAUSKAS, 2003). Na figura 2 podemos analisar como é o diagrama do RNA.

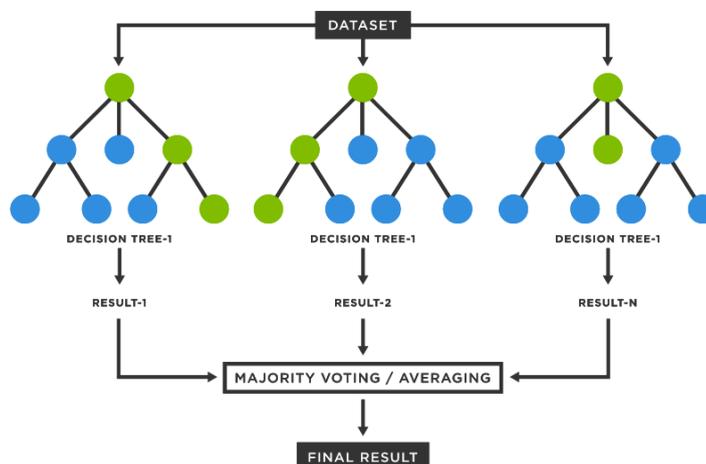
Figura 2: Diagrama Redes Neurais Artificiais.



Fonte: UFPE.

4.6.2 *Random Forest*

Como define (BREIMAN, 2001), *Random Forest* classifica amostras em uma coleção de árvores dispostas aleatoriamente independentes e distribuídas identicamente, e cada uma delas escolhe a que mais julga melhor para a entrada de dados. Este utiliza o mesmo método de *bagging* para o treinamento das amostras. Nele a partir do treinamento original, é construído um novo conjunto, nesse novo subconjunto, relaciona-se com uma seleção aleatória, e assim vai se formando a árvore, a cada “folha” da árvore o melhor atributo é selecionado, tendo um valor fixado para cada “folha”. O que é diferente na *Random Forest* é que os valores das “folhas” são sempre escolhidos aleatoriamente, e no método anterior, todos os atributos são usados. A figura a seguir ilustra uma *Random Forest*.

Figura 3: Diagrama *Random Forest*.

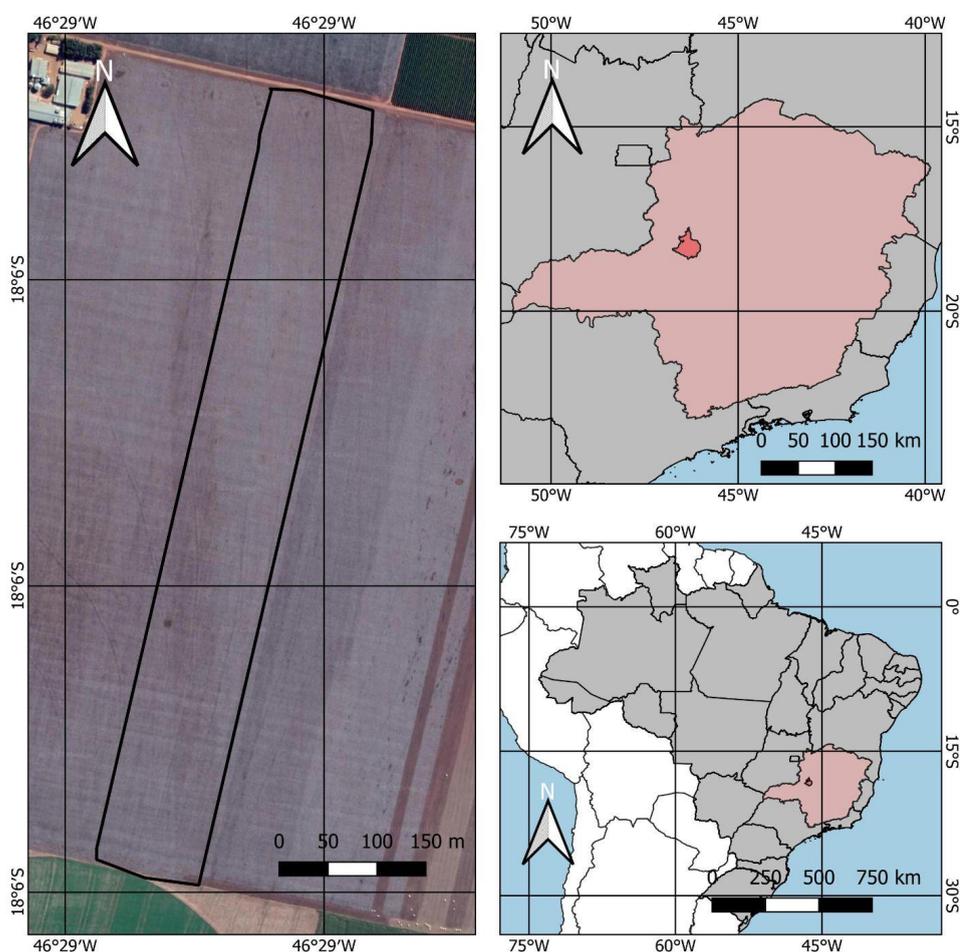
Fonte: Tibco, 2022.

5 MATERIAL E MÉTODOS

5.1 Área de estudo

A presente área onde está a cultura para o estudo, está localizada no município de Presidente Olegário, na mesorregião do Triângulo Mineiro e Alto Paranaíba, estado de Minas Gerais, com coordenadas de $18^{\circ}06'00''\text{S}$ e $46^{\circ}28'54''\text{W}$.

Figura 4: Mapa de Localização da Área de Estudo



Localização da
Área de Estudo

Legenda

- Área de Estudo
- Presidente Olegário
- Minas Gerais
- Brasil
- América do Sul

Sistema de Coordenadas Geográficas
DATUM: WGS84
Fonte de Dados: IBGE 2020
Elaboração: Neves, M. J. F. (2022)

Fonte: o autor.

O milho é uma das três principais culturas plantadas na mesorregião do Triângulo Mineiro e Alto Paranaíba, junto com a soja representa 99% da área de grãos plantada no estado. O milho em segunda safra, que acontece de janeiro a março, vem crescendo muito na região devido a avanços tecnológicos (BORGHI, 2016).

