

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**MELHORIA DO TRATAMENTO DE
OBSTÁCULOS NA ABORDAGEM DE
AGRUPAMENTO DE DADOS ESPACIAIS
SWMU CLUSTERING**

GABRIEL PASSATUTO GALLO

ORIENTADOR: PROF. DR. RICARDO RODRIGUES CIFERRI

São Carlos – SP

20 de agosto de 2020

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**MELHORIA DO TRATAMENTO DE
OBSTÁCULOS NA ABORDAGEM DE
AGRUPAMENTO DE DADOS ESPACIAIS
SWMU CLUSTERING**

GABRIEL PASSATUTO GALLO

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação, área de concentração: Engenharia de Software, Banco de Dados e Interação Humano-Computador (ES-BD-IHC).

Orientador: Prof. Dr. Ricardo Rodrigues Ciferri

São Carlos – SP

20 de agosto de 2020

Gabriel Passatuto Gallo

Melhoria do Tratamento de Obstáculos na Abordagem de Agrupamento de Dados Espaciais SWMU Clustering/ Gabriel Passatuto Gallo. – São Carlos – SP, 20 de agosto de 2020-

72 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Ricardo Rodrigues Ciferri

– Universidade Federal de São Carlos, 20 de agosto de 2020.

1. Dados Espaciais. 2. Agrupamento de Dados. 3. Agricultura de Precisão. I. Prof. Dr. Ricardo Rodrigues Ciferri. II. Universidade Federal de São Carlos. III. Programa de Pós-Graduação em Ciência da Computação, Câmpus São Carlos. IV. Melhorando o Tratamento de Obstáculos na Abordagem de Agrupamento de Dados Espaciais SWMU Clustering.



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência da Computação

Folha de Aprovação

Defesa de Dissertação de Mestrado do candidato Gabriel Passatuto Gallo, realizada em 20/08/2020.

Comissão Julgadora:

Prof. Dr. Ricardo Rodrigues Ciferri (UFSCar)

Prof. Dr. Renato Bueno (UFSCar)

Prof. Dr. Luis Henrique Bassoi (EMBRAPA)

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

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciência da Computação.

Gabriel Passatuto Gallo

Melhoria do Tratamento de Obstáculos na Abordagem de Agrupamento de Dados Espaciais SWMU Clustering

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação, área de concentração: Engenharia de Software, Banco de Dados e Interação Humano-Computador (ES-BD-IHC).

São Carlos – SP, 20 de agosto de 2020:

Prof. Dr. Ricardo Rodrigues Ciferri
(UFSCar/DC)
Orientador

Prof. Dr. Renato Bueno (UFSCar/DC)

**Dr. Luis Henrique Bassoi (Embrapa
Instrumentação Agropecuária)**

São Carlos – SP
20 de agosto de 2020

Dedico esse trabalho à memória de minha querida avó.

AGRADECIMENTOS

Agradeço a Jesus por me proporcionar a saúde e a esperança necessária para superar todos os desafios da minha jornada. Agradeço aos meus pais, Alfredo e Nilza, por sempre estarem ao meu lado, motivando e amparando independente das circunstâncias vivenciadas, pela educação que me proporcionaram, pelos princípios que me ensinaram e pelo exemplo de vida que me inspira a cada dia. Agradeço a minha companheira de jornada, Lorena, por todo amor, paciência e carinho que sempre me dedicou. Agradeço também, ao Professor Ricardo, por ter aceitado a orientação dessa pesquisa, e durante todos esses anos me apoiando e motivando em todos os desafios, e ao Eduardo Speranza pelo auxílio na orientação e por toda dedicação a essa pesquisa. Por fim, agradeço a CAPES por apoiar essa pesquisa de Mestrado.

Só haverá Paz, quando a Fraternidade habitar os corações.
(José Simões de Paiva Netto)

RESUMO

O cenário tecnológico vem se ampliando consideravelmente nos últimos anos, trazendo os mais diversos benefícios para inúmeras áreas de aplicação. Dentre essas áreas, a agricultura teve um grande impulso, conseguindo elevar a sua produção e ao mesmo tempo reduzir custos e impactos ambientais por intermédio de técnicas de gestão da lavoura, praticando assim, os conceitos de Agricultura de Precisão (AP). Um dos métodos usados em AP é o delineamento da área plantada em parcelas menores com valores semelhantes de atributos de solo e de planta, conhecido como zonas de manejo ou unidades de gestão diferenciada (UGDs). Desse modo, o agrupamento de dados espaciais é usado para criação de mapas de UGDs, nos quais retratam a variabilidade do solo. *Spatial Ward's Management Units Clustering* (SWMU Clustering) é uma abordagem de agrupamentos de dados espaciais que possibilita o delineamento de UGDs em AP. A sua principal vantagem com relação a outras abordagens correlatas é a redução significativa da estratificação nos agrupamentos, obtendo-se mapas de UGDs facilmente interpretáveis pelo usuário final. Tendo isso exposto, esta pesquisa de Mestrado investigou como melhorar o gerenciamento de obstáculos espaciais realizado pela abordagem *SWMU Clustering*. Nesse sentido, foram propostas duas novas estratégias: Estratégia de Substituição do Conjunto de Amostras Internas aos Obstáculos e Estratégia de *Buffer*. Essas estratégias foram comparadas com a estratégia original da abordagem *SWMU Clustering*, mostrando que a estratégia de *Buffer* gerou os melhores resultados. Além disso, como resultado desta pesquisa foi desenvolvida uma aplicação *web* para a abordagem *SWMU Clustering*, disponibilizando-a como um serviço para que o usuário final possa interagir e usar a abordagem, desde o envio dos dados de entrada até a visualização dos resultados de delineamento de UGDs.

Palavras-chave: Banco de Dados, Dados Espaciais, Mineração de Dados, Agrupamento de Dados, Agricultura de Precisão, Zonas de Manejo.

ABSTRACT

The technological has been improved considerably in recent years, providing the great benefits to several areas of application. Among these areas, agriculture had a great boost, enabling the increasing of the production and at the same time reducing costs and environmental impacts through crop management techniques, thus in this sense practicing the concepts of Precision Agriculture (AP). One of the methods used in PA is to design the planted area in smaller plots with similar values of soil and plant attributes, known as management zones or differentiated management units (UGDs). In this way, spatial data clustering algorithms are used to create UGD maps, in which they depict soil variability. Spatial Ward's Management Units Clustering (SWMU Clustering) is an approach to spatial data clustering that enables the design of UGDs in AP. Its main advantage over related approaches is the significant reduction of stratification in clusters, obtaining maps of UGDs that are easily interpretable by the end user. This Master's research investigated how to improve the management of spatial obstacles performed by the SWMU Clustering approach. In this sense, two new strategies were proposed: Replacement Strategy for the Set of Internal Samples to Obstacles and Buffer Strategy. These strategies were compared to the original strategy of the SWMU Clustering approach, showing that the Buffer strategy generated the best results. In addition, as a result of this research, an web application was developed for the SWMU Clustering approach, making it available as a service so that the end user can interact with the SWMU Clustering ap, from sending the input data until the visualization of the UGD results.

Keywords: Database, Spatial Data, Data Mining, Data Clustering, Precision Agriculture, Management Zones.

LISTA DE SIGLAS

UGD - Unidade de Gestão Diferenciada

CGD - Classe de Gestão Diferenciada

SIG - Sistema de Informações Geográficas

TIC - Tecnologia de Informação e Comunicação

LISTA DE FIGURAS

Figura 1 – Tipos de dados espaciais básicos.	21
Figura 2 – Tipos de dados espaciais complexos.	21
Figura 3 – Conjunto de dados espaciais representados por pontos no formato matricial.	22
Figura 4 – Ciclo da AP considerando a utilização de TICs.	27
Figura 5 – Mapa de CGD 1	29
Figura 6 – Mapa de CGD 2	30
Figura 7 – Mapa UGD com estratificação	31
Figura 8 – <i>Interface Web SWMU Clustering</i>	33
Figura 9 – Níveis de estratificação; (a) mapa gerado por uma abordagem tradicional; e (b) mapa gerado com a abordagem <i>SWMU Clustering</i>	35
Figura 10 – Mapa com obstáculo espacial de declividade com detalhamento	37
Figura 11 – Obstáculo Espacial Árvore Nativa.	40
Figura 12 – Obstáculo Espacial Mata Nativa.	40
Figura 13 – Obstáculo Espacial Reservatório de Água no Solo.	41
Figura 14 – Obstáculo Espacial Construção.	42
Figura 15 – Obstáculo Espacial Reservatório de Água.	43
Figura 16 – Obstáculo Espacial Estrada.	43
Figura 17 – Obstáculo Espacial Rede Elétrica.	44
Figura 18 – Obstáculo Espacial Poste.	45
Figura 19 – Obstaculo Espacial Torre e Rede Elétrica Alta Tensão.	45
Figura 20 – Obstáculo Espacial Pivô.	46
Figura 21 – Obstáculo Espacial Equipamentos de Irrigação.	46
Figura 22 – Obstáculo Espacial Rio.	47
Figura 23 – Obstáculo Espacial Represa.	47
Figura 24 – Etapas de Tratamento dos Dados.	51
Figura 25 – Estratégia de <i>Buffer</i>	52
Figura 26 – Obstáculo identificado com imagem Google Mapas	54
Figura 27 – Mapas de UGDs gerados pela abordagem <i>SWMU Clustering Original</i>	54
Figura 28 – Mapas de UGDs gerados pela abordagem <i>SWMU Clustering</i> usando a estratégia de substituição do conjunto de amostras internas aos obstáculos.	55

Figura 29 – Mapas de UGDs gerados pela abordagem SWMU <i>Clustering</i> usando a estratégia de <i>buffer</i>	55
Figura 30 – Validação das estratégias propostas com o Critério SD-Spatial.	56
Figura 31 – <i>Interface Web SWMU Clustering</i>	58
Figura 32 – Validação protótipo - mapa contendo 4 CGDs com porcDMU=5% e varTess 0,3%	59
Figura 33 – Validação protótipo - mapa contendo 4 CGDs com porcDMU=5% e varTess 1%.	60
Figura 34 – Validação protótipo - mapa contendo 4 CGDs com porcDMU=5% e varTess 2,5%.	60
Figura 35 – Validação protótipo - mapa contendo 5 CGDs com porcDMU=5%, minDMC=2%, maxDMC = 5% e varTess 0,3%.	61
Figura 36 – Validação protótipo - mapa contendo 5 CGDs com porcDMU=5%, minDMC=2%, maxDMC = 5% e varTess 0,3%.	61
Figura 37 – Validação protótipo - mapa contendo 5 CGDs com porcDMU=10%, minDMC=2%, maxDMC = 5% e varTess 0,3%.	62
Figura 38 – Validação protótipo - mapa contendo 5 CGDs com com porcDMU=5%, varTess=5%.	62
Figura 39 – Validação protótipo - mapa contendo 4 CGDs.	63
Figura 40 – Validação protótipo - mapa contendo 4 CGDs.	63
Figura 41 – Validação protótipo - mapa contendo 4 CGDs.	64
Figura 42 – Validação protótipo - mapa contendo 3 CGDs.	64
Figura 43 – Validação protótipo - mapa contendo 4 CGDs.	65
Figura 44 – Validação protótipo - mapa contendo 3 CGDs.	65
Figura 45 – Validação protótipo - mapa contendo 5 CGDs.	66

LISTA DE TABELAS

Tabela 1 – Classificação dos obstáculos	48
Tabela 2 – Tipos de dados espaciais dos obstáculos	49

SUMÁRIO

CAPÍTULO 1–INTRODUÇÃO	15
1.1 Contextualização	15
1.2 Motivação	16
1.3 Objetivo	17
1.4 Metodologia	18
1.5 Estrutura do Trabalho	18
CAPÍTULO 2–FUNDAMENTAÇÃO TEÓRICA	20
2.1 Dados Espaciais	20
2.1.1 Banco de Dados Espaciais	21
2.1.2 Sistemas de Informações Geográficas	23
2.1.3 Dados Espaciais via <i>Web</i>	24
2.1.3.1 Especificações OGC	24
2.2 Agricultura de Precisão	26
2.2.1 Estratégias de Aplicação e Construção de Mapas em AP	27
2.2.2 Unidades de Gestão Diferenciada (UGDs)	28
2.2.3 Tipos de Dados para Geração de Mapas em AP	31
2.3 R <i>Shiny</i>	33
CAPÍTULO 3–TRABALHOS RELACIONADOS	34
3.1 Considerações Iniciais	34
3.2 Abordagem SWMU <i>Clustering</i>	34
3.3 Critério de Validação Usado na Abordagem SWMU <i>Clustering</i>	37
3.4 Considerações Finais	37
CAPÍTULO 4–OBSTÁCULOS ESPACIAIS	39
4.1 Considerações Iniciais	39
4.2 Descoberta dos Obstáculos Espaciais	39
4.3 Classificação dos Obstáculos Investigados	48
CAPÍTULO 5–ESTRATÉGIAS DE TRATAMENTOS DE OBSTÁCULOS	50
5.1 Considerações Iniciais	50
5.2 Estratégia da Substituição	50
5.3 Estratégia de <i>Buffer</i>	52
5.4 Validação das Estratégias	53

CAPÍTULO 6–PROTÓTIPO DO SERVIÇO <i>Web</i>	57
6.1 Considerações Iniciais	57
6.2 Desenvolvimento e Tecnologias	57
6.3 Validação do Protótipo	59
6.4 Considerações Finais	66
CAPÍTULO 7–CONCLUSÕES E TRABALHOS FUTUROS	67
7.1 Considerações Iniciais	67
7.2 Contribuições	67
7.3 Trabalhos Futuros	68
7.4 Considerações Finais	69
REFERÊNCIAS	70

Capítulo 1

INTRODUÇÃO

Este capítulo introdutório está organizado da seguinte forma: primeiramente, o contexto que o trabalho está inserido conjuntamente com as questões de pesquisa e as hipóteses; em seguida, as justificativas que motivaram o desenvolvimento desta pesquisa de Mestrado são descritas; logo após, os objetivos pretendidos são abordados; na sequência, é apresentado a metodologia adotada; e por fim, a estrutura do trabalho é detalhada.

1.1 Contextualização

O cenário tecnológico vem crescendo e melhorando consideravelmente nos últimos anos possibilitando inúmeros benefícios para as mais diversas áreas de aplicação. Dentre essas áreas, a agricultura teve um grande impulso, podendo elevar a sua produção e ao mesmo tempo reduzir custos e impactos ambientais, prática essa que pode ser considerada como Agricultura de Precisão (AP). AP é definida como um conjunto de técnicas que gerenciam a variabilidade espacial da lavoura para obter retorno econômico e ambiental (BERNARDI *et al.*, 2014). Para (SPERANZA, 2017), a agricultura tradicional realiza a aplicação de insumos de maneira homogênea em toda a área de cultivo, ou seja, não existe um aplicação a taxa variável em função das características do solo e cultura, enquanto que em AP os insumos são aplicados em taxas condizentes com as características de solo e cultura. Também se pode definir AP como um sistema de gerenciamento agrícola, que considera a variabilidade espacial de solo e cultura, com o objetivo de aumentar a produtividade e diminuir impactos ambientais (BERNARDI *et al.*, 2014; SPERANZA, 2017; MOLIN, 2003). Ruß (2012) define AP como um área de pesquisa que se baseia na coleta de grandes volume de dados com objetivo de auxiliarem nas estratégias de tomadas de decisão quanto as operações agrícolas, tendo em foco a variabilidade espacial.

Dessa forma, o uso de tecnologias e procedimentos relacionados à automatização e otimização das culturas com base na variabilidade espacial é fundamental (MOLIN, 2003). Com o acesso facilitado às tecnologias de informação e comunicação (TICs), como GPS, sensores, câmeras que equipam *drones* e satélites, foi viabilizado a criação de muitos sistemas e ferramentas para auxiliar o plantio, a colheita e todo o gerenciamento de uma lavoura (Silva *et al.*, 2014;

[SPERANZA, 2017](#))

Uma dessas soluções é o algoritmo *Spatial Ward's Management Units Clustering (SWMU Clustering)* ([SPERANZA, 2017](#)) que foi desenvolvido especificamente para a tarefa de delimitação de Zonas de Manejo, também conhecidas como Unidades de Gestão Diferenciadas (UGDs). Sua principal vantagem é permitir a obtenção de mapas pouco estratificados e condizentes com a realidade do solo, tornando-os mais claros, legíveis e com possibilidade de uso pelo usuário final no campo. Apesar de sua eficácia em gerar mapas aplicáveis em AP, existem pontos que podem ser melhorados no algoritmo *SWMU Clustering* visando aumentar a sua eficácia no tratamento de áreas de manejo com obstáculos espaciais.

O objetivo principal desta pesquisa de Mestrado, portanto, foi investigar como melhorar o gerenciamento de obstáculos espaciais realizado pela abordagem *SWMU Clustering*. Dado o contexto apresentado, as questões de pesquisa são:

QP1 - "Quais os obstáculos espaciais presentes na agricultura que podem influenciar a eficácia de algoritmos de delimitação de UGDs em AP?".

QP2 - "Como tratar adequadamente os obstáculos espaciais identificados na questão QP1 no algoritmo *SWMU Clustering*?".

Assim, foram investigadas as seguintes hipóteses para que essas questões de pesquisa sejam validadas:

H₁: Diferentes tipos de obstáculos espaciais presentes no contexto de agricultura de precisão necessitam de tratamento específico e diferenciado por algoritmos de delimitação de UGDs.

H₂: O tratamento de diferentes tipos de obstáculos espaciais presentes no contexto de agricultura de precisão afetam de forma distinta a eficácia de algoritmos de delimitação de UGDs.

