

# FATORES CONTRIBUINTES DE ACIDENTES OCORRIDOS EM INTERSEÇÕES DE TRÊS RAMOS EM PORTUGAL: ELEMENTOS DA CONSISTÊNCIA GEOMÉTRICA

Jocilene Otilia Costa<sup>1</sup>, Elisabete Fraga Freitas<sup>2</sup>, Maria Alice Prudêncio Jacques<sup>3</sup>, Paulo António Alves Pereira<sup>4</sup>

<sup>1</sup>Universidade Federal de Pernambuco – Centro Acadêmico do Agreste, Núcleo de Tecnologia, 55002-970 Nova Caruaru, Caruaru, Brasil email: jocilene.mt@gmail.com

<sup>2</sup>Universidade do Minho, Departamento de Engenharia Civil, 4800-058 Azurém, Guimarães, Portugal

<sup>3</sup>Universidade de Brasília, Departamento de Engenharia Civil, Brasília, Brasil

<sup>4</sup>Universidade do Minho, Departamento de Engenharia Civil, 4800-058 Azurém, Guimarães, Portugal

---

## Sumário

*Este artigo apresenta um estudo sobre os acidentes ocorridos em interseções de três ramos inseridas em estradas nacionais da região norte portuguesa (2008-2010). O estudo é feito através do desenvolvimento de modelos que permitem a avaliação do impacto da consistência geométrica do traçado sobre a ocorrência de acidentes. Foram analisados os parâmetros: frequência de acidentes, tráfego e características geométricas de 177 interseções (531 observações). O processo foi baseado em equações de estimação generalizadas, com estrutura de correlação predominante autorregressiva. O estudo confirma a crescente necessidade de ser incorporado na modelação de acidentes o efeito da variação das condições gerais do traçado das estradas.*

---

**Palavras-chave:** Equações de estimação generalizadas; Modelos de previsão de acidentes; Fatores contribuintes de acidentes; Interseções de três ramos; Consistência do traçado.

## 1 INTRODUÇÃO

A Organização Mundial de Saúde – OMS revelou que mais de 1,24 milhões de pessoas morrem e 50 milhões ficam feridas nas estradas mundiais a cada ano [1]. Em Portugal, a Autoridade Nacional de Segurança Rodoviária – ANSR reportou 518 mortos e 38.872 feridos como os principais dados da sinistralidade do ano de 2013. Isto significa que diariamente ocorrem cerca de 2 mortes e aproximadamente 106 feridos devido a acidentes nas estradas portuguesas. A ANSR reporta ainda que nas estradas nacionais - EN's da rede rodoviária nacional, ocorrem cerca de 35% e 30% das vítimas e dos feridos graves, respetivamente[2].

Devido aos enormes custos dos acidentes de viação para a sociedade, o conhecimento dos fatores que afetam a probabilidade de ocorrência dum acidente tem sido uma área de investigação desde há muitas décadas. A maioria dos investigadores que abordou este problema ([3], [4], [5], [6], [7], [8], entre outros) centrou-se na compreensão dos fatores que afetam a frequência de acidentes ocorridos num determinado elemento viário (segmento ou interseção) durante um determinado período de tempo (semana, mês, ano ou grupo de anos).

O desenvolvimento dos modelos de previsão de acidentes tem diferentes propósitos, tais como: investigação das relações entre os acidentes e as principais características do ambiente rodoviário [9]; e previsão quantitativa dos acidentes de viação. Este desenvolvimento também requer cuidados especiais, para que os resultados e as interpretações que lhes estão associadas sejam os mais corretos [10].

Os dados sobre acidentes a modelar serão sempre inteiros não negativos, para os quais são adequados os modelos de regressão de dados de contagens, ou outros que possam lidar com a natureza deste tipo de informação. Lord e Mannering apresentam uma análise sobre diferentes alternativas metodológicas que podem ser empregadas para o desenvolvimento desses modelos, indicando o potencial e as limitações de cada abordagem, tendo em conta as características da base de dados disponível [10]. Uma análise similar, porém focada no estudo de interseções foi realizada por [11]. Entre as abordagens utilizadas, [12], por exemplo,

desenvolveram modelos baseados na técnica de modelação dos modelos lineares generalizados – MLG e com distribuição binomial negativa, para interseções inseridas em estradas rurais, enquanto que [4] utilizaram a técnica de modelação das equações de estimação generalizadas – EEG, também com distribuição binomial negativa para o erro, na modelação da frequência de colisões traseiras em interseções semaforizadas localizadas em estradas interurbanas, levando em conta a correlação temporal e espacial entre os dados. Nos últimos anos outras técnicas de modelação têm sido aplicadas, como por exemplo: [13] utilizaram - o modelo de parâmetros aleatórios tobit para estudar a sinistralidade em segmentos rodoviários.

Na maioria dos estudos revistos foram utilizadas variáveis de dois grupos de fatores: (i) fatores de exposição ao risco de acidentes (usualmente: volume de tráfego e distância percorrida); e (ii) fatores associados à via e ao ambiente viário, que se referem diretamente ao elemento viário em questão (geometria, tipo de controle, etc.).

