

Augusto César Ferreira de Miranda Oliveira

**Sistema web para processamento e análises espaciais:
uma aplicação em dados de triagem neonatal do Rio
Grande do Sul, Brasil**

Recife

30 de março de 2022



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

Sistema web para processamento e análises espaciais: uma aplicação em dados de triagem neonatal do Rio Grande do Sul, Brasil

Tese julgada adequada para obtenção do título de Doutor em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 23/02/2022 pela comissão examinadora

Área de Concentração: Modelagem e Métodos Computacionais

Orientador: Dr. Moacyr Cunha Filho
Coorientador: Dr. Guilherme Vilar

Recife

30 de março de 2022

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

- O48s Oliveira, Augusto César Ferreira de Miranda
Sistema web para processamento e análises espaciais: uma aplicação em dados de triagem neonatal do Rio Grande do Sul, Brasil / Augusto César Ferreira de Miranda Oliveira. - 2022.
111 f.
- Orientador: Moacyr Cunha Filho.
Coorientador: Guilherme Vilar.
Inclui referências, apêndice(s) e anexo(s).
- Tese (Doutorado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Biometria e Estatística Aplicada, Recife, 2022.
1. Processamento e organização de dados. 2. Sistemas de informações geográficas. 3. Análises espaciais. 4. Abordagem web. 5. Neonatologia. I. Filho, Moacyr Cunha, orient. II. Vilar, Guilherme, coorient. III. Título

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

Sistema web para processamento e análises espaciais: uma aplicação em dados de triagem neonatal do Rio Grande do Sul, Brasil

Augusto César Ferreira de Miranda Oliveira

Tese julgada adequada para obtenção do título de Doutor em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 23/02/2022 pela comissão examinadora

Orientador:

Dr. Moacyr Cunha Filho
Orientador

Banca examinadora:

Dr. Guilherme Vilar
Universidade Federal Rural de
Pernambuco

Dr. Edson Hilan Gomes de Lucena
Universidade Federal da Paraíba

Dr. Fábio Henrique Portella Corrêa
de Oliveira
COMPESA

Dr. Victor Casimiro Piscoya
Universidade Federal Rural de
Pernambuco

Dedico este trabalho aos meus avós:

Josefa Maria; e
Francisco Ferreira
(*in memoriam*)

“Eles foram para mim um exemplo muito grande de
esforço, determinação e trabalho duro.”

Agradecimentos

– Primeiramente à Deus, pelo dom da minha vida; pela sabedoria; pela determinação; pela coragem; por também sempre manter firme minha esperança em dias melhores; e por sempre renovar minhas forças quando achei que estava sem condições de seguir adiante.

– Agradeço aos meus pais, Maria Luci e Gilberto Beserra, por todo o apoio que me foi dado; por todo carinho; e atenção nos momentos mais complicados que passei durante os meus estudos.

– Aos meus irmãos, Gustavo Henrique e Marília Gabriela, que me ajudaram nos momentos mais difíceis deste processo.

– Ao prof. Dr. Guilherme Vilar, por todo o aprendizado adquirido desde o mestrado até aqui; por todas as portas abertas; por todos os momentos em que o sr. foi mais que um professor, um pai. Apesar de qualquer situação ocorrida nos últimos anos, minha gratidão pelo sr. é imensurável.

– Ao prof. Dr. Moacyr Cunha Filho que, mesmo com a forma que nos conhecemos kkk, uma grande relação acadêmica foi construída. Não tenho palavras para agradecer o que sr. fez por mim e pela nossa turma. Minha gratidão será para sempre. O sr. é um grande exemplo para mim, pois aprendi que manter a paz e ajudar todos a crescer (o programa) é fundamental.

– Ao prof. Dr. Alexsandro Machado, por todo o suporte nos momentos mais delicados da minha vida acadêmica; pelos conselhos; pelas oportunidades abertas; e por todo o aprendizado.

– As prof^{as}. Dr^{as}. Simone Castro e Dr^a. Cristiane Kopacek, que nos confiaram os seus dados e acreditaram no nosso potencial para construção deste estudo.

– À todo o corpo docente do Programa de Pós-Graduação em Biometria e Estatística Aplicada (PPGBEA) que, com as suas *expertises*, me possibilitaram adquirir o conhecimento necessário e suficiente para realizar este estudo.

– À Bárbara Carvalho, minha colega de doutorado e minha amiga de vida, por todos os momentos em que você teve paciência comigo, me ensinando e me ajudando nas disciplinas mais complexas do doutorado, pois minha base matemática era muito fraca.

– À Lucas Amaral, meu colega de departamento e amigo pessoal, por todos os

momentos em que você me ajudou na compreensão do método fractal abordado neste trabalho.

– Aos vóio (apelido carinhoso kkk) Henrique, Eduardo e João, que foram fundamentais na minha formação e na dos meus colegas, nos dando conselhos e incentivos.

– Aos meus tios, em especial Joseval (Josó), Josemir, Jobel (Joba), Jailton, Luciana e Rogério, por cada palavra incentivo e por toda a atenção que vocês me deram. Vocês foram e ainda são fundamentais na minha formação humana e profissional.

– Aos meus amigos, em especial Tiago Ferreira e Leandro Camboim, por todas as conversas e principalmente os momentos de descontração. Vocês são meus grandes amigos e que se importam com o meu crescimento pessoal e profissional.

– À todos da casa estudantil Residência das Palmeiras, que foi a minha segunda casa durante a minha formação. Agradeço a vocês, em especial João, Sampaio, Luiz, Natacha, Nórthon e Emanuel, por todo o conhecimento compartilhado e os momentos de “zueira”, diversão, cachaça kk e alegria.

– À Universidade Federal Rural de Pernambuco (UFRPE), que foi a minha terceira casa (eu chegava +- às 9h e saía às 19h, todos os dias) e que me deu todo o suporte estrutural necessário, desde o mestrado, para a minha formação acadêmica. Irei sentir muita falta deste espaço, do aconchego e de todas as pessoas nela presentes.

– À CAPES, pelo apoio financeiro. Sem a bolsa de estudos, com certeza, o resultado teria sido diferente. Ela foi fundamental para a concretização deste estudo.

– Aos membros avaliadores deste trabalho, prof. Dr. Edson Hilan; prof. Dr. Fábio Henrique; e prof. Dr. Victor Casimiro, pelas excelentes contribuições para melhoria do texto da tese e da pesquisa em si.

– À todos que também contribuíram indiretamente na concepção deste estudo e na minha formação.

“Sou eternamente grato a todos vocês.
Que Deus derrame bênçãos e mais bênçãos na vida de todos e que
possamos trabalhar juntos para um futuro mais próspero e
justo para as próximas gerações.” – (Augusto César)

“Deus é grande, Deus é forte, quando ele quer não tem quem não queira.”
(Ayrton Senna)

Resumo

A análise de dados espaciais se refere ao processo de encontrar padrões, detectar anomalias ou testar hipóteses e teorias, através da observação de fenômenos associados a uma área geográfica ou localização específica. Compreender a distribuição espacial de fenômenos ainda é um desafio frequentemente relatado em diversas áreas do conhecimento, devido ao acesso a dados espaciais precisos, históricos, oportunos e compatíveis. Outro desafio também está relacionado à organização de dados espaciais de forma inteligente a agilizar o processo de consulta e de construção de análises espaciais. Neste sentido, este trabalho teve como objetivo propor uma abordagem na *web* para processamento de dados e construção de análises espaciais de área e pontos. Para concepção das análises foram construídos dois métodos/algoritmos, um para geração de dados pontuais espaciais e um outro para associação de dados ao seu contexto espacial, ou seja, às suas microrregiões e mesorregiões. A abordagem proposta foi desenvolvida utilizando a linguagem de programação JavaScript. Os métodos de mapeamento coroplético; *kernel density estimation*; e dimensão de correlação, foram utilizados para a construção das análises. A abordagem foi validada utilizando dados de três doenças (fibrose cística, hiperplasia adrenal congênita, e hemoglobinopatias) de um programa de triagem neonatal no sul do Brasil. Os dados foram coletados entre os anos de 2004 e 2020. A abordagem desenvolvida se mostrou relevante no contexto de análise espacial, possibilitando velocidade no processamento, organização de dados e, conseqüentemente, na construção de resultados significativos que possam ser utilizados em políticas públicas que impactem diretamente na qualidade de vida das pessoas e nos desafios da saúde. A abordagem também mostrou elevado potencial de replicação para outros contextos de estudo.

Palavras-chaves: Processamento e organização de dados; Sistemas de informações geográficas; Análises espaciais; Abordagem *web*; Neonatologia.

Abstract

Spatial data analysis refers to the process of finding patterns, detecting anomalies, or testing hypotheses and theories by observing phenomena associated with a specific geographic area or location. Understanding the spatial distribution of phenomena is still a challenge frequently reported in several areas of knowledge due to access to accurate, historical, timely, and compatible spatial data. Another challenge is also related to spatial data organization in an intelligent way to speed up the consultation process and the construction of spatial analyses. In this sense, this work aimed to propose a web approach for data processing and construction of spatial analysis of area and points. To design the analyses, two methods/algorithms were built, one for generating spatial point data and another for associating data to its spatial context, that is, to its micro-regions and meso-regions. The proposed approach was developed using the JavaScript programming language. Choroplectic mapping methods; kernel density estimation; and correlation dimension were used to construct the analyses. The approach was validated using data from three diseases (cystic fibrosis, congenital adrenal hyperplasia, and hemoglobinopathies) from a neonatal screening program in southern Brazil. Data were collected between 2004 and 2020. The approach developed proved to be relevant in the context of spatial analysis, enabling speed in processing, data organization, and, consequently, in the construction of significant results that can be used in public policies that directly impact people's quality of life and health challenges. The approach also showed high replication potential for other study contexts.

Keywords: Data processing and organization; Geographic information systems; Spatial analysis; Web approach; Neonatology.

Lista de figuras

Figura 1 – Mapa de localização do estado do Rio Grande do Sul – Brasil, com as suas mesorregiões.	5
Figura 2 – Mapa de localização da capital Porto Alegre e do Hospital Materno-Infantil Presidente Vargas (HMIPV).	6
Figura 3 – Fluxograma geral da abordagem proposta.	7
Figura 4 – Fluxograma de processamento e organização dos dados.	8
Figura 5 – Interface na <i>web</i> para processamento de dados e geração de dados espaciais < https://spatialworkspace.ddns.net/geocoding >.	10
Figura 6 – Interface na <i>web</i> para validação de dados espaciais < https://spatialworkspace.ddns.net/map >.	11
Figura 7 – Interface na <i>web</i> para geocodificação manual de endereços < https://spatialworkspace.ddns.net/address >.	12
Figura 8 – Fluxograma de associação de região.	13
Figura 9 – Interface na <i>web</i> para associação de dados às suas regiões < https://spatialworkspace.ddns.net/region >.	14
Figura 10 – Interface na <i>web</i> de busca manual de regiões < https://spatialworkspace.ddns.net/search >.	15
Figura 11 – Interface na <i>web</i> para contagem automática de número de casos < https://spatialworkspace.ddns.net/cases >.	16
Figura 12 – Interface na <i>web</i> de exemplo de distância angular entre dois eventos < https://spatialworkspace.ddns.net/distance >.	19
Figura 13 – Interface na <i>web</i> de exemplo do cálculo de agrupamento da dimensão de correlação fractal < https://spatialworkspace.ddns.net/distance >.	20
Figura 14 – Interface (1) na <i>web</i> de cálculo automático da dimensão de correlação fractal < https://spatialworkspace.ddns.net/fractal >.	21
Figura 15 – Interface (2) na <i>web</i> de cálculo automático da dimensão de correlação fractal < https://spatialworkspace.ddns.net/fractal >.	21
Figura 16 – Mapa de distribuição dos casos das doenças por mesorregião.	24
Figura 17 – Mapa de distribuição dos casos das doenças por microrregião.	26
Figura 18 – Mapa coroplético do número total de casos das doenças por mesorregião.	29
Figura 19 – Mapa coroplético do número total de casos das doenças por microrregião.	30
Figura 20 – Mapa coroplético do número total de casos das doenças por município.	31

Figura 21 – Mapa coroplético da prevalência das doenças por mesorregião.	32
Figura 22 – Mapa coroplético da prevalência das doenças por microrregião.	33
Figura 23 – Mapa coroplético da prevalência das doenças por municípios.	34
Figura 24 – Mapa pontual dos casos das doenças sobre o número total de habitantes.	35
Figura 25 – Mapa de KDE dos casos das doenças sobre o número total de habitantes.	37
Figura 26 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Centro Oriental Rio-grandense.	38
Figura 27 – Mapa de localização dos casos das doenças sobre a mesorregião Centro Oriental Rio-grandense.	39
Figura 28 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Nordeste Rio-grandense.	40
Figura 29 – Mapa de localização dos casos das doenças sobre a mesorregião Nordeste Rio-grandense.	41
Figura 30 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Centro Ocidental Rio-grandense.	42
Figura 31 – Mapa de localização dos casos das doenças sobre a mesorregião Centro Ocidental Rio-grandense.	43
Figura 32 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Sudeste Rio-grandense.	44
Figura 33 – Mapa de localização dos casos das doenças sobre a mesorregião Sudeste Rio-grandense.	45
Figura 34 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Metropolitana de Porto Alegre.	46
Figura 35 – Mapa de localização dos casos das doenças sobre a mesorregião Metropolitana de Porto Alegre.	47
Figura 36 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Noroeste Rio-grandense.	48
Figura 37 – Mapa de localização dos casos das doenças sobre a mesorregião Noroeste Rio-grandense.	49
Figura 38 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Sudoeste Rio-grandense.	50
Figura 39 – Mapa de localização dos casos das doenças sobre a mesorregião Sudoeste Rio-grandense.	51
Figura 40 – Gráfico log – log de $C(r)$ por r (km) para todas as mesorregiões.	52
Figura 41 – Gráfico da variação dos valores da dimensão de correlação D_c para todas as mesorregiões.	53

Figura 42 – Mapa das mesorregiões com o maior e o menor índice de dimensão de correlação.	54
Figura 43 – Mapa das microrregiões com os menores índices de dimensão de correlação.	57
Figura 44 – Mapa das microrregiões com os maiores índices de dimensão de correlação.	58

Lista de tabelas

Tabela 1 – Número de casos das doenças por mesorregião	25
Tabela 2 – Número de casos das doenças por microrregião	27
Tabela 3 – Dimensão de correlação por mesorregião	55
Tabela 4 – Dimensão de correlação por microrregião	56

Lista de abreviaturas e siglas

API	<i>Application Programming Interface</i>
AR	<i>Associate Regions</i>
ASSYRIA	<i>dA</i> ta <i>proceSsing SY</i> stem fo <i>R</i> <i>spatIal A</i> nalysis
CEP	Código de Endereçamento Postal
DATASUS	Departamento de Informática do Sistema Único de Saúde
GIS	<i>Geographic Information System</i>
HMIPV	Hospital Materno-Infantil Presidente Vargas
IBGE	Instituto Brasileiro de Geografia e Estatística
IDH	Índice de Desenvolvimento Humano
KDE	<i>Kernel Density Estimation</i>
NaN	<i>Not a Number</i>
RN	Recém-nascidos
RS	Rio Grande do Sul
SIDRA	Sistema IBGE de Recuperação Automática
SHP	<i>Shape File</i>
SRTN	Serviço de Referência em Triagem Neonatal

Sumário

1	Introdução	1
2	Métodos	4
2.1	Área de estudo	4
2.2	Coleta dos dados	7
2.3	Fluxograma da abordagem proposta	7
2.3.1	Processamento e organização dos dados	7
2.3.1.1	Aquisição dos dados	9
2.3.1.2	Processamento e organização dos dados	9
2.3.1.3	Validação dos dados	11
2.3.2	Associação de região	12
2.3.3	Análises espaciais	15
2.3.3.1	Mapeamento coroplético	15
2.3.3.2	<i>Kernel density estimation</i>	17
2.3.3.3	Dimensão de correlação	18
3	Resultados	23
3.1	Georreferenciamento	23
3.2	Associação de região	23
3.3	Análises espaciais	28
3.3.1	Mapeamento coroplético	28
3.3.2	<i>Kernel density estimation</i>	35
3.3.3	Dimensão de correlação	37
4	Discussão	59
4.1	Georreferenciamento	59
4.2	Associação de região	61
4.3	Análises espaciais	62
4.3.1	Mapeamento coroplético	62
4.3.2	<i>Kernel density estimation</i>	63
4.3.3	Dimensão de correlação	64
5	Conclusões	67
	Referências bibliográficas	68

Apêndices	74
APÊNDICE A Spatial Workspace	75
Anexos	76
ANEXO A Parecer Consubstanciado do Comitê de Ética em Pesquisa (CEP)	77
ANEXO B Certificado de Registro de Programa de Computador	78

1 Introdução

Um dado espacial é definido por qualquer tipo de dado que esteja associado a uma área geográfica ou local específico que represente, por exemplo, a ocorrência de fenômenos, como locais de crimes, doenças, ou incêndios [(COX; ISHAM, 1980), (CÂMARA; CASANOVA; MAGALHÃES, 1996), (MONTEIRO et al., 2004)]. Por sua vez, a compreensão da distribuição espacial de fenômenos é dada a partir da análise espacial de dados, utilizando métodos estatísticos responsáveis em encontrar padrões espaciais ou anomalias [(MONTEIRO et al., 2004), (GOODCHILD, 2008)].

No entanto, compreender a distribuição espacial de fenômenos ainda é um desafio comumente relatado em diversas áreas do conhecimento, devido ao acesso a dados espaciais precisos, históricos, oportunos e compatíveis [(SOARES, 2010), (MURAD, 2018)]. Outro desafio também está relacionado à organização de dados espaciais de forma inteligente a agilizar o processo de consulta e de construção de análises espaciais [(LI et al., 2016), (BOQUETT et al., 2018)].

Na literatura é possível identificar diferentes trabalhos que abordam seus fenômenos através da análise espacial de dados, desde a área da física até a área das ciências da saúde, buscando apoiar tomadas de decisão que incidam na melhoria de políticas públicas e sociais em todo o mundo [(TELESCA et al., 2007a), (SHANMUKHAPPA; HO; TSE, 2018), (MURAD, 2018), (LAN et al., 2020), (SCHMIDT; DRÖGE-ROTHAAR; RIENOW, 2021)].

Öncel et al. (1996) e Telesca et al. (2007a), analisaram as variações espaciais de terremotos e o grau de agrupamento espacial da sequência de incêndios florestais, na Turquia e na região norte da Itália, respectivamente. As análises foram modeladas a partir de uma abordagem de análise fractal utilizando o método dimensão de correlação e dados pontuais. Análise fractal consiste em diferentes métodos que buscam atribuir uma dimensão fractal a um conjunto de dados, indicando o grau em que os objetos/eventos preenchem o espaço geográfico disponível [(MANDELBROT, 1967), (MANDELBROT; MANDELBROT, 1982)]. No entanto, os resultados obtidos apresentaram apenas uma análise temporal em nível macro na área delimitada aos países. Realizar análises de agrupamento por microrregiões e mesorregiões, possibilitaria a comparação entre as áreas mais afetadas e traria mais informações para gestores locais nas suas tomadas de decisão ou no planejamento estratégico de ações específicas para cada área em estudo.

[Murad \(2018\)](#), apresentou diferentes aplicações de sistemas de informação geográfica, ou GIS (*Geographic Information System*), para o planejamento de serviços de saúde na cidade de Jeddah, Arábia Saudita. Os problemas identificados foram modelados no *software* ArcGIS utilizando as funções de mapeamento coroplético, densidade de *kernel* e distância euclidiana (linha reta). As abordagens utilizadas se mostraram essenciais para gestores em saúde na tomada de decisão e no planejamento estratégico da cidade. No entanto, o estudo não apresentou formalmente o processo de tratamento e organização realizado sobre os dados a fim de alcançar os resultados obtidos. Descrever o processo de organização dos dados de forma clara, por meio de fluxogramas ou algoritmos, poderia facilitar a replicação do estudo para outras áreas de pesquisa.

[Moraga \(2017\)](#), apresentou uma aplicação *web* para análise de dados espaciais e espaço-temporais de doenças na área da vigilância em saúde pública. Para validar o seu estudo, utilizou dados de mortalidade por câncer de pulmão em Ohio, Estados Unidos. Os resultados obtidos pelo autor possibilitaram ajustar modelos bayesianos para obter estimativas de risco de doenças. Além disso, a disponibilização da aplicação na *web* permitiu que pesquisadores realizassem análises sofisticadas de vigilância sem a necessidade de ter habilidades estatísticas ou de programação avançadas. No entanto, a aplicação foi projetada para trabalhar apenas com dados de área. Dados de área são dados que estão associadas a uma área delimitada no espaço geográfico, por exemplo números ou valores associados a municípios, bairros ou setores censitários [[MONTEIRO et al., 2004](#)]. [Moraga \(2017\)](#) ainda reforça que, sempre que dados pontuais estiverem disponíveis, é recomendado realizar análises neste nível para que informações mais detalhadas não sejam perdidas. Dessa forma, prover meios de acesso ou geração de dados pontuais também é essencial para trazer mais intuições sobre o fenômeno em estudo.

Ainda, é possível identificar na literatura diversos trabalhos que utilizam de métodos disponíveis nos GIS e em *softwares* estatísticos, como o R, para realizar análises espaciais [[VALJAREVIĆ et al., 2018](#)], [[ALGARIN et al., 2018](#)], [[VOUGAT et al., 2019](#)], [[QUININO et al., 2021](#)]. No entanto, apesar destes métodos e ferramentas serem valiosos para o contexto de análise espacial de dados, eles podem ser complexos de serem utilizados por pessoas de baixo nível de instrução estatística ou de programação que, em alguns casos, podem ser profissionais de saúde, gestores ou pesquisadores ([MORAGA, 2017](#)). Dessa forma, facilitar o acesso a esses métodos é essencial para alcançar uma gama maior de usuários. Os trabalhos de [Bui e Pham \(2016\)](#) e [Moraga \(2017\)](#) mostraram que diferentes esforços vêm sendo feitos para desenvolver sistemas na *web* mais ativos, dinâmicos e de fácil uso para o processamento de dados e construção de análises espaciais. De acordo com [Agrawal e Gupta \(2014\)](#), [Zerihun et al. \(2017\)](#) e [Endalew, Shiferaw e Kindie \(2019\)](#), as principais vantagens da utilização de *softwares* de análises baseados na *web* incluem: disponibilização

gratuita; diminuição de custos de licença de software; aumento de segurança e estabilidade; proteção à privacidade; e maior alcance de usuários. [Highfield et al. \(2011\)](#) e [Bui e Pham \(2016\)](#), recomendam também que a construção desses sistemas seja o mais interativo possível, permitindo que os usuários definam dinamicamente seus parâmetros de análises, assim, facilitando o seu uso.

Considerando os problemas e avanços identificados nos estudos discutidos anteriormente, esta tese propõe uma abordagem baseada na *web* para processamento de dados e construção de análises espaciais de área e pontos. Para concepção das análises foram construídos dois métodos/algoritmos, um para geração de dados pontuais espaciais e um outro para associação de dados ao seu contexto espacial, ou seja, às suas microrregiões e mesorregiões. Os métodos de mapeamento coroplético; *Kernel Density Estimation* (KDE); e dimensão de correlação fractal, foram utilizados para a construção das análises. Para validação desta abordagem, foram utilizados os dados de diagnóstico de três doenças da triagem neonatal do estado do Rio Grande do Sul, Brasil, foram elas: fibrose cística; hiperplasia adrenal congênita; e hemoglobinopatias.

A discussão deste trabalho está estruturada em: Seção (2), apresenta a área de estudo e a abordagem proposta. Seção (3), apresenta os resultados obtidos em forma de mapas de área e pontos. Seção (4), discute os benefícios e as limitações da abordagem proposta. Por fim, na Seção (5), conclui-se este estudo, apresentando as indicações de uso da abordagem proposta e as principais contribuições obtidas.

2 Métodos

Esta seção está estruturada em três subseções: Seção (2.1), apresenta a área de concentração do estudo e as suas principais características. Seção (2.2), apresenta os dados, as doenças, e as variáveis estudadas neste trabalho. Por fim, na Seção (2.3), é apresentado o fluxograma da abordagem proposta e o seu funcionamento.

2.1 Área de estudo

Este estudo foi realizado no estado do Rio Grande do Sul (RS), Brasil. Situado na região sul, é o quarto maior estado do Brasil com mais de 11 milhões de habitantes e tem principal descendência italiana e alemã. Atualmente, encontra-se distribuído em 497 municípios e em 7 mesorregiões (IBGE, 2010), como apresentado na Figura (1).



Figura 1 – Mapa de localização do estado do Rio Grande do Sul – Brasil, com as suas mesorregiões.

No ano de 2001, o Serviço de Referência em Triagem Neonatal (SRTN) foi implantado na capital do estado, Porto Alegre, no Hospital Materno-Infantil Presidente Vargas (HMIPV) – Figura (2). SRTN é um serviço responsável pela análise e diagnóstico de recém-nascidos (RN) portadores de seis doenças triadas no “teste do pezinho” da rede pública saúde, são elas: fibrose cística; hipotireoidismo congênito; deficiência de biotinidase; hemoglobinopatias; fenilcetonúria; e hiperplasia adrenal congênita (KOPACEK et al., 2015).

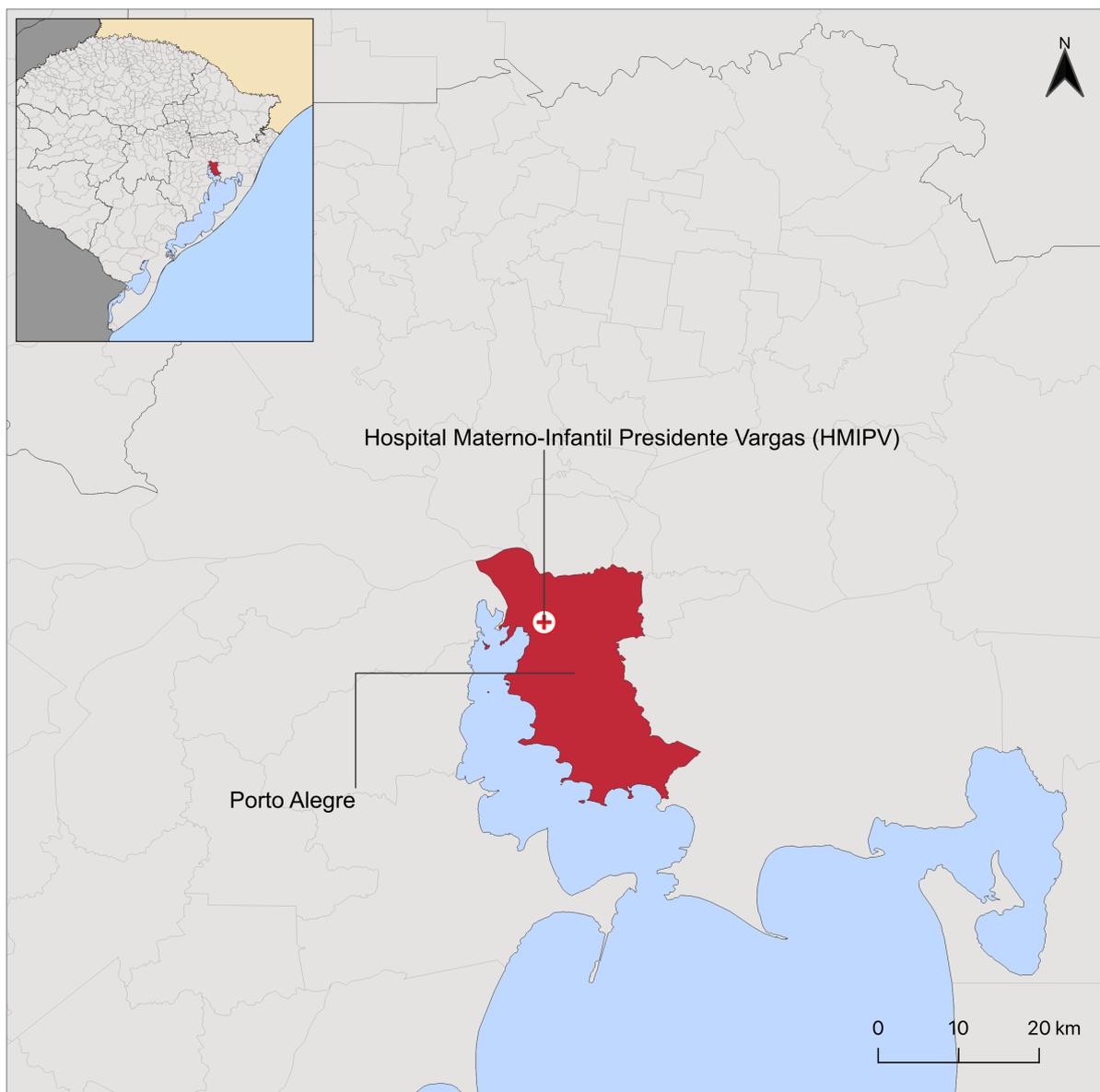


Figura 2 – Mapa de localização da capital Porto Alegre e do Hospital Materno-Infantil Presidente Vargas (HMIPV).

A estratégia da triagem neonatal prevê que o diagnóstico e o início do tratamento ocorra antes da manifestação dos sintomas, reduzindo, desta forma, a morbimortalidade infantil e as sequelas específicas de cada doença (BOTLER, 2010). O SRTN recebe todos os dados do teste do pezinho realizados em 1.307 postos de coleta do estado em um volume que corresponde a aproximadamente 75% dos RN do estado do Rio Grande do Sul, média de 105.000 nascimentos por ano.

2.2 Coleta dos dados

Este estudo incluiu dados de pacientes de três doenças: fibrose cística; hiperplasia adrenal congênita; e hemoglobinopatias, coletados no período de 2004 a 2020. Ao total, foram obtidos 405 registros das três doenças especificadas. As variáveis associadas aos registros e trabalhadas neste estudo foram: endereço e patologia dos pacientes. Os dados foram obtidos em forma de planilha excel (.xlsx) no SRTN.

Este estudo foi aprovado pelo Comitê de Ética do Grupo de Pesquisa e Pós-Graduação do Hospital Materno-Infantil Presidente Vargas de Porto Alegre, Rio Grande do Sul, sob o número 4.397.969.

