



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Identificação dos Pontos com Maior Incidência de Acidentes de Trânsito Fatais no Distrito Federal

Bianca Nogueira Agapito

10/0008283

Brasília

2014

Bianca Nogueira Agapito

10/0008283

Identificação dos Pontos com Maior Incidência de Acidentes de Trânsito Fatais no Distrito Federal

Relatório apresentado à disciplina Estágio Supervisionado II do curso de graduação em Estatística, Departamento de Estatística, Instituto de Exatas, Universidade de Brasília, como parte dos requisitos necessários para o grau de Bacharel em Estatística.

Orientador: Prof. Dr. Alan Ricardo da Silva

Brasília

2014

Agradecimentos

Agradeço primeiramente aos meus pais pelo apoio e incentivo. Agradeço ao meu orientador por toda a empolgação com o trabalho, todas as ideias e o suporte essencial ao desenvolvimento, estando sempre presente para ajudar, e principalmente por ter me apresentado a área de estatística espacial, que hoje considero a área mais interessante do curso de estatística. Agradeço aos meus colegas do DETRAN pela ideia de tema para este trabalho, por toda a ajuda na consistência do banco de dados, e por terem tornado a segurança no trânsito um assunto de tamanho interesse para mim. Agradeço aos amigos que fiz ao longo do curso, por todo o apoio e incentivo nos momentos de desespero, pelas madrugadas, finais de semana e feriados estudando juntos e pelas festas e viagens inesquecíveis. Cada um ajudou de um jeito, mas todos foram essenciais para a conclusão desse trabalho.

*“Até mesmo o acaso não é
impenetrável, tem suas próprias
regras”*

Novalis, F.

Resumo

A violência no trânsito é um tema de interesse geral da população, pois todos estão inseridos de alguma forma nesse meio, seja como condutor, pedestre ou passageiro. Neste trabalho, o objetivo foi tentar localizar locais com maior risco de ocorrência de acidentes de trânsito fatais nas Regiões Administrativas de Brasília, Ceilândia e Taguatinga.

Foram utilizadas as técnicas do Estimador de Intensidade modificado por Zheng et al. (2011) e a Estatística *Scan* de Kulldorff (1997). Os resultados dos métodos não foram muito satisfatórios para os dados utilizados, pois apesar de apontarem para a existência de aglomerados de acidentes, a maior parte deles não foi estatisticamente significativa. Apenas dois *clusters* significativos foram identificados, quando agrupados por áreas, sendo um em Brasília e um em Taguatinga. Uma sugestão para uma melhor análise futura seria a utilização de dados de acidentes não fatais como controle dos acidentes fatais para a Estatística *Scan*.

Palavras-chaves: Espacial, *Cluster*, *Scan*, Kernel, Acidentes, Trânsito

Lista de Tabelas

4.1	Acidentes de Trânsito com Morte por Região Administrativa	20
5.1	Acidentes de Trânsito com Morte por RA e Existência de Fiscalização Eletrônica	26
5.2	Lambda (λ) e P-valor do maior cluster de cada RA	29

Lista de Figuras

2.1	Fluxo de Informações	4
2.2	Cronologia 2011	6
2.3	Acidentes com Morte por Região Administrativa, 2011-2013	7
2.4	Acidentes Fatais 2011	9
2.5	Acidentes Fatais 2012	10
3.1	Exemplo de uso da Estatística <i>Scan</i>	14
3.2	Exemplo de uso da Regressão Geograficamente Ponderada	18
5.1	Acidentes Fatais em Brasília 2011 - 2013	23
5.2	Acidentes Fatais em Ceilândia e Taguatinga 2011 - 2013	24
5.3	Principais Vias de Brasília	24
5.4	Principais Vias de Ceilândia e Taguatinga	25
5.5	Acidentes Fatais por Existência de Fiscalização Eletrônica e Setores	27
5.6	Acidentes Fatais Finais por Setores	28
5.7	Maiores <i>Clusters</i> em cada RA	30
5.8	Acidentes com Morte por Veículos Envolvidos	31
5.9	Análise Pontual	32
5.10	Estimador de Intensidade Original	34
5.11	Vizinho Mais Próximo	35
5.12	Envelope de Simulação	35
5.13	Estimador de Intensidade Modificado	36
6.1	Acidentes com Morte por Existência de Dados de Fiscalização Eletrônica	39

Sumário

Resumo	iv
1 INTRODUÇÃO	1
1.1 OBJETIVOS	2
2 ESTATÍSTICAS DE ACIDENTES DE TRÂNSITO	3
2.1 INTRODUÇÃO	3
2.1.1 ORIGEM DOS DADOS	3
2.2 ACIDENTES DE TRÂNSITO	5
2.2.1 ACIDENTES DE TRÂNSITO FATAIS	6
2.2.2 GEORREFERENCIAMENTO	8
3 ESTATÍSTICA SCAN E ESTIMADOR DE INTENSIDADE	11
3.1 INTRODUÇÃO	11
3.2 ESTATÍSTICA <i>SCAN</i>	11
3.3 Estimador de Intensidade	14
4 MATERIAL E MÉTODOS	19
4.1 INTRODUÇÃO	19
4.2 BANCO DE DADOS DE ACIDENTES DE TRÂNSITO	19

4.2.1	Análise Descritiva dos Dados	20
4.3	MÉTODOS	21
5	ANÁLISE DOS RESULTADOS	22
5.1	INTRODUÇÃO	22
5.2	MAPAS	22
5.3	Estatística <i>Scan</i>	25
5.3.1	Análise por Área	25
5.3.2	Análise Pontual	31
5.4	Estimador de Intensidade	33
5.4.1	Estimador de Intensidade Original de Kernel	33
5.4.2	Estimador de Intensidade Modificado de Zheng et al. (2011) .	35
6	CONCLUSÕES	37
6.1	PROBLEMAS OBSERVADOS	38
	REFERÊNCIAS	41

Capítulo 1

INTRODUÇÃO

Atualmente, os acidentes de trânsito podem ser considerados um problema de segurança pública no Brasil, pois em alguns estados o número de óbitos no trânsito chega a ser maior que o de homicídios (MS, 2012). Esse é um dado preocupante, já que segundo o IPEA (2013) cada acidente com morte pode custar até R\$ 800 mil aos cofres públicos. Esse valor, no entanto, não traduz o custo de uma vida.

De acordo com a ABNT (1989), acidente de trânsito é definido como: “Todo evento não premeditado de que resulte dano em veículo ou na sua carga e/ou lesões em pessoas e/ou animais, em que pelo menos uma das partes está em movimento nas vias terrestres ou áreas abertas ao público. Pode originar-se, terminar ou envolver veículo parcialmente na via pública.”

Os acidentes de trânsito com vítimas são classificados como acidentes com feridos e acidentes com morte, sendo este último todo aquele em que ocorra a morte de, pelo menos, uma pessoa até 30 dias após a data do acidente. Portanto, é possível que em um acidente com morte também existam feridos (GEREST, 2012a).

O Distrito Federal tem um dos menores índices de mortos por 10.000 veículos do Brasil, tendo fechado o ano de 2013 com o valor 2,6 (GEREST, 2013). Esse índice

foi criado pela ONU e o valor recomendado para países em desenvolvimento é de 3, que significa três mortos em acidentes de trânsito para cada dez mil veículos em circulação (DETRAN, 2012). Esse dado leva em consideração o total de veículos registrados no Distrito Federal. Entretanto, a mesma ideia pode ser utilizada para comparar acidentes de trânsito com o fluxo de veículos da via no momento em que o evento ocorreu. Dessa forma, é possível testar a existência de *clusters*, ou aglomerados de acidentes, que indicariam os pontos mais perigosos em relação à esse evento (Kulldorff, 1997).

O foco deste trabalho será analisar as regiões administrativas (RAs) do DF com mais acidentes de trânsito nos anos de 2011 a 2013, ou seja, Brasília, Ceilândia e Taguatinga, e testar a existência de *clusters* nessas regiões.

1.1 OBJETIVOS

O objetivo geral do trabalho é indicar os setores com maior incidência de acidentes de trânsito fatais no Distrito Federal, utilizando análise espacial.

Os objetivos específicos são:

- apresentar os acidentes de trânsito fatais, no Distrito Federal, de 2011 a 2013, em mapas, utilizando o *software* SAS 9.2;
- implementar o algoritmo de Kulldorff para análise espacial e comparar com o Estimador de Intensidade usado por Zheng et al. (2011).

Capítulo 2

ESTATÍSTICAS DE ACIDENTES DE TRÂNSITO

2.1 INTRODUÇÃO

Os dados de acidentes de trânsito constam no banco de dados do Departamento de Trânsito do Distrito Federal (DETRAN/DF). Esses dados são oriundos, basicamente, de três fontes oficiais: Polícia Civil, Instituto Médico Legal (IML) e Secretaria de Saúde.

2.1.1 ORIGEM DOS DADOS

A Polícia Civil do DF é responsável pelo preenchimento das ocorrências de trânsito. Normalmente o acidente é registrado na delegacia mais próxima do local do fato. É a fonte com a maior concentração de informações e, também a mais completa. Do IML são obtidas as informações complementares sobre os mortos no trânsito. Semanalmente é realizada uma listagem nominal de todas as vítimas mortas, que serve como base para a checagem do número de óbitos e para a identificação das vítimas que falecerem após a data do acidente. Uma última compatibilização é feita com o Sistema de Informações sobre Mortalidade (SIM) da Secretaria de Saúde,

por meio das Declarações de Óbito, para complementação de informações pessoais, caso ainda não constem no banco do DETRAN/DF.

Acredita-se que, após o cruzamento das informações dessas fontes, nenhuma vítima fatal deixará de ser considerada. Após consolidação do banco de dados, é feito o tratamento das informações por meio da crítica de acidentes e endereços (GEREST, 2012a). Esse processo está esquematizado na Figura 2.1.

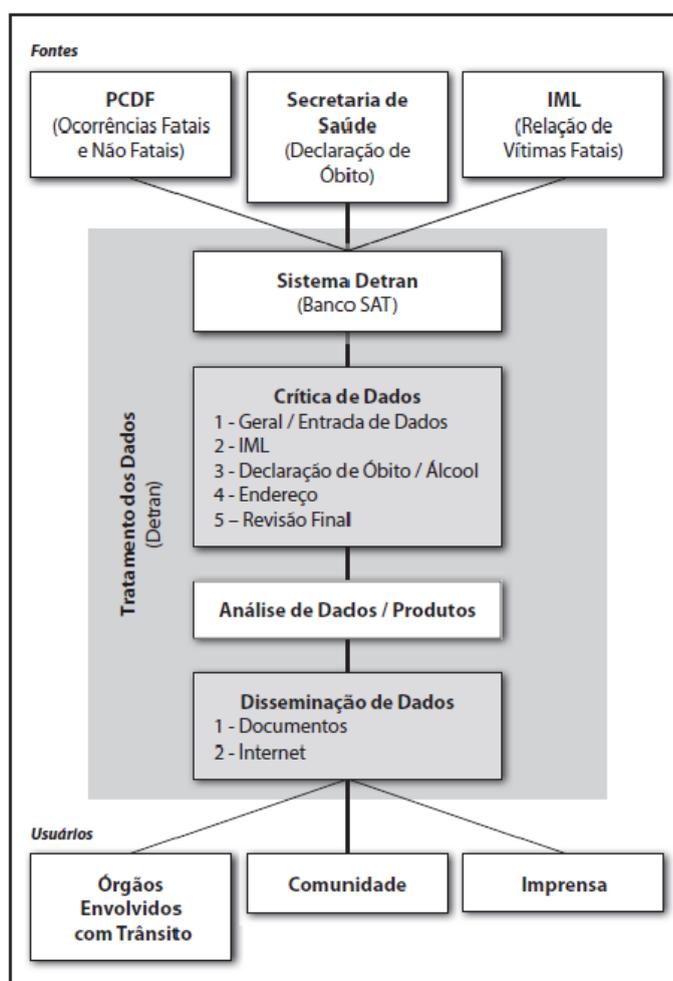


Figura 2.1: Fluxo de Informações
Fonte: GEREST - DETRAN/DF

2.2 ACIDENTES DE TRÂNSITO

Como se pode ver na Figura 2.1, após a crítica dos dados vem a análise e geração de produtos. São feitas análises descritivas dos dados, abordados por vários ângulos. Há 4 bancos de dados principais: acidentes, vítimas, condutores e veículos, cada um com suas variáveis de interesse, e algumas variáveis em comum, como número da delegacia de polícia e número da ocorrência.

