

MINISTÉRIO DA EDUCAÇÃO
INSTITUTO FEDERAL DE EDUCAÇÃO CIÊNCIA E TECNOLOGIA DE GOIÁS
DEPARTAMENTO ACADÊMICO DA ÁREA III – CÂMPUS GOIÂNIA
COORDENAÇÃO DE GEOMÁTICA
CURSO DE ENGENHARIA CARTOGRÁFICA E DE AGRIMENSURA

JOÃO PEDRO DE SOUSA COSTA

TRABALHO DE CONCLUSÃO DE CURSO
PREDIÇÃO DA PRODUTIVIDADE DE CANA-DE-AÇÚCAR A
PARTIR DE PARÂMETROS BIOMÉTRICOS E IMAGENS ORBITAIS
DO SENSOR MSI/SENTINEL-2

Goiânia, GO

2021

JOÃO PEDRO DE SOUSA COSTA

**PREDIÇÃO DA PRODUTIVIDADE DE CANA-DE-AÇÚCAR A PARTIR DE
PARÂMETROS BIOMÉTRICOS E IMAGENS ORBITAIS DO SENSOR
MSI/SENTINEL-2**

Trabalho de Conclusão de Curso (TCC) apresentado à Coordenação de Geomática, do Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Campus Goiânia, como parte dos requisitos necessários para a obtenção do título de Bacharel em Engenharia Cartográfica e de Agrimensura.

Orientador: Prof. Dr. Max Well de Oliveira Rabelo

Goiânia, GO

2021

C837e Costa, João Pedro de Sousa.

Predição da produtividade de cana-de-açúcar a partir de parâmetros biométricos e imagens orbitais do sensor MSI/sentinel-2/ João Pedro de Sousa Costa. – Goiânia : Instituto Federal de Educação, Ciência e Tecnologia de Goiás, 2021.
89f. ; il.

Orientador: Prof. Dr. Max Well de Oliveira Rabelo.

TCC (Trabalho de Conclusão de Curso) – Curso de Engenharia Cartográfica e de Agrimensura, Coordenação de Geomática, Departamento Acadêmico da Área III, Instituto Federal de Educação, Ciência e Tecnologia de Goiás.
Inclui apêndice.

1. Sensoriamento remoto. 2. Produtividade. 3. Cana de açúcar. I. Rabelo, Max Well de Oliveira (orientador). II. Instituto Federal de Educação, Ciência e Tecnologia de Goiás. III. Título.

CDD 526

Ficha catalográfica elaborada pela bibliotecária Lana Cristina Dias Oliveira CRB1/ 2.631
Biblioteca Professor Jorge Félix de Souza,
Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Câmpus Goiânia.

**TERMO DE AUTORIZAÇÃO PARA DISPONIBILIZAÇÃO
NO REPOSITÓRIO DIGITAL DO IFG - ReDi IFG**

Com base no disposto na Lei Federal nº 9.610/98, AUTORIZO o Instituto Federal de Educação, Ciência e Tecnologia de Goiás, a disponibilizar gratuitamente o documento no Repositório Digital (ReDi IFG), sem ressarcimento de direitos autorais, conforme permissão assinada abaixo, em formato digital para fins de leitura, download e impressão, a título de divulgação da produção técnico-científica no IFG.

Identificação da Produção Técnico-Científica

- | | |
|--|---|
| ☐ Tese | ☐ Artigo Científico |
| ☐ Dissertação | ☐ Capítulo de Livro |
| ☐ Monografia – Especialização | ☐ Livro |
| ☒ TCC - Graduação | ☐ Trabalho Apresentado em Evento |
| ☐ Produto Técnico e Educacional - Tipo: _____ | |

Nome Completo do Autor: João Pedro de Sousa Costa

Matrícula: 20161011190075

Título do Trabalho: Predição da produtividade de cana-de-açúcar a partir de parâmetros biométricos e imagens orbitais do sensor MSI/sentinel-2

Autorização - Marque uma das opções

- ☐ Autorizo disponibilizar meu trabalho no Repositório Digital do IFG (acesso aberto);
- ☒ Autorizo disponibilizar meu trabalho no Repositório Digital do IFG somente após a data **10/08/2022** (Embargo);
- ☐ Não autorizo disponibilizar meu trabalho no Repositório Digital do IFG (acesso restrito).

Ao indicar a opção **2 ou 3**, marque a justificativa:

- ☐ O documento está sujeito a registro de patente.
☒ O documento pode vir a ser publicado como livro, capítulo de livro ou artigo.
☐ Outra justificativa: _____

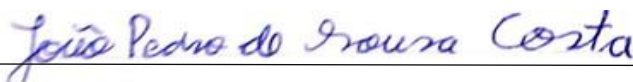
DECLARAÇÃO DE DISTRIBUIÇÃO NÃO-EXCLUSIVA

O/A referido/a autor/a declara que:

- o documento é seu trabalho original, detém os direitos autorais da produção técnico-científica e não infringe os direitos de qualquer outra pessoa ou entidade;
- obteve autorização de quaisquer materiais incluídos no documento do qual não detém os direitos de autor/a, para conceder ao Instituto Federal de Educação, Ciência e Tecnologia de Goiás os direitos requeridos e que este material cujos direitos autorais são de terceiros, estão claramente identificados e reconhecidos no texto ou conteúdo do documento entregue;
- cumpriu quaisquer obrigações exigidas por contrato ou acordo, caso o documento entregue seja baseado em trabalho financiado ou apoiado por outra instituição que não o Instituto Federal de Educação, Ciência e Tecnologia de Goiás.

Goiânia
Local

, 17/04/2021.
Data



Assinatura do Autor e/ou Detentor dos Direitos Autorais



INSTITUTO FEDERAL
Goiás

MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE GOIÁS
CÂMPUS GOIÂNIA

ATA Nº ECA 05 - TCC/2021.1

ATA DA SESSÃO PÚBLICA DE APRESENTAÇÃO E DEFESA DO TRABALHO DE CONCLUSÃO DE CURSO

Aos 19 dias do mês de março do ano de dois mil e vinte e um, às 13:00 horas, na sala virtual meet.google.com/mqp-hhuz-pxb, foi realizada a sessão pública de apresentação e defesa do Trabalho de Conclusão de Curso do Graduando **João Pedro de Sousa Costa** (matrícula 20161011190075) do curso de Bacharelado em Engenharia Cartográfica e de Agrimensura, do Departamento III, Câmpus Goiânia, do Instituto Federal de Goiás, no primeiro semestre do ano de dois mil e vinte e um. A banca foi composta pelos seguintes membros:

- Max Well de Oliveira. Rabelo (Professor-Orientador)
- Lucas Rios do Amaral (Avaliador Externo)
- Édipo Henrique Cremon
- Giovanni de Araújo Boggione

sob a presidência da(o) primeira(o). O projeto de trabalho de conclusão de curso tem como título **“Predição da produtividade de cana-de-açúcar a partir de parâmetros biométricos e imagens orbitais do sensor MSI/Sentinel-2”**, sob orientação do Prof. **Max Well de O. Rabelo**. Após a apresentação do Projeto de Trabalho de Conclusão de Curso, tendo sido o autor arguido pela Banca Examinadora, a nota obtida foi 9,5 (nove virgula cinco).

Encerra-se a sessão às 16 horas e 30 minutos. Eu, Profa. Marina Alberti Macedo, dato e assino a presente ata que segue assinada por todos os membros da Banca Examinadora.

Profa. Marina Alberti Macedo (Assinado eletronicamente)

Prof. Max Well de Oliveira Rabelo (Assinado eletronicamente)

Prof. Lucas Rios do Amaral (Assinado eletronicamente)

Prof. Édipo Henrique Cremon (Assinado eletronicamente)

Prof. Giovanni de Araújo Boggione (Assinado eletronicamente)

Documento assinado eletronicamente por:

- **Marina Alberti Macedo**, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 22/03/2021 13:28:58.
- **Lucas Rios do Amaral**, LUCAS RIOS DO AMARAL - 332105 - PROFESSOR ASSISTENTE DE REGÊNCIA DE CLASSE - UNIVERSIDADE ESTADUAL DE CAMPINAS (46068425000133), em 19/03/2021 15:25:03.
- **Edipo Henrique Cremon**, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 19/03/2021 15:21:59.
- **Giovanni de Araujo Boggione**, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 19/03/2021 15:21:03.
- **Max Well de Oliveira Rabelo**, PROFESSOR ENS BASICO TECN TECNOLOGICO, em 19/03/2021 15:20:45.

Este documento foi emitido pelo SUAP em 04/03/2021. Para comprovar sua autenticidade, faça a leitura do QRCode ao lado ou acesse <https://suap.ifg.edu.br/autenticar-documento/> e forneça os dados abaixo:

Código Verificador: 137303

Código de Autenticação: 41f98baa47



Instituto Federal de Educação, Ciência e Tecnologia de Goiás

Rua 75, nº 46, Centro, GOIÂNIA / GO, CEP 74055-110

(62) 3227-2735 (ramal: 2735)

DEDICATÓRIA

Aos meus pais, que sem eles eu não seria nada e que sempre me apoiaram em tudo.

A minha família, pelo apoio e amor.

Aos meus amigos, que mesmo distante sempre estiveram comigo.

Ao meu orientador, agradeço pela amizade, por tudo que fez e faz por mim.

AGRADECIMENTOS

Aos meus pais e a minha família, por todo apoio, amor e educação. Que sempre estiveram me incentivando durante a graduação, pois, sem eles, eu nunca teria essa oportunidade.

Aos meus amigos, que fui ganhando durante a faculdade e no mercado de trabalho. Mas em especial ao grupo “Os Meno” e o MiniJão que desde do ensino fundamental e Médio, estiveram comigo, me ouvindo, me distraíndo nos momentos de estresse da faculdade e me apoiando. A minha namorada, Myssao, que sempre me apoiou, principalmente nos momentos mais complicados do trabalho e da faculdade.

Ao Instituto Federal de Educação, Ciência e Tecnologia de Goiás (IFG), instituição que contribui para minha evolução, crescimento pessoal e profissional.

Aos professores do Curso de Engenharia Cartografia e de Agrimensura, que sempre estiveram presentes ao esclarecimento de dúvidas, sempre receptivos e dispostos a ajudar nesses 5 anos de curso.

Ao professor e orientador Dr. Max Well de Oliveira Rabelo, que desde 2017, na primeira Iniciação Científica, vem me orientando, me auxiliando, me aconselhando e sendo um grande amigo que o curso me deu. Agradeço por tudo, e se fosse escrever tudo não caberia nesses Agradecimentos.

Agradecer a empresa Sanovita Brasil por cobrir os gastos para realização da parte do campo no Paraná e a Fazenda Miss Brasil por receber, auxiliar e ceder os dados de produtividade e biometria utilizados nesse trabalho.

À todas pessoas que de alguma forma contribuíram de forma direta ou indireta neste trabalho.

RESUMO

Título: Predição de produtividade de cana de açúcar a partir de parâmetros biométricos e imagens orbitais do sensor MSI/Sentinel 2

Autor: João Pedro de Sousa Costa

Orientador: Dr. Max Well de Oliveira Rabelo

Durante toda safra da cana de açúcar o monitoramento da cultura é fundamental para acompanhar os resultados do manejo adotado, identificar problemas que podem ocasionar baixas produtividades, e estabelecer um planejamento logístico de acordo com a produção de cada área cultivada. Para suprir esta demanda, são empregadas técnicas de campo para mensurar as características da cultura e estimar a produtividade dos talhões com antecedência. Contudo a maioria das práticas atualmente empregadas estimam somente a produtividade média, desconsiderando sua variabilidade espacial. Neste sentido o presente trabalho teve como objetivo gerar um modelo de regressão linear múltipla que permita prever a produtividade da cana de açúcar com cinco meses de antecedência à colheita, a partir de dados de sensoriamento remoto e dados biométricos obtidos em campo. Para elaboração deste trabalho, os dados biométricos utilizados foram o diâmetro médio do colmo (DMC), o número de perfilhos industrializáveis (NPI), estatura média dos colmos (EMC) e Tonelada de Cana por Hectares (TCH); imagens orbitais do sensor MSI do satélite Sentinel 2 na data da realização das biometrias; e os dados de produtividade obtidos pelo monitor de produtividade da colhedora CH670 da marca John Deere. Foi criado 16 cenários com diferentes combinações de dados biométricos e índices de vegetação, e um cenário baseado exclusivamente em dados de sensoriamento remoto. Para obtenção dos modelos de regressão linear múltipla utilizou-se o software WEKA. Através da análise de correlação e ocorrência, os parâmetros biométricos EMC e NPI e os índices *Modified Chlorophyll Absorption Ratio Index* (MCARI1) e *Canopy Chlorophyll Content Index* (CCCI), apresentaram maior correlação com a produtividade. Na validação cruzada o cenário que utiliza dos parâmetros DMC, NPI, EMC e CCCI, apresentou o menor RMSE, com valor de 3,62 ton.ha⁻¹ e um R² ajustado de 0,72. Devido à incompatibilidade de escala entre os dados biométricos e os dados de SR apenas os modelos que não utilizam dados biométricos em suas variáveis foram aplicados espacialmente. O melhor cenário somente com dados de SR obteve um RMSE de 8,15 ton.ha⁻¹ nos dados de validação.

Palavras-chave: Sensoriamento Remoto Orbital, Índices de Vegetação, Agricultura de Precisão, Predição de Produtividade, Variabilidade Espacial

ABSTRACT

Title: Sugarcane yield estimation from biometric parameters and orbital images from the MSI sensor /Sentinel 2

Author: João Pedro de Sousa Costa

Advisor: Dr. Max Well de Oliveira Rabelo

During the entire sugarcane harvest, crop monitoring is essential to monitor the results of the management adopted, identify problems that can cause low yields, and establish logistical planning according to the production of each cultivated area. To meet this demand, field techniques are employed to measure the characteristics of the crop and estimate the productivity of the plots in advance. However, most practices currently employed estimate only average productivity, disregarding its spatial variability. In this sense, the present work had as objective to generate a model of multiple linear regression that allows to estimate the productivity of the sugar cane five months before the harvest, from remote sensing data and biometric data obtained in the field. For the elaboration of this work, the biometric data used were the average stem diameter (DMC), the number of industrializable tillers (NPI), average height of the stems (EMC) and Ton of Sugarcane per Hectares (TCH); orbital images from the Sentinel 2 satellite's MSI sensor on the date of the biometrics; and the productivity data obtained by the productivity monitor of the John Deere CH670 harvester. 16 scenarios were created with different combinations of biometric data and vegetation indices, and a scenario based exclusively on remote sensing data. To obtain the multiple linear regression models, the WEKA software was used. Through the correlation and occurrence analysis, the EMC and NPI biometric parameters and the Modified Chlorophyll Absorption Ratio Index (MCARI1) and Canopy Chlorophyll Content Index (CCCI) indices, showed a higher correlation with productivity. In cross-validation, the scenario using the DMC, NPI, EMC and CCCI parameters, presented the lowest RMSE, with a value of 3.62 ton.ha-1 and an adjusted R^2 of 0.72. Due to the incompatibility of scale between the biometric data and the SR data, only the models that do not use biometric data in their variables were applied spatially. The best scenario with SR data alone obtained an RMSE of 8.15 ton.ha-1 in the validation data.

Key words: Orbital Remote Sensing, Vegetation Indexes, Precision Agriculture, Productivity Prediction, Spatial Variability.

LISTA DE FIGURAS

Figura 1 - Ciclo fenológico da cana-de-açúcar	21
Figura 2 - Folhas em diferentes níveis de senescência.	26
Figura 3 - Arquivo vetorial de pontos de produtividade.....	29
Figura 4 - Colhedora de cana de açúcar com a localização dos diferentes pontos de sensoriamento que podem ser usados para gerar dados de fluxo de cana.	30
Figura 5 - Localização da Área de Estudo - Talhão 10 da Fazenda Miss Brasil.	32
Figura 6 - Colhedora CH670/John Deere.	33
Figura 7 - Distribuição das amostras no talhão 10.	34
Figura 8 - Fluxograma Metodológico.....	35
Figura 9 - Monitor de colheita Harvest Monit.	36
Figura 10 – Receptor GNSS GR3 - TOPOCON coletando dados no limite do talhão.	37
Figura 11 - Software Mapfilter	41
Figura 12 – Raios de influência.	42
Figura 13 - Processo de generalização dos dados de produtividade.....	43
Figura 14 - Parâmetros do software Weka.....	44
Figura 15 - Pontos de produtividade filtrados.	47
Figura 16 – Raster dos Índices de Vegetação	48
Figura 17 - Correlações entre os índices de vegetação, parâmetros biométricos e produtividade.	49
Figura 18 - Frequência das variáveis nos modelos de predição.	53
Figura 19 - Boxplot dos dados reais e estimados por cada modelo.....	59
Figura 20 - Dados de produtividade generalizados.	60
Figura 21 - Dados de produtividade medida e estimado.	61
Figura 22 - Gráfico de dispersão: Cenário A.....	62
Figura 23 - Gráfico de dispersão: Cenário B.	63
Figura 24 – Distribuição do erro dos cenários de estimava.	65
Figura 25 - Frequência dos valores de erro dos cenários de produtividade A e B.	66
Figura 26 - Frequência acumulada dos valores de erro dos cenários de produtividade A e B.....	67

LISTA DE TABELAS

Tabela 1 - Bandas do sensor MSI - Sentinel 2.....	24
Tabela 2 - Índices de Vegetação.....	39
Tabela 3 - Cenários para criação dos modelos de predição.....	45
Tabela 4 - Modelos gerados para cada cenário.....	51
Tabela 5 - Variáveis com maior e menor influência na predição da produtividade.	54
Tabela 6 - Frequência das variáveis com maior e menor influência na predição da produtividade.....	55
Tabela 7 - Métricas estáticas dos dados de treinamento.....	56
Tabela 8 - Resultado do teste de normalidade dos resíduos - <i>Shapiro-Wilk</i>	57
Tabela 9 - Estatísticas descritivas	58
Tabela 10 - Estatísticas descritivas do dado real e os estimados.	62
Tabela 11 - Erros dos modelos de predição.....	64
Tabela 12 - Teste de Nemenyi	68

LISTA DE ABREVIATURAS E SIGLAS

ANOVA	Análise de Variância
AP	Agricultura de Precisão
BOA	<i>Bottom Of Atmosphere</i>
CCCI	<i>Canopy chlorophyll content index</i>
CONAB	Companhia Nacional de Abastecimento
DMC	Diâmetro médio do colmo
EMC	Estatura média dos colmos
ESALQ	Escola Superior de Agricultura "Luiz de Queiroz"
EVI	<i>Enhanced Vegetation Index</i>
GCI	<i>Green Chlorophyll Index</i>
GEE	<i>Google Earth Engine</i>
GMSE	<i>Global Monitoring for Environment and Security</i>
GNSS	<i>Global Navigation Satellite System</i>
GPP	Produtividade Primária Bruta
GRVI	<i>Green-Red Vegetation Index</i>
INPE	Instituto Nacional de Pesquisas Espaciais
IV	Índices de vegetação
MCARI1	<i>First Modified Chlorophyll Absorption Ratio Index</i>
MCARI2	<i>Second Modified Chlorophyll Absorption Ratio Index</i>
MSAVI	<i>Modified SAVI</i>
MSI	<i>Multispectral Instrument</i>
MSR	<i>Modified SR</i>
NDRE	<i>Normalized difference red edge index</i>
NDVI	<i>Normalized Difference Vegetation Index</i>
NPI	Número de perfislos industrializáveis
OSAVI	<i>Optimized SAVI</i>
r^2	Coefficiente de Determinação
RDVI	<i>Renormalized Difference Vegetation Index</i>
REM	Radiação eletromagnética
SAVI	<i>Soil Adjusted Difference Vegetation Index</i>

SIPI	<i>Structural Insensitive Pigment Index</i>
SR	<i>Simple Ratio</i>
SR	Sensoriamento Remoto
TCH	Toneladas de Cana por Hectare
TOA	<i>Top of Atmosphere</i>
UNICA	União da Indústria de Cana-de-Açúcar
VANT	Veículo aéreo não tripulado

SUMÁRIO

CDD 526	3
1. INTRODUÇÃO	17
2. OBJETIVOS	19
2.1 OBJETIVO GERAL	19
2.2 OBJETIVOS ESPECÍFICOS	19
3. REVISÃO DA LITERATURA	20
3.1 CANA-DE-AÇÚCAR E TECNOLOGIAS APLICADAS NO SETOR CANAVIEIRO	20
3.1.1 Características gerais da cana de açúcar e da variedade RB867515.....	20
3.2 SENSORIAMENTO REMOTO APLICADO NA AGRICULTURA	22
3.2.1 Sentinel 2	23
3.2.2 Comportamento da vegetação.....	24
3.3 AGRICULTURA DE PRECISÃO NA CULTURA DA CANA DE AÇÚCAR.	27
3.3.1 Monitor de colheita.....	28
3.4 MODELOS DE ESTIMATIVA DE PRODUTIVIDADE	30
4. METODOLOGIA.....	32
4.1 LOCALIZAÇÃO E CARACTERIZAÇÃO DA ÁREA DE ESTUDO	32
4.2 MATERIAIS.....	33
4.3 PROCEDIMENTOS METODOLÓGICOS	35
4.3.1 Levantamento de Campo	36
4.3.1 Aquisição de Dados	37
4.3.2 Processamento Digital dos Dados	38
4.3.3 Processamento dos modelos de regressão	43
4.3.5 Análise da qualidade dos modelos de regressão.....	46

5. RESULTADOS E DISCUSSÕES	47
5.1 RESULTADOS DOS PROCESSAMENTOS DOS DADOS	47
5.2 CORRELAÇÃO DOS PARÂMETROS PARA PREDIÇÃO DA PRODUTIVIDADE	49
5.3 MODELOS DE PREDIÇÃO.....	50
5.4 ANÁLISE DA QUALIDADE DOS MODELOS POR VALIDAÇÃO CRUZADA.....	55
5.4 ANÁLISE DA QUALIDADE DOS MODELOS BASEADOS EXCLUSIVAMENTE EM DADOS DE SENSORIAMENTO REMOTO PELOS DADOS DE TESTE.....	59
6. CONSIDERAÇÕES FINAIS	68
REFERÊNCIAS	69
7. APÊNDICE.....	74

1. INTRODUÇÃO

O Brasil é o maior produtor mundial de cana-de-açúcar, sendo responsável por exportar, em Julho de 2020 um volume de 3,487 milhões de toneladas (CONAB, 2020). A safra referente ao período de 2020/2021, representa uma área estimada a ser colhida de 8.409,8 mil hectares, com uma projeção de produtividade média de 76.348 Mg.ha⁻¹ (CONAB, 2020). Os principais estados produtores de cana de açúcar no Brasil são o estado de São Paulo na região Sudeste, considerado o maior produtor nacional; o estado de Goiás na região Centro-Oeste; na região Sul, o estado do Paraná; na região Nordeste, o estado de Alagoas; e na região Norte, o estado do Tocantins. Essa cultura movimenta anualmente no Brasil um valor bruto de mais de US\$ 100 bilhões, representando cerca de 2% do PIB, ou aproximadamente US\$ 40 bilhões (UNICA, 2019).

Segundo Almeida (20198), a produção de Cana-de-açúcar tem como principal custo, as operações agrícolas mecanizadas, como o preparo do solo (Aração, Terraceamento, Gradagem, Subsolagem, Manutenção de estradas), o plantio e a colheita, que representam um custo médio de 4.884,83 reais por hectare. Ao acrescentar outras operações agrícolas, insumos e a própria gestão, o investimento na produção de cana-de-açúcar pode chegar a um custo total de 7.324,88 reais por hectare (ALMEIDIA, 2019). Dessa forma, devido as extensas áreas de plantio e o elevado valor agregado na produção da cana-de-açúcar ao longo de toda safra, se faz necessário investigar em toda lavoura os resultados do manejo aplicado e identificar os problemas presentes na cultura (ROCHA, 2018).

Uma das soluções para o levantamento de informações espaciais dos cultivos nas áreas de cana-de-açúcar é a utilização das técnicas de Sensoriamento Remoto (SR). De acordo com Jensen (2009), o SR é considerado como a ciência que trata da aquisição, do processamento e da interpretação de imagens, podendo ser adquiridas em nível de terreno (com espectrorradiômetro portátil), de aeronave e orbital. No SR, as imagens são obtidas pelo registro da radiação eletromagnética em interação com a matéria. Um dos produtos mais conhecidos e utilizados do SR são o Índices de Vegetação (IV), que são modelos matemáticos que tem como objetivo ressaltar o comportamento espectral da vegetação em relação a outros alvos da superfície terrestre.

Os dados obtidos por Sensoriamento Remoto são de suma importância, principalmente para Agricultura de Precisão (AP), que tem por premissa compreender a variabilidade espacial dentro de uma lavoura de modo a otimizar o manejo da cultura e o gerenciamento de insumos

agrícolas, com eficiência e economia (BRASIL, 2013). O uso mais comum desses dados dentro da AP é para o monitoramento e detecção do comportamento espectral da cultura durante o período de safra, principalmente com uso dos modelos matemáticos de IV. Um exemplo dessas técnicas é o *Normalized Difference Vegetation Index* (NDVI), que possui alta correlação linear com desenvolvimento das principais culturas agrícolas (BERNARDI *et al.*, 2014). Com a detecção das alterações do comportamento da vegetação dentro de uma região, pode-se identificar os locais específicos e as possíveis causas na redução do vigor da planta, como pragas, doenças, déficit nutricional e hídrico.