2.3 Fluxograma da abordagem proposta

O fluxo da abordagem proposta neste trabalho foi dividida em três etapas, foram elas: **(i)** processamento e organização dos dados; **(ii)** associação de região; e **(iii)** análises espaciais, conforme a Figura (3). A seguir as etapas serão apresentadas e detalhadas.

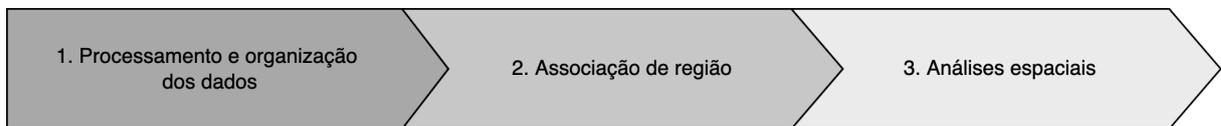


Figura 3 – Fluxograma geral da abordagem proposta.

2.3.1 Processamento e organização dos dados

O processamento realizado sobre os dados foi realizado seguindo o fluxograma apresentado na Figura (4), a partir de três passos: **(i)** aquisição dos dados; **(ii)** processamento e organização dos dados; e **(iii)** validação dos dados.

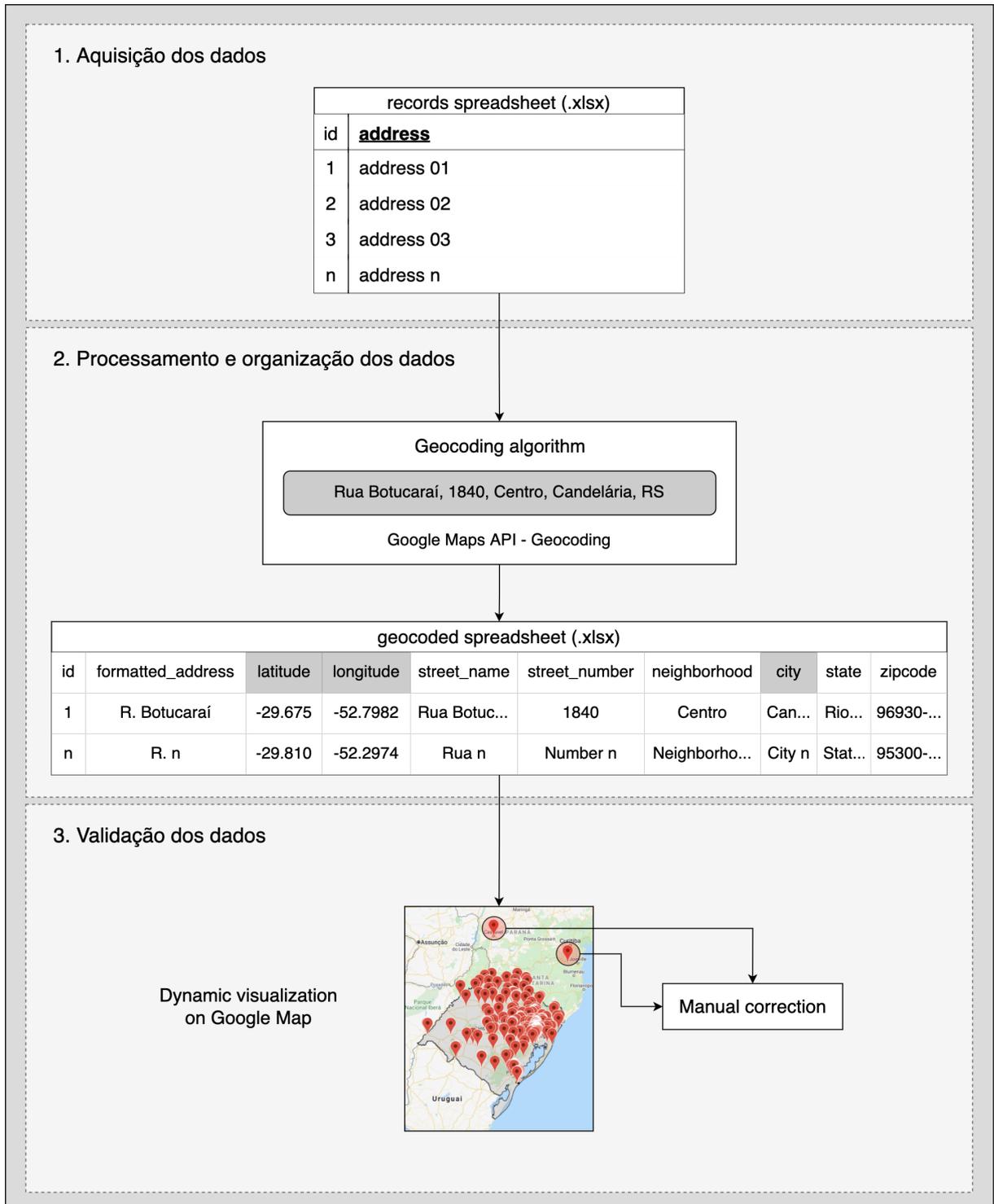


Figura 4 – Fluxograma de processamento e organização dos dados.

2.3.1.1 Aquisição dos dados

Inicialmente, os dados coletados no SRTN em forma de planilha eletrônica (.xlsx) foram importados numa interface *web* de aquisição de dados disponível em <<https://spatialworkspace.ddns.net/geocoding>> e submetidos ao processo de processamento de dados. Neste momento, foram selecionadas duas colunas de dados da planilha eletrônica: **(i)** a coluna que identifica o endereço de localização dos pacientes; e **(ii)** a coluna que identifica o registro. O endereço foi utilizado no processo de georreferenciamento dos registros. Georreferenciamento é o processo de conversão de endereços, como "1600 Anfiteatro Parkway, Mountain View, CA" para coordenadas geográficas, como "latitude 37.423021 e longitude -122.083739" (YANG et al., 2004). O identificador do registro foi uma variável única utilizada para diferenciar um registro dos demais. Além disso, a variável única foi utilizada para associar aos registros seus possíveis atributos relevantes, como exemplo: raça e tipo de doença, em processos subsequentes.

2.3.1.2 Processamento e organização dos dados

Em seguida, os dados foram processados e organizados. Durante o processamento, o algoritmo disponível na *web* realizou uma iteração sobre os registros, percorrendo um por um, realizando uma chamada HTTP para a *Application Programming Interface* (API) de geocodificação do Google *Maps*. A API de geocodificação do Google *Maps* é um serviço responsável em converter endereços em coordenadas geográficas (PLATFORM, 2021). HTTP é um protocolo base para troca de dados na *web* (DOCS, 2021). A chamada para API foi realizada atribuindo a variável de endereço (texto) como parâmetro da requisição. Requisição, na área da computação, é o processo intermediário entre a comunicação do cliente (página na *web* ou app) com o servidor, ou seja, entre a nossa interface na *web* e a API de geocodificação do Google *Maps*.

Ao fim do processamento, foi disponibilizado para *download* uma nova planilha eletrônica com os dados geocodificados. A nova planilha possuía duas folhas, uma de falha e uma outra de sucesso. A folha de falha apresentou os endereços em que a API não retornou resultados de georreferenciamento. Os registros que não conseguiram ser georreferenciados, passaram por intervenção manual para correção de possíveis erros de ortografia e em seguida submetidos novamente ao processo de georreferenciamento automático. Os dados processados e organizados com sucesso foram separados nas seguintes variáveis/colunas na nova planilha eletrônica: endereço formatado; latitude e longitude; nome da rua; número da casa; bairro; cidade; estado; e *zipcode* ou Código de Endereçamento Postal (CEP).

A escolha da API do Google *Maps* de geocodificação para este processo foi baseada na análise e experimentação de seis serviços comerciais de geocodificação, foram eles:

LocationIQ (2021), OpenCage (2021), ArcGIS (2021), HERE (2021), Google (2021) e Bing (2021). A API do Google *Maps* de geocodificação apresentou maior precisão no georreferenciamento, ou seja, correspondência entre o endereço local e as suas coordenadas geográficas, e uma maior variedade de formatos de resposta suportados. Além disso, possui 99% de cobertura no mundo, dados confiáveis e abrangentes de mais de 200 países e territórios (CLOUD, 2021).

O algoritmo foi desenvolvido utilizando a linguagem de programação para *web* JavaScript que possibilitou calcular, manipular e validar dados (W3SCHOOLS, 2021). Além disso, foi utilizada a biblioteca React, construída em JavaScript, que permitiu criar interfaces de usuário em páginas na *web* (REACT, 2021). A Figura (5) apresenta a interface disponível na *web* e o seu <link> de acesso para processamento de dados e geração de dados espaciais.

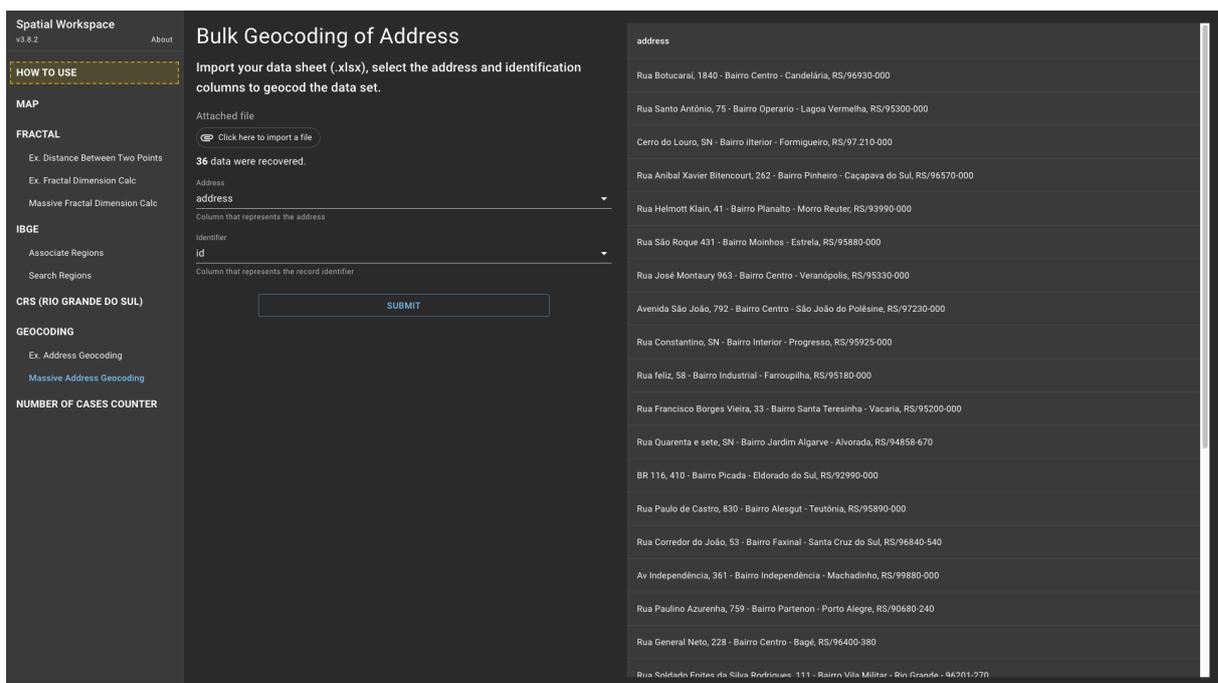


Figura 5 – Interface na *web* para processamento de dados e geração de dados espaciais <<https://spatialworkspace.ddns.net/geocoding>>.

Este método/algoritmo foi denominado ASSYRIA, a partir do acrônimo de “*dAta proceSsing SYstem foR spatIal Analysis*”. O acrônimo foi gerado no *website* Acronymify! (2021).

2.3.1.3 Validação dos dados

Na validação da precisão dos dados resultantes do processo de geocodificação, foi utilizado um processo de correções pós-manuais com base em [Goldberg et al. \(2008\)](#), que propuseram uma correção manual dos dados geocodificados incorretamente por meio de uma abordagem interativa na *web*. Este processo foi realizado a partir da visualização dinâmica dos dados geocodificados no mapa do Google, em que, os dados visualizados fora da área delimitada ao estudo, passaram por correção manual. Uma versão *demo* deste processo pode ser acessada em <https://spatialworkspace.ddns.net/map>, no qual, foi possível importar uma planilha eletrônica de dados georreferenciados e visualizar dinamicamente os dados no mapa do Google. A Figura (6) também exemplifica um cenário em que três coordenadas geográficas estão fora da área de estudo e que devem passar pelo processo de correção manual.

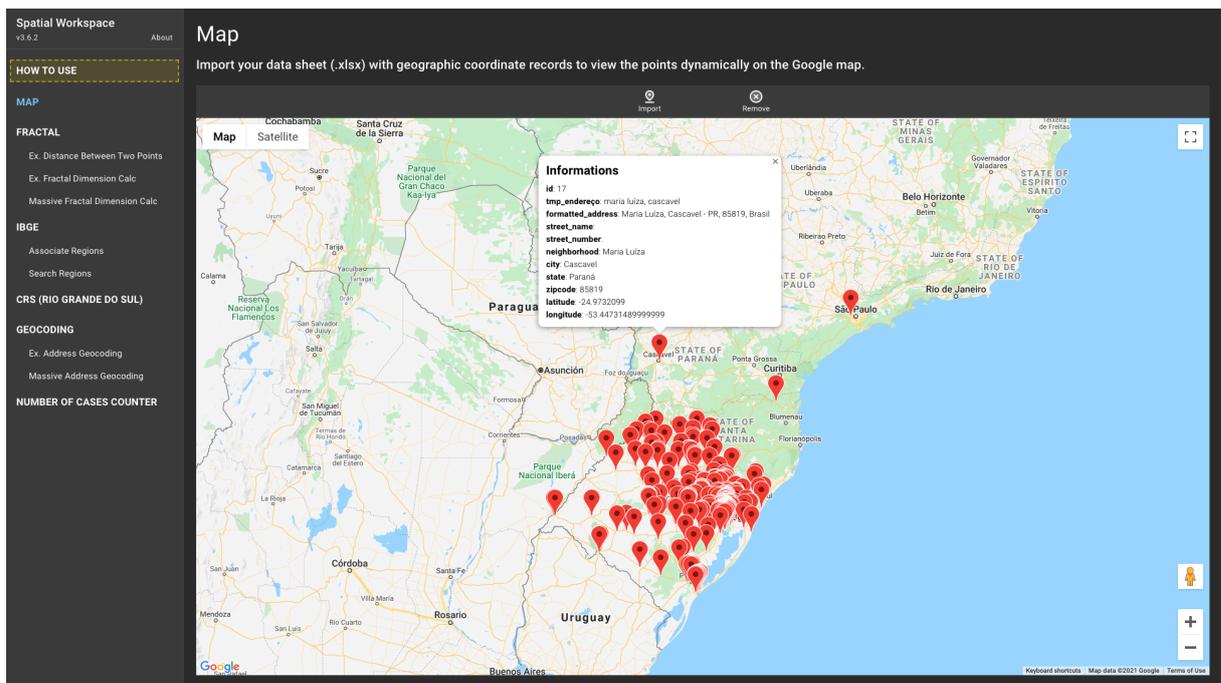


Figura 6 – Interface na *web* para validação de dados espaciais <https://spatialworkspace.ddns.net/map>.

Esta visualização possibilitou observar dados com possíveis erros de posicionamento ou georreferenciamento. Os dados identificados fora da área de estudo ou com possíveis erros de georreferenciamento, passaram por intervenção manual. As correções foram realizadas utilizando uma versão *web demo* de correção manual. A versão *demo* pode ser vista em <https://spatialworkspace.ddns.net/address> e está apresentada na Figura (7). Este processo assegura que os dados que irão para o processo de análise estejam precisos e

dentro do cenário de estudo, evitando possíveis erros de compilação ou interpretação no processo de análise espacial dos dados.

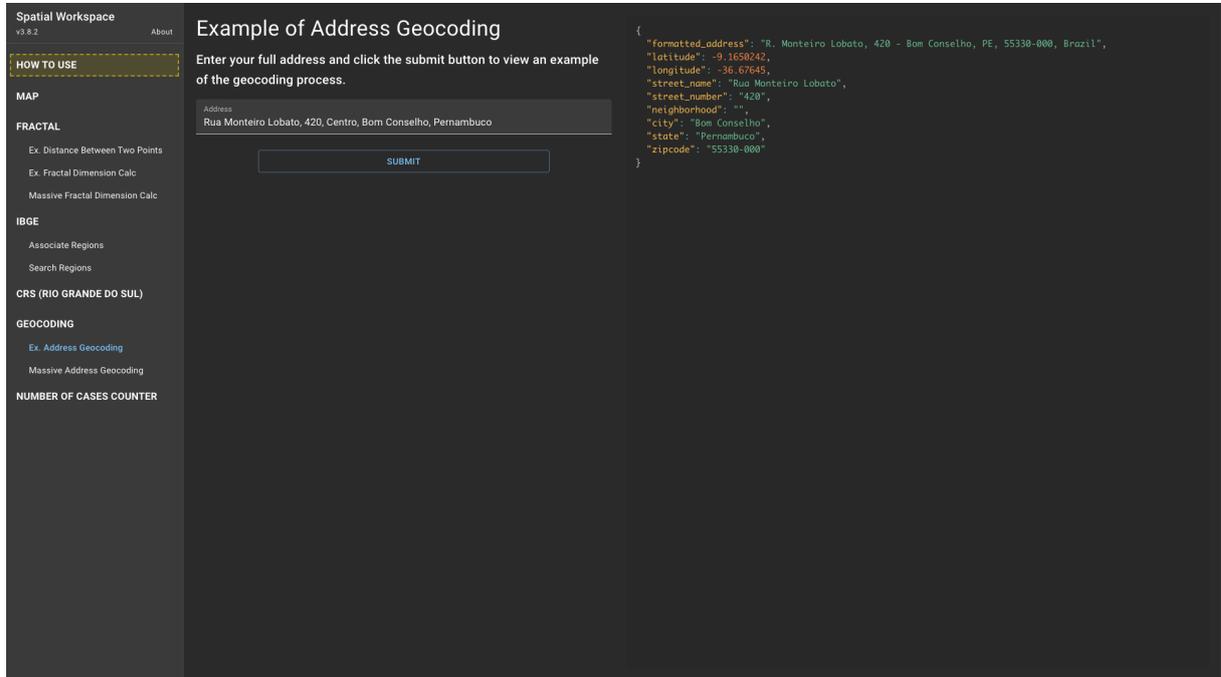


Figura 7 – Interface na *web* para geocodificação manual de endereços <<https://spatialworkspace.ddns.net/address>>.

2.3.2 Associação de região

A associação dos dados às suas regiões administrativas seguiu o fluxo apresentado na Figura (8), a partir de duas etapas: **(i)** importação dos dados; e **(ii)** associação dos dados.

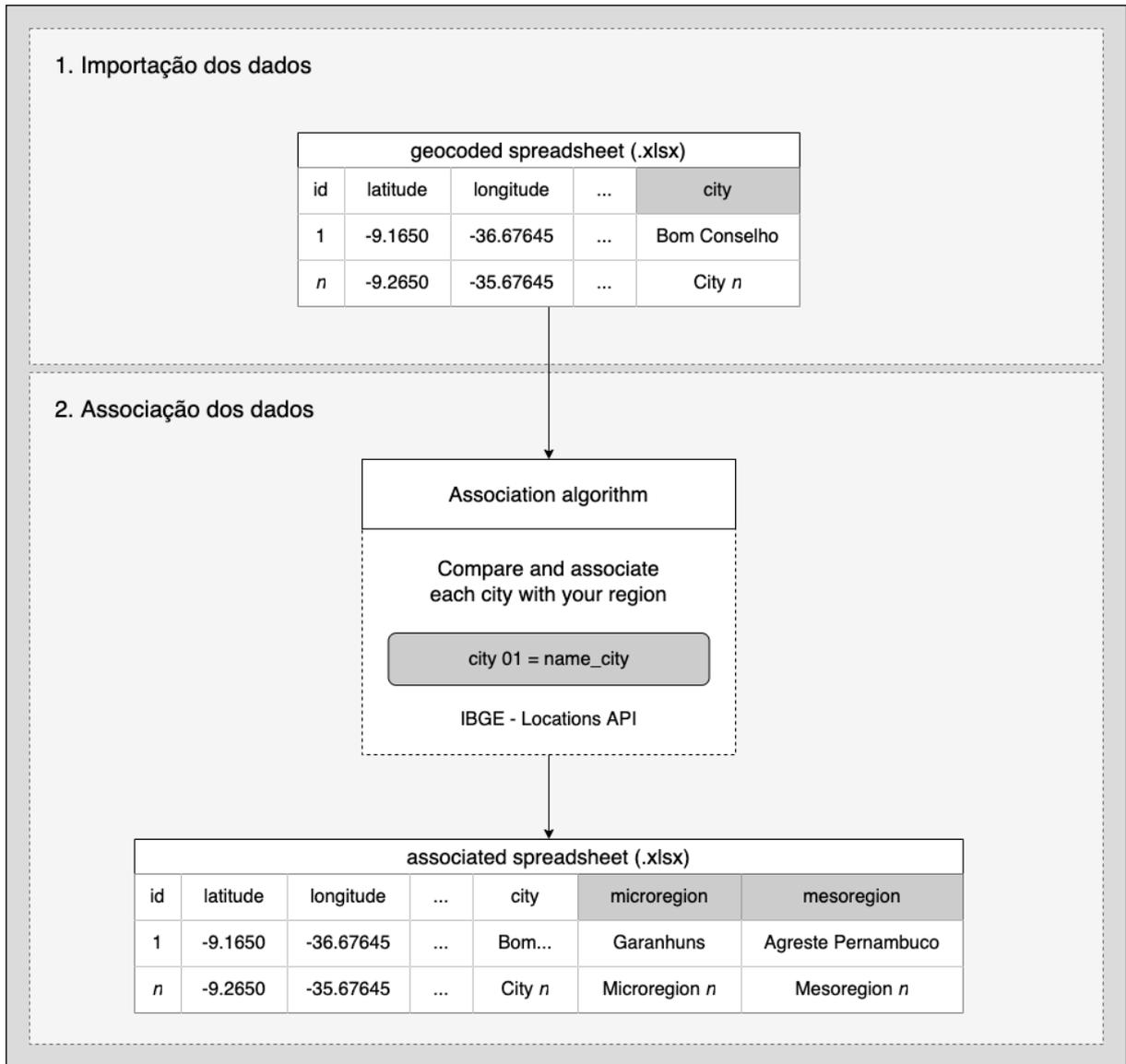


Figura 8 – Fluxograma de associação de região.

No primeiro momento, os dados resultantes da Seção (2.3.1), processamento e organização dos dados, foram importados numa outra interface *web*, disponível em <<https://spatialworkspace.ddns.net/region>>, para associação dos dados às suas mesorregiões e microrregiões da divisão político-administrativa do estado do Rio Grande do Sul. Para realizar a associação foi utilizada a coluna que representava a cidade dos registros.

Posteriormente, foi utilizado o algoritmo denominado *Associate Regions* (AR) para associar o atributo cidade (texto) às suas respectivas regiões. A associação foi feita buscando, de forma automática, as regiões correspondentes à variável cidade na API de localidades do Instituto Brasileiro de Geografia e Estatística (IBGE). A API de localidades do IBGE

é um serviço público do Brasil que disponibiliza na *web* informações das mesorregiões e microrregiões de todas as cidades das 26 unidades federativas brasileiras (IBGE, 2021a).

A associação realizada gerou mais cinco campos associados ao atributo cidade, foram eles: identificador da cidade/município; nome da microrregião; identificador da microrregião; nome da mesorregião; e identificador da mesorregião. Os dados/campos gerados foram salvos em uma nova planilha eletrônica (.xlsx) e disponibilizada para *download*. Uma versão *demo* deste processo pode ser acessada em <<https://spatialworkspace.ddns.net/region>> e visualizada na Figura (9).

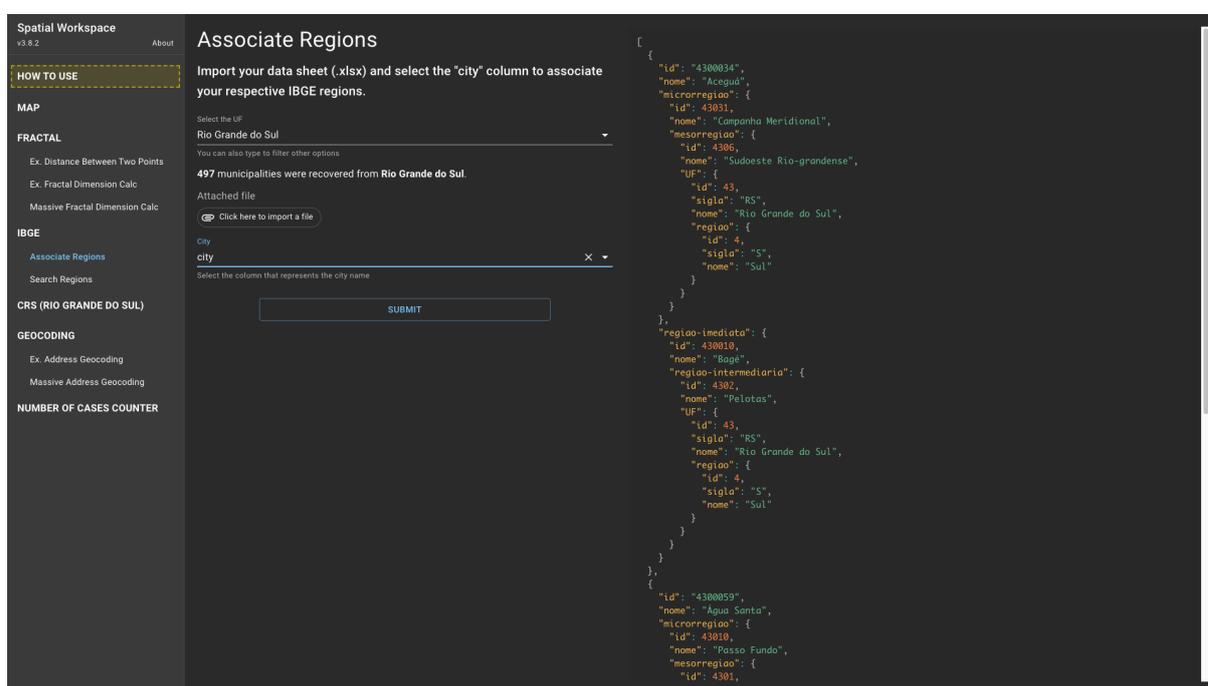


Figura 9 – Interface na *web* para associação de dados às suas regiões <<https://spatialworkspace.ddns.net/region>>.

Os registros em que possivelmente falhassem no processo automático de associação de região, eram disponibilizados na folha de fracassos do arquivo de planilha eletrônica. Os registros que apresentavam falha, por sua vez, eram buscados de forma manual em <<https://spatialworkspace.ddns.net/search>> e, em seguida, incorporados ao conjunto de dados de sucesso do estudo. A Figura (10) demonstra o processo de busca manual de regiões na *web*.

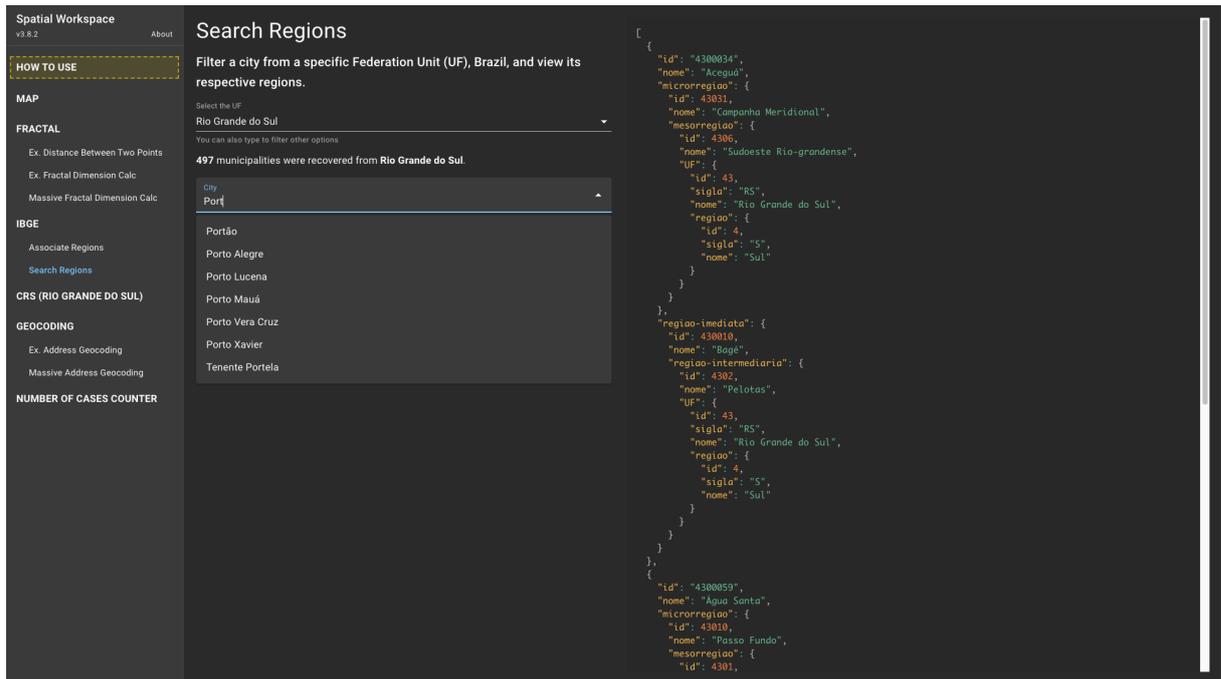


Figura 10 – Interface na *web* de busca manual de regiões <<https://spatialworkspace.ddns.net/search>>.

2.3.3 Análises espaciais

Neste estudo foram utilizados três métodos de análise espacial de dados: **(i)** mapeamento coroplético; **(ii)** *kernel density estimation*; e **(iii)** dimensão de correlação. O mapeamento coroplético foi utilizado para analisar os dados de área. Já os métodos KDE e dimensão de correlação, foram utilizados para analisar os dados de eventos pontuais. A seguir os métodos serão apresentados e descritos.