Em 2011, foram registrados 9.428 acidentes com vítimas no Distrito Federal, sendo desses 418 com morte e 9.010 com feridos. Esses números são referentes a todo o Distrito Federal, incluindo todos os tipos de vias. No total foram 12.358 vítimas, sendo 465 mortas e 11.893 feridas. Isso dá uma média de 33,9 vítimas por dia, 1,3 mortas em acidentes de trânsito. Algumas informações interessantes referentes à esses dados de 2011 estão na Figura 2.2, retirada, assim como as demais informações acima de GEREST (2012a).

É comum realizar mais análises com dados de acidentes com morte, pois estes causam mais impacto social e econômico do que os acidentes com feridos. Além disso são dados mais confiáveis, já que nem todo acidente com feridos tem a obrigatoriedade de um boletim de ocorrência policial, pois dependendo da gravidade do ferimento, os envolvidos podem não ver a necessidade de contato com a polícia.

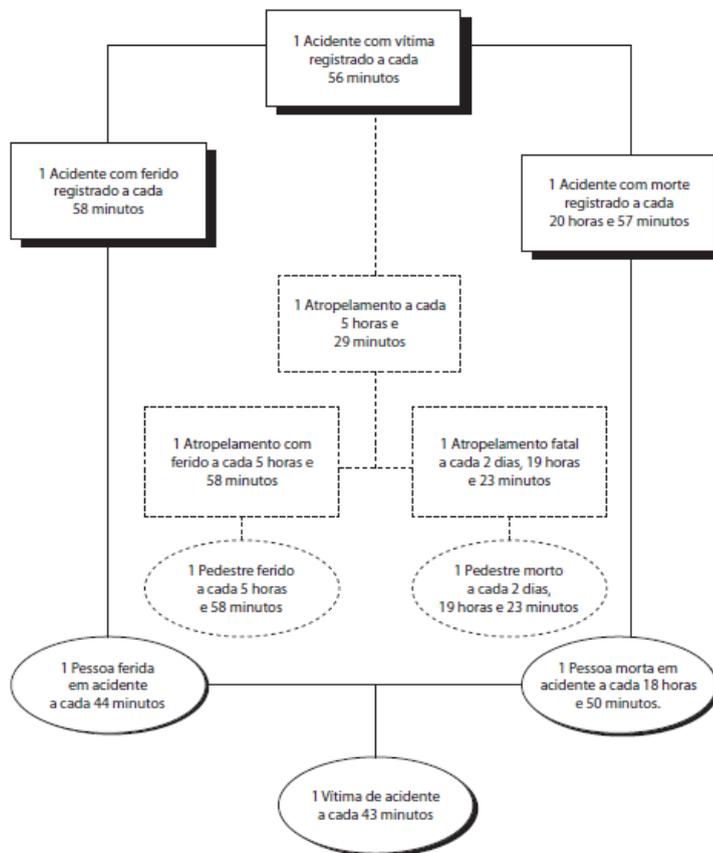


Figura 2.2: Cronologia 2011
 Fonte: GEREST - DETRAN/DF

2.2.1 ACIDENTES DE TRÂNSITO FATAIS

Os acidentes de trânsito com morte são acompanhados semanalmente, devido à sua gravidade, para rápida identificação de fatores de risco corrigíveis.

VIAS URBANAS

Apesar de coletar e analisar dados de todos os tipos de vias do Distrito Federal, o DETRAN/DF atua apenas nas Vias Urbanas, ou seja, todas as vias que não são Rodovias Federais (BRs) ou Rodovias Distritais (DFs). As Vias Urbanas costumam ter velocidade máxima de até 60 km/h e alguns exemplos são: Eixo Monumental

(BSB), W3 (BSB), L2 (BSB), Avenida Hélio Prates (TAG). As estatísticas de acidentes pelas regiões administrativas são feitas usando apenas esse tipo de via, pois as rodovias atravessam várias RAs, e não há uma delimitação estabelecida e aceita de quais trechos pertencem a cada uma delas. A Figura 2.3 mostra a quantidade de acidentes com morte acumulada de 2011 a 2013, para cada uma das regiões do DF. Percebe-se um destaque para as regiões de Ceilândia, Taguatinga e Brasília, com uma quantidade de acidentes bem superior às demais. Essas três RAs juntas tiveram, entre 2011 e 2013, mais da metade dos acidentes em Vias Urbanas do Distrito Federal. Nesse sentido, são boas regiões para estudo, pois representam uma parcela considerável do total de acidentes fatais.

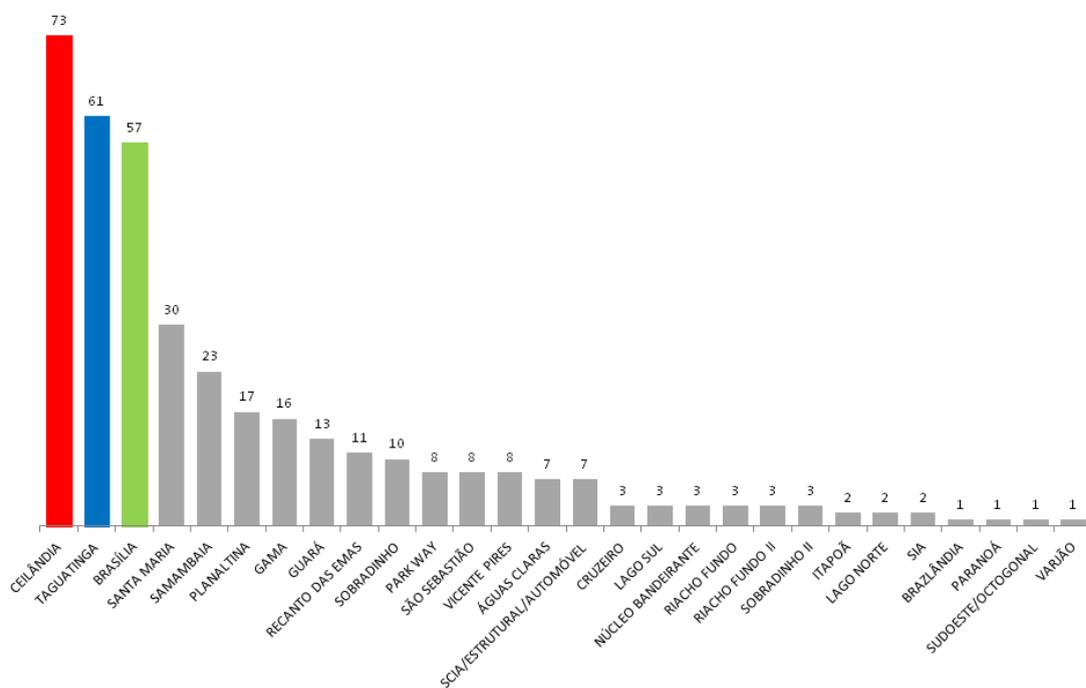


Figura 2.3: Acidentes com Morte por Região Administrativa, 2011-2013

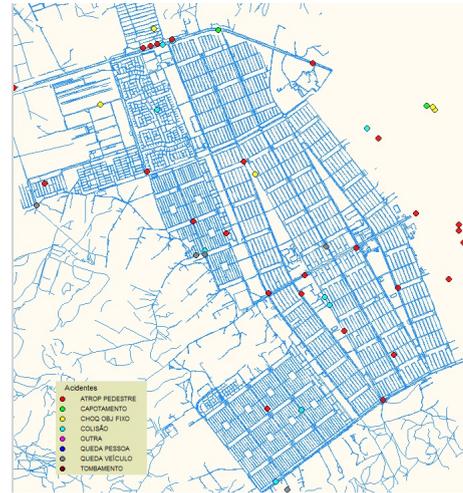
Logo, verifica-se a necessidade de um estudo mais detalhado nessas localidades com maiores índices de acidentes de trânsito, de forma a buscar fatores locais de risco que possam ser corrigidos, evitando futuras mortes no trânsito. Para esse estudo são necessárias as coordenadas geográficas de cada acidente fatal.

2.2.2 GEORREFERENCIAMENTO

Atualmente, as coordenadas geográficas dos locais dos acidentes de trânsito não são coletadas, sendo apenas identificadas por meio dos endereços, que muitas vezes não são precisos. A Central de Investigação e Tratamento de Acidentes de Trânsito (CITA) faz uma pesquisa mais profunda dos acidentes, frequentemente indo aos locais em busca de mais informações e referências. Após esse tratamento os endereços ficam mais precisos, mas ainda sem georreferenciamento. Por conta da falta de coordenadas geográficas, os acidentes fatais são manualmente localizados em um sistema georreferenciado por meio dos endereços, e os mapas são produzidos conforme as Figuras 2.4 e 2.5.



