

Análise de variações temporais na estimação de acidentes em segmentos arteriais urbanos

Lucas Tito Pereira Sobreira¹, Flávio José Craveiro Cunto²

Resumo: Os modelos de previsão de acidentes (MPA) tem se destacado como ferramenta estatística com potencial para aperfeiçoar os estudos de segurança viária. A utilização de séries históricas de acidentes de trânsito na calibração de MPA introduz uma correlação temporal entre as observações de uma mesma entidade representando uma fonte adicional de viés ao processo de modelagem. Este trabalho objetiva avaliar o efeito de tendências temporais no desenvolvimento de MPA para segmentos arteriais urbanos em Fortaleza. Foram desenvolvidos MPA a partir dos modelos lineares generalizados (MLG) e das equações de estimação generalizada (EEG) para uma amostra de 283 segmentos com o total de acidentes para os anos de 2007 a 2011 assumindo diversas estruturas de correlação temporal. Os resultados não corroboram a hipótese da existência de tendência temporal na amostra, entretanto, os MLG calibrados com coeficientes individuais para cada ano tiveram desempenho superior aos modelos temporalmente agregados.

Palavras-chave: Modelos de previsão de acidentes. Equações de estimação generalizada. Segurança viária. Acidentes de trânsito.

Abstract: Accident Prediction Models (APM) have emerged as a statistical tool with the potential to improve road safety studies. The use of time series of traffic accidents in the calibration of APM introduces a temporal correlation between observations of the same entity representing an additional source of bias to the modeling process. This study evaluates the effect of temporal trends in the development of APM for urban arterial segments in Fortaleza. APM were developed from generalized linear models (GLM) and generalized estimating equations (GEE) for a sample of 283 segments with the total number of accidents for the years 2007 to 2011 assuming various structures of temporal correlation. The results do not support the hypothesis of temporal trends in the sample, however, the GLM developed with individual coefficients for each year outperformed the temporally aggregated models.

Keywords: Safety performance functions. Accident predicting models. Generalized estimating equations. Road safety.

1 INTRODUÇÃO

Profissionais dedicados à pesquisa da segurança viária tem procurado voltar seus esforços para a identificação dos principais fatores que contribuem para a ocorrência dos acidentes de trânsito, bem como, na compreensão objetiva do impacto da variação desses fatores no desempenho da segurança viária (Hauer, 2002; Davis, 2004). Nesse sentido, os modelos de previsão de acidentes (MPA) tem se destacado como ferramenta estatística com potencial para aperfeiçoar a seleção de locais críticos de acidentes de trânsito, além de permitir análises do impacto de estratégias para o aumento da segurança viária de forma mais coerente com a natureza aleatória e rara dos acidentes de trânsito (Persaud e Lyon, 2007 e Geedipally e Lord, 2010).

Nos esforços de calibração dos MPA são frequentemente encontrados estudos que utilizam séries históricas de três a seis anos de aci-

dentos ocorridos nas entidades (segmentos, interseções, etc) que compõem a amostra disponível (Bonneson e McCoy, 1993; Greibe, 2003; Persaud *et al*, 2002;). Se, por um lado a utilização de séries históricas mais longas pode ajudar a aumentar o tamanho da amostra disponível para a calibração ou elevar a média de acidentes observados (agregando-se os acidentes por um período maior de tempo), reduzindo o excesso de zeros, por outro lado, acrescenta-se um caráter longitudinal ao experimento que não pode ser negligenciado.

Alguns aspectos que influenciam a tendência temporal na ocorrência de acidentes como a evolução tecnológica dos veículos, implantação de políticas públicas de longo prazo, alterações na estrutura operacional das agências responsáveis pela coleta dos dados, dentre outros, não são incorporados de forma direta nos MPA. Esses fatores contribuem para a existência de uma correlação temporal entre os acidentes de trânsito e têm motivado o desenvolvimento de pesquisas voltadas para responder questões do tipo: 1) Qual a validade temporal dos MPA?; 2) Até que ponto pode-se agregar séries históricas de acidentes no desenvolvimento de MPA?; 3) Como incorporar a variação/correlação temporal no desenvolvimento

¹ Departamento de Engenharia de Transportes, Universidade Federal do Ceará. (lucastito000@gmail.com).

² Departamento de Engenharia de Transportes, Universidade Federal do Ceará (flaviocunto@det.ufc.br).

Manuscrito recebido em 09/07/2014 e aprovado para publicação em 22/09/2014. Este artigo é parte de TRANSPORTES v. 22, n. 3, 2014. ISSN: 2237-1346 (online).

DOI: <http://dx.doi.org/10.14295/transportes.v22i3.818>.

dos MPA?; (Montain et al, 1998; Lord e Persaud, 2000, Claude, 2012).

