



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

LUCAS CAVALCANTE MACHADO

**MODELO DE DECISÃO E ANÁLISE ESPAÇO-TEMPORAL DE ACIDENTES
DE TRÂNSITO**

Recife
2022

LUCAS CAVALCANTE MACHADO

**MODELO DE DECISÃO E ANÁLISE ESPAÇO-TEMPORAL DE ACIDENTES
DE TRÂNSITO**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para obtenção do título de mestre em Engenharia de Produção. Área de Concentração: Gerência da Produção.

Orientadora: Profa. Dra. Caroline Maria de Miranda Mota

Recife
2022

Catálogo na fonte
Bibliotecária Maria Luiza de Moura Ferreira, CRB-4 / 1469

M149m Machado, Lucas Cavalcante.

Modelo de decisão e análise espaço-temporal de acidentes de trânsito
/ Lucas Cavalcante Machado. - 2022.

96 folhas, il.; tab., abr. E sigl.

Orientadora: Profa. Dra. Caroline Maria de Miranda Mota.

Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Programa
de Pós-Graduação em Engenharia de Produção, 2022.

Inclui Referências.

1. Engenharia de Produção. 2. Sinistro viário. 3. Espaço-temporal.
4. MCDM/A. I. Mota, Caroline Maria de Miranda (Orientadora). II. Título.

UFPE

658.5 CDD (22. ed.)

BCTG/2022-151

LUCAS CAVALCANTE MACHADO

**MODELO DE DECISÃO E ANÁLISE ESPAÇO-TEMPORAL DE ACIDENTES DE
TRÂNSITO**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para obtenção do título de mestre em Engenharia de Produção. Área de Concentração: Gerência da Produção.

Aprovada em: 10/ 03 / 2022

BANCA EXAMINADORA

Profª. Dra. Caroline Maria de Miranda Mota (Orientadora)
Universidade Federal de Pernambuco

Prof. Dr. Cristiano Alexandre Virgínio Cavalcante (Examinador Interno)
Universidade Federal de Pernambuco

Prof. Dr. Flávio José Craveiro Cunto (Examinador Externo)
Universidade Federal do Ceará

AGRADECIMENTOS

Agradeço primeiramente a Deus, por estar sempre ao meu lado, e por ter me dado força para concluir essa dissertação e me conceder o grau de mestre em Engenharia de produção.

Agradeço a minha mãe Clovânia, meu pai Marcos e minhas irmãs Larissa e Laíza, e meus sobrinhos Sara e Samuel, por todo incentivo e apoio nessa trajetória. Além disso, agradeço aos demais familiares que de alguma forma deram suporte nessa jornada, em especial minhas vós e vô, e meus tias Cláudia, Lane, Márcia, Rose e Jussara, meus tios Paulo, Maciel e Adriano, e também, Lurdes, Marluce e Jane.

Agradeço também à minha orientadora Caroline Mota, por todo apoio e suporte dado no desenvolvimento da pesquisa, sem ela nada seria possível. Agradeço aos meus amigos e colegas por todo apoio, suporte e compartilhamento de conhecimento. Em especial, Igor, Lucas, Miguel, Maria Luíza, Maria Julha, Paolla, Maricélia, Arthur, Ciro e Amanda, e demais integrantes do PMD. Além disso, deixo minha gratidão aos professores do PPGEF.

Deixo minha gratidão também, aos meus professores de graduação da UFAL, Victor Carvalho, Jonhatan Magno, Aline Thamyres e Robério, por todo apoio e incentivo.

E por fim, agradeço a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro nesses dois anos de mestrado.

RESUMO

Os sinistros de trânsito estão entre as principais causas de morte no mundo, sendo responsável por cerca de 1,35 milhão de vítimas fatais por ano e outros milhares de feridos. Além desses, os distúrbios psicológicos, congestionamento, interrupção de transporte público e transporte de carga, sobrecarga nos atendimentos de emergências, comprometimento no abastecimento de pessoas, são outras consequências que estão associadas a esses eventos. Diante disso, analisar essas ocorrências é de fundamental importância para entender a dinâmica desses sinistros e desenvolver políticas mitigadoras para redução desses sinistros. Com base nisso, está pesquisa propõe um modelo de análise espaço-temporal de sinistros de trânsito, buscando identificar, caracterizar e priorizar pontos críticos de sinistros de trânsito. As informações obtidas com essa aplicação serão essenciais para o desenvolvimento das medidas de controle. Para tanto, o estudo utiliza as abordagens de densidade de Kernel, autocorrelação espacial global e local para análise espaço-temporal dos sinistros, e assim identificar o padrão espaço-temporal. Além do mais, o GWR (*Geographically Weighted Regression*) foi aplicado para verificar o impacto de alguns fatores na variação dos sinistros. Por fim, um modelo de decisão multicritério foi desenvolvido para ordenação de *hotspots* identificados. Após as análises, foi identificado os locais onde possuem uma maior densidade de sinistros, como o bairro Centro de Fortaleza, e que o padrão espacial dos sinistros varia com o tempo. Diante disso, o bairro Centro foi identificado como o local precário, devendo este ser priorizado nas ações das autoridades de trânsito.

Palavras-chave: sinistro viário; espaço-temporal; MCDM/A.

ABSTRACT

Traffic accidents are among the leading causes of death in the world, being responsible for about 1.35 million fatalities per year and thousands more injured. Besides these, psychological disturbances, congestion, interruption of public transport and cargo transportation, overload in emergency care, compromising the supply of people, are other consequences that are associated with these events. Therefore, analyzing these occurrences is of fundamental importance to understand the dynamics of these accidents and develop mitigating policies to reduce these accidents. Based on this, this research proposes a model of spatial-temporal analysis of traffic accidents, seeking to identify, characterize and prioritize accident hot spots. The information obtained from this application will be essential for the development of control measures. To this end, the study uses Kernel density approaches, global and local spatial autocorrelation for spatio-temporal analysis of accidents, and thus identify the spatio-temporal pattern. Furthermore, GWR (Geographically Weighted Regression) was applied to verify the impact of some factors on accident variation. Finally, a multi-criteria decision model was developed to order the identified hotspots. After the analyses, it was identified the places where there is a higher density of accidents, such as the Downtown neighborhood of Fortaleza, and that the spatial pattern of accidents varies with time. In view of this, the Center neighborhood was identified as the precarious location, which should be prioritized in the actions of traffic authorities.

Keywords: traffic accident; spatio-temporal; MCDM/A.

LISTA DE FIGURAS

Figura 1 - Fluxograma da metodologia proposta	31
Figura 2 - Número de sinistros e vítimas feridas entre 2015 e 2019	46
Figura 3 - Número de mortes e atropelamentos entre 2015 e 2019	47
Figura 4 - Mapas de densidade de Kernel de 2015 à 2019	54
Figura 5 - Mapas de densidade de Kernel para os meses de janeiro a abril	55
Figura 6 - Mapas de densidade de Kernel para os meses de maio a outubro	56
Figura 7 - Mapas de densidade de Kernel para os meses de novembro e dezembro	57
Figura 8 - Mapas de densidade de Kernel para os dias de segunda e terça	58
Figura 9 - Mapas de densidade de Kernel para os dias de quarta a domingo	59
Figura 10 - Mapas hotspots para os anos de 2015 a 2018	63
Figura 11 - Mapas hotspots para o ano de 2019	64
Figura 12 - Mapas hotspots para os meses de janeiro a junho	65
Figura 13 - Mapas hotspots para os meses de julho a dezembro	66
Figura 14 - Mapas hotspots para os dias de segunda a sábado	68
Figura 15 - Mapas hotspots para o dia de domingo	69
Figura 16 - Saída do modelo GWR	74
Figura 17 - Valores do R^2 para as variáveis de A até D	74
Figura 18 - Valores do R^2 para as variáveis de E até J	75
Figura 19 - Valores do R^2 para as variáveis de K até P	76
Figura 20 - Valores do R^2 para a variável N° de interseções	77
Figura 21 - Localização dos hotspots no bairro Centro de Fortaleza	83
Figura 22 - Diagrama de Hasse da ordenação dos pontos críticos	86

LISTA DE QUADROS

Quadro 1 - Classificação do atributo densidade de trânsito	82
Quadro 2 - Classificação do atributo sinalização	82
Quadro 3 - Classificação do atributo para tipo de via e cruzamento	82
Quadro 4 - Descrição das alternativas	84

LISTA DE TABELAS

Tabela 1 - Resultado do teste de Kruskal-Wallis	48
Tabela 2 – Similaridade de sinistros entre os anos	48
Tabela 3 - Índice de similaridade de sinistros nos meses de janeiro e fevereiro	49
Tabela 4 - Índice de similaridade de sinistros nos meses de março e abril	49
Tabela 5 - Índice de similaridade de sinistros nos meses de maio e junho	50
Tabela 6 - Índice de similaridade de sinistros nos meses de julho e agosto	50
Tabela 7 - Índice de similaridade de sinistros nos meses de setembro e outubro	50
Tabela 8 - Índice de similaridade de sinistros nos meses de novembro e dezembro	51
Tabela 9 - Índice de similaridade de sinistros para o período do dia	51
Tabela 10 - Índice de similaridade de sinistros entre os dias da semana	52
Tabela 11 - Índice de similaridade de sinistros entre os meses de 2019 (Continua)	52
Tabela 12 - Parâmetros da análise de densidade de kernel para anos	55
Tabela 13 - Parâmetros da análise de densidade de kernel para meses	57
Tabela 14 - Parâmetros da análise de densidade de kernel para os dias da semana	60
Tabela 15 - Resultado da Autocorrelação espacial global para os dias da semana	61
Tabela 16 - Resultado da Autocorrelação espacial global para os meses (Continua)	61
Tabela 17 - Resultado da Autocorrelação espacial global para os anos	62
Tabela 18 - Resultado da análise hotspots para os anos	64
Tabela 19 - Resultado da análise hotspots para os meses de 2019	67
Tabela 20 - Resultado da análise hotspots para dias da semana de 2019	69
Tabela 21 - Dados do processo de elicitação	85

LISTA DE ABREVIATURAS E SIGLAS

AHP	Analytical Hierarchy Process
AMC	Autarquia Municipal de Trânsito e Cidadania
ANOVA	Análise de Variância
DPVAT	Danos Pessoais Causados por Veículos Automotores de Vias Terrestres.
GIS	Geographic Information System
GLONASS	Global Navigation Satellite System
GPS	Global Positioning System
GWNBR	Geographically Weighted Negative Binomial Regression
GWPR	Geographically Weighted Panel Regression
GWR	Geographically Weighted Regression
IBGE	Instituto Brasileiro de Geografia e Estatística
KDE	Kernel Density Estimation
MCDM/A	Multiple Criteria Decision Making/Aid methods (MCDM/A)
ROC	Ranking Ordered Centroid
SAMU	Serviço de Atendimento Móvel de Urgência
SAW	Simple Additive Weighting
SMARTER	Simple Multi-Attribute Rating Technique Exploiting Ranks
TOPSIS	Technique For Order Of Preference By Similarity To Ideal Solution
WHO	World Health Organization

SUMÁRIO

1	INTRODUÇÃO	13
1.1	DESCRIÇÃO DO PROBLEMA	15
1.2	JUSTIFICATIVA E RELEVÂNCIA	15
1.3	OBJETIVO	16
1.3.1	Objetivo Geral	16
1.3.2	Objetivo específico	17
1.4	ESTRUTURA DO TRABALHO	17
2	REVISÃO DA LITERATURA	18
2.1	APLICAÇÕES DE FERRAMENTAS GIS E ANÁLISE ESPAÇO-TEMPORAL EM SINISTROS DE TRÂNSITO.	18
2.2	ESTIMATIVA DE DENSIDADE DE KERNEL	20
2.3	AUTOCORRELAÇÃO ESPACIAL MORAN I GLOBAL	22
2.4	APLICAÇÃO DA ANÁLISE AUTOCORRELAÇÃO ESPACIAL LOCAL EM SINISTROS DE TRÂNSITO	23
2.5	GEOGRAPHICALLY WEIGHTED REGRESSION	25
2.6	ABORDAGENS MCDM/A APLICADAS NA SEGURANÇA VIÁRIA	28
2.7	DISCUSSÃO FINAL DO CAPÍTULO	28
3	DADOS E MÉTODOS	30
3.1	DADOS	30
3.2	ÁREA DE ESTUDO	30
3.3	METODOLOGIA	31
3.4	MÉTODOS	35
3.4.1	Sistema de informação geográfica	35
3.4.2	Modelo de dados espaço-temporal	35
3.4.3	Análise de Cluster	36
3.4.4	Estimativa de densidade de Kernel	37
3.4.5	Autocorrelação espacial	38
3.4.6	Autocorrelação espacial global	39
3.4.7	Análise de Hot Spot	40
3.4.8	Geographically weighted regression	41
3.4.9	Teste de similaridade em padrões espaciais baseados em área	43
3.4.10	FITradeoff para problemática de ordenação	44

4	MODELO DE ANÁLISE ESPACIAL: ESTUDO DE CASO COM DADOS DE SINISTROS DA CIDADE DE FORTALEZA	46
4.1	ANÁLISE DESCRITIVA	46
4.2	ANÁLISE DE VARIÂNCIA	47
4.3	ANÁLISE DE PADRÃO DE PONTOS ESPACIAIS	48
4.4	ESTIMATIVA DE DENSIDADE DE KERNEL	53
4.5	AUTOCORRELAÇÃO ESPACIAL GLOBAL	60
4.6	ANÁLISE <i>HOTSPOTS</i>	62
4.7	DISCUSSÃO FINAL DO CAPÍTULO	70
5	ANÁLISE DE REGRESSÃO ESPACIAL	72
5.1	DADOS	72
5.2	SELEÇÃO DAS VARIÁVEIS	72
5.3	RESULTADO DO GWR	73
5.4	DISCUSSÃO FINAL DO CAPÍTULO	77
6	MODELO DE DECISÃO MULTICRITÉRIO PARA PRIORIZAÇÃO DOS HOTSPOTS	80
6.1	DESCRIÇÃO DO PROBLEMA DE DECISÃO MULTICRITÉRIO	80
6.2	CONSTRUÇÃO DO MODELO DE DECISÃO MULTICRITÉRIO	81
6.3	RESULTADOS	84
6.4	DISCUSSÃO FINAL DO CAPÍTULO	86
7	CONCLUSÃO	88
	REFERÊNCIAS	90

1 INTRODUÇÃO

De acordo com a *World Health Organization* (2018), os sinistros de trânsito são responsáveis por um número elevado de mortes globais. Apenas no ano 2016, cerca de 1,35 milhão de vidas foram perdidas em decorrência disso no mundo, correspondendo a 2,5% das causas de mortes no planeta e estando entre as dez maiores causas de morte, além de contabilizar 50 milhões de feridos.

De acordo com a seguradora Líder (2019), o Brasil está entre os 10 países que possuem os maiores números de mortes por sinistros de trânsito. Os dados do relatório anual de 2019 da seguradora, foram pagos um total de 353.232 indenizações pelo Danos Pessoais Causados por Veículos Automotores de Vias Terrestres (DPVAT) por sinistros nas rodovias brasileiras, sendo 40.721 indenizações por mortes, 235.456 por invalidez permanente e 77.055 com despesas médicas. A maior parte dessas vítimas foram em sinistros envolvendo motocicletas, correspondendo a 77% das indenizações pagas pela seguradora.

A partir desse relatório, a região nordeste do Brasil concentra a maior parte dessas indenizações pagas em 2019, seguida da região sudeste, ambas contabilizando 113.178 e 102.592 indenizações respectivamente. As duas regiões também lideram no número de indenizações pagas por mortes, sendo que a região Sudeste é responsável por 34% das mortes no Brasil enquanto a região Nordeste é responsável por 32%. Na região Nordeste, o estado do Ceará lidera o ranking na região com o acúmulo de 7,21% dos casos do país, seguido dos estados da Bahia (5,20%) e Pernambuco (5%).

Neste contexto, é importante entender que a mortalidade no trânsito precisa de uma atenção mais cuidadosa voltada à saúde pública, por suas altas taxas de ocorrências de sinistros. De acordo com Harirforoush e Bellalite (2019), uma das maiores preocupações das autoridades de transporte é entender onde e quando ocorrem esses sinistros de trânsito, e a identificação desses locais pode auxiliar nas estratégias de redução desses eventos. Dessa forma, conforme Leung, Li e Ji (2017) e Rahman e Kubota (2016) é necessário uma apuração minuciosa desses sinistros para solucionar e desenvolver planos de prevenção em todo mundo e assim, compreender os principais fatores desses acontecimentos.

Portanto, este estudo é relevante por propor um modelo espaço-temporal para detectar esses pontos críticos e encontrar os padrões de ocorrências de sinistros de trânsito. Assim, os resultados encontrados no desenvolvimento deste trabalho serão importantes para as autoridades de transporte pois, com a localização desses pontos e suas principais análises, será

possível elaborar estratégias de melhoria da segurança das vias, para diminuir o número desses sinistros.

Para Le, Liu e Lin (2020a) Le, Liu e Lin (2020b), é necessário identificar os locais com maior frequência de sinistros, chamados de pontos críticos, para redução desses acontecimentos. Segundo esses pesquisadores, existem muitos métodos que podem conduzir a identificação dos pontos críticos de sinistros e uma análise detalhada sobre esses pontos deve ser realizada para o desenvolvimento de estratégias para tornar as vias mais seguras.

Nessa circunstância, a análise espacial e temporal de sinistros de trânsito é relevante para desenvolver estratégias para prevenção dessas ocorrências, porque identificam os pontos críticos de sinistros (LAKES, 2017). Assim, de acordo com Sun et al. (2020), uma análise espaço-temporal pode ser útil para estudar as colisões de estradas pois, podem avaliar a relação entre essas colisões e as características ambientais podendo fornecer evidências empíricas para melhoria de segurança no trânsito, além disso pode encontrar áreas com alto risco de ocorrências de sinistros no trânsito.

Nesse contexto, o *Geographic Information System* (GIS) pode auxiliar na análise espaço-temporal de sinistros de trânsito. Além disso, é uma ferramenta muito útil, uma vez que com ela é possível armazenar atributos e ter uma compreensão visual através de mapas das regiões que concentram as maiores densidades de sinistros, além de ser bastante utilizado na identificação dos pontos críticos de sinistros (LE; LIU; LIN, 2020a).

De acordo com Yildirim e Mert Kantar (2020) as ocorrências de sinistros de trânsito aumentam na maioria dos países por diversos fatores. Os principais fatores que contribuem para essas ocorrências são: fatores humanos (YU et al., 2021; YILDIRIM; MERT KANTAR, 2020; SUN et al., 2020); violação das regras de trânsito (YU et al., 2021); aumento do número de veículos nas estradas (YU et al., 2021); fluxo de veículos (YILDIRIM; MERT KANTAR, 2020); fatores ambientais (YU et al., 2021; SUN et al., 2020); fatores climáticas (YU et al., 2021; YILDIRIM; MERT KANTAR, 2020; SUN et al., 2020); e condição das vias (YU et al., 2021).

Os parâmetros que provocam os sinistros de trânsito têm uma estrutura complexa e por essa razão esse assunto é relevante para os pesquisadores que têm o interesse de analisar esses acontecimentos para entender e modelar suas ocorrências (YILDIRIM; MERT KANTAR, 2020).

1.1 DESCRIÇÃO DO PROBLEMA

Os sinistros de trânsito estão entre as dez principais causas de morte no mundo, de acordo com a WHO (2018). Somente em 2016, foram registrados aproximadamente 1,35 milhão de vítimas fatais, além de milhares de feridos. Se nenhuma estratégia de prevenção for adotada, essas mortes podem se tornar a quinta maior causa de morte no mundo (ZOU; ZHANG; CHENG, 2021). Conforme esses autores, essas ocorrências custam anualmente US \$518 bilhões para o mundo.

Além das perdas por morte que são uma das consequências irreversíveis associadas aos sinistros de trânsito, Marzoug et al. (2021) relata que a perda da qualidade de vida da vítima não fatal, em relação a fatores como lesões e distúrbios psicológicos, é outra consequência impactante dessas ocorrências. Algumas vítimas sofrem lesões irreversíveis que as limitam no desenvolvimento de suas atividades laborais e relacionadas ao seu dia-a-dia.

Ainda de acordo com o autor, os sinistros de trânsito são os principais motivos para os congestionamentos nas zonas urbanas, comprometendo dessa forma a dinâmica do tráfego nas vias das cidades. Consequentemente, esse congestionamento pode afetar o funcionamento de diversos serviços de uma região como por exemplo, o transporte público, transportes de carga, ações da polícia, ocorrências do Serviço de Atendimento Móvel de Urgência (SAMU) ou Corpo de Bombeiros, podendo resultar em prejuízos consideráveis para a economia.

Diante disso, as análises dos sinistros de trânsito são um dos fatores principais no estudo do trânsito, visto que essas ocorrências podem causar um enorme impacto na vida humana (Marzoug et al., 2021). Dessa forma, essas ocorrências podem ser analisadas através de diversos métodos, como a análise espaço-temporal que é bastante utilizada em estudos de sinistros.

1.2 JUSTIFICATIVA E RELEVÂNCIA

De acordo com os dados da plataforma VIDA, em 2019 foram registrados 14.525 sinistros de trânsito nas vias da cidade de Fortaleza. Apenas neste ano 198 pessoas morreram e outras 13880 ficaram feridas. Nesse contexto, pode-se observar que as consequências atreladas a essas ocorrências, como o elevado número de mortes e lesões, deve ser uma preocupação primordial das autoridades de trânsito e também da saúde pública, visto que, as mortes por sinistros estão entre as principais causas de mortes no mundo.

Além dos prejuízos atrelados às vítimas, como mortes e ferimentos, os sinistros causam um impacto financeiro para o governo por conta da demanda por atendimento hospitalares, primeiros socorros com SAMU e corpo de bombeiros. Além disso, os prejuízos financeiros podem se estender para as unidades de serviços e indústrias nos centros urbanos, podendo comprometer o funcionamento da logística e impactar no abastecimento de pessoas, comércio e indústria.

Diante desses números, pode-se notar que a realização desse trabalho se torna relevante porque, com a aplicação do modelo espaço-temporal de análise de sinistros, se pretende identificar os pontos críticos, o padrão e o comportamento dessas ocorrências ao longo do tempo. Além disso, é fundamental entender quais os fatores que influenciam na quantidade de sinistros nas diversas regiões de análise.

Os dados extraídos dessas análises são de extrema importância para auxiliar as autoridades de trânsito no desenvolvimento de estratégias e adoção de intervenções necessárias para redução desses sinistros, diminuir consequentemente o número de mortos e feridos, tornando as vias mais seguras. Dessa forma, os órgãos de trânsito vão estar mais preparados para lidar e evitar esses sinistros, podendo atacar exatamente os pontos de maior concentração e fazer uma alocação de recursos de forma mais eficiente.

O uso do GIS será importante para localizar os pontos com mais sinistros, entender quais são os padrões dessas ocorrências ao longo do tempo e entender quais são os fatores que influenciam no número e fatalidade desses sinistros.

1.3 OBJETIVO

Nesta seção, serão apresentados o objetivo geral e os objetivos específicos que nortearão o desenvolvimento deste trabalho.

1.3.1 Objetivo Geral

O objetivo geral desta pesquisa consiste em propor um modelo de análise espaço-temporal de sinistros de trânsito, para identificar, caracterizar e priorizar pontos críticos de sinistros em vias urbanas, com base em um conjunto de dados históricos dessas ocorrências.

1.3.2 Objetivo específico

Os objetivos específicos desta pesquisa consistem em:

- Identificar os padrões espaciais e temporais dos sinistros de trânsito em vias urbanas.
- Identificar a existências de *clusters* espaciais de sinistros de trânsito em vias urbanas.
- Explorar possíveis associações entre variáveis socioeconômicas e a frequência de sinistros em bairros.
- Estruturar um modelo de decisão multicritério para priorizar locais críticos de sinistros e aplicar o modelo em um estudo de caso utilizando dados espaciais de sinistros de trânsito da cidade de Fortaleza, Brasil.

1.4 ESTRUTURA DO TRABALHO

A dissertação está organizada em sete capítulos. No próximo capítulo será exposto o resultado de uma revisão da literatura, que apresenta diversas aplicações das abordagens proposta na metodologia desse trabalho, e que mostram o quão relevante elas são na aplicação em estudos de sinistros de trânsito. Esses trabalhos, darão suporte nessa aplicação e na obtenção dos resultados esperados.

No terceiro capítulo será apresentado os conceitos básicos utilizados nesta pesquisa, investigado através da realização de uma busca na literatura. Além disso, será explanado sobre a área de abrangência desse estudo, os dados utilizados nessa aplicação, e sobre a metodologia proposta para atingir os abjetivos definidos para o trabalho.

No quarto capítulo será desenvolvido uma modelo de análise espacial com dados de sinistros de trânsito da cidade de Fortaleza. Nesse capitulo, serão apresentados uma análise qualitativa dos dados, e também a aplicação da ferramenta de análise de densidade de kernel, análise de autocorrelação espacial global e local utilizando as ferramentas de autocorrelação espacial com estatística de Moran global e análise *hotspots*, respectivamente.

O quinto capítulo apresenta uma modelo de análise de regressão espacial utilizando a ferramenta *Geographically Weighted Regression*. E no sexto, será desenvolvido um modelo de decisão multicritério para priorização de *hotspots* de sinistros de trânsito, com a aplicação do método FITradeoff para ordenação. E por fim, serão apresentadas as conclusões finais dessa pesquisa no sétimo capítulo.

2 REVISÃO DA LITERATURA

Este capítulo apresenta uma revisão da literatura, que traz um apanhado de trabalhos que mostram a relevância e o impacto da aplicação de algumas abordagens e conceitos nos estudos exploratórios de sinistros de trânsito.

2.1 APLICAÇÕES DE FERRAMENTAS GIS E ANÁLISE ESPAÇO-TEMPORAL EM SINISTROS DE TRÂNSITO.

A literatura está repleta de trabalhos com métodos de análise de sinistros de trânsito utilizando ferramentas GIS e análise espaço-temporal para identificação de tendências, mostrando o quão relevante esse tema é para os pesquisadores e que essas ferramentas são importantes para extrair informações essenciais para a definição medidas para redução de sinistros nas vias.

Diversos pesquisadores desenvolvem métodos e forma de analisar esses sinistros a fim prevenir essas ocorrências, como Yaacob et al. (2020) que utiliza aplicativos GIS para realizar a análise do espaço e do padrão dos sinistros através de técnicas de mapeamento; Rahman, Jamal e Al-Ahmadi (2020) utilizam a ferramenta GIS para análise de pontos de acesso de sinistro e em conjunto com a regressão geograficamente ponderada foi possível examinar os pontos críticos de colisões e suas relações espaciais com o uso do solo.

Hu et al. (2020), realizou uma combinação entre métodos de sistema de informação geográfica, regressão logística e modelos baseados em árvores para observar a distribuição de atropelamentos nas cidades. Enquanto isso, Al-Omari et al. (2020) utiliza a junção das ferramentas GIS e lógica fuzzy para previsão de pontos críticos de sinistros de trânsito em diferentes locais.

Lakes (2017) analisou a distribuição espaço-temporal dos dados de estatísticas de sinistros de pedestres e ciclistas na cidade de Berlin, Alemanha. Já Yu et al. (2021), propõem em seu trabalho um novo tipo de rede convolucional de grafo espacial-temporal profundo, para a previsão de sinistros de trânsito. Yildirim e Mert Kantar (2020) modelaram uma distribuição espacial da taxa de sinistros levando em consideração a ligação espaço e tempo, na Turquia.

Em seu estudo, Ayob et al. (2020) analisou a variação do espaço e tempo de atropelamentos por herpetofauna nas estradas de Langkawi, na Malásia; Kibar e Tuydes-

Yaman (2020) utilizou o agrupamento hierárquico de vizinho mais próximo para analisar a distribuição espaço-temporal dos pontos críticos de sinistros de trânsito na Turquia.

Özcan e Küçükönder (2020) analisou as densidades espaciais dos sinistros, e identificar as mudanças espaço-temporal dessas ocorrências na cidade de Kahramanmaraş, na Turquia. Já o trabalho de Souza et al. (2020), analisou a distribuição espaço-temporal das mortes de sinistro motociclistas um período antes e depois da implantação da Lei Seca no estado de Alagoas, no Brasil.