No que diz respeito à inclusão de variáveis que permitam avaliar o impacto da consistência geométrica do traçado na ocorrência de acidentes (que fazem parte do grupo de fatores ligados à via), são poucos os trabalhos presentes na literatura. Os trabalhos de [7] e [14], por exemplo, que são dedicados a segmentos rodoviários, incorporam a questão da consistência geométrica por meio da estimativa da velocidade operacional dos veículos (V85) em segmentos viários consecutivos. Na revisão da literatura efetuada para a realização desta pesquisa não foram encontrados trabalhos que abordem a questão da consistência geométrica na modelação da frequência de acidentes em interseções. Destaca-se que a consistência geométrica de um traçado refere-se às condições de geometria do traçado, especialmente do seu alinhamento horizontal e vertical, que permitem ao condutor dirigir o seu veículo sem ser surpreendido por uma modificação brusca da geometria que requeira uma mudança substancial da velocidade do veículo. Segundo [15], as similaridades das características de projeto contribuem para a segurança viária. As interseções de nível são locais operacionalmente complexos dentro do sistema de transporte no estudo da sinistralidade devido ao número de conflitos, como realça [16]. No contexto da consistência geométrica, a sua presença já representa em si uma alteração severa do traçado. No entanto, lembrando que a área de influência de interseção compreende segmentos viários com determinadas características geométricas, é importante investigar se a consistência geométrica desses segmentos com relação aos segmentos limítrofes é um fator agravante para a ocorrência de acidentes.

Os modelos obtidos nos estudos de sinistralidade disponíveis na literatura podem constituir uma ferramenta de utilização generalizada, mas exigem, no entanto, uma apropriada adaptação ao contexto rodoviário em causa e ao nível de abordagem a que se destinam. Numa ótica prática, é importante assegurar a utilidade dos modelos face aos fins que justificam o seu desenvolvimento. Diferentes estudos destacam que a utilização direta dos modelos desenvolvidos em realidades distintas daquela onde os dados utilizados foram recolhidos não apresenta resultados satisfatórios ([17], [18]). Nestes estudos foram verificadas diversas incompatibilidades que não recomendam a aplicação de modelos desenvolvidos ao contexto diferente do qual foi obtido. Neste sentido, nos últimos anos diversos estudos de sinistralidade foram realizados em Portugal, envolvendo os diferentes elementos viários, segmentos ou interseções ([8], [19], [20], [21], [22], [23]). No entanto, nenhum desses trabalhos aborda os acidentes ocorridos nos diferentes tipos de interseções das EN's portuguesas.

Este trabalho, portanto, apresenta como objetivo o desenvolvimento de um modelo de previsão de acidentes para interseções de três ramos localizadas em estradas nacionais - EN's da região norte de Portugal. O foco principal da modelação é a identificação dos fatores contribuintes dos acidentes nesses locais. Na seleção das variáveis a serem investigadas como fatores contribuintes dos acidentes, foram incluídas variáveis relacionadas ao tráfego em cada local e variáveis que representam as características físicas do ambiente. Estas últimas incluem tanto variáveis que representam as características da interseção propriamente dita e do segmento viário localizado na sua área de influência, como também variáveis que refletem a diferença entre a geometria desse segmento e a geometria dos segmentos viários limítrofes. Esse último conjunto de variáveis visou permitir a avaliação do impacto da consistência do traçado sobre a ocorrência de acidentes nas interseções que foram objeto do estudo.

O estudo, portanto, parte da hipótese de que, a exemplo do que já foi verificado para os segmentos viários, a ocorrência de acidentes em interseções é afetada por mudanças acentuadas entre as características físicas do segmento incluído na área de influência da interseção e as correspondentes observadas nos segmentos viários contíguos a essa área. Assim, além da contribuição específica para a promoção da segurança viária nas estradas portuguesas, considera-se que a principal contribuição metodológica do presente estudo para a investigação de fatores contribuintes de acidentes em interseções está na investigação do impacto da consistência geométrica sobre a ocorrência de acidentes nestes locais.

## 2 METODOLOGIA

### 2.1 Definição e ajuste dos modelos

Para a modelação de dados longitudinais, diferentes abordagens podem ser empregadas ([24], [25]). Em todos os casos, é preciso definir a distribuição dos dados de contagem disponíveis, sendo a distribuição básica a de Poisson, que deve ser substituída pela binomial negativa quando houver superdispersão nos dados [10]. Portanto, a distribuição binomial negativa foi adotada para o presente estudo.

Definida a distribuição de probabilidade a ser adotada, a estimativa dos parâmetros do modelo pode ser feita com o uso de diferentes métodos. Considerou-se conveniente adotar os resultados produzidos pelas EEG (equações de estimação generalizadas), um dos métodos abordados por [24] e [25]. Portanto, a forma geral adotada neste trabalho para os modelos de previsão é a apresentada na Equação 1. Para o ajuste dos modelos, o período  $t$  é igual a um ano e a variável tráfego refere-se ao valor do tráfego médio diário anual - TMDA.

$$y_{mt} = \alpha \left\{ \left[ \text{Tráfego}_{mt} \right]^\gamma e^{\sum_j \beta_j x_{j,mt}} \right\} \quad (1)$$

em que:  $y_{mt}$  = Número esperado de acidentes para a interseção  $m$  no período de tempo  $t$ ;  
 $\text{Tráfego}_{mt}$  = Função de variáveis relacionadas com o tráfego na interseção  $m$  no período de tempo  $t$ ;  
 $x_{j,mt}$  = Variável explicativa  $i$  observada na interseção de três ramos  $m$  no período de tempo  $t$ ; e  
 $\alpha, \gamma, \beta_j$  = Parâmetros do modelo a serem estimados.

Para a escolha dum modelo de previsão que possa ser reconhecido como útil para a previsão de acidentes, aliou-se a validade do modelo sob o ponto de vista da engenharia de tráfego (um modelo que exclua a variável TMDA certamente não apresenta essa reconhecida utilidade) com o nível de significância desejável e possível, dadas as diversas limitações da base de dados disponível, sendo considerado um nível de significância em torno de 5%. O algoritmo de exclusão sequencial foi utilizado na escolha das variáveis.

Além disso, no uso dos métodos que abordam os dados em painel, a correta identificação do tipo de estrutura de correlação dos dados longitudinais de acidentes é essencial [26]. Estas estruturas permitem a especificação da correlação entre número de acidentes observado em diferentes períodos de tempo. Neste trabalho, foram testadas as seguintes estruturas de correlação: (i) independente; (ii) permutável; e (iii) autorregressiva.