A incorporação de medidas de desempenho de segurança viária no processo de planejamento em nível estratégico vem sendo sistematicamente explorada com o emprego de MPA desenvolvidos separadamente para segmentos e interseções e plataformas de modelagem da demanda em nível macroscópico como o Transcad e o EMME2 (Lord e Persaud, 2004; Tarko, 2006). No cenário brasileiro, o desenvolvimento de MPA para segmentos viários urbanos pode permitir a complementação da maioria dos estudos recentes que tiveram foco na calibração de MPA para interseções urbanas (Cunto et al, 2011; Claude, 2012).

O presente artigo tem como objetivo desenvolver MPA para segmentos arteriais urbanos em Fortaleza que incorporem o efeito de tendências temporais na série histórica dos acidentes de trânsito.

2 MPA PARA SEGMENTOS VIÁRIOS E ASPECTOS TEMPORAIS NA CALIBRAÇÃO DE MODELOS

Os MPA são expressões obtidas com a aplicação de técnicas de modelagem linear generalizada assumindo normalmente a estrutura do erro seguindo a distribuição binomial negativa. Esses modelos estimam o número de acidentes de trânsito em função de variáveis que representam atributos geométricos e operacionais de um determinado segmento (ou interseção) e não devem ser vistos como expressões das relações de causa/efeito entre as variáveis envolvidas e a ocorrência de acidentes de trânsito (Hauer, 2002).

A grande maioria dos esforços de calibração de MPA para segmentos viários concentra-se em modelos para segmentos homogêneos em rodovias rurais e em vias urbanas (Persaud e Mucsi, 1995; Sawalha e Sayed, 2001; Greibe, 2003; Cardoso e Goldner, 2007). O ambiente viário urbano acrescenta um nível de dificuldade ao processo de modelagem em virtude da maior quantidade de variáveis que podem influenciar nos acidentes, como o uso do solo, tipo de estacionamento, presença de canteiro central, influência de usuários vulneráveis, paradas de ônibus, acessos aos lotes, etc, além dos atributos normalmente presentes nos trechos rurais, como o número e largura de faixas

e tipo de acostamento. A Tabela 1 apresenta um resumo das variáveis bem como da série histórica de acidentes utilizada em alguns estudos com foco na calibração de MPA para segmentos viários.

Os modelos desenvolvidos para segmentos viários podem assumir diversas estruturas com componentes aditivos, multiplicativos e exponenciais em função das variáveis consideradas, entretanto, a estrutura geral mais comumente encontrada pode ser expressa por:

$$\#Ac = \alpha \cdot VDMA^{\beta_0} \cdot L^{\beta_1} \cdot e^{\sum_j (\gamma_j A_j)} + (\sum_k \delta_k B_k) \quad (1)$$

em que:

$\#Ac$ = número esperado de acidentes em um intervalo de tempo;

$VDMA$ = volume médio diário anual (exposição);

L = comprimento do trecho j ;

A = vetor de variáveis preditoras presentes ao longo do trecho (número de faixas, largura de acostamentos, etc);

B = vetor de variáveis preditoras presentes em pontos específicos do trecho (entradas/saídas de lotes, ponto de ônibus, pontes, etc);

$\alpha, \beta, \gamma, \delta$ = parâmetros do modelo.

Na grande maioria dos MPA a variável dependente é na realidade uma taxa anual de acidentes estimada a partir de observações feitas ao longo de alguns anos. Desta forma, tendências de crescimento (ou redução) dos acidentes presentes nas observações de uma mesma entidade em anos diferentes acrescentam um caráter longitudinal ao estudo que não está sendo considerado e pode reduzir a qualidade geral desses modelos.

Mountain et al. (1998) investigaram a tendência temporal dos acidentes em uma amostra de 15 anos (1980-1994) em interseções do Reino Unido. O caráter temporal foi incorporado ao MPA com a introdução de um parâmetro multiplicativo γ que representava a variação anual dos acidentes em relação a um ano base para a análise. Os resultados indicaram uma tendência de redução anual de 6% ($\gamma=0,94$) nos acidentes de trânsito que, de acordo com os autores deve ser incorporada aos MPA de modo a não torná-los obsoletos de forma precoce.

Tabela 1 - Modelos de previsão de acidentes de trânsito para segmentos

Autores	#Anos	# Trechos	Entidade	Variáveis exploradas
Persaud e Mucsi, 1995	1998-1999	2.014	Rodovias de duas faixas de tráfego	Volume médio horário, extensão, largura de faixa de tráfego e acostamento.
Bonneson e McCoy (1997)	1991-1993	189	Vias urbanas	VDMA, extensão, uso do solo, tipo de estacionamento, densidade de entradas de garagem, densidade de interseções não semaforizadas, percentual de acidentes sem feridos, presença de faixa de conversão à esquerda, presença de separador central.
Sawalha e Sayed (2001)	1994-1996	392	Rodovias arteriais urbanas semaforizada	VDMA, extensão, número de faixas, densidade de interseções não semaforizadas, densidade de calçadas, densidade de acessos, presença de separador central e uso do solo (comércio).
Greibe (2003)	1990-1994	314	Vias urbanas	VDMA, extensão, número de faixas, densidade de interseções não semaforizadas, tipo de estacionamento e uso do solo.
Cardoso e Golder (2007)	1998-2004	73	Vias urbanas	VDMA, extensão e largura do segmento, percentual de motocicletas e veículos pesados, percentual de pedestres atravessando sobre a faixa, fluxo de pedestres, velocidade no 85%, condição do pavimento, densidade de vegetação, fluxo veicular por aproximação.

