

Metodologia de análise do impacto de obras de melhoria rodoviária sobre a segurança do tráfego, utilizando Sig e regressão logística multinomial

Methodology for analyzing the impacts of highway engineering interventions over traffic safety using Gis and multinomial logistic regression

DOI:10.34117/bjdv7n1-602

Recebimento dos originais: 01/01/2021

Aceitação para publicação: 22/01/2021

Rian das Dores de Miranda

Mestrando em Engenharia Civil

Universidade Federal Fluminense

Endereço: Rua Passo da Pátria, 156 - São Domingos, Niterói - RJ, CEP 24210-240

E-mail: rian_miranda@id.uff.br

Walber Paschoal da Silva

Doutor em Engenharia de Transportes

Universidade Federal Fluminense

Endereço: Rua Passos da Patria, 156, bl.D, sl.563C – São Domingos, Niteroi-RJ, CEP 24210-240

E-mail: walberpaschoal@id.uff.br

Steven Dutt Ross

Doutor em Engenharia de Produção

Universidade Federal do Estado do Rio de Janeiro (UNIRIO)

Endereço: Avenida Pasteur, Urca, Rio de Janeiro – RJ, CEP 22290-240

E-mail: steven.ross@uniriotec.br

RESUMO

Desde os anos 1950, o Brasil tem priorizado o modal rodoviário no transporte de bens e pessoas. Como resultado, sua frota rodoviária chegou a 65,8 milhões de veículos, com 96.366 acidentes registrados em 2016. Este trabalho apresenta uma metodologia para análise do impacto de obras de melhoria rodoviária sobre a segurança do tráfego, utilizando sistema de informação geográfica (SIG) e linguagem R. As análises se baseiam nos dados disponibilizados pela Polícia Rodoviária Federal (PRF), e consistem na geração de mapas de kernel, verificações de aleatoriedade espacial pelo método dos quadrantes, e regressão logística multinomial para a identificação dos fatores que influenciam a ocorrência dos acidentes. Esta metodologia foi aplicada em um trecho de 2,2 km localizado na cidade de Niterói. Os resultados revelam a existência de três pontos críticos na via, e apontam a não aleatoriedade na distribuição espacial dos acidentes. A regressão logística indica que os acidentes mais severos tendem a ocorrer após uma obra de ampliação deste trecho, realizada em 2015.

Palavras-chave: segurança do tráfego, pontos críticos de acidentes, mapa de densidade kernel, regressão logística multinomial.

ABSTRACT

Since the 1950's, Brazil has prioritized the transportation of goods and people through motorways. As a result, Brazil reached the mark of 65.8 million vehicles, and 96,366 crashes in 2016. This study presents a methodology to analyze the impacts of highway improvement constructions over the traffic safety, supported by R language and GIS. The analysis are based on the Federal Highway Police database, and consists on the generation of kernel maps, the spatial randomness verification via quadrat method, and the multinomial logistic regression to identify influent factors that cause vehicle crashes. This methodology was applied to a 2.2 km stretch at the city of Niterói. The results reveal the existence of three hotspots on the road, as well as the non-random spatial distribution of the crashes. The logistic regression indicates that high severity crashes tend to occur after a road widening construction, held in 2015.

Keywords: traffic safety, accident hotspots, kernel map, multinomial logistic regression.

1 INTRODUÇÃO

O modelo desenvolvimentista proposto pelo presidente Juscelino Kubitschek entre 1955 e 1960 transformou o transporte rodoviário no principal modal para a movimentação de bens e pessoas no território brasileiro. Segundo dados do Banco Mundial (Banco Mundial, 2013 apud BBC, 2018), o Brasil é o país que apresenta a maior concentração de transporte rodoviário de cargas e passageiros entre as principais economias do mundo. Os deslocamentos realizados por meio da infraestrutura rodoviária representam 58% das operações de transporte do país, contra 53% da Austrália, 50% da China, 43% da Rússia e 8% do Canadá. Como consequência, a frota brasileira em circulação atingiu a marca de 65,8 milhões de veículos, entre automóveis, comerciais leves, ônibus, caminhões e motocicletas (IBPT, 2018).

Com o constante crescimento da frota de veículos no Brasil, que de 2009 a 2018 passou de 59.361.642 para 100.746.553 (IBGE, 2018), considerando-se todas as categorias, observa-se um elevado número de acidentes de trânsito nos últimos anos. Segundo o estudo realizado pelo Ministério dos Transportes (2018), o número total de acidentes registrados nas rodovias federais em 2016 foi de 96.366, os quais resultaram em 6.398 vítimas fatais e cerca de 87.000 feridos. O estudo ainda aponta que 30,3% dos acidentes fatais ocorreram devido ao desrespeito às normas de trânsito.

Para auxiliar no processo de análise de segurança rodoviária, a linguagem de programação R e os sistemas de informação geográfica (SIG) surgem como importantes ferramentas de apoio. A integração destas duas tecnologias permite a representação de informações espacialmente distribuídas na forma de mapas temáticos, auxiliando a identificação das localidades onde os acidentes de trânsito apresentam maior recorrência

e severidade, denominados de pontos críticos. Ademais, a linguagem R viabiliza o desenvolvimento de modelos lineares generalizados, que auxiliam na compreensão dos fatores influentes à ocorrência dos acidentes.

Assim, o objetivo deste trabalho é apresentar uma metodologia para a investigação da influência de uma obra de melhoria rodoviária sobre a segurança de tráfego da via. Para tanto, os seguintes objetivos específicos serão contemplados:

- Localização dos pontos críticos nos períodos anterior e posterior à obra, por meio da criação de mapas de densidade kernel;
- Identificação da aleatoriedade de distribuição espacial dos acidentes antes e após a obra, através do método de contagem de quadrantes;
- Identificação dos fatores que influenciam a ocorrência dos acidentes no trecho em análise através de modelos de regressão logística multinomial.

Espera-se que esta metodologia contribua com o trabalho de gerenciamento de segurança rodoviária realizado pelas agências de transporte, concessionárias e agentes de segurança pública. Os procedimentos apresentados nesta metodologia buscam auxiliar os órgãos competentes tanto na identificação de pontos críticos, quanto no reconhecimento das causas principais dos acidentes registrados na via. Estas informações são fundamentais para o estabelecimento de políticas que busquem a redução do número de acidentes de trânsito, bem como a atenuação de seus danos, beneficiando assim toda a sociedade civil.

Quanto à organização do artigo, na seção 2 é conduzida a revisão bibliográfica na qual se baseia o desenvolvimento deste trabalho. A seção 3 apresenta a metodologia adotada para a identificação de pontos críticos, e estimativa de impactos da intervenção sobre a segurança da via. Na seção 4 é conduzida a discussão e análise dos resultados obtidos pela aplicação da metodologia em um corredor rodoviário de 2,2 km de extensão em Niterói, no estado do Rio de Janeiro. Por fim, a seção 5 apresenta as conclusões e direções para estudos futuros.

2 REVISÃO BIBLIOGRÁFICA

Atualmente existem diversas normas e manuais, desenvolvidos por agências de transportes de todo o mundo, que auxiliam o reconhecimento de pontos críticos de acidentes de trânsito. Nos Estados Unidos, a principal referência nesta área é o Highway

Safety Manual, que apresenta métodos para a estimação da frequência e da severidade dos acidentes de trânsito (AASHTO, 2014). No Brasil, o Ministério dos Transportes (2002) publicou o manual intitulado “Programa PARE - Procedimentos para o Tratamento de Locais Críticos de Acidentes de Trânsito”. Neste documento, foi realizada uma revisão dos métodos de identificação de pontos críticos apresentados pelo DENATRAN em 1987, como a técnica da severidade de acidentes, com o objetivo de adequar sua aplicação aos dias atuais.