2.3.3.1 Mapeamento coroplético

Segundo MacEachren (1994), o objetivo do mapeamento coroplético é apresentar dados numéricos sobre áreas utilizando cores distintas, em que, cores mais claras representam valores numéricos mais baixos e cores mais escuras representam valores mais altos da variação numérica do fenômeno em estudo.

Normalmente, o mapeamento coroplético é utilizado para explorar visualmente o padrão espacial de distribuição de atributos entre regiões, por exemplo: atributos demográficos, como "densidade populacional e população por sexo", e socioeconômicos, como "renda per capita e Índice de Desenvolvimento Humano (IDH)" [(WEI; TONG; PHILLIPS, 2017), (MURAD, 2018)].

Neste estudo, o mapeamento coroplético foi utilizado para observar: **(i)** a distribuição espacial do número total de casos das três doenças; e **(ii)** a sua prevalência por município, microrregião e mesorregião no estado do Rio Grande do Sul, Brasil.

Para a construção dos mapas de distribuição espacial do número total de casos, foram gerados três indicadores, foram eles: número total de casos por: **(i)** município; **(ii)** microrregião; e **(iii)** mesorregião. Para contagem do número de casos, foi utilizado um algoritmo de contagem automática na *web*. Uma versão *demo* deste algoritmo pode ser acessada em <https://spatialworkspace.ddns.net/cases> e visualizada na Figura (11).

nm_meso	cases
Centro Oriental Rio-grandense	31
Nordeste Rio-grandense	34
Centro Ocidental Rio-grandense	14
Sudeste Rio-grandense	45
Metropolitana de Porto Alegre	215
Noroeste Rio-grandense	42
Sudoeste Rio-grandense	24

Figura 11 – Interface na *web* para contagem automática de número de casos <https://spatialworkspace.ddns.net/cases>.

Já para a construção dos mapas coropléticos de distribuição espacial da prevalência, foi utilizado o Sistema IBGE de Recuperação Automática (SIDRA), censo demográfico de 2010, para obtenção de dados populacionais (SIDRA, 2021). Utilizando o SIDRA foi obtido o número total da população residente em área rural e urbana por município, microrregião e mesorregião. O cálculo da prevalência, apresentado na Equação (2.1), demonstra a proporção da frequência de uma doença sobre a população de uma região em um determinado período de tempo (JOHN, 2001). As constantes populacionais de 10.000, 100.000 e 200.000 foram utilizadas, respectivamente, para município, microrregião e mesorregião. A escolha dos valores destas constantes foi dada a partir de experimentação,

baseando-se na população residente no estado do Rio Grande do Sul, Brasil, e nas características de incidência das doenças.

$$prevalence = \frac{\text{number of existing cases at any given time}}{\text{number of people in the population at the same time}} * constant \quad (2.1)$$

O GIS gratuito e de código aberto QGIS, versão 3.10 (QGIS, 2021), foi utilizado para construção dos mapas. Ambos os mapas, de número de casos e prevalência, foram classificados em três categorias (baixo, moderado e alto) utilizando a classificação de Quebras Naturais de Jenk (DENT; TORGUSON; HODLER, 1993). O método de Quebras Naturais de Jenk é um método de classificação de dados projetado para reduzir a variância dentro das classes e maximizar a variância entre as classes [(JENKS, 1967), (MCMMASTER, 1997)]. Os *Shape File* (SHP) da divisão político-administrativa do estado do Rio Grande do Sul, Brasil, ano de 2020, em nível de mesorregião, microrregião e município, foram obtidos no IBGE (IBGE, 2021b).

2.3.3.2 Kernel density estimation

Apesar do mapeamento coroplético ser uma tradicional forma de observar dados agregados por região ou área, como taxas de contagens e proporções, este tipo de mapeamento ainda é limitado para estudar outros desafios na saúde (AZEVEDO et al., 2020). Pois, os riscos e eventos por área podem mudar diante de diferentes limites espaciais, por exemplo, dentro dos limites de uma grande metrópole.

Então, para complementar as análises deste estudo, foi utilizado o método não paramétrico KDE para explorar a densidade espacial dos nossos dados pontuais. Densidades de *kernel* foram frequentemente utilizadas na análise de eventos pontuais e exploração de focos de calor em uma variedade de domínios, incluindo criminologia (LEVINE, 2006), epidemiologia espacial [(CHAIKAEW; TRIPATHI; SOURIS, 2009), (CROMLEY; MCLAFFERTY, 2011), (SHIODE et al., 2015), (MURAD, 2018)] ou ecologia (WORTON, 1989). Além disso, trata-se de um método de fácil uso e interpretação (MONTEIRO et al., 2004). Formalmente este método é dado pela Equação (2.2).

$$\hat{\lambda}_\tau(u) = \frac{1}{\tau^2} \sum_{i=1}^n k \left(\frac{d(u_i, u)}{\tau} \right), \quad d(u_i, u) \leq \tau \quad (2.2)$$

Em que u_1, \dots, u_n são localizações de eventos pontuais observados numa determinada região. O estimador de densidade é computado a partir dos m eventos u_i, \dots, u_{i+m-1} contidos num raio de tamanho τ em torno de u e da distância d entre a posição e a i -ésima amostra. Assim, um *kernel* circular é formado em torno de cada ponto gerador (por

exemplo, locais de vítimas de assassinato) com uma largura de banda predefinida como seu raio (SHIODE et al., 2015). Sendo assim, a densidade da superfície do mapa é produzida a partir da contagem de todos os pontos dentro do raio de influência, ponderando-os pela distância de cada um à localização de cada ponto gerado.

Os parâmetros básicos deste método, são: a) raio de influência ($\tau \geq 0$); e b) função de estimação do *kernel*. O raio de influência define a área centrada no ponto de estimação u , que indica quantos eventos u_i contribuem para a estimativa da função de densidade λ . Normalmente, a largura de banda, ou tamanho do raio, é decidida de maneira exploratória, em que, o KDE é produzido utilizando diferentes tamanhos de largura de banda (raio) para identificação empírica de um valor adequado para o contexto em estudo (SHIODE et al., 2015). Valores pequenos de raio resultam em superfícies descontínuas, e valores altos demais resultam em superfícies amaciadas. As funções de cálculo de estimação do *kernel* são: `uniform`; `triangle`; `Epanechnikov`; `quartic (biweight)`; `tricube`; `triweight`; `Gaussian`; `quadratic`; e `cosine`. No entanto, normalmente, é utilizado funções de terceira ou quarta ordem ou *kernel* gaussiano para estimação do *kernel* (MONTEIRO et al., 2004).

O GIS gratuito e de código aberto QGIS, versão 3.10 (QGIS, 2021), também foi utilizado para construção desta análise. A função de densidade de *kernel* foi utilizada para visualizar os focos de calor ou áreas com maior densidade de número de casos das três doenças estudadas neste trabalho. Os parâmetros definidos para esta análise foram: a) largura de banda/tamanho do raio = 10 km; e b) função de quarta ordem para estimação do *kernel*. Os registros dos eventos pontuais utilizados nesta análise foram resultantes do processo de geocodificação realizado na Seção (2.3.1), processamento e organização dos dados. O *shape file* da divisão político-administrativa em nível municipal (ano de 2020) e a população residente em cada área, foram utilizados como base para a visualização dos focos de calor gerados. O *shape file* foi obtido no acervo de malhas municipais do IBGE (IBGE, 2021b).

2.3.3.3 Dimensão de correlação

A fim de enriquecer e ampliar a discussão deste trabalho, também foi utilizado o método de dimensão de correlação (*Correlation Dimension*), da análise fractal de dados, para verificar o grau de agrupamento entre os eventos/casos das doenças. O cálculo do grau de agrupamento, Equação (2.3), é baseado na integral de correlação de Grassberger e Procaccia (2004), utilizada para estimar a dimensão de correlação D_c dos eventos.

$$D_c = \lim_{r \rightarrow 0} \frac{\log C(r)}{\log r} u \quad (2.3)$$

Em que, $C(r)$ mede o grau de correlação de agrupamento de um conjunto de pontos, dado uma distância angular r entre os eventos. Formalmente este cálculo é dado pela Equação (2.4):

$$C(r) = \frac{2N_{R<r}}{N(N-1)} \quad (2.4)$$

Onde, $N_{R<r}$ representa o número de pares de eventos separados por uma distância R menor que r ; N representa o número total de eventos, ou seja, casos das doenças. A distância angular r em graus entre dois eventos pontuais é dada pela Equação (2.5):

$$r = R \cos^{-1}[\cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2 \cos(\phi_1 - \phi_2)] \quad (2.5)$$

Em que, (θ_1, θ_2) e (ϕ_1, ϕ_2) são as colatitudes e longitudes dos dois eventos. A Figura (12) exemplifica a distância angular entre dois eventos e também pode ser acessada em <https://spatialworkspace.ddns.net/distance> para fins de ilustração didática.

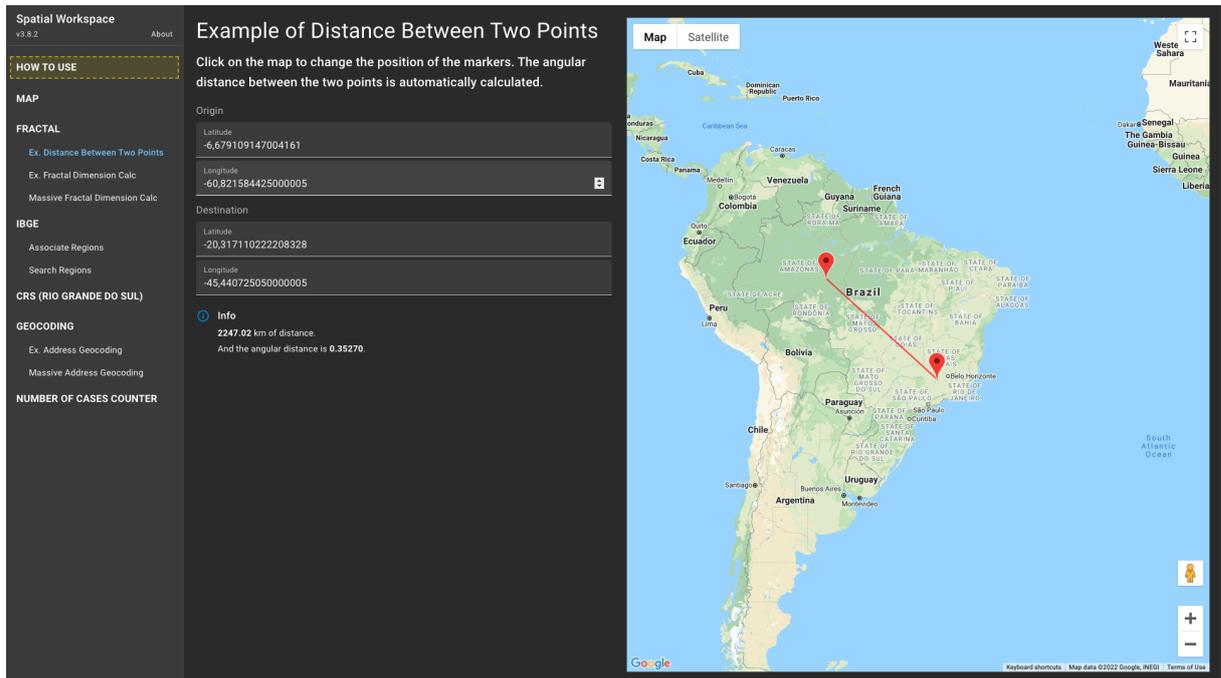


Figura 12 – Interface na *web* de exemplo de distância angular entre dois eventos <https://spatialworkspace.ddns.net/distance>.

Para validar o cálculo da função $C(r)$, que mede o grau de correlação de agrupamento dos eventos, foi criado um exemplo *demo* na *web* que pode ser acessado em <https://spatialworkspace.ddns.net/dimension>. Através de cliques sobre o mapa é possível

adicionar eventos pontuais. O cálculo da dimensão fractal é realizado automaticamente. Variações entre a correlação da distância entre os eventos pode ser realizada no *input* do tipo *slide*. Os valores de $C(r)$ variam entre 0 e 2, considerando um plano bidimensional. Em que, $C(r) = 0$ aponta que todos os eventos estão concentrados (agrupados) em um único ponto; e $C(r) = 2$ indica que todos os eventos cobrem homogeneamente todo o espaço bidimensional (TOSI, 1998). A Figura (13) ilustra o funcionamento do cálculo e também pode ser utilizada para fins educativos de ensino.

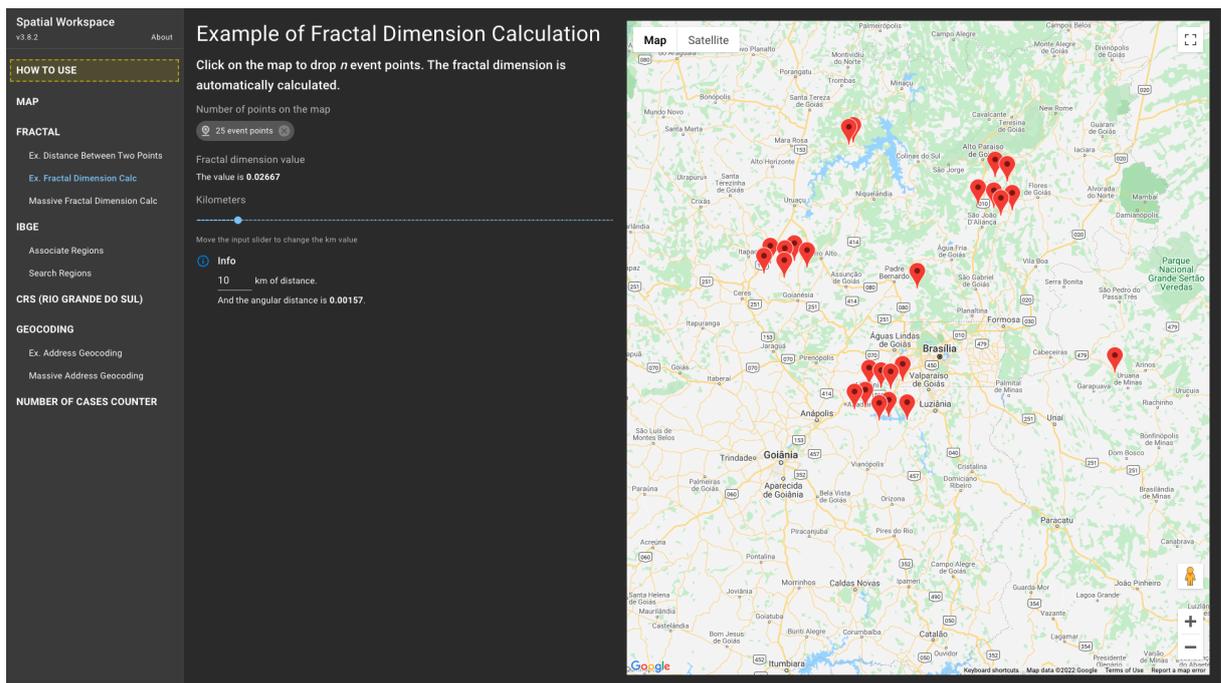


Figura 13 – Interface na *web* de exemplo do cálculo de agrupamento da dimensão de correlação fractal <<https://spatialworkspace.ddns.net/distance>>.

Este método foi desenvolvido e disponibilizado na *web* para a construção automática de análises de dimensão de correlação fractal a partir de um conjunto de dados de eventos pontuais. Uma versão *demo* deste processo pode ser acessada em <<https://spatialworkspace.ddns.net/fractal>>. Em que, os dados podem ser importados em formato de planilha eletrônica (.xlsx) e automaticamente a dimensão de correlação fractal entre os eventos é calculada por região de interesse, conforme exemplificado nas Figuras (14) e (15).

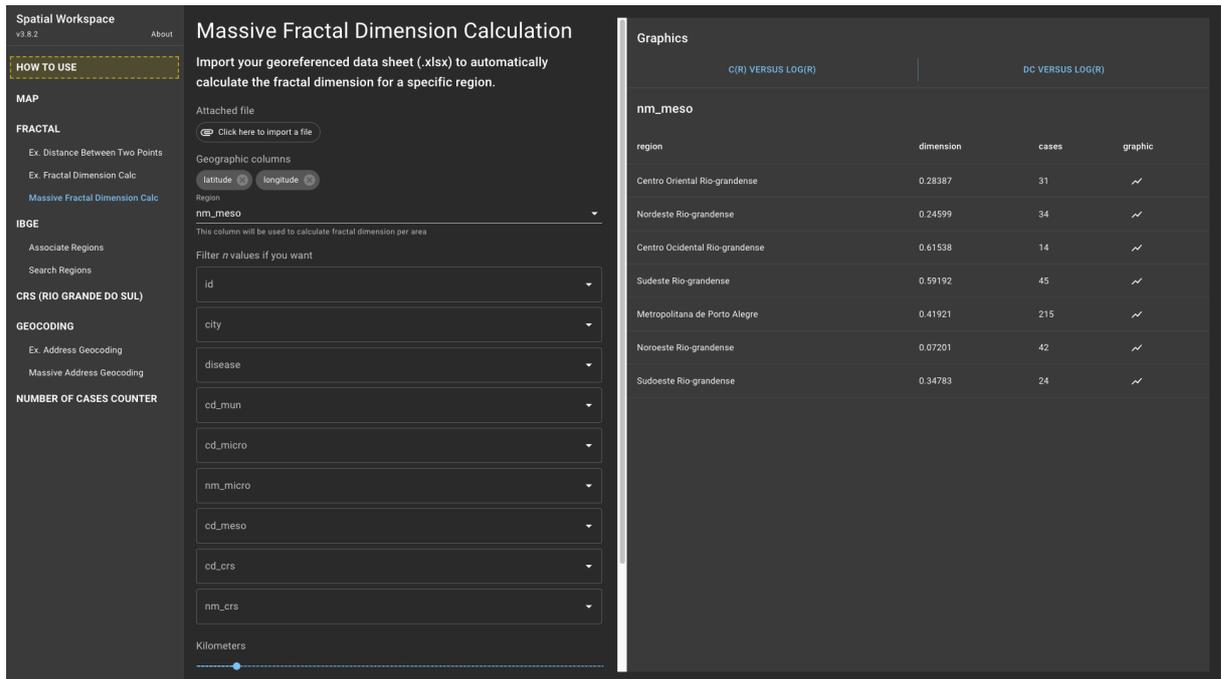


Figura 14 – Interface (1) na *web* de cálculo automático da dimensão de correlação fractal <<https://spatialworkspace.ddns.net/fractal>>.

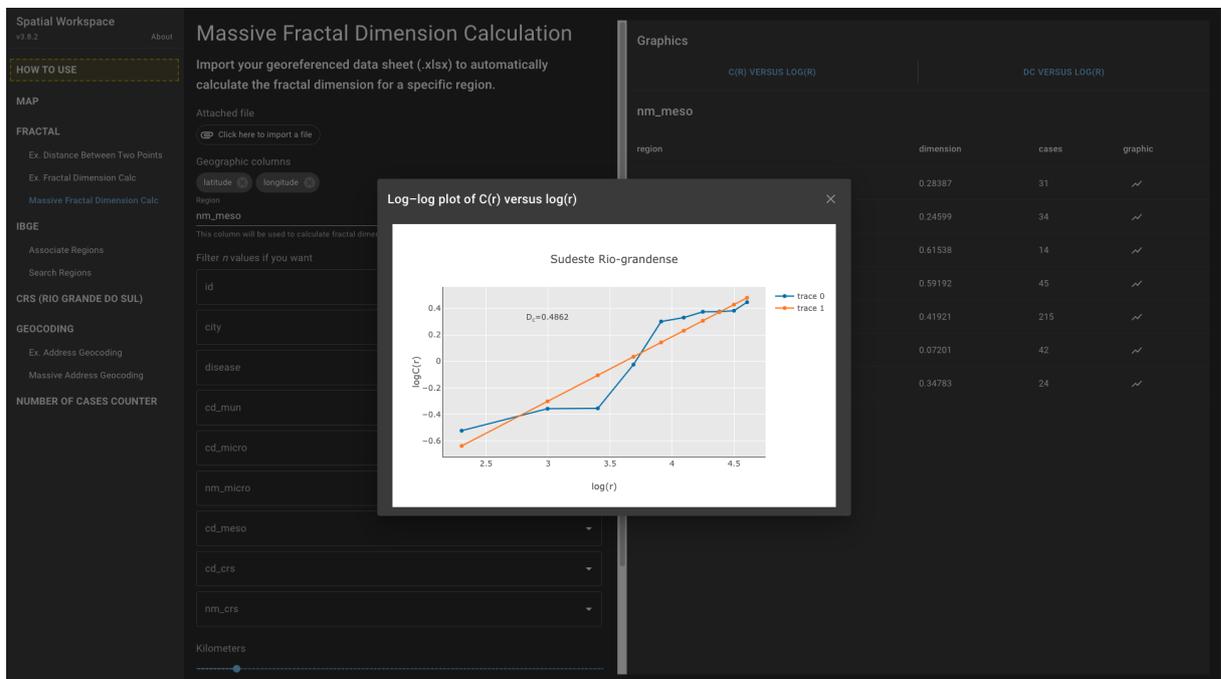


Figura 15 – Interface (2) na *web* de cálculo automático da dimensão de correlação fractal <<https://spatialworkspace.ddns.net/fractal>>.

Para a construção dos gráficos na *web*, foi utilizada a biblioteca JavaScript de construção de gráficos Plotly ([PLOTLY, 2022](#)). Além disso, para consistência desta análise, todos os dados foram analisados numa mesma escala de comprimento ($10 < r < 100$ km) e de modo comparativo entre as microrregiões e mesorregiões do estado.

3 Resultados

Esta seção está estruturada em três subseções: Seção (3.1), apresenta os resultados de sucesso e falha no processo de georreferenciamento dos dados. Seção (3.2), apresenta os resultados obtidos no processo de associação de região. Já a Seção (3.3), apresenta as principais observações percebidas utilizando os métodos de: mapeamento coroplético; KDE; e dimensão de correlação.

3.1 Georreferenciamento

Utilizando a API de geocodificação do Google *Maps*, foi obtido sucesso no georreferenciamento de 401 registros, dos 405 endereços estudados neste trabalho. Dos 4 registros que apresentaram falhas, 2 falharam no processo automático de geocodificação e os outros 2 foram georreferenciados fora da área de estudo. Os registros de falha foram corrigidos, em possíveis erros de ortografia, e novamente submetidos ao processo de georreferenciamento manual. Sendo assim, o processo automático teve 99,01% de sucesso no georreferenciamento e menos de 1% de falha neste conjunto de dados.

3.2 Associação de região

Após o processo de georreferenciamento e organização dos dados, descrito na Seção (2.3.1), as regiões da divisão político-administrativa do estado do Rio Grande do Sul foram associadas aos 405 registros. Destes dados, 398 obtiveram sucesso no processo de associação automática e apenas 7 registros falharam no processo automático. Os registros que falharam no processo de associação foram associados de forma manual. As falhas aconteceram devido a erros de acentuação e presença de caracteres especiais nos registros obtidos na API de localidades do IBGE. Dessa forma, o processo automático de associação teve 98,27% de sucesso e menos de 2% de falha neste conjunto de dados.

A Figura (16) mostra a localização específica dos casos das doenças sobre a área delimitada ao estado do Rio Grande do Sul, evidenciada pelas suas sete mesorregiões. Ao observar a figura foi possível identificar uma maior concentração de eventos na mesorregião Metropolitana de Porto Alegre. Na mesorregião Noroeste Rio-grandense, visualizou-se eventos dispersos por toda a região. Na mesorregião Sudeste Rio-grandense, os eventos

encontravam-se mais agrupados. Já nas outras mesorregiões, visualizou-se eventos mais dispersos e também concentrados em algumas delimitações.

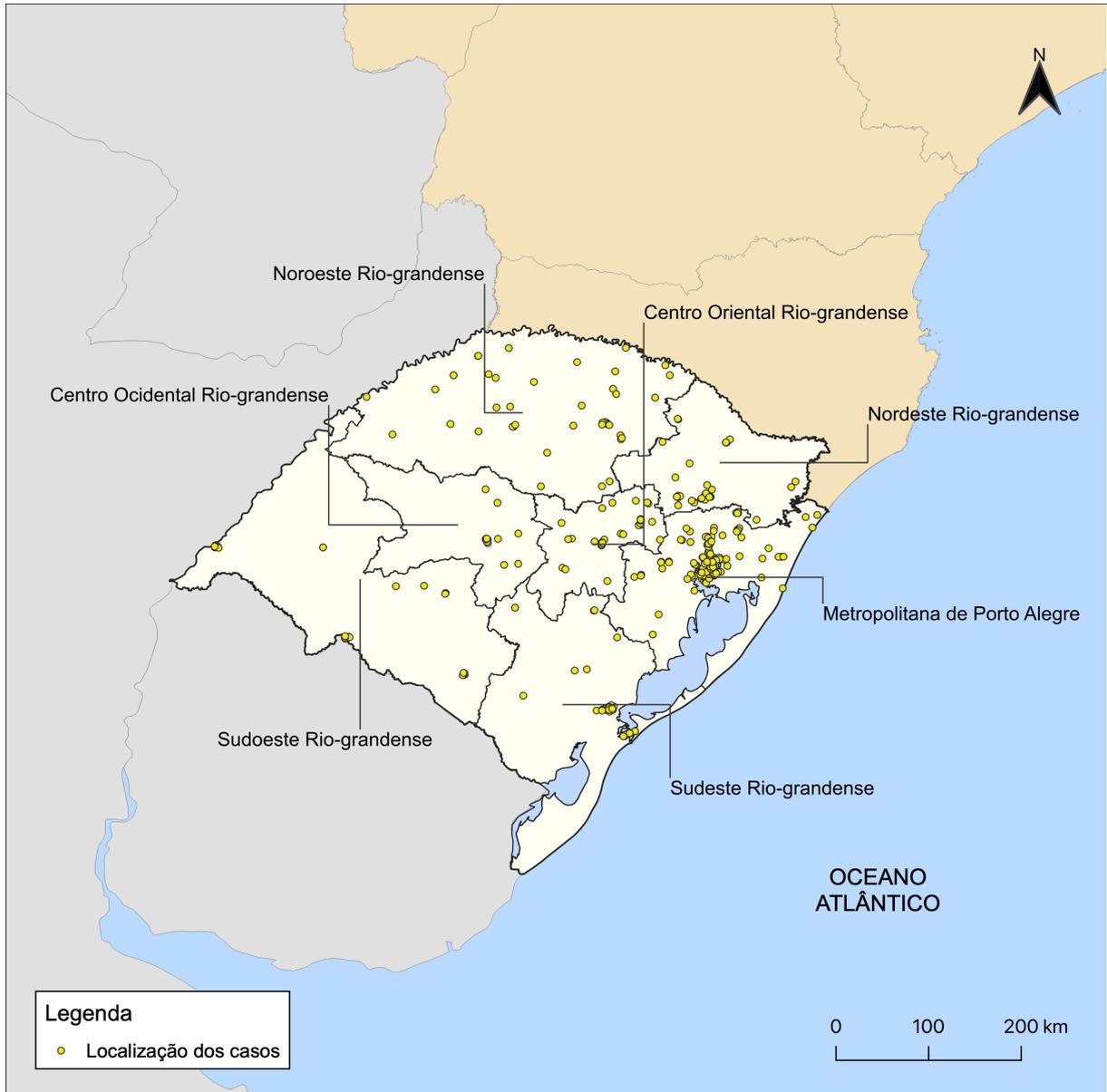


Figura 16 – Mapa de distribuição dos casos das doenças por mesorregião.

A Tabela (1) apresenta o número de casos das doenças associados a cada mesorregião. A tabela se encontra classificada em ordem decrescente pelo número de casos, ou seja, das mesorregiões que apresentaram o maior número de casos até a mesorregião que apresentou o menor número. Quatro das sete mesorregiões apresentaram número de casos que variaram entre 31 e 45, foram elas: Centro Oriental, Nordeste, Noroeste e Sudeste. As mesorregiões

Sudoeste e Centro Ocidental apresentam o menor número de casos, abaixo de 24. Já a mesorregião Metropolitana apresentou o maior número, ou seja, 215 casos.

Tabela 1 – Número de casos das doenças por mesorregião

Mesorregião	Nº de casos	Porcentagem
Metropolitana de Porto Alegre	215	53,09%
Sudeste Rio-grandense	45	11,11%
Noroeste Rio-grandense	42	10,37%
Nordeste Rio-grandense	34	8,40%
Centro Oriental Rio-grandense	31	7,65%
Sudoeste Rio-grandense	24	5,93%
Centro Ocidental Rio-grandense	14	3,46%

A Figura (17) mostra também a localização pontual dos eventos em estudo, porém, evidenciados pelas 35 microrregiões do estado. As microrregiões de Jaguarão e Não-Me-Toque não apresentaram a ocorrência de casos. As microrregiões de Porto Alegre, Pelotas e Litoral Lagunar apresentaram eventos mais concentrados, ou seja, próximos um dos outros. Já as outras microrregiões apresentaram uma diminuição gradativa dos casos e eventos mais dispersos.