Uma outra aplicação para as técnicas de Índices de Vegetação é a quantificação da produtividade das culturas agrícolas em geral. De acordo com Rocha (2018), o uso de IV de forma isolada não permite modelar precisamente informações como a produtividade. Para tanto faz-se necessário correlacionar esses índices com outros parâmetros da cultura. Nos dias atuais, diversos trabalhos têm sido desenvolvidos com o objetivo de correlacionar as informações biofísicas da cultura com a produtividade utilizando imagens de SR. Pode-se citar como exemplos, Oliveira (2015), que relacionou NDVI obtidos por imagens do sensor OLI do Landsat 8 para estimar a biomassa de cana-de-açúcar, onde se obteve um modelo com maior significância estatística aos 8 meses de idade após a colheita. Amaral *et al.* (2012), que utilizaram o sensor de reflectância de dossel para estimar a produtividade na cana-de-açúcar, mostrando que essa pode ser uma ferramenta útil para estimar esses dados, além de apresentar um modelo generalizado baseado no NDVI para estimar a produtividade.

Neste contexto, com uso dos dados biométricos da cana-de-açúcar adquiridos em campo e com os dados adquiridos a partir de imagens de Sensoriamento remoto em nível orbital, espera-se obter modelos de regressão para predição da produtividade da cana de açúcar.

2. OBJETIVOS

2.1 OBJETIVO GERAL

Predizer a produtividade da cultura da cana-de-açúcar a partir de dados de sensoriamento remoto e dados biométricos obtidos em campo.

2.2 OBJETIVOS ESPECÍFICOS

- Identificar os parâmetros biométricos da cana-de-açúcar com maior correlação com os dados de Sensoriamento Remoto.
- Gerar um modelo de regressão linear múltipla que permita prever a produtividade da cana de açúcar.
- Prever a produtividade da cana de açúcar com 5 meses de antecedência utilizando biometrias e IV.
- Aplicar um modelo de forma espacializada.
- Compreender e analisar a variabilidade espacial da produtividade da Cana-de-açúcar utilizando dados de Sensoriamento Remoto e os modelos de regressão obtidos.

3. REVISÃO DA LITERATURA

3.1 CANA-DE-AÇÚCAR E TECNOLOGIAS APLICADAS NO SETOR CANAVIEIRO

A cana de açúcar pertence ao gênero *Saccharum L.* de origem indiana, e chegou no Brasil no Século XVI, trazida pelos portugueses devido a fatores que favoreceram seu desenvolvimento como os solos férteis, disponibilidade de recursos hídricos e temperaturas adequadas (RODRIGUES e ROSS, 2020). Nos dias atuais, devido ao aumento da necessidade de combustíveis ecológicos, crescentes demandas de exportação, grandes áreas cultiváveis, as condições climáticas e o crescimento em investimento e tecnologia no setor agroindústria sucroalcooleira tornaram o Brasil o maior produtor de cana de açúcar do mundo (CONAB, 2020).

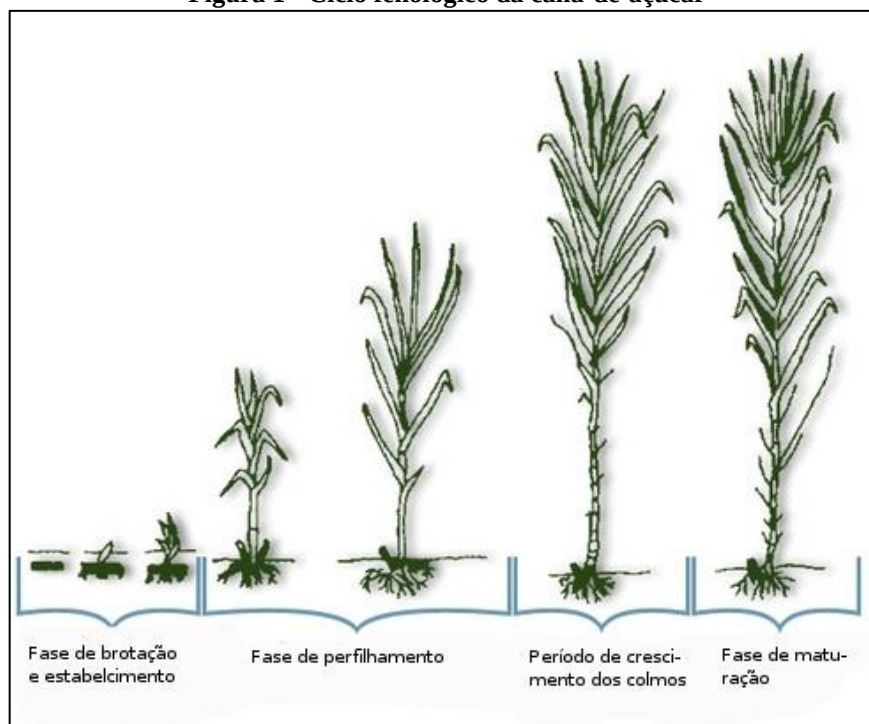
Para o alcance de maior produtor de cana de açúcar do mundo, como apresenta Abarca (1999), o Brasil precisou iniciar um processo de desenvolvimento e inovação de tecnologia que trouxe de outros setores da indústria e os adaptou para produção de cana de açúcar. Segundo esse autor, as principais inovações tecnológicas nesse setor são a automação de processos, melhoramento genético, monitoramento por satélite, fermentação contínua e novos produtos. Mesmo apresentando anos de inovação tecnológica, de acordo com Alves e Paixão (2017), são necessárias mais inovações e alternativas neste setor, principalmente para o aumento de produtividade, que vem apresentando um esgotamento nas melhorias tecnológicas resultando em uma estagnação no crescimento da produtividade. Porém, mesmo com os grandes desafios e limitações do setor, de acordo com Lankriet e Poppe (2018), através de projetos como RenovaBio, que incentiva o aumento do uso de fontes de energias ecológicas, estima-se um consumo em 2030 de até 54 bilhões de litros de etanol, atualmente chegando a valores de 27 bilhões de litros.

3.1.1 Características gerais da cana de açúcar e da variedade RB867515

A cana de açúcar é uma cultura que pertence à família *Poaceae* e ao gênero *Saccharum*, uma cultura perene e do tipo C4, seu processo metabólico de incorporação de CO₂ é de máxima eficiência fotossintética quando há uma maior disponibilidade de radiação solar, umidade no solo e temperatura adequada (THOMAS, 2016). Segundo esse autor, de forma geral

a cana de açúcar é formada por colmos, folhas e inflorescência compondo a parte aérea da planta, enquanto as raízes compõem a parte subterrânea. A cultura da cana de açúcar apresenta quatro estágios, são eles: do plantio à brotação da gemas, da brotação ao final do perfilhamento, do perfilhamento ao início da acumulação de açúcar, e finalmente a maturação (KRUTMAN, 1966).

Figura 1 - Ciclo fenológico da cana-de-açúcar



Fonte: (GASCHO & SHIH, 1983).

O desenvolvimento da cana de açúcar inicia-se com brotação da gema, 20 a 30 dias após o plantio, esta fase depende, da qualidade da muda, variedade, época, manejo do plantio, temperatura e umidade do solo, profundidade de plantio, presença da bainha da folha protegendo a gema do contato com o solo e posição da gema ao longo do colmo. Com o desenvolvimento dos colmos principais inicia-se o crescimento das raízes e dos rizomas (GASCHO *et al.*, 1983; THOMAS, 2016). Após a emergência dos colmos, inicia-se o perfilhamento, no qual formam-se os perfilhos que são uma ramificação originária da base planta ou região axilar da folha basal (SILVA *et al.*, 2007). Esses perfilhos possuem sistema radicular, colmos e folhas próprias, sendo independentes dos demais. O ápice do perfilhamento acontece cerca de 2 a 4 meses após a emergência, tendo como fatores influenciadores a variedade, temperatura, radiação solar, fertilidade do solo, disponibilidade hídrica, densidade

de plantio e competição. Por meio do perfilhamento é formado a touceira (GASCHO *et al.*, 1983; THOMAS, 2016).

Em seguida ao processo de perfilhamento, é iniciado o crescimento da parte aérea, onde ocorre o aumento da massa seca, formação e o alongamento do colmo, início do acúmulo de sacarose, sendo a fase mais importante do cultivo (THOMAS, 2016; JADOSKI *et al.*, 2010). Esse processo depende de fatores, como o perfilhamento, da época de plantio, a variedade, temperatura, radiação solar, fertilidade do solo, disponibilidade hídrica, densidade de plantio e competição. A maturação é o último estágio de desenvolvimento da cana de açúcar, segundo Maldaner (2017), esse processo é o momento de acumulação máxima de sacarose nos colmos da planta.

Em relação às variedades de cana de açúcar, o mercado brasileiro possui uma extensa gama, entre elas a variedade RB867515, bastante utilizada no Paraná. Esta variedade é resultante de um policruzamento, utilizando como variedade progenitora a RB72454. A variedade RB867515 apresenta hábito de crescimento ereto e fácil despalha. O perfilhamento é médio com colmos de diâmetro médio e alta uniformidade. Os colmos possuem entrenós cilíndricos. O anel de crescimento tem largura média. A zona radicular é de largura média, sem enraizamento aéreo. As gemas são de tamanho médio do tipo pentagonal, pouco proeminente ultrapassando o anel de crescimento e com poro germinativo apical. As folhas são de largura média, arqueadas, curvas e bordos com serrilhamento pouco agressivo (DAROS *et al.*, 2010).

3.2 SENSORIAMENTO REMOTO APLICADO NA AGRICULTURA

O SR se trata da ciência de aquisição, processamento e interpretação de imagens, adquirida por sensores que registram uma parte do processo de interação entre a matéria e a radiação eletromagnética (JENSEN, 2019). Dentro da agricultura, o SR é de grande auxílio, pois permite o monitoramento e mapeamento em grandes extensões de produção agrícola, com uma gama de sensores em satélites ou embarcados em máquinas ou em *drones* permitindo maiores características temporais, espaciais e espectrais gerando um imenso conjunto de dados (GREGO *et al.*, (2014).

Existem várias aplicações do SR na agricultura, como estimativa de biomassa, produtividade, estresse hídrico, vigor e estágio fenológico (BRANDÃO, 2019). Para estimativa desses processos de interesse, com objetivo de ressaltar o comportamento espectral da vegetação em relação a outros alvos da superfície terrestre, foram desenvolvidos índices de

vegetação que utilizam de modelos matemáticos entre comprimentos específicos de onda correlacionados com os parâmetros que busca se conhecer da planta (BERNARDI *et al.*, 2014). De acordo com esses autores, o NDVI (*Normalized Difference Vegetation Index*) é o principal IV utilizado para o monitoramento agrícola devido à alta correlação linear com desenvolvimento das culturas. Como apresentando por Abarca (1999) já na década de 1990, uma das principais inovações tecnológicas neste setor da agroindústria sucroalcooleira é o monitoramento da lavoura por imagens orbitais. Através de programas como Canasat, executado pelo Instituto Nacional de Pesquisas Espaciais (INPE), o Centro de Tecnologia Canavieira (CTC), a Escola Superior de Agricultura "Luiz de Queiroz" (ESALQ) e a União da Indústria de Cana-de-Açúcar (UNICA), foi realizado o monitoramento das lavouras de cana de açúcar da região Centro-Sul (Proença, *et al.*, 2009).

Atualmente, o SR com aeronaves, vem ganhando maior utilização nas áreas canavieiras através do uso de veículo aéreo não tripulado (VANT), como apresenta Assaiante e Cavichioli (2020), essa tecnologia vem permitindo uma melhor detecção e controle de pragas, avaliação de falhas de plantio, modelagem de erosão, identificação de déficit hídrico, entre outras aplicações. Contudo, mesmo com os altos benefícios é necessário um investimento em equipamentos, licenças, software de manipulação e processamento e equipe de operações treinada.

As imagens orbitais tem como vantagens a sua disponibilização em muitos casos de forma gratuita, o fato de serem facilmente manipuladas em ambiente SIG e muitas são disponibilizadas com as correções atmosféricas em plataformas de processamento em nuvem como o *Google Earth Engine* (GEE).

3.2.1 Sentinel 2

A missão Sentinel 2 disponibilizou suas imagens a partir de Julho de 2015, criada pelo programa *Global Monitoring for Environment and Security* (GMES) pela Comunidade Europeia e a *European Space Agency* (ESA). A missão é formada por uma constelação de dois satélites idênticos em uma mesma órbita com diferença de 180° entre ambos, Sentinel 2A e 2B, ambos embarcados com o sensor *Multispectral Instrument* (MSI). A órbita é semipolar, heliosíncrona, descendente, 98.5623° de inclinação, altitude de 786 Km e o horário de imageamento é por volta das 10h 30 min. A missão fornece para o usuário dois tipos de produto,

sendo eles o Level-1C, onde as imagens são disponibilizadas ortoretificadas e em valores de reflectância de topo de atmosfera ou *Top of Atmosphere* (TOA) e o produto Level-2A com imagens ortoretificadas e em valores de reflectância de superfície ou *Bottom Of Atmosphere* (BOA) (ESA, 2015). A resolução temporal da constelação Sentinel 2 é de 5 dias no equador e a resolução radiométrica do sensor MSI é de 12 bits, permitindo adquirir imagens com valores de nível de cinza de 0 a 4095. A resolução espectral e espacial são de 4 bandas de 10 metros, 6 bandas de 20 metros e 3 bandas de 60 metros, conforme apresentado na Tabela

Tabela 1 - Bandas do sensor MSI - Sentinel 2.

Número da banda	Banda	Resolução espacial (m)	S2A		S2B	
			Comprimento de onda central (nm)	Largura de banda (nm)	Comprimento de onda central (nm)	Largura de banda (nm)
2	Blue (Azul)	10	492,4	66	492,1	66
3	Green (Verde)		559,8	36	559,0	36
4	Red (Vermelho)		664,6	31	664,9	31
8	NIR (infravermelho Próximo)		832,8	106	832,9	106
5	Red Edge 1	20	704,1	15	703,8	16
6	Red Edge 2		740,5	15	739,1	15
7	Red Edge 3		782,8	20	779,7	20
8a	Red Edge 4		864,7	21	864	22
11	SWIR 1		1613,7	91	1610,4	94
12	SWIR 2	60	2202,4	175	2185,7	185
1	Aerossol		442,7	21	442,2	21
9	Water Vapor		945,1	20	943,2	21
10	Cirrus		1373,5	31	1376,9	30

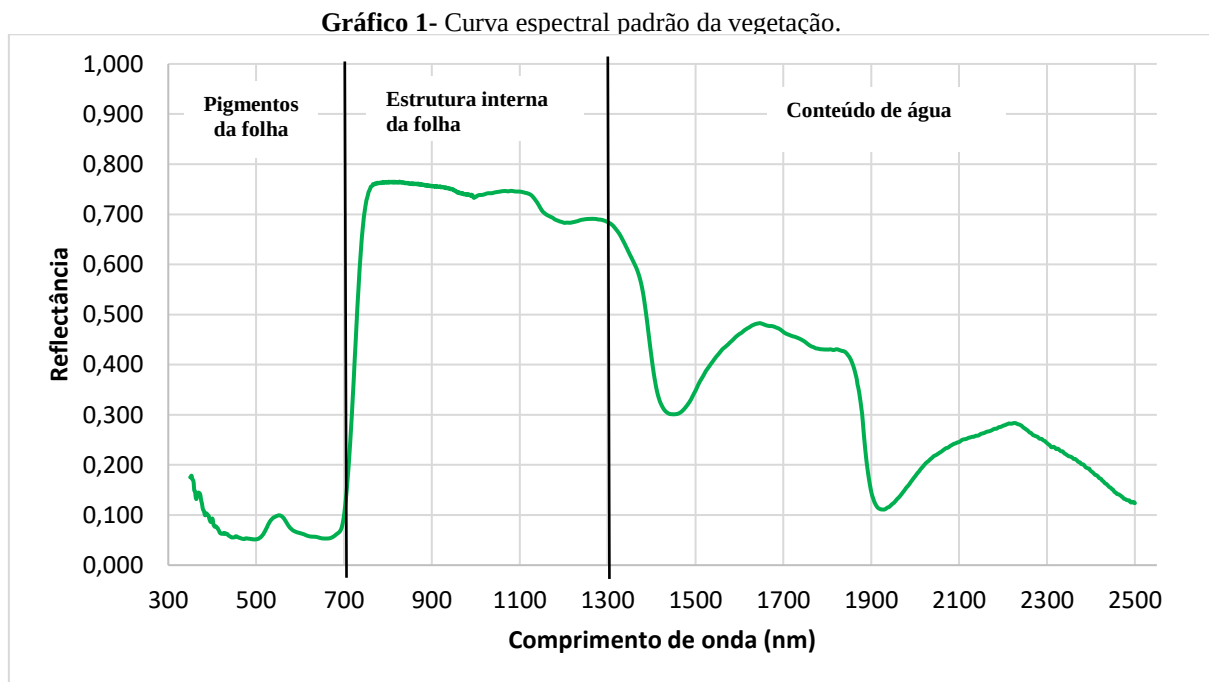
Fonte: ESA (2015)

3.2.2 Comportamento da vegetação.

O sensoriamento remoto da superfície terrestre tem como objetivo extrair informações de diferentes recursos naturais, como água, solo, rocha e a vegetação. Dentro dos alvos de vegetação, o SR analisa desde florestas naturais e plantadas, biomas, culturas agrícolas, entre

outros. Mesmo que cada uma dessas vegetações possuam diferentes espécies de plantas, com diferentes condições ambientais, fechamento de dosséis diferentes, a vegetação possui um comportamento espectral padrão, dentro do intervalo de 400 a 2500 nm (ROSA, 2009).

Considerando somente a análise de uma folha verde sadia, dentro do espectro do visível, existem duas regiões de absorção devido aos pigmentos fotossintetizantes como as carotenos (6%), clorofila (65%) e xantofilas (29%) que absorvem a região do espectro aproximadamente nos 480 nm e 680 nm, sendo essas regiões do espectro eletromagnético correspondentes as bandas do azul e do vermelho, portanto, apresentando um pico de reflectância na banda do verde. Na região do infravermelho próximo, correspondendo na faixa dos 700 a 1300 nm, é uma região de maior espalhamento da REM devido a estrutura interna celular da folha, mais especificamente devido a estrutura do mesófilo. Na faixa correspondente entre 1100 a 3200 nm, região do infravermelho médio, a absorção ocorre pela presença de água, pontualmente nos comprimento de 1100 nm, 1450 nm, 1950 nm e 2700 nm (PONZONI *et al.*, 2012; ROSA, 2009). O Gráfico 1 apresenta uma curva de folha verde sadia media através de um Espectrorradiômetro portátil.



Fonte: O autor (2021).

No sensoriamento remoto orbital da vegetação é necessário levar em consideração outros fatores além do próprio comportamento de uma única folha, que é considerado a função

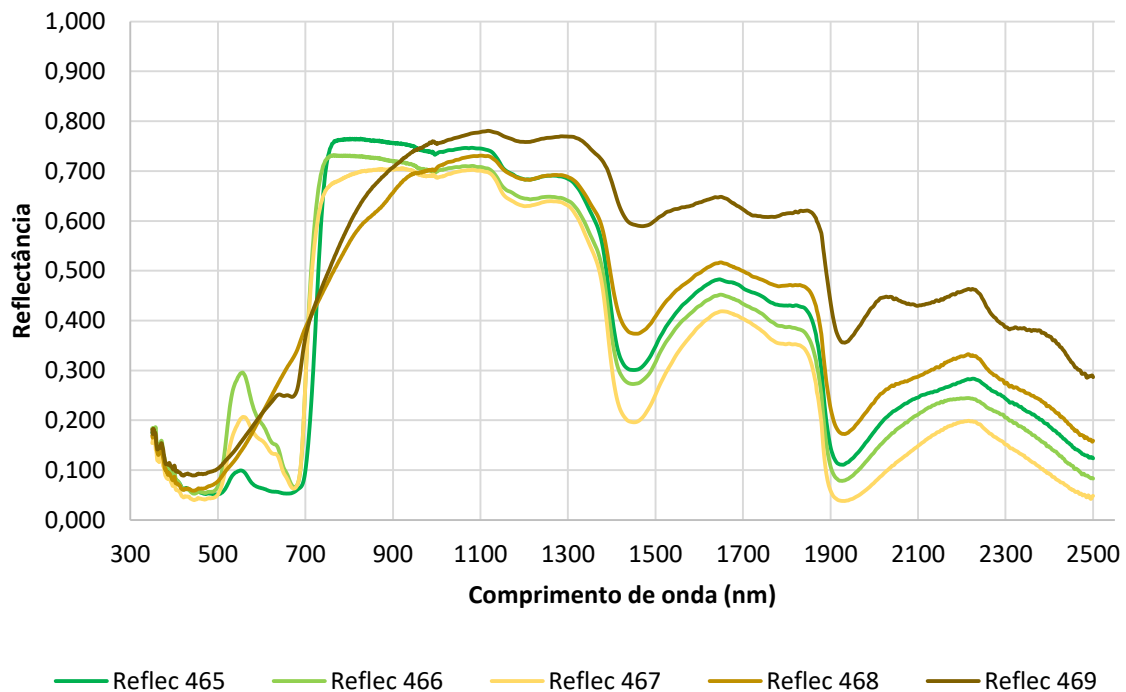
de reflectância bi-direcional hemisférica, pois em condições naturais, onde será imageado os dosséis da vegetação, a medida da reflectância espectral depende de fatores como a geometria de imageamento devido a posição do sensor e da iluminação, no qual ocorre o fator de reflectância bidirecional. Além de fatores como a distribuição espacial da vegetação, densidade, orientação, define a arquitetura do dossel (PONZONI *et al.*, 2012). Diante do exposto, partir desses dados levantados pelos sensores, é possível é extrair informações como estado fenológico, senescência, contaminação por patógenos que afetam a estrutura interna das folhas e necroses, pois esses fatores modificam o comportamento espectral da vegetação. A Figura 1 apresenta as folhas de uma mesma planta em diferentes níveis de senescência, enquanto o Gráfico 2 apresenta a curva espectral de cada uma delas.

Figura 2 - Folhas em diferentes níveis de senescência.



Fonte: O autor (2021).

Gráfico 2 - Curva espectral da folhas em diferentes níveis de senescência.



Fonte: O autor (2021).

3.3 AGRICULTURA DE PRECISÃO NA CULTURA DA CANA DE AÇÚCAR.

De acordo com Brasil (2013), a Agricultura de Precisão (AP) tem como premissa compreender a variabilidade espacial dentro de uma lavoura, através de técnicas, metodologias e tecnologias, com o objetivo de otimizar o manejo da cultura e o gerenciamento de insumos agrícolas, com eficiência e economia. Uma das informações primordiais para uso da AP são os dados espacializados de produtividade da cultura, que fornecem grande subsídio para compreensão da variabilidade espacial da cultura (MOLIN, 2000). Para obtenção desses dados de produtividade utiliza-se um sistema com monitor, equipado com receptor GNSS (*Global Navigation Satellite System*) e sensores, produzidos por diversos fabricantes e principalmente para cultura de grãos (GREGO *et al.*, 2014). Contudo, como apresenta Grego *et al.* (2014), existem poucos modelos de monitores de produtividade para a cana-de-açúcar, sendo um recurso valioso para iniciar os processos de AP na cultura.

Outra forma de obtenção de informações a respeito da variabilidade espacial é o mapeamento dos fatores de que influenciam a produtividade das culturas, como a coleta de amostras de solo georreferenciadas, que possibilitam a geração mapas de fertilidade e permitem

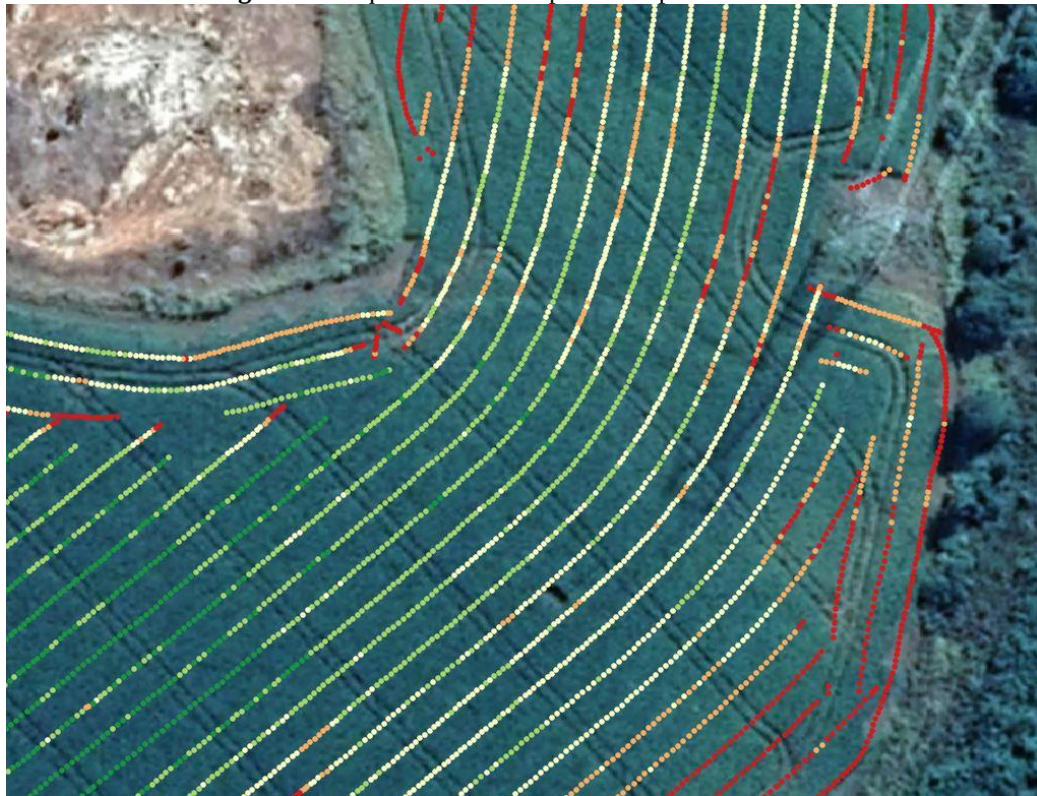
compreender a variabilidade espacial dos elementos do solo dentro do talhão e com isso é possível entre outras ações, a aplicação de insumos em taxa variável, atendendo a demanda de cada região da lavoura (RESENDE e COELHO, 2017). Contudo, a amostragem de solo pode apresentar um alto custo, chegando a valores de até R\$ 33,00 para cada amostra coletada (Gonçalves, 2017), mais gastos laboratoriais, equipamentos, processamentos dos mapas, entre outros custos, que podem desmotivar o produtor a adquirir essas informações.

