

# Análise exploratória para a modelagem da frequência de acidentes de trânsito agregados ao nível de zonas de tráfego

Marcos José Timbó Lima Gomes<sup>1</sup>, Caio Assunção Torres<sup>2</sup>,  
Francisco Moraes de Oliveira Neto<sup>3</sup> e Flávio José Craveiro Cunto<sup>4</sup>

**Resumo:** Este trabalho apresenta uma análise exploratória das variáveis que contribuem na modelagem da segurança viária ao nível estratégico e tático de planejamento. A princípio, uma base de dados geográficos de 126 zonas de tráfego é compilada a partir de dados socioeconômicos disponíveis, características da rede e dos acidentes georreferenciados. Visando separar os efeitos parciais das variáveis sobre a ocorrência de acidentes, a análise compreende: i) uma análise fenomenológica e de correlação que resultou na redução de variáveis preditoras, e ii) uma análise de regressão binomial negativa, que identificou os preditores significativos. Dada a natureza espacial dos dados geográficos, a suposição de independência espacial entre erros modelados é analisada. Os resultados confirmaram que os acidentes estão relacionados com a renda e atividade econômica, no nível estratégico, e a atributos físicos e de operação da rede, no nível tático. A evidência da dependência espacial no nível estratégico aponta direções para trabalhos futuros.

*Palavras-chave:* modelagem de segurança viária, modelos lineares generalizados, autocorrelação espacial.

**Abstract:** This study presents an exploratory analysis of the variables that may contribute to modeling of road safety at the strategic and tactical planning level. At first, a geographic database of 126 traffic zones is compiled from available socio-economic datasets, roadway data and geocoded accidents. Aiming to correctly partial out the effects of each variable on accident frequency, the analysis comprises: i) a phenomenological and correlation analysis that yielded a reduced set of predictor variables, and ii) a Negative Binomial regression analysis that identified the significant predictors. Given the spatial nature of geographic data, the assumption of spatial independence between model errors is analyzed. The results confirmed that accidents are related to income and economic activity, at strategic level, and physical and operation attributes of the roadway system, at the tactical level. The evidence of spatial dependence at the strategic level points to directions for further work.

*Keywords:* road safety modeling, generalized linear models, spatial auto-correlation.

## 1. INTRODUÇÃO

O processo de planejamento de transporte, em nível estratégico, tem estado bastante voltado para a modelagem de cenários e avaliação de alternativas levando em consideração medidas de desempenho relacionadas à fluidez do sistema. Entretanto, observa-se que, apesar de sua reconhecida importância, o Desempenho da Segurança Viária (DSV) não está incorporado de forma objetiva e sistemática nesse processo de planejamento (ITE, 2009).

Acredita-se, pelo menos em parte, que a discrepância entre as dimensões de fluidez e segurança viária no processo de planejamento podem estar associadas aos desafios relacionados à coleta de informações e consolidação dos bancos de dados de acidentes de trânsito e em decorrência da natureza estocástica do fenômeno. Ressalta-se ainda que, os esforços de modelagem do DSV, principalmente no ambiente brasileiro, têm estado voltado para a estimação dos acidentes ao nível de entidades viárias como interseções e segmentos (Cunto et al., 2012; Claude et al., 2013), sendo portanto, pouco compatível com o nível de agregação necessário no planejamento estratégico.

O desenvolvimento de modelos de previsão de acidentes utilizando variáveis agregadas ao nível de zonas de tráfego permite a avaliação do DSV em cenários de longo prazo que envolvam mudança na população e alteração na infraestrutura e espaço construído (Washington et al., 2006). Entretanto, as principais iniciativas desse tipo de modelagem indicam problemas relacionados à etapa de escolha das variáveis explicativas e, sob a ótica fenomenológica, desafios para a estimação e incorporação do efeito da dependência espacial dos acidentes de trânsito no processo (Hedayghi et al., 2003; Washington et al., 2006).

Assim, este trabalho tem por objetivo realizar uma análise exploratória da relação, em nível global espacial, entre variáveis socioeconômicas, da demanda e da oferta dos sistemas de transporte e a frequência dos acidentes de trânsito agregada ao nível de zonas de tráfego. Espera-se que esse estudo prospectivo venha a contribuir na definição de uma metodologia abrangente para a modelagem do DSV no processo de planejamento dos sistemas de transporte aos níveis estratégico e tático.

## 2. MODELAGEM DO DSV NO PLANEJAMENTO DE TRANSPORTES

Os principais esforços para a modelagem do DSV no processo de planejamento podem ser divididos em função do nível de agregação das variáveis envolvidas. Em um nível mais desagregado procura-se o desenvolvimento de modelos de previsão de acidentes (MPA) independentes e calibrados para entidades da rede, como segmentos (arcos) e interseções (nós), os quais são analisados em conjunto ao nível de zonas de tráfego (Lord, 2000; Tarko, 2006; Fer-

<sup>1</sup> Marcos José Timbó Lima Gomes, Universidade Federal do Ceará Departamento de Engenharia de Transportes. (timbo@det.ufc.br)

<sup>2</sup> Caio Assunção Torres, Universidade Federal do Ceará Departamento de Engenharia de Transportes. (caio@det.ufc.br)

<sup>3</sup> Francisco Moraes de Oliveira Neto, Universidade Federal do Ceará Departamento de Engenharia de Transportes. (moraes@det.ufc.br)

<sup>4</sup> Flávio José Craveiro Cunto, Universidade Federal do Ceará Departamento de Engenharia de Transportes. (flaviocunto@det.ufc.br)

**Tabela 2.** Acidentes registrados e georreferenciados

<i>Local do Acidente</i>	<i>Acidentes Registrados</i>	<i>% de Acidentes Registrados</i>	<i>Acidentes Georreferenciados</i>	<i>% de Sucesso do Georreferenciamento</i>
Em interseção	9.741	41,0	9.344	95,9
Em logradouro com numeração	9.547	40,2	8.577	89,8
Em logradouro sem numeração	4.470	18,8	0	0,0
Total	23.758	100,0	17.921	75,4

reira, 2010). Nestes casos, as variáveis mais comumente utilizadas são o volume diário médio anual (VDMA), a extensão do arco e, com menos frequência, outros atributos da rede como, número de faixas de tráfego, presença de canteiro central, etc.