### 2.2 Avaliação da qualidade do ajuste dos modelos

Nos estudos de sinistralidade com as EEG são comumente utilizados alguns elementos para avaliação do ajuste dos modelos, nomeadamente: o teste dos resíduos acumulados - *CURE Plot*,  $R^2$  marginal e o critério de informação quasi-verossimilhança (QIC). [27] refere que com o QIC são avaliadas as estruturas de correlação, sendo a selecionada a que apresentar menor QIC. Além disso, a medida de ajuste introduzida por [28] para os modelos ajustados com as EEG foi o  $R^2$  marginal. O valor de  $R^2$  marginal indica quanto a variância da variável resposta é explicada pela variabilidade do modelo ajustado.

O *CURE Plot* auxilia na seleção do melhor modelo. Segundo [29], na elaboração do *CURE Plot* considera-se a diferença entre o número de acidentes observado e o previsto (o resíduo) como o componente de base para julgar o ajuste, permitindo o exame dos resíduos acumulados em função da variável de interesse, aqui relacionada com o tráfego. Um bom ajuste significa que os resíduos acumulados no *CURE Plot* oscilam em torno de zero e não ultrapassam as duas curvas adicionais dos limites aceitáveis para os resíduos acumulados.

Num estudo de modelação é preciso garantir que os modelos desenvolvidos representam a população geral (generalização) e sejam apropriados às situações nas quais serão usados (transferibilidade), utilizando a teoria da validação. Para a validação dos modelos de previsão desenvolvidos será utilizada a validação cruzada *leave-one-out*, sugerida por [30] como uma alternativa quando a capacidade de recolher novos dados é limitada ou impraticável devido a fatores como o custo e o tempo. Os autores também sugerem que a raiz do erro quadrático médio - REQM é o principal elemento estatístico que deve ser utilizado na análise desta validação.

### 3 DESCRIÇÃO DOS DADOS

#### 3.1 Dados rodoviários

No Quadro 1 é apresentada a distribuição das 177 interseções de três ramos eleitas. Na seleção das interseções considerou-se: (i) via principal pertencente a uma EN com uma faixa de rodagem e duas vias de sentidos opostos; (ii) via secundária também com duas vias de sentidos opostos; (iii) via principal com e sem canalização; (iv) via secundária com e sem canalização; (v) centro das interseções afastados, no mínimo, 76m.

Quadro 1: Distribuição das 177 interseções estudadas

Trecho	Estrada Nacional	Interseções
Braga-Famalicão	EN-14	35
Braga-Guimarães	EN-101	30
Braga-Vila Verde	EN-101	17
Famalicão-Guimarães	EN-206	39
Guimarães-Felgueiras	EN-101	36
Guimarães-Fafe	EN-206	20
Total		177

#### 3.2 Dados de acidentes

Os dados dos acidentes de viação utilizados neste estudo foram fornecidos pela Autoridade Nacional de Segurança Rodoviária – ANSR e abrangem o período de 2008 a 2010. A ANSR mantém uma base de dados com informações recolhidas a partir do Boletim Estatístico de Acidentes de Viação – BEAV, que é preenchido no local da ocorrência do acidente de viação. Nesta base é possível identificar os acidentes de viação desagregados por tipo (colisão, despistes e atropelamentos).

A variável resposta utilizada neste estudo são os acidentes, excluindo os atropelamentos, sendo identificados 112 no período de estudo. Na contabilização dos acidentes ocorridos em cada interseção foi considerado uma área de influência de 76 metros antes e depois do centro da interseção na direção principal, tendo este valor sido recomendado pelo *Crash Prediction Module* - CPM [31]. Os acidentes ocorridos na direção secundária das interseções não foram considerados por indisponibilidade de dados das entidades gestoras. Relativamente às estatísticas descritivas dos acidentes, observou-se que o número médio anual de ocorrências por local foi de 0,21, o desvio padrão é igual a 0,53 e os valores de mínimos e máximos de zero e cinco, respetivamente.

#### 3.3 Dados de tráfego

O tráfego médio diário anual - TMDA nas aproximações das interseções é um indicador de exposição ao risco de acidentes essencial e por isso obrigatoriamente presente nos modelos de previsão. Em Portugal, a entidade Estradas de Portugal - EP é responsável pela recolha de dados de tráfego na rede rodoviária. Portanto, foi necessário obter a série histórica do TMDA das interseções para o período de 2008 a 2010 mediante contagens de tráfego na direção principal e secundária das interseções e também com a análise e tratamento dos dados disponibilizados pela EP. Além disso, no processo de modelação foram definidas cinco variáveis de tráfego em função da forma funcional adotada, sendo elas: (i) TMDAP (tráfego médio diário anual na direção principal); (ii) TMDAS (tráfego médio diário anual na direção secundária); (iii) TMDAT (soma do TMDAP com o TMDAS); (iv) TMDAPMS (multiplicação do TMDAP pelo TMDAS); e (v) TMDASDP (divisão do TMDAS pelo TMDAP). O TMDAP apresentou uma média de 16.457 com desvio padrão de 7.662, enquanto que os valores mínimos e máximos foram 3.160 e 38.753, respetivamente. O TMDAS teve média de 883, desvio padrão de 1.344, mínimo de 15 e máximo 7.896.

