

CONSTRUÇÃO DE UMA BASE DE DADOS PARA DESENVOLVIMENTO DE MODELOS DE ESTIMATIVA DA FREQUÊNCIA DE ACIDENTES

Hudson Carrer Pereira¹, Ana Bastos Silva² e Alvaro Seco²

¹Universidade de Coimbra, Centro de Investigação do Território, Transportes e Ambiente (CITTA), Departamento de Engenharia Civil, Rua Luís Reis Santos – Polo II, 3030-788 Coimbra, Portugal

email: hudcp@yahoo.com.br <https://citta.fe.up.pt>

²Universidade de Coimbra, Centro de Investigação do Território, Transportes e Ambiente (CITTA), Departamento de Engenharia Civil, Rua Luís Reis Santos – Polo II, 3030-788 Coimbra, Portugal

Sumário

O presente artigo descreve a metodologia adotada na recolha e tratamento de dados para apoio ao desenvolvimento de modelos de estimativa da frequência de acidentes aplicados ao Distrito Federal, Brasil. A base de dados incide sobre as vias arteriais e abrange o período de 2009 a 2017, sendo condicionada a acidentes rodoviários com fatalidades e feridos. São apresentados os critérios de divisão dos segmentos; compilação da informação de acidentes e vítimas com posterior georreferenciação; uso das ferramentas box plot e jackknife para tratamento do número de veículos; seleção das variáveis de infraestrutura; registo de velocidade em regime de livre circulação; recomendações ao órgão gestor para melhoria da base de dados.

Palavras-chave: Modelos de estimativa da frequência de acidentes; Técnicas estatísticas; Vias arteriais.

1 INTRODUÇÃO

Os modelos de estimativa da frequência de acidentes (MEFA) relacionam a frequência de acidentes a um conjunto de variáveis que permitam estimar e explicar o fenómeno. Este trabalho centra-se na descrição da metodologia empregue nos trabalhos de recolha de dados de campo, bem como nos trabalhos de compatibilização de dados oriundos de diferentes fontes de informação para o desenvolvimento de MEFA em segmentos de vias arteriais, aplicáveis ao Distrito Federal, Brasil.

O Departamento de Trânsito (Detran/DF) é o órgão responsável [1] no Distrito Federal por recolher e processar os registos de acidentes de trânsito, assim como por fiscalizar o funcionamento da via, dispondo nas vias selecionadas e para o período em estudo (2009 a 2017) de 145 equipamentos de fiscalização eletrónica com contagem de veículos. O acesso restrito aos dados primários de acidentes e ao Sistema de Informações Georreferenciadas onde se armazenam os dados derivados da fiscalização eletrónica limita o desenvolvimento dos MEFA, conhecendo-se apenas a existência de um estudo [2] no Distrito Federal aplicado a interseções de uma única região administrativa (Taguatinga/DF).

Nesta mesma linha de ação, o presente estudo centra-se na apresentação dos trabalhos preparatórios ao desenvolvimento de MEFA aplicáveis a segmentos de vias arteriais do Distrito Federal. São apresentadas as variáveis dependentes (total de acidentes e por tipos de acidentes) e as variáveis explicativas consideradas. É apresentado o referencial teórico, as especificidades características das vias do Distrito Federal, os dados fornecidos pelas diferentes instituições parceiras e a metodologia adotada na construção da base de dados.

2 REFERENCIAL TEÓRICO

O método preditivo aplicado a vias arteriais urbanas [3] passa pela concretização de várias etapas para desenvolvimento de MEFA. Essas etapas integram desde a definição do período de estudo, a determinação do tráfego diário médio anual (TDMA), o levantamento das características geométricas dos trechos, a divisão dos

trechos em segmentos, qualidade e precisão dos acidentes e a respetiva alocação aos locais específicos etc. [3]. A seleção das variáveis explicativas [4] depende da disponibilidade de dados; do efeito previsível que cada uma possa exercer sobre a frequência de acidentes; da forma como podem ser recolhidas de forma fácil e fidedigna; da salvaguarda da independência entre cada uma das variáveis integradas no modelo.

A construção da base de dados tem sua importância reforçada na adoção da abordagem da Análise Exploratória de Dados (AED) [5], de modo que os próprios dados revelem a seleção das variáveis explicativas com influência na frequência de acidentes, sendo parte integrante da construção do ciclo de modelagem. Como exemplo, o autor cita que, se uma determinada característica (variável) afeta os acidentes, então é necessário um MEFA específico para esse tipo de acidente e via. Nesse sentido, as variáveis a serem utilizadas nos modelos dependem do objetivo implícito ao desenvolvimento do MEFA. Para o presente trabalho isto implicará na retirada dos dados de acidentes para períodos e segmentos específicos em que, por exemplo, não há contagem de veículos pelo não funcionamento de um determinado equipamento.