Como uma solução para o custo de amostragem, pode ser a utilização do SR, que se tornou popular nas áreas agrícolas ao mapear os diferentes comportamentos da cultura dentro do talhão. Com a análise do comportamento espectral utilizando mapas de composições coloridas ou mapas de IV, de acordo Molin, Amaral e Colaço (2015), essas informações podem servir como subsídio para criação de zonas de manejo, aplicação de fertilizantes em taxa variável, preparo do solo de forma localizada, monitorar o vigor da cultura, modelar e estimar a produtividade, entre outras ações. Além de monitoramento da cultura da cana de açúcar, como apresenta Grego *et al.* (2014), o mapeamento em AP não pode ser somente em nível de propriedade, necessitando de um mapeamento multiescala, auxiliando planejamentos logísticos de colheita, escoamento de colheita e produção, previsão de safra, clima, entre outros. Segundo os autores, é imprescindível para melhoria no mapeamento para cana de açúcar, a pesquisa, o desenvolvimento e a inovação em geotecnologias.

3.3.1 Monitor de colheita

Com a necessidade de compreender a variabilidade espacial da produtividade, através da integração de um conjunto de sensores embarcado nas colheitadeiras, a partir da década de 80 iniciou-se o desenvolvimento de sistemas para monitorar a produtividade das culturas agrícolas (SANTI *et al.*, 2016). O dado de produtividade gerado a partir desses sistemas é uma malha de pontos contendo informações de parte da lavoura, um exemplo desses dados é apresentado na Figura 2. Com essas informações de produtividade é possível identificar as áreas com alto e baixo potencial produtivo, com isso compreender os fatores que contribuem para o desempenho de cada parte do talhão, afim de criar uma melhor gestão, atendendo todas as necessidades que existem dentro de cada talhão (MALDANER, 2017).

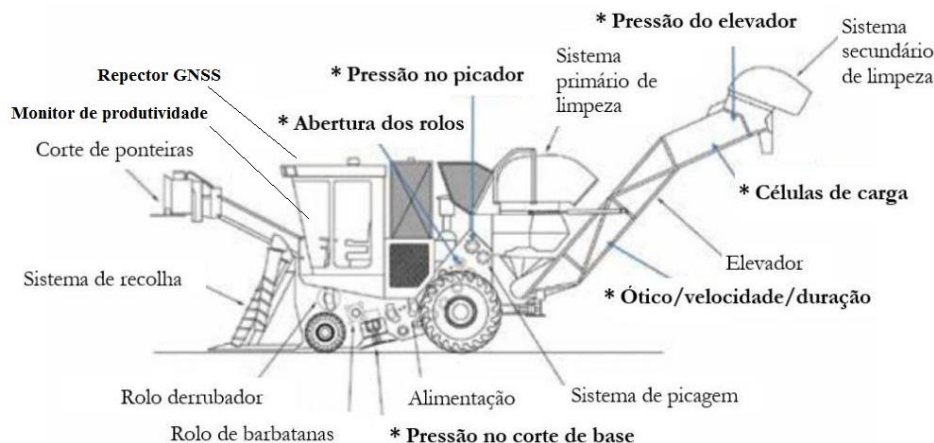
Figura 3 - Arquivo vetorial de pontos de produtividade.



Fonte: O autor (2021).

Existem vários métodos para obtenção dos dados necessários à mensuração da produtividade na cultura da cana de açúcar, e todos eles demandam de sensores que medem fluxo de massa de cana de açúcar que passa pela colhedora. Segundo Maldaner (2017), baseado em Cox (2002), os métodos usados para medir esse fluxo são através da medição direta da massa que passa pelo elevador por meio de células de carga; medição do volume de cana através da distância entre os rolos de alimentação do picador, existem outros métodos volumétricos de medição com o uso de sensores estereoscópicos de alta resolução com reconhecimento de padrões, que permite mensurar o volume; medição da potência hidráulica do picador usada para cortar a cana de açúcar em rebolos; e a medição da potência hidráulica usada para transportar a cana de açúcar no elevador. Além desses sensores, o sistema possui um receptor GNSS para georreferenciamento das informações, sensores de velocidade, entre outros sensores dependendo do fabricante. A Figura 3, apresenta o local de cada sensor para o mapeamento da produtividade em uma colhedora.

Figura 4 - Colhedora de cana de açúcar com a localização dos diferentes pontos de sensoriamento que podem ser usados para gerar dados de fluxo de cana.



Fonte: Adaptado de Maldaner (2017).

Contudo, mesmo com alta tecnologia em sensores embarcados nas colheitadeiras para garantir a geração desses mapas, existem erros sistemáticos inerentes no mapeamento da colheita. Segundo Molin (2000), os principais erros no processo de coleta de dados de produtividade georreferenciados são de suavização, largura incorreta da plataforma, ajuste do interruptor da plataforma, tempo de enchimento da colhedora, posicionamento incorreto, calibração, erro do receptor GNSS, entre outros. Segundo este autor, muitos desses erros sistemáticos podem ser eliminados em algoritmos de remoção de mapeamento de produtividade.

Segundo, Menegatti e Molin (2003), é necessário a classificação, identificação e caracterização dos possíveis erros presentes nos dados de colheita para realização do processo de filtragem, onde os pontos considerados errôneos possam ser removidos com objetivo de aumentar a qualidade dos mapas de produtividade. Para o pós-processamento dos dados de produtividade existe uma gama de software para esse procedimento como SGIS, SST Summit, SMS, InCeres, FalkerMap, Map Filter, CR Campeiro 7, Yield Editor, entre outros.

3.4 MODELOS DE ESTIMATIVA DE PRODUTIVIDADE

Segundo Oliveira *et al.* (2012), os modelos de estimativa de produtividade de cana de açúcar podem ser utilizados para o auxílio no planejamento agrícola das usinas com objetivo de definir as ações nas tomadas de decisão durante a safra, possibilitando uma economia de tempo, trabalho e recursos. Um dos principais processos dentro das usinas são as operações de

corte, carregamento e transporte (CCT), processo que garante o abastecimento de cana na moenda (MUNDIM, 2009). Segundo esse autor, o número de colhedoras, transbordos e caminhões necessários para cada frente de trabalho são determinados pelo responsável do transporte de forma empírica, geralmente baseado em conhecimentos práticos e subjetivos.

Os modelos de estimativa para cana de açúcar podem ser baseados em modelos agrometeorológicos, aprendizado de máquina, entre outros modelos. Moraes (2012) utilizou de modelos agrometeorológicos combinando com dados do sensor MODIS foi possível estimar a produtividade de forma promissora. Picoli (2006), avaliou o uso de redes neurais artificiais, dados de produtividade, características da cultura, do solo e modelo agrônomo-espectral com dados MODIS para estimativa de cana soca, onde se obteve resultados satisfatórios.

Nos modelos de regressão múltipla, Sugawara *et al.* (2007) combinou dados de estágios de corte, agrônomicas, variáveis espectrais e meteorológicas que apresentou resultados promissores com esse modelo de predição. Dias *et al.* (1999), utilizou de variáveis os fatores climáticos, características químicas, físicas e hídricas do horizonte subsuperficial em modelos de regressão múltipla.

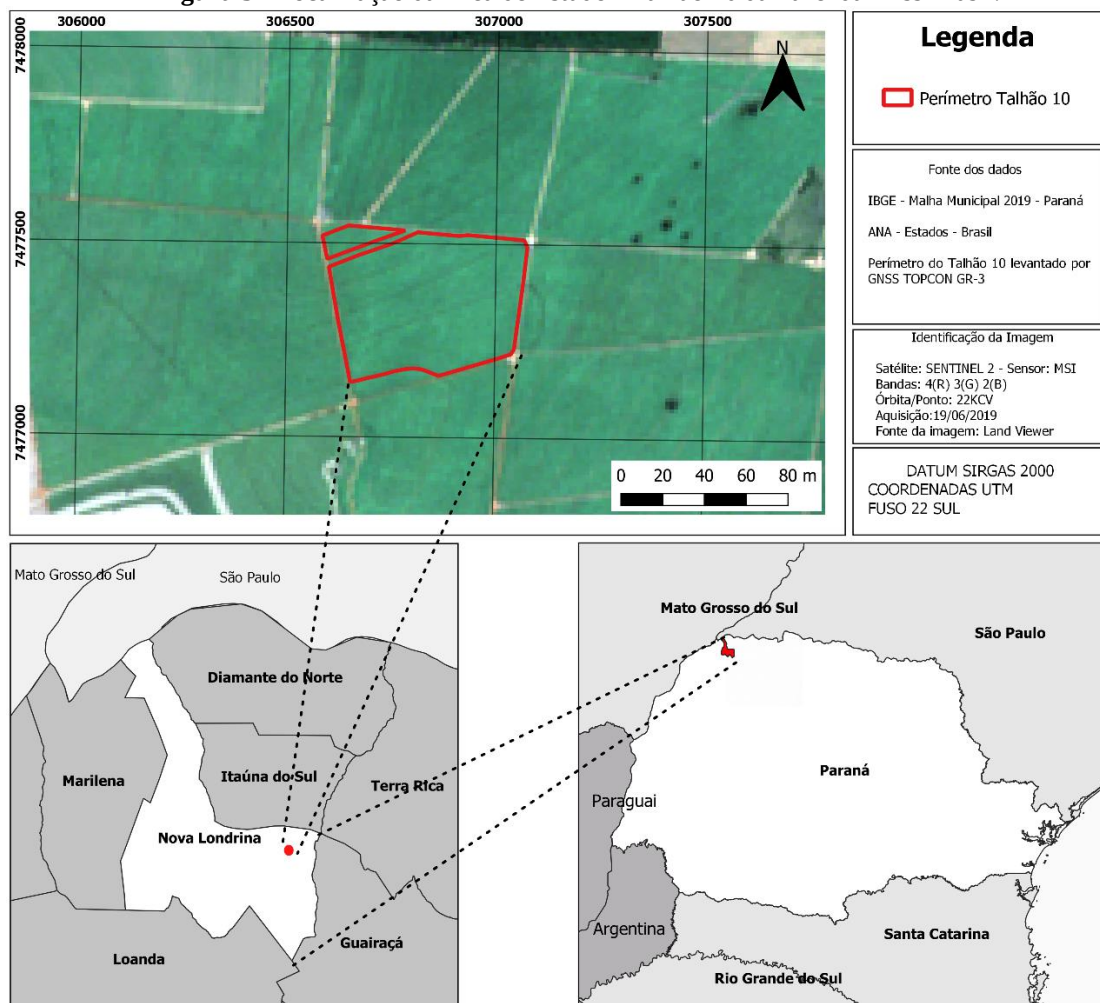
4. METODOLOGIA

Os procedimentos metodológicos para o desenvolvimento do presente trabalho serão descritos a seguir.

4.1 LOCALIZAÇÃO E CARACTERIZAÇÃO DA ÁREA DE ESTUDO

O presente trabalho foi realizado no Talhão 10, da Fazenda Miss Brasil, localizado no Município de Nova Londrina – PR, como mostra a Figura 4. O talhão possui uma área de 15,578 hectares e nele é cultivada a cana-de-açúcar implantada em espaçamento duplo-alternado de 90x120 cm, atualmente com cana soca de quarto corte da variedade RB86-7515 implantada em 21/11/2016.

Figura 5 - Localização da Área de Estudo - Talhão 10 da Fazenda Miss Brasil.



Fonte: O autor (2021).

4.2 MATERIAIS

Para a execução desse estudo foram utilizados os seguintes materiais:

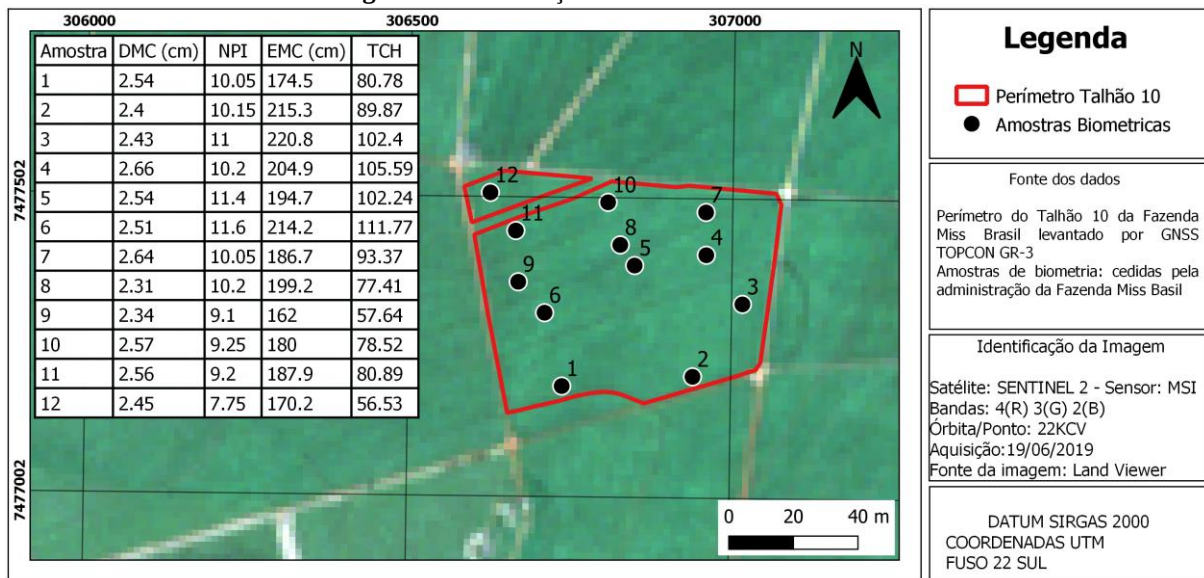
- Dados de colheita de cana de açúcar adquiridos pela colhedora CH670 com monitor de colheita *Harvest Monitor* e o monitor GS3 2630, equipamentos da marca John Deere (Figura 5).

Figura 6 - Colhedora CH670/John Deere.



Fonte: O autor (2021).

- Relatório contendo o levantamento de campo de 12 pontos amostrais com dados sobre os parâmetros biométricos da cana-de-açúcar referente ao dia 19/06/2020, elaborado pela administração da Fazenda Miss Brasil. No referido relatório constam os seguintes dados: o diâmetro médio do colmo (DMC), o número de perfilhos industrializáveis (NPI), estatura média dos colmos (EMC) e Tonelada de Cana por Hectares (TCH), como apresenta a Figura 6.

Figura 7 - Distribuição das amostras no talhão 10.

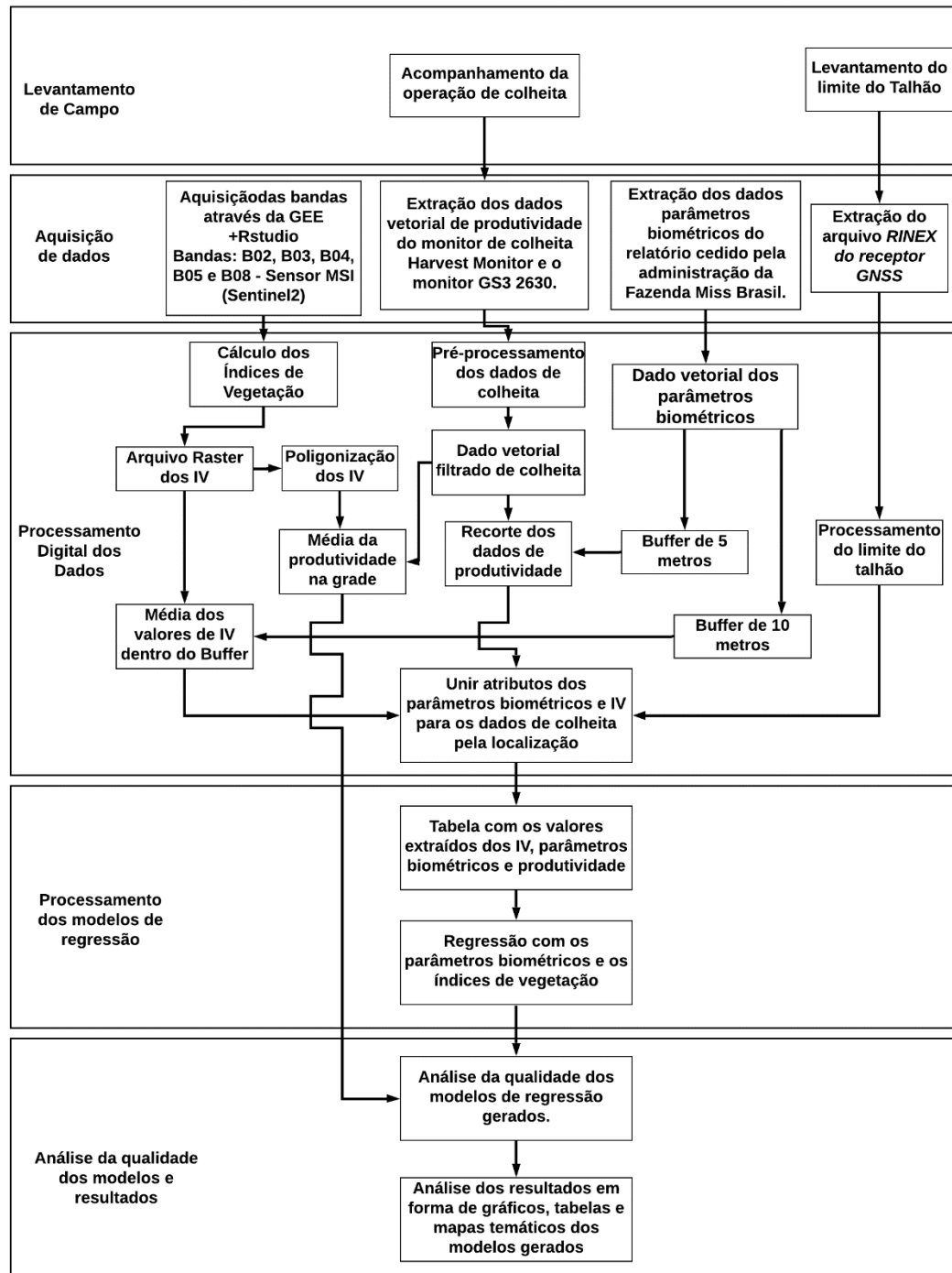
Fonte: O autor (2021).

- Imagens orbitais do sensor MSI do satélite Sentinel 2, bandas 490 nm (B2), 560 nm (B3), 665 nm (B4) e 842 nm (B8), com resolução espacial de 10 metros; e a banda 705 nm (B5), com resolução espacial de 20 metros, todas imagens em valores de reflectância de superfície. Todas as imagens foram obtidas na data de 19/06/2020, ou seja, a data do levantamento de campo que contém os parâmetros biométricos.
- Receptor GNSS marca TOPCON modelo GR-3 tem como especificações: dupla frequência L1 e L2; integrado de receptor, antena, rádio UHF, modem GSM/GPRS e bateria; possui 72 canais universais; rastreia sinais de satélites GPS, GLONASS e Galileo; taxa de atualização de 5 Hz; precisão horizontal de 3mm + 0.5 ppm e vertical de 5mm + 0.5 ppm para levantamentos estáticos e rápido-estáticos e horizontal de 10mm + 1 ppm e vertical de 15mm + 1 ppm para levantamentos cinemáticos e RTK.

4.3 PROCEDIMENTOS METODOLÓGICOS

Os procedimentos metodológicos que foram realizados para a execução do trabalho são apresentados no fluxograma da Figura 7.

Figura 8 - Fluxograma Metodológico.



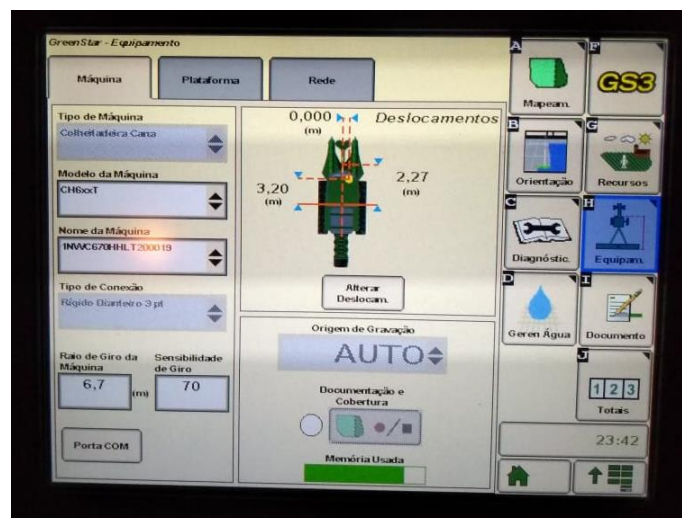
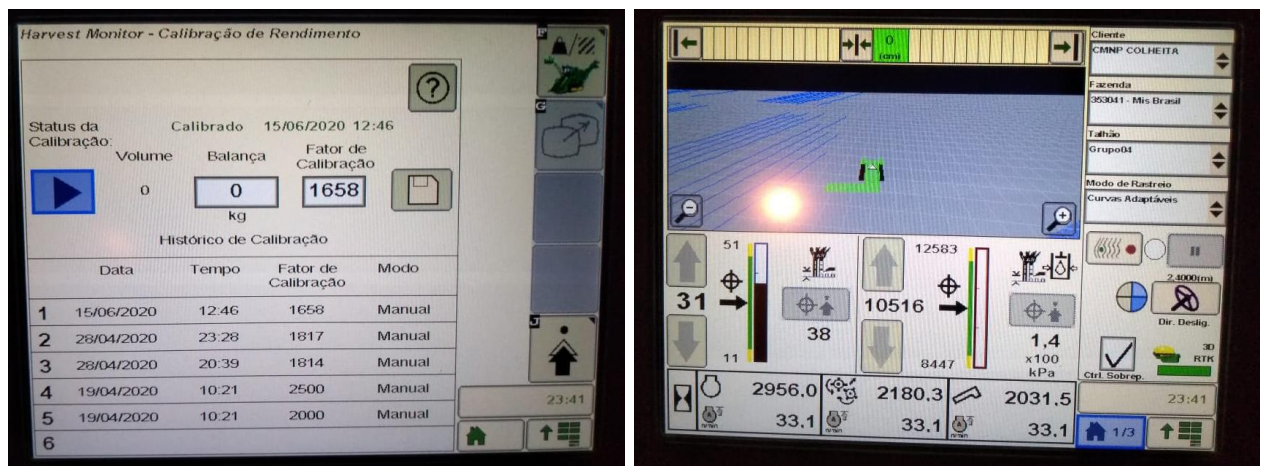
Fonte: O autor (2021).

4.3.1 Levantamento de Campo

A colheita da cana de açúcar de quarto corte do Talhão 10 da Fazenda Miss Brasil, iniciou-se no dia 10/11/2020 e foi finalizada no dia 13/11/2020. O equipamento utilizado foi a colhedora CH 670 John Deere, equipada com o monitor de colheita *Harvest Monitor* e o monitor GS3 2630.

Antes do início da operação de colheita verificou as calibrações do monitor de colheita, toda parte de documentação e informações do implemento da colhedora para garantir a qualidade dos dados adquiridos durante a operação de colheita do talhão, como mostra a Figura 8(a), 8(b) e 8(c), respectivamente.

Figura 9 - Monitor de colheita Harvest Monit.



Fonte: O autor (2021).

Durante os quatro dias de colheita foram realizadas verificações de forma aleatória dos dados através do monitor GS3 2630, afim de garantir que o registro dos valores de colheita estavam sendo realizados pelo equipamento. Durante essa operação foi levantado através do GPS Geodésico GR3 o limite do talhão, Figura 9, através do método relativo estático, no qual se ocupou por 10 minutos cada um dos vértices que delimitavam o perímetro do talhão, para fins de processamento dos dados de produtividade. Esse levantamento foi processado no *software Topcon Tools*, utilizando das observações da Rede Brasileira de Monitoramento Contínuo dos Sistemas GNSS (RBMC), das estações de PRMA – Maringá, PRUR – Umuarama e ROSA – Rosana.