O domínio natural de aplicação dos modelos desagregados concentra-se na etapa de avaliação de cenários de planejamento propostos. Dentre os desafios dessa abordagem, ressaltam-se a dependência de informações obtidas por outro processo de modelagem (ex. VDMA), além do nível de confiabilidade dos modelos em decorrência da aleatoriedade relacionada à ocorrência dos acidentes de trânsito, principalmente quando analisados em um nível tão desagregado (Hadayeghi, et al., 2003; Washington et al., 2006).

Para o nível de modelagem agregado as variáveis são frequentemente relacionadas ao nível de zonas de tráfego (ZT). As variáveis explicativas estão mais relacionadas ao uso do solo e a aspectos socioeconômicos da área em estudo, o que atribui uma natureza mais preditiva e menos explicativa aos modelos (Washington et al., 2006). Além disso, esses autores ressaltam a utilidade dessa abordagem para o estabelecimento de metas de investimentos, políticas públicas para a melhoria do DSV e estimação do número previsto de acidentes em áreas em virtude da variação temporal de fatores socioeconômicos e do ambiente construído.

Hadayeghi et al. (2003) relacionaram o número total de acidentes e os acidentes com feridos graves e vítimas fatais em 463 ZT na cidade de Toronto com 22 variáveis divididas em três grupos. Dentre as principais variáveis preditoras, o estudo apontou como mais importantes o comprimento das vias, quilometragem percorrida pelos veículos e densidade de interseções nas zonas de tráfego. O estudo ressaltou ainda que a natureza espacial da maioria das variáveis preconiza a utilização de bancos de dados georreferenciados e que, o efeito da dependência/correlação espacial entre variáveis pode influenciar significativamente a qualidade dos modelos.

Aguero-Valverde et al. (2006) investigaram o risco de acidente para municípios da Pensilvânia em relação as variáveis sócio-demográficos, condições climáticas, infraestrutura de transporte e quantidade de viagens. O estudo constatou que embora os modelos de acidente em nível agregado tenham sido pensados para estarem em vantagem devido à disponibilidade de vários dados de transporte e socioeconômicas, eles sofrem com o problema da heterogeneidade espacial.

Karlaftis e Tarko (1998) se esforçaram para controlar a heterogeneidade espacial. Eles usaram técnicas de agrupamento para gerar grupos homogêneos com características semelhantes de trânsito, de infra-estrutura e sócio-econômica, para os condados de Indiana. Seus resultados indicaram que os modelos desenvolvidos para grupos homogêneos de municípios foram mais eficientes do que os modelos comuns.

Dada a carência ou inexistência de bancos de dados contendo informações sobre a demanda, alguns pesquisadores avaliaram a utilização de valores do fluxo veicular, obtidos a partir de outro processo de modelagem, como variáveis preditoras na modelagem do DSV. Lord (2000) e Hadayeghi et al. (2003) utilizaram a ferramenta EMME/2 para criar redes digitais para auxiliar a elaboração dos MPA. Os resultados mostraram que é possível prever acidentes em redes de transporte computadorizados, mas que a precisão dos MPA é diretamente relacionada com a precisão das previsões de fluxo de tráfego de planejamento de transporte do aplicativo.

A Tabela 1 apresenta a lista de variáveis envolvidas nos esforços de modelagem revisados nesse artigo. A revisão da literatura reforça a necessidade da utilização frequente de bancos de dados distintos, incluindo informações obtidas por outro processo de modelagem, obviamente sujeito a diversos outros tipos de simplificações e erros. Observou-se, ainda, a presença de diversas variáveis preditoras colineares e com aparente correlação entre si.

Ainda em relação à Tabela 1 é possível classificar as variáveis de acordo com sua utilização no processo tradicional de modelagem da demanda dentro do planejamento. As variáveis socioeconômicas são empregadas normalmente nos modelos de geração de viagens e possuem relação indireta com os acidentes. As variáveis relacionadas às características da rede, demanda de tráfego e meio ambiente, por sua natureza, ensejam a criação de modelos sensível a alterações mais localizadas na rede, compatíveis com a etapa de análise de alternativas.

Considerações mais formais e sistemáticas do caráter espacial das interações podem ser obtidas com a aplicação de modelos locais espaciais a partir de técnicas de Regressão Geograficamente Ponderada (RGP) e Geograficamente Ponderada de Poisson (RPGP). Hadayeghi et al. (2010) compararam os modelos de previsão para a cidade de Toronto com RPGP e Modelos Lineares Generalizados (MLG), assumindo distribuição binomial negativa, e constataram que os coeficientes de correlação linear dos modelos melhoraram com a utilização de RPGP diminuindo, assim, os erros médios quadrados dos mesmos em relação aos MLG.