Dentre as variáveis que merecem destaque no desenvolvimento de MEFA está o Tráfego Diário Médio Anual (TDMA), sendo comumente a variável de exposição utilizada (e única) num modelo simplificado, quando o foco está na estimação da frequência de acidentes, ou em modelos mais complexos sempre que se procura identificar as variáveis explicativas mais significativas [6]. Quando o TDMA não está disponível para todo o período de análise [7], poderá ser inferido por recurso a operações simples de interpolação/extrapolação. Na realidade, o TDMA pode ser estimado tendo por base a série histórica do tráfego ou a partir de uma amostra com contagens classificadas.

Em relação ao sentido do tráfego, considerando-se o perfil pendular do tráfego do Distrito Federal (dada a atração de viagens a Brasília), foram criadas duas bases de dados: uma em que o TDMA corresponde ao tráfego médio diário nos dois sentidos de tráfego, relativos ao ano em análise; outra em que, para cada sentido da via, o TDMA e número de acidentes são introduzidos separadamente. Tal segregação deverá assim potencializar avaliar se, seja nos MEFA, seja na AED, identificam-se alterações significativas na frequência e tipologia dos acidentes, sempre que no mesmo trecho de estrada um sentido ofereça condições livres de circulação e o outro encontre-se congestionado.

Outra questão que afeta a escolha dos modelos e, por conseguinte, a construção das bases é o período de análise. A aleatoriedade de ocorrência dos acidentes e a tendência de regressão à média (curto prazo) em contraste com as alterações das características viárias (longo prazo) influenciam o período de análise, sendo recomendado [3] estimar a frequência de acidentes para as condições específicas para cada ano em um período de estudo. Ainda que em alguns modelos simplificados esse fator forneça bons ajustes na estimação da frequência de acidentes, em contraste com a grande variação ao longo dos dias e semanas que se verifica no tráfego, o TDMA representa um valor médio anual, envolvendo uma perda significativa de informação relativa à variabilidade diária, semanal e mensal. De forma a avaliar a eventual relevância desse efeito, foi desenvolvida uma segunda base em que, além do TDMA, considerou-se o tráfego médio anual por hora e dia da semana.

Destaca-se, por fim, a inclusão da velocidade na tentativa de captar a influência desta variável nos acidentes. Ainda que a velocidade seja um dos principais fatores contribuintes de acidentes, independentemente da robustez dos modelos estatísticos, os resultados científicos incidentes sobre o efeito da velocidade nos acidentes são inconsistentes [8]. Dentre as explicações [9] sobre a relação contraintuitiva (maior velocidade, menos acidentes), destaca-se que os acidentes são eventos multicausais; a falta de dados de velocidade de tráfego adequados (dados de acidentes e velocidade em diferentes períodos no tempo); o viés de variáveis omitidas; modelos estatísticos inadequados. Assim, a variação da velocidade entre veículos numa corrente de tráfego e entre diferentes trechos de vias tenderá a afetar o risco e a gravidade dos acidentes.

3 DESCRIÇÃO DA METODOLOGIA

Este capítulo descreve cada etapa das atividades realizadas no processo de recolha e construção da base de dados.

3.1 Construção da Base de Dados

Com vista a coibir o excesso de velocidade e a reduzir pontualmente a velocidade em locais críticos, o Distrito Federal tem vindo a instalar equipamentos de fiscalização eletrónica em diversas vias arteriais e rodovias. No caso das vias secundárias e locais, o órgão tem optado maioritariamente pela implantação de lombas. Sem questionar a efetividade dessas soluções, o presente trabalho de investigação focou as vias urbanas arteriais pelas seguintes

razões de base: pela disponibilidade de contagens automáticas de tráfego; por constituir uma rede bem delimitada e de extensão moderada; por concentrar um número significativo de acidentes em vias urbanas no Distrito Federal (36% do total de acidentes e 43% dos acidentes fatais).

A base de dados de acidentes rodoviários criada no âmbito do presente trabalho tem por base a informação cedida pela Gerência de Estatística do Detran/DF correspondente aos últimos cinco anos (2011 a 2015). A seleção das vias teve por base as Instruções de Serviço nºs 311/2001 e 251/2004, que caracteriza e define a classe de vias arteriais no Distrito Federal, ainda que outras vias tenham sido construídas ou adquirido características semelhantes desde então. As contagens de veículos oriundas dos equipamentos de fiscalização eletrônica foram obtidas no Sistema de Informações Georreferenciadas (SIG-Detran/DF) do Núcleo de Geoprocessamento do Detran/DF.

Atendendo à disponibilidade dos dados sobre fluxos de veículos e com o objetivo de integrar na análise o maior número possível de regiões administrativas, ou RAs (o Distrito Federal está dividido em 31 RAs), os trabalhos de campo incidiram sobre a recolha de um conjunto alargado de características locais, a saber: número de vias do perfil transversal; separador central; existência (ou não) de estacionamento; tipologia das interseções; existência de via exclusiva para transporte público; presença de peões; existência de passeios; tipo de ambiente rodoviário e velocidade máxima legal estabelecida em cada local. Optou-se, ainda, por considerar a variável dependente como o número total de acidentes, com variações temporais e por tipo de acidente.