A fim de auxiliar o processo de identificação dos pontos críticos e entender suas possíveis causas, a tecnologia SIG vem sendo amplamente empregada nas análises de acidentes de trânsito em todo o mundo. No estudo realizado por Çodur *et al.* (2014), o SIG foi utilizado tanto para detectar o período do ano com a maior ocorrência de acidentes, quanto para verificar a correlação entre as características do traçado e o número de acidentes registrados ao longo da via. Já Truong e Somenahalli (2011) utilizaram o SIG para investigar o risco de atropelamentos de pedestres em pontos de ônibus na região metropolitana de Adelaide, na Austrália.

Desde o início dos anos 2000, o SIG tem sido amplamente adotado no desenvolvimento de mapas de densidade kernel para o auxílio à identificação de pontos críticos. O mapa de densidade kernel constitui de um método estatístico de estimação de curvas de densidades. Anderson (2006) destaca que os mapas de kernel fornecem um resultado satisfatório sob os pontos de vista estatístico e estético, utilizando as densidades para representar o padrão espacial esperado para os acidentes de trânsito. Já Teodoro *et al.* (2014) indicam que os resultados obtidos pelos estimadores de kernel são similares aos apresentados pelas metodologias do Programa PARE. Desta maneira, os autores concluem que os mapas de kernel são uma alternativa viável ao gerenciamento de segurança rodoviária.

Recentemente, grupos de pesquisa têm realizado estudos de segurança rodoviária com auxílio de SIG e regressão linear. Andrade *et al.* (2014) utilizaram mapas de kernel e modelos de regressão para verificar, tanto a localização de pontos críticos no trecho da BR-277 pertencente ao estado do Paraná, quanto o impacto das características da via sobre os acidentes fatais. O estudo apontou a concentração de acidentes em segmentos de pista dupla, revelando que áreas rurais concentram o maior número de colisões entre veículos, enquanto áreas urbanas lideram os atropelamentos de pedestres. Já Machado *et al.* (2015) utilizaram o método de kernel na identificação de pontos críticos de acidentes envolvendo pedestres e usuários de veículos não motorizados nas cidades de São Paulo e Roma. O

estudo ainda apresentou uma análise de regressão múltipla para investigar a relação entre os acidentes e as variáveis que possivelmente contribuíram para suas ocorrências, como polos geradores de viagens e informações demográficas da região.

Conforme destacado por Milton *et al.* (2008), muitas agências de transporte priorizam a compreensão da relação entre as variáveis causadoras de acidentes, e a severidade das ocorrências. Neste contexto, pesquisadores de todo o mundo tem aplicado o método de regressão logística para a modelagem dos acidentes de trânsito. Esta técnica tem aplicação satisfatória aos dados de acidentes rodoviários, já que permitem a construção de modelos matemáticos a partir de variáveis categóricas.

A regressão logística é um dos possíveis modelos lineares generalizados para variável categórica. Além da função de ligação *logit*, destacam-se o modelo *probit* e o modelo complemento *log-log*. O modelo logístico multinomial deve ser utilizado quando temos mais de duas categorias não ordenadas. Em outras palavras, mais do que as tradicionais categorias dicotômicas (também chamada de variáveis *dummys*).

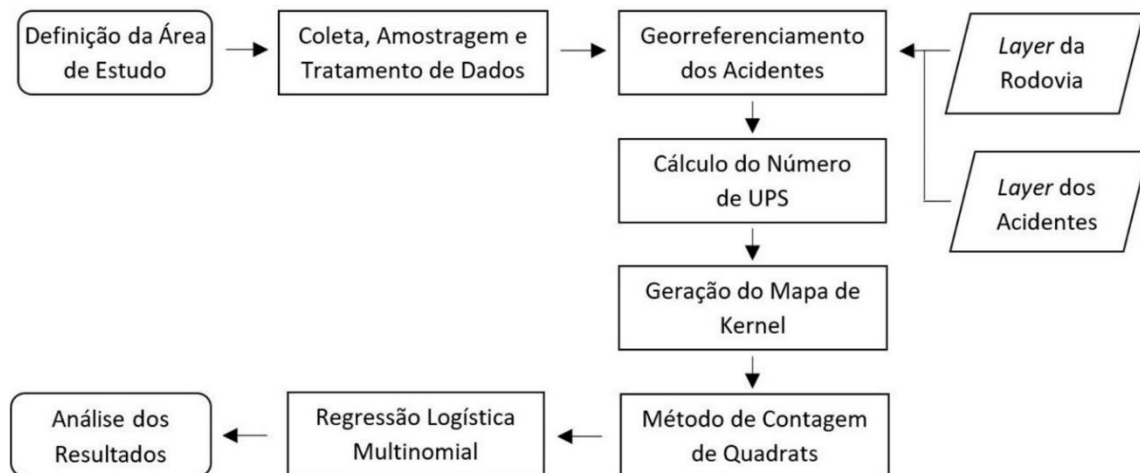
Os modelos de regressão logística múltipla se aplicam aos estudos nos quais a severidade é analisada por meio de apenas duas categorias, enquanto as variáveis explicativas podem assumir duas ou mais categorias (Kong e Yang, 2010; Alvarenga, 2015). Como exemplo, Kong e Yang (2010) adotaram o modelo de regressão logística múltipla para identificar a influência da idade do condutor e da velocidade do veículo, em acidentes envolvendo pedestres que resultem em ferimentos ou fatalidades.

Já a regressão logística multinomial tem aplicação nos casos onde a severidade pode assumir mais de duas categorias (Tay *et al.*, 2011; Coutinho *et al.*, 2015; Wahab e Jiang, 2019). Como exemplo, Tay *et al.* (2011) adotaram a regressão logística multinomial na identificação das causas de acidentes envolvendo pedestres com três níveis de severidade: ferimentos leves, ferimentos e fatalidades. O estudo concluiu que os acidentes fatais estão mais suscetíveis a ocorrência quando os condutores são do sexo masculino, estão sob efeito de bebidas alcoólicas ou conduzem veículos pesados.

3 METODOLOGIA

Este trabalho estabelece uma metodologia para a investigação da influência de uma obra de melhoria rodoviária sobre a segurança de tráfego, utilizando SIG e regressão logística multinomial. O fluxograma apresentado pela figura 1 apresenta os procedimentos adotados neste estudo.

Figure 1 –Fluxograma da metodologia



3.1 COLETA, AMOSTRAGEM E TRATAMENTO DE DADOS

Os bancos de dados dos acidentes ocorridos nas rodovias federais são coletados por meio do site oficial da Polícia Rodoviária Federal (PRF, 2019). Estes dados são contabilizados anualmente desde 2007, e registram informações importantes a respeito de cada acidente, como a data da ocorrência, a localização em função do quilômetro da rodovia, as condições meteorológicas, a fase do dia, tipo de pista, além do número de mortos, feridos e ilesos.

Com o objetivo utilizar apenas os dados referentes ao trecho e ao período em análise, realiza-se a amostragem dos dados através da linguagem de programação R. A linguagem R ainda é utilizada para a identificação e exclusão de registros que apresentem dados faltantes, conforme procedimento adotado por Milton *et al.* (2008).

