

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE GEOGRAFIA

FILIPE VIEIRA DA SILVA

**ANÁLISE DO POTENCIAL DE IMAGENS MULTIESPECTRAIS NA PREDIÇÃO
DA PRODUTIVIDADE DA CULTURA DO TRIGO**

Monte Carmelo - MG

2023

FILIPPE VIEIRA DA SILVA

**ANÁLISE DO POTENCIAL DE IMAGENS MULTIESPECTRAIS NA PREDIÇÃO
DA PRODUTIVIDADE DA CULTURA DO TRIGO**

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 - MG

2023

FILIPE VIEIRA DA SILVA

**ANÁLISE DO POTENCIAL DE IMAGENS MULTIESPECTRAIS NA PREDIÇÃO
DA PRODUTIVIDADE DA CULTURA DO TRIGO**

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.

Aprovado em: 29/11/2023

Banca Examinadora:

Prof. Dr. George Deroco Martins
Universidade Federal de Uberlândia

Eng. Laura Cristina Moura Xavier

Prof. Dr. Odair José Marques
Universidade Federal de Uberlândia

AGRADECIMENTOS

Nenhuma jornada de sucesso é construída sem apoio, por isso, é com muita alegria que venho demonstrar minha gratidão a cada um que de alguma forma se fez presente em minha formação acadêmica.

Primeiramente agradeço a Deus pelo dom da vida e da sabedoria. Obrigado por guiar meus passos e nunca me abandonar nos momentos de aflições.

Agradeço aos meus familiares por acreditarem em mim e sempre me incentivar a seguir o caminho da educação. Em especial, agradeço a minha mãe Sirlene, por ser o meu Norte, por apoiar minhas decisões e estar sempre ao meu lado em cada etapa da minha vida.

À empresa Lallemand Plant Care por fornecer os dados que tornaram possível o desenvolvimento deste trabalho.

Ao meu orientador prof. Dr. George Deroco pelos ensinamentos, paciência, disposição e por uma orientação repleta de oportunidades e leveza.

Por fim, aos meus colegas e amigos que tornaram o caminho mais leve, minha eterna gratidão. Obrigado pelas broncas, auxílio em cada trabalho, seminário e prova, pelo companheirismo e pelas “horas bar”.

RESUMO

A estimativa de produtividade pode ser uma ferramenta significativa para o setor agroindustrial de um país, auxiliando no manejo da cultura, na comercialização, computação de valores, além de direcionar e facilitar o escoamento e armazenamento da colheita. Para estimar a produtividade dos cereais de forma simples, é preciso coletar amostras em campo, contabilizar a quantidade de plantas em um metro quadrado, pesar os grãos e por fim estimar para toda a área cultivada. No método desenvolvido neste trabalho, foi avaliado o potencial de imagens multiespectrais na estimativa da produtividade do trigo com o uso de algoritmos de aprendizado de máquinas. Além disso, foi estabelecido o estágio fenológico ideal para aquisição das imagens para uma predição mais acurada da produtividade e testados modelos regionais e temporais. O modelo temporal consistiu na estimativa de produtividade utilizando todos os dados (do plantio à colheita), enquanto o modelo regional gerou a estimativa para uma região de estudo baseada em outra região, ou seja, foi avaliado se as informações de colheita de uma região eram capazes de gerar um modelo preditivo que estimasse outra região de estudo. Com o uso de imagens multiespectrais do satélite Sentinel 2 e com amostras de produtividade georreferenciadas e coletadas de forma pontual em 3 áreas de estudo localizadas na França, determinou-se que o estágio inicial da maturação (enchimento dos grãos) é o melhor momento para tomada das imagens. Os resultados mostraram que o algoritmo Random Forest apresentou as melhores estimativas de produtividade do trigo com um menor erro, entre 5 e 8% aproximadamente. Por fim, foram detectados 9 atributos que devem compor o modelo de estimativa, sendo 7 bandas originais do satélite e 2 índices derivados, e foi possível gerar mapas de produtividade das 3 áreas de estudo.

Palavras-chave: cereais; monitoramento agrícola; modelos espectrais; estimativa de produtividade; trigo; mapa de produtividade do trigo.

ABSTRACT

Productivity estimation can be a significant tool for a country's agro-industrial sector, helping with crop management, marketing, computing values, as well as directing and facilitating the flow and storage of the harvest. To estimate cereal productivity in a simple way, it is necessary to collect samples in the field, count the number of plants in a square meter, weigh the grains and finally estimate it for the entire planted area. In the method developed in this work, the potential of multispectral images in estimating wheat productivity was evaluated using machine learning algorithms. Furthermore, the ideal phenological stage for image acquisition was established for a more accurate prediction of productivity and regional and temporal models were tested. The temporal model consisted of estimating productivity using all data (from planting to harvest), while the regional model generated the estimate for a study region based on another region, that is, it was evaluated whether the harvest information for a region could generate a predictive model that estimated another study region. Using multispectral images from the Sentinel 2 satellite and georeferenced productivity samples collected punctually in 3 study areas located in France, it was determined that the initial stage of maturation (grain filling) is the best time to take of the images. The results showed that the Random Forest algorithm presented the best estimates of wheat productivity with a smaller error, between 5 and 8%. Finally, 9 attributes were detected that should compose the estimation model, 7 of which were original satellite bands and 2 derived indices, and it was possible to generate productivity maps of the 3 study areas.

Keywords: cereals; agricultural monitoring; spectral models; productivity estimation; wheat; wheat productivity map.

LISTA DE ILUSTRAÇÕES

Figura 1 -	Estádios fenológicos Feekes - Large (1954)	16
Figura 2 -	Relação IV e estágio fenológico	18
Figura 3 -	Mapa de localização da região de estudo A	22
Figura 4 -	Mapa de localização das regiões de estudo B e C	23
Figura 5 -	Pontos de produtividade	25
Figura 6 -	Fluxograma	27
Figura 7 -	Relação pixel e ponto de produtividade	29
Figura 8 -	Relação entre os IVs	30
Figura 9 -	Resultados processamentos	36
Figura 10 -	Resultados das regiões A, B e C	37
Figura 11 -	Erro de estimativa cometido pelo algoritmo de predição	41
Figura 12 -	Mapa de produtividade da região A	42
Figura 13 -	Mapa de produtividade da região B	43
Figura 14 -	Mapa de produtividade da região C	43

LISTA DE QUADROS

Quadro 1 -	Estudos relacionando o trigo com o uso de IV	17
Quadro 2 -	Índices selecionados para a geração dos modelos preditivos	33
Quadro 3 -	Bandas Sentinel 2 utilizada nos cálculos dos IVs	35
Quadro 4 -	Resultados do teste regional e da série temporal	39
Quadro 5 -	Variáveis que devem compor a estimativa	39

SUMÁRIO

1	INTRODUÇÃO.....	12
2	OBJETIVOS.....	14
2.1	Objetivo geral.....	14
2.2	Objetivos específicos	14
3	REVISÃO BIBLIOGRÁFICA	15
3.1	Cultura do trigo	15
3.2	Sensoriamento remoto no trigo	16
3.3	IVs relacionados com a estimativa de produtividade.....	19
3.4	Modelos espectrais para se estimar a produtividade.....	20
4	MATERIAL E MÉTODOS	22
4.1	Área de estudo.....	22
4.2	Material	23
4.3	Métodos.....	26
4.3.1	Aquisição dos dados.....	28
4.3.2	Pré-processamento dos dados	28
4.3.3	Processamento dos dados	30
4.3.4	Validação e melhoramento do modelo	31
4.3.5	Representação Gráfica	32
5	RESULTADOS E DISCUSSÕES.....	33
5.1	IVs selecionados para a modelagem.....	33
5.2	Algoritmo de aprendizado de máquina com melhor desempenho	35
5.3	Estádio fenológico ideal para a estimativa do trigo	36
5.4	Modelo regional e série temporal	38
5.5	IVs / bandas que compõem a melhor estimativa.....	39
5.6	Espacialização da produtividade do trigo	40
6	CONCLUSÃO.....	44
	REFERÊNCIAS	45

1 INTRODUÇÃO

Em novembro de 2022, a população atingiu a marca de 8 bilhões de habitantes em todo o planeta (Ladeira, 2022). Além disso, houve um aumento na esperança de vida para cerca de 72 anos em 2019 e estima-se que esse crescimento deve continuar ocorrendo até atingir 82,1 anos em 2100 (United Nations, 2022?). Com a evolução tanto no número de pessoas quanto no prolongamento da vida humana, a produção de alimentos em quantidade suficiente para alimentar toda a população acaba sendo uma preocupação, uma vez que o número de pessoas aumenta e a área agricultável é limitada (Saath; Fachinello, 2018).

Neste cenário mundial, cerca de 10% da população sofreu com a fome em 2021 e quase 29% das pessoas atingiram a classificação de insegurança alimentar grave ou moderada (FAO, 2022a). A produtividade, por sua vez, graças ao avanço tecnológico e às pesquisas na área, aumentou cerca de 52% para as culturas primárias, como o milho, cana-de-açúcar, o arroz, o trigo, entre outras (FAO, 2022b). As culturas citadas são responsáveis aproximadamente pela metade da produção primária, tendo o milho como o principal cereal produzido e o trigo como o segundo maior (FAO, 2022b; Jacto, 2021).

O trigo pertence ao gênero botânico denominado *Triticum*, no qual há várias espécies, mas a espécie mais cultivada no mundo é a *Triticum aestivum*. Essa espécie é composta por cultivares que passaram por seleção natural e estudos de melhoramento. Através desses processos, a cultura do trigo evoluiu a partir de uma forma primitiva na qual os grãos estavam ligados à parte floral e as espigas eram suscetíveis a quebra, para uma variedade atual com sementes mais resilientes. (Scheeren; Castro; Caierão, 2015).

De acordo com a análise mensal de abril de 2022 realizada pela United States Department of Agriculture - USDA baseado na safra 2021/22, houve um aumento de cerca de 0,13% de área plantada com a cultura do trigo. Enquanto isso, o consumo teve um aumento de 1,38% ou cerca de 788 milhões de toneladas (Conab, 2022). Com isso, é possível perceber que o aumento da produtividade não tem acompanhado o consumo.

Como principais produtores da cultura, podemos citar a União Europeia, a China e a Índia, seguida da Rússia e os Estados Unidos. Enquanto isso, o Brasil ocupa a 15ª posição do ranking. A Rússia é o país que mais exporta e a Ucrânia ocupa a 5ª colocação entre os principais exportadores de trigo globalmente, além de ocuparem a quarta e a sétima colocação, respectivamente, na lista de maiores produtores. Por isso, com o cenário atual em que há uma guerra justamente entre Rússia e Ucrânia, que impacta diretamente na produção e no escoamento, existe uma preocupação mundial com a oferta, a produção, e o preço do trigo

(Conab, 2022). Logo, estimar a produtividade pode ser uma ferramenta de grande valia no manejo da colheita, no planejamento estratégico da comercialização, bem como direcionar o armazenamento e o processamento da produção, mostrando ser um instrumento importante para o setor agroindustrial de cada país. (Letras Ambientais, 2021).

Assim como ocorre com outras culturas agrícolas, a estimativa de produtividade do trigo pode ser realizada por meio da visita ao campo de produção, onde deve-se selecionar aleatoriamente alguns pontos bem espalhados na área para gerar uma amostra que represente o todo. Em seguida, em cada ponto deve-se medir uma área de 1 m² e contar os componentes de produção, no caso do trigo o número de plantas com espigas, o número de espiguetas viáveis (com grãos formados) e o número de grãos por espiguetas, em seguida deve-se pesar os grãos, determinar a umidade dos grãos e ajustar a produtividade para quilogramas de grãos por hectare (kg ha⁻¹). Em dezembro de 2022, o Departamento de Energia, Meio Ambiente e Ação Climática do estado de Victoria, Austrália, apresentou uma fórmula para se calcular a produtividade do trigo de forma similar a esta e acrescenta ainda que é necessário considerar a perda de grãos (entre 5% e 10%) que podem ocorrer antes e durante a colheita (Agriculture Victoria, 2022).

Ademais, pesquisadores vêm trabalhando a estimativa de produtividade do trigo com o uso de sensoriamento remoto na agricultura. Paiva, Tsukahara e França (2013) desenvolveram em seu trabalho uma metodologia que vincula dados meteorológicos com dados espectrais para estimar a produtividade da cultura. Enquanto Souza (2018) trabalhou a estimativa de biomassa do trigo utilizando imagens multiespectrais obtidas através de veículo aéreo não tripulado (VANT) relacionadas com índices de vegetação.

É notório que o método convencional de estimativa apresentado demanda um certo tempo, recurso, conhecimento de campo e da área de estudo, e que outros métodos de estimativa estão sendo desenvolvidos com a tentativa de vincular as características biofísicas da planta e do espaço com a produtividade. Neste trabalho, foi desenvolvida a estimativa por meio de imagens de satélite que pode diminuir o tempo gasto, o esforço físico, o custo e ainda pode ser avaliado durante toda a fase de desenvolvimento da planta, enquanto o método convencional fica limitado à fase reprodutiva do trigo, aprimorando assim o cálculo de produtividade da cultura.

2 OBJETIVOS

2.1 Objetivo geral

O presente trabalho teve como objetivo geral avaliar o potencial de imagens multiespectrais, tomadas por satélites, na estimativa da produtividade do trigo.

2.2 Objetivos específicos

Os objetivos específicos deste trabalho foram:

- Avaliar o melhor estágio fenológico para estimativa da produtividade do trigo.
- Avaliar a viabilidade de um modelo espectral regional para estimar a produtividade do trigo.
- Avaliar o ganho de acurácia na estimativa da produtividade a partir de séries temporais.

3 REVISÃO BIBLIOGRÁFICA

3.1 Cultura do trigo

O trigo é, provavelmente, a espécie vegetal cultivada desde os primórdios da agricultura, há cerca de 10.000 anos, quando surgiu a intensificação da produtividade de cereais devido ao período neolítico que se instaurou na sociedade (Scheeren; Castro; Caierão, 2015; Pinto *et al.*, 2017). Vale ressaltar que a cultura se tornou uma das principais cultivadas no mundo e, por meio do processamento do grão que começou com os egípcios em meados de 4.000 a. C., faz parte da composição de diversos alimentos dando origem a farinha, ao pão, compondo o macarrão, o pastel, dentre outros (Pinto *et al.*, 2017).