Já Ramírez e Valencia (2021) modelou um processo estocástico espaço-temporal para análise de sinistros utilizando o modelo Log-Gaussian Cox, na identificação de fatores que contribuem para o aumento de riscos de sinistros, além de identificar localização desses pontos críticos na cidade de Bogotá, na Colômbia.

Além das GIS e análise de espaço-temporal que podem ser utilizadas em conjunto como apresentadas nas análises realizadas pelos autores citados anteriormente, outras abordagens são muito empregadas na análise de sinistros de trânsito, como a análise de cluster, big data, Fuzzy, Topsis, redes neurais e diversos outros algoritmos, como mostrado nos trabalhos a seguir.

Os autores Zhang, Li e Li (2020) usam o algoritmo *random forest* para construir um modelo de previsão de gravidade de sinistros de trânsito. Iqbal et al. (2020) fez a aplicação do método de peso de pontos quentes de sinistros, para classificar os pontos de ocorrências em uma rodovia do Paquistão, e Čubranić-Dobrodolac et al. (2020) utiliza o método sistema inferência *Fuzzy* tipo 2 intervalar para realizar a comparação entre o impacto da velocidade e percepção do espaço, na ocorrência de sinistros de trânsito.

Liu et al. (2020) utilizou o TOPSIS e um modelo de regressão logística binária para analisar o impacto das características do alinhamento da estrada em diferentes tipos de sinistros de trânsito. Yue et al. (2017) utiliza uma tecnologia de correlação cinza de *big data* para analisar as estradas em sinistros de trânsito, e também correlação entre as características dos sinistros e dos motoristas.

Zou, Zhang e Cheng (2021) utilizam um modelo binomial negativo e um *modelo log-change* para avaliar os impactos do clima e das condições meteorológicas extremas em sinistros de trânsito fatais. Ospina-Mateus et al. (2021) aplica técnicas de mineração de dados e aprendizado de máquina para análise de sinistros com motocicletas nas estradas da cidade de Bogotá, na Colômbia.

Marzoug et al. (2021) desenvolveu um modelo de simulação que analisa como a ocorrência de sinistros de trânsito afeta o tráfego de veículos na interseção, com base no tempo

de evacuação dos veículos envolvidos, taxa de transmissão e a probabilidade de ocorrência com divisórias abertas. Wang et al. (2021) aplica o algoritmo de seleção de recursos de invólucros para prever a gravidade da lesão em sinistros de trânsito.

Diversos autores utilizam abordagens de cluster para prever, agrupar, analisar e mapear sinistros de trânsito, como Pasupathi et al. (2021) que utiliza a abordagem de *clustering* hierárquico aglomerativo para big data de série temporal para analisar as tendências dos sinistros de trânsito. Já Tain e Zhang (2020) propuseram em seu trabalho uma otimização contínua não convexa do algoritmo *k-means* para previsão de sinistros, fazendo a aplicação de algoritmo de agrupamento otimizado de big data em ambiente de computação em nuvem.

Syahputri et al. (2020) utilizou o algoritmo *fuzzy c-means* para agrupar a vulnerabilidade de sinistros de trânsito na cidade de Medan, na Indonésia. Os autores Taamneh, Taamneh e Alkheder (2017) utilizaram o cluster hierárquico juntamente com redes neurais artificiais para propor uma classificação baseada em cluster de sinistros de trânsito para previsão da gravidade das lesões ocorridas nos sinistros de trânsito.

Selvi e Caglar (2018) utilizam métodos de cluster para mapeamento multivariado de sinistros de trânsito. Já o trabalho proposto por Zahid et al. (2020) utiliza métodos de agrupamento de vizinhos próximos em frequências, para o desenvolvimento de mapas de pontos quentes para investigar as violações de tráfego na cidade de Luzhou, na China. E por fim, Sun et al. (2020) utilizou o método de detecção de cluster com base no modelo Bayesiano rápido sem covariáveis para encontrar clusters espaço-temporais de sinistros na cidade de Londres entre os anos de 2010 e 2014.

2.2 ESTIMATIVA DE DENSIDADE DE KERNEL

Há uma diversidade de trabalhos na literatura que utilizam a abordagem de estimativa de densidade de Kernel aplicados a dados de sinistros de trânsito em diversas partes do mundo, para identificar pontos críticos de sinistros de trânsito ao longo de rodovias e vias da cidade. Essa aplicação é realizada em diferentes tipos de estudos como mostrados nos trabalhos dos autores citados a seguir.

Le, Liu e Lin (2020a) utilizaram em seu trabalho a estimativa de densidade de Kernel planar para identificar pontos críticos de sinistros de trânsito, em Hanói no Vietnã. Com isso, obtiveram os mapas de densidade de sinistros com base em um índice de gravidade definido

para cada sinistro, e utilizando uma largura de banda de 1000 metros, e com isso, pode identificar facilmente os pontos de alta densidade na área analisada.

Na pesquisa de Hashimoto et al. (2016), um modelo KDE é utilizado como variável para encontrar estimativas de densidade de sinistros de trânsito, para utilizá-las como variável dependente em uma análise de regressão múltipla, e comparou os resultados com um modelo de contagem bruta de dados. Ele verifica que não há diferença entre os coeficientes dos modelos KDE e os modelos de contagem bruta de dados, além disso, os números estimados através da aplicação dos modelos possuem uma correlação positiva forte com o número reais de sinistros.

Os autores Özcan e Küçükönder (2020), aplicaram a estimativa de densidade de kernel para identificar os locais com maior densidade de sinistros em Kahramanmaraş, na Turquia, com dados de sinistros entre 2008 e 2015, e verificaram as mudanças no comportamento dessas ocorrências, com análise espaço-temporal. Desse modo, os autores identificaram que ocorreram mudanças na densidade desses sinistros em decorrência do desenvolvimento urbano no período da análise, e também, verificou que as densidades mais altas ocorrem no centro da cidade.

Shafabakhsh, Famili e Bahadori (2017), usaram a estimativa de densidade de Kernel para analisar os pontos críticos de sinistros de trânsito e fornecer informações para auxiliar no processo de tomada de decisão para redução de sinistros. Como resultado, os autores identificaram que os locais críticos, com alta probabilidade de sinistros estão localizados próximos à cidade de Mashhad, um dos grandes pontos turísticos do Irã.

No trabalho de Anderson (2009), o mesmo utilizou o KDE para analisar os padrões espaciais de sinistros rodoviários com lesões, na cidade de Londres, Reino Unido, com dados de 1999 a 2003. Ele não apenas utiliza o método para identificação de áreas de grandes densidades, mas também, para a seleção de locais alvos de investigação. A ferramenta permitiu criar a unidade espacial básica para o método de agrupamento de pontos críticos de sinistros.

Já no trabalho proposto por Le, Liu e Lin (2020b), o método foi utilizado para analisar *hotspots* de sinistros de trânsito, e investigou a influência dessas ocorrências nos padrões espaço-temporal, com dados de sinistros de 2015 a 2017 ocorridos na cidade de Hanói, no Vietnã. O método foi aplicado para analisar os padrões em diferentes estações do ano, e diferentes intervalos de tempo (como noturno, horário de pico e outros). Além disso, ele usou dois tipos de análise, uma considerando um índice de gravidade, que reflete a gravidade de cada ocorrência através de uma ponderação, e outra, sem os dados ponderados.

Como resultado, Le, Liu e Lin (2020b) concluíram que a análise realizada considerando o índice de gravidade é mais precisa ao determinar os pontos críticos de sinistros, e também,

mostrou que as interseções de vias próximas das estações Cho Tia e Van Diem na NH-1A, eram locais que tinham sinistros graves em todos os momentos da sua análise.

Kalantari et al. (2021), analisou a distribuição espacial de sinistros de trânsito em Zanjan, no Irã, associando o espaço-temporal com as colisões, e identificando os locais de risco de sinistros rodoviários. Assim, ele obteve os mapas de sinistros para quatro períodos do dia, divididos entre intervalos de seis horas. Além disso, observou que as estradas com alta densidade de tráfego, tinham um agrupamento de colisão mais intenso.

Sandhu et al. (2016) aplica a estimativa de densidade de Kernel para identificar locais com alta densidade de ocorrências de sinistros em uma rodovia entre as cidades Gurgaon e Jaipur, na Índia. A aplicação do método, resultou na identificação de sete pontos críticos de sinistros de trânsito nessa rodovia. E por fim, Machado, Souza e Mota (2021) utilizaram a abordagem a mesma para identificação de locais com grande concentração de sinistros de trânsito, na cidade de Fortaleza. Com essa análise, foi possível identificar os locais mais críticos de sinistros.

2.3 AUTOCORRELAÇÃO ESPACIAL MORAN I GLOBAL

É possível encontrar diversas pesquisas na literatura que aplicam a autocorrelação espacial Moran I global em sinistros de trânsito, para analisar a existência de cluster de sinistros nas áreas averiguadas. O trabalho proposto por Le, Liu e Lin (2020a) é um dos exemplos em que se aplica estatística global de Moran I para avaliar a correlação espacial dos eventos de ponto e como complemento da análise de densidade de kernel testando, assim, a distribuição aleatória dos pontos.

Os autores avaliaram o padrão de 1365 sinistros ocorridos em Hanói, no Vietnã, para determinar a significância estatística dos pontos críticos de sinistros de trânsito identificados através da estimativa de densidade de kernel e com isso, verificaram que os sinistros analisados demonstram ter um padrão agrupado. Além do mais, esse processo foi repetido inúmeras vezes para encontrar a largura de banda que corresponda num agrupamento maximizado.

Já Yildirim e Kantar Mert (2020), aplicaram a autocorrelação global de Moran I local para analisar a existência de *clusters* semelhantes, que expõem taxas elevadas de sinistro nas províncias da Turquia, utilizando registros de sinistros entre os anos de 2013 e 2018. Como resultado dessa aplicação, o autor concluiu que as taxas de sinistros na Turquia, não são

distribuídas aleatoriamente e também a distribuição espacial das províncias com taxas elevadas de sinistros também não são aleatórias.

Butt et al. (2017) em sua pesquisa, realizou a análise de autocorrelação espacial para verificar o padrão global dos dados de sinistros na cidade de Rawalpindi, no Paquistão, com dados entre os anos de 2009 e 2013. O autor conclui que a maior parte dos padrões observados foram aleatórios, e apenas os dados de 2011 apresentam um padrão agrupado.

No trabalho proposto por Manap et al. (2021), o autor usou a autocorrelação global para detectar agrupamentos de sinistros com veículos pesados para diferentes critérios, de uma via expressa interurbana, na Malásia. A análise revela a existência de um padrão aglomerado para os três critérios utilizados: frequência de sinistros com veículos pesados, número de veículos pesados e índice de gravidade.

Soltani e Askari (2017), usaram a análise de autocorrelação espacial global para determinar a concentração espacial de sinistros de trânsito, em Shiraz, a capital da província de Fars no Irã, onde analisou três tipos de sinistros, sendo eles, sinistros fatais, com lesão e apenas danos materiais. Os resultados demonstram que o padrão da distribuição desses sinistros é agrupado, para os diferentes tipos analisados.

Sandhu et al. (2016), utiliza a ferramenta para avaliar a autocorrelação espacial e identificar o padrão dos sinistros ocorridos em uma rodovia entre as cidades Gurgaon e Jaipur, na Índia. A análise revela que o padrão de sinistros no trecho analisado é aleatório, considerando diferentes distâncias. Assim, os locais dos sinistros são distribuídos aleatoriamente.

Ouni e Bellouni (2019), aplicaram essa abordagem para verificar o padrão de distribuição de sinistros em rodovias na Tunísia, com os sinistros registrados entre 2002 e 2013. A autocorrelação espacial global é utilizada para a identificação de tendência de aglomeração nas ocorrências de sinistros na área de estudo. Com isso, verificou-se que os sinistros têm um padrão aleatório na região centro-leste, entre os anos de 2002 e 2005, e na região centro-oeste, esses dados tem um tendências aleatória para o mesmo período.

2.4 APLICAÇÃO DA ANÁLISE AUTOCORRELAÇÃO ESPACIAL LOCAL EM SINISTROS DE TRÂNSITO

Diversos trabalhos na literatura, utilizam a análise autocorrelação local para identificação de *hotspots* e *coldspots* para dados de sinistros. Essa ferramenta é útil para detectar

os pontos críticos dessas ocorrências, e é utilizada em várias aplicações de sinistros ao redor do mundo.

Os Ersen, Büyüklü e Taşabat (2021) usaram a ferramenta de análise de autocorrelação local como proposta em seu trabalho, para detectar os *hotspots* e *coldspots* de sinistros fatais e com ferimentos no trânsito do distrito de Sarıyer em Istambul, na Turquia. Como resultado, os mapas gerados pela aplicação do método, elucidaram a locais com valores altos e baixos, pontos frios e quentes sinistros, mostrando as ruas e estradas que possuem essas características.

Butt et al. (2017), aplicou a análise autocorrelação espacial local para identificar pontos críticos de sinistros em locais específicos da cidade de Rawalpindi, no Paquistão. O mapa resultante da sua análise mostrou, que os *hotspots* ocorrem principalmente nas partes norte, noroeste e nordeste da cidade. Já os *coldspots*, estão localizados na área norte. E assim, o autor identificou os locais que mais necessitam de intervenções.

Manap et al. (2021), utiliza a estatística Getis Ord-Gi * para identificar pontos críticos de sinistros locais, em uma via expressa norte-sul interurbana, na Malásia em 2018. A partir dos mapas de saída, foi possível observar que os *hotspots* estão localizados em diferentes locais para os diferentes critérios da análise. O mapa dos dados de número de veículos pesados envolvidos apresenta a maior quantidade de pontos, 512 no total, enquanto que, para veículos pesados foi de 472, e levando em consideração o índice de gravidade, 475 pontos *hotspots*.

Tola et al. (2021) utilizou essa análise para verificar os pontos críticos de sinistros em cidades e regiões da Etiópia, com dados de sinistros entre 2010 e 2017. Sua análise revelou que os pontos críticos de sinistros de trânsito evidenciam a cidade de Addis Abeba e a região de Oramia, sendo que Addis Abeba, capital e maior cidade da Etiópia, foi identificada como o primeiro *hotspots*. Já os territórios federais Dire-Dawa e Harari, foram classificados como *coldspots*.

Sandhu et al. (2016), utiliza a ferramenta para identificação de pontos críticos de sinistros com base na estatística Getis-Ord Gi *, ao longo de uma rodovia entre as cidades Gurgaon e Jaipur, na Índia. Com isso, o autor identifica diversos *hotspots* de sinistros ao longo da rodovia na seção. Jakimavičius (2018) também utilizou a autocorrelação espacial local para analisar pontos críticos de sinistros nas estradas nacionais da Lituânia, com dados de 2014 até 2017. Com isso, o autor consegue identificar 66 seções de sinistros como *hotspots*, nas estradas do país. E por fim, Machado, Souza e Mota (2021) aplicaram a abordagem de autocorrelação espacial local para identificar os pontos críticos da cidade de Fortaleza, utilizando dados de sinistros de trânsito de 2019.

2.5 GEOGRAPHICALLY WEIGHTED REGRESSION

O modelo de regressão geograficamente ponderado é muito aplicado para verificar fatores que influenciam nas ocorrências de sinistros de trânsito. A literatura está repleta de pesquisas que utilizam o GWR e suas diversas extensões, em várias partes do mundo, para analisar a relação espacial entre os sinistros de trânsito e diversas variáveis associativas.

Iyanda e Osayomi (2021) fizeram a aplicação do da versão do GWR clássico e uma versão GWR multiescala em sua pesquisa, para analisar dados de sinistros de trânsito fatais entre os anos 2010 a 2015, no estado do Texas, nos Estados Unidos, agregados através do setor censitário. A variável dependente utilizada em ambos os modelos, foi a taxa de mortalidade em cada setor censitário e outras doze variáveis econômicas, foram consideradas como variáveis independentes do modelo.

Os autores, inicialmente utilizaram as seguintes variáveis independentes no modelo: porcentagem de população ocupada acima dos 16, porcentagem da população que utilizam o transporte público, bicicletas, táxis ou se deslocavam a pé ao trabalho, porcentagem da população que trabalha em *homeoffice*, porcentagem da população que dirigia sozinha, partilhavam caronas, tempo médio da viagem até o trabalho, nível de ocupação de serviços, nível de ocupação em administração, vendas e ocupação dos escritórios, recursos da construção, naturais e manutenção, ocupação de produção e população do setor censitário.

Na sequência usaram a técnica de rotação Varimax para reduzir as variáveis associativas para quatro (carona para o trabalho, dirigir sozinho, tempo médio de viagem ao trabalho e deslocamento no trabalho). Os autores perceberam que o tempo médio de viagem e caronas, possuem associação positiva em algumas áreas com alta taxa de letalidade. Além disso, verificaram que indivíduos que trabalham com vendas e dirigem sozinho, estão mais propensos às fatalidades de trânsito, principalmente em áreas não urbanas.

Na pesquisa propostas por Rahman, Jamal e Al-Ahmadi (2020), um modelo GWR é utilizado para analisar a relação espacial entre as colisões de tráfego e o uso do solo urbano, para melhorar a eficiência nas previsões de colisão na cidade de Damman, na Arábia Saudita. O modelo utilizado por eles averiguar as relações espaciais existentes entre o número de colisões e as densidades populacionais em cada bairro com base na gravidade (fatal e lesão, e apenas dano à propriedade), causas (distrações, excesso de velocidade, desvios repentinos de faixa, não cedendo, estrada ruim, fadiga de direção e violações de tráfego), e tipos de sinistros (colisões de veículos, atropelamento de objetos fixos e atropelamentos).

Como resultado, os pesquisadores observaram que os sinistros fatais e com feridos são observados principalmente onde há densidade populacionais altas, nos bairros com uso residencial e residenciais e comerciais. As distrações, excesso de velocidade, sinistros de sono e fadiga estão associados com bairros que possuem densidade populacional mais elevada, nas áreas mais residenciais. Já as colisões e colisões em locais fixos ocorrem, principalmente nos bairros com grande densidade populacional, na área noroeste da cidade, onde as ruas são mais estreitas.

No trabalho de Soroori, Mohammadzadeh Moghaddam e Salehi (2020), dados de sinistros de trânsito foram utilizados para verificar o desempenho da regressão binomial negativa geograficamente ponderada GWNBR, e compará-la com a regressão de Poisson geograficamente ponderada GWPR, ambas são extensões do GWR. A frequência dos sinistros com lesões foi estimada por eles, utilizando dados de sinistros da cidade de Mashhad no Irã, entre os anos de 2014 e 2015.

Os autores, analisaram a relação entre o número de colisões e variáveis associativas categorizadas, como fatores socioeconômicos e demográficos, características das estradas, exposição de viagem, transporte público, uso do solo. Como um dos resultados, identificaram um impacto significativo do número de sinistros com feridos para as variáveis independentes: vias arteriais, rampas, número de cruzamentos, pontos de ônibus e funcionários, uso de solo comercial, residencial e agrícola.

O autor Kavousi et al. (2019), utiliza um modelo GWR em sua pesquisa, para verificar os fatores que contribui para a frequência de sinistros com óbitos em 22 distritos do Teerã, no Irã, com registros de ocorrências entre os anos de 2014 e 2016. As variáveis independentes utilizadas são classificadas nas categorias uso do solo, população e encontro, redes de estradas e viagens.

Com isso, ele constatou que o comprimento das rodovias, nível de desenvolvimento, o uso de solo educacional, uso do solo industrial e número de empregados no local de residência, tem uma relação positivamente significativa com a variável dependente utilizada nessa proposta de modelo de regressão.

Saadat et al. (2019), aplicou um modelo GWR para analisar a relação entre a localização dos sinistros de trânsito com mortes motociclistas e fatores ambientais, classificados em três categorias, uso do solo, rede de estradas e viagens, e população e exposição. Esse estudo foi realizado também no Teerã, com dados de sinistros entre os anos de 2011 a 2016, em 22 distritos. O modelo desenvolvido mostrou que há uma associação significativamente estatística

entre o número de sinistros com mortes de motociclistas e as características demográficas, rede de vias e ao tipo de uso do solo.

O trabalho proposto por Zhang et al. (2015), aplica um modelo GWR para investigar os impactos entre os sinistros não motorizados em uma rede rodoviária. Dessa forma, ele considera a média de sinistros com pedestres e ciclistas entre os anos de 2004 e 2006, como variável dependente, e as variáveis associativas classificadas nas categorias medidas estruturais, uso do solo, comportamento de viagem, meios de transporte e características demográficas.

Através desse estudo, o autor conclui que a estrutura da rede viária é fundamental para a segurança de pedestres e ciclistas. Além disso, observou que os fatores variam em toda área de estudo, apontando que, as autoridades devem desenvolver medidas diferentes para cada área, ao invés de uma tomar apenas uma medida para toda a área de estudo.

Liu et al. (2018), analisou os impactos de sinistros de trânsito utilizando um modelo GWR, com dados de sinistros motorizados e atropelamentos de 2015, no sudoeste de Michigan. Como resultado, verificou-se que há correlação espacial entre bater e fugir do local. Os atropelamentos estão mais suscetíveis de ocorrerem em setores censitários com maiores taxas de desemprego, menor renda, menor grau de ensino, e maior densidade populacional.

Bao et al. (2017) propõe em sua pesquisa, o uso do GWR para verificar a relação entre a quantidade de sinistros com diversos fatores. O mesmo comparou o desempenho dos modelos GWR utilizando apenas variáveis tradicionais de tráfego, apenas atividades humanas baseadas na rede social Twitter, e por fim, outro modelo utilizando ambos. Nessa análise os sinistros foram divididos em dois grupos: sinistros com pedestres e envolvendo apenas veículos.

As variáveis associativas utilizadas por ele estavam classificadas nas seguintes categorias: atividade humana, exposição ao tráfego, características das estradas, e demográficas e socioeconômicas. O autor conclui que as atividades humanas impactam significativamente nas ocorrências de sinistros, onde, as variáveis as atividades ligadas a alimentação, recreação e educação, impactaram principalmente os pedestres, enquanto que, atividade ligadas a alimentação, compras e atividades sociais afetaram nos sinistros entre veículos.

Por fim, os autores Gomes, Cunto e Da Silva (2017), realizaram uma comparação entre os modelos espaciais Regressão de Poisson Geograficamente Ponderada GWPR e Regressão Binomial Negativa Ponderada Geograficamente GWNBR, utilizando dados de sinistros de trânsito da cidade de Fortaleza, no Brasil. Para essa análise eles utilizaram variáveis independentes classificadas em quatro categorias, que foram: características das vias, socioeconômicas, uso do sol e exposição.

2.6 ABORDAGENS MCDM/A APLICADAS NA SEGURANÇA VIÁRIA

Diversas abordagens de decisão são propostas na literatura para aplicação no contexto da segurança viária. Uma delas foi apresentada por Martins et al. (2020) que propõe um modelo de decisão multicritério utilizando o método flexível e interativo FITradeoff para problemática de ordenação para priorização de trechos de rodovias federais no estado de Pernambuco. O autor avalia 22 seções de estradas com diferentes características, e realiza um ranqueamento desses trechos de acordo com a criticidade de sinistros que eles apresentam.

Essas alternativas foram avaliadas através de 11 critérios definidos. Com aplicação do método, o autor consegue ordenar as alternativas em quatorze níveis de classificação e, com isso, identificou quais os seis pontos mais essenciais para ações da polícia rodoviária federal.

Martins e Garcez (2021) desenvolveram um modelo de decisão multicritério para priorização de 11 trechos de estradas do estado de Pernambuco, que foram avaliadas através de 11 critérios considerados pelo decisor. Esses critérios englobam diferentes dimensões, como questões sobre os estados da via e a situação do trânsito, consequências aos seres humanos, que podem influenciar na ocorrência de sinistros. Eles aplicaram abordagens de tomada de decisão multiperíodo e multicritério para priorização das estradas, de forma que as autoridades possam direcionar os recursos para os trechos mais críticos.

Kanuganti et al. (2017) apresenta uma metodologia para determinar a prioridade de vias em relação à segurança de trânsito no distrito de Jhunjhunu de Rajasthan, Índia. O autor utilizou as técnicas de decisão multicritério *Additive Weightage* (SAW), *Analytical Hierarchy Process* (AHP) e *Fuzzy AHP*. Os três modelos usados apresentam a mesma seção de estrada como crítica.

Temrungsie et al. (2015) apresenta um modelo de decisão multicritério utilizando o AHP para priorizar informações que venham a contribuir na redução de sinistros. O autor utilizou critérios relacionados a engenharia, economia, ambiental e social, e gestão pública para resolver a problemática. Com isso, os fatores que precisam de maior atenção revelados através da aplicação do AHP são os conhecimentos dos usuários sobre as estradas e regras, e aplicação da lei.

2.7 DISCUSSÃO FINAL DO CAPÍTULO

Os trabalhos explanados anteriormente darão suporte no desenvolvimento do modelo proposto e na utilização das abordagens apresentadas, além de auxiliar na interpretação dos resultados alcançados por essa aplicação. Com isso, esse trabalho possui uma grande contribuição na área da segurança pública, especificamente na segurança das vias de trânsito urbano, pois pretende identificar e priorizar os principais pontos de sinistros de trânsito e com isso, auxiliar as autoridades de trânsito no desenvolvimento de estratégias para redução dessas ocorrências.

3 DADOS E MÉTODOS

Nesta seção serão apresentados os dados, métodos e os conceitos básicos que norteiam o escopo do trabalho, e que são de extrema importância para a aplicação da metodologia proposta para alcançar os objetivos definidos.

3.1 DADOS

O Estudo de caso utiliza um banco de dados com registros de sinistros de trânsito da cidade de Fortaleza-CE para analisar o comportamento dessas ocorrências nas vias dessa localidade. Esses dados são disponibilizados através da plataforma digital VIDA, onde é possível obter informações de sinistros de trânsito da cidade de forma aberta e gratuita, consolidadas pela Autarquia Municipal de Trânsito e Cidadania (AMC).

Esse banco de dados utilizado, contém registros de sinistros de trânsito ocorridos nas vias de Fortaleza desde de 1 de janeiro de 2015 até 31 de dezembro de 2019. São disponibilizadas em formato Excel e contém as informações de parâmetros geográficos do local de ocorrência do sinistro, como o nome da rua, número e localização georreferenciada do mesmo, com a latitude e longitude. Também são disponibilizadas informações temporais, como data, horário de ocorrência.

Outras informações relevantes como natureza do sinistro (Colisão, queda, atropelamento, engavetamento ...), tipo de iluminação, classificação do tempo, tipo de cruzamento, sinalização da pista (Se tinha ou não), superfície da pista (seca, molhada...) e uso de solo. Ainda são disponibilizados números de mortos, feridos, ilesos, e o tipo de veículo envolvido no sinistro, que são informações importantes para as análises.

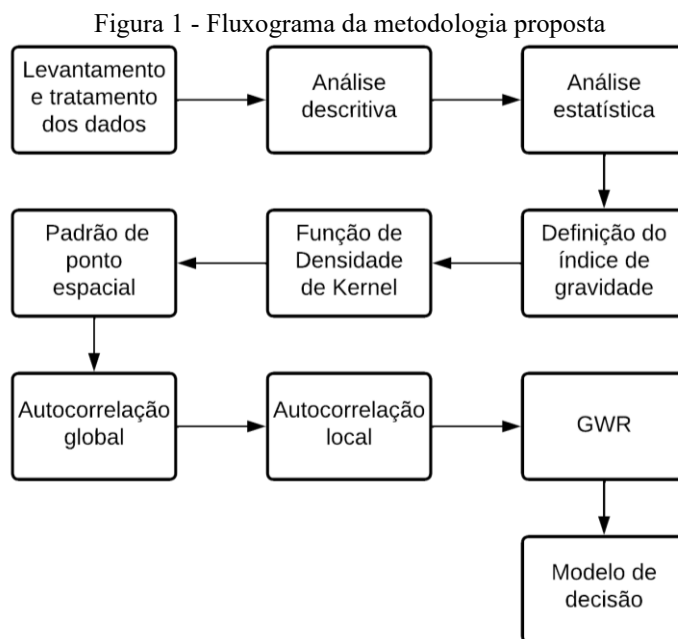
Esses registros de sinistros estão unidos aos segmentos das rodovias através da coordenada geográfica de cada ocorrência, latitude e longitude. Esta pesquisa, realizou a priori uma verificação dos dados para remoção de observações de sinistros com informações georreferenciadas ausentes ou incorretas, como sinistros que estavam localizados fora do limite do município, e também aqueles que não coincidiam com os segmentos de vias.