(a) Brasília

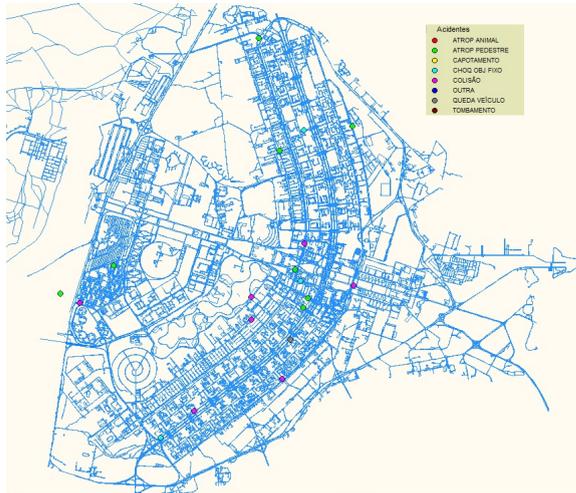


(b) Ceilândia

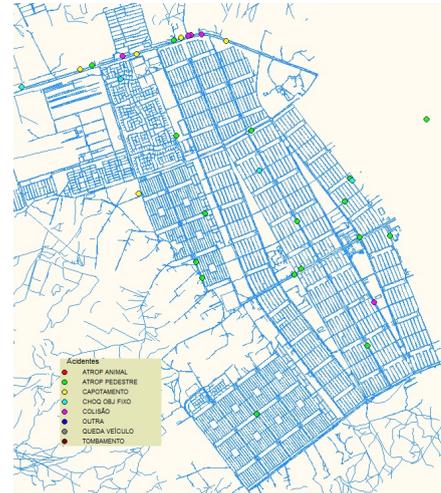


(c) Taguatinga

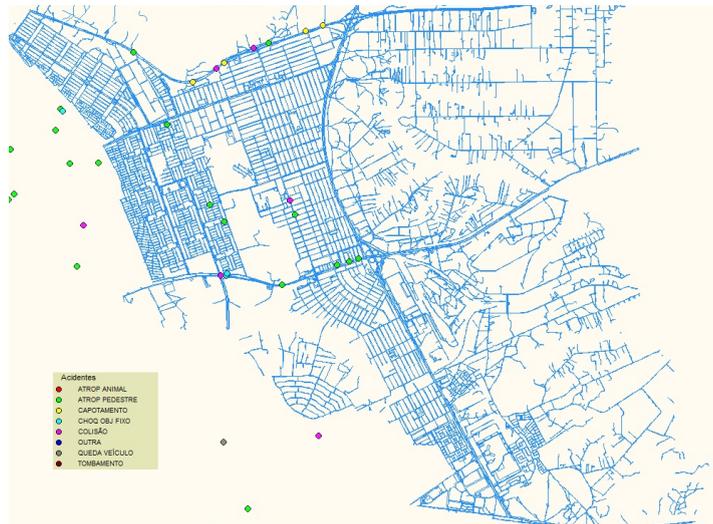
Figura 2.4: Acidentes Fatais 2011
 Fonte: NUGEO - DETRAN/DF



(a) Brasília



(b) Ceilândia



(c) Taguatinga

Figura 2.5: Acidentes Fatais 2012
 Fonte: NUGEO - DETRAN/DF

Capítulo 3

ESTATÍSTICA SCAN E ESTIMADOR DE INTENSIDADE

3.1 INTRODUÇÃO

O Objetivo desse Capítulo é apresentar os métodos *Scan* e Estimador de Intensidade, de forma a identificar *clusters* de acidentes de trânsito. A estatística *Scan* testa *clusters* em áreas e pontos, enquanto o Estimador de Intensidade vai mostrar os locais com maior intensidade de acidentes.

3.2 ESTATÍSTICA *SCAN*

A estatística *scan* é comumente usada para testar se um ponto é puramente aleatório, ou se algum *cluster* pode ser detectado. O maior interesse é em detectar os *clusters* não explicados pelos processos tradicionais de Bernoulli e Poisson, com intensidade proporcional a algumas funções conhecidas.

Denota-se:

- N = População total no mapa

- $C =$ Total de casos no mapa
- $n_Z =$ População na zona Z
- $c_Z =$ Casos na zona Z
- $\mu_Z =$ Número esperado de casos em Z
- $c_{\bar{Z}} = C - c_Z =$ Número de casos fora da zona Z
- $n_{\bar{Z}} = N - n_Z =$ População fora da zona Z

Poisson

Para o modelo de Poisson, os pontos são gerados por um processo heterogêneo de Poisson. As hipóteses são (Kulldorff, 1997):

$$\begin{cases} H_0 : p = q \\ H_1 : p > q \end{cases}$$

Sob H_0 temos que p é a probabilidade de um indivíduo ser um caso dentro da zona Z .

$$p = \frac{C}{N}$$

Sob H_1 temos que q é a probabilidade de um ponto estar fora da zona Z .

$$p = \frac{c_Z}{n_Z}$$

$$q = \frac{C - c_Z}{N - n_Z}$$

Logo, temos que a função de verossimilhança para o modelo de Poisson, segundo Kulldorff (1997), é:

$$L(Z) = \frac{c_Z^{c_Z} (C - c_Z)^{C - c_Z} e^{-C}}{c_Z! (C - c_Z)!} \quad (3.1)$$

e

$$L_0(Z) = \frac{\mu_Z^{c_Z} (C - \mu_Z)^{C - c_Z} e^{-C}}{c_Z! (C - c_Z)!} \quad (3.2)$$

A razão de verossimilhança λ , que é a estatística do teste, pode ser escrita como:

$$\lambda = \frac{L}{L_0} = \left(\frac{c_Z}{\mu_Z} \right)^{c_Z} \left(\frac{C - c_Z}{C - \mu_Z} \right)^{C - c_Z}, \quad \frac{c_Z}{\mu_Z} > \frac{C - c_Z}{C - \mu_Z} \quad (3.3)$$

Propriedades do Teste

A maioria dos métodos estatísticos para análise de *cluster* de um ponto no espaço são ou descritivas no sentido de poder detectar a localização de *clusters* sem nenhuma inferência envolvida, ou fazem inferência sem a habilidade de detectar a localização dos *clusters*. Uma característica importante do teste espacial *Scan* é que ele faz os dois, então quando a hipótese nula é rejeitada, pode-se localizar a área específica do mapa que causa a rejeição (Kulldorff, 1997).

Exemplo

Os modelos foram ilustrados por Kulldorff (1997) usando dados da Síndrome de Mortalidade Infantil Repentina (SIDS, em inglês) no estado da Carolina do Norte. Os casos da síndrome foram ajustados pela distribuição geográfica desigual dos nascimentos. A Figura 3.1 mostra os resultados desse caso.

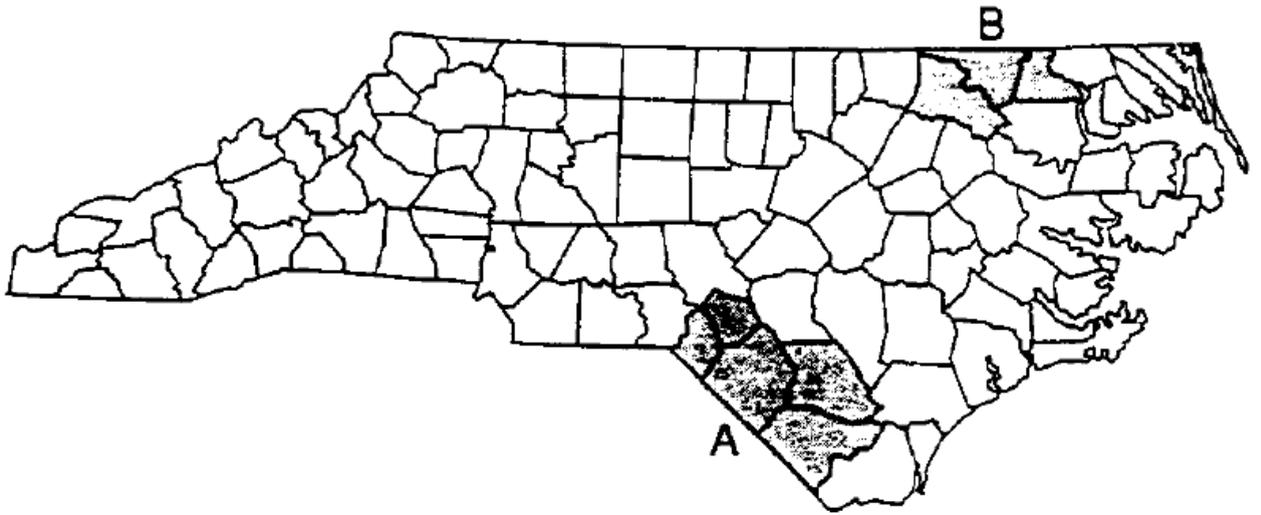


Figura 3.1: Exemplo de uso da Estatística *Scan*
 Fonte: Kulldorff (1997)

3.3 Estimador de Intensidade

Uma forma simples de analisar o comportamento de padrões de pontos é estimar a intensidade pontual do processo em toda a região de estudo. Pode-se ajustar uma função bidimensional sobre os eventos considerados, compondo uma superfície cujo valor será proporcional à intensidade de amostras por unidade de área. Esta função realiza uma contagem de todos os pontos dentro de uma região de influência, ponderando-os pela distância de cada um à localização de interesse. O estimador de intensidade é muito útil para fornecer uma visão geral da distribuição de primeira ordem dos eventos (Druck et al., 2004).

Suponha que u_1, \dots, u_n localizações de n eventos observados em uma região A e que u representa uma localização genérica cujo valor se quer estimar.

O Estimador de Intensidade de kernel é calculado como (Druck et al., 2004):

$$\hat{\lambda}_\tau(u) = \sum_{h_i \leq \tau} \frac{3}{\pi \tau^2} \left(1 - \frac{h_i^2}{\tau^2}\right)^2 \quad (3.4)$$

Onde, τ é o raio de influência de kernel e h_i representa a distância entre a localização em que desejamos calcular a função e o evento observado.

No artigo de Zheng et al. (2011), o Estimador de Intensidade utilizado foi modificado. Ele é descrito nas seguintes equações:

$$\hat{f}(x) = \frac{1}{n\tau} \sum_{i=1}^n K\left(\frac{1}{\tau}(x - X_i)\right) \quad (3.5)$$

Onde,

n = tamanho da amostra

X_i = frequência observada de incidentes no segmento i

$\hat{f}(x)$ = estimativa de intensidade do ponto no espaço medido na localização x

$(x - X_i)$ é a distância h_i

A função $K(x)$ será uma função densidade de probabilidade simétrica, normal, com média zero e variância um.

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad (3.6)$$

O raio de influência define a área centrada no ponto de estimação u que indica quantos eventos u_i contribuem para a estimativa da função de intensidade. Se o raio for muito pequeno, vai gerar uma superfície descontínua, no entanto, se for muito grande a superfície pode ficar muito amaciada (Druck et al., 2004).

Dependência Espacial

O método do vizinho mais próximo estima a função de distribuição cumulativa baseado nas distâncias entre eventos em uma região de análise. A plotagem dos resultados dessa função pode ser usada como um método exploratório para se verificar se existe evidência de interação entre os eventos. Se o gráfico apresentar um crescimento rápido para pequenos valores de distância, esta situação indica uma interação entre os eventos, caracterizando agrupamentos nestas escalas. Se o gráfico crescer apenas em valores maiores de distância, esta situação aponta para uma distribuição mais regular (Druck et al., 2004).

A análise de vizinhança pode ser usada como método formal para se comparar estatisticamente a distribuição dos eventos observados com o que se esperaria da hipótese de aleatoriedade espacial completa (CSR). Serão criados envelopes de simulação para a distribuição CSR, a fim de se acessar a significância dos desvios. Na hipótese de CSR, a função de distribuição $G(w)$ seria uma Poisson da forma (Druck et al., 2004):

$$G(h) = 1 - e^{-\lambda\pi h^2}, h \geq 0 \quad (3.7)$$

A estimação simulada para a distribuição $G(w)$ assumindo CSR é

$$\bar{G}(h) = \frac{\sum_i^k \hat{G}_i(h)}{k} \quad (3.8)$$

Onde $\hat{G}_i(h)$, $i = 1, 2, \dots, k$ são funções de distribuição empíricas, estimadas a partir de k simulações independentes dos n eventos, na hipótese de CSR (n eventos independentes e uniformemente distribuídos). Os envelopes superior e inferior são

definidos como:

$$\begin{cases} U(h) = \max\{\hat{G}_i(h)\}, i = 1, \dots, k \\ L(h) = \min\{\hat{G}_i(h)\}, i = 1, \dots, k \end{cases}$$

O gráfico da distribuição estimada versus a distribuição simulada, com a adição dos envelopes superior e inferior, permite medir a significância dos desvios relativos a aleatoriedade. Se a condição CSR for válida para os dados observados, o gráfico deve ser praticamente linear com um ângulo de 45 graus. Se o dados apresentam tendências para agrupamentos, os traçados estarão acima da linha de 45 graus. Padrões de regularidade serão encontrados em traçados abaixo da linha de 45 graus (Druck et al., 2004).