### 3. PROCEDIMENTO METODOLÓGICO - ANÁLISE EXPLORATÓRIA

A análise exploratória proposta nesse trabalho foi realizada a partir das seguintes etapas sequenciais: i) Consolidação dos bancos de dados existentes; ii) Coleta e georreferenciamento dos acidentes de trânsito; iii) Análise fenomenológica das variáveis; iv) Desenvolvimento e avaliação de modelos exploratórios; v) Análise da dependência espacial das variáveis. A seguir serão detalhadas cada uma das etapas propostas.

**Tabela 1.** Lista de variáveis utilizadas em estudos de modelagem da segurança viária

<i>Autores</i>	<i>Variável dependente</i>	<i>Grupo</i>	<i>Variáveis preditoras</i>
Hadayeghi et al (2003)	#acidentes, # acidentes com vítimas feridas e fatais	Socioeconômico e Demográfico	População, #domicílios, #empregos, #veículos, #veículos/domicílio, densidade (população, domicílios, empregos).
		Características da rede	#Interseções, interseções/km <sup>2</sup> , km de vias principais, km de vias secundárias, km total de vias e Área.
		Demanda de Tráfego	Velocidade limite, V/C, Veículo-km e Fluxo Total.
Aguero-Valverde et al. (2006)	# acidente fatais, # acidades com feridos, # acidentes com feridos e fatais, Taxas de acidente por severidade	Socioeconômico	População por faixa etária, #Homens, População em situação de pobreza, #prisões de trânsito, taxa de população sobre influência de transportes motorizados
		Características da rede	Veículo-km, km total de vias, km vias federais, da malha, Extensão da malha federal e Densidade da malha
		Meio ambiente	Valores médios de precipitação, dias chuvosos, queda de neve e dias com chuva e neve
Lord e Persaud (2004)	#acidentes, # acidentes com vítimas fatais e feridas	Demanda de tráfego	Volume médio diário anual
		Características da rede	# Arcos, Extensão do arco, #interseções semaforizadas de 4 e 3 ramos e interseções não semaforizadas de 3 e 4 ramos.

### 3.1. Consolidação dos bancos de dados existentes

A análise utilizou inicialmente uma base georreferenciada de Fortaleza, a partir do estudo de Henrique (2004), já com informações socioeconômicas atualizadas do Censo 2010 e agregadas em 126 zonas de tráfego. Para a obtenção de variáveis da infraestrutura da rede utilizou-se a base de logradouros georreferenciada disponibilizada pelo órgão gestor do trânsito do município (AMC), que também forneceu a localização de semáforos, equipamentos de fiscalização eletrônica e informações de acidentes de trânsito, obtido do Sistema de Informações de Acidentes de Trânsito de Fortaleza (SIAT/FOR).

De posse desses dados, e utilizando ferramentas típicas de um SIG, como funções que permitem realizar consulta espacial, sobreposição de camadas e operações espaciais com base no relacionamento topológico das entidades geográficas, foi possível consolidar um banco de dados único e georreferenciado, com todas as informações agregadas por zonas de tráfego.

### 3.2. Coleta e georreferenciamento dos acidentes de trânsito

Os dados de acidentes utilizados neste estudo referem-se ao ano de 2011, que foram classificados em dois tipos: número total de acidentes e número total de acidentes com vítimas. Esses dados foram separados em 2 grupos, de acordo com os procedimentos/ferramentas utilizadas para o georreferenciamento: acidentes ocorridos em interseções e acidentes ocorridos em segmentos de logradouros.

Para a espacialização dos acidentes em interseções utilizou o software *TransCAD 5.0* (2008), juntamente com rotina em linguagem GISDK desenvolvido por Queiroz (2003). Com relação aos acidentes em logradouros, o georreferenciamento foi realizado a partir das informações do serviço web disponíveis no *Google Maps*, usando uma rotina *API do Google* gratuita.

A Tabela 2 ilustra o número total de acidentes disponíveis para o ano de 2011 e os resultados finais do processo de georreferenciamento. Apesar dos bons resultados de su-

cesso no georreferenciamento, há ainda um número considerável de acidentes em logradouros sem numeração (aproximadamente 19%), o que demonstra necessidade de melhorias na coleta dessas informações em campo.

### 3.3. Análise fenomenológica das variáveis

Neste trabalho decidiu-se por separar o processo de modelagem nos níveis de análise estratégico e tático. O primeiro diz respeito ao objetivo de avaliar políticas públicas para reduzir acidentes de trânsito. O segundo nível é relacionado ao objetivo de incluir indicadores do DSV na avaliação de cenários de intervenção em planejamento de transportes.

Ao nível estratégico, acredita-se que as variáveis estão mais relacionadas a fatores que geram viagens no meio urbano. Portanto, variáveis demográficas, de uso do solo, socioeconômicas e de acessibilidade seriam as mais indicadas. A hipótese é que em um nível agregado de zona, o número de acidentes tende a ser maior com o aumento da densidade demográfica, do poder aquisitivo, da intensidade de atividade econômica e da acessibilidade do transporte. Já ao nível tático, adotam-se fatores de risco que estão relacionados às características das vias e interseções, comportamento dos condutores e atributos operacionais como o tipo de controle de tráfego. Em relação à exposição são utilizados o volume de veículos na rede e extensão percorrida na rede. Exemplos de variáveis que podem representar bem estes fatores são: forma da rede viária, extensão e hierarquia das vias, número de interseções (controladas ou não), VDMA, equipamentos de fiscalização e nível de saturação das vias.

Devido à limitação da base de dados, algumas variáveis não estão disponíveis e outras não estão na sua forma pura. Por exemplo, renda é combinada com a variável número de domicílios. Além disso, outras variáveis podem ser derivadas de processos de modelagem, como o fluxo veicular por arco. A Tabela 3 apresenta as estatísticas descritivas das variáveis selecionadas dos bancos de dados, classificadas conforme o nível de análise.