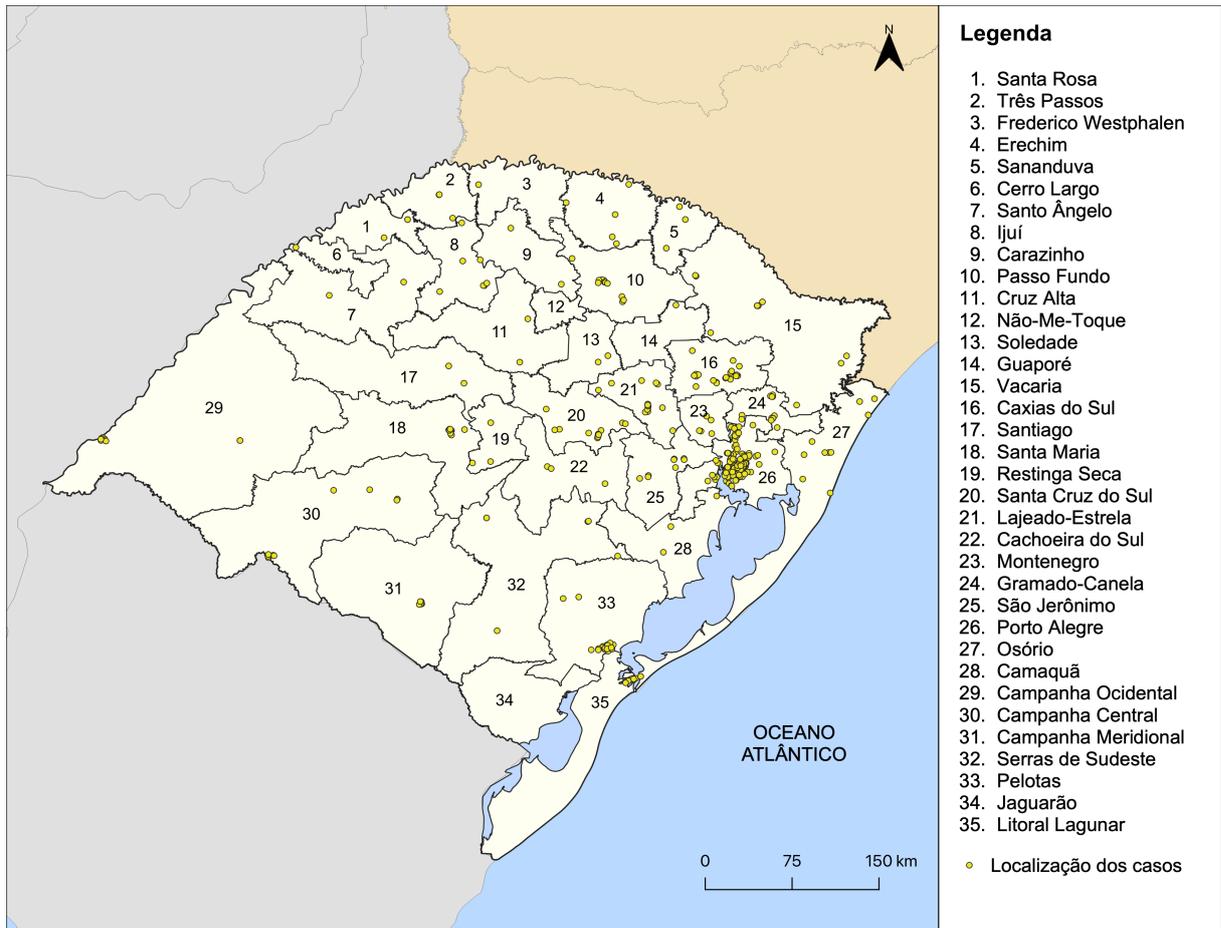


Figura 17 – Mapa de distribuição dos casos das doenças por microrregião.

Na Tabela (2), visualiza-se a distribuição do número de casos por microrregião, também classificados em ordem decrescente. Observando os valores, identificou-se que pelo menos 25 microrregiões possuíam um número de casos abaixo de 10. Obtendo a moda dos valores, verificou-se que 6 microrregiões possuíam o mesmo número de casos, ou seja, 3. Obtendo a mediana, foi visto que o número de 6 casos representou o valor central do conjunto de dados. A microrregião de Porto Alegre apresentou o maior número, 176 casos.

Tabela 2 – Número de casos das doenças por microrregião

Microrregião	Nº de casos	Porcentagem
Porto Alegre	176	43,46%
Pelotas	25	6,17%
Caxias do Sul	23	5,68%
Lajeado-Estrela	16	3,95%
Litoral Lagunar	14	3,46%
Santa Cruz do Sul	12	2,96%
Gramado-Canela	12	2,96%
Passo Fundo	11	2,72%
Osório	11	2,72%
Vacaria	10	2,47%
Santa Maria	9	2,22%
Campanha Central	9	2,22%
São Jerônimo	8	1,98%
Campanha Ocidental	8	1,98%
Ijuí	7	1,73%
Campanha Meridional	7	1,73%
Serras de Sudeste	6	1,48%
Montenegro	5	1,23%
Erechim	5	1,23%
Três Passos	3	0,74%
Santa Rosa	3	0,74%
Sananduva	3	0,74%
Restinga Seca	3	0,74%
Camaquã	3	0,74%
Cachoeira do Sul	3	0,74%
Soledade	2	0,49%
Santo Ângelo	2	0,49%
Santiago	2	0,49%
Cruz Alta	2	0,49%
Carazinho	2	0,49%
Guaporé	1	0,25%
Frederico Westphalen	1	0,25%
Cerro Largo	1	0,25%
Jaguarão	0	0%
Não-Me-Toque	0	0%

3.3 Análises espaciais

3.3.1 Mapeamento coroplético

Nesta seção serão apresentados e descritos os mapas coropléticos de: **a)** número total de casos; e **b)** prevalência – por mesorregião, microrregião e município.

a) Número total de casos

Na Figura (18), visualiza-se o mapa coroplético do número total de casos das três doenças por mesorregião. Foi possível observar uma maior concentração do número de casos na mesorregião Metropolitana, acima de 45 casos.

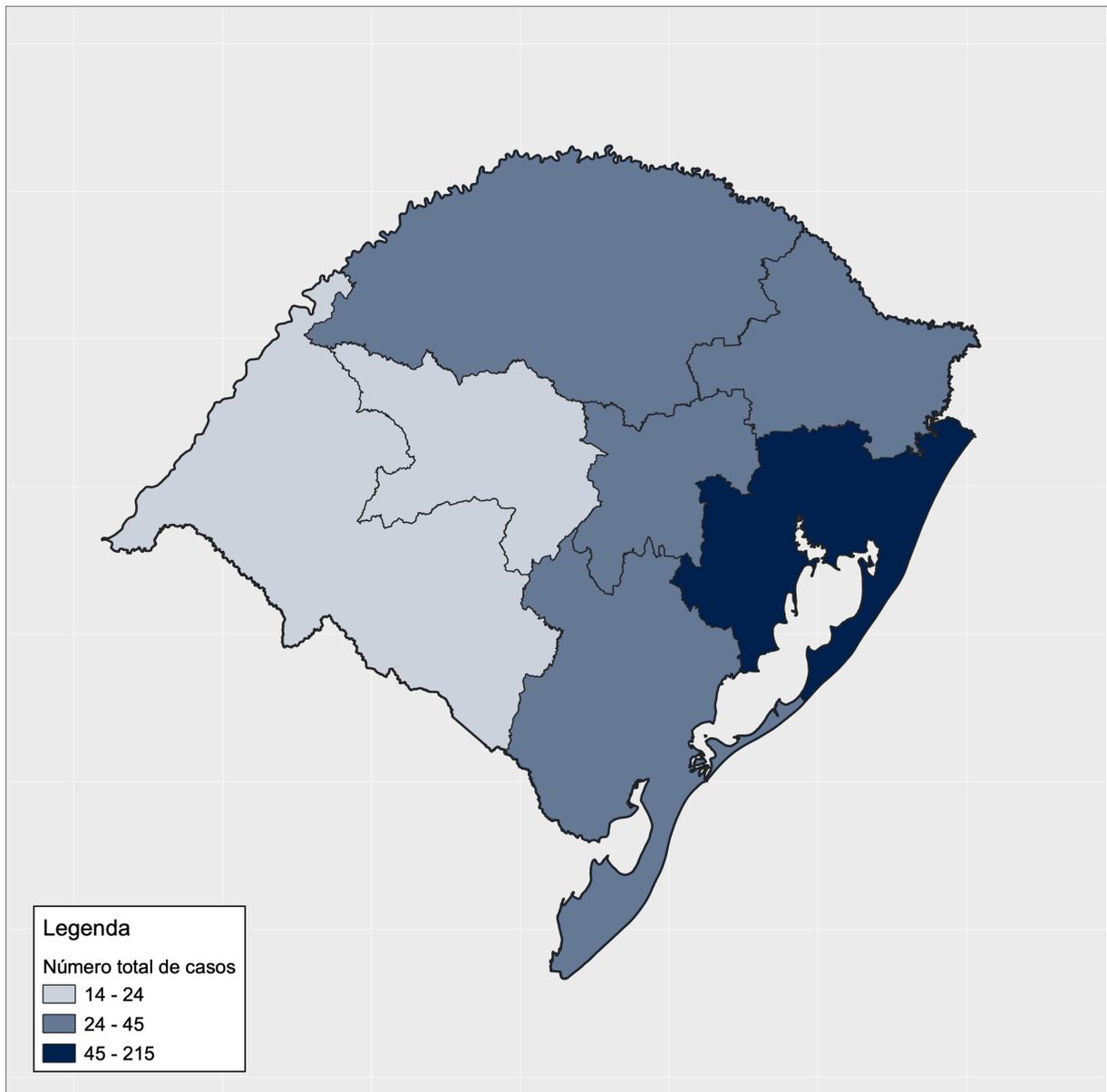


Figura 18 – Mapa coroplético do número total de casos das doenças por mesorregião.

Na Figura (19), foi observado uma diminuição gradativa do número de casos, em nível de microrregião, no Noroeste e Sudeste do estado.

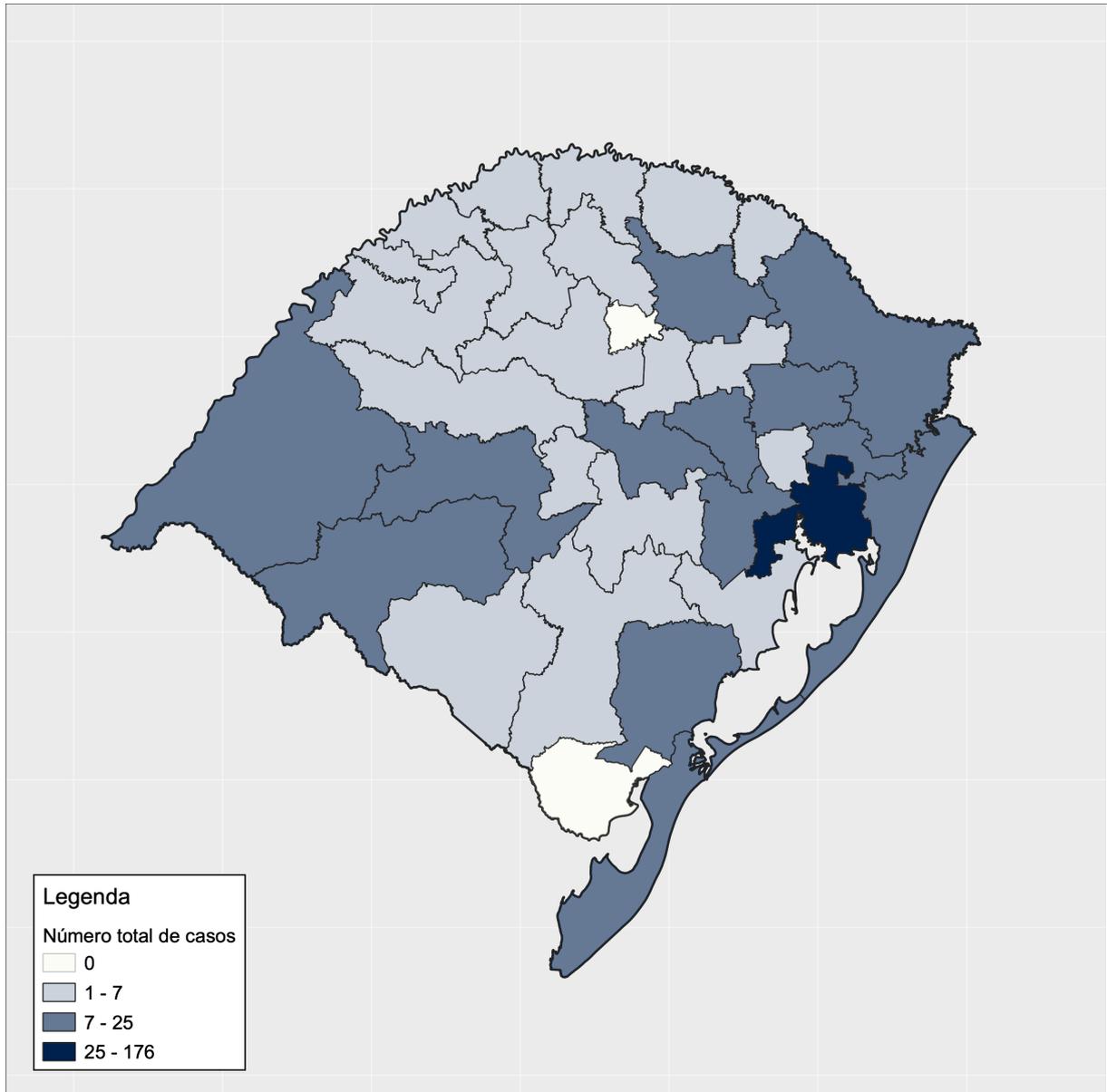


Figura 19 – Mapa coroplético do número total de casos das doenças por microrregião.

Na Figura (20), foi possível visualizar inúmeros municípios sem o registro de número de casos nas mesorregiões Sudeste, Centro Ocidental, Noroeste, e uma forte concentração de casos na capital do estado, Porto Alegre.

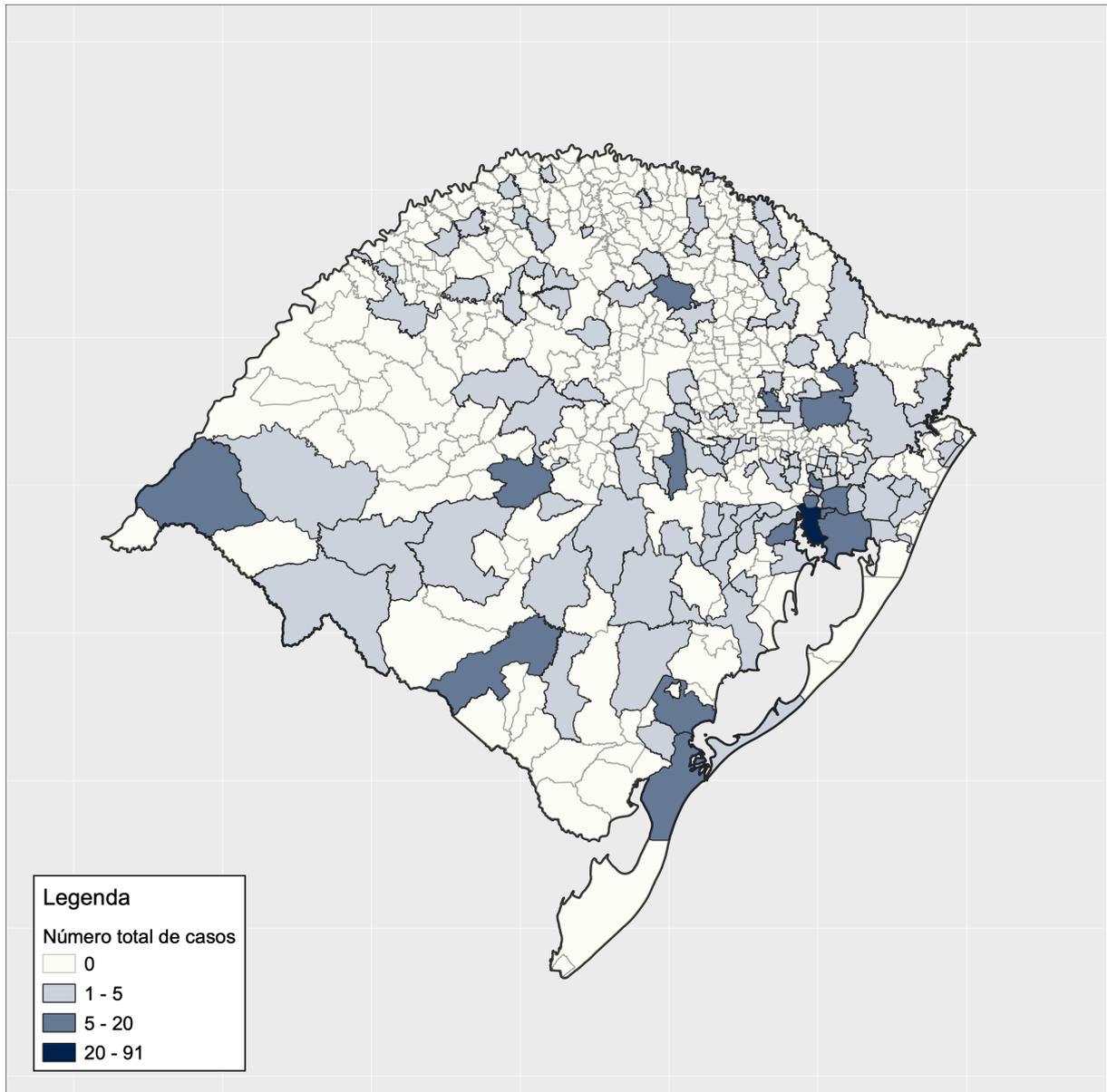


Figura 20 – Mapa coroplético do número total de casos das doenças por município.

b) Prevalência

Na Figura (21), visualiza-se o mapa coroplético da prevalência das três doenças por mesorregião. Foi possível observar uma maior concentração de prevalência, por 200.000 mil habitantes, nas mesorregiões Metropolitana e Sudeste, entre 8 - 10 casos.

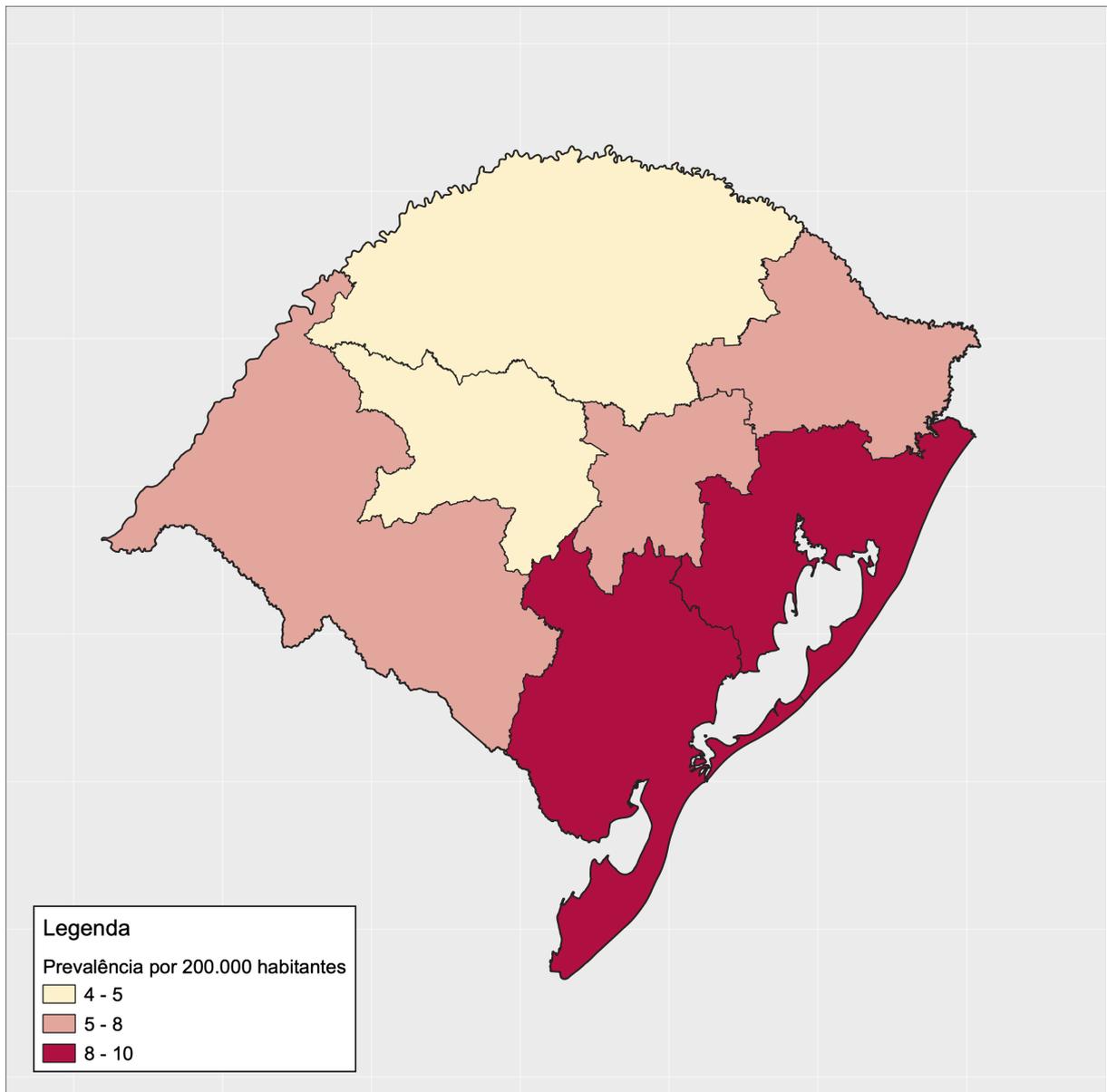


Figura 21 – Mapa coroplético da prevalência das doenças por mesorregião.

Na Figura (22), foi possível visualizar também uma alta concentração, em nível de microrregião, de prevalência por 100.000 mil habitantes, nas regiões Sudoeste, Sudeste, Metropolitana e Nordeste do estado.

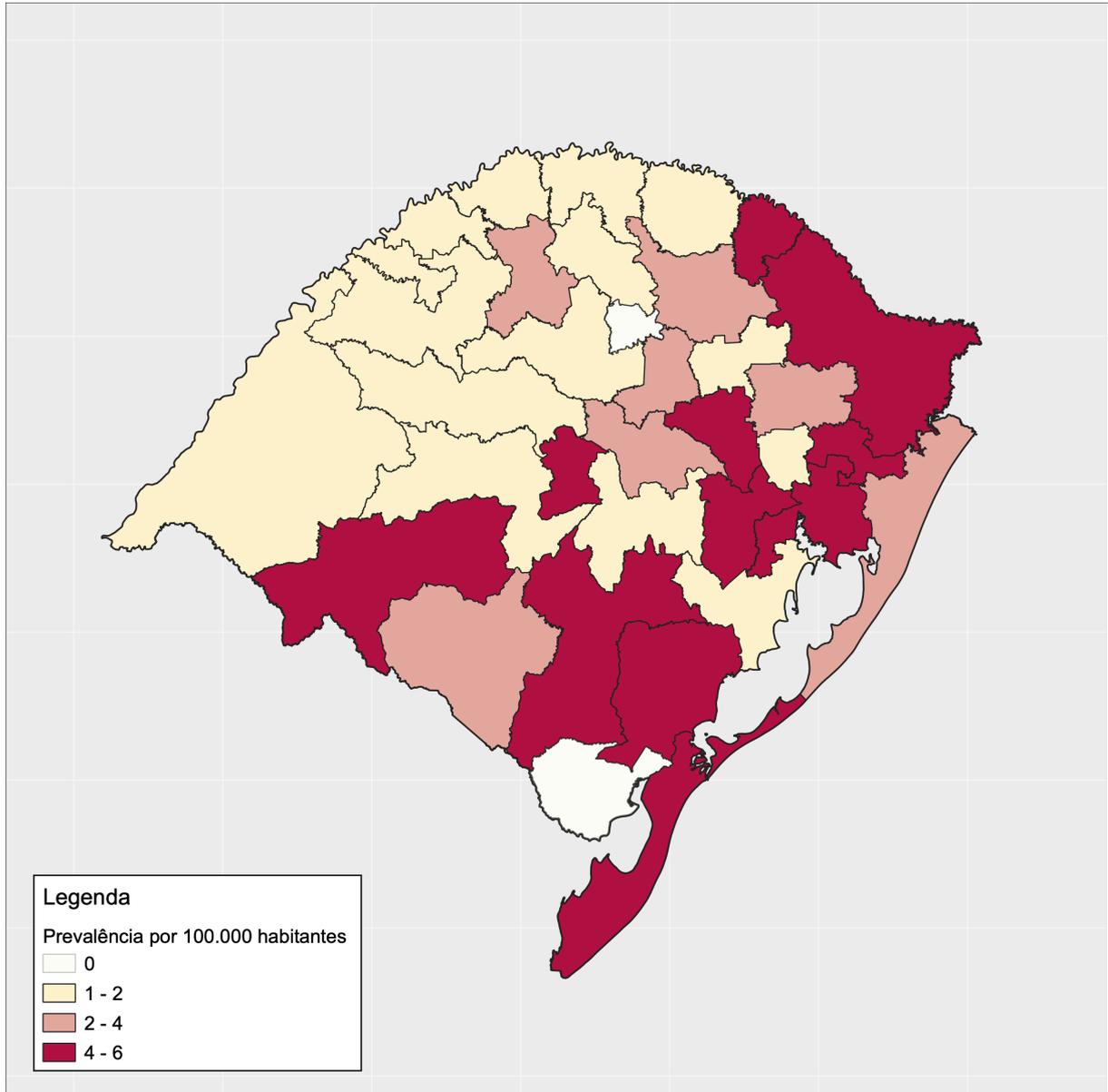


Figura 22 – Mapa coroplético da prevalência das doenças por microrregião.

Na Figura (23), foi visto que a maior parte dos municípios das regiões Sudoeste e Sudeste apresentaram baixo nível de prevalência por 10.000 habitantes, entre 0 - 1 caso, e somente 11 municípios possuíram alta concentração de prevalência, foram eles: Formigueiro, São João do Polêsine, Nova Bréscia, Cambará do Sul, Dom Pedro de Alcântara, São Jorge, Nova Ramada, São Pedro das Missões, Erebangó, Entre Rios do Sul, Mariano Moro.

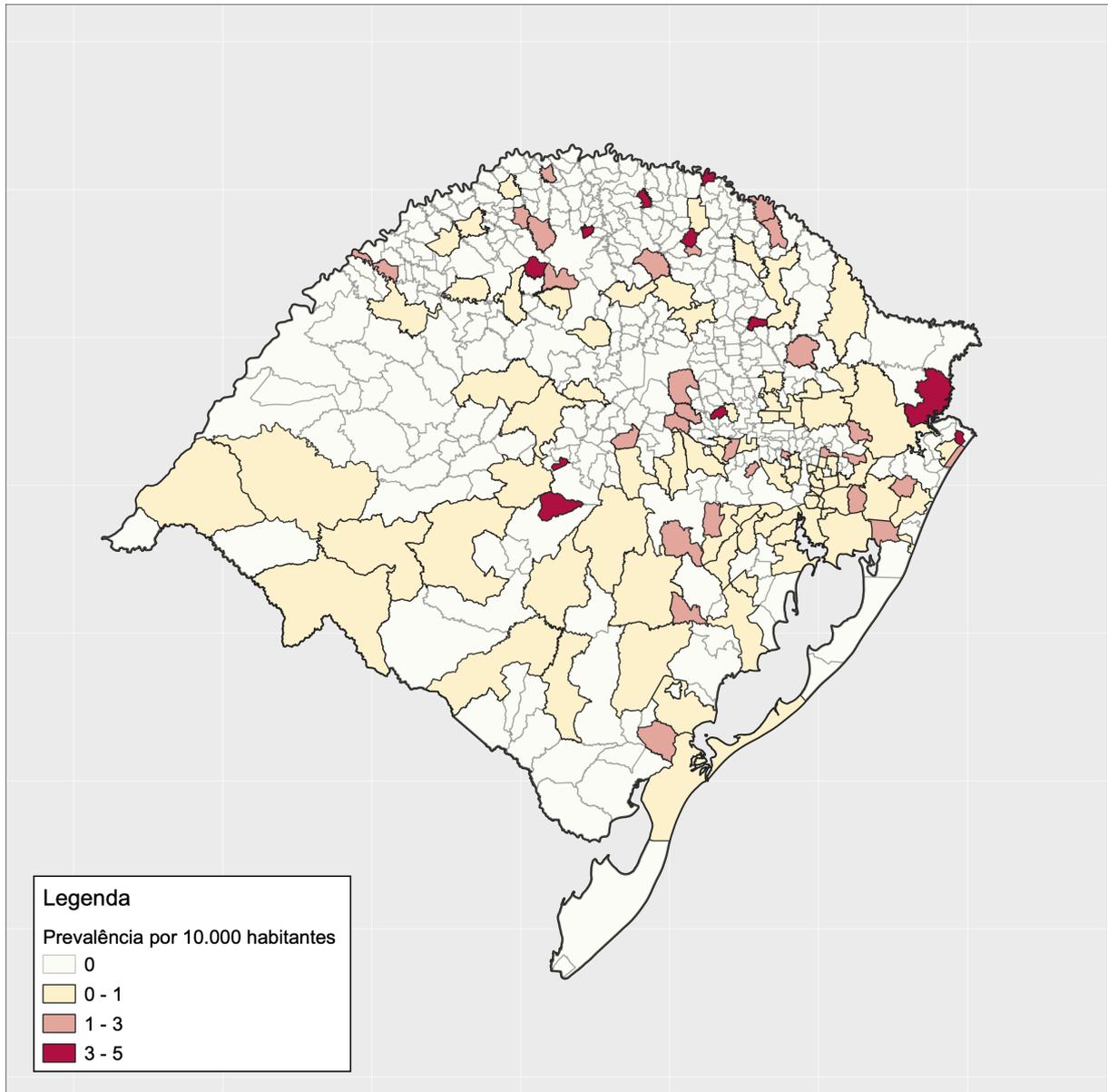


Figura 23 – Mapa coroplético da prevalência das doenças por municípios.

3.3.2 Kernel density estimation

Nesta seção, foi observado como resultado uma visualização ainda mais específica, complementando a visualização da distribuição espacial do número total de casos por região/área, como demonstrado anteriormente nas Figura (18), Figura (19) e Figura (20). Dessa forma, na Figura (24), foi visto com maior precisão a localização dos casos das doenças sobre o número total de habitantes dos municípios do estado, no qual, foram graduados na cor verde.