Figura 10 – Receptor GNSS GR3 - TOPOCON coletando dados no limite do talhão.



Fonte: O autor (2021).

4.3.1 Aquisição de Dados

Inicialmente o relatório cedido pela administração da Fazenda Miss Brasil foi utilizado para obtenção dos dados dos parâmetros biométricos, sendo extraídos os seguintes parâmetros:

- Diâmetro médio do colmo (DMC);
- Número de perfilhos industrializáveis (NPI);
- Estatura média dos colmos (EMC); e
- Tonelada de Cana por Hectares (TCH)

Com valores das biometrias DMC em cm, NPI em quantidade, EMC em cm, TCH em ton.ha⁻¹. O procedimento de coleta de campo dos dados biométricos descrito no relatório consistiam em medidas coletadas manualmente utilizando de equipamento um paquímetro e uma trena. Nas doze amostras distribuídas dentro do talhão, o procedimento adotado de coleta em que cada amostra foi composta por uma fileira dupla de cana de açúcar com espaçamento de 90 cm e comprimento de 10 metros. Realizou-se a contagem do número de perfilhos industrializáveis para obtenção do dado de NPI e em seguida de forma aleatória foram selecionados dez colmos de cana de açúcar para obtenção dos valores de DMC e EMC. O parâmetro Tonelada de Canas por Hectare (TCH) é uma estimativa de volume de cana de açúcar pelos parâmetros biométricos, o modelo utilizado no relatório é descrita por Martins e Landell (1995).

$$TCH = D^2 \times C \times H \times \frac{0,07854}{E} \quad (1)$$

Onde:

D é o diâmetro de colmos (cm);

C é o número de colmos por metro linear;

H é o comprimento médio de colmos (cm); e,

E é o espaçamento entre sulcos (m).

De forma concomitante, foram obtidas as imagens das bandas B02 (BLUE), B03 (GREEN), B04 (RED), B05 (RED EDGE 1) e B08 (NIR) do Sensor MSI do satélite Sentinel 2, cuja resolução espacial das imagens são de 10 e 20 metros. As imagens foram obtidas com correção atmosférica em nuvem através da plataforma *Google Earth Engine* (GEE) integrada ao *software RStudio*.

4.3.2 Processamento Digital dos Dados

Após a obtenção das imagens foi realizado o cálculo dos 15 Índices de Vegetação (IV) mostrados na Tabela 2, utilizando o *software Qgis 3.10.3* aplicando a técnica de álgebra de mapas.

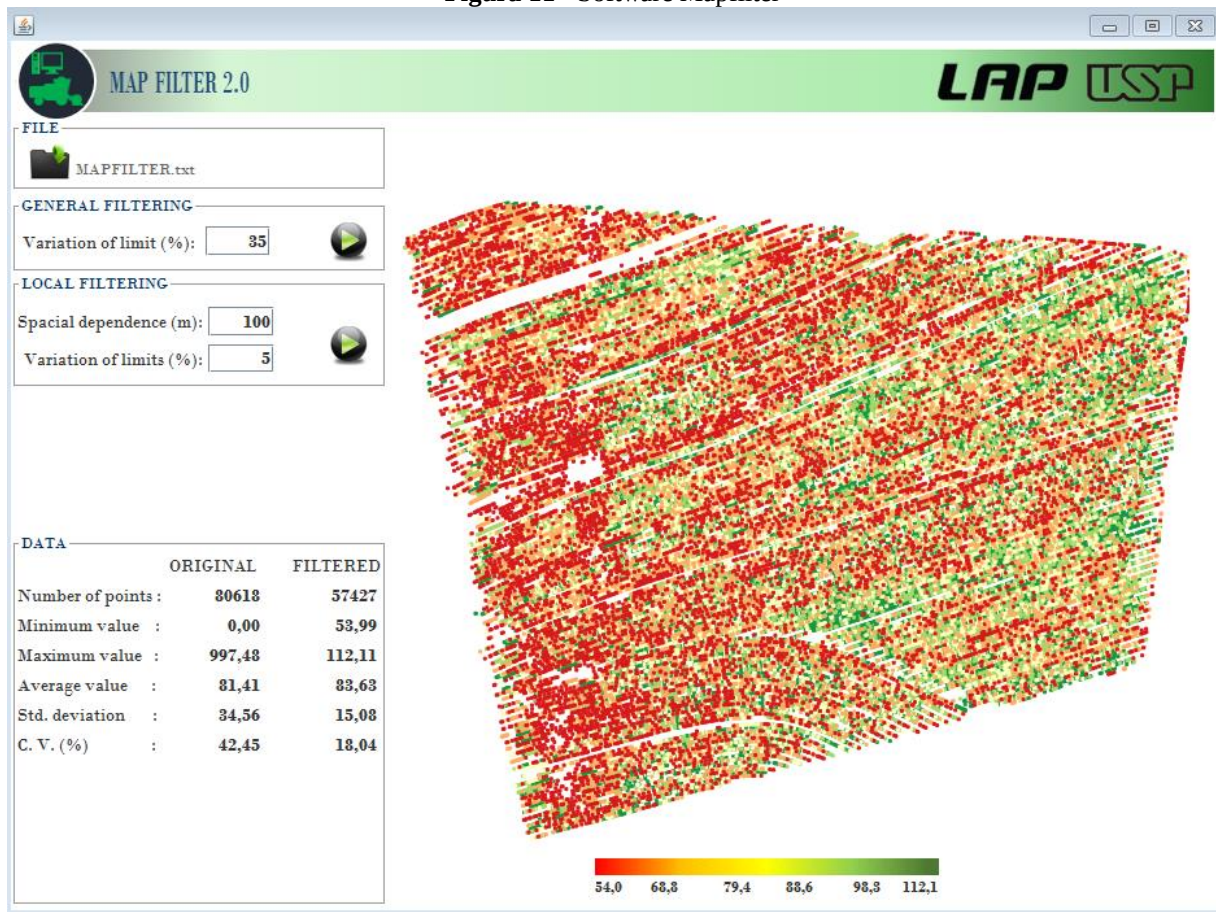
Tabela 2 - Índices de Vegetação.

Índice de vegetação	Modelo	Referência
Canopy chlorophyll content index (CCCI)	$\frac{NDRE}{NDVI}$	Long <i>et al.</i> (2009)
Enhanced Vegetation Index (EVI)	$2,5 \frac{NIR - RED}{(NIR + 6RED - 7,5BLUE) + 1}$	Huete <i>et al.</i> (1997)
Green Chlorophyll Index (GCI)	$\frac{NIR}{GREEN} - 1$	Gitelson <i>et al.</i> (2003)
Green-Red Vegetation Index (GRVI)	$\frac{GREEN - RED}{GREEN + RED}$	Tucker (1979)
First Modified Chlorophyll Absorption Ratio Index (MCARI1)	$1,2(2,5(NIR - GREEN) - 1,3(NIR - GREEN))$	Haboudane <i>et al.</i> (2004)
Second Modified Chlorophyll Absorption Ratio Index (MCARI2)	$1,5 \frac{2,5(NIR - RED) - 1,3(NIR - GREEN)}{\sqrt{(2NIR + 1)^2 - (6NIR - 5\sqrt{RED})} - 0,5}$	Haboudane <i>et al.</i> (2004)
Modified SAVI (MSAVI)	$\frac{2NIR + 1 - \sqrt{(2NIR + 1)^2 - 8(NIR - RED)}}{2}$	Qi <i>et al.</i> (1994)
Modified SR (MSR)	$\frac{\left(\frac{NIR}{RED}\right) - 1}{\sqrt{\left(\frac{NIR}{RED}\right) + 1}}$	
Normalized difference red-edge index (NDRE)	$\frac{NIR - rededge}{NIR + rededge}$	Barnes <i>et al.</i> (2000)
Normalized Difference Vegetation Index (NDVI)	$\frac{NIR - RED}{NIR + RED}$	Rouse <i>et al.</i> (1974)
Optimized SAVI (OSAVI)	$(1 + 0,16) \frac{NIR - RED}{NIR + RED + 0,16}$	Rondeaux <i>et al.</i> (1996)
Renormalized Difference Vegetation Index (RDVI)	$\frac{NIR - RED}{\sqrt{(NIR + RED)}}$	Roujean e Breon (1995)
Soil Adjusted Difference Vegetation Index (SAVI)	$(1 + L) \frac{NIR - RED}{NIR + RED + L}$	Huete (1988)

Structural Insensitive Pigment Index (SIPI)	$\frac{\text{NIR} - \text{BLUE}}{\text{NIR} + \text{RED}}$	Peñuelas <i>et al.</i> (1995)
Simple Ratio (SR)	$\frac{\text{NIR}}{\text{RED}}$	Birth e McVey (1968)
Em que: BLUE, valores de reflectância em 490 nm; GREEN, valores de reflectância em 560 nm; RED, valores de reflectância em 665 nm; NIR, valores de reflectância em 842 nm; rededge, valores de reflectância em 705 nm; L (SAVI), valor de 0,5.		
Fonte: O autor (2021).		

Para obtenção dos dados de produtividade do monitor, foi realizada extração das informações através do descarregamento do cartão do monitor via pendrive. Com esse cartão em formato John Deere, utilizando do programa *Apex*, os dados foram convertidos para o formato Shapefile para manipulação em ambiente SIG. A partir desses dados brutos de produtividade, foi aplicado um pré-processamento de filtragem para remoção de outliers utilizando do *software Mapfilter*, como mostra a Figura 10. Esse *software* inicialmente aplica uma filtragem global, que utiliza a mediana dos valores de produtividade em análise para calcular os limites de corte superior e inferior. O parâmetro *Variation of limits* utilizado foi de 35%. Em seguida é aplicado o filtro local, que é dividido em duas etapas: o filtro anisotrópico e o filtro isotrópico. Os parâmetros utilizados *Spacial dependence* de 100 metros e *Variation of limits* de 5,00%.

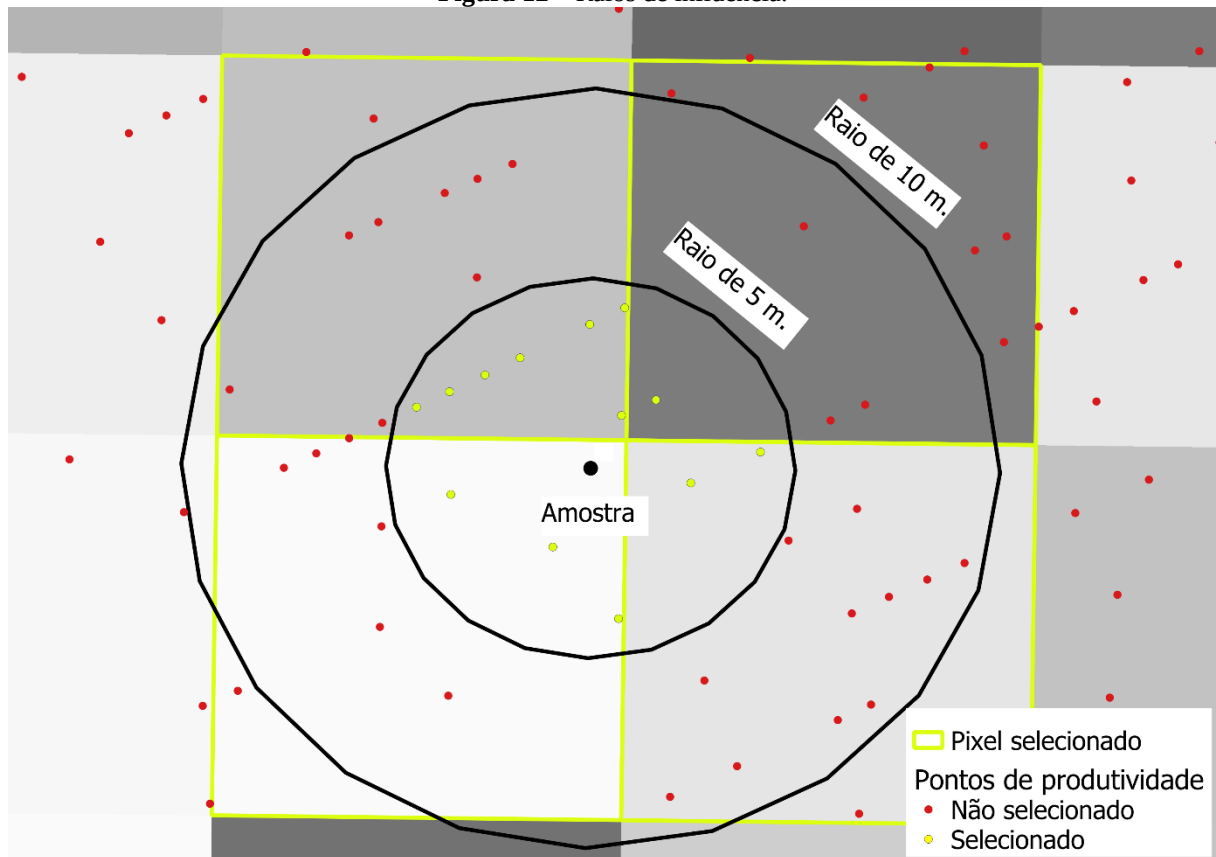
Figura 11 - Software Mapfilter



Fonte: O autor (2021).

Com os dados vetoriais dos parâmetros biométricos, foi criada uma zona de influência com um raio de 5 metros no entorno das amostras, na qual todos os dados de produtividade contidos foram associados ao respectivos valores de biometria. Ainda com a localização das amostras de biometria, foi gerada uma zona de influência com raio de 10 metros, e para cada dado de produtividade selecionado na etapa anterior foi associado o valor médio dos pixels de IV cujo centroide estava a menos de 10 metros da localização da amostra de biometria. Estes procedimentos foram adotados para compatibilização da resolução espacial das variáveis utilizadas neste trabalho.

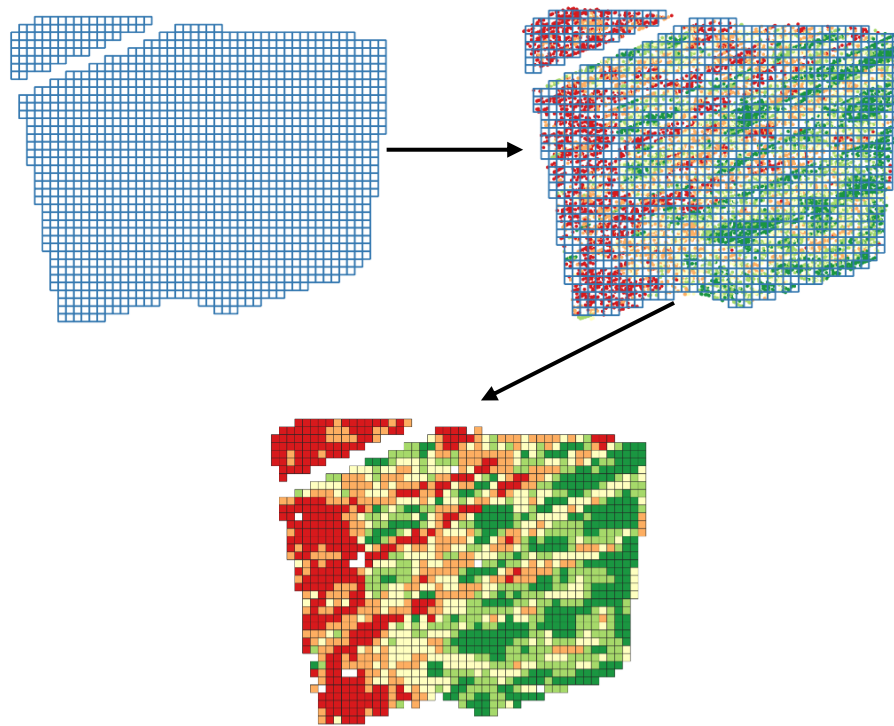
A Figura 11 mostra os raios utilizados para obtenção desses dados. Esses procedimentos foram realizados através do *software SAGA GIS* integrado no *software Qgis 3.10.3*. Após a realização desse procedimento o arquivo vetorial gerado com as informações extraídas foi exportado para o formato de tabela. Esta tabela por sua vez foi manipulada do *software WEKA* para geração dos modelos de regressão.

Figura 12 – Raios de influência.

Fonte: O autor (2021).

Para compatibilizar a resolução espacial dos dados de produtividade coletados pela colhedora com a resolução espacial dos dados de SR, e permitir avaliar a aplicação dos modelos foi realizada a generalização dos dados de produtividade. Inicialmente foi gerada uma grade com resolução de 10x10 metros, compatível com a resolução espacial das imagens MSI, e para cada célula foi calculado o valor médio dos pontos de produtividade ali contidos. A Figura 12 ilustra o processo de generalização aplicado aos dados de produtividade.

Figura 13 - Processo de generalização dos dados de produtividade.



Fonte: O autor (2021).

4.3.3 Processamento dos modelos de regressão

Para geração dos modelos para predição de produtividade foi utilizado o modelo de Regressão Linear Múltipla no *software Weka*. Este modelo utiliza como entrada mais de uma variável, considerando os valores biométricos e os índices de vegetação. O modelo estatístico dessas regressões, como apresentado por Cunha (2018) é:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \epsilon_i \quad (2)$$

Onde:

Y_i é o valor de Y no i-ésimo nível de X (variável independente);

β são os parâmetros desconhecidos que serão estimados;

X é o i-ésimo nível das variáveis independentes; e,

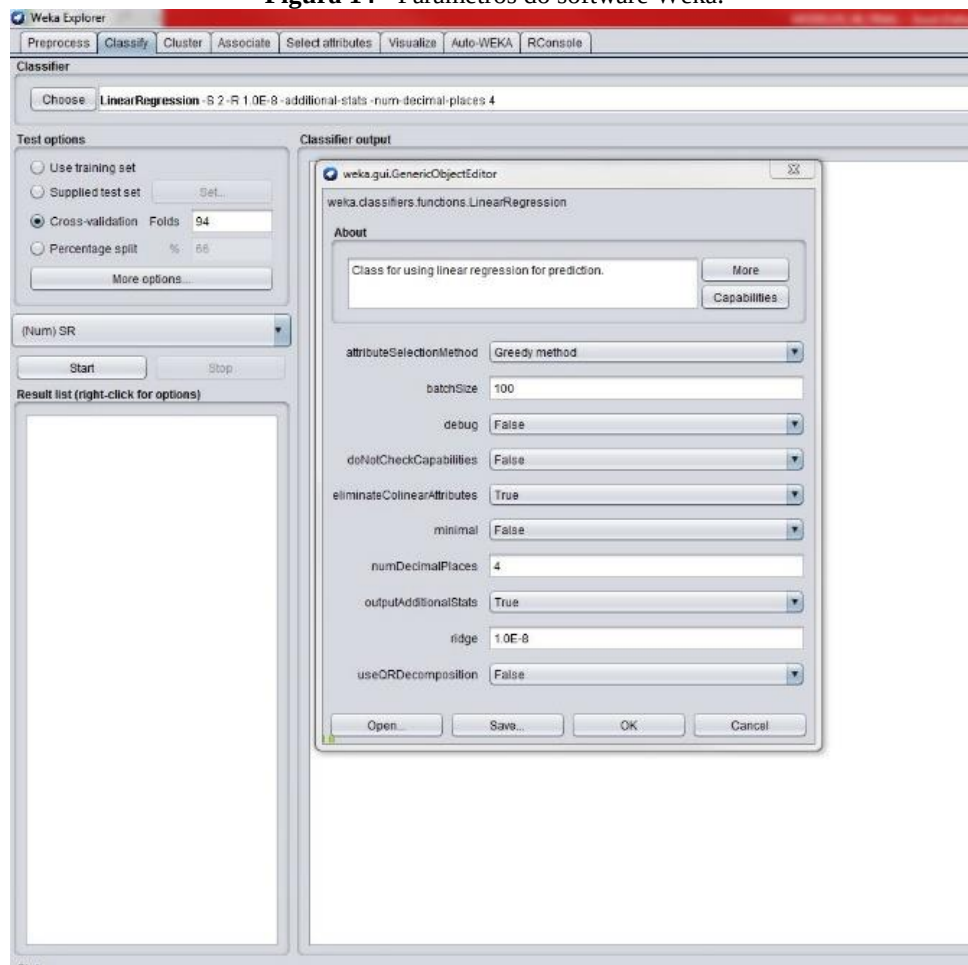
ϵ é o erro associado a regressão. Sendo a distância entre o valor do modelo e o valor real.

No *software* Weka, para desenvolvimento dos modelos de Regressão Linear Múltipla foi adotado o algoritmo de *Greedy* para seleção dos atributos que serão associados às variáveis independentes do modelo. Esse algoritmo tem como objetivo reduzir os efeitos de colinearidade entre as biometrias e os IV, a fim de evitar redundância ou até mesmo à inserção de erros nos modelos de predição.

Para ajuste do modelo de regressão, foi realizada a *Leave-One-Out Cross-Validation* (LOOCV), uma validação cruzada na qual o modelo é recalculado removendo uma amostra por vez, obtendo um valor predito para a amostra removida, até que todas as possíveis combinações tenham sido realizadas.

A Figura 13, apresenta todos os parâmetros utilizados para o processamento dos modelos, sendo o *Cross-validation*, onde se estabeleceu 94 interações (Número de amostras de treinamento para realização da LOOCV) e *Attribute Selection Method* no qual se adotou o método *Greedy*, os demais parâmetros foram mantido as opções padrões do *software* Weka.

Figura 14 - Parâmetros do software Weka.



Fonte: O autor (2021).

Foram considerados 16 cenários para criação dos modelos. Todos os cenários contêm como variáveis os 15 índices de vegetação mencionados anteriormente, e como fator de variação entre cenários, a presença de um, dois, três ou quadro variáveis biométricas, com suas respectivas possibilidades de combinação. A tabela 3 apresenta todos os cenários considerados para criação dos modelos de predição de produtividade. Este conjunto de cenários permitiu avaliar se os dados biométricos têm influência no modelo gerado, e se a utilização de mais de um tipo de biometria possibilita a obtenção de modelos com maior qualidade.

Tabela 3 - Cenários para criação dos modelos de predição.

Id	Cenários
1	Sem Biometrias
2	DMC
3	NPI
4	EMC
5	TCH
6	DMC - NPI
7	DMC - EMC
8	DMC - TCH
9	NPI - EMC
10	NPI - TCH
11	EMC - TCH
12	DMC - NPI - EMC
13	DMC - NPI - TCH
14	NPI - EMC - TCH
15	EMC - TCH - DMC
16	DMC - NPI - EMC - TCH

Fonte: O autor (2021).

Para compreender a influência que cada variável do modelo possui na predição, foi realizada a normalização dos dados de produtividade, IV e biometrias de treinamento, e em seguida os modelos de cada cenário foram recalculados. Essa padronização das escalas das variáveis independentes dos modelos de regressão se faz necessária pois, analisando os valores dos parâmetros β obtidos com os dados normalizados é possível verificar qual parâmetro possui maior influência na predição.

4.3.5 Análise da qualidade dos modelos de regressão.

Para análise da qualidade dos modelos de regressão obtidos em cada cenário, foram utilizadas as seguintes métricas estatísticas: coeficiente de correlação (R), coeficiente de determinação (R^2) e coeficiente de determinação ajustado para múltiplas regressões (R^2 ajustado). Além disso, através da LOOCV foi calculado o erro médio absoluto e a raiz do erro quadrático médio (RMSE).

Inicialmente a análise de qualidade foi realizada com o produto da LOOCV dos dados de treinamento. Como uma das premissas da regressão múltipla determina que os resíduos devem apresentar distribuição normal. Foi utilizado o teste estatístico *Shapiro-Wilk* no programa *RStudio* para avaliação da normalidade dos resíduos.

Os dados de produtividade que não foram utilizados para o treinamento dos modelos serviram como dados de validação da aplicação dos modelos. Entretanto, algumas variáveis possuem incompatibilidade de escala, o que limita a aplicação de alguns modelos.

O procedimento de obtenção dos parâmetros biométricos da cana de açúcar é realizado de forma manual e pontual, uma pratica comum dentro das usinas para gestão e manejo da lavoura, sendo portanto um dado atualizado frequentemente para cada talhão. Entretanto, existe uma problemática no uso destes dados nos modelos do presente trabalho, com relação à compatibilidade de escala com os dados de SR pois geralmente são coletadas poucas amostras de biometria e os dados acabam sendo generalizados para o talhão.

Uma possível utilização dos dados biométricos seria a interpolação para permitir seu uso como variável espacializada, o que atualmente não têm condições de ser realizado com consistência devido à baixa densidade amostral. Outra opção seria utilizar um valor numérico único, porém esta prática desprezaria a variabilidade espacial e possivelmente introduziria erros no processo de predição da produtividade.

Neste contexto, apenas os modelos que não utilizam as variáveis biométricas foram aplicados e os resultados comparados com os dados de produtividade de validação. Foram calculados os valores de discrepância entre os dados de produtividade medidos e os dados estimados por cada modelo em validação, obtendo valores de Erro Médio (ME) e Erro Médio Absoluto (MAE), desvio padrão dos erros e RMSE. Em seguida foi realizado o teste de Friedman para verificar diferença significativa entre as médias dos valores de produtividade medidos e os estimados pelos modelos aplicados; e o teste Nemenyi para verificação de diferença significativa para cada par de médias.