Outro fator que favorece o plantio de milho em segunda safra é o clima, a região está em um clima tropical e tem seu inverno em seca (ÁVILA et al., 2014), com as temperaturas variando no verão de 23 °C a 18 °C e no inverno de 16 °C a 21 °C (BORGHI, 2016).

5.2 Material

Genericamente os materiais utilizados para o desenvolvimento deste trabalho foram obtidos em duas etapas. A primeira foi a obtenção dos dados de produtividade através de um sistema mecanizado das colheitadeiras que estimam a produção por hectare, ou seja, a cada ponto é estimado a produtividade por hectare baseado nos grãos até ali colhidos. A segunda foi o acesso a imagens multiespectrais de alta resolução.

A aquisição das imagens multiespectrais de alta resolução foram obtidas do satélite *Planet* com um intervalo de mais ou menos 15 dias desde o plantio até a colheita. As imagens obtidas possuem 4 bandas, e oferecem uma resolução espacial de 3 metros e resolução radiométrica de 12 bits. A figura a seguir mostra as bandas e seus respectivos comprimentos de onda do satélite *Planet*.

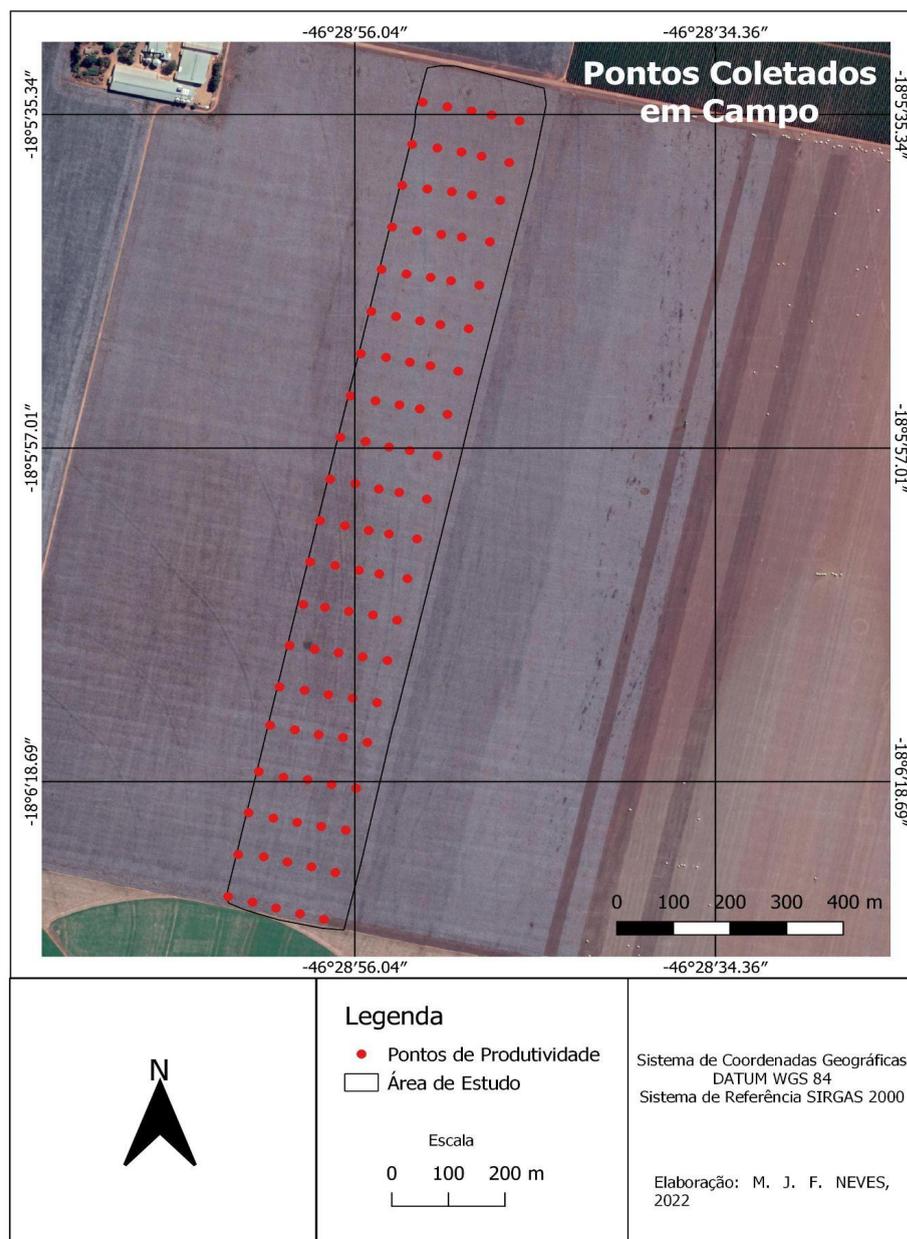
Tabela 1 – Bandas e comprimentos de onda do *Planet*

Nomenclatura	Banda	Comprimento de Onda
B1	Azul	465 - 515 nm
B2	Verde	547 - 583 nm
B3	Vermelho	650 - 680 nm
B4	Infravermelho Próximo	845 - 885 nm

Fonte: Adaptado *Planet* (2022).

Para a extração de valores de reflectância das imagens, foi utilizado o *software* “ENVI”, onde foi possível obter os valores da imagem por cada pixel para imagem de cada estágio fenológico do milho, sendo um total de 100 pontos distribuídos pela área de plantio.

Figura 5: Pontos coletados em campo



Fonte: o autor.