Por fim, é gerada a variável explicativa “Intervenção”, que indica se o acidente ocorreu antes ou após a obra rodoviária. Para tanto, utilizasse a variável “data”, presente no banco de dados da PRF. A variável “Intervenção” assume o valor “Antes” caso o acidente tenha ocorrido no período anterior à obra, e o valor “Depois” caso a ocorrência tenha sido registrada após a obra. O quadro 1 apresenta o dicionário das variáveis adotadas nesta metodologia.

3.2 GEORREFERENCIAENTO DOS ACIDENTES

Após o cumprimento das etapas anteriores, utiliza-se o software ArcGIS para o georreferenciamento dos acidentes de trânsito. Para tanto, cria-se um SIG de trabalho onde são adicionadas as layers referentes às malhas municipal e rodoviária do Brasil, sendo a primeira coletada a partir do portal do IBGE (2015) e a segunda extraída do portal do

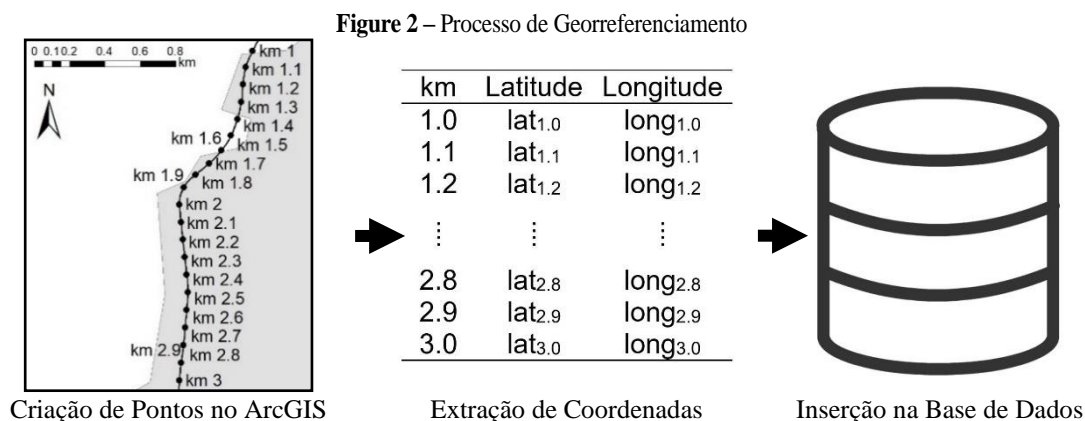
Ministério do Meio Ambiente (2008). Nesta etapa, adota-se um sistema de coordenadas projetadas.

Quadro 11: Dicionário de variáveis

Variáveis	Descrição	Categorias
Mês	Mês da ocorrência	Janeiro a Dezembro
Dia da Semana	Dia da semana da ocorrência	Segunda a Domingo
Fase do dia	Fase do dia no momento do acidente	Amanhecer, Pleno dia, Anoitecer, Plena Noite
Condição Meteorológica	Condição meteorológica no momento do acidente	Sol, Céu claro, Vento, Nublado, Chuva, Nevoeiro/Neblina, Garoa/Chuvisco
Traçado da Via	Descrição do traçado da via	Reta, Curva, Ponte, Interseção de vias, Túnel
Sentido da Via	Sentido da via onde ocorre o acidente	Crescente, Decrescente
Data	Data da ocorrência	N/A
Intervenção	Indica se o acidente ocorreu antes ou após a obra	Antes, Depois

Fonte: PRF (2017) – adaptado pelo autor

O processo de georreferenciamento é realizado com auxílio da ferramenta *Construct Points*, que permite a criação de pontos equidistantes ao longo do trecho rodoviário em análise. Para garantir que os pontos sejam dispostos sobre os locais exatos de ocorrência dos acidentes, a distância entre pontos consecutivos deve ser igual a 100 metros. Tal medida se dá pelo fato de que a PRF registra a localização dos acidentes através da variável “quilômetro”, que apresenta precisão de 100 metros. Estes pontos são utilizados como referenciais para a coleta das coordenadas dos locais de ocorrência dos acidentes, que é realizada por meio da ferramenta *add XY Coordinates* do ArcGIS. Por fim, as coordenadas são inseridas na base de dados. A figura 2 ilustra o processo de georreferenciamento dos acidentes.



3.3 CÁLCULO DO NÚMERO DE UPS

Nesta etapa, utiliza-se a técnica de severidade dos acidentes proposta pelo Programa PARE, para determinar o número de unidades padrão de severidade (UPS) em cada quilômetro do trecho em análise. Esta abordagem consiste na atribuição de pesos a cada acidente registrado na base de dados, de acordo com seu nível de severidade. Segundo o Programa PARE, os pesos devem ser atribuídos com base na seguinte regra: peso igual a 1 para acidentes somente com danos materiais; 4 para acidentes com feridos; 6 para acidentes com feridos envolvendo pedestres; e 13 para acidentes fatais. Em seguida, calcula-se o número de UPS para os locais de ocorrência de cada acidente conforme a equação 1 (BRASIL, 2002):

$$UPS = 1 \times A_{DM} + 4 \times A_{FR} + 6 \times A_{FP} + 13 \times A_{FT} \quad (1)$$

onde A_{DM} : Número de acidentes somente com danos materiais

A_{FR} : Número de acidentes com feridos

A_{FP} : Número de acidentes com feridos envolvendo pedestres

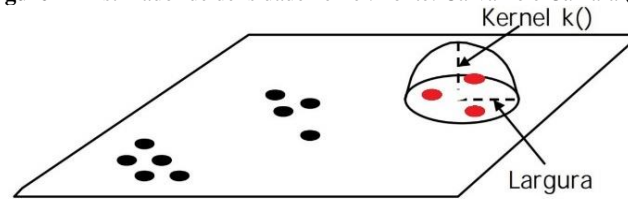
A_{FT} : Número de acidentes fatais

É importante ressaltar que, caso um acidente envolva mais de uma pessoa, sua classificação é realizada com base na vítima que apresente danos mais severos.

3.4 GERAÇÃO DO MAPA DE KERNEL

A potencialidade de ocorrência dos acidentes no trecho em estudo é determinada por meio da abordagem estatístico-espacial do estimador de densidade kernel. Segundo Carvalho e Câmara (2004), é possível descrever uma função kernel bidimensional sobre eventos espacialmente distribuídos, dando origem a uma superfície raster com valores proporcionais à quantidade de eventos observados por unidade de área. Esta função contabiliza todos os pontos situados no interior da área de influência de um local de interesse, ponderando-os com base nas suas distâncias em relação à origem. A figura 2 ilustra o estimador de densidade kernel.