**Tabela 3.** Estatística descritiva das variáveis

Grupos	Variáveis	Média	Mínimo	Máximo	Desvio padrão
Variáveis respostas	Total de Acidentes	140,37	2,00	934,00	152,46
	Total de Acidentes com Vítimas	46,47	1,00	256,00	39,17
Variáveis explicativas (Nível estratégico)	Consumo de Energia – Indústria	12.575,59	19,19	105.441,95	14.096,32
	Consumo de Energia - Comércio e Serviço	271.096,37	6.413,66	4.017.134,92	517.956,80
	População entre 0 e 19 anos	6.080,45	322,00	42.260,00	6.278,31
	População entre 20 e 64 anos	11.853,52	736,00	68.335,00	9.957,10
	População maior que 64 anos	1.266,79	81,00	4.622,00	866,13
	Domicílios com renda até 3 Salários	3.778,60	201,00	25.893,00	3.783,29
	Domicílios com renda acima de 3 Salários	1.087,37	31,00	7.015,00	1.101,60
Variáveis explicativas (Nível tático)	Interseções Semaforizadas	4,79	-	78,00	8,98
	Interseções não Semaforizadas	152,59	5,00	908,00	128,05
	Extensão Vias Expressas	0,75	-	6,25	1,39
	Extensão Vias Arteriais	1,90	-	9,50	1,93
	Extensão Vias Coletoras	0,45	-	4,04	0,78
	Extensão Vias Locais	26,76	2,44	162,93	22,25
	Equipamentos Fiscalização	1,87	-	14,00	2,42

As Tabelas 4 e 5 apresentam as matrizes de correlação entre todas as variáveis para os níveis estratégico e tático, respectivamente. A partir destes resultados e da análise dos gráficos de dispersão entre as variáveis, decidiu-se por agregar ou excluir algumas variáveis, com intuito de reduzir a multicolinearidade nas variáveis explicativas. Este primeiro tratamento nas variáveis visou separar os efeitos parciais de cada variável sobre a ocorrência de acidentes.

No nível estratégico, as duas variáveis de consumo de energia, separadas por tipo de setor, foram agregadas em virtude da alta correlação apresentada. Observe ainda, que as populações por idade são altamente correlacionadas entre si e apresentam alta correlação com o número de domicílios com renda até três salários. Este fato retrata bem a característica da maioria das zonas de tráfego de Fortaleza, onde a maior parte da população tem renda baixa. Assim, decidiu-se por considerar apenas a variável domicílio por renda, que caracteriza bem a realidade demográfica e de variação da renda entre zonas.

No nível tático, existe uma alta correlação entre o número de interseções não-semaforizadas e a extensão de vias locais. É razoável assumir que estas duas variáveis representem os mesmos fatores de rede que explicam os acidentes. Assim, decidiu-se excluir da análise a extensão de vias locais, que apresenta menor correlação com fenômeno de

acidentes, além de ser uma variável desconsiderada nos estudos de modelagem macro da demanda por viagens.

A Figura 1 mostra os gráficos de dispersão para duas variáveis explicativas. Em geral, os gráficos apresentam uma relação positiva entre o fenômeno de acidentes e as variáveis dependentes, como esperado. Além disso, é possível observar uma maior dispersão dos dados para maiores valores da variável dependente, evidenciando a heterocedasticidade do fenômeno de acidentes.

### 3.4. Desenvolvimento e avaliação de modelos exploratórios

Nos MPA com variáveis agregadas, observa-se um direcionamento aos modelos lineares generalizados em virtude da variável resposta ser aleatória discreta e positiva, além dos erros se apresentarem com distribuição compatível com a distribuição de Poisson ou binomial negativa (Hadayeghi, et al., 2003; Lord e Persaud, 2004; Tarko, et al., 2008). A formulação do modelo, como proposta na literatura, adotada neste estudo é apresentada a seguir:

$$E(y|x_1, x_2, \dots, x_k) = \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k) \quad (1)$$

**Tabela 4.** Correlações das variáveis de nível estratégico

	Total de Acidentes	Total de Acidentes com Vítimas	Consumo Energia na Ind.	Consumo Energia no Com. e Serv.	População entre 0 e 19 anos	População entre 20 e 64 anos	População maior que 64 anos	Dom. com renda até 3 SM	Dom. com renda acima de 3 SM
Total de Acidentes	1,00								
Total de Acidentes com Vítimas	0,90	1,00							
Consumo de Energia na Indústria	0,70	0,72	1,00						
Consumo de Energia no Comércio e Serviço	0,78	0,61	0,73	1,00					
População entre 0 e 19 anos	0,05	0,33	0,20	-0,02	1,00				
População entre 20 e 64 anos	0,18	0,43	0,30	0,08	0,98	1,00			
População maior que 64 anos	0,46	0,60	0,45	0,27	0,73	0,85	1,00		
Domicílios com renda até 3 Salários	0,03	0,34	0,23	-0,05	0,99	0,97	0,74	1,00	
Domicílios com renda acima de 3 Salários	0,69	0,50	0,40	0,54	0,12	0,29	0,57	0,05	1,00