Lord e Persaud (2000) utilizaram uma extensão dos modelos lineares generalizados (MLG) denominada Equações de Estimação Generalizada (EEG) para avaliar o impacto da tendência temporal em uma amostra de 868 interseções semaforizadas monitoradas ao longo de 6 anos (1990-1995). As EEG permitem a estimação dos parâmetros do modelo e sua variância a partir de suposições sobre a dependência temporal das observações (Zeger e Liang, 1986). Os autores concluíram que a aplicação das EEG não alterou os coeficientes dos modelos calibrados com a técnica tradicional dos MLG, entretanto, os modelos calibrados com as EEG resultam em um erro padrão maior dos coeficientes em virtude da variação entre os grupos (observações da mesma interseção) ter sido considerada no método de calibração.

Uma aplicação das EEG para a calibração de MPA no cenário brasileiro foi apresentada no trabalho de Claude (2012). Esse trabalho empregou as EEG como ferramenta para permitir o aumento de sua amostra inicial de 32 interseções de Brasília, incluindo observações de vários anos (2005 a 2010) dos mesmos locais sem desprezar o aspecto longitudinal do experimento.

Claude (2012) assumiu as seguintes hipóteses em relação à estrutura de correlação temporal entre as observações: i) Independente, ii) Autoregressiva em que a correlação diminui com o afastamento temporal entre duas observações e iii) Intercambiável, na qual a correlação entre duas observações de uma mesma entidade é assumida constante. Os resultados in-

dicaram que os modelos calibrados utilizando a hipótese de correlação temporal intercambiável tiveram melhor desempenho.

Os estudos sintetizados deixam aparente que, na maioria das vezes, a correlação temporal entre os acidentes de trânsito não deve ser desprezada, e, além disso, sua incorporação aos MPA deve considerar hipóteses sobre sua estrutura de correlação temporal.

A seguir será apresentada uma breve descrição do banco de dados de acidentes de trânsito e dos atributos geométricos e operacionais utilizados na calibração de MPA considerando a tendência temporal nos acidentes de trânsito em segmentos de vias urbanas em Fortaleza/CE.

3 BANCO DE DADOS DO EXPERIMENTO

A amostra utilizada para a calibração dos modelos contém 283 segmentos urbanos de vias arteriais e coletoras da região norte da cidade de Fortaleza. Os segmentos foram selecionados em virtude da disponibilidade de informações de fluxo veicular obtidos com o auxílio de laços indutores que compõem o sistema de controle de tráfego em área de Fortaleza (CTAFOR).

O uso do solo predominante na região de estudo é misto, com a presença de estabelecimentos comerciais e pequeno e médio porte, consultórios médicos e outros serviços, além de prédios de apartamentos residenciais. Os segmentos apresentam ainda um padrão de estaci-

onamento que varia entre proibido estacionar e a presença esporádica de estacionamentos recuados à 45 graus. Os segmentos variam pela presença ou não de canteiro central, pelo número de faixas, sentidos único e duplo de tráfego e por sua extensão com as interseções semaforizadas nas extremidades jusantes. As Figuras 1 e 2 apresentam dois segmentos representativos da amostra utilizada nesse trabalho.

As variáveis preditoras utilizadas na modelagem foram o volume médio diário anual

(VDMA), a extensão dos segmentos, o número de faixas e a presença de canteiro central. O VDMA foi estimado utilizando dados de fluxos veiculares presentes no banco de dados do CTAFOR. Foram coletados os VDMM (Volume Médio Diário Mensal) dos anos de 2007 a 2011 para o mês de outubro os quais foram expandidos para VDMA com a aplicação do fator de expansão médio de 1,01 estimado por Oliveira (2004).



Figura 1 - Avenida Antonio Sales, trecho entre Ruas Silva Paulet e Jose Vilar (Fonte: Google Earth, 2013)



Figura 2 - Avenida Des. Moreira, trecho entre Av. Santos Dumont e R. Des. Leite Albuquerque (Fonte: Google Earth, 2013)

Características geométricas dos segmentos, como o número de faixas e a presença de canteiro central, foram observadas com o auxílio de imagens aéreas do aplicativo Google Earth[®]. A extensão dos trechos foi coletada através de dados informatizados georeferenciados no aplicativo TransCAD.