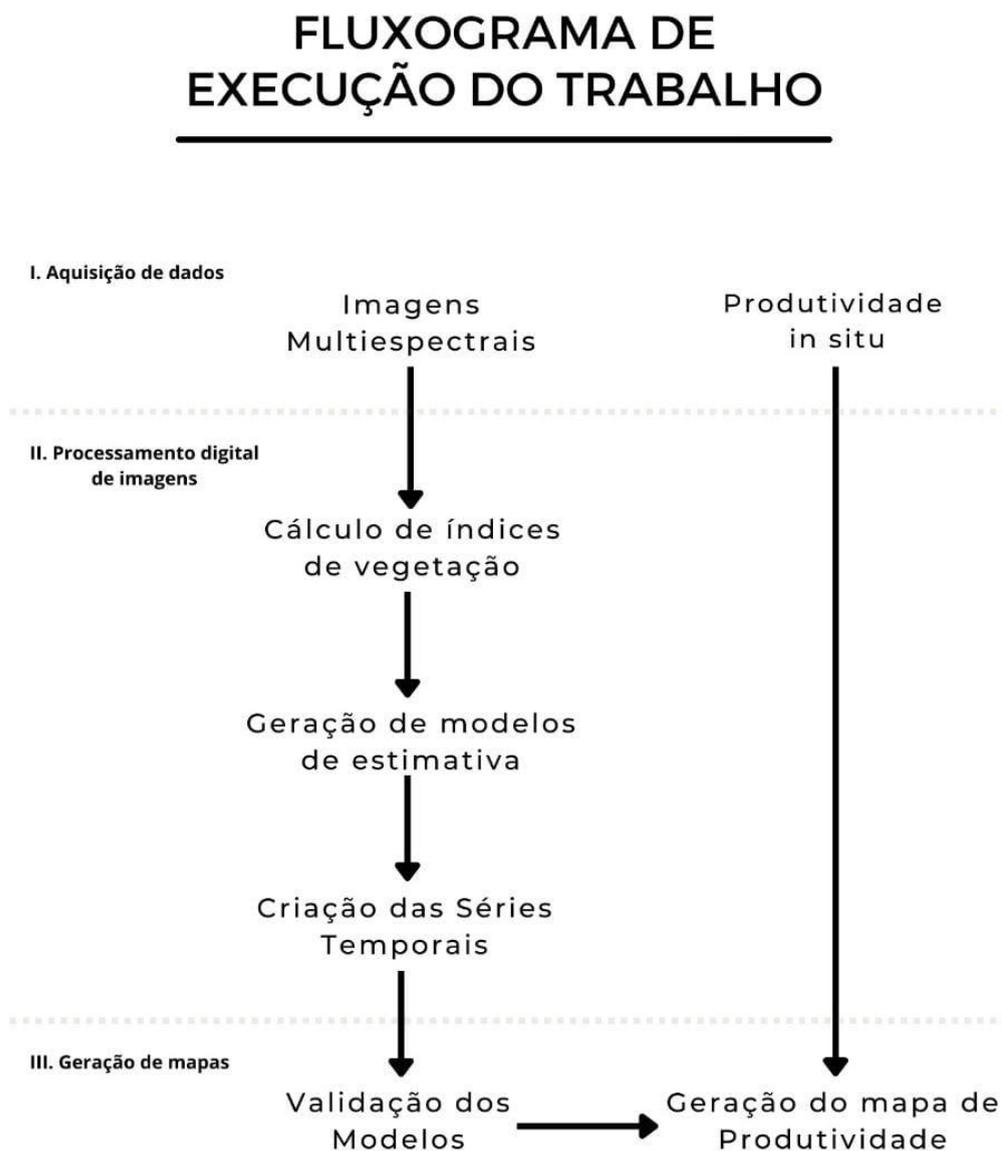
Após obtidos os valores de reflectância, foram eles tratados no *software* “WEKA”, onde através de *machine learning*, foram gerados modelos de estimativa, sendo eles melhorados pela mineração de dados, para que se pudesse obter melhor acurácia e precisão.

Por fim, foi utilizado o *software* “QGIS” para a geração de *layout* dos mapas.

5.3 Métodos

Para que o trabalho fosse organizado e com uma boa interpretação do que seria feito, foi feito um fluxograma para especificar a ordem de realização das tarefas a serem concluídas. Elas foram organizadas em: (I) obtenção de dados, (II) processamento digital de imagens e (III) geração de mapas.

Figura 6. Fluxograma de execução do trabalho.



Fonte: o autor.

5.3.1 Aquisição dos dados de produtividade

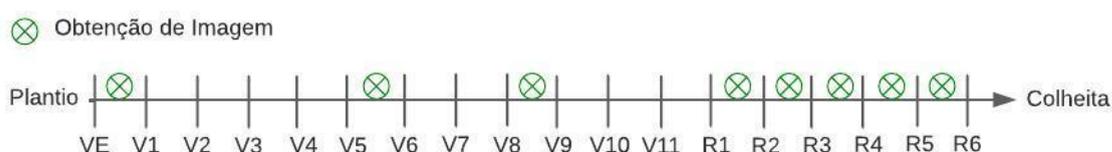
Os dados de produtividade para fim de estimativa foram colhidos *in loco*, por um maquinário, por meio de tecnologias, onde a máquina entende que o produzido pela planta

colhida é homogêneo para todas, estimando assim produtividade por hectare, tendo ela como ponto de referência, podendo assim estimar a produtividade final da lavoura. Este tipo de obtenção de dados é passível de erros, onde os dados coletados podem ser *outliers*, sendo assim, os dados cedidos para o estudo já foram filtrados. O resultado final foi um total de 100 pontos (Imagem 3) distribuídos por toda a área com respectivas coordenadas, que foram colocadas em uma camada *shapefile*.

5.3.2 Aquisição de imagens multiespectrais de alta resolução

Para fim de se estimar a produtividade do milho, foram usadas imagens multiespectrais de alta resolução dos satélites *Planet*, a fim de se descobrir se é possível essa estimativa através da imagem de um específico estágio fenológico ou se seria necessária a criação de uma série temporal. Para isso foram obtidas 11 imagens entre os dias 21/02/2020 e 04/08/2020, a imagem a seguir simplifica as datas de ocorrência dos estádios e em quais houve aquisição de imagens. Uma observação a ser feita é que no estágio fenológico R5 foram obtidas 4 imagens, por se tratar de um estágio mais longo, que antecede a colheita da cultura.

Figura 7: Estádios com imagem adquirida.



Fonte: o autor.

As imagens do satélite *Planet* já vem com todas suas imagens com as correções atmosféricas e geométricas feitas, então, não foi preciso nenhum tipo de tratamento com as imagens, já sendo possível a utilização das imagens para o estudo.