#### 3.4 Tratamentos dos dados geométricos das interseções

Para o estudo das interseções de três ramos foram recolhidas informações das características geométricas, da ocupação marginal do ambiente envolvente da interseção, do sistema de controlo de tráfego e da sinalização. Estas características foram tratadas como variáveis explicativas iniciais na modelação. Também foram recolhidos dados sobre as características geométricas dos dois segmentos de 200 metros de comprimento que fazem limite com a área de influência de cada interseção na direção principal. Dadas as particularidades das interseções estudadas foram definidos dois grupos de variáveis para serem considerados na modelação, nomeadamente: (i) Grupo 1, que inclui características específicas das interseções e respetivas áreas de influência,

remetendo às variáveis utilizadas tradicionalmente no estudo de sinistralidade; e (ii) Grupo 2, referente às diferenças entre as características da área de influência da interseção (representadas por algumas das variáveis do Grupo 1) e dos segmentos que com ela fazem limite<sup>1</sup>, tendo como objetivo estudar o impacto da consistência geométrica do traçado sobre a frequência de acidentes. As estatísticas descritivas dos dois grupos de variáveis são apresentadas nos Quadros 2 e 3.

Quadro 2: Estatísticas descritivas das variáveis do Grupo 1 das 177 interseções

Variável	Unidade	Média	D. P.	Mín	Máx
Largura média das vias nos 152 m da área de influência da direção principal – LVP	m	3,67	0,31	3,03	5,17
Largura média das vias nos 76 m da área de influência da direção secundária – LVS	m	3,62	0,97	2,20	7,76
Largura média da berma nos 152 m da área de influência da direção principal – BVP	m	0,70	0,38	0,00	2,15
Largura média da berma nos 76 m da área de influência da direção secundária – BVS	m	0,10	0,36	0,00	3,20
Desobstrução lateral média nos 152 m da área de influência da direção principal – DLVP	m	2,55	1,29	0,40	8,24
Desobstrução lateral média nos 76 m na área de influência da direção secundária – DLVS	m	1,02	1,29	0,00	8,25
Presença de travessias pedonais na direção principal – TPVP	“dummy”	0,26	0,44	0,00	1,00
Presença de travessias pedonais na direção secundária – TPVS	“dummy”	0,09	0,29	0,00	1,00
Presença de canalização na direção principal – PCVP	“dummy”	0,12	0,32	0,00	1,00
Presença de canalização na direção secundária – PCVS	“dummy”	0,51	0,50	0,00	1,00
Placas de sinalização na direção principal – PSVP	“dummy”	0,87	0,34	0,00	1,00
Placas de sinalização na direção secundária – PSVS	“dummy”	0,97	0,18	0,00	1,00
Semáforos na direção principal – S	“dummy”	0,04	0,20	0,00	1,00
Classificação da área envolvente do segmento onde a interseção está inserida (1 – rural ou periurbana e 0 – urbana) – D1	“dummy”	0,62	0,49	0,00	1,00
Classificação da área envolvente do segmento onde a interseção está inserida (1 – urbana ou periurbana e 0 – rural) – D2	“dummy”	0,91	0,29	0,00	1,00
Proporção do comprimento da reta nos 152 metros de área influente da interseção – PCRI	-	0,71	0,27	0,00	1,00
Inverso do raio médio das curvas nos 152 metros de área influente da interseção – IRHI	1/m	0,01	0,01	0,00	0,03
Sinuosidade em planta dos 152 metros de área influente da interseção – SHI	grau/m	132,84	158,76	0,00	635,75
Proporção do comprimento do tranel nos 152 metros de área influente da interseção – PCTI	-	0,71	0,42	0,00	1,00
Sinuosidade em perfil dos 152 metros de área influente da interseção – SVI	%/m	7,27	10,12	0,00	40,50
Densidade de pontos de acesso nos 152 metros de área influente da interseção – DPAI	nº/km	27,40	14,63	0,00	75,00

Quadro 3: Estatísticas descritivas das variáveis do Grupo 2 das 177 interseções

Variável	Unidade	Média	D. P.	Mín	Máx
Diferença absoluta entre ML e LVP – DLVP	m	0,21	0,24	0,00	2,01
Diferença absoluta entre MB e BVP – DB	m	0,29	0,33	0,00	1,95
Diferença absoluta entre MDL e DLVP – DDL	m	0,77	0,91	0,00	4,82
Diferença absoluta entre MPCR e PCRI – DPCR	-	0,21	0,17	0,00	0,67
Diferença absoluta entre MIRH e IRHI – DIRH	1/m	0,00	0,00	0,00	0,01
Diferença absoluta entre MSH e SHI – DSH	grau/m	120,99	151,47	0,00	764,28
Diferença absoluta entre MPCT e PCTI – DPCT	-	0,17	0,24	0,00	1,00
Diferença absoluta entre MSV e SVI – DSV	%/m	5,42	8,83	0,00	47,90
Diferença absoluta entre MDPA e DPAI – DDPA	nº/km	16,74	14,39	0,00	92,50

## 4 DESENVOLVIMENTO DOS MODELOS DE PREVISÃO

### 4.1 Resultados

Nesta seção são apresentados os modelos de previsão para as 177 interseções de três ramos. Para a modelação foram utilizados diferentes grupos de variáveis, definidos em função das características das interseções e dos

<sup>1</sup>Variáveis que contemplam características médias dos dois segmentos que fazem limite com a área de influência de cada interseção, sendo elas: (i) Média da largura da via nos segmentos que fazem limite com área de influência de cada interseção – ML; (ii) Média da largura da berma nos segmentos que fazem limite com área de influência de cada interseção – MB; (iii) Média da desobstrução lateral dos segmentos que fazem limite com área de influência de cada interseção – MDL; (iv) Média da proporção do comprimento em reta dos segmentos que fazem limite com área de influência de cada interseção – MPCR; (v) Média do inverso do raio das curvas horizontais dos segmentos que fazem limite com área de influência de cada interseção – MIRH; (vi) Média da sinuosidade em planta dos segmentos que fazem limite com área de influência de cada interseção – MSH; (vii) Média da proporção do comprimento em tranel dos segmentos que fazem limite com área de influência de cada interseção – MPCT; (viii) Média da sinuosidade em perfil dos segmentos que fazem limite com área de influência de cada interseção – MSV; (ix) Média da densidade de pontos de acesso dos segmentos que fazem limite com área de influência de cada interseção – MDPA.

segmentos de 200 metros que fazem limite com a área de influência de cada interseção. Assim, foram testadas três combinações destes grupos de variáveis, nomeadamente: (i) Combinação 1: Tráfego; (ii) Combinação 2: Tráfego + Grupo 1; e (iii) Combinação 3: Tráfego + Grupo 1 + Grupo 2. Além disso, nesta modelação foram testadas quatro formas funcionais recomendadas por [32]. Estas formas levam em conta diferentes configurações do tráfego da direção principal e da direção secundária (Quadro 4).