H₃: Os tipos de obstáculos descobertos requerem adaptações no algoritmo *SWMU Clustering* de forma a prover uma boa eficácia no resultado final composto por UGDs.

1.2 Motivação

O volume de dados espaciais vem aumentando nos últimos anos, oriundos das mais diversas fontes. Com isso, a dificuldade de analisá-los também acompanha esse crescimento ([Silva et al., 2014](#)) ([SPERANZA, 2017](#)).

Muitas pesquisas na área da Computação voltadas ao contexto de AP foram realizadas nos últimos anos, dentre essas pesquisas, a abordagem de agrupamento de dados espaciais *SWMU Clustering* proposta por [Speranza \(2017\)](#) para auxiliar no delimitação de UGDs em AP. Essa abordagem é composta por restrições espaciais, tendo como principal objetivo produzir

mapas de Unidades de Gestão Diferenciadas (UGDs) com pouca estratificação. Dentre essas restrições, os obstáculos espaciais foram explorados neste trabalho. Nesse contexto, esse trabalho aprofunda a investigação no quesito restrição espacial de obstáculos, avançando nas soluções propostas originalmente na tese de Doutorado de [Speranza \(2017\)](#) com relação ao tratamento de obstáculos no delineamento de UGDs.

O algoritmo *SWMU Clustering* trata de maneira superficial a questão de obstáculos, desconsiderando as variáveis de espaço de atributos, calculando a distância euclidiana para cada par de amostras, o que torna o processo lento e genérico, considerando que existam uma diversidade de obstáculos presentes no meio agrícola. Portanto, há necessidade de um tratamento mais elaborado para todos os tipos de obstáculos presentes no campo e desta forma permitir uma melhor eficácia do resultado gerado pelo algoritmo *SWMU Clustering* em sua solução final. Em particular, espera-se que os valores usados na representação dos obstáculos sejam calculados de forma mais acurada para que estes representem uma melhor estimativa com relação aos valores usados pelos vizinhos mais próximos aos obstáculos que possam influenciá-los de fato.

1.3 Objetivo

Essa pesquisa objetivou melhorar a abordagem *SWMU Clustering* especificamente quanto ao tratamento de obstáculos espaciais relacionados ao delineamento de UGDs para AP. Além disso, essa pesquisa objetiva desenvolver uma aplicação orientada a serviços para a *web*.

Dados os objetivos gerais, com as investigações feitas, pretende-se atingir os seguintes objetivos específicos:

1. Identificar os possíveis tipos de obstáculo espacial presentes no contexto de AP que possam afetar a tarefa de delineamento de UGDs.
2. Propor soluções para o tratamento eficiente de cada tipo de obstáculo espacial para a abordagem *SWMU Clustering*.
3. Implementar no algoritmo *SWMU Clustering* as soluções propostas.
4. Realizar testes experimentais para analisar a eficácia das soluções propostas que foram incorporadas ao algoritmo *SWMU Clustering*.
5. Desenvolver um protótipo *web* para a abordagem *SWMU Clustering*, disponibilizando em forma de serviço.
6. Validar o protótipo do serviço *Web SWMU Clustering*.

1.4 Metodologia

Para que os objetivos propostos fossem atingidos, a metodologia adotada incluiu constantes estudos aprofundados sobre a abordagem SWMU *Clustering* proposta por Speranza (2017), trabalho esse, fundamental para essa pesquisa de Mestrado que visa incrementar técnicas de tratamentos de obstáculo na abordagem. A fim de descobrir obstáculos que existem no contexto agrícola, e que possam interferir diretamente ou indiretamente na eficácia da abordagem SWMU *Clustering*, ocorreram visitas a áreas agrícolas na região de Vista Alegre do Alto -SP, visando investigar e identificar esses obstáculos em um contexto real.

Reuniões periódicas com o orientador Professor Dr. Ricardo Rodrigues Ciferri, foram fundamentais para o reconhecimento e classificação dos obstáculos investigados, assim como participação nas reuniões periódicas do Grupo de Banco de Dados (GBD) da UFSCar (GBD, 2019).

Após a investigação e descoberta dos tipos de obstáculos espaciais presentes no contexto agrícola, foram propostas duas técnicas que visam tratar esses obstáculos de forma que aumente a qualidade dos mapas gerados pela abordagem de Speranza (2017). Para o desenvolvimento das técnicas de tratamento de obstáculos, ocorreram inúmeras reuniões junto ao autor da abordagem SWMU *Clustering*, Dr. Eduardo Antônio Speranza e o orientador desse trabalho. Nessas reuniões, também foram discutidas formas de implementação dessas técnicas ao algoritmo usado pela abordagem.

Com os obstáculos identificados e classificados, e as propostas de tratamento de obstáculos implementadas, faz-se necessário validar o incremento feito na abordagem SWMU *Clustering*. Para isso, foi usado o critério SD-Spatial, mesmo critério usado na abordagem original de Speranza (2017). Após a implementação e validação das novas técnicas de tratamento de obstáculos, foi-se desenvolvido uma interface web para a abordagem SWMU *Clustering*, disponibilizando-a como um serviço *web*, podendo ser acessada através de uma interface amigável pelo o usuário final.

1.5 Estrutura do Trabalho

Esse trabalho está organizado da seguinte forma:

- O Capítulo 2 apresenta a fundamentação teórica para o entendimento do trabalho, expondo conceitos sobre Dados Espaciais e Agricultura de Precisão.
- O Capítulo 3 apresenta a abordagem SWMU *Clustering*, voltada para o delineamento de mapas de UGDs em AP. Nesta seção também é apresentado o pacote *Shiny*, usado para criar interfaces *web* em R.
- O Capítulo 4 descreve e apresenta obstáculos espaciais presentes no contexto de AP.

-
- O Capítulo 5 propõe estratégias para o tratamento de obstáculos, com base nos obstáculos descoberto na seção 4. Nesta seção também é apresentado a validação das estratégias sugeridas, usando o critério SD-Spatial como métrica de validação.
 - O Capítulo 6 apresenta o protótipo do serviço *Web SWMU Clustering*, assim como as tecnologias usadas para em seu desenvolvimento e a validação dos mapas gerados.
 - O Capítulo 7 encerra essa dissertação apresentando as contribuições, conclusões obtidas com essa pesquisa e indica possíveis trabalhos futuros.

Capítulo 2

FUNDAMENTAÇÃO TEÓRICA

Para o entendimento deste trabalho, esta seção abordará os fundamentos básicos sobre os assuntos que serão tratados ao longo do trabalho. Os conceitos aqui expostos estão alinhados e embasados aos da tese de Doutorado do [Speranza \(2017\)](#), visto que esse trabalho é uma continuação e incremento de sua abordagem. Mais especificamente, esse capítulo está organizado da seguinte forma:

- A Seção 2.1 apresenta conceitos sobre Dados Espaciais.
- A Seção 2.2 apresenta conceitos sobre Agricultura de Precisão.
- A Seção 2.3 apresenta conceitos sobre o pacote Shiny para a linguagem de programação R.

2.1 Dados Espaciais

Existem diferentes tipos de dados espaciais, que por sua vez, a *Open Geospatial Consortium (OGC)*, os formalizou. A OGC é uma organização internacional sem fins lucrativos, na qual seus membros atuam em diversas áreas como, governo, indústria, universidades e institutos de pesquisa, criada para desenvolver padrões relacionados aos tipos de dados espaciais ([OGC, 2020](#)). Segundo [Ciferri \(2002\)](#), os tipos básicos de dados espaciais são ponto, linha e polígono. A Figura 1 apresenta os tipos de dados básicos.

Figura 1 – Tipos de dados espaciais básicos.

Fonte: Adaptado de Ciferri (2002)

Como apresentado na Figura 1, Ciferri (2002) ainda define os pontos como sendo a menor unidade normalmente usada para representar localizações, como exemplo, uma cidade em uma mapa ou uma amostra de solo em uma lavoura, as linhas são comumente usadas para representar rios e estradas, por fim os polígonos simples, representam uma área fechada, por exemplo, um hectare de terra ou uma fazenda. Conforme apresentado, podemos compreender que o uso dos tipos básicos nos permite criar objetos espaciais como os apresentados na Figura 2.

Figura 2 – Tipos de dados espaciais complexos.

Fonte: Adaptados de Ciferri (2002)

Conforme apresentado na Figura 2, para Ciferri (2002), os objetos espaciais são usados para representação do mundo real, sendo o objeto formado por uma dimensão espacial, como coordenadas geográficas, e sua extensão no espaço euclidiano, nos quais podem ser os atributos de tal localidade espacial.

2.1.1 Banco de Dados Espaciais

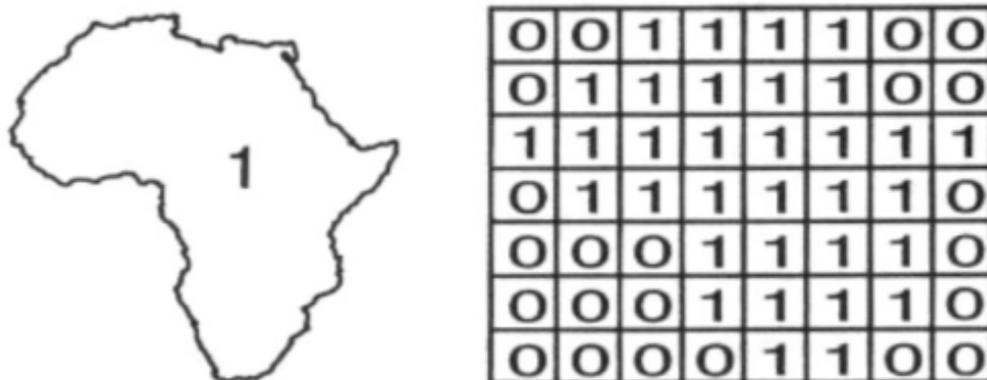
Para que os bancos de dados sejam manipulados, são usados *softwares* conhecidos como Sistemas Gerenciadores de Bancos de Dados (SGBDs). Em alguns desses SGBDs foram

desenvolvidas extensões para a manipulação, armazenamento e recuperação de dados espaciais, sendo que os SGBs mais populares que possuem tal funcionalidade são Oracle (ORACLE, 2020) e o PostgreSQL (POSTGRESQL, 2017) (CIFERRI, 2002) (SPERANZA, 2017). Um SGBD com extensão espacial é denominado Sistema Gerenciador de Banco de Dados Espaciais (SGBDE).

A fim de representar os fenômenos do mundo real, os SGBDEs permitem a representação de dados espaciais tanto no espaço bidimensional quanto no tridimensional, normalmente anexados a uma coordenada geográfica, tal como latitude e longitude (SPERANZA, 2017). De acordo com Ciferri (2002), tradicionalmente os SGBDEs representam dados do tipo espacial por meio de vetores de coordenadas (x,y). Esses dados também podem ser representados como objetos espaciais, categorizados pela OGC como tipos de dados espaciais complexos, conforme apresentados na Figura 2.

No contexto de Agricultura de Precisão, o tipo de dado mais usado é a representação por pontos (SPERANZA, 2017). Isso é devido ao fato das amostras de solo coletadas serem associadas a um ponto com uma localização espacial. Dessa forma é possível realizar agrupamentos de dados representados por células que representam pontos do mapa, para que se crie um conjunto de dados espaciais a fim de apresentar os atributos de solo de forma espacialmente diferenciada. (SPERANZA, 2017) também aponta o uso de representação por linha e polígonos, sendo a linha comumente usada para representar canais de irrigação ou rios, e os polígonos para representar a área de uma propriedade ou talhões. Na Figura 3 é representado um conjunto de dados espaciais representados por pontos no formato matricial.

Figura 3 – Conjunto de dados espaciais representados por pontos no formato matricial.



Fonte: (SPERANZA, 2017)

Como visto na Figura 3, também é possível representar um polígono de uma maneira equivalente no formato matricial, onde os *pixels* da borda e do interior do polígono possuem valor 1, e os externos possuem valor 0 (SPERANZA, 2017). Para Speranza (2017) a manipulação dos dados espaciais, se contrasta com os convencionais, devido sua complexidade, pois não possuem

relação de ordem total, ou seja, operações simples como $>$, $<$, \geq e \leq não podem ser usadas, em vez disso, normalmente são usados operadores de conjunto, topológicos e geométricos disponíveis nos SGBDEs.

2.1.2 Sistemas de Informações Geográficas

Sistemas de Informações Geográficas (SIGs) são *softwares* capazes de manipular e apresentar conjuntos de dados espaciais. Para (SPERANZA, 2017) os SIGs são compostos por módulos nos quais centralizam dados espaciais provindos de fontes heterogêneas com aplicação nas mais diversas áreas. Aronoff (1989) ainda define os SIGs como sendo um *software* de manipulação de dados, no qual representam objetos e fenômenos do mundo real que possuem informações geográficas vinculadas, onde os atributos geográficos são indispensáveis para o processamento e análise da informação.

Com a disponibilização do sinal de GPS na década de 90, os SIGs passaram a ter um papel fundamental para a manipulação e absorção desses volumes de dados gerados, com isso, o mercado passou a desenvolver e incrementar novos recursos nos SIGs, agregando-os funcionalidade para as mais diversas aplicações ligadas ao geoprocessamento (SPERANZA, 2017).

Existem diversos SIGs disponíveis para uso, como a plataforma ArcGIS e o SIG QuantumGIS, sendo o ArcGIS *software* proprietário e o SIG QuantumGIS *Open Souce* (SPERANZA, 2017). Speranza (2017) destaca a plataforma ArcGIS, que foi desenvolvida pela empresa Environmental Systems Research Institute (Esri)¹, a plataforma possui diversas extensões para realizar análises e processamentos de geoestatística, assim como funções de geoprocessamento também presentes em diversos SIGs. O SIG QuantumGIS, é um *software* de código livre e distribuição gratuita. Para Speranza (2017), o SIG QuantumGIS surge como uma ferramenta promissora, dado o grande número de complementos disponíveis para diversas atividades de geoprocessamento.

Em AP o uso de SIGs se torna fundamental, pois os dados espaciais estão intrinsecamente ligados aos processos de gestão da AP. Entretanto, o uso de SIGs como o ArcGIS e o QuantumGIS gera uma certa dificuldade aos usuários finais, dado ao fato de que os dados brutos provindos dos sensores e equipamentos agrícolas proprietários, possuem os mais diferentes formatos (MOLIN et al., 2015; SPERANZA, 2017). Molin et al. (2015) ainda destaca que esta dificuldade de conversão de dados em formatos proprietários dos equipamentos usados para a coleta no campo também ocorre em sentido oposto, na geração de mapas de recomendação que possam ser interpretados pelos equipamentos de automação agrícola proprietários.

¹ Environmental Systems Research Institute, Redlands, Califórnia, EUA, <http://www.esri.com>

2.1.3 Dados Espaciais via Web

O constante crescimento das implementações de SIGs, tal como serviços e ferramentas de informações espaciais via *Web*, fez com que se tornasse possível o compartilhamento de informações espaciais em forma de mapas, assim como a criação de ferramentas de manipulação desses dados de forma interativa, visando facilitar a análise e interpretação dos dados pelo usuário final (MITCHELL, 2005; SPERANZA, 2017). Dessa forma, o tratamento e manipulação dos dados espaciais via *web* podem favorecer e criar novas possibilidades a diversas áreas, tais como a AP.

2.1.3.1 Especificações OGC

A OGC possui algumas especificações para serviços *Web*, a OGC *Web Map Service* (WMS), OGC *Web Feature Service* (WFS) e OGC *Web Processing Service* (WPS). Essas especificações são descritas a seguir:

- **OGC *Web Map Service* (WMS):** Essa especificação provê operações que possibilitam a recuperação e a descrição de *features* geográficas dos mapas em formato de imagens (PERCIVALL et al., 2003; BEAUJARDIERE, 2006; SPERANZA, 2017). A interface cliente da especificação WMS recebe como entrada as *layers* correspondentes aos mapas e a área a ser processada. Como saída o servidor deverá retornar imagens dos mapas em formato JPEG ou PNG, sendo possível a visualização em *browsers* (SPERANZA, 2017).
- **OGC *Web Feature Service* (WFS):** A especificação WFS provê serviços que possibilitam o acesso e a edição de dados geoespaciais por meio de um conjunto de operações que o usuário cliente manipule *features* geográficas (PERCIVALL et al., 2003). Para Speranza (2017), a WFS se contrasta com WMS no ponto em que o retorno das imagem não são estáticas como no WMS. Na WFS é possível manipular de forma refinada as informações geográficas em nível de *feature* e de suas propriedades, sendo possível escolher os dados de interesse, dessa forma o tráfego de dados na rede é reduzido. Para realizar operações de busca e a recuperação dos dados é necessário informar *layers*, áreas de interesse, filtros e formato de saída, tendo como resultado o formato especificado. Além disso, nas operações de transação é necessário informar os dados relacionados ao espaço de atributos e de coordenadas. Já nas operações de inclusão, exclusão ou alteração deve-se informar o identificador da *feature*, como resposta o servidor retornará um documento com o a mensagem de sucesso ou o erro ocorrido durante a operação (SPERANZA, 2017).
- **OGC *Web Processing Service* (WPS):** A especificação WPS provê uma interface para a publicação de processos de geoprocessamento, esses processos podem ser algoritmos, modelos ou cálculos que atuem sobre esses dados, assim como disponibilizá-los (PERCIVALL et al., 2003). Esse processo recebe como entrada parâmetros obtidos por meio de

uma requisição para sua descrição, tendo como saída um documento com a mensagem de sucesso ou o erro ocorrido durante a operação.