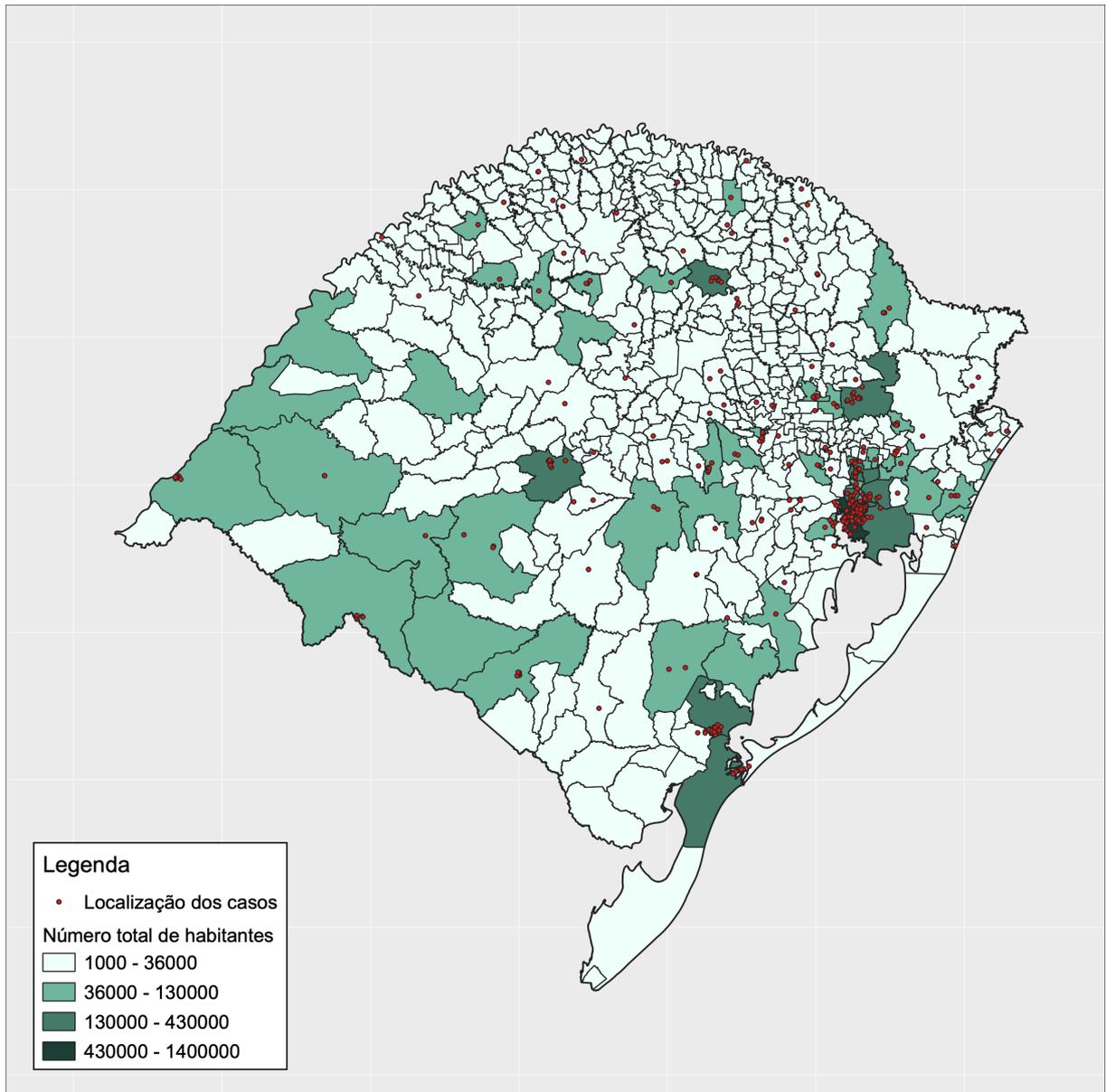


Figura 24 – Mapa pontual dos casos das doenças sobre o número total de habitantes.

Na Figura (25), utilizando o método KDE, em um raio de 10 km, foi possível visualizar as áreas quentes da ocorrência dos casos dentro da extensão territorial de cada município do estado. O mapa mostrou uma elevada concentração de casos, ou seja, acima de 5, em pelo menos 14 municípios dos 497 municípios do estado, foram eles: Pelotas, Rio Grande, São Leopoldo, Canoas, Porto Alegre, Alvorada, Viamão, Caxias do Sul, Bento Gonçalves, Lajeado, Santa Cruz do Sul, Passo Fundo, Santa Maria e Uruguaiana.

Ainda, foi possível observar na Figura (25) que pelo menos 4 municípios das mesorregiões Sudoeste e Centro Oriental do estado apresentaram uma grande concentração de casos sob uma população relativamente pequena, ou seja, em torno de 80 mil habitantes. Os municípios foram: Bagé, Uruguaiana, Santa Cruz do Sul e Lajeado. Também, foi visto que pelo menos 8 municípios possuíram uma grande concentração de casos sob uma população relativamente grande, ou seja, aproximadamente 750 mil habitantes. Os municípios foram: Rio Grande, Pelotas, Santa Maria, Passo Fundo, Caxias do Sul, Porto Alegre, Canoas e São Leopoldo.

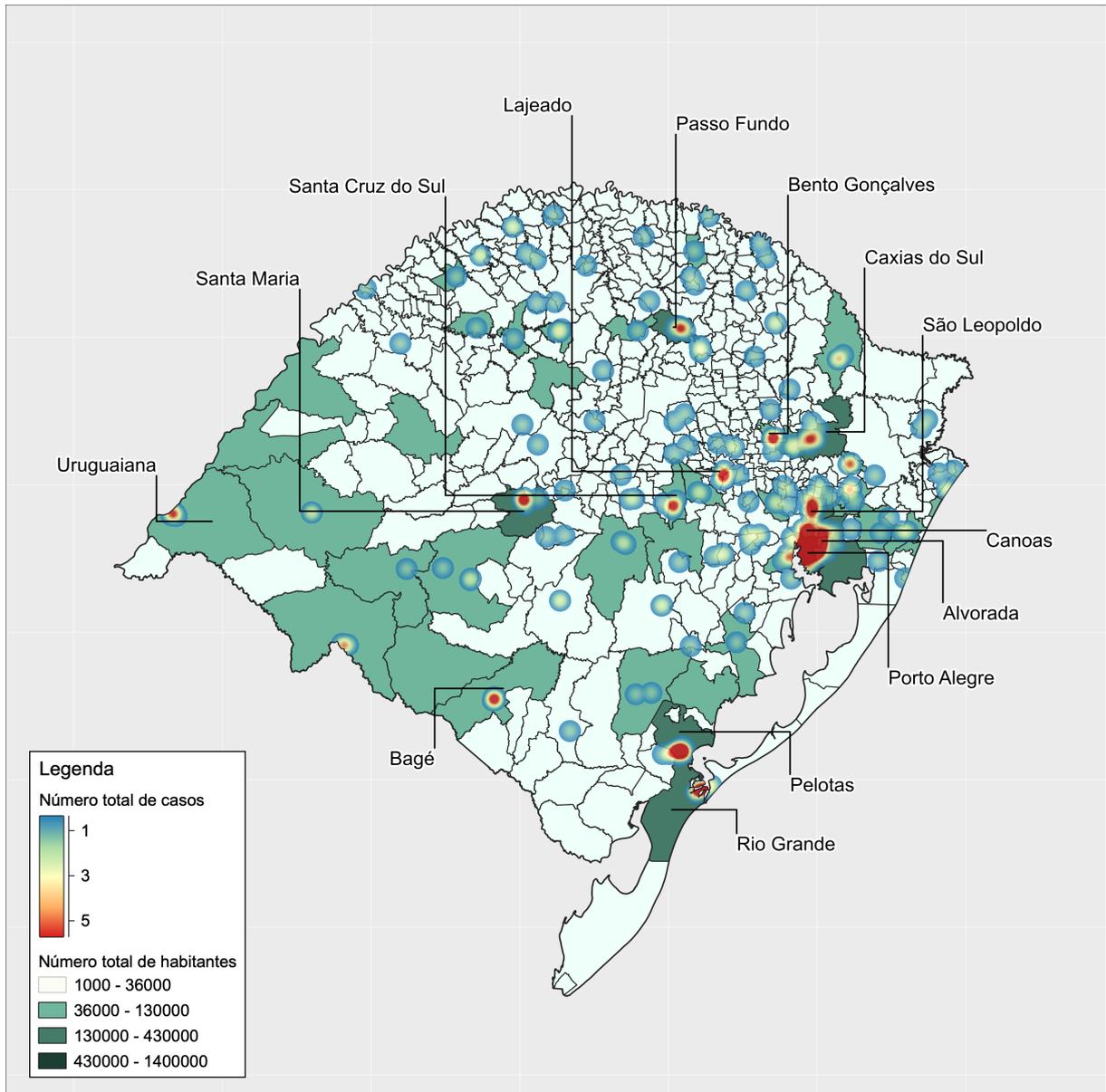


Figura 25 – Mapa de KDE dos casos das doenças sobre o número total de habitantes.

3.3.3 Dimensão de correlação

Nesta seção serão apresentados e descritos os resultados obtidos a partir da aplicação do método de dimensão de correlação por: **a)** mesorregiões; e **b)** microrregiões. Os resultados estão apresentados em forma de gráficos, mapas e tabelas.

a) Mesorregiões

No gráfico exposto na Figura (26), visualiza-se os valores do grau de agrupamento $C(r)$ dos eventos na mesorregião Centro Oriental. Os valores de $C(r)$ estiveram entre 0.283 e 1.823, para cada ponto de distância em quilômetros (km) do eixo x. Entre as distâncias de 10 - 40 km, os valores de $C(r)$ estiveram abaixo de 0.713, indicando a existência de eventos agrupados nesta faixa. A partir de uma distância de 50 - 100 km, houve um crescimento nos valores do grau de agrupamento, chegando até 1.823, indicando que com aumento da área, a distribuição dos pontos tende a ser mais heterogênea, ou seja, sem padrões ou agrupamentos. O valor da dimensão de correlação D_c para todos os pontos de distância (10 - 100 km) foi de ≈ 0.9 , indicando que os eventos estiveram oscilando entre agrupados e dispersos. A Figura (27) mostra a distribuição dos eventos com a mesorregião Centro Ocidental em evidência.

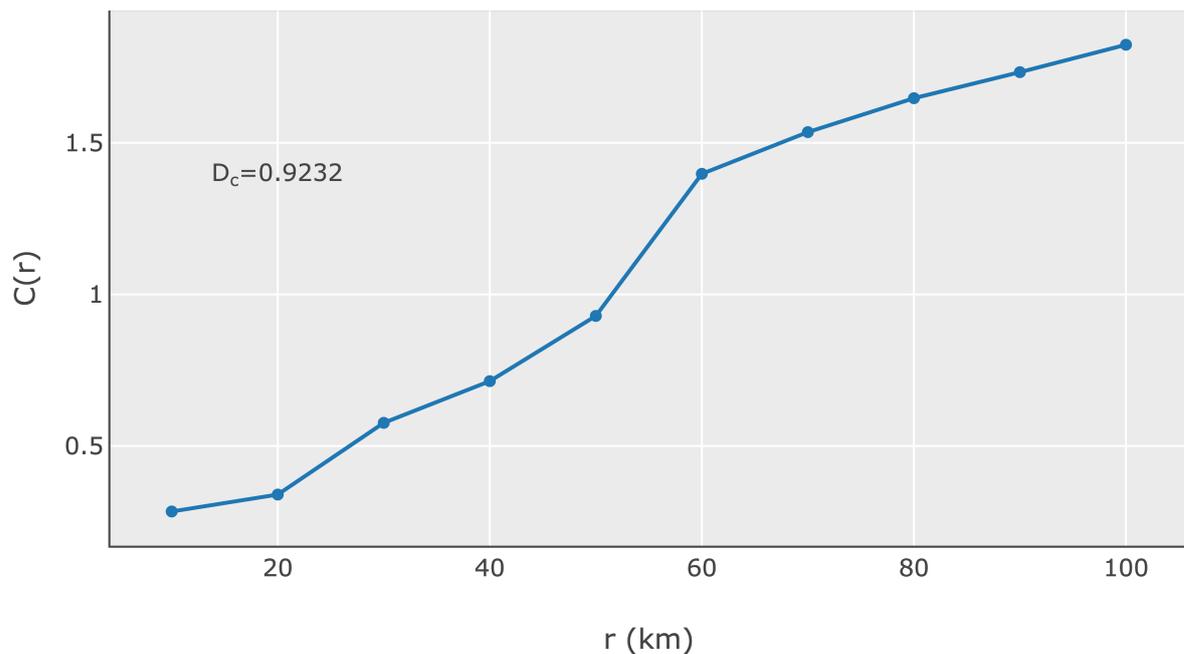


Figura 26 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Centro Oriental Rio-grandense.



Figura 27 – Mapa de localização dos casos das doenças sobre a mesorregião Centro Oriental Rio-grandense.

Na Figura (28), observa-se o gráfico de dimensão de correlação dos eventos na mesorregião Nordeste Rio-grandense. Os valores do grau de agrupamento de $C(r)$ variaram entre 0.245 e 1.632, apontando uma tendência crescente. No entanto, o valor da dimensão de correlação D_c foi de ≈ 0.8 , apontando agrupamento dos eventos entre a área de 10 - 100 km. A Figura (29) mostra a distribuição dos eventos com a mesorregião Nordeste em evidência.

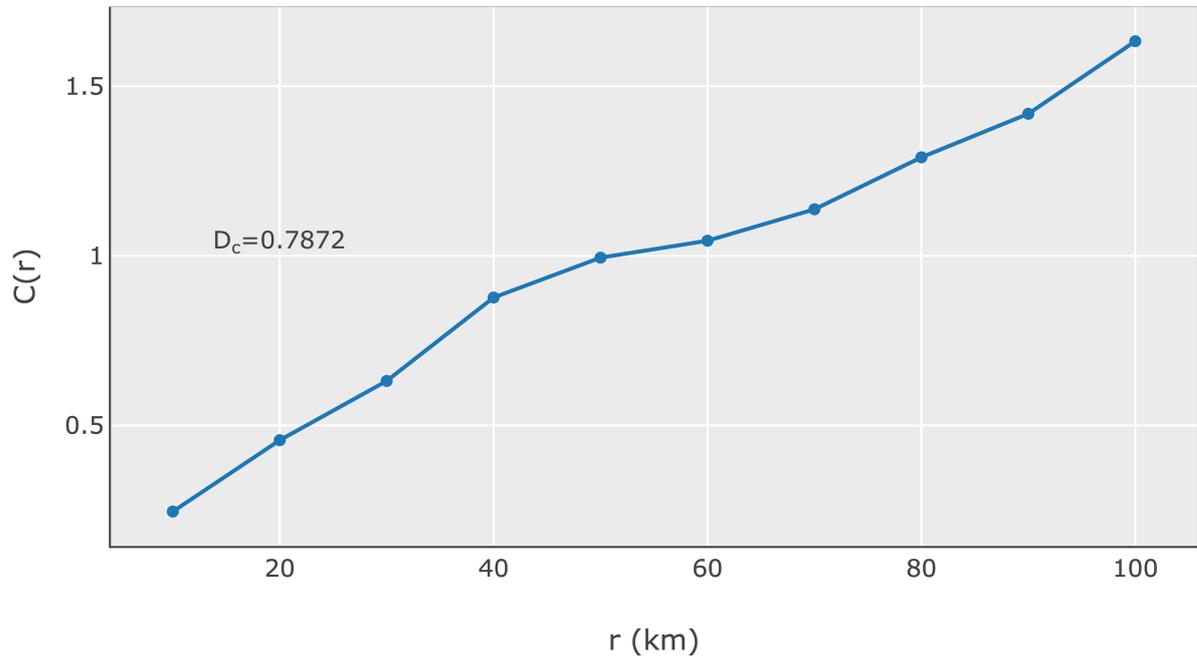


Figura 28 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Nordeste Rio-grandense.



Figura 29 – Mapa de localização dos casos das doenças sobre a mesorregião Nordeste Rio-grandense.

Para a mesorregião Centro Ocidental, os valores do grau de agrupamento $C(r)$ variaram entre 0.615 e 1.956, conforme o gráfico apresentado na Figura (30). Houve uma tendência crescente nos valores. Sendo que, a partir de 70 km de distância entre os eventos, os valores de $C(r)$ foram altos, acima de 1.8, indicando que há eventos heterogeneamente distribuídos nesta área. A Figura (31) apresenta a distribuição dos eventos com a mesorregião Centro Ocidental em destaque.

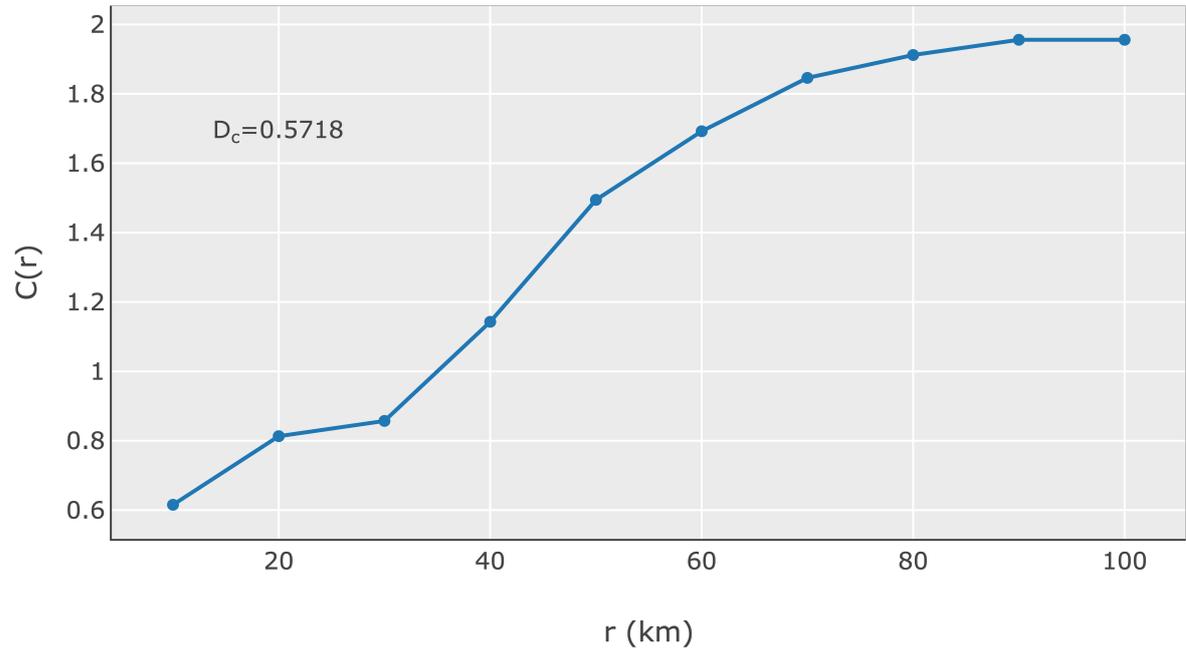


Figura 30 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Centro Ocidental Rio-grandense.



Figura 31 – Mapa de localização dos casos das doenças sobre a mesorregião Centro Ocidental Rio-grandense.

Na Figura (32), gráfico da dimensão de correlação da mesorregião Sudeste, observou-se que entre as distâncias de 50 - 100 km não houve uma alteração significativa nos valores de $C(r)$, sugerindo que todos os eventos estiveram heterogeneamente distribuídos abaixo deste limite. Para toda a faixa de distância (10 - 100 km), o valor da dimensão de correlação D_c foi de ≈ 0.5 , indicando um agrupamento forte de muitos casos. A Figura (33) mostra a distribuição dos eventos sobre a mesorregião Sudeste em evidência.

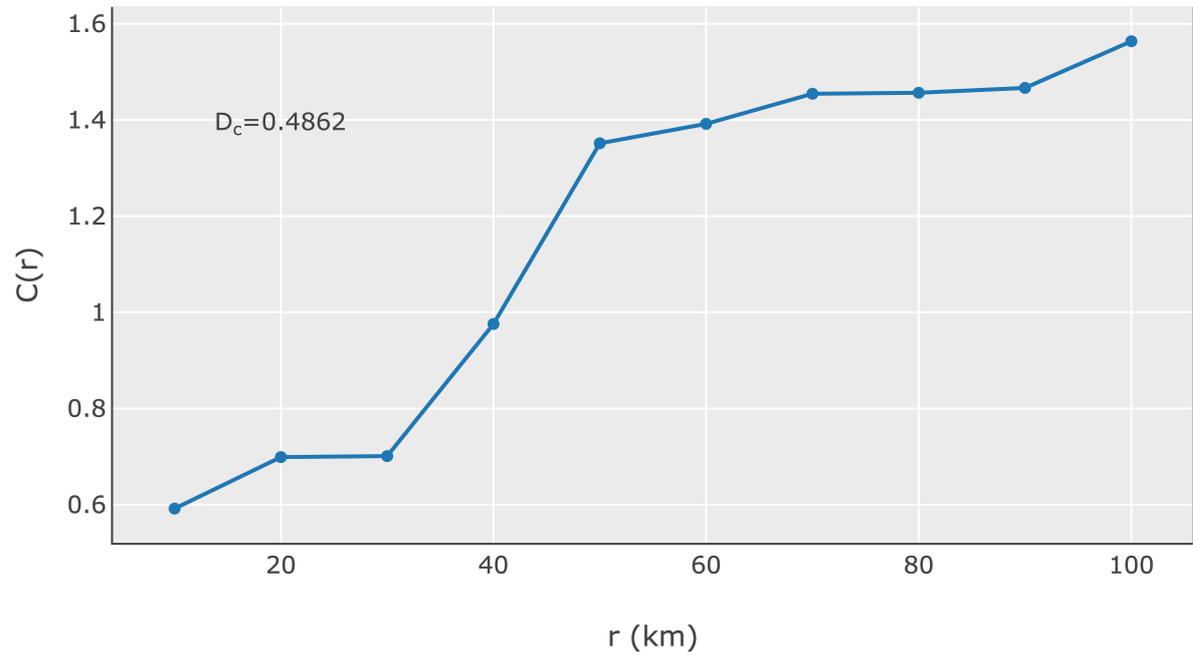


Figura 32 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Sudeste Rio-grandense.



Figura 33 – Mapa de localização dos casos das doenças sobre a mesorregião Sudeste Rio-grandense.

Na mesorregião Metropolitana de Porto Alegre, os valores do grau de correlação entre os eventos variaram entre 0.418 e 1.871, conforme apresentado no gráfico da Figura (34). O valor de $C(r) = 1.871$ no ponto de 100 km, indicou que quase todos os eventos se encontraram uniformemente distribuídos abaixo desta área. Esta mesorregião apresentou um valor de dimensão de correlação $D_c \approx 0.6$, sugerindo um agrupamento forte de muitos casos entre 10 - 100 km. A Figura (35) mostra a distribuição dos eventos sobre a mesorregião Metropolitana em destaque.

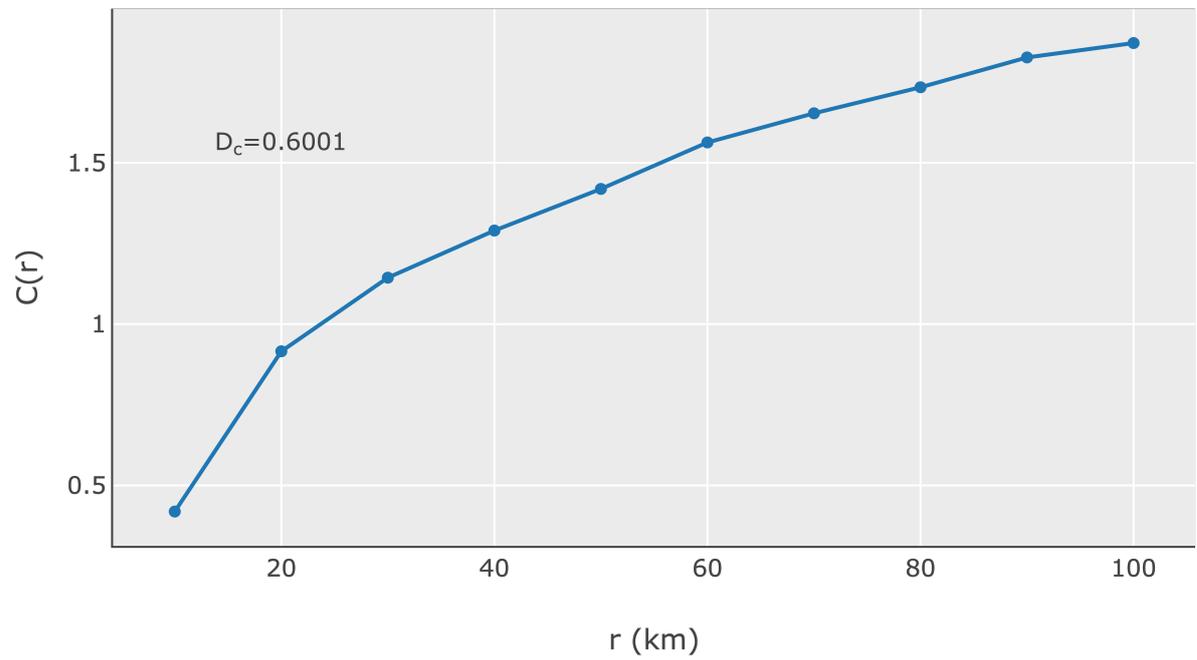


Figura 34 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Metropolitana de Porto Alegre.

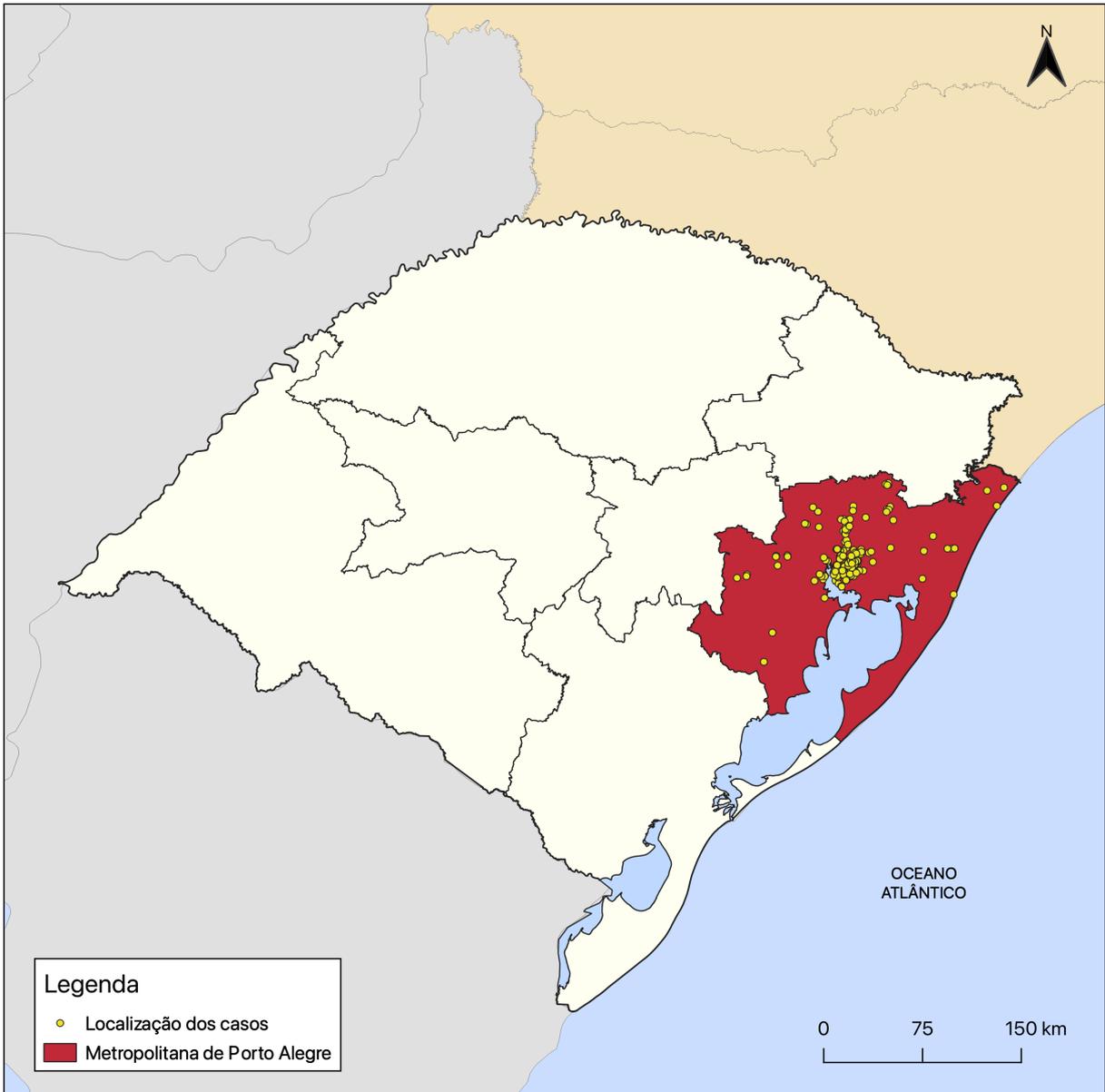


Figura 35 – Mapa de localização dos casos das doenças sobre a mesorregião Metropolitana de Porto Alegre.

No gráfico do grau de correlação da mesorregião Noroeste – Figura (36), visualizou-se um crescimento nos valores de $C(r)$, indicando que a distribuição dos eventos tende a ser dispersa. O valor da dimensão de correlação D_c de ≈ 1.14 , para toda a faixa de distância (10 - 100 km), sugere que os eventos estiveram heterogeneamente distribuídos por toda a região, apresentando poucos agrupamentos e mais dispersão dos casos. A Figura (37) mostra a distribuição dos casos sobre a mesorregião Noroeste em evidência.