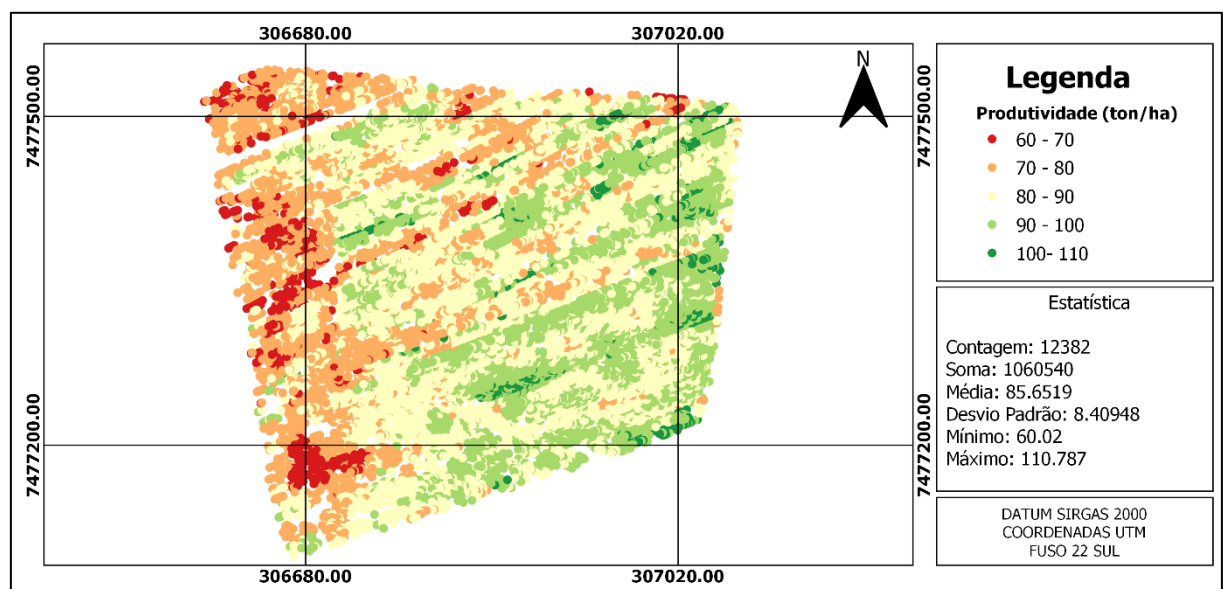
5. RESULTADOS E DISCUSSÕES

5.1 RESULTADOS DOS PROCESSAMENTOS DOS DADOS

Com a extração dos parâmetros biométricos do relatório cedido pela administração da Fazenda Miss Brasil, foram obtidas 12 amostras distribuídas dentro do Talhão 10. O perímetro apresentado na Figura 6, foi gerado através do levantamento geodésico onde foram levantados 21 pontos. Pode se observar a existência de uma descontinuidade do limite, gerando uma divisão em duas partes do Talhão 10, isso ocorreu pois essa faixa foi colhida por uma colhedora que não possui o monitor de produtividade, portanto as informações desta parte da lavoura foram perdidas.

Com o cartão com dados de produtividade extraído da colhedora CH670 em formato John Deere, utilizando do programa *Apex*, foram obtidos 80617 pontos com dados de produtividade em formato Shapefile. Esses dados foram convertidos para formato TXT, para realização do processo de filtragem no *software Mapfilter* e em seguida convertidos novamente para o formato Shapefile. Neste processo foram obtidos 12382 pontos de produtividade, como mostra a Figura 14.

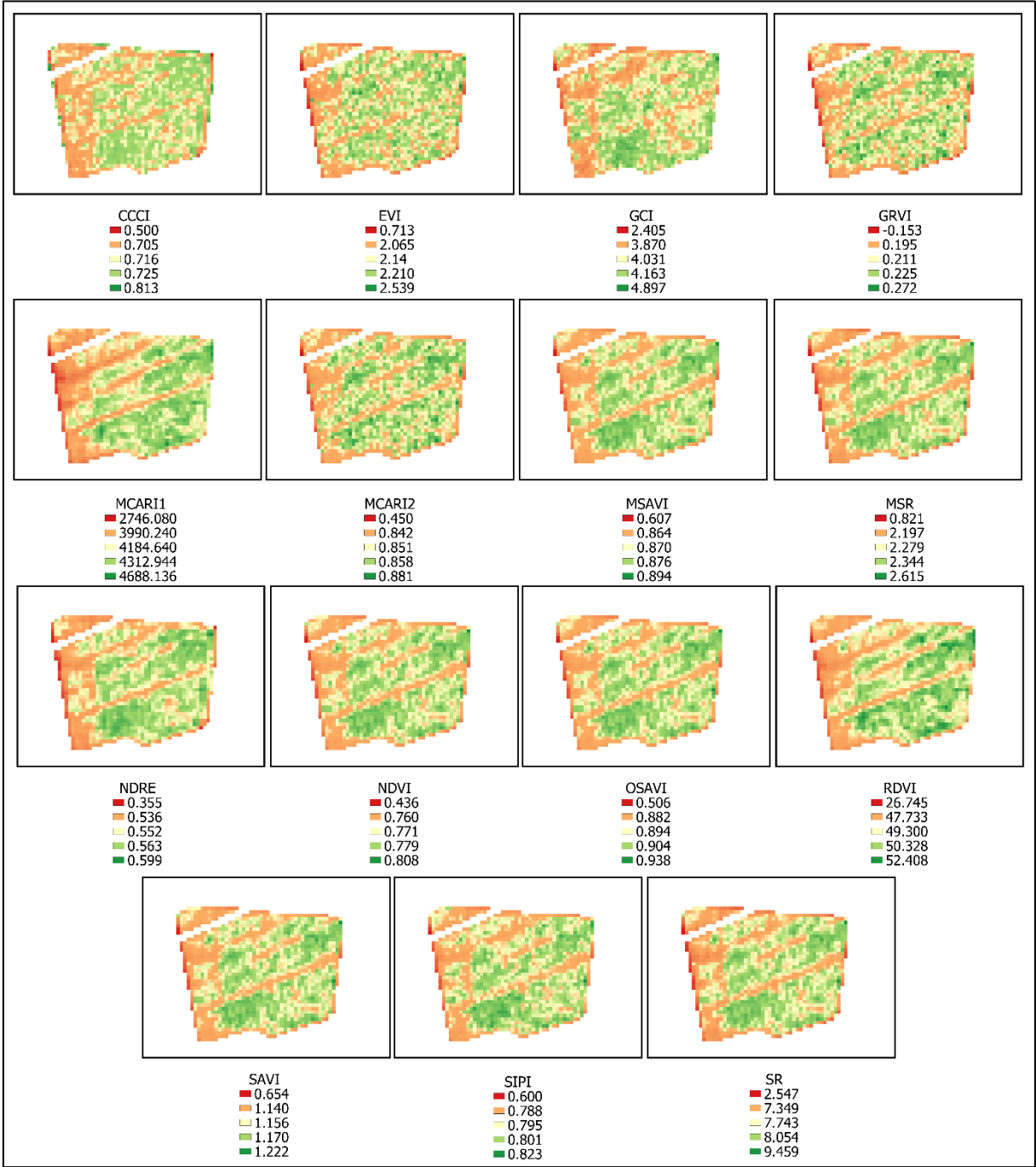
Figura 15 - Pontos de produtividade filtrados.



Fonte: O autor (2021).

Após o tratamento dos dados de produtividade foi realizado o processamento dos dados de sensoriamento remoto, com o cálculo dos índices por álgebra de mapas. A Figura 16 apresenta os índices de vegetação gerados neste trabalho.

Figura 16 – Raster dos Índices de Vegetação



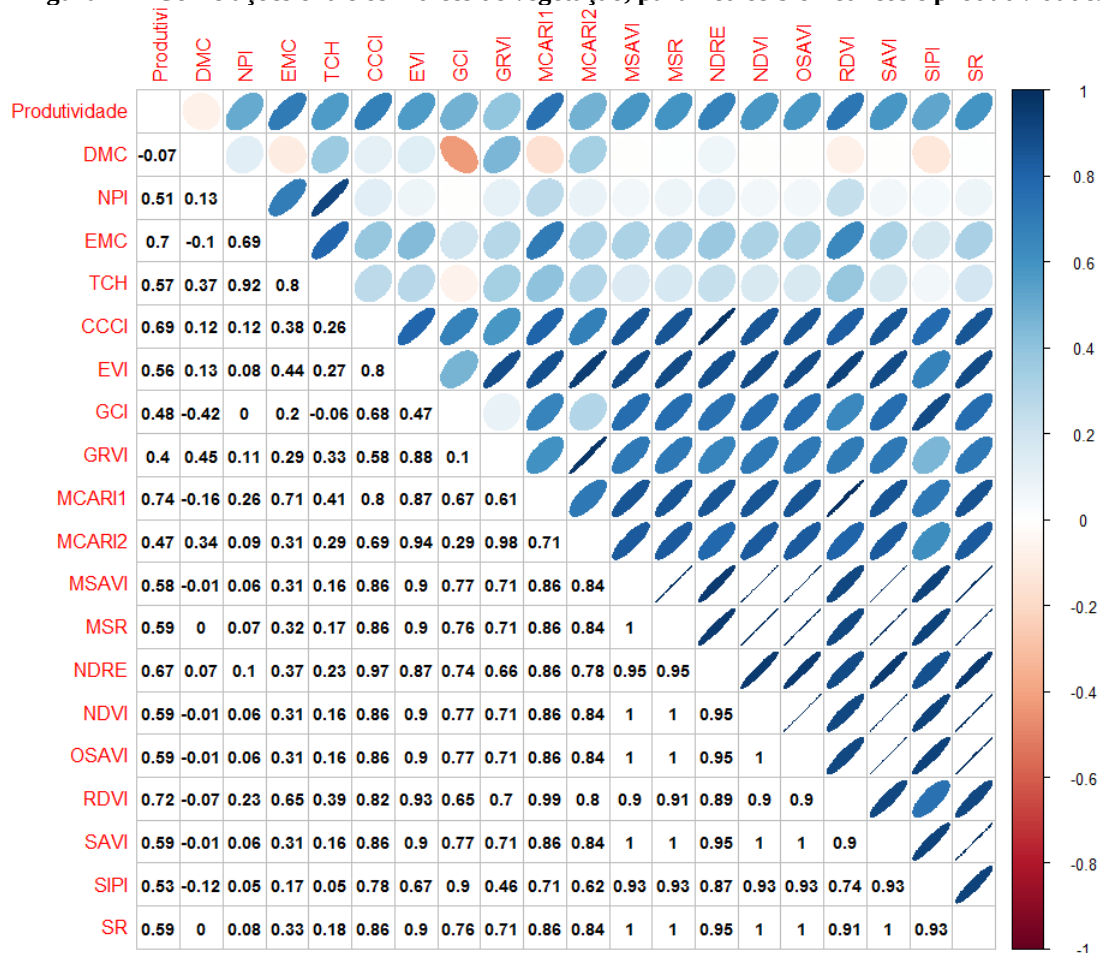
Fonte: O autor (2021).

O procedimento de seleção dos dados de treinamento para obtenção dos modelos teve como resultado 94 amostras de treinamento, cada amostra com valores de produtividade, IV e dados biométricos. Os demais dados de produtividade foram utilizados em etapas subsequentes à geração dos modelos, como dados de avaliação da aplicação dos modelos.

5.2 CORRELAÇÃO DOS PARÂMETROS PARA PREDIÇÃO DA PRODUTIVIDADE

A figura 17 apresenta os valores de correlação entre os índices de vegetação, parâmetros biométricos e a produtividade.

Figura 17 - Correlações entre os índices de vegetação, parâmetros biométricos e produtividade.



Fonte: O autor (2021).

O parâmetro biométrico com maior correlação com a produtividade foi Estatura Média dos Colmos (EMC) com uma correlação de 0,7. Os parâmetros Tonelada de Cana por Hectares (TCH), Numero de Perfilhos Industrializáveis (NPI) e Diâmetro Médio dos Colmos (DMC),

apresentaram 0,57, 0,51 e 0,07, respectivamente. Conforme Miocque (1999), a estatura de colmos tem correlação positiva com a produtividade. De acordo com estudos de Rocha (2018), parâmetro biométrico com maior correlação com a produtividade é o EMC, convergindo com os resultados do presente trabalho.

Com relação aos IV, MCARI1, RDVI e CCCI apresentaram maior correlação com a produtividade. O índice *First Modified Chlorophyll Absorption Ratio Index* (MCARI1) foi desenvolvido para estimar o Índice de Área Foliar (IAF), seu modelo matemático foi elaborado para minimizar os efeitos da clorofila e com a integração do infravermelho próximo para aumentar sensibilidade as mudanças de IAF (HABOUDANE et al., 2004). MCARI1 apresentou uma correlação de 0,74 com a produtividade da cana de açúcar, segundo Clevers (1999), IAF é uma característica das culturas que está mais relacionada com a produção agrícola sendo uma das principais variáveis para predição da produtividade. O *Canopy Chlorophyll Content Index* (CCCI), como o MCARI1, é um índice que quantifica a quantidade de clorofila na vegetação.

O *Renormalized Difference Vegetation Index* (RDVI), elaborado por Roujean e Breon (1995) com objetivo de elaborar um IV mais adaptado para mensurar radiação fotossinteticamente ativa absorvida, combinando as características dos índices DVI e NDVI. A radiação fotossinteticamente ativa absorvida é utilizada em muitos modelos de predição de Produtividade Primária Bruta (GPP), que corresponde a fotossíntese bruta, capaz de mensurar a capacidade de fixação de carbono da cultura (SILVA et al., 2013). A cana de açúcar é uma cultura do tipo C4, portanto apresenta uma alta eficiência em fixação de carbono, o que explica a correlação moderada que este IV apresentou.

Os IV mencionados anteriormente apresentaram uma correlação de 0,74 (MCARI1) e de 0,65 (RDVI) com o parâmetro biométrico EMC, como apresenta Rocha (2018), é possível então utilizar estes IV para mensuração da taxa de brotação, portanto, avaliar locais com falhas de plantio, informação cada vez mais demandada para verificar a disposição e qualidade das mudas depositadas no sulco de plantio em usinas com plantio mecanizado.

5.3 MODELOS DE PREDIÇÃO

Dos 16 cenários, o método para seleção de variáveis *Greedy* obteve como resultados 14 modelos de regressão linear múltipla, uma vez que o cenário Sem Biometrias e o cenário DMC obtiveram como resultado o mesmo modelo, e uma situação semelhante foi verificada nos

modelos obtidos para os cenários DMC-NPI-EMC e DMC-NPI-EMC-TCH. A Tabela 4, apresenta os modelos mais interpretáveis e aplicáveis obtidos pelo algoritmo.

Tabela 4 - Modelos gerados para cada cenário.

Cenários		Variáveis Selecionadas	Parâmetro β	Interseção	Parâmetro β normalizados
Sem Biometrias	DMC	CCCI	205.6814	-145.4984	0.2826
		GRVI	-75.9583		-0.0873
		MCARI1	0.0234		0.4697
	NPI	NPI	1.2661	-681.4635	0.1604
		CCCI	382.1757		0.5251
		EVI	-96.345		-0.6281
		GCI	-36.4603		-0.7259
		MCARI1	0.0462		0.9295
		SIPI	803.468		0.4852
	EMC	EMC	0.1557	-800.6247	0.3012
		CCCI	404.5005		0.5558
		EVI	-145.3082		-0.9473
		MCARI1	0.0267		0.5375
		MCARI2	897.851		0.5664
	TCH	CCCI	406.8982	-888.2205	0.5591
		EVI	-130.3145		-0.8495
		GCI	-52.4399		-1.0441
		MCARI1	0.0599		1.2038
		SIPI	1158.0433		0.6993
	DMC - NPI	DMC	-25.3813	-90.883	-0.2923
		NPI	1.5694		0.1988
		CCCI	545.8827		0.75
		EVI	-151.5088		-0.9877
		GCI	-54.7621		-1.0903
		MCARI1	0.0374		0.7516
		SR	28.5383		0.9389
	DMC - EMC	DMC	-21.0995	-1110.4608	-0.243
		EMC	0.1487		0.2876
		CCCI	522.5631		0.718
		EVI	-86.997		-0.5671
		GCI	-56.0208		-1.1154
		MCARI1	0.026		0.5234
		SIPI	1442.4366		0.8711
	DMC - TCH	DMC	-35.0893	-68.8977	-0.4041
		TCH	0.1452		0.2639
		CCCI	555.8984		0.7638

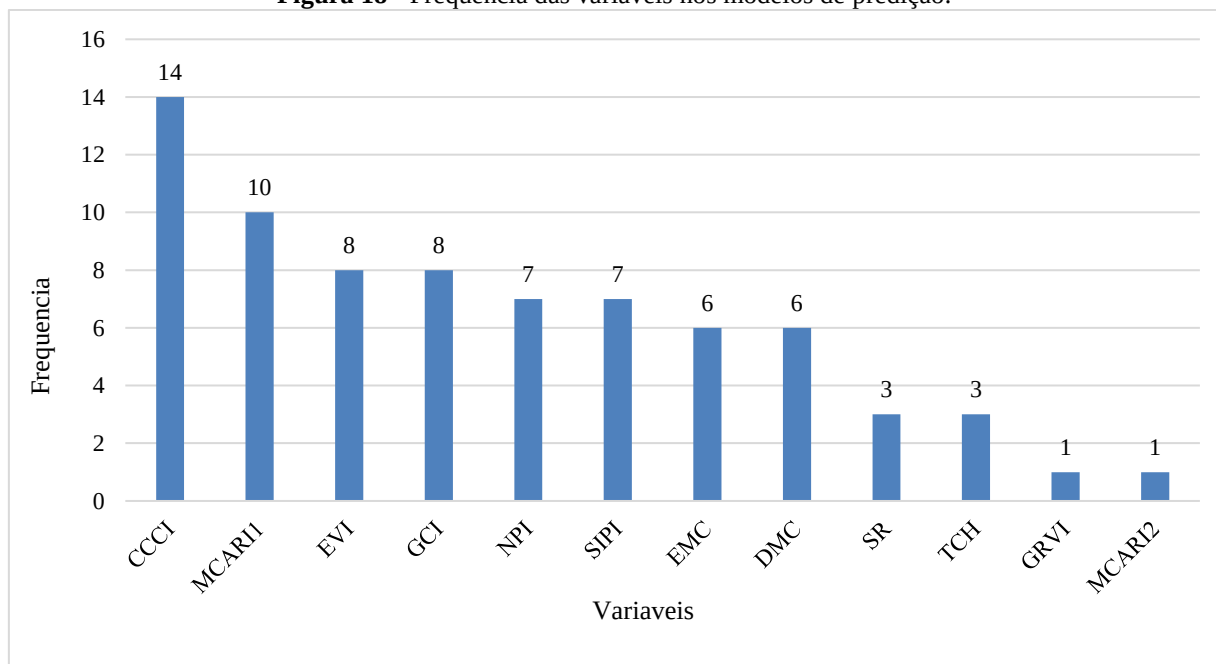
		EVI	-141.8928		-0.925
		GCI	-52.0462		-1.0362
		MCARI1	0.0277		0.558
		SR	29.4169		0.9678
NPI - EMC	NPI	1.2095			0.1532
	EMC	0.1704		-343.3024	0.3296
	CCCI	290.5445			0.3992
	SIPI	218.2266			0.1318
NPI - TCH	NPI	2.967		-274.65	0.3758
	CCCI	457.9003			0.6291
EMC -TCH	EMC	0.1166			0.2255
	CCCI	387.9011			0.533
	EVI	-92.6759		-852.6096	-0.6042
	GCI	-38.6471			-0.7695
	MCARI1	0.0356			0.7152
	SIPI	1058.1964			0.639
DMC - NPI - EMC	DMC	-11.6617			-0.1343
	NPI	1.7523		-218.0813	0.222
	EMC	0.1252			0.2422
	CCCI	403.4793			0.5544
DMC - NPI - TCH	DMC	-11.2299			-0.1293
	NPI	2.7716		-188.0402	0.3511
	CCCI	318.8802			0.4381
	MCARI1	0.0104			0.2086
NPI - EMC - TCH	NPI	3.2485			0.4115
	EMC	0.3472			0.6716
	TCH	-0.3081			-0.56
	CCCI	473.3051		-975.897	0.6503
	EVI	-35.6443			-0.2324
	GCI	-32.9787			-0.6566
	SIPI	1077.2266			0.6505
EMC - TCH - DMC	DMC	-27.1214			-0.3123
	TCH	0.1367			0.2485
	CCCI	507.5902			0.6974
	GCI	-51.4801		-2028.3193	-1.025
	MCARI1	0.0345			0.6944
	SIPI	2734.6226			1.6514
	SR	-39.7875			-1.309

Fonte: O autor (2021).

Neste trabalho foi possível verificar que a variável NPI teve ocorrência em 7 modelos, sendo a variável biométrica mais presente nas predições em seguida da DMC e EMC com 6

ocorrências, e TCH com 3. O NPI foi a terceira biometria com maior correlação com os dados de produtividade, contudo, foi o parâmetro que apresentou menor correlação com IV nenhuma superior a 0,30, o que explica o seu uso em todos os cenários que esteve presente por se tratar de uma variável que introduziu ao algoritmo de regressão novas informações que agregaram aos modelos de predição. A Figura 18 apresenta um gráfico indicando a frequência de ocorrência de cada variável nos modelos de predição gerados. Pode se observar que índices de vegetação mais abordados na literatura para monitoramento agrícola, modelos predições entre outras aplicações, como NDVI, NDRE E SAVI, não foram selecionados em nenhum modelo. Assim, esse resultado mostra a importância de se estudar e combinar índices com diferentes propriedades de campo, para compreender as possíveis aplicações para esses IV.

Figura 18 - Frequência das variáveis nos modelos de predição.



Fonte: O autor (2021).

Analisando os modelos, a biometria DMC apresentou somente parâmetros β negativo, resultado da correlação negativa que essa variável apresentou com a produtividade, portanto, quanto maior o valor de DMC menor será o valor de produtividade. Esses valores condizem com os resultados apresentados por Rocha (2018), no qual se verificou que o DMC obteve um menor impacto nos modelos de predição, devido ao alto coeficiente de variação desta biometria. A Tabela 5 apresenta quais variáveis apresentaram a maior e a menor influência em cada modelo e a Tabela 6 a frequência de cada uma dessas variáveis. A biometria EMC foi a única

que teve maior influência, no cenário NPI - EMC – TCH, enquanto as biometrias NPI e DMC, só tiveram parâmetros negativos nos modelos, tendo menor contribuição nas predições nos modelos que se encontram.

Os cenários onde o índice MCARI1 teve maior influência, o GCI apresentou o menor valor do parâmetro B normalizado, sendo esse índice o que teve menor contribuição nas predições, com uma frequência em 7 modelos. MCARI1 e CCCI, foram os índices com maior frequência no modelo e também foram os índices que mais contribuíram na predição em 4 modelos para ambos IV. Essa frequência está de acordo com os dados apresentando anteriormente na Figura 17, onde essas variáveis apresentaram uma maior correlação com a produtividade.

A frequência desses parâmetros como variáveis nos modelos de predição, tanto biométricos como os IV citados anteriormente, mostraram serem coerentes, pois a produtividade de uma lavoura de cana de açúcar pode ser calculado pelos parâmetros DMC, NPI e EMC; CCCI e MCARI1 são índices de vegetação criados para determinação da quantidade de clorofila, portanto relacionado ao processo de fotossíntese da cultura, além de serem usados para o cálculo de IAF, parâmetro segundo Clevers (1999), que está mais relacionada com a produção agrícola sendo uma das principais variáveis para predição de produtividade.

Tabela 5 - Variáveis com maior e menor influência na predição da produtividade.

Cenários		Menor parâmetro β normalizado	Maior parâmetro β normalizado
Sem Biometrias	DMC	GRVI	MCARI1
	NPI	GCI	MCARI1
	EMC	EVI	MCARI2
	TCH	GCI	MCARI1
	DMC - NPI	GCI	SR
	DMC - EMC	GCI	SIPI
	DMC - TCH	GCI	SR
	NPI - EMC	SIPI	CCCI
	NPI - TCH	NPI	CCCI
	EMC -TCH	GCI	MCARI1
DMC - NPI - EMC	DMC - NPI - EMC - TCH	DMC	CCCI
	DMC - NPI - TCH	DMC	CCCI
	NPI - EMC - TCH	GCI	EMC
	EMC - TCH - DMC	SR	SIPI

Fonte: O autor (2021).

Tabela 6 - Frequência das variáveis com maior e menor influência na predição da produtividade.

Variáveis com menor parâmetro β normalizado	Frequência	Variáveis com maior parâmetro β normalizado	Frequência
GRVI	1	MCARI1	4
GCI	7	MCARI2	1
EVI	1	SR	2
SIPI	1	SIPI	2
NPI	1	CCCI	4
DMC	2	EMC	1
SR	1		

Fonte: O autor (2021).

