



ESTIMAÇÃO DE ACIDENTES DE TRÂNSITO COM ÓBITO ATRAVÉS DE TÉCNICAS DE ANÁLISE ESPACIAL DE DADOS

Monique Martins Gomes

Cira Souza Pitombo

Agência Brasileira do ISBN
ISBN 978-85-87893-17-8



9 788587 893178



ESTIMAÇÃO DE ACIDENTES DE TRÂNSITO COM ÓBITO ATRAVÉS DE TÉCNICAS DE ANÁLISE ESPACIAL DE DADOS

Monique Martins Gomes

Cira Souza Pitombo

Universidade de São Paulo – Escola de Engenharia de São Carlos

RESUMO

O presente texto representa uma síntese do projeto de pesquisa de doutorado em Engenharia de Transportes, iniciado em 2014, cuja proposta é o estudo da distribuição espacial de acidentes com vítimas fatais, no Estado de São Paulo, partindo do pressuposto de que acidentes de trânsito são espacialmente dependentes ou correlacionados. Isto posto, o objetivo geral da pesquisa é comparar duas técnicas de análise espacial de dados, na estimação desses óbitos: Regressão Geograficamente Ponderada e Krigagem com Deriva Externa. O estudo utilizará dados de acidentes de trânsito provenientes do DETRAN – SP e Saúde Pública, associados a dados provenientes do último censo demográfico e de viagens intermunicipais (Plano Nacional de Logística e Transportes - PNLT).

1. INTRODUÇÃO

O acidente de trânsito é, em todo o mundo, visto como um dos problemas mais graves oriundos do tráfego. No Brasil, os acidentes têm assumido números muito elevados, configurando um sério problema ao país, que apresenta índices de fatalidade entre seis e sete vezes maiores que os índices de países desenvolvidos, como Alemanha, Inglaterra e Holanda, além de ser considerado o quarto país do mundo com o maior número de mortes no trânsito, ficando atrás somente da China, Índia e Nigéria (OMS, 2013).

Além das perdas irreparáveis para as famílias das vítimas, os custos oneram toda a sociedade, que sustenta, com o pagamento de impostos e contribuições, o sistema de saúde pública, responsável por grande parte do socorro às vítimas. Segundo informações da Organização Mundial da Saúde - OMS (2013), esses custos, no Brasil, superaram 30 bilhões de reais nos últimos anos, o correspondente a mais de 1,2% do PIB brasileiro. Desta forma, esforços no sentido de buscar soluções mais eficientes que tornem a rede viária mais segura e diminua o número de acidentes, tornam-se cada vez mais necessários.

Como as ocorrências de acidentes representam eventos espaciais aleatórios (alguns locais apresentam um número elevado de ocorrências num período, enquanto outros locais apresentam frequências bem abaixo das médias esperadas), Levine *et al.* (1995), mencionam que deve-se buscar um modelo regional geral para descrever geograficamente o comportamento dos acidentes, já que as suas ocorrências não podem ser representadas apenas por uma distribuição irregular espacial. Tal fato implica que os acidentes, assim como outros fenômenos espaciais, são espacialmente dependentes, ou autocorrelacionados, o que inviabiliza o uso de modelos estatísticos tradicionais, que partem do pressuposto de que os eventos observados são espacialmente independentes.

Nesse sentido, o uso de ferramentas da Estatística Espacial associada a um Sistema de Informações Geográficas – SIG, surge como uma opção, possibilitando maior compreensão de como ocorrem e quais as inter-relações entre os diversos tipos de acidentes de trânsito, que resultam em vítimas fatais.

2. OBJETIVO

O objetivo do trabalho é comparar duas técnicas distintas de análise espacial de dados, na estimação de óbitos oriundos de acidentes de trânsito no Estado de São Paulo. As técnicas a serem utilizadas são Regressão Geograficamente Ponderada e Krigagem com Deriva Externa.

3. ANÁLISE ESPACIAL DE DADOS

A análise espacial refere-se ao estudo, exploração e modelagem de processos que se expressam através da distribuição no espaço. Sua ênfase está em mensurar propriedades e relacionamentos dos dados espaciais, que por sua vez, são definidos como quaisquer dados que possam ser caracterizados no espaço, em função de algum sistema de coordenadas. Essa análise se subdivide, segundo a sua forma geométrica, em análise de superfícies (geoestatística), análise de redes, análise de padrões pontuais e análise de dados em áreas. Neste trabalho, irão ser comparadas técnicas de análise de superfícies (Geoestatística) a ferramentas de análise espacial de áreas (Regressão Geograficamente Ponderada). Tais técnicas serão aplicadas considerando coordenadas geográficas dos 645 municípios do Estado de São Paulo. Um resumo teórico de ambas as técnicas está descrito nas próximas subseções.