A coleta do número de acidentes dos anos de 2007 a 2011 foi realizada pelo Sistema de Informações de Acidentes e Trânsito de Fortaleza (SIATFOR), um banco de dados, atualizado por 10 diferentes fontes de informação, que registra os acidentes de Fortaleza desde o ano de 2000. Nesse estudo foram considerados todos os acidentes (sem vítimas e com vítimas

feridas/fatais) ocorridos nos segmentos. A tabela 2 apresenta uma análise descritiva dos dados utilizados na modelagem.

4 DESENVOLVIMENTO DOS MODELOS DE PREVISÃO DE ACIDENTES

Como uma etapa preliminar à escolha dos modelos, foram realizados testes utilizando o parâmetro de dispersão σ_d (Bonneson e McCoy, 1993) no intuito de verificar entre Poisson e binomial negativa, qual a distribuição mais adequada para a amostra dos acidentes de trânsito. Modelos básicos foram calibrados com as variáveis VDMA e extensão para o ano de

Tabela 2 - Descrição das variáveis utilizadas na modelagem

Ano	Acidentes			VDMA		Extensão (m)		Canteiro Central (cc) 0=não 1=sim	Número de faixas	
	Min	Max	Total	Min	Max	Min	Max		Min	Max
2007	0	22	477	3.584	52.821	44,1	373,5		2	6
2008	0	33	508	4.005	55.398					
2009	0	32	507	4.228	57.630					
2010	0	29	530	2.106	61.243					
2011	0	31	452	2.158	58.509					

2007 resultando em σ_d estimados de 3,84 para a distribuição de Poisson e 1,42 para a binomial negativa, indicando a maior proximidade ao último tipo de distribuição para os acidentes de trânsito ($\sigma_d \approx 1$). A inclusão posterior das variáveis presença de canteiro central e número de faixas ao modelo inicial resultou em coeficientes não significativos ($\alpha = 0,05$).

A avaliação do aspecto temporal nos MPA foi realizada com a aplicação de seis estratégias de calibração dos modelos com a utilização dos modelos lineares generalizados (MLG) e das equações de estimação generalizadas (EEG). A seguir serão descritas as características básicas de cada estratégia avaliada.

Para os modelos tipo I estimou-se inicialmente a expressão para o ano base 2007 ($t = 0$) e a tendência temporal foi representada na calibração dos anos seguintes mantendo-se constante os parâmetros obtidos para $t = 0$ sendo calibrados os coeficientes α para cada ano isoladamente de acordo com as equações abaixo:

$$C_t = \frac{\sum y_t}{\sum \mu_{07}} \quad (2a)$$

$$\alpha_t = C_t \cdot \alpha_0 \quad (2b)$$

$$\#Ac_t = \alpha_t \cdot VDMA^{\beta_1} \cdot L^{\beta_2} \quad (2c)$$

Em que:

y_t = acidentes observados no ano t ;

μ_{07} = acidentes previstos no ano base (2007);

C_t = razão entre os acidentes em anos diferentes;

α_t = intercepto para o ano t .

Os modelos tipo II foram estimados com a soma dos acidentes e a média do VDMA dos cinco anos. Nesse caso, os acidentes esperados para o período de um ano são iguais (para

um mesmo VDMA e extensão), ou seja, os modelos refletem um efeito médio dos fatores relacionados à ocorrência dos acidentes (inclusive o tempo). Conseqüentemente, a existência de tendências temporais não pode ser verificada e poderá ocasionar modelos pouco eficientes principalmente no início ou final do período de análise.

As estratégias dos modelos tipo III a VI empregam as equações de estimação generalizada (EEG) para desenvolver MPA para considerar de forma explícita a existência de correlação temporal entre as observações de uma mesma entidade. As EEG, apresentadas por Zeger e Liang (1986) conseguem estimar os parâmetros de modelos de regressão linear e suas variâncias sem a especificação inicial da distribuição conjunta da variável. Essas equações são consideradas uma extensão dos MLG uma vez que com o emprego de métodos de quase-verossimilhança é possível a obtenção de um modelo marginal do tipo linear generalizado.

