

INVESTIGAÇÃO DA INFLUÊNCIA DE DIFERENTES ÁREAS DE AGREGAÇÃO NOS MODELOS MACROSCÓPICOS DA FREQUÊNCIA DE ACIDENTES DE TRÂNSITO

RESUMO

Os acidentes de trânsito na atualidade constituem um sério desafio global e o enfrentamento dessa problemática deve envolver estudos que permitam uma melhor compreensão dos fatores que influenciam estas ocorrências. Os estudos macroscópicos do desempenho da segurança viária (DSV) utilizam-se dados agregados em nível de área com informações sobre fatores de exposição, características da rede viária, uso do solo, fatores socioeconômicos e demográficos. Um aspecto que tem impacto no processo de modelagem macroscópica do DSV é a utilização de unidades de agregação espaciais arbitrárias em relação aos acidentes de trânsito, o que pode gerar inconsistências estatísticas, conhecida como Problema da Unidade de Área Modificável (PUAM). A identificação e a avaliação dos efeitos do PUAM tornam-se importantes na modelagem do DSV, uma vez que a inferência a partir de dados espaciais pode ser enganosa e induzir ao erro, caso não seja identificado o adequado nível de escala e zoneamento das unidades de análise espacial, de modo a refletir a heterogeneidade do fenômeno estudado. Uma maneira de gerar agregações relevantes ao fenômeno estudado é criar um sistema de zoneamento que minimize as variações intrazonais e maximize as variações interzonais. Para esse objetivo, a regionalização tem sido uma abordagem comumente utilizada, produzindo um novo conjunto de unidades de área agregadas a partir da otimização de uma função objetiva de medição de homogeneidade, mantendo-se a contiguidade dentro das unidades de áreas espaciais. Outra forma de considerar o efeito de escala do PUAM é utilizando os modelos geponderados. Embora qualquer modelagem possa ser afetada em algum grau pela escala de agregação espacial, os modelos geponderados podem ser mais robustos do que outros aos efeitos de mudanças de escala. Diante deste contexto, o objetivo geral deste projeto de pesquisa é investigar a influência de diferentes áreas de agregação nos modelos macroscópicos da frequência de acidentes de trânsito. A metodologia visando o desenvolvimento deste projeto é composta de sete atividades, a saber: 1) Consolidação de bases geográficas de áreas (setor censitário, zonas de tráfego, bairros e áreas de ponderação) com as informações das características operacionais das vias, uso do solo e características socioeconômicas; 2) Avaliação da presença do PUAM em diferentes unidades espaciais a partir da calibração de modelos de regressão Binomial Negativa para as unidades espaciais consolidadas e avaliação da variação dos coeficientes destes modelos; 3) Definição da zona de análise de acidentes (ZAA) a partir da construção de várias configurações de zoneamento usando o processo de regionalização; 4) Calibração e análise dos modelos geponderados com distribuição Poisson e Binomial Negativa para avaliar o impacto do efeito de escala nos modelos geponderados. Espera-se que compreender melhor essa questão melhorará a precisão dos modelos macroscópicos de previsão de acidentes e consequentemente a eficiência na alocação de recursos públicos para à redução do número de acidentes no ambiente urbano.

Palavras-chave

Problema da Unidade de Área Modificável – PUAM; Regionalização; Modelos geponderados.

1. INTRODUÇÃO

Os acidentes de trânsito na atualidade constituem um sério desafio global, sendo responsáveis pela morte ou por sequelas de milhões de pessoas em todo o mundo. De acordo com o relatório da Organização Mundial da Saúde (OMS, 2013), as mortes causadas por este fenômeno ocupam a oitava posição no ranking mundial, com expectativa de vir a ocupar a quinta posição no ano de 2030, tornando-se uma das maiores causas de morte da era moderna. No Brasil, de acordo com dados obtidos do seguro DPVAT, no ano de 2014, houve 763.365 indenizações pagas em decorrência de acidentes de trânsito com vítima, sendo que 52.226 correspondem a vítimas fatais (ONSV, 2014).