**Tabela 5.** Correlações das variáveis de nível tático

	Total de Acidentes	Total de Acidentes com Vítimas	Interseções SemafORIZADAS	Interseções não SemafORIZADAS	Extensão Vias Expressas	Extensão Vias Arteriais	Extensão Vias Coletoras	Extensão Vias Locais	Equipamentos Fiscalização
Total de Acidentes	1,00								
Total de Acidentes com Vítimas	0,90	1,00							
Interseções SemafORIZADAS	0,87	0,74	1,00						
Interseções não SemafORIZADAS	0,19	0,42	0,03	1,00					
Extensão Vias Expressas	0,16	0,30	0,02	0,48	1,00				
Extensão Vias Arteriais	0,76	0,59	0,65	0,13	-0,03	1,00			
Extensão Vias Coletoras	0,29	0,29	0,25	0,13	0,21	0,16	1,00		
Extensão Vias Locais	0,18	0,39	0,03	0,94	0,54	0,09	0,13	1,00	
Equipamentos Fiscalização	0,63	0,61	0,47	0,44	0,30	0,52	0,20	0,39	1,00

Em que  $E(y|x_1, x_2, \dots, x_k)$  é o valor esperado da variável resposta,  $y$ , dado que um conjunto de  $k$  valores das variáveis explicativas é observado. Os  $\beta$ 's são os parâmetros que se deseja estimar. Estes podem ser interpretados aproximadamente como a taxa de variação percentual média em  $y$  para uma dada variação unitária em  $x$ .

Com base nos objetivos da modelagem e nas variáveis coletadas foram investigadas duas categorias de modelos: a primeira utilizando como variável resposta o número total de acidentes de trânsito, e a segunda o número total de acidentes de trânsito com vítimas fatais e feridas como variável dependente, para ambos os níveis de planejamento (estratégico e tático). Ressalta-se que neste trabalho a modelagem é usada como ferramenta de análise exploratória, sem a intenção de estabelecer relação de causa e efeito, pois o nível de agregação das variáveis dificulta tal tipo de inferência.

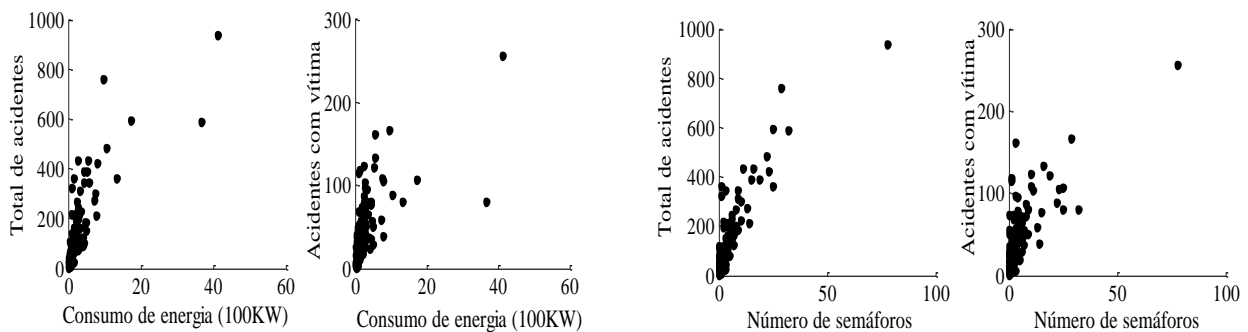
Um teste inicial para avaliar a premissa de superdispersão mais adequada para os dados foi realizado com a estatística amostral  $\sigma^2 = (n - k - 1)^{-1} \sum_i r_i^2 / \hat{y}_i$ . Onde  $r_i$  são os resíduos do modelo calibrado para cada observação  $i$  da amostra de tamanho  $n$ , e  $\hat{y}_i$  são os valores previstos pelo modelo. Valores superiores a 1,00 para este parâmetro indicam dispersão nos dados superior à assumida na distribuição (Poisson ou binomial negativa). Os resultados indicaram valores de  $\sigma^2$  entre 12 e 60 para os modelos assumindo distribuição de Poisson e de 1,11 e 1,14 para os modelos

com a distribuição binomial negativa, sendo portanto, a última premissa mais adequada para os dados. Os parâmetros dos modelos foram estimados pelo método da máxima verossimilhança (Hauer, 2002; Lord, 2006). As Tabelas 6 e 7 apresentam os coeficientes e erros padrões para os modelos com todas as variáveis (a) e com apenas as variáveis significativas (b).

### 3.4.1. Modelos de nível estratégico

Os resultados corroboram a hipótese de que zonas com maior densidade demográfica, população com maior poder aquisitivo e maior intensidade de atividade econômica apresentam maior ocorrência de acidentes de trânsito totais e com vítima.

O efeito da variável “domicílios com renda até 3 salários” é significativo apenas para o modelo de acidentes com vítimas. Uma possível explicação pode estar relacionada ao menor número de viagens motorizadas individuais pelo modo automóvel, nas zonas de menor poder aquisitivo. Desta forma, ao realizar grande parte dos deslocamentos diários por meios não motorizados espera-se uma maior severidade dos acidentes ocorridos nessas zonas. Outra linha de argumentação pode ser levantada em função de diferenças no estado de conservação do pavimento, sinalização horizontal e vertical, presença de fiscalização, iluminação e outros fatores entre as zonas de menor e maior poder aquisitivo.