5.3.3 Cálculo de Índices de Vegetação e extração dos valores de reflectância

Com as imagens em mãos, foi feito o cálculo de dois Índices de Vegetação utilizando o *software* ENVI, sendo eles o NDVI e o GNDVI. O NDVI é um IV que é a relação da diferença entre o infravermelho com a banda do vermelho, com a soma dos mesmos.

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

Já o GNDVI é a relação entre a diferença do infravermelho próximo com a banda do verde e a soma deles.

$$GNDVI = \frac{NIR - GREEN}{NIR + GREEN}$$

Onde: NIR é o infravermelho próximo; Red é a banda do vermelho; Green a banda do verde.

No cálculo do NDVI não é necessário inserir a fórmula, através da aba *Transform* é possível selecionar a opção NDVI, e selecionar as bandas do vermelho (B3) e a banda do NIR (B4), e o *software* realiza o cálculo do índice automaticamente. Para calcular o GNDVI, foi usado um outro recurso do *software*, com a opção *Band Math* é possível que se insira a fórmula e assim é calculado o IV. Os cálculos são feitos em cima dos 100 pontos obtidos através das máquinas colheitadeiras.

Após calculados os índices de vegetação e obtido os valores de reflectância, foi utilizado o recurso *Layer Stacking*, esse recurso faz com que se permita a criação de uma imagem com as quatro bandas provindas do satélite, com mais os dois índices de vegetação (NDVI e GNDVI), ambos com seus valores de reflectância. Esse formato será necessário para a geração de modelos posteriormente.

5.3.4 Geração dos modelos de estimativa

Para a geração dos modelos, foram dos 100 pontos obtidos *in situ*, 80% utilizados para a estimativa e 20% para a validação do modelo. O *software* WEKA gera aleatoriamente quais serão os pontos utilizados para estimativa e quais serão utilizados para a validação dos dados (Tabela 2).

Tabela 2: Dados usados para validação.

Produtividade	Produtividade
3965,5	3998,8
4452,6	3572,7

3287,8	3547
4463,1	2806,7
3811,7	2898,5
4806,6	3695,4
2920,7	2838,5
4549,7	3437
4059,7	3292,7
2747,9	1120,8

Fonte: o autor.

Para gerar esses modelos que possibilitaram a estimativa da produtividade do milho, foi usado o aprendizado de máquinas, através de redes neurais e árvores aleatórias. Esses modelos foram: a *Simple Linear Regression*, que é um modelo estatístico cujo objetivo é indicar o comportamento de uma variável dependente (Y) em função de uma ou mais variáveis independentes (X), ela verifica se a uma regressão entre duas variáveis; a *Linear Regression*, é um modelo analítico que usamos ao modelar a relação linear entre uma variável de resultado contínua e várias variáveis de previsão que podem ser contínuas ou categóricas, ao contrário da *Simple Linear Regression*, essa pode conter mais de duas variáveis; *Support Vector Machine*, tem o objetivo de determinar as restrições de decisão que produzem a melhor separação entre as classes, minimizando o erro. Um SVM é uma técnica de aprendizado computacional para problemas de reconhecimento de padrões. *Multilayer Perceptron*, que segundo SILVA (2019), é uma rede neural com uma ou mais camadas ocultas com um número variável de neurônios. As camadas ocultas são assim chamadas porque a saída desejada das camadas intermediárias não pode ser prevista; e *Random Forest*, podemos pensar em uma árvore de decisão como uma representação gráfica para um determinado processo de decisão. As árvores são formadas por nós, que armazenam informação. O nó raiz é o nó que possui maior nível hierárquico e, a partir dele, ramificam-se os nós filhos. O nó que não possui filhos é conhecido como nó folha ou terminal (LIMA; AMORIM, 2020).

Também foi feita a mineração de dados no intuito de melhorar a precisão e acurácia dos modelos, essa mineração foi feita ainda através do WEKA, onde ele informa qual a banda que está sendo menos significativa para a geração do modelo, assim ela é retirada se verificado se o coeficiente de correlação melhorou ou não, caso não, a mineração é

descartada, se houve melhora, ela é refeita até achar o melhor coeficiente de correlação. O modelo escolhido foi baseado no que melhor apresentou um Coeficiente de Correlação na maioria dos estádios.

Após a geração dos modelos e da mineração de dados, ainda houve um ajuste do modelo pelo método do *Root Mean Square Error* (RMSE), que a partir do erro calculado pelo *software*, ainda foi feito o refinamento do mesmo pela equação a seguir.

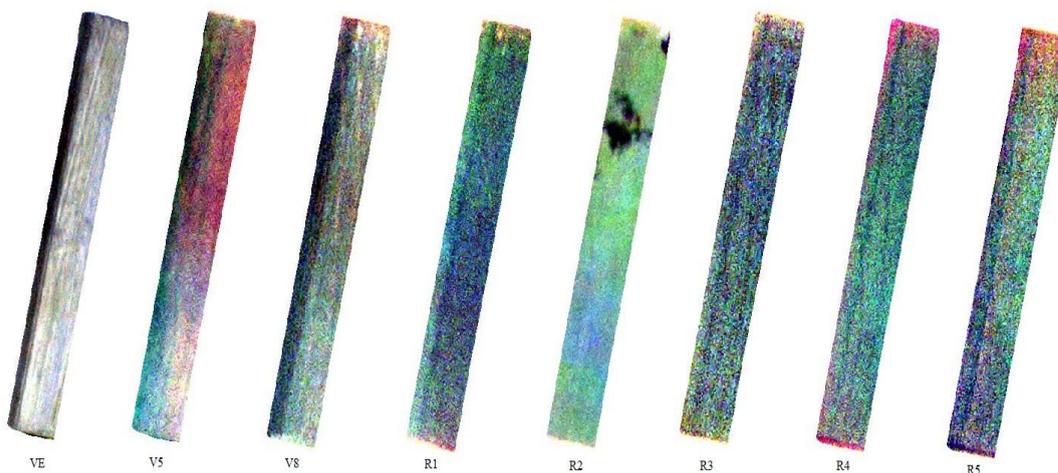
$$RMSE(\%) = \frac{\sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}}{\frac{\sum y_i}{n}} * 100$$

Onde: \hat{y}_i são os valores preditos, y_i são os valores mensurados em campo, n o número total de observações.