A partir da AED, verificou-se que o período de análise inicialmente previsto (2011 a 2015) deveria ser alargado (2009 a 2017) de forma a contemplar dois fenómenos relevantes: troca de equipamentos de fiscalização (que se pode traduzir em diferenças na contagem de veículos) e alteração do método de recolha de informações sobre alcoolemia e drogas, ambos alterados a partir de 2015. A título de exemplo, entre 2009 e 2014 registou-se a presença de álcool/drogas em 15% dos acidentes fatais, mas, após convênio com o Instituto Médico Legal, o percentual de acidentes fatais envolvendo álcool/drogas aumentou para 43% entre 2015 e 2017. Tal alteração deverá justificar modelos específicos para a estimação desse tipo de acidente, e apenas os dados 2015-2017 deverão ser considerados fiáveis.

Ao longo do processo de recolha de dados, várias melhorias foram solicitadas e implementadas no SIG-Detran/DF, destacando-se a segregação dos volumes de tráfego por gama de variação de velocidades e desagregadas por períodos de 15 minutos. A base de dados resultou numa matriz de 78.840 linhas (24h x 365 dias x 9 anos) para cada um dos 145 equipamentos de fiscalização presentes na amostra.

O número de acidentes nas vias urbanas (sem considerar rodovias), registado entre 2009 e 2017 no Distrito Federal, ascendeu a 53.141, dos quais 1.164 envolveram fatalidades (2,2%). Em termos do número de vítimas, registou-se um total de 63.354, das quais 1.207 fatais (1,9%). A amostra criada para suporte do presente estudo, limitada às vias arteriais, engloba um total de 6.319 acidentes, dos quais 176 envolveram fatalidades (2,8%).

3.2 Determinação do TDMA

Existem vários métodos de contagem de tráfego [10], entre os quais os métodos eletrônicos, os manuais e os baseados em filmagens. A ausência ou limitação das séries históricas do tráfego pode ser preenchida por linhas de tendência e recurso a contagens em alguns períodos e uso de fatores de expansão.

A metodologia inicial foi extraída de estudo [2] de MEFA para interseções em Taguatinga/DF, que por sua vez utilizou [11] a ferramenta computacional em *software* SAS para aplicação da ferramenta *jackknife*. Estes trabalhos [2, 11] foram desenvolvidos concomitantemente de forma a superar os períodos de não funcionamento e sub-registo dos equipamentos de fiscalização no Distrito Federal, bem como a calcular fatores de expansão para vias secundárias sem registo de dados. Os dados correspondem a séries temporais do número de veículos que passam por equipamentos de fiscalização eletrônica, detetados por laços indutivos instalados no pavimento.

O programa em SAS foi cedido graciosamente para desenvolvimento deste trabalho de investigação e aplicado depois de introduzidas as devidas adaptações. O objetivo da programação [11] foi executar o método de reamostragem de *jackknife* a fim de identificar sub-registos e valores em falta associados a períodos de inatividade do equipamento (por exemplo, por falta de energia), bem como realizar a imputação de dados nos casos identificados.

Antes de descrever o método de reamostragem, ressalta-se que faz parte da análise dos resultados [11] a presença de valores extremos superiores no banco de dados. Como a ferramenta *jackknife* trata somente dos valores

inferiores, substituindo-os pela média do conjunto, os valores extremos acabam por influenciar o valor a ser imputado.

Foram realizados dois passos prévios ao uso do *software*. Para cada série histórica foram excluídos os valores horários inicial e final relacionados com a inoperacionalidade do equipamento, presumindo que a contagem não se realizou durante todos os 60 minutos. Seguidamente aplicou-se a ferramenta *box plot* para identificar valores superiores e inferiores extremos para cada hora e separados por dia da semana num mesmo ano, os quais, depois de analisados, foram retirados do conjunto. Desse procedimento concluiu-se que, entre os meses de maio e setembro de 2017, foram inseridos de maneira duplicada no SIG-Detran/DF todos os equipamentos do tipo controlador eletrônico de velocidade localizado em segmentos (chamados de “pardais”). O problema foi relatado à entidade responsável e a base de dados foi corrigida e novamente inserida e tratada. O uso do *box plot* não suprimiu a exclusão dos dados de início e fim de paralisação, já que em algumas contagens o equipamento pode ter deixado de funcionar por poucos minutos (contagem incorreta), não sendo por isso considerado um valor discrepante inferior.