O enfrentamento da problemática da segurança viária deve envolver estudos que permitam uma melhor compreensão dos fatores que influenciam a ocorrência dos acidentes de trânsito, na expectativa de que sejam capazes de prever melhor e assegurar um direcionamento para a adoção de políticas e medidas eficazes destinadas à redução do número e da severidade dos acidentes. Uma ferramenta muito utilizada para investigar a segurança viária são os modelos estatísticos aplicados ao desempenho da segurança viária (DSV).

Nos modelos estatísticos do DSV, em um nível mais desagregado, utilizam-se entidades da rede, como segmentos (*links*) e interseções (nós). Esses modelos permitem avaliar o nível de segurança dessas entidades, relacionando a influência dos atributos geométricos/operacionais com a frequência e/ou severidade dos acidentes, possibilitando identificar pontos críticos e avaliar soluções de engenharia.

Recentemente, houve um esforço de pesquisas para o desenvolvimento de modelos de previsão de acidentes em uma escala mais agregada, a fim de fornecer ferramentas para o planejamento da segurança viária (SONG *et al.*, 2015). Esses estudos em nível macroscópico são baseados em várias unidades espaciais e têm como objetivo capturar a relação entre a frequência de acidente e uma série de variáveis explicativas, que podem incluir informações socioeconômicas, demográficas, geração de viagens, características da rede viária, características operacionais, uso do solo, etc. Esse nível de análise destina-se a concepção, implementação e gestão de políticas públicas para a melhoria do DSV, como alterações na legislação de trânsito, intensificação da fiscalização, elaboração de campanhas educativas e intervenções de engenharia em uma área (LEE, *et al.*, 2015).

Conforme Zeng e Huang (2014), a segurança viária é um problema microscópico e as causas diretas podem estar relacionadas a fatores micro associados a uma interseção, segmento de via ou às unidades motorista-veículo envolvidas. Por conta disso, os modelos macroscópicos não são aplicáveis quando se pretende avaliar soluções específicas na rede ou ainda como apoio a seleção das características da entidade viária (TARKO, *et al.*, 2008). Os modelos macroscópicos são de natureza preditiva e pretendem informar ao analista uma estimativa da segurança no futuro, no entanto, eles não são modelos explicativos que descrevem por que certos resultados ocorrem (WASHINGTON, *et al.* 2006).

Apesar disso, Huang *et al.* (2016) afirmam que a análise macroscópica dos acidentes é importante porque eles requerem menos dados detalhados, o que facilitaria uma investigação rápida ao monitorar a segurança em nível regional e permite estimular a redução do número de acidentes nas regiões através de programas específicos de segurança. Uma segunda vantagem dos modelos macroscópicos é que fatores, como características da rede viária, uso do solo, população e nível de renda, podem fornecer informações adicionais para

problemas que não são da engenharia de tráfego e não são facilmente integrados na modelagem microscópica.

Por conta dessas vantagens, uma variedade de unidades geográficas espaciais tem sido explorada nos estudos macroscópicos de segurança viária, como setores censitários (Abdel-Aty *et al.*, 2013), bairros (Rocha, 2015), zonas de tráfego (Hadayeghi *et al.*, 2010; Xu *et al.*, 2014; Lee *et al.*, 2014; Cai *et al.*, 2016), municípios (Aguero-Valverde e Jovanis, 2006; Li *et al.*, 2013), malha de quadriculas uniformes (Amoh-Gyimah *et al.*, 2017), etc. Um aspecto importante e que muitos estudos não levam em consideração é que quando se utilizam dados espaciais agregados, os resultados da análise podem ser influenciados diretamente pelo tamanho e a configuração das unidades espaciais, inconsistência esta usualmente denominada de Problema da Unidade de Área Modificável (PUAM).

As zonas de tráfego (ZT) são as unidades espaciais mais utilizadas nos estudos macroscópicos do DSV, pelo fato de permitir a integração de dados provenientes do processo de planejamento de transportes de forma mais fácil (ABDEL-ATY *et al.*, 2013). Entretanto, os acidentes apresentam uma heterogeneidade espacial/temporal que não são normalmente consideradas na delimitação das ZT (Hung *et al.*, 2013). Conforme O'Sullivan e Unwin (2010), se as unidades de agregação são arbitrárias em relação ao fenômeno espacial investigado, diferentes configurações espaciais da mesma área oferecerão diferentes resultados da análise estatística.