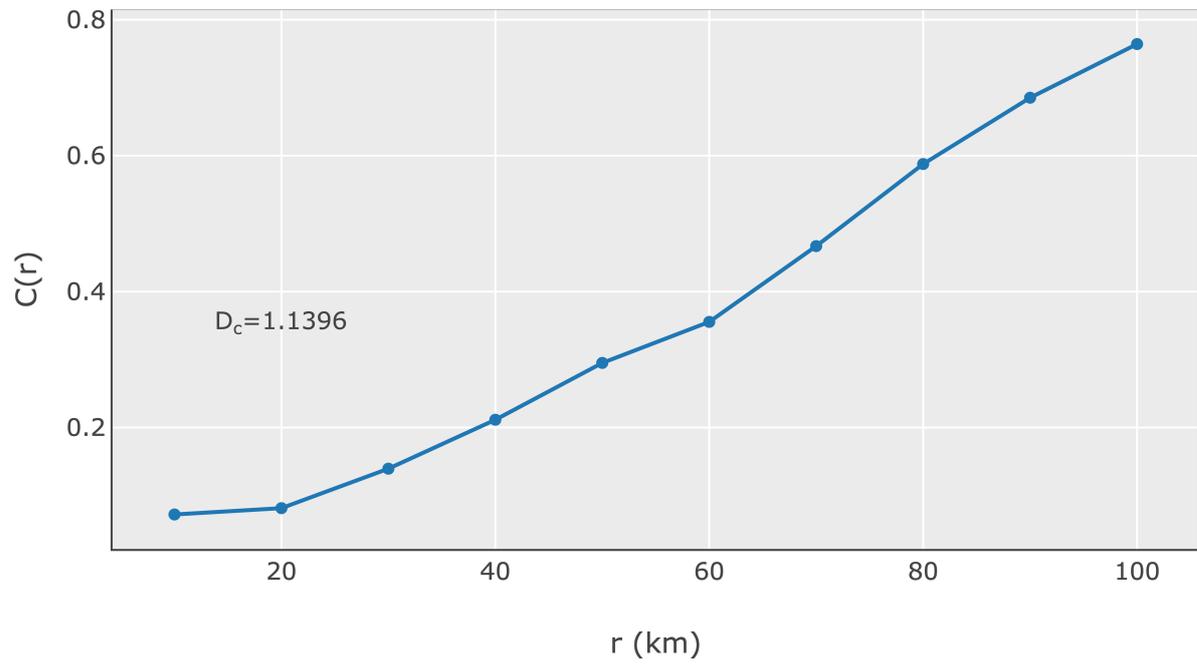


Figura 36 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Noroeste Rio-grandense.



Figura 37 – Mapa de localização dos casos das doenças sobre a mesorregião Noroeste Rio-grandense.

Por fim, observando o gráfico da mesorregião Sudoeste - Figura (38), verificou-se que os valores do grau de agrupamento $C(r)$ apresentaram uma pequena variação, entre 0.347 e 0.434. O valor da dimensão de correlação D_c foi de ≈ 0.08 , um valor muito baixo, indicando que os eventos estiveram extremamente agrupados dentro da faixa de 10 - 100 km. A Figura (39) mostra a distribuição dos eventos sobre a mesorregião Sudoeste em destaque.

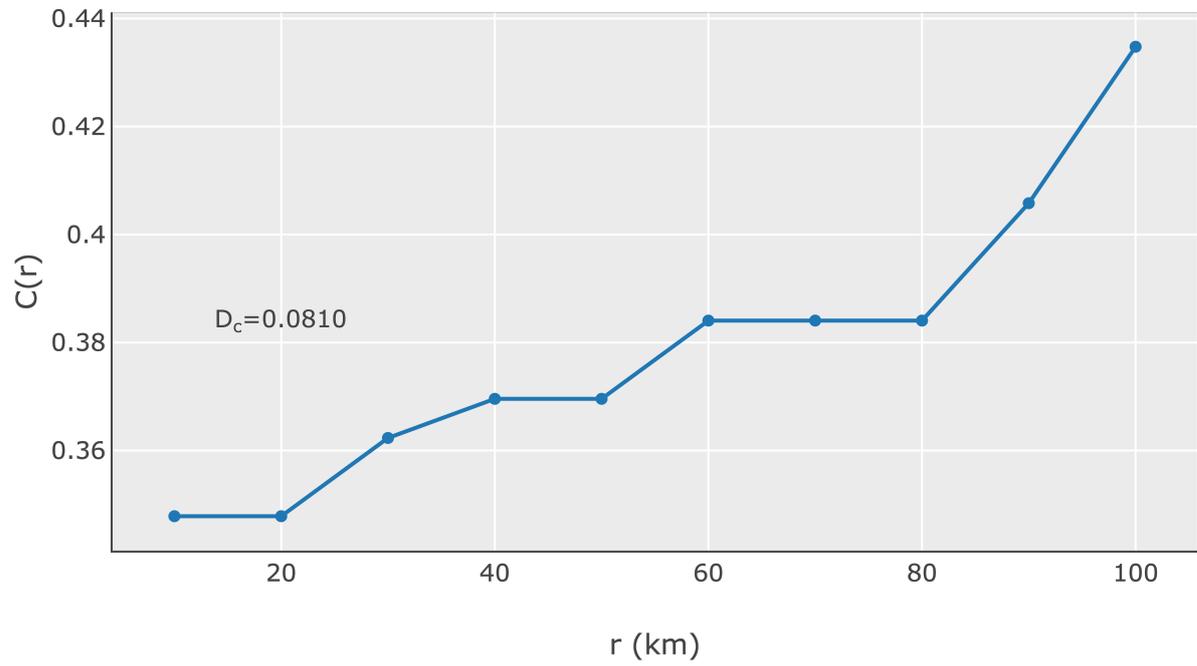


Figura 38 – Gráfico log – log de $C(r)$ por r (km) da mesorregião Sudoeste Rio-grandense.



Figura 39 – Mapa de localização dos casos das doenças sobre a mesorregião Sudoeste Rio-grandense.

O gráfico da Figura (40) sintetiza o grau de correlação para todas as mesorregiões do estado. Verificou-se que cinco mesorregiões possuíam uma tendência crescente nos valores de $C(r)$, convergindo para uma média de ≈ 1.8 , sugerindo que quase todos os pares de eventos estiveram heterogeneamente distribuídos numa distância de até 100 km. As mesorregiões foram: Centro Ocidental, Metropolitana, Centro Oriental, Nordeste e Sudeste. A mesorregião Noroeste também apresentou uma tendência crescente nos seus valores de $C(r)$. Porém, com uma inclinação menor, fixando o seu valor em 0.764 na distância

de 100 km, sugerindo que a maior parte dos eventos nesta área esteve heterogeneamente distribuída numa distância maior que 100 km. Já a mesorregião Sudoeste apresentou uma leve constância nos seus valores de $C(r)$, refletindo um forte agrupamento dos dados já numa distância abaixo de 30 km.

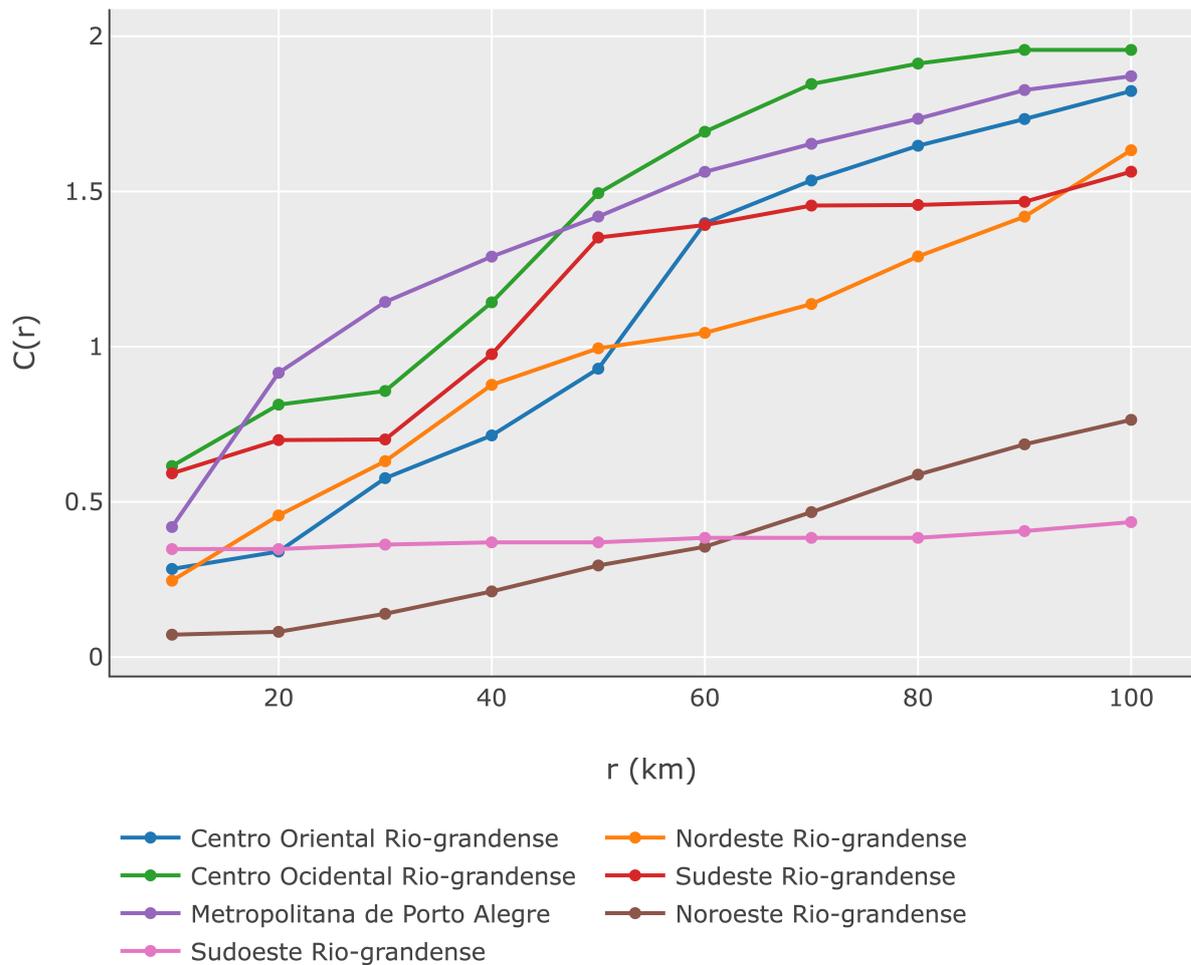


Figura 40 – Gráfico log – log de $C(r)$ por r (km) para todas as mesorregiões.

Na Figura (41), visualiza-se o gráfico da variação dos valores da dimensão de correlação D_c para todas as mesorregiões. Observou-se que a mesorregião Noroeste apresentou o maior valor de dimensão de correlação $D_c \approx 1.15$ em relação às outras, indicando heterogeneidade, maior dispersão, entre seus pares de eventos. Já a mesorregião Sudoeste apresentou o menor valor da dimensão de correlação $D_c \approx 0.08$, tornando-se a mesorregião com o maior índice de agrupamento entre os pares de eventos. A Figura (42) realça as duas mesorregiões em questão e os seus eventos.

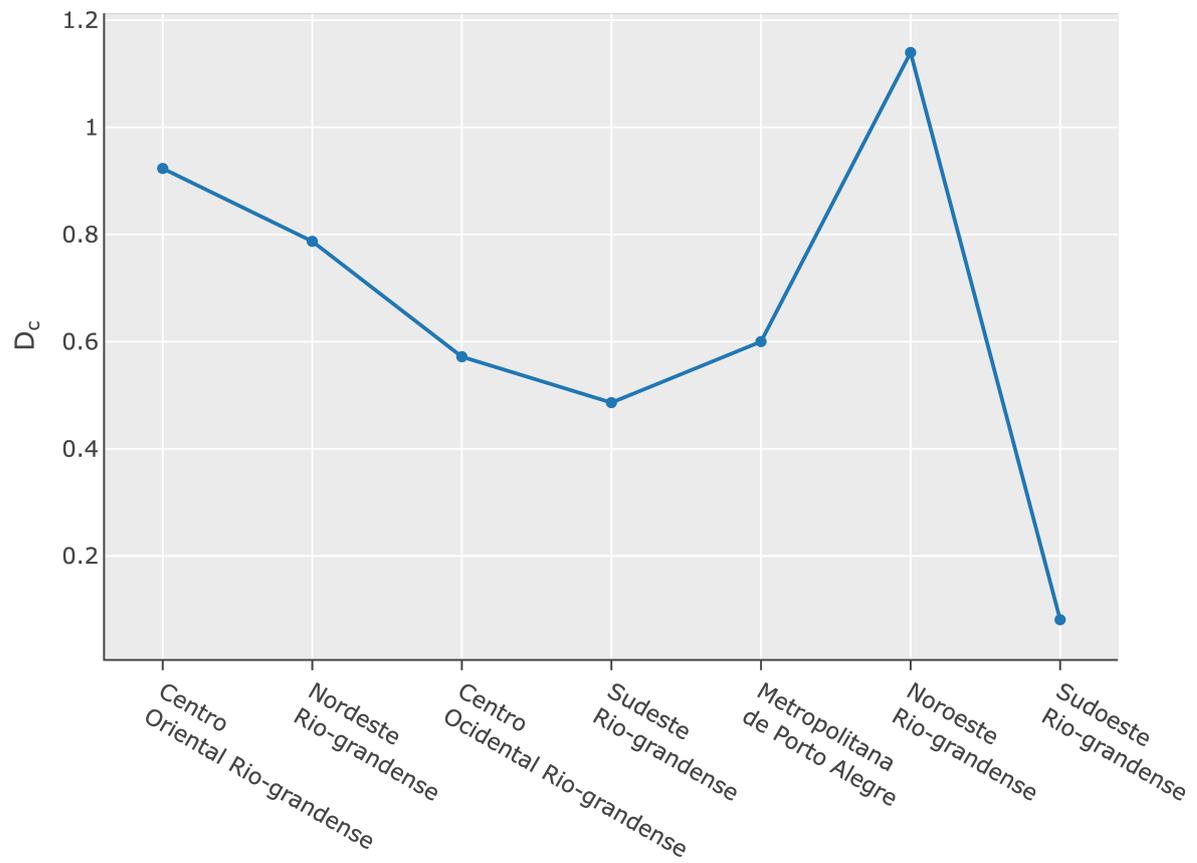


Figura 41 – Gráfico da variação dos valores da dimensão de correlação D_c para todas as mesorregiões.



Figura 42 – Mapa das mesorregiões com o maior e o menor índice de dimensão de correlação.

A Tabela (3) apresenta os valores da dimensão de correlação D_c para todas as mesorregiões, classificadas em ordem crescente, ou seja, das mesorregiões que apresentaram maior índice de agrupamento para as que apresentaram maior dispersão entre seus eventos.

Tabela 3 – Dimensão de correlação por mesorregião

Mesorregião	D_c
Sudoeste Rio-grandense	0.0810
Sudeste Rio-grandense	0.4862
Centro Ocidental Rio-grandense	0.5718
Metropolitana de Porto Alegre	0.6001
Nordeste Rio-grandense	0.7872
Centro Oriental Rio-grandense	0.9232
Noroeste Rio-grandense	1.1396

a) Microrregiões

No que se segue, na Tabela (4), visualiza-se os valores da dimensão de correlação D_c para todas as microrregiões. Os valores foram classificados em ordem crescente. Pelo menos 10 microrregiões apresentaram um valor baixo de dimensão de correlação D_c , abaixo de 0.5, indicando fortes agrupamentos, foram elas: Campanha Meridional; Campanha Ocidental; Litoral Lagunar; Santa Maria; Pelotas; Campanha Central; Passo Fundo; Santa Cruz do Sul; Porto Alegre; e São Jerônimo. A Figura (43) realça as microrregiões citadas e a localização dos seus eventos.

Ainda, analisando a Tabela (4), observou-se também que 13 microrregiões apresentaram valores NaN (*Not a Number*). Isto se deve ao fato de que algumas microrregiões não apresentaram a ocorrência de casos ou, então, apresentaram um pequeno número de eventos, abaixo de três, impossibilitando o cálculo da dimensão fractal. As 13 microrregiões representavam aproximadamente 37% de todas as microrregiões do estado, foram elas: Restinga Seca; Sananduva; Santo Ângelo; Camaquã; Cerro Largo; Cruz Alta; Carazinho; Guaporé; Frederico Westphalen; Santiago; Soledade; Jaguarão; e Não-Me-Toque.

Tabela 4 – Dimensão de correlação por microrregião

Microrregião	D_c
Campanha Meridional	0.0000
Campanha Ocidental	0.0000
Litoral Lagunar	0.0563
Santa Maria	0.0785
Pelotas	0.1420
Campanha Central	0.2677
Passo Fundo	0.4148
Santa Cruz do Sul	0.4354
Porto Alegre	0.4404
São Jerônimo	0.4610
Montenegro	0.5027
Serras de Sudeste	0.5240
Três Passos	0.5288
Santa Rosa	0.5288
Lajeado-Estrela	0.5659
Caxias do Sul	0.6072
Gramado-Canela	0.6247
Cachoeira do Sul	0.6281
Vacaria	0.6395
Ijuí	0.9759
Osório	1.0758
Erechim	1.2340
Restinga Seca	NaN
Sananduva	NaN
Santo Ângelo	NaN
Camaquã	NaN
Cerro Largo	NaN
Cruz Alta	NaN
Carazinho	NaN
Guaporé	NaN
Frederico Westphalen	NaN
Santiago	NaN
Soledade	NaN
Jaguarão	NaN
Não-Me-Toque	NaN

* NaN (*Not a Number*), não é um número.



Figura 43 – Mapa das microrregiões com os menores índices de dimensão de correlação.

Também identificou-se que, entre as 35 microrregiões, três delas apresentaram os maiores valores da dimensão de correlação D_c , com uma média de ≈ 1.1 , sugerindo poucas aglomerações e mais aleatoriedade na distribuição dos seus eventos, foram as microrregiões: Ijuí; Osório; e Erechim. A Figura (44) realça estas microrregiões e os seus eventos.

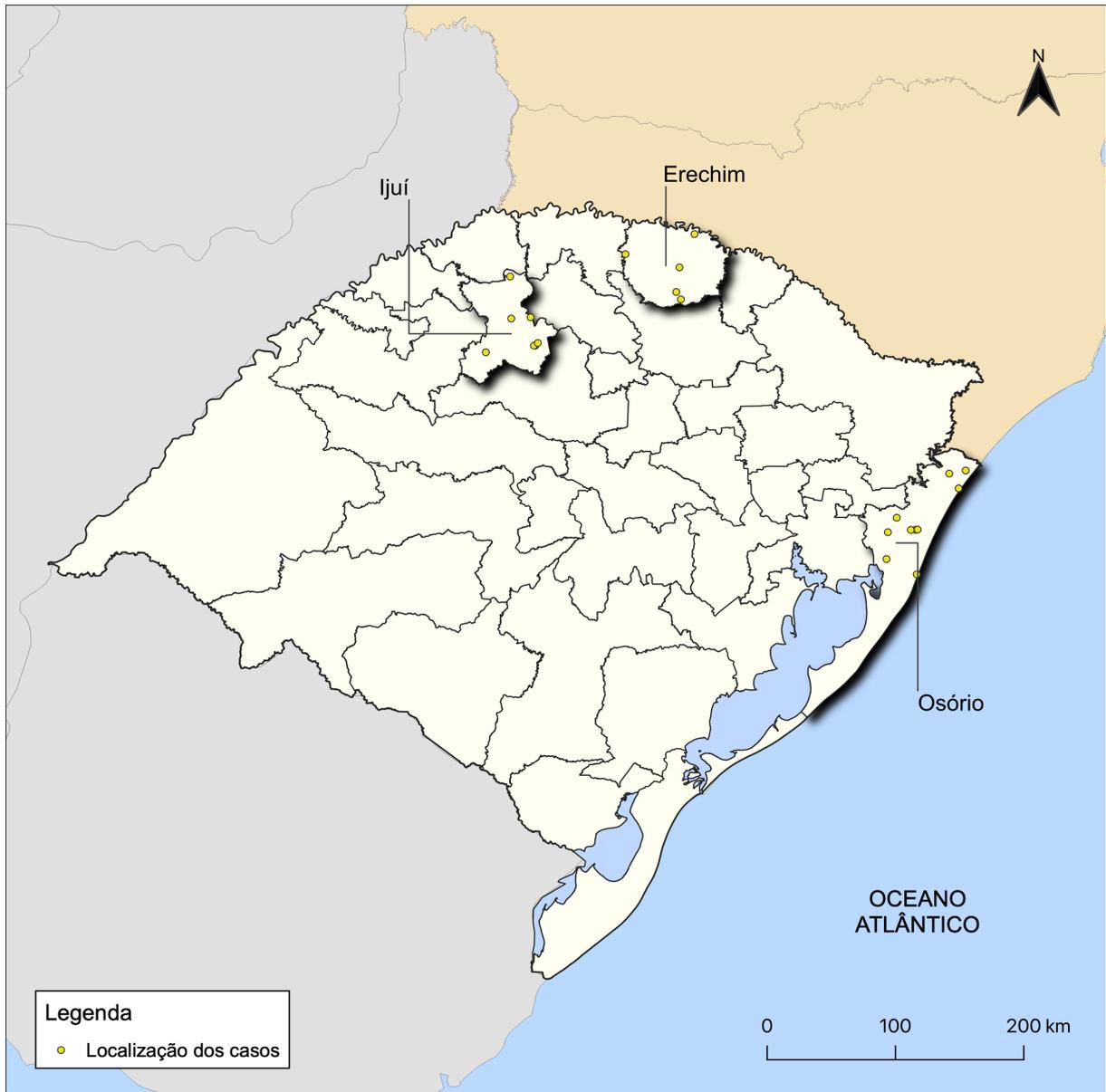


Figura 44 – Mapa das microrregiões com os maiores índices de dimensão de correlação.

4 Discussão

A discussão deste estudo está organizada em três subseções: a Seção (4.1) e a Seção (4.2), discutem as vantagens e limitações dos métodos de georreferenciamento e associação de região. Já a Seção (4.3), discute as principais intuições obtidas a partir da aplicação dos métodos de: mapeamento coroplético; KDE; e dimensão de correlação.

4.1 Georreferenciamento

A construção do método e algoritmo ASSYRIA para georreferenciamento, organização e processamento de dados de endereço, possibilitou manipular os diferentes retornos da API de *geocoding* do *Google Maps*, como: rua, cidade e CEP. Em outros serviços de georreferenciamento, como ArcGIS e QGIS, não foi possível manipular dinamicamente diferentes retornos de dados, sendo utilizado apenas as informações de coordenadas geográficas (latitude; longitude).

Além disso, a construção do algoritmo numa linguagem de programação e tecnologia atual – JavaScript; React (OVERFLOW, 2021), possibilitou a sua disponibilização de forma interativa e gratuita na *web* (<<https://spatialworkspace.ddns.net/geocoding>>). A sua disponibilização permitiu acelerar o processo de organização dos dados; compatibilização com GIS; construção de mapas; planejamento estratégico; e tomada de decisão.

Silveira, Oliveira e Junger (2017), em seu estudo, também utilizaram a API de *geocoding* do *Google Maps* para realizar o georreferenciamento automático de dados de óbitos no Rio de Janeiro, Brasil. No entanto, foi utilizado o *software* livre de computação estatística R (R, 2022) para a construção do algoritmo de georreferenciamento de endereços. Apesar do esforço, algoritmos desenvolvidos em R não são acessíveis na *web*, dificultando, assim, o acesso e utilização por outros profissionais ou pesquisadores. Diferente do trabalho de Silveira, Oliveira e Junger (2017), neste estudo, foi construído um algoritmo baseado na *web* para georreferenciamento em massa de endereços a partir de dados contidos em planilhas eletrônica (.xlsx). Sendo assim, outros pesquisadores e profissionais, sem o conhecimento prévio de programação, poderão utilizar esta abordagem para realizar o processamento e validação de dados de geográficos. Além disso, outros trabalhos científicos poderão utilizar este método para realizar análises comparativas de sucesso e falha em processos de georreferenciamento automático de registros de endereço.

Maas et al. (2019), descreveram os padrões espaciais de encaminhamentos para o programa de saúde mental no oeste de Sydney, Austrália. As taxas de encaminhamentos foram analisadas por meio de índices de autocorrelação espacial e regressão espacial nos softwares ArcGIS e GeoDa. Os resultados e as técnicas utilizadas foram úteis para o monitoramento da desigualdade de atendimento e no planejamento da política de saúde na cidade. Ainda que o estudo tenha utilizado diferentes técnicas estatísticas espaciais, a observação de dados geográficos pontuais não foi contemplada. Observar também a distribuição pontual dos eventos no estudo, poderia complementar as análises, apontando a construção de possíveis pólos de atendimento ou corroborar com os já existentes. Sendo assim, trabalhos semelhantes ao de Maas et al. (2019), que abordaram seus fenômenos com apenas análises de área, podem também utilizar o processo automático de georreferenciamento descrito neste trabalho para trazer mais evidências e discussão aos seus estudos.

Durante o estudo, foi observado que, para registros de endereço completo (rua, número, bairro, cidade, estado, CEP) o georreferenciamento se mostrou preciso, ou seja, correspondente com o endereço físico no mapa global. Já para registros de endereço incompleto, exemplo, apenas cidade e estado, o georreferenciamento apontou localizações na centralidade da área. Dessa forma, é indicado, quando possível, a utilização de registros de endereços completos, maior número de variáveis, para maior fidedignidade dos resultados de georreferenciamento e, conseqüentemente, das análises espaciais.

A API de *geocoding* do Google *Maps* não possui limite diário para o número de solicitações que podem ser feitas aos seus servidores. Porém, existe um limite de uso, no modo gratuito, que esteve relacionado ao número máximo de consultas/requisições por segundo. No máximo 50 requisições por segundo, ou 3.000 requisições por minuto. Se este limite for excedido, um código de status "OVER_QUERY_LIMIT" é retornado como resposta.

O método/ algoritmo desenvolvido neste trabalho para o processo automático de georreferenciamento e organização de dados, realizou requisições de forma simultânea e iterativa. Em alguns testes o limite de uso para o número máximo de solicitações por minuto foi excedido, para contornar esta limitação, foi adicionado um atraso de 200 ms a cada iteração do algoritmo. O tempo de execução das 405 requisições/solicitações do nosso conjunto de dados, para a API de *geocoding*, ocorreu em aproximadamente 6 minutos e 15 segundos, uma média de pelo menos 1 minuto e 30 segundos a cada 100 registros, a partir dessa abordagem adotada.

Uma das características mais importantes deste método disponível na *web* de georreferenciamento, foi a disponibilização dos dados georreferenciados (latitude; longitude)

em arquivo de planilha eletrônica (.xlsx). Além disso, outra característica é a validação destes dados na região da área de estudo de maneira dinâmica, utilizando o mapa do Google e, conseqüentemente, a possibilidade da compatibilização com outros *softwares* de análise espacial de dados.

A API foi testada para um volume de dados de até 5.000 registros, e não foi observado o potencial da plataforma para grandes volumes de dados utilizando a versão gratuita da API do Google *Maps*. Sugere-se também, como trabalho futuro, observar o potencial da API de *geocoding* do Mapbox, que promete até 100.000 solicitações gratuitas de georreferenciamento por mês, limitando-se em até 600 requisições por minuto (MAPBOX... , 2022).

4.2 Associação de região

A associação dos municípios às suas regiões administrativas dependeu do processo descrito na Seção (2.3.1), processamento e organização dos dados. Uma vez que, o método ASSYRIA de processamento de dados, permitiu obter as variáveis de endereço completo (rua, número, bairro, cidade) de forma padronizada e sem erros de ortografia, por exemplo, o seguinte registro: "rua moneiro lobato, 420, bom conselho", foi obtido formatado e padronizado, como: "Rua Monteiro Lobato, 420, Bom Conselho". Letras maiúsculas, minúsculas, espaços no texto ou erros de ortografia, todos foram corrigidos e formatados.

Este processo implica na compatibilização com as malhas municipais dos 26 estados brasileiros para a construção de mapas científicos de área e possíveis cruzamentos de dados com bases públicas do governo, como IBGE e SIDRA. As malhas municipais dos estados brasileiros são padronizadas pelo IBGE e disponibilizadas no formato *shape file* (IBGE, 2021c). Esse formato pode ser lido pela grande maioria dos *softwares* de geoprocessamento e visualizadores de geoinformação.

O algoritmo de associação não possuiu limite de uso, diferente do algoritmo de georreferenciamento. Apenas uma requisição foi feita para o servidor do IBGE, recuperando todos os dados. Os dados foram comparados de forma iterativa e associados uns aos outros, dessa forma, não há restrição ao volume do conjunto de dados.

Quinino et al. (2021), em seu estudo, realizou uma análise espacial e temporal da interiorização da Covid-19 nos municípios do estado de Pernambuco, Brasil, considerando aspectos socioeconômicos correlacionados. No entanto, a discussão da análise ficou limitada aos municípios do estado. Trabalhos futuros e semelhantes ao de Quinino et al. (2021) podem se beneficiar deste processo de associação automática de regiões para acelerar o processo de organização de dados e, conseqüentemente, de construção de análises espaciais.

Além disso, gestores, secretários de saúde ou tomadores de decisão, também poderão ser beneficiados com mais informações e intuições para a priorização de políticas de saúde pública em municípios de uma mesorregião ou microrregião específica.

Neves et al. (2019), observaram a incidência e distribuição espacial da leucemia mielóide crônica, doença hematológica maligna caracterizada por um distúrbio clonal, no estado de Pernambuco, Brasil. A incidência e a distribuição espacial foi observada nos níveis político-administrativos de município e microrregião. No entanto, o processo de organização dos dados a fim de possibilitar a construção das análises espaciais por região administrativa não foi descrito em seu estudo. Descrever todos os processos, inclusive de organização de dados, poderia facilitar a replicação do estudo em outros cenários. O algoritmo de associação desenvolvido neste estudo se mostrou relevante para suprir esta lacuna. Pois, possibilita que outros pesquisadores possam utilizá-lo em seus estudos através de uma interface disponível na *web* (<<https://spatialworkspace.ddns.net/region>>).

A replicação deste método para outros países ficou restrito, pois, a API do IBGE fornece apenas informações sobre a divisão político-administrativa dos estados brasileiros. Porém, acredita-se que o método AR, descrito na Seção (2.3.2), pode ser modelo para construção de outros algoritmos de associação de região em outros países do mundo. Assim, contribuindo também para a observação das suas características sociais ou demográficas, a partir das suas macro e micro áreas.

4.3 Análises espaciais

4.3.1 Mapeamento coroplético

A combinação dos métodos de georreferenciamento e associação de região possibilitaram a construção dos seguintes mapas: **(i)** mapas coropléticos do número total de casos; e **(ii)** os mapas de prevalência das três doenças (fibrose cística, hiperplasia adrenal congênita e hemoglobinopatias) nos níveis administrativos de mesorregião, microrregião e município.