3.2 ÁREA DE ESTUDO

A área de exploração desse estudo é a cidade de Fortaleza, capital do estado do Ceará. Segundo o IBGE (2016), a cidade possui um território de 312,353 km², uma população estimada de 2.703.391 pessoas no ano de 2021, densidade demográfica de 7.786,44 hab/km², e um total de 1.156.849 veículos em 2020. É a quinta maior cidade do Brasil, e a maior do estado do Ceará. A região metropolitana de Fortaleza é composta por 19 municípios, como Fortaleza, Caucaia e Maracanaú. É a quinta cidade mais populosa do país de acordo com os dados do IBGE (2016), e é cortada pelas rodovias BR- 020, que liga a cidade até a Capital Federal e a BR-116, que é uma das mais importantes do país, fazendo a ligação até as cidades do Sudeste e sul do país.

3.3 METODOLOGIA

A metodologia proposta para a realização dessa pesquisa é composta por 10 etapas que são fundamentais para alcançar os objetivos deste trabalho. A Figura 1 esquematiza a sequência dessas etapas, e nos tópicos abaixo, a descrição de cada um deles.



Fonte: O autor (2022)

- a) Levantamento e tratamento dos dados - a primeira etapa dessa metodologia consiste em levantamento de dados de sinistros de trânsito, através dos registros das autoridades de trânsito, para organizar, analisar e explorá-los nas aplicações seguintes.

- b) Análise descritiva - será realizada uma análise descritiva detalhada das informações do banco de dados, a fim de caracterizar esses sinistros, averiguando quais as principais causas, tipo do sinistro, dias e meses de maiores ocorrências, condições meteorológicas e fase do dia no momento da ocorrência.
- c) Análise estatística - A análise estatística será utilizada, a priori, para definir o melhor intervalo de tempo para a realização da análise espaço-temporal. Para isso, serão contabilizados o número de sinistros de cada mês, cada dia da semana e de cada período do dia (Madrugada, manhã, tarde e noite) entre os anos de 2015 e 2019. Para as três contagens, foram obtidas tabelas de dados que foram utilizadas como inputs para os testes estatísticos de significância.

Dessa forma, essa análise estatística tem por objetivo verificar se há diferença estatisticamente significativa entre o número de sinistros em diferentes períodos analisados. Assim, os dados de cada período serão submetidos aos testes de forma individual.

Esta verificação pode ser realizada através da aplicação da análise da variância, se os dados tiverem uma distribuição normal e as amostras possuírem variâncias iguais. Caso esses requisitos não forem satisfeitos, será utilizado o teste não-paramétrico para a análise, como o teste de Kruskal-Wallis que verifica se as medianas são estatisticamente diferentes ou não. Logo, os dados que possuírem diferenças significativas serão utilizados para verificação de possíveis diferenças espaciais.

- d) Análise de padrão de pontos espaciais - Os testes de análise de padrão de pontos espaciais, serão realizados para verificar se os padrões de distribuição de pontos de sinistros, são semelhantes ou não, e com isso, identificar se há variação com o passar do tempo. Para isso, neste estudo será utilizado o teste de similaridade de padrões espaciais de pontos baseados em áreas proposto por Andresen (2009), que visa encontrar o nível de similaridade entre diversos pontos.

Essa análise, consiste em verificar se dois conjuntos de pontos possuem ou não o mesmo padrão espacial. Nesse caso, será analisado se os padrões espaciais entre os meses em diferentes anos, diferentes anos, entre todos os meses dos anos de 2019, todos os dias da semana do ano de 2019 e diferentes horários do dia, no ano de 2019.

Primeiramente, os dados georreferenciados dos sinistros foram divididos em diferentes arquivos no formato *shapefile*. Logo após, os dois bases dados serão introduzidos no sistema, onde uma delas será considerada como camada de referência e a outra com camada de teste. É adicionada também ao sistema, um base de polígonos, que neste caso, será com

as delimitações dos bairros da cidade. Para os dois padrões serem semelhantes, o valor do índice de similaridade deverá ser de pelo menos 0,8, ou seja, pelo menos 80% das áreas analisadas devem ser similares (ANDRESEN, 2009).

- e) Definição do índice de gravidade do sinistro - O índice de gravidade do sinistro será definido para cada ocorrência e é de extrema importância para a realização das análises espaciais, uma vez que essa informação será utilizada no *input* em cada processo. Ele será definido através de uma abordagem multicritério compensatória.

Para isso, será utilizado o método multicritério aditivo *Simple Multi-Attribute Rating Technique Exploiting Ranks* (SMARTER) proposto por Edwards e Barron (1994), para definir o índice para cada registro. Neste caso, as alternativas serão cada registro de sinistro viário do banco de dados, e os critérios serão considerados o número de mortes, feridos e ilesos envolvidos.

Por fim, será utilizado a abordagem de pesos substitutos *Ranking Ordered Centroid* (ROC), para definir os pesos dos critérios, dada a ordenação desses através de informação parcial do decisor, e assim, será obtido cada índice por meio da soma ponderada entre os valores das alternativas e os critérios. Os pesos dos critérios são obtidos através da equação (3.1) do ROC, onde w_i é o peso do critério na posição i , e n é o número de critérios.

$$w_i = \frac{1}{n} \sum_{j=1}^n \frac{1}{j} \quad (3.1)$$

- f) Função de densidade do kernel - A utilização dessa ferramenta tem como objetivo obter mapas de densidade de calor gerados através dos pontos georreferenciados dos sinistros analisados. Com os mapas de calor, será possível identificar os locais com maiores densidades de sinistros de trânsito em diferentes dias, meses e anos; e assim, verificar também a variação dessa densidade no espaço e no tempo.

Essa análise será realizada utilizando o índice de gravidade definido para cada ocorrência, como *input* no processo de produzir os mapas, assim, esse será gerado com base neste atributo. Também será utilizada a medida de distância inversa nessa análise.

- g) Autocorrelação global (Global Moran's I) - O uso da ferramenta de autocorrelação global (Global Moran 's) propicia a identificação de um padrão espacial para os dados analisados, considerando toda área de estudo, mostrando se esses tendem a estar agrupados, dispersos e aleatórios no espaço. Para o cálculo desse índice, será utilizado o índice de gravidade

como *input* no processo e a distância padrão, que considera que cada evento tenha pelo menos um vizinho.

- h) *Análise Hotspots (Getis-Ord Gi*)* - Essa ferramenta *Hotspots analyses* (Getis-Ord Gi*) será utilizada com o objetivo de identificar a autocorrelação espacial local em áreas específicas do espaço. Dessa forma, no output da análise é disponibilizado um mapa, onde é possível verificar a existências de pontos quentes e frios, estatisticamente significativos no âmbito local, se estes existirem. Essa análise será realizada com base no índice de gravidade dos sinistros e a distância padrão será utilizada.
- i) *Geographically Weighted Regression (GWR)* - A regressão geograficamente ponderada será utilizada a fim de analisar a relação existente entre a quantidade de sinistros e algumas variáveis definidas. Esse modelo de regressão local, será construído com base nas variáveis selecionadas. A Priori é necessário investigar se existe correlação entre as variáveis do modelo, pois a ferramenta utilizada não permite que o mesmo seja implementado caso essa exista essa correlação entre as variáveis utilizadas.

O modelo retornará os coeficientes da regressão, e com isso, será possível relacioná-los às ocorrências de sinistros. A partir deles, será analisado o poder explicativo do modelo completo para cada local, e também, o poder explicativo para o modelo geral. A autocorrelação espacial será aplicada aos resíduos do modelo GWR, para verificar se os mesmos estão aleatórios no espaço.

Como forma de comparação, adotou-se a regressão espacial local GWR para permitir que os parâmetros do modelo variem ao longo dos bairros, visto que são calculados para cada bairro *i* separadamente.

- j) *Modelo de decisão multicritério para ordenar pontos de hotspots de sinistros de trânsito*
 - O modelo de decisão multicritério proposto, será para resolver uma problemática de ordenação, e dessa forma, os trechos críticos identificados através das análises espaciais serão ordenados de acordo com os critérios colocados pelo decisor. Assim, as preferências do decisor será elicitada, e uma abordagem adequada para resolução desse problema será então selecionada.

Como resultados do modelo os *hotspots* identificados através das análise espaço-temporal serão então ordenados em relação a sua criticidade, para que os gestores possam dar mais prioridades a esses locais no desenvolvimento de suas ações. A seguir é apresentado uma descrição dos métodos utilizados.

3.4 MÉTODOS

Nesta seção, será apresentado os conceitos básicos e descrição dos métodos que serão utilizados nesta pesquisa, fundamentais para alcançar os objetivos definidos para o trabalho.

3.4.1 Sistema de informação geográfica

Uma variedade de métodos tem sido desenvolvida para o estudo da distribuição espacial de sinistros, e identificação de pontos focais de sinistros de trânsito para uma região definida (AL-OMARI et al., 2020). Nesse contexto, de acordo com o autor, as ferramentas de sistemas de informação geográfica são capazes de realizar essa análise espacial.

De acordo com Sjafrri et al. (2018), o GIS é um sistema desenvolvido para facilitar a análise de dados espaciais, e dessa forma ele é capaz de reter, memorizar, averiguar, integrar, analisar e expor dados geográficos de um local determinado. Yaacob et al. (2020) expõe que o GIS é uma ferramenta que está sendo bastante empregada na análise de sinistros de trânsito, e através de seu uso é possível fornecer resultados rápidos aos tomadores de decisão, facilitando dessa forma suas decisões e solução de problemas.

Nesse cenário, Nasir, Maulud e Yusoff (2016) traz que o GIS pode ser visto como a ferramenta mais adequada para aprimorar as operações de gestão, uma vez que ela pode ser ajustada com a configuração geográfica das redes rodoviárias, tendo bastante utilidade para solução de problemas relacionados a essas redes. Essa abordagem auxilia na obtenção de respostas para problemas permitindo ao usuário uma visualização e interpretação de dados de maneira mais fácil, para que eles possam compreender os resultados facilmente (NASIR; MAULUD; YUSOFF, 2016). De acordo com Burrough et al. (2015) o GIS é relevante para a análise da situação atual de transporte, porque auxilia na identificação de melhorias e na construção de sistemas de apoio à decisão.

3.4.2 Modelo de dados espaço-temporal

De acordo com Dagaeva et al. (2019) o modelo de dados espaço-temporal constitui-se de duas informações importantes. São elas:

Espacial - que consiste em informações geográficas a respeito do acontecimento ou incidente em relação à local onde se encontra a área. Dessa forma, a localização pode ser

fornecida em termos de coordenadas GPS ou GLONASS (latitude e longitude) pelas vítimas do sinistro.

Temporal - que consiste em informações relacionadas ao momento de ocorrência do evento como o dia e hora do acontecimento.

Ainda segundo o autor, em relação a análise de dados de emergência, esses podem ser evidenciados através das seguintes abordagens para a mineração de dados espaço-temporal:

- (i) Agrupamento espaço-temporal a fim de identificar as áreas de alto interesse;
- (ii) Mineração de padrões de co-localização espaço-temporal a fim de identificar novas relações espaço-temporal através de um alto volume de dados;
- (iii) Identificação de outliers espaço-temporais;
- (iv) Análise de autocorrelação espacial para identificar a relação entre eventos.

As estatísticas policiais reúnem os dados a respeito do evento em nível de ponto espacialmente localizados e a partir dessas é possível realizar uma análise espaço-temporal com a integração do GIS (LAKES, 2017). Ainda com esse o autor aprendemos que as análises espaciais e temporais com a ajuda do GIS, auxiliam na identificação de possíveis focos e causas dos sinistros de trânsito, assim como na formação de recomendações para tal.

3.4.3 Análise de Cluster

A análise de cluster consiste em uma ferramenta estatística que permite a exploração de agrupamentos não aleatórios que ocorrem em um resultado, conforme Francis et al. (2020). Segundo Pasupathi et al. (2021), essa análise é realizada utilizando objetos semelhantes que podem ser agrupados apenas em um agrupamento ou em diversos grupos, de acordo com suas propriedades e atributos. Assim, esse agrupamento classifica os objetos semelhantes no mesmo cluster, enquanto que, os objetos diferentes são categorizados em clusters diferentes.

A análise de cluster tem como objetivos realizar o agrupamento de coleções padrão, baseado na medida de similaridade, e o resultado do cluster será definido através dessa medida (TAIN; ZHANG, 2020). Assim, os dados que estão agrupados no mesmo cluster terão uma similaridade maior do que os dados que estão agrupados em outro cluster.

Os métodos de análise de cluster podem ser divididos em hierárquicos, que são aqueles que fazem o agrupamento de objetos de dados na forma de uma estrutura de árvore; e os métodos não hierárquicos, que classificam os n dados em k cluster pré-estabelecidos, de forma

que em cada cluster tenha pelo menos objeto, e que cada objeto esteja pelo menos em um cluster, conforme citado por Selvi e Caglar (2018).

A análise de cluster é apropriada para explorar doenças infecciosas ou que são capazes de sofrer influência de fatores ambientais (FRANCIS et al., 2020). Entretanto, ele é bastante utilizado na previsão de sinistros de trânsito.

3.4.4 Estimativa de densidade de Kernel

A estimativa de densidade de Kernel (KDE) é utilizada para prever função de densidade de probabilidade (LI; LIANG, 2018). O método consiste em adicionar uma superfície simétrica sobre cada ponto, e examinar a distância entre esse ponto e um local considerado como referências, através de uma função, e com isso, os resultados de todas as superfícies são somados para esse local de referência, e o processo é repetido para os demais pontos consecutivos (ANDERSON, 2009; SHAFABAKHSH; FAMILI; BAHADORI, 2017),

Dessa forma, a função de estimativa de densidade geral é calculada pela expressão (3.2) apresentada a seguir (LE; LIU; LIN, 2020b; ANDERSON, 2009; SHAFABAKHSH; FAMILI; BAHADORI, 2017):

$$f(x, y) = \frac{1}{nh^2} \sum_{i=1}^n K\left(\frac{d_i}{h}\right) \quad (3.2)$$

Onde n representa o número total de observações, K a função de Kernel, $f(x, y)$ é a estimativa de densidade em (x, y) , h a largura de banda, d_i é a distância (x, y) e o local da observação h . Os principais parâmetros que determinam o resultado do método são a largura de banda e o tamanho de célula (LE; LIU; LIN 2020b).

A principal vantagem da utilização da estimativa de densidade de kernel é que ele consegue determinar a extensão do risco de um sinistro (ANDERSON, 2009; SHAFABAKHSH; FAMILI; BAHADORI, 2017). A propagação desse risco, diz respeito a área que está ao redor de um cluster crítico para as ocorrências, onde existe uma grande probabilidade de ocorrer sinistros com base na dependência espacial Anderson (2009). Além disso, o método é aplicado na identificação de padrões espaciais, mostrando mapas de superfície contínua e discretas evidenciando os locais de alto risco (HASHIMOTO et al., 2016).

3.4.5 Autocorrelação espacial

A autocorrelação espacial é o grau de dependência espacial entre o valor de uma única variável em um determinado local e os valores da mesma variável em uma localização vizinha. Segundo a primeira lei da Geografia, os objetos que estão em uma vizinhança mais próxima, têm a tendência de serem mais semelhantes e interativos, em relação às observações que estão a uma distância maior (DALE; FORTIN, 2014).

Antes mesmo de realizar qualquer análise estatística mais avançada com dados espaciais, deve ser realizado um teste de autocorrelação espacial global ou local, e dessa forma verificar se existe alguma relação entre os valores das observações próximas e se elas configuram um padrão espacial (DALE; FORTIN, 2014).

Para identificar a autocorrelação espacial pode-se utilizar de medidas de autocorrelação espacial global e local. As medidas globais avaliam esse grau de associação através de um valor único, estimado para toda área de estudo e, dessa forma, avalia se a distribuição tem uma tendência de agrupamento, dispersão ou aleatória. As medidas mais utilizadas para esse fim, são: *Moran's I index*, *General G-Statistic* e *Geary's C index* (DALE; FORTIN, 2014).

Entretanto, estimando apenas a autocorrelação global não é suficiente para afirmar que essa autocorrelação espacial ocorrerá em toda área do estudo. Ele pode ser mais evidente em algumas áreas do que em outras, em virtude da heterogeneidade do espaço. Desse modo, pode-se utilizar medidas locais para avaliar o grau de dependência no nível local que caracterizam a heterogeneidade com a identificação de pontos quentes ou frios, *clusters* e *outliers*. As medidas mais comuns utilizadas para esse fim, são: *Local Moran's I index* e *Getis-Ord Gi and G* i statistics* (DALE; FORTIN, 2014).

A autocorrelação espacial pode ser positiva, que ocorre quando uma variável de atributo do espaço analisado está associada com ela própria em uma distância determinada, e dessa forma os dados que estão mais próximos um dos outros possivelmente são mais parecidos quando se comparados aos dados mais distantes, mostrando que a localização influencia na variável formando agrupamentos em algumas regiões do estudo. Enquanto isso, na correlação negativa os valores espaciais próximos tendem a ser diferentes. Se aparentemente as evidências apontam para a não associação entre os dados, há uma distribuição aleatória em relação à localização (DALE; FORTIN, 2014).

3.4.6 Autocorrelação espacial global

O índice global de Moran mensura a autocorrelação espacial de atributos utilizando a estatística de Moran e analisa o padrão dos eventos mostrando se são agrupados, dispersos ou aleatórios. A ferramenta de autocorrelação espacial realiza o cálculo desse índice e retorna escore Z e valor de p que revelam se o valor do índice encontrado é estatisticamente significativo ou não (SANDHU et al., 2016).

O valor desse índice varia entre -1 e +1. Valores positivos mais próximos de 1 indica um grau maior de agrupamento espacial, enquanto que valores negativos revelam um padrão espacial mais disperso e o valor 0 caracteriza um padrão espacial distribuído de forma aleatória (SOLTANI; ASKARI, 2017).

A estatística de I de Moran é calculada de acordo com a equação (3.3) (FISCHER; GETIS, 2010):

$$I = \frac{N \sum_{i=1}^N \sum_{j=1, j \neq i}^N w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{S_0 \sum_{i=1}^N (x_i - \bar{x})^2} \quad (3.3)$$

Onde S_0 é dado por:

$$S_0 = \sum_{i=1}^N \sum_{j=1, j \neq i}^N w_{ij} \quad \forall_i = 1, \dots, n; \quad \forall_j = 1, \dots, n; \quad (3.4)$$

Onde w_{ij} é uma matriz de contiguidade binária espacial onde estão os pesos espaciais que representam a relação de proximidade entre o local i e seu vizinho j ; x_i é o valor da variável em um determinado local i ; x_j representa o valor da variável em outro local (i/j); N representa a totalidade das localizações; $\bar{x}$ a média da variável e S_0 é a soma de todos elementos que estão na matriz de contiguidade binária espacial W (LI; LIANG, 2018).

A padronização da estatística I de Moran global para valores de escore Z , é representada através da expressão (3.5) seguinte:

$$Z = \frac{I - E(I)}{\sqrt{\text{var}(I)}} \quad (3.5)$$

Onde $E(I)$ é o valor do índice esperado para um padrão aleatório e $\text{var}(I)$ corresponde à variância (SOLTANI; ASKARI, 2017).

O escore Z e valor p são estatisticamente significativos e mostram se a hipótese nula que geralmente indica que o padrão observado é aleatório, será aceita ou rejeitada (MANAP et al., 2021).

3.4.7 Análise de Hot Spot

A análise de Hot Spot é um recurso utilizado para localizar clusters espaciais e identificar pontos quentes e frios estatisticamente significativos utilizando a estatística de Getis-Ord G_i^* . Através do valor dessa estatística será possível identificar os pontos quentes que são indicados por um alto valor desse índice mostrando que os atributos com alto valor estão agrupados, e os pontos frios que são indicados por agrupamentos de valores menores desse índice (SOLTANI; ASKARI, 2017).

Essa estatística é calculada para cada observação do conjunto de dados levando em conta a relação de cada ponto com as informações das observações da vizinhança, e assim o método gera esses valores através da comparação entre esses recursos (ERSEN; BÜYÜKLÜ; TAŞABAT, 2021). Com isso, o método pode identificar clusters com valores altos ou baixos. Essa ferramenta apresenta uma pontuação Z e valores de p , que são medidas de significância estatística de agrupamento (SANDHU et al., 2016). Assim, o valor da estatística Getis-Ord G_i^* pode ser determinada através da equação (3.6) (MANAP et al., 2021; ERSEN; BÜYÜKLÜ; TAŞABAT, 2021; TOLA et al., 2021).

$$G_i^* = \frac{\sum_{j=1}^n w_{ij}x_j - \bar{X} \sum_{j=1}^n w_{ij}}{S \sqrt{\frac{n \sum_{j=1}^n w_{ij}^2 - (\sum_{j=1}^n w_{ij})^2}{n-1}}} \quad i \neq j \quad (3.6)$$

Onde $\bar{X}$ e S são dados através das equações (3.7) e (3.8), respectivamente:

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n} \quad (3.7)$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2} \quad (3.8)$$

Onde w_{ij} corresponde à matriz de peso espacial para todas as ligações entre as observações que estão dentro de uma determinada distância d do local j , com j ; x_j é o valor do atributo para a observação j , d é o raio de banda fixa ao redor de i , e n é o número de observações ponderadas.

Uma pontuação positiva de G_i^* indica, a existência de *hotspots*, enquanto que valores negativos se referem a *clodspots*, e assim, quanto mais altos forem os valores de G_i^* , mais intenso será o agrupamento espacial de valores altos, e quando menor o valor de G_i^* , a tendência

é de um grupamento espacial intenso de valores baixos na determinada região (MANAP et al., 2021; ERDOGAN, 2009).

3.4.8 Geographically weighted regression

A regressão geograficamente ponderada é um tipo de modelo de regressão linear local utilizado para verificar heterogeneidade e as relações espaciais variáveis (FOTHERINGHAM; BRUNSDON; CHARLTON, 2002). É uma técnica que consiste da extensão da estrutura de regressão tradicional, e permite variação local das taxas de mudanças, de forma que, os coeficientes do modelo são especificados para cada local i (BRUNSDON; FOTHERINGHAM; CHARLTON, 1996). Segundo Liu, Wu e Wang (2019), o modelo GWR é eficaz para verificar a não estacionariedade espacial.

A equação (3.9) representa a equação da regressão, que é dada por:

$$y_i = a_{i0} \sum_{k=1,m} a_{ik} x_{ik} + \varepsilon_i \quad (3.9)$$

Onde, i é o número total de locais observados, y_i corresponde ao valor da k -ésima variável dependente na localidade i , a_{i0} é interceptação do modelo em i , a_{ik} é o coeficiente de regressão da k -ésima variável independente no local i , x_{ik} é o valor da k -ésima variável independente em i , m é o número de variáveis dependentes do modelo, e ε_i o erro aleatório no local i .

O modelo GWR, explica uma variável dependente com base em uma função que é variável de variáveis associativas do modelo, e assim, em cada observação da superfície analisada, serão definidos os coeficientes utilizando a ponderação das variáveis independentes de acordo com a distância (RENNERMALM; BRING; MOTE, 2012). O modelo GWR é um modelo de regressão espacial que gera coeficientes espacialmente contínuos de todas as variáveis na área de estudo (JIANG et al., 2017). As funções de decaimento de distância de Poisson, Gaussiana e Logística, podem ser utilizadas no GWR (MOHAMMADINIA et al., 2019).

Em geral, nos modelos de regressão ponderados são considerados os valores de w_i como constantes, e assim, a calibração para encontrar o conjunto de estimativas de coeficientes, necessita de ser realizada apenas uma vez (BRUNSDON; FOTHERINGHAM; CHARLTON, 1996). No modelo GWR, a observação será ponderada através da proximidade que esta tem do ponto i , de tal forma que, a ponderação de um local observado seja variável, de acordo com o

local i , e assim não terá mais a calibração como constante (FOTHERINGHAM; CHARLTON; BRUNSDON, 1998). Além disso, os dados das observações que estiverem mais próximas de i , serão mais ponderados, do que as demais que estiverem mais distantes. Dado isso, tem-se que:

$$a_i = (X^T W_i X)^{-1} X^T W_i Y \quad (3.10)$$

Onde a_i é o coeficiente de cada local, X corresponde a matriz com os valores da variável independente, Y um vetor n por um da observação dependente, W_i representa uma matriz n por n que possui os valores da ponderação geográfica no ponto i localizadas na diagonal dessa matriz, e o restante dos elementos zero.

$$W_i = \begin{bmatrix} w_{i1} & 0 & 0 & 0 \\ 0 & w_{i2} & 0 & 0 \\ 0 & 0 & w_{ij} & 0 \\ 0 & 0 & 0 & w_{in} \end{bmatrix} \quad (3.11)$$

Onde i corresponde ao ponto no espaço no qual será estimado os parâmetros entre o local j , que por sua vez, é outro ponto no espaço cujos dados são observados. w_{ij} é o peso entre o ponto i e j e varia entre $[0,1]$. Diversos sistemas de ponderação podem ser utilizados, por exemplo, como iniciar essa ponderação baseada na localidade, retirando do modelo as observações de calibração que não estão dentro do limite de distância d do ponto (FOTHERINGHAM; BRUNSDON; CHARLTON, 2002). Assim:

$$w_{ij} = 1, \text{ se } d_{ij} < d \quad (3.12)$$

$$w_{ij} = 0, \text{ se } d < d_{ij} \quad (3.13)$$

Neste caso é um sistema binário, e se a distância d_{ij} for maior que d , $w_{ij} = 0$, e se d_{ij} for menor que d , $w_{ij} = 1$.

De acordo com Fotheringham, Brunson e Charlton (2002), os núcleos espaciais são utilizados frequentemente no cálculo desses pesos. Dessa forma, o kernel espacial utilizado para realizar a ponderação geográfica deve ser designado antes mesmo da calibração do modelo GWR. Segundo o autor, as funções de Kernel fixa e adaptativa, são as mais utilizadas no modelo GWR.

O tipo de Kernel fixo mais utilizado é o gaussiano, para calcular os pesos locais (IBEAS et al., 2011). Assim, ele é calculado de acordo com a equação (3.14) de acordo com Fotheringham, Brunson e Charlton (2002):

$$w_{ij} = \exp \left[-\frac{1}{2} \left(\frac{d_{ij}}{h} \right)^2 \right] \quad (3.14)$$

onde d_{ij} corresponde à distância existente entre os locais i e j , e h largura de banda não variável do Kernel. Já no kernel adaptativa, a função bi-quadrado pode ser utilizada para cálculo dos pesos locais com base em N , que corresponde o número de vizinhos mais próximos, dado por Fotheringham, Brunsdon e Charlton (2002):

$$w_{ij} = \exp \left[-1 - \left(\frac{d_{ij}^2}{d_{iN}^2} \right)^2 \right] \quad (3.15)$$

Caso o ponto j está entre os N -ésimos pontos mais próximos de i , ou $w_{ij} = 0$ se isso não acontece (FOTHERINGHAM; BRUNSDON; CHARLTON, 2002).

3.4.9 Teste de similaridade em padrões espaciais baseados em área

O teste de similaridade de padrões espaciais de pontos baseados em áreas, proposto por Andresen (2009), está fundamentado através de um teste de padrão de pontos espaciais que utiliza uma abordagem de Monte Carlo não paramétrica. Com esse teste, é possível mensurar o grau de similaridade existente entre dois padrões de pontos espaciais

O teste descrito por Andresen (2009) é composto de 11 etapas. Primeiramente, duas bases de dados de pontos são obtidos a fim de verificar a similaridade entre eles. Essas bases serão classificadas, uma como o conjunto de dados de referências e a outra, como conjunto de dados de teste. Logo após, os pontos da base de referência serão contados dentro de cada unidade espacial definida para a análise.

A posteriori, serão selecionados aleatoriamente uma amostra contendo 85% dos pontos da base de dados de teste, afim de preservar o padrão espacial dos pontos dessa base e então esses dados serão agregados por unidade de fronteira, e o processo repetido por 200 vezes. Os dados gerados através das amostras aleatórias serão coletados, e então, será definido um intervalo de confiança não paramétrico de 95%, para cada uma das unidades espaciais estabelecidas. Por fim, o índice de similaridade pode ser calculado através da Equação (3.16) e a saída dos dados pode ser mapeada (ANDRESEN, 2009).

$$S = \frac{\sum_{i=1}^n s_i}{n} \quad (3.16)$$

Quando s_i for 1, as duas bases utilizadas são consideradas semelhantes na unidade espacial i , e será 0, caso o contrário ocorra, e n será o número de unidades espaciais utilizadas. S será o índice de similaridade, medindo o grau de similaridade no nível local, indicando a

porcentagem de unidades espaciais semelhantes para os dois conjuntos de dados (ANDRESEN, 2009).