Fazendo parte da família gramínea, o trigo possui as mesmas características marcantes das culturas de inverno e carrega o propósito de produzir grãos, assim como a cevada, o centeio, triticale e aveia. Além disso, podemos dividir a planta em 4 componentes: raízes, colmo, folhas e inflorescência (Scheeren; Castro; Caierão, 2015).

O sistema radicular da planta pode ser dividido em 3 partes: raízes adventícias aéreas, adventícias permanentes e seminais. As raízes seminais surgem no início do desenvolvimento da planta, vindo diretamente da semente e estabelecendo a plântula, tem um papel importante de suprir a mesma de nutrientes e água até atingir a fase de afilamento. Posteriormente, com a formação da coroa surgem as raízes adventícias permanentes que levam um desenvolvimento lento durante o estágio de alongamento e se apresenta completamente formadas já na fase de espigamento. Por fim, as raízes adventícias aéreas podem surgir no primeiro e segundo nó da planta acima do solo. (Scheeren; Castro; Caierão, 2015).

As folhas desenvolvem um papel importante na produtividade de grãos e na caracterização dos cultivares de trigo. A quantidade de folhas (de 3 a 8), a forma, o tamanho, dentre outras características podem variar de planta para planta e ser fundamental no desenvolvimento e rendimento. (Scheeren; Castro; Caierão, 2015).

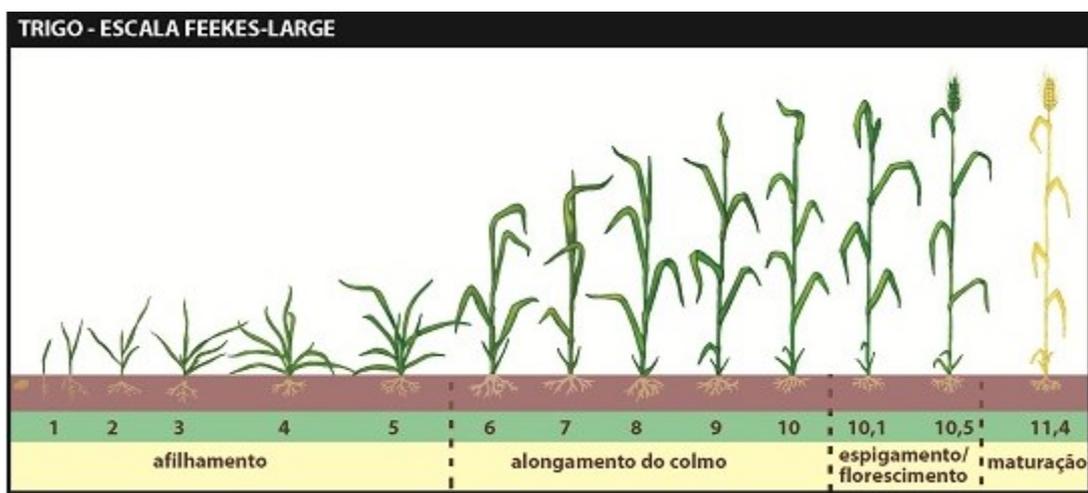
Por sua vez, o colmo é uma estrutura cilíndrica oca capaz de armazenar nutrientes que servirão para preencher os grãos, além disso é o colmo que vai ditar a altura da planta que pode ser variável entre os genótipos e até mesmo para genótipos iguais cultivados em situações ambientais diferentes. O colmo recebe um alto índice de desenvolvimento após a fase de afilamento (Scheeren; Castro; Caierão, 2015).

Por fim, a inflorescência está relacionada com a fase de espigamento e reprodutiva da cultura, onde aparece uma espiga que dá origem aos grãos do trigo. Essa espiga é composta por

diversas estruturas que são capazes de garantir a fecundação e a formação dos grãos, além disso tais estruturas podem distinguir os cultivares e determinar o formato básico da espiga. (Scheeren; Castro; Caierão, 2015).

Para compreender melhor o desenvolvimento da planta e para monitorar estudos e ações realizadas durante o cultivo, são utilizadas as escalas fenológicas que ditam o momento de desenvolvimento em que a planta se encontra, a fim de determinar os tratos culturais, como adubação em cobertura, controle de pragas e doenças. Na cultura do trigo são conhecidas duas escalas fenológicas: a de Feekes (1940) e a de Zadoks, Chang, e Konzac de 1974. A primeira é mais resumida e basicamente detalha o desenvolvimento da planta em 4 estádios: afilhamento, alongamento, espigamento e maturação. Enquanto isso, a escala de Zadoks *et al.* traz um maior detalhamento do desenvolvimento contando com 10 estádios fenológicos (em ordem cronológica): germinação, crescimento da plântula, afilhamento, alongamento, emborrachamento, emergência da inflorescência, antese, desenvolvimento do grão leitoso, desenvolvimento do grão em massa e maturação. Vale ressaltar que a escala de Feekes teve sua proposta modificada por Large em 1954 e é também bastante conhecida e utilizada para estudo e compreensão da evolução do trigo (Scheeren; Castro; Caierão, 2015). Na figura 1 é possível observar os estádios fenológicos descritos por Feekes adaptado por Large.

Figura 1 - Estádios fenológicos Feekes - Large (1954)



Fonte: Santos, Pires e Fontaneli (2014).

3.2 Sensoriamento remoto no trigo

Nos últimos anos tem surgido bastante pesquisa em busca da tentativa de se relacionar os índices de vegetação (IV) com o desenvolvimento da planta, bem como a análise de

produtividade e o uso das informações geoespaciais para identificar estratégias de manejo. No Quadro 1 é apresentada uma síntese de trabalhos que evidenciam aplicações do sensoriamento remoto na cultura do trigo.

Quadro 1 - Estudos relacionando o trigo com o uso de IV

Autores	Título	Ano
Pires <i>et al.</i>	Uso de sensor óptico ativo para caracterização do perfil de NDVI em dosséis de trigo submetidos a diferentes estratégias de manejo.	2014
Magney <i>et al.</i>	Proximal NDVI derived phenology improves in-season predictions of wheat quantity and quality.	2016
Souza	Estimativa de biomassa de trigo usando imagens de alta resolução espacial a partir de veículo aéreo não tripulado.	2018
Arangure, Catéllon, Aizpururua	Wheat yield estimation with NDVI values using a proximal sensing tool.	2020
Carvalho	Índices de vegetação e produtividade de cultivares de trigo submetidas à diferentes épocas de semeadura.	2021

Fonte: O autor.

Pires *et al.* (2014), apresentaram em sua pesquisa que a utilização de um sensor ativo com o uso do índice Normalized Difference Vegetation Index (NDVI) pode auxiliar nas práticas de manejo e na comparação de genótipos da cultura do trigo, além de proporcionar uma avaliação não destrutiva da planta. Já Magney (2016) utilizou o índice para avaliar a fenologia do trigo e associar os dados à produtividade da lavoura, mostrando ser um preditor ruim para as fases iniciais de desenvolvimento da planta.

Souza (2018) evidencia a aplicação de IVs para avaliar a biomassa na cultura do trigo em dois estádios de desenvolvimento da planta. A autora utilizou 10 IVs (dentre eles o NDVI, o *Soil Adjusted Vegetation Index* – SAVI e o *Optimized Soil-Adjusted Vegetation Index* – OSAVI) e concluiu que os índices geraram um modelo com alto valor de correlação com a biomassa, sendo possível a sua utilização para a predição da variabilidade espacial da biomassa seca dos cultivares testados.

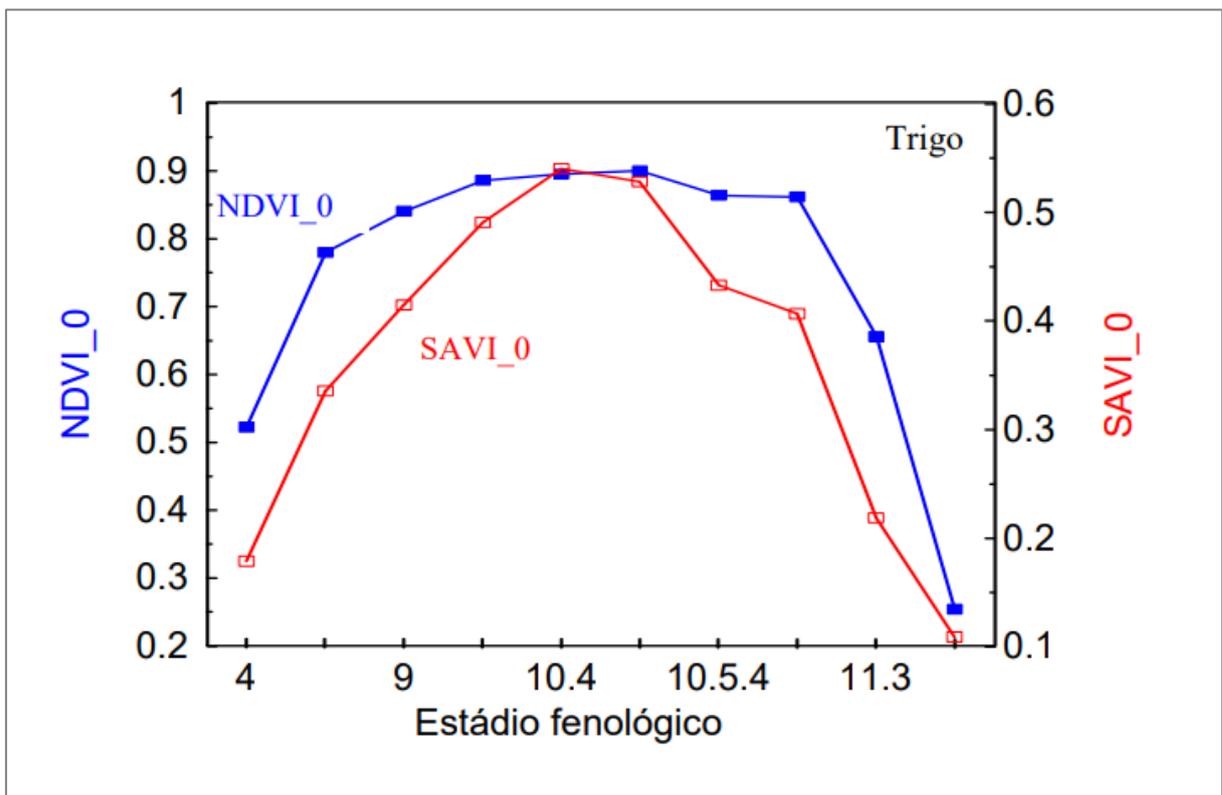
Em 2020, Arangure, Catéllon e Aizpururua, detectaram em seu estudo que o uso do NDVI pode, além de ser um indicador da biomassa da plantação, ser uma ferramenta que auxilia

na identificação do conteúdo de nitrogênio da planta, facilitando o ajuste da adubação do trigo com o uso das faixas de NDVI calculadas. Um ano depois, Carvalho expôs em seu trabalho que a época de semeadura pode ser determinante na produção de matéria seca e de grãos, sendo avaliada com o uso de sensores multiespectrais e com o auxílio de alguns IVs.

Logo, pode-se notar o quanto o índice NDVI é uma ferramenta cada vez mais explorada, que diz muito sobre a saúde da planta e do rendimento dela. Com isso, atrelar o estágio vegetativo com os IVs pode auxiliar na determinação do estágio fenológico que melhor estima a produtividade.

Na Figura 2 é apresentado um gráfico que descreve o comportamento dos índices NDVI e SAVI ao longo dos estádios especificados por Large. É possível notar que entre os estádios 10.4 e 10.5.4 são apresentados os maiores valores para os índices, mostrando que são os estádios em que a planta está em seu alto vigor vegetativo e que possivelmente é o ponto em que há mais biomassa verde ativa.

Figura 2 - Relação IV e estágio fenológico



Fonte: Epiphaneio; Almeida Júnior; Formaggio (1996).

É notório que os parâmetros agrônômicos do trigo podem ser evidenciados por meio do processamento de imagens multiespectrais, fazendo o uso das técnicas de sensoriamento remoto

(SR) aplicados à agricultura de precisão. Assim, a correlação entre o SR e esses parâmetros pode ocorrer através do cálculo e aplicações dos IVs (Epiphaneio; Almeida Júnior; Formaggio, 1996; Shiratsuchi *et al.*, 2019).

3.3 IVs relacionados com a estimativa de produtividade

O NDVI, um dos IVs mais usado e conhecido, possui uma certa deficiência quando se trata da relação com as características biofísicas, sendo afetado com a influência do solo, da atmosfera, com a saturação, calibração e até mesmo pelas condições de visada e iluminação (Epiphaneio; Almeida Júnior; Formaggio, 1996). Este índice demonstra o nível de absorção de energia e clorofila da planta, já que trabalha as bandas do vermelho e do infravermelho próximo (Myneni; Ramakrishna; Running, 1997), evidenciando a saúde e vigor da planta.

Através do quociente da diferença entre as bandas do infravermelho próximo (NIR – do inglês Near InfraRed) e do vermelho (Red) pela soma da reflectância das mesmas bandas, a equação 1 demonstra o cálculo do NDVI.

$$NDVI = (NIR - RED)/(NIR + RED) \quad (1)$$

Com isso, novos índices têm surgido com o intuito de corrigir estes problemas ou melhorar os IVs, além de ressaltar características diversas da planta como a clorofila, estresse hídrico, biomassa, senescência, dentre outras. Como exemplo podemos citar os índices SAVI e o OSAVI, que otimizam os resultados do NDVI.

