



**UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
ENGENHARIA CIVIL
DEPARTAMENTO DE ENGENHARIA DE TRANSPORTES**

MATHEUS GOMES DE LIMA

**AVALIAÇÃO DO PROBLEMA DA UNIDADE DE ÁREA MODIFICÁVEL (PUAM)
NA TAXA DE ACIDENTES DE TRÂNSITO AGREGADOS EM ÁREAS**

FORTALEZA

2017

MATHEUS GOMES DE LIMA

AVALIAÇÃO DO PROBLEMA DA UNIDADE DE ÁREA MODIFICÁVEL (PUAM) NA
TAXA DE ACIDENTES DE TRÂNSITO AGREGADOS EM ÁREAS

Monografia apresentada ao Departamento de Engenharia de Transportes do Curso de Engenharia Civil da Universidade Federal do Ceará, como requisito parcial à obtenção do título de bacharel em Engenharia Civil.

Orientador: Prof. Flávio José Craveiro Cunto, PhD.

FORTALEZA

2017

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- L699a Lima, Matheus.
Avaliação do Problema de Unidade de Área Modificável (PUAM) na taxa de acidentes de trânsito agregados em áreas / Matheus Lima. – 2017.
80 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Tecnologia, Curso de Engenharia Civil, Fortaleza, 2017.
Orientação: Prof. Dr. Flávio José Craveiro Couto.
1. Problema de Unidade de Área Modificável. 2. Análise Exploratória Espacial. 3. Análise Exploratória Não Espacial. 4. Modelos de Regressão Espacial. I. Título.

CDD 620

MATHEUS GOMES DE LIMA

AVALIAÇÃO DO PROBLEMA DA UNIDADE DE ÁREA MODIFICÁVEL (PUAM) NA
TAXA DE ACIDENTES DE TRÂNSITO AGREGADOS EM ÁREAS

Monografia apresentada ao Departamento de Engenharia de Transportes do Curso de Engenharia Civil da Universidade Federal do Ceará, como requisito parcial à obtenção do título de bacharel em Engenharia Civil.

Aprovada em: 01/12/2017.

BANCA EXAMINADORA

Prof. Flávio José Craveiro Cunto, PhD. (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Me. Marcos José Timbó Lima Gomes
Universidade Federal do Cariri (UFCA)

Prof. Francisco Moraes de Oliveira Neto, PhD.
Universidade Federal do Ceará (UFC)

À minha família, por todo o cuidado durante
esses anos e pela formação da pessoa que sou
hoje.

AGRADECIMENTOS

Aos meus pais, Agar e João, que me apoiaram com todo o seu cuidado e coragem sem medidas ao longo desse período.

A Saskia, minha irmã e amiga, por todo o apoio e incentivo.

A Dona Edith e Francisco Amadeu *in memorian*, meus avós, também pelo apoio e cuidado.

A Elizabeth e Amaury, por todo o cuidado de mãe e pai que me deram.

Aos meus irmãos de outras mães, Erisson, Leodécio, Evandro, Alison, Erly Júnior, Emy, Wendel, Jefferson, pelos momentos vividos durante esses anos.

Aos meus amigos, Breno, Madson e Mateus, pela experiência de ter dividido o mesmo teto.

Aos amigos que a Engenharia me deu, Rodrigo, Manoel, Lucas, Yan, João Pedro, José Aragão, Carlos Henrique, Felipe do Nascimento, Artur Holanda, Artur Costa, Webert, Israel, por tornarem mais leve a jornada da graduação.

A todos os meus amigos do Grupo de Estudantes Adventistas da UFC (GEA UFC), por me fortalecer espiritualmente nesse período.

Ao Professor Flávio Cunto, pelas experiências compartilhadas durante grande parte da minha graduação através das orientações.

Ao Professor Marcos Timbó, por ter feito parte de etapas importantes durante as pesquisas na graduação.

Ao Professor Francisco Moraes por estar fazendo parte da banca examinadora.

Ao Eng^o Francisco Eudes de Aguiar Costa, ao Senhor Oberdan Paiva e a todos os membros da empresa Construtec Engenharia e IE Construções pelas oportunidades e confiança dadas a mim.

RESUMO

A observação e avaliação dos efeitos do Problema de Unidade de Área Modificável tornam-se uma questão importante na modelagem do Desempenho da Segurança Viária, pois, caso não sejam identificados os adequados níveis de escala geográfica e de configuração de zonas a modelagem estatística baseada em dados espaciais pode induzir a conclusões enganosas. O objetivo geral desse trabalho é avaliar a influência do Problema de Unidade de Área Modificável (PUAM) na frequência de acidentes de trânsito agregados em áreas. As etapas metodológicas definidas para este trabalho de conclusão de curso são: i) Consolidação da base de dados de acidentes de trânsito; ii) Agregação espacial dos dados em configurações diferentes de áreas; iii) Execução de análises exploratórias (espaciais e não espaciais) dos dados agregados; iv) Comparação dos indicadores resultantes das análises. Através da análise não espacial, foi percebido que o grau de correlação entre as variáveis aumentava com a elevação do nível de agregação. Foi observado também que para a taxa de acidentes de trânsito, indicadores como a média também sofrem alterações quando há a mudança na configuração da unidade de agregação. Em relação a análise espacial, há um padrão de aumento da correlação espacial para a taxa de acidentes a medida que o nível de agregação também aumenta, com exceção para a configuração em zonas de tráfego, cujo valor diminui em relação ao nível de agregação anterior. Por fim, com a construção dos modelos de regressão que consideram os efeitos espaciais de forma global (SAR e SEM), foi percebido o aumento no poder de estimação desses com a elevação do nível de agregação.

Palavras-chave: Problema de Unidade de Área Modificável. Análise Exploratória Espacial. Análise Exploratória Não-Espacial. Modelos de Regressão Espacial

ABSTRACT

The observation and evaluation of the effects of the Modifiable Area Unit Problem become an important issue in the Road Safety Performance modeling, because if the appropriate levels of geographic scale and configuration of the zones are not identified, the statistical modeling in spatial data may lead to misleading conclusions. The general objective of this work is to evaluate the influence of the Modifiable Areal Unit Problem (MAUP) on the frequency of traffic accidents in areas. The methodological steps defined for the course completion work are: i) Consolidation of the database of traffic accidents; ii) Spatial aggregation of data in different areas; (iii) Exploratory analysis (spatial and non-spatial) of aggregate data; iv) Comparison of the indicators resulting from the analyzes. Through the non-spatial analysis, it was noticed that the degree of correlation between variables increased as the level of aggregation raise. It was also observed that for the traffic accident rate, indicators, such as average also suffer when there is a change in the configuration of the aggregation unit. Regarding to the spatial analysis, there is a pattern of increase of the spatial correlation for the accident rate the aggregation measure also increases, except for a configuration in traffic zones which value decreases if compared to the previous aggregation volume. Finally, with a construction of the regression models that consider the spatial products of the global form (SAR and SEM), it was noticed the increase in their power of estimation with the increase of the level of aggregation.

Keywords: Modifiable Areal Unit Problem. Exploratory Spatial Analysis. Non-Spatial Exploratory Analysis. Spatial Regression Models

LISTA DE FIGURAS

Figura 1 - Proporção na contribuição de ocorrência dos acidentes de trânsito	15
Figura 2 - Representação da distribuição de Poisson para os acidentes com ciclistas.	16
Figura 3 - Exemplo de gráfico de dispersão de Moran para os resíduos de um modelo de previsão de acidentes.	19
Figura 4 - LISA para resíduos de modelos de previsão de acidentes para o município de Fortaleza	21
Figura 5 - Agregação a nível de: (a) Setores Censitários; (b) block groups.....	25
Figura 6 - Variáveis para análise do efeito de zoneamento.	26
Figura 7 - Configurações de zoneamento aplicadas nas variáveis.	26
Figura 8 - Fluxograma das etapas metodológicas.	29
Figura 9 - Base de logradouros do município de Fortaleza.	31
Figura 10 - Exemplo de configurações de áreas: (a) Setores Censitários; (b) Zonas de Tráfego.	32
Figura 11 - Malha de unidades de áreas sobreposta aos Setores Censitários.	33
Figura 12 - Exemplo de seleção de setores a partir da malha de áreas.....	34
Figura 13 - Distribuição espacial dos acidentes georreferenciados.....	36
Figura 14 - Configurações finais de áreas: (a) Configuração 1 - Setores Censitários; (b) Configuração 2; (c) Configuração 3; (d) Configuração 4; (e) Configuração 5 - Zonas de Tráfego.....	38
Figura 15 - Mapa Coroplético do Total de Acidentes por Setores Censitários.....	40
Figura 16 - Mapa Coroplético do Total de Acidentes por zonas da Configuração 2.....	40
Figura 17 - Mapa Coroplético do Total de Acidentes por zonas da Configuração 3.....	41
Figura 18 - Mapa Coroplético do Total de Acidentes por zonas da Configuração 4.....	41
Figura 19 - Mapa Coroplético do Total de Acidentes por Zonas de Tráfego.	42

LISTA DE TABELAS

Tabela 1 - Estatística descritiva dos resultados para os dois níveis de agregação.	25
Tabela 2 - Estatística descritiva dos resultados para as variáveis 1 e 2.	26
Tabela 3 - Parâmetros das malhas de referência de construção das configurações de áreas.	33
Tabela 4 - Qualidade do georreferenciamento de acidentes da base viária “antes” e “depois” das atualizações.	36
Tabela 5 - Precisão do georreferenciamento antes e depois da atualização.	37
Tabela 6 - Total de unidades de áreas por configuração de agregação.	39
Tabela 7 - Matriz de Correlação entre as variáveis - Configuração 1.	43
Tabela 8 - Matriz de Correlação entre as variáveis - Configuração 2.	44
Tabela 9 - Matriz de Correlação entre as variáveis - Configuração 3.	45
Tabela 10 - Matriz de Correlação entre as variáveis - Configuração 4.	46
Tabela 11 - Matriz de Correlação entre as variáveis - Configuração 5.	47
Tabela 12 - Resultados da estatística descritiva das variáveis.	50
Tabela 13 - Índices de Moran Globais para Taxa de Acidentes por quilômetro de via.	52
Tabela 14 - Índices de Moran bivariado da variável taxa de acidentes por quilômetro de via.	53
Tabela 16 - Coeficientes e Parâmetros dos modelos de regressão espacial SAR.	55
Tabela 17 - Índice de Moran dos Resíduos dos Modelos de Regressão Espacial.	57

SUMÁRIO

1	INTRODUÇÃO	11
1.1	Problema e questões de pesquisa	12
1.2	Objetivos	13
<i>1.2.1</i>	<i>Objetivo Geral</i>	13
<i>1.2.2</i>	<i>Objetivos Específicos</i>	13
2	REVISÃO BIBLIOGRÁFICA	14
2.1	Modelagem do Desempenho da Segurança Viária	14
2.2	Análise Exploratória de Dados	17
<i>2.2.1</i>	<i>Análise exploratória espacial dos dados</i>	18
<i>2.2.1.1</i>	<i>Índice Global de Moran</i>	18
<i>2.2.1.2</i>	<i>Diagrama de dispersão de Moran e Local Indicator of Spatial Association (LISA)</i> 19	
<i>2.2.1.3</i>	<i>Modelos de regressão espacial</i>	21
2.3	Problema da Unidade de Área Modificável	23
3	METODOLOGIA	29
3.1	Consolidação da base de dados de acidentes de trânsito	29
3.2	Realização da agregação espacial dos dados em áreas	31
3.3	Execução de análises exploratórias (espaciais e não espaciais) dos dados agregados	34
3.4	Comparação dos indicadores resultantes das análises	35
4	RESULTADOS E DISCUSSÃO	36
4.1	Consolidação da base de dados	36
4.2	Realização da agregação espacial dos dados	37
4.3	Análises exploratórias	40
<i>4.3.1</i>	<i>Análise Não Espacial</i>	40
<i>4.3.2</i>	<i>Análise Espacial</i>	51
5	CONCLUSÕES	58
	REFERÊNCIAS	60
	APÊNDICE A – DIAGRAMAS DE DISPERSÃO DE MORAN	63
	APÊNDICE B – MAPA DE MORAN DOS RESÍDUOS DOS MODELOS DE REGRESSÃO ESPACIAL	75

1 INTRODUÇÃO

Segundo o relatório da Organização Mundial da Saúde (OMS, 2015), mais de um milhão de pessoas morrem nas estradas em todo o mundo e as lesões causadas por acidentes de trânsito são a nona principal causa de morte, podendo ser a sétima até o ano de 2030. Diante dessa situação, o desenvolvimento de estudos que permitam uma melhor compreensão dos fatores que influenciam a ocorrência dos acidentes de trânsito se faz necessário para assegurar um melhor direcionamento na adoção de medidas eficazes em reduzir o número e a severidade desses acontecimentos.

O desenvolvimento de modelos de previsão de acidentes pode permitir a avaliação do Desempenho da Segurança Viária (DSV) em cenários de longo prazo que envolvam mudança na população e alteração na infraestrutura e espaço construído, tornando-se um estudo relevante para o procedimento de planejamento de transportes.

Considerando que o fenômeno acidente de trânsito possui uma dimensão geográfica, ou seja, existe uma dependência espacial entre as observações e também ocorre uma heterogeneidade espacial entre essas, a utilização de modelos estatísticos para estimar o DSV que ignoram esses efeitos espaciais possui um poder de estimação menos eficiente e, conseqüentemente, ocorre uma diminuição da precisão na estimação desses parâmetros. Agüero-Valverde e Jovanis (2006) comentam que a consideração dos efeitos espaciais dos acidentes de trânsito nos modelos de DSV pode ser um substituto para variáveis desconhecidas e relevantes, melhorando assim a estimativa do modelo.

Para o desenvolvimento dessa modelagem citada anteriormente, a agregação dos dados sobre as características do tráfego, de configuração da rede viária e as informações socioeconômicas às informações de acidentes de trânsito se faz necessária. Comumente, esse processo é feito em nível de áreas que permitam maior facilidade de incorporação ao planejamento de transportes, essas zonas são conhecidas como Zonas de Tráfego (ZT). (ABDEL-ATY et al., 2013). Mas o grande problema da utilização dessas áreas está em torno da sua desconsideração da heterogeneidade espacial dos acidentes de trânsito para a sua delimitação (HUANG et al., 2013).

A utilização de diferentes definições espaciais de áreas pode levar a uma inconsistência nos resultados, como coeficientes de correlação e de regressão, além da

possibilidade de gerar uma dependência espacial artificial (XU et al., 2014). Essa questão é conhecida como o Problema da Unidade de Área Modificável (PUAM).

A ideia principal do PUAM, ou como originalmente o problema foi formulado, *Modifiable Areal Unit Problem* (MAUP), consiste em que as múltiplas formas que uma mesma área pode ser delimitada ou criada irão resultar em diferentes resultados e, conseqüentemente, diferentes análises (WONG, 2004). Desta forma, a observação e avaliação dos efeitos do PUAM tornam-se uma questão importante na modelagem do DSV, pois caso não sejam identificados os adequados níveis de escala geográfica e de configuração de zonas a modelagem estatística baseada em dados espaciais, isso pode induzir a conclusões enganosas.

1.1 Problema e questões de pesquisa

Baseado na contextualização feita acima, o problema de pesquisa pode ser definido como a influência do PUAM na geração dos modelos de previsão de acidentes e as inconsistências acarretadas em sua análise, tanto na escolha de diferentes configurações de zonas quanto na utilização de diferentes escalas geográficas.

A partir da definição do problema de pesquisa, podem ser definidas a seguintes questões de pesquisa que fundamentarão os objetivos, geral e específicos, desse estudo:

- a) Quais são as principais diferenças entre as análises exploratórias espaciais e não-espaciais?
- b) Quais são os principais aspectos da espacialidade dos acidentes de trânsito?
- c) Como realizar a consolidação de uma base de dados com variáveis socioeconômicas e de acidentes de trânsito em diferentes agregações espaciais?
- d) Qual a influência de diferentes níveis de agregação espacial na relação entre acidentes de trânsito e variáveis associadas a sua ocorrência?

1.2 Objetivos

1.2.1 Objetivo Geral

O objetivo geral desse trabalho é avaliar a influência de diferentes níveis de agregações na estatística de acidentes de trânsito agregados em áreas.

1.2.2 Objetivos Específicos

A partir da definição do objetivo geral da pesquisa, podem ser definidos os seguintes objetivos específicos:

- a) Caracterizar as principais técnicas de análises exploratórias espaciais e não espaciais;
- b) Identificar os principais atributos estocásticos espaciais e não espaciais que podem ser associados à ocorrência dos acidentes de trânsito em áreas urbanas;
- c) Consolidar uma base de dados com variáveis socioeconômicas e de acidentes de trânsito em diferentes agregações espaciais;
- d) Efetivar um estudo de caso na área urbana da cidade de Fortaleza para avaliar o efeito de diferentes níveis de agregação nos acidentes de trânsito e em variáveis independentes que possam estar correlacionadas com a taxa de acidentes de trânsito.

2 REVISÃO BIBLIOGRÁFICA

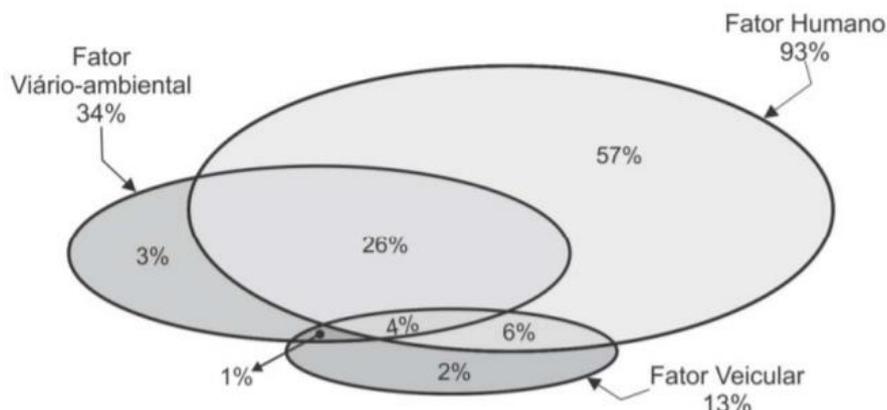
Ao se abordar a questão do PUAM para os acidentes de trânsito faz-se necessário um melhor entendimento do fenômeno e também do próprio problema o qual baseia esse trabalho, além de saber as devidas ferramentas e técnicas utilizadas para sua avaliação, tornando claro os conceitos que serão abordados. Logo, os objetivos dessa seção serão comentar as principais características dos acidentes de trânsito e fatores relacionados a sua ocorrência, conceituar as importantes técnicas de análise exploratória espaciais e não-espaciais e os conceitos as quais estão interligadas e investigar estudos e abordagens aplicadas ao PUAM.

2.1 Modelagem do Desempenho da Segurança Viária

No âmbito de modelagem da segurança viária, é indispensável a abordagem de características importantes relacionadas aos acidentes de trânsito, pois elas servirão de fundamentação no processo de definição dos modelos que, por sua vez, influenciam diretamente nos resultados obtidos, principalmente as características espaciais, que, para esse fenômeno, apresentam-se de forma expressiva.

Segundo ABNT (1989), um acidente de trânsito é definido como um evento não antevisto que sucede em algum dano material (veículo, carga e etc.) ou pessoal (lesões ou óbitos) sendo que uma das partes deve estar em movimento nas vias terrestres ou em áreas abertas ao público. Ainda segundo a norma, os elementos que contribuem para a ocorrência de um acidente são classificados individualmente em quatro grupos: Fator Humano, associado ao comportamento humano como o contribuinte para a ocorrência de acidentes; Fator Via, atrelado às condições da via e de sua sinalização; Fator Veículo, correspondente às falhas mecânicas do veículo que prejudicam a segurança; o Fator Meio Ambiente, associados à características da natureza e do meio ambiente que afetam a segurança.

Figura 1 - Proporção na contribuição de ocorrência dos acidentes de trânsito



Fonte: Silva (2006).

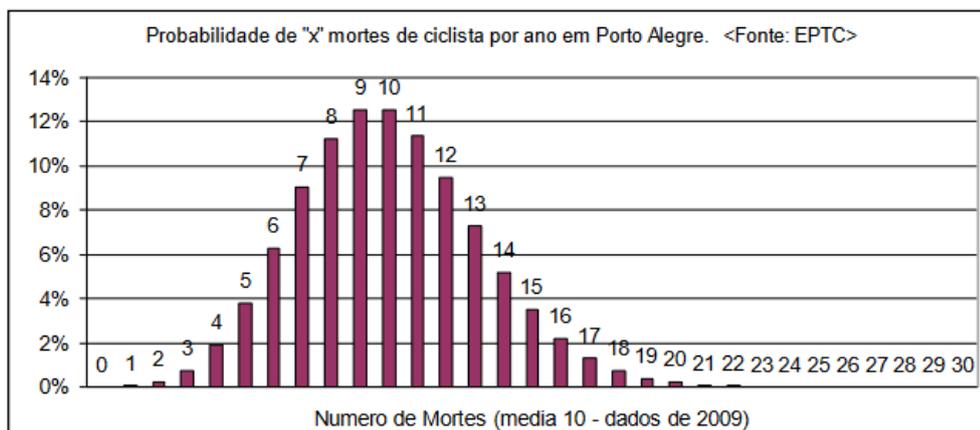
Comumente, para a construção de Modelos de Previsão de Acidentes, utilizam-se variáveis relacionadas às condições da via e também variáveis relacionadas ao tráfego, onde o critério de escolha baseado na disponibilidade de informações é determinante. Como foi mostrado por Cunto et al. (2012).

Os mesmos autores destacam também a dificuldade na disponibilidade, qualidade e precisão das informações provenientes dos bancos de dados de acidentes, associados a continuidade de alimentação desses dados, influenciaram no início tardio do processo de modelagem aqui no Brasil comparados a países como Canadá e Estados Unidos, por exemplo.

Estudos realizados por Ceder e Livneh (1982), mostram a utilização da variável de exposição volume do tráfego em modelos de previsão de regressão, mas Jovanis e Chang (1986) mostraram que certas propriedades da ocorrência de acidentes são incompatíveis com as premissas da regressão linear tradicional, como por exemplo, a não-normalidade da distribuição dos resíduos.

Segundo Cardoso e Goldner (2006), pelo fato de que os acidentes são independentes, imprevisíveis e que a quantidade de ocorrências é de certa forma constante em um período, é que a distribuição de Poisson é empregada para explicar o número de acidentes em uma determinada extensão de tempo. A figura abaixo mostra um exemplo para os acidentes com ciclistas para a cidade de Porto Alegre.