3.4.10 FITradeoff para problemática de ordenação

O método de elicitação FITradeoff para problemática de ordenação proposto de Frej et al. (2019), é uma abordagem para resolver problemas de ordenação com base nas informações parciais do decisor, através de um processo de elicitação interativo e flexível. Segundo a autora, a problemática de ordenação corresponde à alocação das alternativas dos problemas em uma ordem crescente de preferência, por meio de um modelo de preferência.

Nesse método, são realizadas comparações par-a-par entre as alternativas e em cada comparação as relações de dominâncias entre elas são verificadas. Um modelo de programação linear será realizado a fim de buscar uma relação de dominância entre cada par de alternativas comparado (FREJ et al., 2019).

A modelagem de preferência do decisor é realizada através de um sistema de apoio à decisão FITradeoff. O processo inicia com o input da matriz de consequências no sistema. Na primeira etapa de elicitação do FITradeoff, o decisor deverá ordenar as constantes de escala dos critérios, através de uma avaliação holística ou mecanismo de comparação par a par (FREJ et al., 2019).

Com o ranqueamento dos critérios, é formado um espaço de peso inicial, que já incorpora uma restrição no problema de programação linear, para conseguir obter a relação de dominância entre as alternativas. Com essa etapa, já é possível obter uma ordenação completa ou parcial das alternativas através das relações de dominância encontradas, e o resultado pode ser visualizado através de um diagrama (FREJ et al., 2019).

Neste momento, o decisor pode verificar se o resultado obtido já é satisfatório ou não, podendo parar o processo ou dar continuidade. Caso o decisor continue o processo suas preferências serão avaliadas por meio da comparação par a par entre as alternativas, considerando a troca entre diferentes critérios. O processo pode ser encerrado quando a ordem completa for alcançada, se a recomendação em uma determinada interação for o suficiente para o decisor (FREJ et al., 2019).

A cada comparação par-a-par efetivada novas restrições são incorporadas ao problema de programação linear e através da relação de dominância definida por ele o ranqueamento das alternativas é então formado (FREJ et al., 2019). E assim o espaço de peso é atualizado a cada

interação, ficando mais apertado. Dessa forma, para cada par de alternativas comparadas (A_i, A_k) , o seguinte problema de programação linear é executado:

$$\max D(A_i, A_k) = \sum_{j=1}^m w_j v_j(A_i) - \sum_{j=1}^m w_j v_j(A_k) \quad (3.17)$$

$$\forall i \text{ e } j, \text{ onde } i \neq j$$

s. t.

$$w_1 > w_2 > \dots > w_m \mid \sum_{j=1}^m w_j = 1 \quad (3.18)$$

$$w_j v_j(x'_j) > w_{j+1} \quad j = 1 \text{ to } m - 1 \quad (3.19)$$

$$w_j v_j(x''_j) < w_{j+1} \quad j = 1 \text{ to } m - 1 \quad (3.20)$$

$$w_j \geq 0, \quad j = 1 \dots m \quad (3.21)$$

Onde A_i corresponde a alternativa i , e A_k a alternativa k ; w_j é a constante de escala (ou pesos) do critério j ; $v_j(A_i)$ é o valor da alternativa A_i no critério j ; $v_j(A_k)$ é o valor da alternativa A_k no critério j ; m é o número de critérios; x'_j e x''_j são valores das consequências no critério j (FREJ et al., 2019).

Na equação (3.16) é calculado a diferença máxima entre os valores das duas alternativas. A equação (3.17) mostra a ordenação das constantes de escala dos critérios. A (3.17) mostra a equação que representa a normalização das constantes de escalas, a equação (3.18) mostra o limite máximo da constante de escala e a equação (3.19) o limite mínimo (FREJ et al., 2019).

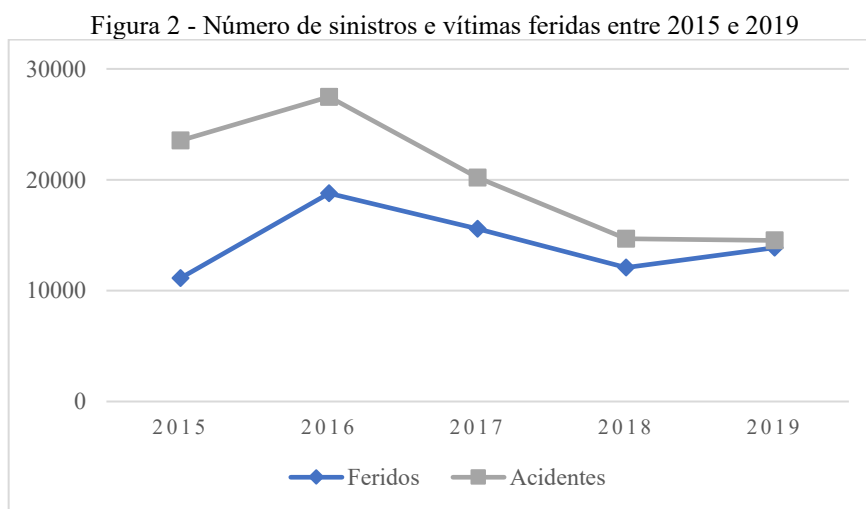
4 MODELO DE ANÁLISE ESPACIAL: ESTUDO DE CASO COM DADOS DE SINISTROS DA CIDADE DE FORTALEZA

Neste capítulo será realizada uma investigação minuciosa dos dados de sinistros de trânsito da cidade de Fortaleza. Eles serão analisados descritivamente para melhor compreender a dinâmica dessas ocorrências ao longo do tempo e espaço. Além do mais, foram realizadas algumas análises exploratórias com a utilização de ferramentas estatísticas para entender o padrão espacial dessas ocorrências em diferentes períodos, identificar os locais com maior densidade dos sinistros e também, os *hotspots* de sinistros.

Uma análise do padrão de pontos espacial será realizada para identificar a existência de similaridade entre a distribuição espacial dos sinistros de trânsito em diferentes períodos, e assim, entender se a distribuição dessas ocorrências é semelhante ou não, entre os períodos analisados. Com essas informações, será possível desenvolver políticas de retenção mais efetivas para essas localidades.

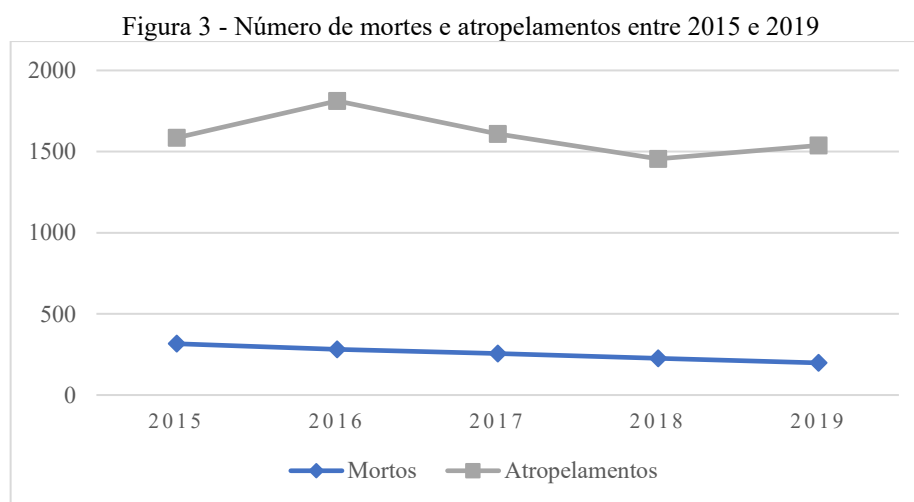
4.1 ANÁLISE DESCRITIVA

Os dados de sinistros de trânsito analisados nesta pesquisa compreendem o período de 2015 a 2019. Entre o intervalo analisado, o ano de 2016 possui a maior quantidade de registro, com um total de 27.471 sinistros, como mostrado na Figura 2. O número de sinistros vem sendo menor a cada ano desde de 2017, e o número de mortes vem se reduzindo desde 2015, como elucidado no Figura 3.



Fonte: O autor (2022)

A maior parte desses sinistros tem automóveis e motocicletas envolvidos, sendo que, nos anos de 2018 e 2019 a maior parcela desses sinistros envolvem motociclistas. As colisões entre veículos, é a natureza de sinistros mais comum em ambos os anos. Além disso, as quedas e atropelamentos são expressivos nesses períodos da análise. A maior parte das ocorrências ocorrem em meio de quadra, entretanto, o número de ocorrências em interseções é muito elevado.



Fonte: O autor (2022)

4.2 ANÁLISE DE VARIÂNCIA

A fim de verificar a existência de diferenças significativas entre as observações de sinistros, foi aplicado o teste de análise de variância não paramétrico de Kruskal-Wallis. A utilização deste teste é justificada pois, algumas amostras dos dados analisados não atendiam o critério de normalidade, além disso, para os dados de meses e dias da semana, as variâncias não eram iguais, dessa forma a ANOVA não poderia ser utilizada, pois seus critérios fundamentais não estavam sendo verificados nos dados dessa análise, assim, os dados analisados atendem os critérios de utilização do teste de Kruskal-Wallis.

O resultado da análise de variância mostrado na Tabela 1, indica que não existe uma diferença estatisticamente significativa para as ocorrências de sinistros em diferentes períodos do dia nos anos analisados, apesar do período da noite possuir o maior número de registros. Entretanto, para os dias da semana e meses, foi verificada a existência de uma variação estatisticamente significativa, mostrando que existe diferença entre a quantidade de ocorrência

nos diferentes meses e dias da semana. Os testes estatísticos foram realizados com o auxílio do software *past*.

Tabela 1 - Resultado do teste de Kruskal-Wallis

Parâmetros	Meses	dias	Período do dia
$H (X^2)$	51,13	28,69	5,414
p	0,00001	0,00001	0,2474

Fonte: O autor (2022)

4.3 ANÁLISE DE PADRÃO DE PONTOS ESPACIAIS

Os testes de análise de padrão de pontos espaciais, foram realizados para verificar a existência de similaridade entre os padrões de pontos espaciais de sinistros viário para diferentes períodos. Essa análise foi realizada através das áreas dos bairros da cidade de Fortaleza.

O processo foi iniciado verificando a similaridade par-a-par entre a distribuição dos sinistros de 2015 a 2019 para detectar ou não, a existência de um padrão similar entre os anos. Em seguida a mesma análise foi realizada para verificar esse padrão para cada mês em diferentes anos e, assim, concluir se determinado mês tem uma distribuição similar ao longo desses anos.

O resultado dessa análise, mostrado na Tabela 2, indica o valor do índice calculado para as comparações par-a-par entre os anos. Esse índice, mostra a proporção dos bairros que possuem um padrão de pontos semelhante nessa comparação, e com isso, a similaridade de padrão de pontos espaciais é assumida entre eles se o valor desse índice for de pelo menos 0,8 ou seja, se pelo que menos 80% dos bairros tenham um padrão similar.

Tabela 2 – Similaridade de sinistros entre os anos

Ano	2015	2016	2017	2018
2015				
2016	0,3223			
2017	0,2975	0,2810		
2018	0,2149	0,2149	0,3140	
2019	0,1240	0,1736	0,1736	0,2893

Fonte: O autor (2022)

Os resultados da análise realizada para os anos, demonstra que não há similaridade significativa entre a padrão da distribuição dos sinistros nesses anos. A maior similaridade foi

registrada entre os anos de 2015 e 2016, onde 32,23% dos bairros tem um padrão semelhante, entretanto, não foi o bastante para a área inteira ser significativa. A Tabela 3, mostra os resultados dos índices calculados para os meses de janeiro e fevereiro.

Tabela 3 - Índice de similaridade de sinistros nos meses de janeiro e fevereiro

Ano	2015	2016	2017	2018	2019
2015	-	0,3306	0,3140	0,3802	0,3719
2016	0,3140	-	0,4132	0,3802	0,3967
2017	0,4050	0,3554	-	0,3802	0,3223
2018	0,3719	0,3471	0,3388	-	0,3306
2019	0,3554	0,2975	0,3140	0,4298	-

Fonte: O autor (2022)

Os resultados para o mês de janeiro, apresentados no lado inferior esquerdo da tabela, revelam que não há similaridade significativa entre nenhum dos anos, e dessa forma o mês em análise possui um padrão diferente para cada ano de estudo. A similaridade máxima encontrada foi de 42,98%, entre os anos de 2018 e 2019.

O mês de fevereiro também não obteve nenhum índice próximo a 0,80, mostrando ter um padrão diferente para cada ano, como pode ser visualizado através dos resultados no lado superior direito da tabela.

A Tabela 4, mostra os resultados para os meses de março (lado inferior esquerdo) e abril (lado superior direito). Em março, nenhum padrão foi considerado significativo, e este obteve o maior índice entre os anos de 2017 e 2018. Para abril, o maior índice foi observado entre os anos de 2016 e 2018, não tendo similaridade entre o padrão de pontos entre os anos.

Tabela 4 - Índice de similaridade de sinistros nos meses de março e abril

Ano	2015	2016	2017	2018	2019
2015	-	0,3471	0,3554	0,3636	0,2975
2016	0,3802	-	0,3471	0,4298	0,3884
2017	0,3884	0,3802	-	0,3719	0,3719
2018	0,3719	0,4132	0,4380	-	0,3058
2019	0,3058	0,3884	0,2893	0,3719	-

Fonte: O autor (2022)

Os meses de maio e junho também não apresentaram nenhum padrão de pontos de sinistro de trânsito similares entre os anos analisados. Para maio, o maior valor do índice foi entre os anos de 2017 e 2018, com 38,84% de semelhança, como pode ser visto do lado inferior

esquerdo da Tabela 5. Já para junho, a maior similaridade foi entre os anos 2015 e 2018, como mostrado na parte superior direita da mesma tabela.

Tabela 5 - Índice de similaridade de sinistros nos meses de maio e junho

Ano	2015	2016	2017	2018	2019
2015	-	0,3058	0,3140	0,4132	0,3306
2016	0,3388	-	0,3306	0,3967	0,3223
2017	0,3719	0,3223	-	0,3388	0,2893
2018	0,3554	0,2645	0,3884	-	0,3223
2019	0,2893	0,3140	0,3388	0,2810	-

Fonte: O autor (2022)

Para os meses de julho e agosto os índices registrados nas comparações também não foram significativos. Em julho, esses índices variam entre 29,75% e 43,80%, como é ilustrado na parte inferior esquerda da Tabela 6. Já em agosto, a variação é entre 28,93% e 41,32%, como mostra o lado superior direito da tabela abaixo.

Tabela 6 - Índice de similaridade de sinistros nos meses de julho e agosto

Ano	2015	2016	2017	2018	2019
2015	-	0,3471	0,3223	0,3884	0,4050
2016	0,3471	-	0,3967	0,2893	0,3884
2017	0,3306	0,2975	-	0,3967	0,4132
2018	0,4380	0,4132	0,4628	-	0,3140
2019	0,3884	0,3223	0,2975	0,3554	-

Fonte: O autor (2022)

Já a Tabela 7, mostra os resultados para os meses de setembro situados na parte inferior esquerda e de outubro no lado superior direito. Ambos os meses, não possuem índice de similaridade significativo. O setembro possui o maior valor de similaridade entre os dois meses, entre os anos de 2016 e 2016, registrando uma semelhança de 44,63%.

Tabela 7 - Índice de similaridade de sinistros nos meses de setembro e outubro

Ano	2015	2016	2017	2018	2019
2015	-	0,3223	0,3388	0,3719	0,3471
2016	0,3388	-	0,3471	0,3802	0,3554
2017	0,3802	0,4463	-	0,3140	0,3306
2018	0,3636	0,4132	0,4050	-	0,2893
2019	0,2727	0,2975	0,3554	0,3140	-

Fonte: O autor (2022)

E para finalizar a análise de similaridade de padrão de pontos de sinistros, os meses de novembro e dezembro também não registraram nenhum índice significativo. Para novembro, esses valores estão sempre entre 26,45% a 38,84%, como mostrado na parte inferior esquerda da Tabela 8. Já em dezembro, esse índice varia entre 31,40% e 42,15%, como mostrado no lado superior direito da tabela abaixo.

Tabela 8 - Índice de similaridade de sinistros nos meses de novembro e dezembro

Ano	2015	2016	2017	2018	2019
2015	-	0,3884	0,3719	0,3306	0,3223
2016	0,3223	-	0,4215	0,4132	0,3967
2017	0,3884	0,3967	-	0,3140	0,3554
2018	0,3636	0,3719	0,3388	-	0,3554
2019	0,2645	0,3058	0,3719	0,3719	-

Fonte: O autor (2022)

Outra análise foi realizada considerando apenas os dados georreferenciados de sinistros de trânsito do ano de 2019. Assim, a verificação do padrão espacial foi feita para diferentes períodos do dia (madrugada, manhã, tarde e noite), meses e dias da semana. A Tabela 9, mostra os índices de similaridades calculados para diferentes períodos do dia.

Tabela 9 - Índice de similaridade de sinistros para o período do dia

Período	Madrugada	Manhã	Tarde	Noite
Madrugada				
Manhã	0,1818			
Tarde	0,1322	0,3306		
Noite	0,2066	0,3058	0,3388	
Total	0,1322	0,2810	0,3802	0,3471

Fonte: O autor (2022)

Na comparação dos períodos do dia, verificou-se que não há similaridade entre nenhum dos pares, mostrando que os padrões de pontos são diferentes em cada um dos horários. Nessa análise, o padrão de pontos do período da tarde e o do total de pontos obteve um índice resultante de 0,3802, mostrando que 38,02% dos bairros tem similaridade no padrão de pontos espaciais, sendo este o maior índice observado. A Tabela 10, mostra os resultados dos índices calculados entre os dias da semana.

Tabela 10 - Índice de similaridade de sinistros entre os dias da semana

Dia da semana	Segunda	Terça	Quarta	Quinta	Sexta	Sábado	Domingo
Segunda	-						
Terça	0,41322						
Quarta	0,29752	0,42975					
Quinta	0,43802	0,40496	0,38017				
Sexta	0,33884	0,40496	0,41322	0,33058			
Sábado	0,28099	0,33058	0,28099	0,23967	0,37190		
Domingo	0,30579	0,34711	0,32231	0,27273	0,37190	0,40496	
Todos	0,20661	0,25620	0,22314	0,18182	0,19835	0,19835	0,24793

Fonte: O autor (2022)

Na análise de similaridade entre os pares de dias da semana, foi observado também que não há nenhum índice igual ou superior 0,80, mostrando que cada dia da semana possui um padrão espacial de pontos diferentes. As comparações entre os dias segunda-feira e quinta-feira obteve o maior índice de similaridade, revelando que 41,32% dos bairros possuem semelhança entre o padrão de pontos espaciais. Ainda nessa análise, o índice mínimo observado foi entre a quinta-feira e todo conjunto de sinistros, com uma equivalência apenas de 18,18% no padrão espacial de pontos.

Considerando os meses, constatou-se que não existe nenhum valor com pelo menos 0,8, mostrando que os meses possuem um padrão de pontos diferentes. O maior índice observado na nesta análise, foi registrado entre os meses de janeiro e abril, elucidando uma equivalência do padrão espacial de pontos entre 44,63% dos bairros analisados.

Além disso, os meses de fevereiro e agosto possuem uma semelhança de 43,80%. Ainda contextualizando essa análise, o mês de fevereiro e os dados gerais, possuem a menor similaridade entre as observações de pares como pode ser observado na Tabela 11.

Portanto, conclui-se que não é possível confirmar a existência de similaridade entre as análises de pares realizadas. Dessa forma, os sinistros de trânsito registrados na cidade de Fortaleza possuem um padrão de pontos específico para diferentes anos, meses, dias da semana e períodos do dia. Com isso, conclui-se que não é possível agregar esses dados, visto que os padrões observados diferentes.

Tabela 11 - Índice de similaridade de sinistros entre os meses de 2019 (Continua)

Meses	Mês 01	Mês 02	Mês 03	Mês 04	Mês 05	Mês 06	Mês 07	Mês 08	Mês 09	Mês 10	Mês 11	Mês 12
Mês 01												
Mês 02	0,4215											

Tabela 11 - Índice de similaridade de sinistros entre os meses de 2019 (Conclusão)

Meses	Mês 01	Mês 02	Mês 03	Mês 04	Mês 05	Mês 06	Mês 07	Mês 08	Mês 09	Mês 10	Mês 11	Mês 12
Mês 03	0,4050	0,3471										
Mês 04	0,4463	0,3636	0,2810									
Mês 05	0,3388	0,3636	0,4298	0,2810								
Mês 06	0,3884	0,3636	0,3223	0,3636	0,4132							
Mês 07	0,3140	0,4050	0,3884	0,3554	0,3967	0,3884						
Mês 08	0,3471	0,4380	0,3223	0,3554	0,3636	0,3471	0,3636					
Mês 09	0,2727	0,2645	0,3223	0,3388	0,4132	0,2975	0,3884	0,2893				
Mês 10	0,3967	0,4298	0,3388	0,3223	0,3306	0,3058	0,3058	0,3306	0,3388			
Mês 11	0,3554	0,3471	0,2893	0,2645	0,3554	0,4050	0,3636	0,2975	0,3223	0,3471		
Mês 12	0,3554	0,3306	0,3967	0,3471	0,3471	0,3636	0,3884	0,3388	0,3884	0,3306	0,3884	
Total	0,1736	0,1074	0,1488	0,1901	0,1818	0,1818	0,2066	0,1240	0,1653	0,1322	0,1983	0,1653

Fonte: O autor (2022)

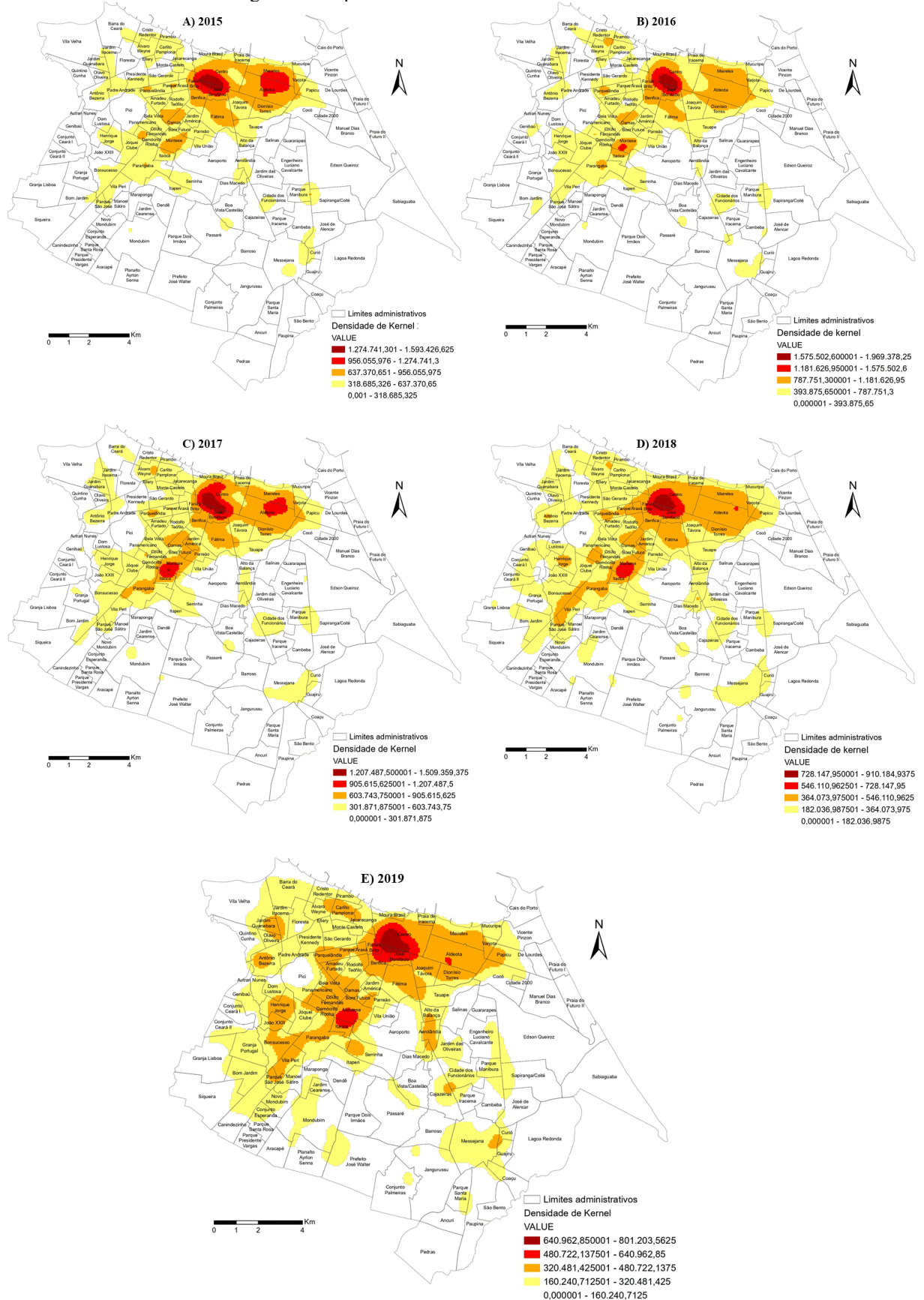
4.4 ESTIMATIVA DE DENSIDADE DE KERNEL

Para esta análise foi utilizada a ferramenta densidade de Kernel disponível no ARCGIS PRO, para auxiliar na identificação de possíveis aglomerados de sinistros nos períodos analisados. Foram submetidos a essa análise, apenas os dados que tiveram uma diferença significativa no teste estatístico de Kruskal-Wallis, e que não obteve semelhança significativa nas comparações pareadas no teste de similaridade, simultaneamente.

A distância planar entre as feições foi utilizada para essa análise, pois se trata de um padrão do GIS utilizado, e é adequado para a aplicação de uma escala local. O índice de gravidade definido para cada ocorrência foi usado como *input* nas análises, e como saída dessa aplicação foram obtidos os mapas de densidade de Kernel no formato *raster* para todos os meses e dias da semana do ano de 2019. Além disso, essa análise também foi realizada para os anos de 2015 a 2019, como mostra a Figura 4.

Foi utilizado um raio de pesquisa padrão especificado para cada um dos conjuntos de dados de sinistros através uma variante da regra de Silverman (1986). Além disso, foi determinado um valor de tamanho da célula de saída específico para cada conjunto de dados. As especificações desses parâmetros para os dados de sinistros dos anos são mostradas na Tabela 12.