2.2 Agricultura de Precisão

A Agricultura de Precisão e seus conceitos foram restabelecidos e disseminados na década de 90, entretanto o primeiro registro dos conceitos de AP foi em 1929 quando Lins e Bauer geraram um mapa de acidez do solo para tratar os níveis de calcário de forma espacialmente diferenciada (SPERANZA, 2017). Apesar desse primeiro registro ocorrer em 1929, demorou mais de seis décadas para a popularização da AP, devido as dificuldade tecnológicas para a elaboração desses mapas com variabilidade espacial e sua aplicação prática no campo.

Com o surgimento de novas tecnologias e o desenvolvimento de automação agrícola, assim como a disponibilização do sinal de GPS, permitiram que a AP retomasse na década de 90, entretanto apesar de sua popularização, esse termo ainda era pouco compreendido, sendo associado a soluções milagrosas que resolveriam todos os problemas e desafios presentes em uma lavoura de forma instantânea (MOLIN, 2003; SPERANZA, 2017). Outra atribuição errônea ao termo Agricultura de Precisão é quanto a sua adoção, sendo ela somente viável para grandes produtores, fazendo uso de máquinas e implementos agrícolas com tecnologia embarcada e específica para o manejo localizado, em conjunto com o GPS e SIGs. Dessa forma, desde que haja uma gestão na lavoura ou plantio, independentemente da tecnologia usada para tal fim, é válido atribuir o termo Agricultura de Precisão.

Com o desenvolvimento e a consolidação das técnica de AP, diversos autores à definiram. Srinivasan (2006) define AP como sendo uma estratégia que contribui com o meio ambiente, podendo ser utilizados diversos insumos e métodos de cultivo para identificar variações de solo e de cultura. Já para BERNARDI et al. (2014), AP é um conjunto de técnicas que gerenciam a variabilidade espacial da lavoura para obter retorno econômico e ambiental. Enquanto na agricultura tradicional a aplicação de insumos em uma cultura é realizada de maneira homogênea para toda a área de cultivo, a AP permite o gerenciamento espacialmente diferenciado da área de cultivo (SPERANZA, 2017). Ainda Speranza (2017) define AP com um sistema de gerenciamento para a agricultura, que considera a variabilidade de solo e de planta, tendo como objetivo, maior produtividade e uso adequado de insumos, proporcionando também, um ganho econômico e menor impacto ambiental.

Dadas as definições, podemos compreender que AP é a gestão da lavoura de forma espacialmente diferenciada, permitindo a aplicação de insumos (fertilizantes, pesticidas, irrigação, entre outros) nos locais corretos e na quantidade apropriada, tendo por consequência o aumento de produção e a redução de impactos ambientais.

A fim de estabelecer uma ligação entre computação e AP, Ruß (2012) define AP como um campo de pesquisa que proporciona grandes volumes de dados para coleta, podendo assim serem usados estrategicamente para realizar operações agrícolas considerando a variabilidade de solo. Baseado neste contexto, Speranza (2017) propôs um modelo para o ciclo de AP considerando os grandes volumes de dados coletados por TICs, conforme exibido na Figura 4.

Figura 4 – Ciclo da AP considerando a utilização de TICs.

Fonte: (SPERANZA, 2017)

Na Figura 4 podem ser observadas três fases no ciclo de AP: coleta de dados, interpretação e aplicação. Para Speranza (2017), na fase de coleta de dados são usados sensores e equipamentos embarcados no maquinário agrícola, já na interpretação estão presentes algoritmos de mineração e agrupamentos de dados e geoprocessamento. Por fim, na fase de aplicação, equipamentos com tecnologia embarcada faz as devidas correções e intervenções. Speranza (2017) aponta que para o sucesso de uma boa gestão da lavoura é desejável que haja um tripé consistente, formado por informação, tecnologia e tomada de decisão, sendo que a informação está relacionada a uma consistente obtenção e disponibilidade dos dados, a tecnologia esta relacionada a forma com que os dados são coletados, e a tomada de decisão envolve a decisão crítica feita por um especialista para definir ações a serem tomadas.

2.2.1 Estratégias de Aplicação e Construção de Mapas em AP

Para Speranza (2017), existem basicamente dois tipos de estratégias para a implementação dos conceitos de AP em uma lavoura, a escolha de qual caminho estratégico seguir está intimamente ligado aos objetivos que se deseja alcançar. A primeira estratégia, chamada de iniciação, pode ser considerada mais simples e com menor custo. Sua técnica de aplicação consiste em coletar amostras de solo em pontos específicos, levando-os para análise laboratorial a fim de que os resultados possibilitem a aplicação de fertilizantes e corretivos à taxa variada. A estratégia de iniciação visa apenas a economia de insumos, levando a uma redução de custos investidos na lavoura. Para ir além da economia de insumos e realizar um aumento na produtividade, é necessário que se adote a segunda técnica que é conhecida como especializada. A técnica de manejo especializada consiste em analisar dados de solo em uma resolução maior, ou seja, são coletados amostras de solo em um intervalo espacial menor. Além das características do solo também é

levado em consideração a cultura em produção e a produtividade em anos anteriores. Dessa forma é possível realizar uma análise histórica da lavoura, permitindo encontrar deficiências no solo para que possam ser realizadas as devidas intervenções corretivas. (SPERANZA, 2017)

A adoção da estratégia especializada permite ao especialista uma tomada de decisão baseada em um conjunto de dados melhor detalhados, tornando mais assertivas as ações tomadas. Entretanto, um conjunto de dados com maior resolução, aumenta a complexidade da atividade de agrupamento e análise (MOLIN, 2004). No Brasil, a estratégia de iniciação ainda é predominante, no entanto, com a disponibilidade e um acesso facilitado aos recursos e ferramentas que possam medir e coletar dados de solo e cultura em uma lavoura, essa técnica pode se tornar mais abrangente podendo realizar uma gestão considerando produtividade, solo, planta, infestações, doenças e pragas (BERNARDI et al., 2014) (SPERANZA, 2017).

A fim de criar indicadores de fertilidade do solo que apoiem essas tradicionais técnicas, desenvolveu-se mapas com amostragem de solo em grade, também bastante popular em AP (SPERANZA, 2017). Essa técnica tem como objetivo apresentar as características do solo em uma grade, de modo que um talhão é dividido em porções quadradas menores que representam uma unidade de amostragem. Para Speranza (2017), quando existe um prévio conhecimento da variabilidade da área, deve-se usar a amostragem por ponto, que consistem em retirar uma amostra central para cada quadrado da grade e outras subamostras em torno da amostra, respeitando a distância em relação a amostra central para que não haja interferência na variabilidade espacial da área em questão.

Para a construção de mapas que representam a variabilidade espacial do solo, estes podem ser obtidos por meio de outras formas que não sejam por amostras do solo, como exemplo, imagens de satélite ou *drones*. Dessa forma, para cada valor obtido em uma localização espacial deverá existir uma recomendação para aplicação de corretivos nessa mesma localização (SPERANZA, 2017). Sendo assim, todos os dados obtidos para a elaboração do mapa deverão pertencer a mesma grade espacial para que exista coesão na variabilidade espacial.

2.2.2 Unidades de Gestão Diferenciada (UGDs)

Para Speranza (2017), a gestão da lavoura embasada na variabilidade de solo e planta, ou seja, a variabilidade espacial, é a essência da AP. Sendo assim, quando aplicadas as técnicas de AP em uma lavoura, espera-se que o tratamento da área de plantio seja feito espacialmente diferenciado, de forma a analisar porções menores da área. Dessa forma quanto maior o número de divisões feitas na área de plantio analisada, maior será a resolução do mapa gerado. A fim de tratar essas divisões de forma padronizada, existem algumas estratégias que podem ser utilizadas, tais como soluções que englobam sensores, mapas de recomendação e o tratamento da área plantada baseado em unidades de gestão diferenciada (UGDs) (MOLIN et al., 2015) (SPERANZA, 2017).

Considerando o âmbito agrícola, [Molin et al. \(2015\)](#) define UGDs como áreas espacialmente delimitadas em uma lavoura com uma variabilidade interna desprezível, ou seja, as características do solo na área em questão são semelhantes, com isso o manejo espacialmente distinto torna essa área possivelmente mais produtiva. Dessa forma, sobre um olhar voltado a dados espaciais, [Córdoba et al. \(2013\)](#) e [Pedroso et al. \(2010\)](#) definem UGDs como áreas espacialmente contíguas nas quais possuem características semelhantes, possibilitando um tratamento específico para cada uma delas. Já para [Ruß \(2012\)](#), que define UGDs na perspectiva de mineração de dados espaciais, UGDs são coleções de dados espaciais com atributos semelhantes que podem ser usados para um agrupamento que indique ou auxilie a tomada de decisão.

Com o intuito de se gerar mapas de UGDs, é necessário agrupar um conjunto de UGDs com características semelhantes, dessa forma, podemos considerar que um grupo de UGDs com variabilidade interna desprezível, ou seja, com as mesmas características, formam uma classe de gestão diferenciada (CGD). Para [Speranza \(2017\)](#), uma CGD é um grupo de UGDs com variabilidade de solo semelhantes que podem ou não serem adjacentes umas às outras. A fim de consolidar as definições, a Figura 5 apresenta um mapa de CGDs.

Figura 5 – Mapa de CGD 1



Fonte: ([SPERANZA, 2017](#))

Na Figura 5 pode-se observar quatro CGDs contíguas representadas por C1, C2, C3 e C4, sendo cada uma representada por UGDs com rótulos U1, U2, U3 e U4 respectivamente. Conforme apresentado na Figura 6, pode existir CGDs não contíguas, ou seja, grupos de UGDs espacialmente não adjacentes.

Figura 6 – Mapa de CGD 2

Fonte: (SPERANZA, 2017)

Na Figura 6 as CGDs C1 e C3 podem ser consideradas não contíguas por serem compostas pelas UGDs (U1, U3) e (U4,U5) respectivamente. Já as CGDs C3 e C4 são contíguas, sendo cada uma composta por uma única UGD. Com isso, pode-se entender que um conjunto de UGDs com características semelhantes podem formar uma CGD, que por sua vez, podem ou não serem contíguas.

Considerando as estratégias de divisão espacial do solo por UGDs e CGDs, torna-se possível realizar diversas intervenções na lavoura, de modo espacialmente diferenciado considerando variabilidade espacial. Para que essas intervenções agronômicas aconteçam, esses conjuntos de dados espaciais devem ser compostos por atributos relacionados com as características do solo e de planta (SPERANZA, 2017). Com isso, o uso da técnica de UGDs permite uma gestão mais acurada da lavoura, possibilitando o tratamento e o manejo da cultura de forma específica, ou seja, as estratégias de aplicação de insumos passam a considerar o solo e a cultura de forma espacialmente diferenciada, respeitando a variabilidade natural do campo, permitindo um possível aumento na produtividade e, por conseguinte reduzindo o custo e o impacto ambiental.

Com o intuito de complementar os conceitos de UGD e CGD assim como embasar essa fundamentação para os objetivos deste trabalho, a Figura 7 apresenta conceitos de estratificação em mapas de UGDs.

Figura 7 – Mapa UGD com estratificação

Fonte: Adaptado de [Speranza \(2017\)](#)

Conforme observado na Figura 7 (a), o mapa possui 4 CGDs não contíguas, formando diversos pontos de estratificação indicados pelos círculos vermelhos. Uma estratificação ocorre quando existe uma amostra com variabilidade diferente de todas as amostras vizinhas e cuja área é pequena em relação à área total, tornando inviável o manejo diferenciado dessa UGD estratificada. Para [Speranza \(2017\)](#), uma das principais causas de estratificação no delineamento de mapas de UGDs é o agrupamento dos dados levando em consideração apenas o espaço de atributos. Dessa forma, a Figura 7 (b) apresenta um mapa gerado usando a dimensão de atributos aplicada às restrições do espaço de coordenadas, permitindo a construção de CGDs contíguas. Um dos principais problemas dos mapas com excesso de estratificação está relacionado à inviabilidade do manejo espacialmente diferenciado pelo agricultor no campo.

2.2.3 Tipos de Dados para Geração de Mapas em AP

Existem muitas formas e meios de coleta de dados no contexto de AP. Nos últimos anos, surgiram inúmeras soluções que contribuíram para o uso e adoção das técnicas de AP. Junto a essas soluções surgiram diversos equipamentos com tecnologia embarcada e sensores capazes de gerar dados relevantes para a AP. Existem algumas características em comum pertencentes a esses dados que devem ser levado em consideração previamente ao desenvolvimento de soluções aplicadas a AP, sendo que, uma das principais características está relacionada aos atributos presentes nos conjuntos de dados, que são comumente representado por coordenadas geográficas como latitude e longitude do globo ([SPERANZA, 2017](#)). Dessa forma, é necessário considerar que os dados de cada amostra possuem um espaço multidimensional de atributos que contém informações sobre determinada amostra, assim como os atributos referente a sua localização geográfica.

[Casanova et al. \(2005\)](#) destaca outras duas características nos dados produzidos em AP, os quais podem ser de natureza vetorial ou matricial. Os dados de natureza matricial são aqueles oriundos de imagens, tais como imagens de sensoriamento remoto, imagens de satélites ou de *drones*. Nessas imagens, cada pixel da imagem possui informações em bandas espectrais, sendo

possível identificar dados relativos ao solo e à cultura (SPERANZA, 2017). Por outro lado, nos dados de natureza vetorial estão aqueles dados que são coletados por meio de sensores em pontos específicos da lavoura, contendo atributos da característica do solo ou planta e a georreferencia da amostra (SPERANZA, 2017).

Outra característica dos dados usados na AP está relacionada à densidade amostral. Para Speranza (2017), as formas de coleta de dados em uma lavoura ocorrem por meio de sensores que percorrem toda a área em velocidade constante ou por meio de técnicas tradicionais de amostragem. Dessa forma, é possível concluir que existe um padrão na quantidade de amostras por hectare. Entretanto, existem técnicas que divergem a esse método, como a de Tangerino (2014), onde a densidade amostral pode variar nas regiões da área analisada, tendo uma densidade amostral maior nos lugares que o solo apresenta maior variabilidade, e densidade menor nos locais em que as amostras apresentam menor variabilidade (SPERANZA, 2017). Entretanto, para a maioria das abordagens de delineamento de UGDs que analisam e agrupam esses dados, assim como *SWMU Clustering*, é usado a densidade amostral constante.

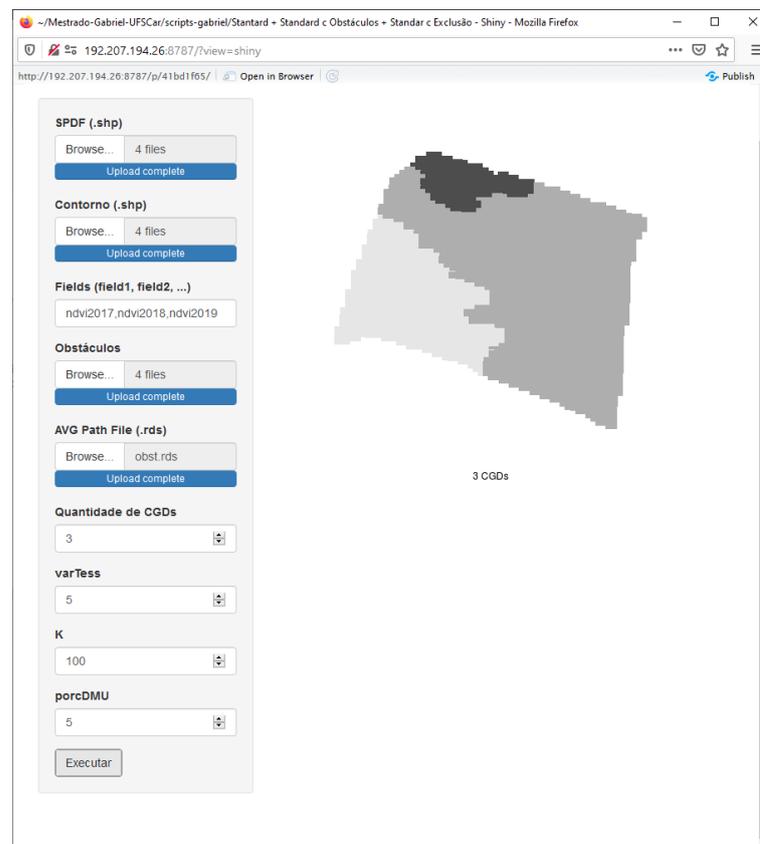
2.3 R Shiny

R é uma linguagem de programação voltada para computação estatística e geração de gráficos, tornando-se possível a implementação de técnicas estatísticas em um ambiente computacional (R-PROJECT, 2020). A linguagem encontra-se na versão 4.0.2. O código fonte do projeto R é disponibilizado como *software* livre com licença GPL (GPL, 2020).

O ambiente de desenvolvimento R pode ter suas capacidades e recursos estendidos a partir de pacotes criados pelos seus usuários, a fim de incrementar funções e técnicas para atender os mais diversos domínios (R-PROJECT, 2020). Para o desenvolvimento do protótipo proposto nesse trabalho, foi utilizado o pacote Shiny (2020).

O pacote Shiny possibilita ao usuário desenvolver *interfaces Web* a partir da linguagem R, tendo a interatividade com sistemas estatísticos computacionais complexos de maneira simples e também remota (via Web) (R-STUDIO, 2020). Sendo assim, o pacote shiny foi usado no desenvolvimento do protótipo da aplicação web para a abordagem SWMU Clustering. A Seguir, na Figura 8 é ilustrado um exemplo de *interface Web* construída com o pacote Shiny.

Figura 8 – Interface Web SWMU Clustering



Fonte: Imagem Elaborada Pelo Autor

O exemplo ilustrado pela Figura 8, apresenta a *interface Web* desenvolvida com o pacote Shiny e sua execução em um navegador de internet.