O SAVI é baseado no NDVI e busca ajustar o índice sob o efeito de reflectância do solo, assim é acrescido à fórmula do NDVI uma constante L que representa o nível de cobertura do solo, variando de 0,25 a 1 (Gameiro *et al.*, 2017). A equação 2 define o cálculo do SAVI, onde a constante L é adicionada ao denominador da fórmula do NDVI e ao numerador é acrescentada a função (1+L).

$$SAVI = (1 + L) (NIR - RED)/(NIR + RED + L) \quad (2)$$

Enquanto isso, o índice OSAVI definido pela equação 3, possui uma constante “a” adicionada ao denominador do NDVI, que recebe o valor de 0,16. Esse valor consegue diminuir o ruído causado pelo solo de forma satisfatória tanto para as regiões que possuem uma alta taxa

de cobertura vegetal, quanto para as áreas com baixa cobertura vegetal (Rondeaux, Steven, Baret, 1996).

$$OSAVI = (NIR - RED)/(NIR + RED + a) \quad (3)$$

Analisando as equações apresentadas, pode-se perceber que os índices SAVI e OSAVI modificam de forma simples a fórmula do NDVI, acrescentando uma constante. Logo, é possível imaginar as inúmeras modificações que podem ser feitas em torno de um único índice e, além disso, é possível melhorar um índice já modificado, que seria o caso do *Modified Soil Adjusted Vegetation Index* (MSAVI) que modifica o índice SAVI. Para Qi *et al* (1994) a constante L presente no SAVI, deve ser substituída por uma função variável que se dá por meio da indução ou pelo produto entre o NDVI e outro índice de diferença ponderada, resultando na formulação do MSAVI. O autor apresenta em seu trabalho os resultados testados sobre o dossel de uma plantação de algodão.

O Index Data Base lista mais de 500 índices existentes e apresenta um mecanismo de busca que permite ao usuário separar os IVs pela aplicação ou por meio do sensor selecionado, além de disponibilizar uma lista de literatura onde o usuário consegue encontrar mais detalhes sobre os IVs (IDB, 2023). Com tantos IVs disponíveis, é preciso filtrar e selecionar alguns para se trabalhar. Os IVs citados anteriormente (SAVI, OSAVI e MSAVI) por exemplo, são baseados no NDVI e, portanto, ressaltam a mesma propriedade biofísica da planta. Sendo assim, não se faz necessário a utilização de todos os índices para a geração do modelo de estimativa.

3.4 Modelos espectrais para se estimar a produtividade

Os modelos buscam representar parte do mundo real por meio de formulações matemáticas (Caixeta-Filho, 2001). Através dos algoritmos de aprendizado de máquinas, esses modelos podem ser utilizados para estimar e prever a produtividade, e são baseados em métodos paramétricos e não paramétricos. No método paramétrico temos como exemplo os modelos de Regressão Linear, enquanto no segundo método os modelos gerados com o Support Vector Machine (SVM), Random Forest (RF), Neural Net (NN) ou por meio da Random Tree (RT) são alguns exemplos de algoritmos disponíveis (Richards; Jia, 2006).

O RT e RF possuem as mesmas características funcionais, porém, enquanto o RT considera uma única árvore de decisão o algoritmo do RF baseia-se nas múltiplas árvores de decisão. As árvores são geradas durante o treinamento dividindo os dados em subconjuntos a

serem avaliados em cada árvores e por fim, o resultado se dá por votação majoritária ou por média para regressão (Budu, 2023; Breiman, 2001). Além disso, não é exigido que os dados admitam distribuição normal. Por este motivo, no RF o processamento de dados advindos de fontes diferentes é facilitado, como é o caso da pesquisa onde temos dados mensurados em campos e outros derivados de imagens de satélites (Budu, 2023). Com isso, o RF tem se mostrado uma grande ferramenta para a estimativa de produtividade, além de seu uso para estimar outras propriedades do solo e da produção, a fim de auxiliar na solução de problemas na área agrícola. (Khanal *et al.*, 2018).

A NN é considerada uma máquina ou um processador modelado para trabalhar como o cérebro humano, o qual têm a capacidade de armazenar conhecimento, processar as informações e retornar um resultado baseados nas experiências passadas. Para isso, a NN passa por um processo repetitivo de interações durante o treinamento do algoritmo com os dados na camada dos neurônios de entrada, adiante eles são processados pelas camadas ocultas e geram os neurônios de saída, que são os resultados. A NN traz como benefício a não linearidade, adaptabilidade, o mapeamento de entrada-saída, informação contextual, resposta a evidências etc. (Haykin, 2001). Argenta *et. al* (2020) demonstram em seu trabalho a capacidade da NN em simular a produtividade do trigo envolvendo informações ambientais, nutricionais e biológicas.

Por fim temos o SVM que faz o uso de métodos de aprendizado de máquina para classificação, detecção de outliers e regressão. Para isso, o algoritmo utiliza-se de vetores suporte (ou registros essenciais) do conjunto de treino para criar hiperplanos que separam os dados em duas classes, além de calcular as margens dos hiperplanos para identificar o melhor dentre todos os criados (Silva, Zullo Júnior, Romani, 2017). Quando se trata do desempenho do algoritmo na predição de produtividade agrícola, Michelon *et al.* (2017) evidenciaram em sua pesquisa que o uso do SVM foi capaz de estimar a produtividade da soja utilizando apenas 3 variáveis.

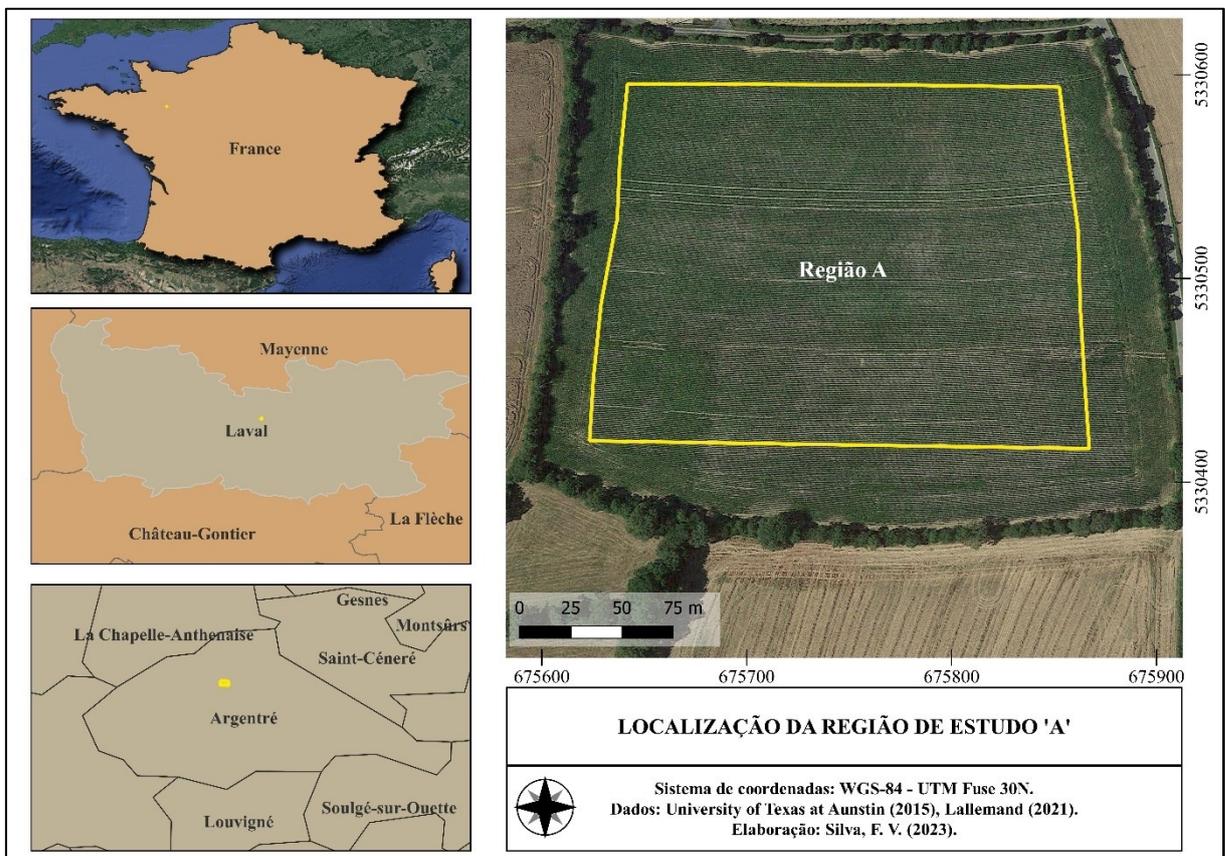
4 MATERIAL E MÉTODOS

4.1 Área de estudo

Os dados de produtividade geolocalizados, bem como as regiões de plantio do trigo utilizadas para este estudo estão localizadas na França. Foram analisadas três áreas cultivadas: a primeira presente na comuna de Argentré do distrito de Laval e as outras duas regiões pertencentes à comuna de Neulliac do distrito de Pontivy.

A região A (Figura 3) foi tomada como a principal área de estudo por contar com mais imagens disponíveis, ao menos uma para cada mês de desenvolvimento da planta, com um montante de 13 pacotes de imagens. A lavoura colhida no dia 23 de julho de 2021, possuía uma área plantada de 4,05 hectares, um perímetro de 809,85 metros e cerca de 3500 pontos amostrados com dados de produtividade. Além disso, atingiu média de produtividade de 8,13 t ha⁻¹, tendo o valor mínimo de 5,91 t ha⁻¹ e a produtividade máxima de 10,42 t ha⁻¹. Enquanto isso, 329 pixels por imagem foram selecionados para representar o dossel da lavoura.

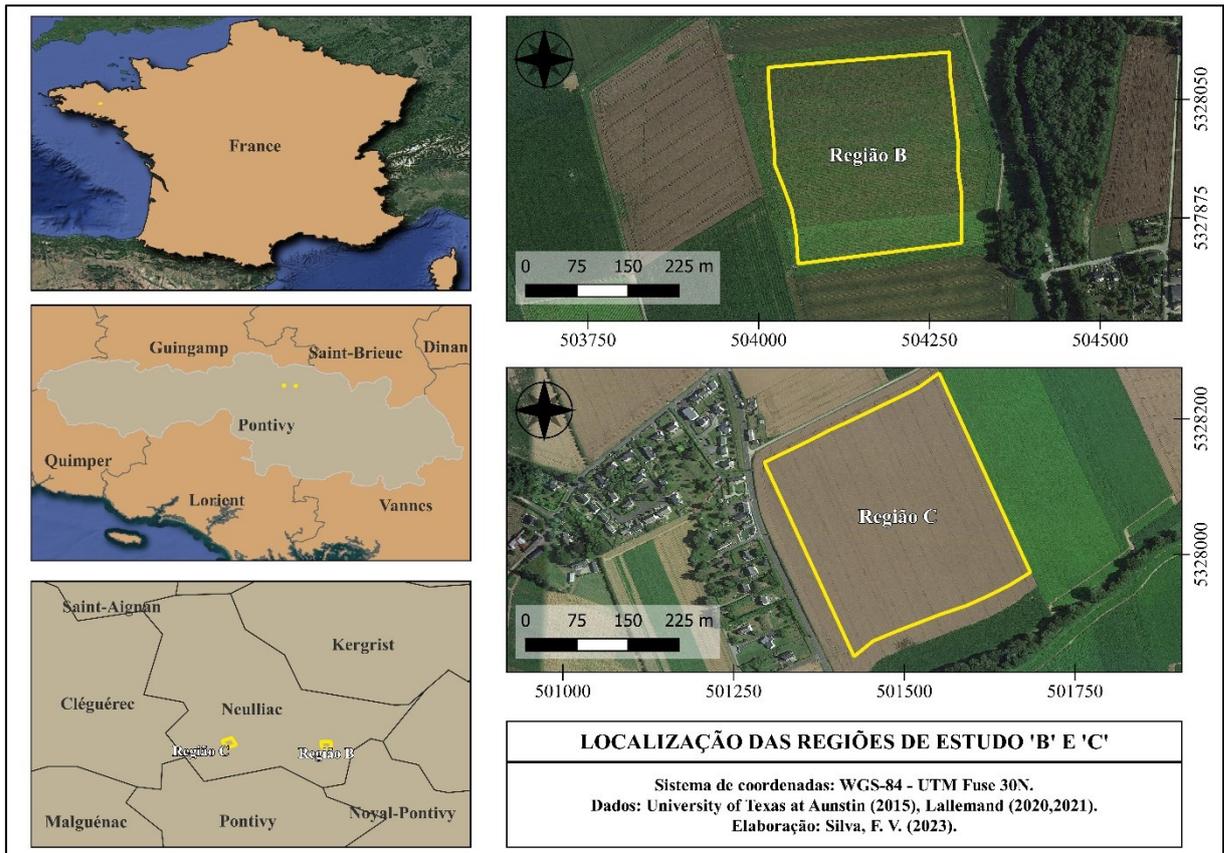
Figura 3 - Mapa de localização da região de estudo A



Fonte: o autor.

A região B (Figura 4) foi colhida no dia 22 de julho de 2021, obteve produtividade média de $7,98 \text{ t ha}^{-1}$, 4100 pontos amostrados, área de 7,46 hectares e 750 pixels foram selecionados para gerar os modelos multiespectrais. Por fim, a região C (Figura 4) que foi colhida no dia 20 de julho de 2020 contou com 8241 pontos amostrais de produtividade, 927 pixels selecionados e produtividade média de $8,05 \text{ t ha}^{-1}$, com área cultivada de 9,06 hectares.

Figura 4 - Mapa de localização das regiões de estudo B e C



Fonte: o autor.

4.2 Material

Para a execução do trabalho, foram utilizadas imagens orbitais do satélite Sentinel 2 em diferentes épocas da evolução das lavouras, desde o plantio à colheita. O intuito foi estimar a produtividade a cada fase fenológica e encontrar aquela que mais se aproxima do valor real absoluto, por isso a aquisição dos dados (imagens) mostrou-se tão importante, uma vez que quanto maior o número de imagens, maior é a resolução temporal e a quantidade de informações. Dessa forma, tornou-se possível a estimativa de produtividade a cada estágio fenológico e encontrou-se o estágio ideal para aquisição das imagens multiespectrais.