Quadro 4: Formas funcionais testadas nos modelos de previsão das interseções

Nº	Formas funcionais	Observações
1	$\ln y_{mt} = \beta_0 + \beta_1 \times \ln(TMDAT_{1,mt})$	$TMDAT = (TMDAP + TMDAS)$
2	$\ln y_{mt} = \beta_0 + \beta_1 \times \ln(TMDAP_{1,mt}) + \beta_2 \times \ln(TMDAS_{2,mt})$	-
3	$\ln(y_{mt}) = \beta_0 + \beta_1 \times \ln(TMDAPMS_{1,mt})$	$TMDAPMS = (TMDAP \times TMDAS)$
4	$\ln y_{mt} = \beta_0 + \beta_1 \times \ln(TMDAT_{1,mt}) + \beta_2 \times \ln(TMDASDP_{1,mt})$	$TMDASDP = \left( \frac{TMDAS}{TMDAP} \right)$

$y_{mt}$ : Número esperado de acidentes no elemento  $m$  no ano  $t$ ; TMDAP: TMDA na direção principal; TMDAS: TMDA na direção secundária

Nos modelos de previsão ajustados foi considerado o período de 2008 a 2010, por ser o período em que provavelmente não houve intervenções nas interseções, sendo que esta informação não foi confirmada pelas autoridades competentes. Todavia, no Quadro 5 são apresentados os resultados da modelação. Como pode ser observado, quando utilizada somente a variável de tráfego, Combinação 1, o modelo de previsão que apresentou melhor resultado para um nível de significância de 5% foi o da forma funcional 3. Neste quadro também pode ser observado que a melhor estrutura de correlação de trabalho a foi autorregressiva. Relativamente ao *CURE Plot*, verificou-se que a curva dos resíduos acumulados não ultrapassa os limites, inferior e superior. Quanto ao  $R^2$  marginal o modelo selecionado expôs um valor em torno de 3%.

Quadro 5: Resumo dos modelos

Combinação	Parâmetros	Estrutura de Correlação								
		Independente			Permutável			Autorregressiva		
		Coef.	D. P.	$p$	Coef.	D. P.	$p$	Coef.	D. P.	$p$
1	Interceto	-5,13	1,50	<0,01	-5,13	1,50	<0,01	-5,10	1,48	<0,01
	$\ln(TMDAPMS)$	0,23	0,09	0,02	0,23	0,09	0,02	0,22	0,09	0,02
	QIC	549,84			549,85			549,30		
	$R^2$ marginal	0,03			0,03			0,03		
2	Interceto	-6,48	2,33	0,01	-6,49	2,33	0,01	-6,39	2,31	0,01
	$\ln(TMDAT)$	0,46	0,24	0,06	0,46	0,24	0,06	0,45	0,24	0,06
	TPVP	0,48	0,24	0,05	0,48	0,24	0,05	0,49	0,24	0,04
	D2	0,48	0,25	0,06	0,48	0,25	0,06	0,48	0,25	0,06
	QIC	533,24			533,26			531,60		
	$R^2$ marginal	0,02			0,02			0,02		
3	Interceto	-4,56	1,34	<0,01	-4,57	1,34	<0,01	-4,56	1,33	<0,01
	$\ln(TMDAPMS)$	0,16	0,08	0,05	0,16	0,08	0,05	0,16	0,08	0,05
	PSVP	-0,59	0,29	0,04	-0,59	0,29	0,04	-0,59	0,29	0,04
	D2	0,60	0,23	0,01	0,60	0,23	0,01	0,60	0,23	0,01
	DB	0,76	0,26	<0,01	0,76	0,26	<0,01	0,76	0,26	<0,01
	DSV	0,04	0,01	<0,01	0,04	0,01	<0,01	0,04	0,01	<0,01
	QIC	570,90			570,90			570,89		
	$R^2$ marginal	0,08			0,08			0,08		

Relativamente aos resultados dos modelos ajustados com a Combinação 2, composta pelas variáveis de tráfego mais as variáveis do Grupo 1, a forma funcional para o tráfego que melhor se adaptou foi a número 1 (Quadro 5). Para esta combinação foram obtidos modelos com nível de significância em torno de 5%. As variáveis selecionadas apresentaram resultados em conformidade com a engenharia de tráfego, sendo elas: (i) travessias pedonais na direção principal – TPVP; (ii) classificação do ambiente envolvente do segmento onde se encontra inserida a interseção – D2, ambos com contribuição negativa para a segurança viária (seus coeficientes apresentam valores positivos no modelo). Para a variável TPVP este resultado justifica-se com a perturbação no tráfego que ocorre na presença de travessias pedonais contribuindo para ocorrência de acidentes. Quanto à variável D2, verifica-se que há maior ocorrência de acidentes somente com danos materiais em zonas urbanas, porém, quando analisado somente os acidentes com vítimas, objeto do presente estudo, a zona urbana influi de forma positiva na segurança viária. Além disso, a melhor estrutura de correlação de trabalho foi a autorregressiva. No que diz respeito ao *CURE Plot* deste modelo, observa-se na Figura 1(a) que a curva dos

resíduos acumulados oscila em torno de zero e não toca o limite inferior e superior. O  $R^2$  marginal deste modelo previsão foi próximo a 2%.

