

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
Escola de Engenharia  
Programa de Pós-Graduação em Engenharia Elétrica

Lucas Lima Carneiro

**Metodologia sistemática para avaliar impactos de desastres e de epidemias  
na população atingida**

Belo Horizonte  
2022

Lucas Lima Carneiro

**Metodologia sistemática para avaliar impactos de desastres e de epidemias na população atingida**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Prof. Dr. Walmir Matos Caminhas

Coorientador: Prof. Dr. Ed Wilson Rodrigues Vieira

Belo Horizonte

2022

C289m	<p>Carneiro, Lucas Lima.  Modelagem sistemática para avaliar impactos de desastres e de epidemias na população atingida [recurso eletrônico] / Lucas Lima Carneiro. - 2022.  1 recurso online (66 f. : il., color.) : pdf.</p> <p>Orientador: Walmir Matos Caminhas.  Coorientador: Ed Wilson Rodrigues Vieira.</p> <p>Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Engenharia.</p> <p>Bibliografia: f. 61-66.  Exigências do sistema: Adobe Acrobat Reader.</p> <p>1. Engenharia elétrica - Teses. 2. Aprendizado do computador - Teses.  3. Catástrofes naturais - Teses. 4. Estatística - Análise - Teses.  I. Caminhas, Walmir Matos. II. Vieira, Ed Wilson Rodrigues.  III. Universidade Federal de Minas Gerais. Escola de Engenharia.  IV Título.</p> <p style="text-align: right;">CDU: 621.3(043)</p>
-------	--



UNIVERSIDADE FEDERAL DE MINAS GERAIS  
ESCOLA DE ENGENHARIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

## FOLHA DE APROVAÇÃO

### "METODOLOGIA SISTEMÁTICA PARA AVALIAR IMPACTOS DE DESASTRES E DE EPIDEMIAS NA POPULAÇÃO ATINGIDA"

**LUCAS LIMA CARNEIRO**

Dissertação de Mestrado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do grau de Mestre em Engenharia Elétrica. Aprovada em 20 de setembro de 2022. Por:

Prof. Dr. Walmir Matos Caminhas  
DELT (UFMG)

Prof. Dr. Ed Wilson Rodrigues Vieira  
Departamento de Enfermagem Materno-Infantil e Saúde Pública

Prof. Dr. Frederico Gadelha Guimarães  
DEE (UFMG)

Prof. Dr. Jorge Gustavo Velásquez Meléndez  
Enfermagem materno infantil e saúde pública (UFMG)

Prof. Dr. Mário Neto Borges  
DELE (UFSJ)



Documento assinado eletronicamente por **Walmir Matos Caminhas, Professor do Magistério Superior**, em 20/09/2022, às 11:26, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Frederico Gadelha Guimaraes, Professor do Magistério Superior**, em 20/09/2022, às 14:49, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Mario Neto Borges, Usuário Externo**, em 20/09/2022, às 18:34, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

Documento assinado eletronicamente por **Ed Wilson Rodrigues Vieira, Professor do Magistério**



**Superior**, em 21/09/2022, às 13:57, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Jorge Gustavo Velasquez Melendez, Professor do Magistério Superior**, em 21/09/2022, às 14:54, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



A autenticidade deste documento pode ser conferida no site [https://sei.ufmg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **1767105** e o código CRC **EAA6A416**.

---

À meu pai Fábio (*in memoriam*) que, do seu jeito particular, torcia por mim.  
À meu amigo Rafael (*in memoriam*) por todo seu carinho, brincadeiras, tempo de jogatina e principalmente pela importância que ele dava à minha educação.

## AGRADECIMENTOS

Agradeço, de coração, ao meu orientador Walmir Caminhas por ter me aceitado como orientando e por todo apoio durante minha trajetória na pós-graduação; por todas as orientações, todos os projetos e oportunidades; por ser um braço amigo, alguém com quem eu posso contar. Sem ele, este trabalho não existiria.

Não tenho palavras para agradecer a Érica Freire, minha namorada, que veio comigo para Belo Horizonte, nessa aventura incerta, por todo o amor e dedicação à nossa vida juntos, por sempre estar comigo nos momentos difíceis, por me erguer nos momentos de fraqueza, por trazer à tona o melhor de mim.

Sou eternamente grato a minha mãe Dilma por “ter me deixado voar”, apesar de todas as dificuldades e tudo que passamos juntos; por ter suportado toda a dor para que eu estivesse onde estou hoje. Agradeço também aos meus irmãos Kauan e Laura por sempre me manterem feliz com suas birras e brincadeiras, e ao meu padastro Alessandro por cuidar dessas pessoas que amo tanto.

Agradeço aos meus amigos Victor e Raquel por todas as conversas, todo o apoio e todo o carinho; a Rafael por todas as resenhas; a Pedro por me ajudar nesse mundo acadêmico desde a graduação com suas orientações; e em especial a Thiago e Sarah por, além de tudo, terem compartilhado moradia comigo e com Érica durante os primeiros anos nessa nova cidade. Agradeço também ao professor Thiago Chagas, da UESC, por ter me guiado no caminho da pesquisa e dado o seu melhor para que eu crescesse como aluno.

Agradeço a todos os envolvidos no Projeto Brumadinho UFMG, em especial aos professores Ed Wilson Vieira e Gustavo Velasquez-Melendez por todos os ensinamentos, parcerias e por terem aceitado participar da banca deste trabalho. Agradeço também aos demais professores que aceitaram participar da banca, Mário Neto Borges e Frederico Guimarães, por seu tempo e paciência.

Por fim, agradeço a Deus por ter colocado todas essas pessoas maravilhosas em minha vida.

*“It’s not about changing the world. It’s about doing our best to leave the world...  
The way it is”  
(Big Boss)*



## RESUMO

Esta dissertação propõe uma metodologia para a análise do impacto de eventos, como desastres naturais, tecnológicos e epidemias, em um dado conjunto de localidades. O método proposto utiliza técnicas de aprendizado de máquina e ferramentas estatísticas para investigar os efeitos do evento, em uma ou mais variáveis. Propõe-se realizar três diferentes análises, onde duas são univariáveis e uma é multivariável. O Gráfico de Resultantes Vetoriais representa uma nova técnica para visualizar resultados de diversos diagramas de controle em uma única figura e permite comparar os dados após a ocorrência do evento com seus limites históricos; a comparação estatística, por testes  $t$  pareados, permite comparar a mudança de comportamento das localidades atingidas com um conjunto de localidades de controle; já a análise multivariável, por meio do algoritmo de agrupamento *Fuzzy c-means*, observa a mudança de comportamento das localidades atingidas ao investigar as relações entre as variáveis estudadas. Ademais, a metodologia apresenta uma forma de utilizar o algoritmo *Fuzzy c-means* para determinar as localidades de controle. Por fim, o método proposto é aplicado em dois estudos de caso: “Rompimento da barragem de Fundão em Mariana–MG” e “Impacto da pandemia da COVID-19 nos atendimentos para acompanhamento do crescimento e desenvolvimento de crianças no Brasil”.

**Palavras-chave:** Aprendizado de máquina; Gestão de desastres; Análise estatística.

## ABSTRACT

This work aims at developing a new methodology for assessing the impact of an event, such as epidemics and natural disasters, in a given set of places. The proposed method uses machine learning techniques and statistical tools to investigate effects of the event, in one or more variables. Three distinct analyses are performed where two are monivariate and one is multivariate. The Resultant Vectors Graph presents a new technique to visualize results from several control charts in only one diagram and it also allows comparison between data after the event had occurred and its historical limits; the statistical comparison, through paired  $t$  tests, allows to compare the change in behavior between possibly affected locations and its control locations; in turn, the multivariate analysis, through Fuzzy c-means clustering algorithm, observes the behavior change in possibly affected locations while investigating the relationship between studied variables. In addition, the methodology presents a way of using the Fuzzy c-means algorithm to determine the set of control locations. Lastly, the proposed methodology is tested in two case studies: the “Fundão dam failure in Mariana – MG” and the “Impact of COVID-19 pandemics in visits for monitoring child growth and development in Brazil”.

**Keywords:** Disaster management; Machine learning; Statistical analysis.

## LISTA DE FIGURAS

Figura 3.1 – Ilustração da região crítica de uma distribuição $t$ com $n = 5$ graus de liberdade e nível de significância $\alpha = 0,05$ . . . . .	25
Figura 4.1 – Ilustração de um diagrama de controle . . . . .	31
Figura 4.2 – Ilustração da técnica de Gráficos de Resultantes Vetoriais utilizada para sintetizar as informações dos diagramas de controles, indicando os vetores unitários atribuídos ao gráfico: 4.2a acima do limite superior esperado; 4.2b abaixo do limite inferior esperado; 4.2c dentro dos limites esperados; e 4.2d vetor resultante. . . . .	32
Figura 4.3 – Campos de interesse no código JavaScript destacados para fácil visualização. . . . .	37
Figura 4.4 – Fluxograma dos passos para implementação da metodologia proposta. . . . .	38
Figura 5.1 – Gráfico de resultantes vetoriais – Municípios com média anual acima do esperado. . . . .	43
Figura 5.2 – Gráfico de resultantes vetoriais – Demais municípios estudados. . . . .	44
Figura 5.3 – Taxas mensais de atendimentos (por 1.000 crianças menores de 5 anos de idade) para acompanhamento do crescimento e desenvolvimento de crianças nos serviços de Atenção Primária à Saúde no Brasil, segundo Regiões. . . . .	54
Figura 5.4 – Taxas mensais de atendimentos (por 1.000 crianças menores de 5 anos de idade) para acompanhamento do crescimento e desenvolvimento de crianças nos serviços de Atenção Primária à Saúde no Brasil, segundo estados e Distrito Federal no Brasil. . . . .	55
Figura 5.5 – Gráfico de resultantes vetoriais das taxas de atendimentos (por 1.000 crianças menores de 5 anos de idade) para acompanhamento do crescimento e desenvolvimento de crianças nos serviços de Atenção Primária à Saúde, segundo estados e Distrito Federal no Brasil. . . . .	56

## LISTA DE TABELAS

Tabela 5.1 – Localidades mais similares às atingidas, ao considerar taxas de atendimento anual por dengue e IDHm, obtidas por meio de <i>clusterização fuzzy</i> – são exibidas até as segundas localidades mais similares para simplificação da tabela. . . . .	40
Tabela 5.2 – Número de atendimentos observados antes e após o evento para os municípios atingidos e seus controles. . . . .	41
Tabela 5.3 – Resultante do teste <i>t</i> pareado para os municípios atingidos e seus respectivos controles. . . . .	44
Tabela 5.4 – Pertinências <i>fuzzy</i> antes e após a ocorrência do evento, considerando os atributos Taxa de atendimento por dengue e PIB municipal. . . . .	47
Tabela 5.5 – Significado linguístico dos grupos obtidos na análise multivariável. . . . .	49
Tabela 5.6 – Distribuição do número de atendimentos para acompanhamento do crescimento e do desenvolvimento estudados (x 1.000) nos períodos pré (abril de 2017 a março de 2020) e durante a pandemia (abril de 2020 a março de 2021), segundo Regiões, Estados e Distrito Federal. . . . .	52
Tabela 5.7 – Diferenças entre as médias das taxas de atendimentos para acompanhamento do crescimento e do desenvolvimento de crianças menores de 5 anos de idade (por mil habitantes) nos períodos pré (abril de 2017 a março de 2020) e durante a pandemia da COVID-19 (abril de 2020 a março de 2021), segundo Regiões, Estados e Distrito Federal. . . . .	53

## LISTA DE ABREVIATURAS

<b>FCM</b>	<i>Fuzzy c-means</i>
<b>GRV</b>	Gráfico de Resultantes Vetoriais
<b>PIB</b>	Produto Interno Bruto
<b>IBGE</b>	Instituto Brasileiro de Geografia e Estatística
<b>IDHm</b>	Índice de Desenvolvimento Humano Municipal
<b>ACD</b>	Acompanhamento do Crescimento e Desenvolvimento
<b>APS</b>	Atenção Primária à Saúde
<b>SISAB</b>	Sistema de Informação em Saúde para a Atenção Básica

## LISTA DE SÍMBOLOS

$C_T$	Conjunto total de localidades observadas
$C_{LA}$	Conjunto de localidades atingidas
$C_{LC}$	Conjunto de localidades de controle
$C_A$	Conjunto de atributos possivelmente afetados pelo evento
$\mu_{ij}$	Pertinência da amostra $j$ ao grupo $i$
$m$	Parâmetro de <i>fuzzyficação</i> do <i>Fuzzy c-means</i>
$k_i$	Valor de centro do grupo $i$
$c$	Número de grupos
$r_a$	Raio de vizinhança do centro de <i>cluster</i> no Agrupamento Sub-trativo

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>16</b>
1.1	Objetivos	17
1.1.1	Objetivo Geral	17
1.1.2	Objetivos Específicos	17
1.2	Organização do Texto	18
<b>2</b>	<b>FORMULAÇÃO DO PROBLEMA DE PESQUISA</b>	<b>19</b>
2.1	Formulação 1	19
2.2	Formulação 2	20
2.3	Conclusão do Capítulo	21
<b>3</b>	<b>REVISÃO TEÓRICA</b>	<b>22</b>
3.1	Análise Estatística	22
3.1.1	Diagramas de Controle	22
3.1.2	Teste $t$ Pareado	23
3.2	Análise por Inteligência Computacional	25
3.2.1	Agrupamento ( <i>Clusterização</i> )	25
3.2.2	Agrupamento Nebuloso ( <i>Clusterização Fuzzy</i> )	26
3.2.2.1	O Algoritmo <i>Fuzzy c-means</i>	27
3.2.3	Agrupamento Subtrativo	29
3.3	Conclusão do Capítulo	30
<b>4</b>	<b>METODOLOGIA PROPOSTA</b>	<b>31</b>
4.1	Diagramas de Controle	31
4.2	Gráfico de Resultantes Vetoriais	32
4.3	Formulação 1	33
4.3.1	Definição das Localidades de Controle	33
4.3.2	Análise Univariável	33
4.3.3	Análise Multivariável	34
4.4	Formulação 2	35
4.5	Extração e Tratamento de Dados	36
4.5.1	Extração de Dados Automatizada	36
4.6	Conclusão do Capítulo	37
<b>5</b>	<b>ESTUDOS DE CASO</b>	<b>39</b>
5.1	Formulação 1 – Impacto do rompimento da barragem de Fundão, em Mariana–MG, nos casos de Dengue e no PIB municipal	39
5.1.1	Definição dos Controles	40
5.1.2	Análise Univariável	41
5.1.3	Análise Multivariável	46

5.1.4	Conclusões . . . . .	49
5.2	Formulação 2 – Impacto da pandemia da COVID-19 nos atendimentos para acompanhamento do crescimento e desenvolvimento de crianças no Brasil .	50
5.2.1	Tratamento de Dados . . . . .	51
5.2.2	Análises e Discussão . . . . .	51
5.2.3	Conclusões . . . . .	58
<b>6</b>	<b>PRINCIPAIS CONCLUSÕES . . . . .</b>	<b>59</b>
6.1	Trabalhos Futuros . . . . .	60
	<b>BIBLIOGRAFIA . . . . .</b>	<b>61</b>



## 1 INTRODUÇÃO

Eventos catastróficos como epidemias, desastres naturais e tecnológicos ocorrem frequentemente em todo o mundo, apenas na última década foram registrados mundialmente mais de 5,7 mil desastres [1]. Quando eventos dessa magnitude ocorrem, diferentes formas de assistência são providenciadas. Uma resposta imediata serve para controlar os danos e minimizar as perdas, porém, após o evento, um tipo de assistência igualmente importante é a avaliação de impactos para amparar os moradores das localidades atingidas.

Os impactos podem ser de diferentes naturezas, sejam impactos ambientais, na saúde ou até mesmo econômicos. Impactos ambientais são modificações nas propriedades físicas, químicas ou biológicas do meio ambiente, provocadas direta ou indiretamente por ação humana, que afetam a segurança da população, as atividades socioeconômicas, a biota, as condições estéticas e sanitárias do meio ambiente ou a qualidade dos recursos ambientais; alguns exemplos são os diversos tipos de poluição, desmatamentos e redução do número de animais silvestres [2]. Impactos na saúde humana estão associados a ações que podem propiciar o surgimento e proliferação de doenças, assim como às próprias enfermidades. Já impactos socioeconômicos estão relacionados a alterações na distribuição de renda e indicadores de desenvolvimento, por exemplo. Para investigar esses impactos, diferentes metodologias podem ser utilizadas.

Como exemplo de estudos de impacto de desastres [Becquart \*et al.\* \[3\]](#) podem ser citados. Eles estudaram as hospitalizações diárias de idosos por problemas cardiovasculares antes, durante e após a passagem do furacão Katrina. Para isso, regressão segmentada foi empregada para suavizar os dados da série temporal e identificar se para cada período houve tendência de aumento ou diminuição da taxa de hospitalizações. A regressão segmentada é uma abordagem supervisionada, isto é, ela utiliza conjuntos de entradas e saídas conhecidas para gerar um modelo capaz de prever saídas para entradas desconhecidas. Assim, a incerteza da predição está associada à qualidade do modelo gerado.

Outra forma de regressão foi utilizada por [Kishi \*et al.\* \[4\]](#), que investigou a qualidade de vida relacionada à saúde bucal da população atingida pelo Grande Terremoto do Leste do Japão. Este estudo utilizou dados coletados por meio de formulários enviados aos sobreviventes do desastre. Os indivíduos que aceitaram participar da pesquisa foram submetidos a uma avaliação de higiene bucal. Esta forma de coleta de dados, apesar de ser planejada especificamente para responder às questões de pesquisa, está sujeita ao aceite de participação pelos voluntários e necessita de um maior suporte financeiro.

Estudos como o realizado por [Nishijima & Rocha \[5\]](#) utilizam bancos de dados secundários, ou seja, dados que já foram coletados e estruturados para alguma outra

finalidade que não a da pesquisa em questão. Esse estudo investigou a relação entre os casos de dengue e o rompimento da barragem de Fundão em Mariana–MG. Para isso, as localidades atingidas foram agrupadas e um grupo de controle foi formado por localidades em um raio de 200 Km. Então, a técnica *Difference-in-Difference* foi empregada para avaliar as diferenças entre os grupos antes e após o evento.

Vista a variedade de caminhos (métodos) que os pesquisadores podem seguir para avaliação de impactos, e as limitações como as encontradas nos estudos supracitados, esta dissertação propõe uma nova metodologia para avaliar impactos de desastres e epidemias a partir do emprego de técnicas de aprendizado de máquina (*machine learning*) em conjunto com análises estatísticas. Dentre as técnicas utilizadas, destaca-se o algoritmo de Agrupamento Nebuloso *Fuzzy c-means*. Esta classe de algoritmos é utilizada em diversos problemas por ser capaz de modelar a incerteza do mundo real, tornando-os mais robustos [6]. Além disso, o emprego do *Fuzzy c-means* em conjunto com o Agrupamento Subtrativo [7] permite que os centros iniciais sejam definidos previamente, o que promove a repetibilidade dos resultados. Além do *machine learning*, o uso de ferramentas estatísticas permite observar se a diferença entre dados de dois períodos distintos é ou não significativa, o que permite comparar os períodos antes e após o evento.

Portanto, a metodologia proposta terá as vantagens de utilizar bancos de dados secundários, possibilitar a repetibilidade dos experimentos, proporcionar maior interpretabilidade dos resultados obtidos por uso de lógica *fuzzy*, além de ser robusta, pois possibilita a análise de um mesmo evento por diferentes “pontos de vista”, ou seja, por diferentes métodos. Após a metodologia ser formulada, ilustra-se sua forma de utilização em dois estudos de caso. Essas aplicações permitem observar as vantagens supracitadas assim como limitações do método.

## 1.1 Objetivos

Esta seção apresenta os objetivos geral e específicos desta dissertação.

### 1.1.1 Objetivo Geral

Desenvolver uma metodologia para avaliar impactos de desastres e epidemias na população atingida.

### 1.1.2 Objetivos Específicos

Os objetivos específicos desta dissertação são:

- a) Estudar o comportamento das localidades observadas antes e após o evento;
- b) Avaliar se a mudança de comportamento de uma localidade após o evento é significativa;

- c) Desenvolver um método robusto para determinação de localidades de controle;
- d) Quando mais de uma variável for estudada, examinar interações entre elas;
- e) Implementar a metodologia desenvolvida.

## 1.2 Organização do Texto

O texto é organizado da seguinte forma:

- **Capítulo 2:** Apresenta a formulação do problema de pesquisa. Define a ocorrência do evento, determinação de localidades observadas, atingidas e de controle em termos matemáticos, e explica brevemente os passos para empregar a metodologia, esta que é dividida em duas formulações.
- **Capítulo 3:** Descreve os principais conceitos teóricos necessários para o entendimento das técnicas usadas para aplicar a metodologia.
- **Capítulo 4:** Explica como as técnicas abordadas no Capítulo 3 serão utilizadas em conjunto com os passos citados no Capítulo 2. Em suma, descreve detalhadamente os passos para aplicar a metodologia.
- **Capítulo 5:** A metodologia é aplicada em dois diferentes estudos de caso. O primeiro estudo de caso aborda a Formulação 1 enquanto o segundo aborda a Formulação 2 da metodologia.
- **Capítulo 6:** Apresenta as principais conclusões do trabalho.