Cada imagem multiespectral baixada contém 13 bandas espectrais com resolução espacial de 10, 20 e 60 m, e com a resolução espectral variando de 443 a 2190 nanômetros (Embrapa, 2018). Por isso as imagens Sentinel se mostraram válidas para o uso desde trabalho. A quantidade de banda nos permite explorar o cálculo de diversos índices correlacionados com a produtividade e saúde da planta, agregando mais informações aos modelos, melhorando a estimativa. Além disso, o satélite conta com uma alta resolução espacial e um curto tempo de revisita de 5 dias (Embrapa, 2018), o que permitiu a avaliação a cada estágio fenológico do trigo. Vale acrescentar que todas as imagens são fornecidas gratuitamente no site da Copernicus. (<https://dataspace.copernicus.eu/browser>)

Além das imagens, foram necessários os dados de produtividade georreferenciados, bem como o limite da área de plantio. Os pontos de produtividade foram obtidos por meio de uma colhedora com um sistema georreferenciador acoplado, disponibilizados pela empresa Lallemand Plant Care. Já para o processamento e espacialização dos dados vetoriais foi utilizado um software SIG. A distribuição espacial dos pontos de produtividade e a representação gráfica das áreas de estudo são apresentados na Figura 5.

Priorizando o uso de softwares livres, para o desenvolvimento do trabalho foi utilizado o QGIS na versão Desktop 3.22.8 e o WEKA em sua versão 3.9.6. O QGIS é um aplicativo de sistemas de informação geográfica de código aberto da Open Source Geospatial Foundation (OSGeo). Este software, trata-se de uma ferramenta que possibilita a criação e a manipulação de dados vetoriais (ex.: pontos, linhas, polígonos) e matriciais (ex.: raster) (QGIS, 2023). Por sua vez, o software WEKA (Waikato Environment for Knowledge Analysis) trata-se de uma plataforma desenvolvida para a mineração de dados. Ele contém um conjunto de algoritmos e ferramentas que permitem a preparação de dados, agrupamento, classificação, regressão, entre outras funções. O WEKA também detém em suas características o código aberto e foi emitido pela GNU – General Public License. (Waikato, 201?). Ademais, foram utilizadas planilhas eletrônicas para cálculo dos IVs e tabulação dos resultados encontrados.

Figura 5 - Distribuição espacial de pontos de produtividade

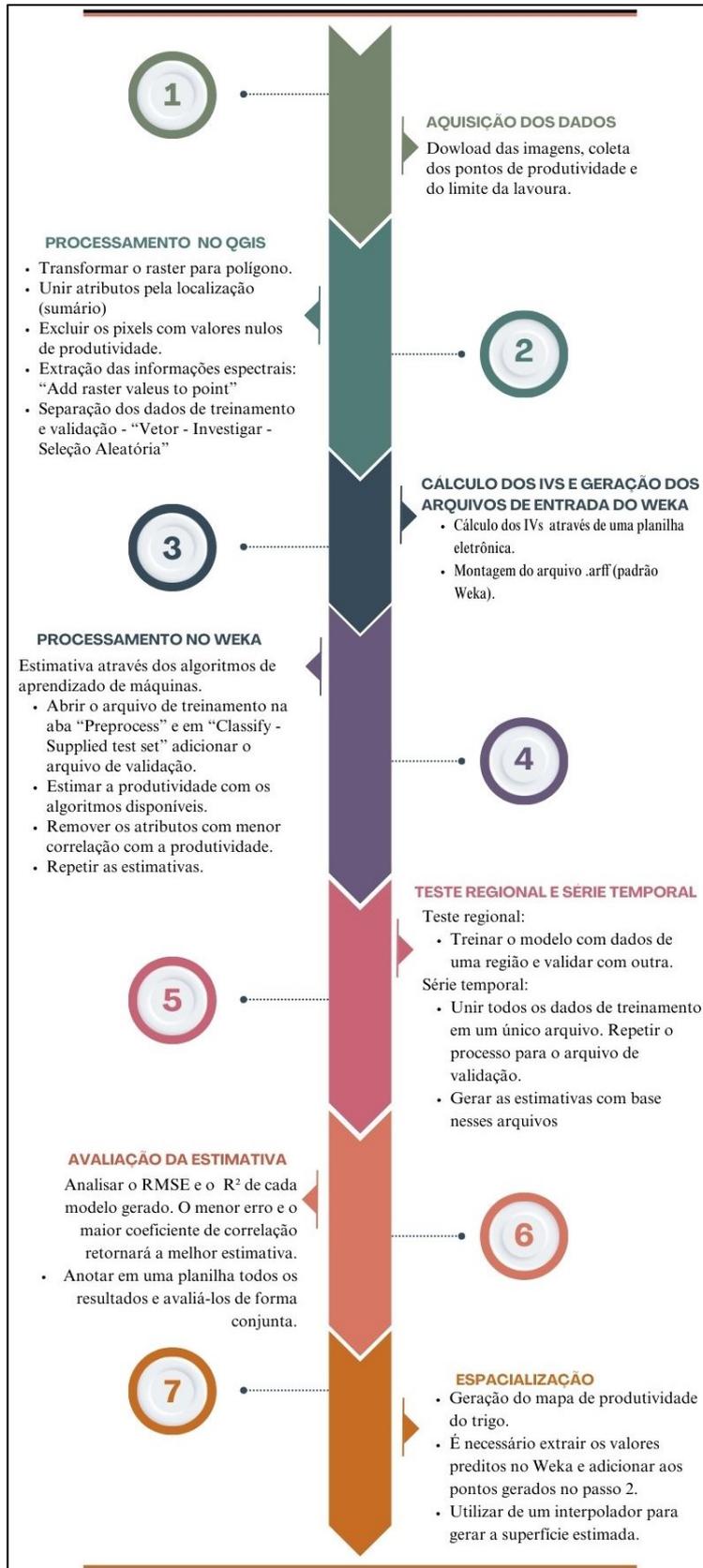


Fonte: o autor.

4.3 Métodos

A metodologia empregada para o desenvolvimento pleno deste trabalho pode ser dividida em 4 partes: aquisição dos dados, pré-processamento dos dados, processamento dos dados e a representação gráfica. Na Figura 6 é apresentado um fluxograma que explicita o desenvolvimento da metodologia aplicada.

Figura 6 - Fluxograma



Fonte: o autor.

4.3.1 Aquisição dos dados.

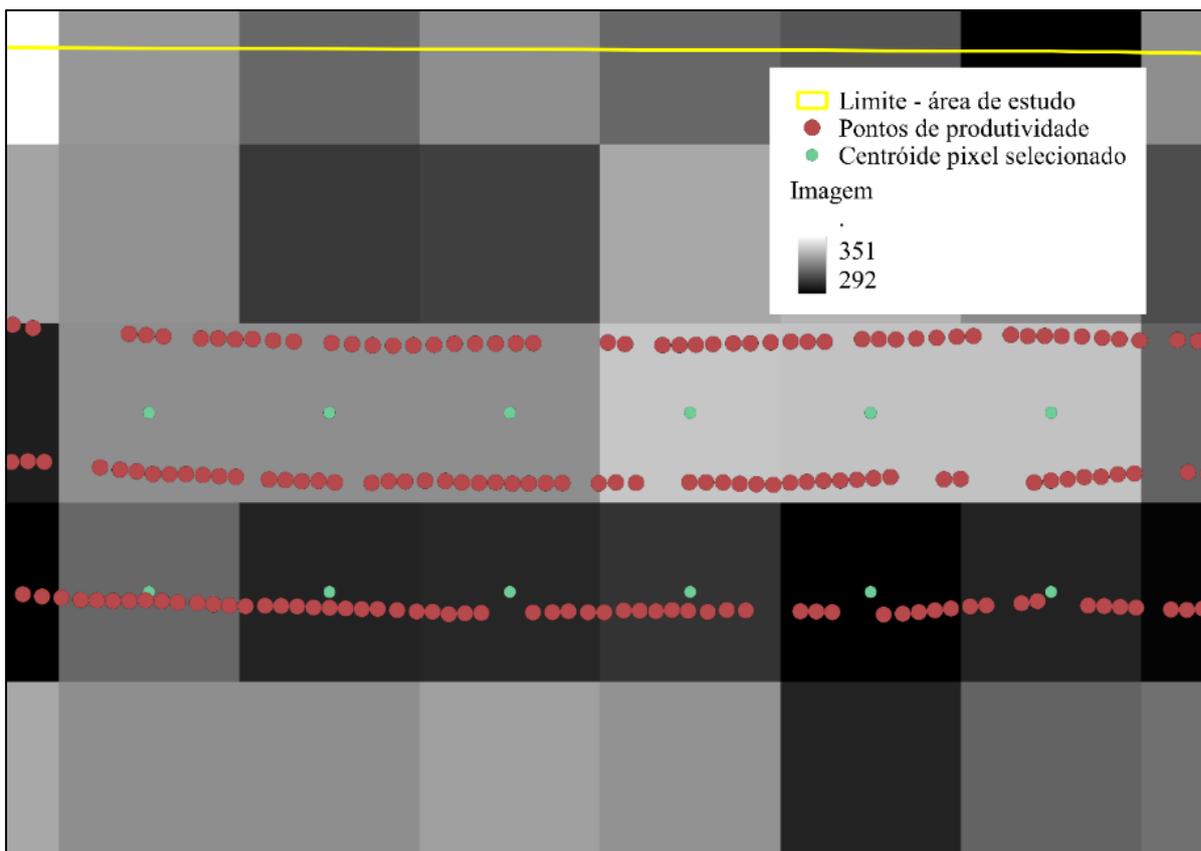
Para iniciar o trabalho foram baixadas imagens Sentinel 2 do site da Copernicus, criando um polígono envolvendo cada área de estudo, para identificar as imagens que cobrem a região, e aplicando filtros de seleção, como a data, o satélite, o nível de processamento e de cobertura de nuvens. Foi baixada ao menos uma imagem para cada estágio fenológico e imagens com nível de processamento 2A, que possui correção de reflectância na base da atmosfera (ESA, 2017).

4.3.2 Pré-processamento dos dados

O pré-processamento consiste na preparação do arquivo de entrada do WEKA que possibilita dar início à terceira fase. Este arquivo deve conter as medidas de produtividade e os valores espectrais de cada banda relacionado aos pontos de produtividade.

Visto que cada pixel cobre uma área de 10 m x 10 m, tiveram áreas com vários pontos amostrais dentro de um mesmo pixel e pixels que não tinham nenhum ponto de produtividade em seu interior (Figura 7). Em média 867, 550 e 910 pontos por hectare foram amostrados para as regiões A, B e C respectivamente. Com isso, foram selecionados somente os pixels que intersectam os pontos de produtividade para garantir que a resposta espectral seja referente ao dossel da lavoura e esteja vinculada com os valores de produtividade coletados.

Figura 7 - Relação pixel e ponto de produtividade



Elaboração: o autor.

Para realizar a seleção dos pixels, primeiro transformou-se o raster para polígono e em seguida, por meio da ferramenta “unir atributos pela localização (sumário)”, adicionou-se a média da produtividade como um atributo à tabela da camada vetorial gerada. Em seguida, foram selecionados e excluídos os pixels com a produtividade nula, ou seja, aqueles que não tinham vínculo com os pontos amostrais de produtividade foram descartados. Feito isto com uma das bandas espectrais da imagem, extraiu-se os valores das demais bandas através da ferramenta “Add raster values to points” e transformou-se o polígono para ponto.

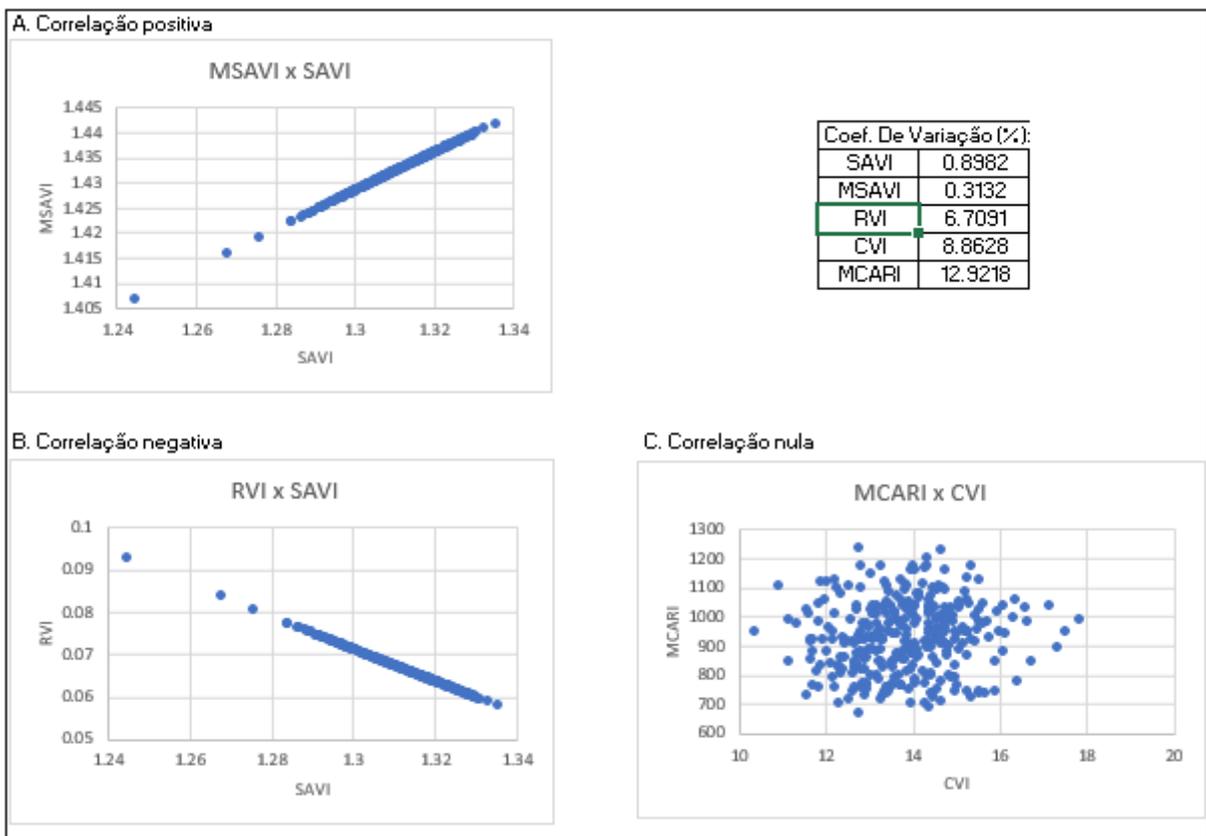
Os pontos foram divididos em dois arquivos, um para o treinamento e outro para validação do modelo, por meio de uma ferramenta de seleção aleatória dos pontos disponível no QGIS. Para a execução deste trabalho foi utilizado o método Holdout, com os dados sendo separados em um conjunto contendo 80% dos pontos para a geração do modelo e outro acervo com os 20% restante dos pontos para a validação do modelo gerado, sendo esta relação 80/20 muito comum entre os usuários do método (Allibhai, 2018). Por se tratar de amostras com mais de 3500 pontos avaliados, temos que um subconjunto com 20% dos pontos pode ser considerado representativo para toda a lavoura, além de ser um método com baixo esforço computacional, tornando a utilização do método uma escolha viável.