6 RESULTADOS E DISCUSSÕES

Para efeito de comparação, a figura 7 mostra a evolução do dossel do milho nos estádios fenológicos com imagens adquiridas, estando elas em RGB.

Figura 8: Imagens RGB dos estádios fenológicos do milho.



Fonte: o autor.

As tabelas de 3 a 13, apresenta a acurácia (RMSE) dos modelos de regressão para estimativa da produtividade e o coeficiente de correlação entre a produtividade medida e a estimada para os principais estádios fenológicos do milho.

Ao analisar os modelos obtidos para tomada de imagens no estágio fenológico VE (Tabela 3), observa-se a capacidade de estimar a produtividade média com um RMSE de até 25%, a partir do modelo baseado na regressão linear simples, porém os baixos valores do Coeficiente de Correlação denotam que não é possível acertar a distribuição espacial das áreas mais ou menos produtivas.

Tabela 3. Coeficiente de correlação e RMSE em VE

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	-0,1238	25%
Linear Regression	-0,0779	25,15%
Multilayer Perceptron	0,1611	27,36%
Support Vector Machine	-0,0633	25,93%
Random Forest	0,0902	30,06%

Fonte: o autor.

Na Tabela 4, resultados que remetem ao estágio fenológico V5: Surgimento das primeiras folhas, ou seja, a sensibilidade das bandas espectrais dos modelos com a pigmentação presente na folha pode ter sido um fator que tenha contribuído para se estimar a produtividade haja visto que a equação composta pela banda do verde fora a mais propícia para estimar a produtividade em um modelo de regressão linear simples ($\text{Produtividade} = -17,8 * B2 + 13504,67$). Já temos uma capacidade para definir as áreas mais e menos produtivas, uma vez que o coeficiente de correlação se aproxima de 0,3.

Tabela 4. Coeficiente de correlação e RMSE em V5

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	0,2697	24,05%
Linear Regression	0,1925	27,35%
Multilayer Perceptron	0,1382	32,94%
Support Vector Machine	0,159	26,61%
Random Forest	0,3017	32,94%

Fonte: o autor.

Na tabela 5, refere-se aos modelos criados para o estágio fenológico V8: Surgimento de 8 folhas, ou seja, maior taxa de fotossíntese que os estágios inferiores. Ao analisar o modelo de regressão linear simples, nota-se que a banda do azul (Produtividade = $-18,71 * B1 + 8436,68$) foi a mais sensível a produtividade, uma vez que, o modelo de estimativa constituído pela banda mais sensível a fotossíntese consegue estimar a produtividade com uma acurácia $\approx 77\%$ além de, melhorar significativamente, a capacidade de definir as regiões com mais e menos produtividade a partir do Coeficiente de Correlação do modelo de *Linear Regression* (0,47).

Tabela 5. Coeficiente de correlação e RMSE em V8

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	0,4076	22,83%
Linear Regression	0,4767	21,52%
Multilayer Perceptron	0,3467	27,27%
Support Vector Machine	0,4639	20,74%
Random Forest	0,3623	23,02%

Fonte: o autor.

Além do mais, esta condição demonstra que o vegetal em pleno desenvolvimento, ou seja, com alta taxa de fotossíntese reflete a futura produtividade. Para corroborar esta constatação, o modelo de *Linear Regression* (Produtividade = $-35,9641 * B1 + 9,2072 * B3 + 9484,394$) mostra que exatamente as bandas do azul e vermelho, ou seja, as mais sensíveis a clorofila e que são as mais propícias para compor os modelos.

Porém, dada a complexibilidade nutricional do vegetal, devido ao seu crescimento, os modelos de *data learning* por terem sido mais assertivos, demonstram que a estimativa da produtividade passa a ser mais acurada com relações não lineares entre a reflectância das imagens a variável de estudo.

Existem aquelas relações não lineares, que não são as mais acuradas na estimativa da produtividade, mas estruturalmente são as que melhor conseguem estimar as regiões mais e menos produtivas, que é o caso das Redes Neurais, as quais são menos acuradas quando comparadas com o algoritmo *Support Vector Machine*.

No estágio R1(Surgimento das Espigas)(Tabela 6), observou-se que não foi o período ideal para estimar a distribuição espacial da produtividade. Os modelos de regressão

apontaram que a banda do vermelho ($\text{Produtividade} = -3,23 * B3 + 4254,79$) é a que melhor estima a produtividade, ou seja, a senescência do vegetal e ainda faixa de fotossíntese são determinantes para se prever o que irá acontecer com a produtividade, porém por se tratar de época em que todas as plantas apresentam espiga em estágio inicial, a homogeneidade, ou baixa variabilidade de biomassa dificultam a conclusão de modelos mais assertivos.

Tabela 6. Coeficiente de correlação e RMSE em R1

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	-0,2258	25,48%
Linear Regression	0	23,97%
Multilayer Perceptron	0,2084	29,13%
Support Vector Machine	-0,0868	24,79%
Random Forest	0,0479	30,25%

Fonte: o autor.

Para o estágio R2 o desempenho dos modelos (Tabela 7) repete-se o caso de R1, o dossel ainda se apresenta muito homogêneo dada a formação inicial das espigas, ou seja, o modelo é acurado para estimar a produtividade média, porém imprecisa para estimar a distribuição espacial da produtividade, dados os baixos valores do Coeficiente de Correlação.

Tabela 7. Coeficiente de correlação e RMSE em R2

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	-0,0754	25,40%
Linear Regression	-0,1271	26,45%
Multilayer Perceptron	-0,1548	32,17%
Support Vector Machine	-0,1382	25,90%
Random Forest	-0,1246	33,94%

Fonte: o autor.

Nos estádios de R1 e R2 os parâmetros de sensibilidade das bandas do visível, ou seja, clorofila e pigmentos são determinantes para estimar a produtividade.

R3: B1 e B2 (Tabela 8) ainda é determinante para se estimar a produtividade, porém com aumento das espigas, ou seja, aumento da variabilidade da resposta espectral do dossel,

os modelos são mais acurados para estimar a produtividade média e conseqüentemente a distribuição espacial da produtividade.