## 2 FORMULAÇÃO DO PROBLEMA DE PESQUISA

Uma localidade atingida é definida como uma localidade que possivelmente foi afetada pelo evento, por exemplo, a localidade onde o evento ocorreu. Em um desastre como o rompimento de uma barragem, as localidades atingidas podem ser obtidas ao analisar o curso da lama, ou seja, quais cidades foram afetadas pelos dejetos. Contudo, em alguns casos como uma epidemia, a definição das localidades atingidas pode não ser tão direta. Nesse exemplo, pode não ser possível determinar onde a doença surgiu e para quais cidades ela se propagou. Assim, a metodologia proposta é dividida em duas formulações:

- Formulação 1 – Dada a ocorrência de um evento, as localidades atingidas são bem definidas;
- Formulação 2 – Dada a ocorrência de um evento, as localidades atingidas não são bem definidas ou todas as localidades observadas foram atingidas.

As formulações serão detalhadas nas seções a seguir.

### 2.1 Formulação 1

Dada a ocorrência de um evento, sejam  $N_{LA}$  o número de localidades atingidas e  $N_T$  o número total de localidades observadas, define-se

$$C_T = \{L_1, L_2, \dots, L_{N_T}\}, \quad (2.1)$$

como o conjunto de localidades observadas e

$$C_{LA} = \{L_i \in C_T \mid C_{LA} \subseteq C_T\} \quad (2.2)$$

como o conjunto de localidades atingidas; onde  $|C_{LA}| = N_{LA}$  é a cardinalidade de  $C_{LA}$ . Por sua vez, o conjunto  $C_T$  compreende todas as localidades imediatamente próximas às atingidas que tenham a mesma granularidade. Alguns exemplos são dados a seguir:

- Todos os bairros e regiões vizinhas para um evento que afete apenas uma região de um município (ex.: apagões, deslizamento de terra, problemas de mobilidade urbana);
- Todos os municípios de um estado para um evento que possa afetar uma ou mais cidades (ex.: terremotos, tsunamis, problemas de abastecimento, rompimentos de barragem);

- Todos os estados de um país para um evento que possa afetar um estado ou região (ex.: epidemia, crise hídrica).

Busca-se um conjunto de localidades de controle, estas devem possuir características similares às atingidas. Os controles serão utilizados para estabelecer o comportamento padrão<sup>1</sup> que uma localidade atingida deveria obedecer caso o evento não tivesse ocorrido. Assim, com um número  $N_{LC}$  de localidades de controle, tem-se

$$C_{LC} = \{L_i \in C_T \mid C_{LC} \subset C_T; C_{LC} \cap C_{LA} = \emptyset\} \quad (2.3)$$

como o conjunto de localidades de controle.

Para avaliar o efeito do evento nas localidades, estuda-se  $N_a$  características que possivelmente foram impactadas.<sup>2</sup> Assim, define-se

$$C_A = \{A_1, A_2, \dots, A_{N_a}\} \quad (2.4)$$

como o conjunto de atributos que possivelmente foram afetados pelo evento. Assume-se que os dados estejam estruturados em séries temporais, onde  $T_e$  marca o instante da ocorrência do evento. Assim, as etapas desta formulação são descritas a seguir:

- a) Definir o conjunto das localidades de controle;
- b) Análise univariável
  - Descrever os efeitos do evento, ou falta destes, observados em cada atributo, por localidade atingida;
  - Avaliar se o efeito do evento em cada atributo é significativo nas localidades atingidas e nos respectivos controles;
- c) Análise multivariável
  - Avaliar se o evento provoca mudança no padrão de comportamento das localidades atingidas;
  - Investigar interações entre as variáveis.

Ao final das etapas, torna-se possível inferir quais localidades mais sofreram impactos decorrentes do evento em estudo.

## 2.2 Formulação 2

Dada a ocorrência de um evento, busca-se 1) descobrir quais localidades mais sofreram impacto, ou 2) avaliar o efeito do evento em todas as localidades observadas.

<sup>1</sup> Comportamento padrão se refere ao estado dos atributos antes do evento.

<sup>2</sup> Estas são as características que serão avaliadas na metodologia para análise de um provável impacto do evento, elas podem ser diferentes das características do conjunto para seleção dos controles.

Para isso, define-se o conjunto  $C_T$  de localidades imediatamente próximas ao local do evento de acordo com (2.1). Então, estuda-se um conjunto  $C_A$  de  $N_a$  atributos possivelmente afetados, de acordo com (2.4).

Para investigar o impacto do evento e encontrar o conjunto  $C_{LA}$  (2.2) de localidades atingidas são executadas as etapas desta formulação, como descritas a seguir:

- a) Análise univariável
  - Descrever os efeitos do evento, ou falta destes, observados em cada atributo, por localidade observada;
  - Avaliar se o efeito do evento em cada atributo é significativo nas localidades observadas;
- b) Análise multivariável
  - Avaliar se o evento provoca mudança no padrão de comportamento das localidades observadas;
  - Investigar interações entre as variáveis;
- c) Se necessário, definir o conjunto das localidades atingidas;

Na etapa final, o conjunto de localidades atingidas é constituído pelas localidades que apresentarem indicativos de impacto do evento em ao menos duas das análises realizadas.

### 2.3 Conclusão do Capítulo

O conjunto de análises uni e multivariáveis permite observar o efeito do evento por diferentes perspectivas. Após observar os distintos resultados da aplicação da metodologia o pesquisador pode decidir se houve ou não impacto do evento nas localidades atingidas (Formulação 1) ou selecionar as localidades nas quais houve impacto (Formulação 2). Ao comparar as formulações, observa-se que a Formulação 1 é mais específica, pois ela analisa somente as localidades atingidas e seus controles. Dessa forma, para um conjunto de localidades atingidas relativamente pequeno, ela possui uma implementação mais breve. Já a Formulação 2 possui caráter exploratório, por isso, para um conjunto de localidades observadas relativamente grande, sua implementação pode ser mais custosa. Por fim, vale ressaltar que esta metodologia não assume relação de causalidade entre o evento e o efeito observado nas variáveis, apenas estabelece, ou não, correlação entre eles; e destaca-se que para a aplicação do método os dados devem estar estruturados em forma de séries temporais.

### 3 REVISÃO TEÓRICA

Este capítulo aborda detalhadamente os conceitos essenciais para aplicar a metodologia proposta nesta dissertação. A análise estatística inclui conceitos de diagramas de controle e testes estatísticos, estes que são fundamentais para a proposta das análises univariáveis. Já a análise por aprendizado de máquina apresenta dois principais algoritmos utilizados para as etapas da análise multivariável e da definição das localidades de controle.

#### 3.1 Análise Estatística

As ferramentas estatísticas aqui abordadas são empregadas na etapa de análise univariável da metodologia proposta. Os diagramas de controle permitem comparar uma determinada variável com relação ao seus limites históricos de uma forma visual (por meio de gráficos) enquanto que o emprego do teste estatístico permite inferir se há ou não diferença significativa entre os períodos; esses conceitos são explanados a seguir.

##### 3.1.1 Diagramas de Controle

Diagramas de controle são uma ferramenta estatística amplamente utilizada em epidemiologia para estudar variações nos níveis endêmicos sazonais; são uma forma simples de descrever o comportamento e os resultados de um processo, cronologicamente, em formato de série temporal. Fundamentado na teoria de probabilidades, permitem comparar a incidência observada (dados atuais do processo) com os limites de incidência esperada (dados históricos). Dessa forma, permitem diferenciar processos com características de variação espontânea daqueles que podem apresentar motivos para variação [8, 9, 10, 11].

Um diagrama é comumente constituído de uma linha central, um limite de controle superior e um limite de controle inferior, todos gerados a partir de dados históricos (dados anteriores ao período em estudo); por fim, os dados atuais são apresentados em contraste com os limites já mencionados. A linha central corresponde ao comportamento típico do processo e é usada para detectar se houve alguma mudança estatística, esta linha é normalmente dada pela média aritmética. Já os limites superiores e inferiores são, normalmente, dados pelo valor da média somados e subtraídos de 3 vezes o desvio padrão do universo do estudo, respectivamente [9]. Contudo, nesta dissertação é utilizado 1,96 desvios padrão para construção dos limites, pois esse valor corresponde a 95% da distribuição.

### 3.1.2 Teste $t$ Pareado

Uma hipótese estatística pode ser definida como uma afirmação sobre algum aspecto de um modelo, por exemplo, se duas diferentes configurações de um algoritmo apresentam o mesmo desempenho em uma tarefa [12]. De maneira geral, escreve-se:

$$\begin{cases} H_0 : \theta_1 = \theta_2 \\ H_1 : \theta_1 \neq \theta_2 \end{cases}, \quad (3.1)$$

onde, neste exemplo,  $\theta_1$  e  $\theta_2$  são as médias aritméticas do desempenho das configurações 1 e 2, respectivamente. A afirmação  $H_0 : \theta_1 = \theta_2$  é chamada de hipótese nula e afirma que, em média, as configurações resultam em igual desempenho; já a  $H_1 : \theta_1 \neq \theta_2$  é chamada de hipótese alternativa e afirma que há uma diferença bilateral entre os desempenhos, isto é, a Configuração 1 pode ter desempenho superior ou inferior à Configuração 2.

Para verificar as hipóteses é realizado o teste de hipótese, que é executado ao calcular uma estatística de teste (uma equação definida de acordo com o teste aplicado) apropriada e então em rejeitar ou falhar em rejeitar a hipótese nula. Durante o teste, dois tipos de erro podem ser mensurados, são eles o Erro Tipo I ( $\alpha$ ): também chamado de nível de significância, que ocorre quando a hipótese nula é rejeitada mesmo sendo verdadeira; e o Erro Tipo II ( $\beta$ ): quando a hipótese nula não é rejeitada mesmo esta sendo falsa [12]. Normalmente um teste estatístico é projetado para possuir um erro tipo II baixo para um nível de significância definido *a priori*.

Neste contexto, assume-se que as configurações serão testadas em instâncias distintas, assim pode-se escrever um modelo estatístico para descrever o experimento:

$$y_{ij} = \theta_i + \beta_j + \epsilon_{ij} \quad (3.2)$$

onde  $y_{ij}$  é o desempenho da configuração  $i$  na instância  $j$ ,  $\theta_i$  é a média real do desempenho da configuração  $i$ ,  $\beta_j$  é o efeito da instância  $j$  no desempenho das configurações, e  $\epsilon_{ij}$  é um erro aleatório com média 0 e variância  $\sigma_i^2$ , isto é, a variância de desempenho das configurações  $i$ .

Se for computada a  $j$ -ésima diferença pareada  $d_j = y_{1j} - y_{2j}$ , ao assumir que as variâncias de ambas as configurações são iguais, o valor esperado dessa diferença é:

$$\begin{aligned} \theta_d &= E(d_j) \\ &= E(y_{1j} - y_{2j}) \\ &= \theta_1 + \beta_j - \theta_2 - \beta_j \\ &= \theta_1 - \theta_2 \end{aligned} \quad (3.3)$$



Assim, pode-se verificar que o efeito das instâncias  $\beta_j$  é cancelado quando as observações são pareadas [12]. Dessa forma, é possível inferir sobre a diferença em desempenho médio das configurações  $\theta_1 - \theta_2$  ao inferir sobre a média das diferenças  $\theta_d$ , ou seja, a hipótese 3.1 é equivalente a:

$$\begin{cases} H_0 : \theta_d = 0 \\ H_1 : \theta_d \neq 0 \end{cases} . \quad (3.4)$$

O teste de hipótese da formulação 3.4 é realizado ao calcular a seguinte estatística de teste:

$$t_0 = \frac{\bar{d}}{S_d/\sqrt{n}}, \quad (3.5)$$

onde  $n$  é o número de amostras<sup>1</sup>,

$$\bar{d} = \frac{1}{n} \sum_{j=1}^n d_j \quad (3.6)$$

é a média amostral das diferenças e

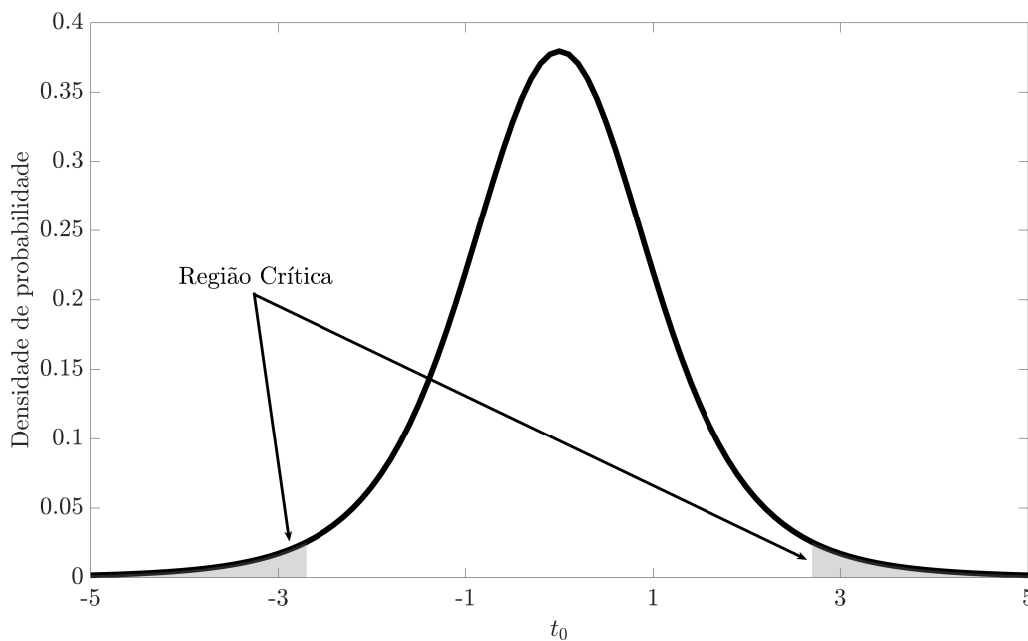
$$\begin{aligned} S_d &= \sqrt{\frac{\sum_{j=1}^n (d_j - \bar{d})^2}{n-1}} \\ &= \sqrt{\frac{\sum_{j=1}^n d_j^2 - \frac{1}{n} (\sum_{j=1}^n d_j)^2}{n-1}} \end{aligned} \quad (3.7)$$

é o desvio padrão amostral das diferenças.

Após este cálculo, compara-se  $t_0$  com os valores limites da região crítica de uma distribuição  $t$  com  $n - 1$  graus de liberdade, denotado de  $t_{\alpha/2, n-1}$  para um nível de significância  $\alpha$ . A Figura 3.1 apresenta um exemplo de região crítica para  $\alpha = 0,05$  e  $n = 5$ . Desse modo, infere-se que  $H_0 : \theta_d = 0$  será rejeitada se  $|t_0| > t_{\alpha/2, n-1}$ ; caso contrário, o teste falha em rejeitar a hipótese nula. Este teste estatístico é denominado teste  $t$  pareado.

Apesar de ser possível inferir se a hipótese nula será ou não rejeitada, a estatística de teste não permite deduzir o quão perto ou longe a hipótese estava de ser, ou não, rejeitada. Para isso, comumente utiliza-se de uma variável denominada  $p$ -valor. O  $p$ -valor é definido como o menor nível de significância que faz com que a hipótese nula seja rejeitada. Dessa forma, ele indica a quantidade de informação contra a hipótese nula, além de ser um medidor da significância dos dados [12].

<sup>1</sup> Normalmente o Erro Tipo II  $\beta$  diminui conforme o número de amostras é incrementado.



**Figura 3.1** – Ilustração da região crítica de uma distribuição  $t$  com  $n = 5$  graus de liberdade e nível de significância  $\alpha = 0,05$ .

## 3.2 Análise por Inteligência Computacional

A análise por inteligência computacional será empregada em duas etapas distintas da metodologia. Para a primeira etapa da Formulação 1, busca-se uma ferramenta que seja capaz de agrupar as localidades de acordo com suas semelhanças, para que seja possível definir localidades de controle. Já a segunda etapa requer uma ferramenta que possa investigar o efeito conjunto de todas as variáveis estudadas em uma dada localidade. Dada essas restrições, algoritmos de agrupamento se mostram capazes de realizar estas tarefas, pois conseguem detectar padrões e separar um conjunto de dados de acordo com uma função de similaridade [13].

### 3.2.1 Agrupamento (*Clusterização*)

A análise por agrupamento, ou *clusterização*, busca particionar um conjunto de dados de acordo com o grau de similaridade de seus elementos [13]. Assim, um grupo (*cluster*) contém dados similares entre si e dissimilares entre dados de outros grupos. Essa similaridade deve ser mensurada para que o particionamento ocorra. Para isso, comumente utiliza-se uma função de distância [14].

Uma forma de particionar um conjunto é utilizar algoritmo de agrupamento *hard*, isto é, dado um conjunto  $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \in \mathbb{R}^{n \times p}$  com  $n$  amostras e  $p$  atributos, particioná-lo em  $c$  grupos mutuamente exclusivos. Dessa forma, um dado pertence a apenas um único grupo. Uma partição *hard* pode ser representada através de funções de

pertinência da seguinte forma:

**Definição 3.1** (Função de Pertinência (adaptado de [15])). *Seja  $B_i$  uma partição de um conjunto  $X$ . Então, a função de pertinência da partição hard  $B_i \subset X$  é definida por:*

$$\mu_{ij} = \begin{cases} 1, & \mathbf{x}_j \in B_i \\ 0, & \mathbf{x}_j \notin B_i \end{cases}. \quad (3.8)$$

A função de pertinência  $\mu_{ij}$  expressa o grau com que um dado  $\mathbf{x}_j$  pertence à partição  $B_i$ . Portanto, infere-se que, para um agrupamento *hard*, cada amostra pertence exclusivamente a um único grupo (*cluster*). Em algumas aplicações, essa interpretação advinda do agrupamento *hard* pode levar a representações que não descrevem bem o comportamento do conjunto de dados [13].

Uma forma de contornar esse problema, é a partir da suavização da forte restrição imposta por (3.8). Ao tornar a função de pertinência contínua, surge uma nova classe de algoritmos: o agrupamento nebuloso (*clusterização fuzzy*). Métodos dessa classe particionam o conjunto de forma que um elemento possa pertencer a vários grupos com diferentes graus de pertinência [13]. Essa nova forma de interpretar o agrupamento possibilita fazer inferências a respeito da similaridade de um dado a todos os grupos. Desse modo, uma partição *fuzzy* atende às seguintes condições:

$$\mu_{ij} \in [0, 1], \quad 1 \leq i \leq c, \quad 1 \leq j \leq n, \quad (3.9a)$$

$$\sum_{i=1}^c \mu_{ij} = 1, \quad 1 \leq j \leq n, \quad (3.9b)$$

$$0 < \sum_{j=1}^n \mu_{ij} < n, \quad 1 \leq i \leq c. \quad (3.9c)$$

A condição (3.9a) estabelece que um dado pode pertencer a cada um dos grupos com graus de pertinência entre 0 e 1; a (3.9b) restringe a soma das pertinências de um dado a todos os grupos a ser unitária; já a (3.9c) garante que existam no mínimo dois grupos e que nenhum deles seja vazio.

### 3.2.2 Agrupamento Nebuloso (*Clusterização Fuzzy*)

Os algoritmos de agrupamento nebuloso são baseados em função objetivo. Deles, diversos são fundamentados em minimizar o funcional *Fuzzy c-Means* dado por [15]:

$$\mathcal{J}(X; U, K) = \sum_{i=1}^c \sum_{j=1}^n \mu_{ij}^m d_{ijA}^2 \quad (3.10a)$$

onde  $U$  é uma partição nebulosa de  $X$ ,

$$K = \{\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_c\} \in \mathbb{R}^{c \times p} \quad \mathbf{k}_i \in \mathbb{R}^p \quad (3.10b)$$

é o conjunto de centros dos grupos,

$$d_{ijA}^2 = \|\mathbf{x}_j - \mathbf{k}_i\|_A^2 = (\mathbf{x}_j - \mathbf{k}_i)^T A (\mathbf{x}_j - \mathbf{k}_i) \quad (3.10c)$$

é qualquer produto escalar que induz uma norma  $A$  em  $\mathbb{R}^p$  (função de distância) e,

$$m \in (1, \infty) \quad (3.10d)$$

é um expoente chamado de parâmetro de *fuzzyficação* que determina quão nebulosos serão os grupos resultantes. Essa é uma formulação generalizada para o problema de agrupamento, visto que quando  $m \rightarrow 1$ , os grupos tendem a se tornar similares aos resultantes do agrupamento *hard*.