O passo seguinte incidiu sobre a aplicação da técnica de reamostragem de *jackknife* detalhada em [11]. De forma simplificada, a técnica consiste em calcular médias para cada conjunto de valores com critério de identificação por um limite inferior. Caso uma das médias seja menor que o limite inferior calculado, um novo valor com base na média e no desvio padrão do conjunto é imputado. O termo conjunto é considerado como o volume horário de determinado dia da semana em um mesmo mês e ano para um dado equipamento. Por exemplo, para janeiro de 2017, um dos conjuntos é o número de veículos que passou de 0h a 1h para cada segunda-feira, correspondendo a cinco observações. Para cada conjunto são realizadas no máximo três imputações, não se aplicando tratamento caso sejam necessárias mais que três imputações. Se isso ocorrer, sub-registos permanecem e os volumes são calculados sem a devida qualidade [11].

Para o conjunto de dias da semana, este trabalho considerou, em um único grupo, domingos, feriados e pontos facultativos. Foi calculada a média do volume para cada hora e adotado o TDMA como a somatória das horas médias calculadas. Isso permitiu considerar os dados em que o equipamento tenha funcionado parcialmente durante o dia. Essa disposição dos dados também permitiu criar uma base das médias horárias anuais para cada conjunto dos dias da semana, que por sua vez foram separadas em dois grupos: volume horário durante a semana e no final de semana, com diferenças significativas que podem sugerir modelos específicos.

Caso determinado segmento tenha a presença de dois ou mais equipamentos, adotou-se a média dos valores para as bases horárias e diárias anuais. Conforme citado anteriormente, a fim de captar o efeito pendular do fluxo no Distrito Federal, para cada base (horária e diária) foram criadas outras duas, em que se considera o fluxo em um único sentido ou em ambos os sentidos de circulação. Nesse último caso, o segmento deve ter contagem em ambos os sentidos, que corresponderá à somatória das médias, caso contrário os dados do segmento não farão parte da modelagem.

3.3 Divisão dos Segmentos

Conforme [12] apud [13], o critério para determinação dos segmentos viários é fundamental em MEFA, dada a influência potencialmente significativa nos resultados. [14] apud [13] propõe duas formas para a determinação do segmento: comprimento fixo ou homogeneidade. Considere-se um segmento viário homogêneo [15] como um trecho de via sujeito a um volume de tráfego diário médio relativamente constante e que não apresenta variações significativas nas características geométricas. Dentre os critérios para definir os pontos de início e fim de um segmento tem-se [3]:

- (i) Número de vias por sentido;
- (ii) Densidade de acesso;
- (iii) Intervalos de volumes de tráfego;
- (iv) Tipo ou largura do separador central, ou ambos;
- (v) Velocidade de operação ou velocidade regulamentada;
- (vi) Uso adjacente do solo (urbano, suburbano, rural);
- (vii) Terreno (plano, ondulado, montanhoso); e
- (viii) Classificação funcional (arterial, coletora, local).

Especificamente para vias arteriais urbanas, as características consideradas na divisão de segmentos são, entre outras [3]: comprimento do segmento; TDMA; separador central; estacionamento; densidade de objetos fixos; iluminação viária; presença de controle automático de velocidade.

Sob o aspecto da velocidade em relação aos acidentes, [16] recomenda a análise por segmentos homogêneos que permitam avaliar a qualidade da via, o que justificaria maiores velocidades e menor número de acidentes para vias com melhores atributos.

As vias arteriais no Distrito Federal possuem algumas características semelhantes. Com início ou término em rodovias e com uso do solo menos denso (o que pode influenciar maiores velocidades), acabam por assumir um aumento do TDMA à medida que avançam para a área central da respectiva região administrativa. O aumento gradual de densidade dificulta o estabelecimento claro dos limites, e a partir disso considerou-se secundariamente a extensão do segmento. A descrição dos locais de acidentes também influenciou os limites a atribuir aos segmentos. É o caso de acidentes que ocorrem em logradouros com indicação limitada ao número da quadra, o qual teve de ser integralmente associado a um único segmento, sob o risco de referenciar o acidente no segmento incorreto.

Como resultado final, os comprimentos dos segmentos são similares, mas não um valor constante, o que não deixa de estar de acordo com outros estudos da especialidade. A título de síntese, o Quadro 1 apresenta as dimensões adotadas em referências consultadas.

Quadro 1. Comprimento aproximado do segmento (m)

Referência	Tipo via	Mínimo	Médio	Máximo
[6]	Rodovia	50	200	500
[9]	Rodovia	não informado	4.860	não informado
[17]	Rodovia	62	747	1.208
[18]	Urbana	não informado	450 (DP=285)	não informado
[19]	Urbana	65	856	2.830
[20]	Urbana	9,15	128	1.008
[21]	não informado	150	1.470	9.070

Em síntese, dadas a intensidade do tráfego expressa pelos dados disponibilizados por equipamentos de fiscalização e a variabilidade do uso do solo ao longo das vias arteriais, além dos demais critérios de divisão, a segmentação procurou uniformizar o comprimento dos segmentos para uma mesma via, respeitando a alocação correta dos acidentes com endereços incompletos com início e término em interseções. A diferença de comprimento dos segmentos entre vias pressupõe cuidados na adoção da variável dependente por taxa de exposição ou definição de variáveis explicativas que considerem a densidade de ocorrências por comprimento. Os segmentos variam de 600 m a 1.560 m, com comprimento médio de 1.153 m.