As EEG consideram o caráter longitudinal pela inclusão de uma matriz $R(\lambda)$ (chamada de matriz de trabalho ou *working matrix*) que representa o tipo de correlação entre pares de observações de uma mesma entidade na estimação da matriz de covariância. Desta forma, no processo de estimação do modelo os principais componentes são:

a) Preditor linear e Função de ligação (logarítmica para esse estudo) representada por

$$g(\mu_{it}) = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (3)$$

Em que β são os coeficientes do modelo, x as covariáveis e μ_{it} é o valor esperado dos acidentes de trânsito.

b) Especificação da distribuição da variável dependente (binomial negativa) empregada

na estimação da matriz de covariância expressada por

$$V_i = \phi A_i^{1/2} R_i(\lambda) A_i^{1/2} \quad (4)$$

Em que ϕ é o parâmetro de escala da distribuição escolhida (família exponencial) e A_i é uma matriz ($n_i \times n_i$) com diagonal $[V(u_{i,1}), \dots, V(u_{i,T_i})]$.

c) Especificação da estrutura de correlação da variável dependente (matriz de trabalho). No modelo III a suposição é a de independência temporal entre as observações de uma mesma entidade, ou seja $R(\lambda)$ é a matriz identidade ($I_{5 \times 5}$ nesse estudo). Ressalte-se que o EEG produz os mesmos coeficientes que os MLGs tradicionais nesse tipo de correlação sendo que os erros-padrão são normalmente maiores com o procedimento EEG pela incorporação da variação entre elementos do mesmo grupo (Ballinger, 2004; Claude, 2012).

A estrutura de correlação assumida para o modelo IV foi a intercambiável (*compound symmetry*) na qual as observações de uma mesma entidade (segmento) apresentam correlação constante A estrutura de correlação nesse caso pode ser expressa por (Claude, 2012):

$$\text{Corr}(y_{ij}, y_{ik}) = \begin{cases} \lambda, & j = k \\ 0, & j \neq k \end{cases} \quad (5)$$

e a matriz de trabalho para essa pesquisa definida por:

$$R_{i,j} = \begin{bmatrix} 1 & \lambda & \dots & \lambda \\ \lambda & 1 & \dots & \lambda \\ \vdots & \vdots & \ddots & \vdots \\ \lambda & \lambda & \dots & 1 \end{bmatrix} \quad (6)$$

Para o modelo do Tipo V foi testada a do tipo autoregressiva em que a correlação entre as variáveis decresce com o afastamento temporal entre elas (Claude, 2012), podendo ser expressa por:

$$\text{Corr}(y_{ij}, y_{ij+t}) = \lambda^t, \quad t = 0, 1, 2, \dots, t_i \quad (7)$$

e a matriz de trabalho definida por:

$$R_{i,j} = \begin{bmatrix} 1 & \dots & \lambda^t \\ \vdots & \ddots & \vdots \\ \lambda^t & \dots & 1 \end{bmatrix} \quad (8)$$

Para o modelo tipo VI a estrutura de correlação testada foi a "não estruturada", ou seja, nesse caso, a matriz de trabalho estima as possíveis correlações entre observações da mesma entidade e as inclui no cálculo das variâncias (Fitzmaurice *et al*, 1993). Nesse caso a matriz de trabalho seria definida por:

$$R_{i,j} = \begin{bmatrix} 1 & \lambda_{12} & \dots & \lambda_{1t} \\ \lambda_{21} & 1 & \dots & \lambda_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{t1} & \lambda_{t2} & \dots & 1 \end{bmatrix} \quad (9)$$

Maiores detalhes sobre a estimação dos valores λ podem ser encontrados em Liang e Zeger (1986). Os modelos com a estrutura EEG foram calibrados com a rotina *genmod* para do aplicativo computacional SAS. Os MLG por sua vez foram calibrados com a utilização do aplicativo R função *glm.nb*.

A qualidade global e relativa dos modelos bem como a avaliação da tendência temporal nos dados será realizada com a utilização das métricas AIC (Critério de Informação de Akaike) para os modelos I e II e o QIC (Critério de Informação de Quase-verossimilhança) para os modelos III e VI. Adicionalmente serão analisados os gráficos de resíduos acumulados (*CURE Plots*) para avaliar o desempenho dos modelos ao longo do domínio da variável preditora VDMA.

O Akaike (AIC) expressa a diferença entre a variância e o viés do modelo proposto, desta forma, deve-se então privilegiar aquele modelo que equilibra viés/variância, i.e., menor AIC. Considerando que os modelos testados apresentam amostras de tamanhos diferentes, será utilizada a seguinte expressão (Claude, 2012):

$$AIC = \frac{-2 \ln L(M_k) + 2P}{N} \quad (10)$$

Em que:

$\ln L(M_k)$ = logaritmo da verossimilhança do modelo k ;

P = número de parâmetros utilizados;

N = número de observações.

O critério de informação de Quase-verossimilhança (QIC) é uma modificação do critério AIC desenvolvida para ser empregado em modelos para dados longitudinais (EEG) (Pan, 2001). Assim como o AIC, os modelos mais indicados são aqueles com os menores valores de QIC, entretanto, não é possível comparar os valores de AIC e QIC diretamente, desta forma, o AIC será utilizado para verificar os modelos I e II e o QIC indicará o melhor dos modelos desenvolvidos com o EEG.