### 3.2.2.1 O Algoritmo *Fuzzy c-means*

O algoritmo *Fuzzy c-means* (FCM), inicialmente proposto por Dunn [16] para  $m = 2$  e generalizado para  $m \in (1, \infty)$  por Bezdek [15], utiliza a norma euclidiana como medida de dissimilaridade. Nele, os centros são encontrados através da minimização da função objetivo (3.10a). Este é em um problema de otimização não linear que é resolvido por meio de uma iteração de Picard simples através das condições de primeira ordem para pontos estacionários de (3.10a). Devido à alta dimensionalidade da função objetivo, sua solução é obtida encontrando as passagens por zero das derivadas parciais.

Como

$$\min_{\mathbf{k}} \left\{ \sum_{i=1}^c \sum_{j=1}^n \mu_{ij}^m d_{ijA}^2 \right\} = \sum_{j=1}^n \min_{\mathbf{k}} \left\{ \sum_{i=1}^c \mu_{ij}^m d_{ijA}^2 \right\},$$

se  $\sum_{i=1}^c \mu_{ij}^m d_{ijA}^2$  é minimizado, (3.10a) também é. Assim, os pontos estacionários de (3.10a) são encontrados ao adicionar a restrição (3.9b) por meio de multiplicadores de Lagrange  $\lambda \in \mathbb{R}$ . Então, para todo  $\mathbf{x} \in X$ , o Lagrangiano é:

$$\mathcal{J}_j(\lambda, \boldsymbol{\mu}_j) := \sum_{i=1}^c \mu_{ij}^m d_{ijA}^2 - \lambda \left( \sum_{i=1}^c \mu_{ij} - 1 \right) \quad (3.11)$$

O par  $(\lambda, \boldsymbol{\mu}_j)$  é estacionário apenas se  $\nabla_{\lambda, \boldsymbol{\mu}_j} \mathcal{J}_j(\lambda, \boldsymbol{\mu}_j) = (0, \mathbf{0} \in \mathbb{R}^c)$ . Igualando este gradiente a zero, obtém-se

$$\frac{\partial}{\partial \lambda} \mathcal{J}_j(\lambda, \boldsymbol{\mu}_j) = \left( \sum_{i=1}^c \mu_{ij} - 1 \right) = 0 \quad (3.12a)$$

$$\frac{\partial}{\partial \mu_{ij}} \mathcal{J}_j(\lambda, \boldsymbol{\mu}_j) = m \cdot \mu_{ij}^{m-1} \cdot d_{ijA}^2 - \lambda = 0. \quad (3.12b)$$

Ao considerar que  $d_{ijA}^2 \neq 0$ , obtém-se a partir de (3.12b):

$$\mu_{ij} = \left( \frac{\lambda}{m \cdot d_{ijA}^2} \right)^{\frac{1}{m-1}}. \quad (3.13)$$

Ao substituir (3.13) em (3.12a) tem-se:

$$\begin{aligned}
 1 &= \sum_{k=1}^c \mu_{ij} \\
 &= \sum_{k=1}^c \left( \frac{\lambda}{m} \right)^{\frac{1}{m-1}} \left( \frac{1}{d_{kjA}^2} \right)^{\frac{1}{m-1}} \\
 &= \left( \frac{\lambda}{m} \right)^{\frac{1}{m-1}} \sum_{k=1}^c \left( \frac{1}{d_{kjA}^2} \right)^{\frac{1}{m-1}}.
 \end{aligned} \tag{3.14}$$

Portanto,

$$\left( \frac{\lambda}{m} \right)^{\frac{1}{m-1}} = \frac{1}{\sum_{k=1}^c \left( \frac{1}{d_{kjA}^2} \right)^{\frac{1}{m-1}}}. \tag{3.15}$$

Ao utilizar (3.15) em (3.13), obtém-se a função de pertinência para o *Fuzzy c-means*:

$$\begin{aligned}
 \mu_{ij} &= \frac{1}{\sum_{k=1}^c \left( \frac{1}{d_{kjA}^2} \right)^{\frac{1}{m-1}}} \cdot \left( \frac{1}{d_{ijA}^2} \right)^{\frac{1}{m-1}} \\
 &= \frac{1}{\sum_{k=1}^c \left( \frac{d_{ijA}}{d_{kjA}} \right)^{\frac{2}{m-1}}}.
 \end{aligned} \tag{3.16}$$

Para que a condição (3.9c) seja satisfeita, fazer  $n > c$  é suficiente.

Por fim, os centros  $\mathbf{k}_i$  devem minimizar (3.10a). Para isso, o gradiente de  $\mathcal{J}$  em relação a  $\mathbf{k}_i$ ,  $i \in \mathbb{N}_{\leq c}$ , é necessariamente zero. Portanto, para todo  $\xi \in \mathbb{R}^p$  com  $t \in \mathbb{R}$  e  $A = I$ ,

$$\begin{aligned}
 0 &= \frac{\partial}{\partial k_i} \sum_{j=1}^n \sum_{l=1}^c \mu_{ij}^m \|x_j - k_l\|^2 \\
 &= \sum_{j=1}^n \mu_{ij}^m \frac{\partial}{\partial k_i} \|x_j - k_i\|^2 \\
 &= \sum_{j=1}^n \mu_{ij}^m \lim_{t \rightarrow 0} \frac{\|x_j - (k_i + t\xi)\|^2 - \|x_j - k_i\|^2}{t} \\
 &= \sum_{j=1}^n \mu_{ij}^m \lim_{t \rightarrow 0} \frac{1}{t} \left[ ((x_j - k_i) - t\xi)^T ((x_j - k_i) - t\xi) - (x_j - k_i)^T (x_j - k_i) \right] \\
 &= \sum_{j=1}^n \mu_{ij}^m \lim_{t \rightarrow 0} \frac{-2t(x_j - k_i)^T \xi + t^2 \xi^T \xi}{t} \\
 &= -2 \sum_{j=1}^n \mu_{ij}^m (x_j - k_i)^T \xi,
 \end{aligned}$$

a partir disto, segue que

$$\begin{aligned}
& \frac{\partial}{\partial k_i} \mathcal{J} = 0 \\
& \iff \sum_{j=1}^n \mu_{ij}^m (x_j - k_i) = 0 \\
& \iff k_i = \frac{\sum_{j=1}^n \mu_{ij}^m x_j}{\sum_{j=1}^n \mu_{ij}^m}.
\end{aligned} \tag{3.17}$$

No caso em que  $m = 1$  o algoritmo não é definido devido à indeterminação que ocorre em (3.16).

O algoritmo *Fuzzy c-means* se resume em computar iterativamente os valores dos centros e, em seguida, das pertinências<sup>2</sup> até que a condição de parada  $\|K^{(t)} - K^{(t-1)}\| < \epsilon$  seja satisfeita, onde  $t$  é a iteração atual e  $\epsilon > 0$  um valor de tolerância. Para ser inicializado, o FCM requer a definição da quantidade de grupos e uma estimativa da matriz de pertinências  $U$  ou dos centros  $K$ .

### 3.2.3 Agrupamento Subtrativo

O Agrupamento Subtrativo (*Clusterização Subtrativa*), inicialmente proposto por Chiu [7] como uma técnica de *clusterização*, consiste em um método utilizado para a estimação do número de grupos e de seus respectivos centros comumente empregado para inicializar algoritmos de agrupamento como o *Fuzzy c-means*.

Neste método, cada amostra do conjunto de dados é vista como um potencial centro de grupo. As amostras que apresentam maior potencial são sequencialmente escolhidas como centro até que um critério de parada seja atingido. O algoritmo é fundamentado como a seguir: seja  $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \in \mathbb{R}^{n \times p}$  um conjunto de dados, define-se a medida do potencial de um ponto  $\mathbf{x}_i$  ser centro de grupo como:

$$f_i = \sum_{j=1}^n e^{-\alpha \|\mathbf{x}_i - \mathbf{x}_j\|^2}, \tag{3.18}$$

onde  $\alpha = 4/r_a^2$  e  $r_a$  é uma constante positiva que define o raio de vizinhança para cada centro de *cluster*. Esse raio determina a influência que cada centro terá nos dados vizinhos e, a partir de (3.18), tem-se que quanto maior é esta influência, uma menor quantidade de grupos é obtida [17].

Após o potencial de todos os dados serem computados, a amostra com maior potencial é escolhida como primeiro centro de grupo. Seja  $\mathbf{x}_1^*$  a posição do primeiro centro e  $f_1^*$  seu valor de potencial, cada ponto  $\mathbf{x}_i$  restante tem seu potencial ajustado por

$$f_i \leftarrow f_i - f_1^* e^{-\beta \|\mathbf{x}_i - \mathbf{x}_1^*\|^2}, \tag{3.19}$$

<sup>2</sup> Ou iniciar com a computação das pertinências e, em seguida, computar os centros

onde  $\beta = 4/r_b^2$  e  $r_b$  é uma constante positiva que representa o raio de vizinhança sobre o qual haverá redução de potencial;  $r_b = 1,25r_a$  é um valor comumente utilizado [17, 18].

Após o potencial de todas as amostras serem ajustados por (3.19), o dado com o novo maior potencial é selecionado como segundo centro. Assim, após o  $k$ -ésimo centro ser selecionado, o potencial das amostras restantes é ajustado por

$$f_i \leftarrow f_i - f_k^* e^{-\beta \|\mathbf{x}_i - \mathbf{x}_k^*\|^2}, \quad (3.20)$$

onde  $\mathbf{x}_k^*$  é a posição do  $k$ -ésimo centro e  $f_k^*$  o seu valor de potencial. Este processo é repetido iterativamente até que o critério de parada seja alcançado.

### 3.3 Conclusão do Capítulo

Este capítulo apresentou conceitos dos diagramas de controle e do teste  $t$  pareado, ambos métodos usados para comparação de séries históricas de um dado processo. Os algoritmos de agrupamento abordados nesta dissertação foram explicitados, sendo eles o FCM, utilizado para particionar conjuntos de dados em grupos dissimilares entre si, e o Agrupamento Subtrativo, que aqui tem o papel de inicializar o FCM.

## 4 METODOLOGIA PROPOSTA

Este capítulo apresenta os métodos propostos para a análise de impactos de desastres e epidemias. Estes que são utilizados nos passos das formulações que foram discutidos no Capítulo 2. Propõe-se uma nova técnica intitulada Gráfico de Vetores Resultantes, desenvolvida para uma melhor visualização e síntese dos resultados. Também é demonstrado como as técnicas de aprendizado de máquina podem ser utilizadas tanto para a definição de localidades de controle quanto para a análise multivariável. Por fim, uma seção é dedicada ao processo de automatização da extração e tratamento de dados.

### 4.1 Diagramas de Controle

Dado um evento ocorrido em  $T_e$  e uma série temporal de um atributo contido em  $C_A$ , dado por (2.4), referente a uma localidade contida em  $C_{LA}$ , como em (2.2), um diagrama de controle pode ser construído ao calcular a média aritmética e o desvio padrão de todos os dados anteriores ao momento de ocorrência do evento  $T_e$ , obedecendo à granularidade dos dados. Por exemplo, suponha que seja necessário avaliar a evolução mensal de uma variável e que o intervalo de tempo estudado anterior a  $T_e$  seja de três anos. Assim, calcula-se a média e o desvio padrão correspondente a cada mês e é desenhado o limite superior, dado por média + 1,96 desvios padrão, e o limite inferior, dado por média - 1,96 desvios padrão [8]. Esses limites são sobrepostos com os dados do período em estudo (dados após  $T_e$ ) para gerar o diagrama, como exemplificado na Figura 4.1:

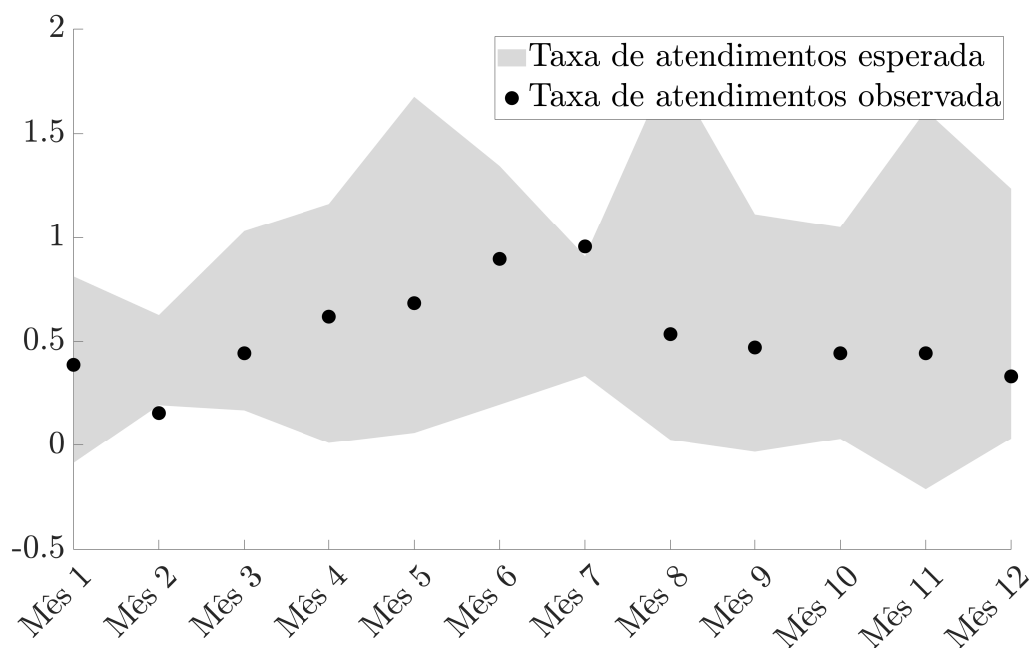
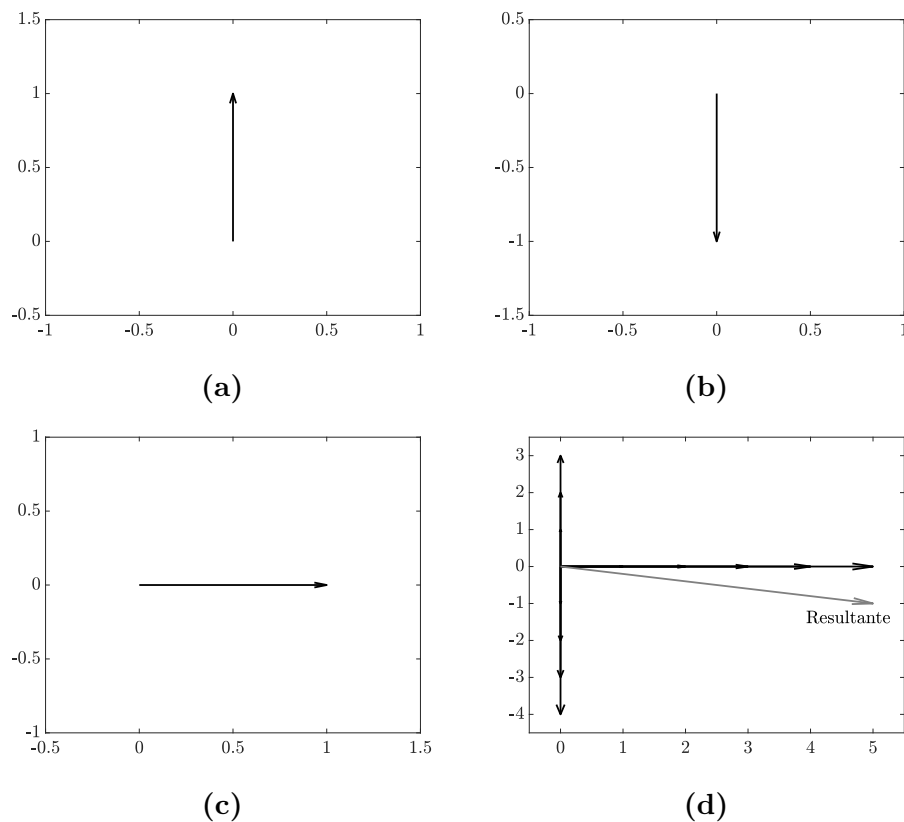


Figura 4.1 – Ilustração de um diagrama de controle



## 4.2 Gráfico de Resultantes Vetoriais

O Gráfico de Resultantes Vetoriais (GRV) é uma técnica proposta com a intenção de sintetizar as informações dos diagramas de controles em apenas um gráfico. Os GRVs contemplam, simultaneamente, três informações dos diagramas: atributos acima, dentro ou abaixo dos limites históricos. A partir dos diagramas, cada vez que o atributo ultrapassa o limite superior esperado, é atribuído um vetor unitário na direção de crescimento do eixo das ordenadas (Figura 4.2a); quando o atributo fica abaixo do limite inferior esperado, um vetor unitário é atribuído na direção de diminuição do eixo das ordenadas (Figura 4.2b); e quando o atributo fica dentro dos limites esperados, é atribuído um vetor unitário no sentido de crescimento do eixo das abcissas (Figura 4.2c). Por fim, após “percorrer” todo o diagrama, os vetores são somados para gerar um resultante (Figura 4.2d).



**Figura 4.2** – Ilustração da técnica de Gráficos de Resultantes Vetoriais utilizada para sintetizar as informações dos diagramas de controles, indicando os vetores unitários atribuídos ao gráfico: 4.2a acima do limite superior esperado; 4.2b abaixo do limite inferior esperado; 4.2c dentro dos limites esperados; e 4.2d vetor resultante.

Quando mais de um diagrama de controle é utilizado para gerar o GRV, os resultantes vetoriais são exibidos em uma única figura. No GRV, se o resultante estiver no primeiro quadrante, indicará um aumento na variável para o período estudado; se estiver

no quarto quadrante, a variável terá estado aquém do esperado; e se estiver próximo do eixo das abcissas, a variável estará dentro dos limites históricos.

### 4.3 Formulação 1

Esta formulação pressupõe que o conjunto de localidades atingidas  $C_{LA}$  seja conhecido. Assim, determina-se o conjunto de localidades observadas  $C_T$  e a partir dele busca-se encontrar o conjunto de localidades de controle  $C_{LC}$ . As variáveis utilizadas para encontrar  $C_{LC}$  podem ser as mais diversas, desde variáveis socioeconômicas, para obter localidades com semelhantes perfis socioeconômicos, até variáveis pertencentes ao conjunto  $C_A$ . Contudo, a escolha destas variáveis depende do tipo e objetivo do estudo e aconselha-se que sejam escolhidas pelo pesquisador que utilizar esta metodologia. Para exemplificar, algumas possíveis variáveis são exibidas no Capítulo 5. Ressalta-se que nesta etapa utiliza-se apenas dados anteriores ao evento.

#### 4.3.1 Definição das Localidades de Controle

A busca pelo conjunto das localidades de controle  $C_{LC}$  é realizada com base na medida de similaridade *fuzzy* (3.16). O algoritmo *Fuzzy c-means* particiona o conjunto de dados, com  $m = 2$ , onde os centros são as amostras correspondentes às localidades contidas em  $C_{LA}$ . Após apenas 1 iteração, as pertinências das demais localidades ( $C_T \cap C_{LA}$ ) são ordenadas de forma decrescente para a escolha dos controles. Aqui, a pertinência exprime o quão similar as localidades observadas são das localidades atingidas (centros), de acordo com um conjunto de características. Desse modo, as localidades com maior pertinência a cada centro são escolhidas como seu respectivo controle. Por fim, deve-se verificar se não há controles repetidos entre as localidades atingidas. Caso isso ocorra, deve-se substituir um dos controles pela segunda localidade mais similar.

#### 4.3.2 Análise Univariável

A análise univariável é realizada individualmente para cada atributo contido em  $C_A$  e é composta por três etapas:

- Análise descritiva;
- Análise por GRV;
- Análise estatística (comparação com controles).

A análise descritiva tem o objetivo de apresentar os dados, informar proporções relativas a dados antes e após o evento para cada localidade e identificar possíveis problemas com a base de dados, tal como dados faltantes.

Na segunda etapa, o GRV é aplicado nos dados das localidades atingidas. O GRV é uma comparação dos dados do estado atual da localidade atingida (após o evento) com os dados históricos (antes do evento), ou seja, é uma auto-comparação. Procura-se estabelecer quais foram os principais impactos sofridos nos atributos, se houve aumento, queda, ou permanência dos valores observados na série histórica. Para auxiliar na interpretação, recomenda-se verificar quais localidades apresentam vetores mais próximos do eixo das ordenadas.

Já a comparação com o controle se dá ao avaliar as séries temporais antes e após o evento por meio do teste  $t$  pareado formulado como em (3.4), ou seja, a hipótese nula é de que não há diferença entre os períodos anterior e posterior a  $T_e$  para o atributo em estudo; já a hipótese alternativa é de que existe uma diferença.

