

# MODELOS ESPACIAIS DE ACIDENTES DE TRÂNSITO COM ÓBITOS

Murilo Castanho dos Santos

Cira Souza Pitombo

Universidade de São Paulo

Escola de Engenharia de São Carlos, Departamento de Engenharia de Transportes

## RESUMO

Neste relatório é apresentado o projeto de pesquisa de mestrado na área de Engenharia de Transportes. É descrito o método do trabalho, que se baseia na comparação de modelagem não espacial confirmatória (Regressão Linear Múltipla) e modelagem espacial (Regressão Geograficamente Ponderada) para previsão de acidentes com vítimas fatais no ano de 2010 no Estado de São Paulo. São apresentados resultados preliminares referentes à abordagem não espacial e mapas que corroboram a hipótese de incrementos de tais modelos a partir da incorporação do fator espacial.

## 1. INTRODUÇÃO

Desde o advento do automóvel, no início do século 20, até o ano de 2012, estima-se que da ordem de 40 milhões de pessoas no mundo tenham morrido em razão dos acidentes de trânsito (Ferraz *et al.*, 2012). De acordo com essa perspectiva, é fundamental tratar com devida atenção a identificação dos locais com maior número de acidentes de trânsito de forma a mitigar os seus efeitos.

Silva (2012) avaliou os resultados da aplicação do modelo de previsão de acidentes apresentado no *Highway Safety Manual* – HSM, publicado pela *American Association of State Highway and Transportation Officials* (AASHTO) no ano de 2010, para trechos de rodovias de pista simples, localizados no interior do Estado de São Paulo. Os valores resultantes da aplicação do método do HSM calibrado e do método de Bayes, em todas as análises realizadas, aproximaram-se dos valores observados. Em outras regiões do país esses valores podem ser empregados como primeira aproximação, pois é recomendada a realização de estudos específicos regionais para a aplicação do modelo.

Através da distribuição espacial dos locais de ocorrência dos acidentes de trânsito fatais em Campinas (São Paulo), construída mediante análise da densidade de pontos com a técnica de *Kernel*, Marín-León *et al.* (2012) observaram, entre os acidentes fatais ocorridos em rodovias, trechos com concentração de seis a dez acidentes fatais por quilômetro quadrado. No caso de acidentes não rodoviários, identificaram duas regiões que registraram densidade mais elevada (6 a 11 óbitos por quilômetro quadrado), sendo uma delas a região central da cidade.

Técnicas de análise espacial de dados podem aprimorar pesquisas em que a variável de estudo possui dependência espacial. Em geral, tais análises são sensíveis ao tipo de distribuição, à presença de valores extremos e à ausência de estacionariedade. Na análise espacial é importante também investigar *outliers* não só no conjunto dos dados, mas também em relação aos vizinhos (Câmara *et al.*, 2004).

De forma geral, acidentes de trânsito são fenômenos que possuem uma tendência ou padrão espacial. Tal fenômeno, em tese, poderia ser analisado a partir de técnicas de análise espacial de dados, tal como a Regressão Geograficamente Ponderada (RGP). Rodrigues (2012) explica que a ideia da técnica é realizar um ajuste local para cada ponto da região de estudo com base nas observações mais próximas.

## 2. OBJETIVO

O objetivo do presente trabalho é propor modelos espaciais de previsão de acidentes de trânsito com óbitos, considerando variáveis censitárias agregadas, informações relativas a fluxo veicular de viagens intermunicipais e de frota veicular.

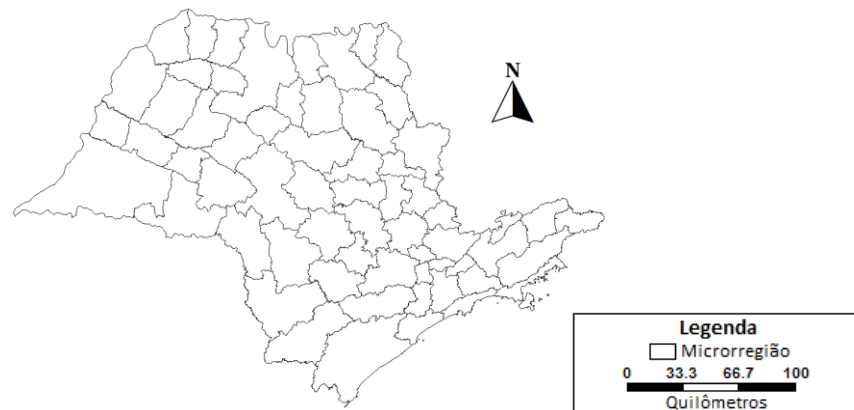
## 3. MÉTODO

### 3.1 Tratamento do banco de dados

O estudo tem como referência dados do ano de 2010 do Estado de São Paulo, agregados por microrregiões, ilustrado na Figura 1. As informações de acidentes com óbitos foram retiradas do Sistema de Informações sobre Mortalidade (SIM); a frota foi obtida do Departamento Nacional de Trânsito (DENATRAN). Os dados socioeconômicos são provenientes do Instituto Brasileiro de Geografia e Estatística (IBGE).