A identificação e a avaliação dos efeitos do PUAM tornam-se importantes na modelagem do DSV, uma vez que a inferência a partir de dados espaciais pode ser enganosa e induzir ao erro, caso não seja identificado o adequado nível de escala e zoneamento das unidades de análise espacial, de modo a refletir a heterogeneidade do fenômeno estudado (Huang *et al.*, 2013). Além disso, segundo Darmofal (2006), o PUAM é particularmente problemático porque implica que uma escolha inadequada da escala geográfica pode gerar dependência espacial artificial.

Sabe-se que não há uma solução global que permita eliminar os efeitos do PUAM quando se usa dados espaciais agregados, porém é sensato questionar se os resultados estatísticos mantêm-se quando as unidades espaciais têm suas agregações alteradas, ao invés de simplesmente ignorar este aspecto. Compreender melhor essa questão melhorará a precisão dos modelos macroscópicos de previsão de acidentes e consequentemente a eficiência na alocação de recursos públicos para à redução do número de acidentes no ambiente urbano.

2. OBJETIVOS DA PESQUISA

O objetivo geral deste projeto de pesquisa é investigar a influência de diferentes áreas de agregação nos modelos macroscópicos da frequência de acidentes de trânsito. Com vistas a atingir o objetivo geral estabelecido, foram definidos os seguintes objetivos específicos:

- a) Avaliar a presença do PUAM na modelagem da frequência de acidentes de trânsito em diferentes unidades espaciais existentes;
- b) Determinar o zoneamento mais adequado para a modelagem da frequência de acidentes trânsito;

- c) Calibrar e avaliar o efeito de escala em modelos geoponderados para a frequência de acidentes de trânsito;

3. REVISÃO BIBLIOGRÁFICA

3.1. O efeito PUAM em estudos do DSV em áreas

O PUAM é composto pelo efeito de escala e de zoneamento. No efeito de escala, para um mesmo conjunto de dados, diferentes resultados estatísticos podem ser obtidos quando os tamanhos das unidades espaciais mudam devido às agregações. No efeito de zoneamento, resultados estatísticos diferentes podem ocorrer quando se alteram as fronteiras das unidades espaciais, mantendo-se o mesmo número de áreas (FOTHERINGTHAM *et al.*, 2002). De acordo com Briant *et al.* (2010), o efeito de escala é reconhecidamente o componente mais problemático, enquanto o zoneamento é de menor importância. Para abordar o PUAM, Manley (2013) recomenda um procedimento no qual o primeiro passo é definir a adequada escala geográfica e em seguida avaliar possíveis configurações de zoneamento desta escala escolhida.

Anselin (1988) afirma que a menos que exista um processo espacial homogêneo dos dados, qualquer agregação tenderá a ser enviesada. Por conta disso, é necessário que ao se utilizar dados agregados em unidades geográficas de área, o analista adote procedimentos que permitam mitigar os impactos ocasionados pela agregação espacial nos resultados do estudo.

De acordo com Páez e Scoot (2004), a única forma para que os resultados de uma análise espacial não sofram influência do sistema de zoneamento é utilizando dados individuais. Estes dados diferem dos dados agregados em área e possuem suas posições representadas pelas coordenadas geográficas (latitude e longitude), mas, estas informações nesse nível são raramente disponíveis. Viegas *et al.* (2009) afirmam que uma forma de mitigar os efeitos do PUAM é utilizar dados a partir da menor divisão espacial disponível e realizar agregações dessas unidades espaciais de uma forma relevante para o fenômeno estudado, avaliando repetitivamente os resultados para várias configurações de agregações.

Uma maneira de gerar agregações relevantes ao fenômeno estudado é criar um sistema de zoneamento que minimize as variações intrazonais e maximize as variações interzonais (WONG, 2009). Para esse objetivo, a regionalização tem sido uma abordagem comumente utilizada, produzindo um novo conjunto de unidades de área agregadas a partir da otimização de uma função objetiva de medição de homogeneidade, mantendo-se a contiguidade dentro das unidades de áreas espaciais.