O teste é realizado para cada localidade, atingida e controle, com o nível de significância  $\alpha$ . Devido à restrição de que os dados a serem comparados devem possuir o mesmo número de observações, sugere-se calcular a média dos dados anteriores ao evento, analogamente ao processo de construção do diagrama de controle descrito na Seção 4.1. O  $p$ -valor resultante do teste permite verificar se a diferença observada é significativa para o nível de significância especificado, ou para qualquer  $\alpha$  desejado. Após a realização dos testes, os possíveis resultados são descritos a seguir:

- **Diferença significativa apenas na localidade atingida:** Há evidências de impacto do evento, visto que houve alteração apenas em uma das localidades.
- **Diferença significativa apenas no controle:** Há evidências de impacto do evento, visto que houve alteração apenas em uma das localidades.
- **Diferença significativa na localidade atingida e no controle:** Caso as diferenças tenham mesmo sinal, não há evidências de impacto do evento; se as diferenças têm sinais opostos, há evidências de impacto do evento.
- **Não há diferença significativa:** Não há evidências de impacto do evento.

Contudo, as análises não devem se limitar aos casos mencionados acima. O  $p$ -valor pode indicar o quão forte é a evidência contra a hipótese nula. Deste modo,  $p$ -valores que excedem mas ainda estão próximos à  $\alpha$  podem indicar um provável impacto do evento, ainda que com um nível de significância diferente [12].

### 4.3.3 Análise Multivariável

A análise multivariável é utilizada caso  $N_a > 1$  e é executada por meio do algoritmo *Fuzzy c-means* inicializado com agrupamento subtrativo. Um conjunto de dados  $\mathbf{X}^* = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{2N_{LA}}\} \in \mathbb{R}^{2N_{LA} \times N_a}$  é criado a partir da junção de dados anteriores e posteriores

ao evento. Dessa forma, tem-se o termo  $2N_{LA}$  que corresponde a  $N_{LA}$  observações pré e  $N_{LA}$  observações pós-evento; cada uma com  $N_a$  características. A forma empregada para adequar o conjunto a este formato é utilizar a média dos dados para cada período (pré e pós evento). Por exemplo, ao tomar a média anual para os dados da Figura 4.1, cada localidade apresenta apenas um valor para cada período, ao invés de 36 dados pré e 12 pós evento (um para cada mês). Ressalta-se que, como nesta etapa as amostras são  $N_a$ -dimensionais, a média deve ser realizada para cada dimensão individualmente.

O conjunto  $\mathbf{X}^*$  é particionado em 3 grupos. Para isso, varia-se o  $r_a$  até que este valor de  $c$  seja alcançado.<sup>1</sup> Após o particionamento, verifica-se a pertinência de cada observação aos grupos resultantes. Desse modo, é possível comparar a qual grupo uma localidade mais pertencia antes do evento com a qual ela mais pertence após o evento.

O objetivo desta análise é verificar se houve mudança de grupo, ou seja, se houve mudança de comportamento para uma dada localidade. Para simplicidade da análise, considera-se que uma observação pertence ao grupo ao qual esta apresenta maior pertinência. Mas, como esta análise utiliza de lógica *fuzzy*, a pertinência pode indicar o quanto que uma localidade se distanciou do seu comportamento padrão, isto é, se a mudança de grupo corresponde a uma pequena ou grande variação no valor de pertinência.

#### 4.4 Formulação 2

Esta formulação não pressupõe que o conjunto de localidades atingidas  $C_{LA}$  seja conhecido, seja pelo desconhecimento das áreas de provável impacto ou por o evento ser de uma magnitude suficientemente grande para que todas as localidades observadas tenham sido afetadas. Assim, busca-se estudar o conjunto de localidades observadas  $C_T$  para avaliar o impacto do evento em cada localidade e, se necessário, definir um conjunto de localidades atingidas  $C_{LA}$  a partir dos resultados. Como as localidades atingidas não são definidas, esta formulação não necessita das localidades de controle, então esta etapa é omitida.

A análise univariável segue idêntica à da Seção 4.3.2, com exceção da comparação com controles. A análise descritiva apresenta os dados e apresenta informações relevantes sobre eles; e a análise por Gráfico de Resultantes Vetoriais permite a auto-comparação dos dados atuais das localidades com seus limites históricos. Já a análise estatística é realizada para todas as localidades; como não há localidades de controle nessa formulação, esta análise se torna uma segunda forma de auto-comparação.

Análise multivariável segue os mesmos passos apresentados na Seção 4.3.3, com a mesma forma de interpretação dos resultados. Por fim, caso seja necessário selecionar o

<sup>1</sup> Nenhum valor de  $r_a$  é definido nesta metodologia visto que o  $c$  resultante depende do conjunto de dados.

conjunto  $C_{LA}$ , sugere-se ver quais localidades apresentam evidências de impacto do evento para ao menos dois dos três tipos de análise utilizados.<sup>2</sup>

## 4.5 Extração e Tratamento de Dados

O estudo de desastres e epidemias como os investigados nesta dissertação podem necessitar da extração e tratamento de um grande volume de dados advindos dos sistemas de informação (ver Capítulo 5). A extração dos dados provenientes do Sistema de Informação em Saúde para a Atenção Básica (SISAB), utilizado em um dos estudos de caso apresentados no Capítulo 5, foi realizada de forma automatizada como descrito a seguir.

### 4.5.1 Extração de Dados Automatizada

O código de automação da extração de dados foi escrito em linguagem JavaScript. O *software* Node.js se fez necessário para instalar o pacote *node-fetch* que dá acesso à função *fetch*. Essa função permite requisitar informações aos servidores de páginas na internet, o que possibilita salvá-las em um arquivo, como planilhas.

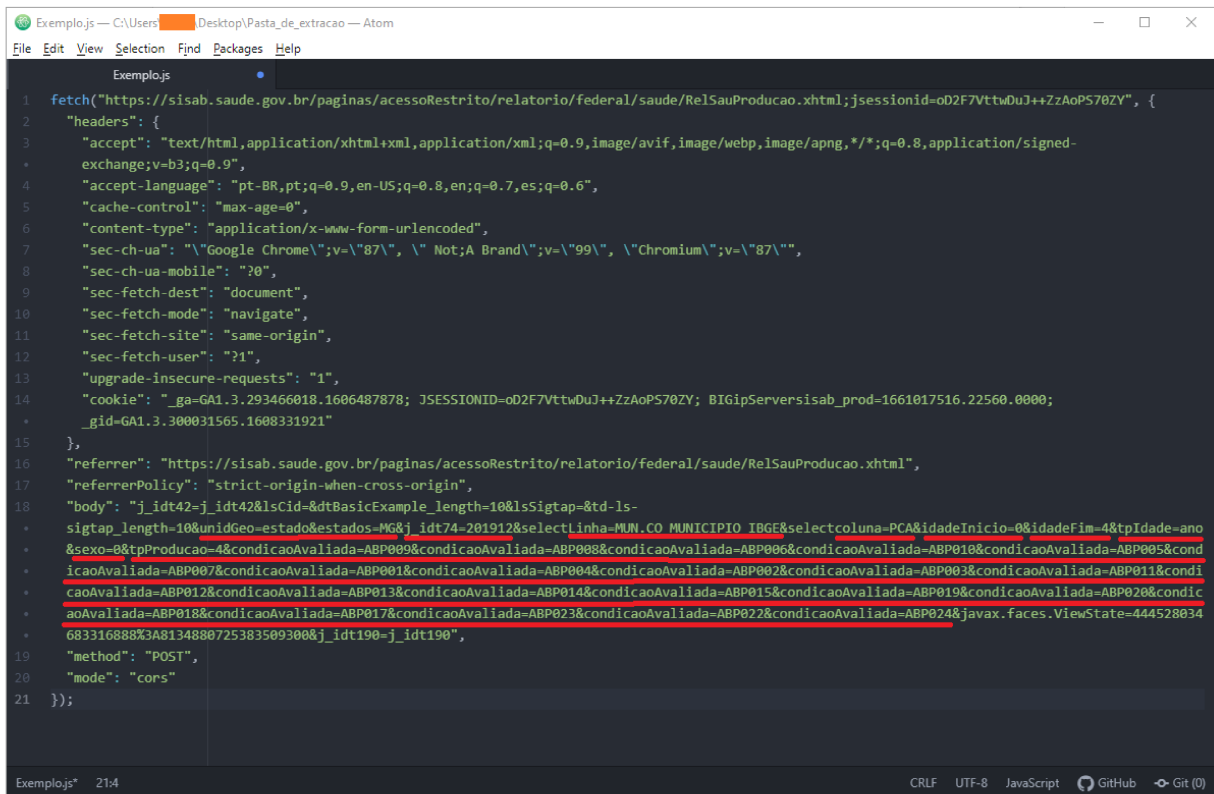
A instalação do pacote é feita através do *Prompt* de Comando do *Windows* ao executar o comando *npm install node-fetch*. Este pacote deve ser instalado na pasta onde a extração será realizada. O Google Chrome foi utilizado como navegador de internet e suas Ferramentas de Desenvolvedor foram utilizadas para inspecionar o site do SISAB. Na aba *Network* desta ferramenta foi possível visualizar o código que é enviado para o servidor do SISAB no momento que o *download* de uma planilha é requisitado. Esta informação foi então copiada em formato Node.js *fetch*.

Com o auxílio de um *software* editor de texto, foi possível identificar as partes do texto que se referem a cada informação de interesse: unidade geográfica, competência (ano e mês), linha e coluna do relatório (granularidade da localidade e atributos em estudo, respectivamente) e filtros (faixa etária, sexo, entre outros); como pode ser visto na Figura 4.3. Assim, a construção do método de extração consistiu em criar um código que envia a solicitação para o servidor do *site* e salva a planilha recebida em formato *.csv*. Esse código é constituído por um laço de repetição onde um conjunto de valores desejados para cada uma das informações de interesse supracitadas é solicitado a cada iteração.

O código de extração deve ser salvo em um arquivo com o nome do formato *nome\_do\_código.js*. Para executá-lo, deve-se utilizar o *Prompt* de Comando do *Windows*. Após mudar o diretório para a pasta onde as planilhas serão salvas, executa-se o seguinte comando: *node nome\_do\_código.js*. Assim, inicia-se o processo de extração automatizada dos dados.

---

<sup>2</sup> Três tipos de análise ao excluir a análise descritiva.



```

1 fetch("https://sisab.saude.gov.br/paginas/acesoRestrito/relatorio/federal/saude/RelSauProducao.xhtml;jsessionid=oD2F7VttwDuJ++ZzAoP570ZY", {
2   "headers": {
3     "accept": "text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,image/webp,image/apng,*/*;q=0.8,application/signed-exchange;v=b3;q=0.9",
4     "accept-language": "pt-BR,pt;q=0.9,en-US;q=0.8,en;q=0.7,es;q=0.6",
5     "cache-control": "max-age=0",
6     "content-type": "application/x-www-form-urlencoded",
7     "sec-ch-ua": "\"Google Chrome\";v=\"87\"\", \"Not;A Brand\";v=\"99\"\", \"Chromium\";v=\"87\"\"",
8     "sec-ch-ua-mobile": "?0",
9     "sec-fetch-dest": "document",
10    "sec-fetch-mode": "navigate",
11    "sec-fetch-site": "same-origin",
12    "sec-fetch-user": "?1",
13    "upgrade-insecure-requests": "1",
14    "cookie": "_ga=GA1.3.293466018.1606487878; JSESSIONID=oD2F7VttwDuJ++ZzAoP570ZY; BIGipServersisab_prod=1661017516.22560.0000; _gid=GA1.3.300031565.1608331921"
15  },
16  "referrer": "https://sisab.saude.gov.br/paginas/acesoRestrito/relatorio/federal/saude/RelSauProducao.xhtml",
17  "referrerPolicy": "strict-origin-when-cross-origin",
18  "body": "j_idt42=j_idt42&lsCid=&dtBasicExample_length=10&lsSigtap=&td-ls-sigtap_length=10&unidGeo=estado&estados=MG&j_idt74=201912&selectLinha=MUN.CO MUNICIPIO IBGE&selectcoluna=PCA&idadeInicio=0&idadeFim=4&tpIdade=ano&sexo=0&tpProducao=4&condicaoAvaliada=ABP009&condicaoAvaliada=ABP008&condicaoAvaliada=ABP006&condicaoAvaliada=ABP010&condicaoAvaliada=ABP005&condicaoAvaliada=ABP007&condicaoAvaliada=ABP001&condicaoAvaliada=ABP004&condicaoAvaliada=ABP002&condicaoAvaliada=ABP003&condicaoAvaliada=ABP011&condicaoAvaliada=ABP012&condicaoAvaliada=ABP013&condicaoAvaliada=ABP014&condicaoAvaliada=ABP015&condicaoAvaliada=ABP019&condicaoAvaliada=ABP020&condicaoAvaliada=ABP018&condicaoAvaliada=ABP017&condicaoAvaliada=ABP023&condicaoAvaliada=ABP022&condicaoAvaliada=ABP024&javax.faces.ViewState=444528034683316888%3A8134880725383509300&j_idt190=j_idt190",
19  "method": "POST",
20  "mode": "cors"
21 });

```

**Figura 4.3** – Campos de interesse no código JavaScript destacados para fácil visualização.

Dados advindos de sistemas de informação como o SISAB podem apresentar informações desnecessárias como cabeçalhos e rodapé. Em problemas com um grande volume de dados, retirar essas informações manualmente se torna inviável. Para contornar este problema, foi desenvolvido um código em MATLAB para “limpeza” do banco de dados.

O código consiste em um laço de repetição que utiliza a função *xlsread* para importar as planilhas para o *software*. Então, as linhas referentes ao cabeçalho e rodapé, que sempre são fixas, são excluídas. Após terminado o processo de limpeza, as planilhas são exportadas por meio da função *xlswrite*.

## 4.6 Conclusão do Capítulo

Este capítulo apresentou as técnicas utilizadas em cada passo da metodologia proposta, para ambas as formulações, assim como um método de extração automatizada de bancos de dados. Ressalta-se que, com a exceção do MATLAB, todos os *softwares* utilizados para extração de dados são gratuitos. Ainda assim, a limpeza das planilhas pode ser realizada por meio de outra linguagem de programação. O fluxograma apresentado na Figura 4.4 sumariza os passos para a implementação da metodologia.

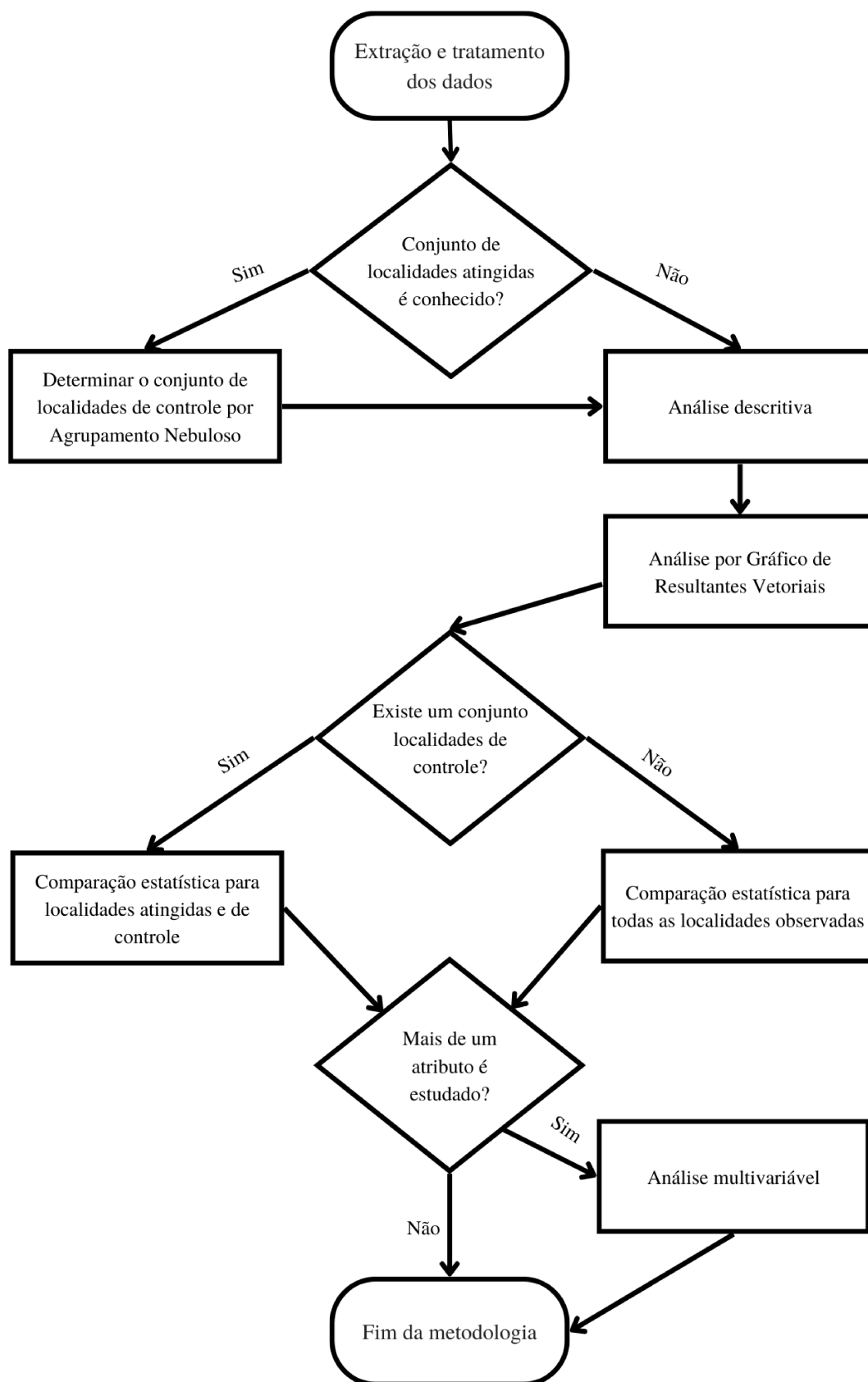


Figura 4.4 – Fluxograma dos passos para implementação da metodologia proposta.

## 5 ESTUDOS DE CASO

Neste capítulo serão apresentados estudos de caso nos quais a metodologia proposta foi utilizada para analisar o impacto de alguns eventos. A Formulação 1 é aplicada ao rompimento da barragem de Fundão, um desastre tecnológico. Especificamente, a variável de atendimentos por dengue foi selecionada devido à natureza do evento e consequente susceptibilidade da área afetada a apresentar poças de água parada, além do possível armazenamento incorreto de água fresca pela população atingida [5, 19]. A Formulação 2 é aplicada à pandemia de COVID-19, onde o atributo de interesse selecionado foi o número de atendimentos para acompanhamento do crescimento e desenvolvimento da criança (puericultura). Porém, qualquer atributo seria suficiente para aplicar a metodologia proposta.

### 5.1 Formulação 1 – Impacto do rompimento da barragem de Fundão, em Mariana–MG, nos casos de Dengue e no PIB municipal

Neste estudo de caso, serão avaliados os impactos do rompimento da barragem de Fundão em Mariana no estado de Minas Gerais ocorrido em 5 de novembro de 2015, identificado como o maior desastre socioambiental do país no setor de mineração [20]. Foram considerados como anteriores ao evento dados de novembro de 2010 a outubro de 2015 e como posteriores ao evento dados de novembro de 2015 a outubro de 2016.

Define-se  $C_T$  como todos os  $N_T = 853$  municípios de Minas Gerais e  $C_{LA}$  como os  $N_{LA} = 36$  municípios de Minas Gerais que foram direta ou indiretamente atingidos pela lama decorrente do rompimento da barragem de Fundão que foram contemplados no Termo de Transação e Ajustamento de Conduta [21]. O conjunto  $C_A$  é composto por  $A_1 =$  Taxa mensal de atendimentos por dengue e por  $A_2 =$  Produto Interno Bruto (PIB) municipal anual.

Dados mensais de atendimentos por dengue durante o período de 2010 a 2016 foram extraídos da plataforma Tabnet [22]. Por meio desta plataforma, o DATASUS disponibiliza informações para subsidiar análises da situação sanitária, tomadas de decisão baseadas em evidências e elaboração de programas de ações de saúde [22]. Os dados são então convertidos em taxa de atendimentos por mil habitantes utilizando estimativas de população do Instituto Brasileiro de Geografia e Estatística (IBGE) [23].

Por sua vez, os dados de PIB municipal dos anos de 2015 e 2016 para os municípios de Minas Gerais foram extraídos do Cidades@, o sistema agregador de informações do IBGE sobre os municípios e estados do Brasil [24]. Além disso, uma base de dados secundária contendo o Índice de Desenvolvimento Humano Municipal (IDHm) para os



municípios de Minas gerais foi obtida em [25], os valores mais recentes datam de 2010.

### 5.1.1 Definição dos Controles

Para a definição dos controles foram definidos como atributos a taxa de atendimento anual por dengue e o valor de IDHm. Essas variáveis foram selecionadas para incluir o efeito da sazonalidade dos casos de dengue na escolha dos controles assim como para considerar o porte socioeconômico do município. A metodologia para determinação dos controles foi aplicada e o resultado é exibido na Tabela 5.1.

**Tabela 5.1** – Localidades mais similares às atingidas, ao considerar taxas de atendimento anual por dengue e IDHm, obtidas por meio de *clusterização fuzzy* – são exibidas até as segundas localidades mais similares para simplificação da tabela.