A análise entre os modelos desenvolvidos com diferentes técnicas de modelagem será complementada com a utilização dos gráficos de resíduos acumulados (*cure plots*). Os *cure plots* permitem a avaliação visual dos resíduos modelos ao longo de suas variáveis preditoras, identificando possíveis *outliers* e regiões de superestimação e subestimação dos modelos. Maiores detalhes a respeito da confecção e interpretação dos *cure plots* podem ser encontrados em Hauer e Bamfo (1997).

5 ANÁLISE DOS RESULTADOS

A Tabela 3 apresenta um resumo dos coeficientes estimados com seus respectivos indicadores de variação (erros-padrão), parâmetro ϕ da distribuição binomial negativa, coeficiente de dispersão e indicadores AIC/QIC. Ressalte-se que todos os coeficientes estimados para os modelos foram significativos (valores de $p < 0,001$).

Os valores de AIC para os modelos I e II indicam uma superioridade para o modelo I dentre os modelos desenvolvidos com o MLG. Observe-se que o parâmetro de dispersão σ_d do modelo II é o maior dentre os modelos testados, indicando que a agregação temporal acentuou a dispersão entre os dados, tornando a suposição de distribuição binomial negativa menos consistente. No caso específico do modelo I, a Tabela 4 apresenta os coeficientes α para os anos de 2007 a 2011. Os valores desses coeficientes reforçam a hipótese de uma relativa indefinição em termos de tendência temporal nos acidentes de trânsito para amostra analisada.

Em relação aos modelos III a VI desenvolvidos com as EEG, os resultados em termos do indicador QIC indicam que a hipótese de independência temporal (modelo III) resultou em modelos mais adequados e a superioridade significativa do modelo tipo III. Uma análise do comportamento dos resíduos de cada modelo foi realizada com a utilização dos gráficos de resíduos acumulados apresentados na Figura 3.

Em relação aos modelos I e II os gráficos oscilam razoavelmente em torno de zero e estão contidos na envoltória $\pm 2\sigma_d$. Observa-se entretanto uma variação residual relativa bastante acentuada no modelo II, provavelmente em virtude da escala da variável dependente, que nesses modelos representa os acidentes de trânsito ocorridos em cinco anos. Ressalta-se ainda que o *cure plot* para o modelo I utilizou somente os valores para o ano de 2007, entretanto, todos os gráficos para os anos subsequentes apresentaram comportamento similar.

Tabela 3 - Resumo dos parâmetros estimados para os MPA

	Modelos					
	I(*)	II	III	IV	V	VI
$\ln(\alpha)$	-17,60 (1,60)	-14,16 (1,12)	-15,60 (1,56)	-14,27 (1,55)	-14,66 (1,57)	-13,72 (1,65)
β_1	1,35 (0,16)	1,25 (0,12)	1,18 (0,13)	1,01 (0,14)	1,04 (0,14)	0,92 (0,16)
β_2	0,92 (0,20)	0,78 (0,14)	0,88 (0,17)	0,94 (0,18)	0,95 (0,16)	1,01 (0,16)
ϕ	1,38	1,49	1,47	1,47	1,47	1,47
σ_d	1,42	1,55	1,28	1,29	1,32	1,36
AIC	3,22	5,85	-	-	-	-
QIC	-	-	45,21	51,62	47,00	56,30
Modelo	MLG	MLG	EEG	EEG	EEG	EEG

(*) coeficiente α para o modelo do ano 0 (2007).