Figura 2 - Representação da distribuição de Poisson para os acidentes com ciclistas.



Fonte: Rêgo e Pena (2012)

Mountain e Jarret (1996) aplicaram essa distribuição em um estudo abrangendo 7 distritos do Reino Unido, aproximadamente quatro quilômetros de rodovia, obtendo valores de significância dos modelos de regressão menores do que 10%.

Sawalha e Sayed (2011) estudaram 58 vias arteriais de duas cidades no Canadá, também com a utilização da distribuição de Poisson, sendo que essa escolha se deu a partir de uma metodologia adicional, calculando um parâmetro de dispersão, o qual indica a aplicação de uma distribuição Binomial ou de Poisson para os resíduos.

No Brasil, Cardoso e Goldner (2006), realizaram uma análise espacial dos acidentes dos anos de 1998 a 2004, para a cidade de Porto Alegre, Rio Grande do Sul, baseando sua opção na distribuição de Poisson nos estudos já realizados.

Cunto et al. (2012) também calcularam o parâmetro de dispersão para todas as categorias de acidentes que foram definidas para a construção dos modelos, cujos resultados fizeram os autores optarem pela utilização da distribuição Binomial Negativa, consistindo em uma possibilidade para se tratar o padrão disperso dos dados utilizados.

De fato, a consideração dessas variáveis citadas anteriormente possui sua importância na estimação da taxa de acidentes de trânsito de uma determinada região, mas se faz necessária uma maior sofisticação para a utilização desses modelos, pois quando não são levadas em consideração características como a heterogeneidade e dependência espacial desse fenômeno, esse poder de estimação é comprometido (AGUERO-VALVERDE e JOVANIS, 2006).

Segundo Gomes (2017) quando os efeitos espaciais são pertinentes (dependência e heterogeneidade espacial) eles acabam por contrariar duas premissas básicas dos modelos tradicionais de regressão, que são a independência entre as observações e a variância constante dos erros das observações (homocedasticidade).

A correlação espacial ou também chamada de dependência espacial, é a relação que o valor de uma variável para uma determinada unidade de área, depende do valor dessa mesma variável das regiões vizinhas, ou seja, a correlação dos valores ao longo do espaço. (ANSELIN, 1993).

Ainda segundo Gomes (2017) a heterogeneidade ocorre quando há uma não estacionariedade nas relações entre as variáveis explicativas e a variável dependente ao longo das diversas unidades que subdividem uma região de estudo podendo levar a uma modelagem com previsões menos eficientes, de modo que, ambas as situações (dependência e heterogeneidade) não podem ser resolvidas através da estatística clássica. A complicação ocorre quando ambas ocorrem juntas, tornando o problema ainda mais complexo necessitando de uma abordagem espacial específica.

:

2.2 Análise Exploratória de Dados

Uma Análise Exploratória de Dados (AED), tem como objetivo extrair informações a partir de técnicas estatísticas e gráficas, de modo a se ter um entendimento básico dos dados, avaliando a relação entre as variáveis e também o padrão de comportamento das mesmas.

A análise exploratória não-espacial é a mais comumente utilizada, com ela são observados parâmetros da estatística clássica, como média, mediana, desvio padrão, máximos e mínimos, coeficiente de variação e dentre outros. A análise feita a partir desses parâmetros supracitados podem ser enganosas muitas vezes, para caso de uma avaliação descritiva dos acidentes de trânsito, visto que a variável analisada se trata de um evento geograficamente dependente. Por isso que as ferramentas da análise exploratória espacial se aplicam melhor para a obtenção de informações para o fenômeno acidente de trânsito (ERDOGAN et al., 2008).

2.2.1 Análise exploratória espacial dos dados

Segundo Anselin e Bera (1998) *apud*. Sobreira (2016) a análise exploratória espacial tem como objetivo a verificação de padrões e comportamentos anômalos, assim como a análise exploratória não espacial, mas focada justamente na visualização ao longo do espaço das variáveis e a sua interdependência (correlação espacial).

Esses parâmetros de observação podem ser classificados em globais e locais. Os parâmetros globais não são subdivididos, consistindo em um único valor para toda a região de estudo, enquanto que os locais possuem o objetivo de mostrar as dependências espaciais para cada sub-região de interesse (GOMES, 2017).

Dentre os indicadores globais mais utilizados, têm-se o Índice Global de Moran, Diagrama de Dispersão de Moran, e para representação de parâmetros locais têm-se o Índice Local de Moran ou como é mais comumente denominado *LISA (Local Indicator of Spatial Association)*.

2.2.1.1 Índice Global de Moran

Como citado anteriormente, o índice global de moran é um parâmetro que indica a medida global de autocorrelação espacial global de uma variável, o mesmo é representado pela equação 2.1:

$$I = \left[\frac{n}{\sum_{i=1}^n (y_i - \bar{y})^2} \right] \cdot \left[\frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y}) \cdot (y_j - \bar{y})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \right] \quad (2.1)$$

Onde n é o número de regiões, y_i é o valor da variável considerada na área i , y_j é o valor da variável do vizinho a área i , \bar{y} é o valor médio da variável na região de estudo e w_{ij} são os elementos da matriz de ponderação espacial normalizada.

Esse índice possui valores variando entre -1 e 1, tal que a resultando em valor nulo, indica uma inexistência de correlação espacial, de tal forma que para valores positivos significa uma maior semelhança entre zonas vizinhas, podendo ser valores superiores ou inferiores a média, e para valores negativos menor a semelhança e maior a dispersão dos dados.

No estudo realizado por Queiroz (2003) o autor observou com a utilização desse parâmetro, uma dispersão maior entre os acidentes ocorridos no ano de 2002 para a cidade de

Fortaleza, Ceará em comparação com os dados do ano anterior (2001), concluindo que esse aumento foi devido a medidas mais incisivas adotadas pela gestão do município em áreas mais específicas em detrimento de outras regiões que necessitavam de intervenção.

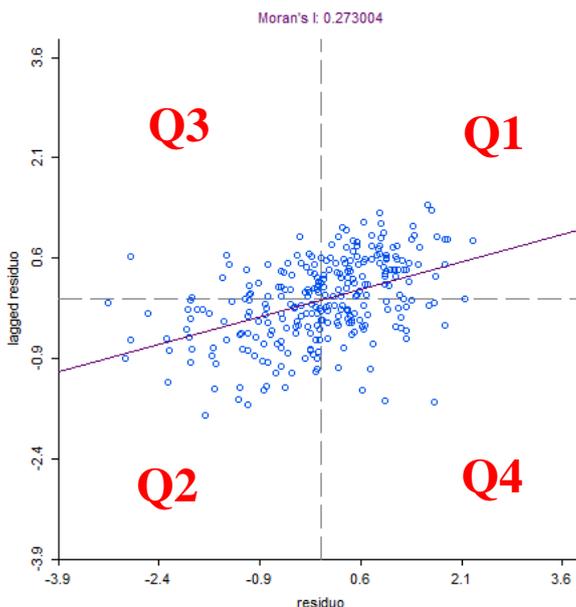
Um caso semelhante foi relatado por Santos e Raia Júnior (2006), onde os mesmos através do índice de Moran conseguiram identificar uma maior dispersão para os acidentes de 2003 na cidade de São Carlos, São Paulo, apontando também que essa situação dificulta o estudo e a prevenção desses acidentes.

Siddiqui et al. (2012) propondo uma análise do efeito da correlação espacial para acidentes com pedestres e ciclistas de 5 cidades do estado da Flórida, Estados Unidos da América, observou esses efeitos no âmbito macroscópico através do índice global de Moran para os modelos de previsão. Os seus resultados mostraram que tanto para o número total de acidentes com pedestres e o total de acidentes com ciclistas, apresentou correlação espacial semelhante entre si.

2.2.1.2 Diagrama de dispersão de Moran e Local Indicator of Spatial Association (LISA)

A grande vantagem do diagrama de dispersão de Moran está na identificação de valores anômalos que fogem do padrão de comportamento apresentado pelos outros valores (GOMES, 2017). Esses *outliers* podem ser resultado da má especificação da matriz de adjacência ou com a escala espacial de observação das variáveis (QUEIROZ, 2003).

Figura 3 - Exemplo de gráfico de dispersão de Moran para os resíduos de um modelo de previsão de acidentes.



Fonte: Elaborada pelo Autor (2015)

Segundo Anselin (1992) *apud*. Sobreira (2016) a construção do diagrama se dá a partir da comparação dos valores de uma variável em uma determinada zona com a média normalizada de seus vizinhos, resultando pois na divisão em quatro setores, sendo que o primeiro e o segundo quadrantes representam uma correlação positiva (valores positivos na zona e valores positivos para a média da vizinhança ou valores negativos na zona e valores negativos da média da vizinhança), e o restante representam uma correlação negativa (sinais contrários para os valores da variável na zona e os valores da média da vizinhança).

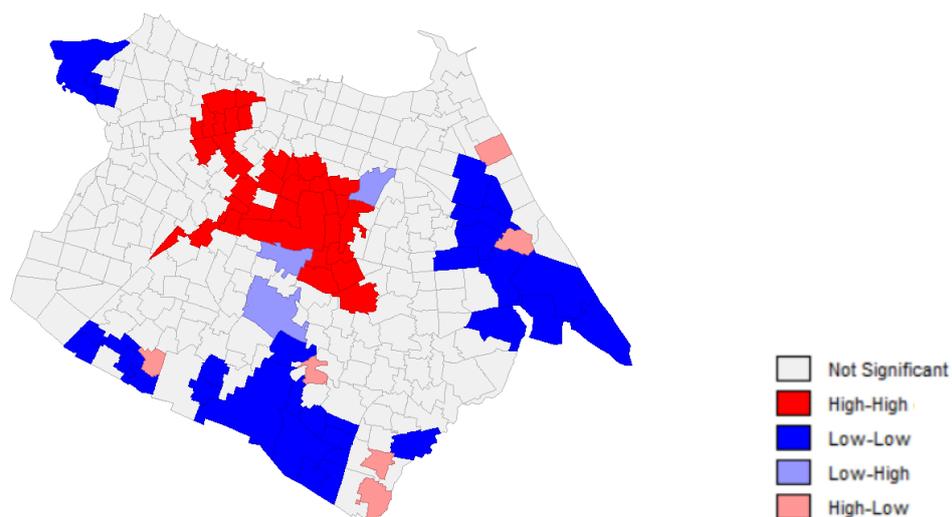
Ainda segundo Gomes (2017), a utilização de indicadores locais podem revelar padrões locais de associação, diferentemente quando se faz uma abordagem global. Por isso que além da observação do valor global do índice de Moran, é importante o cálculo do valor individual de cada zona, sendo destacadas aquelas em que apresentam correlação local significantes, essas áreas são conhecidas como “bolsões” de não-estacionariedade e possuem uma dinâmica espacial própria.

O cálculo do índice local de Moran é feito de acordo com a equação 2.2.

$$I_i = \frac{z_i \sum_{j=1}^n w_{ij} z_j}{\sum_{j=1}^n z_j^2} \quad (2.2)$$

A figura 2 mostra um exemplo de um mapa gerado a partir dos índices locais mais significantes, para as observações de um resíduo de um modelo de previsão de acidentes para o município de Fortaleza.

Figura 4 - LISA para resíduos de modelos de previsão de acidentes para o município de Fortaleza



Fonte: Elaborada pelo Autor (2015)

Os valores representados como “*High-High*” e “*Low-Low*” são aqueles com correlação positiva representados nos quadrantes 1 e 2 do diagrama de dispersão de Moran, respectivamente os classificados como “*Low-High*” e “*High-Low*” são os que apresentaram correlação negativa, alocados nos quadrantes 3 e 4 do diagrama de dispersão de Moran respectivamente.

A análise feita por Carvalho (2011) pôde-se demonstrar de forma visual que os municípios que apresentaram correlação positiva da mortalidade por acidentes de motocicletas no estado de Pernambuco se tratam de regiões em desenvolvimento do agreste e do sertão e as regiões de correlação negativa estão mais próximas da região metropolitana da capital do estado.

Através do mapa de espalhamento de Moran, Sobreira (2016) mostrou que, para os atropelamentos ocorridos nos anos de 2009 a 2011, bairros localizados na região sudeste da cidade de Fortaleza e que apresentavam na época baixa concentração de domicílios a aglomeração foi baixa com uma correlação positiva.

2.2.1.3 Modelos de regressão espacial

Partindo do pressuposto de que o acidente de trânsito é um evento dependente espacialmente, a utilização de modelos de regressão convencionais desconsidera a dependência espacial entre as observações, necessitando de uma melhora da modelagem no quesito em acrescentar essa espacialidade nos Modelos de Previsão de Acidentes (MPA). Os modelos de regressão espaciais se diferem dos modelos de regressão convencionais pelo fato da incorporação da correlação espacial entre as áreas, tanto que essa dependência altera o poder explicativo do modelo (RÊGO E PENA, 2012).

Um dos papéis importantes da modelagem espacial é indicar quais variáveis explicativas, a partir de seus coeficientes, influenciam, mais significativamente, nas observações previstas para a variável dependente, sendo uma importante ferramenta para o planejamento a longo prazo para os gestores do tráfego (CÂMARA et al., 2002).

Ainda segundo o mesmo autor, ao se fazer uma análise de regressão espacial a hipótese padrão é que as observações não são correlacionadas e que os resíduos do modelo também são independentes e não-correlacionados com a variável dependente , têm variância constante, e apresentam distribuição normal com média zero.

A consideração dos efeitos espaciais através dos modelos de regressão pode ser feita de forma mais simples, como fazem os modelos com efeitos espaciais globais, resumindo a correlação espacial em um único parâmetro na equação, ou de forma mais complexa, como fazem os modelos com efeitos espaciais locais, os quais podem variar os parâmetros ao longo do espaço ou modelar a variação espacial de forma discreta com várias regressões para cada uma das várias regiões.

Para efeito de simplificação, será tratado apenas os modelos espaciais com efeitos espaciais globais, mais especificamente os modelos Espacial Autorregressivo Misto (SAR) e o modelo do Erro Espacial (SEM).

Segundo Rêgo e Pena (2012), o modelo SAR considera a dependência espacial através do acréscimo de um termo na regressão, de acordo com a equação X:

$$Y = \rho WY + X\beta + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I) \quad (2.3)$$

Onde:

Y é a variável dependente;

ρ parâmetro responsável pela estimação do grau de dependência espacial da variável a ser explicada com os seus vizinhos;

X representa as variáveis explicativas;

β representa os coeficientes de regressão;

ϵ é o erro aleatório;
 σ^2 é a variância do modelo;
 I é a matriz identidade.

Já o modelo do tipo SEM, se difere na questão da consideração da matriz de vizinhança como sendo parte do erro aleatório, obedecendo as seguintes equações:

$$y = X\beta + u \quad (2.4)$$

$$u = \lambda Wu + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I) \quad (2.5)$$

Onde:

u é o erro aleatório;
 λ é um parâmetro espacial.

Lima e Cunto (2015), realizaram uma regressão espacial para acidentes, tanto em números totais, quanto para acidentes com vítimas, utilizando os dois tipos de modelos espaciais com efeitos globais apresentados acima, para zonas da região de Fortaleza. Seus resultados mostraram que, para ambas as variáveis e ambos os modelos, a espacialidade não foi eliminada, diferentemente do que foi apresentado na modelagem utilizando os modelos com efeitos espaciais locais, apresentando-se para valores do índice de Moran de -0.003 e -0.006.

Apesar de sua abordagem simples, os modelos SAR e SEM podem mostrar de forma clara a influência da mudança na configuração espacial através da mudança dos coeficientes de regressão, atendendo aos objetivos apresentados para esse trabalho.

2.3 Problema da Unidade de Área Modificável

Segundo Wong (2004), quando se trabalha com modelagem espacial têm-se uma tentativa de se representar um meio contínuo, que é a superfície terrestre através de elementos de áreas com suas fronteiras bem definidas. A definição dessas áreas se baseia em agregar regiões com características semelhantes para um fenômeno no qual se quer analisar, o Problema da Unidade de Área Modificável (PUAM) ou do original em inglês, *Modifiable Areal Unit Problem (MAUP)*, se resume nas inconsistências estatísticas que são geradas a

partir das diferentes formas de agregação espacial a qual está submetida uma determinada observação.

O PUAM se trata de um problema comum quando se trabalha com análises estatísticas espaciais, não só para o planejamento de transportes como também para as áreas relacionadas à saúde pública, economia e também agricultura. Gehlke e Biehl (1934), correlacionaram os índices de delinquência juvenil masculina com o aluguel equivalente mensal por cada setor censitário da cidade de Cleveland, no estado de Ohio nos Estados Unidos. A partir do momento em que essas unidades de áreas eram agregadas formando outras maiores os coeficientes de correlação entre as variáveis mudavam de forma que os menores valores de correlação eram observados para unidades de áreas menores.

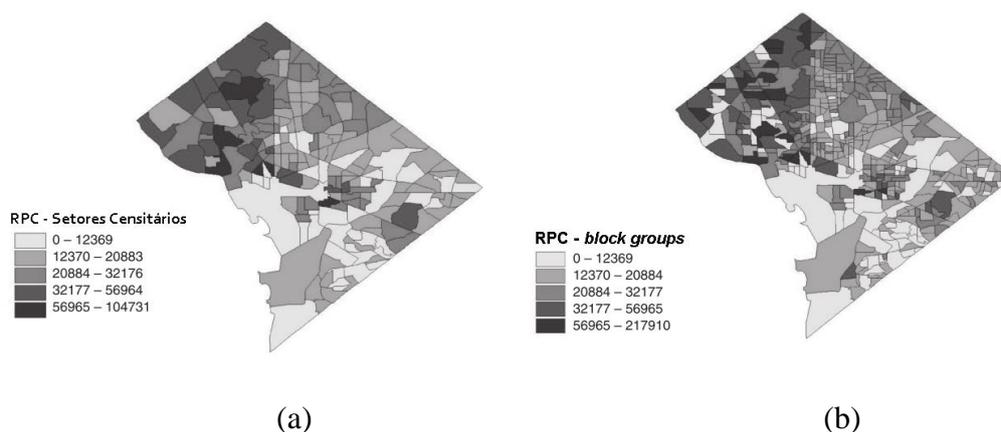
Fotheringham e Wong (1991), já no final do século XX, retrataram de forma mais extensa o efeito desse problema para outras variáveis socioeconômicas, avaliando-o através de uma análise sensitiva com diversas configurações espaciais obtidas através da agregação de *block groups*, que são unidades censitárias que representam uma divisão estatística dos setores censitários. Seus resultados mostraram, por exemplo, que um aumento de 0.1 na proporção de idosos para uma configuração espacial de 25 zonas proporciona um decréscimo em torno de quatro vezes maior caso utilizasse uma configuração espacial de 800 zonas.

Wang et al. (2011), fez a análise desse problema através da geração de números aleatórios assumindo que esses valores são observações de uma variável qualquer chamada de “x” então, adotando que os mesmos obedecem uma distribuição Poisson, realizou uma modelagem utilizando o valor de “x” como variável independente para a estimação de outro termo chamado de “xp”. As observações foram agregadas em 300 grupos aleatórios e depois em 80 grupos que correspondem a combinação da primeira forma de configuração utilizada. Seus resultados mostraram que o coeficiente da variável independente do modelo de regressão cresce com o crescimento do número de zonas utilizadas.

Os efeitos do PUAM podem ser observados a partir de outros dois problemas. O primeiro se refere ao que pode ser chamado de efeito de escala e o segundo é o chamado efeito de zoneamento.

Para ilustrar o efeito de escala, será utilizado o exemplo demonstrado por Wong, (2004) onde foi utilizada a variável renda per capita (RPC) para negros na cidade de Washington, D.C utilizando-se dois níveis de agregação: a nível de setores censitários (Figura 3a) e *block groups* (Figura 3b).

Figura 5 - Agregação a nível de: (a) Setores Censitários; (b) *block groups*



Fonte: Adaptado de Wong, (2004)

Feita a agregação, o autor realizou uma análise exploratória e obteve os seguintes parâmetros representados na Tabela 1.

Tabela 1 - Estatística descritiva dos resultados para os dois níveis de agregação.

Renda Per Capita para negros	Média	Mínimo	Máximo	Desvio Padrão
Setores Censitários	21879	0	104731	15053
Block Groups	23390	0	217910	20073

Fonte: Wong, (2004).

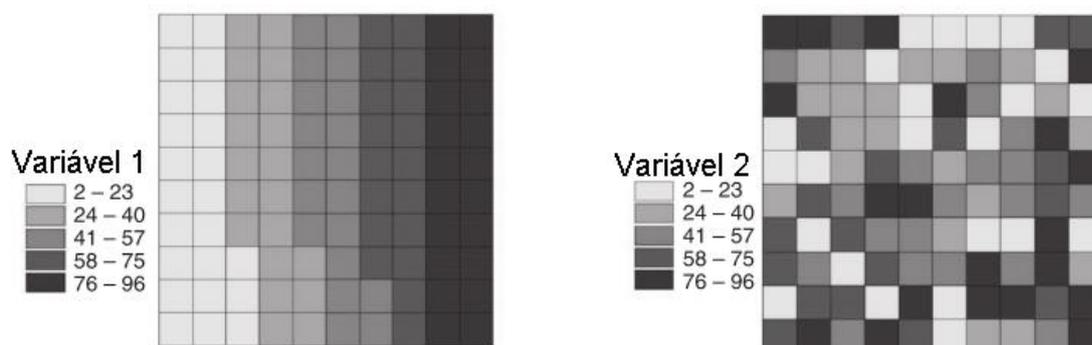
A utilização de unidades maiores, como no caso dos setores censitários, ocasionou uma homogeneização dos dados podendo mascarar de certa forma valores extremos das unidades menores e assim houve uma suavização dos valores observados.

Em um estudo do âmbito do planejamento de transportes, Thomas (1996), estudou o efeito da escala através da influência do tamanho do segmento da rodovia na avaliação estatística no número e densidade de acidentes que ocorreram no ano de 1990 em 65 subseções de rodovias localizadas na Bélgica. Seus resultados mostraram que com o aumento do comprimento de segmento utilizado, a média e a variância do número de acidentes também aumentava, diferentemente da densidade de acidentes, cuja variância diminuía com o aumento do segmento e a média não era afetada.

Para a ilustração do efeito de zoneamento, será utilizado também outro exemplo de Wong (2004), onde foram analisadas duas variáveis as quais foram distribuídas de formas

diferentes ao longo do espaço, sendo a variável 1 localizada de forma mais uniforme, e a variável 2 de forma mais desordenada (Figura 4).