4.3.3 Processamento dos dados

Dando início à fase três, os IVs foram calculados em uma planilha eletrônica, além disso foi computado o coeficiente de variação e o gráfico de dispersão para os IVs que evidenciam a mesma característica da planta (Figura 8). Aqueles que apresentaram graficamente uma correlação positiva, demonstraram grande correlação entre os dados e por tanto optou-se por utilizar um IV ou o outro, a fim de evitar a redundância, visto que ambos estão agregando a mesma informação ao modelo. Por fim, foi escolhido o IV que apresentou o maior valor de coeficiente de variação, para garantir a variabilidade dos dados.

Enquanto isso, para os IVs que não demonstraram uma correlação positiva, foi admitido o uso de ambos os IVs. Uma vez que não existe correlação entre os IVs ou há uma correlação negativa entre os dados, isso indica que mesmo ressaltando a mesma característica biofísica da planta eles estão agregando informações diferentes ao modelo de predição.

Figura 8 - Relação entre os IVs



Fonte: o autor.

O arquivo de entrada do tipo .arff (Attribute – Relation File Format), padrão do WEKA, foi gerado com os dados já separados e os IVs calculados, dando sequência à fase de processamento. O arquivo é composto por um cabeçalho, listando os tipos de atributos e seus valores, e uma seção com os dados separados por vírgula.

O software WEKA permite a utilização de algoritmos de aprendizado de máquina para estimar uma variável de interesse, nesse caso a produtividade. A regressão linear, NN, RF e o SVM são alguns algoritmos implementados, os quais foram utilizados no formato padrão disponibilizado no software para a geração do modelo multiespectral. Além da estimativa, o software retorna ao usuário o erro, o coeficiente de correlação (R), o Root mean square error (RMSE), dentre outros parâmetros para a avaliação do modelo treinado. Além disso, através da ferramenta “Select attributes” é possível identificar os atributos que possuem uma baixa correlação com a produtividade, o que pode confundir o algoritmo e gerar uma estimativa ruim.

Com isso, na página inicial do WEKA foi adicionado o arquivo de treinamento e na aba “Classify” através da ferramenta “Supplied test set” foi selecionado o arquivo de validação. Em seguida, foram testados os algoritmos de aprendizado de máquina e anotado em uma planilha os valores de RMSE e R de cada estimativa computada. Os parâmetros de avaliação R e RMSE são calculados segundo as equações 4 e 5 respectivamente.

$$R = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2 \sum_i^n (y_i - \bar{y})^2}} \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_i^n (y_i - \bar{y})^2} \quad (5)$$

Onde: x_i e y_i são os valores medidos; \bar{x} e \bar{y} são os valores estimados.

As variáveis foram removidas uma a uma, seguindo a classificação de R dos atributos com a produtividade, e gerou-se novas estimativas com todos os algoritmos testados anteriormente. A remoção dos atributos foi interrompida quando sobraram apenas cinco variáveis para gerar os modelos preditivos. Dessa forma, foi possível identificar as bandas e IVs que devem compor o modelo para uma estimativa mais acurada.

4.3.4 Validação e melhoramento do modelo

A validação dos modelos ocorreu por parte da análise do RMSE, onde o modelo com menor erro foi determinado como aquele que obteve maior desempenho para os dados de treinamento. Ao mesmo tempo, foi avaliado o R dos modelos, tomando o maior valor como referência para a escolha do melhor modelo. Assim, o estágio fenológico que melhor estima a produtividade do trigo foi determinado de acordo com a fase do modelo escolhido, contando com o menor RMSE e o maior R.

Uma vez que foi determinado o estágio fenológico, iniciou-se o processo de tentativa de melhoramento dos modelos gerados, o qual ocorreu com o teste de uma série histórica e com a avaliação de modelos regionais, buscando atingir os objetivos específicos.

Logo, para a série temporal foi avaliado o comportamento da resposta espectral e dos índices ao longo da época de cultivo gerando um modelo utilizando todos estes dados e não somente para uma certa data do desenvolvimento como foi feito anteriormente. Já o modelo regional foi avaliado de forma simples, treinando o modelo sobre uma das regiões de estudo e validando a estimativa da produtividade com dados de outra região.

4.3.5 Representação Gráfica

Com as análises realizadas, foram adicionados os valores estimados a uma camada vetorial que continham todos os pontos (pontos de treinamento, validação e aqueles com a produtividade nula que foram retirados para geração dos modelos) para gerar a espacialização da produtividade de toda a lavoura. Estes pontos são correspondentes aos centroides dos pixels que cobriram toda a lavoura, gerando assim uma grade regular 10 x 10 m. Os pontos foram interpolados com o uso da krigagem e gerou-se um layout de impressão para gerar o mapa de produtividade do trigo.

5 RESULTADOS E DISCUSSÕES

5.1 IVs selecionados para a modelagem

Diante de todos os procedimentos realizados, pode-se considerar a escolha dos IVs como primeiro resultado encontrado no trabalho. Dessa forma, foram calculados 23 IVs e sendo estabelecido segundo a metodologia proposta o uso de 10 IVs mais o NDVI devido a sua importância e presença em diversas pesquisas. Tais IVs destacam a biomassa, a clorofila, área foliar e, a qualidade e quantidade da vegetação, atributos que influenciam diretamente na produtividade da cultura. Os IVs utilizados na modelagem da estimativa são apresentados no Quadro 2.

Quadro 2 - Índices selecionados para a geração dos modelos preditivos

Índice	Equação	Contribuição	Referência
RVI – Ratio Vegetation Index.	$\frac{R}{Nir}$	Correlação com a biomassa e para monitorar e estimar a cobertura vegetal.	Fitz (2020).
SAVI- Soil Adjusted Vegetation Index	$\frac{(1 + L) * (Nir - R)}{(Nir + R + L)}$	Indicador da quantidade e qualidade da vegetação, ajustado ao solo.	Gameiro <i>et al.</i> (2017).
EVI – Enhanced Vegetation Index	$2,5 * \frac{G * (Nir - R)}{(Nir + C1 * R * C2 * B + L)}$	Reduz a influência atmosférica e do solo, vigor da vegetação e alta resposta a variação fenológica.	Justice <i>et al.</i> (1997).

TVI – Transformed Vegetation Index	$\sqrt{NDVI + 0,5}$	Medição da biomassa e monitorar a progressão da vegetação.	Rouse <i>et al.</i> (1974)
MCARI – Modified Chlorophyll Absorption Ratio Index	$[(Re1 - R) - 0,2 * (Re1 - G)] * \frac{Re1}{R}$	Indicador de clorofila da planta e preditor do índice de área foliar.	Daughtry <i>et al.</i> (2000)
MTVI2 – Modified Triangular Vegetation Index 2	$\frac{1,5[2,5(Nir - B) - 2,5(R - G)]}{\sqrt{(2Nir + 1)^2 - 6Nir - 5\sqrt{(Nir - 0,5)}}$	Preditor de índice de área foliar.	Haboudane <i>et al.</i> (2003)
TGI – Triangular Greenness Index	$-0,5 * ((665 - 490) * (R - G) - (665 - 560) * (R + B))$	Sensibilidade à clorofila.	Hunt Junior <i>et al.</i> (2004)
CVI – Chlorophyll Vegetation Index	$Nir * \frac{R}{G^2}$	Estimador de clorofila foliar.	Vincini; Frazzi; D'Alessio (2008)
GLI – Green Leaf Index	$\frac{2 * G - R - B}{2 * G + R + B}$	Cobertura vegetal da cultura.	Louhaichi; Borman; Johnson (2001)
MTCI – MERIS Terrestrial Chlorophyll Index	$\frac{Nir - Re2}{Re2 - R}$	Sensibilidade à clorofila.	Dash; Curran (2007)

Fonte: o autor.

Onde:

Quadro 3 - Bandas Sentinel 2 utilizada nos cálculos dos IVs

Siglas – Significado	Banda Sentinel 2	Comprimento de onda Central (μm)
B – Banda do Azul	Banda 2	0,490
G – Banda do Verde	Banda 3	0,560
R – Banda do Vermelho	Banda 4	0,665
Re1 – Banda Red Edge 1	Banda 5	0,705
Re2 – Banda Red Edge 2	Banda 6	0,740
Nir – Banda do Infravermelho Próximo	Banda 8	0,842

Fonte: o autor.

5.2 Algoritmo de aprendizado de máquina com melhor desempenho

A Figura 9 a seguir evidencia os resultados obtidos para um dos períodos analisados, o qual conta com imagens que foram mensuradas no dia 28 de maio, cerca de 2 meses antes da colheita. Nela é possível observar a forma que se sucedeu a remoção das variáveis, bem como os algoritmos que trouxeram o melhor resultado e em destaque o modelo que gerou a melhor estimativa para a data analisada.

Figura 9 - Resultados processamentos

Região A - Período estimado: 28/05/2021									
Testes	Correlation	RMSE	Modelo	Variáveis	Testes	Correlation	RMSE	Modelo	Variáveis
1	0.536	0.6366	Multilayer Perceptron	Todas	35	0.6704	0.5654	Linear Regression	Sem MTVI2
2	0.7114	0.5423	SMDreg	Todas	36	0.7413	0.5337	Random Forest	Sem MTVI2
3	0.6733	0.5645	Linear Regression	Todas	37	0.5999	0.5967	Multilayer Perceptron	Sem SAVI
4	0.7207	0.5518	Random Forest	Todas	38	0.7082	0.5508	SMDreg	Sem SAVI
5	0.4544	0.6874	Multilayer Perceptron	Sem B	39	0.6844	0.5567	Linear Regression	Sem SAVI
6	0.7118	0.5426	SMDreg	Sem B	40	0.7377	0.5386	Random Forest	Sem SAVI
7	0.6313	0.5666	Linear Regression	Sem B	41	0.6168	0.5937	Multilayer Perceptron	Sem RVI
8	0.7207	0.5467	Random Forest	Sem B	42	0.7088	0.5512	SMDreg	Sem RVI
9	0.6616	0.5668	Multilayer Perceptron	Sem EVI	43	0.6844	0.5567	Linear Regression	Sem RVI
10	0.7091	0.5435	SMDreg	Sem EVI	44	0.7317	0.5353	Random Forest	Sem RVI
11	0.6874	0.5555	Linear Regression	Sem EVI	45	0.6249	0.5878	Multilayer Perceptron	Sem SWIR2
12	0.7252	0.5467	Random Forest	Sem EVI	46	0.6642	0.5656	SMDreg	Sem SWIR2
13	0.6067	0.5996	Multilayer Perceptron	Sem TVI	47	0.6376	0.583	Linear Regression	Sem SWIR2
14	0.7097	0.5434	SMDreg	Sem TVI	48	0.7538	0.5274	Random Forest	Sem SWIR2
15	0.6916	0.5532	Linear Regression	Sem TVI	49	0.4845	0.6475	Multilayer Perceptron	Sem G
16	0.7052	0.5602	Random Forest	Sem TVI	50	0.5877	0.6054	SMDreg	Sem G
17	0.5372	0.6387	Multilayer Perceptron	Sem NDVI	51	0.5932	0.6103	Linear Regression	Sem G
18	0.7093	0.5441	SMDreg	Sem NDVI	52	0.7031	0.5497	Random Forest	Sem G
19	0.6916	0.5532	Linear Regression	Sem NDVI	53	0.6296	0.6113	Multilayer Perceptron	Sem R
20	0.7241	0.5467	Random Forest	Sem NDVI	54	0.5444	0.6233	SMDreg	Sem R
21	0.6517	0.5869	Multilayer Perceptron	Sem MCARI	55	0.5354	0.6314	Linear Regression	Sem R
22	0.7109	0.5445	SMDreg	Sem MCARI	56	0.6916	0.5505	Random Forest	Sem R
23	0.6879	0.5548	Linear Regression	Sem MCARI	57	0.5798	0.6185	Multilayer Perceptron	Sem RE1
24	0.7392	0.5384	Random Forest	Sem MCARI	58	0.5582	0.6186	SMDreg	Sem RE1
25	0.6155	0.5874	Multilayer Perceptron	Sem TGI	59	0.5794	0.6151	Linear Regression	Sem RE1
26	0.7108	0.5453	SMDreg	Sem TGI	60	0.6508	0.572	Random Forest	Sem RE1
27	0.6844	0.5567	Linear Regression	Sem TGI	61	0.449	0.6655	Multilayer Perceptron	Sem RE2
28	0.7316	0.5456	Random Forest	Sem TGI	62	0.5565	0.6191	SMDreg	Sem RE2
29	0.6497	0.5739	Multilayer Perceptron	Sem CVI	63	0.5794	0.6151	Linear Regression	Sem RE2
30	0.7112	0.5489	SMDreg	Sem CVI	64	0.6398	0.5791	Random Forest	Sem RE2
31	0.6833	0.5583	Linear Regression	Sem CVI	65	0.5921	0.6125	Multilayer Perceptron	Sem RE3
32	0.7336	0.5408	Random Forest	Sem CVI	66	0.5255	0.6356	SMDreg	Sem RE3
33	0.6435	0.5798	Multilayer Perceptron	Sem MTVI2	67	0.5567	0.6282	Linear Regression	Sem RE3
34	0.7091	0.5499	SMDreg	Sem MTVI2	68	0.6091	0.593	Random Forest	Sem RE3

Fonte: o autor.