A abordagem da regionalização foi utilizada no campo da segurança viária por alguns autores para desenvolver sistemas de divisão espaciais mais eficazes para a modelagem dos acidentes de trânsito a partir das ZT (HUANG *et al.*, 2013; XU *et al.*, 2014 e LEE *et al.*, 2014). Estes pesquisadores realizaram análises de sensibilidade dos coeficientes dos modelos de regressão calibrados para diferentes conjuntos de zoneamento gerado por um processo de regionalização utilizando a zona de tráfego como a menor unidade espacial de agregação. Foi avaliada a autocorrelação espacial dos resíduos dos modelos em várias escalas, sendo considerada a agregação adequada para a modelagem da segurança viária, aquela que apresentou menores variações dos coeficientes e melhores medidas de ajustes do modelo.

Conforme Xu *et al.* (2018), outra forma de considerar o efeito de escala do PUAM é utilizando os modelos geoponderados. Embora qualquer modelagem possa ser afetada em algum grau pela escala de agregação espacial, os modelos geoponderados podem ser mais robustos do que outros aos efeitos de mudanças de escala. Isso não implica que o modelo geoponderado resolva o PUAM, apenas sugere que a variação local do modelo pode não ser influenciada por questões de escala como ocorre nos modelos globais tradicionais. (FOTHERINGHAM *et al.*, 2002). Segundo os autores, esta questão deve ser aprofundada e na área da segurança viária não foi encontrado nenhum estudo que avaliasse esse aspecto.

Para os municípios brasileiros ainda não há uma indicação de qual unidade espacial é mais adequada para a modelagem macroscópica da segurança viária e a utilização de processo de agregação a partir da zona de tráfego pode gerar o inconveniente do desenvolvimento de modelos estatísticos a partir de pequenas amostras. Desta forma, mesmo com a progressão constante de inovações metodológicas utilizadas pelos modelos em nível macroscópico, que possam atenuar os efeitos do PUAM, estes aspectos não podem fundamentalmente resolver todo o problema, se o próprio sistema de zoneamento tiver limitações.

3.2. Modelos de regressão geograficamente ponderado

3.2.1. Regressão de Poisson Geograficamente Ponderada (RPGP)

Dentre os modelos geoponderados que se adequam ao fenômeno acidentes de trânsito, têm-se os Modelos Lineares Generalizados Geograficamente Ponderados (MLGGP), que conforme Chen e Yang (2012) expandem o conceito da Regressão Geograficamente Ponderada (RGP) no contexto dos Modelos Lineares Generalizados (MLG). Com os MLGGP é possível a aplicação de modelos de regressão locais, onde a variável dependente não necessita ser uma medida contínua e os termos do erro não seguem necessariamente uma distribuição normal.

Um exemplo de MLGGP é a Regressão de Poisson Geograficamente Ponderada (RPGP), desenvolvida por Nakaya *et al.* (2005), que propuseram essa técnica para relacionar o número de mortes em zonas com variáveis socioeconômicas. Especificamente para a área de segurança viária, alguns estudos já utilizaram a RPGP para investigar as variações espaciais locais em relação ao número de acidentes em áreas (HADAYEGHI *et al.*, 2010; LI *et al.*, 2013; PIRDAVANI *et al.*, 2014). O modelo RPGP proposto é definido pela Equação 1:

$$y_j \sim \text{Poisson}[\exp(\sum_k \beta_k(u_j, v_j)x_{jk})] \quad (1)$$

onde (u_j, v_j) são os locais (coordenadas) dos pontos j , para $j = 1, \dots, n$, β_k é o parâmetro relacionado a variável explicativa x_k , para $k = 1, \dots, K$ e y_j é a j -ésima variável dependente.

A ideia básica da regressão geoponderada é que os dados observados próximos ao ponto u tenham mais influência sobre a estimativa dos β_k 's do que dados localizados mais longe de u . Esta influência em torno do ponto u é descrita por uma função de ponderação. O método tenta capturar a variação espacial ajustando um modelo de regressão em cada ponto individualmente, ponderando todas as observações vizinhas por uma função da distância, denominada função de ponderação *kernel*.