Figure 2 – Estimador de densidade kernel. Fonte: Carvalho e Câmara (2004)



Desta maneira, o cálculo do estimador de densidade kernel ao longo da área de estudo é realizado por meio da equação 2 (Vargas, 2015):

$$\hat{f} = \frac{1}{nh_x h_y} \sum_{i=1}^n K\left(\frac{x - X_i}{h_x}, \frac{y - Y_i}{h_y}\right) \quad (2)$$

onde n: tamanho da amostra;

(h_x, h_y) : largura de banda nas direções x e y;

(x, y) : coordenadas do local de interesse (origem da área de influência);

(x_i, y_i) : coordenadas do evento i da amostra;

K: função de interpolação kernel do tipo gaussiana

Segundo Sheather (2004), o estimador de densidades de kernel é influenciado principalmente pelo valor da largura de banda. Vargas (2015) destaca que valores elevados de largura de banda levam à suavização excessiva das densidades. Entretanto, larguras reduzidas minimizam a área de influência, reduzindo a quantidade de eventos captados pela função kernel. Como forma de auxílio à seleção da largura de banda, Silverman (1986) introduziu a fórmula da largura de banda ótima, apresentada pela equação 3. Nos mapas de kernel, esta fórmula deve ser aplicada nas direções x e y.

$$h = 1,06 \times \min(\sigma, AIQ) n^{-1/5} \quad (3)$$

onde σ : desvio padrão das coordenadas na direção analisada;

AIQ: amplitude interquartil das coordenadas na direção analisada;

Neste estudo, são gerados mapas de kernel para a apresentação do padrão de distribuição espacial dos acidentes nos períodos anterior e posterior à execução da obra. Esta análise objetiva tanto a identificação de pontos críticos, quanto a comparação entre

suas intensidades e localizações em ambos os cenários. Destaca-se ainda que os acidentes são ponderados pelo número de UPS, a fim de amplificar os resultados do estimador kernel nas localidades impactadas por acidentes de severidade elevada.

3.5 MÉTODO DOS QUADRANTES

O método dos quadrantes é utilizado para a investigação da aleatoriedade espacial dos acidentes ao longo da rodovia. Este método consiste na divisão da área de estudo em uma grade, composta por um conjunto de quadrantes. Em seguida, é realizada a contagem do número de UPS identificado no interior de cada quadrante. Caso o processo de ocorrência de acidentes seja completamente aleatório ao longo do espaço, as contagens devem seguir uma distribuição de Poisson (Loo e Anderson, 2015). A hipótese de completamente aleatório ao longo do espaço é verificada por meio do teste qui-quadrado, ao nível de significância de 5%. A estatística de teste qui-quadrado é expressa pela equação 4, e as hipóteses do teste são apresentadas a seguir:

H_0 : Os acidentes são completamente aleatórios ao longo do espaço.

H_1 : Os acidentes não são completamente aleatórios ao longo do espaço.

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (4)$$

onde O_i : Valor da contagem no quadrante i ;

E_i : Valor esperado para a contagem no quadrante i .

Loo e Anderson (2015) destacam que o tamanho e a quantidade de quadrantes influenciam no resultado do teste qui-quadrado. A adoção de uma grade composta por poucos quadrantes de tamanho elevado, tende a concentrar um alto número de ocorrências por quadrante. Por outro lado, a escolha por muitos quadrantes de tamanho reduzido pode resultar em uma elevada quantidade de quadrantes vazios. A fim de selecionar a quantidade e o tamanho adequado dos quadrantes, esta metodologia considera a divisão da área de estudo em grades de configuração 1x3, 1x4 e 1x5. Por fim, os testes qui-quadrado são conduzidos nos cenários anterior e posterior à obra.

3.6 REGRESSÃO LOGÍSTICA MULTINOMIAL

Nesta etapa, utiliza-se a regressão logística multinomial para o ajuste de um modelo matemático que descreva o nível de severidade dos acidentes, a partir das variáveis explicativas registradas na amostra. A variável resposta adotada é o peso dos acidentes, que apresenta quatro níveis: 1) Peso = 1 (Ilesos), 2) Peso = 4 (Feridos), 3) Peso = 6 (Feridos envolvendo pedestres) e 4) Peso = 13 (Mortos). Cada acidente só pode ser classificado em uma dessas categorias, sendo a categoria Peso = 1 adotada como referência. O objetivo desta análise é identificar a influência da variável explicativa “Intervenção” sobre a severidade dos acidentes observados na via. O modelo de regressão multinomial é expresso pela equação 5 (Hosmer *et al.*, 2013).

$$g_k(x) = \ln \left[\frac{\Pr(k|x)}{\Pr(0|x)} \right] = \beta_{k0} + \beta_{k1}x_1 + \dots + \beta_{kp}x_p \quad (5)$$
$$\forall (k,i) / k \in \{0,1,2\dots K\} \cap i \in \{0,1,2\dots p\}$$

onde k: severidade do acidente;

K: número de categorias da variável severidade;

p: número de variáveis explicativas;

x: vetor que contém os valores das variáveis explicativas;

Pr (k|x): Probabilidade condicional da severidade k dado o vetor x;

Pr (0|x): Probabilidade condicional da severidade referência dado o vetor x;

β_{ki} : Parâmetro do modelo de severidade k, referente à variável explicativa i.

A construção do modelo é realizada por meio de quatro etapas apresentadas por Alvarenga (2015): 1) Método *stepwise*; 2) Teste de Razão de Verossimilhança; 3) Teste de Hosmer-Lemeshow; 4) Teste Wald. O nível de significância adotado para os testes estatísticos é igual a 5%. Assim, a hipótese nula é rejeitada caso o p-valor das estatísticas de teste seja inferior a 0,05, ou aceita caso contrário.

O método *stepwise* consiste no desenvolvimento de modelos de regressão a partir da inclusão ou exclusão de variáveis por etapas. A cada etapa do *stepwise*, o ajuste do modelo é avaliado por meio do Critério de Informação de Akaike (AIC). A estatística AIC busca penalizar modelos que apresentem um elevado número de variáveis, priorizando

modelos de complexidade reduzida e com qualidade preditiva satisfatória (Alvarenga, 2015; Murphy, 2012). A equação 6 apresenta a medida AIC:

$$AIC = -2[Log(L) - p] \quad (6)$$

onde L: Valor da verossimilhança do modelo estimado;

p: Número de parâmetros do modelo.

Segundo Alvarenga (2015), os modelos de regressão apresentam um melhor ajuste quanto menor for o valor de AIC. Deste modo, o critério de parada do método *stepwise* é satisfeito quando um modelo apresenta o valor mínimo de AIC em comparação aos modelos estimados em etapas anteriores (Burnham e Anderson, 2004; Hastie *et al.*, 2017). Neste contexto, uma nova adição ou remoção de variáveis aumentaria o AIC do modelo resultante, reduzindo assim a qualidade do ajuste.

A fim de verificar a significância estatística do conjunto de variáveis selecionadas pelo método *stepwise*, realiza-se o Teste da Razão de Verossimilhanças. Neste teste, é realizada a comparação entre a qualidade do ajuste do modelo estimado em relação ao modelo nulo. As hipóteses do teste são apresentadas a seguir, e a estatística de teste é indicada pela equação 7 (Hosmer *et al.*, 2013; Washington *et al.*, 2003).

H₀: As variáveis no modelo não são significativas.

H₁: As variáveis no modelo são significativas.

$$G = -2 \times [LL(\beta_R) - LL(\beta_U)] \quad (7)$$

onde LL(β_R): Logaritmo de verossimilhança do modelo sem as variáveis;

LL(β_U): Logaritmo de verossimilhança do modelo com as variáveis.

A qualidade do ajuste do modelo em relação aos dados da amostra é avaliada através do teste de Hosmer-Lemeshow. A estatística adotada neste teste é o qui-quadrado de Pearson, apresentado pela equação 8, e as hipóteses do teste são apresentadas a seguir (HOSMER *et al.*, 2013; Fagerland *et al.*, 2008).

H_0 : O modelo se ajusta bem aos dados.

