

O USO DE DADOS GEOESPACIAIS EM MODELOS DE REGRESSÃO LOGÍSTICA NA DETERMINAÇÃO DE ACIDENTES POR ATROPELAMENTO EM ÁREAS URBANAS

Lucas Israel Damasceno Silva¹, Andersonn Magalhães de Oliveira², José Gabriel Vieira Santos³, Erison Rosa de Oliveira Barros⁴

^{1,3} Departamento de Cartografia, discente, Universidade Federal de Pernambuco, Brasil

² Pós-Graduação em Ciências Geodésicas e Tecnologias da Geoinformação, discente, UFPE, Brasil

⁴ Departamento de Cartografia, docente, Universidade Federal de Pernambuco, Brasil

CT06 - Sistemas de Informações Geográficas e Infraestrutura de Dados Espaciais

RESUMO

O propósito do presente estudo é de se elaborar um modelo de regressão logística, cujo vetor resposta de probabilidade possa indicar a ocorrência de acidentes por atropelamento na área urbana da cidade do Recife - Pernambuco, para as áreas em que a BR 101 secciona. A proposta é de se utilizar de dados geoespaciais, tanto os referentes à ocorrência dos acidentes dessa natureza, como também para se relacionar informações socio-econômicas locais com a probabilidade de ocorrência desse tipo de evento. Com o produto espacial gerado a partir desse modelo, o que se espera é que a análise de áreas críticas se dê de forma mais eficiente, tanto para a implantação da infraestrutura adequada de segurança aos pedestre, como para ações de conscientização do trânsito.

Palavras chave: Regressão Logística, Acidentes, Dados Geoespaciais

ABSTRACT

The purpose of the present study is to elaborate a logistic regression model, whose probability response vector may indicate the occurrence of road accidents in the urban area of the city of Recife - Pernambuco, for the areas in which the BR 101 section. The proposal is to use geospatial data, both those related to the occurrence of accidents of this nature, as well as to relate local socio-economic information with the probability of occurrence of this type of event. With the spatial product generated from this model, what is expected is that the analysis of critical areas is given more efficiently, both for the implementation of adequate infrastructure for pedestrian safety and for actions to raise awareness of traffic.

Keywords: Logistic Regression, Accidents, Geospatial Data

1- INTRODUÇÃO

O espaço faz parte da nossa vida cotidiana, pois as pessoas e o ambiente, tanto o natural quanto o construído, ocupam um lugar na Terra, ou seja, tudo está em algum lugar. Poderíamos dizer de uma maneira simplista que o espaço nada mais é do que a superfície do nosso planeta. Superfície esta que, apesar das dimensões, a cada dia que passa se torna menos relevante se considerarmos o avanço das tecnologias de transporte e comunicação (MOUW, 2000). Mas não podemos concluir que o espaço é importante apenas porque tudo acontece em algum lugar, mas porque saber onde as coisas acontecem é crucial para o

entendimento de como e porque elas acontecem (WARF; ARIAS, 2009).

Muitas teorias sociais relacionam o comportamento individual ou de um grupo familiar com o contexto da localização espacial deste indivíduo ou grupo familiar (RINDFUSS; STERN, 1998), ou seja, procuram explicar esse comportamento tendo em conta um contexto social e espacial (ENTWISLE, 2007). Seguindo este raciocínio, para termos uma visão completa de um fenômeno precisamos compreender a relação entre a ocorrência deste fenômeno em uma determinada região e as diversas dimensões da realidade que se apresentam naquele espaço, desde os

aspectos físicos e ambientais, até os aspectos humanos e sociais.

Diante disso, e segundo o relatório da Organização das Nações Unidas sobre estado global da segurança viária (WHO, 2009), as mortes por acidentes de trânsito são a maior causa de morte na população entre 15 e 29 anos e serão a 5ª maior causa de morte no mundo considerando todas as idades em 2030, sendo que o Brasil consta como o 5º país com maior quantidade de acidentes de trânsito no mundo. Por outro lado, os custos envolvidos com os acidentes também revelam valores expressivos. Segundo o relatório sobre impactos sociais e econômicos dos acidentes de trânsito nas rodovias brasileiras do Instituto de Pesquisa Econômica Aplicada – (IPEA, 2006), o custo dos acidentes nas rodovias brasileiras atingiu cerca de 1,2% do Produto Interno Bruto (PIB) brasileiro em 2005. No relatório do mesmo Instituto, elaborado em 2003, sobre impactos sociais e econômicos dos acidentes de trânsito nas aglomerações urbanas (IPEA, 2003), o custo dos acidentes nas áreas urbanas atingiu um valor que correspondia a aproximadamente 0,5% do PIB brasileiro em 2001. Tendo em vista tal realidade, a Organização das Nações Unidas elegeu o período de 2011-2020 como sendo a Década de Ação pelo Trânsito Seguro, na qual governos de todo o mundo se comprometem a tomar novas medidas para prevenir os acidentes no trânsito. No Brasil, o governo brasileiro vem realizando diversas ações, dentre as quais lançou o Projeto Vida no Trânsito, onde se busca, dentre outros objetivos, o de identificar fatores de riscos e grupos de vítimas mais vulneráveis a acidentes de transportes terrestres.