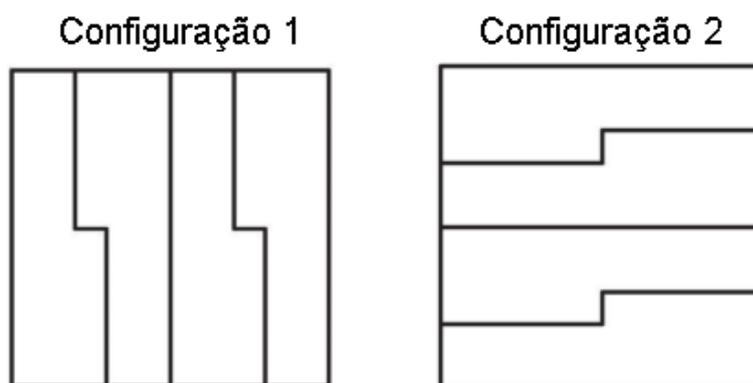
Figura 6 - Variáveis para análise do efeito de zoneamento.



Fonte: Adaptado de Wong, (2004).

Ambas foram submetidas a duas configurações diferentes de zoneamento (Figura 5) e assim foi realizado a análise exploratória. Os resultados obtidos estão na Tabela 2, de tal modo que V1 – C1 representa os resultados para a variável 1 na configuração de zoneamento 1; V1- C2 representa os resultados para a variável 1 na configuração de zoneamento 2 e assim por diante.

Figura 7 - Configurações de zoneamento aplicadas nas variáveis.



Fonte: Adaptado de Wong, (2004).

Tabela 2 - Estatística descritiva dos resultados para as variáveis 1 e 2.

Variáveis (V1 e V2) e Configurações (C1 e C2)	Média	Mínimo	Máximo	Desvio Padrão
V1 e V2	49.00	2.00	96.00	27.00
V1-C1	48.60	15.00	84.00	29.70
V1-C2	48.60	42.80	54.76	4.98
V2-C1	48.60	40.68	53.60	5.55
V2-C2	48.60	43.24	59.60	7.45

Fonte: Adaptado de Wong, (2004)

Os resultados da tabela acima mostram os valores obtidos considerando os valores médios como representativos de cada zona e a primeira linha representa os resultados originais das duas variáveis. Percebe-se que a configuração V1-C1 obteve resultados semelhantes, se tratando do valor máximo e do desvio padrão, esse comportamento é devido ao fato de que a configuração 1 de zoneamento foi propositadamente elaborada para seguir o padrão de distribuição da variável 1. Em contrapartida, a variável 2 não conseguiu apresentar resultados condizentes por conta de que ambas as situações não contemplam seu padrão totalmente aleatório de distribuição ao longo do espaço.

Xu et al. (2014), com um estudo mais direcionado para a área de engenharia e planejamento de transportes, utilizaram dados reais do total de acidentes dos anos de 2005 a 2007 do condado de Hillsborough, localizado no estado da Flórida, Estados Unidos da América. Nesse estudo, os pesquisadores investigaram quantitativamente os impactos da configuração espacial com o método REDCAP (*Regionalization with dynamically constrained agglomerative clustering and partitioning*) cujo objetivo é construir regiões homogêneas agregando áreas com atributos de valores semelhantes. Utilizando um total de 738 Zonas de Tráfego e um total de 14 esquemas de zonas gerados a partir do método REDCAP, os autores utilizaram de modelos de previsão de acidentes cujas variáveis explicativas se tratavam de variáveis socioeconômicas e também relacionadas com o uso do solo. Para uma configuração de 50 zonas, o coeficiente para a variável explicativa densidade populacional foi de 0,53, já para a configuração de 738 zonas, o valor do coeficiente foi de 0,14, isso para o modelo cuja variável dependente era frequência de acidentes severos. Ainda segundo o mesmo estudo, a estimação dos parâmetros que compõem o grupo de variáveis explicativas é mais sensível à variação quando o número de zonas é baixo.

Davis (2004) e Xu et al. (2016) colocaram como possível solução para esse problema, a utilização de dados desagregados, confrontando com a atual realidade na qual

devido a natureza dos dados utilizados para a previsão de acidentes e também devido a razões legais, a utilização dos dados nesse formato não é possível, cabendo a avaliação do PUAM a uma análise sensitiva que obtenha resultados em vários níveis de agregação de dados e com diferentes configurações de áreas.

3 METODOLOGIA

As etapas metodológicas definidas para este trabalho de conclusão de curso são: i) Consolidação da base de dados de acidentes de trânsito; ii) Agregação espacial dos dados em configurações diferentes de áreas; iii) Execução de análises exploratórias (espaciais e não espaciais) dos dados agregados; iv) Comparação dos indicadores resultantes das análises. A figura 8 retrata a sequência metodológica proposta.

Figura 8 - Fluxograma das etapas metodológicas.



Fonte: Elaborada pelo autor (2017)

3.1 Consolidação da base de dados de acidentes de trânsito

Nessa etapa em específico, foi construída uma base de dados que contém informações sobre aspectos socioeconômicos, dados da infraestrutura urbana, aspectos do tráfego urbano e também dados sobre os acidentes de trânsito, a fim de se encontrar as variáveis que possuem maior correlação com a ocorrência dos acidentes.

Primeiramente, foi escolhida uma base geográfica com a representação da rede viária do município de Fortaleza, a base também contém informações sobre a numeração das edificações e também informações identificadoras dos logradouros a fim de ser realizado o georreferenciamento dos acidentes de trânsito. A figura 9 mostra um exemplo de base de geográfica que será utilizada.

A localização de um determinado acidente de trânsito é feita a partir do seu endereço de correspondência, determinando o logradouro e a numeração conforme sua posição sendo necessária a realização do processo de geocodificação, transformando esses endereços em coordenadas geográficas por um processo de interpolação a partir dos dados espaciais de primeira e última numeração do lado esquerdo e direito de cada segmento de via (*links*). Já para os acidentes em interseções, esse processo é feito apenas identificando-se o nó correspondente ao local na base de dados.

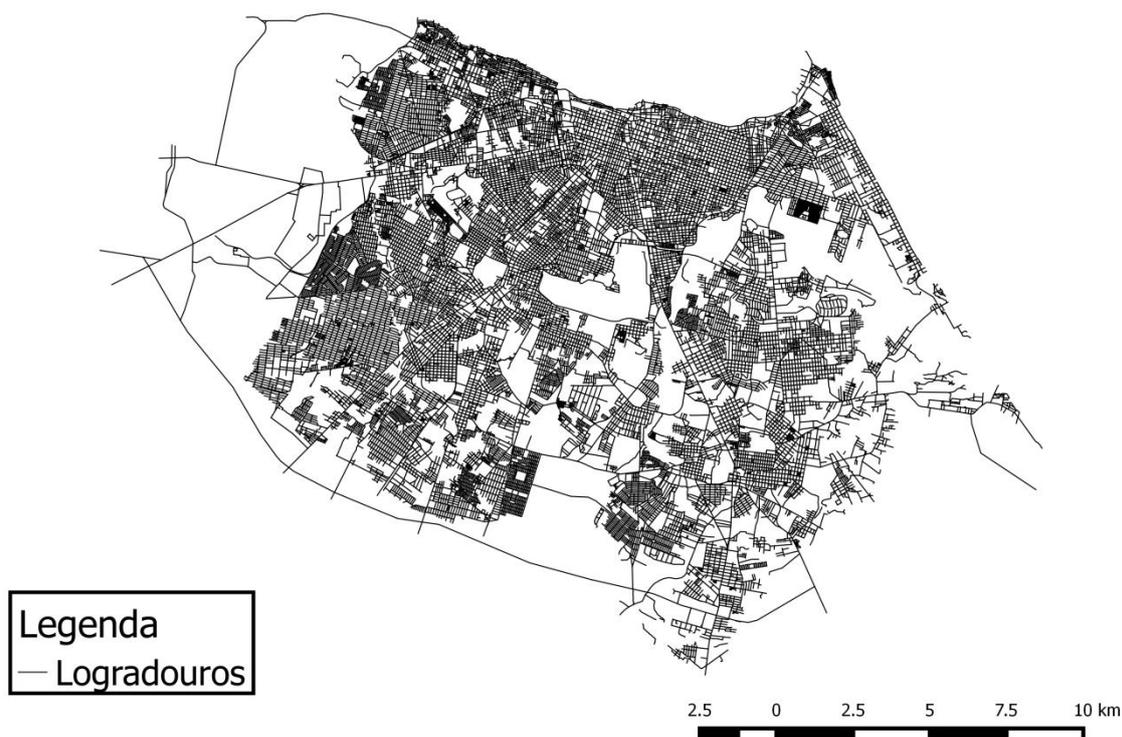
Devido a certos problemas como: ausência ou incoerência de faixas de numeração, erros gramaticais, mudança de nome do logradouro, logradouros homônimos, erros na coleta de dados e numerações de edificações que não obedecem ao padrão seguido em um determinado logradouro, a geocodificação pode ser comprometida, necessitando-se de um processo contínuo de atualização dos dados existentes.

A base de dados utilizada, foi consolidada e atualizada por Lima et al. (2015), obedecendo as seguintes etapas: 1) Escolha da base geográfica de logradouros inicial, escolhendo-se aquela que possuía maior número de informações a ser utilizada para o processo de geocodificação, no caso, a base da Autarquia Municipal de Trânsito e Cidadania (AMC); 2) Verificou-se e corrigiu-se a conectividade e a topologia dos *links* existentes, eliminando os pseudo nós e associando a topologia ao sentido de crescimento da numeração da via, contendo números pares à direita e números ímpares à esquerda; 3) Atualização da numeração e inserção de atributos físicos da via como número de pistas e faixas, sentido de circulação utilizando-se a ferramenta *Street View* do *Google*; 4) Realização de testes de qualidade do processo de atualização.

Os dados relacionados aos acidentes foram disponibilizados pelo Sistema Informações de Acidentes de Trânsito do Município de Fortaleza (SIAT/FOR) e foram utilizadas as informações mais recentes que já estão consolidadas. No caso, correspondem aquelas dos anos de 2009, 2010 e 2011.

As características do tráfego urbano como número de interseções semaforizadas, número de aparelhos de fiscalização eletrônica foram coletadas a partir de banco de dados também disponibilizados pela AMC. Os dados socioeconômicos e de uso solo serão utilizados aqueles disponibilizados pela Secretaria de Finanças (SEFIN) com base no Censo de 2010.

Figura 9 - Base de logradouros do município de Fortaleza.



Fonte: Elaborada pelo autor (2017).

3.2 Realização da agregação espacial dos dados em áreas

Os diversos dados citados anteriormente estão disponibilizados em diferentes formas de agregação, essa etapa tem como intuito compatibilizar as informações, nesse trabalho em específico essa compatibilização será feita no nível de áreas.

Como o objetivo geral está na análise do PUAM, se faz necessária a utilização de diferentes configurações de áreas para a agregação dos dados, para isso serão utilizadas dois tipos de configurações: As Zonas de Tráfego (ZT) (Figura 8b) e os Setores Censitários (SC) (Figura 8a), compostos, respectivamente, de 126 e 3043 unidades, sendo as ZT's uma forma mais agregada de configuração dos SC's.

Figura 10 - Exemplo de configurações de áreas: (a) Setores Censitários; (b) Zonas de Tráfego.



Fonte: Elaborada pelo autor (2017).

Os erros existentes em se utilizar bases de dados de áreas são relacionados a não correspondência dos limites das unidades de áreas (que, geralmente, são as ruas da rede viária local) com a base de dados viária utilizada. Devido a isso, as unidades de áreas devem ser reconfiguradas de forma a coincidir os seus limites de acordo com o eixo das vias. Para corrigir esse problema, com o auxílio do Sistema de Informações Geográficas (SIG) *Quantum GIS* (QGIS), versão 2.10 foi realizada essa correção manual.

Serão utilizadas um total de cinco níveis de configurações de áreas. O primeiro nível será a configuração em ZT (126 áreas) e o último nível, mais desagregado possível, será a configuração em SC (3043 áreas), sendo os níveis intermediários construídos com base no número de unidades de áreas que desejasse analisar, compreendendo valores intermediários às unidades de Setores censitários e de zonas de tráfego.

Dessa forma, foram escolhidos, além das 126 zonas de tráfego e dos 3043, configurações que compreendessem aproximadamente 2000, 1000 e 500 unidades de áreas. Para isso tomou-se a área total da cidade (316 km²) e dividiu-se pelo número de agregações obtendo-se o valor desejado de área para cada unidade pertencente a uma determinada configuração com um número desejado de unidades conforme a tabela abaixo:

Tabela 3 - Parâmetros das malhas de referência de construção das configurações de áreas.

Área de Fortaleza (km ²)	Número de Agregações	Área de uma unidade (km ²)	Dimensão do lado de uma unidade de área (km)
316	2000	0,158	0,397
	1000	0,316	0,562
	500	0,632	0,795

Fonte: Elaborada pelo autor (2017).

Para facilitar, os valores obtidos foram arredondados, e foram construídos, a partir do software AutoCad, um arquivo no formato .dxf, contendo uma série de unidades quadradas (Figura 11) de tamanho estabelecido conforme a Tabela 3. Essa malha foi importada para o software Transcad v4.5, onde foram as malhas foram confrontadas com os Setores Censitários, e foram selecionadas as unidades de áreas que tocavam ou continham alguma feição da camada de SC.

Figura 11 - Malha de unidades de áreas sobreposta aos Setores Censitários.

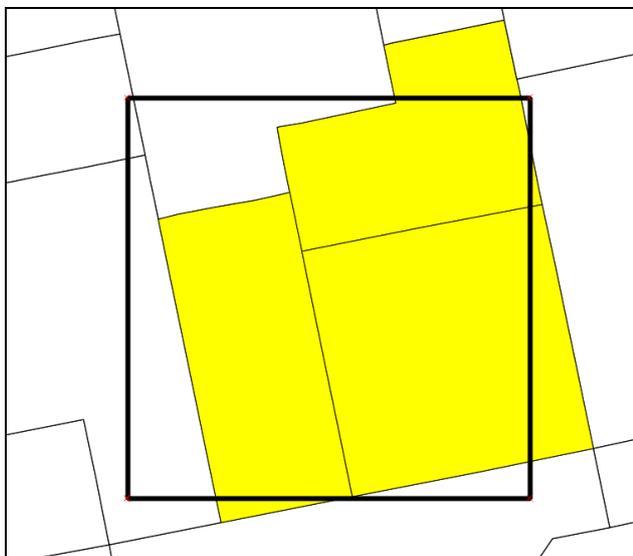


Fonte: Elaborada pelo autor (2017).

Terminada essa etapa, os setores foram agregados a partir de sua localização em relação à malha, o critério utilizado para confirmar se um determinado setor pertencia a uma determinada área da malha, era se o setor estava contido em mais de 50% de sua área dentro dessa determinada unidade de área da malha, assim como indica a figura 12, onde os setores indicados em amarelo, são os que fazem parte da determinada unidade de área. Dessa forma

as informações contidas para as variáveis utilizadas para análise, estavam agregadas por Setores Censitários, foram agregadas para cada configuração de área a partir do software QGis v.2.18.14.

Figura 12 - Exemplo de seleção de setores a partir da malha de áreas.



Fonte: Elaborada pelo autor (2017).

3.3 Execução de análises exploratórias (espaciais e não espaciais) dos dados agregados

Com os dados já agregados, foram selecionadas quatro variáveis, a partir do seu grau de correlação com a variável Taxa de acidentes totais por quilômetro, para que fossem realizadas as devidas avaliações espaciais. O critério utilizado foi de um grau de correlação maior ou igual a 0,4.

A partir dessas quatro variáveis e mais a taxa de acidentes totais por quilômetro de via foi realizada uma análise exploratória tanto de indicadores espaciais como não espaciais. Primeiramente, para a análise não espacial, foram feitas estatísticas descritivas das variáveis além de gerar matrizes de correlação para as mesmas para cada um dos cinco níveis de agregação, para que dessa forma, o PUAM fosse identificado a partir de padrões de variação conforme eram mudadas as configurações de áreas, tanto em indicadores como a média, desvio padrão, máximos e mínimos como também na correlação entre as variáveis.

Antes de serem iniciadas as análises espaciais, definiu-se como critério de vizinhança o critério de Queen, de forma a incorporar melhor a espacialidade presente na realidade para o ambiente estudado.

Para as todas as cinco variáveis, foram observados o índice global de moran univariado para cada uma das cinco agregações, além da observação do índice global de moran bivariado da taxa de acidentes totais por quilômetro de via com cada uma das outras cinco variáveis escolhidas a partir do critério de correlação.

A análise espacial foi continuada a partir da construção dos modelos de regressão espacial com efeitos globais SAR e SEM, mas, para que fossem obedecidas as hipóteses iniciais dos modelos, conforme citado anteriormente, a variável dependente, taxa de acidente de trânsito por quilômetro, necessitou de um tratamento inicial:

1. Primeiramente, para que os resíduos obedeçam uma distribuição normal, será necessário que a variável dependente, taxa de acidentes totais por quilômetro de via, também obedeça a distribuição normal, por isso foi necessário a aplicação do logaritmo natural dessa taxa;
2. Como existem diversas áreas com nenhum acidente e/ou nenhuma via existente, portanto, com taxa de acidentes nula, resultando em uma indeterminação na aplicação do logaritmo natural, foi necessário que o valor utilizado para a taxa não fosse o observado e sim o estimado a partir do estimado empírico bayesiano.

Utilizando-se o logaritmo natural da taxa estimada de acidentes por quilômetro de via como variável dependente e as demais selecionadas pelo teste de correlação como variáveis explicativas, os modelos foram construídos e foram observados indicadores relacionados ao poder de previsão do modelo, como o coeficiente de determinação, além do teste de normalidade para a comprovação das hipóteses iniciais estabelecidas. Além disso, avaliou-se os coeficientes de regressão de cada modelo de forma a entender a influência de cada variável dentro do modelo para cada uma das agregações.

3.4 Comparação dos indicadores resultantes das análises

Com a obtenção dos resultados da etapa anterior, foi feita a comparação dos mesmos entre as áreas delimitadas, e assim identificou-se a variação do padrão de comportamento das mesmas entre os diferentes layouts de delimitação das áreas, observando-se a influência do PUAM para acidentes de trânsito e as demais variáveis relacionadas com a ocorrência dos mesmos no município de Fortaleza.

4 RESULTADOS E DISCUSSÃO

4.1 Consolidação da base de dados

Segundo os resultados obtidos por Lima et al. (2015) foram atualizadas as informações necessárias para o georreferenciamento (nome, código e numeração) de 12.169 links da base viária da AMC de um total de 36.663 links. Além disso, foram corrigidas as conectividades de 746 cruzamentos e a topologia dos arcos de 7.092 links.

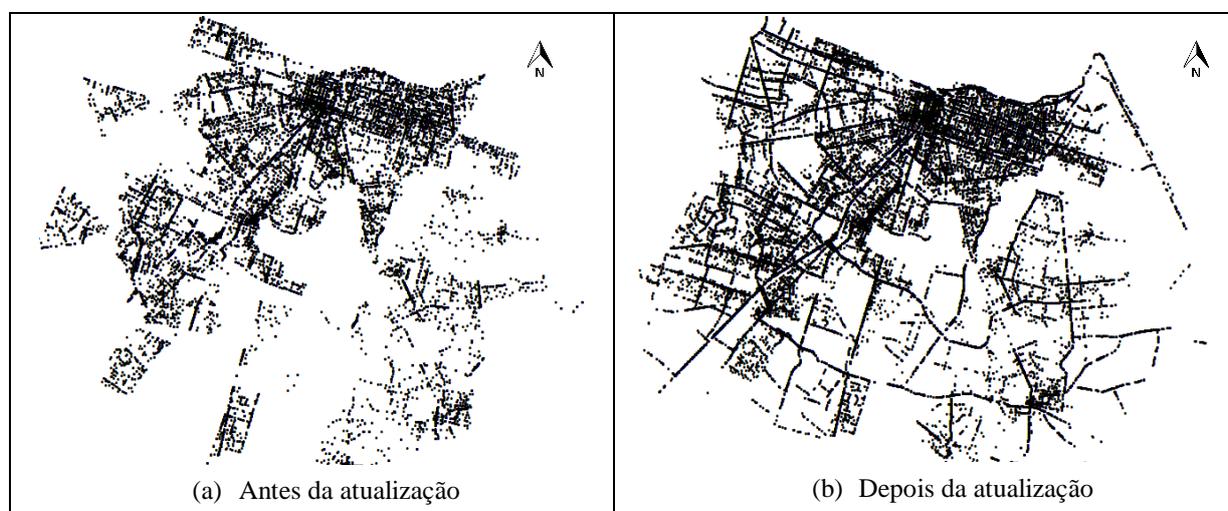
Essa atualização foi avaliada quanto a qualidade do georreferenciamento dos acidentes, através da geocodificação dos acidentes para os anos de 2004 a 2011 disponibilizados pelo SIAT/FOR, foram comparados os resultados obtidos para a base de dados “antes” e “depois” da atualização. Esses resultados estão resumidos na Tabela 4:

Tabela 4 - Qualidade do georreferenciamento de acidentes da base viária “antes” e “depois” das atualizações.

Base Viária	Acidentes em Interseções		Acidentes em segmentos	
	#Acidentes georreferenciados	% Sucesso do georreferenciamento	#Acidentes georreferenciados	% Sucesso do georreferenciamento
“Antes”	64.666	91,3%	38.741	48,0%
“Depois”	70.257	99,2%	66.507	82,4%

Fonte: Lima et al (2015).

Figura 13 - Distribuição espacial dos acidentes georreferenciados.



Fonte: Lima et al (2015).

Observa-se que para as interseções, o processo de atualização da base viária gerou um aumento de 7,9% no número de acidentes. Já para os segmentos, esse acréscimo foi de 34,4%. Outra grande contribuição dos autores está na avaliação da qualidade quanto a precisão do georreferenciamento. Foram georreferenciados os acidentes do ano de 2011 e esses foram comparados com os pontos classificados como “ROOFTOP”, classificação que é resultado de uma aplicação de uma rotina API (Application Programming Interfaces) gratuita do Google Maps. Ao todo foi testada uma amostra de 2275 acidentes, os resultados estão resumidos na Tabela 5:

Tabela 5 - Precisão do georreferenciamento antes e depois da atualização.