H_1 : O modelo não se ajusta bem aos dados.

$$\chi^2 = \sum_{j=1}^J \sum_{k=0}^{K-1} \frac{(O_{jk} - E_{jk})^2}{E_{jk}} \quad (8)$$

onde J: Número de grupos no qual a amostra é dividida;

O_{jk} : Frequência observada para a severidade k no grupo j;

E_{jk} : Frequência esperada para a severidade k no grupo j.

Por fim, realiza-se o teste Wald com o objetivo de identificar os parâmetros β_{ki} , bem como a significância estatística de cada variável do modelo. A estatística de teste segue a distribuição normal padrão, e é apresentada pela equação 9. As hipóteses do teste são apresentadas a seguir:

H_0 : $\beta_{ki} = 0$

H_1 : $\beta_{ki} \neq 0$

$$W_{ki} = \frac{\hat{\beta}_{ki}}{SE(\hat{\beta}_{ki})} \quad (9)$$

onde $\hat{\beta}_{ki}$: Parâmetro estimado referente à severidade k e à variável explicativa i;

$SE(\hat{\beta}_{ki})$: Erro padrão de β_{ki} .

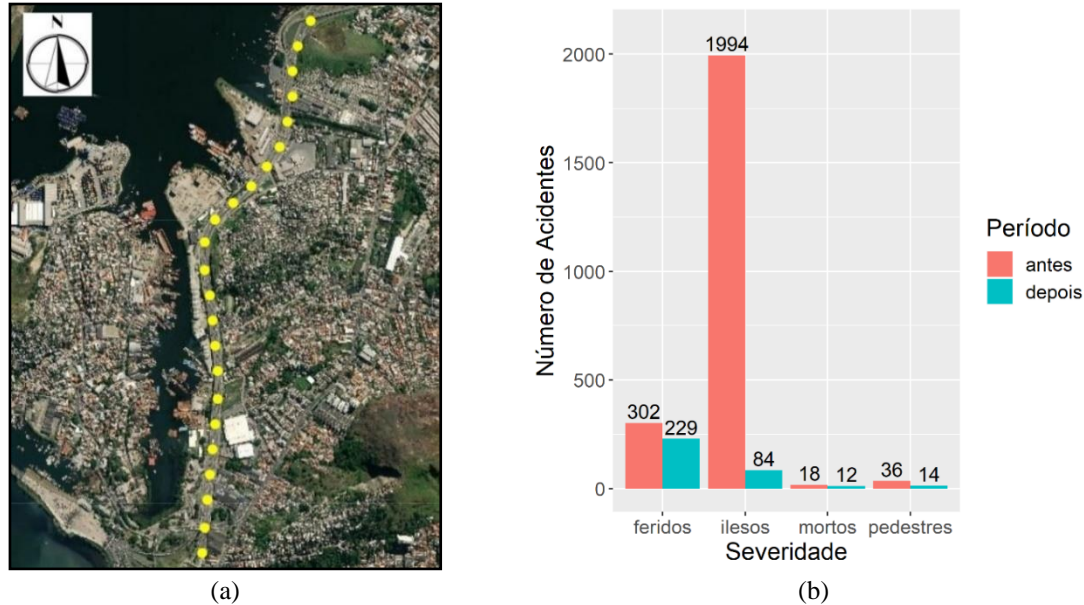
4 RESULTADOS E DISCUSSÃO

A metodologia apresentada na seção 3 é aplicada no trecho entre os quilômetros 319,8 e 322 da Avenida do Contorno, na cidade de Niterói, no estado do Rio de Janeiro. Este trecho é pertencente à BR-101, e passou por uma obra de ampliação inaugurada em 7 de agosto de 2015. Esta obra consistiu no aumento do número de faixas de duas para três em ambos os sentidos da via, além da construção de acostamentos e da instalação de luminárias de LED (GRANDES CONSTRUÇÕES, 2015; ARTERIS, 2015). A figura 4a apresenta o trecho de estudo com os acidentes georreferenciados.

Os dados da PRF apontam a ocorrência de 2728 acidentes no período entre janeiro de 2011 e novembro de 2019, dos quais 39 foram excluídos da base de dados por apresentarem informações faltantes. A avaliação dos 2689 registros restantes indica que 2350 acidentes ocorreram no período anterior à obra, enquanto 339 foram observados após a inauguração da mesma. A distribuição de severidades dos acidentes no período anterior

à obra é de 1994 ilesos, 302 feridos, 36 feridos envolvendo pedestres e 18 mortos. Após a obra, foram registrados 84 ilesos, 229 feridos, 14 feridos envolvendo pedestres e 12 mortos.

Figura 4: (a) Acidentes georreferenciados no trecho em estudo; (b) Frequência dos acidentes antes e após a obra



Na sequência, aplica-se a Técnica da Severidade dos Acidentes, com o objetivo de identificar os locais que acumulam os maiores e os menores números de UPS. Esta técnica revela que os quilômetros 320, 321 e 322 apresentam os maiores números de UPS durante o período de estudo, com valores iguais a 854, 1322 e 2410 respectivamente. Nota-se ainda que o número de UPS das três localidades sofreu reduções após a execução da obra, variando de 611 para 243 no quilômetro 320, de 911 para 411 no quilômetro 321, e de 1914 para 496 no quilômetro 322. Estas reduções se devem principalmente pela queda no número de acidentes com vítimas ilesas no período posterior à obra. Por fim, os menores números de UPS foram registrados nos quilômetros 321,3 e 321,4, que contabilizaram apenas 1 unidade cada. Ambas ocorrências foram observadas no período anterior à obra.

Com base nas UPS computadas anteriormente, inicia-se o processo de geração dos mapas kernel. Para o período anterior à obra, a equação 3 resulta em valores de largura de banda iguais a 7,36m na direção x, e 147,70m na direção y, enquanto no período posterior os valores são de 10,48m na direção x e 183,57m para a direção y. Nota-se que as larguras obtidas na direção x são consideravelmente inferiores aos valores da direção y. Este fato se deve à baixa variância nas longitudes dos acidentes registrados entre os quilômetros 321 e 322, que concentram a maior parcela do total de UPS do trecho. É importante

salientar que valores baixos de largura de banda inibem o bom funcionamento da função kernel, limitando a quantidade de eventos capturados por sua área de influência, e gerando descontinuidades no mapa.

A fim de garantir a continuidade das densidades do mapa kernel, adota-se o valor de 50m de largura de banda para a direção x, tanto no período anterior, quanto no posterior à obra. Os mapas resultantes são apresentados pela figura 5.

Os resultados indicam que os pontos críticos se localizam nos quilômetros 320, 321 e 322, tanto no período anterior quanto no período posterior à ampliação. Entretanto, a intensidade do estimador kernel assume valores superiores no período anterior à obra, indicando um maior potencial para a ocorrência de acidentes em relação ao período posterior. Ambos os mapas ainda revelam que, no quilômetro 322, o estimador kernel atinge seu valor máximo, registrando valores de 0,6 no período anterior à obra, e 0,12 após a sua execução.

Na sequência, aplica-se o método dos quadrantes para a identificação da aleatoriedade espacial dos acidentes. Após a divisão da área de estudo em grades de padrão 1x3, 1x4 e 1x5, nota-se a concentração de UPS no quadrante sul da área de estudo, alavancada principalmente pelas ocorrências no quilômetro 322. Este fato é observado tanto antes quanto após a intervenção, conforme ilustrado pela figura 6. A aplicação do teste qui-quadrado resulta no p-valor igual a $2,2e^{-16}$, independentemente do padrão de grade adotado ou do período da análise. Assim, rejeita-se a hipótese nula e conclui-se que os acidentes não ocorrem de forma aleatória ao longo do trecho em estudo.