**Figura 1.** Gráficos de dispersão

**Tabela 6.** Modelos de Previsão de Acidente - nível estratégico

Parâmetro	Acidentes totais				Acidentes com vítimas			
	1(a)		2(b)		1(a)		2(b)	
	$\beta$	EP	$\beta$	EP	$\beta$	EP	$\beta$	EP
Constante	3,980	0,106	4,077	0,088	3,000	0,097	3,307	0,087
Consumo de Energia no comércio, serviço e indústria ( $10^5Wh$ )	0,061	0,014	0,061	0,014	0,049	0,012	0,045	0,013
Domicílios com renda até 3 Salários (1000 dom.)	0,027*	0,016			0,078	0,015		
Domicílios com renda acima de 3 Salários (1000 dom.)	0,417	0,067	0,427	0,067	0,251	0,061	0,277	0,066

\*parâmetro não significativo ( $\alpha=0,05$ )**Tabela 7.** Modelos de Previsão de Acidente - nível tático

Parâmetro	Acidentes Totais				Acidentes com vítimas			
	1(a)		2(b)		1(a)		2(b)	
	$\beta$	EP	$\beta$	EP	$\beta$	EP	$\beta$	EP
Constante	3,800	0,101	3,935	0,087	2,895	0,095	2,928	0,095
Interseções Semaforizadas	0,032	0,009	0,038	0,008	0,022	0,007	0,027	0,008
Interseções não Semaforizadas	0,001*	0,001	-	-	0,002	0,001	0,002	0,001
Extensão Vias Expressas	0,048	0,047	0,099	0,044	0,056*	0,044	-	-
Extensão Vias Arteriais	0,188	0,041	0,179	0,042	0,114	0,038	0,096	0,038
Extensão Vias Coletoras	0,145*	0,075	-	-	0,108*	0,069	-	-
Equipamentos Fiscalização	0,056	0,031	0,071	0,030	0,050	0,029	0,060	0,029

\*parâmetro não significativo ( $\alpha=0,05$ )

Pelos coeficientes da Tabela 6 pode-se afirmar que uma variação de 1000 domicílios com renda acima de 3 salários representa uma variação média de 42% nos acidentes totais e 25% em acidentes com vítimas. Estes valores altos podem ser o resultado de um viés positivo causado pela omissão de outras variáveis correlacionadas positivamente com as variáveis consideradas nos modelos. Tais variáveis omitidas podem estar relacionadas a fatores de geração das viagens nas zonas, como por exemplo, a acessibilidade do sistema de transportes.

### 3.4.2. Modelos de nível tático

Os resultados para o nível tático corroboram a hipótese de que os fatores relacionados a geometria e operação da rede de tráfego, e comportamento dos condutores afetam a ocorrência e severidade dos acidentes nas zonas. Vale ressaltar que o volume por veículo- quilômetro, que é o principal variável de exposição neste nível de análise não foi considerado. Todas as variáveis consideradas na análise devem ter correlação com esta variável já que qualquer intervenção na rede é relacionada com a intensidade de tráfego. Assim, é possível que os coeficientes apresentados estejam superestimados ou não seriam significativos caso fatores de exposição fossem considerados. Portanto, os coeficientes das variáveis na Tabela 7 devem ser interpretados com cautela.

A variável “número de interseções não-semaforizadas” foi significativa apenas para os modelos de acidentes com vítimas, corroborando os estudos que indicam que o perfil mais usual dos acidentes nestes locais (colisões transversais) tende a gerar ocorrências de maior severidade nesse tipo de interseção. O número de semáforos é considerado uma variável *proxy* para o fluxo e nesse sentido o seu coeficiente indica, como esperado, um crescimento dos acidentes (totais ou com vítimas) em virtude do crescimento na exposição a conflitos de colisões traseira e transversais à medida que o fluxo cresce (levando a instalação de semáforos).

O efeito da extensão das vias mostra que aquelas classificadas como arteriais são as que mais contribuem para ocorrência de acidentes em geral. É interessante investigar quais são os critérios adotados para classificar estas vias. Em geral estas vias são corredores importantes de transporte público e do tráfego geral, apresentam considerável movimento de estacionamento devido ao uso comercial do solo de entorno e são controladas por semáforos. Portanto, são vias que apresentam consideráveis conflitos veiculares e de pedestres.

Os equipamentos de fiscalização apresentaram efeitos significativos, ao nível de 5%, em especial para os modelos de acidentes com vítimas. É importante investigar os critérios para instalação destes dispositivos. Em geral eles são localizados em semáforos para controle de velocidade e avanço no vermelho. Portanto, levanta-se a hipótese de que eles são localizados, na maioria das vezes, em locais com uma maior exposição e conseqüentemente risco de acidentes.

### 3.5. Análise da dependência espacial

Os modelos calibrados na seção anterior foram para toda a cidade de Fortaleza, por isso são denominados de modelos de regressão global. Segundo Hadayeghi et al. (2003), esses modelos assumem que o processo a ser examinado é constante em toda a área de estudo e a relação entre a variável dependente e as variáveis independentes não mudam ao longo do espaço. A premissa desse tipo de modelo é que as observações individuais (neste caso, dados das zonas) devem ser independentes espacialmente uma das outras.

Já que nesse estudo utiliza-se dados eminentemente espaciais, a suposição de independência espacial pode não estar correta. Sendo assim, não só as variáveis do modelo global podem apresentar dependência espacial (ou seja, locais próximos terão valores similares), mas também os resíduos, e estas características dos dados espaciais têm implicações na estimativa dos parâmetros, uma vez que altera o poder explicativo do modelo.