Um exemplo típico de função de ponderação é a *kernel* adaptativa (Figura 1), dada pela Equação 2.

$$\text{Biquadrática: } w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{G_i}\right)^2\right]^2 & \text{se } d_{ij} < G_i \\ 0 & \text{caso contrário} \end{cases} \quad (2)$$

em que w_{ij} é o valor do peso de uma observação em j para a estimação do coeficiente em i e d_{ij} é a distância euclidiana entre os pontos i e j . O parâmetro G_i é a distância do i até o N -ésimo vizinho, obtido a partir do tamanho da largura de banda otimizada (medidas em vizinhos). A função *kernel* biquadrática adaptativa proporciona uma função de ponderação contínua semelhante à Gaussiana até a uma distância menor que G_i do ponto de regressão e depois disso, o peso considerado para os coeficientes será zero em qualquer ponto com distância maior ou igual a G_i .

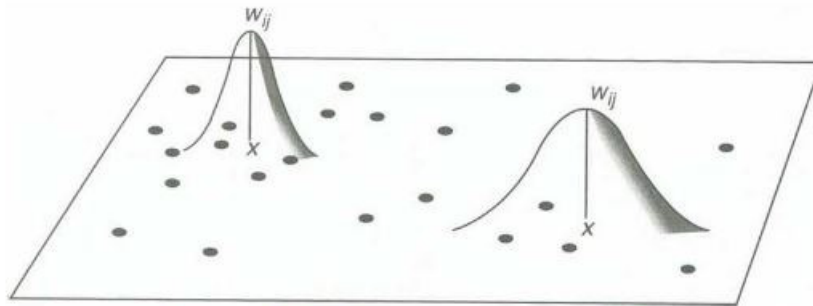


Figura 1 - Função *kernel* adaptativa

Fonte: Fotheringham (2009)

Os resultados da regressão geponderada são relativamente indiferentes à forma da função *kernel* escolhida, sendo a largura de banda o aspecto mais importante na calibração do modelo. A largura de banda pode ser considerada como um parâmetro de suavização: quanto maior a banda, maior a suavização dos coeficientes locais; por sua vez, quanto menor a banda, maior será a heterogeneidade nas respostas, pois menos observações serão usadas ao redor do ponto de observação. A largura de banda envolve um equilíbrio entre viés e variância, sendo que a escolha de uma largura de banda muito pequena leva a uma variância grande nas estimativas locais, enquanto que uma banda muito larga traz viés às estimativas locais (FOTHERINGHAM *et al.*, 2002).

Para a seleção da banda ótima é preferível utilizar o Critério de Informação de Akaike Corrigido (AICc) (FOTHERINGHAM *et al.*, 2002; NAKAYA *et al.*, 2005), pois o mesmo penaliza o número de parâmetros e consequentemente, a complexidade do modelo. Por conta disso, esse critério também pode ser utilizado para comparar modelos de diferentes variáveis explicativas ou comparar os modelos de regressão geograficamente ponderados com outros modelos candidatos.

3.2.2. Regressão Binomial Negativa Geograficamente Ponderada

Uma limitação do modelo RPGP é que não considera a superdispersão presente nos dados de acidentes. Hadayeghi *et al.* (2010), Li *et al.* (2013), Pirdavani *et al.* (2014), Xu e

Huang (2015) comentam que, apesar de ser benéfico o uso de um modelo de regressão geograficamente ponderada com uma estrutura binomial negativa, utilizaram a RPGP em seus estudos, porque a Regressão Binomial Negativa Geograficamente Ponderada (RBNGP) não estava disponível. Lord e Mannering (2010) e Xu e Huang (2015) argumentam que, embora o uso da regressão de Poisson em vez da Binomial Negativa não produza em geral estimativas imprecisas, seria possível subestimar a variância dos coeficientes, produzindo assim, mais variáveis significativas.

Para permitir a modelagem de dados discretos de forma não estacionária e considerando a superdispersão dos dados, Silva e Rodrigues (2014) desenvolveram RBNGP. A Equação 3 mostra a forma geral do modelo RBNGP.

$$y_j \sim NB[t_j \exp(\sum_k \beta_k(u_j, v_j)x_{jk}), \alpha(u_j, v_j)] \quad (3)$$

onde (u_j, v_j) são as coordenadas de localização dos pontos dos dados j , para $j=1, \dots, n$, t_j é a variável *offset*, α é o parâmetro de superdispersão, β_k são os coeficientes relacionados às variáveis explicativas x_k , para $k=1, \dots, k$, e y_j é a j -ésima variável dependente.