Figura 5: (a) Mapa Kernel antes da obra; (b) Mapa Kernel após a obra

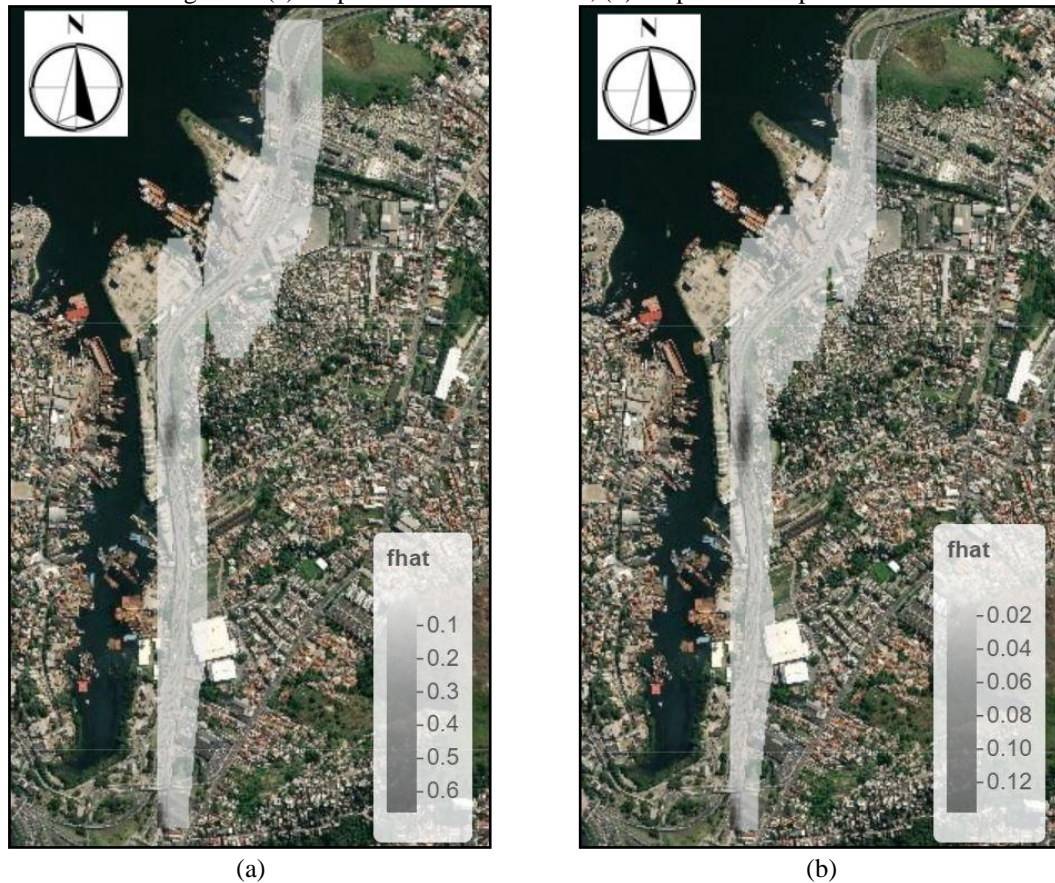
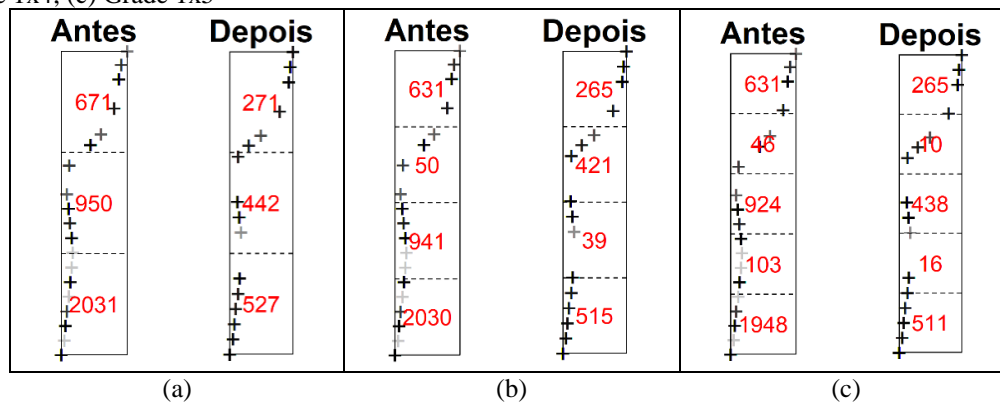


Figura 6: Método dos Quadrantes aplicado antes e depois da obra em três configurações: (a) Grade 1x3; (b) Grade 1x4; (c) Grade 1x5



Finalmente, aplica-se a regressão logística multinomial para a identificação dos efeitos da variável “Intervenção” sobre a severidade dos acidentes registrados no trecho. Seguindo a metodologia, a regressão logística é iniciada pela seleção de variáveis significativas para o modelo a partir do método *stepwise*. A aplicação deste método nas direções *forward*, *backward* e *both* resulta em modelos iguais, nos quais as variáveis explicativas são “Intervenção”, “Sentido da via” e “Fase do dia”. A tabela 1 apresenta o processo de geração do modelo através do método *stepwise forward*. Conforme discutido

na metodologia, o modelo adotado foi atingido na iteração 3, uma vez que este apresenta o menor AIC entre os modelos analisados.

Tabela 1: Resultados do Método *Stepwise Forward*

Iteração	Modelo	AIC
1	Peso ~ Intervencao	2966.09
2	Peso ~ Intervencao + sentido da via	2927.26
3	Peso ~ Intervencao + sentido da via + fase do dia	2918.55
4	Peso ~ Intervencao + sentido da via + fase do dia + traçado da via	2920.376

Na sequência, são aplicados o Teste da Razão de Verossimilhanças e o Teste de Hosmer-Lemeshow, com o objetivo de verificar, respectivamente, a significância estatística do conjunto de variáveis selecionadas e a qualidade de ajustamento do modelo. Os resultados do Teste da Razão de Verossimilhanças apontam que a estatística de teste atinge o valor $G = 579,6$, com o p -valor $= 2,2e^{-16}$. Assim, a hipótese nula é rejeitada, e conclui-se que o conjunto de variáveis selecionadas para o modelo é estatisticamente significativo. No Teste de Hosmer-Lemeshow, a estatística qui-quadrado resulta em $\chi^2 = 3,367$, com o p -valor $= 0,7615$. Desta maneira, não se rejeita a hipótese nula de que o modelo se ajusta bem aos dados da amostra.

Por fim, é realizado o Teste de Wald para a estimativa dos parâmetros do modelo, bem como para a avaliação da significância estatística individual das variáveis. Considerando o nível de significância de 5% e os resultados apresentados na tabela 2, observa-se que a variável “Intervenção/Depois” é estatisticamente significativa para os acidentes de peso 4, 6 e 13. Nos três casos, os parâmetros β_{ki} desta variável assumem valores positivos. Estes resultados sugerem que o período pós-obra tende a provocar mais acidentes com feridos, feridos envolvendo pedestres e mortos, quando comparado ao período antes da obra. Uma possível explicação para este fato reside na melhoria do fluxo de tráfego promovida pela ampliação da via, possibilitando velocidades de condução superiores em relação ao período anterior à obra, e por consequência, ocasionando acidentes de maior severidade.