Exemplo

O Estimador de Intensidade foi aplicado por Zheng et al. (2011) em dados referentes a danos de acidentes de trânsito na região de Hampton Roads na Virgínia. Os dados foram ponderados pelos custos econômicos dos acidentes. A Figura 3.2 mostra os resultados obtidos, em que as áreas em vermelho são as mais perigosas em quantidade de acidentes, e as áreas em verde as menos perigosas.

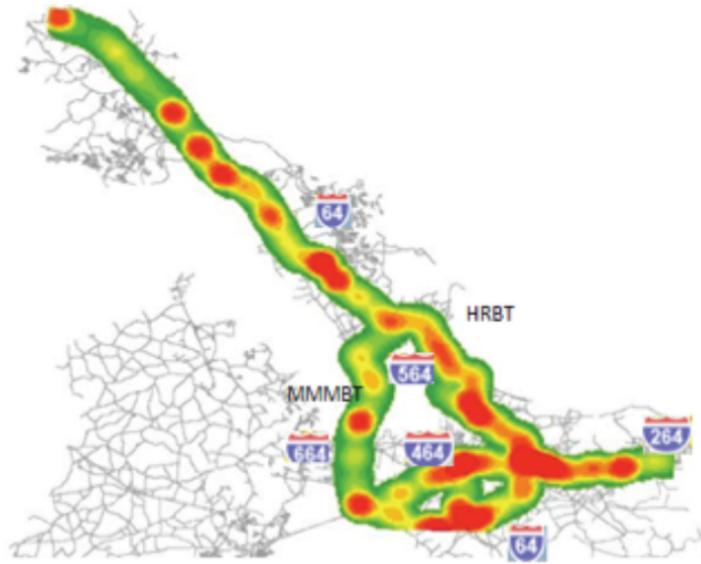


Figura 3.2: Exemplo de uso da Regressão Geograficamente Ponderada
Fonte: Zheng et al. (2011)

Capítulo 4

MATERIAL E MÉTODOS

4.1 INTRODUÇÃO

A seguir são apresentados os principais materiais utilizados nesse trabalho, assim como os métodos que serão aplicados aos dados em questão.

4.2 BANCO DE DADOS DE ACIDENTES DE TRÂNSITO

Os dados de acidentes de trânsito foram obtidos com a Gerência de Estatística de Acidentes de Trânsito (GEREST) do DETRAN/DF. Esses dados são armazenados no Sistema de Informações de Acidentes de Trânsito (SAT), e são oriundos, basicamente, de três fontes oficiais: Polícia Civil, Instituto Médico Legal (IML) e Secretaria de Saúde.

Desde 2000, as cidades do Distrito Federal com maior número de acidentes de trânsito com morte foram Brasília, Ceilândia e Taguatinga. Nos anos em estudo, de 2011 a 2013, Ceilândia foi a cidade com mais acidentes, como pode-se observar na Tabela 4.1:

Tabela 4.1: Acidentes de Trânsito com Morte por Região Administrativa

RA	2011	2012	2013	TOTAL
Brasília	21	17	19	57
Ceilândia	27	20	26	73
Taguatinga	21	17	23	61

Serão utilizados no trabalho os dados de acidentes de trânsito com morte nas Regiões Administrativas de Brasília, Ceilândia e Taguatinga, de janeiro de 2011 a dezembro de 2013. Essas regiões foram selecionadas por terem tido as maiores quantidades de acidentes dos últimos anos no Distrito Federal, e o período foi escolhido de forma que a análise fosse atual, porém não restrita a um único ano, para minimizar a possibilidade de viés em decorrência de alguma diferença específica do ano.

De forma a se ter uma visão geral dos dados trabalhados, alguns aspectos dos acidentes de trânsito nessas regiões foram analisados a seguir.

4.2.1 Análise Descritiva dos Dados

Analisando o banco de dados, observa-se que o perfil dos acidentes nessas três regiões é relativamente similar. No período estudado houve uma concentração dos acidentes nos dias de semana (de segunda-feira à sexta-feira). As naturezas mais frequentes foram atropelamento de pedestres e colisões, sendo que em Brasília também houve destaque nos choques com objeto fixo. Quanto ao período do dia, Brasília e Ceilândia tiveram pouco mais da metade dos acidentes durante o dia (06h-17h59), e Taguatinga teve 65% dos acidentes com morte no período noturno (18h-05h59).

Essas informações podem ser indicativos de falhas em algum dos diversos fatores que são usados para modelar a ocorrência de acidentes de trânsito. Nesse trabalho

será feito um estudo mais aprofundado do fator localização geográfica, buscando identificar os locais onde o fator espaço é mais significativo, ou seja, locais mais perigosos para acidentes de trânsito.

4.3 MÉTODOS

As coordenadas geográficas dos dados utilizados serão manualmente localizadas no *Google Earth* por meio dos endereços dos acidentes. Os pontos serão plotados em mapas georreferenciados das Regiões Administrativas, e o método da Estatística *Scan* de Kulldorff (1997) será utilizado para procurar identificar a existência de *clusters*. Para a análise de área de Kulldorff, serão pesquisados os fluxos de veículos no local, dia e hora dos acidentes, e os veículos não envolvidos nos acidentes serão usados como caso controle em cada um dos setores estabelecidos. Esses dados são registrados pelos pontos de fiscalização eletrônica e serão obtidos no sistema SIG DETRAN. A fiscalização eletrônica registra o número de veículos que passaram pelo ponto de 15 em 15 minutos. No caso dessa análise, foram usados os dados dos 15 minutos imediatamente anteriores ao horário do acidente. Para a análise pontual dos acidentes serão procurados clusters de acidentes com morte envolvendo motos, sendo o caso controle todos os acidentes com morte em que não houve envolvimento de motos. Em seguida, serão plotados mapas de intensidade de acidentes com morte nas regiões estudadas, de forma a visualizar melhor os locais mais intensos. Após essa comparação poderão ser identificados alguns dos locais mais perigosos para acidentes de trânsito nas Regiões de Brasília, Taguatinga e Ceilândia, levando em consideração o fator espaço.

Capítulo 5

ANÁLISE DOS RESULTADOS

5.1 INTRODUÇÃO

A Análise Espacial, neste caso, consiste em estudar a distribuição espacial dos acidentes de trânsito com morte, representados por pontos, testando hipóteses sobre o padrão observado: se é aleatório, ou se apresenta-se em aglomerados, ou regularmente distribuído.

5.2 MAPAS

A Análise Espacial começa com a observação visual da distribuição dos pontos no espaço. Os acidentes de trânsito com morte dos anos de 2011 a 2013 nas regiões administrativas de Brasília, Ceilândia e Taguatinga foram plotados nas Figuras 5.1 e 5.2.

Observa-se que em Brasília há uma maior concentração de pontos na Asa Sul e no centro do Plano Piloto, enquanto a Asa Norte teve poucas ocorrências nesses três anos (Figura 5.3). Os pontos do centro estão bem próximos, sugerindo possíveis aglomerados, que serão testados posteriormente pela estatística *Scan* de Kulldorff.

Em Taguatinga ocorreu o contrário, a parte norte da RA apresentou muito mais

ocorrências que a parte sul. Vias como a LJ 1 e LJ 2 registraram muitos acidentes nesse período (Figura 5.4).

Já em Ceilândia, os pontos não aparentam estar concentrados em nenhuma região específica, com exceção da Via MN1 (Hélio Prates), que divide a RA em norte e sul (Figura 5.4).