<b>Localidade Atingida</b>	<b>1ª Mais Similar</b>	<b>2ª Mais Similar</b>
Aimorés	Quartel Geral	São Gonçalo do Pará
Alpercata	Morro da Garça	Santa Maria de Itabira
Barra Longa	Caparaó	Rio Pardo de Minas
Belo Oriente	Elói Mendes	Itamarati de Minas
Bom Jesus do Galho	Cônego Marinho	Jacinto
Bugre	Riacho dos Machados	Mendes Pimentel
Caratinga	Santa Juliana	Santa Bárbara
Conselheiro Pena	Lajinha	Januária
Córrego Novo	Lagoa dos Patos	Martins Soares
Dionísio	Guaranésia	Presidente Olegário
Fernandes Tourinho	Capitão Enéas	Machacalis
Galiléia	Guarará	Santo Hipólito
Governador Valadares	Além Paraíba	Santo Antônio do Monte
Iapu	Carmo da Cachoeira	Capelinha
Ipaba	Miradouro	Volta Grande
Ipatinga	Uberaba	Lavras
Itueta	Oratórios	Conceição do Mato Dentro
Mariana	Caxambu	Bicas
Marliéria	Monte Azul	Brasília de Minas
Naque	Centralina	Jaguaraçu
Periquito	Gameleiras	Rio Manso
Pingo-d'Água	Ibiaí	Malacacheta
Ponte Nova	Ibiá	Ijaci
Raul Soares	Pescador	Águas Formosas
Resplendor	Claro dos Poções	Água Comprida
Rio Casca	São João do Oriente	Itanhomi
Rio Doce	São Brás do Suaçuí	Mercês
Santa Cruz do Escalvado	São João do Pacuí	Piedade dos Gerais
Santana do Paraíso	Antônio Prado de Minas	Pimenta
São Domingos do Prata	Manhuaçu	Astolfo Dutra
São José do Goiabal	Uruana de Minas	Chácara
São Pedro dos Ferros	Conceição do Rio Verde	Inimutaba
Sem-Peixe	Abre Campo	Cana Verde
Sobrália	Urucânia	Tarumirim
Timóteo	Sete Lagoas	Coronel Fabriciano
Tumiritinga	Santana de Pirapama	Espinosa

Como não houve localidades repetidas na lista dos municípios mais similares, estes

formaram o conjunto das localidades de controle.

### 5.1.2 Análise Univariável

Os números de atendimento por dengue, antes e após o evento, para os municípios atingidos e seus controles são exibidos na Tabela 5.2. Em um comportamento normal, esperava-se que todos os municípios apresentassem maior porcentagem de atendimentos antes do evento (Nov/2010 a Out/2015) do que após (Nov/2015 a Out/2016); porém, é observado que alguns municípios como Mariana apresentaram alta porcentagem após o evento, o que pode indicar um provável impacto. Além disso, alguns municípios apresentaram quantidade de atendimentos extremamente baixas, como Barra Longa que apresentou apenas 1 atendimento antes do evento e seu controle, Caparaó, que apresentou apenas 1 atendimento para o total de 6 anos estudados, o que pode indicar um possível problema no registro de atendimentos daquela cidade.

**Tabela 5.2** – Número de atendimentos observados antes e após o evento para os municípios atingidos e seus controles.

Localidade	Número de atendimentos		Total
	Antes do evento (Nov/2010 - Out/2015)	Depois do Evento (Nov/2015 - Out/2016)	
Abre Campo	253 (48,47%)	269 (51,53%)	522
Aimorés	1861 (89,26%)	223 (10,70%)	2085
Além Paraíba	1289 (42,44%)	1747 (57,52%)	3037
Alpercata	55 (98,21%)	1 (1,79%)	56
Antônio Prado de Minas	29 (76,32%)	9 (23,68%)	38
Barra Longa	1 (0,58%)	171 (99,42%)	172
Belo Oriente	204 (41,89%)	283 (58,11%)	487
Bom Jesus do Galho	126 (64,62%)	69 (35,38%)	195
Bugre	37 (36,63%)	64 (63,37%)	101
Caparaó	1 (100%)	0 (0%)	1
Capitão Enéas	1120 (52,68%)	1005 (47,27%)	2126
Caratinga	831 (34,61%)	1569 (65,35%)	2401
Carmo da Cachoeira	60 (57,14%)	44 (41,90%)	105
Caxambu	105 (54,97%)	86 (45,03%)	191
Centralina	539 (98,54%)	8 (1,46%)	547
Claro dos Poções	239 (70,09%)	102 (29,91%)	341
Conceição do Rio Verde	607 (75,50%)	197 (24,50%)	804
Cônego Marinho	67 (84,81%)	12 (15,19%)	79
Conselheiro Pena	388 (60,44%)	254 (39,56%)	642
Córrego Novo	66 (66,00%)	34 (34,00%)	100
Dionísio	128 (57,40%)	95 (42,60%)	223
Elói Mendes	175 (90,67%)	17 (8,81%)	193
Fernandes Tourinho	206 (97,17%)	6 (2,83%)	212
Galiléia	87 (98,86%)	1 (1,14%)	88

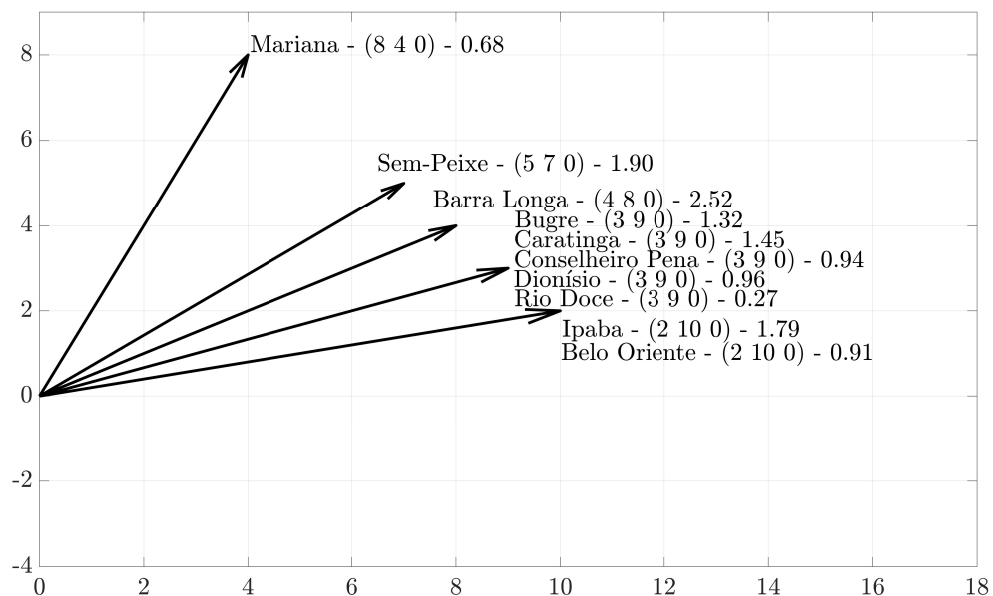
Tabela 5.2 continuada da página anterior

Localidade	Número de atendimentos		Total
	Antes do evento (Nov/2010 - Out/2015)	Depois do Evento (Nov/2015 - Out/2016)	
Gameleiras	136 (93,79%)	9 (6,21%)	145
Governador Valadares	10279 (90,80%)	1035 (9,14%)	11320
Guaranésia	289 (30,71%)	652 (69,29%)	941
Guarará	56 (36,84%)	96 (63,16%)	152
Iapu	73 (85,88%)	12 (14,12%)	85
Ibiá	1029 (89,09%)	126 (10,91%)	1155
Ibiaí	501 (92,61%)	40 (7,39%)	541
Ipaba	445 (53,23%)	389 (46,53%)	836
Ipatinga	18007 (75,34%)	5884 (24,62%)	23901
Itueta	217 (82,51%)	46 (17,49%)	263
Lagoa dos Patos	80 (72,07%)	31 (27,93%)	111
Lajinha	291 (53,49%)	253 (46,51%)	544
Manhuaçu	2381 (56,89%)	1804 (43,11%)	4185
Mariana	339 (41,14%)	485 (58,86%)	824
Marliéria	89 (64,03%)	50 (35,97%)	139
Miradouro	275 (63,07%)	161 (36,93%)	436
Monte Azul	460 (69,70%)	200 (30,30%)	660
Morro da Garça	22 (9,28%)	215 (90,72%)	237
Naque	370 (99,46%)	1 (0,27%)	372
Oratórios	172 (83,90%)	33 (16,10%)	205
Periquito	203 (62,08%)	124 (37,92%)	327
Pescador	235 (94,00%)	15 (6,00%)	250
Pingo-d'Água	262 (71,98%)	102 (28,02%)	364
Ponte Nova	2528 (82,51%)	533 (17,40%)	3064
Quartel Geral	261 (88,18%)	35 (11,82%)	296
Raul Soares	1375 (89,58%)	160 (10,42%)	1535
Resplendor	553 (68,10%)	259 (31,90%)	812
Riacho dos Machados	97 (95,10%)	5 (4,90%)	102
Rio Casca	463 (81,80%)	102 (18,02%)	566
Rio Doce	3 (27,27%)	8 (72,73%)	11
Santa Cruz do Escalvado	13 (61,90%)	8 (38,10%)	21
Santa Juliana	120 (47,62%)	132 (52,38%)	252
Santana de Pirapama	294 (49,33%)	302 (50,67%)	596
Santana do Paraíso	490 (56,32%)	380 (43,68%)	870
São Brás do Suaçuí	5 (33,33%)	10 (66,67%)	15
São Domingos do Prata	492 (78,97%)	131 (21,03%)	623
São João do Oriente	246 (50,00%)	246 (50,00%)	492
São João do Pacuí	11 (61,11%)	7 (38,89%)	18
São José do Goiabal	359 (67,35%)	174 (32,65%)	533
São Pedro dos Ferros	409 (76,16%)	128 (23,84%)	537
Sem-Peixe	57 (47,90%)	62 (52,10%)	119
Sete Lagoas	23197 (66,07%)	11914 (33,93%)	35111

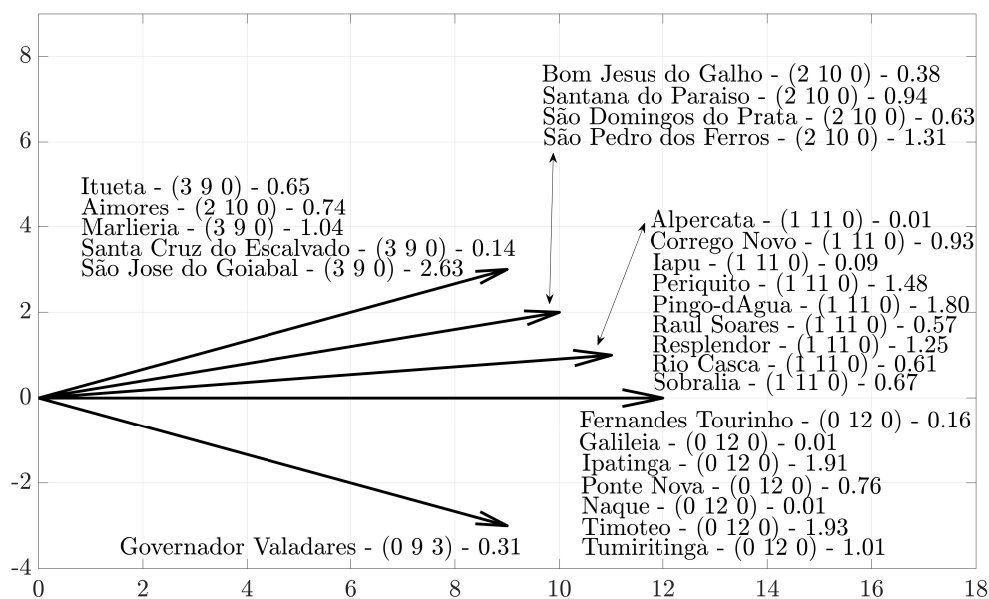
Tabela 5.2 continuada da página anterior

Localidade	Número de atendimentos		Total
	Antes do evento (Nov/2010 - Out/2015)	Depois do Evento (Nov/2015 - Out/2016)	
Sobrália	132 (74,16%)	46 (25,84%)	178
Timóteo	7945 (79,73%)	2020 (20,27%)	9965
Tumiritinga	248 (75,61%)	80 (24,39%)	328
Uberaba	19334 (88,78%)	2436 (11,19%)	21778
Uruana de Minas	205 (80,08%)	51 (19,92%)	256
Urucânia	251 (46,06%)	294 (53,94%)	545

Para a auto-comparação, o Gráfico de Resultantes Vetoriais foi aplicado aos municípios atingidos e os resultados podem ser visualizados nas Figuras 5.1, municípios com média anual acima do esperado, e 5.2, demais municípios. Os rótulos entre parênteses contêm a quantidade de meses acima, dentro e abaixo do esperado, respectivamente; e separado por um travessão é exibido valor da média anual da taxa de atendimentos de dengue. Destaca-se que por a média dos atendimentos antes do evento ser tomada mês a mês, o efeito da sazonalidade de dengue é considerado.



**Figura 5.1** – Gráfico de resultantes vetoriais – Municípios com média anual acima do esperado.



**Figura 5.2** – Gráfico de resultantes vetoriais – Demais municípios estudados.

Destacam-se os municípios de Barra Longa, Sem-peixe e Mariana por apresentarem alto número de meses com taxas de atendimento acima do esperado. A maior parte dos demais municípios apresentou taxas dentro dos limites históricos ou com poucos meses acima da taxa esperada. Contudo, o município de Governador Valadares apresentou uma diminuição da taxa de atendimentos por dengue em 3 dos 12 meses estudados.

A segunda parte da análise compreende os testes estatísticos e comparação com os controles. Os resultados são exibidos na Tabela 5.3.

**Tabela 5.3** – Resultante do teste *t* pareado para os municípios atingidos e seus respectivos controles.

Municípios <sup>a</sup>	Dif. das Médias <sup>b</sup>	Erro Padrão	<i>p</i> -valor <sup>c</sup>
<b>Municípios Atingidos</b>			
Aimorés	-0,4854	0,2598	0,0886
Alpercata	-0,1140	0,0543	0,0595
Barra Longa	2,5221	1,3895	0,0968
Belo Oriente	0,7697	0,6342	0,2503
Bom Jesus do Galho	0,2453	0,2860	0,4094
Bugre	1,1655	0,8634	0,2042
Caratinga	1,2934	0,7737	0,1228
Conselheiro Pena	0,6501	0,3125	0,0617
Córrego Novo	0,5747	0,7612	0,4661
Dionísio	0,7163	0,4489	0,1388
Fernandes Tourinho	-0,9471	0,2583	<b>0,0037</b>
Galiléia	-0,1956	0,0773	<b>0,0280</b>

Tabela 5.3 continuada da página anterior

Municípios	Dif. das Médias	Erro Padrão	p-valor
Governador Valadares	-0,3171	0,1036	<b>0,0108</b>
Iapu	-0,0212	0,1160	0,8586
Ipaba	1,3650	0,9002	0,1576
Ipatinga	0,7038	0,6918	0,3308
Itueta	0,0377	0,2779	0,8947
Mariana	0,5808	0,2786	0,0612
Marliéria	0,6712	0,3876	0,1113
Naque	-0,9213	0,4463	0,0634
Periquito	1,0060	1,0580	0,3621
Pingo-d'Água	0,8471	1,1232	0,4666
Ponte Nova	0,0379	0,1262	0,7693
Raul Soares	-0,3939	0,2726	0,1764
Resplendor	0,7186	0,3368	0,0562
Rio Casca	0,0675	0,2151	0,7595
Rio Doce	0,2483	0,1505	0,1271
Santa Cruz do Escalvado	0,0943	0,0617	0,1543
Santana do Paraíso	0,6688	0,4482	0,1638
São Domingos do Prata	0,1633	0,1140	0,1796
São José do Goiabal	1,5570	1,5437	0,3349
São Pedro dos Ferros	0,5042	0,4102	0,2447
Sem-Peixe	1,5593	1,0412	0,1624
Sobrália	0,2971	0,1698	0,1079
Timóteo	0,3685	0,5892	0,5445
Tumiritinga	0,3734	0,2698	0,1937
Municípios de Controle			
Quartel Geral	-0,4213	0,7025	0,5608
Morro da Garça	6,7341	4,2682	0,1429
Caparaó	-0,0031	0,0031	0,3388
Elói Mendes	-0,0562	0,0469	0,2552
Cônego Marinho	-0,0173	0,0700	0,8087
Riacho dos Machados	-0,1256	0,0927	0,2027
Santa Juliana	0,6531	0,3193	0,0655
Lajinha	0,8176	0,4702	0,1099
Lagoa dos Patos	0,2967	0,4529	0,5259
Guaranésia	2,6151	1,5070	0,1106
Capitão Enéas	4,3295	2,0320	0,0565
Guarará	1,8132	1,0811	0,1217
Além Paraíba	3,5573	1,8385	0,0791
Carmo da Cachoeira	0,2197	0,1540	0,1815
Miradouro	0,8330	0,4637	0,0999
Uberaba	-0,4066	0,1147	<b>0,0046</b>
Oratórios	-0,0293	0,3563	0,9359
Caxambu	0,2528	0,1311	0,0800
Monte Azul	0,4343	0,2473	0,1068

Tabela 5.3 continuada da página anterior

Municípios	Dif. das Médias	Erro Padrão	p-valor
Centralina	-0,7971	0,2839	<b>0,0171</b>
Gameleiras	-0,2926	0,2223	0,2149
Ibiaí	-0,6248	0,5710	0,2972
Ibiá	-0,2785	0,1517	0,0936
Pescador	-0,6352	0,2184	<b>0,0142</b>
Claro dos Poções	0,5979	0,6297	0,3627
São João do Oriente	2,1377	1,2451	0,1140
São Brás do Suaçuí	0,2037	0,1129	0,0987
São João do Pacuí	0,0926	0,0723	0,2269
Antônio Prado de Minas	0,1704	0,3035	0,5859
Manhuaçu	1,2340	0,5470	<b>0,0454</b>
Uruana de Minas	0,2642	0,2863	0,3759
Conceição do Rio Verde	0,4671	0,2715	0,1133
Abre Campo	1,3641	0,8493	0,1365
Urucânia	1,9652	1,5459	0,2299
Sete Lagoas	2,5435	1,4226	0,1013
Santana de Pirapama	2,6167	1,5295	0,1151

<sup>a</sup> Municípios ordenados para que a ordem dos Municípios de Controle corresponda à ordem dos Municípios Atingidos.

<sup>b</sup> Uma diferença de média positiva indica que houve aumento da média da taxa de atendimentos.

<sup>c</sup>  $p$ -valores menores que 0,05 são destacados em negrito.

Os municípios de Fernandes Tourinho, Galiléia e Governador Valadares apresentaram diferença significativa entre as médias enquanto seus respectivos controles não apresentaram. Como a diferença das médias possui sinal negativo, isto indica que houve uma diminuição da taxa de atendimento por dengue nestes municípios, o que corrobora com os resultados observados na Figura 5.2. Os controles de Ipatinga, Naque e Raul Soares apresentaram diferença negativa e significativa, o que pode indicar que estas localidades atingidas permaneceram com nível de atendimentos estável enquanto que seus prováveis comportamentos, não havendo a ocorrência do evento, seriam de diminuir a taxa de atendimentos. Além disso, alguns municípios como Conselheiro Pena e Mariana apresentaram  $p$ -valor próximo ao nível de significância de 0,05 utilizado como referência. Assim, pode-se inferir que o evento impactou Conselheiro Pena e Mariana no sentido de aumento da média de atendimentos por dengue, corroborando com os resultados da Figura 5.1.

### 5.1.3 Análise Multivariável

A análise multivariada foi empregada em um conjunto de dados composto pelos atributos  $C_A$ , que foi normalizado para que os valores de cada variável estivessem no intervalo  $[0, 1]$ . A quantidade  $c = 3$  grupos foi obtida com  $r_a = 0, 1$ , e as pertinências estão apresentadas na Tabela 5.4. A Tabela 5.5 explica o significado de cada grupo em termos

de níveis das variáveis. Os centros dos grupos foram transformados de valor numérico para valor linguístico por meio da seguinte conversão: foi atribuído “Nível baixo” para valores no intervalo  $[0, 0,3)$ , “Nível médio” para valores no intervalo  $[0,3, 0,6)$  e “Nível alto” para valores em  $[0,6, 1]$ .

**Tabela 5.4** – Pertinências *fuzzy* antes e após a ocorrência do evento, considerando os atributos Taxa de atendimento por dengue e PIB municipal.