5.4 ANALISE DA QUALIDADE DOS MODELOS POR VALIDAÇÃO CRUZADA

A Tabela 7, apresenta as métricas estatísticas da validação cruzada de cada cenário. Os melhores modelos avaliados pelo R^2 ajustado foram dos cenários DMC – NPI, DMC – TCH e EMC - TCH – DMC, com um R^2 ajustado de 0,76, sendo que esse último cenário apresentou o menor Erro Médio Absoluto de 2,94 ton.ha⁻¹. O maior coeficiente de correlação de 0,88 foi do cenário DMC - NPI - EMC/DMC - NPI - EMC – TCH que também apresentou o menor RMSE de 3,62 ton.ha⁻¹. Os melhores modelos apresentaram algum parâmetro biométrico, indo de encontro com os estudos realizados por Rocha (2018). Segundo o autor, a integração de parâmetros biométricos e IV pode ser utilizada para criação de modelos preditivos mais assertivos.

Tabela 7 - Métricas estáticas dos dados de treinamento

Cenários		R ²	R ² ajustado	Coefficiente de correlação	Erro médio absoluto	RMSE
Sem Biometrias	DMC	0,59	0,57	0,73	4,22	5,16
	NPI	0,76	0,74	0,84	3,23	4,04
	EMC	0,74	0,73	0,84	3,23	4,09
	TCH	0,74	0,72	0,83	3,41	4,14
	DMC - NPI	0,78	0,76	0,87	3,05	3,73
	DMC - EMC	0,77	0,75	0,85	3,12	3,92
	DMC - TCH	0,78	0,76	0,87	3,00	3,70
	NPI - EMC	0,72	0,71	0,82	3,43	4,32
	NPI - TCH	0,66	0,65	0,79	3,91	4,63
	EMC - TCH	0,74	0,73	0,83	3,41	4,23
DMC - NPI - EMC	DMC - NPI - EMC - TCH	0,73	0,72	0,88	2,95	3,62
	DMC - NPI - TCH	0,71	0,70	0,83	3,36	4,23
	NPI - EMC - TCH	0,75	0,73	0,85	3,08	3,93
	EMC - TCH - DMC	0,78	0,76	0,87	2,94	3,70

Fonte: O autor (2021).

Os cenários Sem Biometrias e DMC obtiveram os piores resultados dentro da validação cruzada, o modelo gerado por esse cenário não incluiu o parâmetro DMC na predição tendo como variáveis somente os índices de vegetação. O cenário TCH também não apresentou biometrias como parâmetros para predição, contudo, seu modelo apresentou melhores resultados na validação cruzada, apresentando assim que somente uso de IV foi possível ter uma predição assertiva contradizendo o que foi apresentado anteriormente pelos resultados dos melhores modelos e por Rocha (2018).

A Tabela 8 apresenta o resultado do teste de normalidade *Shapiro-Wilk*, que foi aplicado no conjunto dos valores dos resíduos dos modelos nos dados de treinamento. Exceto pelo cenário NPI – TCH, todos os modelos apresentaram normalidade em nível de significância de 0,05, sendo portanto modelos que atendem a premissa da regressão linear, que determina que os resíduos devem apresentar distribuição normal.

Tabela 8 - Resultado do teste de normalidade dos resíduos - *Shapiro-Wilk*

Cenários	Valor-P
Sem Biometria	0.6874
DMC	0.6874
NPI	0.551
EMC	0.9121
TCH	0.3478
DMC - NPI	0.2477
DMC - EMC	0.4588
DMC - TCH	0.2764
NPI - EMC	0.4387
NPI - TCH	0.004033
EMC - TCH	0.7799
DMC - NPI - EMC	0.2168
DMC - NPI - TCH	0.2417
NPI - EMC - TCH	0.1072
EMC - TCH - DMC	0.1804
DMC - NPI - EMC - TCH	0.2168

Fonte: O autor (2021).

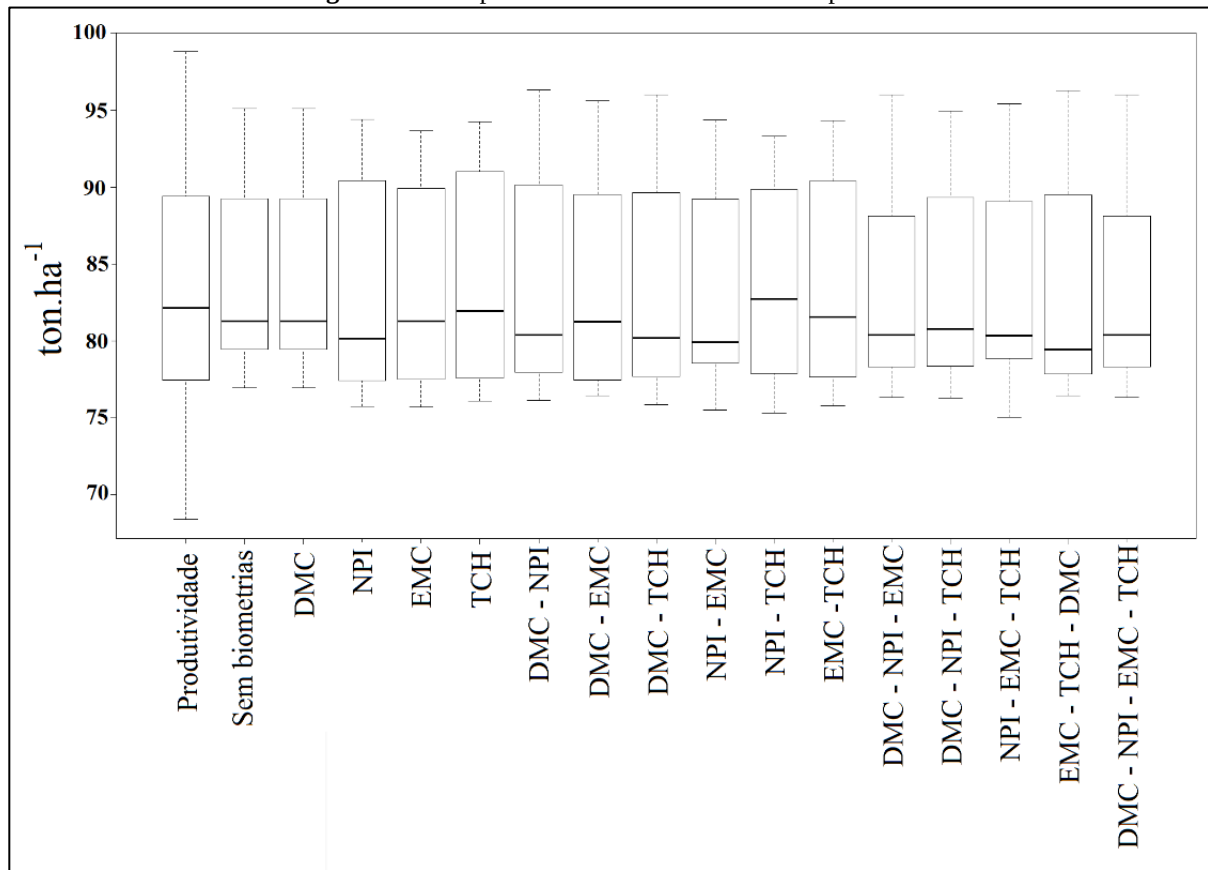
A Tabela 9 apresenta as estatísticas descritivas de cada modelo e dos dados de produtividade real. Pode se observar que nenhuma média dos cenários apresentou uma diferença maior que uma ton.ha^{-1} em relação a produtividade real, sendo a maior diferença entre as médias de $-0,20 \text{ ton.ha}^{-1}$ nos cenários Sem Biometrias, DMC e DMC-TCH. A variância e os valores de desvio padrão dos modelos das predições foram menores que os apresentados pela produtividade real, portanto os dados estimados mostraram maior uniformidade. No Apêndice 1, é apresentado os gráficos de dispersão do modelo de cada cenário.

Tabela 9 - Estatísticas descritivas

Cenários	Média	Variância	Desvio Padrão
Produtividade	83.57	56.97	7.55
Sem biometrias	83.77	33.55	5.79
DMC	83.77	33.55	5.79
NPI	83.50	43.10	6.56
EMC	83.47	42.31	6.50
TCH	83.75	42.14	6.49
DMC - NPI	83.69	44.35	6.66
DMC - EMC	83.46	43.57	6.60
DMC - TCH	83.37	44.06	6.64
NPI - EMC	83.57	41.03	6.41
NPI - TCH	83.57	37.32	6.11
EMC - TCH	83.74	42.52	6.52
DMC - NPI - EMC	83.57	41.57	6.45
DMC - NPI - TCH	83.69	40.79	6.39
NPI - EMC - TCH	83.58	42.97	6.56
EMC - TCH - DMC	83.45	44.42	6.66
DMC - NPI - EMC - TCH	83.57	41.57	6.45

Fonte: O autor (2021).

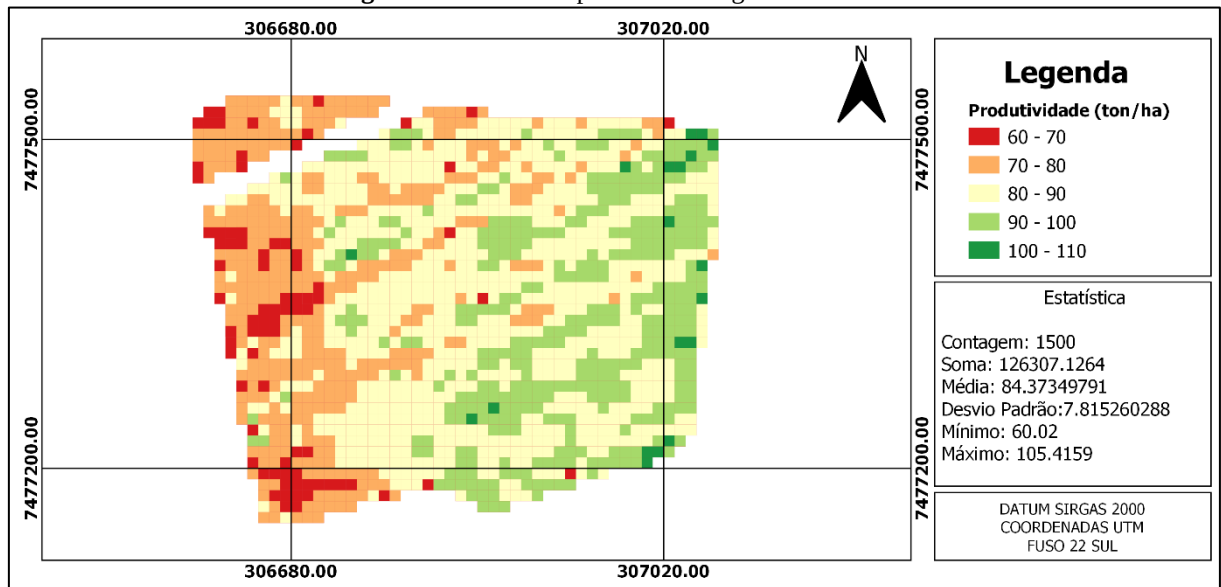
Para melhor análise da variação dos resultados da validação cruzada, foi gerado o gráfico *boxplot* da Figura 19. Pode se observar que as caudas dos modelos não se aproximaram dos dados reais de produtividade, permitindo inferir que os modelos subestimaram os dados de maior produtividade e superestimaram os dados de menor produtividade. A mediana dos modelos se aproximou do dado real, apresentando simetria semelhante, quando a mediana se encontra mais próxima da parte inferior da caixa (mais próxima do primeiro quartil).

Figura 19 - Boxplot dos dados reais e estimados por cada modelo.

Fonte: O autor (2021).

5.4 ANALISE DA QUALIDADE DOS MODELOS BASEADOS EXCLUSIVAMENTE EM DADOS DE SENSORIAMENTO REMOTO PELOS DADOS DE TESTE

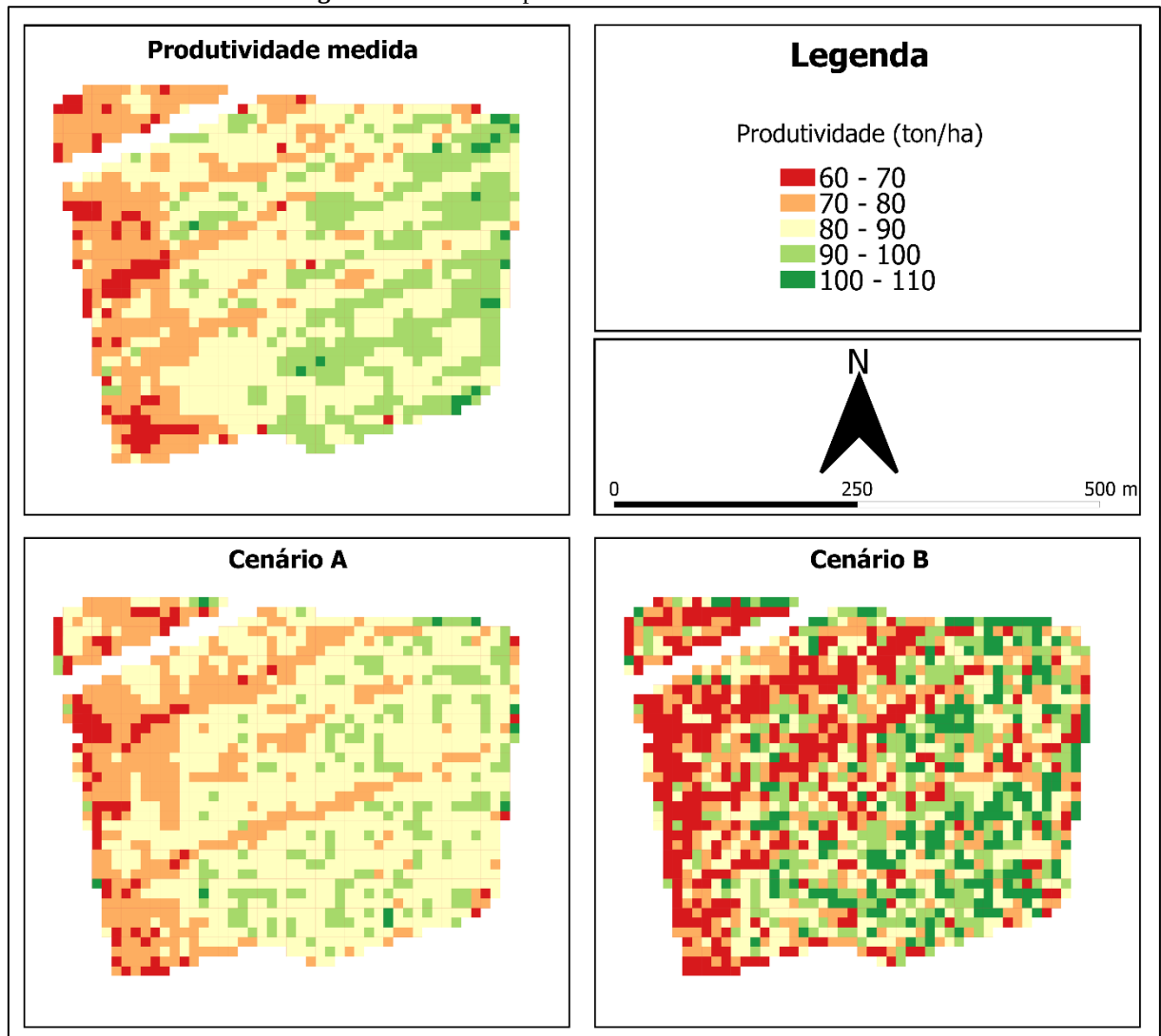
Para validação dos modelos sem a utilização dos dados de treinamento, foi realizada a generalização dos dados de produtividade filtrados, compatibilizando-os com a resolução espacial dos dados de sensoriamento remoto. A Figura 15, apresenta o resultado da generalização dos dados de produtividade para grade de 10 metros. Em seguida, os modelos baseados em dados de sensoriamento remoto foram aplicados e os resultados comparados com estes dados de produtividade.

Figura 20 - Dados de produtividade generalizados.

Fonte: O autor (2021).

Para a aplicação dos modelos faz-se necessária a disponibilização de dados espacializados para cada parâmetro. Entretanto, existe uma problemática no uso de dados biométricos nestes modelos por se tratar de uma informação pontual. No momento de aplicação do modelo seria necessário interpolar os parâmetros biométricos para seu uso como variável ou utilizar como valor numérico único, o que não levaria em conta a variabilidade espacial desses parâmetros. Vale ressaltar, como verificado na validação LOOCV, os dados biométricos tem grande relevância pois sua combinação com os dados de SR é possível melhorar a qualidade dos modelos de predição, sendo necessário portanto melhorar as técnicas de amostragem e mensuração destes dados biométricos em campo.

Os modelos dos cenários Sem Biometrias/DMC e TCH não apresentaram biometrias como variável em seus modelos de predição, portanto, foram aplicados para predição da produtividade. Estes cenários serão nomeados como cenário A e B, respectivamente, como apresenta a Figura 20, que também apresenta a representação espacial dos dados de produtividade medidos.

Figura 21 - Dados de produtividade medida e estimado.

Fonte: O autor (2021).

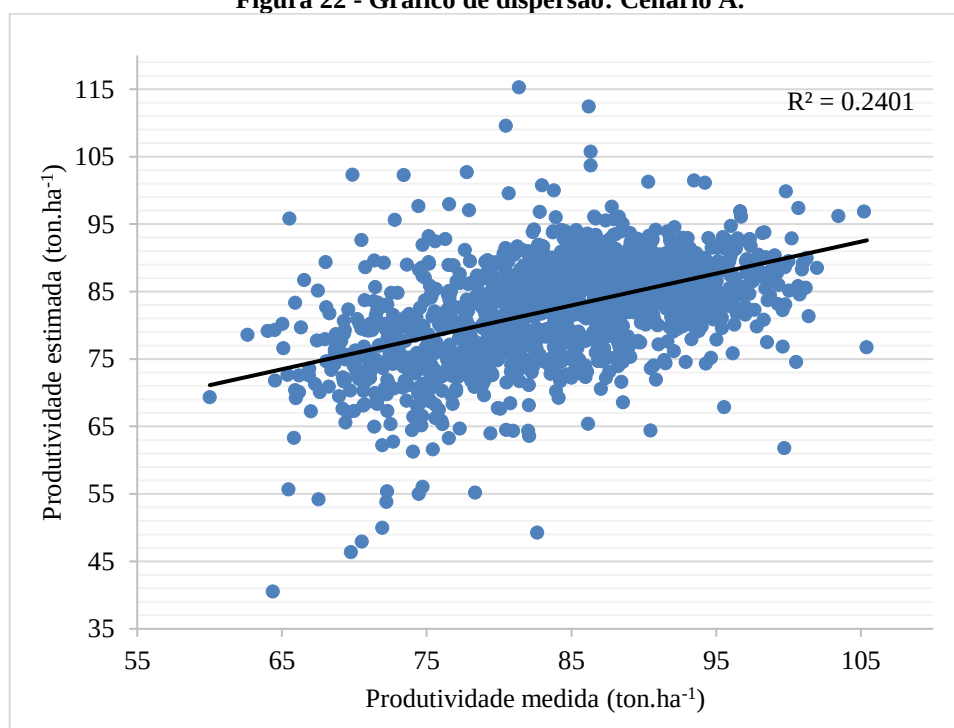
A Tabela 10, apresenta as estatísticas descritivas dos três cenários (Produtividade medida, A e B), analisando a média e comparado aos dados de validação do campo, o cenário A subestimou em 1,73 ton.ha⁻¹ em quanto o cenário B teve uma diferença de 3,17 ton.ha⁻¹. Os cenários A e B, subestimaram o valor mínimo e superestimaram o valor máximo de produtividade comparado ao dado medido, destacando o cenário B que obteve as maiores diferenças, chegando a valores de mais de 50 ton.ha⁻¹. Os dois cenários apresentaram um intervalo maior do que os dados medidos em campo, reflexo da subestimação dos valores mínimos e superestimação dos valores máximos de produtividade.

Tabela 10 - Estatísticas descritivas do dado real e os estimados.

	<i>Medido</i>	<i>Cenário A</i>	<i>Cenário B</i>
Média	84.37	82.65	81.20
Mediana	84.83	83.82	81.94
Modo	88.09	89.34	75.09
Desvio padrão	7.82	7.54	18.47
Variância da amostra	61.08	56.87	341.09
Curtose	-0.42	2.27	0.91
Assimetria	-0.25	-0.70	-0.30
Intervalo	45.40	74.74	154.76
Mínimo	60.02	40.51	3.79
Máximo	105.42	115.25	158.55
Soma	126307.13	123721.05	121554.62

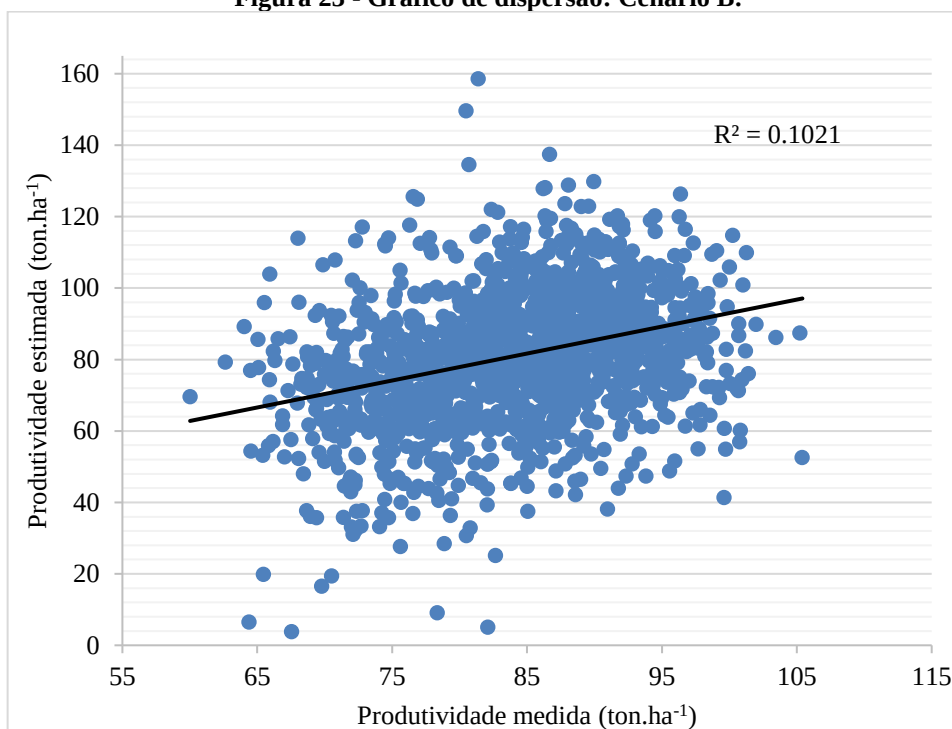
Fonte: O autor (2021).

A Figura 22 e 23, apresenta respectivamente o gráfico de dispersão do cenário A e B, dos valores estimados com os valores medidos de produtividade. Analisando os gráficos pode se observar que a relação entre a produtividade estimada e a medida não possuem uma relação linear. Portanto, para futuras predição de produtividade, pode se adotar a utilização de modelos não lineares para uma melhor predição da produtividade.

Figura 22 - Gráfico de dispersão: Cenário A.

Fonte: O autor (2021).

Figura 23 - Gráfico de dispersão: Cenário B.



Fonte: O autor (2021).

Analisando os modelos apresentados na Tabela 4, mencionada anteriormente, o índice MCARI1, nos dois cenários, apresentou maior coeficiente β , mostrando-se então como índice de vegetação bastante promissor para predição de produtividade comparado a índices de vegetação tradicionais como NDVI e NDRE, bastante consolidados na literatura. Analisando a Tabela 11, tanto no Erro Médio (ME) quanto no Erro Médio Absoluto (MAE), o cenário A apresentou os menores valores, de $1,73 \text{ ton.ha}^{-1}$ e $6,16 \text{ ton.ha}^{-1}$, respectivamente. Além de ter apresentado menores valores de desvio padrão para os dois erros, portanto, esse modelo além de apresentar uma maior acurácia comparado ao modelo do cenário B, obteve uma maior precisão na predição.

Sugawara *et al.* (2007), também utilizando da técnica de análise de regressão linear múltipla, apresentou em seu modelo um erro absoluto de $10,22 \text{ ton.ha}^{-1}$ para predição de produtividade de cana, enquanto o modelo do cenário A apresentou um erro de $6,16 \text{ ton.ha}^{-1}$. O modelo proposto por esses autores, leva em consideração a quantidade de safras pela qual a cultura implantada passou, e no estudo citado o dado de entrada foi cana soca de 4 corte, sendo a mesma situação dos dados utilizados neste trabalho. O modelo do cenário A apresentou o menor RMSE, de $8,15 \text{ ton.ha}^{-1}$, comparado ao modelo do cenário TCH que obteve $18,50 \text{ ton.ha}^{-1}$.

Picoli (2006), utilizando de redes neurais artificiais, dados de produtividade e modelo agrônomo-espectral com dados MODIS, obteve um modelo que apresentou um RMSE de 13,9 ton.ha⁻¹ para predição de cana soca. Como pode se observar na comparação entre outras metodologias, o modelo cenário A mostrou-se ser bastante promissor na predição de produtividade de cana de açúcar.