Capítulo 3

TRABALHOS RELACIONADOS

3.1 Considerações Iniciais

Este capítulo descreve os trabalhos relacionados a esse projeto de pesquisa. Também serão apresentadas as motivações e limitações de cada trabalho analisado. Esse capítulo está fortemente baseado na tese de Doutorado do Speranza (2017), pois essa pesquisa de Mestrado visa uma continuidade e melhoria de sua abordagem proposta. Dessa forma, os conceitos e definições devem estar alinhados com os apresentados inicialmente por (SPERANZA, 2017). O capítulo está organizado da seguinte forma:

- A Seção 3.2 apresenta o algoritmo *SWMU Clustering* usada para o delineamento de UGDs em AP, bem como, restrições espaciais implementadas no algoritmo.
- A Seção 3.3 apresenta o critério de validação SD-Spatial proposto por (SPERANZA, 2017).
- A Seção 3.4 conclui o capítulo com as considerações finais.

3.2 Abordagem *SWMU Clustering*

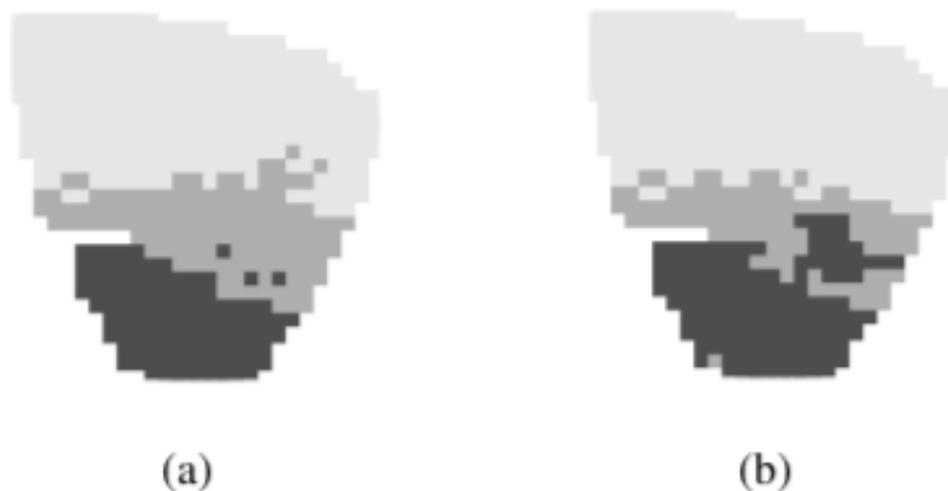
A abordagem *Spatial Ward's Management Units Clustering* (*SWMU Clustering*) é um algoritmo de agrupamento de dados espaciais baseado no algoritmo de Ward (1963). O *SWMU Clustering* é voltado para o contexto de agricultura de precisão, para realizar a tarefa de delineamento de UGDs. Sua principal vantagem, com relação a outras abordagens de objetivos semelhantes está na produção de mapas de UGDs com variância interna desprezível, ou seja, os mapas gerados pela abordagem possuem baixo nível de estratificação. Outra característica da abordagem *SWMU Clustering* é que a tarefa de agrupamento de dados espaciais usa o espaço de coordenadas geográficas para permitir ou restringir os agrupamento entre as amostras, conforme relação com o espaço de atributos (SPERANZA, 2017). Com isso, a abordagem realiza o agrupamento dos dados com base no relacionamento entre o espaço de coordenadas e o espaço

de atributos. Ao longo do trabalho, o termo algoritmo também será usado para referenciar a abordagem SWMU *Clustering*.

Com o intuito de reduzir a estratificação ao gerar mapas de UGDs para AP, a abordagem SWMU *Clustering* implementa algumas restrições espaciais que contribuem para essa redução. Uma dessas restrições espaciais aplicadas na solução, está relacionada a identificação e posicionamento dos centroides de cada grupo. A abordagem SWMU *Clustering* calcula a dissimilaridade entre os grupos baseando-se na variabilidade dos atributos das amostras em relação ao centroide, de modo que considera as restrições do espaço de coordenadas para definir uma amostra como centroide em cada grupo e em todas as fases de construção do dendrograma (SPERANZA, 2017). Com o objetivo de diminuir a estratificação, a abordagem criou um novo cálculo para a soma dos erros quadráticos de cada grupo, considerando as restrições espaciais. Com isso, a abordagem constrói o dendrograma com base nas restrições espaciais, utilizando apenas o conjunto de dados fornecidos inicialmente, sem necessitar que o usuário forneça novos parâmetros para realizar a operação (SPERANZA, 2017).

O cálculo diferenciado para obtenção dos centroides espaciais e a soma dos erros quadráticos para cada grupo são etapas fundamentais para a abordagem SWMU *Clustering* gerar mapas de UGDs com baixo nível de estratificação. A Figura 9 apresenta a redução da estratificação quando aplicado a restrição do centroide espacial.

Figura 9 – Níveis de estratificação; (a) mapa gerado por uma abordagem tradicional; e (b) mapa gerado com a abordagem SWMU *Clustering*.



Fonte: (SPERANZA, 2017)

Conforme apresentado na Figura 9, a abordagem SWMU *Clustering* possui um menor nível de estratificação com relação às abordagens tradicionais. Conforme ilustrado na Figura 9 (a), a CGD no tom de cinza mais escuro possui uma UGD de maior área e três de menor área, podendo assim, serem consideradas como estratos. Na Figura 9 (b), não ocorreu a estratificação

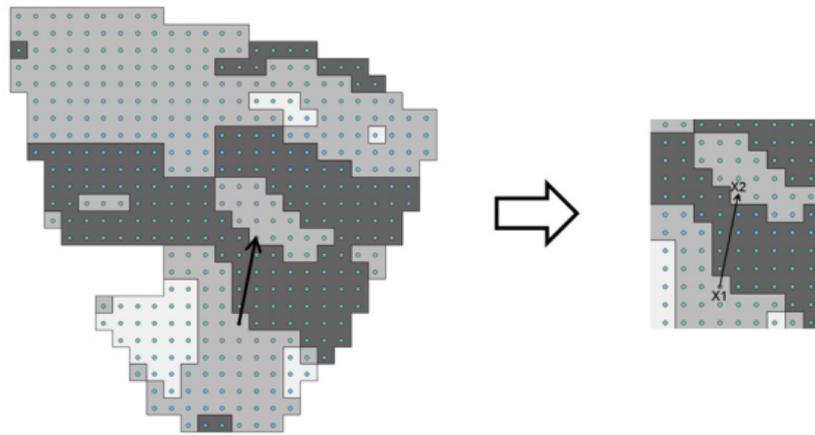
como na Figura 9 (a), no entanto, na área inferior do mapa foi gerado uma UGD muito pequena, que pode ser considerado um ponto de estratificação.

Com a restrição do centroide espacial, a abordagem *SWMU Clustering* tem uma considerável redução nos níveis de estratificação. Entretanto os resultados ainda são indesejáveis por possuírem UGDs estratificadas. Com isso a abordagem adota três restrições parametrizadas e opcionais, as quais são fornecidas pelo usuário final. Essas três restrições são a tesselação inicial, área mínima de cada UGD e obstáculos espaciais, nas quais serão descritas a seguir:

- A tesselação inicial consiste em desconsiderar determinadas UGDs no início da construção do dendrograma, as quais ficarão isoladas e serão consideradas estratos. Para isso, o algoritmo possibilita ao usuário informar o percentual da variância média para o agrupamento por meio do parâmetro *varTess*. Dessa forma, é desconsiderada toda a variância média que for menor do que a informada pelo usuário, possibilitando assim, desconsiderar UGDs que se tornaria pontos de estratificação logo no início da construção do dendrograma (SPERANZA, 2017).
- Área mínima para UGDs é uma restrição espacial opcional presente na abordagem *SWMU Clustering* que possibilita ao usuário final determinar o tamanho mínimo de cada UGD por meio do parâmetro *porcCMU*. Nesse contexto, existem outros dois parâmetros que são o *minDCM* e o *maxDCM*, que permite ao usuário determinar a quantidade mínima e máxima de CGDs. Durante o processo hierárquico do algoritmo, se a quantidade de CGDs for igual ao parâmetro *maxDCM*, e houver uma UGD menor do que o parâmetro *porcCMU*, essa UGD é absorvida pela UGD cujo centroide está espacialmente mais próximo (SPERANZA, 2017). Dessa forma a abordagem garante que não existam UGDs com área pequena suficiente para ser considerado ponto de estratificação.
- Os obstáculos espaciais, assim como as outras, é uma restrição opcional que verifica a existência de obstáculos ou barreiras entre duas amostras, usado como fator de ponderação para executar os cálculos de dissimilaridade (SPERANZA, 2017). Dessa forma, a restrição de obstáculos apenas pode ser usada a partir de um conjunto de dados espaciais que identifique a área do obstáculo, para que o algoritmo possa verificar a área com a barreira espacial entre duas amostras (SPERANZA, 2017). Para calcular a dissimilaridade entre duas amostras, a abordagem realiza a soma da dissimilaridade média da primeira amostra com todas as amostras internas ao obstáculo (ou barreira) com a dissimilaridade média da segunda amostra e todas as mesmas amostras interna ao obstáculo. Para tanto, a abordagem considera apenas o espaço de atributos e a distância Euclidiana para realizar o cálculo de dissimilaridade entre amostras (SPERANZA, 2017). Outro ponto a ser considerado, é quanto ao tipo de obstáculo usado na avaliação experimental do algoritmo, que foi usado apenas a declividade de uma área.

A estratégia de tratamento com a restrição espacial feita pela abordagem é ilustrada na Figura 10.

Figura 10 – Mapa com obstáculo espacial de declividade com detalhamento



Fonte: (SPERANZA, 2017)

Conforme observado na Figura 10, pode-se notar a área considerada como obstáculo espacial destacada em um tom de cinza mais escuro entre as duas amostras X1 e X2 (SPERANZA, 2017).

3.3 Critério de Validação Usado na Abordagem SWMU *Clustering*

Para a validação da abordagem SWMU *Clustering*, foi desenvolvido por (SPERANZA, 2017) um critério baseado no critério SD (HALKIDI et al., 2000), chamado SD-Spatial. O SD-Spatial é um critério de minimização para verificar a qualidade dos *clusters* dos mapas gerados. A principal evolução do critério SD-Spatial em relação ao critério SD está em considerar adicionalmente o espaço de coordenadas geográficas das amostras, ou seja, o SD-Spatial considera tanto o espaço de atributos quanto o espaço de coordenadas.

Sendo assim, o critério SD-Spatial é um critério de validação interna e qualitativa, verificando a coesão e a separação das CGDs, considerando tanto o espaço de atributos quanto o espaço de coordenadas (SPERANZA, 2017).

3.4 Considerações Finais

A abordagem SWMU *Clustering* é eficaz em gerar mapas de UGDs para o contexto de AP. Entretanto, a restrição de obstáculos espaciais pode ser melhor trabalhada, tanto para

identificar os possíveis obstáculos quanto para melhorar a qualidade do algoritmo no ponto em questão.

O Critério de validação interna SD-Spatial, terá grande relevância para comprovar a eficácia dos mapas gerados com as novas soluções propostas, sendo portanto, indispensável a sua descrição conceitual nessa seção.

Os trabalhos abordados possuem fundamental relevância aos objetivos pretendidos com essa pesquisa, dado que a abordagem SWMU *Clustering* é a base para as contribuições desse trabalho, e o critério SD-Spatial será a métrica de eficácia dos resultados a serem obtidos. No próximo capítulo serão apresentados os obstáculos espaciais investigados em um contexto agrícola real.

Capítulo 4

OBSTÁCULOS ESPACIAIS

Nessa seção serão apresentados os tipos de obstáculos investigados em um contexto agrícola real. Os obstáculos descobertos são classificados entre os que afetam e os que não afetam o delineamento de UGDs e em particular a abordagem SWMU *Clustering*.

4.1 Considerações Iniciais

Essa seção está organizada da seguinte forma:

- A seção 5.2 apresenta a descoberta dos obstáculos no contexto de AP que possam afetar de maneira distinta a abordagem SWMU *clustering*.
- A seção 5.3 classifica os obstáculos que interferem e que não interferem na abordagem SWMU *Clustering*.

4.2 Descoberta dos Obstáculos Espaciais

A fim de realizar uma investigação em um cenário agrícola real, surgiu a iniciativa de visitar propriedades rurais visando a descoberta de obstáculos espaciais que poderiam vir a interferir na abordagem SWMU *Clustering*. Para a realização das investigações dos obstáculos espaciais, foram definidos como busca, todos os elementos existentes na agricultura que poderiam inviabilizar de alguma forma o cultivo e o manejo da cultura.

As visitas foram realizadas na região do município de Vista Alegre do Alto - SP, localizado a 372 km da capital do estado de São Paulo. Os Obstáculos encontrados, portanto, são comuns para essa região, deixando a possibilidade de existirem outros tipos de obstáculos particulares de outras regiões, gerados por um diferente clima, técnicas de manejo ou plantio, elevação, localização ou outras variáveis naturais que não estão presentes na região investigada.

Dado o contexto da investigação dos obstáculos espaciais em um cenário agrícola real, a seguir serão apresentados os obstáculos descobertos. A Figura 11 apresenta uma árvore nativa

como obstáculo espacial fixo, desde que o referido obstáculo impossibilita de realizar o plantio nessa área em questão.

Figura 11 – Obstáculo Espacial Árvore Nativa.



Fonte: Imagem Elaborada Pelo Autor

Conforme apresentado na Figura 11, a árvore nativa é um ponto espacial dentro de uma área, onde não é possível o cultivo, desta forma foi considerada como obstáculo espacial. A Figura 12 ilustra uma área de mata nativa.

Figura 12 – Obstáculo Espacial Mata Nativa.



Fonte: Imagem Elaborada Pelo Autor

A Figura 12, apresenta ao centro da imagem uma mata nativa, região na qual não é feito o cultivo nem o manejo, por se tratar de uma área preservada. Assim como o obstáculo espacial árvore nativa, a mata nativa também é um obstáculo fixo, que pode interferir nos resultados gerados pela abordagem SWMU *Clustering*. A mata nativa é representada como um polígono para delimitar a sua área.

Figura 13 – Obstáculo Espacial Reservatório de Água no Solo.



Fonte: Imagem Elaborada Pelo Autor

A Figura 13 apresenta outro tipo comum de obstáculo espacial fixo. Conforme ilustrado na Figura 13, o reservatório de água no solo pode ser considerado um obstáculo fixo por se tratar de uma área inviabilizada para o plantio. Também é notada na imagem em questão, a presença do obstáculo espacial árvore nativa.

Figura 14 – Obstáculo Espacial Construção.

Fonte: Imagem Elaborada Pelo Autor

A próxima imagem apresentada trata-se do obstáculo espacial construção. A Figura 14 ilustra o obstáculo espacial construção. Esse tipo de obstáculo abrange todos os tipos de construções, alguns exemplos: galpões, casas, casa de máquinas, celeiro, entre outros com características semelhantes. O obstáculo espacial construção também pode ser considerado um obstáculo espacial fixo com sua área inviabilizada para o cultivo. Na Figura 14 também pode-se observar outros tipos de obstáculo como mata nativa, árvore nativa e reservatório de água (construção branca para caixa d'água).

Figura 15 – Obstáculo Espacial Reservatório de Água.

Fonte: Imagem Elaborada Pelo Autor

A Figura 15 apresenta três tipos de obstáculos espaciais fixos, sendo eles, construção, reservatório de água e poço artesiano, ambos com área inviabilizada para o plantio e manejo. O obstáculo espacial reservatório de água, também pode ser encontrado nas Figuras 12 e 14. A Figura 15 ilustra o obstáculo espacial poço artesiano, entretanto na imagem em questão é apresentado um poço semi-artesiano.

Figura 16 – Obstáculo Espacial Estrada.

Fonte: Imagem Elaborada Pelo Autor

Conforme podemos notar na Figura 16, são ilustrados os obstáculos espaciais construção, reservatório de água, reservatório de água no solo, estrada, poste e rede elétrica. Os obstáculos espaciais estrada, poste e rede elétrica podem ser identificados como obstáculos espaciais fixos. Entretanto, o obstáculo espacial rede elétrica permite o cultivo e o manejo em sua área, porém com restrição de altura da cultura. Os obstáculos estrada e poste, não permitem o plantio em sua área espacial. Exposto isso, o obstáculo espacial rede elétrica pode ser considerado um obstáculo, mas não afeta diretamente a abordagem SWMU *Clustering* por não interferir no plantio a ponto de impossibilitá-lo. Para algumas culturas, que exigem maquinário específico ou uso de técnicas invasivas (como queimadas), isso pode ser, no entanto, um fator limitante para o cultivo.

Na Figura 17 podemos identificar os obstáculos espaciais rede elétrica, construção, estrada e mata nativa. O obstáculo rede elétrica pode ser observado com maior nível de detalhamento na Figura 17.

Figura 17 – Obstáculo Espacial Rede Elétrica.



Fonte: Imagem Elaborada Pelo Autor

Assim como na Figura 17, na Figura 18 pode-se observar com maior evidência o obstáculo espacial poste.

Figura 18 – Obstáculo Espacial Poste.

Fonte: Imagem Elaborada Pelo Autor

A Figura 19 ilustra os obstáculos torre e rede elétrica alta tensão.

Figura 19 – Obstáculo Espacial Torre e Rede Elétrica Alta Tensão.

Fonte: Imagem Elaborada Pelo Autor

Semelhante a classificação feita ao obstáculo espacial poste e rede elétrica, os obstáculos torre e rede elétrica alta tensão compartilham da mesma classificação. Ambos podem ser considerados obstáculos espaciais fixos. O obstáculo espacial torre impossibilita o plantio em sua área espacial, afetando a abordagem SWMU *Clustering*.