Em 90,5% dos modelos gerados para a região A, considerando todos as datas analisadas, o algoritmo RF obteve o melhor desempenho. Logo, este pode ser considerado o segundo resultado da pesquisa, a descoberta de que o algoritmo RF fornece a melhor estimativa para a cultura do trigo. Diante disso, as regiões B e C, os testes regionais e a série temporal foram realizadas apenas com o algoritmo RF.

5.3 Estádio fenológico ideal para a estimativa do trigo

Figura 10 – Resultados das regiões A, B e C



Fonte: o autor.

Na Figura 10 são demonstrados os resultados dos melhores modelos de cada região de estudo. É notório que em todas as estimativas da região A o valor do R foi maior que 0,5, chegando a atingir 0,75. Fica evidente que o R se manteve relativamente alto, sendo assim o RMSE foi determinante para a escolha do melhor modelo. Com isso, o período de 31 de maio determina o melhor momento para a tomada das imagens e estimativa de produtividade da região A.

De forma análoga foram analisadas as regiões B e C. O intuito foi verificar se o padrão dos resultados da região A seria replicado nas outras regiões. Foi possível perceber que, apesar de os parâmetros de avaliação piorarem (R chegando a 0,28 e RMSE a 0,76), a data de aquisição da imagem para a melhor estimativa foi correspondente.

Vale ressaltar que o solo exposto tende a saturar alguns índices, como o NDVI, bem como influenciar nas respostas espectrais das imagens (Epiphaneu Almeida Júnior; Formaggio, 1996). Com isso, os valores obtidos tendem a influenciar a estimativa, confundindo o algoritmo e admitindo bons resultados, porém tais resultados não são confiáveis devido a essa influência. Por isso, as estimativas realizadas para o mês de março foram desconsideradas, visto que se trata do estágio de afilhamento da cultura, onde a planta conta com poucas folhas e nós, e com isso tem-se a presença e influência de solo exposto na imagem.

Como foi evidenciado, na região B a melhor estimativa se deu na data do dia 30 de março e, além disso, é possível observar que os resultados são semelhantes para a data de 01 de abril e 31 de maio. Ao descartar a estimativa do mês de março, fica claro que o RMSE será determinante para a escolha da melhor estimativa, visto que os dois modelos gerados possuem um R de 0,48. Dessa forma, mais uma vez o dia 31 de maio se mostrou o ideal para aquisição das imagens multiespectrais.

Com isso, pode-se concluir que o estágio fenológico que melhor estima a produtividade do trigo é o estágio referente à data 31 de maio. Nesta data, nas três regiões de estudo as lavouras se encontravam possivelmente na fase de enchimento dos grãos do estágio maturação.

5.4 Modelo regional e série temporal

Com o estágio determinado, foram realizados os testes regionais e a série temporal na tentativa de melhorar a estimativa. Os resultados deixaram evidente que não é possível gerar um modelo regional com os parâmetros de avaliação melhores que aqueles da estimativa local, ou seja, o modelo regional não foi capaz de melhorar a estimativa de produtividade (vide Quadro 4). Isso deixa claro que para este tipo de trabalho o sensoriamento remoto ainda é uma ferramenta de análise local, com dependência e influência espacial.

Com a série temporal não foi diferente, agravou-se os parâmetros de validação e consequentemente houve uma piora na estimativa de produtividade do trigo. Porém a análise reafirmou os resultados que surgiram na fase de determinação do estágio fenológico, pois quando os dados do dia 31/05 foram removidos da geração do modelo, os resultados pioraram consideravelmente (o RMSE chegou a 0,54 que antes era de 0,45 e o R que antes era de 0,53 chegou a atingir 0,13), evidenciando que este período mantém a estimativa estável e assertiva. Importante ressaltar que a série temporal foi realizada somente para a principal área de estudo, região A.

Quadro 4 - Resultados do teste regional e da série temporal

	Treinamento	Validação	R	RMSE
Teste regional	Região A	Região B	0,29	0,77
	Região B	Região A	0,08	0,54
	Região C	Região B	0,25	0,82
	Região B	Região C	0,23	0,6
Série temporal da região de estudo A			0,5355	0,4555

Fonte: o autor.

5.5 IVs / bandas que compõem a melhor estimativa

Diante de todos os processos e avaliações foi possível notar que alguns atributos se repetiam na composição dos melhores modelos. Com base nisso, foi gerado o Quadro 5 que contém as variáveis que devem compor a estimativa de produtividade do trigo, um resultado importante visto que existem diversos índices, bandas e dados que destacam as características biofísica da planta. Vale destacar que o IV MTCI foi o atributo com maior correlação com a produtividade para as 3 regiões de estudo, contudo não foi possível ranquear as variáveis em uma ordem de influência, já que em cada região as demais variáveis ocuparam posições diferentes quando avaliou-se esta correlação entre a produtividade e os atributos.

Quadro 5 - Variáveis que devem compor a estimativa

Variáveis presentes nos melhores modelos das regiões A, B e C.
MTCI
G
R
GLI
RE1
RE2
RE3
NIR
SWIR2

Fonte: o autor.

Quando se trata da reflectância avaliada no dossel de uma lavoura, a região do comprimento de onda visível (bandas R, G e B) é influenciada pelos pigmentos existentes nas folhas, como clorofila, carotenos e xantofilas. Somente a clorofila absorve a radiação compreendida na região da banda do vermelho (Ponzoni; Shimabukuro; Kuplich, 2012). Com isso, a resposta espectral na banda R é um forte indicador da presença de clorofila da planta, diretamente relacionada com a fotossíntese.

Moreira (2011) descreve em seu trabalho fatores que interferem na reflectância da folha e a interação da radiação solar com o dossel da vegetação. Nele é apresentado o comportamento de uma curva de reflectância para a cultura do trigo avaliada com dois níveis de adubação nitrogenada. O autor destaca que as respostas das bandas do verde (banda G) e infravermelho próximo (correspondente às bandas NIR, RE1, RE2, RE3) são mais acentuadas quando a planta possui maior quantidade de nutrientes e fitomassa. Dessa forma, as informações espectrais das bandas citadas são indicadores da presença de nutrientes e fitomassa da vegetação, evidenciando o vigor vegetativo.

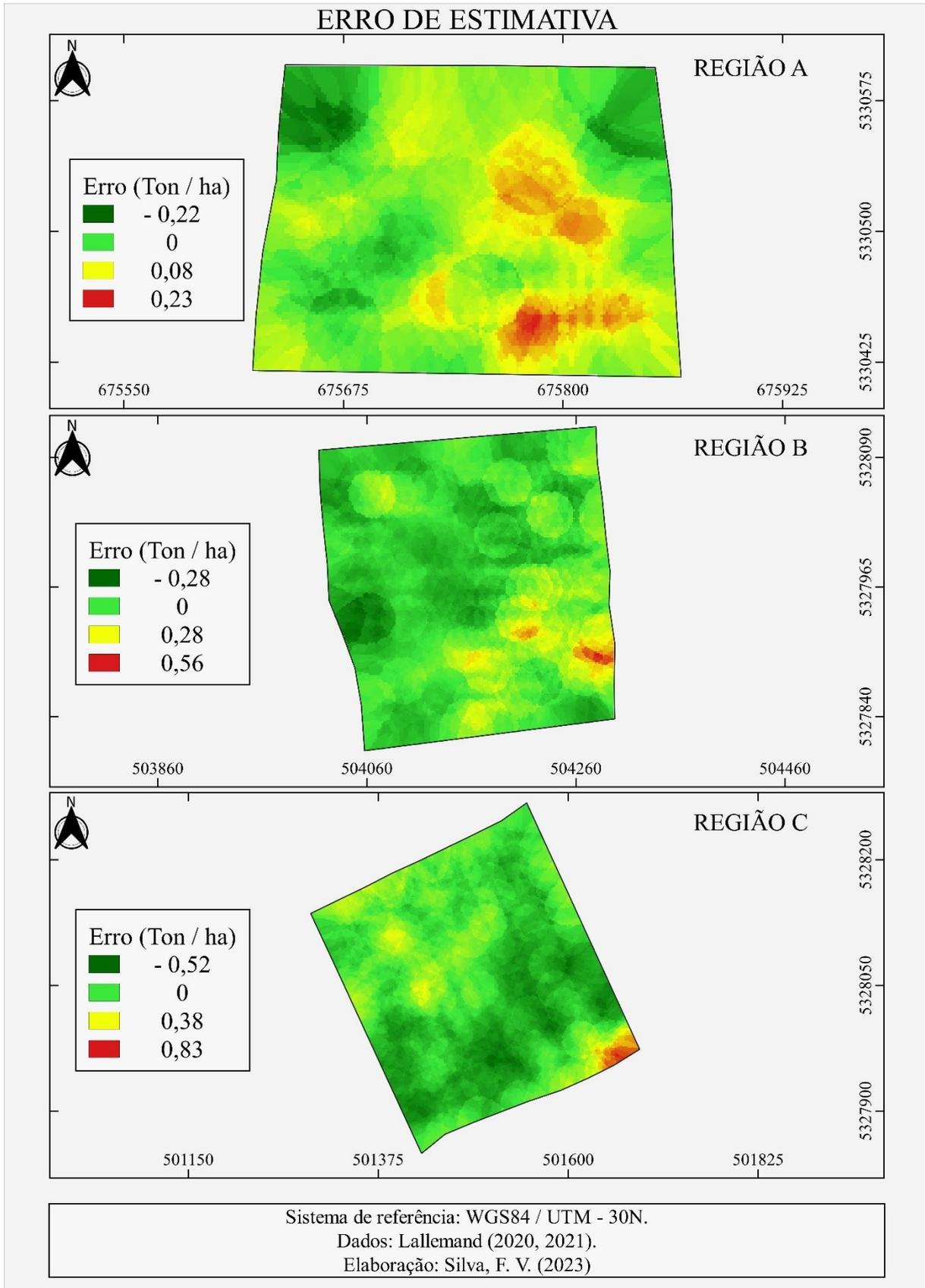
Enquanto isso, a curva espectral das bandas do infravermelho médio (correspondente às bandas SWIR1 e SWIR2) está relacionada com a presença e influência da água na planta. Os valores de reflectância tende a aumentar conforme as folhas perdem água e a diminuir quando a presença de água se torna maior. Com isso, fica claro que a alteração do equilíbrio hídrico afeta diretamente na reflectância da região do infravermelho médio. (Ponzoni, Shimabukuro, Kuplich 2012).

Como foi estabelecido, os melhores modelos de estimativa ocorreram na fase de enchimento dos grãos, período em que a planta ainda possui alta atividade fotossintética, o que justifica a presença das bandas e IVs presentes nos modelos. Os IVs MTCI e GLI são indicadores de clorofila, além disso o MTCI foi identificado como um IV capaz de estimar a produção primária bruta do trigo (Chaoyang et. al, 2009). Enquanto isso, as bandas presentes agregam informações a respeito dos nutrientes, fitomassa, água e pigmentação da lavoura.

5.6 Espacialização da produtividade do trigo

Para finalizar, como último resultado desta pesquisa tem-se a espacialização da produtividade estimada. Uma vez que há os valores preditos e os valores mensurados em campo, é possível computar a diferença entre as medidas e identificar as regiões em que o algoritmo errou, seja superestimando, seja subestimando a produtividade. Dessa forma, utilizando da krigagem foi possível especializar o erro cometido pelo algoritmo na estimativa (Figura 11).

Figura 11- Erro de estimativa cometido pelo algoritmo de predição

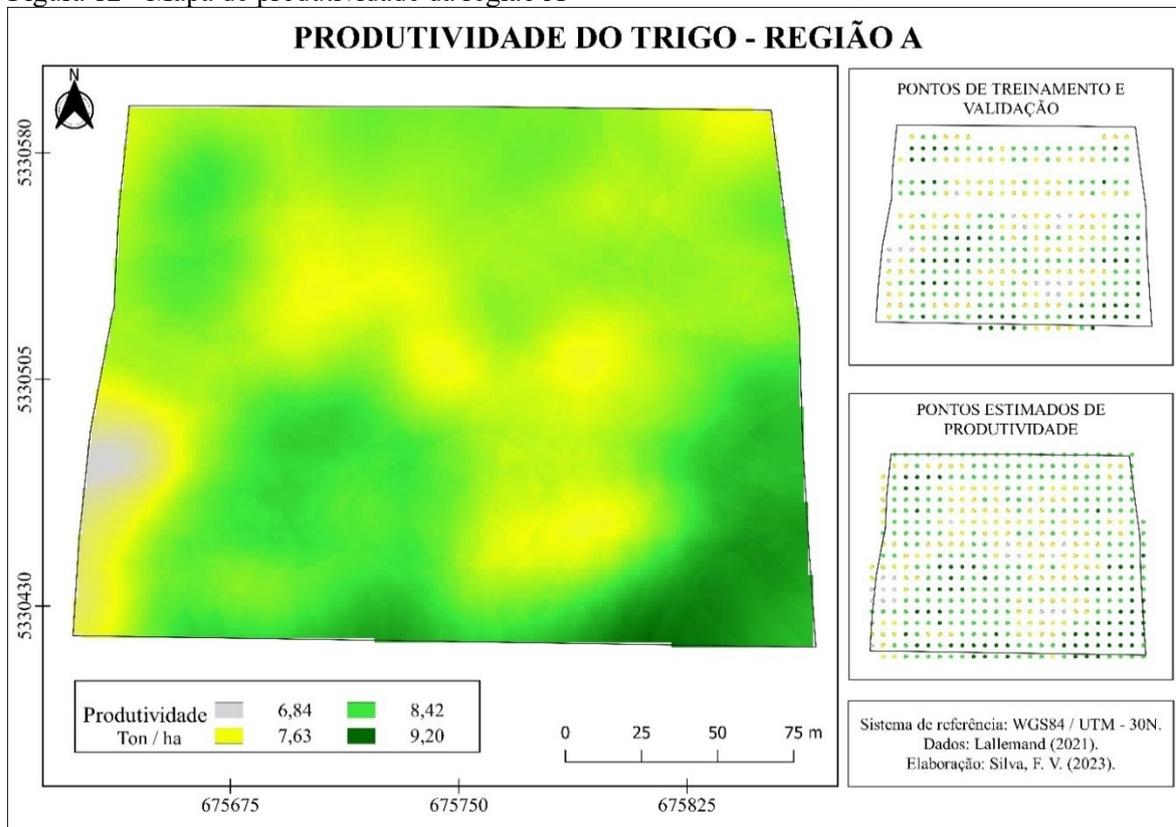


Fonte: o autor.