3.4 Compilação da Informação sobre Acidentes

Selecionadas as vias arteriais que constituem objeto de estudo, o passo seguinte passou pela imputação dos acidentes a cada um dos segmentos. A falta de georreferenciação dos acidentes dificultou substancialmente essa tarefa, pelo que deve ser apontada como uma potencial melhoria a ser implementada. Os principais problemas encontrados foram: indicação de pontos de referência que se alteram ao longo dos anos; dados incompletos; falta de indicação do sentido da via em que ocorreu o acidente; divergência entre a quadra e o sentido da via; divergência entre a quadra e o ponto de referência.

Cada um dos 6.319 acidentes que compõem a base de estudo tiveram os endereços conferidos e alocados, excluindo os acidentes não pertencentes ao local de estudo. Registaram-se 54 acidentes onde não foi possível definir com precisão o segmento onde ocorreram e outros 113 que permanecem sob análise detalhada com base no boletim de ocorrência. Esses 167 acidentes não considerados representam 2,6% do total da amostra, sendo o maior índice na via Hélio Prates (4,5%) e o menor índice na via M4 (0%).

A falta de indicação do sentido da via em que ocorreu o acidente prejudicou a alocação nos Eixos W (norte e sul) e L (norte e sul). Para essas vias, torna-se inviável considerar os segmentos separados por sentido, pelo que apenas foram integrados na base de dados agregada.

O tipo de entidade onde ocorreu o acidente foi definido com base na descrição do endereço (quando há descrição de que ocorreu em seção, rotunda ou interseção) ou em coluna específica de pesquisa dos acidentes. Essa segunda informação é gerada por descrição do boletim de ocorrência ou pela equipe de investigação da Gerência de Estatística do Detran-DF. Não havendo essa indicação, todos os acidentes foram considerados como tendo ocorrido no segmento.

A disposição dos dados desta forma permite separar os acidentes por entidade (segmento, interseção em nível semaforizadas e não semaforizadas). Assim, pode-se analisar somente os acidentes ocorridos nos segmentos, retirando todos os acidentes que envolvem interseções, ou por troços, onde se consideram conjuntamente os acidentes ocorridos nos segmentos e nas interseções, com possível perda de ajuste (uma vez que não há dados disponíveis do tráfego das vias secundárias), porém úteis para fins de comparação. Nesse último tipo de modelo, é expectável que os acidentes em ambas as entidades (segmento e interseção) sejam considerados, incluindo como variáveis explicativas a densidade e tipos de controle das interseções. Para a análise por troços, não havendo descrição exata do local, convencionou-se alocar os acidentes ocorridos na interseção para o segmento à montante, conforme o sentido do tráfego.

Destaca-se a tentativa de detalhar os dados de feridos em leves e graves (diferenciadas por tempo de internamento hospitalar inferior ou superior a 24h), porém essa informação não estava disponível nos órgãos competentes. A sua obtenção é recomendada, bem como avanços nesse sentido podem ser direcionados para a implantação da classificação MAIS [22].

3.5 Variáveis Explicativas

Enquanto os modelos simplificados utilizam somente o TDMA para estimar a frequência de acidentes, há modelos mais complexos que se focam na identificação das variáveis explicativas e seus respectivos coeficientes [6], com o objetivo, entre outros, de estimar o benefício de segurança de uma intervenção [3].

O presente trabalho procurou integrar um número significativo de variáveis, designadamente aquelas de presença frequente nas referências bibliográficas da especialidade. Algumas dessas variáveis não foram passíveis de ser recolhidas para o período do estudo (2009-2017), como é o caso da contagem de peões nas travessias pedonais, dado que 25,2% dos acidentes são atropelamentos. A visibilidade é outra variável relevante, mas difícil de medir ao longo do tempo. Por não ser o foco deste trabalho, não foram recolhidos dados do tráfego das vias secundárias das interseções.

Também os níveis de precipitação revelam-se fundamentais à estimação dos acidentes ocorridos em pavimento molhado (dados da Gerest-Detran/DF). Como variável explicativa, alguns estudos [23] consideram o número de dias chuvosos no mês e a precipitação máxima. No âmbito do presente estudo e embora tenha sido obtida a precipitação diária no Instituto Nacional de Meteorologia (Inmet), a estação meteorológica de Brasília serve como informação para todo o Distrito Federal, não discriminando cada uma das RAs. Outras informações disponibilizadas pelo Inmet são: direção e velocidade do vento, temperatura, umidade relativa e nebulosidade.