A Figura 20 ilustra o obstáculo espacial pivô.

Figura 20 – Obstáculo Espacial Pivô.

Fonte: Imagem Elaborada Pelo Autor

Conforme ilustrada pela Figura 20, o equipamento de irrigação pivô pode ser identificado como obstáculo espacial móvel. Devido ao fato do obstáculo espacial pivô permitir o cultivo e o manejo em sua área, esse obstáculo em questão não afeta a abordagem SWMU *Clustering*. O obstáculo espacial pivô, ilustrado na Figura 20 por um pivô central, pode estender sua classificação para o pivô rebocável e outros equipamentos de irrigação, assim como o apresentado na Figura 21.

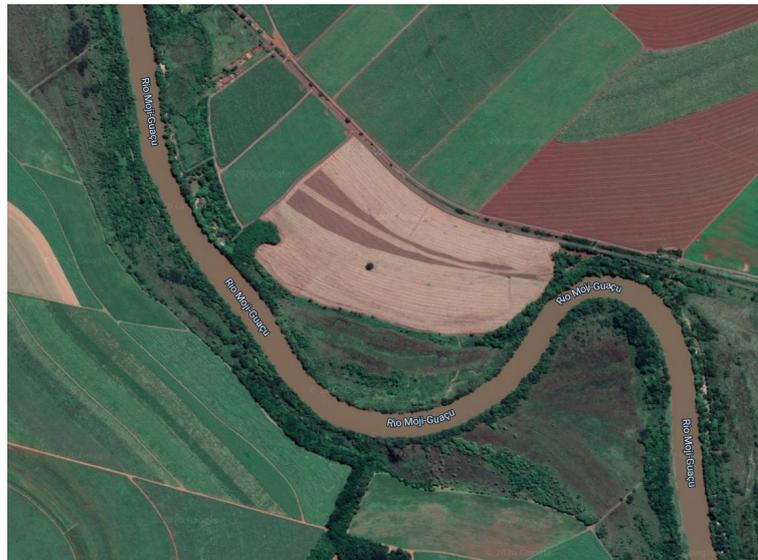
Figura 21 – Obstáculo Espacial Equipamentos de Irrigação.

Fonte: Imagem Elaborada Pelo Autor

Assim como exposto anteriormente, a Figura 21, apesar de apresentar um obstáculo espacial fixo, não possui tamanho relevante para interferir no plantio e manejo da cultura. Sendo assim, não afeta na abordagem SWMU *Clustering*. Essa classificação pode ser estendida para todos os equipamento de irrigação de tamanho que não inviabilize a cultura.

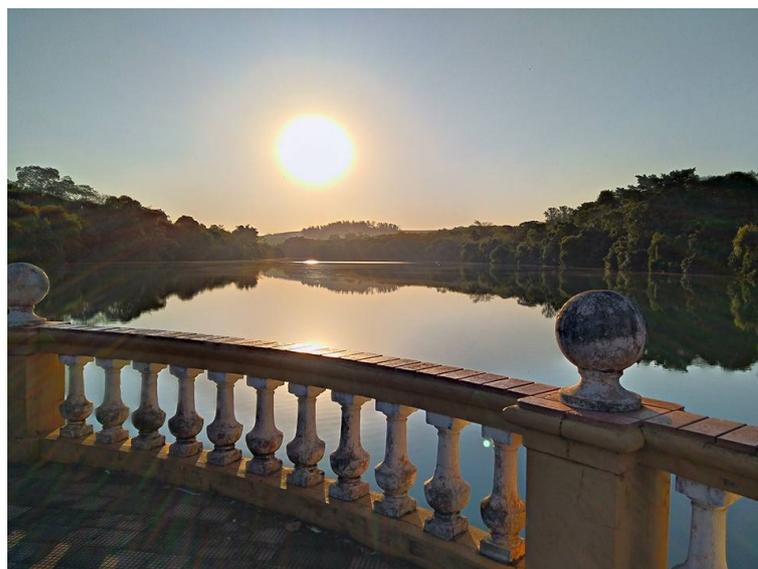
Outros obstáculos espaciais fixos a serem considerados, são rio, riacho, represa e outros similares a esses, pois inviabilizam o plantio em sua área. As Figuras 22 e 23 exemplificam esses obstáculos espaciais.

Figura 22 – Obstáculo Espacial Rio.



Fonte: (GOOGLE, 2020)

Figura 23 – Obstáculo Espacial Represa.



Fonte: (USP, 2020)

Os obstáculos espaciais rio, riacho, represa e similares são apresentados nas Figuras 22 e

23, tendo como exemplo o Rio Mogi-Guaçu localizado no interior do estado de São Paulo e a represa situada no Campus Ribeirão Preto da Universidade de São Paulo (USP).

4.3 Classificação dos Obstáculos Investigados

Exposto os obstáculos espaciais investigados na seção 5.2, se faz necessário classificá-los entre os que afetam e os que não afetam a abordagem SWMU *Clustering*. Sendo Assim, é apresentada a tabela a seguir com a classificação de todos os obstáculos espaciais investigados.

Tabela 1 – Classificação dos obstáculos

Obstáculos	Afetam a Abordagem SWMU <i>Clustering</i>	Não Afetam a Abordagem SWMU <i>Clustering</i>
Arvore Nativa	X	-
Mata Nativa	X	-
Estrada e Vias de Veículos	X	-
Reservatório de Água	X	-
Reservatório de Água no Solo	X	-
Poço Artesiano e Semi-Artesiano	X	-
Poste	X	-
Torre	X	-
Rede Elétrica de Baixa Tensão	X	-
Rede Elétrica de Alta Tensão	X	-
Construções (Casa, Galpão e similares)	X	-
Rio, Riacho e Represa	X	-
Equipamentos de Irrigação de Pequeno Porte	-	X
Pivô Central e Pivô Rebocável	-	X

Fonte: Imagem Elaborada Pelo Autor

Conforme visto na Tabela 1, os obstáculos espaciais considerados que afetam a abordagem SWMU *Clustering* foram os que inviabilizaram o plantio e o manejo da cultura. Já os obstáculos que possibilitaram o plantio, não serão considerados obstáculos espaciais que afetam a abordagem SWMU *Clustering*. Na Tabela 2 é apresentado os tipos de dados espaciais dos obstáculos investigados.

Tabela 2 – Tipos de dados espaciais dos obstáculos

Obstáculos	Tipos de Dados Espaciais
Arvore Nativa	Ponto
Mata Nativa	Polígono
Estrada e Vias de Veículos	Linha
Reservatório de Água	Ponto
Reservatório de Água no Solo	Ponto
Poço Artesiano e Semi-Artesiano	Ponto
Poste	Ponto
Torre	Ponto
Rede Elétrica de Baixa Tensão	Linha
Rede Elétrica de Alta Tensão	Linha
Construções (Casa, Galpão e similares)	Polígono
Rio, Riacho e Represa	Linha
Equipamentos de Irrigação de Pequeno Porte	Ponto
Pivô Central e Pivô Rebocável	Linha

Fonte: Imagem Elaborada Pelo Autor

Conforme apresentado na Tabela 2, foi relacionado os tipos de dados espaciais com os obstáculos investigados. Na próxima seção é apresentada as estratégias de tratamento de obstáculos propostas por essa pesquisa de Mestrado.

Capítulo 5

ESTRATÉGIAS DE TRATAMENTOS DE OBSTÁCULOS

Nessa seção serão apresentadas as estratégias adicionais para a melhoria do tratamento de obstáculos espaciais propostas para a abordagem SWMU *Clustering*.

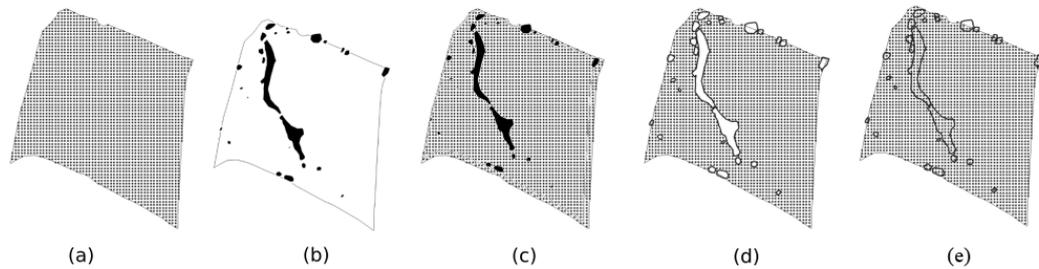
5.1 Considerações Iniciais

Essa seção está organizada da seguinte forma:

- A seção 6.2 apresenta a estratégia de substituição do conjunto de amostras internas aos obstáculos espaciais.
- A seção 6.3 apresenta a estratégia de *buffer* para o tratamento de obstáculos espaciais.
- A seção 6.4 apresenta a validação com o critério SD-Spatial das estratégias propostas em comparação com a proposta original do algoritmo SWMU *Clustering*, com o intuito de comprovar o quão eficaz são as estratégias propostas com relação à abordagem original.

5.2 Estratégia da Substituição

Com o intuito de propor estratégias para o gerenciamento de obstáculos espaciais que possam afetar a abordagem SWMU *Clustering* no delineamento de mapas de UGDs, é apresentado nesta seção a estratégia de substituição do conjunto de amostras internas aos obstáculos. Essa estratégia pode ser considerada uma estratégia de tratamento de dados. A Figura 24 apresenta as etapas de tratamento dos obstáculos nesta estratégia.

Figura 24 – Etapas de Tratamento dos Dados.

Fonte: Imagem Elaborada Pelo Autor

Como se pode observar na Figura 24, a estratégia é dividida em cinco etapas para realizar o tratamento dos dados. A seguir estão descritas as etapas:

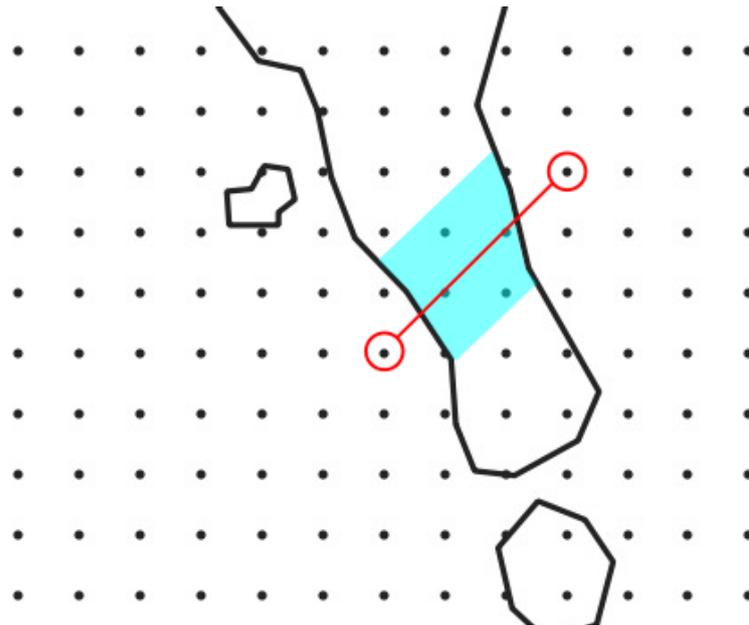
- Na primeira etapa, representada pela Figura 24 (a), é considerado todo o conjunto de dados de uma área contendo os obstáculos e as áreas de plantio, ou seja, o conjunto de dados é coletado sem distinção entre áreas de cultivo e obstáculos espaciais.
- Na segunda etapa, ilustrada pela Figura 24 (b), é feita a identificação dos obstáculos espaciais, considerando como obstáculos, os que afetam a abordagem *SWMU Clustering* que foram identificados e classificados no Capítulo 4 desse trabalho.
- Na terceira etapa, apresentada pela Figura 24 (c), é sobreposto o conjunto de dados dos obstáculos espaciais (representado pela Figura 24 (b)) ao conjunto de dados geral (representado pela Figura 24 (a)).
- Na quarta etapa, apresentada pela Figura 24 (d), são identificadas e marcadas todas as amostras internas aos obstáculos espaciais.
- Na quinta e última etapa, apresentada pela figura 24 (e), é feita a substituição dos valores das amostras marcadas na quarta etapa pelos valores interpolados das amostras mais próximas externas ao obstáculo. Para isso, é usado o algoritmo *Inverse Distance Weighted* (IDW) (SHEPARD, 1968) com potência quadrática. Assim, os valores das amostras internas aos obstáculos serão aproximados aos valores das áreas próximas externas aos obstáculos.

Com a aplicação dessa estratégia pretende-se aproximar os valores das amostras internas aos obstáculos aos valores das amostras externas ao obstáculo. Dessa forma, os obstáculos podem vir a interferir de maneira menos intensa no delineamento de mapas de UGDs feito pela abordagem *SWMU Clustering*.

5.3 Estratégia de *Buffer*

Em contraste com a estratégia anterior, apresentada na seção 6.2, a estratégia de *buffer* foi implementada na abordagem SWMU *Clustering*. A estratégia consiste em gerar uma área que envolve os obstáculos, área esta comumente denominada de zona de *buffer* em sistemas de informações geográficas. A Figura 25 ilustra a funcionalidade dessa estratégia.

Figura 25 – Estratégia de *Buffer*.



Fonte: Imagem Elaborada Pelo Autor

Conforme Ilustrado na Figura 25, foi criado um *buffer* de tamanho fixo de três vezes a resolução espacial dos dados ao redor da reta entre os quais se deseja verificar a presença de obstáculos. Em seguida, a dissimilaridade entre esse par de amostras é calculada somando-se a dissimilaridade média da primeira amostra com relação a todas as amostras internas à intersecção entre o *buffer* gerado e o obstáculo que separa as duas amostras, com a dissimilaridade média da segunda amostra com relação a essas mesmas amostras.

A principal diferença em relação à estratégia original da abordagem SWMU *Clustering*, é que o cálculo da dissimilaridade média de uma amostra externa ao obstáculo não é feita com todas as amostras internas ao obstáculo, e sim com as amostras internas a intersecção gerada entre o *buffer* e o obstáculo presente entre o par de amostras. Esta estratégia visa reduzir a variabilidade das amostras geradas pelo cálculo médio da dissimilaridade, considerando que as amostras internas ao obstáculo estão espacialmente mais próximas, diferente da estratégia original que considera todas as amostras internas ao obstáculo, podendo existir amostras espacialmente distantes.

5.4 Validação das Estratégias

Nesta seção é descrita a forma de validação das estratégias propostas para o tratamento de obstáculos, a qual foi realizada com o uso do critério SD-Spatial. Os experimentos foram realizados utilizando-se dados de uma área de cultivo de citros localizada no município de Pirangi-SP, com área total de aproximadamente 25 hectares.

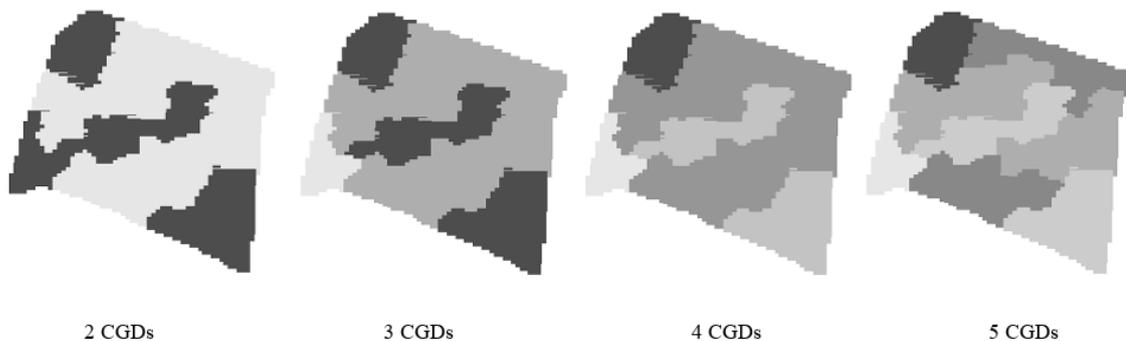
O ambiente de desenvolvimento da pesquisa foi em um *notebook* com processador intel core i5 e 8GB de memória RAM. Para a realização dos testes de validação das técnicas propostas e da aplicação *web* foi usada a infraestrutura de nuvem da Embrapa, na sua unidade de Informática em Campinas, na qual Eduardo Speranza está lotado (SPERANZA, 2017). A infraestrutura dos experimentos está instalada em ambiente OpenStack (2020) com 256 GB de memória RAM e 64 CPUs de processamento com o sistema operacional Ubuntu Server 18.04. O ambiente possui R-Studio (2020) versão 1.2.1335 e RServer versão 3.5.3 com o Shiny (2020) versão 1.5.0.

Para a realização dos experimentos foram usados dados de composições do índice NDVI (*Normalized Difference Vegetation Index*) por meio de imagens multiespectrais do satélite Sentinel-2 com resolução espacial de dez metros. No total foram 2346 amostras georreferenciadas dos anos 2017, 2018 e 2019. Para obter os dados consolidados de 2017, 2018 e 2019, foi utilizado o sistema SATVeg (2020) da Embrapa para verificação do período de pico vegetativo da cultura em cada ano. Foi obtido então, para cada pixel, o valor máximo de NDVI no período em cada um dos anos, com o intuito de reduzir a interferência de nuvens e verificar com mais precisão o potencial vegetativo da cultura. Os obstáculos para este experimento foram identificados por formas poligonais, manualmente delineadas a partir da utilização de imagem de fundo atualizada, disponibilizada pela interface do Google Mapas (GOOGLE, 2020) (Figura 26). A existência dos obstáculos delineados foi verificada de maneira presencial na área de estudo.