De acordo com o IBGE, o Estado de São Paulo é dividido em 63 microrregiões. As demais informações extraídas do IBGE foram: área, população, Produto Interno Bruto (PIB), pessoas ocupadas, pessoas com rendimento e pessoas que realizam deslocamentos intermunicipais para o trabalho. Foram utilizados óbitos por microrregiões do SIM do Departamento de Informática do SUS (DATASUS). Os óbitos de acidentes de transportes foram agregados nas seguintes categorias: pedestre, ciclista, motociclista, triciclo, automóvel, caminhonete, veículo pesado, ônibus, outros acidentes e total de acidentes. Com relação à frota foram extraídos os números de automóvel, caminhão, ônibus, motocicleta, micro ônibus.

As microrregiões do Estado de São Paulo com maior número relativo de habitantes são: São Paulo, Campinas, Osasco e Santos conforme o IBGE. A partir dessa informação utilizou-se o *software* TransCAD para obter a distância de cada microrregião para as mais populosas, através do centroide de cada uma delas. Espera-se observar uma correlação do número de acidentes com a distância até as microrregiões mais populosas.



**Figura 1:** Microrregiões do Estado de São Paulo

Barbetta (2012) explica que a distribuição normal com média 0 (zero) e desvio padrão 1 (um) é conhecida como distribuição normal padrão. Para transformar um valor  $x$ , de uma distribuição normal com média  $\mu$  e desvio padrão  $\sigma$ , em um valor  $z$  da distribuição normal padrão, basta fazer a seguinte operação:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

em que  $z$ : valor da distribuição normal padrão;  $x$ : valor observado;  $\mu$ : média;  $\sigma$ : desvio padrão.

O valor  $z$  conhecido como valor padronizado é uma medida relativa. Mede o quanto  $x$  se afasta da média ( $\mu$ ), em unidade de desvio padrão ( $\sigma$ ). Com base nessa informação, todas as variáveis foram padronizadas com a intenção de deixar todos os dados relativos, sobretudo considerando a alta variabilidade das variáveis do banco de dados.

### 3.2. Aplicação da Regressão Linear Múltipla *stepwise*: modelos não espaciais

Um propósito fundamental da Regressão Linear Múltipla (RLM) é prever a variável dependente a partir de um conjunto de variáveis independentes numéricas. A RLM fornece um meio de avaliar o poder preditivo de um conjunto de variáveis independentes (Hair *et al.*, 2009). O método *stepwise* avalia as variáveis individualmente quanto à sua contribuição na previsão da variável dependente e acrescenta ao modelo de regressão ou elimina do mesmo com base em sua contribuição relativa.

Para a análise de RLM *stepwise* desse trabalho, todas as variáveis foram padronizadas. O total de acidentes foi utilizado como variável dependente para a primeira análise. As variáveis independentes foram os dados socioeconômicos do IBGE, os dados de frota do DENATRAN e as distâncias para as microrregiões com maior número de habitantes. O modelo não espacial obtido possui muito alto poder preditivo ( $R^2 = 0,987$ ) sem intercepto. Seus principais resultados são apresentados na Tabela 1.

**Tabela 1:** Resultados da Análise de RLM *stepwise*

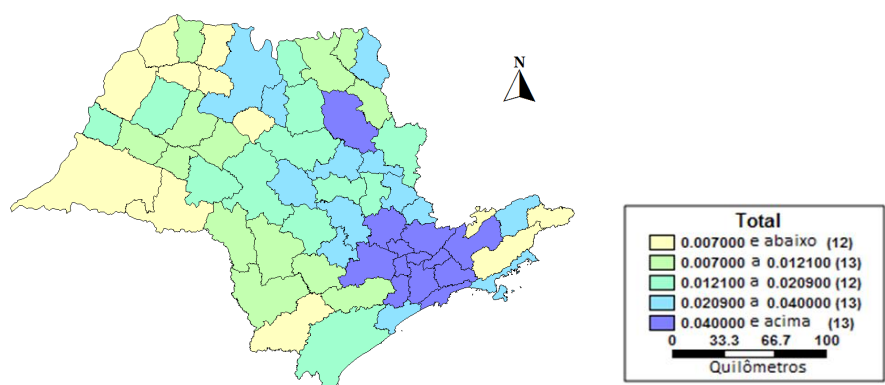
Modelo	Coeficientes		R quadrado	t	Sig.
	B	Erro Padrão			
Motocicleta	0,975	0,15	0,987	64,345	0,000
Dist_Santos	-0,067	0,016		-4,101	0,000
Área	0,050	0,016		3,206	0,002

Motocicleta: valor padronizado da frota de motocicleta da microrregião; Dis\_Santos: valor padronizado da distância a Santos baseado no centroide da microrregião; Área: valor padronizado da área da microrregião.