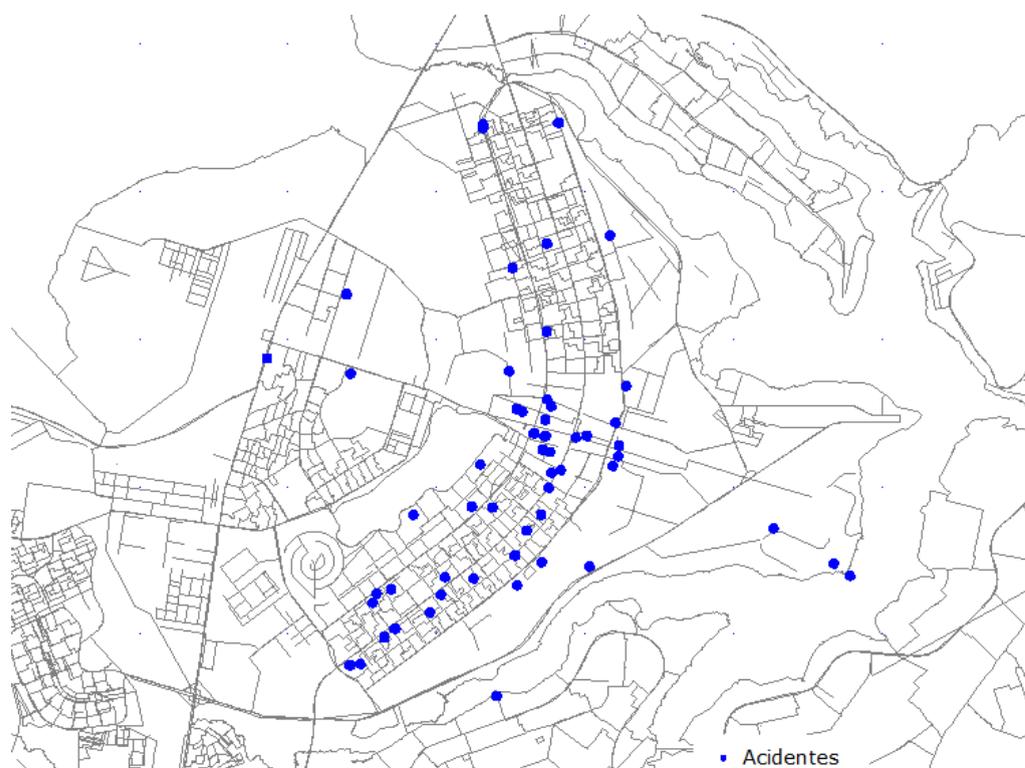


Figura 5.1: Acidentes Fatais em Brasília 2011 - 2013

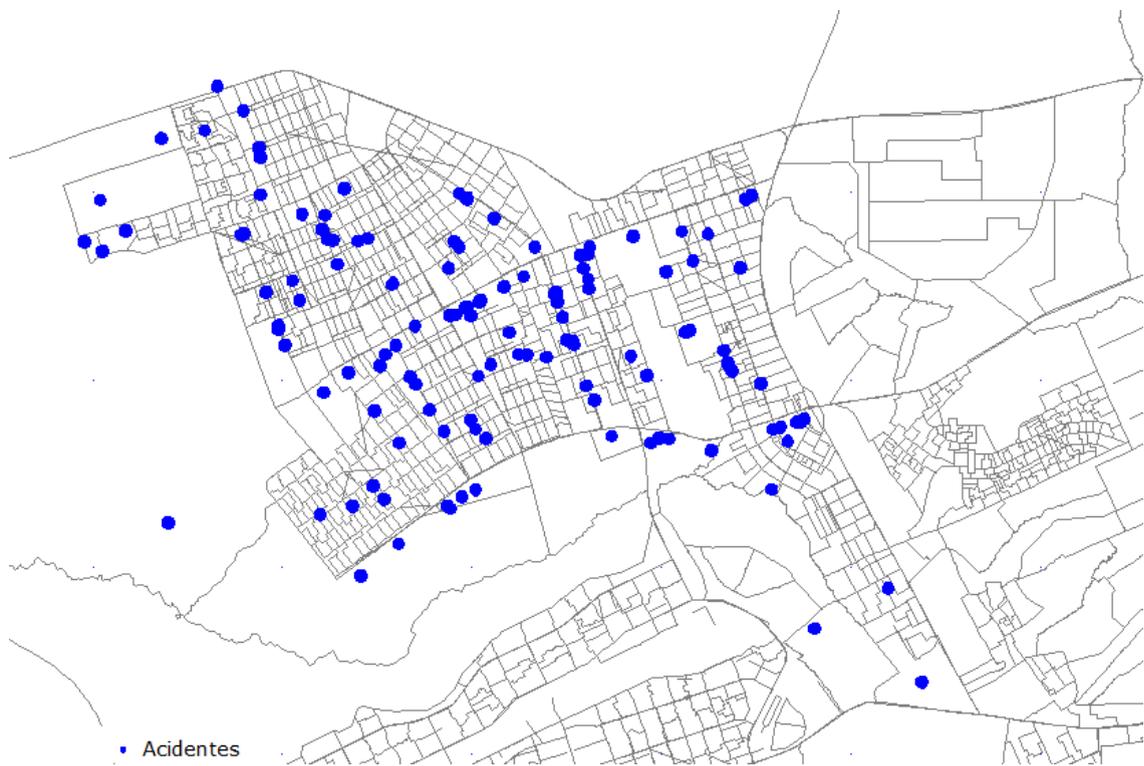


Figura 5.2: Acidentes Fatais em Ceilândia e Taguatinga 2011 - 2013

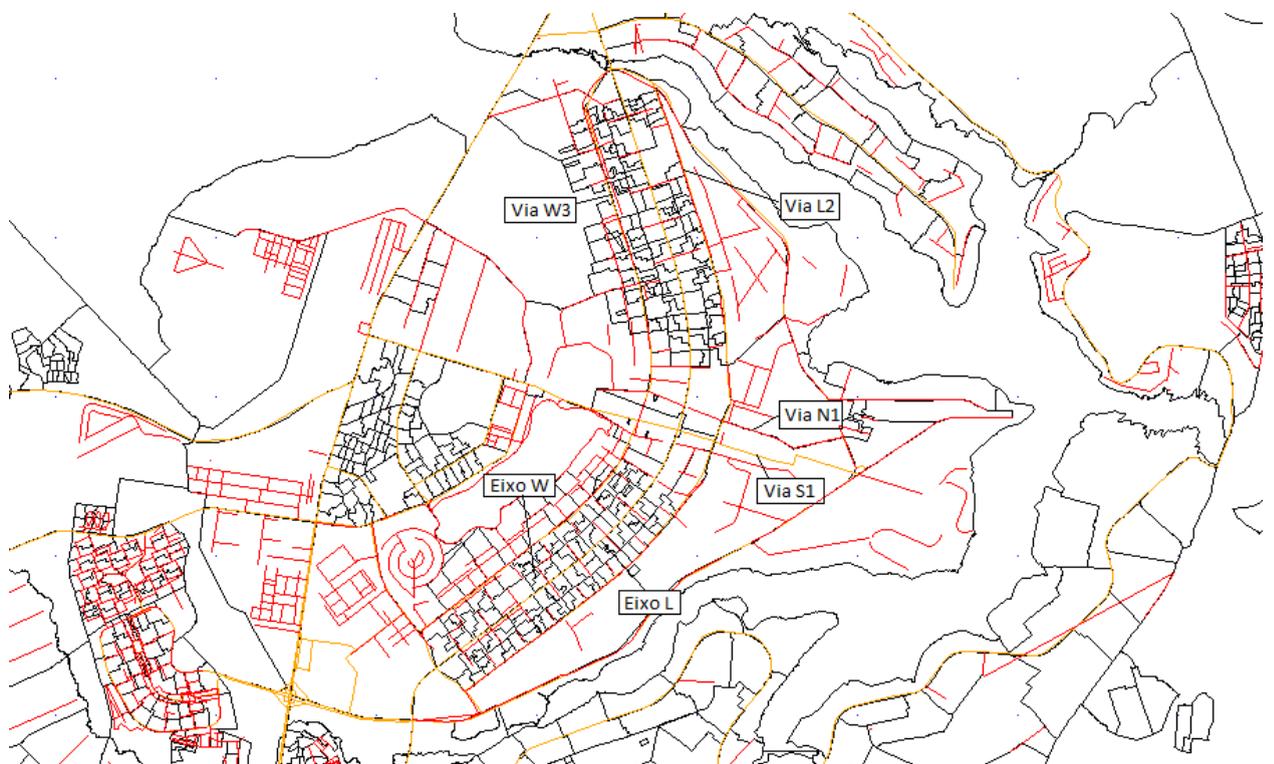


Figura 5.3: Principais Vias de Brasília

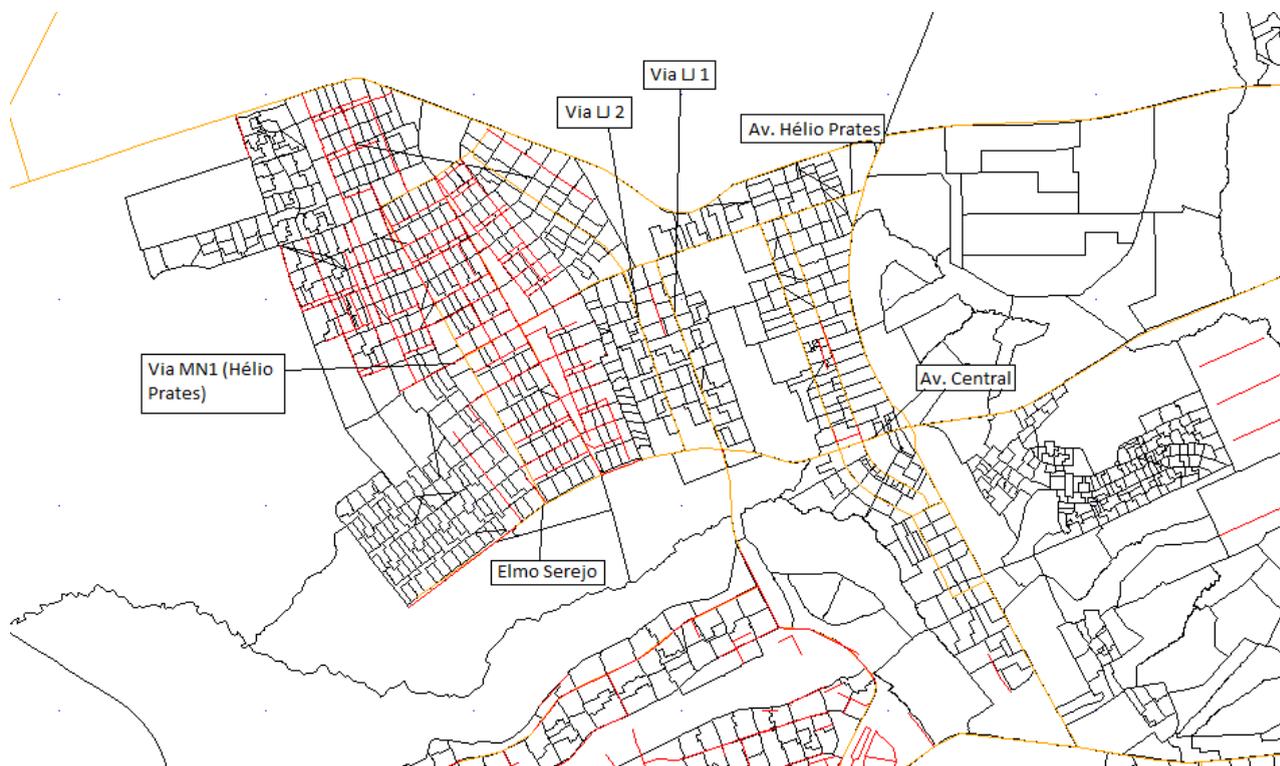


Figura 5.4: Principais Vias de Ceilândia e Taguatinga

5.3 Estatística *Scan*

A Estatística *Scan* de Kulldorff pode ser usada tanto para pontos quanto para áreas. Essa seção apresentará as análises para esses dois tipos de dados.

5.3.1 Análise por Área

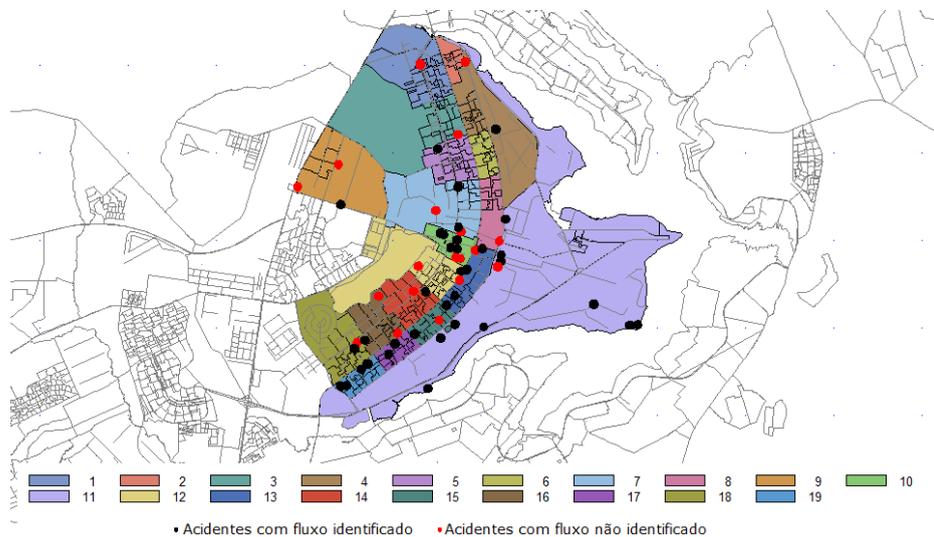
Para utilizar a estatística *Scan* de Kulldorff para áreas, é necessário ter o mapa dividido em setores, que foram testados para a possível existência de *clusters*. Inicialmente, as três RAs foram divididas em 50 setores, sendo 19 em Brasília, 11 em Ceilândia e 20 em Taguatinga. As coordenadas de alguns pontos foram ajustadas nessa etapa, pois estavam fora dos setores demarcados. Para a estatística *Scan*, utilizou-se os dados de fluxo dos pardais como controle dos casos de acidentes com

morte. Entretanto, nem todos os fluxos foram obtidos, devido à falta de fiscalização eletrônica próxima ao ponto do acidente. Nesses casos, o ponto foi descartado para esse teste. A Tabela 5.1 mostra o percentual de pontos que foram utilizados em cada RA. A Figura 5.5 ilustra esse problema.