Tabela 8. Coeficiente de correlação e RMSE em R3

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	0,6552	20,66%
Linear Regression	0,5454	20,39%
Multilayer Perceptron	0,5657	22,29%
Support Vector Machine	0,5418	21,89%
Random Forest	0,1743	33,69%

Fonte: o autor.

R4 (Tabela 9) em diante, ou seja, período de maturidade até a colheita, neste período o vegetal começa a secar, ou seja, diminui-se a sensibilidade das propriedades da planta com as bandas multiespectrais, mitigando a capacidade de estimar a produtividade e sua distribuição espacial a partir de imagens.

Tabela 9. Coeficiente de correlação e RMSE em R4

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	0,3776	22,21%
Linear Regression	0,3753	22,61%
Multilayer Perceptron	0,2986	26,42%
Support Vector Machine	0,3813	21,75%
Random Forest	0,5245	20,45%

Fonte: o autor.

As tabelas a seguir (10, 11, 12 e 13) mostram os resultados das 4 imagens obtidas durante o estágio de R5.

Tabela 10. Coeficiente de correlação e RMSE na primeira imagem em R5

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	0,2278	22,98%

Linear Regression	0,2278	22,98%
Multilayer Perceptron	0,201	27,62%
Support Vector Machine	0,2557	23,10%
Random Forest	0,0804	27,48%

Fonte: o autor.

Tabela 11. Coeficiente de correlação e RMSE na segunda imagem em R5

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	0,1459	24,62%
Linear Regression	0,3687	21,77%
Multilayer Perceptron	-0,0355	27,72%
Support Vector Machine	0,2137	23,76%
Random Forest	0,1543	25,40%

Fonte: o autor.

Tabela 12. Coeficiente de correlação e RMSE na terceira imagem em R5

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	-0,0256	40,81%
Linear Regression	-0,0256	40,81%
Multilayer Perceptron	0,0359	32,96%
Support Vector Machine	-0,0497	42,51%
Random Forest	0,2675	26,29%

Fonte: o autor.

Tabela 13. Coeficiente de correlação e RMSE na quarta imagem em R5

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	-0,4389	24,94%
Linear Regression	0	23,97%
Multilayer Perceptron	0,3903	23,69%
Support Vector Machine	0,0526	23,52%
Random Forest	0,2237	24,28%

Fonte: o autor.

Na série temporal, o melhor RMSE e índice de correlação, foi o modelo de *Random Forest*, porém o índice de correlação chegou a 0,5508, sendo inferior ao *Multilayer Perceptron* em R3. Mesmo que seu RMSE tenha sido de 20%, foi dada a prioridade para o melhor índice de correlação.

Tabela 14. Coeficiente de correlação e RMSE na série temporal

Modelo	Coeficiente de Correlação	RMSE
Simple Linear Regression	-0,0256	40,81%
Linear Regression	0,4701	26,66%
Multilayer Perceptron	0,4151	33,7%
Support Vector Machine	0,3152	27,55%
Random Forest	0,5508	20,13%

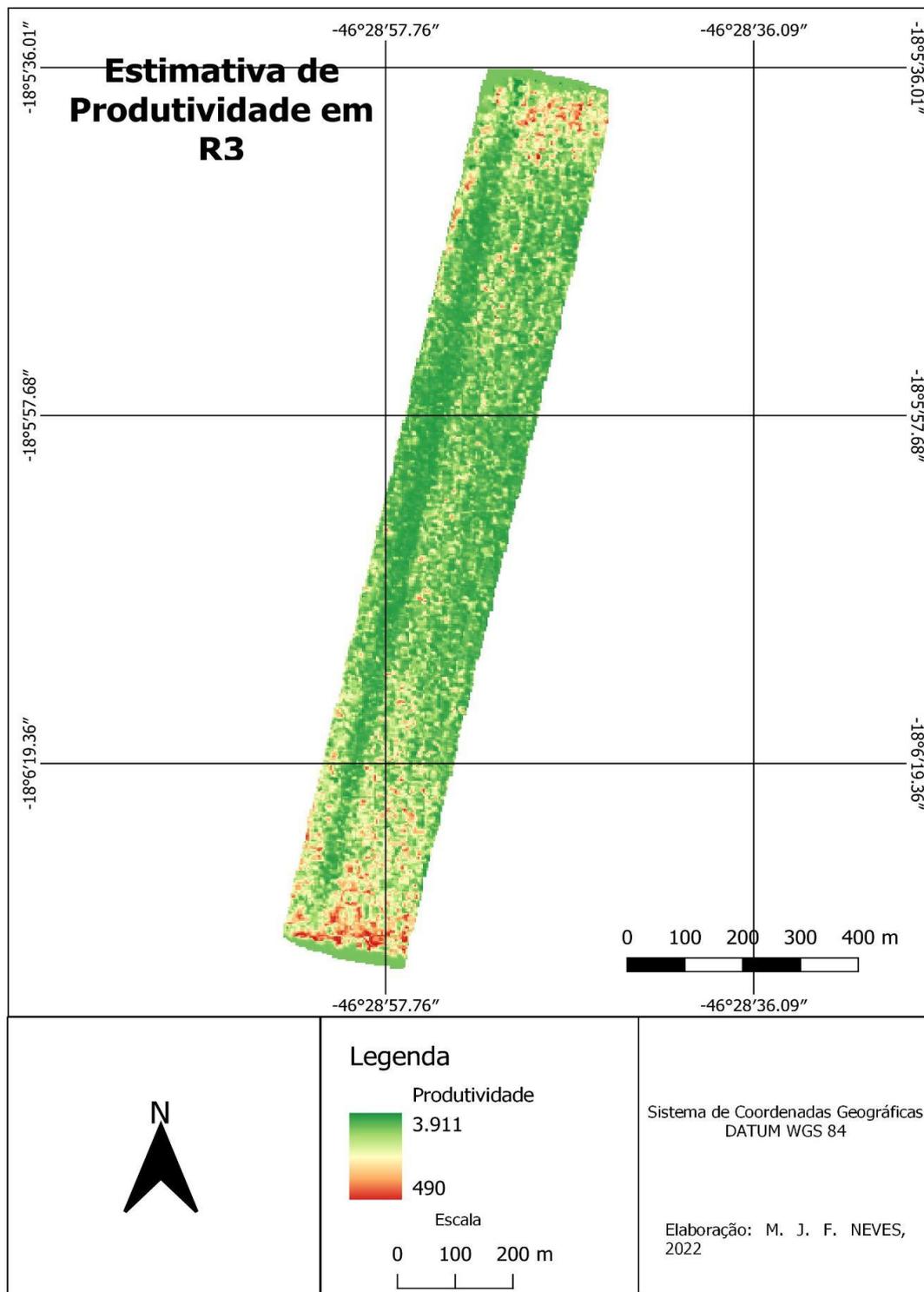
Fonte: o autor.