Figura 4 - Mapas de densidade de Kernel de 2015 à 2019



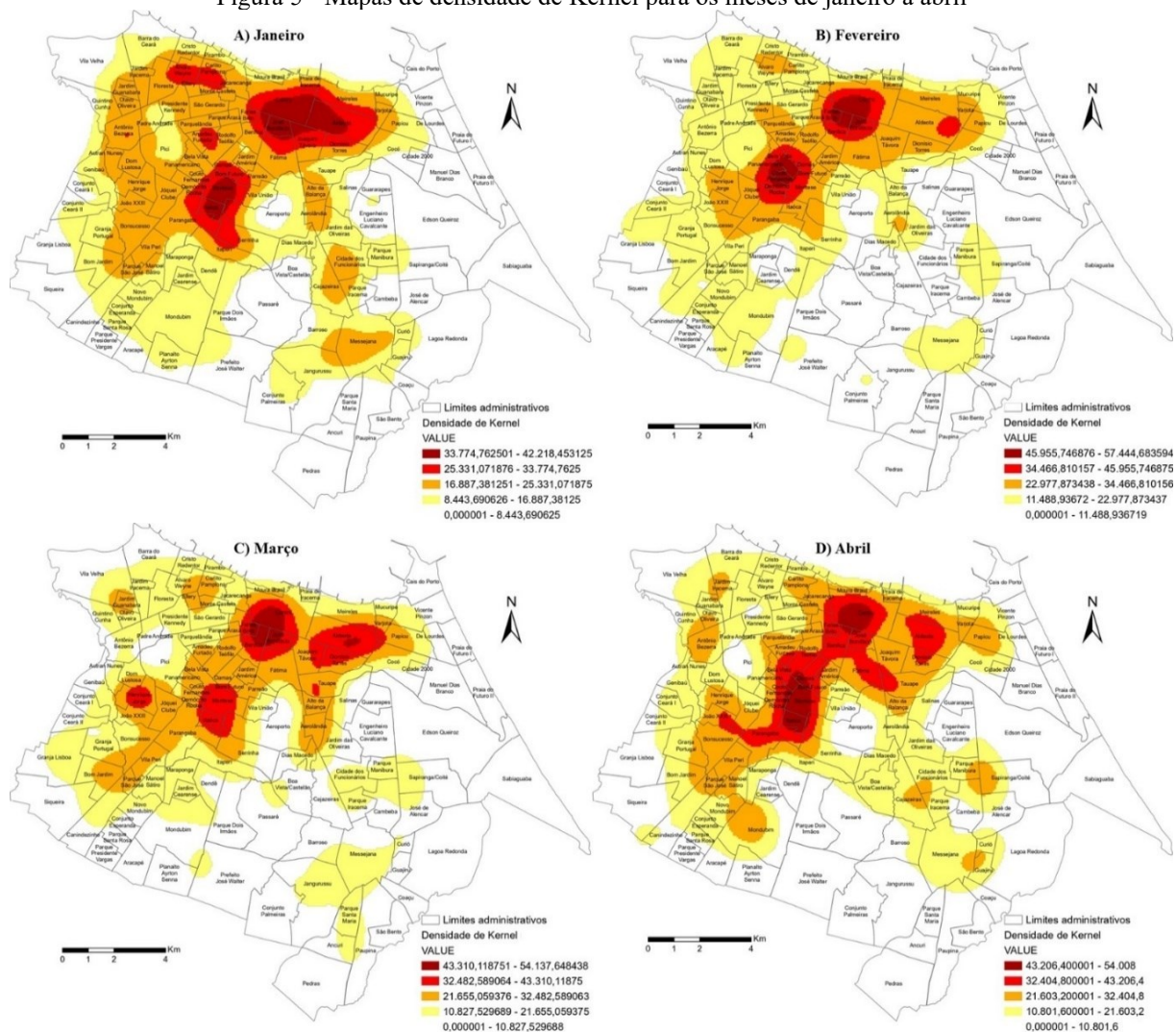
Fonte: O autor (2022)

Tabela 12 - Parâmetros da análise de densidade de kernel para anos		
Ano	Tamanho da célula de saída (Graus)	Raio (Graus)
2015	7,45E-04	0,008337014
2016	7,57E-04	0,008106479
2017	7,87E-04	0,008370149
2018	7,78E-04	0,009282907
2019	7,52E-04	0,009470268

Fonte: O autor (2022)

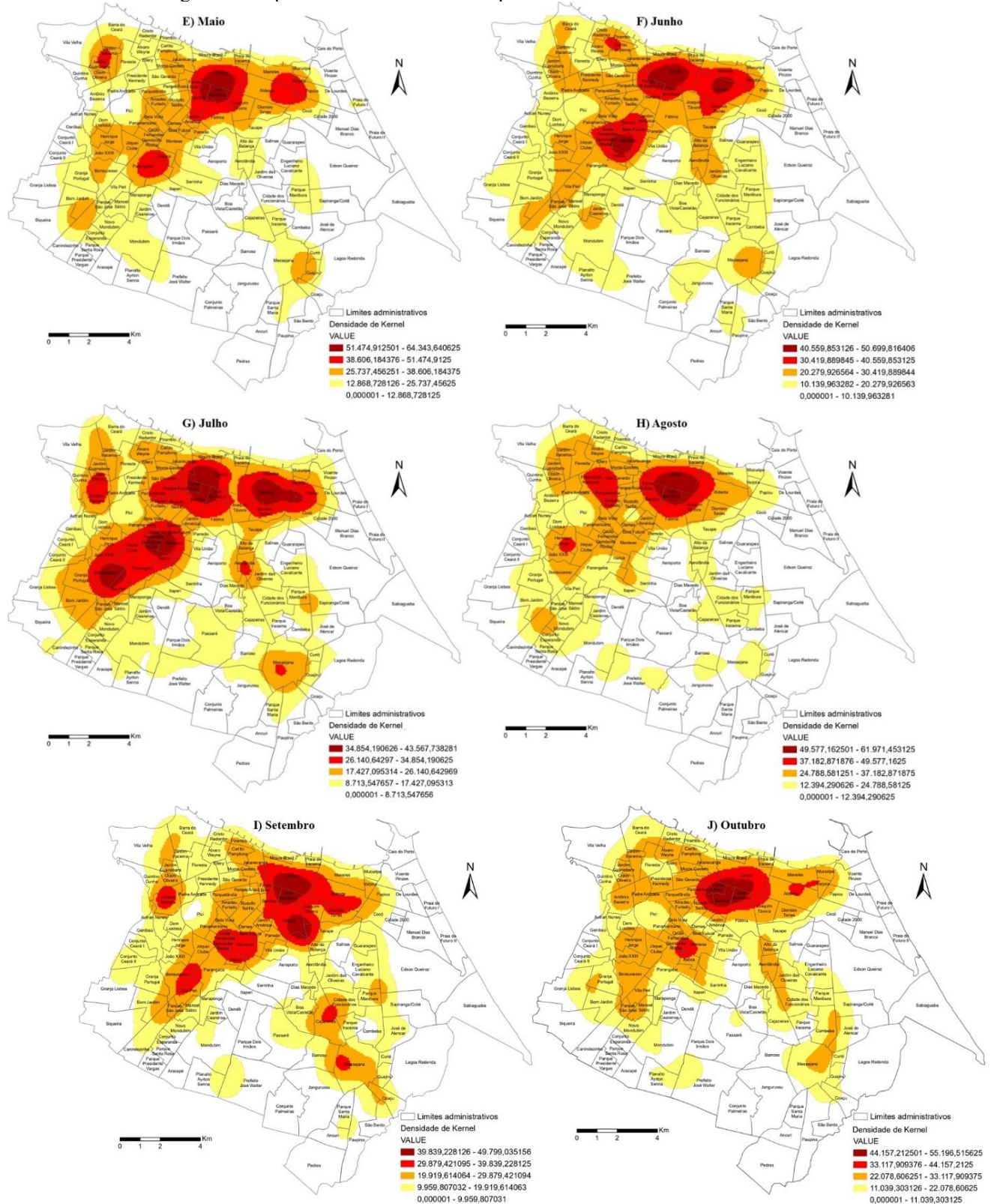
Os mapas de densidade produzidos com os dados dos anos, apresentam um padrão semelhante em relação a região onde se concentra a alta densidade dos sinistros de trânsito. Ambas as figuras, apresentam a região ao norte da cidade, como local de alto risco. As Figuras 5, 6 e 7 mostram os mapas para os diferentes meses do ano de 2019.

Figura 5 - Mapas de densidade de Kernel para os meses de janeiro à abril



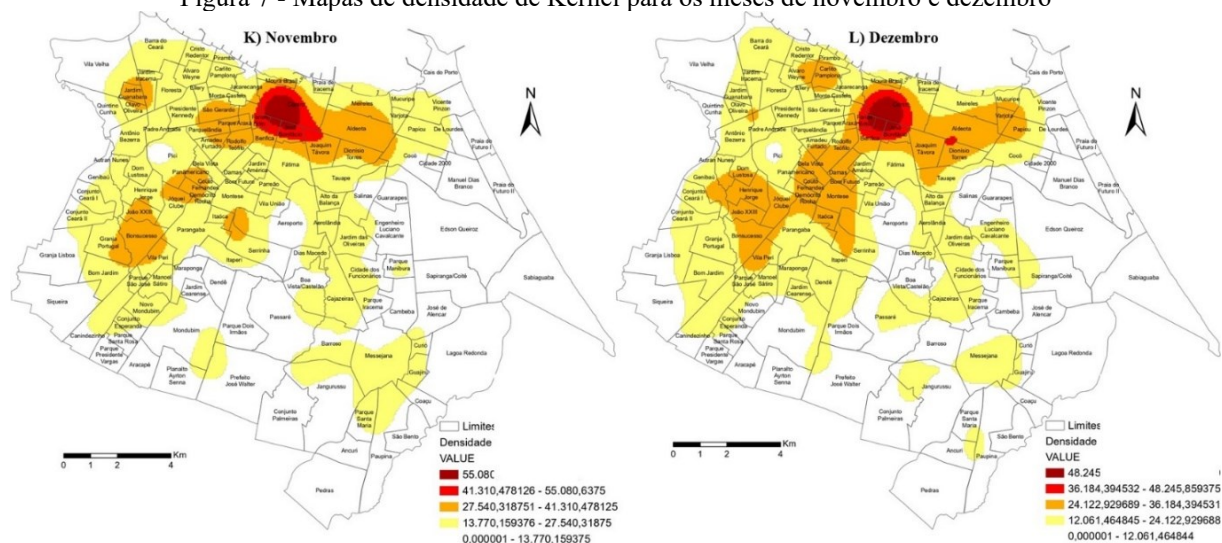
Fonte: O autor (2022)

Figura 6 - Mapas de densidade de Kernel para os meses de maio à outubro



Fonte: O autor (2022)

Figura 7 - Mapas de densidade de Kernel para os meses de novembro e dezembro



Fonte: O autor (2022)

Assim como na análise realizada anteriormente para os diferentes anos, foram determinados tamanho do raio da análise e tamanho da célula de saída para cada conjunto de dados de sinistros para os meses de 2019. Os parâmetros especificados para esses conjuntos são mostrados na Tabela 13, abaixo.

Tabela 13 - Parâmetros da análise de densidade de kernel para meses

Mês	Tamanho da célula de saída (Graus)	Raio (Graus)
Janeiro	6,47E-04	0,015544327
Fevereiro	6,73E-04	0,01535309
Março	7,03E-04	0,015521439
Abril	7,25E-04	0,015131442
Maio	7,21E-04	0,01523714
Junho	6,97E-04	0,015611307
Julho	7,44E-04	0,015958801
Agosto	7,10E-04	0,01514603
Setembro	7,46E-04	0,015851678
Outubro	7,34E-04	0,015730456
Novembro	6,95E-04	0,015695353
Dezembro	7,01E-04	0,015813442

Fonte: O autor (2022)

Assim como observado nos resultados dos índices de similaridade, é possível verificar através dos mapas de densidade de Kernel que o padrão espacial dos sinistros de trânsito é distinto em cada um dos meses analisados, elucidando que os sinistros estão aglomerados em diferentes locais com o passar dos meses.

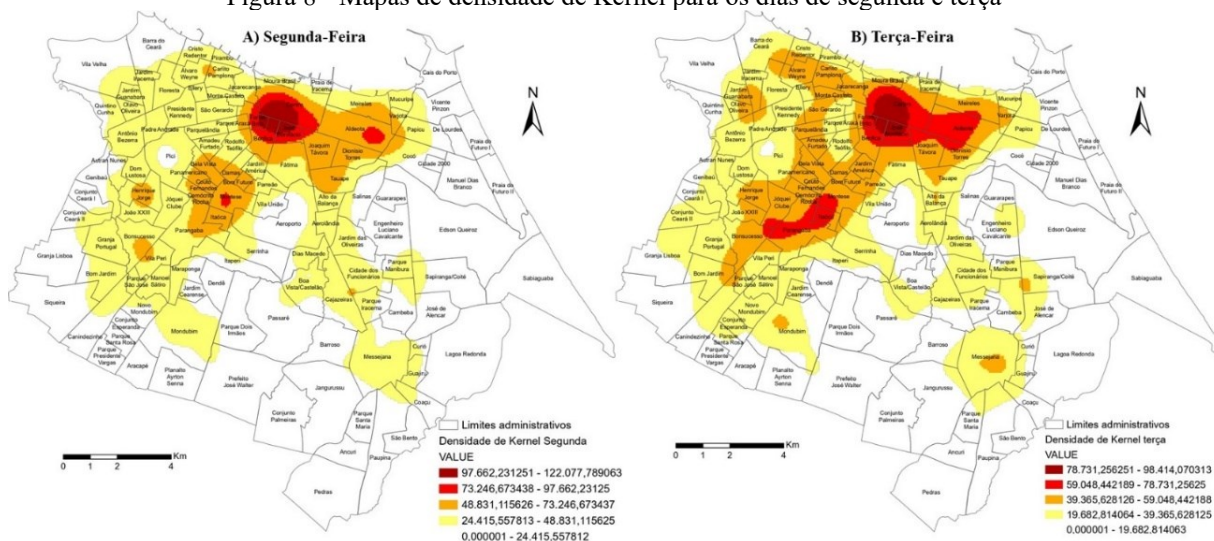
O resultado dessa análise, mostra que o mês de janeiro apresenta uma alta intensidade de sinistros na região norte de Fortaleza, nas imediações dos bairros Centro, José Bonifácio, Aldeoda, Joaquim Távora, Meireles e Benfica. Outro ponto de grande densidade pode ser observado na área oeste da cidade, nas imediações dos bairros Montese, Aeroporto, Itaóca e Serrinha. Os meses de fevereiro e abril apresentam um padrão de densidade semelhante ao exposto no mês de janeiro.

Os meses de março, maio, agosto, setembro, outubro, novembro e dezembro, apresentam região de aglomeração semelhante localizada na zona norte da cidade, nas imediações dos bairros Centro, José Bonifácio, Aldeoda, Joaquim Távora, Meireles e Benfica. Mas, o tamanho da área de abrangência é diferente em cada situação.

Já o mês de junho, apresenta uma grande densidade nas regiões norte, oeste e nordeste, enquanto que o mês de julho, apresenta uma grande densidade de sinistros de trânsito sobre as regiões norte, oeste, nordeste e sudoeste de Fortaleza. As Figuras 8 e 9 mostram os mapas de densidade de Kernel resultantes da análise para cada dia da semana.

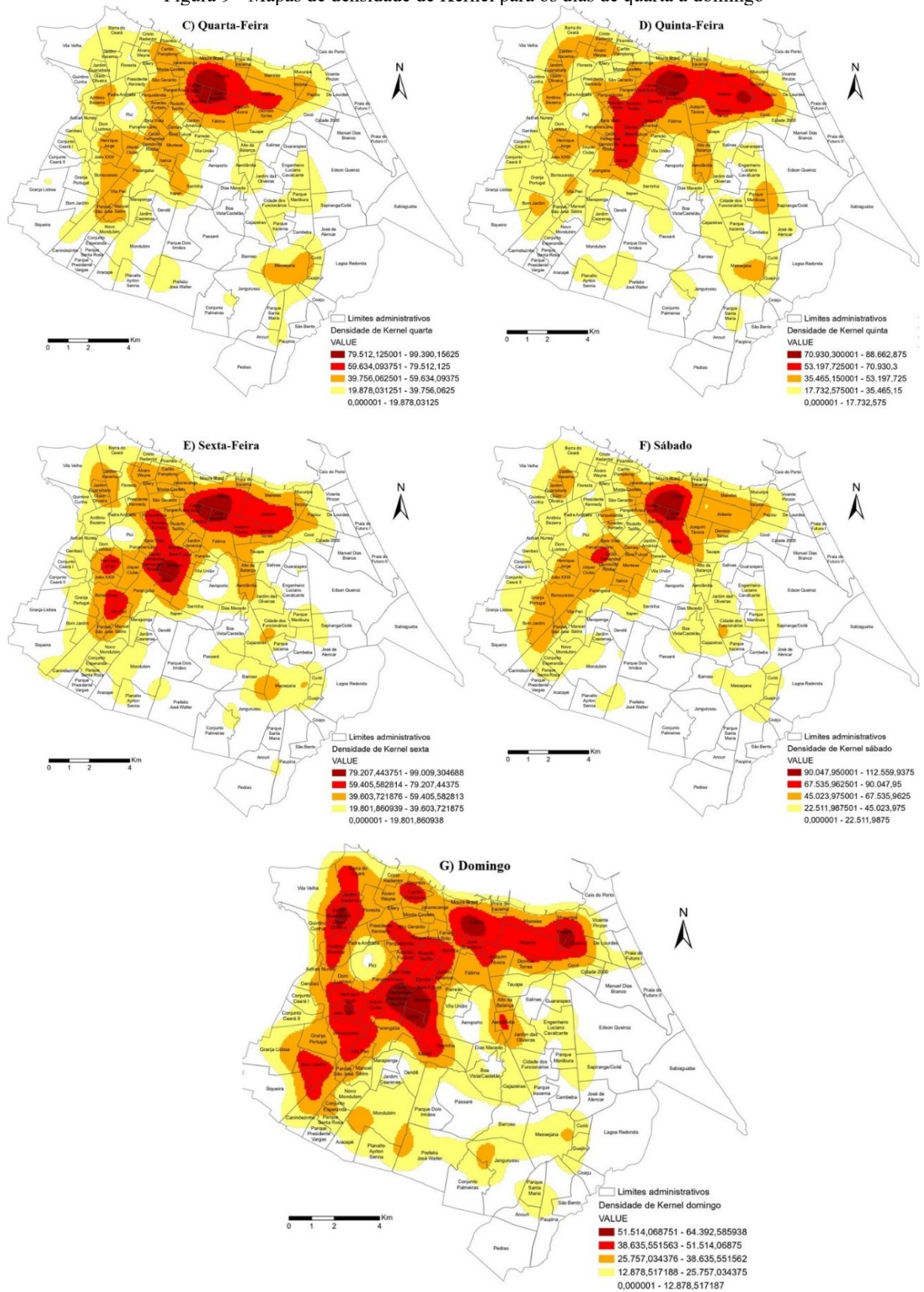
Para as análises realizadas para os dias da semana, os mapas apresentam uma região de alta intensidade de sinistros comum para os dados de sinistros da segunda-feira, terça-feira, quarta-feira e sábado, compreendendo a zona norte da cidade. Já para quinta-feira, o mapa ilustra pontos de alta densidade na região norte e nordeste. Para a sexta-feira, norte e sudoeste, e por fim, no domingo, norte, nordeste, oeste e noroeste de Fortaleza. A Tabela 14, mostra os parâmetros usados nas análises dos dias da semana de 2019.

Figura 8 - Mapas de densidade de Kernel para os dias de segunda e terça



Fonte: O autor (2022)

Figura 9 - Mapas de densidade de Kernel para os dias de quarta à domingo



Fonte: O autor (2022)

Para as análises realizadas para os dias da semana, os mapas apresentam uma região de alta intensidade de sinistros comum para os dados de sinistros da segunda-feira, terça-feira, quarta-feira e sábado, compreendendo a zona norte da cidade. Já para quinta-feira, o mapa ilustra pontos de alta densidade na região norte e nordeste. Para a sexta-feira, norte e sudoeste, e por fim, no domingo, norte, nordeste, oeste e noroeste de Fortaleza. A Tabela 14, mostra os parâmetros usados nas análises dos dias da semana de 2019.

Tabela 14 - Parâmetros da análise de densidade de kernel para os dias da semana

Dia da semana	Tamanho da célula de saída (Graus)	Raio (Graus)
Segunda-Feira	7,33E-04	0,01369468
Terça-Feira	7,34E-04	0,011276131
Quarta-Feira	7,45E-04	0,011337984
Quinta-Feira	7,46E-04	0,011215106
Sexta-Feira	7,35E-04	0,010989686
Sábado	7,44E-04	0,011312164
Domingo	7,10E-04	0,01176109

Fonte: O autor (2022)

Apesar da diferença entre a padrão de distribuição espacial dos sinistros de trânsito de Fortaleza nos diferentes meses e dias da semana analisados, pode-se observar que a região norte da cidade possui um alta densidade de ocorrências para todos os períodos da análise. Entretanto, cada mapa de densidade mostra um padrão específico para cada período, sinalizando que os sinistros tendem a se concentrar em regiões diferentes com o passar dos meses e dias da semana.

Com a aplicação dessa abordagem foi possível observar nos mapas onde se concentra as maiores densidades de sinistros de acordo com o índice de gravidade nos anos de 2015 a 2019, nos meses de e dias das semanas do ano de 2019. Com isso, foi notado que a região do Centro da cidade aparece em evidência em todos os mapas dos períodos analisados.

4.5 AUTOCORRELAÇÃO ESPACIAL GLOBAL

Essa análise foi realizada para verificar o padrão espacial dos sinistros de trânsito em diferentes períodos e com isso, identificar se o padrão desses dados tende a ser aleatório, disperso ou agrupado. Para isso, foi utilizado a estatísticas I de Moran Global, com apoio da ferramenta Autocorrelação Espacial disponível no ARCGIS PRO.

O índice de gravidade definido para cada sinistro é utilizado como *input* para avaliação da autocorrelação espacial, e uma distância padrão foi especificada para cada conjunto de pontos, de forma que, todas as observações tenham pelo menos um vizinho. A distância padrão de cada análise é mostrada nas tabelas seguintes.

Além disso, a relação entre os recursos espaciais foi definida através da distância inversa, que considera que as feições mais próximas da observada tem maior influência sobre ela. A ferramenta GIS utilizada retorna os valores do Índice I de Moran, Índice Esperado, Variância, valor *z*, valor *p*, para os diferentes meses e dias da semana do ano de 2019. A Tabela 15, mostra o resultado da análise de autocorrelação espacial que foi realizada para verificar o padrão espacial dos sinistros em cada dia da semana e a distância padrão usada em cada situação.

Tabela 15 - Resultado da Autocorrelação espacial global para os dias da semana

Dia da semana	Índice I de Moran	Índice Esperado	Variância	z-score	p-value	Distância padrão (M)
Segunda	0,0178	-0,0006	0,0002	1,3533	0,1759	2082,0802
Terça	0,0210	-0,0006	0,0002	1,3759	0,1689	1321,355
Quarta	-0,0061	-0,0006	0,0003	-0,3243	0,7457	993,4141
Quinta	0,0120	-0,0006	0,0002	0,8676	0,3856	1428,7755
Sexta	0,0218	-0,0005	0,0002	1,6483	0,0993	1953,5726
Sábado	0,0106	-0,0006	0,0001	0,9974	0,3186	1546,0464
Domingo	0,0061	-0,0007	0,0002	0,4918	0,6229	1909,1992

Fonte: O autor (2022)

O resultado dessa análise revela que o padrão espacial global das ocorrências de sinistros de trânsito em Fortaleza é significativamente aleatório para todos os dias da semana. Já, em relação aos meses, o resultado indica que o padrão espacial é diferente ao longo dos meses, mostrando uma autocorrelação positiva para alguns e uma distribuição aleatória nos demais, assim como é mostrado na Tabela 16, que apresenta os valores resultantes da análise para os meses.

Tabela 16 - Resultado da Autocorrelação espacial global para os meses (Continua)

Meses	Índice I de Moran	Índice Esperado	Variância	z-score	p-value	Distancia Padrão (M)
Janeiro	-0,0031	-0,0012	0,0004	-0,0912	0,9274	1912,84
Fevereiro	-0,0038	-0,0011	0,0003	-0,1673	0,8671	2411,75
Março	0,0343	-0,0011	0,0003	2,0105	0,0444	2464,87
Abril	0,0648	-0,001	0,0004	3,2054	0,0013	1938,61

Tabela 16 - Resultado da Autocorrelação espacial global para os meses (Conclusão)

Meses	Índice I de Moran	Índice Esperado	Variância	<i>z-score</i>	<i>p-value</i>	Distancia Padrão (M)
Maio	-0,0047	-0,0009	0,0004	-0,186	0,8525	1105,266
Junho	0,0174	-0,001	0,0003	0,9961	0,3192	1985,998
Julho	0,0028	-0,001	0,0004	0,1885	0,8505	1471,804
Agosto	0,0092	-0,001	0,0003	0,5909	0,5546	1839,023
Setembro	0,027	-0,001	0,0002	2,0168	0,0437	1690,417
Outubro	-0,0028	-0,001	0,0003	-0,1063	0,9154	3656,035
Novembro	0,0081	-0,001	0,0003	0,5601	0,5754	4024,095
Dezembro	0,0407	-0,001	0,0004	2,1087	0,035	1402,486

Fonte: O autor (2022)

No contexto global, os meses de março, abril, setembro e dezembro possuem uma autocorrelação espacial positiva, ou seja, os dados de sinistros tendem a estar mais agrupados no espaço. Para os demais meses, os dados tendem a ser distribuídos aleatoriamente no espaço, sem qualquer tipo de autocorrelação. A análise foi também realizada para os diferentes anos da análise, onde foi verificado um padrão agrupado, como mostrado na Tabela 17.

Tabela 17 - Resultado da Autocorrelação espacial global para os anos

Ano	Índice I de Moran	Índice Esperado	Variância	<i>z-score</i>	<i>p-value</i>	Distância Padrão (M)
2015	0,0364	-0,0001	0	10,8646	0	1786,18
2016	0,0429	-0,0001	0	11,564	0	869,572
2017	0,0348	-0,0001	0	13,0423	0	1653,22
2018	0,0286	-0,0001	0	5,987	0	990,463
2019	0,0172	-0,0001	0	3,7242	0,0002	1490,812

Fonte: O autor (2022)

Entretanto, isso não significa que esse padrão aleatório estará em todos os locais da área analisada, pois o espaço é dinâmico. Com isso, pode-se utilizar a análise de autocorrelação espacial no nível local, para verificar se há autocorrelação em algumas regiões da área de estudo.

4.6 ANÁLISE HOTSPOTS

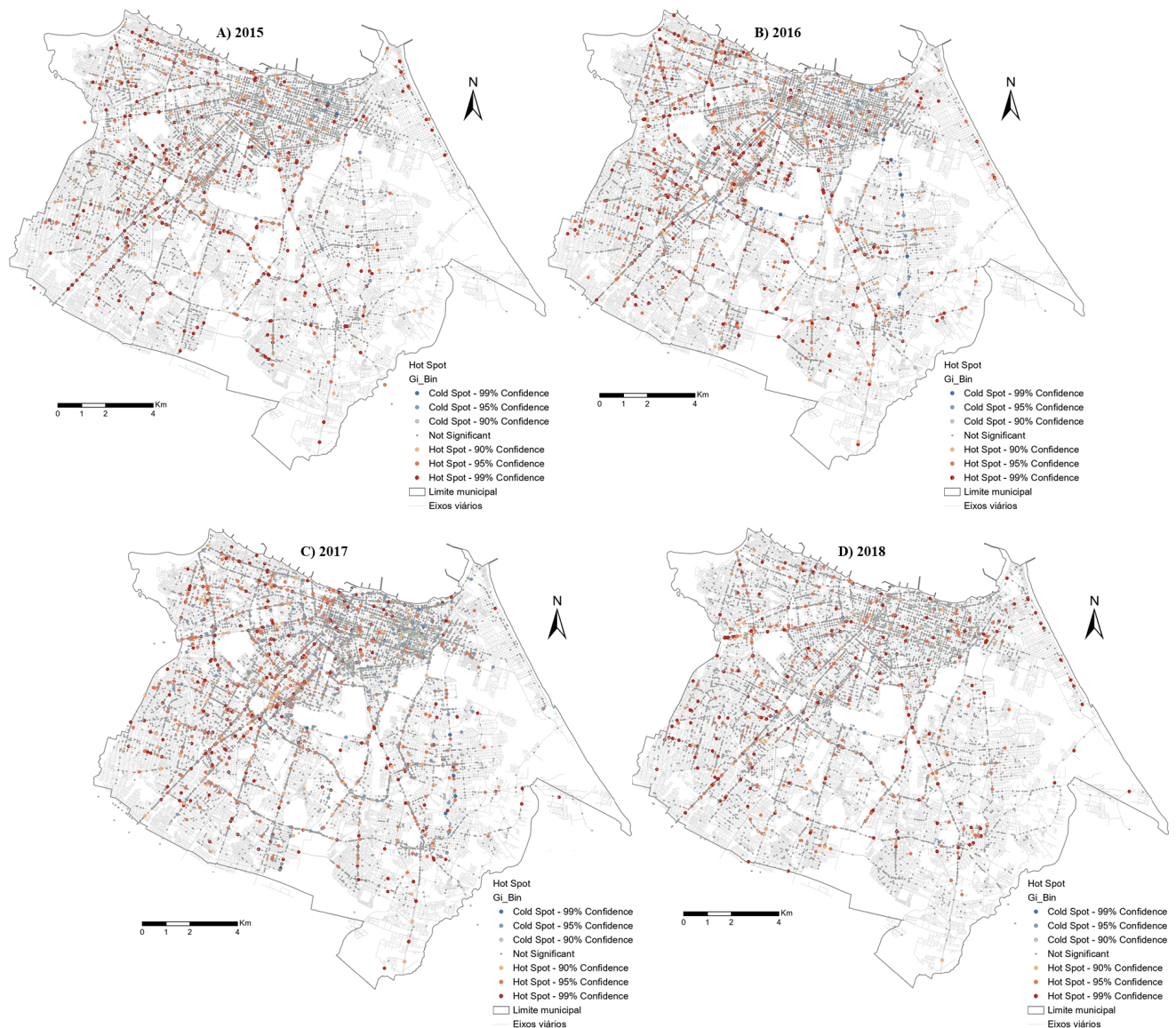
A autocorrelação espacial local foi realizada através da ferramenta análise *hotspots* disponível no ARCGIS PRO, utilizando a estatística de Getis-Ord G_i^* para verificar a existência de *hotspot* e *coldspots* de sinistros de trânsito significativos estatísticos. Com isso,

foram gerados os mapas *raster* para dados dos meses e dias do ano de 2019, e dos anos individualmente.