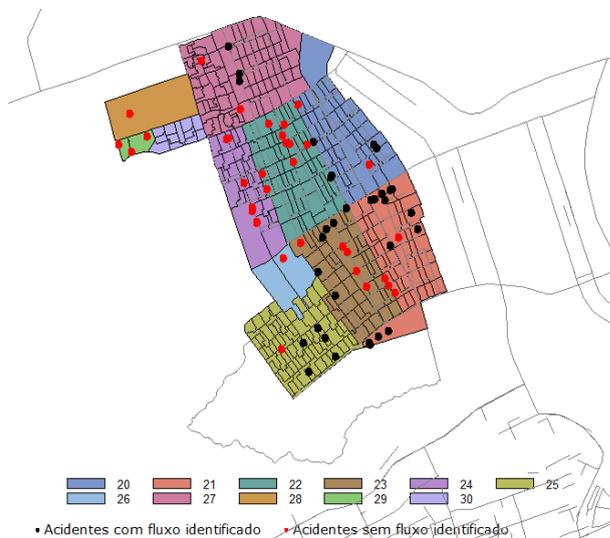
Tabela 5.1: Acidentes de Trânsito com Morte por RA e Existência de Fiscalização Eletrônica

RA	Total de Pontos	Pontos Utilizados	Percentual de Pontos Utilizados
Brasília	57	37	65%
Ceilândia	73	33	45%
Taguatinga	61	41	67%
Total	191	111	58%

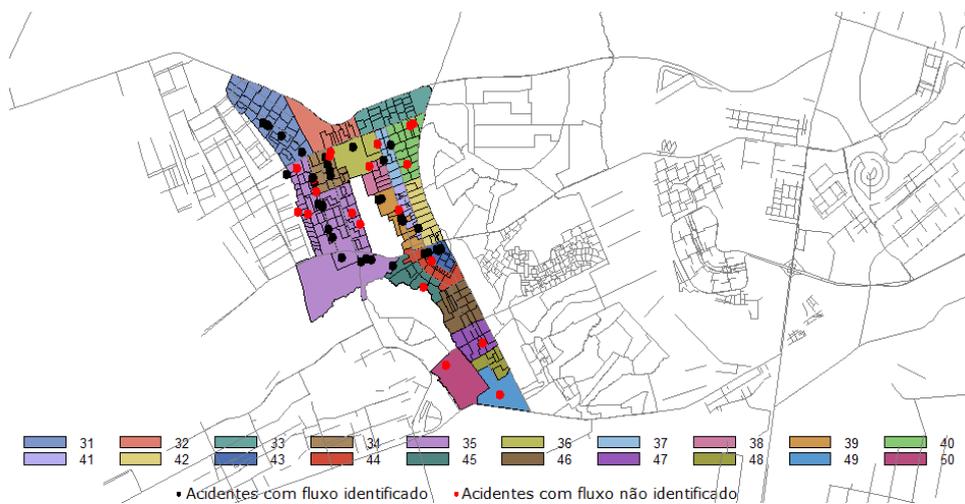
Devido aos pontos descartados, muitos setores ficaram sem casos, e consequentemente, sem casos-controle. Os setores foram, então, ajustados de forma que todos tivessem pelo menos um acidente fatal com dados de fiscalização eletrônica, e não houvesse buracos sem setores no mapa. Sobraram 32 setores, sendo 15 em Brasília, 6 em Ceilândia e 11 em Taguatinga. Os setores e pontos finais podem ser vistos na Figura 5.6.



(a) Brasília

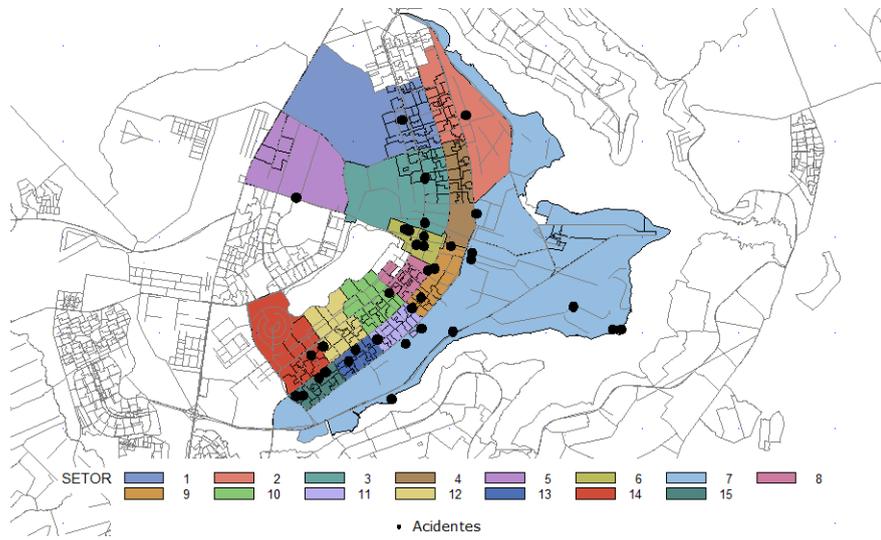


(b) Ceilândia

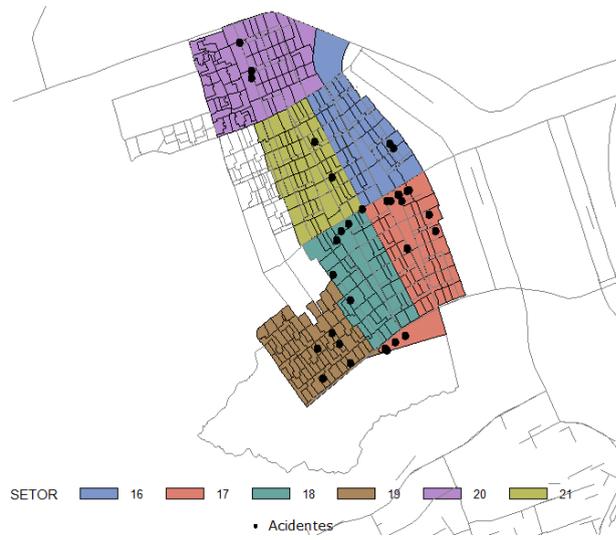


(c) Taguatinga

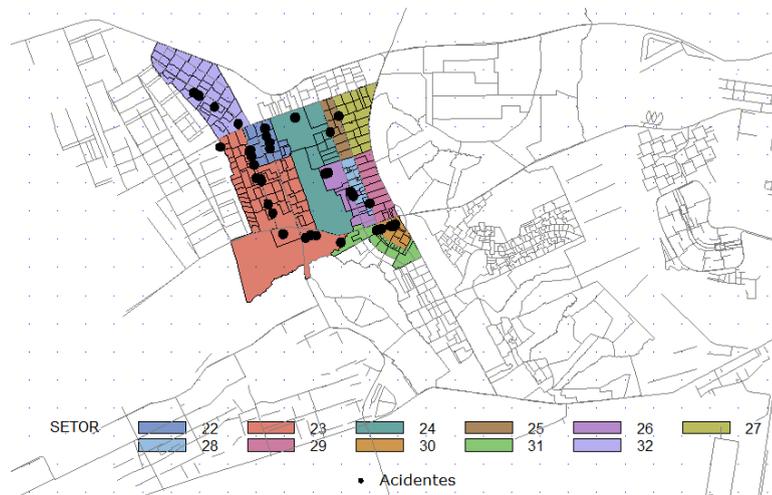
Figura 5.5: Acidentes Fatais por Existência de Fiscalização Eletrônica e Setores



(a) Brasília



(b) Ceilândia



(c) Taguatinga

Figura 5.6: Acidentes Fatais Finais por Setores

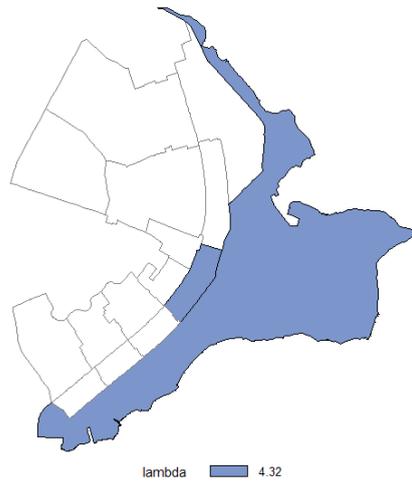
Os pontos e setores finais foram, então, utilizados para calcular os *clusters* e seus respectivos lambdas (λ) e p-valores. A Tabela 5.2 apresenta o maior *cluster* de cada RA. Em Brasília, o maior λ foi encontrado no agrupamento dos setores 7 e 9, com 12 pontos no total. Em Ceilândia, o maior valor foi encontrado no aglomerado dos setores 16, 20 e 21, somando 8 casos, e em Taguatinga, foi o setor 29, com apenas 1 caso. Isto foi possível devido à ponderação da quantidade de acidentes com morte e o número de veículos que passaram pelo ponto no momento do acidente. No caso de Taguatinga, por exemplo, apenas 2 veículos foram registrados na fiscalização eletrônica próxima ao acidente do setor 29. A Figura 5.7 ilustra esses setores nos mapas.

Para um nível de significância de 5%, tem-se que a hipótese nula foi rejeitada para o maior *cluster* de Brasília e Taguatinga. Ou seja, Brasília possui um *cluster* estatisticamente significativo nos setores 7 e 9, e Taguatinga no setor 29. No entanto, o aglomerado dos setores 16, 20 e 21, em Ceilândia, não é significativo.

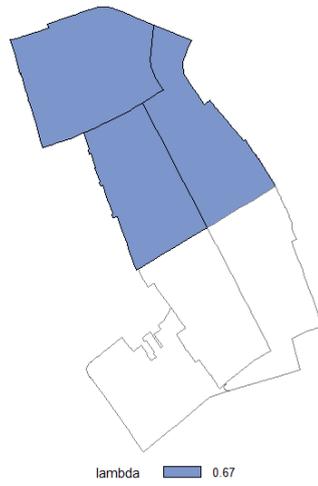
Tabela 5.2: Lambda (λ) e P-valor do maior cluster de cada RA

RA	Maior <i>Cluster</i>	λ	p-valor
Brasília	7, 9	4,32	0,042
Ceilândia	16, 20, 21	0,67	0,303
Taguatinga	29	3,57	0,049

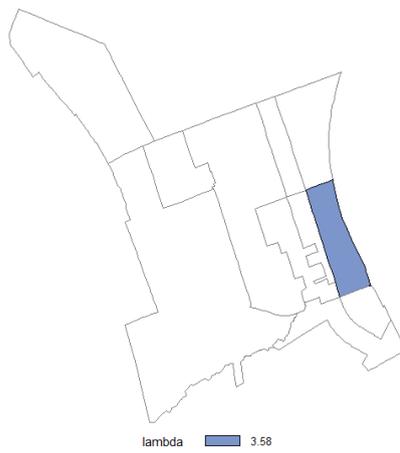
Pode-se concluir, segundo o teste *Scan* de Kuldorff por áreas, que foram encontrados dois *clusters*, sendo um em Brasília e um em Taguatinga.



(a) Brasília



(b) Ceilândia



(c) Taguatinga

Figura 5.7: Maiores *Clusters* em cada RA

5.3.2 Análise Pontual

A Figura 5.8 mostra os acidentes por veículos envolvidos. Definiu-se como “Automóvel” todo veículo que não seja moto ¹. Os acidentes envolvendo motos são aqueles marcados como “Moto Única” e “Automóvel x Moto” (ilustrados de laranja e verde no mapa). Observando a Figura 5.8, percebeu-se que haviam muitos acidentes com morte envolvendo motos. Isso levantou a suspeita de que pudessem haver *clusters* de acidentes com motocicletas em relação ao total de acidentes com outros veículos. A análise *scan* foi usada para verificar essa hipótese.