Segundo Wei, Tong e Phillips (2017), o mapeamento coroplético é uma importante técnica exploratória de análise de dados espaciais que tem sido amplamente utilizada para explorar visualmente o padrão espacial de distribuições de atributos entre regiões. Além disso, é um dos métodos mais amplamente utilizados para explorar visualmente as distribuições espaciais de dados demográficos e socioeconômicos (MURAD, 2018).

Esta análise possibilitou a visualização de padrões espaciais, ou seja, locais com maior concentração de números de casos e principalmente áreas em que podem estar

desassistidas pela saúde pública local. Este trabalho se limitou à observação dos indicadores de número total de casos e prevalência. No entanto, a partir da abordagem utilizada, os resultados apontaram que outros indicadores demográficos e sociais também podem ser utilizados. Por exemplo, indicadores de: população branca ou negra; educação; IDH; e trabalho e rendimento, podem ser utilizados para a construção de outros mapas que possam trazer informações relevantes para planejamento estratégico e tomadas de decisão pelo governo ou órgão competente local.

Outra limitação deste estudo foi não ter observado de forma individual a distribuição espacial de cada doença. Pois, o propósito da construção desta análise, utilizando as ferramentas de georreferenciamento e associação de região, foi evidenciar o potencial de replicação dos métodos para outros contextos de estudo. Além disso, apontar possibilidades de construção de outros mapas coropléticos utilizando outros indicadores. Também não foi possível realizar uma avaliação das interfaces dos métodos que possibilitaram a construção destas análises com usuários reais. Sugere-se, como trabalho futuro, uma avaliação das interfaces através de heurísticas ou testes de usabilidade.

Sugere-se também, como trabalho futuro, a implementação deste método na *web* através da biblioteca *JavaScript* Plotly de mapeamento coroplético ([CHOROPLETH...](#), 2022), com os pacote de dados vetorizados por mesorregião, microrregião e municípios no formato GeoJSON ([GEOGRAPHIC...](#), 2022). Tal sugestão poderia trazer contribuições significativas para a área de análise espacial de dados. Além disso, praticidade para profissionais e pesquisadores que não possuem habilidade de construção de mapas científicos utilizando sistemas de informação geográfica, como ArcGIS e QGIS.

4.3.2 *Kernel density estimation*

O método KDE utilizado nesta análise possibilitou observar áreas de focos de calor, ou seja, locais com forte concentração de números de casos. Esta análise complementou a análise espacial de área e regiões, descrita na Seção (2.3.3.1). Dessa forma, foi agregado mais informações ao estudo, trazendo um conceito de visualização do mais amplo, casos por área, para o mais específico, local pontual dos casos.

Outros tamanhos de raio do *kernel* também foram testados, como 15 km e 20 km. Porém, em alguns casos, eles ficavam maiores que a extensão territorial de alguns municípios, dificultando a visualização da distribuição espacial. Logo, o tamanho do raio em 10 km, neste estudo, se mostrou ideal para complementar o conceito de visualização espacial do mais amplo para o mais específico.

A utilização do método KDE também trouxe outras intuições para utilização de

outros métodos de análise espacial pontual. Por exemplo, utilização de métodos inferenciais para prever a construção de um futuro hospital ou centro de atendimento em uma área específica. Verificar se os centros de atendimento existentes estão distribuídos corretamente, diante da distribuição espacial dos casos, ou a identificação de áreas que ainda possam estar carentes de atendimentos de saúde.

4.3.3 Dimensão de correlação

Nesta seção serão discutidos os resultados e as intuições obtidas a partir da aplicação do método de dimensão de correlação. A seção foi organizada em: **a)** aplicação do método; e **b)** abordagem na *web*.

a) Aplicação do método

Neste trabalho foi utilizado uma abordagem fractal, através do método de dimensão de correlação, para comparar as variações de agrupamento espacial de eventos da área das ciências da saúde. As análises foram realizadas nos níveis de mesorregiões e microrregiões no estado do Rio Grande do Sul, Brasil. Para tanto, foram utilizados dados de três doenças triadas durante o teste do pezinho, foram elas: fibrose cística; hiperplasia adrenal congênita e hemoglobinopatias.

A análise comparativa possibilitou observar mesorregiões e microrregiões com os maiores e menores índices de dimensão de correlação, ou seja, áreas com maior dispersão entre seus eventos e áreas com maiores agrupamentos. A construção desta análise só foi possível a partir do desenvolvimento do algoritmo de associação de região, descrito na Seção (2.3.2).

Os trabalhos identificados na literatura, até o momento, utilizaram do método de dimensão de correlação fractal para realizar análises comparativas ao longo do tempo numa mesma área, seja na área delimitada à um país ou local específico. Entre os trabalhos, podemos citar as obras de Öncel et al. (1996); Telesca et al. (2007a); Telesca et al. (2007b); Soares (2010); e Mondal, Roy e Behera (2017). Diferente dos trabalhos citados, neste trabalho, utilizou-se uma abordagem comparativa entre diferentes áreas da divisão político-administrativa de um estado brasileiro.

De acordo com o IBGE, as micro e mesorregiões brasileiras foram especificadas para estabelecer uma maior compreensão dos territórios com base em sua extensão, clima e demografia (IBGE, 1990). Em que, a compreensão destas regiões geográficas proporciona uma visão mais intimista e capaz de aferir as diferenças e peculiaridades de um determinado território, diferentemente do que ocorre na análise de macrorregiões, que é mais generalista.

Sendo assim, a abordagem utilizada neste trabalho, traz também intuições de possíveis correlação com dados disponíveis em bases de dados abertas, como IGBE (IBGE, 2022) e Departamento de Informática do Sistema Único de Saúde (DATASUS) (DATASUS, 2022), para melhor compreender características de um fenômeno específico. Por exemplo, por que os maiores índices de agrupamentos estão na região sul do estado, próximos à fronteira com o Uruguai?

Durante o estudo não foi identificado a utilização do método de dimensão de correlação em periódicos e conferências da área das ciências da saúde. Dessa forma, este trabalho também buscou trazer um método comumente utilizado na área da física estatística para a área da saúde. Esta contribuição traz uma alternativa complementar a outros métodos frequentemente utilizados na área da saúde para analisar dados de eventos pontuais, como o método KDE [(MONTEIRO et al., 2004); (GOMES et al., 2012); (MURAD, 2018); (REX; BORGES; KÄFER, 2020)].

Gestores, secretários de saúde e pesquisadores podem se beneficiar da utilização desta abordagem e dos resultados obtidos neste trabalho. Tais benefícios incluem: planejamento estratégico; e tomadas de decisão. Por exemplo, a construção de polos de atendimento especializado em áreas com maior índice de agrupamento; distribuição de verba para unidades de saúde baseado em resultados numéricos; otimização de recursos públicos em áreas menos carentes de atenção ao cuidado humano; e investigação de causa de determinados acontecimentos.

Além disso, trabalhos futuros e semelhantes ao de (TELESCA et al., 2007a) e (SOARES, 2010), na área da física estatística, que estudaram a variabilidade temporal de focos de incêndio em macroáreas, podem também se beneficiar das intuições obtidas neste trabalho para estudar a distribuição de focos de incêndio por mesorregiões e microrregiões. Trazendo, assim, mais evidências aos seus estudos e uma maior previsibilidade de riscos de queima de vegetação em áreas específicas.

Não foi possível realizar as análises fractais sobre os municípios, devido ao pequeno número de ocorrência de casos.

b) Abordagem na web

Diferentes avanços na construção de sistemas baseados na *web* para concepção de análises espaciais foram identificados na literatura. Bui e Pham (2016), propuseram uma estrutura técnica baseada na *web* para visualização de dados de saúde e construção de análises espaciais. Dados de malária no Vietnã e os métodos de vizinho mais próximo, função K e autocorrelação espacial foram utilizados para validar a estrutura proposta. A estrutura baseada na *web* se mostrou relevante para auxiliar profissionais de saúde em

processos de tomada de decisão.

Moraga (2017), apresentou o SpatialEpiApp, uma aplicação *web* que integrava duas das abordagens mais comuns em vigilância em saúde: mapeamento de doenças e detecção de agrupamentos. A aplicação foi validada utilizando dados de mortalidade por câncer de pulmão em Ohio, Estados Unidos. Métodos bayesianos foram utilizados para obter estimativa de risco da doença. A aplicação se mostrou ser de fácil uso e não dependia de nenhum conhecimento prévio de programação.

Lan et al. (2020), construíram um sistema de apoio à decisão espacial baseado na *web* para facilitar o monitoramento da contaminação da água em poços privados em Gaston County, Carolina do Norte (EUA). Para tanto, utilizaram uma abordagem de interpolação espacial utilizando os métodos de *Inverse Distance Weighting* e *Kriging*. O sistema teve por objetivo permitir que tomadores de decisão, especialistas em saúde ambiental ou o público em geral realizasse processamento automatizado de dados espaciais e análises de dados sob demanda.

Schmidt, Dröge-Rothaar e Rienow (2021), desenvolveram um sistema de informação geográfica baseado na *web* para detecção e análise de casos de Covid-19 na cidade de Colônia, Alemanha. Para tal, a aplicação na *web* solicitava de forma voluntária o compartilhamento de informações sobre a infecção pelo SARS-CoV-2.

Apesar destes avanços, não foi identificado na literatura uma abordagem na *web* de análise fractal de dados utilizando o método de dimensão de correlação. Inclusive, Moraga (2017) e Lan et al. (2020) incentiva a construção de mais funcionalidades de análise espacial de dados na *web* para melhorar intuições, discussões e estimativas de riscos em estudos na área da saúde.

Dessa forma, este trabalho contribui para com a literatura com mais um método disponível na *web* para análise de agrupamento espacial de eventos pontuais. Além disso, esta abordagem busca trazer mais intuições sobre a distribuição de eventos pontuais, a partir da análise exploratória de dados espaciais. Por exemplo, determinando se os eventos observados exibem algum padrão sistemático, em oposição à uma distribuição aleatória.

Pesquisas em outros cenários também são incentivadas, através da abordagem utilizada neste trabalho. Por exemplo, o estudo temporal e comparativo da dimensão de correlação dos casos de Covid-19, e outras doenças, por mesorregião e microrregião.

5 Conclusões

Neste estudo, foi desenvolvida uma abordagem baseada na *web* para processamento de dados e construção de análises espaciais de área e pontos. Para tanto, foram construídos dois métodos automatizados para concepção das análises: a) ASSYRIA, para processamento e organização de dados espaciais; e b) AR, para associação de dados ao seu contexto espacial, ou seja, às suas microrregiões e mesorregiões. Até onde se sabe, esta é a primeira abordagem na *web* que processa, organiza, valida e disponibiliza dados geográficos em formato de planilha eletrônica (.xlsx) para construção de mapas científicos de área e pontos utilizando GIS ou semelhantes.

Este estudo também trouxe uma aplicação do método de dimensão de correlação (*Correlation Dimension*), da análise fractal de dados, comumente utilizado na área da física estatística, para a área das ciências da saúde. Além disso, o método foi disponibilizado na *web* para utilização por outros pesquisadores ou estudiosos para construção de análises de agrupamento espacial de eventos pontuais. Até então, esta também é a primeira abordagem na *web* do método.

Os resultados obtidos destacaram o potencial de replicação da abordagem para outros contextos de estudo. Os métodos/algoritmos utilizados e desenvolvidos se mostraram relevantes no contexto de análise espacial, possibilitando velocidade no processamento, organização de dados e, conseqüentemente, na construção de resultados significativos que possam ser utilizados para conhecer melhor a realidade local e tomar decisões mais direcionadas em cada situação, assim, impactando diretamente na qualidade de vida das pessoas e nos desafios da saúde.

Acredita-se também que a abordagem desenvolvida neste estudo pode ser útil de ser utilizada por usuários de baixo nível de instrução estatística ou de programação que, em algumas situações, podem ser profissionais de saúde, gestores ou pesquisadores. Por fim, é recomendado a continuidade de pesquisas envolvendo a abordagem proposta em outras áreas do conhecimento e a disponibilização de outros métodos de análise espacial de dados na *web*.

Referências bibliográficas

- ACRONYMIFY! **Automatically generate fun acronyms for your project**. 2021. <<https://acronymify.com/>>. (Accessed on 10/11/2021).
- AGRAWAL, S.; GUPTA, R. D. Development and comparison of open source based web gis frameworks on wamp and apache tomcat web servers. **The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences**, Copernicus GmbH, v. 40, n. 4, p. 1, 2014.
- ALGARIN, A. B. et al. Spatial distribution of partner-seeking men who have sex with men using geosocial networking apps: epidemiologic study. **Journal of medical Internet research**, JMIR Publications Inc., Toronto, Canada, v. 20, n. 5, p. e9919, 2018.
- ARCGIS. **Geocoding | Documentation**. 2021. <<https://developers.arcgis.com/documentation/mapping-apis-and-services/search/geocoding/>>. (Accessed on 08/13/2021).
- AZEVEDO, L. et al. Geostatistical covid-19 infection risk maps for portugal. **International Journal of Health Geographics**, BioMed Central, v. 19, n. 1, p. 1–8, 2020.
- BING. **Find a Location by Address | Microsoft Docs**. 2021. <<https://docs.microsoft.com/en-us/bingmaps/rest-services/locations/find-a-location-by-address>>. (Accessed on 08/13/2021).
- BOQUETT, J. A. et al. Spatial analyzes of hla data in rio grande do sul, south brazil: genetic structure and possible correlation with autoimmune diseases. **International journal of health geographics**, Springer, v. 17, n. 1, p. 1–12, 2018.
- BOTLER, J. **Avaliação de desempenho do Programa de Triagem Neonatal do Estado do Rio de Janeiro**. Tese (Doutorado) — Escola Nacional de Saúde Pública Sergio Arouca, Rio de Janeiro, 2010.
- BUI, T. Q.; PHAM, H. M. Web-based gis for spatial pattern detection: application to malaria incidence in vietnam. **SpringerPlus**, SpringerOpen, v. 5, n. 1, p. 1–14, 2016.
- CÂMARA, G.; CASANOVA, M. A.; MAGALHÃES, G. C. Anatomia de sistemas de informação geográfica. 1996.
- CHAIKAEW, N.; TRIPATHI, N. K.; SOURIS, M. Exploring spatial patterns and hotspots of diarrhea in chiang mai, thailand. **International Journal of Health Geographics**, BioMed Central, v. 8, n. 1, p. 1–10, 2009.
- CHOROPLETH Mapbox in JavaScript. 2022. <<https://plotly.com/javascript/mapbox-county-choropleth/>>. (Accessed on 02/04/2022).

- CLOUD, G. **APIs de geolocalização | Google Maps Platform**. 2021. <<https://cloud.google.com/maps-platform/?hl=pt-PT>>. (Accessed on 09/22/2021).
- COX, D. R.; ISHAM, V. **Point processes**. [S.l.]: CRC Press, 1980. v. 12.
- CROMLEY, E. K.; MCLAFFERTY, S. L. **GIS and public health**. [S.l.]: Guilford Press, 2011.
- DATASUS, D. de Informática do Sistema Único de S. **Sobre o DATASUS**. 2022. <<https://datasus.saude.gov.br/sobre-o-datasus/>>. (Accessed on 02/05/2022).
- DENT, B. D.; TORGUSON, J.; HODLER, T. Cartography: thematic map design. dubuque: Wm. c. **C. Brown Geographic Designs**, 1993.
- DOCS, M. W. **An overview of HTTP**. 2021. <<https://developer.mozilla.org/en-US/docs/Web/HTTP/Overview>>. (Accessed on 12/23/2021).
- ENDALEW, M.; SHIFERAW, W.; KINDIE, A. Development of web mapping application for spatial ecotourism information using qgis plugin and freely available web platforms in north west highlands of ethiopia a case study: Chokie mountain watersheds. **Advances in Internet of Things**, Scientific Research Publishing, v. 9, n. 3, p. 50–61, 2019.
- GEOGRAPHIC Information Systems (GIS) Dataset Brasil - Coleção de shapefiles, GeoJSON e TopoJSON prontas para uso. 2022. <<https://github.com/fititnt/gis-dataset-brasil>>. (Accessed on 02/04/2022).
- GOLDBERG, D. W. et al. An effective and efficient approach for manually improving geocoded data. **International journal of health geographics**, Springer, v. 7, n. 1, p. 1–20, 2008.
- GOMES, E. et al. Schistosomiasis transmission and environmental change: a spatio-temporal analysis in porto de galinhas, pernambuco-brazil. **International journal of health geographics**, BioMed Central, v. 11, n. 1, p. 1–11, 2012.
- GOODCHILD, M. F. Data analysis, spatial. In: _____. **Encyclopedia of GIS**. Boston, MA: Springer US, 2008. p. 200–203. ISBN 978-0-387-35973-1. Disponível em: <https://doi.org/10.1007/978-0-387-35973-1_236>.
- GOOGLE. **Geocoding Service | Maps JavaScript API**. 2021. <<https://developers.google.com/maps/documentation/javascript/geocoding>>. (Accessed on 08/13/2021).
- GRASSBERGER, P.; PROCACCIA, I. Measuring the strangeness of strange attractors. In: **The theory of chaotic attractors**. [S.l.]: Springer, 2004. p. 170–189.
- HERE. **Geocoding & Search API**. 2021. <https://developer.here.com/documentation/geocoding-search-api/dev_guide/topics/quick-start.html>. (Accessed on 08/13/2021).
- HIGHFIELD, L. et al. Interactive web-based mapping: bridging technology and data for health. **International Journal of Health Geographics**, BioMed Central, v. 10, n. 1, p. 1–13, 2011.

- IBGE. **Divisão do Brasil em Mesorregiões e Microrregiões Geográficas**. [S.l.]: Ministério da Economia, Fazenda e Planejamento Rio de Janeiro, 1990.
- IBGE. **Pesquisa Nacional por Amostra de Domicílios e Contagem da População**. [S.l.]: Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística (IBGE), 2010.
- IBGE. **API de Localidades**. 2021. <<https://servicodados.ibge.gov.br/api/docs/localidades#api-Municipios-estadosUFMunicipiosGet>>. (Accessed on 08/13/2021).
- IBGE. **Malha Municipal: Downloads**. 2021. <<https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/15774-malhas.html?=&t=downloads>>. (Accessed on 08/13/2021).
- IBGE. **Malha Municipal: O que é?** 2021. <<https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/15774-malhas.html?=&t=o-que-e>>. (Accessed on 08/13/2021).
- IBGE. **Portal Brasileiro de Dados Abertos**. 2022. <<https://dados.gov.br/organization/instituto-brasileiro-de-geografia-e-estatistica-ibge>>. (Accessed on 02/05/2022).
- JENKS, G. F. The data model concept in statistical mapping. **International yearbook of cartography**, v. 7, p. 186–190, 1967.
- JOHN, M. **A dictionary of epidemiology**. [S.l.]: Oxford university press Oxford, UK, 2001.
- KOPACEK, C. et al. Evolução e funcionamento do programa nacional de triagem neonatal no rio grande do sul de 2001 a 2015. **Boletim Científico de Pediatria-Vol**, v. 4, n. 3, p. 71, 2015.
- LAN, Y. et al. A web-based spatial decision support system for monitoring the risk of water contamination in private wells. **Annals of GIS**, Taylor & Francis, v. 26, n. 3, p. 293–309, 2020.
- LEVINE, N. Crime mapping and the crimestat program. **Geographical analysis**, Wiley Online Library, v. 38, n. 1, p. 41–56, 2006.
- LI, S. et al. Geospatial big data handling theory and methods: A review and research challenges. **ISPRS journal of Photogrammetry and Remote Sensing**, Elsevier, v. 115, p. 119–133, 2016.
- LOCATIONIQ. **Fast Geocoding and Reverse Geocoding service from Unwired Labs**. 2021. <<https://locationiq.com/docs#search-forward-geocoding>>. (Accessed on 08/13/2021).
- MAAS, C. et al. A spatial analysis of referrals to a primary mental health programme in western sydney from 2012 to 2015. Thomson Reuters, 2019.
- MACEACHREN, A. M. **Some truth with maps: A primer on symbolization and design**. [S.l.]: Assn of Amer Geographers, 1994.

- MANDELBROT, B. How long is the coast of britain? statistical self-similarity and fractional dimension. **science**, American Association for the Advancement of Science, v. 156, n. 3775, p. 636–638, 1967.
- MANDELBROT, B. B.; MANDELBROT, B. B. **The fractal geometry of nature**. [S.l.]: WH freeman New York, 1982. v. 1.
- MAPBOX | Geocoding API. 2022. <https://docs.mapbox.com/api/search/geocoding/?utm_source=mapboxcom&utm_medium=pricing&utm_content=temporary-api#geocoding-api-pricing>. (Accessed on 02/04/2022).
- MCMMASTER, R. In memoriam: George f. jenks (1916-1996). **Cartography and Geographic Information Systems**, Taylor & Francis, v. 24, n. 1, p. 56–59, 1997.
- MONDAL, D.; ROY, P.; BEHERA, P. Use of correlation fractal dimension signatures for understanding the overlying strata dynamics in longwall coal mines. **International Journal of Rock Mechanics and Mining Sciences**, Elsevier, v. 91, p. 210–221, 2017.
- MONTEIRO, A. M. V. et al. Análise espacial de dados geográficos. **Brasília: Embrapa**, 2004.
- MORAGA, P. Spatialepiapp: A shiny web application for the analysis of spatial and spatio-temporal disease data. **Spatial and spatio-temporal epidemiology**, Elsevier, v. 23, p. 47–57, 2017.
- MURAD, A. Planning and location of health care services in jeddah city, saudi arabia: Discussion of the constructive use of geographical information systems. **Geospatial health**, v. 13, n. 2, 2018.
- NEVES, W. B. d. et al. Incidence and spatial distribution of chronic myeloid leukemia by regions of economic development in the state of pernambuco, brazil. **Hematology, transfusion and cell therapy**, SciELO Brasil, v. 41, p. 212–215, 2019.
- ÖNCEL, A. O. et al. Spatial variations of the fractal properties of seismicity in the anatolian fault zones. **Tectonophysics**, Elsevier, v. 257, n. 2-4, p. 189–202, 1996.
- OPENCAGE. **Geocoding API Documentation**. 2021. <<https://opencagedata.com/api#forward-resp>>. (Accessed on 08/13/2021).
- OVERFLOW, S. **Developer Survey, 2020**. 2021. <<https://insights.stackoverflow.com/survey/2020#technology-programming-scripting-and-markup-languages-professional-developers>>. (Accessed on 08/13/2021).
- PLATFORM, G. M. **Geocoding Service**. 2021. <<https://developers.google.com/maps/documentation/javascript/geocoding>>. (Accessed on 12/23/2021).
- PLOTLY. **Line Plots**. 2022. <<https://plotly.com/javascript/line-charts/>>. (Accessed on 02/02/2022).
- QGIS. **User Guide, version 3.10**. 2021. <https://docs.qgis.org/3.10/en/docs/user_manual/>. (Accessed on 08/13/2021).

QUININO, L. R. d. M. et al. Spatial and temporal aspects and factors associated with the spread of covid-19 to the interior of the state of pernambuco, brazil. **Ciência & Saúde Coletiva**, SciELO Brasil, v. 26, p. 2171–2182, 2021.

R. **The R Project for Statistical Computing**. 2022. <<https://www.r-project.org/>>. (Accessed on 02/04/2022).

REACT. **A JavaScript library for building user interfaces**. 2021. <<https://reactjs.org/>>. (Accessed on 08/13/2021).

REX, F. E.; BORGES, C. A. d. S.; KÄFER, P. S. Spatial analysis of the covid-19 distribution pattern in são paulo state, brazil. **Ciência & Saúde Coletiva**, SciELO Public Health, v. 25, p. 3377–3384, 2020.

SCHMIDT, F.; DRÖGE-ROTHAAR, A.; RIENOW, A. Development of a web gis for small-scale detection and analysis of covid-19 (sars-cov-2) cases based on volunteered geographic information for the city of cologne, germany, in july/august 2020. **International Journal of Health Geographics**, BioMed Central, v. 20, n. 1, p. 1–24, 2021.

SHANMUKHAPPA, T.; HO, I. W.-H.; TSE, C. K. Spatial analysis of bus transport networks using network theory. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 502, p. 295–314, 2018.

SHIODE, N. et al. The mortality rates and the space-time patterns of john snow’s cholera epidemic map. **International journal of health geographics**, BioMed Central, v. 14, n. 1, p. 1–15, 2015.

SIDRA. **Tabela 3175: População residente**. 2021. <<https://sidra.ibge.gov.br/Tabela/3175>>. (Accessed on 08/13/2021).

SILVEIRA, I. H. d.; OLIVEIRA, B. F. A. d.; JUNGER, W. L. Utilização do google maps para o georreferenciamento de dados do sistema de informações sobre mortalidade no município do rio de janeiro, 2010-2012. **Epidemiologia e Serviços de Saúde**, SciELO Brasil, v. 26, p. 881–886, 2017.

SOARES, M. F. M. **CORRELAÇÕES ESPACIAIS DOS FOCOS DE CALOR NO BRASIL**. Dissertação (Mestrado) — Universidade Federal Rural de Pernambuco, Departamento de Estatística e Informática, Recife, 2010.

TELESCA, L. et al. Identifying spatial clustering properties of the 1997–2003 liguria (northern italy) forest-fire sequence. **Chaos, Solitons & Fractals**, Elsevier, v. 32, n. 4, p. 1364–1370, 2007.

TELESCA, L. et al. Space–time fractal properties of the forest-fire series in central italy. **Communications in Nonlinear Science and Numerical Simulation**, Elsevier, v. 12, n. 7, p. 1326–1333, 2007.

TOSI, P. Seismogenic structure behaviour revealed by spatial clustering of seismicity in the umbria-marche region (central italy). **Annals of Geophysics**, v. 41, n. 2, 1998.

VALJAREVIĆ, A. et al. A geographical information systems-based approach to health facilities and urban traffic system in belgrade, serbia. **Geospatial health**, Univ Naples Federico II, Naples, v. 13, n. 2, p. 308–313, 2018.

VOUGAT, R. R. N. et al. Using google earth™ and geographical information system data as method to delineate sample domains for an urban household surveys: the case of maroua (far north region-cameroon). **International Journal of Health Geographics**, BioMed Central, v. 18, n. 1, p. 1–12, 2019.

W3SCHOOLS. **What is JavaScript?** 2021. <https://www.w3schools.com/whatis/whatis_js.asp>. (Accessed on 08/13/2021).

WEI, R.; TONG, D.; PHILLIPS, J. M. An integrated classification scheme for mapping estimates and errors of estimation from the american community survey. **Computers, Environment and Urban Systems**, Elsevier, v. 63, p. 95–103, 2017.

WORTON, B. J. Kernel methods for estimating the utilization distribution in home-range studies. **Ecology**, Wiley Online Library, v. 70, n. 1, p. 164–168, 1989.

YANG, D.-H. et al. Improving geocoding practices: evaluation of geocoding tools. **Journal of medical systems**, Springer, v. 28, n. 4, p. 361–370, 2004.

ZERIHUN, M. E. et al. Web based gis for tourism development using effective free and open source software case study: Gondor town and its surrounding area, ethiopia. **Journal of Geographic Information System**, Scientific Research Publishing, v. 9, n. 01, p. 47, 2017.

Apêndices

APÊNDICE A – Spatial Workspace

Spatial Workspace

Spatial Workspace é uma plataforma *web* automatizada para gerar e organizar dados espaciais de pontos e áreas a partir de endereços (texto). Além disso, a plataforma permite a análise da dimensão fractal de dados georreferenciados utilizando o método *Correlation Dimension* e a compatibilidade com Sistemas de Informação Geográfica (GIS) para a construção de mapas coropléticos e outros tipos de análises espaciais, como mapas de kernel. ter-los

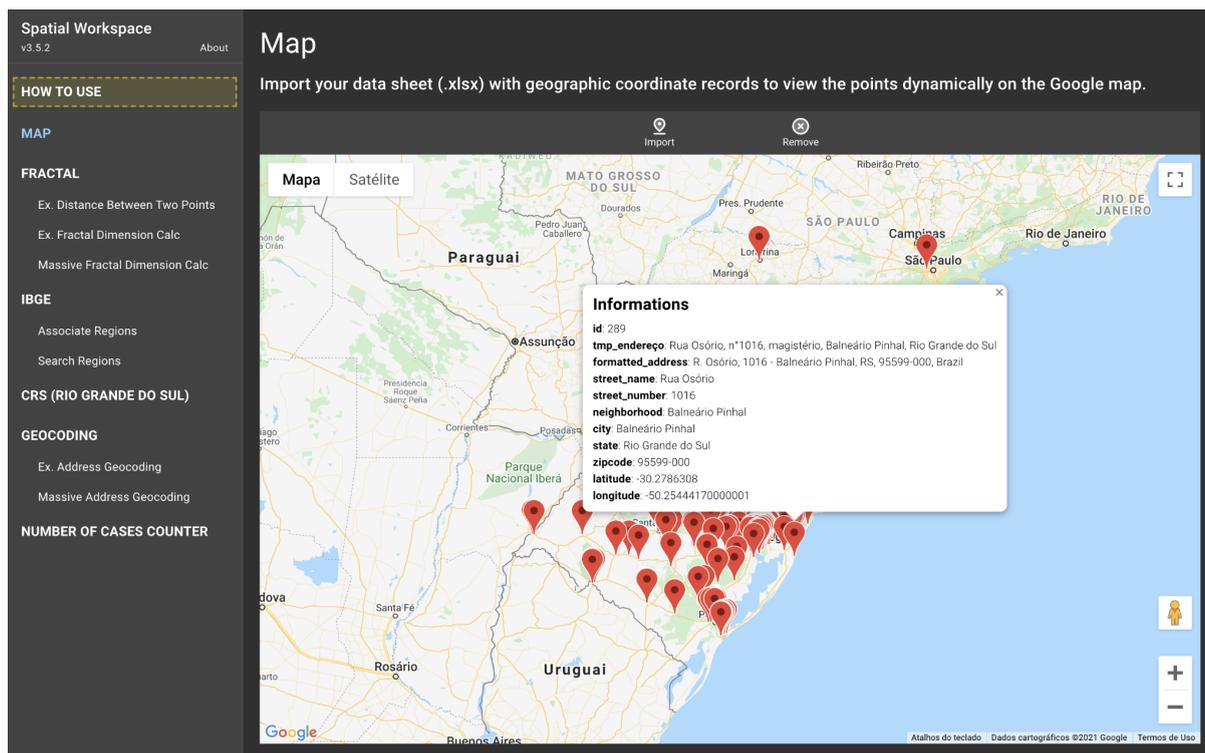
Sumário

MAP	1
FRACTAL	2
Ex. Distance Between Two Points	2
Ex. Fractal Dimension Calc	3
Massive Fractal Dimension Calc	4
IBGE	5
Associate Regions	5
Search Regions	6
CRS (RIO GRANDE DO SUL)	7
GEOCODING	8
Ex. Address Geocoding	8
Massive Address Geocoding	9
NUMBER OF CASES COUNTER	10

MAP

Nesta interface é possível importar dados georreferenciados (latitude e longitude) de um arquivo de planilha eletrônica (.xlsx) para visualização dinâmica dos eventos no mapa do Google. Ao clicar nos pontos sobre o mapa, é possível visualizar as informações associadas a eles.