Municípios	Período	Pertinências aos Grupos		
		1	2	3
Aimorés	Pré-Evento	0,0983	<b>0,7787</b>	0,1230
	Pós-Evento	0,0915	<b>0,8937</b>	0,0148
Alpercata	Pré-Evento	<b>0,9732</b>	0,0229	0,0039
	Pós-Evento	<b>0,9305</b>	0,0579	0,0116
Barra Longa	Pré-Evento	<b>0,9257</b>	0,0618	0,0125
	Pós-Evento	0,1183	0,2297	<b>0,6520</b>
Belo Oriente	Pré-Evento	<b>0,8124</b>	0,1574	0,0302
	Pós-Evento	0,0727	<b>0,8856</b>	0,0417
Bom Jesus do Galho	Pré-Evento	<b>0,9807</b>	0,0165	0,0028
	Pós-Evento	<b>0,9051</b>	0,0860	0,0089
Bugre	Pré-Evento	<b>0,9774</b>	0,0194	0,0032
	Pós-Evento	0,1312	<b>0,6842</b>	0,1845
Caratinga	Pré-Evento	<b>0,7966</b>	0,1710	0,0324
	Pós-Evento	0,0847	0,3577	<b>0,5576</b>
Conselheiro Pena	Pré-Evento	<b>0,9888</b>	0,0100	0,0013
	Pós-Evento	0,0370	<b>0,9487</b>	0,0142
Córrego Novo	Pré-Evento	<b>0,9272</b>	0,0654	0,0074
	Pós-Evento	0,0771	<b>0,8961</b>	0,0268
Dionísio	Pré-Evento	<b>0,9873</b>	0,0112	0,0016
	Pós-Evento	0,0707	<b>0,9010</b>	0,0283
Fernandes Tourinho	Pré-Evento	0,0972	<b>0,8357</b>	0,0671
	Pós-Evento	<b>0,9789</b>	0,0181	0,0029
Galiléia	Pré-Evento	<b>0,9899</b>	0,0088	0,0013
	Pós-Evento	<b>0,9306</b>	0,0578	0,0115
Governador Valadares	Pré-Evento	0,2985	<b>0,3636</b>	0,3379
	Pós-Evento	0,3511	<b>0,3625</b>	0,2864
Iapu	Pré-Evento	<b>0,9718</b>	0,0241	0,0042
	Pós-Evento	<b>0,9648</b>	0,0299	0,0053
Ipaba	Pré-Evento	<b>0,8402</b>	0,1462	0,0136
	Pós-Evento	0,1152	0,3309	<b>0,5539</b>



Tabela 5.4 continuada da página anterior

Municípios	Período	Pertinências aos Grupos		
		1	2	3
Ipatinga	Pré-Evento	0,2431	<b>0,3033</b>	0,4535
	Pós-Evento	0,2008	<b>0,2708</b>	0,5284
Itueta	Pré-Evento	0,4043	<b>0,5663</b>	0,0294
	Pós-Evento	0,3194	<b>0,6515</b>	0,0292
Mariana	Pré-Evento	<b>0,5094</b>	0,3528	0,1377
	Pós-Evento	0,2835	<b>0,6018</b>	0,1147
Marliéria	Pré-Evento	<b>0,9155</b>	0,0762	0,0084
	Pós-Evento	0,0839	<b>0,8712</b>	0,0449
Naque	Pré-Evento	0,0709	<b>0,9035</b>	0,0255
	Pós-Evento	<b>0,9304</b>	0,0580	0,0116
Periquito	Pré-Evento	<b>0,7283</b>	0,2514	0,0203
	Pós-Evento	0,1359	<b>0,5450</b>	0,3191
Pingo-d'Água	Pré-Evento	0,0764	<b>0,8947</b>	0,0290
	Pós-Evento	0,1203	0,3400	<b>0,5397</b>
Ponte Nova	Pré-Evento	0,1624	<b>0,7962</b>	0,0414
	Pós-Evento	0,1464	<b>0,8066</b>	0,0470
Raul Soares	Pré-Evento	0,0409	<b>0,9418</b>	0,0173
	Pós-Evento	0,4846	<b>0,4889</b>	0,0264
Resplendor	Pré-Evento	<b>0,5879</b>	0,3874	0,0247
	Pós-Evento	0,1106	<b>0,7515</b>	0,1378
Rio Casca	Pré-Evento	<b>0,5583</b>	0,4158	0,0259
	Pós-Evento	0,3828	<b>0,5892</b>	0,0280
Rio Doce	Pré-Evento	<b>0,9316</b>	0,0571	0,0113
	Pós-Evento	<b>0,9784</b>	0,0191	0,0026
Santa Cruz do Escalvado	Pré-Evento	<b>0,9421</b>	0,0486	0,0093
	Pós-Evento	<b>0,9752</b>	0,0213	0,0036
Santana do Paraíso	Pré-Evento	<b>0,9846</b>	0,0136	0,0018
	Pós-Evento	0,0128	<b>0,9817</b>	0,0055
São Domingos do Prata	Pré-Evento	<b>0,7461</b>	0,2352	0,0187
	Pós-Evento	0,3220	<b>0,6510</b>	0,0270
São José do Goiabal	Pré-Evento	0,0877	<b>0,8598</b>	0,0524
	Pós-Evento	0,1242	0,2329	<b>0,6429</b>
São Pedro dos Ferros	Pré-Evento	0,0960	<b>0,8835</b>	0,0205
	Pós-Evento	0,1265	<b>0,6898</b>	0,1837
Sem-Peixe	Pré-Evento	<b>0,9361</b>	0,0573	0,0066
	Pós-Evento	0,1148	0,2998	<b>0,5854</b>

Tabela 5.4 continuada da página anterior

Municípios	Período	Pertinências aos Grupos		
		1	2	3
Sobralia	Pré-Evento	<b>0,9028</b>	0,0879	0,0094
	Pós-Evento	0,2779	<b>0,6930</b>	0,0292
Timóteo	Pré-Evento	0,0385	0,1038	<b>0,8577</b>
	Pós-Evento	0,0024	0,0056	<b>0,9921</b>
Tumiritinga	Pré-Evento	0,3548	<b>0,6153</b>	0,0299
	Pós-Evento	0,0784	<b>0,8836</b>	0,0380

Tabela 5.5 – Significado linguístico dos grupos obtidos na análise multivariável.

Variável	Grupo 1	Grupo 2	Grupo 3
Dengue	Baixo	Médio	Alto
PIB	Baixo	Baixo	Médio

A maior parte dos municípios da Tabela 5.4 apresentou uma mudança de comportamento no sentido de aumento dos níveis de atendimento por dengue. Ressalta-se que alguns municípios como Caratinga, Ipaba e Sem-peixe apresentaram mudança de grupo para um de nível superior no PIB; contudo, com níveis de pertinência baixos ( $\simeq 0,55$ ) o que pode indicar que o aumento não muito expressivo. O município de Fernandes Tourinho apresenta diminuição do nível de atendimentos, o que corrobora com análises anteriores.

Pela Tabela 5.5, infere-se uma possível interação entre as variáveis estudadas. O Grupo 3 é o único que apresenta uma diferença nos níveis de PIB, de baixo para médio, e está associado a um nível alto de atendimentos por dengue. Assim, verifica-se que municípios com maior PIB normalmente apresentam maiores níveis de atendimento por dengue.

#### 5.1.4 Conclusões

A nova metodologia para avaliação de impactos de eventos foi aplicada ao “Rompiemento da barragem de Fundão em Mariana–MG”, que configura um desastre tecnológico. As análises uni e multivariáveis foram aplicadas para avaliar os atributos “Taxa mensal de atendimentos por dengue” e “PIB municipal anual”. O algoritmo *Fuzzy c-means* foi capaz de auxiliar a seleção de localidades de controle baseando-se em similaridade *fuzzy* e dentre os resultados obtidos destacam-se os municípios de Mariana e Conselheiro Pena, pois apresentaram aumento nos níveis de atendimento em quase todas as análises.

## 5.2 Formulação 2 – Impacto da pandemia da COVID-19 nos atendimentos para acompanhamento do crescimento e desenvolvimento de crianças no Brasil

Em fevereiro de 2020, foi registrado o primeiro caso da COVID-19 no Brasil, causada pelo vírus SARS-CoV-2, e atualmente, junho de 2022, mais de 32 milhões de casos e mais de 670 mil mortes no país [26]. Concorrente ao número de infectados e de mortos, tem-se o impacto econômico, social, cultural, político e na saúde pública provocado por essa situação.

Embora os impactos definitivos da pandemia nos sistemas de saúde ainda não tenham sido revelados, em muitos países efeitos têm sido apontados, com destaque para redução na utilização de serviços de saúde para cuidados eletivos, incluindo redução nas taxas de atendimentos clínicos individuais a crianças em serviços de atenção primária [27, 28, 29, 30, 31, 32, 33]. Atendimentos para o acompanhamento do crescimento e desenvolvimento (ACD) de crianças também foram comprometidos durante a pandemia [34, 35, 36].

A princípio, tanto fatores político-organizacionais e de saúde pública quanto decisões individuais contribuíram para esses impactos nos atendimentos. Do ponto de vista individual, o medo de contrair a doença pode ter sido decisivo nas intenções de procura por cuidados, e conseqüentemente na utilização dos serviços [31, 33, 37, 38]. Do ponto de vista político-organizacional e de saúde pública, as medidas para controlar a propagação da contaminação e garantir resposta aos casos mais graves convergiram para desestimular a oferta e a procura por cuidados de rotina e eletivos, incluindo os atendimentos nos programas de ACD de crianças nos serviços de atenção primária à saúde (APS) [39, 40, 41].

O ACD, faz parte de um dos sete eixos estratégicos da Política Nacional de Atenção Integral à Saúde da Criança, além de ser uma das ações que contribuem para o alcance de desafios globais como os Objetivos do Desenvolvimento Sustentável [42]. Na prática, se constitui de atendimentos periódicos nos quais são realizadas ações para a promoção da saúde, aleitamento materno, desenvolvimento, imunização, rastreamento de condições patológicas, prevenção de acidentes e o acompanhamento do crescimento e do peso corporal de acordo com a idade das crianças, e em muitas situações favorecem o acesso ao diagnóstico de doenças agudas e também crônicas [43].

À semelhança de outros cuidados eletivos na APS, é esperado que os atendimentos para ACD de crianças tenham sofrido redução. Contudo, ainda não se tem conhecimento da dimensão desse impacto nem da sua distribuição regional. Neste sentido, o objetivo deste estudo foi analisar o impacto da pandemia da COVID-19 no número de atendimentos para ACD de crianças nos serviços de APS no Brasil nos primeiros 12 meses de pandemia.

### 5.2.1 Tratamento de Dados

Este é um estudo descritivo e analítico, com delineamento ecológico transversal, que utiliza dados do Sistema de Informação em Saúde para a Atenção Básica (SISAB), do Ministério da Saúde brasileiro. O SISAB é de utilização obrigatória em todo o país desde junho de 2015 e integra a estratégia e-SUS Atenção Básica (e-SUS AB) [44].

Foram considerados dados sobre o número mensal de atendimentos individuais para ACD de crianças menores de 5 anos de idade realizados nos serviços de APS de todo o Brasil no período de abril de 2017 a março de 2021. A extração dos dados foi realizada em junho de 2021 de forma automatizada e diretamente do SISAB, pelo processo conhecido como *web scrapping* ou *data scrapping* [45, 46]. O processo de extração foi realizado utilizando o *software* Node.js, com um código em linguagem JavaScript para acessar a página <<https://sisab.saude.gov.br>>, preencher os formulários segundo protocolo previamente definido e fazer o download dos relatórios referentes a problemas ou condições avaliadas pelos profissionais de saúde em cada um dos meses incluídos no período estudado. Dados do Distrito Federal referentes a abril, maio e junho de 2017 não constavam no SISAB e por isso não foram estudados.

Para este estudo de caso considera-se que todas as localidades foram atingidas. Assim, define-se  $C_T$  como composto por todos os 26 estados do Brasil mais o Distrito Federal ( $N_T = 27$ ), estes que podem ser agrupados por regiões. Foram observadas as frequências absolutas e relativas dos atendimentos. As taxas de atendimentos para ACD foram calculadas para cada mil habitantes menores de 5 anos considerando estimativas populacionais por faixas etárias para o Brasil, regiões, estados e Distrito Federal [47]. Desse modo, conjunto  $C_A$  é composto por apenas  $A_1 =$  Taxa mensal de atendimentos de puericultura. Conseqüentemente, a análise multivariada não se aplica.

Por este se tratar de um estudo que utiliza dados de domínio público, de acesso irrestrito e sem a identificação de indivíduos, a apreciação foi dispensada pelo Comitê de Ética em Pesquisas da Universidade Federal de Minas Gerais (CAAE: 46914221.5.0000.5149).

### 5.2.2 Análises e Discussão

Foram contemplados 39.599.313 atendimentos para ACD de crianças ocorridos entre abril de 2017 e março de 2021 (Tabela 5.6). Atendimentos entre abril de 2017 e março de 2020 foram considerados como antes da pandemia, e atendimentos entre abril de 2020 e março de 2021 como durante a pandemia. A proporção de atendimentos estudados ocorridos durante a pandemia variou de 13,5% na região Nordeste a 21,8% na região Sul. Dentre os estados, variou de 11,9% no Ceará a 25,0% no Rio Grande do Sul.

**Tabela 5.6** – Distribuição do número de atendimentos para acompanhamento do crescimento e do desenvolvimento estudados (x 1.000) nos períodos pré (abril de 2017 a março de 2020) e durante a pandemia (abril de 2020 a março de 2021), segundo Regiões, Estados e Distrito Federal.

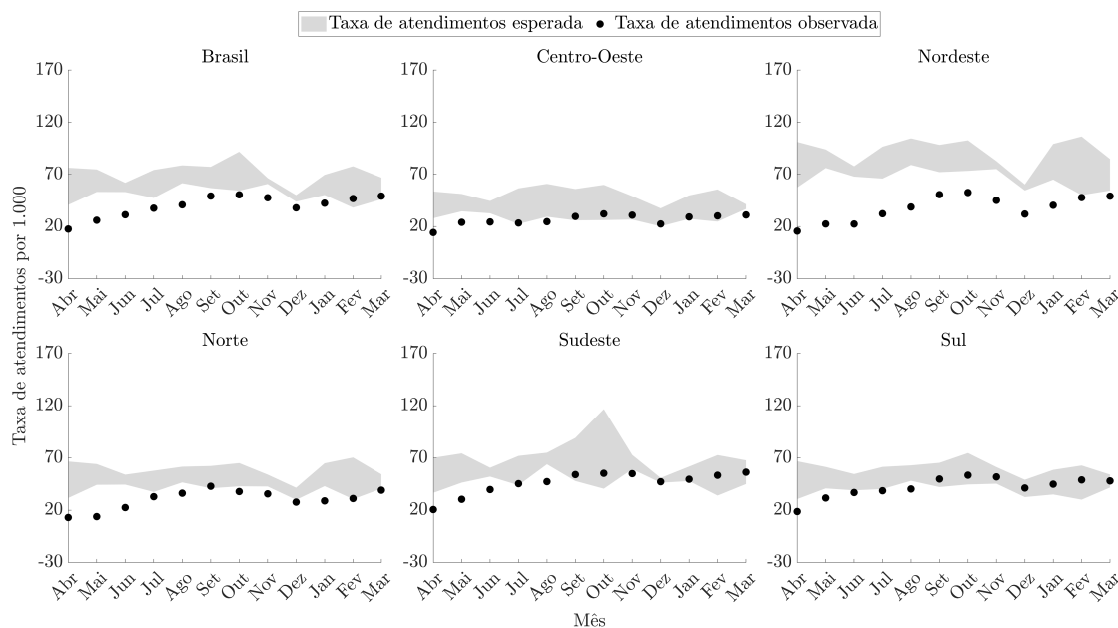
Localidade	Número de atendimentos estudados (x1.000)		
	Antes da pandemia	Durante a pandemia	Total
<b>Brasil</b>	32.586,13 (82,3%)	7.013,18 (17,7%)	39.599,31
<b>Centro-Oeste</b>	1.717,84 (81,6%)	386,29 (18,4%)	2.104,13
Distrito Federal	358,13 (79,3%)	93,59 (20,7%)	451,71
Goiás	583,12 (81,4%)	132,86 (18,6%)	715,98
Mato Grosso	484,82 (85,0%)	85,73 (15,0%)	570,54
Mato Grosso do Sul	291,77 (79,7%)	74,12 (20,3%)	365,90
<b>Nordeste</b>	11.832,95 (86,5%)	1.854,22 (13,5%)	13.687,17
Alagoas	983,10 (86,9%)	148,68 (13,1%)	1.131,78
Bahia	2.465,78 (87,7%)	345,92 (12,3%)	2.811,70
Ceará	1.770,85 (88,1%)	238,74 (11,9%)	2.009,59
Maranhão	1.264,01 (82,6%)	266,92 (17,4%)	1.530,92
Paraíba	1.085,42 (87,0%)	162,71 (13,0%)	1.248,13
Pernambuco	2.290,60 (85,6%)	384,99 (14,4%)	2.675,60
Piauí	943,30 (86,7%)	144,55 (13,3%)	1.087,86
Rio Grande do Norte	698,58 (86,7%)	107,60 (13,3%)	806,18
Sergipe	331,30 (86,0%)	54,11 (14,0%)	385,41
<b>Norte</b>	2.880,28 (83,2%)	580,92 (16,8%)	3.461,20
Acre	77,78 (82,3%)	16,71 (17,7%)	94,49
Amapá	78,18 (82,4%)	16,69 (17,6%)	94,87
Amazonas	716,55 (81,1%)	166,54 (18,9%)	883,09
Pará	1.520,76 (83,9%)	290,90 (16,1%)	1.811,66
Rondônia	180,20 (85,0%)	31,91 (15,0%)	212,11
Roraima	72,69 (82,4%)	15,51 (17,6%)	88,21
Tocantins	234,12 (84,6%)	42,66 (15,4%)	276,78
<b>Sudeste</b>	12.581,89 (79,8%)	3.194,09 (20,2%)	15.775,98
São Paulo	7.026,56 (77,9%)	1.992,16 (22,1%)	9.018,73
Espírito Santo	440,36 (82,5%)	93,28 (17,5%)	533,65
Minas Gerais	2.671,56 (80,4%)	650,43 (19,6%)	3.322,00
Rio de Janeiro	2.443,40 (84,2%)	458,21 (15,8%)	2.901,61
<b>Sul</b>	3.573,18 (78,2%)	997,66 (21,8%)	4.570,84
Paraná	1.339,47 (79,4%)	348,40 (20,6%)	1.687,87
Rio Grande do Sul	1.187,00 (75,0%)	395,59 (25,0%)	1.582,59
Santa Catarina	1.046,70 (80,5%)	253,67 (19,5%)	1.300,38

A taxa mensal média de atendimentos para ACD no Brasil (por mil habitantes) caiu de 61,34 antes da pandemia para 39,70 nos 12 primeiros meses da pandemia ( $p < 0,001$ ), uma queda de 35,28% (Tabela 5.7). Em todos os estados, exceto Rio Grande do Sul, houve redução significativa na taxa mensal de atendimentos, com diferenças variando de -14,21% em São Paulo a -59,66% no Ceará (Tabela 5.7).

**Tabela 5.7** – Diferenças entre as médias das taxas de atendimentos para acompanhamento do crescimento e do desenvolvimento de crianças menores de 5 anos de idade (por mil habitantes) nos períodos pré (abril de 2017 a março de 2020) e durante a pandemia da COVID-19 (abril de 2020 a março de 2021), segundo Regiões, Estados e Distrito Federal.