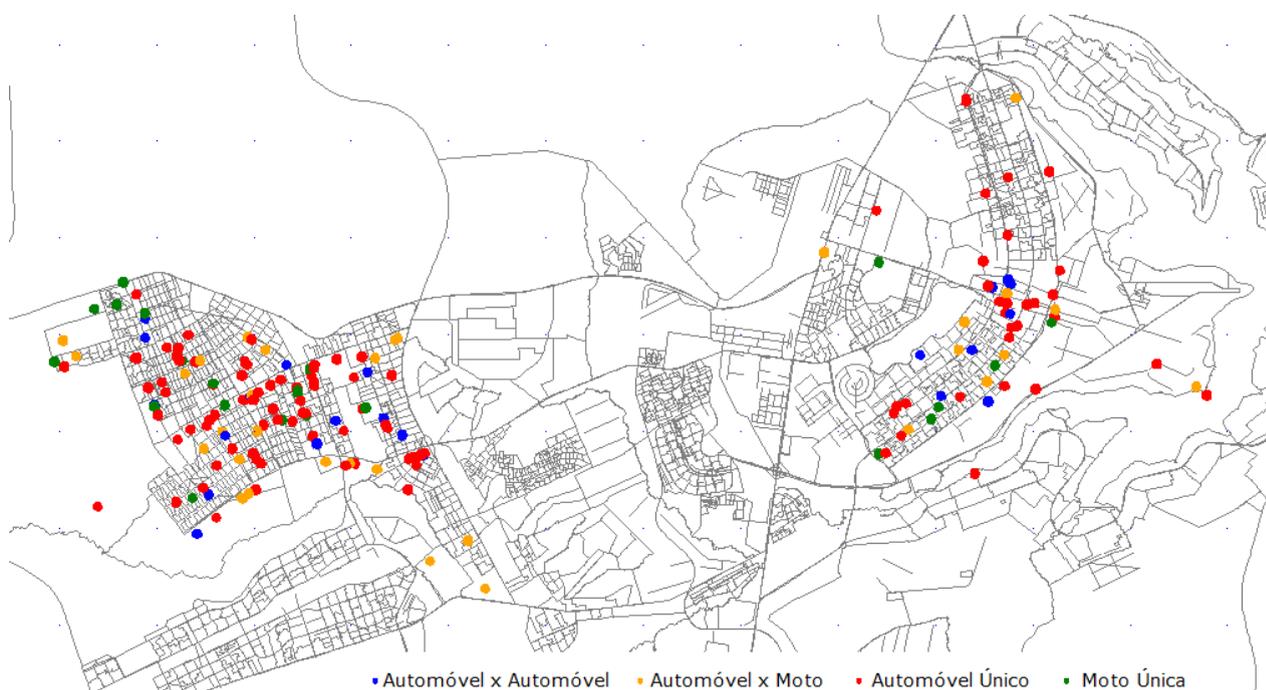
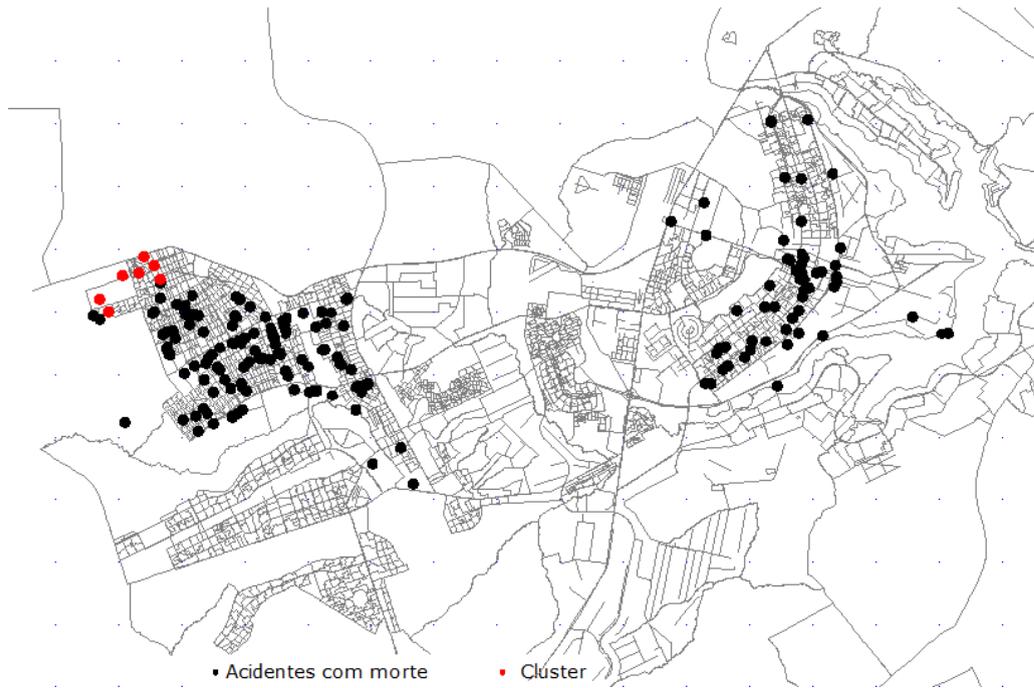


Figura 5.8: Acidentes com Morte por Veículos Envolvidos

Para o uso da Estatística *Scan*, foram considerados como casos os acidentes envolvendo motos e como controle os acidentes que não envolviam motos. Como se pode ver na Figura 5.9, foi encontrado um *cluster* na parte norte de Ceilândia.

¹São consideradas motos: motocicletas, motonetas, ciclomotores e triciclos

Formaram esse *cluster* os pontos: 25, 102, 22, 146, 137, 164 e 136. O λ foi de 2,49 e o p-valor de 0,961. Logo, o *cluster* não é estatisticamente significativo.



(a) Visão Geral



(b) Ceilândia

Figura 5.9: Análise Pontual

5.4 Estimador de Intensidade

Como citado no Capítulo 3, pode-se calcular o estimador de intensidade pela fórmula original de kernel ou pelo estimador modificado usado por Zheng et al. (2011). Nesta seção, esses dois métodos foram comparados.

5.4.1 Estimador de Intensidade Original de Kernel

Pode-se observar, pela Figura 5.10, que há uma concentração mais fraca na parte Sul de Brasília e uma concentração forte entre Ceilândia e Taguatinga. Verifica-se também que nenhum dos *clusters*, significativos ou não, encontrados pela estatística *scan* de Kulldorff estão localizados nas áreas mais escuras do mapa. Isso pode ser explicado pela quantidade de pontos descartados ao fazer a análise por setores. No caso da análise pontual, foram procurados *clusters* de acidentes com motos, logo era esperado que os focos fossem diferentes da análise de intensidade, que considera todos os pontos.



Figura 5.10: Estimador de Intensidade Original

O método do Vizinho mais Próximo (Figura 5.11) mostrou que a curva $G(h)$ teve um crescimento rápido até um raio de 7000, quase estabilizando após esse valor. Há uma indicação de que os dados apresentam tendência para agrupamentos, pelo crescimento rápido dessa curva. O método dos Envelopes de Simulação (Figura 5.12) indica existência de agrupamentos, pois a curva está acima da linha de 45 graus, no entanto os intervalos tocam a linha, o que pode ser um indício de que os *clusters* não são estatisticamente significativos.

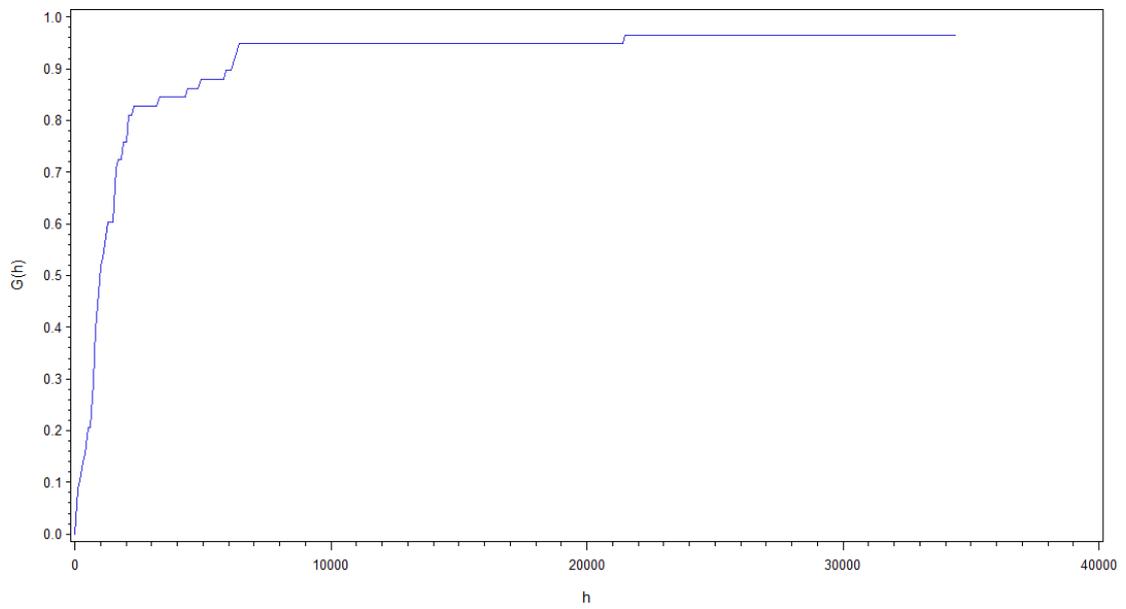


Figura 5.11: Vizinho Mais Próximo

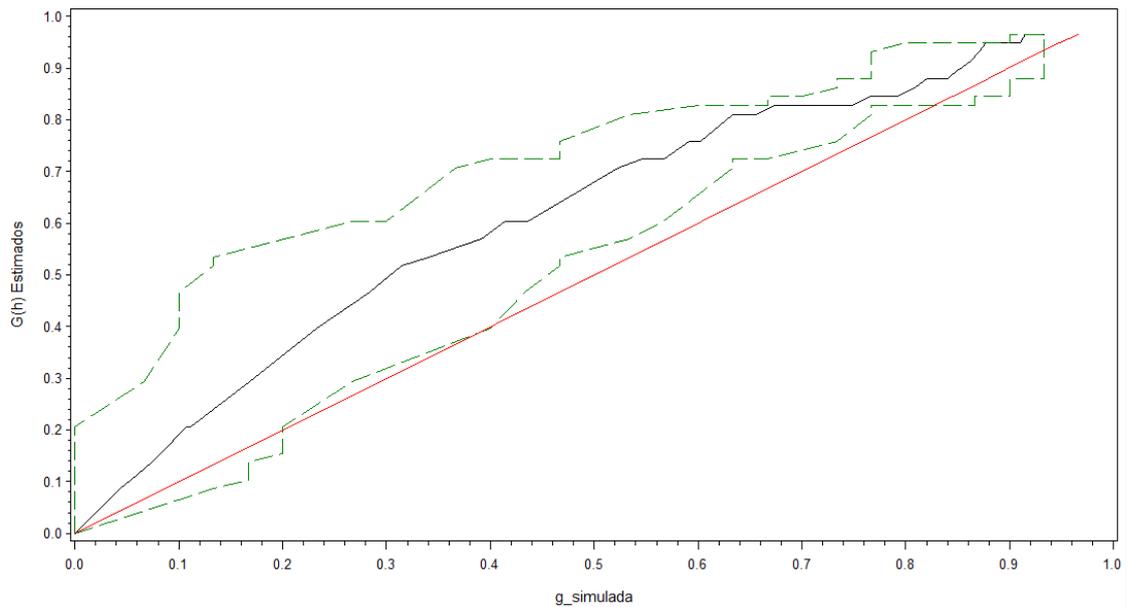


Figura 5.12: Envelope de Simulação

5.4.2 Estimador de Intensidade Modificado de Zheng et al. (2011)

No artigo de Zheng et al. (2011) é utilizada uma fórmula diferente para o cálculo do estimador de kernel. O resultado pode ser observado na Figura 5.13.