3.1 Geoestatística

A geoestatística faz parte de uma categoria da análise espacial de dados, que surgiu na década de 1950, quando Daniel Krige observou que a estatística básica não corresponderia às necessidades de análises de fenômenos espaciais. Diferentemente da estatística clássica, que supõe independência espacial entre os valores das variáveis, a geoestatística supõe a existência de dependência espacial entre as variáveis, ou seja, correlação dos valores das variáveis com a distância dos pontos de onde eles foram obtidos. Baseado nas publicações de Krige (1951), Matheron (1963) formulou a teoria da geoestatística, que trata as variáveis como regionalizadas.

O procedimento de Krigagem é semelhante ao de qualquer interpolação, exceto que na krigagem os pesos são determinados a partir de uma análise espacial, baseada no variograma. O variograma é um método de cálculo de variação do valor de uma variável em relação às restantes da mesma amostragem. A krigagem com deriva externa é uma extensão da krigagem com tendência, e é aplicada no caso em que a variável principal apresenta relação de dependência com variável externa. O método utiliza variáveis externas para estimar um modelo de tendência. Estas devem ser conhecidas em todo o domínio onde a variável principal será estimada, ou seja, a (s) variável (eis) auxiliar (es) deve (m) ser multicolocalizada (s). A proposta do trabalho é realizar interpolação espacial por Krigagem com Deriva Externa de óbitos provenientes de acidentes de trânsito nos municípios do Estado de São Paulo (variável primária), a partir de variáveis agregadas do último censo demográfico e dados de fluxos veiculares intermunicipais do PNLT – (variáveis secundárias ou auxiliares).

3.2 Regressão Geograficamente Ponderada

Proposto por Brunson *et al.* (1996), o modelo de Regressão Geograficamente ponderada, é basicamente uma extensão da regressão convencional, que surgiu da ideia de estudar a não estacionaridade espacial, condição na qual um modelo de regressão espacial global, não consegue explicar adequadamente. Desta forma, parâmetros locais, em vez de globais podem ser estimados (Fotheringham *et al.*, 1997). A ideia básica desse modelo é ajustar um modelo de regressão para

cada ponto no conjunto de dados, ponderando as observações por uma função de distância a este ponto. Como resultado, obtêm-se um conjunto de parâmetros ajustados para cada ponto na região geográfica analisada. A Equação 1 (Regressão Geograficamente ponderada) representa o modelo, que é reescrito em função do modelo não espacial, tornando as estimativas específicas para cada localização de i :

$$y_i = a_{0,i} + \sum_{k=1}^p a_{k,i} * x_{k,i} + \varepsilon_i \quad (1)$$

$a_{0,i}$: intercepto local específico da localização i ; $a_{k,i}$: valor do k -ésimo coeficiente angular local específico da localização i .

4. MÉTODO

Para que sejam alcançados os objetivos dessa pesquisa, pretende-se seguir uma proposta metodológica ilustrada sinteticamente na Figura 1.