A superfície de erro gerada possibilitou realizar uma compensação na estimativa, dessa forma ela foi adicionada à superfície de produtividade gerando uma representação mais verdadeira.

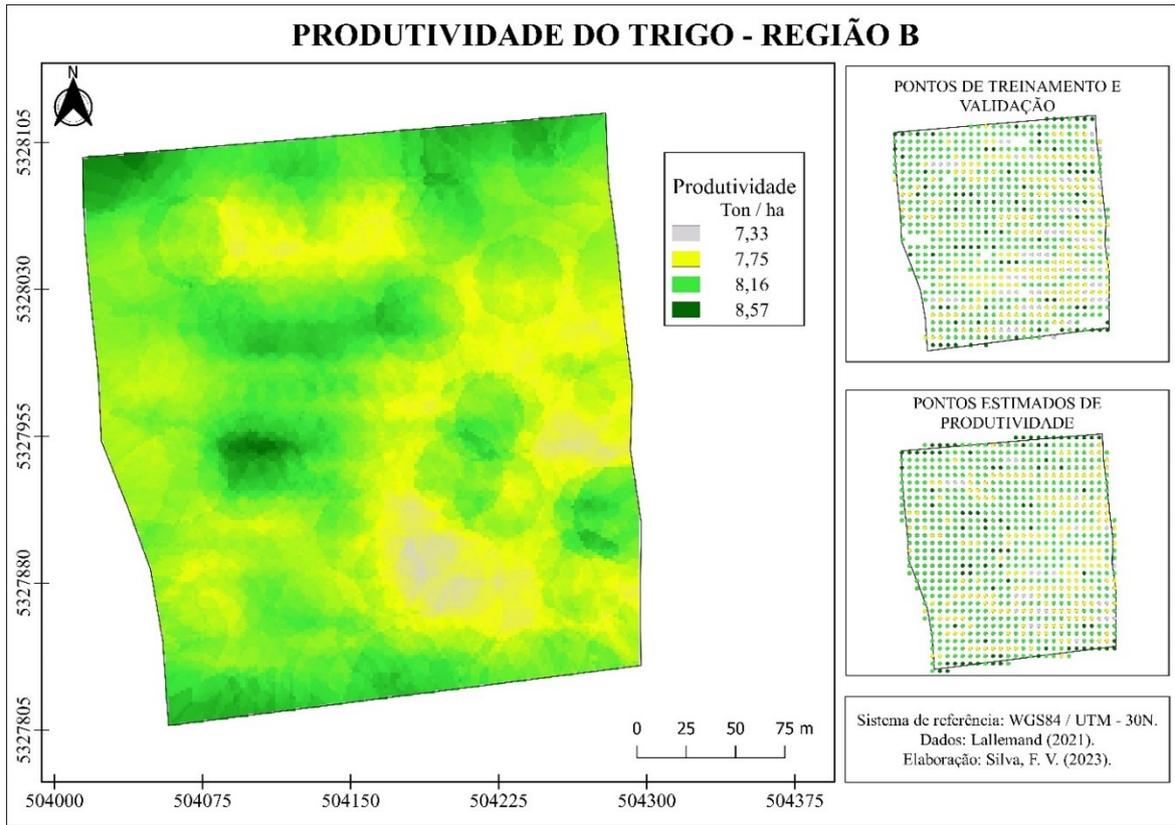
A geração de mapas de produtividade torna possível a análise espacial da distribuição produtiva da lavoura, possibilitando a identificação das áreas mais e menos produtivas. Pode-se notar por exemplo, que a região sudeste da lavoura A obteve os maiores valores de produtividade, ao passo que nas áreas B e C essa mesma região anotou os valores mais baixos de produtividade. As Figuras 12, 13 e 14 demonstram o mapa de produtividade do trigo para cada região de estudo.

Figura 12 - Mapa de produtividade da região A



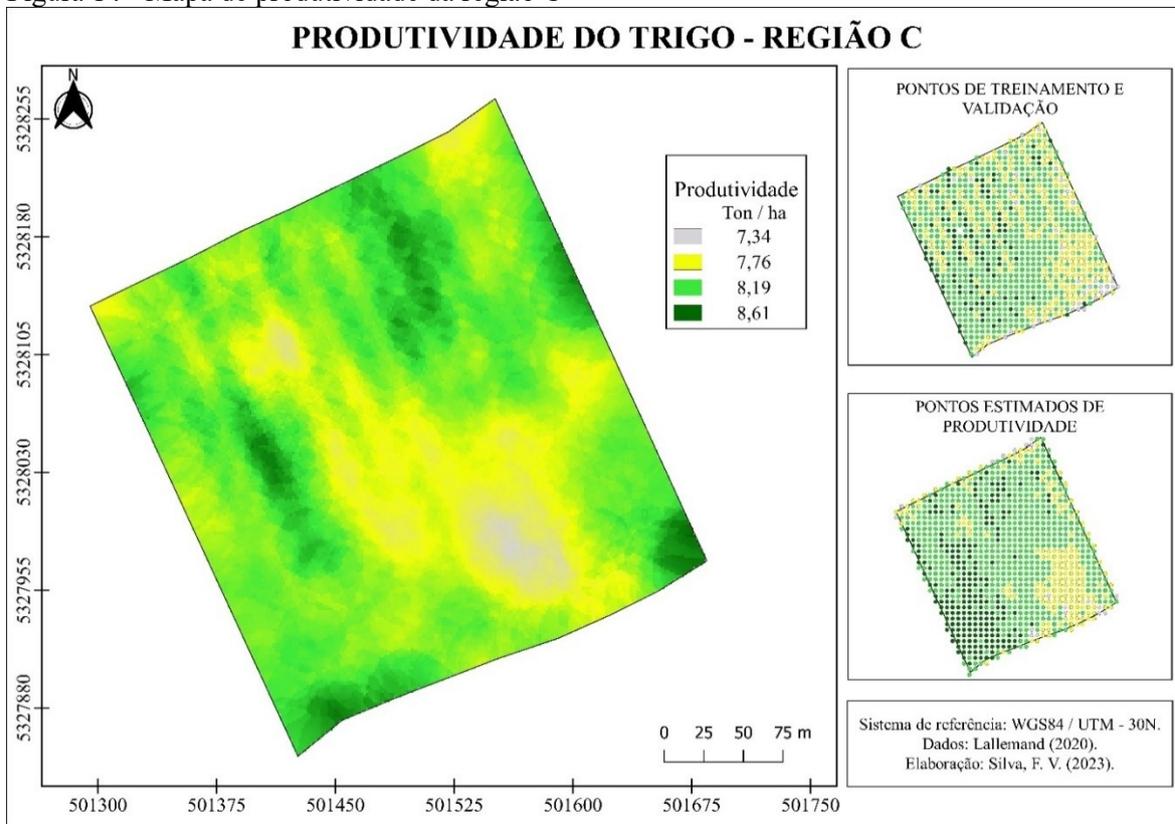
Fonte: o autor.

Figura 13 - Mapa de produtividade da região B



Fonte: o autor.

Figura 14 - Mapa de produtividade da região C



Fonte: o autor.

6 CONCLUSÃO

A estimativa de produtividade do trigo a partir de imagens multiespectrais com o uso de algoritmos de aprendizado de máquinas se mostrou bastante válida e promissora frente a estimativa tradicional. Visto que são acrescentados dados espectrais ao modelo, são adicionadas informações importantes quanto a saúde da planta, gerando estimativas assertivas com um menor esforço físico.

Após a análise de todos os processamentos foi possível identificar o estágio fenológico ideal para aquisição das imagens, os índices e as bandas que devem compor a estimativa, o melhor algoritmo e descartar a possibilidade de modelos regionais e série temporal.

Diante disso, ficou determinado o estágio de maturação ainda na fase de enchimento dos grãos, mais precisamente 51 dias antes da colheita, como o estágio ideal para aquisição das imagens. A estimativa gerada para as áreas de estudo obteve uma acurácia que variou entre 91 e 94%.

Com relação aos algoritmos, o RF se destacou e trouxe a melhor estimativa em 90% dos testes, já com relação aos modelos finais com os melhores desempenhos o algoritmo se fez presente em todas as estimativas, tornando assim o algoritmo de aprendizado de máquina que melhor estima a produtividade do trigo.

Com relação aos índices e bandas multiespectrais que devem compor o modelo de estimativa, 9 variáveis estiveram presentes nos melhores modelos, sendo eles: os índices MTCI e GLI, e as bandas do Red, Green, Red Edge 1, Red Edge 2, Red Edge 3, Near InfraRed e a banda Short Wave InfraRed 2. Dessa forma, para a aplicação do trabalho, futuramente não será necessário a realização de diversos testes e avaliações, basta adicionar estes atributos aos dados de produtividade e gerar o modelo, agilizando todo o processo de estimativa de produtividade da cultura.

À frente de tantos resultados positivos, a série temporal e o modelo regional se mostraram falhos na tentativa de melhorar os modelos de estimativa, onde houve uma piora nos parâmetros de avaliação R e RMSE. O modelo regional foi capaz de estimar a produtividade de uma área com os dados de treinamento de outra, porém com um RMSE de 0,82 e o R chegando a 0,08, mostrando ser uma abordagem inviável.

Com isso, podemos concluir que o sensoriamento remoto aplicado à agricultura de precisão tende a melhorar a estimativa de produtividade do trigo, com uma alta acurácia. Além disso, ficou evidente que esta é uma ótima ferramenta para a análise de produtividade local, mas que no âmbito regional e global ainda não é possível atingir resultados satisfatórios.

REFERÊNCIAS

AGRICULTURE VICTORIA. **A brief guide to estimating crop yields**. Australia: Agriculture Victoria, 2022. Disponível em: <https://agriculture.vic.gov.au/crops-and-horticulture/grains-pulses-and-cereals/crop-production/general-agronomy/a-brief-guide-to-estimating-crop-yields>. Acesso em: 22 maio 2023.

ALLIBHAI, E. **Hold-out vs. cross-validation in machine learning**. [S. l.], 2018. Disponível em: <https://medium.com/@ejaz/holdout-vs-cross-validation-in-machine-learning-7637112d3f8f>. Acesso em: 30 maio 2023.

ARANGUREN, M.; CASTELLON, A.; AIZPURUA, A. Wheat yield estimation with NDVI values using a proximal sensing tool. **Remote Sensing**, [S. l.], v. 12, n. 17, p. 1-17, 2020. DOI: <https://doi.org/10.3390/rs12172749>. Disponível em: <https://www.mdpi.com/2072-4292/12/17/2749>. Acesso em: 22 maio 2023.

ARGENTA, C. V. *et al.* Inteligência artificial na previsão da produtividade de grãos de trigo pela dose de fornecimento de nitrogênio com indicadores biológicos e ambientais. **Salão do conhecimento**, [S. l.], v. 6, n. 6, 2020. Disponível em: <https://publicacoeseventos.unijui.edu.br/index.php/salaconhecimento/article/view/17770>. Acesso em: 04 dez. de 2023.

BREIMAN, L. **Random Forests**. Statistics Department - University of California, Berkeley, CA 94720, 2001.

BUDU, E. **Random Forest vs. extremely: randomized trees**. [S. l.], 2023. Portal: Baeldung. Disponível em: <https://www.baeldung.com/cs/random-forest-vs-extremely-randomized-trees>. Acesso em 22 maio 2023.

CAIXETA-FILHO, V. J. **Pesquisa operacional: técnicas de otimização aplicadas a sistemas agroindustriais**. São Paulo: Atlas, 2001.

CARVALHO, S. R. de. **Índices de vegetação e produtividade de cultivares de trigo submetidas à diferentes épocas de semeadura**. Trabalho de Conclusão de Curso (Bacharelado em Agronomia) – Universidade Estadual do Rio Grande do Sul, Vacaria, RS, 2021. Disponível em: https://repositorio.uergs.edu.br/xmlui/bitstream/handle/123456789/1542/_indices_de_vegetacao_e_produtividade_de_cultivares_de_trigo_submetidas.pdf?sequence=-1&isAllowed=y. Acesso em: 22 maio 2023.

CHAOYANG, W. *et al.* **Remote estimation of gross primary production in wheat using chlorophyll-related vegetation indices**. DOI: <https://doi.org/10.1016/j.agrformet.2008.12.007>. Acesso em: 31 out. 2023.

CONAB. **Trigo: análise mensal – abril 2022**. Brasília, DF: CONAB, 2022. Disponível em: <https://www.conab.gov.br/info-agro/analises-do-mercado-agropecuario-e-extrativista/analises-do-mercado/historico-mensal-de-trigo/item/18197-trigo-analise-mensal-abril-2022>. Acesso em: 22 maio 2023.

DASH, J.; CURRAN, J. P. Evaluation of the MERIS terrestrial chlorophyll index (MTCI). **Advances in Space Research**, [S. l.], v. 39, n. 1, p. 100-104, 2007. DOI: <https://doi.org/10.1016/j.asr.2006.02.034>. Disponível em: https://www.sciencedirect.com/science/article/pii/S0273117706000834?casa_token=X8ukKjL_ipYAAAAA:W1GxXIDvAZaLqbRPlclo1AXEisQVPAL3VNfrX8eir5m5h87iFXILYq79JX3zrnfvW0OqqFqht7I. Acesso em: 22 maio 2023.