	#Acidentes Testados	# Acidentes Precisos	% Acidentes Precisos
Antes da Atualização	1.208	602	49,8%
Depois da Atualização	2.205	2.074	94,1%

Fonte: Lima et al (2015).

Em relação ao valor absoluto de acidentes precisos, percebe-se que a base de dados mais que dobrou a quantidade, com um total de 2074 acidentes cujas coordenadas foram condizentes com a referência utilizada, apresentado em torno de 6% da amostra sem apresentar resultados precisos para a localização.

4.2 Realização da agregação espacial dos dados

De acordo com as etapas descritas no item 3.2, foram obtidas mais três configurações de áreas, além das duas existentes (SC e ZT). A partir da malha com unidades de 0,158 km² (400 metros de lado) foram obtidas um total de 1272 áreas. Já utilizando a malha com unidades de área de 0,316 km² (550 metros de lado) obteve-se um total de 716 unidades de área. Por fim, para a malha com unidades de área de 0,632 km² obteve-se 380 unidades de área. A tabela 6 resume os resultados obtidos e as figuras 14 mostram a configuração final das áreas.

Figura 14 - Configurações finais de áreas: (a) Configuração 1 - Setores Censitários; (b) Configuração 2; (c) Configuração 3; (d) Configuração 4; (e) Configuração 5 - Zonas de Tráfego.

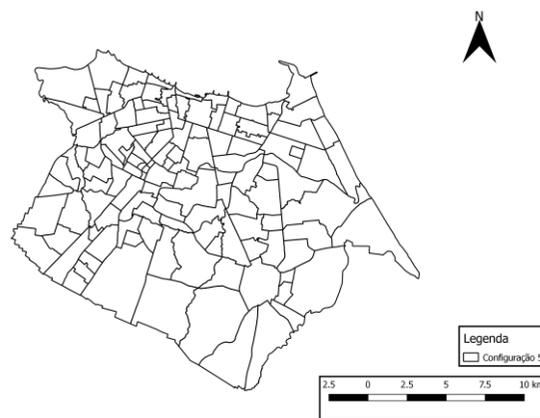
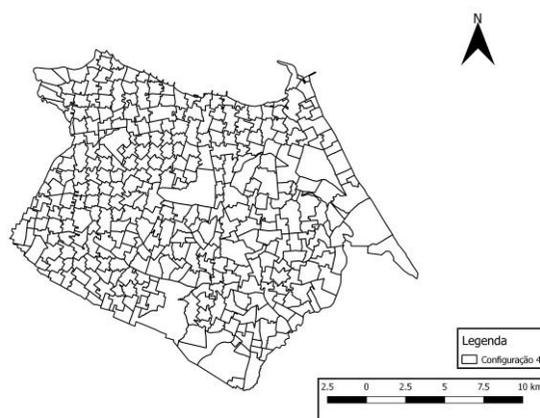
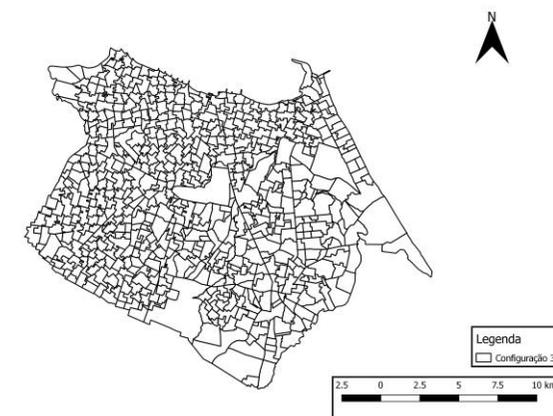
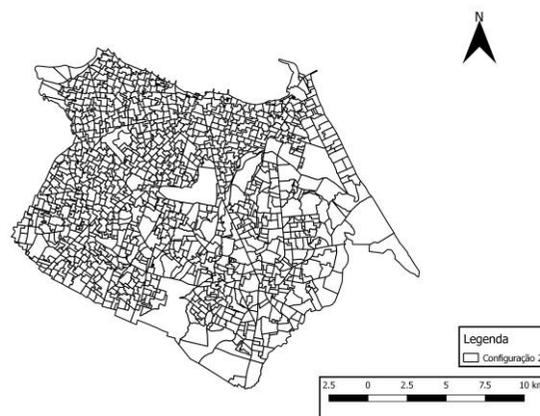
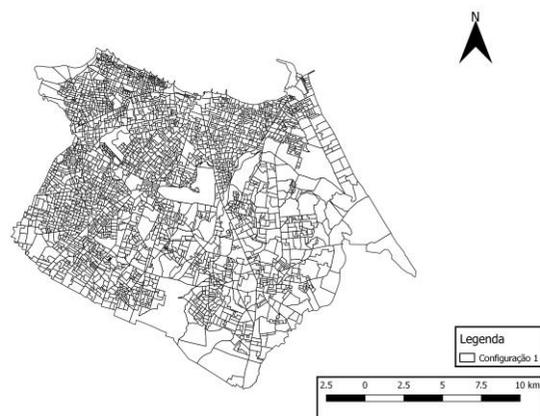


Tabela 6 - Total de unidades de áreas por configuração de agregação.

Item	Referência	Total de Unidades de Área
Configuração 1	Zonas de Tráfego	127
Configuração 2	Grids 800mx800m	380
Configuração 3	Grids 550mx550m	716
Configuração 4	Grids 400mx400m	1272
Configuração 5	Setores Censitários	2928

Fonte: Elaborada pelo autor (2017).

Com as configurações de áreas definidas, os dados puderam ser agregados para cada tipo, permitindo realizar as análises exploratórias para cada uma das configurações definidas.

A base de dados consolidada apresenta as variáveis descritas no Quadro 1:

Quadro 1 - Legenda das variáveis que fazem parte do banco de dados.

Variável	Descrição
AREA	Área da zona (m ²)
AREA_KM	Área da zona (km ²)
AC_T09_11	Número de Acidentes totais no período de 2009 a 2011
AC_V09_11	Número de Acidentes com vítimas no período de 2009 a 2011
TX_ACT	Taxa de Acidentes totais por km de via
TX_ACV	Taxa de Acidentes com vítimas por km de via
EXT_TOT	Extensão total de vias na zona (km)
POP_TOT	População total da zona
DOMICILIOS	Número de Domicílios da zona
DEN_I_SEM	Densidade de Interseções semaforizadas por km de via
DEN_I_NSEM	Densidade de Interseções não semaforizadas por km de via
D_EQUI_FE	Densidade de Equipamentos de fiscalização eletrônica por km de via
P_0_17	Proporção de residentes entre 0 e 17 anos de idade
P_18_64	Proporção de residentes entre 18 e 64 anos de idade
P_M64	Proporção de residentes acima de 64 anos de idade
P_D_A3SM	Proporção de domicílios com renda familiar até 3SM
P_D_M3SM	Proporção de domicílios com renda familiar maior que 3SM
U_RES_A	Uso do solo do tipo residencial (m ²) por m ² da zona
U_COPS_A	Uso do solo do tipo comercial (m ²) por m ² da zona

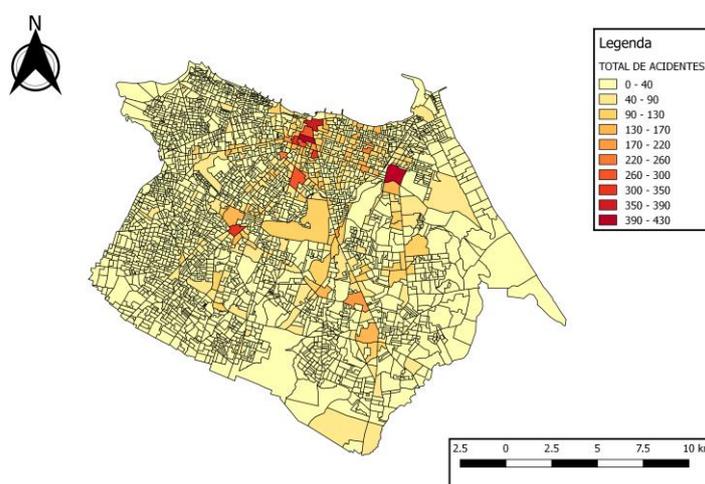
Fonte: Elaborado pelo autor (2017).

4.3 Análises exploratórias

4.3.1 Análise Não Espacial

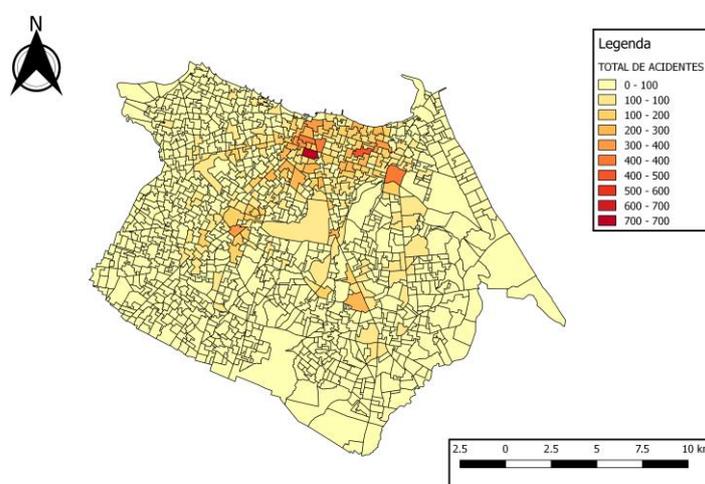
Com a base consolidada, primeiramente, foi feita uma análise visual, a partir de Mapas Coropléticos do total de Acidentes ocorridos por zona para cada agregação, representados pelas Figuras 14 a 18.

Figura 15 - Mapa Coroplético do Total de Acidentes por Setores Censitários.



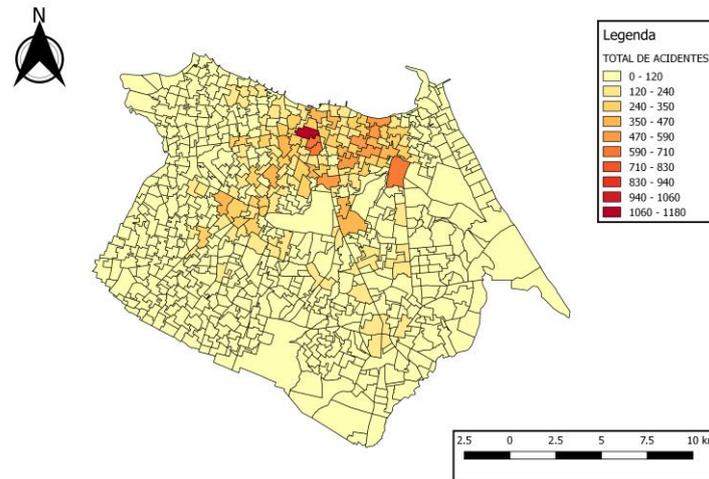
Fonte: Elaborado pelo autor (2017).

Figura 16 - Mapa Coroplético do Total de Acidentes por zonas da Configuração 2.



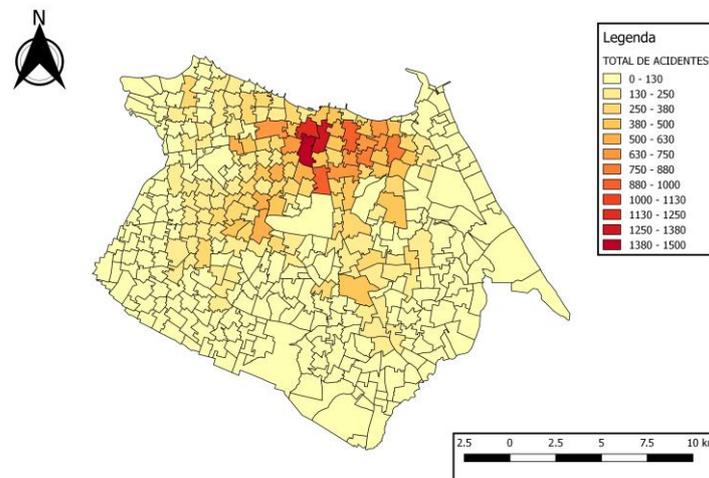
Fonte: Elaborado pelo autor (2017).

Figura 17 - Mapa Coroplético do Total de Acidentes por zonas da Configuração 3.



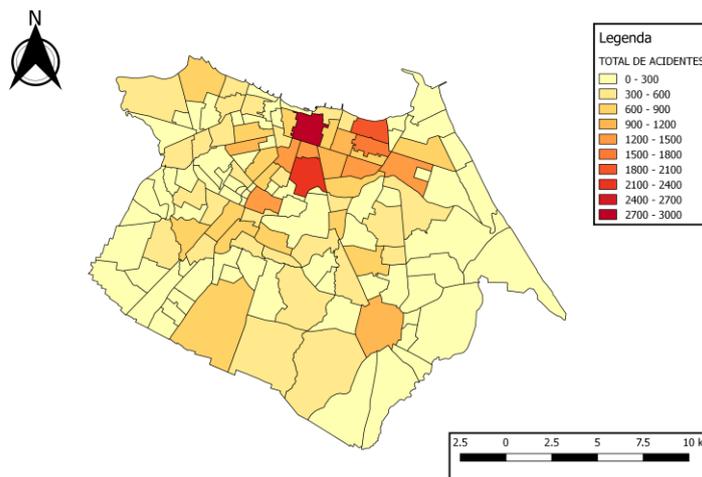
Fonte: Elaborado pelo autor (2017).

Figura 18 - Mapa Coroplético do Total de Acidentes por zonas da Configuração 4.



Fonte: Elaborado pelo autor (2017).

Figura 19 - Mapa Coroplético do Total de Acidentes por Zonas de Tráfego.



Fonte: Elaborado pelo autor (2017).

Primeiramente, mais explicitamente através dos setores censitários, nota-se que há uma maior ocorrência dos acidentes de trânsito nas proximidades de grandes corredores da cidade, como a Avenida Washington Soares, Avenida Bezerra de Menezes e as Rodovias BR-222 e BR-116.

Pode-se perceber também uma diferença muito maior entre as classes, enquanto que o número máximo de acidentes para a configuração em setores censitário é de até 430, para a configuração em zonas de tráfego esse número está na ordem de grandeza de 3000, isso se deve ao tamanho das unidades de áreas, quanto maiores as zonas, mais agregados são os dados consequentemente, maior a quantidade de acidentes.

Apesar dos mapas com configurações diferentes, é possível perceber que a ocorrência dos acidentes é maior na mesma região da cidade, próximo ao bairro do centro. Mas esse maior número de ocorrências é maior devido a valores maiores também da variável expositiva extensão total de via, cujos valores também são maiores na mesma região, sendo, dessa forma, mais razoável a observação da taxa de acidentes de trânsito e não dos valores absolutos de ocorrência dos mesmos.

A partir dessa definição, foram feitas as análises de correlação para todas as configurações de áreas, os resultados estão nas tabelas de 7 a 11.

Tabela 7 - Matriz de Correlação entre as variáveis - Configuração 1.

	AC_T09_11	AC_V09_11	TX_ACT	TX_ACV	EXT_TOT	POP_TOT	DOMICILIOS	DEN_I_SEM	DEN_I_NSEM	D_EQUI_FE	P_0_17	P_18_64	P_M64	P_D_A3SM	P_D_M3SM	U_RES_A	U_COPS_A
AC_T09_11	1,00																
AC_V09_11	0,90	1,00															
TX_ACT	0,80	0,72	1,00														
TX_ACV	0,65	0,77	0,85	1,00													
EXT_TOT	0,30	0,33	0,01	-0,03	1,00												
POP_TOT	-0,04	0,02	-0,17	-0,12	0,35	1,00											
DOMICILIOS	0,03	0,08	-0,09	-0,06	0,34	0,96	1,00										
DEN_I_SEM	0,56	0,46	0,72	0,55	-0,01	-0,20	-0,13	1,00									
DEN_I_NSEM	-0,29	-0,29	-0,26	-0,20	-0,31	-0,05	-0,10	-0,28	1,00								
D_EQUI_FE	0,23	0,20	0,25	0,21	0,03	-0,04	-0,01	0,21	-0,10	1,00							
P_0_17	-0,40	-0,33	-0,50	-0,36	-0,06	0,25	0,13	-0,45	0,18	-0,10	1,00						
P_18_64	0,28	0,25	0,34	0,26	0,06	-0,08	0,03	0,29	-0,12	0,09	-0,62	1,00					
P_M64	0,37	0,30	0,47	0,33	0,00	-0,24	-0,17	0,44	-0,11	0,07	-0,75	0,39	1,00				
P_D_A3SM	-0,42	-0,32	-0,49	-0,30	-0,14	0,20	0,07	-0,46	0,29	-0,12	0,81	-0,55	-0,59	1,00			
P_D_M3SM	0,42	0,32	0,49	0,30	0,14	-0,18	-0,05	0,46	-0,28	0,12	-0,78	0,64	0,61	-0,98	1,00		
U_RES_A	0,11	0,02	0,34	0,19	-0,17	-0,15	-0,05	0,36	-0,09	0,04	-0,52	0,41	0,43	-0,60	0,61	1,00	
U_COPS_A	0,54	0,42	0,56	0,39	0,04	-0,18	-0,11	0,55	-0,21	0,12	-0,44	0,30	0,43	-0,44	0,45	0,27	1,00

Fonte: Elaborado pelo autor (2017).

Tabela 8 - Matriz de Correlação entre as variáveis - Configuração 2.

	AC_T09_11	AC_V09_11	TX_ACT	TX_ACV	EXT_TOT	POP_TOT	DOMICILIOS	DEN_I_SEM	DEN_I_NSEM	D_EQUI_FE	P_0_17	P_18_64	P_M64	P_D_A3SM	P_D_M3SM	U_RES_A	U_COPS_A
AC_T09_11	1,00																
AC_V09_11	0,90	1,00															
TX_ACT	0,86	0,76	1,00														
TX_ACV	0,73	0,82	0,88	1,00													
EXT_TOT	0,24	0,28	-0,03	-0,06	1,00												
POP_TOT	0,07	0,19	-0,09	-0,02	0,44	1,00											
DOMICILIOS	0,13	0,24	-0,03	0,03	0,44	0,99	1,00										
DEN_I_SEM	0,67	0,56	0,76	0,62	-0,02	-0,12	-0,07	1,00									
DEN_I_NSEM	-0,24	-0,20	-0,22	-0,16	-0,20	0,17	0,14	-0,25	1,00								
D_EQUI_FE	0,26	0,23	0,32	0,29	-0,05	-0,07	-0,05	0,29	-0,14	1,00							
P_0_17	-0,50	-0,42	-0,54	-0,42	-0,07	0,12	0,05	-0,49	0,13	-0,17	1,00						
P_18_64	0,33	0,29	0,38	0,32	0,06	-0,06	0,00	0,33	-0,11	0,12	-0,56	1,00					
P_M64	0,53	0,44	0,53	0,40	0,07	-0,03	0,03	0,49	-0,04	0,13	-0,75	0,44	1,00				
P_D_A3SM	-0,48	-0,34	-0,52	-0,35	-0,09	0,20	0,13	-0,50	0,26	-0,17	0,83	-0,52	-0,60	1,00			
P_D_M3SM	0,49	0,35	0,52	0,35	0,10	-0,19	-0,11	0,51	-0,25	0,17	-0,77	0,64	0,63	-0,97	1,00		
U_RES_A	0,35	0,24	0,41	0,27	-0,04	0,09	0,16	0,38	-0,07	0,09	-0,58	0,40	0,58	-0,59	0,61	1,00	
U_COPS_A	0,64	0,52	0,66	0,50	0,03	-0,10	-0,05	0,71	-0,18	0,17	-0,50	0,34	0,51	-0,48	0,49	0,40	1,00

Fonte: Elaborado pelo autor (2017).

Tabela 9 - Matriz de Correlação entre as variáveis - Configuração 3.

	AC_T09_11	AC_V09_11	TX_ACT	TX_ACV	EXT_TOT	POP_TOT	DOMICILIOS	DEN_I_SEM	DEN_I_NSEM	D_EQUI_FE	P_0_17	P_18_64	P_M64	P_D_A3SM	P_D_M3SM	U_RES_A	U_COPS_A
AC_T09_11	1,00																
AC_V09_11	0,91	1,00															
TX_ACT	0,87	0,78	1,00														
TX_ACV	0,75	0,83	0,90	1,00													
EXT_TOT	0,34	0,39	0,05	0,03	1,00												
POP_TOT	0,20	0,33	0,00	0,08	0,60	1,00											
DOMICILIOS	0,26	0,39	0,06	0,13	0,60	0,99	1,00										
DEN_I_SEM	0,69	0,58	0,80	0,66	0,02	-0,07	-0,02	1,00									
DEN_I_NSEM	-0,17	-0,10	-0,17	-0,10	-0,06	0,29	0,27	-0,21	1,00								
D_EQUI_FE	0,25	0,22	0,32	0,29	-0,06	-0,06	-0,04	0,34	-0,10	1,00							
P_0_17	-0,52	-0,44	-0,58	-0,46	-0,09	0,06	0,00	-0,52	0,10	-0,23	1,00						
P_18_64	0,34	0,29	0,39	0,32	0,10	-0,04	0,01	0,35	-0,11	0,19	-0,56	1,00					
P_M64	0,58	0,51	0,62	0,51	0,14	0,06	0,11	0,54	0,00	0,20	-0,78	0,48	1,00				
P_D_A3SM	-0,46	-0,31	-0,52	-0,34	-0,04	0,21	0,15	-0,47	0,24	-0,23	0,83	-0,54	-0,60	1,00			
P_D_M3SM	0,47	0,33	0,53	0,36	0,06	-0,19	-0,13	0,48	-0,23	0,23	-0,79	0,68	0,62	-0,98	1,00		
U_RES_A	0,46	0,34	0,51	0,36	0,06	0,17	0,24	0,44	-0,06	0,21	-0,61	0,43	0,61	-0,61	0,62	1,00	
U_COPS_A	0,61	0,50	0,67	0,52	0,04	-0,07	-0,02	0,76	-0,16	0,22	-0,53	0,37	0,53	-0,47	0,48	0,44	1,00

Fonte: Elaborado pelo autor (2017).

Tabela 10 - Matriz de Correlação entre as variáveis - Configuração 4.