Tabela 11 - Erros dos modelos de predição

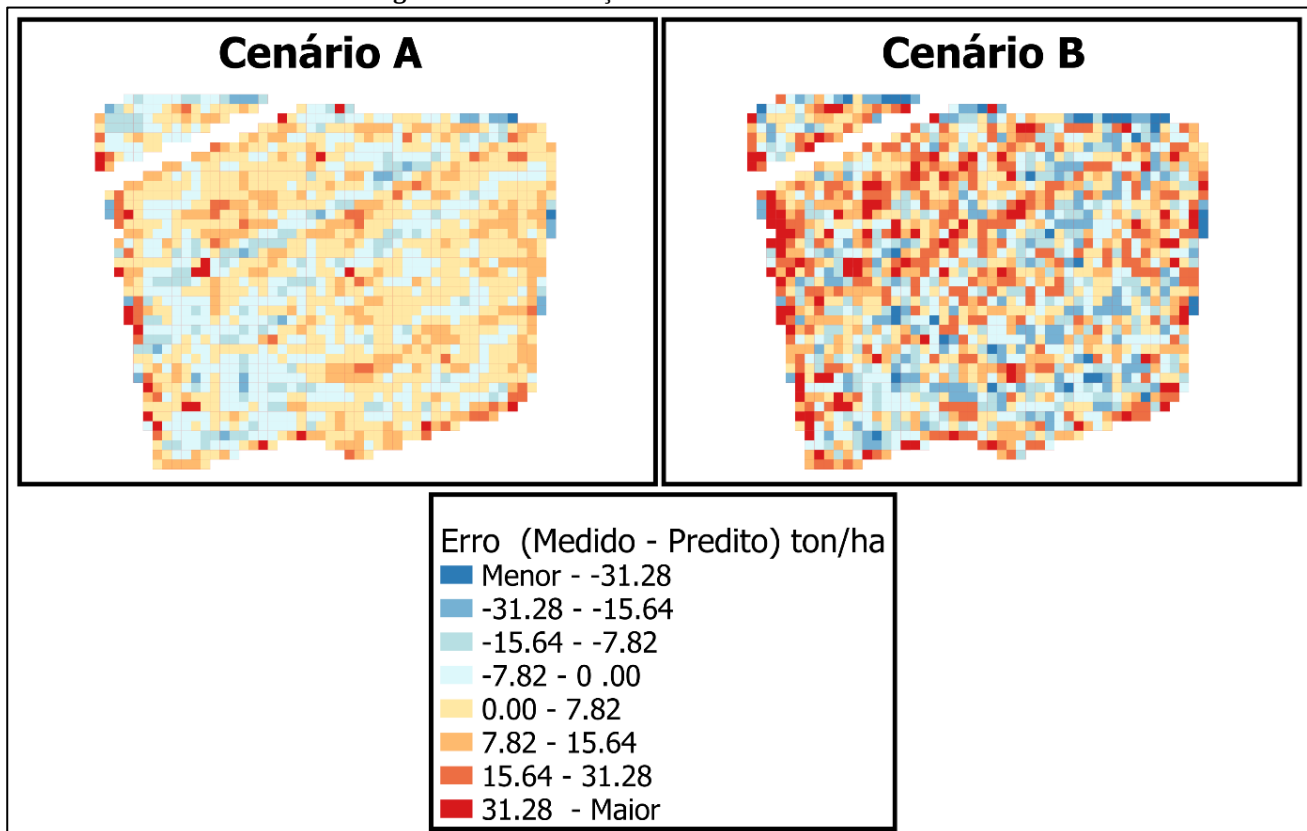
Cenários	Erro Médio (ME)	Desvio Padrão - ME	Erro Médio Absoluto (MAE)	Desvio Padrão - MAE	RMSE
Cenário A	1.73	7.76	6.16	5.01	8.15
Cenário B	3.17	17.60	13.80	11.38	18.50

Fonte: O autor (2021).

Utilizando desvio padrão, no qual se trata de uma medida de dispersão, dos dados de produtividade medidos em campo, como critério para avaliação dos erros dos modelos. Para análise dos erros foi calculada a diferença entre os dados de campo e os valores estimados por cada modelo, representados espacialmente na Figura 21 utilizando classes com intervalos de 7,82 ton.ha⁻¹ acima e abaixo de zero de um desvio padrão.

O cenário A resultou no intervalo variando de -7,82 a +7,82 ton.ha⁻¹, uma área de 10,89 ha o que corresponde a 71,34% da área do talhão. Esse cenário subestimou uma área de 2,42 ha, no intervalo de +7,82 a +15,64 ton.ha⁻¹ e superestimou no intervalo de -7,82 a -15,64 ton.ha⁻¹, 1,17 ha, esses dois intervalos corresponde a 23,51% da área deste talhão. Enquanto o cenário B, no intervalo para ser considerado eficaz seu modelo, obteve somente uma área de 5,55 ha, sendo 36,34% da área, considerando o mesmo intervalo descrito anteriormente, o cenário B subestimou uma área de 2,44 ha e superestimou uma área de 2,04 ha.

Figura 24 – Distribuição do erro dos cenários de estimava.



Fonte: O autor (2021).

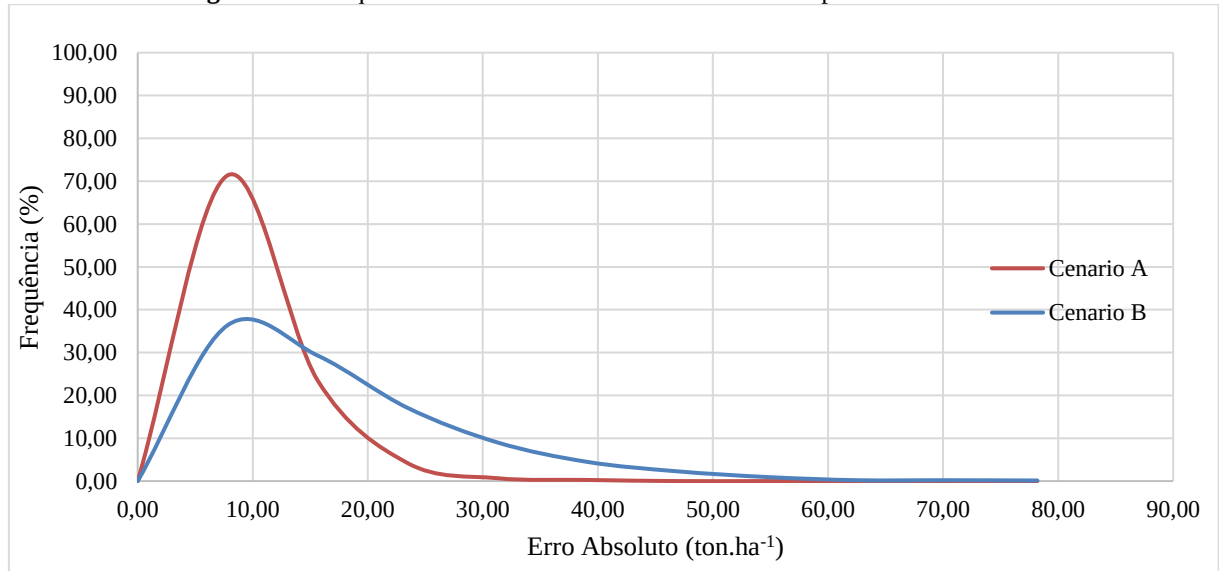
A Figura 22 apresenta o comportamento dos erros absolutos dos cenários predição de produtividade, que resultou nas maiores frequência abaixo de 7,82 ton.ha⁻¹, tanto no cenário A e no B, nos valores de 71,76% e 36,38%, respectivamente para cada cenário. Mais de 50% dos erros do Cenário A se encontram há um desvio padrão, enquanto o cenário B só atingiu valores acima dessa frequência a dois desvio padrões, com uma frequência de 29,31% no intervalo de 15,64 ton.ha⁻¹.

O cenário A se obteve uma frequência de 23,50% com erros entre 7,82 a 15,64 ton.ha⁻¹ sendo a segundo maior frequência neste cenário, tendo uma redução para 4,21% na frequência nos erros no intervalo de 15,64 a 23,46 ton.ha⁻¹. Enquanto o cenário B, apresentou uma frequência de 17,09% até o intervalo de 15,64 a 23,46 ton.ha⁻¹, tendo redução para uma frequência de 9,01% no intervalo de erro acima de 23,46 ton.ha⁻¹.

O cenário B somente apresentou frequências abaixo de 1% a partir do intervalo de erro de 54,74 ton.ha⁻¹ enquanto o cenário A apresentou valores baixos de frequência a partir de 31,38 ton.ha⁻¹. Esses resultado ressaltam o comportamento espacial da Figura 21, apresentada

anteriormente, onde os cenários que apresentaram maiores áreas com erros menores que $7,82 \text{ ton.ha}^{-1}$, possuem seus intervalos de maior frequência abaixo do limite estabelecido.

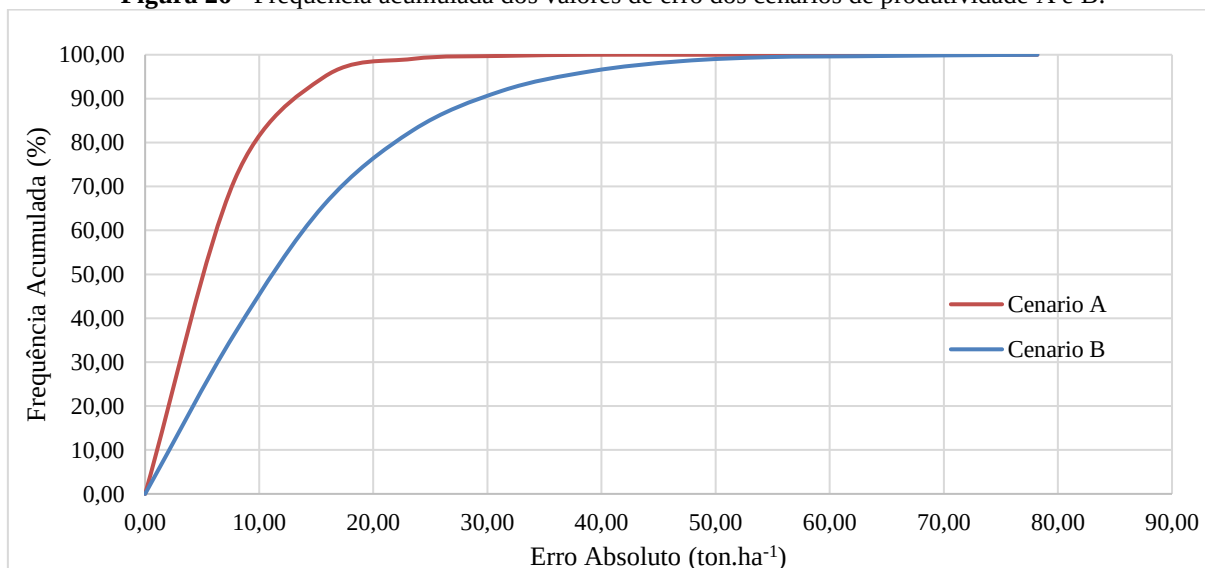
Figura 25 - Frequência dos valores de erro dos cenários de produtividade A e B.



Fonte: O autor (2021).

Pode-se observar na Figura 23, que representa as frequências acumuladas dos erros absolutos, que nos erros abaixo de $7,82 \text{ ton.ha}^{-1}$ somente o Cenário A atingiu frequências acima de 50% o que representa que maioria dos seus erros estão em torno a um desvio padrão dos valores de produtividade medido em campo.

Enquanto o cenário B apresentou somente 36,38% dentro deste mesmo intervalo, ou seja, nem metade dos valores dos erros desse modelo. A frequência acumulada desse cenário só atingiu o patamar dos 50% no intervalo de $7,82$ a $15,64 \text{ ton.ha}^{-1}$, valores de erro 2 vezes maiores que o parâmetro estabelecido.

Figura 26 - Frequência acumulada dos valores de erro dos cenários de produtividade A e B.

Fonte: O autor (2021).

Para os testes estatísticos de comparação de médias, inicialmente testou-se a normalidade dos dados de produtividade (medida e estimada por cada modelo) com teste de *Shapiro-Wilk* em um nível de significância de 0,05. O valor de P calculado foi de $2,2 \times 10^{-16}$, portanto os valores de produtividade não apresentaram distribuição normal.

Por não apresentarem uma distribuição normal e se tratarem de dados pareados, uma vez que os dados de produtividade de cada tratamento e os dados medidos referem-se à mesma unidade amostral, aplicou-se o teste Friedman para verificar a hipótese nula, de que não existe diferença entre as médias de produtividade. O teste de Friedman obteve como resultando um valor P-Valor de $3,229 \times 10^{-14}$, portanto, rejeita-se a hipótese nula e pode-se interpretar que os valores médios de produtividade apresentam diferença significativa entre si.

Para uma verificação de diferenças par-a-par, aplicou-se o teste de Nemenyi com nível de significância de 0,05. A tabela 12 apresenta o resultado do teste de Nemenyi, médias seguidas de mesma letra não são significativamente diferentes. Com esses resultados, a Produtividade Real com cenários apresentaram ter uma diferença significativa entre seus valores por apresentarem o P-Valor menor que 0,05. Enquanto, o Cenário A e B, entre si, apresentaram não ter diferença significativa entre seus dados de produtividade, portanto, podem ser considerados estatisticamente iguais.

Tabela 12 - Teste de Nemenyi

Tratamento	Produtividade média (ton.ha⁻¹)
Produtividade medida	84,37 b
Cenário A	82,64 a
Cenário B	81,20 a

Médias seguidas de mesma letra não são significativamente diferentes ao nível de significância de 5%.

Fonte: O autor (2021).

6. CONSIDERAÇÕES FINAIS

Os parâmetros biométricos que mais contribuíram para determinação da produtividade foram a Estatura Média dos Colmos (EMC) e o Número de Perfilhos Industrializáveis (NPI). Entre os índices de vegetação, os mais importante foram *Canopy Chlorophyll Content Index* (CCCI) e *First Modified Chlorophyll Absorption Ratio Index* (MCARI1), mostraram-se ser índices bastante promissores não só para predição de produtividade mais também para predição da biometria EMC. Para futuros estudos de predição da cultura da cana de açúcar esses IV podem contribuir positivamente nos modelos de predição.

O uso de biometrias possibilitou melhoria no desempenho dos modelos na validação cruzada. Portanto, destaca-se a necessidade de maiores estudos com relação à obtenção de variáveis biométricas em campo com perspectiva de espacialização compatível com a escala dos dados de sensoriamento remoto orbital.

O teste do modelo baseado exclusivamente em dados de sensoriamento (cenário A) que apresentou os melhores resultados permitiu prever a produtividade com mais de 5 meses de antecedência da colheita, apresentando um RMSE de 8,15 ton.ha⁻¹

REFERÊNCIAS

- ABARCA, C. D. G. **Inovações tecnológicas na agroindústria da cana-de-açúcar no Brasil**. Disponível em http://www.agencia.cnptia.embrapa.br/Repositorio/ENEGEP1999_A0105_000fxgg417302wyiv80soht9h4yxjyhn.pdf. Acessado em 10 set. 2020.
- ALMEIDA, Pablo Roger Fernandes de. **Determinantes dos custos na produção de cana-de-açúcar no Brasil**. 2019. 21 f. Trabalho de Conclusão de Curso (Graduação em Ciências Contábeis) – Universidade Federal de Uberlândia, Uberlândia, 2019.
- ALVES, M. B. N.; PAIXÃO, A. E. A. **Indicadores de inovação em cultivares de cana-de-açúcar**. Proceeding of ISTI/SIMTEC –Aracaju/SE. v. 3, n. 1, 2016. Disponível em: <<http://www.api.org.br/conferences/index.php/ISTI2016/ISTI2016/paper/viewFile/81/76>>. Acesso em: 21 set. 2020.
- AMARAL, L. R.; PORTZ, G.; ROSA, H. J. A.; MOLIN, J. P. Estimativa de produtividade de cana-de-açúcar a partir de dados de sensor de refletância do dossel. In: **CONGRESSO BRASILEIRO DE AGRICULTURA DE PRECISÃO**, 5., 2012, Ribeirão Preto. Anais... Jaboticabal: SBEA, 2012.
- BARNES, E.M.; T.R. Clarke; S.E. Richards; P.D. Colaizzi; J. Haberland; M. Kostrzewski;. Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. In: **Proceedings of the 5th International Conference on Precision Agriculture**, 2000.
- Bernardi, A. D. C., Naime, J. D. M., Resende, A. D., Bassoi, L. H., & Inamasu, R. Y. (2014). **Agricultura de precisão: resultados de um novo olhar**. Embrapa Instrumentação-Livro técnico (INFOTECA-E).
- BIRTH, G. S.; MCVEY, G. Measuring the colour of growing turf with a reflectance spectrophotometer. **Agronomy Journal**, v. 60, n. 6, p. 640-643, 1968. <http://dx.doi.org/10.2134/agronj1968.00021962006000060016x>
- BRANDÃO, Z. N. **Estimativa da produtividade e estado nutricional da cultura do algodão irrigado via técnicas de sensoriamento remoto**. 2009. 152 f. Tese (Doutorado em Recursos Naturais)-Universidade Federal de Campina Grande, Campina Grande, 2009.
- BRASIL. Agricultura de Precisão. Ministério da Agricultura, Pecuária e Abastecimento, 2013. Disponível em: <https://www.gov.br/agricultura/pt-br/assuntos/sustentabilidade/tecnologia-agropecuaria/agricultura-de-precisao-1/arquivos-de-agricultura-de-precisao/boletim-tecnico-agricultura-de-precisao-2013.pdf> . Acesso: 15 set. 202
- Clevers, J. G. P. W. 1999. The use of imaging spectrometry for agricultural applications. **ISPRS J. Photogramm. Remote Sens.** 54: 299-304.
- CONAB. **Acompanhamento da safra brasileira de cana-de-açúcar - V. 7 - SAFRA 2020/21 N.2 - Segundo levantamento - Agosto 2020**

COX, Graeme J. **A yield mapping system for sugar cane chopper harvesters**. 2002. Tese de Doutorado. University of Southern Queensland.

CUNHA, J. V. A. & COELHO, A. C. (2009). Regressão Linear Múltipla. In Corrar, J. L., Paulo, E. & Dias Filho, J. M. (Coord.). **Análise multivariada: para os cursos de administração, ciências contábeis e economia. Fundação Instituto de Pesquisas Contábeis, Atuariais e Financeiras**. São Paulo: Atlas, 2009, p. 131-231.

Daros, E., OLIVEIRA, R. D., Zambon, J. L. C., & BESPALHOK FILHO, J. C. (2010). **Catálogo nacional de variedades “RB” de cana-de-açúcar**. Curitiba: RIDESA BRASIL.

DE SOUZA ASSAIANTE, Bianca Adriana; CAVICHIOLI, Fabio Alexandre. A UTILIZAÇÃO DE VEÍCULOS AÉREOS NÃO TRIPULADOS (VANT) NA CULTURA DA CANA-DE-AÇÚCAR. **Revista Interface Tecnológica**, v. 17, n. 1, p. 444-455, 2020.

Dias, F. L. F. et al. Produtividade da cana-de-açúcar em relação a clima e solos da região noroeste do estado de São Paulo. **Revista Brasileira de Ciência do Solo. Sociedade Brasileira de Ciência do Solo**, v. 23, n. 3, p. 627-634, 1999. Disponível em: <<http://hdl.handle.net/11449/114481>>.

ESA. European Space Agency . **Sentinel-2 User Handbook**. Disponível em: https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_User_Handbook. Acesso: 15 dez. 2020.

GASCHO, G. J.; SHIH, S. F. Sugarcane. In: TEARE, I. D.; PEET, M. M. (Ed.). **Crop-water relations**. New York: Wiley-Interscience, 1983. p. 445-479.

GITELSON, A. A.; GRITZ, Y.; MERZLYAK, M. N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. **Journal of Plant Physiology**, v. 160, n. 3, p. 271-282, 2003.

GONÇALVES, Jose Roberto Moreira Ribeiro. **Análise Ambiental e Econômica Comparativa de Métodos de Amostragem do Solo Utilizadas em Agricultura de Precisão**. 2017. 93 f. Dissertação (Mestrado em Ciências, Engenharia Agrícola e Ambiental). Instituto de Tecnologia, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ, 2017.

GREGO, C. R., de ARAUJO, L. S., Vicente, L. E., Nogueira, S. F., Magalhães, P. S. G., Vicente, A. K., ... & Bolfe, E. L. (2014). **Agricultura de precisão em cana-de-açúcar**. Embrapa Territorial-Capítulo em livro científico (ALICE).

HABOUDANE, D.; MILLER, J. R.; PATTEY E.; ZARCOTEJADA, P. J.; STRACHAN, I. B. 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**, v. 90, n. 3, p.337-352, 2004. <http://dx.doi.org/10.1016/j.rse.2003.12.013>

HUETE, A. R. A soil adjusted vegetation index (SAVI). **Remote Sensing of Environment**, v. 25, n. 3, p. 295-309, 1988. [http://dx.doi.org/10.1016/0034-4257\(88\)90106-X](http://dx.doi.org/10.1016/0034-4257(88)90106-X)

HUETE, A. R.; LIU, H. Q.; BATCHILY, K.; VAN LEEUWEN, W. A comparison of vegetation indices global set of TM images for EOS-MODIS. **Remote Sensing of Environment**, v. 59, n. 3, p. 440-451, 1997. [http://dx.doi.org/10.1016/S0034-4257\(96\)00112-5](http://dx.doi.org/10.1016/S0034-4257(96)00112-5).

Jadoski, C. J., Toppa, E. V. B., Julianetti, A., Hulshof, T., Ono, E. O., & Rodrigues, J. D. (2011). Fisiologia do desenvolvimento do estágio vegetativo da cana-de-açúcar (*Saccharum officinarum* L.). *Pesquisa Aplicada & Agrotecnologia*, 169-175.

JENSEN, John R. **Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres**. São José dos Campos: Parêntese Editora, 2009.

KRUTMAN, Sarah. Do grande período de variedades de cana. **Pesquisa Agropecuária Brasileira**, v. 1, n. 1, p. 395-400, 1966.

LANKRIET, Edouard; POPPE, Marcelo. Sistema de Inovação Tecnológica da agroenergia da cana-de-açúcar no Brasil—da sua gênese à transição agroecológica atual. **Parcerias Estratégicas**, v. 23, n. 46, p. 49-88, 2019.

LONG, D. S., EITEL, J. U. H., and HUGGINS, D. R. Assessing nitrogen status of dryland wheat using the canopy chlorophyll content index. **Crop Management**. Online. doi:10.1094/CM-2009-1211-01-RS.

MALDANER, Leonardo Felipe. **Processamento de dados de monitores de produtividade de cana-de-açúcar**. Tese de Doutorado. Universidade de São Paulo.

MARTINS, A. L. M.; LANDELL, M. G. A. Conceitos e critérios para avaliação experimental em cana-de-açúcar utilizados no programa Cana IAC. **Pindorama: Instituto Agrônomo**, 1995, p. 2-14.

MENEGATTI, Leonardo A. Angeli; MOLIN, José P. Metodologia para identificação e caracterização de erros em mapas de produtividade. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 7, n. 2, p. 367-374, 2003.

MIOCQUE, J. Avaliação de crescimento e de produtividade de matéria verde da cana-de-açúcar na região de Araraquara SP. **Stab**, v. 17, p. 45-47, 1999.

MOLIN, J. P.; AMARAL, L. R.; COLAÇO, A. **Agricultura de precisão**. São Paulo: Oficina de Textos, 2015

MOLIN, José P. Geração e interpretação de mapas de produtividade para agricultura de precisão. BORÉM, A.; GIÚDICE, MP; QUEIROZ, DM; MANTOVANI, EC, p. 237-58, 2000.

MORAES, Rafael Aldighieri . **Monitoramento e estimativa da produção da cultura de cana-de-açúcar no estado de São Paulo por meio de dados espectrais e agrometeorológicos**. 2012.

MUNDIM, João Umbirucu Campos. Uso de simulação de eventos discretos para o dimensionamento de frota para colheita e transporte de cana-de-açúcar. 2009. **Tese de Doutorado**. Universidade de São Paulo.

OLIVEIRA, Gildriano Soares de. **Produtividade de biomassa de cana-de-açúcar em função dos índices de vegetação utilizando técnicas de sensoriamento remoto**. 2015. vii, 65 p. Dissertação (mestrado) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal, 2015. Disponível em: <<http://hdl.handle.net/11449/136733>>. Acesso: 03 set. 2020

OLIVEIRA, R. A., SANTOS, R. S. D., RIBEIRO, A., Zolnier, S., & Barbosa, M. H. (2012). Estimativa da produtividade da cana-de-açúcar para as principais regiões produtoras de Minas Gerais usando-se o método ZAE I. **Revista Brasileira de Engenharia Agrícola e Ambiental**, 16(5), 549-557.

PEÑUELAS, J.; BARET, F.; FILELLA, I. Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. **Photosynthetica** v. 31, p. 221-230, 1995.

PICOLI, Michelle Cristina Araújo. **Estimativa da produtividade agrícola da cana-de-açúcar utilizando agregados de redes neurais artificiais**: estudo de caso usina Catanduva. São José dos Campos: INPE, 2006.