Valores entre parêntesis = erro padrão

Tabela 4 - Coeficientes α para os modelos Tipo I

Ano	2007	2008	2009	2010	2011
α_t ($\times 10^{-8}$)	2,27	2,36	2,30	2,37	2,20

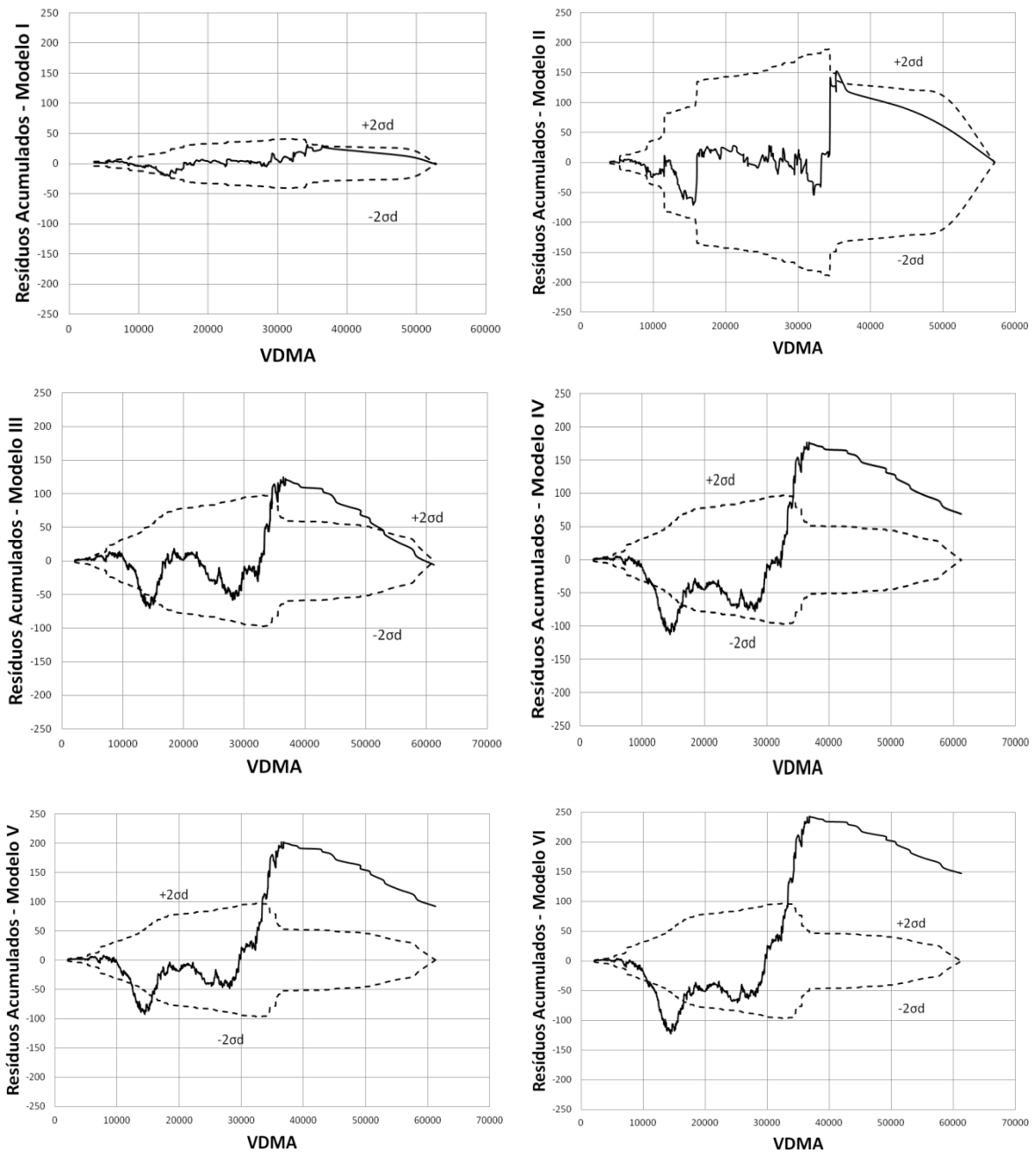


Figura 3 - Gráficos de resíduos acumulados para os modelos I a VI

Dentre os modelos III a VI os gráficos sugerem de forma bastante significativa, o melhor desempenho do modelo III, principalmente, na região de fluxo veicular acima de 33 mil veículos por dia, onde todos os quatro modelos desenvolvidos com o EEG devem ser vistos com cautela por estarem além da envoltória $\pm 2\sigma$. A análise residual sugere finalmente que os modelos desenvolvidos individualmente para cada ano (Modelo I) apresentam desempenho superior quando comparados ao modelo III.

6 CONSIDERAÇÕES FINAIS

A utilização de séries históricas de acidentes de trânsito no desenvolvimento de modelos estatísticos de previsão de acidentes pode introduzir uma correlação temporal entre as observações de uma mesma entidade. Esse aspecto longitudinal da modelagem do desempenho da segurança viária representa uma fonte adicional de viés aos MPA que deve ser considerado e minimizado.

Este artigo investigou a tendência tempo-

ral nos modelos de previsão de acidentes e segmentos viários urbanos de Fortaleza/CE. Foram desenvolvidos MPA para uma amostra de 283 segmentos urbanos utilizando o número total de acidentes para os anos de 2007 a 2011. Os modelos foram calibrados com utilização dos modelos lineares generalizados (MLGs) e das equações de estimação generalizada (EEG). Dentre as estruturas assumidas para a correlação temporal ressaltam-se a independência temporal, a intercambiável, auto-regressiva e a "indefinida" (não estruturada).

Realizou-se uma análise comparativa do desempenho dos modelos propostos com a utilização das métricas AIC/QIC e ainda com o gráfico de resíduos acumulados. Os resultados indicaram a superioridade relativa dos modelos desenvolvidos com os MLGs individuais para cada ano (modelos I). Dentre os modelos desenvolvidos com a técnica dos EEG resalta-se que a suposição da estrutura de correlação independente resultou em modelos satisfatórios, apesar de ligeiramente inferiores aos modelos do tipo I. Desta forma, para a amostra analisada, não foi possível a definição de um padrão temporal de redução ou crescimento da frequência dos acidentes de trânsito. Os resultados indicaram ainda que mesmo sem haver uma tendência temporal definida, a calibração de modelos específicos para anos diferentes melhora as estimativas das frequências dos acidentes de trânsito.