Recentemente Gomes *et al.* (2017) compararam o desempenho dos modelos RPGP e a RBNGP aplicados ao desempenho da segurança viária em 126 zonas de tráfego no município de Fortaleza. Os resultados mostraram que os modelos geoponderados foram superiores ao modelo tradicional Binomial Negativo, e o modelo RBNGP foi mais capaz de reduzir a autocorrelação espacial dos resíduos da frequência de acidentes que a RPGP. Entretanto, os autores não fizeram nenhuma consideração do impacto do efeito de escala sobre estes modelos geoponderados.

4. METODOLOGIA

A pesquisa a ser realizada neste projeto de pesquisa consistirá das etapas descritas a seguir.

1. Consolidação de bases geográficas de áreas: esta pesquisa utilizará dados de acidentes de trânsito, georreferenciados para os anos de 2009 a 2011, informações relativas às características físicas/geométricas das vias, características operacionais, uso do solo, disponibilizados pela Prefeitura Municipal de Fortaleza e dados socioeconômicos do Censo 2010. Estes informações serão consolidadas nas seguintes bases geográficas de áreas: setor censitário, zonas de tráfego, bairro e áreas de ponderação (IBGE).
2. Avaliação da presença do PUAM em diferentes unidades espaciais: calibrar um modelo de regressão Binomial Negativa para cada uma das unidades espaciais existentes e fazer uma análise de sensibilidade da variação dos coeficientes destes modelos. Comparar medidas de ajustes (Raiz do Erro Quadrático Médio) e a autocorrelação espacial dos resíduos dos modelos, para identificar o zoneamento que apresenta menores variações destes parâmetros. Como as medidas de ajuste são baseadas em configurações de números de zonas diferentes, para realizar uma adequada comparação, estas medidas serão transformadas em uma base espacial comum de mesmo tamanho de amostra, usando malhas de quadrículas uniformes (grid) e então comparadas.

3. Definição do zoneamento adequado para a modelagem dos acidentes: este zoneamento será denominado de zona de análise de acidentes (ZAA) e será obtido a partir da agregação de setores censitários, a partir do *software* SKATER (*Spatial “K”luster Analysis Through Edge Removal*), desenvolvido por Assunção *et al.*, (2002). O SKATER também foi implementado em biblioteca do *software* de estatística R e utiliza um conjunto de funções que permite criar clusters espaciais contíguos e internamente homogêneos em relação a algum atributo de interesse, e ao mesmo tempo heterogêneos entre as demais áreas. Para minimizar o efeito de escala a delimitação das áreas deve considerar a frequência dos acidentes, mas também será investigado o impacto de considerar variáveis independentes relacionadas às características operacionais, socioeconômica e uso do solo.
4. Calibração e análise os modelos geponderados: nesta etapa será avaliado o efeito de escala na sensibilidade dos coeficientes estimados pelo modelo geponderado com distribuição do tipo Poisson e Binomial Negativa, a partir dos seguintes passos: i) Serão escolhidos três níveis de agregação espacial (baixo, médio e alto) do ZAA; ii) Será determinado a largura de banda ótima de cada agregação com a banda *kernel* adaptativa; iii) Serão elaborados gráficos do tipo *box-plot* para mostrar a distribuição dos coeficientes e então, serão mapeadas cada variável explicativa dos modelos calibrados para analisar a variabilidade espacial.

5. PRINCIPAIS CONTRIBUIÇÕES CIENTÍFICAS

Considerando-se os acidentes de trânsito em sua dimensão social como questão de saúde pública, esse projeto apresentará impactos sociais e econômicos em nível indireto ou secundário. Os resultados da pesquisa proposta deverão contribuir significativamente para embasar o estabelecimento de políticas públicas para a gestão da mobilidade urbana mais racional e abrangente, considerando de forma objetiva a dimensão associada a segurança viária.

Ainda como contribuições específicas dessa pesquisa ressaltam-se:

- O desenvolvimento de modelos macroscópicos do DSV mais eficientes, com base nas características da rede viária e operacionais, características socioeconômicas e uso do solo.
- Aperfeiçoar a integração do componente segurança viária no processo de planejamento do sistema de transportes.
- Contribuir para a utilização, junto à comunidade técnico-científica, de ferramentas de análise espacial aplicada à problemática da segurança viária.