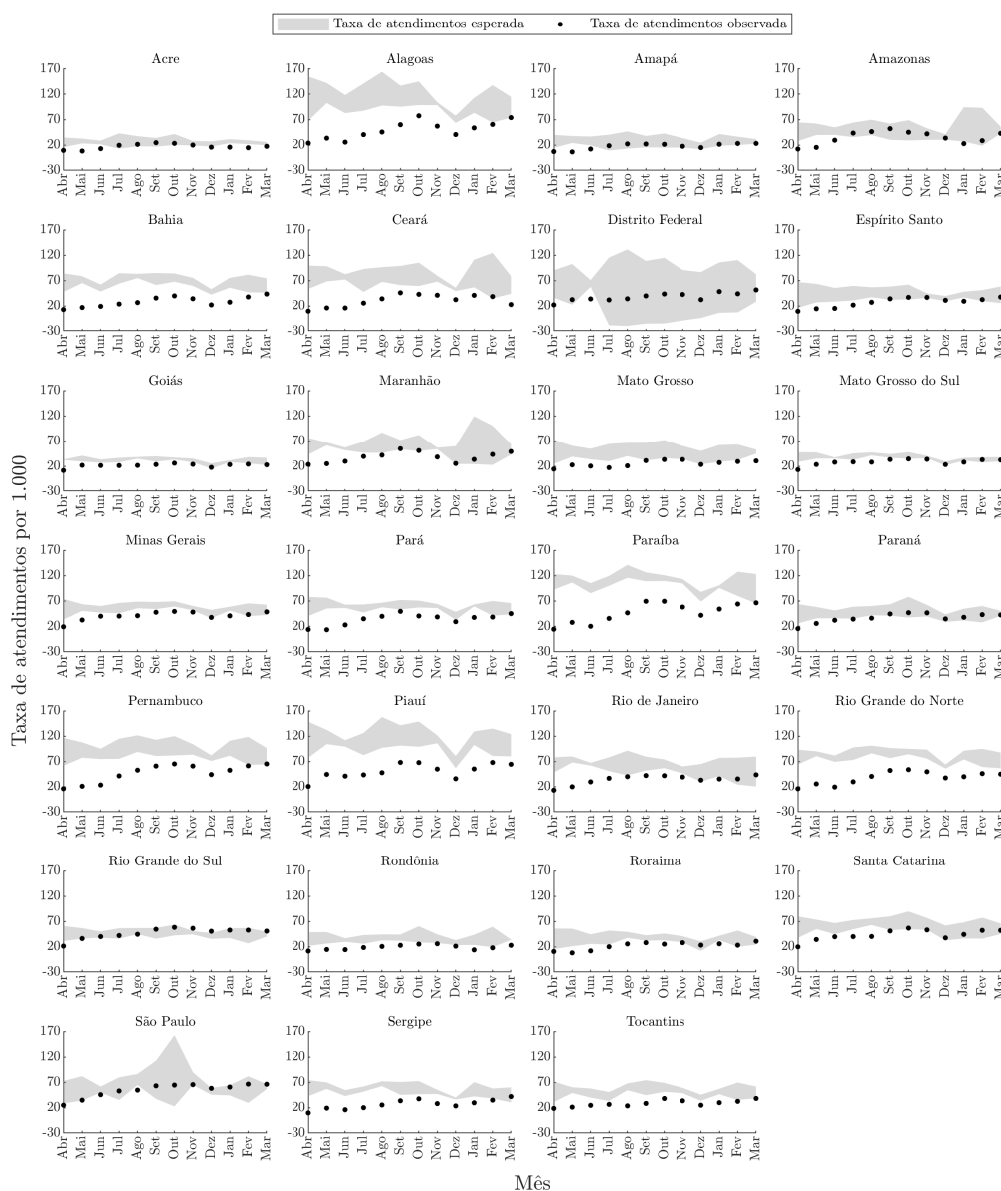
Localidades	Médias das taxas (x1.000 hab.)		Diferença entre as médias (%)	Erro Padrão	p-valor *
	Pré-pandemia	Durante a pandemia			
<b>Brasil</b>	61,34	39,70	-35,28	3,09	<0,001
<b>Centro-Oeste</b>	39,58	26,29	-33,57	1,78	<0,001
Distrito Federal	53,58	37,86	-29,35	3,71	0,001
Goiás	31,56	21,47	-31,96	1,37	<0,001
Mato Grosso	47,70	25,20	-47,15	2,04	<0,001
Mato Grosso do Sul	36,99	28,18	-23,82	2,17	0,002
<b>Nordeste</b>	79,16	37,37	-52,78	4,06	<0,001
Alagoas	106,95	49,25	-53,95	6,87	<0,001
Bahia	66,63	28,13	-57,78	3,53	<0,001
Ceará	75,14	30,31	-59,66	4,11	<0,001
Maranhão	59,68	37,88	-36,53	3,28	<0,001
Paraíba	105,98	47,46	-55,22	6,33	<0,001
Pernambuco	91,75	46,89	-48,90	5,15	<0,001
Piauí	110,13	50,77	-53,90	5,03	<0,001
Rio Grande do Norte	80,78	37,75	-53,27	3,99	<0,001
Sergipe	54,18	26,57	-50,97	3,97	<0,001
Tocantins	52,05	28,28	-45,67	2,37	<0,001
<b>Norte</b>	49,57	30,07	-39,34	3,08	<0,001
Acre	25,82	16,77	-35,05	1,53	<0,001
Amapá	26,87	17,37	-35,34	1,98	0,001
Amazonas	48,85	34,36	-29,66	4,47	0,008
Pará	58,57	33,79	-42,31	3,85	<0,001
Rondônia	35,96	18,99	-47,20	1,52	<0,001
Roraima	35,62	21,59	-39,39	2,78	<0,001
<b>Sudeste</b>	60,24	46,12	-23,44	3,27	0,001
São Paulo	63,79	54,73	-14,21	3,96	0,043
Espírito Santo	42,67	27,02	-36,67	3,07	<0,001
Minas Gerais	55,83	40,84	-26,85	2,45	<0,001
Rio de Janeiro	60,29	34,00	-43,61	4,12	<0,001
<b>Sul</b>	50,05	41,95	-16,18	2,80	0,015
Paraná	47,11	36,82	-21,84	2,45	0,002
Rio Grande do Sul	46,23	46,66	0,93	3,06	0,891
Santa Catarina	60,55	43,43	-28,28	3,22	<0,001

A Figura 5.3 exibe os diagramas de controle por regiões do Brasil. A taxa mensal média de atendimentos na região Nordeste foi a mais impactada, ficando abaixo do esperado nos 12 meses do primeiro ano da pandemia. No Centro-oeste, Sudeste e Sul as taxas mensais médias de atendimentos foram menos impactadas, ficando abaixo das esperadas por 5 meses durante os 12 primeiros da pandemia.



**Figura 5.3** – Taxas mensais de atendimentos (por 1.000 crianças menores de 5 anos de idade) para acompanhamento do crescimento e desenvolvimento de crianças nos serviços de Atenção Primária à Saúde no Brasil, segundo Regiões.

Em relação às taxas mensais de atendimentos para ACD por Estados e Distrito Federal, a Figura 5.4 apresenta os diagramas de controle e a Figura 5.5 exibe a versão simplificada em GRV. Infere-se que os maiores impactos foram identificados nos estados das Regiões Norte (Pará, Rondônia e Tocantins), Nordeste (Alagoas, Bahia, Ceará, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe) e Centro-Oeste (Mato Grosso do Sul), com taxas de atendimentos abaixo das esperadas por 11 ou mais meses durante a pandemia. No Rio Grande do Sul identificou-se o menor impacto nas taxas, tendo sido encontrado três meses abaixo do esperado e iguais três meses acima do esperado. Distrito Federal, Amapá e São Paulo também foram Estados onde a taxa de atendimentos sofreu menos impacto, tendo apresentado dois ou três meses com atendimentos abaixo do esperado. Cabe destacar que as mais intensas reduções no número de atendimentos foram observadas de abril a junho de 2020.



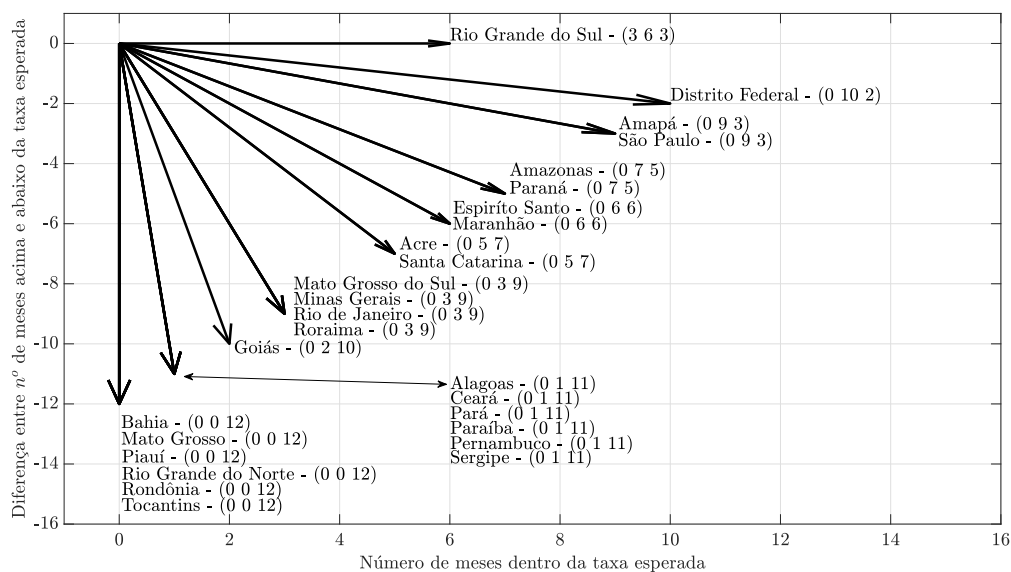
**Figura 5.4** – Taxas mensais de atendimentos (por 1.000 crianças menores de 5 anos de idade) para acompanhamento do crescimento e desenvolvimento de crianças nos serviços de Atenção Primária à Saúde no Brasil, segundo estados e Distrito Federal no Brasil.

Durante os 12 meses do primeiro ano da pandemia da COVID-19 a utilização de serviços de APS para o acompanhamento do crescimento e desenvolvimento de crianças foi significativamente menor que em anos anteriores no Brasil. Essa redução não foi uniforme e ocorreu de forma significativa em praticamente todas as Regiões e Estados. Reconhecendo que algum grau de redução era esperado, visto que a restrição de atendimentos eletivos foi uma das medidas para o controle da pandemia da COVID-19 no Brasil, destaca-se que este estudo avança ao apresentar importantes dimensões dessa redução, como a sua



distribuição geográfica, suas diferentes intensidades e duração no primeiro ano.

A redução média no número de atendimentos nos 12 primeiros meses da pandemia foi similar à encontrada por outros estudos realizados com períodos mais curtos (até os três primeiros meses) [36, 48, 49]. Corroborando também com outro estudo a redução foi mais intensa no período de abril a maio de 2020 [48]. Quanto às diferenças regionais na redução dos atendimentos, devemos considerar a interdependência e indissociabilidade dos aspectos políticos, econômicos e geográficos entre as Regiões brasileiras. Com sua grande extensão territorial, as diferenças e disparidades regionais no Brasil são muitas vezes agravadas por diferentes formas de comando político-econômico e pela força de trabalho em saúde disponível. Ainda a este respeito, em Ruanda um estudo também identificou diferenças regionais na redução de atendimentos a crianças durante a pandemia, em como no Brasil, as medidas de restrição e a capacidade instalada de serviços de saúde variam de região para região e não podemos desconsiderar o número de casos COVID-19 em cada área [50].



**Figura 5.5** – Gráfico de resultantes vetoriais das taxas de atendimentos (por 1.000 crianças menores de 5 anos de idade) para acompanhamento do crescimento e desenvolvimento de crianças nos serviços de Atenção Primária à Saúde, segundo estados e Distrito Federal no Brasil.

A respeito do número de meses durante os quais o número de atendimentos ficou abaixo do esperado antes de se recuperarem, em muitos estados brasileiros identificou-se que a retomada foi muito mais lenta do que observado em outros países [36, 49, 51]. Assim, supõe-se que a flexibilização das medidas de controle da pandemia não se refletiu de imediato na retomada dos atendimentos para ACD no Brasil. Nos locais em que a retomada foi mais rápida, os atendimentos híbridos, ou seja, presencial e virtual, foram uma importante estratégia [52].

As consultas regulares para o ACD de crianças, fortemente recomendadas a partir dos primeiros anos da década de 1980 como Política Pública, tiveram grande importância para a redução da mortalidade infantil no país [53]. Neste sentido, presume-se que a redução identificada no número de atendimentos poderá determinar impactos negativos nos indicadores de mortalidade [54]. Além disso, esses impactos poderão ampliar as já preocupantes desigualdades regionais nas taxas de mortalidade de crianças menores de cinco anos de idade, uma vez que as regiões Norte e Nordeste persistem com as taxas mais elevadas do país [53].

A redução nos atendimentos representa uma barreira aos diagnósticos de problemas com desenvolvimento infantil e para as intervenções precoces quando necessário. E isso ocorre em um cenário no qual a própria pandemia da COVID-19 carrega consigo o potencial de afetar profundamente o desenvolvimento de crianças [35, 55]. Com a redução, também poderá haver o aumento da insegurança alimentar, implicando em mais casos de desnutrição e obesidade [56, 57, 58].

É necessário considerar também que os atendimentos para o ACD representam oportunidade de cuidados regulares à saúde de crianças com condições crônicas. Assim, pressupõe-se que interrupções no acompanhamento poderão fragilizar o cuidado a crianças portadoras de condições crônicas, pois as visitas regulares aos profissionais de saúde permitem a identificação precoce de condições que comprometem o cuidado [59].

O contexto da pandemia e das medidas impostas pelas recomendações sanitárias impuseram novos determinantes e condições de saúde às crianças, com destaque para os efeitos na saúde mental, maior tempo de exposição a telas e jogos eletrônicos [60, 61, 62]. Com um declínio tão importante no acompanhamento regular das crianças, muitas das condições de ausência de saúde e de riscos podem não ter recebido o cuidado que demandariam. Ressalta-se o aumento nos casos de violência contra crianças durante a pandemia que, num contexto de redução do acesso às consultas regulares, podem não ter sido diagnosticados, uma vez que parte importante dos casos são diagnosticados durante consultas rotineiras não motivadas pelos atos em si [63].

Os resultados obtidos alertam para a necessidade do Sistema Único de Saúde se preparar para garantir o acompanhamento não presencial quando num contexto que impeça a proximidade física entre crianças e profissionais da APS. De um modo geral, o uso da Telessaúde foi uma importante estratégia para contornar as barreiras do distanciamento social imposto pela pandemia e favorecer o acesso de crianças aos cuidados de rotina em muitos cenários [64]. No entanto, no Brasil, requer mais investimentos em infraestrutura tecnológica e humana para sua implementação [62].

Dentre as limitações deste estudo, destaca-se que pesquisa foi realizada com dados secundários, ou seja, os dados analisados não foram coletados especificamente para responder à questão investigada. Além disso, as análises foram realizadas com dados agregados

por estados, regiões e país, o que não permite apreciação de eventuais diferenças entre municípios. Novos estudos que avaliem o impacto da pandemia no nível dos sistemas municipais de saúde serão importantes visto que muitas decisões durante a pandemia foram descentralizadas para os municípios. Novos estudos que considerem analisar os impactos no ACD sob a perspectiva das diferenças socioeconômicas também serão importantes, uma vez que as diferenças regionais encontradas no nosso estudo podem ter tido influência dessas desigualdades [52].

Com relação à metodologia, a análise descritiva apresentou a frequência absoluta e relativa dos atendimentos de puericultura nos estados, regiões e país. Neste primeiro momento já foi possível observar diferenças entre as frequências relativas dos estados, indicando uma possível redução de atendimentos. Em seguida, o emprego do teste  $t$  pareado permitiu observar se as reduções nas taxas mensais médias de atendimento eram estatisticamente significativas, o que foi revelado verdadeiro para todos os estados e regiões, com exceção do Rio Grande do Sul. Dessa forma, a presunção levantada na análise descritiva foi corroborada.

As análises por diagramas de controle e GRV permitiram estudar o impacto da pandemia de COVID-19 com relação a cada mês do ano subsequente ao início da pandemia. Aqui foram observados os limites históricos, estes que foram comparados com os dados atuais de forma a ser possível observar quais estados passaram mais meses com atendimentos aquém do esperado. Neste estudo, um grande número de localidades atingidas foi observado, o que resultou em diversos diagramas de controle, como visto na Figura 5.4. Dessa forma, o GRV conseguiu condensar as informações pertinentes e apresentar os resultados em uma única figura. Contudo, uma limitação do GRV é a não especificidade, por exemplo, não foi possível identificar os meses com atendimentos dentro dos limites históricos comuns a todos os estados (abril a junho).

### 5.2.3 Conclusões

A pandemia da COVID-19 parece ter representado barreira ao acesso a consultas para o ACD de crianças menores de cinco anos nos serviços de APS no Brasil, com impactos geográfica e temporalmente desiguais. Embora a restrição de serviços eletivos na APS tenha sido considerada necessária para minimizar o risco de transmissão da COVID-19, o impacto dessas restrições na saúde das crianças poderá ser duradouro. A retomada dos atendimentos para ACD pelos profissionais da APS é uma necessidade urgente, e para tal recomenda-se redução de barreiras administrativas aos atendimentos e a adoção de soluções inovadoras com o uso de tecnologias.

## 6 PRINCIPAIS CONCLUSÕES

A metodologia proposta para avaliação de impactos de epidemias e desastres naturais na população atingida foi formulada, detalhada e empregada em dois estudos de caso distintos. Foi proposto um novo método, intitulado Gráfico de Vetores Resultantes, para síntese e apresentação das informações de múltiplos diagramas de controle em uma única figura. Esse método permite, por exemplo, inferir quantos meses um determinado atributo passou acima, abaixo ou dentro dos limites históricos ao observar a direção e os componentes de um vetor. Contudo, sua limitação é a não especificidade, ou seja, não é possível identificar quais foram os meses em que o atributo apresentou variação de comportamento. Uma outra novidade da metodologia é o emprego de agrupamento nebuloso para definição do conjunto de localidades de controle e para análise multivariável dos atributos.

Com relação às aplicações, o estudo de caso “Rompimento da barragem de Fundão em Mariana–MG” se enquadrou no problema de pesquisa descrito pela Formulação 1 do método, por isso, foi utilizado o algoritmo *Fuzzy c-means* para determinação dos controles. Além disso, foram empregadas as análises uni e multivariáveis visto que este se tratava de um problema com duas variáveis: “Taxa mensal de atendimentos por dengue” e “PIB municipal anual”. Neste estudo, o método evidenciou que algumas localidades sofreram mudança de comportamento após a ocorrência do evento. Contudo, nem todas as mudanças foram no sentido de aumento dos casos de dengue. Por sua vez, o estudo de caso “Impacto da pandemia da COVID-19 nos atendimentos para acompanhamento do crescimento e desenvolvimento de crianças no Brasil” se enquadrou no problema de pesquisa descrito pela Formulação 2. Assim, não se fez necessário determinar localidades de controle. Ademais, por ser um problema de apenas um atributo, apenas a análise univariável foi aplicada. Neste estudo as diferenças entre os períodos pré e durante a pandemia se mostraram bem evidentes. Tanto observando a granularidade estadual quanto regional, a maioria dos estados sofreu algum efeito (negativo) após a ocorrência do evento. Portanto, a metodologia proposta é capaz de apontar os principais efeitos sentidos nas localidades após o evento, dado um conjunto de variáveis, por meio das diferentes análises que agregam robustez ao método.

Em suma, as principais contribuições do trabalho foram:

- Método de sintetização das informações de diversos diagramas de controle: Gráfico de Resultantes Vetoriais;
- Método robusto, baseado em similaridade *fuzzy*, para a escolha das localidades de controle;

- Método para extração automatizada de dados e para “limpeza” de planilhas.

Já as principais vantagens da metodologia proposta são:

- **Velocidade de implementação:** a metodologia utiliza bancos de dados secundários, ou seja, dados já disponíveis. Além disso, inclui uma forma de automatizar a extração de dados;
- **Usabilidade:** pode orientar entidades não acadêmicas, como gestores em saúde, em tomada de decisão. Além de outras aplicações, visto que diversas classes de eventos podem ser analisadas;
- **Monitoramento:** a metodologia pode ser reaplicada para monitorar a evolução das variáveis.

### 6.1 Trabalhos Futuros

Os sistemas de informação em saúde, como o Tabnet e o SISAB utilizados nesta dissertação, são sensíveis a detectar impactos de eventos que possam acontecer agudamente nas populações, embora limitados intrinsecamente à implementação da tecnologia, relação entre paciente e profissional, além de fatores técnicos e sociais. Desse modo, sugere-se utilizar esses sistemas para implementação da metodologia em novos estudos de caso de forma a monitorar os impactos de novos desastres e epidemias, com atenção às interações entre diferentes eventos que possam afetar uma mesma localidade.

Vê-se também a necessidade de generalizar a metodologia para abordar mais de um período após a ocorrência do evento. Por exemplo, investigar a evolução dos atendimentos por dengue ao passar dos anos no estudo do Rompimento da barragem de Fundão em Mariana–MG. Algumas sugestões são utilizar cores para diferenciar os vetores do GRV por ano após o evento; estudar a possibilidade de inserir a magnitude das variáveis no GRV; usar ANOVA para comparar mais de um período de tempo com o comportamento padrão da localidade; remover tendências em séries históricas antes de empregar a metodologia, quando aplicável; e adicionar dados dos próximos períodos de tempo ao conjunto de dados utilizado para o agrupamento nebuloso.

Ademais, algumas possibilidades de aplicação da metodologia proposta são:

- Análise de eventos políticos como eleições e *impeachments*;
- Análise de eventos e outros indicadores (variáveis) socioeconômicos;
- Análise de tragédias climáticas.