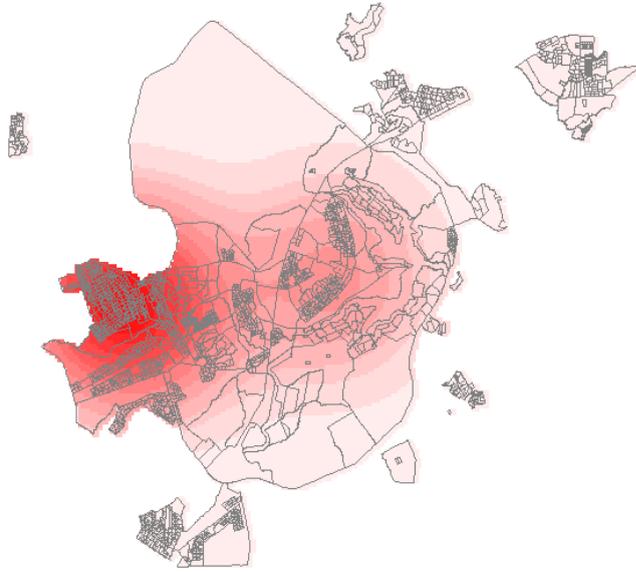


Figura 5.13: Estimador de Intensidade Modificado

Observa-se que o mapa, no geral, é mais homogêneo. Entretanto, os aglomerados mais escuros se encontram nos mesmos locais da análise pela fórmula original do estimador de kernel. As análises de dependência espacial se comportaram da mesma forma que no estimador original.

Capítulo 6

CONCLUSÕES

O trabalho mostrou que os acidentes de trânsito com morte nas Regiões Administrativas de Brasília, Ceilândia e Taguatinga, no geral, apresentaram tendências de agrupamentos, porém apenas dois *clusters* foram identificados na análise de Kull-dorff por áreas, sendo um em Brasília e outro em Taguatinga. Os acidentes foram ponderados pela quantidade de veículos que foram registrados pelos pontos de fiscalização eletrônica no local e hora do acidente. Isso pode explicar o fato da Ceilândia não ter nenhum *cluster* significativo, pois apesar de ser a RA com maior número de acidentes, também tem um grande fluxo de veículos. Além disso, mais da metade dos pontos dessa região foram descartados por falta de dados de fiscalização eletrônica. Este foi o pior problema encontrado nessa análise. No total, apenas 58% dos pontos das 3 RAs foi aproveitado para a análise por áreas da estatística *scan*. A análise pontual foi restrita ao interesse em acidentes envolvendo motocicletas e, apesar de ter sido encontrado um *cluster*, este não era significativo.

O Estimador de Intensidade de Zheng et al. (2011) é bastante visual, no entanto não é um teste, por isso não pode ser considerado que existe um *cluster* nas áreas mais escuras, apenas uma indicação de aglomerado. Os resultados foram de acordo

com o esperado ao observar o mapa de acidentes, com aglomerados muito grandes e que não eram significativos. Na Ceilândia, por exemplo, é difícil identificar pontos mais perigosos, pois toda a região está em uma área mais escura.

Como sugestão, no caso da Estatística *Scan*, o interessante seria ter utilizado os dados de acidentes não fatais como controle para os casos de acidentes fatais. Essa análise não foi feita devido à impossibilidade de pesquisar as coordenadas geográficas do elevado número de acidentes não fatais manualmente no tempo disponível.

6.1 PROBLEMAS OBSERVADOS

Dados dos Acidentes

- A falta de coordenadas geográficas dos acidentes no banco de dados do DETRAN e, conseqüentemente, a necessidade de fazer essa pesquisa manualmente no *Google Earth*;
- A falta de precisão na descrição de muitos endereços, dificultando a localização do ponto exato da ocorrência;
- Atualizações e problemas no sistema do DETRAN, que levaram a modificações no banco de dados após o início do trabalho;

Fluxo de Veículos

- A falta de pontos de fiscalização eletrônica em todos os locais de acidentes. A Figura 6.1 ilustra esse problema. Os pontos em azul são os acidentes que possuem dados de fluxo de veículos, e os pontos em vermelho são os que não

possuem, seja por falta de fiscalização eletrônica na via, ou porque a fiscalização existente não estava funcionando no momento do acidente. Os pontos em vermelho não puderam ser aproveitados na estatística *scan*;

- Muitos endereços não especificavam o sentido da via que o acidente ocorreu, e havia fiscalização eletrônica nos dois sentidos da via. Nesses casos foi feita a média da quantidade de veículos registrada pelos dois aparelhos;
- Em alguns casos, a fiscalização eletrônica mais próxima do local do acidente era muito distante, com vários pontos de entrada e saída de veículos entre eles. Isso poderia causar grande diferença no valor do fluxo de veículos;

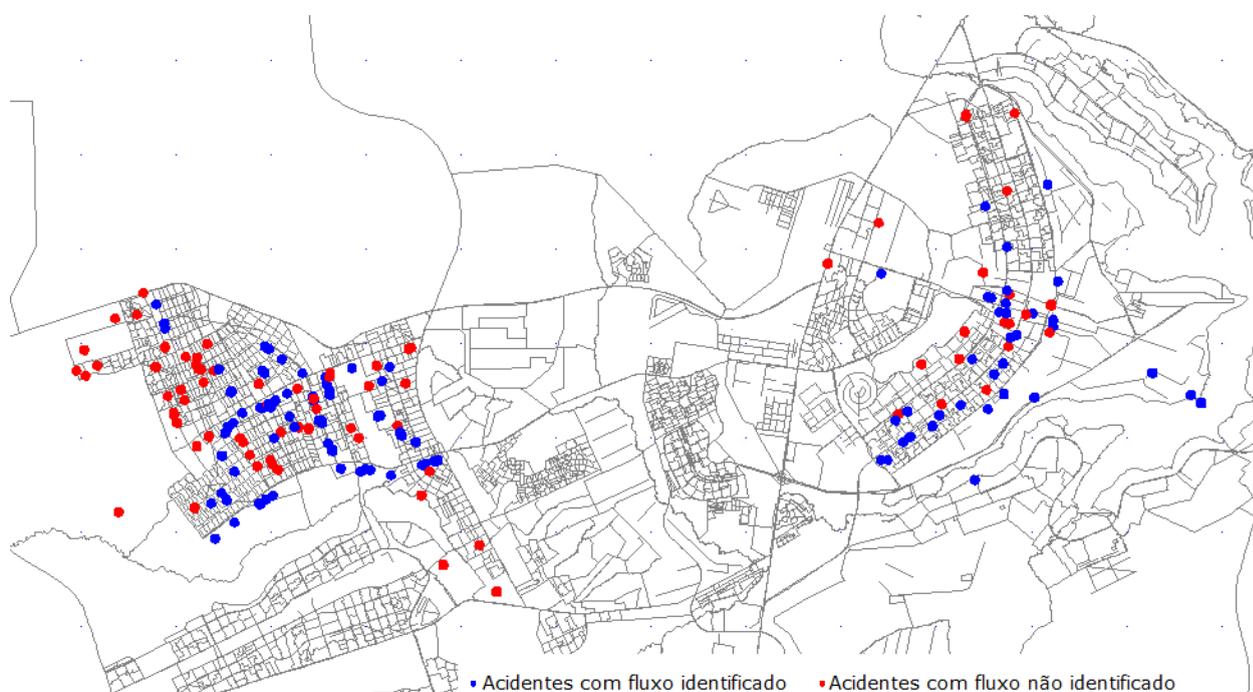


Figura 6.1: Acidentes com Morte por Existência de Dados de Fiscalização Eletrônica

- A fiscalização eletrônica não registra fluxo de motocicletas, que são responsáveis por muitos acidentes nas RAs estudadas. A Figura 5.8 mostrou

os acidentes por veículos envolvidos. Foram considerados como “Automóveis” todos os veículos, com exceção das motocicletas.

Estatística *Scan*

- A grande quantidade de pontos descartados do estudo para esse teste. Dos 191 acidentes, apenas 111 foram utilizados. Em Ceilândia, mais da metade dos acidentes foi desconsiderada. Apenas 33 dos 73 acidentes da RA foram considerados. Isso pode ter sido prejudicial às conclusões sobre essa RA, já que nenhum *cluster* significativo foi encontrado nessa região na análise por áreas.

Referências Bibliográficas

- ABNT (1989). Pesquisa de acidentes de trânsito - terminologia. Technical report.
- Council, F., Harkey, D., Nabors, D., Khattak, A., & Mohamedshah, Y. (2003). Examination of fault, unsafe driving acts, and total harm in car-truck collisions. *Transportation Research Record*, (1830):63–71.
- DETRAN (2012). Mortes por atropelamento caem 25%. Technical report. URL <http://www.detran.df.gov.br/noticias/item/2065-mortes-por-atropelamento-caem.html>. Acesso em 22 ago. 2013.
- Druck, S., Carvalho, M. S., Câmara, G., & Monteiro, A. M. V. (2004). *Análise Espacial de Dados Geográficos*. EMBRAPA.
- Fotheringham, A., Brunson, C., & Charlton, M. (2002). Geographically weighted regression: The analysis of spatially varying relationships. *Chichester: Wiley*.
- GEREST (2012a). Anuário estatístico de acidentes de trânsito de 2011. Technical report, DETRAN-DF.
- GEREST (2012b). Boletim 2012. Technical report, DETRAN-DF. URL <http://www.detran.df.gov.br/images/Boletim%20Anual%20-%202013.pdf>. Acesso em 16 ago. 2013.
- GEREST (2012c). Informativo de cidades. Technical report, DETRAN-DF.
- GEREST (2013). Boletim mensal - dezembro de 2013. Technical report, DETRAN-DF.
- IPEA (2013). O alto custo gerado pela violência no trânsito. Technical report, Instituto de Pesquisa Econômica Aplicada. URL http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=18536&catid=159&Itemid=75. Acesso em 22 ago. 2013.

Kulldorff, M. (1997). A spatial scan statistic. *Commun. Statist. - Theory Meth.*, 26(6):1481–1496.

MS (2012). Indicadores de mortalidade. Technical report, Ministério da Saúde. URL <http://189.28.128.178/sage/>. Acesso em 22 ago. 2013.

Zheng, L., Robinson, R. M., Khattak, A., & Wang, X. (2011). All accidents are not equal: Using geographically weighted regressions models to assess and forecast accident impacts. *3rd International Conference on Road Safety and Simulation*.