Analisando assim os resultados individuais comparados a série temporal, têm se o resultado de que não é preciso a série temporal para conseguir estimar a produtividade do milho, sendo ela possível estimada em um estágio apenas. Tendo em conta isso, o modelo que mais atingiu um Coeficiente de Correlação aceitável foi *Multilayer Perceptron*, tendo em R3 seu melhor resultado.

A figura 4 apresenta um mapa de distribuição espacial da produtividade a partir do modelo espectral gerado pelo algoritmo *MultilayerPerceptron* sob uma imagem no estágio fenológico R3. O mapa apresenta a produtividade variando 490 a 3.911 kg/ha, onde a variabilidade produtividade é notada por um padrão de faixas intercaladas definidas por alta e baixa produtividade.

Deve se ressaltar essa variabilidade ocorrera pelo fato da área ter sido utilizada com teste produtos químicos biológicos, ou seja, foram dispostos tratamentos na forma de faixas ao longo do talhão, onde priorizou-se diversos manejos que influenciaram diretamente na produtividade.

Figura 6. Estimativa de Produtividade em R3



Fonte: o autor.

7 CONCLUSÃO

Através de imagens multiespectrais de alta resolução, associadas com os estádios fenológicos do milho, é possível que se estime a produtividade da lavoura sem a necessidade do uso de uma série temporal, usando apenas uma imagem do determinado estágio. Dentre os modelos de Aprendizado de Máquinas, o *Multilayer Perceptron* foi o que obteve a melhor estimativa, apresentado um RMSE de aproximadamente 22%, ou seja, um modelo que nos proporciona um acerto de 78%, com um índice de correlação igual a 0,5657, no estágio fenológico de R3, tendo em vista que é uma época onde a planta apresenta um alto vigor vegetativo por preceder o início dos estádios onde a planta começa a secar para que possa ser feita a colheita dos grãos.

REFERÊNCIAS

- Avila, L. F.; Mello, C. R. de; Yanagi, S. de N. M.; Neto, O. B. S. **Tendências de temperaturas mínimas e máximas do ar no Estado de Minas Gerais: Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 49, n. 4, p. 247-196, 2014.
<https://doi.org/10.1590/S0100-204X2014000400002>
- Borghetti, E.; Resende, A. V. de; Neto, M. M. G.; Santana, D. P.; Souza, G. F. de; Silva, M. A.; Souza, C. H. E. de. **Sistemas de cultivo de soja e milho na região do Alto Paranaíba-MG e resultados de avaliações na safra 2014/15: Embrapa Milho e Sorgo**, Sete Lagoas, MG, 2016.
- Jensen, J. R. **Sensoriamento Remoto do Ambiente: Uma Perspectiva em Recursos Terrestres**. 2. ed. University of South Carolina. 2011.
- Jensen, J. R. **Introductory Digital Image Sensing of Biophysical Variables: An Overview of the Special Issue**. Remote Sensing of Environment. 2005.
- Hunt, E., C. Daughtry, J. Eitel, and D. Long. "Remote Sensing Leaf Chlorophyll Content Using a Visible Band Index." *Agronomy Journal* 103, No. 4, 2011.
<https://doi.org/10.2134/agronj2010.0395>
- Birth, G., and G. McVey. "Measuring the Color of Growing Turf with a Reflectance Spectrophotometer." *Agronomy Journal* 60, 1968.
<https://doi.org/10.2134/agronj1968.00021962006000060016x>
- Rouse, J., R. Haas, J. Schell, and D. Deering. **Monitoring Vegetation Systems in the Great Plains with ERTS**. Third ERTS Symposium, NASA, 1973.
- Gitelson, A., and M. Merzlyak. "Remote Sensing of Chlorophyll Concentration in Higher Plant Leaves." *Advances in Space Research* 22, 1998.
[https://doi.org/10.1016/S0273-1177\(97\)01133-2](https://doi.org/10.1016/S0273-1177(97)01133-2)
- Contini, E. et al.. **Milho: caracterização e desafios tecnológicos**. Embrapa - Desafios do Agronegócio Brasileiro, 2, 2019.
- Garcia, J., Mattoso, J., Duarte, J.. **Importância do milho em Minas Gerais**. Embrapa - Informe Agropecuário, 27, 2006.
- Kleinschmitt, E.. **Produtividade da cultura do milho (Zea mays) em resposta à inoculação de Axospirillum brasilense em combinação com fertilizantes bioindutores**. Universidade Federal de Santa Catarina. 2016.
- Melo, R. W., Fontana, D. C., Berlato, M. A., Ducati, J. R.. **An agrometeorological-spectral model to estimate soybean yield, applied to southern Brazil**. *International Journal of Remote Sensing*, 29, 2008.
<https://doi.org/10.1080/01431160701881905>

Leal, A., Miguel, E., Baio, F., Neves, D., Leal, U.. **Redes neurais artificiais na predição da produtividade de milho e definição de sítios de manejo diferenciado por meio de atributos do solo.** Solos e Nutrição de Plantas, 74, 2015.

<https://doi.org/10.1590/1678-4499.0140>

Rodrigues, V. N., Pinho, R. G. V., Paglis, C. M., Bueno Filho, J. S. S., Brito, A. H.. **Comparação entre métodos para estimar a produtividade de grãos de milho.** Ciência e Agrotecnologia, 2004.

<https://doi.org/10.1590/S1413-70542005000100004>

Bertolin, N. O., Filgueiras, R., Venancio, L. P., Mantovani, E. C.. **Predição da produtividade de milho irrigado com auxílio de imagens de satélite.** Revista Brasileira de Agricultura Irrigada, 2017.