## BIBLIOGRAFIA

- [1] EM-DAT. **EM-DAT, the International Disaster Database**. 2022. Acesso em: 10/05/2022. Disponível em: <<https://public.emdat.be/data>>. Página 16.
- [2] IBAMA. **RESOLUÇÃO CONAMA Nº 001, de 23 de janeiro de 1986**. 1983. Página 16.
- [3] BECQUART, Ninon A; NAUMOVA, Elena N; SINGH, Gitanjali; CHUI, Kenneth KH. Cardiovascular disease hospitalizations in louisiana parishes' elderly before, during and after hurricane katrina. **International journal of environmental research and public health**, Multidisciplinary Digital Publishing Institute, v. 16, n. 1, p. 74, 2019. Página 16.
- [4] KISHI, Mitsuo; AIZAWA, Fumie; MATSUI, Miki; YOKOYAMA, Yukari; ABE, Akiko; MINAMI, Kentaro; SUZUKI, Ruriko; MIURA, Hiroyuki; SAKATA, Kiyomi; OGAWA, Akira. Oral health-related quality of life and related factors among residents in a disaster area of the great east japan earthquake and giant tsunami. **Health and quality of life outcomes**, BioMed Central, v. 13, n. 1, p. 1–11, 2015. Página 16.
- [5] NISHIJIMA, Marislei; ROCHA, Fabiana Fontes. An economic investigation of the dengue incidence as a result of a tailings dam accident in brazil. **Journal of environmental management**, Elsevier, v. 253, p. 109748, 2020. Páginas 16 e 39.
- [6] BEDREGAL, Benjamín RC *et al.* A comparative study between fuzzy c-means and ckmeans algorithms. In: IEEE. **2010 Annual Meeting of the North American Fuzzy Information Processing Society**. [S.l.], 2010. p. 1–6. Página 17.
- [7] CHIU, Stephen L. Fuzzy model identification based on cluster estimation. **Journal of Intelligent & fuzzy systems**, IOS Press, v. 2, n. 3, p. 267–278, 1994. Páginas 17 e 29.
- [8] SELICK, John A. The use of statistical process control charts in hospital epidemiology. **Infection Control and Hospital Epidemiology**, v. 14, p. 649–656, 11 1993. ISSN 0899823X. Páginas 22 e 31.
- [9] BENNEYAN, James C. Statistical quality control methods in infection control and hospital epidemiology, part i introduction and basic theory. **Infection Control & Hospital Epidemiology**, Cambridge University Press, v. 19, n. 3, p. 194–214, 1998. Página 22.
- [10] BIRNBAUM, David; BENNEYAN, James C. Statistical quality control methods in infection control and hospital epidemiology, part ii: chart use, statistical properties, and research issues. **Infection Control & Hospital Epidemiology**, Cambridge University Press, v. 19, n. 4, p. 265–283, 1998. Página 22.
- [11] ARANTES, Aglai; CARVALHO, Eduardo da Silva; MEDEIROS, Eduardo Alexandrino Servolo; FARHAT, Calil Kairalla; MANTESE, Orlando César. Uso de diagramas de controle na vigilância epidemiológica das infecções hospitalares. **Revista de Saúde Pública**, SciELO Brasil, v. 37, n. 6, p. 768–774, 2003. Página 22.

- [12] MONTGOMERY, Douglas C. **Design and analysis of experiments**. [S.l.]: John Wiley & sons, 2017. Páginas 23, 24 e 34.
- [13] BABUŠKA, Robert. **Fuzzy modeling for control**. [S.l.]: Springer Science & Business Media, 2012. v. 12. Páginas 25 e 26.
- [14] ABONYI, Janos; FEIL, Balazs. **Cluster analysis for data mining and system identification**. [S.l.]: Springer Science & Business Media, 2007. Página 25.
- [15] BEZDEK, James C. **Pattern Recognition with Fuzzy Objective Function Algorithms**. [S.l.]: Springer US, 1981. (Advanced Applications in Pattern Recognition). Páginas 26 e 27.
- [16] DUNN, Joseph C. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. Taylor & Francis, 1973. Página 27.
- [17] YANG, Qing; ZHANG, Dongxu; TIAN, Feng. An initialization method for fuzzy c-means algorithm using subtractive clustering. In: IEEE. **2010 third international conference on intelligent networks and intelligent systems**. [S.l.], 2010. p. 393–396. Páginas 29 e 30.
- [18] XIONG, Sheng-Wu; NIU, Xiao-Xiao; LIU, Hong-Bing. Support vector machines based on subtractive clustering. In: IEEE. **2005 International Conference on Machine Learning and Cybernetics**. [S.l.], 2005. v. 7, p. 4345–4350. Página 30.
- [19] SCHMIDT, Wolf-Peter; SUZUKI, Motoi; THIEM, Vu Dinh; WHITE, Richard G.; TSUZUKI, Ataru; YOSHIDA, Lay-Myint; YANAI, Hideki; HAQUE, Ubydul; THO, Le Huu; ANH, Dang Duc; ARIYOSHI, Koya. Population density, water supply, and the risk of dengue fever in vietnam: Cohort study and spatial analysis. **PLoS Medicine**, v. 8, p. e1001082, 8 2011. ISSN 1549-1676. Página 39.
- [20] IBAMA. **Rompimento da Barragem de Fundão: Documentos relacionados ao desastre da Samarco em Mariana/MG**. 2020. Acesso em: 08/05/2022. Disponível em: <<http://www.ibama.gov.br/cites-e-comercio-externo/cites?id=117>>. Página 39.
- [21] Fundação Renova. **Sobre o termo**. 2018. Acesso em: 08/05/2022. Disponível em: <<https://www.fundacaorenova.org/sobre-o-termo/>>. Página 39.
- [22] DATASUS. **TABNET**. 2018. Acesso em: 08/05/2022. Disponível em: <<https://datasus.saude.gov.br/informacoes-de-saude-tabnet/>>. Página 39.
- [23] IBGE. **Estimativas da População**. 2022. Acesso em: 08/05/2022. Disponível em: <<https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html?=&t=downloads>>. Página 39.
- [24] IBGE. **Produto Interno Bruto dos Municípios**. 2017. Acesso em: 08/05/2022. Disponível em: <<https://cidades.ibge.gov.br/brasil/mg/uniao-de-minas/pesquisa/38/47001>>. Página 39.
- [25] IBGE. **Índice de Desenvolvimento Humano**. 2017. Acesso em: 08/05/2022. Disponível em: <<https://cidades.ibge.gov.br/brasil/mg/uniao-de-minas/pesquisa/37/30255>>. Página 40.



- [26] ORGANIZATION, Geneva: World Health. **WHO COVID-19 Dashboard**. 2020. Disponível em: <<https://covid19.who.int/>>. Página 50.
- [27] CHOU, Yi Chang; YEN, Yung Feng; FENG, Rung Chuang; WU, Meng Ping; LEE, Ya Ling; CHU, Dachen; HUANG, Sheng Jean; CURTIS, J. Randall; HU, Hsiao Yun. Impact of the covid-19 pandemic on the utilization of hospice care services: A cohort study in taiwan. **Journal of Pain and Symptom Management**, Elsevier Inc., v. 60, p. e1–e6, 9 2020. ISSN 18736513. Página 50.
- [28] LIYA, Guo; YUGUANG, Wang; JIAN, Liu; HUAIPING, Yuan; XUE, Han; JIANWEI, Huo; JIAJU, Ma; YOURAN, Lu; CHEN, Ming; YIQING, Jiao. **Studies on viral pneumonia related to novel coronavirus SARS-CoV-2, SARS-CoV, and MERS-CoV: a literature review**. [S.l.]: Blackwell Munksgaard, 2020. 423-432 p. Página 50.
- [29] LUCENA, Edson Hilan Gomes de; FREIRE, Aldelany Ramalho; FREIRE, Deborah Ellen Wanderley Gomes; ARAÚJO, Elza Cristina Farias de; LIRA, Gabriela Nazaré Wanderley; BRITO, Arella Cristina Muniz; PADILHA, Wilton Wilney Nascimento; CAVALCANTI, Yuri Wanderley. Offer and use of oral health in primary care before and after the beginning of the covid-19 pandemic in brazil. **Pesquisa Brasileira em Odontopediatria e Clínica Integrada**, Association of Support to Oral Health Research (APESB), v. 20, p. 1–6, 2020. ISSN 19834632. Página 50.
- [30] LUO, Zujin; LI, Shunli; LI, Nan; LI, Yichong; ZHANG, Ying; CAO, Zhixin; MA, Yingmin. Assessment of pediatric outpatient visits for notifiable infectious diseases in a university hospital in beijing during covid-19. **JAMA network open**, NLM (Medline), v. 3, p. e2019224, 8 2020. ISSN 25743805. Página 50.
- [31] PAHUJA, Meera; WOJCIKEWYCH, Devon. Systems barriers to assessment and treatment of covid-19 positive patients at the end of life. **Journal of Palliative Medicine**, Mary Ann Liebert Inc., v. 24, p. 302–304, 2 2021. ISSN 15577740. Página 50.
- [32] WOODRUFF, Alex; FRAKT, Austin B. Covid-19 pandemic leads to decrease in emergency department wait times. **JAMA Health Forum**, American Medical Association (AMA), v. 1, p. e201172, 9 2020. Página 50.
- [33] HEINTZMAN, John; O'MALLEY, Jean; MARINO, Miguel; TODD, Jonathan V.; STANGE, Kurt C.; HUGUET, Natalie; GOLD, Rachel. **SARS-CoV-2 Testing and Changes in Primary Care Services in a Multistate Network of Community Health Centers during the COVID-19 Pandemic**. [S.l.]: American Medical Association, 2020. 1459-1462 p. Página 50.
- [34] AKSEER, Nadia; KANDRU, Goutham; KEATS, Emily C.; BHUTTA, Zulfiqar A. Covid-19 pandemic and mitigation strategies: Implications for maternal and child health and nutrition. **American Journal of Clinical Nutrition**, Oxford University Press, v. 112, p. 251–256, 8 2020. ISSN 19383207. Página 50.
- [35] SHUMBA, Constance; MAINA, Rose; MBUTHIA, Gladys; KIMANI, Rachel; MBUGUA, Stella; SHAH, Sweta; ABUBAKAR, Amina; LUCHTERS, Stanley; SHAIBU, Sheila; NDIRANGU, Eunice. **Reorienting nurturing care for early childhood development during the covid-19 pandemic in kenya: A review**. [S.l.]: MDPI AG, 2020. 1-19 p. Páginas 50 e 57.



- [36] VOGEL, Mandy; BEGER, Christoph; GAUSCHE, Ruth; JURKUTAT, Anne; PFAEFFLE, Roland; KÖRNER, Antje; MEIGEN, Christof; POULAIN, Tanja; KIESS, Wieland. Covid-19 pandemic and families' utilization of well-child clinics and pediatric practices attendance in germany. **BMC Research Notes**, v. 14, p. 140, 12 2021. ISSN 1756-0500. Páginas 50 e 56.
- [37] CHANG, Hong-Jen; HUANG, Nicole; LEE, Cheng-Hua; HSU, Yea-Jen; HSIEH, Chi-Jeng; CHOU, Yiing-Jenq. The impact of the sars epidemic on the utilization of medical services: Sars and the fear of sars. **American Journal of Public Health**, v. 94, p. 562–564, 4 2004. ISSN 0090-0036. Página 50.
- [38] PENNA, Gerson Oliveira; SILVA, José Agenor Alvares da; NETO, José Cerbino; TEMPORÃO, José Gomes; PINTO, Luiz Felipe. Pnad covid-19: A powerful new tool for public health surveillance in brazil. **Ciência e Saúde Coletiva**, Associação Brasileira de Pós - Graduação em Saúde Coletiva, v. 25, p. 3567–3571, 2020. ISSN 16784561. Página 50.
- [39] ABRASCO, Rede de Pesquisa em APS da. **Desafios da APS no SUS no enfrentamento da Covid-19**. 2020. Disponível em: <<https://www.youtube.com/watch?v=EcfJXwZdAsI&feature=youtu.be>>. Página 50.
- [40] ENGSTROM, Elyne; MELO, Eduardo; GIOVANELLA, Lígia; MENDES, Adelyne; GRABOIS, Victor; MENDONÇA, Maria Helena Magalhães de. **Recomendações para a organização da atenção primária à saúde no SUS no enfrentamento da COVID-19**. 2020. Disponível em: <<https://portal.fiocruz.br/documento/recomendacoes-para-organizacao-da-aps-no-sus-no-enfrentamento-da-covid-19>>. Página 50.
- [41] MEDINA, Maria Guadalupe; GIOVANELLA, Lígia; BOUSQUAT, Aylene; MENDONÇA, Maria Helena Magalhães de; AQUINO, Rosana. Atenção primária à saúde em tempos de covid-19: o que fazer? **Cadernos de Saúde Pública**, v. 36, 2020. ISSN 1678-4464. Página 50.
- [42] ESTRATÉGICAS., Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Ações Programáticas. **Política Nacional de Atenção Integral à Saúde da Criança : orientações para implementação**. 2018. Disponível em: <<https://central3.to.gov.br/arquivo/494643/>>. Página 50.
- [43] BÁSICA., Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção. **Saúde da criança: crescimento e desenvolvimento**. 2012. Disponível em: <[https://bvsm.s.saude.gov.br/bvs/publicacoes/saude\\_crianca\\_crescimento\\_desenvolvimento.pdf](https://bvsm.s.saude.gov.br/bvs/publicacoes/saude_crianca_crescimento_desenvolvimento.pdf)>. Página 50.
- [44] SAÚDE, Brasil: Ministério da. **Portaria GM/MS n 1.412, de 10 de julho de 2013 - Institui o Sistema de Informação em Saúde para a Atenção Básica (SISAB)**. 2013. Disponível em: <[https://bvsm.s.saude.gov.br/bvs/saudelegis/gm/2013/prt1412\\_10\\_07\\_2013.html](https://bvsm.s.saude.gov.br/bvs/saudelegis/gm/2013/prt1412_10_07_2013.html)>. Página 51.
- [45] SAURKAR, Anand V; GODE, Shweta A. **An Overview On Web Scraping Techniques And Tools**. 2018. Disponível em: <<http://www.ijfrcsce.org>>. Página 51.

- [46] GLEZ-PEÑA, Daniel; LOURENÇO, Anália; LÓPEZ-FERNÁNDEZ, Hugo; REBOIRO-JATO, Miguel; FDEZ-RIVEROLA, Florentino. Web scraping technologies in an api world. **Briefings in Bioinformatics**, v. 15, p. 788–797, 9 2014. ISSN 1477-4054. Página 51.
- [47] FREIRE, Flávio Henrique Miranda de Araújo; GONZAGA, Marcos Roberto; GOMES, Marília Miranda Forte. Projeções populacionais por sexo e idade para pequenas áreas no brasil. **Revista Latinoamericana de Población**, v. 14, p. 124–149, 12 2019. ISSN 2393-6401. Página 51.
- [48] JENSEN, C; MCKERROW, N H. Child health services during a covid-19 outbreak in kwazulu-natal province, south africa. **South African Medical Journal**, v. 111, p. 114, 12 2020. ISSN 2078-5135. Página 56.
- [49] NAVSARIA, P H; NICOL, A J; PARRY, C D H; MATZOPOULOS, R; MAQUNGO, S; GAUDIN, R. The effect of lockdown on intentional and nonintentional injury during the covid-19 pandemic in cape town, south africa: A preliminary report. **South African Medical Journal**, v. 111, p. 110, 12 2020. ISSN 2078-5135. Página 56.
- [50] WANYANA, D.; WONG, R.; HAKIZIMANA, D. Rapid assessment on the utilization of maternal and child health services during covid-19 in rwanda. **Public Health Action**, v. 11, p. 12–21, 3 2021. ISSN 2220-8372. Página 56.
- [51] SIEDNER, Mark J; KRAEMER, John D; MEYER, Mark J; HARLING, Guy; MNGOMEZULU, Thobeka; GABELA, Patrick; DLAMINI, Siphephelo; GARETA, Dickman; MAJOZI, Nomathamsanqa; NGWENYA, Nothando; SEELEY, Janet; WONG, Emily; IWUJI, Collins; SHAHMANESH, Maryam; HANEKOM, Willem; HERBST, Kobus. Access to primary healthcare during lockdown measures for covid-19 in rural south africa: an interrupted time series analysis. **BMJ Open**, v. 10, p. e043763, 10 2020. ISSN 2044-6055. Página 56.
- [52] BHAT, Anjana. Analysis of the spark study covid-19 parent survey: Early impact of the pandemic on access to services, child/parent mental health, and benefits of online services. **Autism Research**, v. 14, p. 2454–2470, 11 2021. ISSN 1939-3792. Páginas 56 e 58.
- [53] LEAL, Maria do Carmo; SZWARCOWALD, Celia Landmann; ALMEIDA, Paulo Vicente Bonilha; AQUINO, Estela Maria Leão; BARRETO, Mauricio Lima; BARROS, Fernando; VICTORA, Cesar. Saúde reprodutiva, materna, neonatal e infantil nos 30 anos do sistema Único de saúde (sus). **Ciência Saúde Coletiva**, v. 23, p. 1915–1928, 6 2018. ISSN 1678-4561. Página 57.
- [54] ROBERTON, Timothy; CARTER, Emily D; CHOU, Victoria B; STEGMULLER, Angela R; JACKSON, Bianca D; TAM, Yvonne; SAWADOGO-LEWIS, Talata; WALKER, Neff. Early estimates of the indirect effects of the covid-19 pandemic on maternal and child mortality in low-income and middle-income countries: a modelling study. **The Lancet Global Health**, v. 8, p. e901–e908, 7 2020. ISSN 2214109X. Página 57.
- [55] CADWGAN, Jill; GOODWIN, Jane; ARICHI, Tomoki; PATEL, Anita; TURNER, Susie; BARKEY, Sinead; MCDONALD, Aoife; LUMSDEN, Daniel E.; FAIRHURST, Charlie. Care in covid: A qualitative analysis of the impact of covid-19 on the health and care of children and young people with severe physical neurodisability and their families. **Child: Care, Health and Development**, 12 2021. ISSN 0305-1862. Página 57.

- [56] MAYURASAKORN, Korapat; PINSAWAS, Bonggochpass; MONGKOLSUCHARITKUL, Pichanun; SRANACHAROENPONG, Kittti; DAMAPONG, Sa-nga. School closure, covid-19 and lunch programme: Unprecedented undernutrition crisis in low-middle income countries. **Journal of Paediatrics and Child Health**, v. 56, p. 1013–1017, 7 2020. ISSN 1034-4810. Página [57](#).
- [57] TESTER, June M.; ROSAS, Lisa G.; LEUNG, Cindy W. Food insecurity and pediatric obesity: a double whammy in the era of covid-19. **Current Obesity Reports**, v. 9, p. 442–450, 12 2020. ISSN 2162-4968. Página [57](#).
- [58] YOSHIKAWA, Hirokazu; WUERMLI, Alice J.; BRITTO, Pia Rebello; DREYER, Benard; LECKMAN, James F.; LYE, Stephen J.; PONGUTA, Liliana Angelica; RICHTER, Linda M.; STEIN, Alan. Effects of the global coronavirus disease-2019 pandemic on early childhood development: Short- and long-term risks and mitigating program and policy actions. **The Journal of Pediatrics**, v. 223, p. 188–193, 8 2020. ISSN 00223476. Página [57](#).
- [59] HEARST, Mary O.; HUGHEY, Lauren; MAGOON, Jamie; MUBUKWANU, Elizabeth; NDONJI, Mulemba; NGULUBE, Esther; MAKHOUL, Zeina. Rapid health impact assessment of covid-19 on families with children with disabilities living in low-income communities in lusaka, zambia. **PLOS ONE**, v. 16, p. e0260486, 12 2021. ISSN 1932-6203. Página [57](#).
- [60] ELSAYED, Walaa. Covid-19 pandemic and its impact on increasing the risks of children’s addiction to electronic games from a social work perspective. **Heliyon**, v. 7, p. e08503, 12 2021. ISSN 24058440. Página [57](#).
- [61] GATELL-CARBÓ, Anna; ALCOVER-BLOCH, Elena; BALAGUER-MARTÍNEZ, Josep Vicent; PÉREZ-PORCUNA, Tomás; ESTELLER-CARCELLER, María; GARCIA, Patrícia Álvarez; FORTEA-GIMENO, Emilio. State of child and adolescent mental health during the first wave of the covid-19 pandemic and at the beginning of the 2020–2021 school year. **Anales de Pediatría (English Edition)**, v. 95, p. 354–363, 11 2021. ISSN 23412879. Página [57](#).
- [62] GILBERT, Amanda S.; SCHMIDT, Laurel; BECK, Alan; KEPPER, Maura M.; MAZZUCCA, Stephanie; EYLER, Amy. Associations of physical activity and sedentary behaviors with child mental well-being during the covid-19 pandemic. **BMC Public Health**, v. 21, p. 1770, 12 2021. ISSN 1471-2458. Página [57](#).
- [63] MARQUES, Emanuele Souza; MORAES, Claudia Leite de; HASSELMANN, Maria Helena; DESLANDES, Suely Ferreira; REICHENHEIM, Michael Eduardo. A violência contra mulheres, crianças e adolescentes em tempos de pandemia pela covid-19: panorama, motivações e formas de enfrentamento. **Cadernos de Saúde Pública**, v. 36, 2020. ISSN 1678-4464. Página [57](#).
- [64] DEMAURO, Sara B.; DUNCAN, Andrea F.; HURT, Hallam. Telemedicine use in neonatal follow-up programs – what can we do and what we can’t – lessons learned from covid-19. **Seminars in Perinatology**, v. 45, p. 151430, 8 2021. ISSN 01460005. Página [57](#).