Figura 26 – Obstáculo identificado com imagem Google Maps

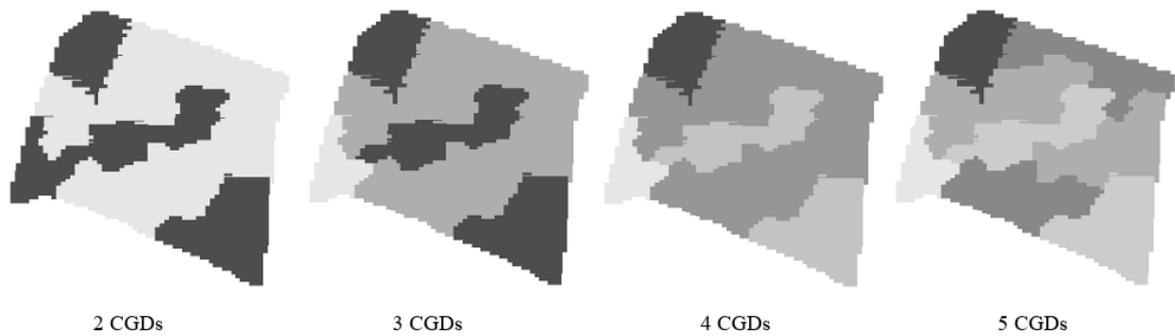
Fonte: Adaptada de (GOOGLE, 2020)

A Figura 27 apresenta os mapas de delineamento de UGDs gerados com a abordagem original sem a aplicação das técnicas propostas. A Figura 28 apresenta os mapas de delineamento de UGDs utilizando a estratégia de tratamento de dados de substituição do conjunto de amostras internas aos obstáculos.

Figura 27 – Mapas de UGDs gerados pela abordagem SWMU *Clustering* Original.

Fonte: Imagem Elaborada Pelo Autor

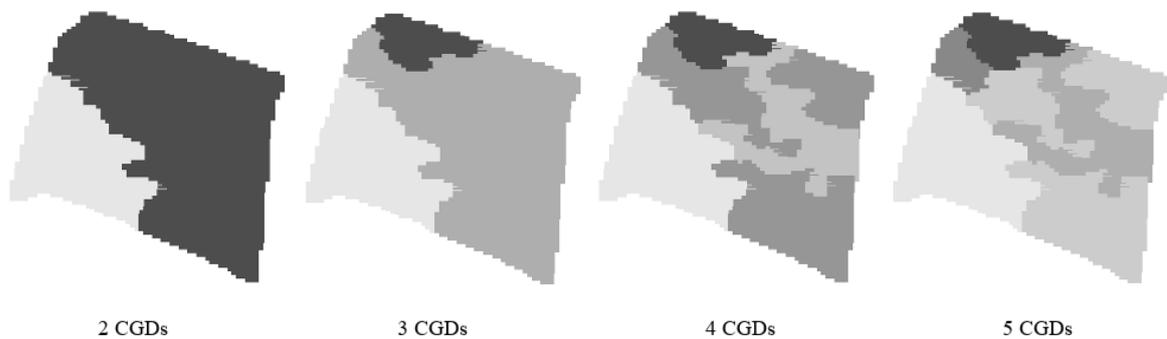
Figura 28 – Mapas de UGDs gerados pela abordagem SWMU *Clustering* usando a estratégia de substituição do conjunto de amostras internas aos obstáculos.



Fonte: Imagem Elaborada Pelo Autor

Conforme foi apresentado nas Figuras 27 e 28, o uso da estratégia de substituição do conjunto de amostras internas aos obstáculos não obteve diferença visual significativa em relação ao uso da solução original da abordagem SWMU *Clustering* sem o tratamento dos dados. A Figura 29, apresenta os mapas de UGDs gerados usando a estratégia de *buffer*.

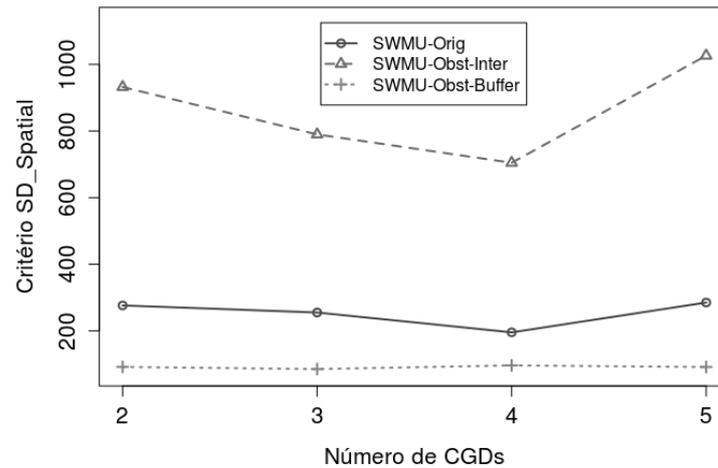
Figura 29 – Mapas de UGDs gerados pela abordagem SWMU *Clustering* usando a estratégia de *buffer*.



Fonte: Imagem Elaborada Pelo Autor

O uso da estratégia de *buffer* (Figura 29) apresentou mapas visualmente diferentes dos gerados pela solução original da abordagem SWMU *Clustering* e pela proposta da estratégia de substituição do conjunto de amostras internas aos obstáculos.

A fim de validar os resultados obtidos com as estratégias propostas foi usado o critério SD-Spatial. A Figura 30 apresenta o gráfico comparativo dos mapas gerados pelas três estratégias (original e as 2 propostas).

Figura 30 – Validação das estratégias propostas com o Critério SD-Spatial.

Fonte: Imagem Elaborada Pelo Autor

Os resultados apontados pelo critério de minimização (menor valor representa um melhor resultado) SD-Spatial, indicam uma melhoria na qualidade dos *clusters* considerando dados espaciais e convencionais utilizando a estratégia de *buffer* em relação às demais estratégias. Além disso, verifica-se uma piora da qualidade utilizando a estratégia de substituição do conjunto de amostras internas aos obstáculos.

Capítulo 6

PROTÓTIPO DO SERVIÇO *Web*

Nessa seção será apresentado o protótipo do serviço *web* para o algoritmo SWMU *Clustering*.

6.1 Considerações Iniciais

Esse capítulo está organizado da seguinte forma:

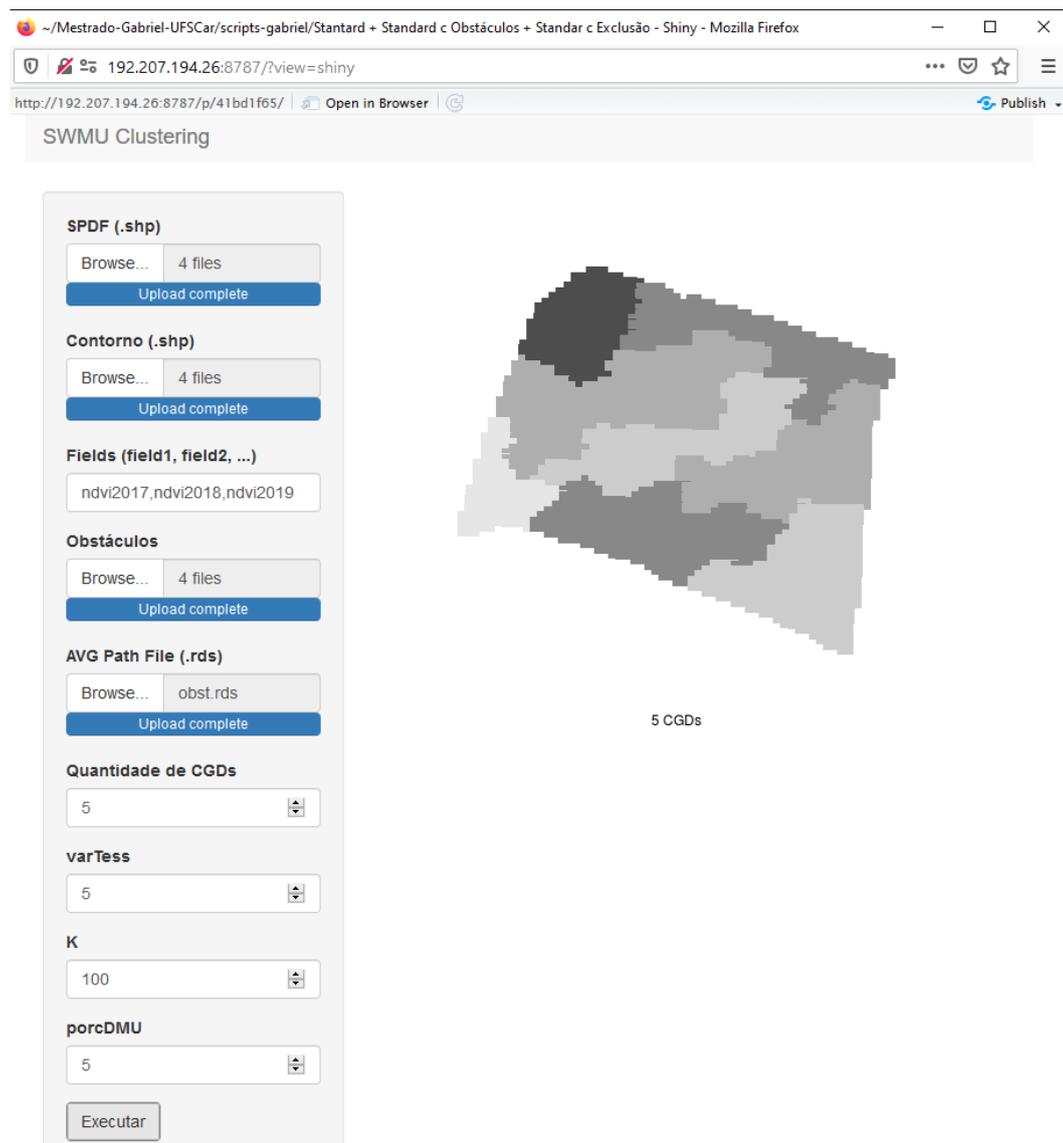
- A seção 7.2 descreve o desenvolvimento do protótipo e as tecnologias usadas.
- A seção 7.3 apresenta a validação do protótipo.
- A seção 7.4 encerra o capítulo com as considerações finais

6.2 Desenvolvimento e Tecnologias

Com o intuito proporcionar o acesso e a utilização da abordagem SWMU *Clustering* de uma maneira mais amigável ao usuário final, foi desenvolvido um protótipo de uma aplicação *web*. Considerando que a linguagem de programação usada no desenvolvimento da abordagem SWMU *Clustering* foi o R, o protótipo incrementou o projeto inicial adicionando o pacote *Shiny* para o desenvolvimento da *interface web*.

A seguir, na Figura 31 é apresentada a *interface* do protótipo desenvolvido.

Figura 31 – Interface Web SWMU Clustering



Fonte: Imagem Elaborada Pelo Autor

Como se pode observar na Figura 31, ao lado esquerdo encontra-se o painel dedicado para a entrada dos dados necessários para o funcionamento da abordagem SWMU Clustering, assim como o botão executar. A direita é reservado para a impressão dos resultados. O protótipo está sendo executado no navegador da Mozilla (2020) na versão 79.0.

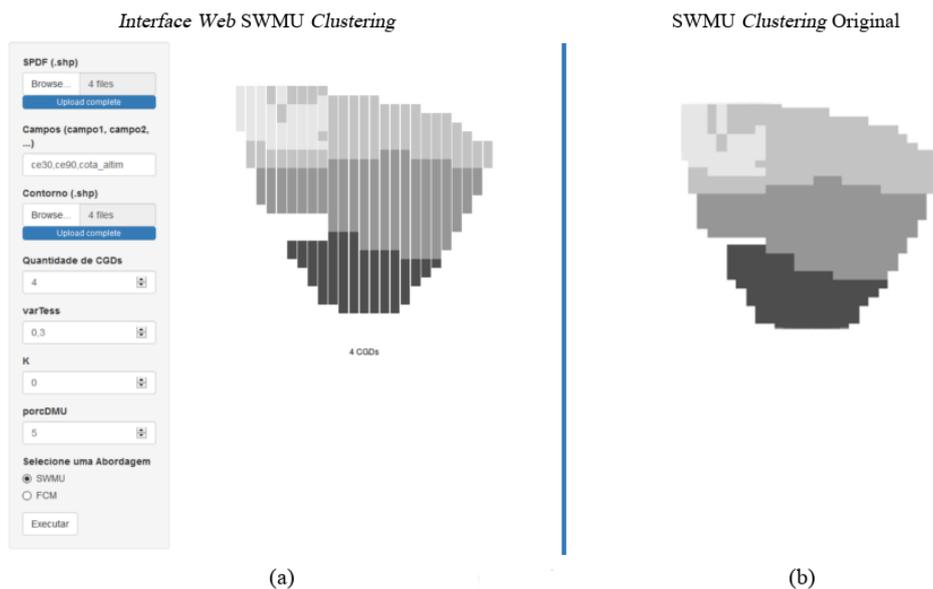
A interface permite que de maneira simples, o usuário final possa usufruir dos recursos oferecidos pela abordagem SWMU Clustering. O seu acesso é feito pelo navegador, portanto a aplicação pode ser hospedada em um servidor web com o R Server instalado, e ser acessado remotamente.

6.3 Validação do Protótipo

A fim de validar a aplicação *web* da abordagem *SWMU Clustering*, comprovando que o delineamento de mapas de UGDs requisitados por meio do protótipo não interfere nas funcionalidades do algoritmo, foram gerados os mesmos mapas de CGDs apresentados na tese de Doutorado de (SPERANZA, 2017), onde a abordagem *SWMU Clustering* é apresentada.

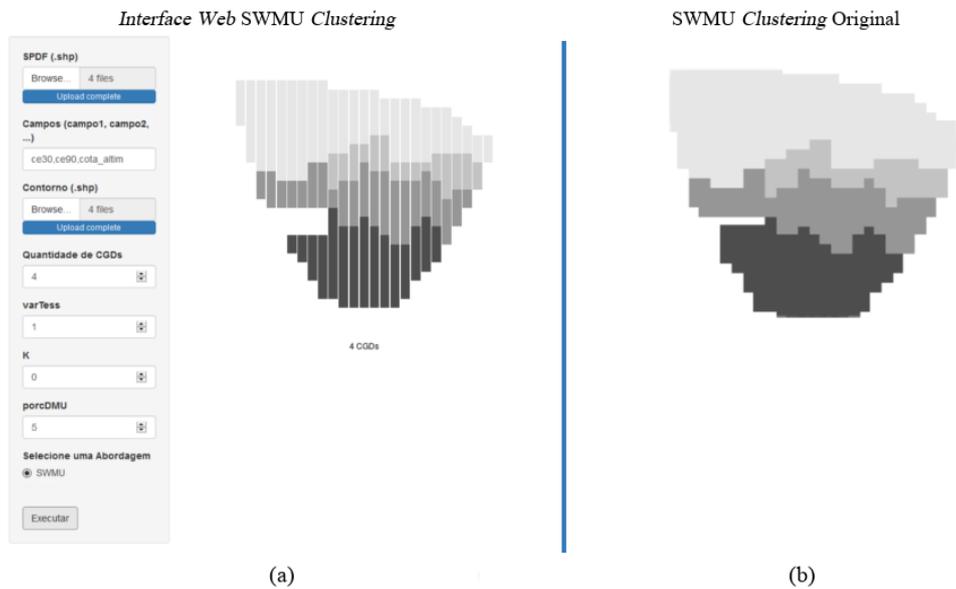
Os experimentos ilustrados pelas Figuras 32, 33, 34, 35, 36, 37, 38 e 39, apresentam no item (a) o mapa gerado pela aplicação *web*, e no item (b) o mapa gerado localmente pela abordagem proposta por (SPERANZA, 2017). Os testes apresentados pelas Figuras 32, 33, 34, 35, 36, 37, 38, 39, 40, 41 e 42 no item (a), possuem faixas brancas verticais devido a resolução do *plot* gerado pelo *R Server*, todavia, isso está relacionado com a camada de visualização do protótipo, não interferindo nos reais resultados gerados pelo algoritmo *SWMU Clustering*. Contudo, essa questão será resolvida em uma versão final da aplicação *web*.

Figura 32 – Validação protótipo - mapa contendo 4 CGDs com $\text{porcDMU}=5\%$ e varTess 0,3%



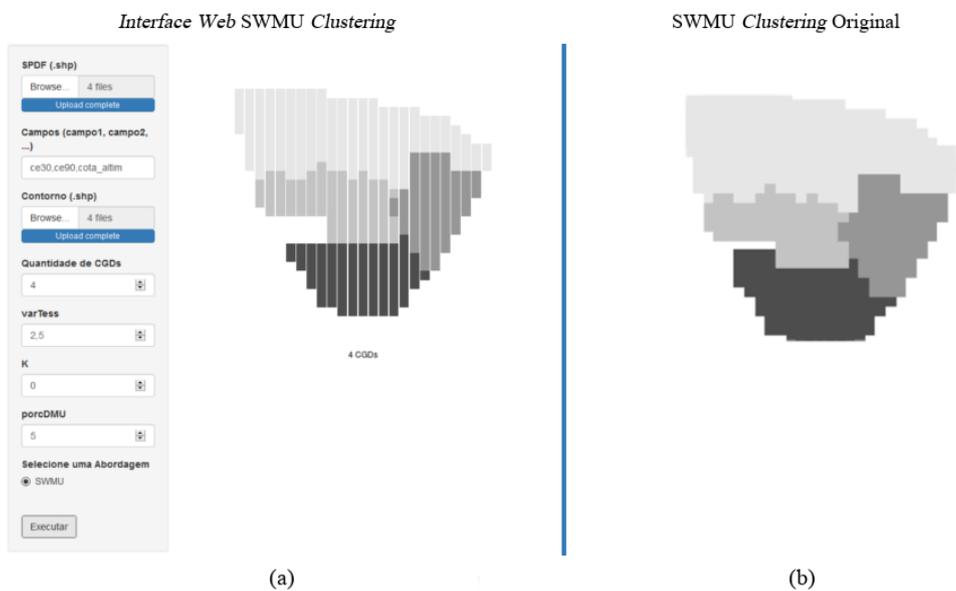
Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Figura 33 – Validação protótipo - mapa contendo 4 CGDs com porcDMU=5% e varTess 1%.