Essa análise foi realizada utilizando índice de gravidade como *input* para avaliação da autocorrelação espacial local, e as distâncias padrão definidas na análise de autocorrelação espacial global, mostradas na seção anterior para cada conjuntos de pontos. A relação entre os recursos espaciais foi definida através da distância inversa.

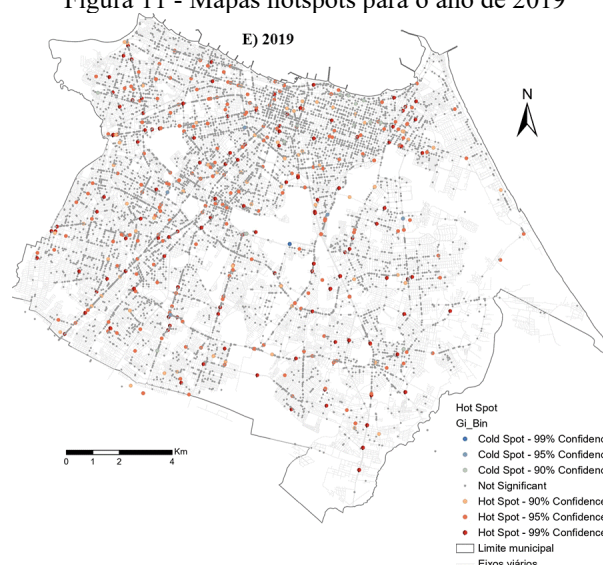
As Figuras 10 e 11 mostram os mapas gerados a partir da análise de *hotspots* nos diferentes anos considerados.

Figura 10 - Mapas *hotspots* para os anos de 2015 à 2018



Fonte: O autor (2022)

Figura 11 - Mapas hotspots para o ano de 2019



Fonte: O autor (2022)

Para ambos os anos da análise, a maior parte das ocorrências são classificadas como aleatórias, assim como mostrado da Tabela 18. Todos os mapas apresentam uma grande quantidade de *hotspots*, sendo que o ano de 2015 apresenta uma maior quantidade para o hotspots com 99% de significância.

Tabela 18 - Resultado da análise hotspots para os anos

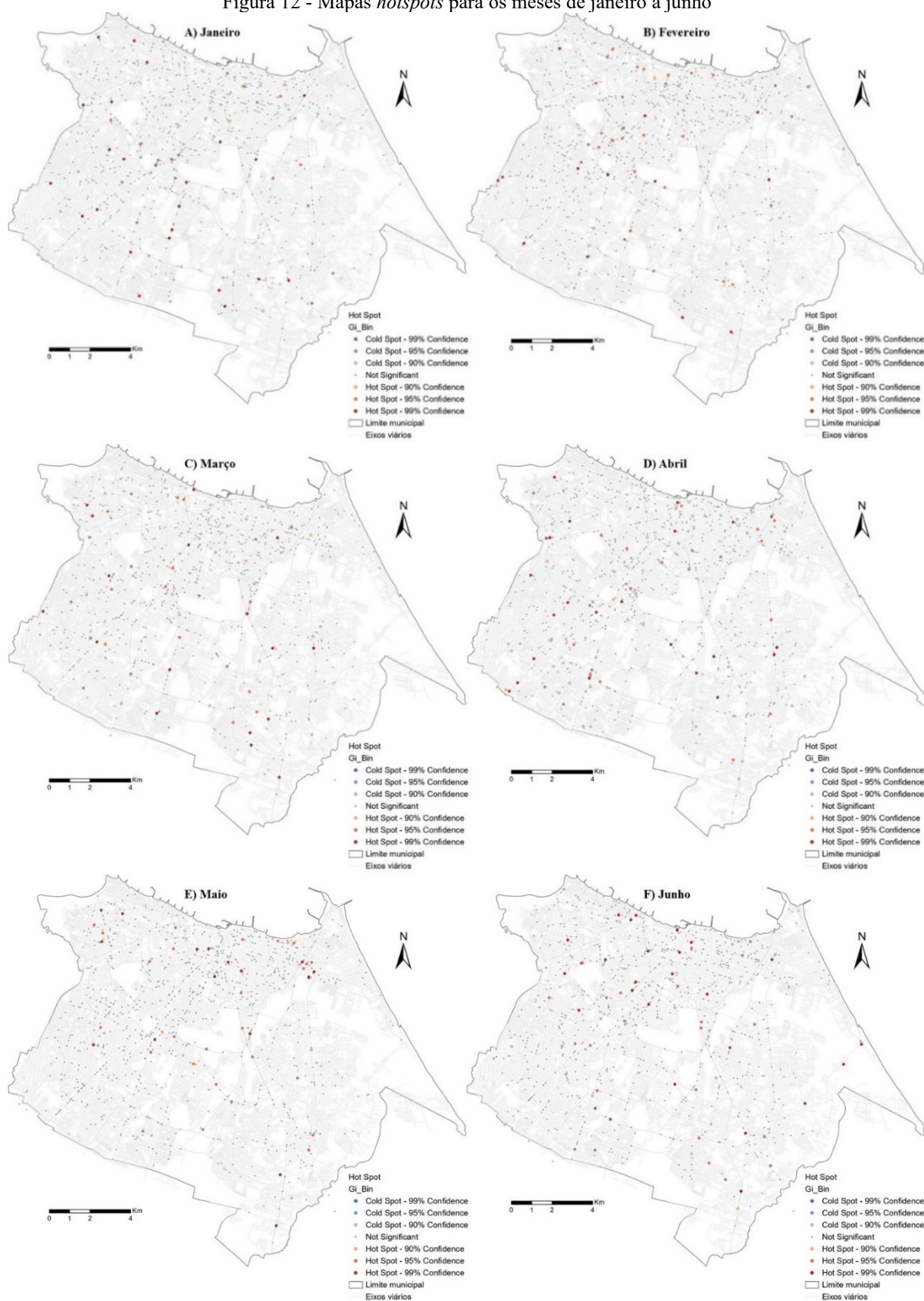
Ano	Hotspots 90%	Hotspots 95%	Hotspots 99%	Aleatório	Coldspots 90%	Coldspots 95%	Coldspots 99%
2015	257	396	744	14880	232	241	63
2016	424	507	449	16121	365	520	430
2017	331	509	497	14633	176	372	231
2018	183	290	393	10828	119	76	5
2019	126	364	378	10731	123	35	38

Fonte: O autor (2022)

Já os *coldspots* são também estão em grande quantidade, mostrando que há vários pontos onde a maior parte dos sinistros são leves. O ano de 2016 apresentou a maior quantidade de *coldspots* entre os anos da análise.

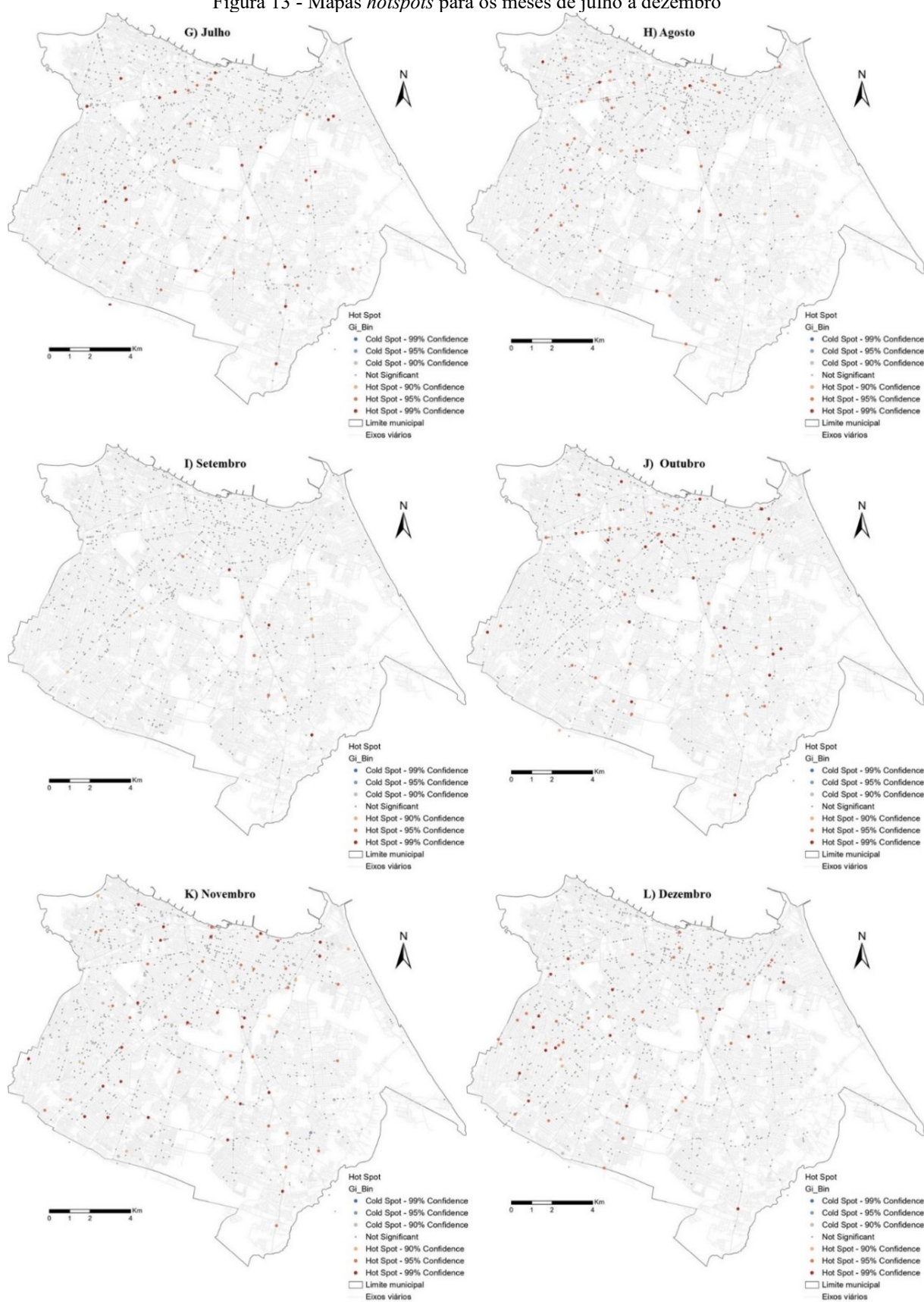
De maneira geral, analisando os mapas obtidos, os mesmos apresentam vários *hotspots* em todos os meses, mostrando uma autocorrelação dos sinistros naquela área. Entretanto, a maior parte dos pontos não apresenta um padrão significante, ou seja, estão formando um arranjo aleatório, como mostrado pelos pontos em cinza. As Figuras 12 e 13 mostram essa mesma análise realizada para dados de sinistros em diferentes meses de 2019.

Figura 12 - Mapas *hotspots* para os meses de janeiro à junho



Fonte: O autor (2022)

Figura 13 - Mapas *hotspots* para os meses de julho à dezembro



Fonte: O autor (2022)

Os pontos destacados em vermelho nos mapas, mostram um *hotspot*, indicando que no entorno desses pontos há outros com características semelhantes, ou seja, pontos de sinistros viários que também possuem um valor alto significativamente estatístico de Z , e por isso o *hotspots* é formado naquele local.

Já os pontos em azul, mostram os *coldspots*, que representa um aglomerado de pontos que possuem o valor muito baixo de Z . Com isso, é possível observar esses valores altos e baixos no espaço através dos valores de p e Z , onde p é a probabilidade de o arranjo ser aleatório.

A Tabela 19 mostra o número *hotspots* com significância igual ou superior a 95% em cada mês, ou seja, pontos possuem que um valor alto de Z e que formam *clusters* significativamente estatísticos. Assim, esses valores expressam o número de áreas perigosas em cada mês para a Cidade, porque esses locais apresentam um grande número de ocorrências com valores altos de gravidade.

Tabela 19 - Resultado da análise hotspots para os meses de 2019

Mês	Hotspots 90%	Hotspots 95%	Hotspots 99%	Aleatório	Coldspots 90%	Coldspots 95%
Janeiro	4	17	24	790	24	0
Fevereiro	7	33	14	844	1	6
Março	3	23	25	811	23	0
Abril	13	26	17	872	42	0
Mai	10	33	17	1061	0	0
Junho	5	27	34	933	15	0
Julho	5	17	26	900	30	0
Agosto	5	37	9	991	0	0
Setembro	23	12	9	967	0	0
Outubro	7	35	28	952	0	0
Novembro	8	37	30	894	30	5
Dezembro	13	30	19	881	27	5

Fonte: O autor (2022)

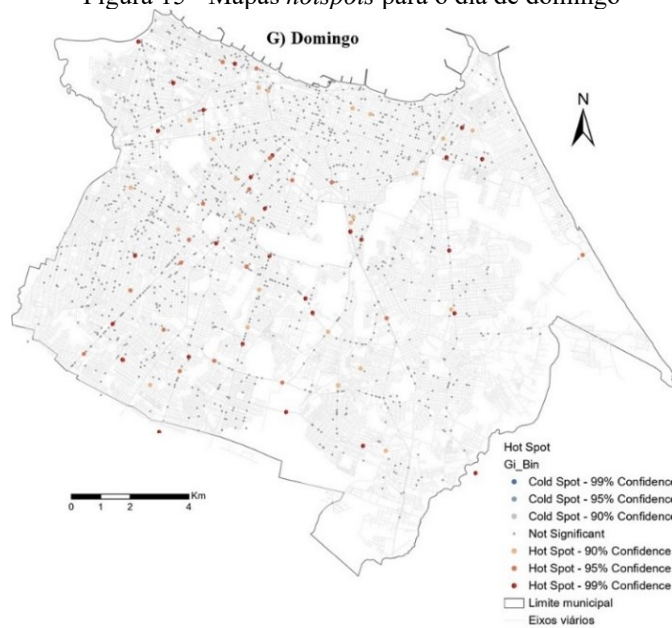
O mês de novembro apresenta o maior número de *hotspots* do ano, com 67 locais com altas taxas de sinistros de acordo com o índice de gravidade, seguido dos meses de outubro e junho respectivamente. Além disso, o mês de setembro obteve apenas 20 pontos com aglomerados de valores significativamente altos. Já o mesmo de maio, apresenta a maior quantidade de pontos de ocorrências classificadas com aleatórias pelo resultado da ferramenta. As Figuras 14 e 15 mostram o resultado da análise de *hotspots* para os dias da semana.

Figura 14 - Mapas *hotspots* para os dias de segunda à sábado



Fonte: O autor (2022)

Figura 15 - Mapas *hotspots* para o dia de domingo



Fonte: O autor (2022)

Os mapas dos dias da semana apresentam também vários *hotspots*. Ainda assim, a maior parte dos pontos não apresenta um padrão significativo. A Tabela 20, apresenta o número de *hotspots* para os dias da semana considerando uma significância maior que ou igual a 95% expressando, dessa maneira, o número de locais com elevado índice de sinistros.

Tabela 20 - Resultado da análise hotspots para dias da semana de 2019

Dia	Hotspots 90%	Hotspots 95%	Hotspots 99%	Aleatório	Coldspots 90%	Coldspots 95%
Segunda	15	44	56	1515	72	15
Terça	7	29	43	1444	53	10
Quarta	18	42	40	1528	7	0
Quinta	11	32	37	1576	24	0
Sexta	17	38	69	1670	52	5
Sábado	27	22	8	1739	0	0
Domingo	24	23	34	1449	0	0

Fonte: O autor (2022)

O dia de sexta-feira apresentou a maior quantidade de *hotspots* em relação aos demais dias da semana, apresentando 107 pontos de alta concentração de acordo com a gravidade dos sinistros. Além disso, a segunda-feira também apresentou uma grande quantidade de *hotspots*, registrando um total de 100 pontos.

Com a aplicação dessa ferramenta foi identificado os principais pontos de sinistros de trânsito nos diferentes períodos da análise. A distribuição desses pontos no espaço é destinto

para cada período, e a maior parte dos sinistros ocorrem aleatoriamente no espaço de acordo com a análise. Além disso, as análises realizadas para os anos mostram uma quantidade maior de *hotspots* em relação aos meses e dias da semana de 2019 pois, o número de sinistros considerados é maior.

4.7 DISCUSSÃO FINAL DO CAPÍTULO

Através das análises espaciais realizadas, como a densidade de kernel e *hotspots*, foi constatado que a maior parte dos sinistros graves ocorre na região norte da cidade de Fortaleza. Além disso, a mesma registra uma quantidade expressiva de mortes e feridos gravemente e leves. Dessa forma, essa região deve ser priorizada no desenvolvimento e adoção de medidas de retenção de sinistros.

Foi constatado que no bairro Centro, a maior parte dos registros de sinistros têm o veículo do tipo motocicleta envolvido, cerca de 65% dos registros analisados. Já os automóveis, também têm participação expressiva nos sinistros, cerca de 63% dos sinistros envolvem esse tipo de veículo. Além disso, mais de 50% desses registros estão localizados sobre algum tipo de interseção, na maioria delas em cruzamentos do tipo cruz.

Já no bairro Montese, que registra um grande número de sinistros e de vítimas feridas, foi verificado que quase 80% dos sinistros do banco de dados são do tipo colisão, sendo essas de diversos tipos. Cerca de 50% dessas ocorrências estão localizadas em algumas interseções, sendo a maioria em cruzamento do tipo cruz.

Foi constatado que bairro Mondubim registrou a maior quantidade de mortos nessa análise. Os sinistros envolvendo motocicleta mostraram ser mais recorrentes nessa localidade, onde mais de 70% das ocorrências registradas, envolve esse tipo de veículo. Já os sinistros envolvendo carros, corresponde a pouco mais de 60%. Nesse bairro, foi averiguado que grande parte desses registros estão situados na CE-060, Avenida Presidente Costa e Silva, Avenida Benjamin Brasil e Rua Monte Líbano.

No bairro Aldeota detectou-se a segunda maior quantidade de sinistros. Mais de 75% das ocorrências analisadas possuem carros envolvidos. Já no bairro Messejana, verificou-se que as vias com maior quantidade de sinistros são: Avenida Frei Cirilo, avenida Washington Soares (CE-040), rua Padre Pedro de Alencar, rua José Hipólito e avenida Doutor Joaquim Bento.

No Bom Sucesso, constatou-se que as ocorrências com motocicletas estão presentes em quase 80% dos sinistros registrados no bairro. As colisões são a natureza mais frequente desses,

e as vias que possuem alto número de sinistros são Avenida General Osório de Paiva (CE-065), rua Maria Júlia Rocha, e avenida Oscar Araripe.

5 ANÁLISE DE REGRESSÃO ESPACIAL

Este capítulo apresenta um modelo de análise de regressão espacial local aplicado a dados de sinistros de trânsito da cidade de Fortaleza. O mesmo será utilizado para verificar a heterogeneidade e as relação entre as ocorrências de sinistros e algumas variáveis selecionadas. Dessa forma, será possível observar quais e o quanto esses fatores podem contribuir para as ocorrências de sinistros de trânsito na cidade. O modelo foi desenvolvido através do GWR, tendo como suporte ARCGIS PRO.

5.1 DADOS

Os dados utilizados nessa análise foram extraídos da plataforma Fortaleza em Mapas, disponibilizada através da prefeitura municipal. Com isso, foram extraídas informações sobre os bairros, sobre mobilidade, educação, meio ambiente, ocupação de solo, demografia, índice de desenvolvimento humano, saúde e outros. Os dados foram compilados em um único *shapefile*, juntamente com a variável dependente que foi considerada o número de sinistros por bairro.

5.2 SELEÇÃO DAS VARIÁVEIS

A regressão espacial foi utilizada a fim de verificar a relação entre o número de registros de sinistros nos bairros de Fortaleza com um conjunto de outras variáveis associativas. Estas variáveis associativas abrangem fatores socioeconômicos e locais que atraem um grande fluxo de pessoas.

O conjunto de variáveis socioeconômicas selecionadas para a análise foram os fatores de renda média de cada bairro, porcentagem de coleta de lixo, porcentagem de captação de esgoto, porcentagem de abastecimento de água, porcentagem de habitantes alfabetizados, área comercial, área industrial, área serviços, área residencial e densidade populacional.

Essas informações foram extraídas do portal Fortaleza em mapas, que dispõe de mapas com informações sociais e financeiras da cidade disponibilizada pelo município. Além disso, essas variáveis socioeconômicas foram selecionadas com base nos trabalhos apresentados na literatura que analisam a relação entre elas e as ocorrências de sinistros de trânsito nas cidades.

Nas áreas que atraem grandes fluxos de pessoas, foram consideradas a quantidade de unidades educacionais de cada bairro, unidades de saúde, agências bancárias, patrimônio tombado, terminais de ônibus, áreas verdes na cidade, shoppings, feiras e locais com plano de arborização. Além disso, foram consideradas também o número de intersecção de vias em cada local, podendo esse ser um cruzamento, intersecção em “T”, via lateral, bifurcação ou confluência.

Foi realizada uma análise de correlação entre as variáveis associativas para identificar possíveis correlações entre elas e em caso afirmativo, retirar essa correlação removendo uma das variáveis correlacionadas. Assim, foi possível selecionar as seguintes variáveis associativas para o modelo GWR: renda média, densidade populacional, coleta de lixo, unidades de saúde, captação de esgoto, abastecimento de água, alfabetizados, porcentagem de área comercial, área industrial, área serviços, unidades educacionais, agências bancárias, patrimônio tombado, terminais de ônibus, áreas verdes na cidade, shoppings, locais com plano de arborização e nº de intersecções de vias. Como variável dependente do modelo será utilizado o número de sinistros viários por bairro.

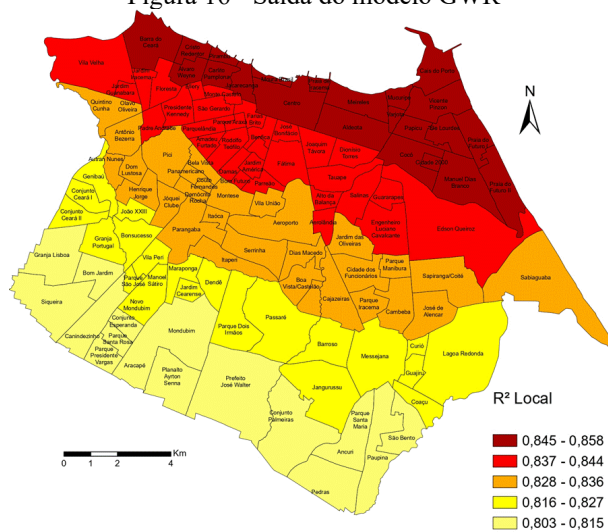
5.3 RESULTADO DO GWR

O modelo GWR desenvolvido assume que a distribuição da variável resposta e dos erros como normais. Além disso, considera também que os erros tenham variância constante. Com isso, o diagnóstico do modelo GWR indicou um valor de R^2 e do R^2 ajustado de 0,8336 e 0,7796 respectivamente. O índice de ajuste do modelo AIC foi de 1248,848. A Figura 17 mostra os mapas de saídas do modelo GWR para o R^2 .

O do R^2 varia entre 0,803 e 0,858 nos bairros da cidade, indicando que as variáveis utilizadas de certa forma estão associadas entre 80,3% a 85,8% das ocorrências de sinistros nesses locais. Além disso, o teste de autocorrelação espacial dos resíduos do GWR (escore $z = -0,372929$; $p = 0,709201$; I de Moran = - 0,020747), mostrou que o escore z é insignificante, apontando eu que não há correlação entre os resíduos, e assim, as variáveis utilizadas no modelo são adequadas para a análise.

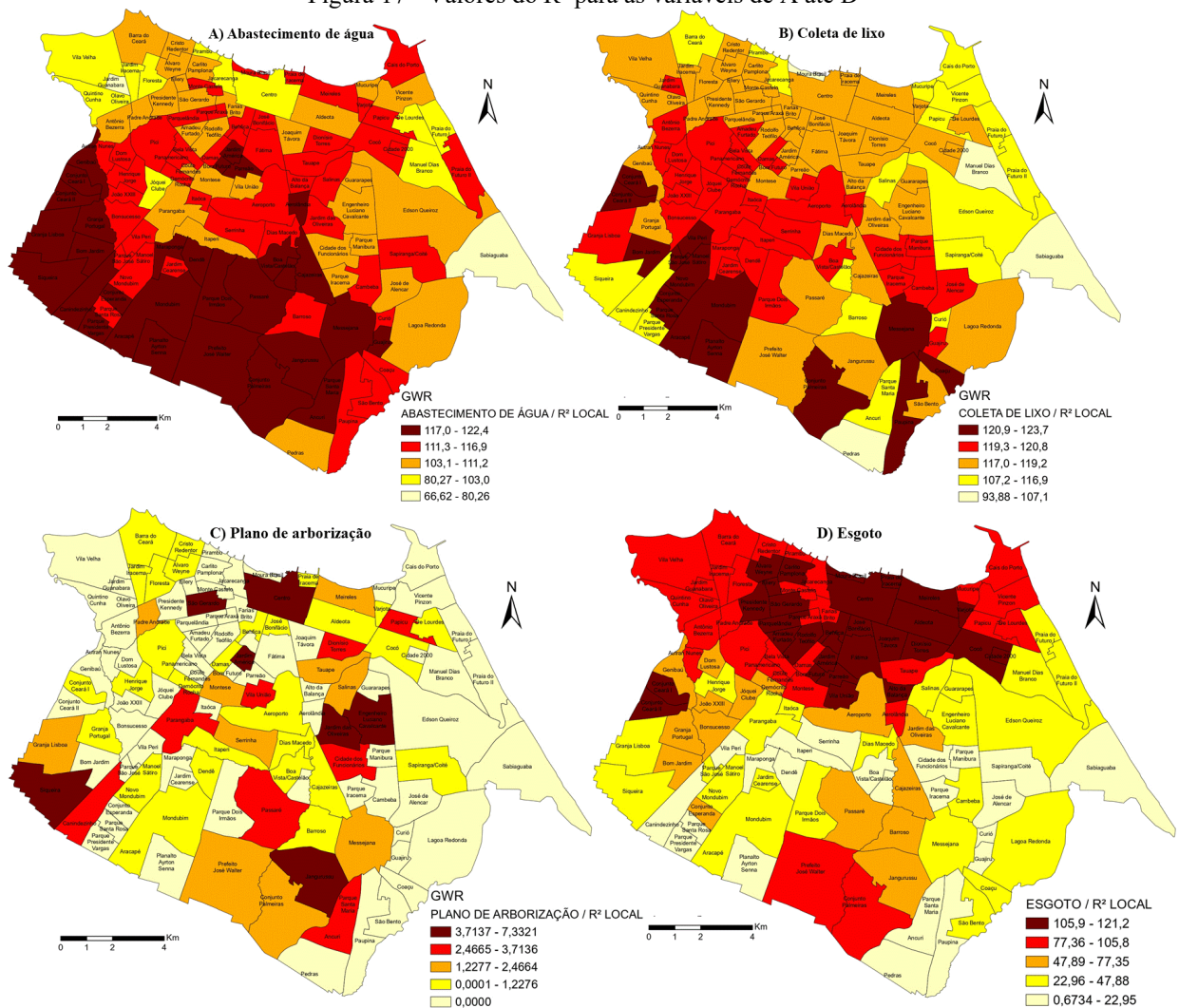
A Figura 16 mostra a distribuição dos valores dos R^2 do modelo para cada local, e as Figuras 17, 18, 19 e 20 elucidam os valores do R^2 de cada variável em cada unidade de bairro.