Por fim, como pode ser observado no Quadro 5, quando utilizada a variável de tráfego juntamente com as variáveis do Grupo 1 e do Grupo 2 (Combinação 3), o modelo de previsão também teve melhores resultados para forma funcional 3 e nível de significância de 5%. Entre os modelos ajustados, a estrutura de correlação de trabalho autorregressiva também foi a que apresentou melhor desempenho. Para este modelo foram identificados quatro fatores contribuintes para ocorrência dos acidentes além do tráfego, nomeadamente: (i) a variável “dummy” presença de sinalização na direção principal - PSVP; (ii) a variável “dummy” do ambiente envolvente do segmento onde se encontra inserida a interseção, D2; (iii) a variável DB, composta pela diferença absoluta entre BVP (largura média da berma na direção principal) e MB (média da largura da berma nos segmentos de 200 metros que fazem limite com a área de influência da interseção); (iv) a variável DSV, composta pela diferença entre SVI (sinuosidade em perfil da interseção) e MSV (média da sinuosidade em perfil dos segmentos de 200 metros que fazem limite com a área de influência da interseção). Estas variáveis apresentaram diferentes contribuições, sendo que D2, DB e DSV mostraram-se como fatores contribuintes negativos para a segurança viária (coeficientes positivos) e a PSVP um fator contribuinte positivo (coeficiente negativo). A contribuição positiva de PSVP pode estar associada ao fato de que a presença de sinalização nas estradas cumpre o dever de chamar a atenção dos condutores para diferentes situações de risco de acidentes e assim o condutor poder assumir uma condução defensiva. A contribuição negativa apresentada pelos fatores DB e DSV pode ser explicada pelo fato de que alterações acentuadas no traçado ao longo da estrada exigem, em geral, frequentes ajustes na forma de condução do veículo e nas velocidades praticadas, o que favorece a ocorrência de acidentes de viação. Além disso, o *CURE Plot* do modelo selecionado (Figura 1.b) mostra que a curva dos resíduos acumulados oscila em torno de zero e não ultrapassam os limites, inferior e superior. Este modelo selecionado apresentou um  $R^2$  marginal de cerca de 8%.

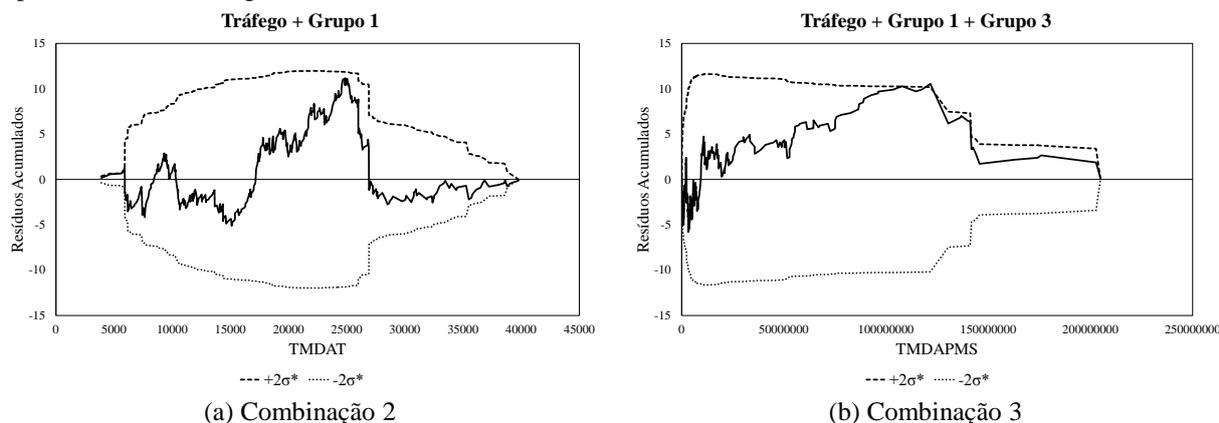


Figura 1: *CURE Plot* dos modelos de previsão de acidentes ajustado com as três combinações

## 4.2 Análise dos resultados

No Quadro 6 apresentam-se as formas funcionais dos modelos selecionados como os mais adequados para cada Combinação utilizada (1, 2 ou 3) e os elementos estatísticos de análise destes modelos (nível de significância, QIC, estrutura de correlação,  $R^2$  marginal e raiz do erro quadrático médio - REQM). O elemento REQM é apresentado para uma análise complementar dos modelos e também para uma posterior comparação com os elementos calculados no processo de validação.

Como pode ser observado na Quadro 6 os modelos finais apresentam um nível de significância de 5%. Relativamente ao QIC dos modelos ajustados para as Combinações 1 e 3 dos acidentes, houve um aumento deste valor. Segundo [27], um aumento do número de variáveis tende a aumentar o valor de QIC, o que se confirma com um aumento de cerca de 4% no valor do QIC dos modelos finais das Combinações 1 e 3, que apresentam a mesma forma funcional. Em todos os modelos, a estrutura de correlação de trabalho dominante foi a autorregressiva. Relativamente ao  $R^2$  marginal, os modelos finais apresentaram resultados similares a outros estudos desenvolvidos no contexto da sinistralidade, como por exemplo, os modelos apresentados por [4].