	AC_T09_11	AC_V09_11	TX_ACT	TX_ACV	EXT_TOT	POP_TOT	DOMICILIOS	DEN_I_SEM	DEN_I_NSEM	D_EQUI_FE	P_0_17	P_18_64	P_M64	P_D_A3SM	P_D_M3SM	U_RES_A	U_COPS_A
AC_T09_11	1																
AC_V09_11	0,92	1,00															
TX_ACT	0,86	0,76	1,00														
TX_ACV	0,78	0,84	0,89	1,00													
EXT_TOT	0,45	0,54	0,16	0,18	1,00												
POP_TOT	0,29	0,47	0,09	0,21	0,72	1,00											
DOMICILIOS	0,36	0,52	0,15	0,26	0,73	0,99	1,00										
DEN_I_SEM	0,80	0,69	0,85	0,73	0,15	0,04	0,09	1,00									
DEN_I_NSEM	-0,12	-0,04	-0,13	-0,04	0,03	0,32	0,29	-0,17	1,00								
D_EQUI_FE	0,37	0,32	0,51	0,46	0,01	-0,01	0,01	0,46	-0,13	1,00							
P_0_17	-0,62	-0,54	-0,65	-0,56	-0,22	-0,08	-0,14	-0,57	0,03	-0,31	1,00						
P_18_64	0,50	0,43	0,53	0,46	0,17	0,01	0,06	0,46	-0,07	0,27	-0,93	1,00					
P_M64	0,64	0,58	0,66	0,58	0,24	0,16	0,22	0,59	0,03	0,29	-0,87	0,63	1,00				
P_D_A3SM	-0,55	-0,41	-0,61	-0,45	-0,14	0,08	0,02	-0,56	0,17	-0,33	0,87	-0,84	-0,72	1,00			
P_D_M3SM	0,55	0,41	0,61	0,45	0,14	-0,08	-0,02	0,56	-0,17	0,33	-0,87	0,84	0,72	-1,00	1,00		
U_RES_A	0,52	0,41	0,58	0,44	0,17	0,23	0,29	0,51	-0,03	0,25	-0,67	0,54	0,69	-0,66	0,66	1,00	
U_COPS_A	0,75	0,65	0,76	0,65	0,18	0,07	0,12	0,84	-0,15	0,40	-0,60	0,50	0,60	-0,56	0,56	0,59	1,00

Fonte: Elaborado pelo autor (2017).

Tabela 11 - Matriz de Correlação entre as variáveis - Configuração 5

	AC_T09_11	AC_V09_11	TX_ACT	TX_ACV	EXT_TOT	POP_TOT	DOMICILIOS	DEN_I_SEM	DEN_I_NSEM	D_EQUI_FE	P_0_17	P_18_64	P_M64	P_D_A3SM	P_D_M3SM	U_RES_A	U_COPS_A
AC_T09_11	1,00																
AC_V09_11	0,90	1,00															
TX_ACT	0,79	0,60	1,00														
TX_ACV	0,65	0,60	0,89	1,00													
EXT_TOT	0,20	0,40	-0,27	-0,34	1,00												
POP_TOT	0,14	0,42	-0,26	-0,24	0,88	1,00											
DOMICILIOS	0,19	0,45	-0,22	-0,20	0,88	1,00	1,00										
DEN_I_SEM	0,64	0,43	0,84	0,72	-0,28	-0,30	-0,26	1,00									
DEN_I_NSEM	-0,24	-0,13	-0,16	-0,02	-0,12	0,12	0,11	-0,15	1,00								
D_EQUI_FE	0,47	0,36	0,53	0,49	-0,11	-0,09	-0,06	0,37	-0,19	1,00							
P_0_17	-0,55	-0,33	-0,75	-0,61	0,31	0,34	0,30	-0,70	0,04	-0,32	1,00						
P_18_64	0,50	0,29	0,62	0,48	-0,25	-0,32	-0,28	0,59	-0,14	0,32	-0,92	1,00					
P_M64	0,51	0,31	0,74	0,64	-0,33	-0,31	-0,27	0,68	0,09	0,27	-0,90	0,65	1,00				
P_D_A3SM	-0,50	-0,22	-0,66	-0,43	0,24	0,35	0,31	-0,65	0,24	-0,32	0,90	-0,90	-0,72	1,00			
P_D_M3SM	0,50	0,22	0,66	0,43	-0,24	-0,35	-0,31	0,65	-0,24	0,32	-0,90	0,90	0,72	-1,00	1,00		
U_RES_A	0,42	0,21	0,62	0,46	-0,24	-0,16	-0,12	0,59	-0,04	0,27	-0,74	0,63	0,72	-0,73	0,73	1,00	
U_COPS_A	0,74	0,54	0,82	0,67	-0,19	-0,23	-0,19	0,86	-0,17	0,42	-0,68	0,56	0,67	-0,59	0,59	0,58	1,00

Fonte: Elaborado pelo autor (2017).

Tomando-se como critério de classificação de alto grau de correlação o valor de 0,4, observa-se que quanto mais agregado o nível de configuração for, ou seja, quanto menos unidades de áreas, mais variáveis apresentam uma correlação alta com a variável taxa de acidentes de trânsito por quilômetro de via (TX_ACT). Enquanto que no nível de Setores Censitários, além das variáveis que representem as quantidades e as taxas de acidentes de trânsito totais e com vítimas, 4 variáveis foram classificadas dessa forma, no nível de Zonas de Tráfego, 7 variáveis apresentaram valores de correlação superiores a 0,4, como está representado no Quadro 2. Isso ocorre por conta de um processo de suavização dos dados que é maior a medida que se aumenta a agregação. Esse efeito já vem sendo observado desde a década de 30 e que está relacionado com o efeito de escala.

Quadro 2 - Lista das variáveis com alta correlação por configuração.

Variável	Config. 1	Config. 2	Config. 3	Config. 4	Config. 5
AC_T09_11	✓	✓	✓	✓	✓
AC_V09_11	✓	✓	✓	✓	✓
TX_ACT	✓	✓	✓	✓	✓
TX_ACV	✓	✓	✓	✓	✓
EXT_TOT					
POP_TOT					
DOMICILIOS					
DEN_I_SEM	✓	✓	✓	✓	✓
DEN_I_NSEM					
D_EQUI_FE				✓	✓
P_0_17					
P_18_64				✓	✓
P_M64	✓	✓	✓	✓	✓
P_D_A3SM					
P_D_M3SM	✓	✓	✓	✓	✓
U_RES_A		✓	✓	✓	✓
U_COPS_A	✓	✓	✓	✓	✓

Fonte: Elaborado pelo autor (2017).

Não somente a quantidade de variáveis, mas os índices de correlação elevaram-se, pode-se tomar como exemplo a variável Densidade de Interseções semaforizadas, que para as Configurações de número 1, 2, 3 e 4 apresentou, respectivamente, índices de 0,72, 0,76, 0,80, 0,85. Nota-se que o padrão não foi seguido com a Configuração 5, houve uma diminuição do valor para 0,84, explicado provavelmente por conta de que esse tipo específico de layout não foi construído de acordo com a mesma metodologia das Configurações 2,3 e 4, apesar de também ser uma forma mais agregada da Configuração 1.

Essa inconformidade com o padrão de crescimento dos índices de correlação não foi observada para as variáveis demográficas e de uso do solo, População de Residentes acima de 64 anos (P_M64) e Uso do Solo do Tipo Comercial por metro quadrado (U_COPS_A) por exemplo, as quais seguiram o padrão de crescimento da correlação a medida que decrescia o número de unidades de áreas.

Como foi citado anteriormente, foram selecionadas as variáveis que apresentaram um grau de correlação maior que 0,4. Dessa forma, foram selecionadas aquelas que se repetiam em todos os níveis de agregação, são elas: Densidade de Interseções Semaforizadas por quilômetro de via (DEN_I_SEM); Proporção de Residentes Acima de 64 anos de idade (P_M64); Proporção de Domicílios com Renda Maior que 3 Salários Mínimos (P_D_M3SM) e Uso do Solo do Tipo Comercial por metro quadrado (U_COPS_A). A variável Taxa de Acidentes com Vítimas por Quilômetro de Via (TX_ACV) também apresentou alta correlação em todas as configurações com a variável TX_ACT, mas não foi utilizada para a análise pois uma informação sobre acidente de trânsito já consiste na análise.

A Tabela 12 mostra indicadores estatísticos descritivos das variáveis supracitadas, obtidos através do software GeoDa.

Tabela 12 - Resultados da estatística descritiva das variáveis.

Agregações/Variáveis		TX_ACT	DEN_I_SEM	P_M64	P_D_M3SM	U_COPS_A
Configuração 1	Média	9,900	0,280	0,069	0,364	0,077
	Mínimo	0,000	0,000	0,000	0,000	0,000
	Máximo	116,349	4,381	0,345	1,000	2,541
	Desv. Padrão	13,586	0,603	0,039	0,275	0,125
Configuração 2	Média	11,159	0,262	0,066	0,378	0,078
	Mínimo	0,000	0,000	0,000	0,000	0,000
	Máximo	106,960	4,235	0,186	1,000	1,673
	Desv. Padrão	14,138	0,494	0,033	0,262	0,118
Configuração 3	Média	11,181	0,227	0,065	0,381	0,077
	Mínimo	0,000	0,000	0,000	0,000	0,000
	Máximo	91,845	3,242	0,169	1,000	1,673
	Desv. Padrão	13,139	0,412	0,031	0,253	0,117
Configuração 4	Média	11,014	0,188	0,063	0,361	0,070
	Mínimo	0,000	0,000	0,019	0,001	0,000
	Máximo	81,338	2,875	0,154	0,968	1,025
	Desv. Padrão	13,152	0,327	0,029	0,229	0,101
Configuração 5	Média	15,836	0,300	0,073	0,414	0,091
	Mínimo	0,269	0,000	0,029	0,075	0,000
	Máximo	65,041	1,917	0,141	0,919	0,724
	Desv. Padrão	13,308	0,353	0,028	0,216	0,098

Fonte: Elaborado pelo autor (2017).

Os valores de desvio padrão para a variável TX_ACT em todas as configurações foram bastante elevados quando comparados com sua medida de tendência central, significando que, ao longo de toda a cidade não existe uma distribuição uniforme para a taxa de acidentes, existindo locais com valores de até 116 acidentes por quilômetro de via, como mostrado para a Configuração 1.

Interessante ressaltar que, ainda sobre a variável TX_ACT, apenas a configuração de número 5 apresentou valor mínimo maior que zero, significando que todas as zonas desse tipo de layout há ocorrência de acidente de trânsito e também possuem vias, diferentemente das demais, mostrando que quanto mais desagregada for a configuração maior a chance de ocorrer observações com valores nulos, influenciando nos resultados obtidos para a média e também para desvio padrão. Esse comportamento pode ser observado nas demais variáveis, de tal forma que, para a configuração de número 5, apresentou as maiores médias para todas as variáveis observadas, devido a inexistência de observações com valores nulos ou até

mesmo apresentando valores nulos em menor quantidade, causando uma superestimação do resultado obtido para a média.

Com dados mais agregados, além da diminuição de valores nulos, percebe-se também a diminuição dos valores máximos observados, pois todas as variáveis são medidas proporcionais, ficando atreladas a extensão total de vias, como ocorre para a taxa de acidentes, ou a área de cada zona, e com o aumento da unidade de área conforme aumenta o nível de agregação, essas variáveis expositivas também aumentam, ocorrendo uma suavização dos dados, como observado também para o índice de correlação e, conseqüentemente, uma diminuição do valor do desvio padrão.

As médias obtidas para as variáveis não apresentaram um padrão de comportamento influenciadas pelo nível de agregação, como pode-se perceber para a variável P_D_3SM, que, para configurações de áreas intermediárias, ocorreu que a medida de tendência central apresentou um decréscimo, quando se esperava um aumento, baseado no padrão observado para as outras configurações mais desagregadas.

O mesmo comportamento é observado para a taxa de acidentes, cujos valores de média crescem com o aumento da agregação dos dados até a Configuração 3, e depois ocorre uma diminuição para a Configuração de número 4, que novamente aumenta para a Configuração 5, é razoável pensar que o crescimento deveria ser observado ao aumentar-se o nível de agregação, mas como se trata da taxa de acidentes, deve ser observado também, o comportamento da variável expositiva ao longo dos níveis de agregação.

4.3.2 Análise Espacial

Através do software GeoDa, utilizando-se como consideração de vizinhança o critério Queen, foram obtidos os índices de Moran globais para a variável TX_ACT, para as 5 configurações de áreas determinadas anteriormente, a significância dos mesmos foram verificadas aplicando-se o método com 999 permutações. Os resultados estão mostrados na Tabela 13 e os diagramas de dispersão de Moran estão no Apêndice B.

Tabela 13 - Índices de Moran Globais para Taxa de Acidentes por quilômetro de via.

	Índice de Moran	Significância
Configuração 1	0,61	0,001
Configuração 2	0,65	0,001
Configuração 3	0,68	0,001
Configuração 4	0,71	0,001
Configuração 5	0,65	0,001

Fonte: Elaborado pelo autor (2017).

Os índices nos mostram que há uma correlação espacial alta entre as observações dessa variável, todos acima de 0,6, até mesmo na forma mais desagregada.

Para as configurações de 1 a 4 percebe-se que a dependência espacial segue um padrão de crescimento, sendo que esse padrão não é obedecido para a Configuração 5, cujo valor da dependência espacial para essa variável diminui em relação a configuração mais desagregada imediatamente superior e quase de mesmo valor a Configuração 1, forma mais desagregada. Possivelmente, esse comportamento em relação aos demais layouts está atrelado ao método de criação das configurações 2,3 e 4 o qual é diferente do método utilizado para a criação das Zonas de Tráfego (Configuração 5) que considera alguns aspectos urbanos para sua delimitação.

Além disso, pode-se perceber que a taxa de acidentes por quilômetro de via, através da Tabela 14, que apresenta os resultados para o Índice de Moran Bivariado, tomando-se a variável Taxa de acidentes por quilômetro de via, que existe correlação espacial da ocorrência de acidentes com as demais variáveis, de tal forma que essa correlação cresce à medida que o nível de agregação aumenta, ocorrendo da mesma forma que o Índice de Moran Univariado, onde há uma quebra do padrão de crescimento na Configuração 5, que corresponde às Zonas de Tráfego.

Tabela 14 - Índices de Moran bivariado da variável taxa de acidentes por quilômetro de via.

	DEN_I_SEM		P_M64		P_D_M3SM		U_COPS_A	
	Moran Bivariado	Signif.						
Configuração 1	0,524	0,001	0,468	0,001	0,458	0,001	0,453	0,001
Configuração 2	0,597	0,001	0,517	0,001	0,467	0,001	0,544	0,001
Configuração 3	0,623	0,001	0,584	0,001	0,467	0,001	0,555	0,001
Configuração 4	0,701	0,001	0,633	0,001	0,543	0,001	0,645	0,001
Configuração 5	0,652	0,001	0,583	0,001	0,502	0,001	0,612	0,001

Fonte: Elaborado pelo autor (2017).

De forma a complementar a análise, os modelos SAR e SEM foram construídos, os coeficientes e os demais parâmetros de cada modelo, estão especificados nas Tabelas 15 e 16.

Tabela 15 - Coeficientes e Parâmetros dos modelos de regressão espacial SEM.

		Configuração 1	Configuração 2	Configuração 3	Configuração 4	Configuração 5
DEN_I_SEM	Coef.	0,711	0,907	1,124	1,262	0,582
	Desv. Padrão	0,041	0,085	0,135	0,248	0,306
	Prob.	0,000	0,000	0,000	0,000	0,057*
P_M64	Coef.	2,788	3,408	7,197	8,187	8,389
	Desv. Padrão	0,725	1,261	1,698	2,634	3,261
	Prob.	0,000	0,007	0,000	0,002	0,010
P_D_M3SM	Coef.	0,816	0,736	0,625	0,553	0,871
	Desv. Padrão	0,101	0,160	0,206	0,334	0,425
	Prob.	0,000	0,000	0,002	0,098*	0,040
U_COPS_A	Coef.	0,737	0,922	0,542	1,591	1,303
	Desv. Padrão	0,169	0,317	0,433	0,707	0,984
	Prob.	0,000	0,004	0,211*	0,024*	0,185*
CONSTANTE	Coef.	0,607	0,694	0,601	0,541	0,989
	Desv. Padrão	0,072	0,114	0,139	0,194	0,233
	Prob.	0,000	0,000	0,000	0,005	0,000
λ	Coef.	0,682	0,702	0,677	0,694	0,659
	Desv. Padrão	0,018	0,026	0,036	0,046	0,083
	Prob.	0,000	0,000	0,000	0,000	0,000
R ²		0,625	0,657	0,665	0,704	0,710

(*) Não-significativos

Fonte: Elaborado pelo autor (2017).

Tabela 16 - Coeficientes e Parâmetros dos modelos de regressão espacial SAR

		Configuração 1	Configuração 2	Configuração 3	Configuração 4	Configuração 5
DEN_I_SEM	Coef.	0,504	0,549	0,670	0,564	0,371
	Desv. Padrão	0,035	0,073	0,121	0,231	0,289
	Prob.	0,000	0,000	0,000	0,015	0,200*
P_M64	Coef.	2,247	3,576	6,330	6,053	6,489
	Desv. Padrão	0,591	1,003	1,401	2,177	2,869
	Prob.	0,000	0,000	0,000	0,005	0,024
P_D_M3SM	Coef.	0,399	0,447	0,389	0,328	0,513
	Desv. Padrão	0,080	0,124	0,158	0,254	0,368
	Prob.	0,000	0,000	0,014	0,197*	0,128*
U_COPS_A	Coef.	0,555	0,435	-0,075	1,022	0,833
	Desv. Padrão	0,163	0,299	0,409	0,729	0,970
	Prob.	0,001	0,146	0,855	0,161*	0,391*
CONSTANTE	Coef.	0,041	-0,009	-0,048	-0,062	0,141
	Desv. Padrão	0,036	0,055	0,074	0,100	0,171
	Prob.	0,244*	0,000	0,515*	0,539*	0,410*
ρ	Coef.	0,613	0,624	0,596	0,622	0,558
	Desv. Padrão	0,018	0,026	0,035	0,046	0,079
	Prob.	0,000	0,000	0,000	0,000	0,000
R ²		0,627	0,656	0,660	0,698	0,704

(*) Não-significativos

Fonte: Elaborado pelo autor (2017).

Primeiramente, através da observação da probabilidade dos coeficientes de espacialidade presentes nos modelos (λ e ρ), confirma-se a espacialidade da variável dependente, pois os mesmos possuem significância estatística para maior que seja o grau de confiança adotado.

Ainda em relação a significância das variáveis, a constante ρ obteve resultados muito bons na modelagem do tipo SEM, enquanto que nos modelos do tipo SAR só se obteve resultados satisfatórios em um nível de agregação, a configuração 2.

Vale ressaltar que para a construção dos modelos de regressão, é recomendado que se faça uma seleção das variáveis pelo grau de correlação das mesmas. Nesse estudo em específico, os modelos de regressão são utilizados com o intuito de comparar o comportamento das variáveis selecionadas, mas, claramente, para cada nível de agregação, seriam construídos modelos com variáveis explicativas diferentes. Um exemplo de variável que seria excluída dos modelos com o aumento do nível de agregação, é a variável Uso do solo comercial por metro quadrado de zona (U_COPS_A), pois a mesma apresenta elevado

grau de correlação com outras variáveis, como densidade de interseções semaforizadas (DEN_I_SEM) e acaba perdendo sua significância dentro dos modelos.

Quanto as demais variáveis, se tratando das que foram analisadas, percebeu-se que densidade de interseções semaforizadas, proporção de domicílios com renda familiar maior que 3 salários mínimos, e uso do solo do tipo comercial por metro quadrado de zona, perderam sua significância ao se elevar o nível de agregação, ocorrendo de forma mais assinalada nos modelos do tipo SAR. Esses resultados mostram como que, apesar de certas variáveis explicativas apresentarem bons graus de correlação com a variável dependente, a configuração espacial pode tornar ou não uma determinada variável como não significativa para o modelo de previsão e, conseqüentemente, para o planejamento de transportes.

Outro ponto interessante a ser destacado é o aumento do valor do coeficiente R^2 para todos os dois tipos de modelos, a justificativa para esse aumento pode estar relacionada ao fato de que, como a correlação entre as variáveis aumenta também com o nível de agregação, como já foi apresentado anteriormente, o poder explicativo do modelo é afetado diretamente de forma positiva. Mas, esse aumento do grau de correlação também afeta de forma negativa dentro dos modelos, pois percebe-se que há uma quantidade maior de variáveis com probabilidade elevada, indicando um baixo nível de significância das mesmas dentro do modelo, pois como sua correlação aumenta, acaba por criar uma espécie de redundância dentro do modelo, o ideal seria a utilização de variáveis o minimamente correlacionadas possível, que também possuem um alto grau de correlação com a variável dependente.

Observa-se também que, em ambos os modelos, os coeficientes de cada variável são diferentes quando são abordadas configurações de áreas diferentes, sem seguir nenhum padrão de variação. Essa mudança também pode ser identificada nos desvios padrões, tanto para os modelos do tipo SAR quanto para os modelos do tipo SEM, esse indicador foi maior quanto maior o nível de agregação.

A incorporação da espacialidade dos acidentes foi bem sucedida através dos dois modelos, como indica o Mapa de Moran de Resíduos dos modelos presentes no Apêndice C, no qual os resíduos foram distribuídos espacialmente de forma aleatória e apresentando valores próximos de zero para o índice de Moran, como apresenta a Tabela 17.

Tabela 17 - Índice de Moran dos Resíduos dos Modelos de Regressão Espacial

	Índice de Moran dos Resíduos	
	SAR	SEM
Configuração 1	0,011	-0,018
Configuração 2	0,008	-0,022
Configuração 3	0,009	-0,021
Configuração 4	0,023	-0,017
Configuração 5	0,050	-0,015

Fonte: Elaborado pelo autor (2017).

5 CONCLUSÕES

Esse Trabalho de Conclusão de Curso apresentou a influência do Problema de Unidade de Área Modificável para variáveis relacionadas a acidentes de trânsito e a sua ocorrência, através de análises exploratórias tanto espacial como não espacial em áreas.

Primeiramente, com a atualização da base de dados, para que se realizasse o georreferenciamento dos acidentes de trânsito, obteve-se uma melhora de 43% na quantidade de acidentes georreferenciados e, com a seleção de uma amostra, a base melhorou a precisão desse georreferenciamento em quase 45%.

Também foram construídas, além das configurações em setores e em zonas de tráfego, mais três layouts de áreas. Obtendo-se um total de unidades de áreas do nível mais desagregado para o nível mais agregado: 1272, 716 e 380 respectivamente.