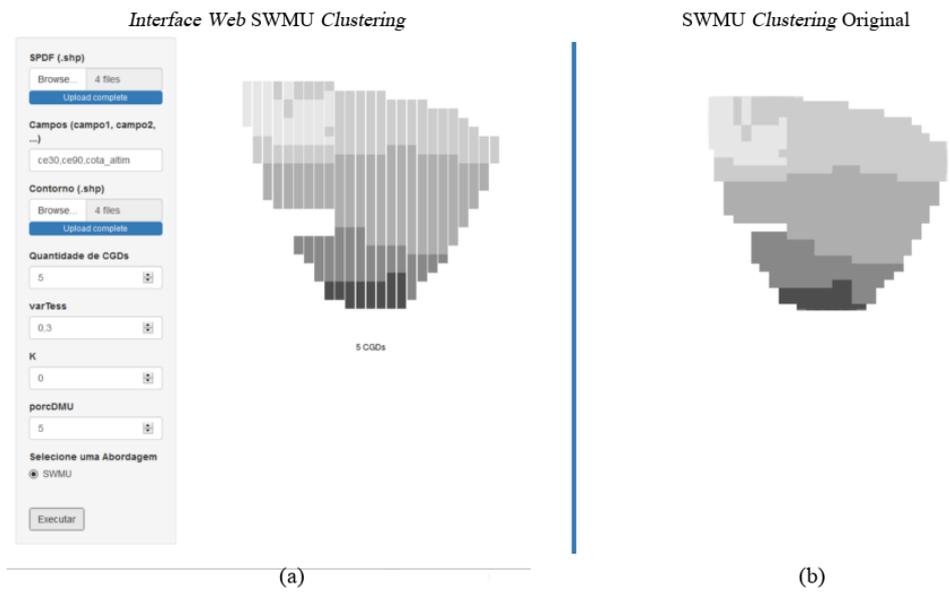
Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Figura 34 – Validação protótipo - mapa contendo 4 CGDs com porcDMU=5% e varTess 2,5%.



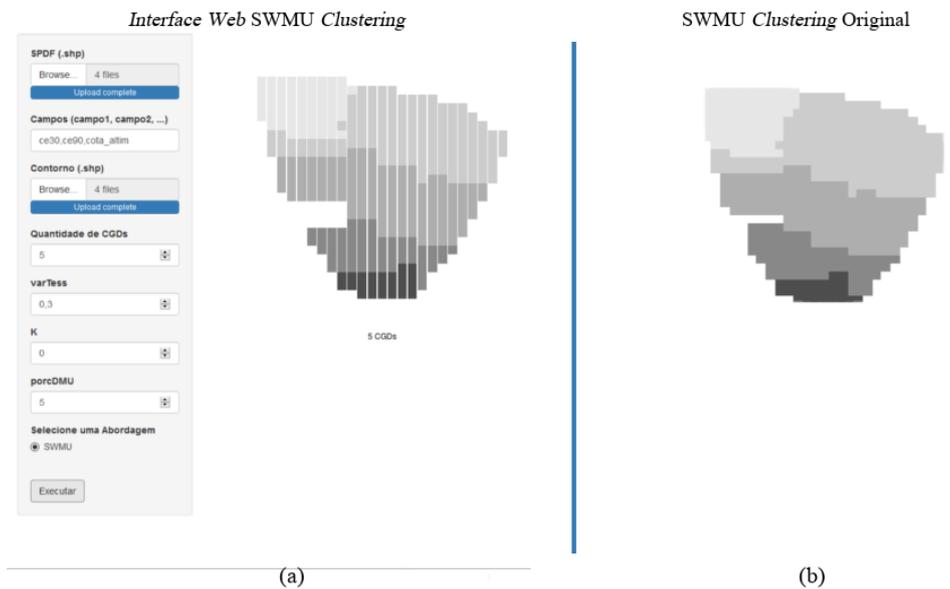
Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Figura 35 – Validação protótipo - mapa contendo 5 CGDs com porcDMU=5%, minDMC=2%, maxDMC = 5% e varTess 0,3%.



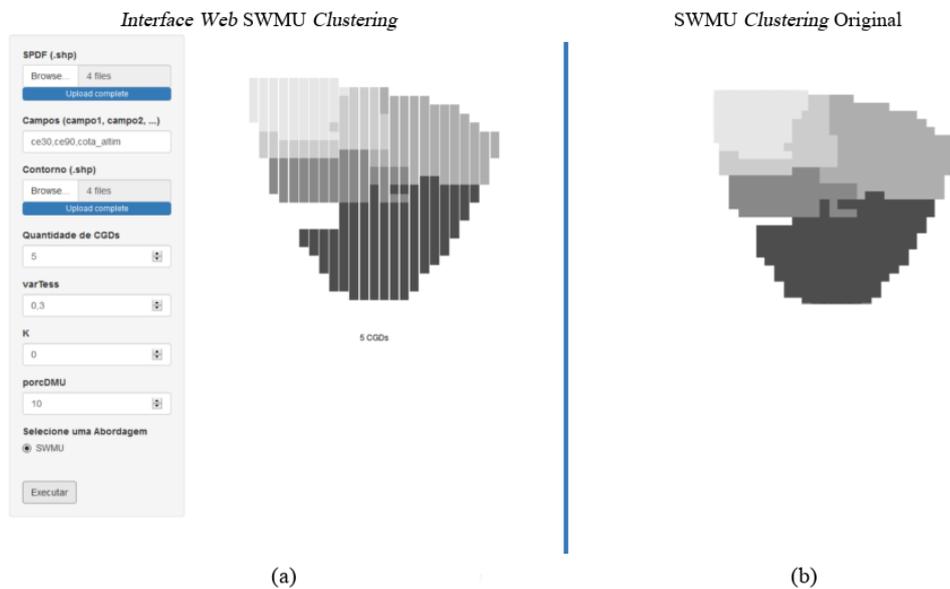
Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Figura 36 – Validação protótipo - mapa contendo 5 CGDs com porcDMU=5%, minDMC=2%, maxDMC = 5% e varTess 0,3%.



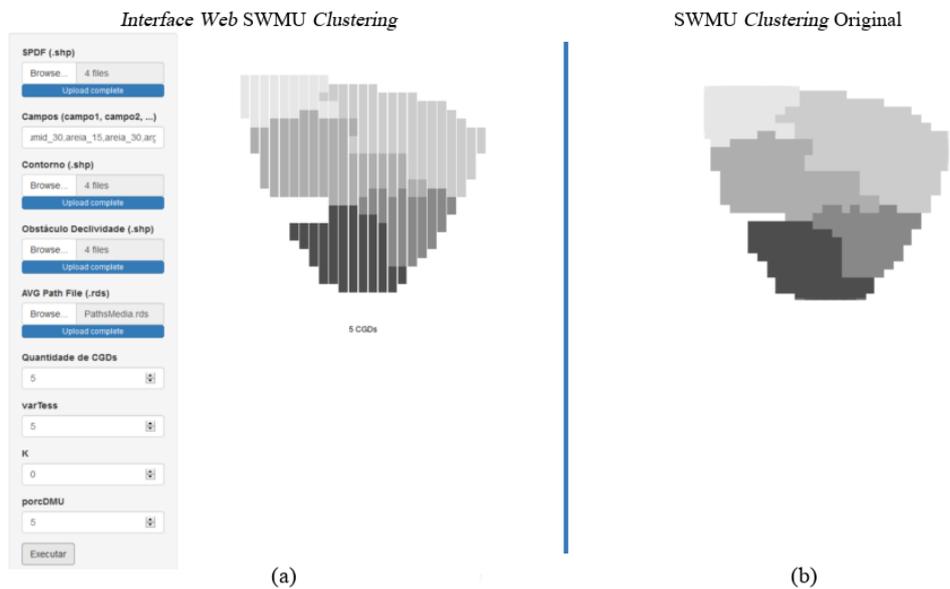
Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Figura 37 – Validação protótipo - mapa contendo 5 CGDs com porcDMU=10%, minDMC=2%, maxDMC = 5% e varTess 0,3%.



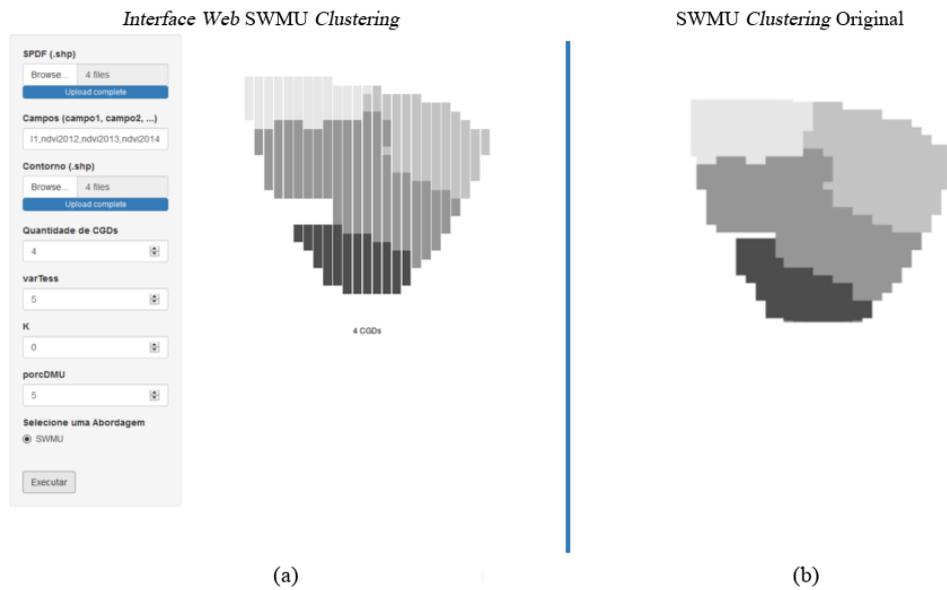
Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Figura 38 – Validação protótipo - mapa contendo 5 CGDs com com porcDMU=5%, varTess=5%.



Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

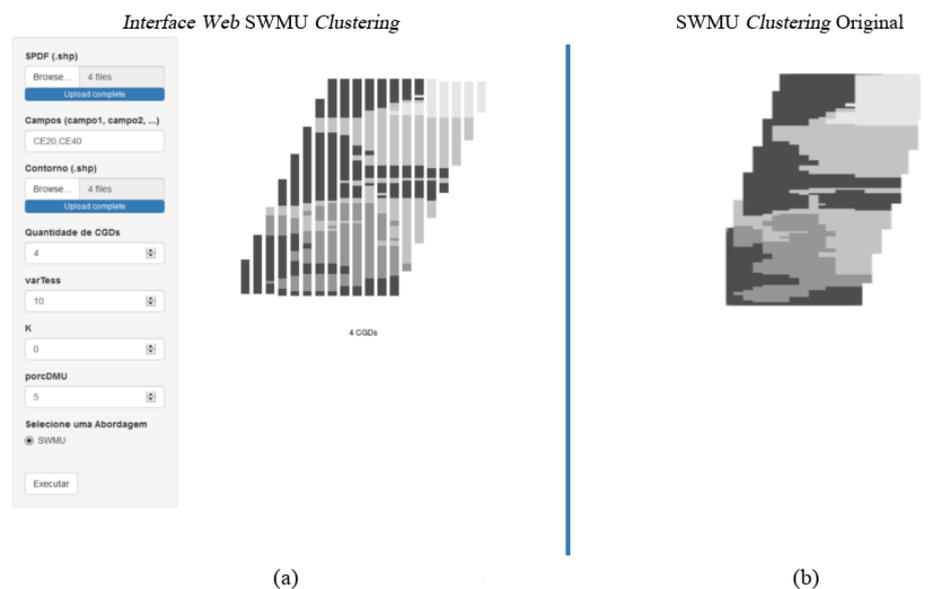
Figura 39 – Validação protótipo - mapa contendo 4 CGDs.



Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

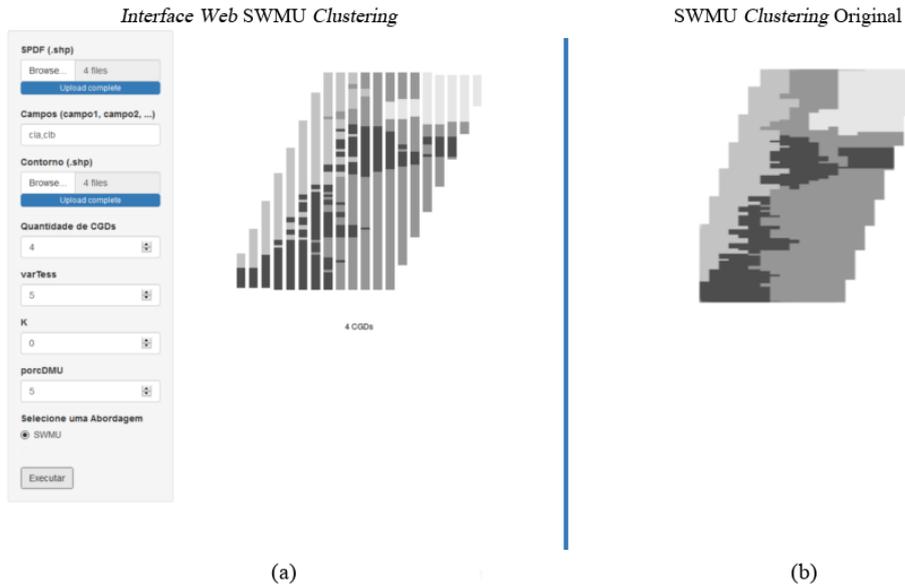
Seguindo os experimentos, será apresentada a validação para um segundo conjunto de dados. As Figuras 40, 41, 42, 43, 44 e 45, apresentam no item (a) o mapa gerado pela aplicação *web* e no item (b) gerado localmente pela abordagem proposta por (SPERANZA, 2017).

Figura 40 – Validação protótipo - mapa contendo 4 CGDs.



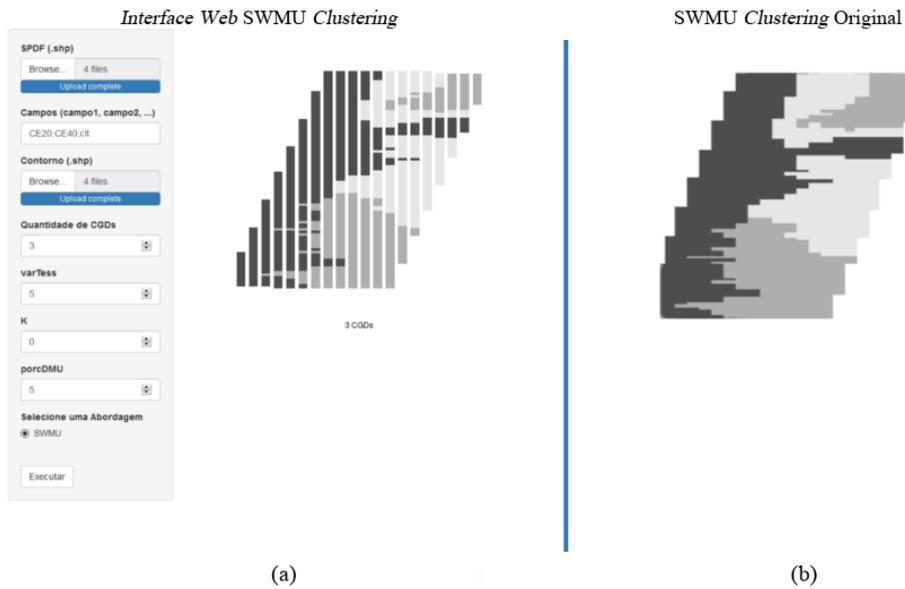
Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Figura 41 – Validação protótipo - mapa contendo 4 CGDs.



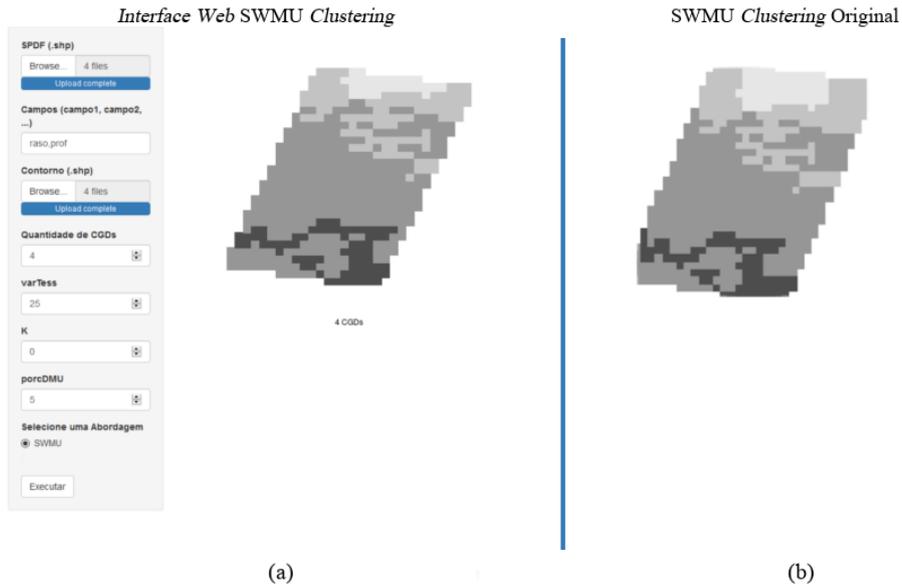
Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Figura 42 – Validação protótipo - mapa contendo 3 CGDs.