O teste ainda revela que acidentes ocorridos no sentido São Gonçalo (decrecente) tendem a apresentar severidades reduzidas em relação a sentido Rio de Janeiro (crescente), sendo este resultado estatisticamente significativo para os pesos 4, 6 e 13. Já os acidentes ocorridos ao pleno dia tendem acarretar em menos mortes em relação aos acidentes ocorridos ao amanhecer.

5 CONCLUSÃO E SUGESTÕES

Este trabalho apresenta uma metodologia para a investigação dos impactos de uma obra rodoviária sobre a segurança de tráfego da via. Os objetivos do trabalho foram alcançados por meio de um estudo realizado no trecho da Avenida do Contorno, em Niterói, que passou por uma ampliação inaugurada no dia 7 de agosto de 2015. O estudo revelou a eficiência da metodologia na realização de análises de segurança rodoviária, apresentando três resultados principais: os mapas de densidade kernel antes e após a obra, os diagramas resultantes do método dos quadrantes, e o modelo de regressão logística multinomial.

Tabela 22: Resultados do Teste Wald

Peso (referência:1)	Variável	β_{ki}	$SE(\beta_{ki})$	W_{ki}	P-Valor
4	(Intercept)	-1.25561	0.2725	-4.6078	4.07e-06
	fase do dia (ref: Amanhecer)				
	Anoitecer	-0.37468	0.35748	-1.0481	0.2945849
	Plena Noite	-0.38824	0.28746	-1.3506	0.1768329
	Pleno dia	-0.12721	0.27721	-0.4589	0.646325
	sentido da via (ref: Crescente)				
	Decrescente	-0.63261	0.11512	-5.4954	3.90e-08
	Intervencao (ref: Antes)				
	Depois	2.84567	0.14332	19.8556	< 2.2e-16
	(Intercept)	-3.5451	0.73735	-4.8079	1.53e-06
6	fase do dia (ref: Amanhecer)				
	Anoitecer	0.80434	0.82843	0.9709	0.3315849
	Plena Noite	-0.39771	0.78426	-0.5071	0.6120765
	Pleno dia	-0.03467	0.74814	-0.0463	0.9630374
	sentido da via (ref: Crescente)				
	Decrescente	-0.64556	0.29669	-2.1759	0.0295634
	Intervencao (ref: Antes)				
	Depois	2.22008	0.33663	6.5951	4.25e-11
	(Intercept)	-3.08846	0.62848	-4.9141	8.92e-07
	fase do dia (ref: Amanhecer)				
13	Anoitecer	-1.24269	1.18158	-1.0517	0.2929286
	Plena Noite	-0.1926	0.65824	-0.2926	0.769832
	Pleno dia	-1.64679	0.7026	-2.3439	0.0190857
	sentido da via (ref: Crescente)				
	Decrescente	-1.39284	0.38842	-3.586	0.0003358
	Intervencao (ref: Antes)				
	Depois	2.70067	0.39637	6.8135	9.53e-12

Os mapas kernel revelaram que os pontos críticos do trecho analisado se localizam nos quilômetros 320, 321 e 322, tanto antes quanto após a realização da obra. O padrão de densidades apresentado pelos mapas revela que, nos pontos críticos, o estimador kernel atinge valores superiores no período anterior à obra quando comparado ao período posterior. Este resultado indica que a intervenção contribuiu para a queda no número de UPS registrado nos locais críticos, dispersando a distribuição de UPS ao longo do trecho. Já o método dos quadrantes revelou que a distribuição espacial dos acidentes ocorre de forma não aleatória, tanto antes, como após a obra. Os diagramas apresentados pela figura 6 sugerem uma concentração de UPS nos quadrantes que contém os pontos críticos.

A regressão logística multinomial apontou que as variáveis “Intervenção”, “sentido da via” e “fase do dia” são estatisticamente significativas para a modelagem da severidade dos acidentes. Esta análise identificou que, após a intervenção, houve uma tendência de crescimento no número de acidentes com feridos, feridos envolvendo pedestres e mortos. Possivelmente, o aumento na severidade dos acidentes ocorreu devido a uma elevação na velocidade de condução dos veículos, promovida pela melhoria no fluxo de tráfego após a implantação da faixa de rolamento adicional em ambos sentidos da via. Ademais, o sentido São Gonçalo tende a registrar menos acidentes de peso 4, 6 ou 13 em relação ao sentido Rio de Janeiro, indicando um comportamento mais arriscado entre os condutores que trafegam em direção à capital do estado. Já a fase “Pleno dia” tende a reduzir a ocorrência de mortes em comparação com o amanhecer, o que pode ser explicado pelos diferentes níveis de iluminação natural em ambos os períodos do dia.

Com base nestes resultados, sugere-se a atenção das autoridades de tráfego para a fiscalização dos pontos críticos, em especial no quilômetro 322, por apresentar o maior número de UPS registrado no trecho. Para os estudos futuros, é recomendada a coleta de dados relativos à velocidade de tráfego nos períodos anterior e posterior à obra. Esta medida possibilitaria a avaliação da importância da velocidade dos veículos sobre a severidade dos acidentes observados no trecho, contribuindo para tomadas de decisão por parte das concessionárias, como a instalação de radares ou redução do limite de velocidade.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior -Brasil (CAPES) - Código de Financiamento 001.

REFERÊNCIAS

- AASHTO. Highway Safety Manual User Guide. **National Cooperative Highway Research Program 17-50**. Washington, DC. 2014. Disponível em: <http://onlinepubs.trb.org/onlinepubs/nchrp/docs/nchrp17-50_userguide.pdf> Acesso em: 28 de abril de 2019.
- Alvarenga A. M. T. **Modelos lineares generalizados: aplicação a dados de acidentes rodoviários**. Dissertação de Mestrado. Faculdade de Ciências, Universidade de Lisboa. Lisboa, Portugal. 2015.
- Anderson, T. K. Comparison of spatial methods for measuring road accident 'hotspots': a case study of London. **Journal of Maps**. v. 16, n. 2, p. 55-63. 2006. <https://doi.org/10.1080/jom.2007.9710827>
- Andrade L.; Vissoci J. R. N.; Rodrigues C. G.; Finato K.; Carvalho E; Pietrobon R.; Souza E. M.; Nihei O. K; Lynch C. e Carvalho M. D. B. Brazilian Road Traffic Fatalities: A Spatial and Environmental Analysis. **PLoS ONE**: e87244, v. 9, n. 1, p. 1-9. 2014. <https://doi.org/10.1371/journal.pone.0087244>
- ARTERIS. Av. do Contorno de Niterói | BR-101. Rio de Janeiro. 2015. **Youtube**. Disponível em: <https://www.youtube.com/watch?v=G-nCt7KV9_A>. Acesso em: 16/03/2020
- Mesquita, L. Crise revela dependência de transporte rodoviário que é 'mais barato e dá voto'. **BBC Brasil**, Londre, 24 Maio 2018. Disponível em: <<https://www.bbc.com/portuguese/brasil-44247460>> Acesso em: 20 de abril de 2019.
- Brasil. Manual para o Tratamento de Locais Críticos de Acidentes de Trânsito. Ministério dos Transportes. **Programa PARE**. Brasília, DF. 2002
- Brasil. Rodovias Federais, Estaduais e Municipais do Brasil - PNLT (2008). Plano Nacional de Logística de Transportes (PNLT). **Shapefile**. Ministério dos Transportes. 2008. Disponível em: <<http://mapas.mma.gov.br/i3geo/datadownload.htm#>> Acesso em: 15 de abril de 2019
- Brasil. Avaliação das Políticas Públicas de Transportes, Segurança nas Rodovias Federais. **Ministério dos Transportes**. Brasília, DF. 2018. Disponível em: <https://www.infraestrutura.gov.br/images/2018/documentos/APT_Seguranca_Rodovias_Federals_Final_8.pdf> Acesso em: 28 de abril de 2019.
- Burnham K. P. e Anderson D. R. Multimodel Inference: Understanding AIC and BIC in Model Selection. **Sociological Methods & Research**. v. 33, n. 2, p.261-304. 2004. <https://doi.org/10.1177/0049124104268644>
- Çodur M.; Atalay A.; Tortum A.; e Doğru A. Geographical Information Systems Aided Traffic Accident Analysis Case Study: City Of Erzurum North Ring Road. In: International Conference on Traffic and Transport Engineering. Belgrade. 2014, **Anais Eletrônicos...** p. 336-342. 2014. Disponível em: < <https://trid.trb.org/view/1409523>>. Acessado em: 08/04/2019