Embora os equipamentos de fiscalização do Distrito Federal discriminem o tipo de veículo, ao longo dos anos essa informação não foi inserida no sistema por falta de recursos computacionais (memória insuficiente), mas estão armazenadas em mídia digital para aproveitamento futuro.

A base de dados das operações de fiscalização por agentes de trânsito foi obtida para os anos de 2015 a 2017, estando divididas por RA, tipo de operação (bafômetro, fiscalização de documentos, escolar, carro de som etc.) e tipo de infração (alcoolemia, inabilitado, dirigir utilizando telemóvel, cinto de segurança, estacionamento etc.). A sua integração nos modelos deverá exigir a sua discriminação para cada via da amostra. A sistematização das informações de fiscalização direcionadas aos MEFA insere-se nas recomendações ao órgão de trânsito.

As características específicas do Distrito Federal para vias arteriais propiciam a inserção de outras variáveis, como: estacionamento e inversão de sentido no separador central; distância entre a via secundária até o acesso para inversão de sentido (e vice-versa) em relação ao número de vias da arterial principal; dados de alcoolemia e drogas. As variáveis de infraestrutura foram levantadas com filmagens feitas em 2017 e para os anos anteriores pelo uso de imagens de satélite [20, 24], que incluiu a ferramenta Street View. Pelo fato dos separadores centrais disporem

muitas vezes de vegetação, a presença lateral ou central da vegetação e dos postes de iluminação foram consideradas.

3.6 Velocidade

A legislação brasileira define dois tipos de equipamentos fixos de fiscalização: o controlador eletrônico de velocidade (conhecido como “pardal”), que tem por objetivo fiscalizar o limite máximo regulamentado para a via; o redutor eletrônico de velocidade (conhecido como “barreira” ou “lombada eletrônica”), destinado a fiscalizar a redução pontual de velocidade em trechos críticos, com limite diferenciado da velocidade máxima regulamentada para a via, integrando um dispositivo que mostre aos condutores a velocidade medida [25].

O Distrito Federal utiliza redutores de velocidade nos segmentos e controladores de velocidade em segmentos e interseções. Quando em interseções, dispõem de função não metrológica que fiscaliza também o avanço de sinal vermelho e de paragem sobre a passadeira.

As pesquisas disponíveis no Sig-Detran/DF não permitem calcular diretamente o desvio padrão da velocidade, estando dispostas na forma do número de veículos (exposição) por intervalo de velocidade (a cada 10 km/h) [26]. Tendo em vista que o TDMA mereceu tratamento (eliminação de valores discrepantes e reamostragem), os dados de velocidade foram transformados em percentuais, que foram multiplicados pelo tráfego tratado.

O uso de equipamentos fixos de fiscalização eletrônica posicionados de maneira ostensiva propicia dados de velocidade reais e com variação registada ao longo do dia e semana, porém com valores assumidamente inferiores aos locais sem fiscalização. Nessa linha de ação, foi realizada uma recolha de dados complementar em que se registou a velocidade real praticada, sem que os condutores se apercebessem que estavam a ser controlados. Atendendo a que a diferença entre a velocidade de projeto em rodovias e a velocidade regulamentada tende a desempenhar papel importante no número de acidentes [27], a diferença entre a velocidade operacional e a velocidade regulamentada pode fornecer informações igualmente relevantes, bem como comparações entre a velocidade operacional e a velocidade registada pelos equipamentos de fiscalização.

A recolha de dados foi realizada de maio a julho de 2017. Utilizou-se o método descrito em [28] para dimensionamento das amostras para caracterizar a velocidade operacional em fluxo livre em cada segmento. As amostras foram recolhidas em pelo menos dois horários diferentes, com o mínimo de 200 veículos por segmento, sendo medida a velocidade dos veículos para cada uma das vias de trânsito alternadamente, com o objetivo de não enviesar a amostra. O ponto de recolha foi selecionado como sendo o de maior velocidade dentro do segmento, com a menor influência possível dos diversos tipos de limitação da velocidade (interseções, passadeiras, semáforos, lombas e equipamentos de fiscalização). Todas as medições foram realizadas em regime livre de circulação e procurou-se registar, proporcionalmente à sua representatividade na corrente de tráfego, todos os tipos de veículos (motocicletas, automóveis, caminhões e autocarros).