<https://doi.org/10.7127/rbai.v11n400567>

Shanahan, J.. **Use of Remote-Sensing Imagery to Estimate Corn Grain Yield.** Agronomy Journal, 93 (3), 583-589, 2001.

<https://doi.org/10.2134/agronj2001.933583x>

Meneses, P., Almeida T.. **Introdução ao processamento de imagens de sensoriamento remoto.** Universidade de Brasília. 2012.

Pádua, L. et al.. **Utilização de imagens térmicas adquiridas por veículos aéreos não tripulados em aplicações agrícolas.** In: Proceedings of the IX Conferência Nacional de Cartografia e Geodesia, Lisbon, Portugal, 29-30. 2018.

Amaral, F. C. S.. **Sistema Brasileiro de Classificação de Terras para Irrigação: enfoque na região semiárida.** Embrapa Solos-Livro técnico. 2011.

Formaggio, A. R., Sanches, I. D. **Sensoriamento remoto em agricultura.** Oficina de Textos. 2017.

Wahab, I., Hall, O., Jirström, M.. **Remote sensing of yields: Application of uav imagery-derived ndvi for estimating maize vigor and yields in complex farming systems in sub-saharan africa.** Drones. 2018.

<https://doi.org/10.3390/drones2030028>

Ahmad, I. et al.. **Yield forecasting of spring maize using remote sensing and crop modeling in Faisalabad-Punjab Pakistan.** Journal of the Indian society of remote sensing. 2018.

<https://doi.org/10.1007/s12524-018-0825-8>

Mekonnen, Y., Namuduri, S., Burton, L., Sarwat, A., Bhansali, S.. **Machine learning techniques in wireless sensor network based precision agriculture.** Journal of the Electrochemical Society. 2019.

<https://doi.org/10.1149/2.0222003JES>

Treboux, J., Genoud, D.. **Improved machine learning methodology for high precision agriculture.** Global Internet of Things Summit (GIoTS). 2018.

<https://doi.org/10.1109/GIOTS.2018.8534558>

Chlingaryan, A., Sukkarieh, S., Whelan, B.. **Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review.** Computers and electronics in agriculture. 2018.

<https://doi.org/10.1016/j.compag.2018.05.012>

Magalhães, P. C.; Durães, F. O. M.; Carneiro, N. P.; Paiva, E. **Fisiologia do milho.** Sete Lagoas: EMBRAPA-CNPMS, 2002.

Barzin, R.; Pathak, R.; Lotfi, H.; Varco, J.; Bora, G.C. **Use of UAS Multispectral Imagery at Different Physiological Stages for Yield Prediction and Input Resource Optimization in Corn.** Remote Sens. 2020.

<https://doi.org/10.3390/rs12152392>

Sunoj, S; Cho, Jason; Guinness, Joe; Jan van Aardt; Czymbek, Karl J; et al. **Remote Sensing. Basel Vol. 13,** Ed. 19, 2021.

<https://doi.org/10.3390/rs13193948>

Figueiredo, D. **Conceitos Básicos de Sensoriamento Remoto.** Brasília: Companhia Nacional de Abastecimento, 2005.

RSI (Research Systems, Inc). **ENVI User's Guide.** ENVI Version 4.0 September, 2003.

Weiers, S.; Bock, M.; Wissen, M.; G. Rossner.. **"Mapping and Indicator Approaches for the Assessment of Habitats at Different Scales using Remote Sensing and GIS Methods"**, Landscape and Urban Planning. 2004.

[https://doi.org/10.1016/S0169-2046\(03\)00028-8](https://doi.org/10.1016/S0169-2046(03)00028-8)

Shippert, P.. **Spotlight on Hyperspectral.** Boulder : Research Systems. 2004.

Running, S. W.; Justice, C. O.; Solomonson, V.; Hall, D.; Barker, J.; Kaufmann, Y. J.; Strahler, A. H.; Huete, A. R.; Muller, J. P.; Vanderbilt, v. Wan, Z. M.; Teillet, P.; D. Carneggie. **Terrestrial Remote Sensing Science and Algorithms Planned for EOS/MODIS.** Intl. Journal of Remote Sensing, 1994.

<https://doi.org/10.1080/01431169408954346>

Huete, A., et al. **Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices.** Remote Sensing of Environment 83. 2002.

[https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)

Gitelson, A., M. Merzlyak. **Remote Sensing of Chlorophyll Concentration in Higher Plant Leaves.** Advances in Space Research 22. 1998.

[https://doi.org/10.1016/S0273-1177\(97\)01133-2](https://doi.org/10.1016/S0273-1177(97)01133-2)

Rouse, J., R. Haas, J. Schell, and D. Deering. **Monitoring Vegetation Systems in the Great Plains with ERTS.** Third ERTS Symposium, NASA. 1973.

Huete, A. **A Soil-Adjusted Vegetation Index (SAVI).** Remote Sensing of Environment 25. 1988.

[https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)

Birth, G., and G. McVey. **Measuring the Color of Growing Turf with a Reflectance Spectrophotometer.** Agronomy Journal 60. 1968.
<https://doi.org/10.2134/agronj1968.00021962006000060016x>

Silva, D. F. S. da.; Flauzino. R. A. "**Redes neurais artificiais para engenharia e ciências aplicadas.**" São Paulo: Artliber (2010).

Lima, M; Amorim, F. "**Random Forest.**" LAMFO - UNB. 2020.

Mitchell, T., et al. "**Annu. Rev. Comput. Sci.**" University Pittsburgh, Pennsylvania. 1990.

McCulloch, W. S., Pitts, W., "**A logical calculus of the ideas immanent in nervous activity.**" Bulletin of Mathematical Biophysics. 1943.
<https://doi.org/10.1007/BF02478259>

Rauber, T. W. "**Redes Neurais Artificiais.**" Universidade Federal do Espírito Santo. 2014.

Monard, M. C.; Baranauskas, J. A. "**Sistemas Inteligentes.**" cap4. 2003.

Hanson, R.; Stutz, J.; Cheeseman, P. "**Bayesian Classification Theory.**" Technical Report. 1990.

Breima, L. "**Random forests.**" Machine Learning. 2001.