PONZONI, F.J.; SHIMABUKURO, Y.E.; KUPLICH, T.M. **Sensoriamento Remoto da Vegetação**. 2. ed. São Paulo: Oficina de Textos, 2012. v. 1, 176p.

PROENÇA, E. R.; Tarsitano, M. A. A.; Bolonhezi, A. C.; Rapassi, R. M. A.; Costa, S. M. A. L. (2009). Inovações Tecnológicas adotadas por Usinas do Setor Sucroalcooleiro na regional de Andradina, SP. **Sociedade Brasileira de Economia, Administração e Sociologia Rural - SOBER, 47º Congresso**, Porto Alegre.

QI, J.; CHEHBOUNI, A.; HUETE A.R.; KERR, Y. H.; SOROOSHIAN, S. A modified soil adjusted vegetation index. **Remote Sensing Environmental**, v. 48, p. 119-126, 1994.

RESENDE, AV; COELHO, A (2017) Amostragem para mapeamento e manejo da fertilidade do solo na abordagem de agricultura de precisão. **International Plant Nutrition Institute (IPNI): Informações Agronômicas**, 156: 1-8. Disponível em: <http://www.ipni.net/publication/ia-brasil.nsf> . Acesso em 10 set. 2020

ROCHA, Murillo Grespan. **Influência de variáveis agronômicas na produtividade de cana-de-açúcar e potencial uso de sensoriamento proximal**. 2018. 1 recurso online (93 p.). Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Agrícola, Campinas, SP. Disponível em: <<http://www.repositorio.unicamp.br/handle/REPOSIP/331897>>. Acesso em: 05 set. 2020.

RODRIGUES, Gelze Serrat de Souza Campos; ROSS, Jurandyr Luciano Sanches. **A trajetória da cana-de-açúcar no Brasil**: perspectivas geográfica, histórica e ambiental. 2020.

RONDEAUX, G.; STEVEN, M.; BARET, F. Optimization of soiladjusted vegetation indices. **Remote Sensing of Environment**, v. 55, n. 2, p. 95-107, 1996.

ROSA, R. **Introdução ao Sensoriamento Remoto**. 7.ed. Uberlândia: EDUFU, 2009. 48 p

ROUJEAN, J. L.; BREON, F. M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. **Remote Sensing of Environment**, v. 51, n. 3, p. 375-384, 1995

ROUSE, J. W.; HAAS, R. H.; SCHELL, J. A.; DEERING, D. W.; HARLAN, J. C. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. Texas: Texas A&M University, 1974. (Type III final rep. NASA/GSFC).

SANTI, A. L., Giotto, E., Sebem, E., & Amado, T. J. C. (2016). Agricultura de Precisão no Rio Grande do Sul. Santa Maria: CESPOL.

SILVA, B. B. D., GALVÍNCIO, J. D., MONTENEGRO, S. M. G. L., MACHADO, C. C. C., OLIVEIRA, L. M. M. D., & Moura, M. S. B. D. (2013). Determinação por sensoriamento remoto da produtividade primária bruta do perímetro irrigado São Gonçalo-PB. **Revista Brasileira de Meteorologia**, 28(1), 57-64.

SILVA, M. D. A., GAVA, G. J. D. C., CAPUTO, M. M., PINCELLI, R. P., JERÔNIMO, E. M., & CRUZ, J. C. S. (2007). **Uso de reguladores de crescimento como potencializadores do perfilhamento e da produtividade em cana-soca**. *Bragantia*, 66(4), 545-552.

Sugawara, L. M., Rudorff, B. F. T., Freitas, C. C., Picoli, M. C. A., & Adami, M. (2007). Estimativa de produtividade de cana-de-açúcar (*Saccharum officinarum* L.) por meio de técnica de análise de regressão linear múltipla. **Simpósio Brasileiro de Sensoriamento Remoto**, 13, 435-442.

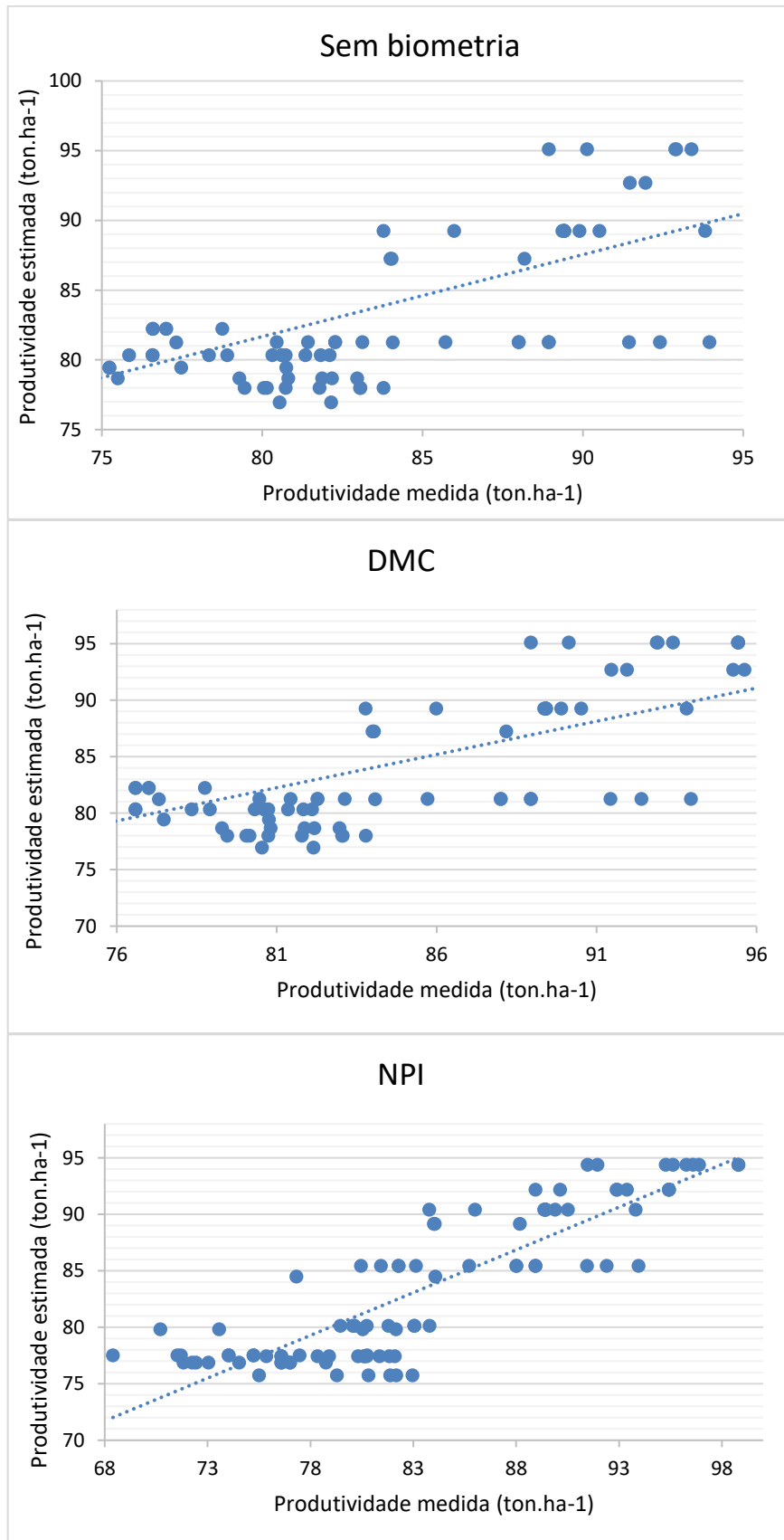
SUGUITANI, C. **Entendendo o crescimento e produção da cana-de-açúcar: avaliação do modelo Mosicas**. 2006. 60 p. 2006. Tese de Doutorado. Tese (Doutorado em Fitotecnia)– Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba.

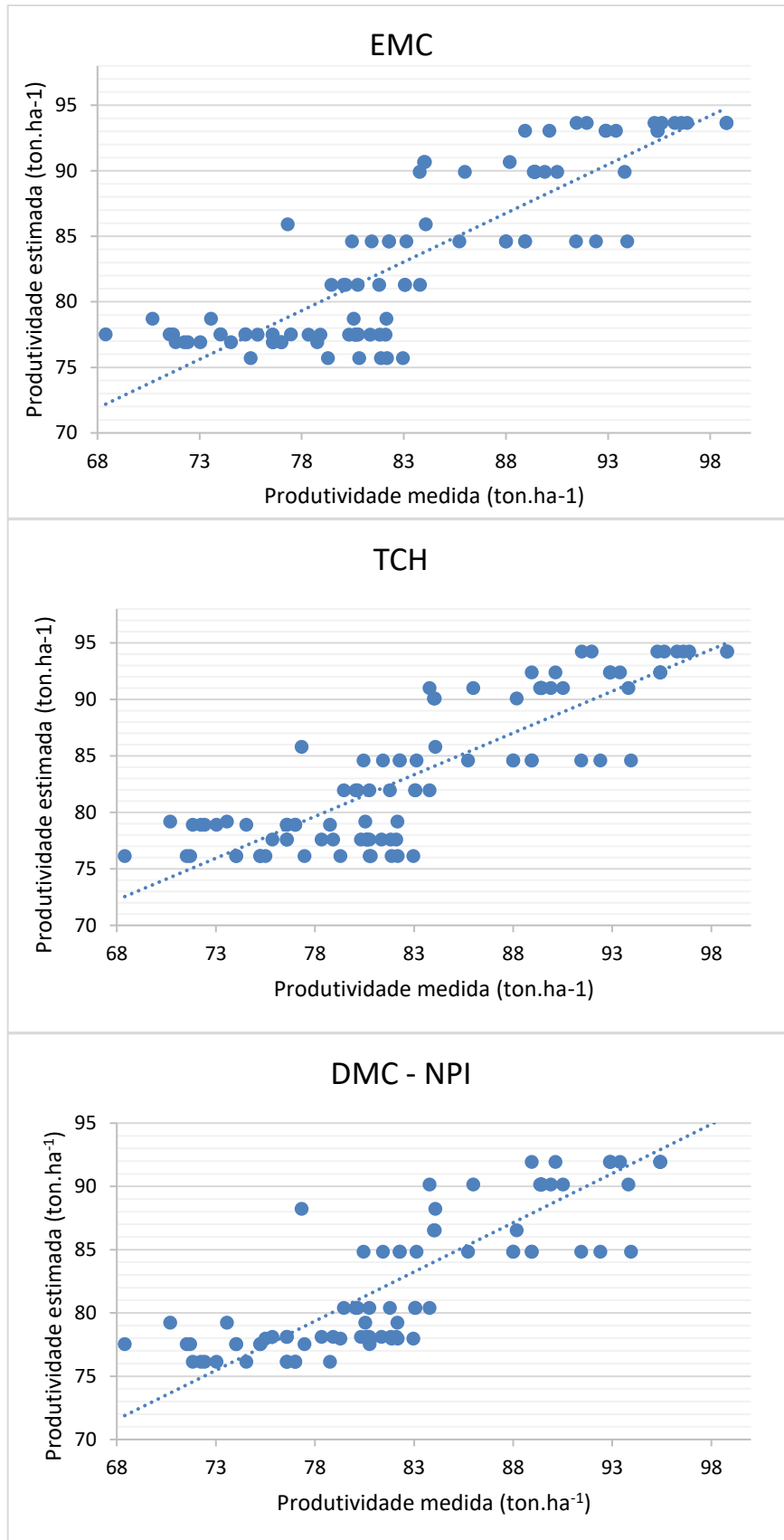
THOMAS, André Luís. **Desenvolvimento da Planta de Cana-de-açúcar. Desenvolvimento das plantas de batata, mandioca, fumo e cana-de-açúcar** [recurso eletrônico]. Porto Alegre: UFRGS, 2016. p. 55-75, 2016.

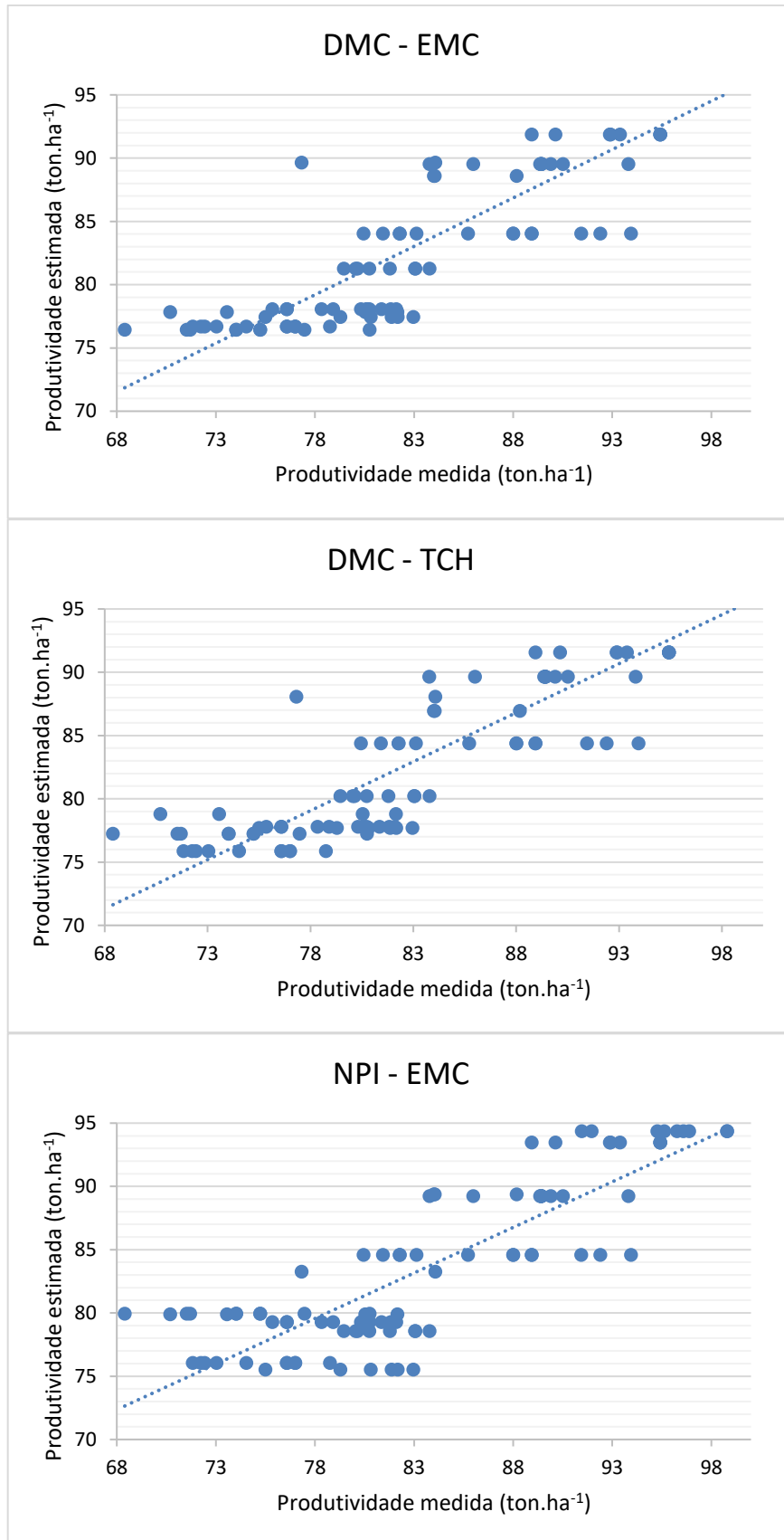
TUCKER, C. J. Red and photographic infrared linear combinations for monitoring vegetation. **Remote Sensing of Environment**, v. 8, n. 2, p. 127-150, 1979.
[http://dx.doi.org/10.1016/0034-4257\(79\)90013-0](http://dx.doi.org/10.1016/0034-4257(79)90013-0)

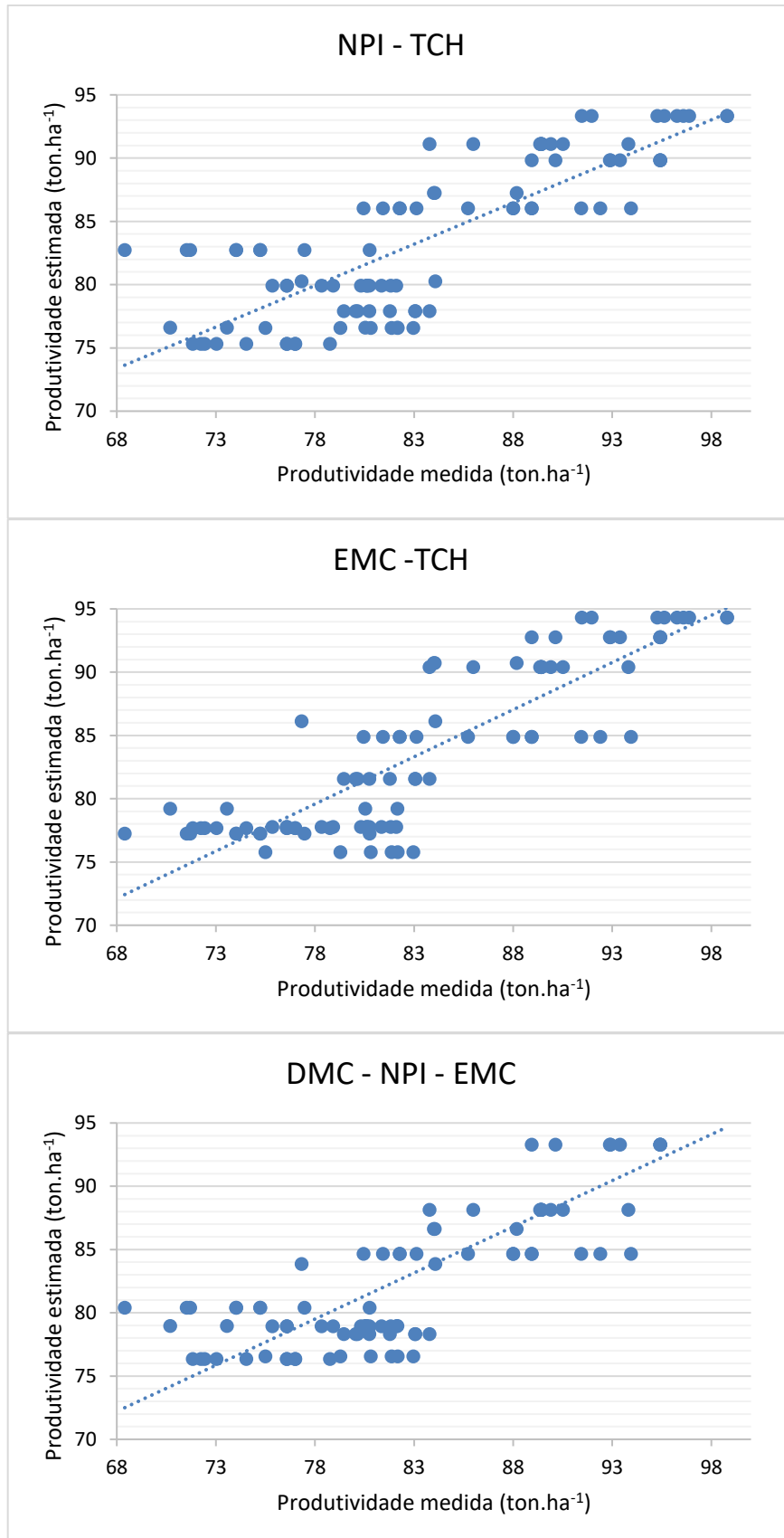
UNICA. União da Indústria de Cana-de-Açúcar. Balanços de atividades 2012/13 a 2018/10. Disponível em: <https://www.unica.com.br/wp-content/uploads/2019/06/Relatorio-Atividades-201213-a-201819.pdf> . Acesso: 18 set. 2020

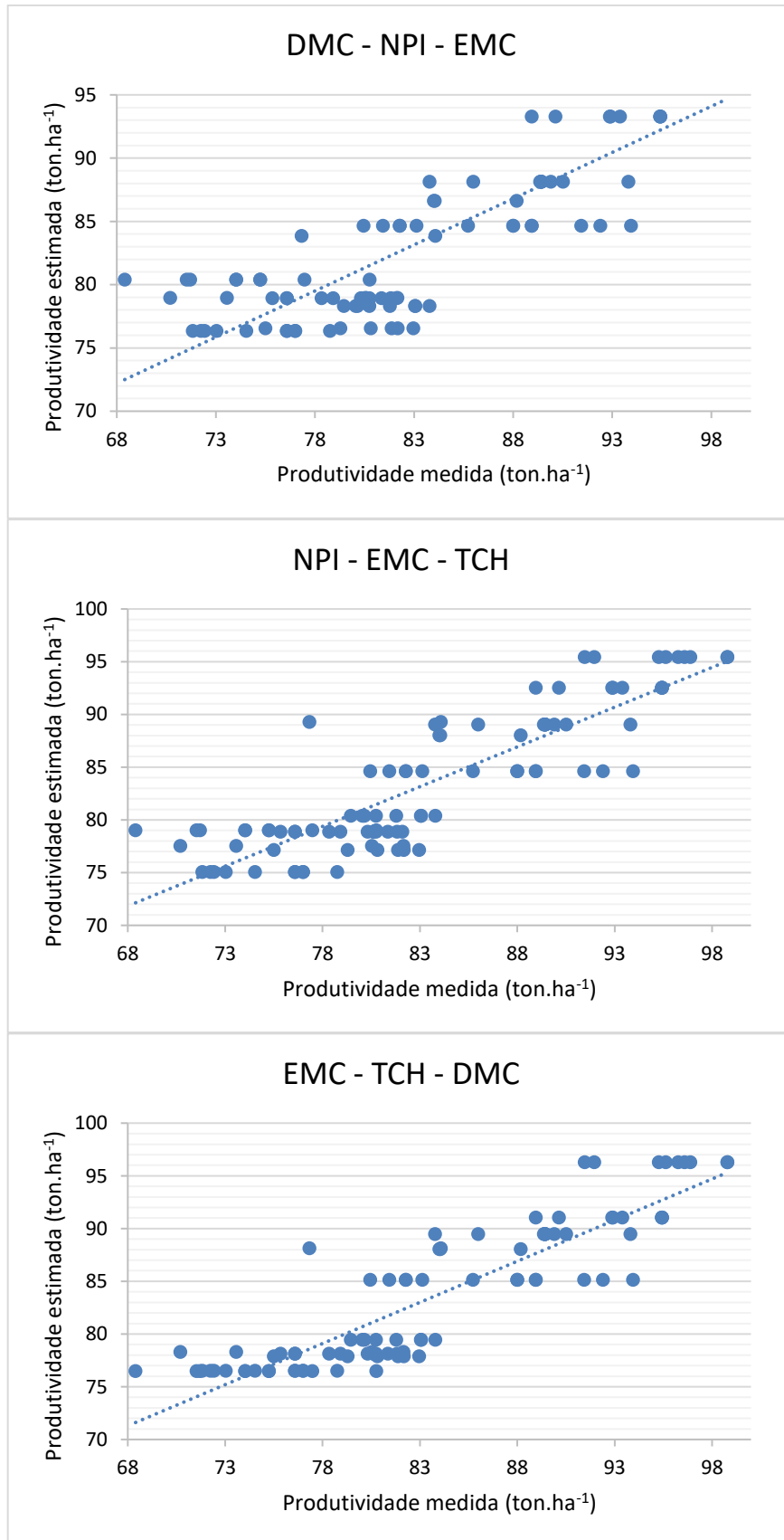
7. APÊNDICE

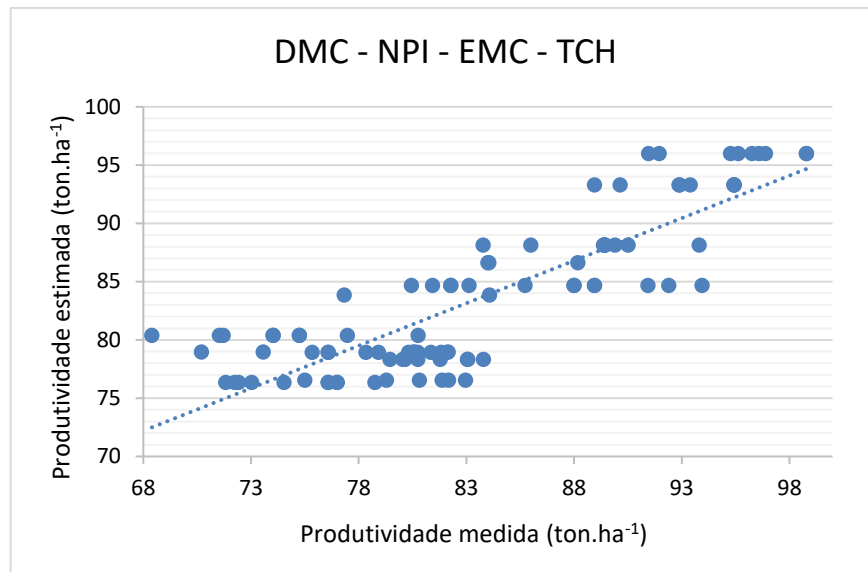












Apêndice 1 - Gráfico de dispersão do modelos - Estimada x Medida. Fonte: Autor (2021).