Faz-se necessária a investigação dessa tendência temporal de forma mais desagregada, i.e., em relação à usuários específicos (motociclistas ou pedestres) bem como para outros tipo de entidades como interseções semaforizadas e não semaforizadas.

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 e fluxo veicular utilizadas nesse trabalho.

REFERÊNCIAS

- Ballinger, G. A. (2004) Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods*, v. 7, p. 127-150. DOI: 10.1177/1094428104263672
- Bonneson, J. A. e McCoy, P. T. (1993) Estimation of safety at two-way stop-controlled intersections on rural highways. *Transportation Research Record*, No. 1401, TRB, Washington, D.C., p. 83-99.
- Cardoso, G. e Goldner, L. (2007) Desenvolvimento e aplicação de modelos para previsão de acidentes de trânsito. *Transportes*, v.XV, p.43-51.
- Claude, G. F. M. (2012) Previsão da ocorrência de acidentes de trânsito em interseções de vias arteriais urbanas – O caso de Taguatinga/DF. Dissertação de Mestrado.
- Cunto, F. J. C., Neto, M. M. de C. e Barreira, D. S. (2011) Modelos de previsão de acidentes de trânsito em interseções semaforizadas de Fortaleza. ANPET, 2011.
- Davis, G. A. (2004) Possible Aggregation Biases in Road Safety Research and a Mechanism Approach to Accident Modeling. *Accident Analysis and Prevention*, v. 36, p. 1119-1127. DOI: 10.1016/j.aap.2004.04.002
- Fitzmaurice, G. M., Laird, N. M. e Rotnitzky, A. (1993). Regression models for discrete longitudinal responses. *Statistical Science*, vol. 8, 284-309. DOI: 10.1214/ss/1177010899
- Geedipally, S. R. e Lord, D. (2010) Investigating the effect of modeling single-vehicle and multi-vehicle crashes separately on confidence intervals of Poisson-gama models. *Accidents Analysis and Prevention*, v. 42, p. 1273-1282. DOI: 10.1016/j.aap.2010.02.004
- Greibe, P. (2003) Accident prediction models for urban roads. *Accident Analysis and Prevention*, v.35,p.273-285. DOI: 10.1016/S0001-4575(02)00005-2
- Hauer, E. (2002) *Observational Before-after Studies in Road Safety*. (1a ed.). Pergamon.
- Hauer, E. e Bamfo, F. (1997) Two tools for finding what function links the dependent variable to the explanatory variables. *ICTCT Conference*, Lund, Sweden. Proceedings of ICTCT 97.
- Lord, D., Persaud, B. N. (2004) Estimating the safety performance of urban transportation networks. *Accidents Analysis and Prevention*, v. 36, p. 609-620. DOI: 10.1016/S0001-4575(03)00069-1
- Lord, D. e Persaud, B.N. (2000) Application of the generalized estimating equations procedure. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1717, TRB, National Research Council, Washington, D.C., p. 102-108. DOI: 10.3141/1717-13
- Mountain, L., Maher, M. e Fawaz, B. (1998) The Influence of Trend on Estimates of Accidents at Junctions. *Accident Analysis and Prevention*, v.30,p.641-649, 1998. DOI: 10.1016/S0001-4575(98)00009-8

Oliveira, M. V. (2004) A Natureza dos Padrões de Variação Espaço-temporal do Volume Veicular em Ambiente Urbano: Estudo de Caso em Fortaleza. Dissertação de

Mestrado. Programa de Pós-graduação em Engenharia de Transportes – PETRAN. Universidade Federal do Ceará, Fortaleza.

Pan, W. (2001) Akaike's information criteria in generalized estimating equations. *Biometrics*, v. 57, p. 120-125.

Persaud, B., Lord, D. e Palmisano, J. (2002) Calibration and transferability of accident prediction models for urban intersections. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1784, pp. 57–64, TRB, National Research Council, Washington, D.C., 2002. DOI: 10.3141/1784-08

Persaud, B. e Mucsi, K. (1995) Microscopic accident potential models for two-lane rural roads. *Transportation*

Research Record: Journal of the Transportation Research Board, No. 1485, p. 134-139.

Sawalha, Z. e Sayed, T. (2001) Evaluating safety of urban arterial roadways. *Journal of Transportation Engineering*. v.127, n.2, p. 151-158. DOI: 10.1061/(ASCE)0733-947X(2001)127:2(151)

Tarko, T. (2006) Calibration of Safety Prediction Models for Planning Transportation Networks. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1950, pp 83-91, TRB, National Research Council, Washington, D.C., 2002. DOI: 10.3141/1950-10

Zeger, S. L. e Liang, K. Y. (1986) Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, v. 42, p. 121-130. DOI: 10.2307/2531248