Figura 16 - Saída do modelo GWR



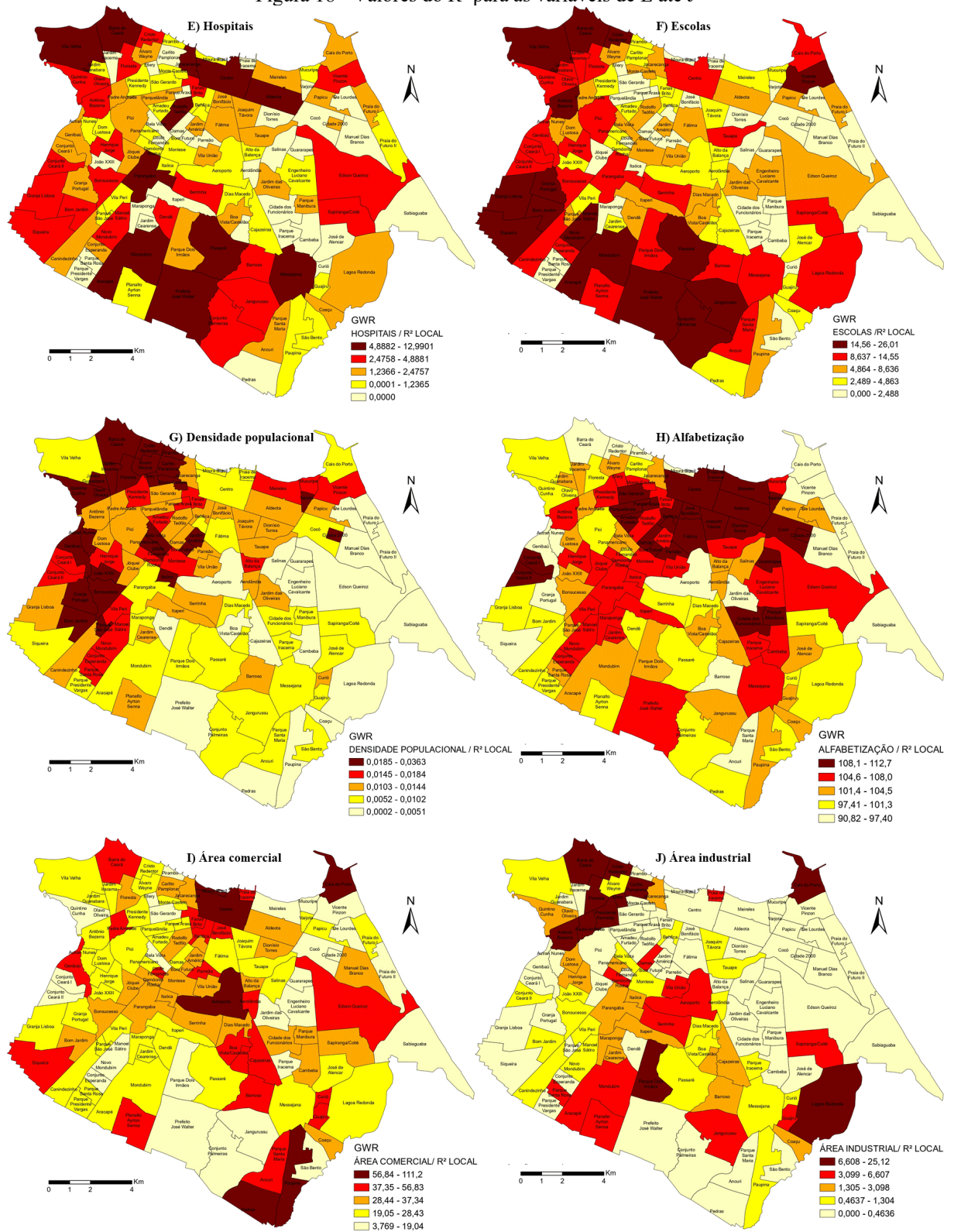
Fonte: O autor (2022)

Figura 17 - Valores do R² para as variáveis de A até D



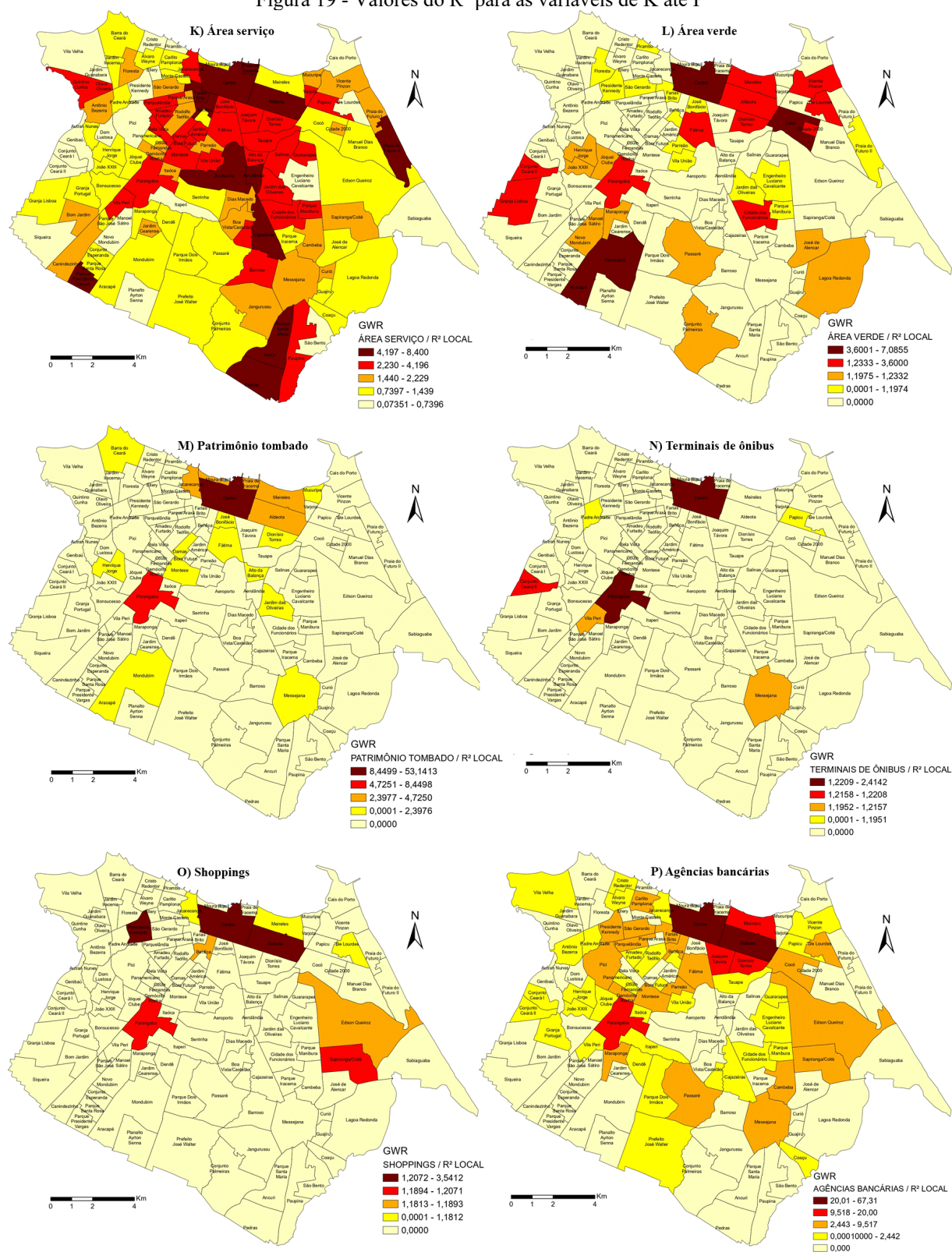
Fonte: O autor (2022)

Figura 18 - Valores do R^2 para as variáveis de E até J



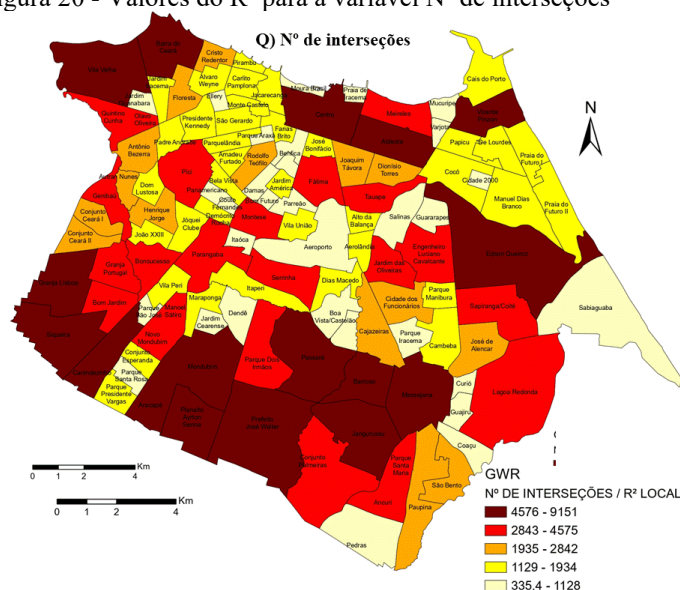
Fonte: O autor (2022)

Figura 19 - Valores do R^2 para as variáveis de K até P



Fonte: O autor (2022)

Figura 20 - Valores do R^2 para a variável N° de interseções



Fonte: O autor (2022)

As áreas compostas por cores de vermelho mais intenso, mostram os locais onde tal variável tem um maior poder explicativo. A maior parte das variáveis explicam fortemente as ocorrências de sinistros no bairro Centro. Com a aplicação desse modelo foi constatado que

As variáveis hospitalais, coleta de lixo, alfabetização, porcentagem de área indústria, serviço e industrial, área verde, escolas, patrimônio tombado, plano de arborização, terminais de ônibus, número de interseções e agências bancárias, tem uma relação positiva com a variável dependente. Já as variáveis shoppings, estação de metrô, abastecimento de água, renda média e esgoto, tem uma relação inversa com a variável associativa.

5.4 DISCUSSÃO FINAL DO CAPÍTULO

As variáveis esgoto e abastecimento de água possuem uma relação negativa com o número de sinistros. Essas variáveis não afetam diretamente as ocorrências, mais provavelmente os locais onde possui saneamento básico instalado dispõe de uma maior infraestrutura viária, e possivelmente pode oferecer mais segurança aos usuários do trânsito. O mapa “D” da Figura 11 mostra a região sudoeste da cidade como área de maior influência para este fator, onde destaca-se os bairros de Mondubim, Granja Lisboa, Granja Portugal e também Messejana.

Foi verificado que a variável hospital, mantém uma relação positiva com os números de sinistros. Isso ocorre, possivelmente por conta da grande procura por serviços de atendimento

médico, há o movimento maior de pessoas e veículos tornando esses locais. Assim, os bairros Centro, Aldeota, Barra do Ceará, Messejana e Mondubim são exemplos de onde essa relação é mais expressiva.

O modelo constatou que o percentual de área de serviço, comercial e industrial possuem uma relação positiva com a variável dependente. Uma das prováveis razões, é que essas áreas atraem grande quantidade de pessoas, pois concentram diversos serviços essenciais à população, como por exemplo o bairro Centro.

O número de patrimônios tombados tem relação positiva com os sinistros. Isso talvez possa ser explicado pelo potencial turístico desses locais que podem atrair grande número de pessoas e também onde as vias possivelmente são mais estreitas, como ocorrem no bairro Centro onde esse fator é mais expressivo.

A densidade populacional mantém uma correlação positiva com o número de sinistros, e assim, maiores densidades populacionais estão associadas a maiores números de sinistros, como ocorre nos bairros Barra do Ceará, Itaoca, Granja Lisboa e Bonsucesso. A renda média possui uma relação negativa com essas ocorrências, mostrando que quanto maior a renda do bairro, menor será a quantidade de sinistros.

O número de shoppings e agências bancárias também estão positivamente relacionados aos sinistros. Esses locais obviamente atraem muitas pessoas porque são ou concentram serviços essenciais para a população, além disso, tendem a estar concentrados junto com outros serviços, e possivelmente por esses motivos, essas variáveis possuem uma relação positiva com a variável dependente.

Áreas verdes e com planos de arborização, também mantêm uma relação positiva com as ocorrências. Esses locais talvez possam atrair diversas pessoas por conta das praças e áreas de lazer. Como exemplo o bairro Centro, que possui uma relação mais forte para ambas as variáveis utilizadas.

O número de escolas se relaciona positivamente com o número de sinistros nos bairros. Isso pode ser explicado provavelmente pelo grande fluxo de pessoas nas escolas e universidades. Os bairros Barra do Ceará, Mondubim, Granja Lisboa e Bom Jardim são exemplos de locais onde há uma grande relação entre o número de instituições de ensino e as ocorrências.

Já a variável alfabetização, possui uma relação negativa com o número de sinistros, possivelmente porque as pessoas com alto nível de educação conheçam mais as leis de trânsito

e dirija com mais consciência, e essa relação pode ser observada mais forte nos bairros Centro, Aldeota, José Bonifácio e Praia de Iracema, por exemplo.

6 MODELO DE DECISÃO MULTICRITÉRIO PARA PRIORIZAÇÃO DOS HOTSPOTS

Esse capítulo apresenta um modelo de decisão multicritério para a ordenação de locais críticos de sinistros de trânsito de acordo com a criticidade. Os resultados obtidos nas análises espaciais realizadas anteriormente serão fundamentais para a aplicação desse modelo e assim, ordenar os *hotspots* mais críticos considerando esses como prioridades no desenvolvimento de ações mitigadoras e alocação dos recursos.

6.1 DESCRIÇÃO DO PROBLEMA DE DECISÃO MULTICRITÉRIO

Através das análises realizadas nos capítulos anteriores com as ferramentas de estatística espacial, foi possível verificar alguns locais críticos de ocorrências de sinistros de trânsito na cidade de Fortaleza, e que devem ser alvos de intervenções. Além disso, a análise de regressão, mostrou quais os principais fatores, dentre os utilizados, impactam na ocorrência desses sinistros em cada bairro. Essa análise revelou que os coeficientes de regressão são diferentes em cada local da análise, indicando que é necessária uma medida de intervenção específica para cada área de estudo.

Na análise de densidade de Kernel, os mapas de densidade para diferentes anos mostram que a região correspondente ao bairro Centro da cidade, sempre está em evidência, mostrando ser uma região que contém um grande número de sinistros. Mas além dela, outros bairros aparecem em destaque quando as análises são realizadas com os dados de meses e dias da semana separadamente. Já a análise *hotspots* revela diversos *hotspots* por toda área do mapa.

Assim, a análise espacial resultou na identificação de diversos pontos críticos que precisam ser analisados e solucionados através de ações estratégicas desenvolvidas pelas autoridades de trânsito para atenuar as ocorrências de sinistros de trânsito no local de estudo. Entretanto, por conta da limitação de recursos e tempo, muitas vezes não é possível intervir em todos os pontos identificados.

Logo é necessário um modelo de decisão multicritério que possa ordenar esses locais de acordo com a criticidade apresentadas por eles, para que os pontos mais relevantes sejam priorizados e os recursos possam ser alocados de forma a maximizar o impacto de sua aplicação na redução das ocorrências de sinistros de trânsito na cidade de Fortaleza.

6.2 CONSTRUÇÃO DO MODELO DE DECISÃO MULTICRITÉRIO

Esse problema consiste em uma situação de decisão individual. O decisor será o responsável por avaliar a criticidade dos locais identificados como críticos na análise espacial, determinando os objetivos, critérios, e alternativas do problema, além de avaliar as alternativas em cada critério definido e determinar a melhor abordagem para solucionar do problema.

O objetivo principal desse modelo, é obter quais são os locais com maior criticidade de sinistros de trânsito por ordem de urgência, que necessitam de estratégias imediatas para a redução dessas ocorrências. Com isso, esses locais são considerados críticos de acordo com o risco de envolvimento de sinistros com veículos, motocicletas e bicicletas, impacto aos motoristas e pedestres, densidade de trânsito e infraestrutura de trânsito.

Para avaliar a criticidade das vias, foram definidos 14 critérios capazes de analisar os objetivos definidos. O objetivo risco de sinistros, será analisado pelos critérios taxa de sinistros na seção de via, índice de atropelamentos, índices de sinistros com motocicletas, automóveis, ônibus e bicicletas.

O objetivo relacionado aos impactos as vítimas, será avaliado pelos critérios índice de mortes, feridos e ilesos. O objetivo densidade de trânsito será analisado pelo critério definido também como densidade de trânsito. Já a situação da infraestrutura viária será avaliada através dos critérios sinalização, tipo de via, cruzamento e velocidade da via.

A taxa de sinistro é um critério que será avaliado através do número de sinistros na seção de via que o *hotspots* está localizado. Já os critérios índices de mortes, feridos, ilesos, atropelamentos, sinistros com motociclistas e bicicletas, serão avaliados pela porcentagem de sinistros nas vias que registra pelo menos uma dessas consequências. Esses critérios são do tipo construído e de maximização, e dessa forma, quanto maior o valor da alternativa nesses critérios maior será a criticidade dela.

A densidade de trânsito será avaliada através da ferramenta *Google Maps*, que fornece dados do comportamento de trânsito típico para diversas localidades. Essa avaliação diz respeito a rapidez do fluxo nas ruas, classificando-os de rápido para muito lento em diferentes horários do dia, e a classificação do atributo será de acordo com o Quadro 1. Esse critério é do tipo construído discreto e de maximização, ou seja, quanto maior o congestionamento no trânsito, maior o risco de sinistros.

Quadro 1 - Classificação do atributo densidade de trânsito

Nível do atributo	Situação do trânsito
1	Rápido
2	Moderado
3	Lento
4	Muito lento

Fonte: O autor (2022)

O critério sinalização, avaliará as condições de sinalização existente no local, e assim, verifica a existência de sinalização horizontal (como faixas de pedestres, linhas de divisão, marcas de canalização por exemplo) e sinalização vertical (como placas, semáforo e outros), e a qualidade desses. Esses critérios são do tipo construídos discretos e de minimização, ou seja, quando menor o nível do atributo, maior as chances de ocorrências de sinistros, como mostrado no Quadro 2.

Quadro 2 - Classificação do atributo sinalização

Nível do atributo	Situação da sinalização
1	Ausente
2	Ruim
3	Regular
4	Bom
5	Ótima

Fonte: O autor (2022)

No critério tipo de via, será avaliado se o *hotspots* está sobre uma avenida ou rua. Geralmente, a velocidade das avenidas é maior e também o fluxo de veículos são maiores, com isso, a probabilidade de ocorrência de sinistros consequentemente são maiores também. No tipo de cruzamento, os pontos serão classificados como meio de quadra ou cruzamento (podendo ser do tipo cruz, do tipo “T”), como mostra o Quadro 3. Ambos os critérios são do tipo construídos discretos e de maximização.

Quadro 3 - Classificação do atributo para tipo de via e cruzamento

Nível do atributo	Tipo de via	Cruzamento
0	Rua	Meio de quadra
1	Avenida	Cruzamento

Fonte: O autor (2022)

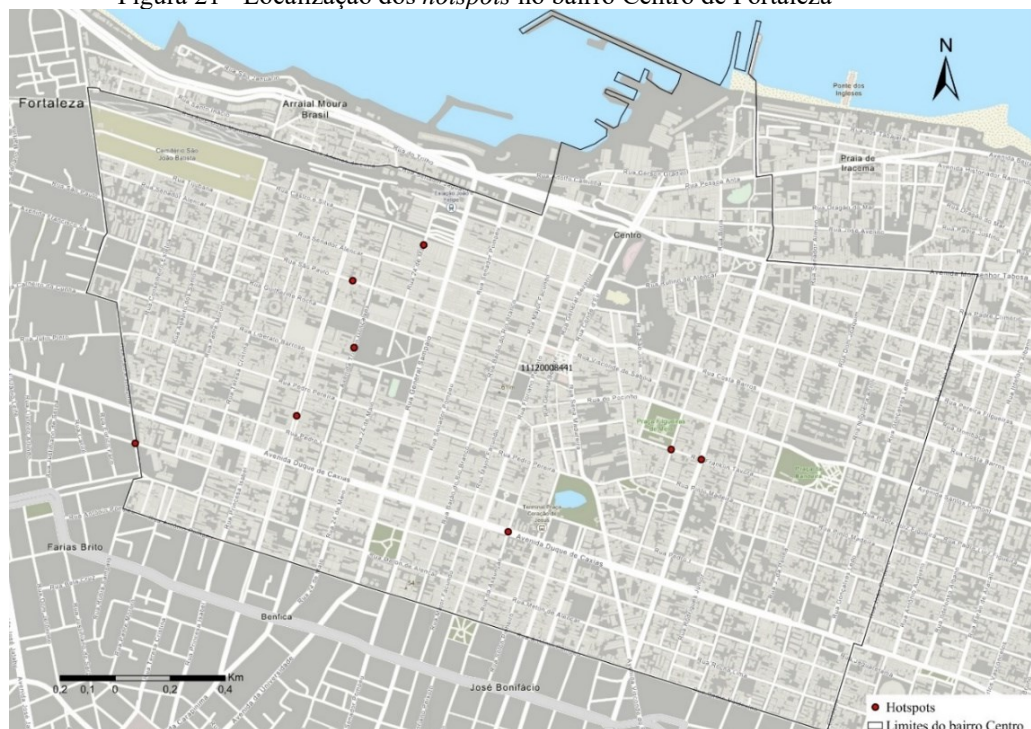
Neste caso, o (0) não indica a ausência de algo, mas sim, que a alternativa é menos relevante do que a classificada como (1). Por fim, o critério velocidade, vai avaliar a velocidade da via, e será um critério do tipo natural.

O FITradeoff de ordenação disponível na *web* foi utilizado para ordenação das alternativas do problema. Os dados foram inseridos na plataforma na *web* (critérios, alternativas, consequência e limiar de indiferença), e logo em seguida, a mesma apresenta as informações colocadas, estando pronta para realizar a primeira etapa da elicitación das preferências do decisor.

O decisor definiu um limiar de indiferença de $\varepsilon = 0,01$ para ser considerado no FITradeoff. De acordo com Frej et al. (2019), esse limiar corresponde a diferença máxima que o valor global entre duas alternativas pode assumir para serem classificadas como indiferenças entre si.

Para simplificação do modelo de decisão, apenas os *hotspots* localizados no bairro Centro serão utilizados como alternativas a serem ranqueadas. Este bairro é o principal foco de sinistros na cidade como mostrado nas análises realizadas nos capítulos anteriores. Com isso, é necessário priorizá-lo. A Figura 21 mostra um mapa do bairro Centro que ilustra localização de cada *hotspots*.

Figura 21 - Localização dos *hotspots* no bairro Centro de Fortaleza



Fonte: O autor (2022)

A Figura acima elucida oito *hotspots* de sinistros de trânsito que foram identificados na análise espacial. Alguns desses pontos estão localizados sobre algumas vias principais do Bairro Centro. Dessa forma, cada ponto identificado e mostrado acima, será considerado como uma alternativa neste problema de decisão. O Quadro 4 apresenta uma descrição de referência para cada ponto de analisado, e a via que está localizado.

Quadro 4 - Descrição das alternativas

Alternativa	Referência
Ponto 1	Cruzamento entre a rua general Clarindo Queiroz e rua Padre Ibiapina
Ponto 2	Avenida imperador
Ponto 3	Avenida Tristão Gonçalves
Ponto 4	Cruzamento da Rua da Assunção com Avenida Duque de Caxias
Ponto 5	Cruzamento entre a rua Franklin Távora e Avenida Dom Manoel
Ponto 6	Rua São Paulo
Ponto 7	Cruzamento entre as ruas Franklin Távora e Rua 25 de março
Ponto 8	Rua 24 de maio

Fonte: O autor (2022)

6.3 RESULTADOS

A primeira etapa da elicitação consistia numa avaliação holística para classificar os critérios, ordenando assim, as constantes de escala. As constantes de escala foram colocadas na seguinte ordem: índice de mortes (C1), índice de feridos (C2), índice de sinistros (C3), índice de atropelamentos (C4), índice de sinistros com bicicletas (C5), índice de sinistros com motocicleta (C6), densidade de trânsito (C7), índice de sinistros com automóveis (C8), índice de sinistros com ônibus (C9), índice de ilesos (C10), cruzamento (C11), sinalização (C12), velocidade (C13) e tipo de via (C14).

Realizada a avaliação, é iniciado a etapa da elicitação flexível, onde a plataforma irá apresentar a comparação entre duas consequências hipotéticas e o decisor responde de acordo com suas preferências. Além disso, ele pode optar por não responder. A cada pergunta respondida, restrições serão inseridas no problema de programação linear, e o ranqueamento vai sendo formado através das relações de dominâncias entre as alternativas.

Para esse problema, um total dezessete perguntas foram respondidas pelo decisor. As duas consequências hipotéticas apresentadas pela plataforma, as respectivas respostas do

decisor, e o número de níveis gerados que são apresentados através das preferências do decisor naquela etapa são apresentadas Tabela 21.

Nesse problema, o decisor conseguiu responder diretamente a dez questionamentos escolhendo uma das consequências hipotéticas mostrada. O mesmo ficou indiferente em seis das dezessete situações lhe apresentada, e não quis responder uma das perguntas realizadas.

Tabela 21 - Dados do processo de elicitación

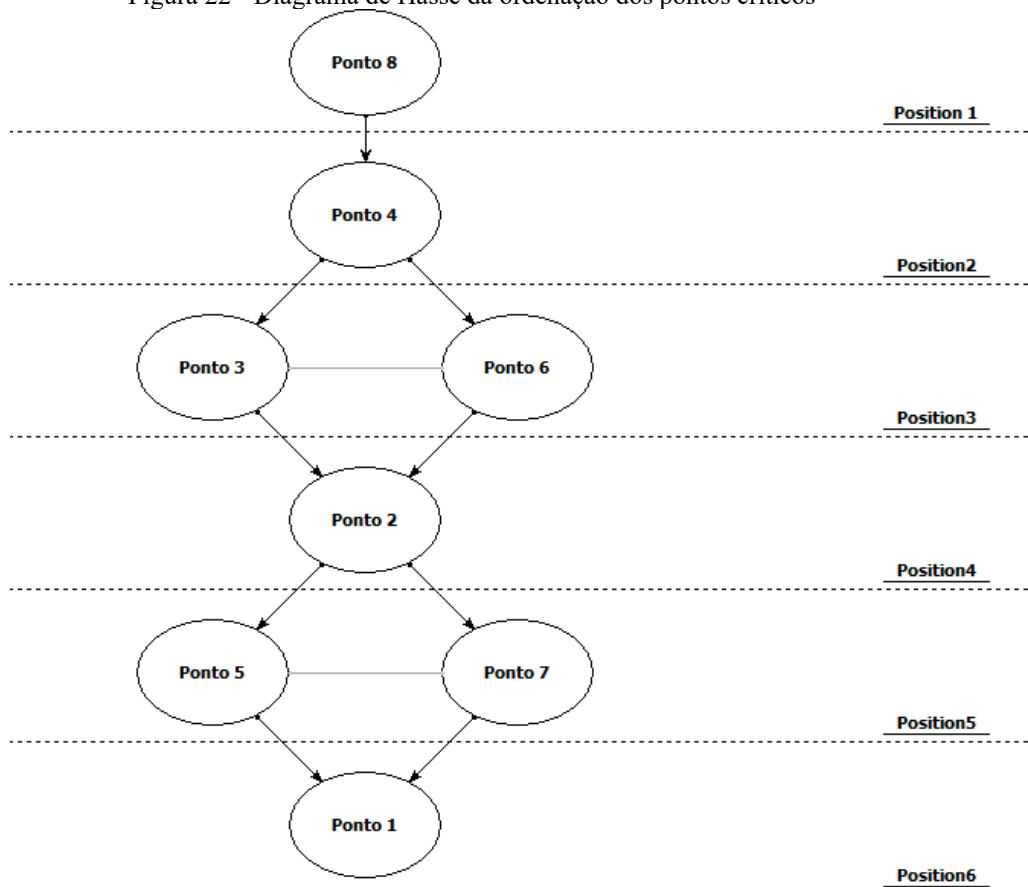
Ciclo	consequência A	consequência B	Resposta	Número de níveis
0	-	-	Ordenação	1
1	(0.125) de C1	(1) de C14	Consequência A	1
2	(0.125) de C1	(1) de C2	Consequência A	1
3	(0.5) de C2	(45.76) de C3	Consequência A	1
4	(25.396) de C3	(0.3409) de C4	Consequência B	1
5	(0.212) de C4	(0.25) de C5	Indiferente	3
6	(0.125) de C5	(0.67647) de C6	Consequência B	3
7	(0.463) de C6	(4) de C7	Consequência A	4
8	(3) de C7	(0.7894) de C8	Consequência B	4
9	(0.603) de C8	(0.333) de C9	Indiferente	4
10	(0.167) de C9	(0.25) de C10	Consequência A	5
11	(289473.625) de C10	(1) de C11	Não respondeu	5
12	(3) de C12	(50) de C13	Consequência B	5
13	(40) de C13	(1) de C14	Indiferente	5
14	(0.063) de C1	(1) de C2	Indiferente	5
15	(0,25) de C2	(45,76) de C3	Indiferente	6
16	(35,578) de C3	(0,3409) de C4	Consequência B	6
17	(0,188) de C5	(0,67647) de C6	Indiferente	7

Fonte: O autor (2022)

Como resultado, a pré-ordenação das alternativas é então alcançada, e pode ser facilmente visualizada através do diagrama de *hasse* mostrado na Figura 22. O diagrama mostra a classificação de cada alternativa nos níveis, e a relação que essa alternativa possui com as demais.