4 RESULTADOS

Foram construídas quatro bases considerando a agregação temporal (média diária anual e média horária anual) e o sentido da via (segregado ou agregado). Cada uma dessas bases conta com as seguintes informações por segmento:

- (i) Acidentes: gravidade (fatal e ferido), hora, data, dia da semana, natureza (colisão, atropelamento, capotamento etc.), local do acidente (segmento, semáforo, passadeira, rotunda ou seção de inversão de marcha), estado do pavimento, condição do tempo, presença de álcool ou drogas, quantidade e tipo de veículos envolvidos, vítima envolvida (peão, condutor, passageiro etc.) e idade da vítima. País, estado, cidade e endereço foram os dados suficientes para uso da ferramenta MMQGIS (*software* QGIS), permitindo o georreferenciamento automático dos acidentes, com precisão suficiente de inserção no segmento correto.
- (ii) Tráfego: quantidade e percentual de veículos diário médio anual e horário médio anual total, por dia da semana e por fim de semana (sábado, domingo, feriado e pontos facultativos). Havendo dois ou mais equipamentos em um mesmo segmento, adotou-se a média de ambos.
- (iii) Velocidade: regulamentada, fiscalizada e amostra fiscalizada sem conhecimento do condutor. Quantidade e percentual de veículos por intervalo de velocidade a cada 10 km/h total, por dia da

semana e por fim de semana. Dado que a maioria das vias possui velocidade regulamentada de 60 km/h, as informações foram agregadas em três grupos de 1-49 km/h, 50-70 km/h e acima de 70 km/h. A percentagem de veículos por faixa de velocidade pode ser multiplicada pela contagem do tráfego com e sem tratamento (*box plot* e *jackknife*).

- (iv) Outras variáveis explicativas: comprimento do segmento (exposição), quantidade/largura de vias; quantidade e densidade de interseções (por tipo de controle), acessos laterais (entrada e saída), parada de autocarro (com e sem recuo), passadeiras (subterrâneas e semaforizadas ou não), lombas e equipamentos de fiscalização; existência de passeio e ciclovia (lateral ou no separador central); largura do separador central; inversão de marcha no separador central com e sem via de aceleração/desaceleração; estacionamento lateral, no separador central (com e sem segregação) ou com via paralela; distância entre a via secundária até o acesso para inversão de sentido (e vice-versa) em relação ao número de vias da arterial principal; iluminação e vegetação lateral ou central; aclive/declive; uso do solo (residencial, comércio e serviço de pequeno, médio e grande porte); precipitação (número de dias no ano). Por não utilizar segmentos de comprimento constante, algumas das variáveis são inseridas no modelo em forma de densidade (quantidade por comprimento do segmento).

5 CONCLUSÕES

O presente trabalho descreve a metodologia utilizada na construção de uma base de dados para apoio ao desenvolvimento de MEFA. A importância do TDMA justificou a aplicação de técnicas de amostragem [11] e eliminação de valores extremos, tendo por base os dados recolhidos de forma automática por recurso a equipamentos de fiscalização eletrônica. As informações relativas a acidentes e a vítimas foram agregadas e associadas aos volumes de tráfego segregado por horas e dia da semana, de forma a potenciar o desenvolvimento de modelos específicos. Foram apresentados os critérios para a divisão dos trechos estudados em segmentos viários homogêneos e as variáveis explicativas consideradas, tendo por base a bibliografia da especialidade e as especificidades associadas ao Distrito Federal. Como propostas de melhoria aplicadas à base de dados do órgão de trânsito, assinalam-se, dentre outras, a falta de coordenadas geográficas para localização fidedigna do acidente, a necessidade de acompanhamento contínuo dos feridos em acidentes, a relevância de serem integrados dados relativos aos peões, aos períodos em que os equipamentos se encontram desligados e às datas em que se registam alterações ao funcionamento da via.

6 AGRADECIMENTOS

Os autores agradecem ao Centro de Investigação do Território, Transportes e Ambiente (Citta), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e ao Departamento de Trânsito do Distrito Federal (Detran-DF).

7 REFERÊNCIAS

1. Brasil, Lei nº 9.503, de 23 de setembro de 1997. Institui o Código de Trânsito Brasileiro. *Diário Oficial da União*, Brasília, DF, 1997. http://www.planalto.gov.br/ccivil_03/leis/19503.htm
2. G. F. de M. Claude, *Previsão da Ocorrência de Acidentes de Trânsito em Interseções de Vias Arteriais Urbanas – O Caso de Taguatinga/DF*, Dissertação de Mestrado, Universidade de Brasília, Brasília, DF, 2012.
3. American Association of State and Highway Transportation Officials – AASHTO, *Highway Safety Manual*, Washington, DC, 2010.
4. R. Eenink, M. Reurings, R. Elvik, J. Cardoso, S. Wichert, & C. Stefan, *Accident Prediction Models and Road Safety Impact Assessment: Recommendations for using these tools*. 2005. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.463.9976&rep=rep1&type=pdf>
5. E. Hauer, *The art of regression modeling in road safety*. Springer, Cham, 2015.