Quadro 6: Elementos estatísticos dos modelos finais ajustados para interseções

Combinação	Forma Funcional	Nível de significância	QIC	Estrutura de correlação	R <sup>2</sup> marginal	REQM
1	3	5%	549,30	Autorregressiva	0,03	0,5173
2	1	5%	531,60	Autorregressiva	0,02	0,5195
3	3	5%	570,90	Autorregressiva	0,08	0,5049

Para validação dos modelos de previsão foi utilizada a validação cruzada K-fold do tipo *leave one out*, visto que o tamanho das amostras dos modelos permite este tipo de validação sem grandes custos computacionais. No Quadro 7 são ilustrados os elementos estatísticos REQM da validação cruzada e dos modelos de previsão ajustados. Os resultados apresentados mostram que a variação do REQM se compreendeu entre 5,17% e 5,70%. Ou seja, pela técnica da validação cruzada os modelos apresentaram melhor desempenho na validação do que no ajuste, e diferenças percentuais entre as estatísticas inferiores a 10%. Assim, em geral, os modelos podem ser considerados válidos, especialmente o modelo das Combinação 3, que apresentou o menor valor para as diferenças percentuais entre as REQM da validação e de ajuste.

Quadro 7: Elementos estatísticos da validação cruzada *leave one out* para os modelos das interseções

Combinação	Média dos Acidentes	REQM		
		Validação Cruzada	Ajuste	Variação
1	0,21	0,4878	0,5173	-5,70%
2	0,21	0,4916	0,5195	-5,37%
3	0,21	0,4788	0,5049	-5,17%

## 5 CONCLUSÕES

O objetivo principal do presente estudo foi a identificação dos principais fatores que contribuem para a ocorrência dos acidentes de viação, através do desenvolvimento de modelos de previsão para as interseções de três ramos inseridas em EN's de uma faixa de rodagem com uma via em cada sentido e localizados na região norte de Portugal. A contribuição deste trabalho para a sociedade em geral é disponibilizar uma ferramenta fundamental para a promoção da segurança no sistema nacional rodoviário do norte de Portugal, que liga muitas cidades e zonas industriais. Para a sua realização foram utilizados dados de acidentes, de tráfego e das características geométricas e de ambiente envolvente das interseções. Assim, os modelos de previsão foram desenvolvidos com o auxílio das equações de estimação generalizadas - EEG, considerando a componente de erro uma distribuição binomial negativa.

Para o desenvolvimento dos modelos de previsão foi elaborada uma base de dados em que foram observados os dados dos acidentes, o tráfego (representado pelo tráfego médio diário anual – TMDA da direção principal e da direção secundária) e as características geométricas para o período de estudo (2008 a 2010). Relativamente às variáveis que compõem as características geométricas foram utilizadas diferentes combinações dos grupos de variáveis (Grupo 1 e Grupo 2), descritos na seção 3. Estes grupos foram definidos em função das características das interseções e dos segmentos de 200 metros de comprimento que fazem limite com a área de influência de cada interseção, buscando com a inclusão das variáveis destes grupos, incluir nos modelos de previsão a consistência do traçado da estrada.

Nos modelos de previsão obtidos verificou-se que a variável tráfego, representada pelo TMDA na direção principal e na secundária das interseções, produziu um bom ajuste em todos os casos e tem uma contribuição negativa para a segurança viária, correspondendo a um incremento no número de acidentes com o aumento do tráfego. Os valores dos coeficientes das variáveis de tráfego foram diferentes conforme a combinação utilizada.

Para o ajuste dos modelos de previsão foram consideradas três Combinações (Tráfego, Tráfego + Grupo 1 e Tráfego + Grupo 1 + Grupo 2). Nos modelos finais em que foram consideradas além do tráfego as características geométricas das interseções e dos segmentos que fazem limite com a área de influência das interseções compostas pelos dois grupos apresentados, as variáveis independentes identificadas como estatisticamente significativas e associadas a acréscimos na frequência de acidentes foram: a variável travessias pedonais na direção principal – TPVP; a variável do ambiente envolvente do segmento onde se encontra inserida a interseção – D2; a variável composta pela diferença absoluta entre BVP (largura média da berma na direção principal) e MB (média da largura da berma nos segmentos de 200 metros que fazem limite com a área de influência da interseção) – DB; a variável composta pela diferença entre SVI (sinuosidade em perfil da interseção) e MSV (média da sinuosidade em perfil dos segmentos de 200 metros que fazem limite com a área de influência da

interseção) – DSV. Apenas uma variável apresentou sinal negativo, que está associado ao decréscimo da frequência dos acidentes, sendo esta variável a presença de sinalização na direção principal – PSVP.

Relativamente à validação dos modelos obtidos, com os resultados da validação cruzada *leave one out* foi possível verificar que os modelos apresentam um bom desempenho, traduzidos por baixos valores dos elementos estatísticos dos modelos obtidos no ajuste e os valores alcançados com a validação.

Por fim, o estudo mostra através da análise conjunta do *CURE Plot*, do  $R^2$  marginal e dos elementos estatísticos da validação cruzada dos modelos desenvolvidos que a Combinação 3 apresenta melhores resultados (Quadro 6, Quadro 7 e Figura 1). Este resultado confirma a crescente necessidade de ser incorporado aos modelos de previsão o efeito da variação das condições gerais do traçado das estradas. Essa variação, especialmente se for acentuada, faz com que o condutor seja surpreendido ao longo do percurso pelas alterações da geometria, o que pode dificultar sua pronta reação para ajustar a condução do veículo às condições prevalecentes da via.

Portanto, considera-se que o presente trabalho estabelece uma contribuição importante para a melhoria da segurança rodoviária nas EN's da região do Minho no norte de Portugal, visto que os modelos de previsão desenvolvidos podem ser utilizados diretamente pelos responsáveis da gestão da infraestrutura rodoviária no sentido de fundamentar intervenções na mesma, direcionadas para a mitigação da sinistralidade. No entanto, destaca-se que a principal limitação do presente estudo foi a ausência da localização dos acidentes na direção secundária das interseções.

## 6 AGRADECIMENTOS

Os autores agradecem o suporte financeiro concedido pela Fundação para a Ciência e a Tecnologia (FCT), Portugal, através do financiamento da bolsa de doutoramento SFRH/BD/62458/2009 e do projeto UI 4047 – 2014: PEst-OE/ECI/UI4047/2014.