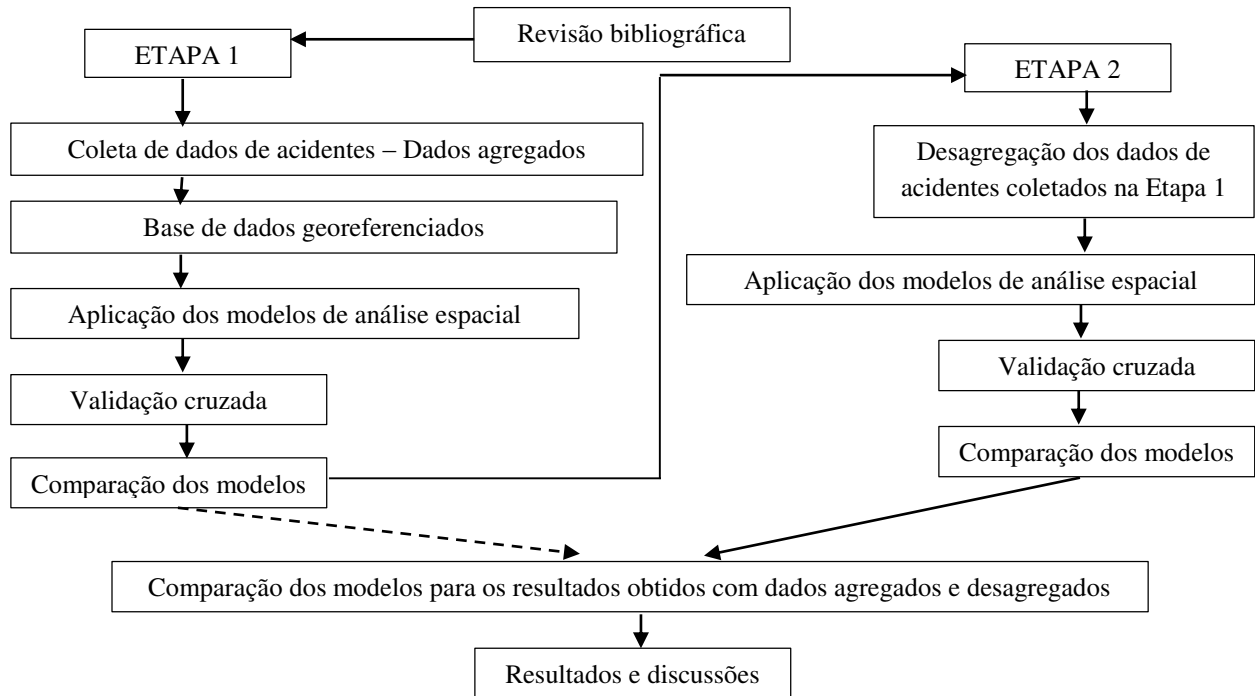


Figura 1: Método de pesquisa proposto

Inicialmente, será realizado o referencial teórico relativo às técnicas utilizadas bem como revisão da literatura subjacente à previsão de acidentes de trânsito. Em seguida, serão coletados e armazenados em um banco de dados, os totais de óbitos por acidentes para cada um dos 645 municípios que compreendem o Estado de São Paulo. Para essa primeira fase da pesquisa, pretende-se trabalhar com os dados agregados, associando a posição geográfica dos centroides dos municípios aos dados de acidentes (provenientes do DETRAN – SP e Saúde Pública) e variáveis provenientes de dados censitários – IBGE/2010, formando ao final uma base de dados georeferenciados, para municípios, em uma plataforma SIG.

Posteriormente, serão aplicados os modelos de Regressão Geograficamente Ponderada e de Krigagem com Deriva Externa, a partir da determinação das variáveis primárias (nº de óbitos por acidentes) e variáveis secundárias (dados censitários). A aplicação dos modelos de análise espacial será feita para diferentes variáveis de interesse, tais como óbitos por atropelamentos, óbitos por colisão, óbitos por gênero, etc. Como ferramenta, será utilizado um software de Regressão Geograficamente Ponderada, como GWR3. Já para a técnica de Krigagem com Deriva Externa, para que seja possível a interpolação dos dados espaciais, primeiramente serão gerados variogramas experimentais, caracterizando o comportamento das variáveis em questão. Posteriormente, esses variogramas serão ajustados, e será realizada interpolação espacial, por meio de um *software*, como o geoMS - *Geoestatistical Modelling Software*.

A acurácia de estimação das técnicas será indicada aplicando-se uma validação cruzada, com 30% dos dados, mostrando o erro de estimação considerando os pontos observados e estimados.

A primeira etapa do método será concluída através da análise de resíduos em coordenadas geográficas cujos valores de óbitos por acidentes de trânsito são conhecidos e estimados pelas duas técnicas de análise espacial de dados. Posteriormente, os mesmos procedimentos serão aplicados, em uma segunda etapa. No entanto, nessa fase da pesquisa, pretende-se trabalhar com os dados dos acidentes desagregados, de forma a aumentar a precisão do trabalho, estimando a localização de ocorrência dos acidentes segundo a proporção dos fluxos de viagens intermunicipais, concedidos pelo PNLT. Finalmente, vantagens e desvantagens do emprego de ambas as técnicas, bem como comparações com modelos não espaciais para previsão de óbitos por acidentes de trânsito, serão apontadas.

Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio e financiamento à pesquisa.

REFERÊNCIAS BIBLIOGRÁFICAS

- Brunsdon, C., A. S. Fotheringham e M. Charlton (1996) Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, 28 (4), 281-289.
- DATASUS (2010) *Informações de Saúde - Estatísticas Vitais-2010*. Disponível em: < <http://www2.datasus.gov.br> >.
- DETRAN (2010) *Acidentes de Transportes Terrestres*. Disponível em: <<http://www.observatorio.detran.sp.gov.br>>.
- Fotheringham, A. S.; M. Charlton e C. Brunsdon (1997) Recent developments in spatial analysis, Chapter measuring spatial variations in relationships with geographically weighted regression. *Springer*, New York, EUA, pages 60- 82.
- IBGE (2010) *Censo Demográfico 2010. Características urbanísticas em torno dos domicílios*. Rio de Janeiro: Ministério do Planejamento, Orçamento e Gestão, 2010.
- Krige, D.G (1951) A statistical approach to some basic mine evaluation problems on the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society*. S. Afri., Johannesburg, n.52, p.119-139, 1951.
- Levine, N; K. E. Kim e L. H. Nitz (1995) Spatial Analysis of Honolulu Motor Vehicle Crashes: I. Spatial Patterns. *Accidents Analysis and Prevention*, v.27, n.5. p. 663-674, 1995a. Disponível em: <<http://www.elsevier.com/locate/aap>>. Acesso em: 15 mai. 2014.
- Matheron, G (1963) Principles of Geostatistics. *Economic Geology*, v. 58, p. 1246–1266, 1963.
- OMS (2013) *Global Status Report on Road Safety 2013*. Supporting a Decade of Action. Luxembourg: WHO, 2013. 303 p. Disponível em <http://www.who.int/violence_injury_prevention/road_safety_status/2013/en/>. Acesso em 17 fev. 2014.