6. J. Ambros, & J. Sedoník, A feasibility study for developing a transferable accident prediction model for Czech regions, *Transportation Research Procedia*, 14, 2054-2063, 2016.
7. K. C. R. Silva, *Aplicação do modelo de previsão de acidentes do HSM em rodovias de pista simples do estado de São Paulo*, Dissertação de Mestrado, Universidade de São Paulo, São Paulo, 2011.
8. M. M. Imprialou, M. Quddus, D. E. Pitfield, & D. Lord, Re-visiting crash-speed relationships: A new perspective in crash modelling, *Accident Analysis & Prevention*, 86, 173-185, 2016.
9. M. Quddus, Exploring the relationship between average speed, speed variation, and accident rates using spatial statistical models and GIS, *Journal of Transportation Safety & Security*, 5(1), 27-45, 2013.
10. J. O. Costa, M. A. P. Jacques, E. F. Freitas, & P. A. A. Pereira, Acidentes rodoviários do tipo colisões nas estradas nacionais na Região do Minho: modelos de previsão para segmentos rodoviários, 2010. http://repositorium.sdum.uminho.pt/bitstream/1822/26158/1/Artigo_7CRP_26.pdf
11. G. G. Mota, *Previsão, por meio de técnicas de séries temporais e imputação de dados, do volume total de veículos em interseções de Taguatinga para o ano de 2012*, Monografia de Bacharelado, Universidade de Brasília, Brasília, DF, 2012.
12. C. T. Nodari, *Método de avaliação da segurança potencial de segmentos rodoviários rurais de pistas simples*, Tese de Doutorado, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2003.
13. G. H. Boffo, *Formatos e técnicas de modelos de previsão de acidentes de trânsito*, Dissertação de Mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011.
14. S. P. Miaou, P. S. Hu, T. Wright, S. C. Davis, & A. K. Rathi, *Development of relationships between truck accidents and highway geometric design. Phase I*, Washington, DC, 1991.
15. National Cooperative Highway Research Program – NCHRP, *User's guide to develop highway safety manual safety performance function calibration factors*, 2014. [http://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP20-07\(332\)_FinalGuide.pdf](http://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP20-07(332)_FinalGuide.pdf)
16. M. C. Taylor, A. Baruya, & J. V. Kennedy, The relationship between speed and accidents on rural single-carriageway roads, *TRL Report*, 511, 2002.
17. J. Aguero-Valverde, K. Wu, & E. T. Donnell, A multivariate spatial crash frequency model for identifying sites with promise based on crash types, *Accident Analysis and Prevention*, 87, 8-16, 2016.
18. P. Greibe, Accident prediction models for urban roads, *Accident Analysis and Prevention*, 35, 273-285, 2003.
19. Q. Zeng, H. Wen, H. Huang, & M. Abdel-Aty, A Bayesian spatial random parameters Tobit model for analyzing crash rates on roadway segments, *Accident Analysis and Prevention*, 100, 37-43, 2017.
20. C. A. Torres, *Avaliação do desempenho da segurança viária em redes urbanas de transportes com utilização de modelos de previsão de acidentes*, Dissertação de Mestrado, Universidade Federal do Ceará, Fortaleza, 2016.
21. Q. Zeng, H. Huang, X. Pei, & S. C. Wong, Modeling nonlinear relationship between crash frequency by severity and contributing factors by neural networks. *Anal. Methods Accid. Res*, 10, 12-25, 2016.
22. S. Ferreira, A. Couto, J. P. Tavares, & C. Rodrigues, O paradigma da classificação MAIS na segurança viária, *8º Congresso Rodoviário Português*, Lisboa, 2016.
23. A. Theofilatos, & G. Yannis, A review of the effect of traffic and weather characteristics on road safety, *Accident Analysis and Prevention*, 72, 244-256, 2014.
24. S. A. Gargoum, & K. El-Basyouny, Exploring the association between speed and safety: A path analysis approach, *Accident Analysis and Prevention*, 93, 32-40, 2016.
25. Conselho Nacional de Trânsito – CONTRAN, Resolução Contran nº 396, de 13 de dezembro de 2011. Dispõe sobre requisitos técnicos mínimos para a fiscalização da velocidade de veículos automotores, reboques e semirreboques, conforme o Código de Trânsito Brasileiro. *Diário Oficial da União*, Brasília, DF, 2011. <https://www.legisweb.com.br/legislacao/?id=114859>

26. S. Heydari, L. F. Miranda-Moreno, & F. Liping, Speed limit reduction in urban areas: A before-after study using Bayesian generalized mixed linear models, *Accident Analysis and Prevention*, 73, 252-261, 2014.
27. L. Aarts, & I. Van Schagen, Driving speed and the risk of road crashes: A review, *Accident Analysis and Prevention*, 38(2), 215-224, 2006.
28. A. B. Gonçalves, *Estudo da velocidade operacional dos veículos em trechos viários de rodovias rurais de pista simples*, Dissertação de Mestrado, Universidade de Brasília, Brasília, DF, 2011.