Além disso, esta visualização possibilita observar quais os pontos de eventos que estão dentro da sua área de estudo ou não.



FRACTAL

Ex. Distance Between Two Points

Esta interface exemplifica a distância angular entre 2 pares de eventos. Clicando sobre o mapa é possível alterar o local dos eventos (origem e destino). A distância angular e em quilômetros entre os eventos é calculada automaticamente após os cliques sobre o mapa. Esta interface pode ser utilizada para fins educativos.

Mais informações sobre o cálculo da distância angular pode ser vista em: [clique aqui](#).

Spatial Workspace
v3.5.2 About

HOW TO USE

MAP

FRACTAL

- Ex. Distance Between Two Points
- Ex. Fractal Dimension Calc
- Massive Fractal Dimension Calc

IBGE

- Associate Regions
- Search Regions

CRS (RIO GRANDE DO SUL)

GEOCODING

- Ex. Address Geocoding
- Massive Address Geocoding

NUMBER OF CASES COUNTER

Example of Distance Between Two Points

Click on the map to change the position of the markers. The angular distance between the two points is automatically calculated.

Origin

Latitude	-6,329810195253717
Longitude	-58,009084425000005

Destination

Latitude	-17,154741828578917
Longitude	-43,595021925000005

Info

1975.47 km of distance.
And the angular distance is 0.31007.

Mapa Satélite

Google

Atalhos do teclado Dados cartográficos ©2021 Google, INEGI Termos de Uso

Ex. Fractal Dimension Calc

Esta interface exemplifica o cálculo da dimensão fractal por meio do método *Correlation Dimension* entre n pares de eventos. Clique sobre o mapa para adicionar novos pontos de evento. O cálculo da dimensão fractal é calculada automaticamente após os cliques sobre o mapa. Variações entre a distância dos eventos pode ser realizada no input slide. Esta interface pode ser utilizada para fins educativos.

Mais informações sobre o método *Correlation Dimension* pode ser vista no artigo científico "[Identifying spatial clustering properties of the 1997–2003 Liguria \(Northern Italy\) forest-fire sequence](#)".

Spatial Workspace v3.5.2 About

HOW TO USE

MAP

FRACTAL

Ex. Distance Between Two Points

Ex. Fractal Dimension Calc

Massive Fractal Dimension Calc

IBGE

Associate Regions

Search Regions

CRS (RIO GRANDE DO SUL)

GEOCODING

Ex. Address Geocoding

Massive Address Geocoding

NUMBER OF CASES COUNTER

Example of Fractal Dimension Calculation

Click on the map to drop n event points. The fractal dimension is automatically calculated.

Number of points on the map: 19 event points

Fractal dimension value: The value is **0.80702**

Kilometers: 70

Move the input slider to change the km value.

Info
70 km of distance.
And the angular distance is **0.01099**.

Mapa Satélite

Google Maps

Massive Fractal Dimension Calc

Nesta interface é possível calcular a dimensão fractal, por meio do método *Correlation Dimension*, de um conjunto de dados georreferenciados. Para calcular a dimensão fractal você deve importar um arquivo de dados no formato .xlsx e selecionar as colunas que representam as coordenadas geográficas (lat & lng). Após isso, você deve escolher uma área de interesse para que o cálculo possa ser realizado de modo comparativo entre as áreas.

Também, é possível realizar um filtro sobre os dados, exemplo: registros de uma determinada raça ou doença. Dinamicamente é calculada a dimensão fractal e apresentada na tabela ao lado. Variações entre a distância dos eventos pode ser realizada no input slide. Clicando no botão "submit" é possível realizar o download das informações apresentadas na tabela.

Spatial Workspace
v3.5.2 About

HOW TO USE

MAP

FRACTAL

- Ex. Distance Between Two Points
- Ex. Fractal Dimension Calc
- Massive Fractal Dimension Calc

IBGE

- Associate Regions
- Search Regions

CRS (RIO GRANDE DO SUL)

GEOCODING

- Ex. Address Geocoding
- Massive Address Geocoding

NUMBER OF CASES COUNTER

Massive Fractal Dimension Calculation

Import your georeferenced data sheet (.xlsx) to automatically calculate the fractal dimension for a specific region.

Attached file

[Click here to import a file](#)

Geographic columns

latitude longitude

Region

nm_mesol

This column will be used to calculate fractal dimension per area

Filter n values if you want

id

city

disease

cd_mun

cd_micro

nm_micro

region	dimension	cases
Centro Oriental Rio-grandense	0.28387	31
Nordeste Rio-grandense	0.24599	34
Centro Ocidental Rio-grandense	0.61538	14
Sudeste Rio-grandense	0.59192	45
Metropolitana de Porto Alegre	0.41921	215
Noroeste Rio-grandense	0.07201	42
Sudoeste Rio-grandense	0.34783	24

IBGE

Associate Regions

Nesta interface é possível associar as regiões político-administrativas do Instituto Brasileiro de Geografia e Estatística (IBGE) às cidades do seu conjunto de dados. Para realizar a associação você deve selecionar a Unidade Federativa (UF) em estudo. Ao selecionar, dinamicamente é recuperada, a partir da API de Localidades do IBGE, todas as cidades da UF selecionada com as regiões em que elas pertencem.

Após selecionada a Unidade Federativa, importe um arquivo de dados no formato .xlsx e selecione a coluna que representa a cidade dos seus registros para realizar a associação de regiões. Clicando no botão "submit" é feito o download de um arquivo .xlsx com a associação gerada entre as cidades e regiões. Os registros que obtiveram sucesso na associação de região, estarão na folha de "sucessos". Já os registros que, possivelmente, obtiveram falha na associação de região, estarão na folha de "falhas". Para os registros que obtiveram falha, é indicado observar se existe alguma diferença entre o nome do registro no arquivo .xlsx e o nome listado pela API de Localidades do IBGE.

Obs.: as associações geradas facilitam a construção de mapas temáticos no QGIS e ArcGIS. Mais informações sobre a API de Localidades do IBGE pode ser vista em: [clique aqui](#).

```
[
  {
    "id": 4300034,
    "nome": "Aceguá",
    "microrregiao": {
      "id": 43031,
      "nome": "Campanha Meridional",
      "mesorregiao": {
        "id": 43006,
        "nome": "Sudoeste Rio-grandense",
        "UF": {
          "id": 43,
          "sigla": "RS",
          "nome": "Rio Grande do Sul",
          "regiao": {
            "id": 4,
            "sigla": "S",
            "nome": "Sul"
          }
        }
      }
    }
  },
  {
    "regiao-imediata": {
      "id": 430010,
      "nome": "Bagé",
      "regiao-intermediaria": {
        "id": 4302,
        "nome": "Pelotas",
        "UF": {
          "id": 43,
          "sigla": "RS",
          "nome": "Rio Grande do Sul",
          "regiao": {
            "id": 4,
            "sigla": "S",
            "nome": "Sul"
          }
        }
      }
    }
  }
]
```

Search Regions

Nesta interface é possível filtrar cidades de uma Unidade Federativa (UF) específica e visualizar suas regiões. Além disso, esta visualização possibilita encontrar possíveis erros de ortografia entre os seus dados e os dados disponíveis na API de localidades do IBGE.

Spatial Workspace
v3.5.2 About

HOW TO USE

MAP

FRACTAL
Ex. Distance Between Two Points
Ex. Fractal Dimension Calc
Massive Fractal Dimension Calc

IBGE
Associate Regions
Search Regions

CRS (RIO GRANDE DO SUL)

GEOCODING
Ex. Address Geocoding
Massive Address Geocoding

NUMBER OF CASES COUNTER

Search Regions

Filter a city from a specific Federation Unit (UF), Brazil, and view its respective regions.

Select the UF
Rio Grande do Sul

You can also type to filter other options

497 municipalities were recovered from Rio Grande do Sul.

City
Viamão

You can also type to filter other options

```
[
  {
    "id": 4300034,
    "nome": "Açegua",
    "microrregiao": {
      "id": 43031,
      "nome": "Campanha Meridional",
      "mesorregiao": {
        "id": 4306,
        "nome": "Sudoeste Rio-grandense",
        "UF": {
          "id": 43,
          "sigla": "RS",
          "nome": "Rio Grande do Sul",
          "regiao": {
            "id": 4,
            "sigla": "S",
            "nome": "Sul"
          }
        }
      }
    }
  },
  "regiao-imediata": {
    "id": 43010,
    "nome": "Bagé",
    "regiao-intermediaria": {
      "id": 4302,
      "nome": "Pelotas",
      "UF": {
        "id": 43,
        "sigla": "RS",
        "nome": "Rio Grande do Sul",
        "regiao": {
          "id": 4,
          "sigla": "S",
          "nome": "Sul"
        }
      }
    }
  }
}
]
```

CRS (RIO GRANDE DO SUL)

Nesta interface é possível associar as Coordenadorias Regionais de Saúde do estado do Rio Grande do Sul, Brasil, às cidades do seu conjunto de dados. Para realizar a associação é necessário importar um arquivo de dados no formato .xlsx e selecionar a coluna que representa a cidade dos seus registros. Clicando no botão “submit” é feito o download de um arquivo .xlsx com a associação realizada.

The screenshot displays the 'Spatial Workspace' application interface. The left sidebar contains navigation options: 'HOW TO USE', 'MAP', 'FRACTAL' (with examples like 'Distance Between Two Points'), 'IBGE' (with 'Associate Regions'), 'CRS (RIO GRANDE DO SUL)', 'GEOCODING', and 'NUMBER OF CASES COUNTER'. The main content area is titled 'Associates CRS with City' and includes instructions: 'Import your data sheet (.xlsx) and select the "city" column to associate your respective Coordenadoria Regional de Saúde of Rio Grande do Sul.' It states that '497 municipalities were recovered from Rio Grande do Sul with their Coordenadorias Regionais de Saúde.' Below this is an 'Attached file' section with a 'Click here to import a file' button and a 'SUBMIT' button. On the right, a JSON object is displayed, listing various municipalities in Rio Grande do Sul, such as 'Porto Alegre', 'Araticá', 'Barão', 'Brochier', 'Cambará do Sul', 'Campo Bom', 'Canoas', 'Capela de Santana', 'Dois Irmãos', 'Estância Velha', 'Esteio', 'Harmonia', 'Igrejinha', 'Ivoti', 'Lindolfo Collor', 'Maratá', 'Montenegro', 'Morro Reuter', 'Nova Hartz', 'Nova Santa Rita', 'Novo Hamburgo', 'Pareci Novo', 'Parobé', 'Portão', 'Presidente Lucena', 'Riozinho', 'Rolante', 'Salvador do Sul', 'Santa Maria do Herval', 'São Francisco de Paula', 'São José do Hortêncio', 'São José do Sul', 'São Leopoldo', 'São Pedro da Serra', 'São Sebastião do Cai', and 'Sapiranga'.

GEOCODING

Ex. Address Geocoding

Esta interface exemplifica o georreferenciamento de um registro de endereço. Para realizar o georreferenciamento preciso é necessário o endereço completo de uma localidade. Ao digitar o endereço e clicar no botão "submit", é enviada uma requisição ao serviço de Geocodificação do Google Maps API. Os dados retornados da API são tratados localmente e apresentados aos usuários de forma organizada na seguinte estrutura: endereço formatado, latitude, longitude, nome da rua, número da casa, bairro, cidade, estado e cep.

Mais informações sobre a API de Geocodificação do Google Maps pode ser acessada em: [clique aqui](#).

The screenshot displays the 'Spatial Workspace' web application interface. On the left is a navigation sidebar with categories: 'HOW TO USE', 'MAP', 'FRACTAL' (with sub-items: 'Ex. Distance Between Two Points', 'Ex. Fractal Dimension Calc', 'Massive Fractal Dimension Calc'), 'IBGE' (with sub-items: 'Associate Regions', 'Search Regions'), 'CRS (RIO GRANDE DO SUL)', 'GEOCODING' (with sub-items: 'Ex. Address Geocoding', 'Massive Address Geocoding'), and 'NUMBER OF CASES COUNTER'. The main content area is titled 'Example of Address Geocoding' and contains the instruction: 'Enter your full address and click the submit button to view an example of the geocoding process.' Below this is a text input field with the address 'Rua Monteiro Lobato, 420, Centro, Bom Conselho, Pernambuco' and a 'SUBMIT' button. To the right of the input field, a JSON object displays the geocoding results: { "formatted_address": "R. Monteiro Lobato, 420, Bom Conselho - PE, 55330-000, Brasil", "latitude": -9.1650242, "longitude": -36.67645, "street_name": "Rua Monteiro Lobato", "street_number": "420", "neighborhood": "", "city": "Bom Conselho", "state": "Pernambuco", "zipcode": "55330-000" }

Massive Address Geocoding

Nesta interface é possível realizar o georreferenciamento em massa de registros de endereço. Para realizar o georreferenciamento é necessário importar um arquivo de dados no formato .xlsx, selecionar a coluna que representa o endereço dos registros e selecionar a coluna que identifica os registros (esta coluna tem por finalidade a organização dos registros). Ao clicar no botão "submit", o georreferenciamento dos registros é iniciado. Após o término, automaticamente será realizado o download de um arquivo .xlsx com os registros georreferenciados. No arquivo .xlsx haverá 2 folhas, uma com os registros que possivelmente obtiveram falha no georreferenciamento, devido a falta de informação ou erros ortográficos, e outra folha com os registros que obtiveram sucesso.

Spatial Workspace
v3.5.2 About

HOW TO USE

MAP

FRACTAL
Ex. Distance Between Two Points
Ex. Fractal Dimension Calc
Massive Fractal Dimension Calc

IBGE
Associate Regions
Search Regions

CRS (RIO GRANDE DO SUL)

GEOCODING
Ex. Address Geocoding
Massive Address Geocoding

NUMBER OF CASES COUNTER

Bulk Geocoding of Address

Import your data sheet (.xlsx), select the address and identification columns to geocod the data set.

Attached file
[Click here to import a file](#)

36 data were recovered.

Address
address

Column that represents the address

Identifier
id

Column that represents the record identifier

SUBMIT

address
Rua Botucaraí, 1840 - Bairro Centro - Candelária, RS/96930-000
Rua Santo Antônio, 75 - Bairro Operario - Lagoa Vermelha, RS/95300-000
Cerro do Louro, SN - Bairro Interior - Formigueiro, RS/97.210-000
Rua Anibal Xavier Bitencourt, 262 - Bairro Pinheiro - Caçapava do Sul, RS/96570-000
Rua Helmut Klain, 41 - Bairro Planalto - Morro Reuter, RS/93990-000
Rua São Roque 431 - Bairro Moinhos - Estrela, RS/95880-000
Rua José Montaury 963 - Bairro Centro - Veranópolis, RS/95330-000
Avenida São João, 792 - Bairro Centro - São João do Polésine, RS/97230-000
Rua Constantino, SN - Bairro Interior - Progresso, RS/95925-000
Rua feliz, 58 - Bairro Industrial - Farroupilha, RS/95180-000
Rua Francisco Borges Vieira, 33 - Bairro Santa Teresinha - Vacaria, RS/95200-000
Rua Quarenta e sete, SN - Bairro Jardim Algarve - Alvorada, RS/94858-670
BR 116, 410 - Bairro Picada - Eldorado do Sul, RS/92990-000
Rua Paulo de Castro, 830 - Bairro Alesgut - Teutônia, RS/95890-000
Rua Corredor do João, 53 - Bairro Faxinal - Santa Cruz do Sul, RS/96840-540

NUMBER OF CASES COUNTER

Nesta interface é possível contar automaticamente o número de ocorrências de um texto ou valor numérico num conjunto de dados. Para realizar a contagem é necessário importar um arquivo de dados no formato .xlsx e selecionar a variável/coluna de interesse para contagem. Os dados de contagem serão apresentados na tabela ao lado. Se necessário, também é possível realizar uma filtragem sobre o conjunto de dados. Se clicado em "submit", é realizado o download dos dados apresentados na tabela em um arquivo .xlsx.

The screenshot shows the 'Number of Cases Counter' interface. The sidebar on the left contains navigation options: 'HOW TO USE', 'MAP', 'FRACTAL' (with examples like 'Distance Between Two Points'), 'IBGE' (with 'Associate Regions'), 'CRS (RIO GRANDE DO SUL)', 'GEOCODING' (with 'Address Geocoding'), and 'NUMBER OF CASES COUNTER' (highlighted in blue). The main panel has a title 'Number of Cases Counter' and instructions: 'Import your data sheet (.xlsx) to automatically count the total number of records for a specific region.' It shows 'Attached file' and '405 data were recovered.' Below this, there are filter dropdowns for 'Region' (set to 'nm_meso'), 'id', 'disease', 'city', 'cd_mun', 'cd_micro', 'nm_micro', and 'cd_meso'. A note states: 'This column will be used to count the number of cases per area'. On the right, a table displays the results:

nm_meso	cases
Centro Oriental Rio-grandense	31
Nordeste Rio-grandense	34
Centro Ocidental Rio-grandense	14
Sudeste Rio-grandense	45
Metropolitana de Porto Alegre	215
Noroeste Rio-grandense	42
Sudoeste Rio-grandense	24

Anexos

ANEXO A – Parecer Consubstanciado do Comitê de Ética em Pesquisa (CEP)



HOSPITAL MATERNO
INFANTIL PRESIDENTE
VARGAS - HMIPV - RS



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Tecnologias integradas para instrumentalizar e humanizar o cuidado em saúde de crianças diagnosticadas em um programa de triagem neonatal no sul do Brasil

Pesquisador: CRISTIANE KOPACEK

Área Temática:

Versão: 3

CAAE: 36380020.8.0000.5329

Instituição Proponente: Hospital Materno Infantil Presidente Vargas - HMIPV - RS

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 4.397.969

Apresentação do Projeto:

Projeto de Pesquisa do Serviço de Referência em Triagem Neonatal (SRTN) do Hospital Materno Infantil Presidente Vargas com as instituições co-participantes UNIVERSIDADE FEDERAL DE CIÊNCIAS DA SAÚDE DE PORTO ALEGRE, UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO e UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL. A triagem neonatal, compreendida pelo Serviço de Referência em Triagem Neonatal (SRTN) é o serviço responsável pela análise e diagnóstico de Recém Nascidos (RN) portadores das seis doenças triadas pelo teste do pezinho, nomeadamente: fibrose cística, hipotireoidismo congênito, deficiência de biotinidase, hemoglobinopatias, fenilcetonúria e hiperplasia adrenal congênita. O tratamento e acompanhamento adequado destes pacientes depende do acesso de informações por diferentes núcleos profissionais em diferentes unidades de saúde. Dessa forma, a coleta, armazenamento e análise de dados dos processos envolvidos se torna fundamental para garantir uma maior efetividade e implementação, proporcionando o fortalecimento da gestão pública, otimização de recursos, comunicação, controle social e organização de banco de dados com segurança para a realização de investigações científicas na área e humanização dos atendimentos. Essa pesquisa pretende criar uma Plataforma Informacional, armazenada em nuvem, para Criação de Aplicativos para Aquisição, Processamento

Endereço: Av. Independência 661- Bl. C 7º andar

Bairro: Independência

CEP: 90.035-076

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3289-3348

E-mail: hmipv.cep@hmipv.prefpoa.com.br



HOSPITAL MATERNO INFANTIL PRESIDENTE VARGAS - HMIPV - RS



Continuação do Parecer: 4.397.969

e Análises de Dados. Além disso, tal Plataforma será integrada a Ambientes Virtuais de Rede Social com Foco coerentes com a Política Nacional de Humanização – Humaniza SUS (PNH). Mediante a utilização de Tecnologias inovadoras que organizem e integrem dados quantitativos e qualitativos do SRTN, espera-se instrumentalizar uma linha de cuidado mais eficiente, otimizada e humanizada para o Programa de Triagem Neonatal.

Objetivo da Pesquisa:

Objetivo Geral:

Criar e implementar uma Plataforma Informacional com tecnologias integradas para coleta georreferenciada, armazenamento e análises de dados de usuários e equipes de saúde envolvidos na triagem neonatal do estado do RS.

Objetivos Específicos:

1. Criação de aplicativos para coleta georreferenciada, armazenamento de banco de dados em servidor próprio em "Nuvem" de dados clínicos georreferenciados por profissionais da saúde envolvidos na PNTN/MS no estado do RS.
2. Criação de aplicativos de Ambientes Virtuais de Rede Social com Foco (Banco de Vidas) para adequada fonte de comunicação e informação em Saúde para profissionais da saúde, usuários e familiares envolvidos na PNTN/MS no estado do RS.
3. Criação de Interface de Gerenciamento e Análises de Dados quantitativos (Banco de Dados) e qualitativos (Banco de Vidas) da PNTN/MS no estado do RS.
4. Estimar áreas de risco para ocorrência de doenças triadas pelo SRTN-RS, usando técnicas de Geoprocessamento e análise espacial.
5. Realizar investigação qualitativa a partir das narrativas autobiográficas de cuidadores de usuários diagnosticados pelo SRTN-RS sob uma perspectiva metodológica transdisciplinar, fundamentada na reflexão teórico-crítica e sensibilização intuitivo-reflexiva.
6. Compreender os processos sócio-culturais, emocionais e educacionais envolvidos no diagnóstico e adesão aos tratamentos indicados.
7. Apresentar importantes dados a serem aproveitados em processos de formação continuada dos profissionais da saúde que atuam na área.
8. Formar recursos humanos na área de saúde e biologia molecular, abrangendo alunos de iniciação científica (alunos de graduação em áreas correlatas), especialização, mestrado e doutorado, auxiliando na capacitação humana e técnico-científica.

Endereço: Av. Independência 661- Bl. C 7º andar

Bairro: Independência

CEP: 90.035-076

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3289-3348

E-mail: hmipv.cep@hmipv.prefpoa.com.br



Continuação do Parecer: 4.397.969

9. Divulgar os resultados obtidos em Congressos, Reuniões Científicas e em Periódicos Indexados.

Avaliação dos Riscos e Benefícios:

Riscos:

Segundo a pesquisadora, a pesquisa se enquadra na modalidade de risco mínimo de acordo com a Resolução CNS 466/12. Por se tratar de uma pesquisa de base populacional, após aprovação do CEP, a coleta de dados se dará através de consulta às informações nos bancos de dados, não sendo necessária a aplicação de Termo de Consentimento Livre e Esclarecido (TCLE) nesta etapa. Para reduzir o risco de identificação individual dos participantes, os dados serão criptografados e guardados em servidor seguro e a divulgação dos resultados será feita de forma a não identificar os voluntários. Os dados coletados nesta pesquisa, ficarão armazenados em servidor, sob a responsabilidade do pesquisador principal, pelo período de mínimo 5 (cinco) anos.

No seu âmbito qualitativo, este estudo poderá apresentar algum pequeno desconforto ao poder, em algum momento, constranger o participante. Todavia, o mesmo poderá abandonar o estudo e/ou pedir para retirar sua narrativa autobiográfica da plataforma a qualquer momento. Para a análise qualitativa, todos os envolvidos assinarão Termo de Consentimento Livre e Esclarecido, contido em Projeto de Pesquisa. O estudo poderá ainda trazer risco em caso de perda ou extravio dos dados dos relatórios coletados. No entanto, para minimizar tais riscos, os dados serão mantidos em sigilo absoluto sendo utilizados para sua coleta e análise, armazenado em nuvem, sob responsabilidade da coordenação desta pesquisa, não gerando prejuízos para a atuação profissional ou para o usuário ao serviço.

Benefícios:

Os resultados da pesquisa ajudarão muito a qualificar a coleta, armazenamento e análise de dados que envolvem as condições de saúde dos usuários da PNTN do RS e a informar qualitativamente, acolhendo os usuários e seus familiares diagnosticados com alguma patologia de maneira mais humanizada com fontes de informações mais fidedignas em ambiência virtual adequada. Além disso, a criação da Plataforma Informacional oferecerá acesso seguro às informações ali depositadas, proporcionando a criação de redes de conexões com o intuito de promover a integração e a aproximação do usuário, das equipes de saúde e dos serviços especializados através de um acesso rápido e personalizado, de ações educacionais, de apoio e monitoramento, podendo desenvolver uma atenção primária mais forte que ofereça um cuidado local ágil, integral, efetivo e humanizado aos pacientes, mesmo que distantes dos centros de

Endereço: Av. Independência 661- Bl. C 7º andar

Bairro: Independência

CEP: 90.035-076

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3289-3348

E-mail: hmipv.cep@hmipv.prefpoa.com.br



HOSPITAL MATERNO INFANTIL PRESIDENTE VARGAS - HMIPV - RS



Continuação do Parecer: 4.397.969

referência.

Comentários e Considerações sobre a Pesquisa:

O tema da pesquisa é relevante, pois a criação de uma Plataforma Informacional para o SRTN pode instrumentalizar uma linha de cuidado mais eficiente, otimizada e humanizada. Além disso, possibilitará a geração de relatórios que auxiliem a tomada de decisão dos gestores para melhorias na assistência aos pacientes, assim como a realização de futuras pesquisas com base nos dados armazenados na plataforma.

Considerações sobre os Termos de apresentação obrigatória:

Todos os termos foram incluídos adequadamente.

Conclusões ou Pendências e Lista de Inadequações:

Projeto aprovado

Considerações Finais a critério do CEP:

1. Informamos que toda e qualquer alteração do projeto deverá ser comunicada imediatamente ao CEP HMIPV.
2. O pesquisador deverá apresentar relatórios semestrais de acompanhamento do projeto, bem como relatório final quando do término do mesmo.
3. Para o ingresso nas dependências do hospital o pesquisador responsável deverá solicitar ao CEP HMIPV a confecção de crachá para toda a equipe de pesquisa.
4. Para o início da pesquisa, o investigador deverá apresentar à chefia do serviço onde será realizada a pesquisa o Parecer Consubstanciado de aprovação do protocolo pelo CEP.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1237199.pdf	29/10/2020 16:00:50		Aceito
Outros	autorizacaodirecaohmipv.pdf	29/10/2020 15:59:20	Simone Martins de Castro	Aceito
Outros	relatorios.pdf	29/10/2020 15:57:37	Simone Martins de Castro	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	ANEXOII_TCLE_profissionalv2.pdf	29/10/2020 15:55:52	Simone Martins de Castro	Aceito
TCLE / Termos de Assentimento /	ANEXO_I_TCLE_responsavel_corrigido.pdf	29/10/2020 15:55:34	Simone Martins de Castro	Aceito

Endereço: Av. Independência 661- Bl. C 7º andar

Bairro: Independência

CEP: 90.035-076

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3289-3348

E-mail: hmipv.cep@hmipv.prefpoa.com.br



HOSPITAL MATERNO
INFANTIL PRESIDENTE
VARGAS - HMIPV - RS



Continuação do Parecer: 4.397.969

Justificativa de Ausência	ANEXO1_TCLEresponsavel_corrigido.pdf	29/10/2020 15:55:34	Simone Martins de Castro	Aceito
Declaração de Pesquisadores	termo_de_compromisso_para_entrega_relatorio.doc	15/10/2020 22:53:27	Simone Martins de Castro	Aceito
Projeto Detalhado / Brochura Investigador	Projeto_Georeferenciamento_final_v2cep.pdf	12/10/2020 17:18:30	Simone Martins de Castro	Aceito
Declaração de Instituição e Infraestrutura	termochefprontu.pdf	05/08/2020 22:59:09	Simone Martins de Castro	Aceito
Declaração de Pesquisadores	termopesquisa.jpg	05/08/2020 22:54:48	Simone Martins de Castro	Aceito
Folha de Rosto	folharosto.pdf	05/08/2020 22:52:50	Simone Martins de Castro	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

PORTO ALEGRE, 13 de Novembro de 2020

Assinado por:
Angela Ester Ruschel
(Coordenador(a))

Endereço: Av.Independência 661- Bl. C 7º andar

Bairro: Independência

CEP: 90.035-076

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3289-3348

E-mail: hmipv.cep@hmipv.prefpoa.com.br

ANEXO B – Certificado de Registro de Programa de Computador



REPÚBLICA FEDERATIVA DO BRASIL
MINISTÉRIO DA ECONOMIA
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
DIRETORIA DE PATENTES, PROGRAMAS DE COMPUTADOR E TOPOGRAFIAS DE CIRCUITOS INTEGRADOS

Certificado de Registro de Programa de Computador

Processo Nº: **BR512021003233-0**

O Instituto Nacional da Propriedade Industrial expede o presente certificado de registro de programa de computador, válido por 50 anos a partir de 1º de janeiro subsequente à data de 05/10/2021, em conformidade com o §2º, art. 2º da Lei 9.609, de 19 de Fevereiro de 1998.

Título: Sistema web para processamento e análises espaciais

Data de criação: 05/10/2021

Titular(es): UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO

Autor(es): AUGUSTO CÉSAR FERREIRA DE MIRANDA OLIVEIRA; MOACYR CUNHA FILHO; GUILHERME VILAR

Linguagem: JAVA SCRIPT

Campo de aplicação: AG-01; IF-02; IF-10; SD-01

Tipo de programa: FA-01; FA-03; GI-01; GI-06; SO-02

Algoritmo hash: SHA-256

Resumo digital hash: 170e4a887952d616c7c024c5937eb1a7c14276bcc1cac08aa845554f79e40f49

Expedido em: 04/01/2022



Aprovado por:

Joelson Gomes Pequeno

Chefe Substituto da DIPTO - PORTARIA/INPI/DIRPA Nº 02, DE 10 DE FEVEREIRO DE 2021