6. CRONOGRAMA DE EXECUÇÃO

O cronograma a seguir apresenta a sequência das atividades a serem realizadas ao longo de um ano (quatro trimestres) com início previsto para Julho de 2019.

Atividades	Semestre			
	1	2	3	4
1. Consolidação de bases geográficas de áreas	XXX			
2. Avaliar a presença do PUAM em diferentes unidades espaciais		XXX	XXX	
3. Definição do zoneamento adequado para a modelagem dos acidentes			XXX	
4. Calibrar e avaliar os modelos geoponderados				XXX

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ABDEL-ATY, M., LEE, J., SIDDIQUI, C. e CHOI, K. Geographical unit based analysis in the context of transportation safety planning. *Transportation Research Part A*, v. 49, p.62-75. 2013.
- AGUERO-VALVERDE, J.; JOVANIS, P. P. Spatial Analysis of fatal and injury crashes in Pennsylvania. *Accident Analysis and Prevention*, v.38, p.618-625. 2006.
- AMOH-GYIMAH, R., SABERI, M., SARVI, M. The effect of variations in spatial units on unobserved heterogeneity in macroscopic crash models. *Analytic Methods in Accident Research*, v.13, p.28-51. 2017.
- ANSELIN, L. *Spatial econometrics: methods and models*. Kluwer Academic Publishers, Dordrecht. 1988.
- ASSUNÇÃO, R. M.; LAGE, J. P.; REIS, E. A. Análise de conglomerados espaciais via árvore geradora mínima. *Revista Brasileira de Estatística*, v. 63, p. 7-24, 2002.
- BRIANT, A.; COMBES, P. P.; LAFOURCADE, M. Dots to boxes: do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics*, v. 67, p. 287-302, 2010.
- CAI, Q., LEE, J., ELURU, N., ABDEL-ATY, M. Macro-level pedestrian and bicycle crash analysis: Incorporating spatial spillover effects in dual state count models. *Analysis and Prevention*, v. 93, p. 14-22. 2016.
- CHEN, V. Y. J.; YANG, T. C. SAS macro programs for geographically weighted generalized linear modeling with spatial point data: Applications to health research. *Computer Methods and Programs in Biomedicine*, v. 107, n. 2, p. 262-273, 2012.
- DARMOFAL, D. *Spatial econometrics and political science*. Annual Meeting of the Southern Political Science Association, Atlanta, GA. 2006.
- FOTHERINGHAM, A.S.; BRUNSDON, C.; CHARLTON, M. *Geographically weighted regression*. Wiley, New York. 2002.
- GOMES, M. J. T. L., CUNTO, F. J. C., SILVA, A. R. Geographically weighted negative binomial regression applied to zonal level safety performance models. *Analysis and Prevention*, v. 106, p. 254-261. 2017.
- FOTHERINGHAM, A. S. Geographically weighted regression. In A. Stewart Fotheringham e Peter A. Rogerson. *The SAGE Handbook of Spatial Analysis*. SAGE Publications, Londres. 2009.
- HADAYEGHI, A., SHALABY, A.S., PERSAUD, B. Development of planning level transportation safety tools using Geographically Weighted Poisson Regression. *Accident Analysis and Prevention*, v. 42, p. 676-688. 2010.
- HUANG, H.; SONG, B.; XU, P.; ZENG, Q.; LEE, J.; ABDEL-ATY, M. Macro and micro models for zonal crash prediction with application in hot zones identification. *Journal of Transport Geography*, v.54, p. 248–256, 2016.