As oito alternativas, foram ordenadas em 6 níveis. A alternativa ponto 8, está na primeira posição do ranking, mostrando ser o ponto que mais necessita de urgência, pois domina as demais alternativas. No nível dois, está a alternativa 4. A alternativa ponto 3 e ponto 6 estão no nível três, e são indiferentes entre si. Em seguida, ficou a alternativa ponto 2 no nível quatro, ponto 5 e 7 são indiferentes e estão no nível cinco, e no último, a alternativa ponto 1, sendo está dominada pelas demais alternativas.

Figura 22 - Diagrama de Hasse da ordenação dos pontos críticos



Fonte: O autor (2022)

6.4 DISCUSSÃO FINAL DO CAPÍTULO

Com a aplicação de modelo de decisão, o ponto *hotspot* localizado na Rua 24 de Maio foi ordenado como o mais crítico de todos, e dessa forma, este local e seu entorno deve ser prioridade para ser investigado e na implementação medidas de contenção de sinistros.

O segundo ponto mais importante é o cruzamento da Rua da Assunção com avenida Duque de Caxias. Por mais que o local já tenha toda infraestrutura adequada para a sinalização horizontal e vertical, as autoridades devem direcionar seus esforços para propor uma medida adicional mais efetiva ou revisar as medidas existentes nessa localidade. Em terceiro lugar, estão os pontos 3 e 6, que correspondem a um ponto na Avenida Tristão Gonçalves e outro na Rua São Paulo, respectivamente.

O desenvolvimento do modelo de decisão foi crucial para identificação dos *hotspots* mais críticos do bairro considerado. Esse resultado alcançado é uma base de referência

importante para que as autoridades possam direcionar seus esforços para analisar primeiramente esses pontos ordenados de acordo com a sua criticidade.

E assim, com essa ordenação disponível, é necessário investigar as principais falhas nos diferentes tipos de sinalização existentes nessas localidades, pois, a maioria dos locais tem sinalizações verticais e horizontais bem definidas. Assim, será possível entender o porquê do elevado número de sinistros nesses pontos críticos e seu entorno, mesmo já tendo sinalizações de trânsito implantadas e dessa forma, propor melhorias para redução dessas ocorrências.

A abordagem do FITradeoff de ordenação foi eficiente e trouxe um resultado satisfatório para o problema. Através do diagrama de Hasse foi possível verificar graficamente de maneira mais fácil a ordenação final e a ordenação parcial obtida a cada resposta do decisor ao longo do processo. A elicitación das preferências de forma interativa e flexível, facilitou o andamento desse procedimento.

7 CONCLUSÃO

As abordagens espaço-temporal utilizadas, neste estudo, para dar suporte à análise do padrão espacial dos sinistros de trânsito, foram úteis na identificação de informações relevantes para o desenvolvimento de políticas de mitigação das ocorrências dos sinistros de trânsito na cidade de Fortaleza.

Com a análise de densidade de kernel, foi possível identificar a região norte de Fortaleza como o local mais crítico para os acontecimentos desses sinistros. Além disso, observou-se que a distribuição dessas ocorrências é diferente em cada região da cidade. Para complementar essa análise, o resultado da autocorrelação global mostra que para os dados dos anos, há uma tendência de agrupamento, e a autocorrelação local com a análise *hotspots* mostra de forma mais pontual os locais mais críticos, facilitando ainda mais na análise desses acontecimentos.

Através da análise de ponto padrão, foi constatado que não existe similaridade entre os padrões de pontos analisados, identificando um padrão específico para cada período. Assim, essas análises devem ser realizadas periodicamente, para identificar possíveis diferenças nos padrões e poder reestruturar os planos de ações adotados, já que há mudanças a cada período.

As variáveis socioeconômicas utilizadas no modelo GWR foram adequadas e apresentaram um resultado satisfatório ao explicar grande parte das variações dos sinistros viários. Entretanto, a falta de dados relacionados aos fatores de trânsito foi uma limitação para esse trabalho. Não foi possível a utilização desse tipo de dado no modelo desenvolvido por conta da indisponibilidade dos mesmos.

Esse fato de certa forma, impactou o resultado do modelo de regressão geograficamente ponderado, por não investigar a relação dos sinistros com variáveis de trânsito, como por exemplo o fluxo de veículos que passam em uma determinada via, fluxo de pedestres, qualidade da via, qualidade da sinalização e etc. Entretanto, essa limitação pode ser explorada em trabalhos futuros, podendo adequar essa análise à modelos de regressão utilizando Poisson ou binomial negativa, de acordo com a distribuição observada dos dados.

O modelo de decisão multicritério desenvolvido com a utilização do FITradeoff para problemática de ordenação foi eficiente ao realizar a ordenação dos pontos *hotspots* do bairro Centro da Fortaleza de maneira interativa e flexível. A ordem das alternativas obtidas, é fundamental para a priorização da aplicação de estratégias de contenção. Entretanto, o modelo proposto tem uma limitação em relação aos dados e as partes interessadas. Não foi possível realizar essa aplicação com as autoridades de trânsito da cidade.

Para trabalhos futuros, podem ser considerados métodos de estruturação de problemas para auxiliar no desenvolvimento do modelo multicritério, e utilizando as informações das principais partes interessadas na questão, que são responsáveis pela segurança viária no local de aplicação da pesquisa e que possui um vasto conhecimento nesse tema.

Contudo, os resultados obtidos com a pesquisa são essenciais para que as autoridades de trânsito possam desenvolver as medidas de intervenções adequadas, e direcionar seus esforços para atender os locais nos diferentes meses e dias da semana, e assim, conseguir resolver a problemática, e realizar a alocação de recursos de forma mais eficientes.

Com isso, o modelo espaço-temporal para a análise de sinistros de trânsito desenvolvido com os ferramentais utilizadas foram fundamentais para alcançar os objetivos definidos para esse trabalho, e dessa forma, foi possível identificar o padrão das ocorrências ao longo do espaço e do tempo, além de priorizar os *hotspots* de sinistros com o modelo de decisão multicritério proposto.

O desenvolvimento dessa pesquisa foi muito desafiador pelo fato de envolver diversas abordagens até então não conhecidas pelo autor, que foram exploradas para trazer um resultado satisfatório para a pesquisa. Além disso, foi uma satisfação enorme trabalhar com um tema que envolve a área de transporte e que é tão relevante para a sociedade, e poder gerar resultados satisfatórios para a solução da problemática analisada.

Além de oferecer um arcabouço metodológico para análise de sinistros de trânsito, esse trabalho contribui para a literatura, oferecendo uma revisão detalhada de abordagens que podem ser aplicadas na identificação de locais críticos de sinistros de trânsito que podem ser essenciais para o desenvolvimento de futuros trabalhos.

REFERÊNCIAS

- AL-OMARI, A., SHATNAWI, N., KHEDAYWI, T., & MIQDADY, T. Prediction of traffic accidents hot spots using fuzzy logic and GIS. *Applied Geomatics*, v. 12, n. 2, p. 149-161, 2020.
- ANDERSON, Tessa K. Kernel density estimation and K-means clustering to profile road accident hotspots. *Accident Analysis & Prevention*, v. 41, n. 3, p. 359-364, 2009.
- ANDRESEN, Martin A. Testing for similarity in area-based spatial patterns: A nonparametric Monte Carlo approach. *Applied Geography*, v. 29, n. 3, p. 333-345, 2009.
- AYOB, N., MUSTAPHA, M. A., SENAWI, J., & AHMAD, N. Herpetofauna Roadkills on Langkawi Island, Peninsular Malaysia: The Influence of Landscape and Season on Mortality Distribution. *Sains Malaysiana*, v. 49, n. 10, p. 2373-2382, 2020.
- BAO, J., LIU, P., YU, H., & XU, C. Incorporating twitter-based human activity information in spatial analysis of crashes in urban areas. *Accident analysis & prevention*, v. 106, p. 358-369, 2017.
- BRUNSDON, Chris; FOTHERINGHAM, A. Stewart; CHARLTON, Martin E. Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis*, v. 28, n. 4, p. 281-298, 1996.
- BURROUGH, Peter A.; MCDONNELL, Rachael A.; LLOYD, Christopher D. *Principles of geographical information systems*. Oxford university press, 2015.
- BUTT, A., AHMAD, S. S., SHABBIR, R., & ERUM, S. GIS based surveillance of road traffic accidents (RTA) risk for Rawalpindi city: A geo-statistical approach. *Kuwait Journal of Science*, v. 44, n. 4, 2017.
- ČUBRANIĆ-DOBRODOLAC, M., ŠVADLENKA, L., ČIČEVIĆ, S., TRIFUNOVIĆ, A., & DOBRODOLAC, M. Using the Interval Type-2 Fuzzy Inference Systems to Compare the Impact of Speed and Space Perception on the Occurrence of Road Traffic Accidents. *Mathematics*, v. 8, n. 9, p. 1548, 2020.
- DAGAEVA, M., GARAeva, A., ANIKIN, I., MAKHMUTOVA, A., & MINNIKHANOV, R. Big spatio-temporal data mining for emergency management information systems. *IET Intelligent Transport Systems*, v. 13, n. 11, p. 1649-1657, 2019.
- DALE, Mark RT; FORTIN, Marie-Josée. *Spatial analysis: a guide for ecologists*. Cambridge University Press, 2014.
- EDWARDS, Ward; BARRON, F. Hutton. SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement. *Organizational behavior and human decision processes*, v. 60, p.306-325, 1994.
- ERDOGAN, Saffet. Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. *Journal of safety research*, v. 40, n. 5, p. 341-351, 2009.

ERSEN, Mert; BÜYÜKLÜ, Ali Hakan; TAŞABAT, Semra Erpolat. Analysis of Fatal and Injury Traffic Accidents in Istanbul Sarıyer District with Spatial Statistics Methods. *Sustainability*, v. 13, n. 19, p. 11039, 2021.

FISCHER, Manfred M.; GETIS, Arthur (Ed.). Handbook of applied spatial analysis: software tools, methods and applications. Berlin: *springer*, 2010.

FOTHERINGHAM, A. Stewart; BRUNSDON, Chris; CHARLTON, Martin. *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons, 2002.

FOTHERINGHAM, A. Stewart; CHARLTON, Martin E.; BRUNSDON, Chris. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and planning A*, v. 30, n. 11, p. 1905-1927, 1998.

FRANCIS, S. S., ENDERS, C., HYDE, R., GAO, X., WANG, R., MA, X., ... & METAYER, C. Spatial–Temporal Cluster Analysis of Childhood Cancer in California. *Epidemiology*, v. 31, n. 2, p. 214-223, 2020.

FREJ, Eduarda Asfora; DE ALMEIDA, Adiel Teixeira; COSTA, Ana Paula Cabral Seixas. Using data visualization for ranking alternatives with partial information and interactive tradeoff elicitation. *Operational Research*, v. 19, n. 4, p. 909-931, 2019.

GOMES, Marcos José Timbó Lima; CUNTO, Flavio; DA SILVA, Alan Ricardo. Geographically weighted negative binomial regression applied to zonal level safety performance models. *Accident Analysis & Prevention*, v. 106, p. 254-261, 2017.

HARIRFOROUSH, Homayoun; BELLALITE, Lynda. A new integrated GIS-based analysis to detect hotspots: a case study of the city of Sherbrooke. *Accident Analysis & Prevention*, v. 130, p. 62-74, 2019.

HASHIMOTO, S., YOSHIKI, S., SAEKI, R., MIMURA, Y., ANDO, R., & NANBA, S. Development and application of traffic accident density estimation models using kernel density estimation. *Journal of traffic and transportation engineering (English edition)*, v. 3, n. 3, p. 262-270, 2016.

HU, L., WU, X., HUANG, J., PENG, Y., & LIU, W. Investigation of clusters and injuries in pedestrian crashes using GIS in Changsha, China. *Safety science*, v. 127, p. 104710, 2020.

IBEAS, Á., CORDERA, R., DELL'OLIO, L., & MOURA, J. L. Modelling demand in restricted parking zones. *Transportation Research Part A: Policy and Practice*, v. 45, n. 6, p. 485-498, 2011.

IBGE – Instituto Brasileiro de Geografia e Estatística (2016). Censo Demográfico - Resultados do universo: características da população e domicílios. <http://www.ibge.gov.br>. Acessado: 2021-05-05.

IQBAL, A., UR REHMAN, Z., ALI, S., ULLAH, K., & GHANI, U. Road Traffic Accident Analysis and Identification of Black Spot Locations on Highway. *Civil Engineering Journal*, v. 6, n. 12, p. 2448-2456, 2020.

IYANDA, Ayodeji E.; OSAYOMI, Tolulope. Is there a relationship between economic indicators and road fatalities in Texas? A multiscale geographically weighted regression analysis. *GeoJournal*, v. 86, n. 6, p. 2787-2807, 2021.

JAKIMAVIČIUS, Marius. Analysis and Assessment of Lithuanian Road Accidents by AHP Method. *The Baltic Journal of Road and Bridge Engineering*, v. 13, n. 3, p. 238-260, 2018.

JIANG, M., SUN, W., YANG, G., & ZHANG, D. Modelling seasonal GWR of daily PM_{2.5} with proper auxiliary variables for the Yangtze River Delta. *Remote Sensing*, v. 9, n. 4, p. 346, 2017.

KALANTARI, M., ZANGANEH SHAHRAKI, S., YAGHMAEI, B., GHEZELBASH, S., LADAGA, G., & SALVATI, L. Unraveling urban form and collision risk: the spatial distribution of traffic accidents in Zanjan, Iran. *International journal of environmental research and public health*, v. 18, n. 9, p. 4498, 2021.

KANUGANTI, S., AGARWALA, R., DUTTA, B., BHANEGAONKAR, P. N., SINGH, A. P., & SARKAR, A. K. Road safety analysis using multi criteria approach: A case study in India. *Transportation research procedia*, v. 25, p. 4649-4661, 2017.

KAVOUSHI, A., MORADI, A., SOORI, H., RAHMANI, K., ZEINI, S., & BONAKCHI, H. Environmental factors affecting the frequency of traffic accidents leading to death in 22 districts of Tehran during 2014–2016. *Archives of Trauma Research*, v. 8, n. 1, p. 33-40, 2019.

KIBAR, Funda Ture; TUYDES-YAMAN, Hediye. GIS-based evaluation of the speed limit increase on urban arterial traffic safety in Ankara. *Arabian Journal of Geosciences*, v. 13, n. 12, p. 1-16, 2020.

LAKES, Tobia. A spatially explicit analysis of traffic accidents involving pedestrians and cyclists in Berlin. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, v. 60, n. 12, p. 1328-1335, 2017.

LE, Khanh Giang; LIU, Pei; LIN, Liang-Tay. Determining the road traffic accident hotspots using GIS-based temporal-spatial statistical analytic techniques in Hanoi, Vietnam. *Geo-spatial Information Science*, v. 23, n. 2, p. 153-164, 2020b.

LE, Khanh Giang; LIU, Pei; LIN, Liang-Tay. Traffic accident hotspot identification by integrating kernel density estimation and spatial autocorrelation analysis: a case study. *International Journal of Crashworthiness*, p. 1-11, 2020a.

LEUNG, Yee; LI, Rongrong; JI, Nannan. Application of extended Dempster–Shafer theory of evidence in accident probability estimation for dangerous goods transportation. *Journal of Geographical Systems*, v. 19, n. 3, p. 249-271, 2017.

LI, Yafei; LIANG, Chen. The analysis of spatial pattern and hotspots of aviation accident and ranking the potential risk airports based on gis platform. *Journal of Advanced Transportation*, v. 2018, 2018.

- LIU, J., KHATTAK, A. J., CHEN, C., WAN, D., MA, J., & HU, J. Revisiting hit-and-run crashes: a geo-spatial modeling method. *Transportation research record*, v. 2672, n. 38, p. 81-92, 2018.
- LIU, Z., HE, J., ZHANG, C., XING, L., & ZHOU, B. The impact of road alignment characteristics on different types of traffic accidents. *Journal of Transportation Safety & Security*, v. 12, n. 5, p. 697-726, 2020.
- MACHADO, L. C.; SOUZA, A. P. G.; MOTA, C. M. M. Análise espacial de locais críticos para segurança de tráfego. In: SBPO 2021 - LII Simpósio de Pesquisa Operacional, 2021, João Pessoa. *Anais do LIII Simpósio Brasileiro de Pesquisa Operacional*. Rio de Janeiro: SOBRAPO, 2021.
- MANAP, N., BORHAN, M. N., YAZID, M. R. M., HAMBALI, M. K. A., & ROHAN, A. Identification of hotspot segments with a risk of heavy-vehicle accidents based on spatial analysis at controlled-access highway. *Sustainability*, v. 13, n. 3, p. 1487, 2021.
- MARTINS, M. A., GARCEZ, T. V., DE GUSMÃO, A. P. H., SILVA, L. G., & DE ALMEIDA, J. A Multicriteria Model Based on FITradeoff Method for Prioritizing Sections of Brazilian Roads by Criticality. *Mathematical Problems in Engineering*, v. 2020, 2020.
- MARTINS, Mateus Alves; GARCEZ, Thalles Vitelli. A. Multidimensional and multi-period analysis of safety on roads. *Accident Analysis & Prevention*, v. 162, p. 106401, 2021.
- MARZOUG, R., BAMAAROUF, O., LAKOUARI, N., CASTILLO-TÉLLEZ, B., TÉLLEZ, M. C., & OUBRAM, O. Traffic intersection characteristics with accidents and evacuation of damaged cars. *Physica A: Statistical Mechanics and its Applications*, v. 561, p. 125217, 2021.
- MOHAMMADINIA, A., SAEIDIAN, B., PRADHAN, B., & GHAEMI, Z. Prediction mapping of human leptospirosis using ANN, GWR, SVM and GLM approaches. *BMC infectious diseases*, v. 19, n. 1, p. 1-18, 2019.
- NASIR, Nor AM; MAULUD, Khairul NA; YUSOFF, Nur IM. Geospatial analysis of road distresses and the relationship with the slope factor. *Journal of Engineering Science and Technology*, v. 11, n. 5, p. 655-665, 2016.
- OSPINA-MATEUS, H., QUINTANA JIMÉNEZ, L. A., LOPEZ-VALDES, F. J., BERRIO GARCIA, S., BARRERO, L. H., & SANA, S. S. Extraction of decision rules using genetic algorithms and simulated annealing for prediction of severity of traffic accidents by motorcyclists. *Journal of Ambient Intelligence and Humanized Computing*, p. 1-22, 2021.
- OUNI, Fedy; BELLOUMI, Mounir. Pattern of road traffic crash hot zones versus probable hot zones in Tunisia: A geospatial analysis. *Accident Analysis & Prevention*, v. 128, p. 185-196, 2019.
- ÖZCAN, Mehmet; KÜÇÜKÖNDER, Muhterem. Investigation of Spatiotemporal Changes in the Incidence of Traffic Accidents in Kahramanmaraş, Turkey, Using GIS-Based Density Analysis. *Journal of the Indian Society of Remote Sensing*, v. 48, n. 7, p. 1045-1056, 2020.

PASUPATHI, S., SHANMUGANATHAN, V., MADASAMY, K., YESUDHAS, H. R., & KIM, M. Trend analysis using agglomerative hierarchical clustering approach for time series big data. *The Journal of Supercomputing*, p. 1-20, 2021.

RAHMAN, Farzana; KUBOTA, Hisashi. Point scoring system to rank traffic calming projects. *Journal of traffic and transportation engineering (English edition)*, v. 3, n. 4, p. 324-335, 2016.

RAHMAN, Muhammad Tauhidur; JAMAL, Arshad; AL-AHMADI, Hassan M. Examining Hotspots of Traffic Collisions and their Spatial Relationships with Land Use: A GIS-Based Geographically Weighted Regression Approach for Dammam, Saudi Arabia. *ISPRS International Journal of Geo-Information*, v. 9, n. 9, p. 540, 2020.

RAMÍREZ, Andrés Felipe; VALENCIA, Carlos. Spatiotemporal correlation study of traffic accidents with fatalities and injuries in Bogota (Colombia). *Accident Analysis & Prevention*, v. 149, p. 105848, 2021.

RENNERMALM, Asa K.; BRING, Arvid; MOTE, Thomas L. Spatial and scale-dependent controls on North American Pan-Arctic minimum river discharge. *Geographical Analysis*, v. 44, n. 3, p. 202-218, 2012.

SAADAT, S., RAHMANI, K., MORADI, A., AD-DIN ZAINI, S., & DARABI, F. Spatial analysis of driving accidents leading to deaths related to motorcyclists in Tehran. *Chinese journal of traumatology*, v. 22, n. 3, p. 148-154, 2019.

SANDHU, H. A. S., SINGH, G., SISODIA, M. S., & CHAUHAN, R. Identification of black spots on highway with kernel density estimation method. *Journal of the Indian Society of Remote Sensing*, v. 44, n. 3, p. 457-464, 2016.

Seguradora líder DPVAT. Relatório anual 2019. Disponível em: <https://www.seguradoralider.com.br/Documents/Relatorio-Anual-2019.pdf?#zoom=65%>. Acessado em 25/01/2021.

SELVI, Huseyin Zahit; CAGLAR, Burak. Using cluster analysis methods for multivariate mapping of traffic accidents. *Open Geosciences*, v. 10, n. 1, p. 772-781, 2018.

SILVERMAN, B. W. Density Estimation for Statistics and Data Analysis, Chapman and Hall, London, 1986. *Crossref*, á, 1986.

SHAFABAKHSH, Gholam Ali; FAMILI, Afshin; BAHADORI, Mohammad Sadegh. GIS-based spatial analysis of urban traffic accidents: Case study in Mashhad, Iran. *Journal of traffic and transportation engineering (English edition)*, v. 4, n. 3, p. 290-299, 2017.

SJAFRI, F. S., ABDUL MAULUD, K. N., WAN MOHD JAAFAR, W. S., ARIF, F., AB RAHMAN, A. A., & MUKHLISIN, M. Development of road maintenance inventory in UKM by using aerial images. *Jurnal Kejuruteraan*, n. 2, p. 65-75, 2018.

SOLTANI, Ali; ASKARI, Sajad. Exploring spatial autocorrelation of traffic crashes based on severity. *Injury*, v. 48, n. 3, p. 637-647, 2017.

SOROORI, Emad; MOHAMMADZADEH MOGHADDAM, Abolfazl; SALEHI, Mahdi. Modeling spatial nonstationary and overdispersed crash data: Development and comparative analysis of global and geographically weighted regression models applied to macrolevel injury crash data. *Journal of Transportation Safety & Security*, p. 1-25, 2020.

SOUZA, C. D. F. D., PAIVA, J. P. S. D., LEAL, T. C., SILVA, L. F. D., MACHADO, M. F., & ARAÚJO, M. D. P. D. Mortality in motorcycle accidents in Alagoas (2001-2015): temporal and spatial modeling before and after the “lei seca”. *Revista da Associação Médica Brasileira*, v. 65, n. 12, p. 1482-1488, 2020.

SUN, Y., WANG, Y., YUAN, K., CHAN, T. O., & HUANG, Y. Discovering Spatio-Temporal Clusters of Road Collisions Using the Method of Fast Bayesian Model-Based Cluster Detection. *Sustainability*, v. 12, n. 20, p. 8681, 2020.

SYAHPUTRI, K., SARI, R. M., RIZKYA, I., TARIGAN, U., SIREGAR, I., & FARHAN, T. A. Clustering the vulnerability of traffic accidents in Medan city with fuzzy c-means algorithm. In: *IOP Conference Series: Materials Science and Engineering*. IOP Publishing, 2020. p. 012030.

TAAMNEH, Madhar; TAAMNEH, Salah; ALKHEDER, Sharaf. Clustering-based classification of road traffic accidents using hierarchical clustering and artificial neural networks. *International journal of injury control and safety promotion*, v. 24, n. 3, p. 388-395, 2017.

TEMRUNGSIE, W., RAKSUNTRON, W., NAMEE, N., CHAYANAN, S., & WITCHAYANGKON, B. AHP-based prioritization on road accidents factors: A case study of Thailand. *International Transaction Journal of Engineering, Management, Applied Sciences & Technologies*, v. 6, n. 4, p. 135-144, 2015.

TIAN, Zhun; ZHANG, Shengrui. Application of big data optimized clustering algorithm in cloud computing environment in traffic accident forecast. *Peer-to-Peer Networking and Applications*, p. 1-13, 2020.

TOLA, A. M., DEMISSIE, T. A., SAATHOFF, F., & GEBISSA, A. Severity, Spatial Pattern and Statistical Analysis of Road Traffic Crash Hot Spots in Ethiopia. *Applied Sciences*, v. 11, n. 19, p. 8828, 2021.

WANG, S., LI, Z., ZHANG, J., YUAN, Y., & LIU, Z. The crash injury severity prediction of traffic accident using an improved wrappers feature selection algorithm. *International Journal of Crashworthiness*, p. 1-12, 2021.

World Health Organization. Leading causes of death and disability. Disponível em: <https://www.who.int/data/stories/leading-causes-of-death-and-disability-2000-2019-a-visual-summary>. Acessado em 25/01/2021.

YAACOB, N. F. F., RUSLIB, N., BOHARIC, S. N., YAZIDA, M. R. M., & DASD, A. M. Integrated GIS Tool for Investigating the Relationship of Road Characteristics with Road Traffic Accidents. *Jurnal Kejuruteraan* 32(4) 683-691, 2020.

- YILDIRIM, Vural; MERT KANTAR, Yeliz. Spatial analysis of the road traffic accident statistics in the provinces of Turkey. *Sigma: Journal of Engineering & Natural Sciences/Mühendislik ve Fen Bilimleri Dergisi*, v. 38, n. 4, , p. 1667-1680, 2020.
- YU, L., DU, B., HU, X., SUN, L., HAN, L., & LV, W. Deep spatio-temporal graph convolutional network for traffic accident prediction. *Neurocomputing*, v. 423, p. 135-147, 2021.
- YUE, X., Ding, Y., Hu, H., & Zhao, D. Analysis on the association between driver's macroscopic characteristics and accident type of urban traffic accidents. In: *2017 36th Chinese Control Conference (CCC)*. IEEE, p. 5658-5662, 2017.
- ZAHID, M., CHEN, Y., KHAN, S., JAMAL, A., IJAZ, M., & AHMED, T. Predicting risky and aggressive driving behavior among taxi drivers: do spatio-temporal attributes matter?. *International journal of environmental research and public health*, v. 17, n. 11, p. 3937, 2020.
- ZHANG, Chaohui; LI, Yijing; LI, Tian. A road traffic accidents prediction model for traffic service robot. *Library Hi Tech*, 2020.
- ZHANG, Y., BIGHAM, J., RAGLAND, D., & CHEN, X. Investigating the associations between road network structure and non-motorist accidents. *Journal of transport geography*, v. 42, p. 34-47, 2015.
- ZOU, Yajie; ZHANG, Yue; CHENG, Kai. Exploring the Impact of Climate and Extreme Weather on Fatal Traffic Accidents. *Sustainability*, v. 13, n. 1, p. 390, 2021.