Coutinho C. H. L.; Cunto F. J. C. e Ferreira S. M. P. Análise da severidade dos acidentes com motocicletas utilizando modelos probit e logit ordenados. **Revista Transportes**. v. 23, n. 4, p. 60-66. 2015. <https://doi.org/10.14295/transportes.v23i4.926>

Carvalho, M.S. e Câmara, G. Análise de Eventos Pontuais. In: Druck, S. e Monteiro, A.V.M. (eds.) Análise Espacial de Dados Geográficos. Brasília, **EMBRAPA**. 2004. Disponível em: <<http://www.dpi.inpe.br/gilberto/livro/analise/>> Acesso em : 20 de Mai 2019

Fagerland M. W.; Hosmer D. W. e Bofin A. M. Multinomial goodness-of-fit tests for logistic regression models. **Statistics in Medicine**. v. 27, n. 21, p. 4238-4253. 2008. <https://doi.org/10.1002/sim.3202>

GRANDES CONSTRUÇÕES. Rodovias: Nova Safra de Concessões Promete Bons Frutos. **Grandes Construções: Construção, Infraestrutura, Concessões e Sustentabilidade**. n. 62, p. 20-54. 2015. Disponível em: <https://issuu.com/sobratema/docs/gc_62/52>. Acesso em: 08/04/2019

Hastie T.; Tibshirane R. e Friedman J. **The Elements of Statistical Learning**. Ed. 2. Stanford, California: Spring. 2017.

Hosmer D. W.; Lemeshow S. e Sturdivant R. X. **Applied Logistic Regression**. Ed. 3. Hoboken, New Jersey: John Wiley & Sons, Inc. 2013

IBGE. BR_Municípios. **Shapefile**. 2015. Disponível em: <ftp://geofp.ibge.gov.br/organizacao_do_territorio/malhas_territoriais/malhas_municipais/municipio_2015/Brasil/BR/> Acesso em: 15 de abril de 2019

IBGE. Cidades – **Frota de veículos 2018**. 2018. Disponível em: <<https://cidades.ibge.gov.br/brasil/pesquisa/22/28120?ano=2018>> Acesso em: 15 de abril de 2019

IBPT. Real Frota Circulante no Brasil é de 65,8 Milhões de Veículos, Indica Estudo. **Instituto Brasileiro de Planejamento e Tributação**. São Paulo, SP. 2018. Disponível em: <http://www.automotivebusiness.com.br/abinteligencia/pdf/estudo_frota_completo.pdf> Acesso em: 20 de abril de 2019

Kong C. e Yang J. Logistic regression analysis of pedestrian casualty risk in passenger vehicle collisions in China. **Accident Analysis and Prevention**. v. 42, n. 4, p. 987-993. 2010. <https://doi.org/10.1016/j.aap.2009.11.006>

Loo B. P. Y. e Anderson T. K. **Spatial Analysis Methods of Road Traffic Collisions**. Ed. 1. Florida: CRC Press. 2015.

Machado C. A. S.; Giannotti M. A.; Neto F. C.; Tripodi A.; Persia L. e Quintanilha J. A. Characterization of Black Spot Zones for Vulnerable Road Users in São Paulo (Brazil) and Rome (Italy). **ISPRS International Journal of Geo-Information**. v. 4, n. 2, p. 858-882. 2015. <https://doi.org/10.3390/ijgi4020858>

Milton C. J.; Shankar V. N. e Mannering F. L. Highway accident severities and the mixed logit model: An explanatory empirical analysis. **Accident Analysis and Prevention**. v. 40, n. 1, p. 260-266. 2008. <https://doi.org/10.1016/j.aap.2007.06.006>

Murphy K. P. *Machine Learning: A Probabilistic Perspective*. Ed. 3. The MIT Press Cambridge, Massachusetts. London, England. 2012.

PRF. **Dados Abertos: 2011-2019**. 2019. Disponível em: <<https://portal.prf.gov.br/dados-abertos-acidentes>> Acesso em: 10 de fevereiro de 2020

PRF. **Dicionário de Variáveis – Acidentes**. 2017. Disponível em: <<https://arquivos.prf.gov.br/arquivos/index.php/s/mheQDcXUx2veHnU#pdfviewer>> Acesso em: 10 de fevereiro de 2020

Sheather S. J. Density Estimation. **Statistical Science**. v. 19, n. 4, p. 588-597. 2004. doi:10.1214/088342304000000297

Silverman B. W. **Density Estimation for Statistics and Data Analysis**. London: Chapman and Hall. 1986.

Tai R.; Choi J.; Kattan L. e Khan A. A Multinomial Logit Model of Pedestrian–Vehicle Crash Severity. **International Journal of Sustainable Transportation**. v. 5, n. 4, p. 233-249. 2011. <https://doi.org/10.1080/15568318.2010.497547>

Teodoro A. B.; Alcantara F. A. e Barbosa H. M. Comparação entre dois Métodos para Identificação de Locais Críticos de Acidentes de Trânsito. In: XXVIII Congresso de Ensino e Pesquisa em Transportes (ANPET). Curitiba. 2014, **Anais Eletrônicos...** Curitiba: ANPET 2014. p. 1-13. 2014. Disponível em: <<http://redpgv.coppe.ufrj.br/index.php/es/produccion/articulos-cientificos/2014-1/866-comparacao-entre-dois-metodos-para-identificacao-de-locais-criticos-de-acidentes-de-transito/file>>. Acesso em: 08/04/2019

Truong L. T. e Somenahalli S. V. C. Using GIS to Identify Pedestrian-Vehicle Crash Hot Spots and Unsafe Bus Stops. **Journal of Public Transportation**. v. 14, n. 1, p. 99-114. 2011. <http://doi.org/10.5038/2375-0901.14.1.6>

Vargas S. A. V. **Previsão da distribuição da densidade de probabilidade da Geração de Energia Eólica usando técnicas não paramétricas**. Tese de Doutorado. Departamento de Engenharia Elétrica, PUC-Rio. 2015.

Wahab L. e Jiang H. A multinomial logit analysis of factors associated with severity of motorcycle crashes in Ghana. **Traffic Injury Prevention**. v. 20, n. 5, p. 521-527. 2019. <https://doi.org/10.1080/15389588.2019.1616699>

Washington S. P.; Karlaftis M. G. e Mannering F. L. **Statistical and Econometric Methods for Transportation Analysis**. Ed. 2. London: Chapman and Hall/CRC. 2003.