- HUANG, H., XU, P. e ABDEL-ATY, M. Transportation Safety Planning: a Spatial Analysis Approach. In: TRB 92th Annual Meeting (TRB 2013-1855). 2013.
- LEE, J.; ABDEL-ATY, M.; JIANG, X. Development of zone system for macro-level traffic safety analysis. *Journal of Transport Geography*, v.38, p. 13–21, 2014.
- LEE, J.; ABDEL-ATY, M.; JIANG, X. Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. *Accident Analysis e Prevention*, v.78, p. 146-154, 2015.
- LI, Z.; WANG, W.; LIU, P.; BIGHAM, J.M.; RAGLAND, D. R. Using geographically weighted poisson regression for county-level crash modeling in California. *Safety Science*, v. 58, p. 89–97, 2013.
- LORD, D.; MANNERING, F. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, v. 44, n. 5, p. 291–305, 2010.
- MANLEY, D. Scale, aggregation, and the modifiable areal unit problem. In Manfred M. Fischer e Peter Nijkamp. *Handbook of Regional Science*. Springer Heidelberg, New York. 2013.
- NAKAYA, T.; FOTHERINGHAM, A.S.; BRUNSDON, C.; CHARLTON, M. Geographically weighted poisson regression for disease association mapping. *Statistics in Medicine*. v. 24, p. 2695–2717, 2005.
- OMS – Organização Mundial da Saúde. Global Status Report on Road Safety 2013. 2013. Disponível em: www.who.int/iris/bitstream/10665/78256/1/9789241564564_eng.pdf. Acesso em: 19 de junho de 2013.
- ONSV. Relatório Anual do Observatório Nacional de Segurança Viária - 2014. Disponível em: <http://www.onsv.org.br/relatorios-anuais-do-observatorio>. Acesso em: 08 de agosto de 2016.
- O’SULLIVAN, D.; UNWIN, D. J. *Geographic information analysis*. John Wiley and Sons, 2a. Edição, New Jersey, 2010.
- PÁEZ, A.; SCOTT, D. M. Spatial statistics for urban analysis: A review of techniques with examples. *GeoJournal*, v. 61, n. 1, p. 53-67, 2004.
- PIRDAVANI, A.; BELLEMANS, T.; BRIJS, T.; WETS, G. Application of geographically weighted regression technique in spatial analysis of fatal and injury crashes. *Journal of Transportation Engineering*, v. 140, n. 8, 2014.
- ROCHA, M. M. Proposta metodológica de análise espacial de acidentes de trânsito no contexto de variáveis agregadas em áreas. Tese (Doutorado em Engenharia de Transportes) - COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, 2015.
- SILVA, A. R.; RODRIGUES, T.C.V. Geographically weighted negative binomial regression-incorporating overdispersion. *Statistics and Computing*, v. 24, p. 769-783, 2014.
- SONG, B.; HUANG, H.; ZENG, Q.; DENG, Q.; ABDEL-ATY, M. A. Comparative analysis of macro and micro models for zonal crash prediction. In: TRB 94th Annual Meeting. 2015.
- TARKO, A. P.; INEROWICZ, M.; RAMOS, J.; LI, W. Tool with road-level crash prediction for transportation safety planning. A tool for transportation safety planning with road-level crash prediction. In: Transportation Research Board Annual Meeting, Washington, D.C., 2008.
- VIEGAS, J. M.; MARTINEZ, L. M.; SILVA, E. A. Effects of the modifiable areal unit problem on the delineation of traffic analysis zones. *Environment and Planning B: Planning and Design*, v. 36, n. 4, p. 625-643, 2009.

- WASHINGTON, S., SCHALKYK, I.V., MEYER, M., DUMBAUGH, E., ZOLL, M.,. Incorporating safety into long-range transportation planning. NCHRP Report n. 546, TRB. National Cooperative Highway Research Program, Washington D.C. 2006.
- WONG, D. The Modifiable Areal Unit Problem (MAUP). In A. Stewart Fotheringham e Peter A. Rogerson. The SAGE Handbook of Spatial Analysis. SAGE Publications, Londres. 2009.
- XU, P.; HUANG, H.; DONG, N.; ABDEL-ATY, M. Sensitivity analysis in the context of regional safety modeling: Identifying and assessing the modifiable areal unit problem. *Accident Analysis and Prevention*, v. 70, p. 110–120, 2014.
- XU, P.; HUANG, H. Modeling crash spatial heterogeneity: random parameter versus geographically weighting. *Accident Analysis e Prevention*, v. 75, p. 16-25, 2015.
- XU, P., HUANG, H., DONG, N. The modifiable areal unit problem in traffic safety: Basic issue, potential solutions and future research. *Journal of Traffic and Transportation Engineering (English Edition)*, v.5 (1), p. 73-82, 2018.
- ZENG, Q.; HUANG, H. Bayesian spatial joint modeling of traffic crashes on an urban road network. *Accident Analysis and Prevention*, v. 67, p. 105–112, 2014.