## REFERÊNCIAS

1. WHO - World Health Organization, 2013. Global status report on road safety 2013: supporting a decade of action. Geneva, World Health Organization.
2. Autoridade Nacional de Segurança Rodoviária – ANSR, Sinistralidade Rodoviária, *Observatório de Segurança Rodoviária*, Lisboa, 2013.
3. D. Lord, B. Persaud, Accident Prediction Models With and Without Trend – Application of the Generalized Estimating Equations Procedure. *Transportation Research Record*, 1717(1) 102-108, 2000.
4. Wang, X., Abdel-Aty, M. Temporal and spatial analysis of rear-end crashes at signalized intersections. *Accident Analysis and Prevention*, 38(1) 1137-1150, 2006.
5. P. C. Anastasopoulos, F. L. Mannering, A note on modeling vehicle accident frequencies with random-parameters count models, *Accident Analysis and Prevention*, 41(1) 153-159, 2009.
6. D. Lord, M. Mahlawat, Examining the application of aggregated and disaggregated Poisson-gamma models subjected to low sample mean bias, *Transportation Research Record*, 2136(1) 1-10, 2009.
7. S. Cafiso, A. Di Graziano, G. Di Silvestro, G. La Cavaa, B. Persaud, Development of comprehensive accident models for two-lane rural highways using exposure, geometry, consistency and context variables. *Accident Analysis and Prevention*, 42(1) 1072-1079, 2010.
8. S. Gomes, The influence of the infrastructure characteristics in urban road accidents occurrence. *Accident Analysis and Prevention*, 60(1) 289-297, 2013.
9. K. Fitzpatrick, D. Lord, B. Park, Accident modification factors for medians on freeways and multilane rural highways in Texas. *Transportation Research Record*, 2083(1) 62-71, 2008.
10. D. Lord, F. Mannering, The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A*, 44(1) 291-305, 2010.

11. C. Bhat, K. Born, R. Sidharthan, P. Bhat. A count data model with endogenous covariates: formulation and application to roadway crash frequency at intersections. *Analytic Methods in Accident Research*, 1(1) 53-71, 2014.
12. D. Kim, S. Washington, The significance of endogeneity problems in crash models: An examination of left-turn lanes in intersection crash models. *Accident Analysis and Prevention*, 38(1) 1094-1100, 2006.
13. P. Anastasopoulos, F. Mannering, V. Shankar, J. Haddock, A study of factors affecting highway accident rates using the random-parameters tobit model, *Accident Analysis and Prevention*, 45(1) 628-633, 2012.
14. J. Mayora, R. Rubio. Relevant variables for crash rate prediction in Spain's two lane rural roads. In: Proceedings of the 82<sup>nd</sup> Transportation Research Board Annual Meeting, Washington DC, USA. 2003.
15. A. Polus, C. Mattar-Habib, New Consistency Model for Rural Highways and Its Relationship to Safety, *Journal of Transportation Engineering*, 130(3) 286-293, 2004.
16. S. Mitra, S. Washington, On the significance of omitted variables in intersection crash modeling, *Accident Analysis and Prevention*, 49(1) 439-448, 2012.
17. Z. Sawalha, T. Sayed, Traffic accident modeling: some statistical issues, *Canadian Journal of Civil Engineering*, 33(9) 1115-1124, 2006.
18. E. Sacchi, B. Persaud, M. Bassani, Assessing International Transferability of the Highway Safety Manual Crash Prediction Algorithm and its Components. Proceedings of the 91<sup>st</sup> Annual Meeting of Transportation Research Board. Washington: TRB. 2012.
19. A. Couto, S. Ferreira, A note on modeling road accident frequency: A flexible elasticity model, *Accident Analysis and Prevention*, 43(1) 2104-2111, 2011.
20. S. Ferreira, A. Couto, Traffic flow-accidents relationship for urban intersections on the basis of the translog function, *Safety Science*, 60(1) 115-122, 2011.
21. S. Gomes, J. Cardoso, Safety effects of low-cost engineering measures. An observational study in a Portuguese multilane road. *Accident Analysis and Prevention*, 48(1) 346-352, 2012.
22. S. Gomes, S. Geedipally, D. Lord, Estimating the safety performance of urban intersections in Lisbon, Portugal, *Safety Science*, 50(1) 1732-1739, 2012.
23. J. Costa, E. Freitas, P. Pereira, M. Jacques. Portuguese two-lane highways: modeling crash frequencies for different temporal and spatial aggregations of crash data. *Transport (Lithuania)*. 2015.
24. J. Hilbe, Negative Binomial Regression. Cambridge University Press, New York. 2011.
25. W. Greene, Econometric Analysis. Prentice Hall, New York. 2007.
26. K. Liang, S. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73 13-22, 1986.
27. Pan, W. Akaike's Information Criterion in Generalized Estimating Equations. *Biometrics*, 57 120-125, 2001.
28. Zheng, B., Summarizing the goodness of fit of generalized linear models for longitudinal data. *Statistics in Medicine*, 19(10) 1265-1275, 2000.
29. E. Hauer, Statistical Road Safety Modeling. *Transportation Research Record*, 1897 81-87, 2004.
30. T. Hastie, R. Tibshirani, J. Friedman. The Elements of Statistical Learning: Prediction, Inference and Data Mining (2<sup>nd</sup> ed.). EUA: Springer Verlag. 2009.
31. FHWA - Federal Highway Administration, TFHRC - Turner-Fairbank Highway Research Center, *Crash Prediction Module (CPM) Engineer's Manual*, EUA: IHSDM, 2003.
32. D. Lord, P. Park, Investigating the effects of the fixed and varying dispersion parameters of Poisson-gamma models on empirical Bayes estimates. *Accident Analysis and Prevention*, 40(4) 1441-1457, 2008.