Através da análise não espacial, percebeu-se que o grau de correlação entre as variáveis aumentava com a elevação do nível de agregação, de tal forma que, 5 variáveis apresentaram uma correlação boa com a variável taxa de acidente de trânsito por quilômetro de via quando agregadas na configuração de setores censitários. Para a configuração de Zonas de Tráfego esse número foi de 8, incluindo as mesmas 5 que apresentaram correlação alta na agregação em setores, mas com graus de correlação melhores. Por essa análise de correlação foram escolhidas as variáveis que fizeram parte das análises espaciais, foram elas: Densidade de interseções semaforizadas, Proporção de residentes acima de 64 anos de idade, Proporção de domicílios com renda familiar acima de 3 salários mínimos e Uso do solo do tipo comercial por metro quadrado da zona.

Foi observado também que, para a taxa de acidentes de trânsito indicadores como a média, também sofre alterações quando há a mudança da configuração da unidade de agregação, as mesmas não seguiram um padrão de aumento ou decréscimo diretamente proporcional ao nível de agregação, mas ocorreu que, houveram pequenas distorções do padrão de crescimento para configurações de áreas intermediárias.

Sobre a análise espacial, sendo mais completa por conta da incorporação da espacialidade dos acidentes, há um padrão de aumento da correlação espacial para a taxa de acidentes a medida que o nível de agregação também aumenta, com exceção para a configuração em zonas de tráfego, cujo valor diminui em relação ao nível de agregação anterior. Esse comportamento foi observado também com a análise do índice de Moran

bivariado da mesma variável taxa de acidentes com cada uma das variáveis escolhidas pelo critério de correlação.

Por fim, com a construção dos modelos de regressão que consideram os efeitos espaciais de forma global (SAR e SEM), pôde-se perceber o aumento no poder de estimação dos mesmos com o aumento do nível de agregação, apesar de que algumas variáveis apresentaram um valor p alto para níveis de agregação mais altos, como foi observado de forma mais acintosa nos modelos do tipo SAR para a variável relacionada ao uso do solo comercial por metro quadrado de zona, mostrando a influência das configurações de agregação na escolha das variáveis a serem analisadas para estudos de previsão de acidentes e para o planejamento nos transportes.

Devido a forma de como as áreas foram construídas, de forma arbitrária, diferentemente da forma de como as zonas de tráfego, as anomalias nos padrões encontrados podem ser explicadas a partir disso, mas de forma geral pôde-se perceber a influência da mudança da configuração utilizada, mesmo que pequenas em ordem de grandeza mas significativas com vista o planejamento dos transportes, tanto no nível tático em relação a locais com prioridade intervenção como no nível estratégico, quando se trata das variáveis que são correlacionadas com a ocorrência dos acidentes.

Conforme foi citado, as zonas de tráfego foram construídas a partir de um método diferente das demais configurações de áreas, ficando como sugestão a tentativa de construção das configurações de áreas levando em conta a variável acidente, numa tentativa de minimizar a influência do PUAM nas análises.

Quanto a modelagem, foram testados nesse trabalho apenas modelos que consideram a espacialidade globalmente, recomendando-se a utilização de modelos que consideram a espacialidade de forma local, avaliando-se também o comportamento em outro âmbito, obtendo-se informações mais precisas sobre a influência do PUAM, ajudando assim no processo de tomada de decisão.

REFERÊNCIAS

- ABDEL-ATY, M., LEE, J., SIDDIQUI, C. e CHOI, K. **Geographical unit based analysis in the context of transportation safety planning**. *Transportation Research Part A*, v. 49, p.62-75. 2013.
- AGUERO-VALVERDE, J.; JOVANIS, P. P. **Spatial analysis of fatal and injury crashes in Pennsylvania**. v. 38, n. December 2005, p. 618–625, 2006.
- ANSELIN, L. **A Workbook for Using**. 1992.
- ANSELIN, L. **Spatial Regression**. 1993.
- ANSELIN, L.; BERA, A. **Spatial dependence in linear regression models with an introduction to spatial econometrics**. *Handbook of Applied Economic Statistic*, p. 237–389, 1998.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 10697**: Pesquisa de acidentes de trânsito. Rio de Janeiro, 1989.
- CÂMARA, G.; CARVALHO, M. S.; CRUZ, O. G.; CORREA, V. **Análise espacial de áreas**. *Análise espacial de dados geográficos*. São José dos Campos: INPE, p. 44, 2002.
- CARDOSO, G.; GOLDNER, L. G. **Desenvolvimento e aplicação de modelos para previsão de acidentes de trânsito**. p. 43–51, 2006.
- CARVALHO, T. **Regressão Binomial Negativa geograficamente ponderada: Modelando superdispersão espacial**. Dissertação de Mestrado - Departamento de Estatística, Universidade de Brasília, 2011.
- CEDER, A.; LIVNEH, M. **Relationships between road accidents and hourly traffic flow**. In: *Accident Analysis & Prevention*, v. 14, Issue 1, p. 19-34, ISSN 0001-4575, 1982.
- CUNTO, F. J. C.; DE CASTRO NETO, M. M.; BARREIRA, D. S. **Modelos de previsão de acidentes de trânsito em interseções semaforizadas de Fortaleza**. *Transportes*, v. 20, n. 2, p. 55–62, 2012.
- DAVIS, G. A. **Possible aggregation biases in road safety research and a mechanism approach to accident modeling**. v. 36, p. 1119–1127, 2004.
- ERDOGAN, S. et al. **Geographical information systems aided traffic accident analysis system case study : city of Afyonkarahisar**. v. 40, p. 174–181, 2008.
- FOTHERINGHAM, A. S.; WONG, D. W. S. (1991). **The modifiable areal unit problem in multivariate statistical analysis**. In: *Environment & Planning A*. 7 ed. Vol. 23, pp. 1025-1044.

- GEHLKE, C. E.; BIEHL, K. **Certain effects of grouping upon the size of the correlation coefficient in census tract material.** Journal of the American Statistical Association. 29 (185A): 169–170. doi:10.2307/2277827, 1934.
- HUANG, H., XU, P. e ABDEL-ATY, M. **Transportation Safety Planning: a Spatial Analysis Approach.** In: TRB 92th Annual Meeting (TRB 2013-1855). 2013.
- JOVANIS, P.P.; CHANG, H.L. **Modeling the relationship of accidents to miles traveled.** Transportation Research Record, p. 42-51, 1986.
- LIMA, M. G., MOREIRA, R. M, GOMES, M. J. T. L, CUNTO, F. J. **Consolidação de base de dados georreferenciada para a modelagem espacial do desempenho da segurança viária.** In: Anais do XXIX ANPET - Congresso de Pesquisa e Ensino em Transportes. Ouro Preto, MG. 2015.
- LIMA, M. G., CUNTO, F. J. **Avaliação da dependência espacial da taxa de acidentes em zonas de tráfego com modelos espaciais globais e locais** In: Anais do XXX ANPET - Congresso de Pesquisa e Ensino em Transportes. Rio de Janeiro, RJ. 2016.
- GOMES, M. J. T. L. **Modelagem do desempenho da segurança viária com a incorporação dos efeitos espaciais.** 2017.
- MOUNTAIN, L.; JARRETT, D. **Accident prediction models for roads with minor junctions.** v. 28, n. 6, p. 695–707, 1996.
- OMS - Organização Mundial da Saúde. **Global Status Report on Road Safety 2015. WHO Library Cataloguing-in-Publication Data Global,** p. 340, 2015.
- QUEIROZ, M. P. **Análise espacial dos acidentes de trânsito do município de Fortaleza. Dissertação de Mestrado** - Departamento de Engenharia de Transportes. Universidade Federal do Ceará, 2003.
- REGÔ, C. V.; PENA, M. G. **Análise dos modelos de regressão espacial SAR, SEM e SAC.** 2012. 76 f., il. Monografia (Bacharelado em Estatística). Universidade de Brasília, Brasília, 2012.
- SANTOS, L.; RAIÁ JÚNIOR, A. A. **Distribuição espacial dos acidentes de trânsito em São Carlos (SP):** Identificação de tendências de deslocamento através da técnica de elipse de desvio padrão. Caminhos de Geografia, [S.l.], v. 7, 2006.
- SAWALHA, Z.; SAYED, T. **Traffic accident modeling: Some statistical issues.** Canadian Journal of Civil Engineering, p. 33, 2011.
- SIDDIQUI, C.; ABDEL-ATY, M.; CHOI, K. **Macroscopic spatial analysis of pedestrian and bicycle crashes. Accident Analysis and Prevention,** v. 45, p. 382–391, 2012.
- SILVA, A. R.; **Avaliação de modelos de regressão espacial para análise de cenários do transporte rodoviário.** 2006. 138 f. Dissertação (Mestrado em Transportes). Universidade de Brasília, Brasília, 2006.

SOBREIRA, L. T. P. **Análise Exploratória de Atropelamentos em Zonas de Tráfego de Fortaleza**. 2016.

THOMAS, I.; **Spatial data aggregation**: Exploratory analysis of road accidents. In: *Accident Analysis & Prevention*, v. 28, p. 251-264, 1996.

WANG, C.; QUDDUS, M.A.; RYLEY, T.; ENOCH, M.; DAVIDSON, L. Spatial models in transport: a review and assessment of methodological issues. **91st Annual Meeting of the Transportation Research Board**, v. 750, n. November, 2011.

WONG, D. W. S. The Modifiable Areal Unit Problem (MAUP). **WorldMinds: Geographical Perspectives on 100 Problems**, p. 571–575, 2004.

XU, P. et al. Sensitivity analysis in the context of regional safety modeling : Identifying and assessing the modifiable areal unit problem. **Accident Analysis and Prevention**, v. 70, p. 110–120, 2014.

XU, P.; HUANG, H.; DONG, N. The modifiable areal unit problem in traffic safety: Basic issue, potential solutions and future research. **Journal of Traffic and Transportation Engineering (English Edition)**, n. December, p. 1–10, 2016.

APÊNDICE A – DIAGRAMAS DE DISPERSÃO DE MORAN

Diagrama de dispersão de Moran Univariado para a variável Taxa de Acidentes totais por quilômetro de via na configuração 1

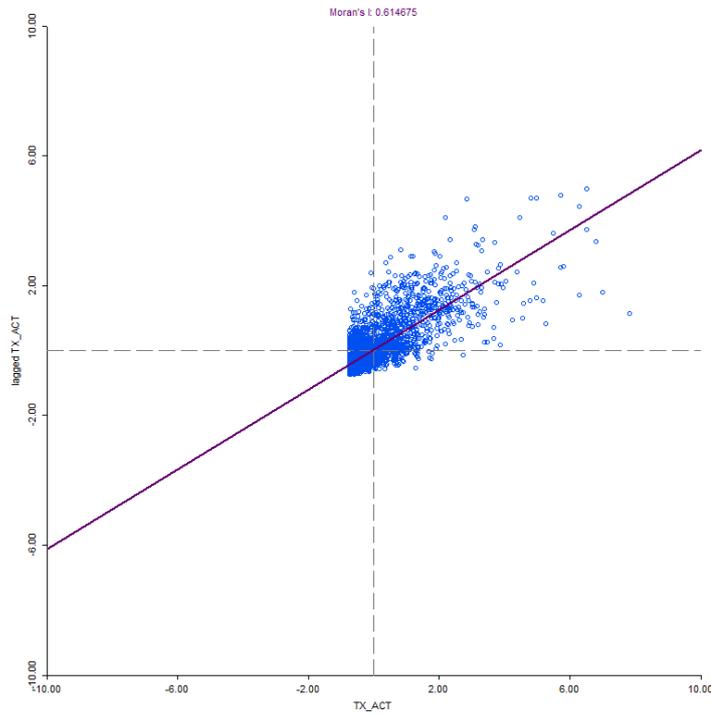


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e densidade de interseções semaforizadas na configuração 1

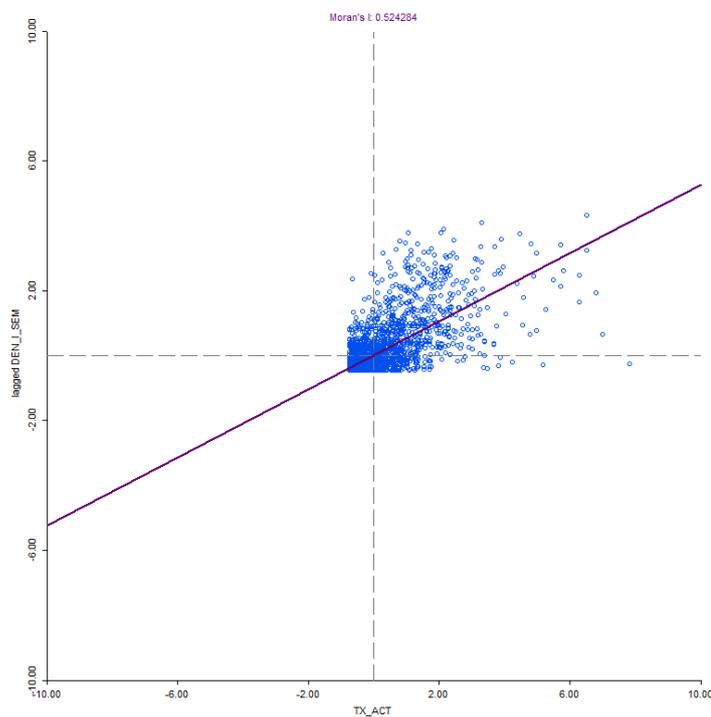


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e proporção de residentes maior que 64 anos na configuração 1

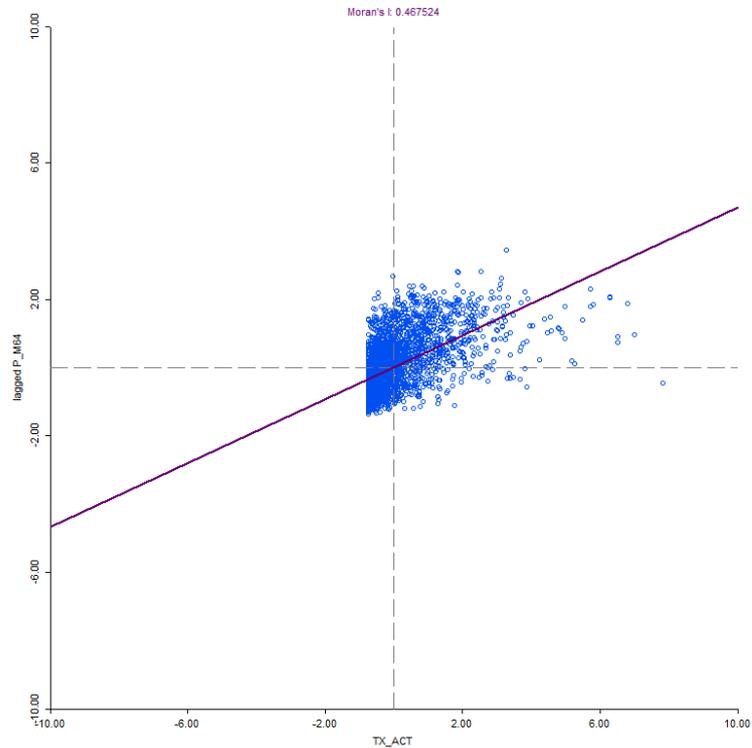


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e proporção de domicílios com renda familiar maior que 3 salários mínimos na configuração 1

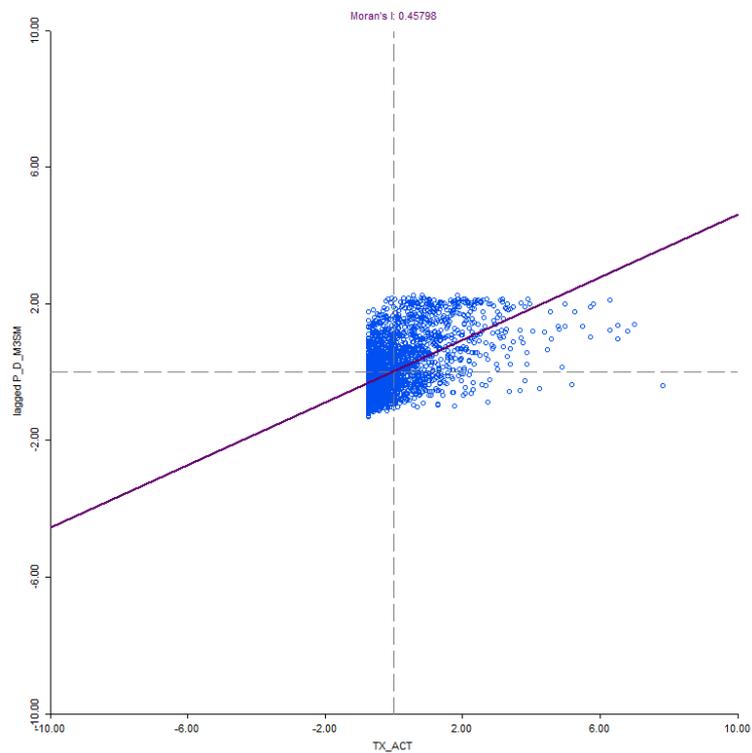


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e uso do solo comercial por metro quadrado de zona na configuração 1

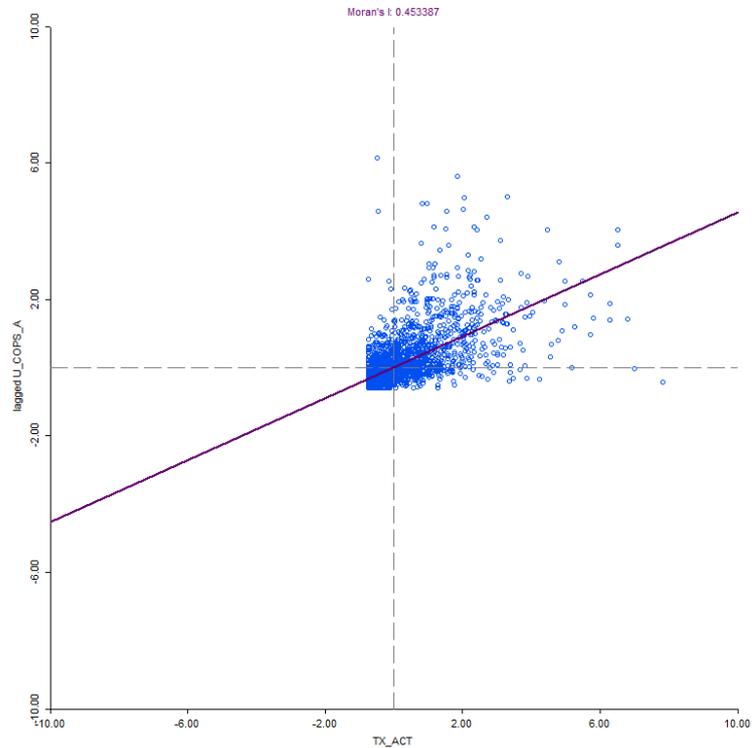


Diagrama de dispersão de Moran Univariado para a variável Taxa de Acidentes totais por quilômetro de via na configuração 2

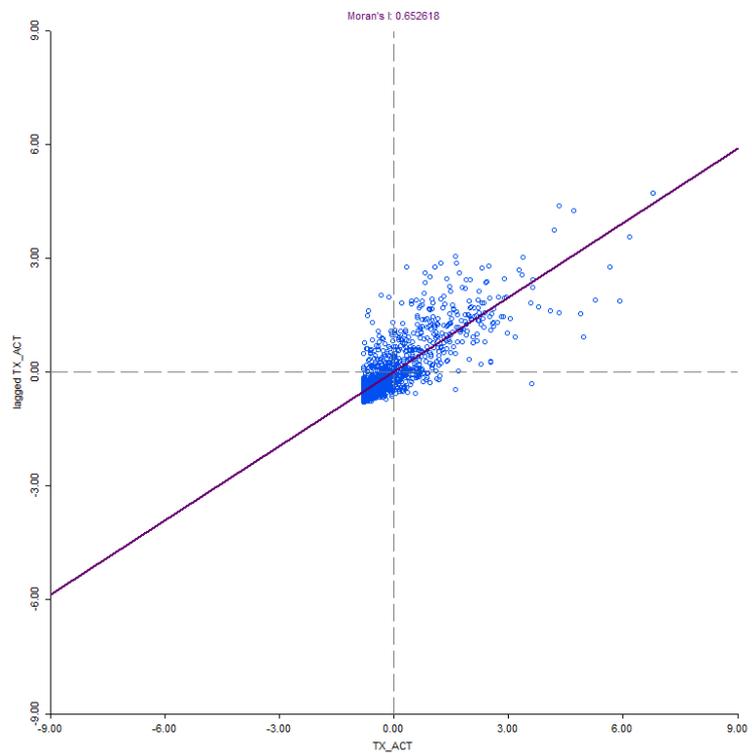


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e densidade de interseções semaforizadas na configuração 2

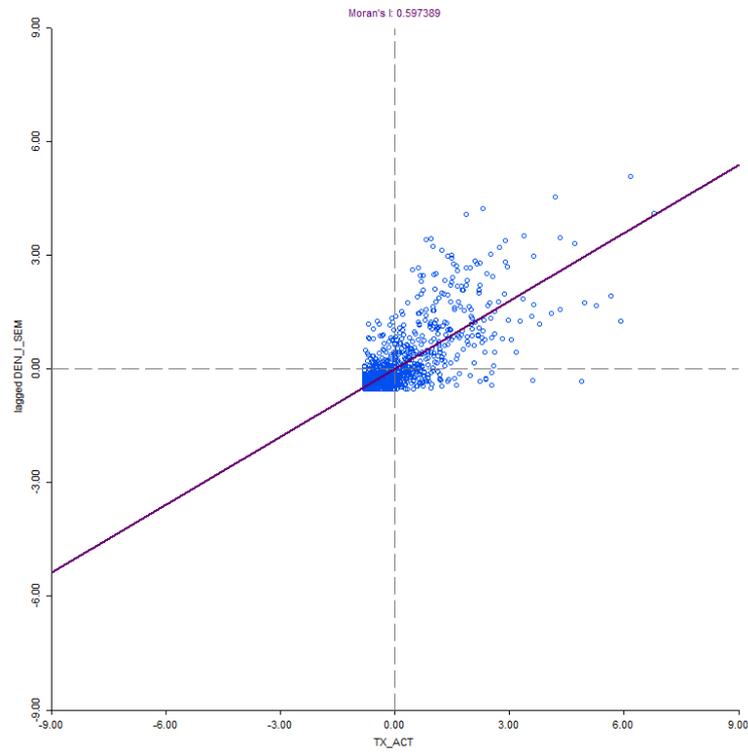


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e proporção de residentes maior que 64 anos na configuração 2