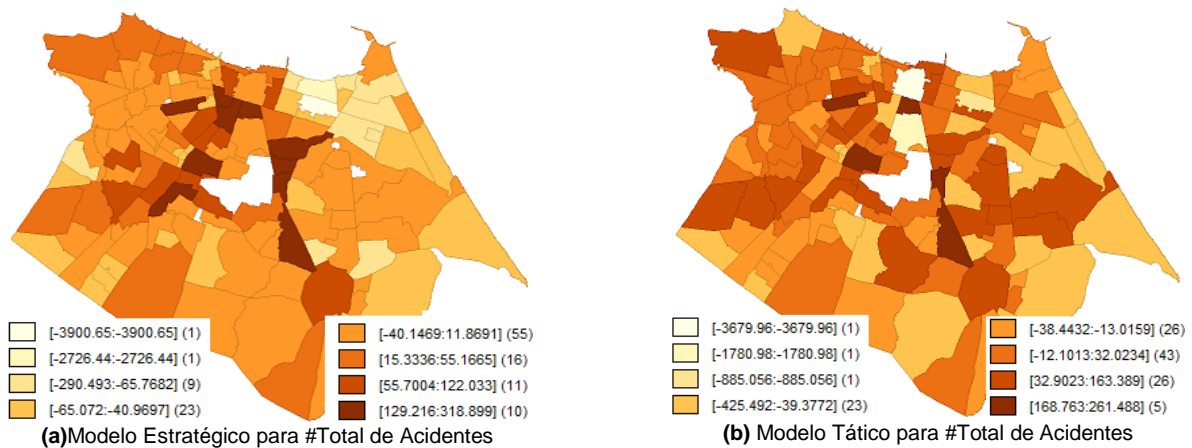


Figura 2. Distribuição espacial dos resíduos

A verificação da dependência espacial dos resíduos foi realizada utilizando o Índice Global de Moran ( $I$ ). Esse índice estima a autocorrelação de uma variável, porém medido em áreas vizinhas e terá um valor varia entre -1 e 1. Assim, quanto mais próximo de 1 (um), maior a semelhança entre vizinhos. O valor 0 (zero) indica inexistência de correlação, e valores negativos indicam dessemelhança. Para determinar o Índice de Moran inicialmente foi definida uma matriz de proximidade espacial, considerando zonas vizinhas que compartilham fronteiras comuns. Para esses cálculos foi utilizado o software GeoDA (SAL, 2005).

Os modelos de nível estratégico apresentaram valor do Índice de Moran de 0,22 e 0,26 confirmando a hipótese de que os resíduos desses modelos possuem uma dependência espacial. Já os modelos no nível tático não apresentaram evidências de autocorrelação espacial (índice de Moran = 0).

Observando o mapa temático na Figura 2(a) percebem-se altos valores negativos ou positivos dos resíduos em zonas próximas, de modo que caso esse modelo global sejam utilizado, estas zonas agrupadas terão valores superestimados ou subestimados para a variável dependente. Já a Figura 2(b) ilustra o modelo cuja variável dependente é o número total de acidentes para o nível tático, e pode-se verificar um padrão considerado aleatório, sem a presença da dependência espacial, o que torna o uso desse modelo global consistente.

A existência da dependência espacial para o nível estratégico sugere a inclusão de uma variável explicativa que não foi considerada nesses modelos e que poderia eliminar o padrão de correlação espacial observado entre os resíduos. As variáveis dos modelos estratégico (consumo de energia e nº de domicílio com renda maior que 3 SM) estão diretamente relacionadas a atração/produção de viagens entre zonas, entretanto estas variáveis não permitem captar o fluxo de tráfego de passagem entre as origens/destinos, o que pode gerar a autocorrelação espacial.

As variáveis existentes nos modelos do nível tático (semáforos, extensão de vias, fiscalização eletrônica, etc) por outro lado, permitiram (mesmo que de forma indireta) considerar medidas de exposição relativas ao volume de tráfego conseguindo uma representação adequada para a estru-

tura espacial dos dados observados. Uma das recomendações para reduzir o efeito espacial nos modelos estratégicos seria a utilização de modelos de Regressão Geograficamente Ponderada (RGP).

#### 4. CONSIDERAÇÕES FINAIS

Esse artigo apresentou uma análise exploratória de variáveis agregadas, ao nível de zona de tráfego, que podem estar relacionadas à ocorrência e severidade de acidentes de trânsito no município de Fortaleza. A análise exploratória é dividida nos níveis estratégico e tático do processo de planejamento. Portanto, no nível estratégico variáveis relacionadas a geração de viagens e interação espacial entre as zonas são selecionadas, enquanto no nível tático, trabalha-se com variáveis relacionadas a fatores de exposição e risco. Devido a limitação das bases de dados somente parte das variáveis estão disponíveis.

Os resultados indicam que as variáveis “consumo de energia” e “domicílios por nível de renda”, selecionadas para o nível estratégico, apresentaram efeitos significativos e positivos, ao nível de significância de 1%. Em especial, a variável número de domicílios com baixa renda apresentou efeito significativo apenas nos modelos de acidentes com vítimas, indicando que nas zonas de menor poder aquisitivo uma eventual maior severidade dos acidentes pode estar associada a maior intensidade de descolamentos não motorizados ou estrutura viária precária. As seguintes variáveis, selecionadas para o nível tático, apresentaram efeitos significativos e positivos: “extensão de vias arteriais e expressas”, “número de interseções semaforizadas e não semaforizadas”, e “número de equipamentos de fiscalização”. Argumenta-se que estas variáveis estão relacionadas a fatores que afetam diretamente os acidentes, como operação e monitoramento do tráfego, comportamento dos condutores, fluxo veicular e de pedestres, e atributos da rede viária. Em particular, a variável “número de interseções não-semaforizadas” foi significativa apenas para os modelos de acidentes com vítimas, corroborando a hipótese de maior severidade dos acidentes em interseções não controladas.

A análise da independência espacial dos modelos mostrou indícios de violação desta premissa ao nível estratégico, indicando que as variáveis utilizadas nos modelos

estratégicos não conseguiram capturar o padrão espacial dos acidentes de trânsito observado, o qual pode estar relacionado ao fluxo de tráfego entre zonas.