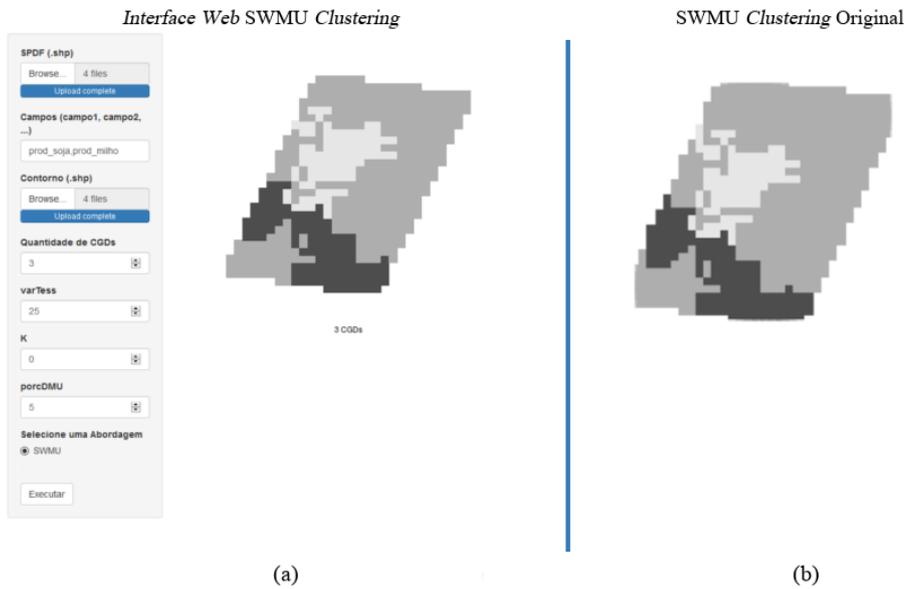
Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Figura 43 – Validação protótipo - mapa contendo 4 CGDs.

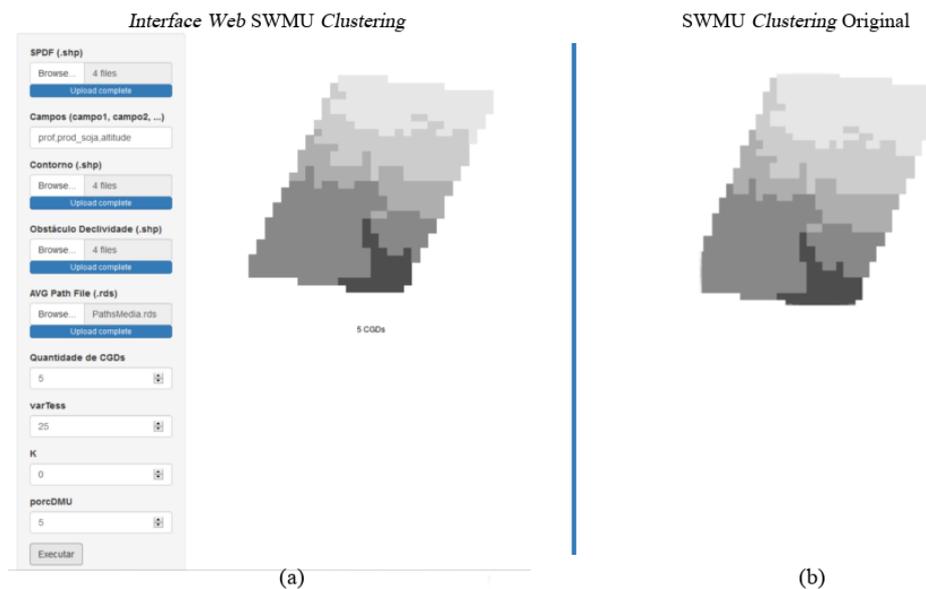


Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Figura 44 – Validação protótipo - mapa contendo 3 CGDs.



Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Figura 45 – Validação protótipo - mapa contendo 5 CGDs.

Fonte: (a) Imagem Elaborada Pelo Autor, (b) (SPERANZA, 2017)

Os experimentos apresentaram os mesmos mapas gerados pela abordagem *SWMU Clustering* sendo executada localmente sem o uso de aplicação *web*, portanto o uso do protótipo não interfere no delineamento de mapas de UGDs gerados pela abordagem.

6.4 Considerações Finais

Nesse capítulo foi apresentado de forma resumida o protótipo de *interface web* para a abordagem *SWMU Clustering*, assim como uma breve validação de suas funcionalidades por meio de experimentos.

Capítulo 7

CONCLUSÕES E TRABALHOS FUTUROS

Essa seção apresenta as conclusões obtidas com o desenvolvimento desta pesquisa de Mestrado e também apresenta algumas sugestões de trabalhos futuros.

7.1 Considerações Iniciais

Este capítulo está organizado da seguinte forma:

- A seção 8.2 apresenta as contribuições obtidas nessa pesquisa.
- A seção 8.3 sugere possíveis trabalhos futuros que podem ser desenvolvidos a partir dos resultados obtidos com essa pesquisa de Mestrado.
- A seção 8.4 conclui essa pesquisa apresentando as considerações finais.

7.2 Contribuições

Os avanços científicos obtidos com essa pesquisa de Mestrado contemplam a abordagem SWMU *Clustering* com novas estratégias de tratamento de obstáculos espaciais. A área de AP também se beneficia desses resultados, sendo considerada a obtenção de mapas de UGDs mais condizentes com o cenário real. As principais contribuições foram a identificação e classificação de obstáculos espaciais presentes no contexto agrícola, assim como o desenvolvimento de estratégias de tratamento de obstáculos e a implementação na abordagem SWMU *Clustering*. Como contribuição técnica, foi desenvolvido uma aplicação *Web* para ser disponibilizada como um serviço *web*.

Conforme exposto no Capítulo 4, pode-se notar que no contexto de AP existem diferentes tipos de obstáculos que podem interferir no delineamento de mapas de UGDs, sendo assim, necessitam de um tratamento específico para gerar mapas condizentes com o cenário real. Porém, apesar da necessidade de se tratar de forma específica os obstáculos, não foi identificado a necessidade do tratamento particular para cada tipo de obstáculo. De fato, as estratégias

propostas são aplicadas a todos os tipos de obstáculos identificados, sem distinção. Com isso, não foi possível comprovar a hipótese H_1 .

H_1 (Não Comprovada): Diferentes tipos de obstáculos espaciais presentes no contexto de agricultura de precisão necessitam de tratamento específico e diferenciado por algoritmos de delineamento de UGDs.

A validação feita com o critério SD-Spatial, apresentada no capítulo 3, comprovam a hipótese H_2 de que o tratamento de diferentes tipos de obstáculo interferem na eficácia do delineamento de mapas de UGDs. Dessa forma, a hipótese H_3 foi comprovada pela validação com o critério SD-Spatial, pois mostra a necessidade de adaptações na abordagem SWMU *Clustering* para prover uma melhor eficácia nos mapas gerados a partir de conjuntos de dados que contenham os obstáculos espaciais investigados.

H_2 (Comprovada): O tratamento de diferentes tipos de obstáculos espaciais presentes no contexto de agricultura de precisão afetam de forma distinta a eficácia de algoritmos de delineamento de UGDs.

H_3 (Comprovada): Os tipos de obstáculos descobertos requerem adaptações no algoritmo SWMU *Clustering* de forma a prover uma boa eficácia no resultado final composto por UGDs.

Entre as estratégias de tratamento de obstáculos propostas nesse trabalho, apresentadas no Capítulo 5, a estratégia de *buffer* comprovou ser a mais eficaz no delineamento de mapas de UGDs, comprovado pelo critério SD-Spatial, enquanto a estratégia de substituição do conjunto de amostras internas ao obstáculo, se mostrou pelo mesmo critério, ter menor eficácia comparada com a estratégia de *buffer* e a abordagem original. Ou seja, a estratégia de substituição do conjunto de amostras internas ao obstáculo piorou as soluções, mesmo aparentemente tendo sido conceitualmente uma ideia preliminarmente interessante. A validação, no entanto, demonstrou o contrário.

Portando, pode-se concluir que a estratégia de substituição do conjunto de amostras internas ao obstáculo não é viável ser incorporada na abordagem SWMU *Clustering*. A estratégia de *buffer* mostrou-se efetiva para contribuir para o delineamento de mapas de UGDs de melhor qualidade, portando é uma estratégia viável para uso.

7.3 Trabalhos Futuros

Visando a continuidade da pesquisa científica a partir dos resultados obtidos com essa pesquisa de Mestrado, abaixo são listados exemplos de possíveis trabalhos futuros.

- Incorporação da proposta da estratégia de *buffer* em outros algoritmos de delineamento de mapas de UGDs, de forma a investigar a sua efetividade em diferentes soluções.

- Realizar novos experimentos com as estratégias propostas, considerando novos cenários de aplicação agrícola, com dados de outras culturas e outros tipos de obstáculos espaciais no campo.
- Criação de uma API para ser incorporada aos serviços da Embrapa, de forma que a solução implementada usando a linguagem R e o pacote *Shiny* possa ser usada pelos usuários finais por meio do portal da Embrapa.

7.4 Considerações Finais

Nesse capítulo foram apresentadas as principais contribuições obtidas com o desenvolvimento desta pesquisa de Mestrado, assim como a comprovação ou refutação das hipóteses. Também foram sugeridos trabalhos futuros que podem partir dos resultados gerados nesse trabalho, a fim de permitir a continuidade da pesquisa científica no assunto investigado de delineamento de UGDs com o tratamento específico de obstáculos espaciais.

REFERÊNCIAS

- ARONOFF, S. Geographic information systems: A management perspective. *Geocarto International*, v. 4, n. 4, p. 58, 1989. Disponível em: <<http://dx.doi.org/10.1080/10106048909354237>>. Citado na página 23.
- BEAUJARDIERE, J. d. L. OpenGIS Web Map Server Implementation Specification. *Open Geospatial Consortium Inc., OGC*, p. 6–42, 2006. Disponível em: <http://portal.opengeospatial.org/files/?artifact_id=14416>. Citado na página 24.
- BERNARDI, A. d. C.; NAIME, J. d. M.; RESENDE, A. d.; INAMASU, R.; BASSOI, L. Agricultura de precisão: resultados de um novo olhar. *Embrapa Instrumentação-Livro técnico (INFOTECA-E)*, Brasília, DF: Embrapa, 2014. 596 p., 2014. Citado 3 vezes nas páginas 15, 26 e 28.
- CASANOVA, M. A.; CÂMARA, G.; DAVIS, C. A.; VINHAS, L. e; QUEIROZ, G. *Bancos de Dados Geográficos*. [S.l.]: MundoGeo, 2005. Citado na página 31.
- CIFERRI, R. R. *Análise da Influência do Fator Distribuição Espacial dos Dados no Desempenho de Métodos de Acesso Multidimensionais*. 279 p. Tese (Doutorado em Ciência da Computação) — Universidade Federal de Pernambuco, 2002. Citado 3 vezes nas páginas 20, 21 e 22.
- CÓRDOBA, M.; BRUNO, C.; COSTA, J.; BALZARINI, M. Subfield management class delineation using cluster analysis from spatial principal components of soil variables. *Computers and Electronics in Agriculture*, v. 97, p. 6–14, 2013. ISSN 0168-1699. Disponível em: <<http://dx.doi.org/10.1016/j.compag.2013.05.009>>. Citado na página 29.
- GBD: Grupo de pesquisa de banco de dados. 2019. Disponível em: <<http://gbd.dc.ufscar.br/site/>>. Acesso em: 12 mai. 2019. Citado na página 18.
- GOOGLE: Google Maps. 2020. Disponível em: <<http://maps.google.com/>>. Acesso em: 01 ago. 2019. Citado 3 vezes nas páginas 47, 53 e 54.
- GPL: GNU General Public License, url = <https://www.gnu.org/>, urlaccessdate = 01 ago. 2020. 2020. Citado na página 33.
- HALKIDI, M.; VAZIRGIANNIS, M.; BATISTAKIS, Y. Quality Scheme Assessment in the Clustering Process. In: ZIGHED, D.; KOMOROWSKI, J.; ŻYTKOW, J. (Ed.). *Principles of Data Mining and Knowledge Discovery*. Springer Berlin Heidelberg, 2000, (Lecture Notes in Computer Science, v. 1910). p. 265–276. ISBN 978-3-540-41066-9. Disponível em: <http://dx.doi.org/10.1007/3-540-45372-5_26>. Citado na página 37.
- MITCHELL, T. *Web mapping illustrated: using open source GIS toolkits*. [S.l.]: O'Reilly Media, Inc., 2005. Citado na página 24.

MOLIN, J. P. Agricultura de precisão: situação atual e perspectivas. *Milho: Estratégias de Manejo para Alta Produtividade*, Piracicaba: ESALQ/USP/LPV, p. 89–98, 2003. Citado 2 vezes nas páginas 15 e 26.

MOLIN, J. P. Tendências da Agricultura de Precisão no Brasil. In: *Congresso Brasileiro de Agricultura de Precisão*. [S.l.: s.n.], 2004. p. 1–10. Citado na página 28.

MOLIN, J. P.; AMARAL, L. R. do; COLAÇO, A. *Agricultura de Precisão*. [S.l.]: Oficina de Textos, 2015. ISBN 9788579752148. Citado 3 vezes nas páginas 23, 28 e 29.

MOZILLA: Navegador Web Mozilla Firefox. 2020. Disponível em: <<https://www.mozilla.org>>. Acesso em: 01 ago. 2020. Citado na página 58.

OGC. *The Open Geospatial Consortium*. 2020. Disponível em: <<http://www.opengeospatial.org>>. Acesso em: 09 out. 2019. Citado na página 20.

OPENSTACK: Open Source Cloud Software. 2020. Disponível em: <<https://www.openstack.org/>>. Acesso em: 01 ago. 2020. Citado na página 53.

ORACLE. *Oracle Documentation*. 2020. Disponível em: <<http://docs.oracle.com>>. Acesso em: 09 out. 2019. Citado na página 22.

PEDROSO, M.; TAYLOR, J.; TISSEYRE, B.; CHARNOMORDIC, B.; GUILLAUME, S. A segmentation algorithm for the delineation of agricultural management zones. *Computers and Electronics in Agriculture*, v. 70, n. 1, p. 199–208, jan. 2010. ISSN 0168-1699. Disponível em: <<http://dx.doi.org/10.1016/j.compag.2009.10.007>>. Citado na página 29.

PERCIVALL, G.; REED, C.; LEINENWEBER, L.; TUCKER, C.; CARY, T. *OGC Reference Model*. Open Geospatial Consortium Inc., 2003. 108 p. Disponível em: <<http://www.opengeospatial.org/standards/orm>>. Citado na página 24.

POSTGRESQL: PostgreSQL Documentation. 2017. Disponível em: <<http://www.postgresql.org/docs/9.4/static/docguide.html>>. Acesso em: 09 out. 2019. Citado na página 22.

R-PROJECT: Projeto R para Computação Estatística, url = <https://www.r-project.org/>, urlaccess-date = 01 ago. 2020. 2020. Citado na página 33.

R-STUDIO: . 2020. Disponível em: <<https://rstudio.com/>>. Acesso em: 01 ago. 2020. Citado 2 vezes nas páginas 33 e 53.

RUSS, G. *Spatial Data Mining in Precision Agriculture*. 251 p. Tese (Doktoringenieur) — Otto-von-Guericke-University of Magdeburg, 2012. Citado 3 vezes nas páginas 15, 26 e 29.

SATVEG: SATVeg. 2020. Disponível em: <<https://www.satveg.cnptia.embrapa.br>>. Acesso em: 01 ago. 2019. Citado na página 53.

SHEPARD, D. A Two-dimensional Interpolation Function for Irregularly-spaced Data. In: *Proceedings of the 1968 23rd ACM National Conference*. New York, NY, USA: ACM, 1968. (ACM '68), p. 517–524. Disponível em: <<http://dx.doi.org/10.1145/800186.810616>>. Citado na página 51.

SHINY: R Studio package to build interactive web apps straight from R. 2020. Disponível em: <<https://shiny.rstudio.com/>>. Acesso em: 01 ago. 2020. Citado 2 vezes nas páginas 33 e 53.

Silva, T. H.; S. Vaz De Melo, P. O.; Almeida, J. M.; F. Loureiro, A. A. Large-scale study of city dynamics and urban social behavior using participatory sensing. *IEEE Wireless Communications*, v. 21, n. 1, p. 42–51, February 2014. ISSN 1536-1284. Citado 2 vezes nas páginas 15 e 16.

SPERANZA, E. A. *Mineração de Dados Espaciais Aplicada no Delineamento de Unidades de Gestão Diferenciada em Agricultura de Precisão*. 340 p. Tese (Doutorado) — Universidade Federal de São Carlos, São Carlos, 2017. Citado 28 vezes nas páginas 15, 16, 17, 18, 20, 22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 34, 35, 36, 37, 53, 59, 60, 61, 62, 63, 64, 65 e 66.

TANGERINO, G. T. *Método de amostragem de área agrícola com sensores embarcados: uma abordagem que leva em conta a variabilidade do campo*. 131 p. Tese (Doutorado) — Universidade de São Paulo, São Carlos, 2014. Citado na página 32.

USP: Usp Câmpus Ribeirão Preto. 2020. Disponível em: <<https://www.facebook.com/usprp/>>. Acesso em: 01 ago. 2020. Citado na página 47.

WARD, J. H. J. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, Taylor Francis, v. 58, n. 301, p. 236–244, 1963. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/01621459.1963.10500845>>. Citado na página 34.