DAUGHTRY T. S. C. *et al.* Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. **Remote Sensing of Environment**, [S. l.], v. 74, n. 2, p. 229-239, 2000. DOI: [https://doi.org/10.1016/S0034-4257\(00\)00113-9](https://doi.org/10.1016/S0034-4257(00)00113-9). Disponível em: <https://reader.elsevier.com/reader/sd/pii/S0034425700001139?token=3BD7075DF3BC1B50907EAAF3F7B11525A888931506846AD96B6C63BCA15C3057C53748A0A17336E971919A04DDB0D76B&originRegion=us-east-1&originCreation=20230512195209>. Acesso em: 22 maio 2023.

EMBRAPA TERRITORIAL. **Sentinel**: missão Sentinel. Campinas, 2018. Disponível em: <https://www.embrapa.br/satelites-de-monitoramento/missoes/sentinel#:~:text=Possui%20resolu%C3%A7%C3%A3o%20de%20500m%20para,1%20km%20de%20resolu%C3%A7%C3%A3o%20espacial>. Acesso em: 22 maio 2023.

EPIPHANEO, J. C. N.; ALMEIDA JÚNIOR, A. C. de.; FORMAGGIO, A. R. Desenvolvimento do trigo avaliado com dois índices de vegetação. *In*: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 8., 1996, Salvador, BA. **Anais** [...]. São José dos Campos: INPE. p. 19-24. Disponível em: <http://mar.tecnico.ipe.br/col/sid.inpe.br/deise/1999/02.01.11.34/doc/T158.pdf>. Acesso em: 22 maio 2023.

ESA. **Sentinel-2 - Data products - sentinel handbook**. [S. l.], 201?. Disponível em: <https://sentinel.esa.int/web/sentinel/missions/sentinel-2/data-products>. Acesso em: 5 jun. 2023.

FAO. **The state of food security and nutrition in the world 2022**: Chapter 2 - food security and nutrition around the world. Food and Agriculture Organization of the United Nations, [S. l.], 2022a. Disponível em: <https://www.fao.org/3/cc0639en/online/sofi-2022/food-security-nutrition-indicators.html>. Acesso em: 22 maio 2023.

FAO. Chapter 2: production, trade and prices of commodities. *In*: FAO. **World food and agriculture 2022 statistical yearbook**. Rome: FAO, 2022b. p. 154-250. Disponível em: <https://www.fao.org/3/cc2211en/cc2211en.pdf>. Acesso em: 22 maio 2023.

FITZ, P. R. Classificação de imagens de satélite e índices espectrais de vegetação: uma análise comparativa. Geosul, Florianópolis, 2020. v. 35, n. 76, p. 171-188, set./dez. 2020. DOI: <https://doi.org/10.5007/2177-5230.2020v35n76p171>. Disponível em: <https://periodicos.ufsc.br/index.php/geosul/article/view/67414/44385>. Acesso em: 22 maio 2023.

GAMEIRO, S. *et al.* Avaliação da cobertura vegetal por meio de índices de vegetação (NDVI, SAVI e IAF) na Sub-Bacia Hidrográfica do Baixo Jaguaribe, CE. **Terrae**, [S. l.], v. 13, n. 1-2, p. 15-22, 2016. Disponível em: <https://www.ige.unicamp.br/terrae/V13/PDFv13/TE067-2.pdf>. Acesso em: 22 maio 2023.

HABOUDANE, D. *et al.* Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. **Remote sensing of environment**, [S. l.], v. 90, n. 3, p. 337-352, 2004. DOI: <https://doi.org/10.1016/j.rse.2003.12.013>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0034425704000264?via%3Dihub>. Acesso em: 22 maio 2023.

HAYKIN, S. **Redes neurais: princípios e práticas**. Porto Alegre, RS. ed. Bookman, 2001. cap 1, p 27-74.

HUNT JUNIOR, R. E. *et al.* Remote sensing leaf chlorophyll content using a visible band index. **Agronomy Journal**, [S. l.], v. 103, n. 4, p. 1090-1099, 2011. DOI: <https://doi.org/10.2134/agronj2010.0395>. Disponível em: <https://naldc.nal.usda.gov/download/53737/PDF>. Acesso em: 22 maio 2023.

IDB. **Index database**: a database for remote sensing. [S. l.], 2023. Disponível em: <https://www.indexdatabase.de/db/i.php?offset=1>. Acesso em: 22 maio 2023.

JACTO. **Acompanhe as principais estatísticas da agricultura mundial!**. [S. l.], 2021. Portal: JACTO. Disponível em: <https://blog.jacto.com.br/agricultura-mundial/>. Acesso em: 22 maio 2023.

JUSTICE, C. O. *et al.* The moderate resolution imaging spectroradiometer (MODIS): land remote sensing for global change research. **IEEE Transactions on Geoscience and Remote Sensing**, [S. l.], v. 36, n. 4, p. 1228-1249, 1997. DOI: 10.1109/36.701075. Disponível em: https://www.researchgate.net/publication/312896120_The_moderate_resolution_imaging_spectroradiometer_MODIS_land_remote_sensing_for_global_change_research Acesso em: 22 maio 2023.

KHANAL, S. *et al.* Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. **Computers and electronics in agriculture**, [S. l.], v. 153, p. 213-225, 2018. DOI: <https://doi.org/10.1016/j.compag.2018.07.016>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0168169918300334?via%3Dihub>. Acesso em: 22 maio 2023.

LADEIRA, S. **Somos 8 bilhões de pessoas e não vamos parar por aí**. [S. l.], 2022. Portal: G1. Disponível em: <https://especiais.g1.globo.com/mundo/2022/8-bilhoes-de-pessoas/#:~:text=Mundo&text=Em%202022%2C%20a%20popula%C3%A7%C3%A3o%20mundial,15%20de%20novembro%20de%202022>. Acesso em: 22 maio 2023.

LETRAS AMBIENTAIS. Os 9 passos para se estimar a produtividade agrícola por satélites. **Letras Ambientais**, [S.l.], 2021. Disponível em: <https://www.letrasambientais.org.br/posts/os-9-passos-para-se-estimar-a-productividade-agricola-por-satelites#>. Acesso em: 22 maio 2023.

LOUHAICHI, M.; BORMAN, M. M.; JOHNSON, D. E. Spatially located platform and aerial photography for documentation of grazing impacts on wheat. **Geocarto International**, Hong Kong, v. 16, n. 1, 2001. DOI: <https://doi.org/10.1080/10106040108542184>. Disponível em:

<https://www.tandfonline.com/doi/epdf/10.1080/10106040108542184?needAccess=true&role=button>. Acesso em: 22 maio 2023.

MAGNEY, T. S. *et al.* Proximal NDVI derived phenology improves in-season predictions of wheat quantity and quality. **Agriculture and Forest Meteorology**. [S. l.], v. 217, n. 15, p. 46-60, 2016. DOI: <https://doi.org/10.1016/j.agrformet.2015.11.009>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0168192315007601?via%3Dihub>. Acesso em: 22 maio 2023.

MICHELON, G. K. *et al.* MÁQUINA DE VETORES DE SUPORTE PARA ESTIMAR A PRODUTIVIDADE DA SOJA. **Revista Engenharia na Agricultura - REVENG**, [S. l.], v. 25, n. 3, p. 240–248, 2017. DOI: 10.13083/reveng.v25i3.745. Disponível em: <https://periodicos.ufv.br/reveng/article/view/669>. Acesso em: 4 dez. 2023.

MOREIRA, M. A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**. 4. ed. atual. e ampl Viçosa: Ed. da UFV; 2011.

MYNENI, R. B.; NEMANI, R. R.; RUNNING, S. W. Estimation of global leaf area index and absorbed par using radiative transfer models. **IEEE Transactions on Geoscience and Remote Sensing**, [S. l.] v. 35, n. 6, p. 1380-1393, 1997. Disponível em: https://sites.bu.edu/cliveg/files/2013/12/Myneni_IEEE_1997.pdf. Acesso em: 22 maio 2023.

PAIVA, M. C.; TSUKAHARA, Y. R.; FRANÇA, B. G. Estimativa de produtividade do trigo via sensoriamento remoto no município de Pirai do Sul, no Estado do Paraná. **XVI Simpósio Brasileiro de Sensoriamento Remoto**, Foz do Iguaçu, PR, 2013. Disponível em: <http://marte2.sid.inpe.br/col/dpi.inpe.br/marte2/2013/05.29.00.34.23/doc/p1220.pdf>. Acesso em: 26 maio de 2023.

PINTO, A. R. da *et al.* Trigo: alimentação e tecnologia. *In*: OLIVEIRA NETO, A. A. de O.; SANTOS, C. M. R. (org.). **A cultura do trigo**. Brasília: Conab, 2017. p. 15-26. Disponível em: https://www.conab.gov.br/uploads/arquivos/17_04_25_11_40_00_a_cultura_do_trigo_versao_digital_final.pdf. Acesso em: 22 maio 2023.

PIRES, F. L. J. *et al.* Uso de sensor óptico ativo para caracterização do perfil de NDVI em dosséis de trigo submetidos a diferentes estratégias de manejo. *In*: BERNARDI, A. C. de C. *et al.* (ed.). **Agricultura de precisão: resultados de um novo olhar**. Brasília, DF: Embrapa, 2014. p. 279-286. Disponível em: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1006307/uso-de-sensor-optico-ativo-para-caracterizacao-do-perfil-de-ndvi-em-dosseis-de-trigo-submetidos-a-diferentes-estrategias-de-manejo>. Acesso em: 22 maio 2023.

PONZONI, F. J.; SHIMABUKURO, Y. E.; KUPLICH, T. M. **Sensoriamento remoto da vegetação**. 2ª edição – atualizada e ampliada. São Paulo: Oficina de Textos, 2012.

QGIS. A liderança do SIG de código aberto. [S. l.], 2023. Disponível em: https://www.qgis.org/pt_BR/site/about/index.html. Acesso em: 25 maio 2023.

QI, J. *et al.* A modified soil adjusted vegetation index. **Remote Sensing of Environment**, [S. l.], v. 48, n. 2, p. 119-126, 1994. DOI: [https://doi.org/10.1016/0034-4257\(94\)90134-1](https://doi.org/10.1016/0034-4257(94)90134-1). Disponível em:

<https://www.sciencedirect.com/science/article/pii/0034425794901341?via%3Dihub>. Acesso em: 22 maio 2023.

RICHARDS, A. J.; JIA X. **Remote sensing digital images analysis: an introduction**. Germany, 2006. 4^o edition. ed: Springer.

RONDEAUX, G.; STEVEN, M.; BARET, F. Optimization of soil-adjusted vegetation indices. **Remote Sensing of Environment**, [S. l.], v. 55, n. 2, p. 95-107, 1996. DOI: [https://doi.org/10.1016/0034-4257\(95\)00186-7](https://doi.org/10.1016/0034-4257(95)00186-7). Disponível em: <https://www.sciencedirect.com/science/article/pii/0034425795001867?via%3Dihub>. Acesso em: 22 maio 2023.

ROUSE, W. J. *et al.* Monitoring vegetation systems in the great plains with ERTS. **NASA** [S. l.], v. 1, 1974. Disponível em: <https://ntrs.nasa.gov/citations/19740022614>. Acesso em: 22 maio 2023.

SAATH, O. C. K. de; FACHINELLO, L. A. Crescimento da demanda mundial de alimento e restrições do fator terra no Brasil. *Revista de Economia e Sociologia Rural*, Brasília, DF, 6 (2), 2018. DOI: <https://doi.org/10.1590/1234-56781806-94790560201>. Acesso em 26 maio 2023.

SANTOS, H. P. dos; PIRES, J. L. F.; FONTANELI, R. S. Cultivo de trigo: semeadura e rotação de culturas. **Embrapa Trigo**, Brasília, DF, 2014. Disponível em: https://www.spo.cnptia.embrapa.br/conteudo?p_p_id=conteudoportlet_WAR_sistemasdeproducao16_1gal1ceportlet&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-1&p_p_col_count=1&p_r_p_-76293187_sistemaProducaoId=3704&p_r_p_-996514994_topicoId=3047. Acesso em: 22 maio 2023.

SCHEEREN, P. L.; CASTRO, R. L. de; CAIERÃO, E. Botânica, morfologia e descrição fenotípica. *In*: BOREM, A.; SCHEEREN, P. L. (ed.). **Trigo: do plantio à colheita**. Viçosa, MG: Ed. UFV, 2015. p. 35-55.

SHIRATSUCHI, S. L. *et al.* Sensoriamento remoto: conceitos básicos e aplicações na agricultura de precisão. *In* **Agricultura de Precisão: resultados de um novo olhar**. Brasília, DF: ed. RB Gráfica Digital Eirelli. Cap 4, p 58-73.

SILVA, J. P. da.; ZULLO JÚNIOR, J.; ROMANI, L. A. S. Active learning e sua aplicação no monitoramento da cana-de-açúcar utilizando o algoritmo SVM. *In*: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 11., 2017, Campinas-SP, **Anais** [...]. Campinas: Editora da Unicamp, 2017. p. 119-128. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/169620/1/Active-SBIAgro.pdf>. Acesso em: 22 maio 2023.

SOUZA, M. R. Q. de. **Estimativa de biomassa de trigo usando imagens de alta resolução espacial a partir de veículo aéreo não tripulado**. 2018. Dissertação (Mestrado em Engenharia Elétrica) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/189943/001090104.pdf?sequence=1>. Acesso em: 22 maio 2023.

UNITED NATIONS. **UN population division data portal**: Interactive access to global demographic indicators. [S. l.], [2022?]. Portal: Population Division Data. Disponível em: <https://population.un.org/dataportal/home>. Acesso em: 22 maio 2023.

VINCINI, M.; FRAZZI, E.; D'ALESSIO, P. A broad-band leaf chlorophyll vegetation index at the canopy scale. **Precision Agric**, [S. l.], v. 9, p. 303-319, 2008. DOI: <https://doi.org/10.1007/s11119-008-9075-z>. Disponível em: <https://link.springer.com/article/10.1007/s11119-008-9075-z>. Acesso em: 22 maio 2023.

WAIKATO. **Weka 3**: Software de aprendizado de máquina em Java. [S. l.], 201?. Disponível em: <https://www.cs.waikato.ac.nz/ml/weka/index.html>. Acesso em 25 maio 2023.