Uma limitação deste trabalho é que considerável parcela da população de acidentes não é georreferenciada e, portanto, não considerada na análise. Outra limitação é a ausência de variáveis importantes nos bancos de dados existentes, como a matriz origem-destino e o fluxo veicular.

Em trabalhos futuros, recomenda-se a investigação do caráter espacial nos modelos de previsão, visando uma melhor estimativa dos efeitos. Neste sentido, quaisquer dos dois caminhos a seguir podem ser adotados. A estimação de modelos globais considerando atributos espaciais, como por exemplo, a proximidade entre as zonas. Ou a estimação de modelos locais, conforme sugerido na literatura, por meio da regressão geograficamente ponderada, com a premissa de que os parâmetros dos modelos variam espacialmente. O primeiro caminho está apoiado na possibilidade de interação espacial entre zonas próximas ou distantes. Já o segundo está apoiado na observação empírica de que entidades espaciais próximas tendem a apresentar características similares.

## AGRADECIMENTOS

Os autores agradecem o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) para o desenvolvimento dessa pesquisa, além das Bolsas de Iniciação de Pesquisa e Produtividade em Pesquisa concedidas. Os autores agradecem ainda a Autarquia Municipal de Trânsito, Serviços Públicos e Cidadania de Fortaleza - AMC pelas informações disponibilizadas sobre os acidentes de trânsito neste trabalho.

## REFERÊNCIAS

- Aguero-Valverde, J. e Jovanis, P.P. (2006). Spatial analysis of fatal and injury crashes in Pennsylvania. *Accident Analysis and Prevention*, v. 38, n. 3, p. 618-625, 2006. DOI:10.1016/j.aap.2005.12.006.
- Caliper (2008) *TransCAD - Transportation Workstation Software, User's Guide, Version 5.0 for Windows*. Caliper Corporation, Newton, USA.
- Claude, G. F. de M. (2012) *Previsão da ocorrência de acidentes de trânsito em interseções de vias arteriais urbanas – O caso de Taguatinga/DF*. Dissertação Mestrado. Programa de Pós-Graduação em Transportes, Universidade de Brasília.
- Cunto, F. J. C.; Neto, M. M .C.; Barreira, D.S. (2012) Modelos de previsão de acidentes de trânsito em interseções semaforizadas de Fortaleza. *Transportes*, v. 20, n. 2, p. 55-64. DOI: 10.4237/transportes.v20i2.558.
- Ferreira, S. (2010). *A Segurança Rodoviária no processo de planejamento de redes de transportes*. Tese de Doutorado. Universidade do Porto.
- Hadayeghi, A., Shalaby, A.S. e Persaud, B. (2003). Macro-level accident prediction models for evaluating safety of urban transportation systems. *Transportation Research Record*, v. 1840, n. 1, p. 87-95. DOI: 10.3141/1840-10.
- Hadayeghi, A., Shalaby, A. e Persaud, B. (2010). Development of planning level transportation safety tools using Geographically Weighted Poisson Regression. *Accident Analysis and Prevention*, v. 42, p. 676-688. DOI:10.1016/j.aap.2009.10.016.
- Hauer, E (2002). *Observational Before-after Studies in Road Safety*. (1ª ed.). Pergamon. Toronto, Canadá.
- Henrique, C. S.(2004). *Diagnóstico Espacial da Mobilidade e da Acessibilidade dos Usuários do Sistema Integrado de Transportes de Fortaleza*. Dissertação de Mestrado, Programa de Mestrado em Engenharia de Transportes, Universidade Federal do Ceará, Fortaleza, CE.
- ITE (2009). *Transportation Planning Handbook*, 3ª Edição. Institute of Transportation Engineers.
- Karlaftis, M. G. e Tarko, A. P. (1997). Heterogeneity considerations in accident modeling. *Accident Analysis and Prevention*, v. 30, p. 425-433. DOI:10.1016/S0001-4575(97)00122-X.
- Levine, N., Kim, K.E. e Nitz, L.H. (1995). Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators. *Accident Analysis and Prevention* v. 27, n. 5, p. 675-685. DOI: 10.1016/0001-4575(95)00018-U.
- Lord, D. (2000). *The prediction of accidents on digital networks: characteristics and issues related to the application of accident prediction models*. Tese de Doutorado, Universidade de Toronto, Canadá.
- Lord, D. e Persaud, B. N. (2004). Estimating the safety performance of urban road transportation networks. *Accident Analysis and Prevention*, v. 36, n. 4, p. 609-620. DOI: 10.1016/S0001-4575(03)00069-1.
- Lord, D. (2006) Modeling motor vehicle crashes using poisson-gamma models: Examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. *Accident Analysis and Prevention*, v. 38, n. 4, p. 751-766. DOI:10.1016/j.aap.2006.02.001.
- SAL (2005) *GeoDA – Geodata Analysis Software*. Version 1.4.6 for Windows. Spatial Analysis Laboratory, Illinois, USA.
- Tarko, A. P. (2006). Calibration of safety prediction models for planning transportation networks. *Transportation Research Record: Journal of the Transportation Research Board*, v. 1950, n. 1, p. 83-91. DOI: 10.3141/1950-10
- Tarko, A. P., Inerowicz, M., Ramos, J. e LI, W. (2008). Tool with road-level crash prediction for transportation safety planning. *Transportation Research Record: Journal of Transportation Research Board*, v. 2083, n. 1, p. 16-25. DOI: 10.3141/2083-03



Washington, S., Schalkwyk, I.V., Meyer, M., Dumbaugh, E. e Zoll, M., (2006). Incorporating safety into long-range transportation planning. *NCHRP Report n. 546*, TRB. National Cooperative Highway Research Program, Washington D.C.