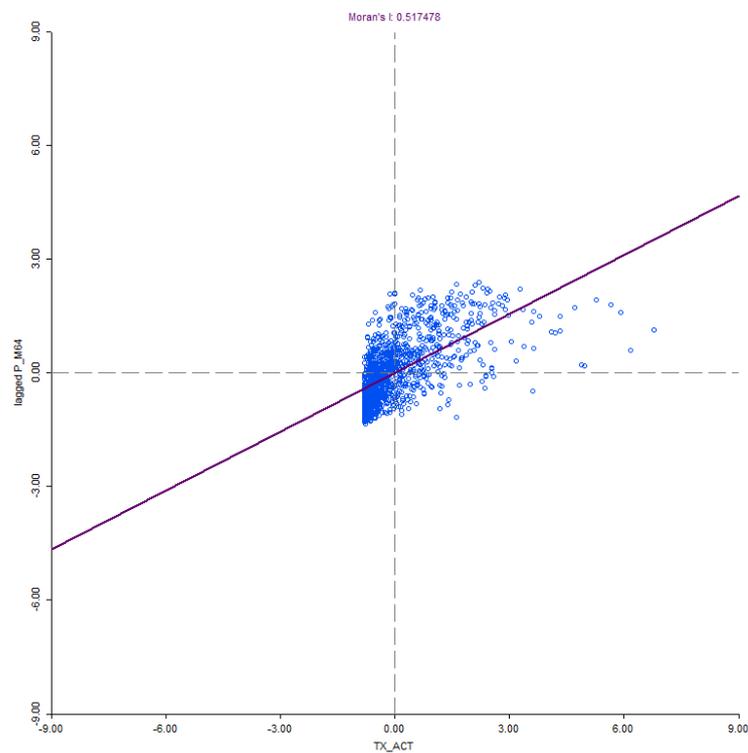


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e proporção de domicílios com renda familiar maior que 3 salários mínimos na configuração 2

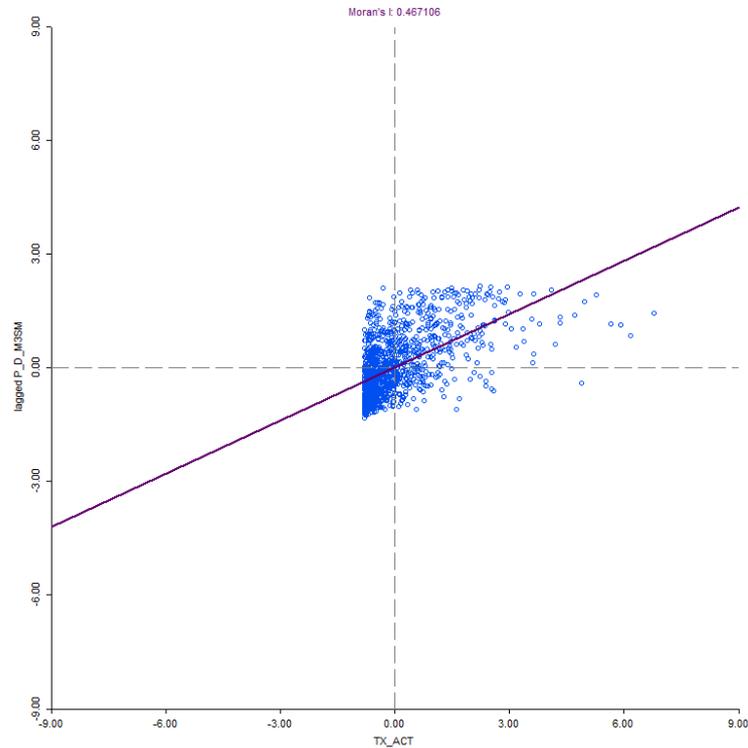


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e uso do solo comercial por metro quadrado de zona na configuração 2

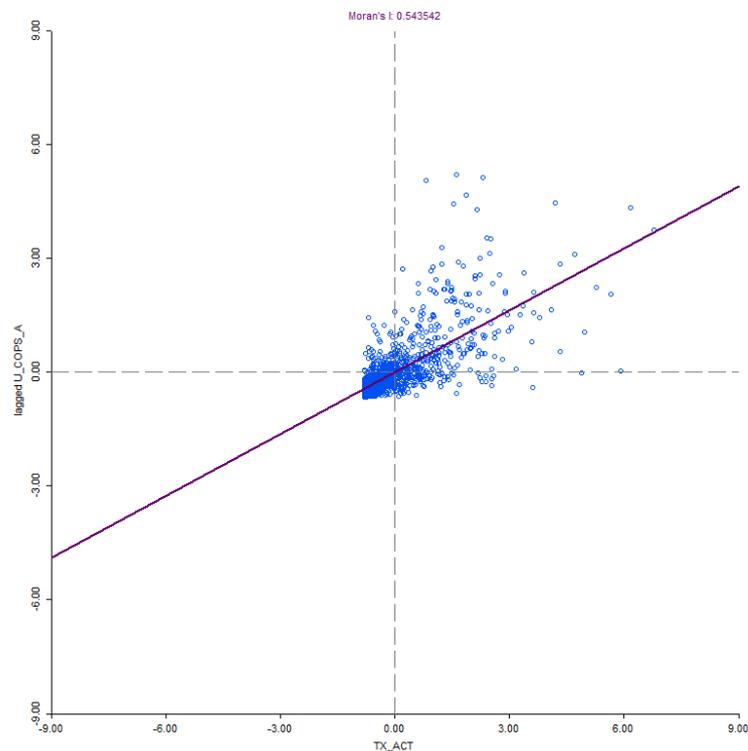


Diagrama de dispersão de Moran Univariado para a variável Taxa de Acidentes totais por quilômetro de via na configuração 3

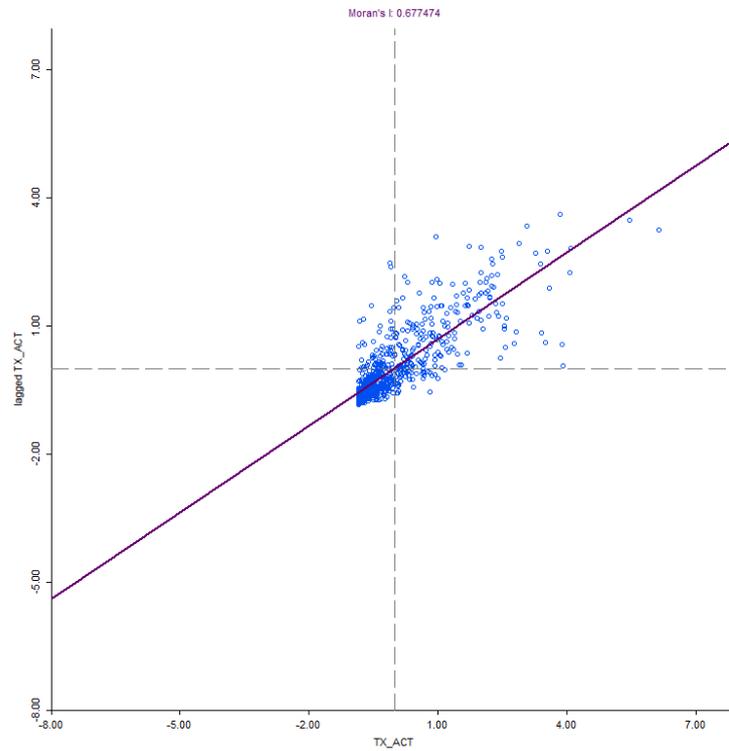


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e densidade de interseções semaforizadas na configuração 3

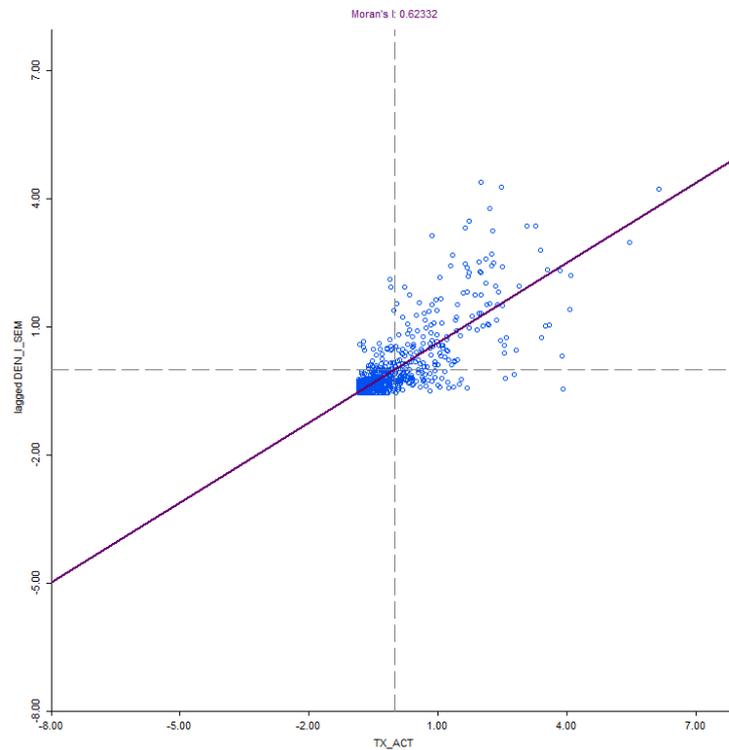


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e proporção de residentes maior que 64 anos na configuração 3

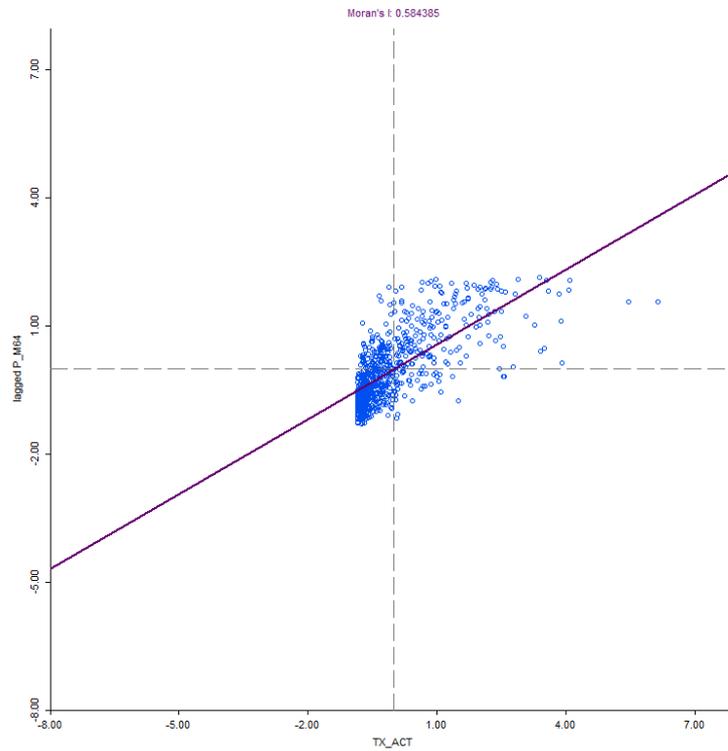


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e proporção de domicílios com renda familiar maior que 3 salários mínimos na configuração 3

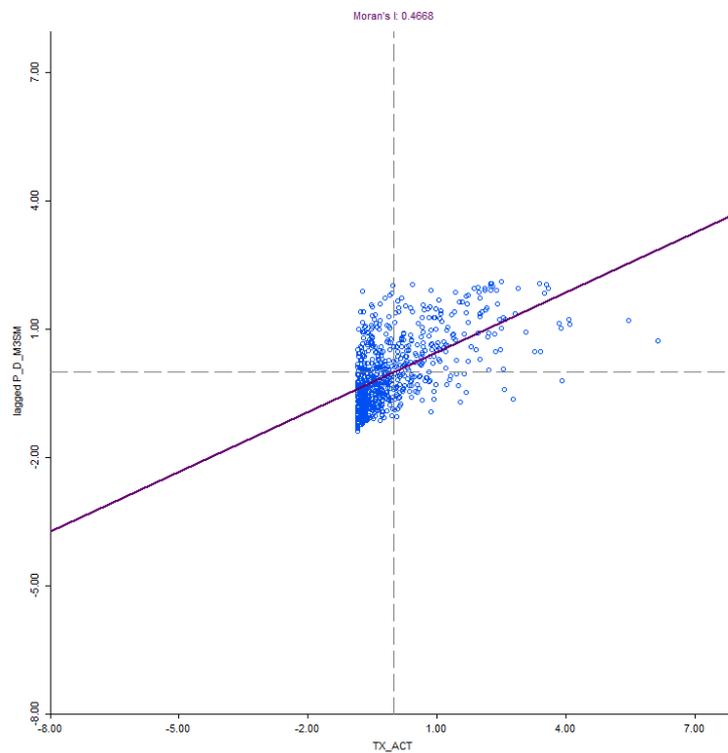


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e uso do solo comercial por metro quadrado de zona na configuração 3

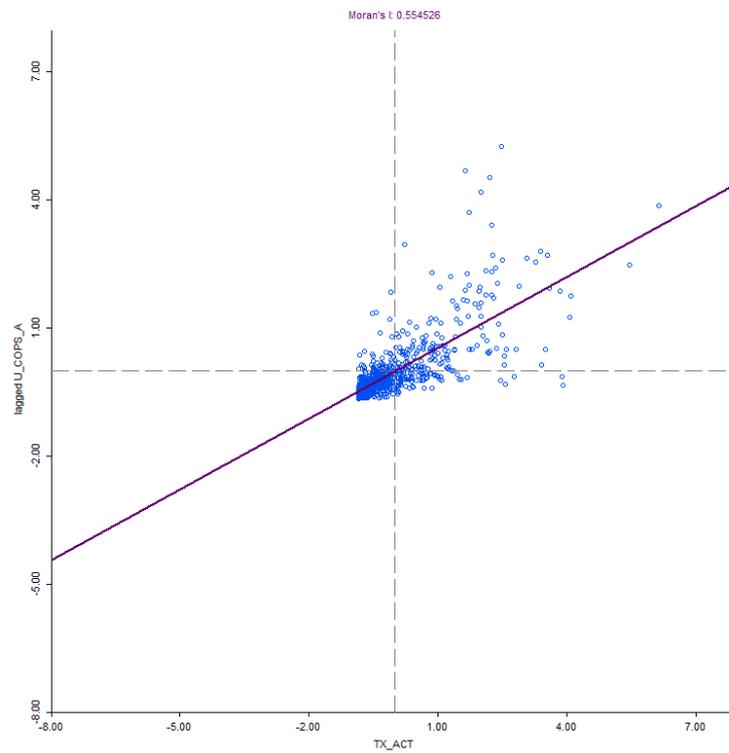


Diagrama de dispersão de Moran Univariado para a variável Taxa de Acidentes totais por quilômetro de via na configuração 4

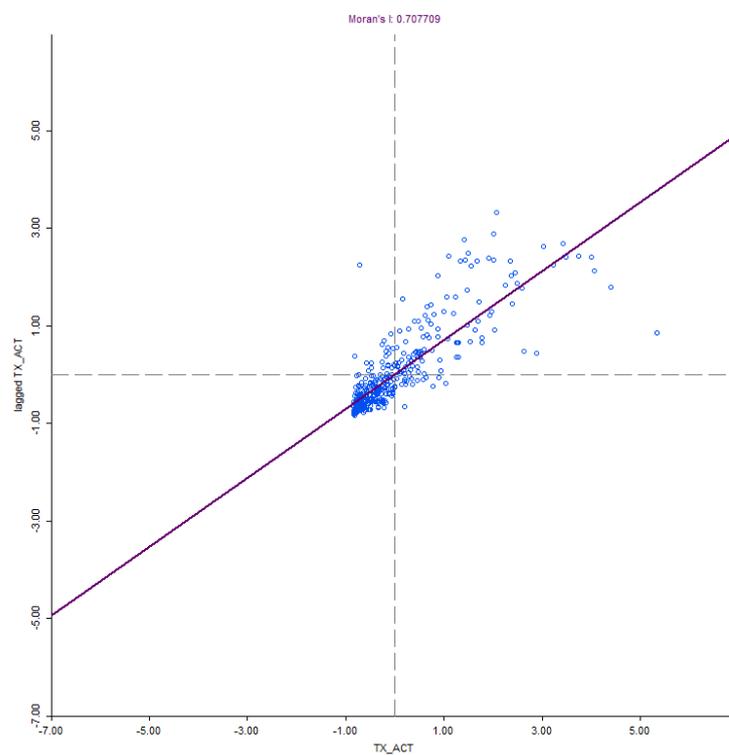


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e densidade de interseções semaforizadas na configuração 4

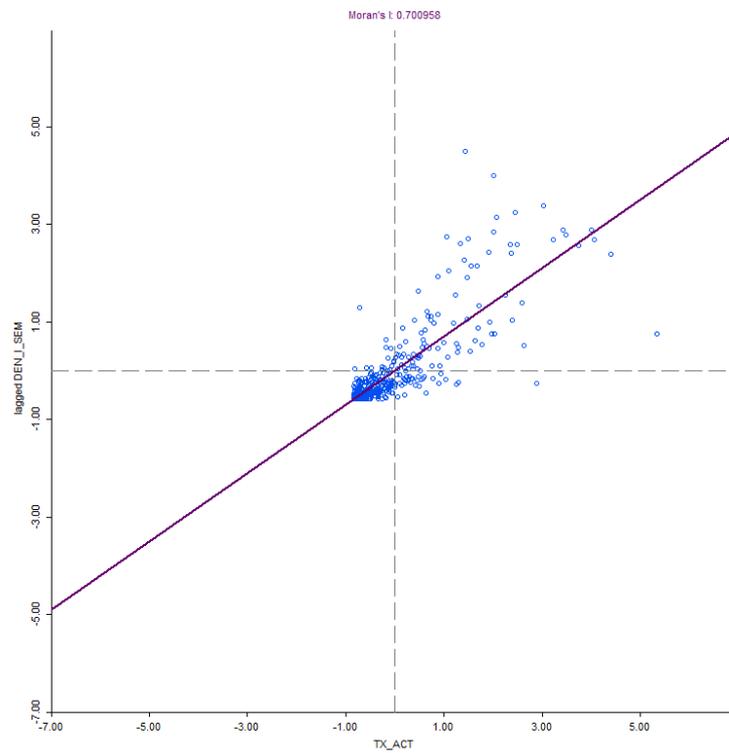


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e proporção de residentes maior que 64 anos na configuração 4

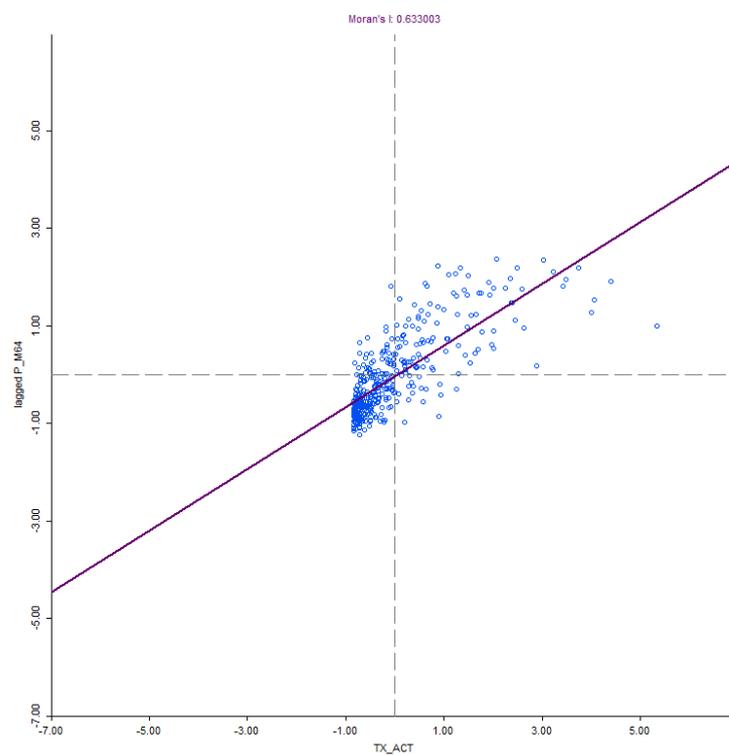


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e proporção de domicílios com renda familiar maior que 3 salários mínimos na configuração 4

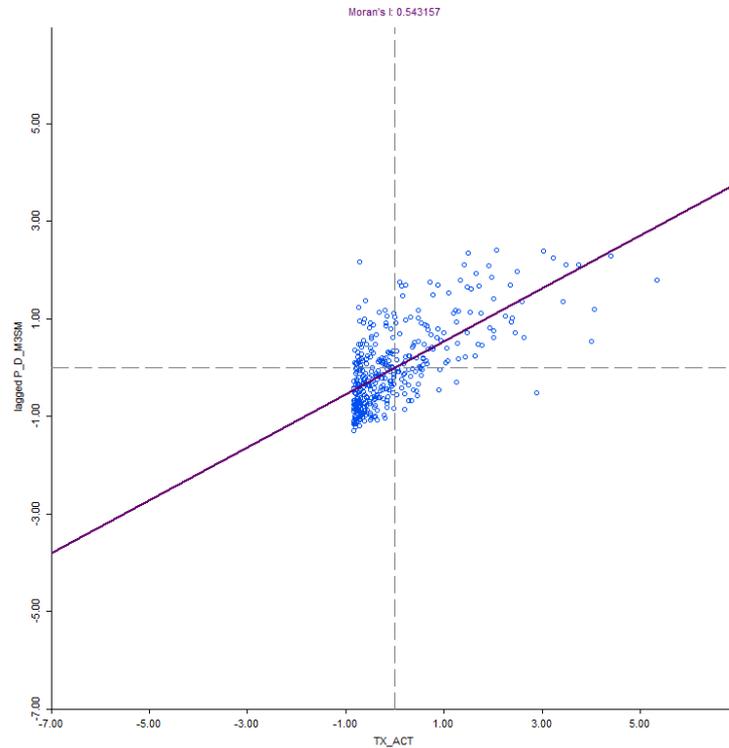


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e uso do solo comercial por metro quadrado de zona na configuração 4