Os resultados mostram que os acidentes com óbitos estão diretamente ligados à frota de motocicletas, à distância para Santos e à área de cada microrregião. Verifica-se ainda que a seleção da variável distância para Santos já é um indício de tendência espacial do fenômeno. Observa-se, pelos valores de parâmetro calibrado e estatística  $t$ , que o número de acidentes aumenta, conforme proximidade da microrregião de Santos (SP). Vale ressaltar que foi realizada análise de multicolinearidade e demais suposições da técnica de RLM.

### 3.3. Análise exploratória espacial

A relação do número de acidentes de trânsito com óbitos e sua localização espacial é a hipótese principal que justifica esse projeto. O mapa apresentado na Figura 2 representa a distribuição espacial do total de acidentes por microrregião. Observa-se uma tendência do aumento do número de acidentes em direção à Microrregião correspondente a São Paulo.



**Figura 2:** Relação de acidentes com óbitos pela área

É previsto que a análise espacial exploratória seja complementada pelos cálculos de índice de Moran e gráfico de espalhamento de Moran, por exemplo.

### 3.4. Aplicação da Regressão Geograficamente Ponderada (RGP)

A ideia básica da RGP é ajustar um modelo de regressão para cada ponto no conjunto de dados, ponderando as observações por uma função de distância a este ponto. Isto corresponde a considerar que pontos mais próximos ao ponto em estudo tenham maior influência nos parâmetros estimados da regressão do que observações obtidas em pontos mais distantes (Ribeiro, 2012).

Considerando as variáveis utilizadas na RLM *stepwise* e as coordenadas dos centroides das microrregiões, a RGP será aplicada a fim de verificar a autocorrelação espacial da variável acidentes do Estado de São Paulo.

### 3.5. Comparação dos resultados

Para comparação das técnicas espaciais e não espaciais, será realizada, com 30% da amostra total, a validação dos modelos. Calculou-se, então, a correlação entre valores observados e estimados, a média dos erros e a variância dos resíduos dos modelos analisados, mostrando a qualidade do ajuste, além de prováveis incrementos através da RGP. A Tabela 2 apresenta a validação já realizada através da RLM *stepwise*.

**Tabela 2:** Resultados da Regressão Linear Múltipla

Correlação	Média dos erros	Variância dos erros
0,994	0,000	0,013

### AGRADECIMENTOS

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

### REFERÊNCIAS BIBLIOGRÁFICAS

- Barbetta, P. A. (2012) *Estatística aplicada às Ciências Sociais* (8ª ed. rev.). Editora da UFSC, Florianópolis.
- Câmara, G.; Druck, S.; Carvalho, M. S.; Monteiro, A. V. M. (2004) *Análise Espacial de Dados Geográficos*. EMBRAPA, Brasília.
- DEPARTAMENTO NACIONAL DE TRÂNSITO. (2010) *Frota de veículos*. Disponível em: <www.denatran.gov.br>. Acesso em: 29 abr. 2014.
- Ferraz, A. C. P.; Raia Júnior, A. A.; Bezerra, B. S.; Bastos, J. T.; Silva, K. C. R. (2012) *Segurança Viária* (1ª ed.). Suprema Gráfica e Editora, São Carlos, SP.
- Hair Jr., J. F.; Anderson, R. E.; Tatham, R. L.; Black, W. C. (2009) *Análise Multivariada de Dados* (6ª ed.). Bookman, Porto Alegre.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. (2010) *Censo Demográfico 2010*. Disponível em: <www.ibge.gov.br>. Acesso em: 15 abr. 2014.
- Marín-León, L.; Belon, A. P.; Barros, M. B. A.; Almeida, S. D. M.; Restitutti, M. C. (2012) *Tendência dos acidentes de trânsito em Campinas, São Paulo, Brasil: importância crescente dos motociclistas*. Cad. Saúde Pública, Rio de Janeiro, v.28, n.1, p. 39-51.
- Ribeiro, V. C. (2012) *Análise de demanda por transportes de passageiros via modelos de regressão georeferenciados*. Dissertação, Universidade Federal do Espírito Santo, Vitória, ES.
- Rodrigues, T. C. V. (2012) *Regressão Binomial Negativa Geograficamente Ponderada: Modelando Superdispersão Espacial*. Dissertação, Universidade de Brasília, Brasília.
- Silva, K. C. R. (2012) *Aplicação do modelo de previsão de acidentes do HSM em rodovias de pista simples do estado de São Paulo*. Dissertação, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos.
- SISTEMA DE INFORMAÇÕES SOBRE MORTALIDADE. (2010) *Estatísticas Vitais*. Disponível em: <tab-net.datasus.gov.br>. Acesso em: 15 abr. 2014.