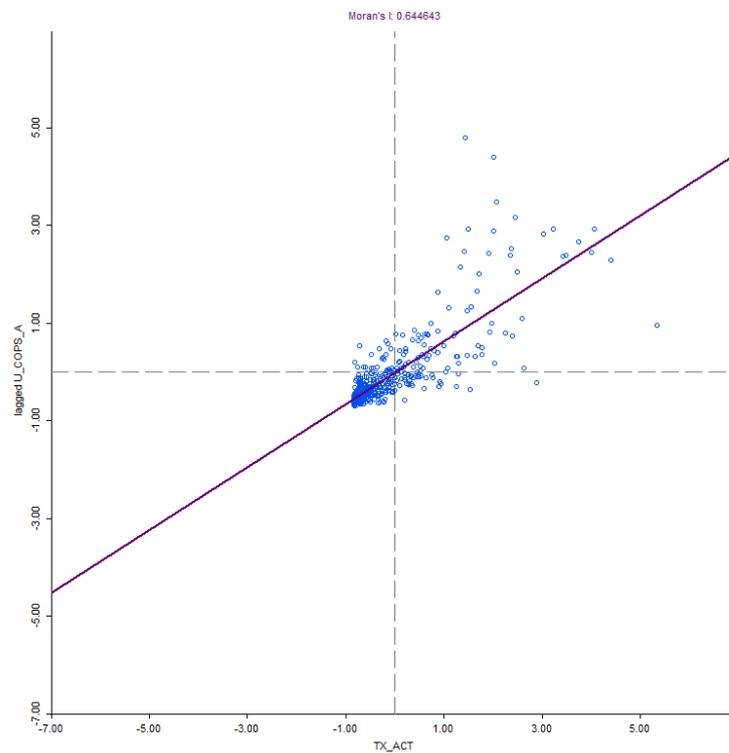


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e densidade de interseções semaforizadas na configuração 5

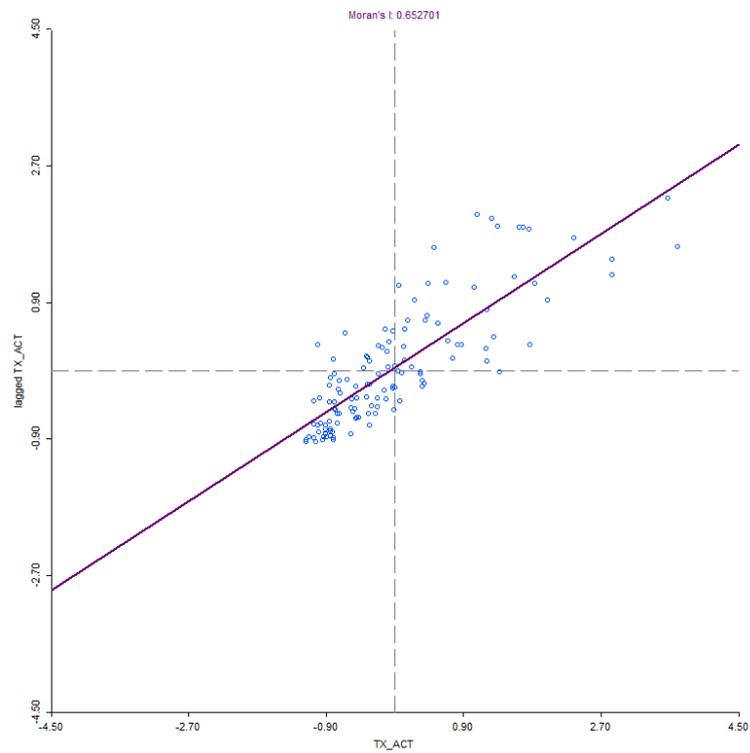


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e proporção de residentes maior que 64 anos na configuração 5

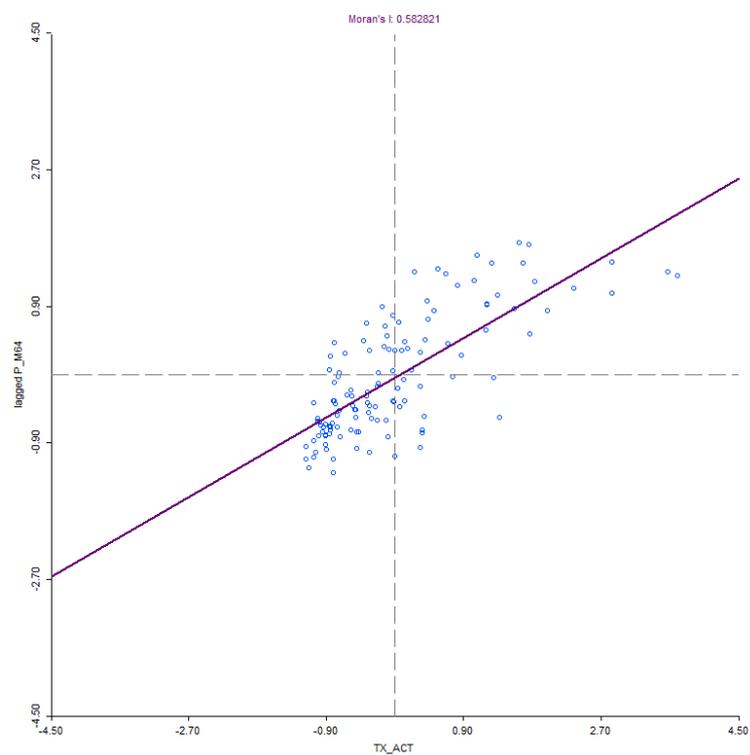


Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e proporção de domicílios com renda familiar maior que 3 salários mínimos na configuração 5

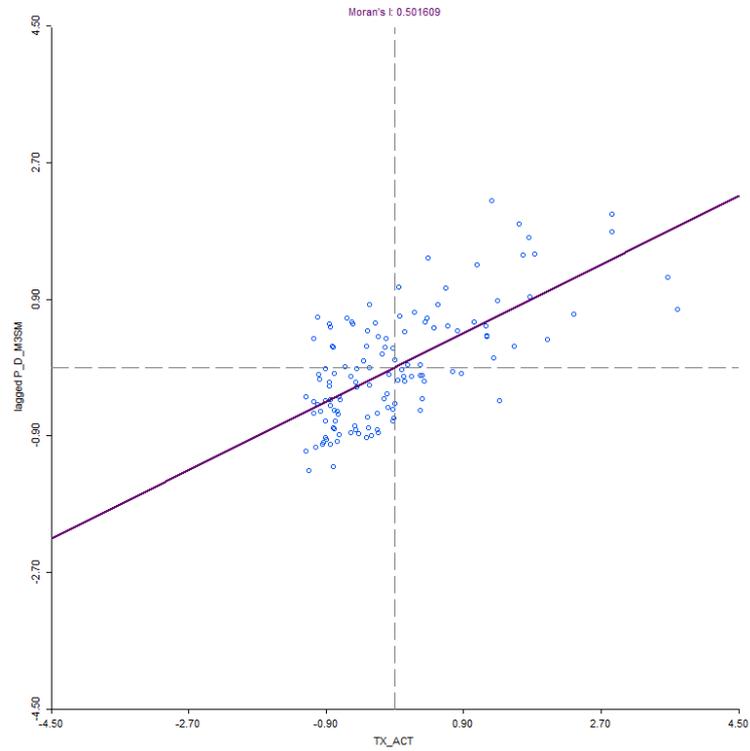
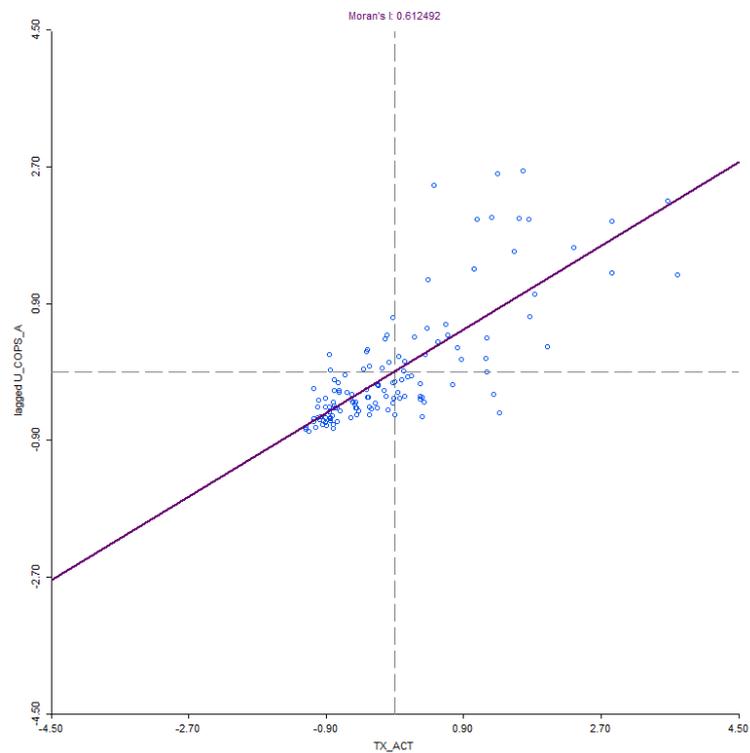


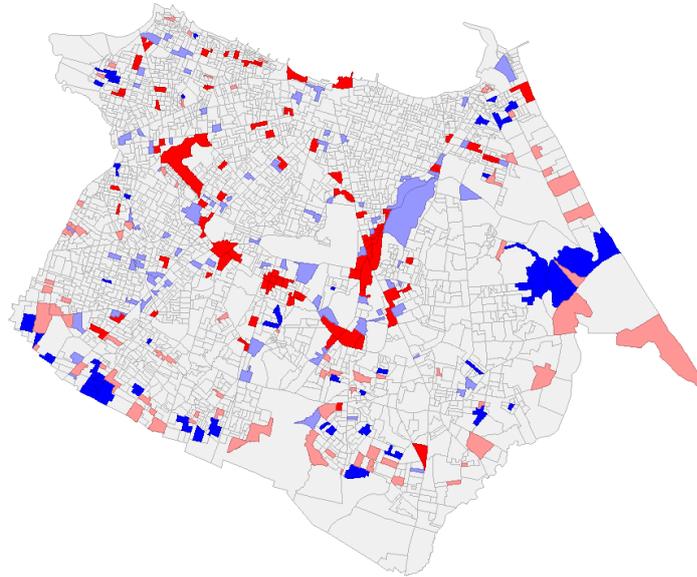
Diagrama de dispersão de Moran Bivariado para a variável Taxa de Acidentes totais por quilômetro de via e uso do solo comercial por metro quadrado de zona na configuração 5



APÊNDICE B – MAPA DE MORAN DOS RESÍDUOS DOS MODELOS DE REGRESSÃO ESPACIAL

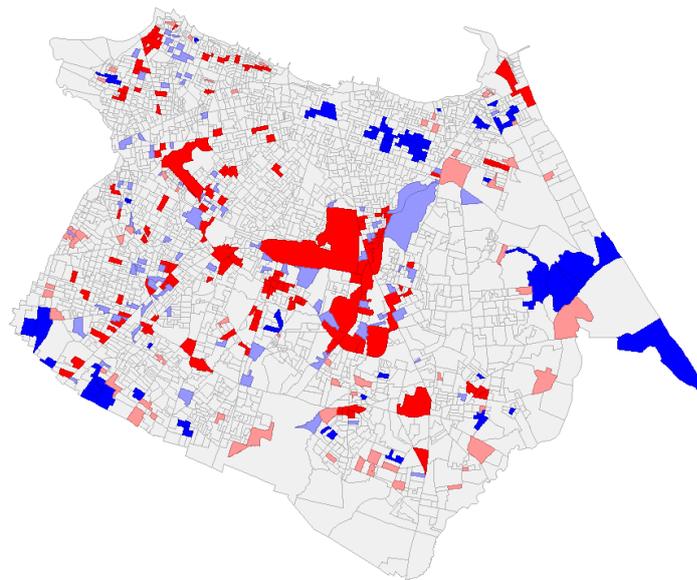
Mapa de Moran do modelo SEM para a Configuração 1

LISA Cluster Map
□ Not Significant (2628)
■ High-High (86)
■ Low-Low (55)
■ Low-High (86)
■ High-Low (73)



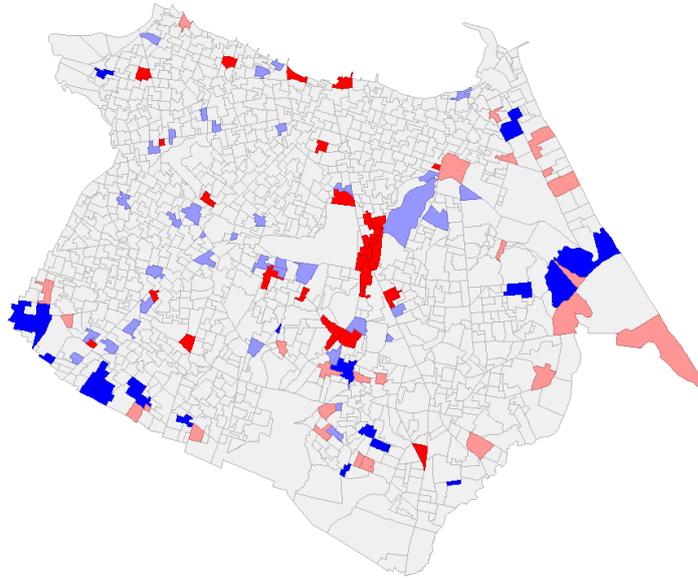
Mapa de Moran do modelo SAR para a Configuração 1

LISA Cluster Map
□ Not Significant (2577)
■ High-High (134)
■ Low-Low (67)
■ Low-High (91)
■ High-Low (59)



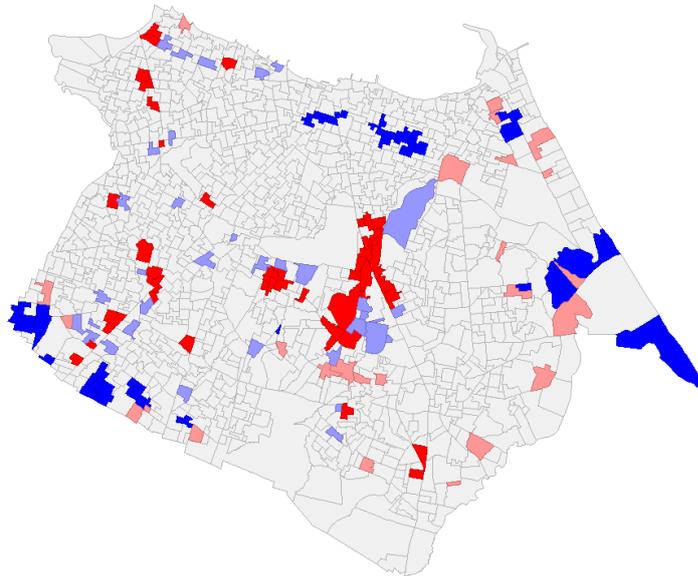
Mapa de Moran do modelo SEM para a Configuração 2

LISA Cluster Map
□ Not Significant (1152)
■ High-High (30)
■ Low-Low (26)
■ Low-High (38)
■ High-Low (26)



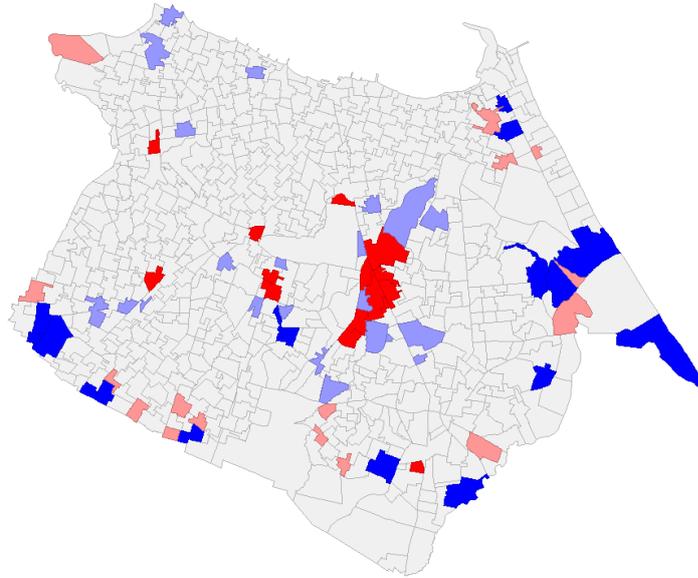
Mapa de Moran do modelo SAR para a Configuração 2

LISA Cluster Map
□ Not Significant (1137)
■ High-High (46)
■ Low-Low (29)
■ Low-High (35)
■ High-Low (25)



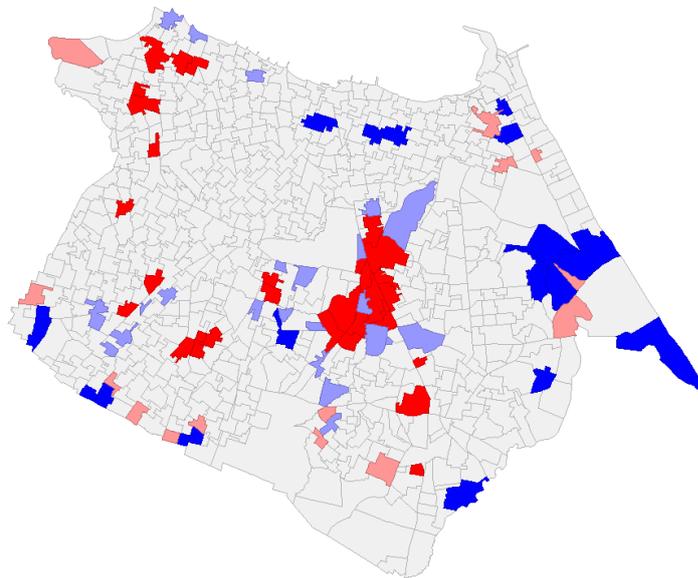
Mapa de Moran do modelo SEM para a Configuração 3

LISA Cluster Map
□ Not Significant (646)
■ High-High (15)
■ Low-Low (17)
■ Low-High (22)
■ High-Low (16)



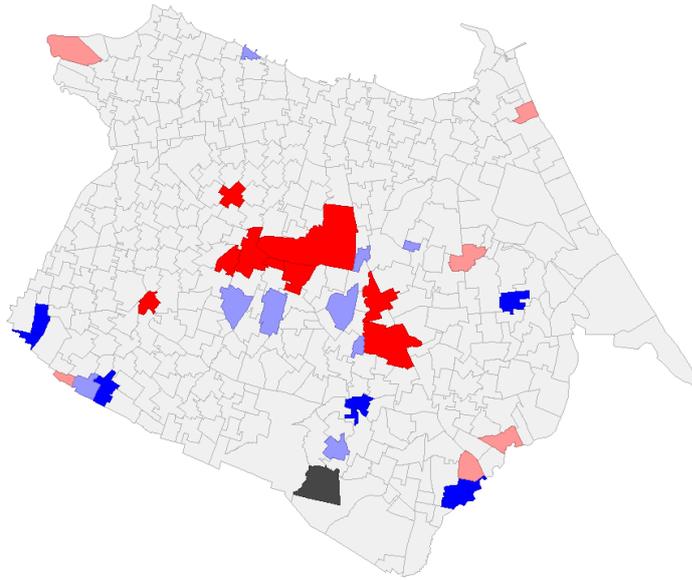
Mapa de Moran do modelo SAR para a Configuração 3

LISA Cluster Map
□ Not Significant (629)
■ High-High (31)
■ Low-Low (19)
■ Low-High (23)
■ High-Low (14)



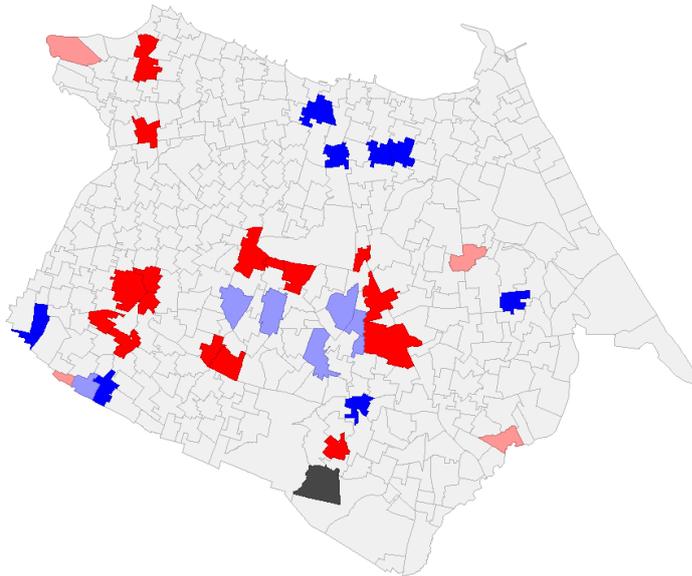
Mapa de Moran do modelo SEM para a Configuração 4

LISA Cluster Map
Not Significant (349)
High-High (9)
Low-Low (6)
Low-High (9)
High-Low (6)
Undefined (1)



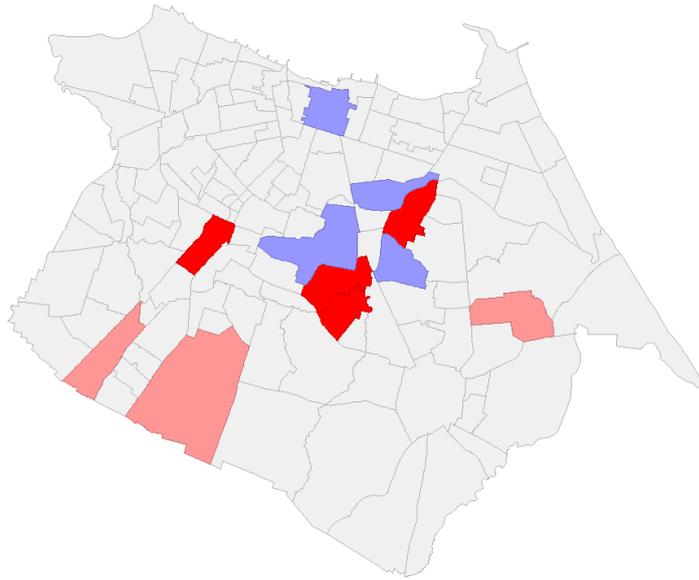
Mapa de Moran do modelo SAR para a Configuração 4

LISA Cluster Map
Not Significant (343)
High-High (16)
Low-Low (9)
Low-High (7)
High-Low (4)
Undefined (1)



Mapa de Moran do modelo SEM para a Configuração 5

LISA Cluster Map
Not Significant (116)
High-High (4)
Low-Low (0)
Low-High (4)
High-Low (3)



Mapa de Moran do modelo SAR para a Configuração 5

LISA Cluster Map
Not Significant (109)
High-High (10)
Low-Low (0)
Low-High (5)
High-Low (3)