Então a partir do problema apresentado, procurou-se propor uma metodologia de análise dos acidentes de trânsito baseado no georreferenciamento dos acidentes e nas especificidades dos dados geográficos, empregados. Aplicando estas variáveis agregadas em áreas e relacionadas à geometria e conectividade das vias, às características demográficas, socioeconômicas e de uso do solo. O conhecimento das relações entre as pessoas e seus ambientes físicos e sociais possibilitou novas visões e entendimentos dessas complexas interações, levando a grandes avanços científicos que, provavelmente, vão continuar acontecendo ainda por muito tempo. Diante disso, será visto em detalhes na revisão bibliográfica, o uso de dados georreferenciados de acidentes em trechos urbanos, segurança viária, travessias urbanas, perfil socioeconômico dos locais onde ocorreram os acidentes, conceitos de agregação, Censo Demográfico e regressão logística.

2- REVISÃO BIBLIOGRÁFICA

Segurança viária pode ser definida como uma “medida” do comportamento do sistema viário, em relação à unidade de viagens, distância percorrida no

sistema viário ou veículos registrados (ITE, 1995; apud Cardoso, 1999). Dessa forma, se pode reduzir a maioria dos acidentes de tráfego se a segurança viária for tratada de forma eficiente em todas as etapas de uma via, ou seja, desde o planejamento de sua construção até durante sua utilização. (Hildebrand & Wilson, 1999; apud Nodari & Lindau, 2001).

Com isso, é necessário que projetos de via que possam evitar mudanças repentinas em seu trajeto e uma sinalização adequada sejam estabelecidos, permitindo uma boa distância para visualização e reação do usuário, com o objetivo de reduzir o risco de erros e permitir escolhas corretas do usuário durante a condução do veículo (Roess & McShane, 1990; apud Cardoso, 2006). Assim, visando diminuir as chances de ocorrência de acidentes no trânsito, empresas responsáveis pela construção, manutenção e operação de rodovias, passam a utilizar, um pouco mais das ASV's (Auditoria da Segurança Viária), a qual fiscaliza periodicamente e faz uma avaliação da probabilidade de acontecer um acidente em determinada via de transporte (Bornsztein, 2001; apud Nodari & Lindau, 2001). Entretanto, como não é possível catalogar todas as possíveis causas dos acidentes em uma tabela de padrões, é impossível afirmar que seguir determinados padrões de segurança viária resulte em garantia total de ambientes viários 100% seguros (Morgan, 1999; apud Nodari & Lindau, 2001).

As causas dos acidentes de trânsito geralmente são divididas em três grupos: humanos, viário-ambientais e veiculares. Alguns autores consideram um quarto fator, conhecido como “fator institucional/social”. Esta causa está associada a falta de regulamentação e fiscalização, como a precariedade da sinalização do trânsito, das condições de trabalho, ausência de fiscalização, e etc. (Gold, 1998; apud Cardoso, 2006). Dos quatro fatores citados anteriormente, a causa humana é a maior responsável pelos acidentes. No entanto, boa parte desses acontecimentos, são resultado da interação das causas humanas com o ambiente viário e uma parcela um pouco menor da relação entre os seres humanos e o fator veicular. (Austroads, 1994; apud Cardoso, 2006). Também deve ser considerado que essa análise sobre as causas dos acidentes de tráfego é um procedimento bastante complexo, visto que ocorrem de forma dependente, ou seja, sempre relacionado com outro fator, e em grande número. Dessa forma, segundo o Department for Transport, Local Government and the Regions of London (2001) a melhor forma para se entender o que influencia os acidentes é analisar os dados a partir do cadastro de registros de dados (Cardoso, 2006).

Os dados de acidentes registrados são os principais recursos para se avaliar as causas dos acidentes no trânsito e procurar corrigi-los, melhorando a segurança do sistema viário (Ibrahim e Silcock, 1992; apud Cardoso, 1999). Assim, para que se obtenha

sucesso na análise dos dados sobre acidentes é necessário que os dados primários possuam boa qualidade de informações. Também é necessário um detalhado conhecimento das variadas formas de apresentação, ordenação e armazenamento, pelo fato de diferentes tipos de usuários necessitarem para diversos objetivos (Baginski, 1995; apud Cardoso, 1999). No entanto, a indisponibilidade de um banco de dados com padrões específicos na metodologia de coleta e armazenamento de dados torna difícil e até impossível de se trabalhar com a identificação e análises na prevenção dos acidentes. A tarefa de avaliar a eficiência de medidas de contenção de acidentes de trânsito (Vieira et al, 1995; apud Cardoso, 1999). Visto isso, é fundamental que os órgãos responsáveis com o gerenciamento e segurança no trânsito, estabeleçam uma integração, como por exemplo, com a base de dados da rede hospitalar (Vieira et al, 1995; apud Cardoso, 1999).

Podemos perceber que os assuntos Geoespaciais tomaram proporções inimagináveis na atualidade, o que levou Taylor (2010) a dizer que estamos testemunhando o surgimento da Era da Localização. Apesar do interesse pelo espaço nos estudos analíticos ter se iniciado há mais de um século, esta nova era parece coincidir com a proliferação do uso de aparelhos com sistemas de localização (TAYLOR, 2010), como automóveis e telefones celulares, bem como com a disponibilização de serviços de mapas na internet (RUMSEY, 2009), como o Google Earth e o Google Maps, que surgiram em 2005.

Espacializar o fenômeno em estudo abre espaço para o emprego de diversas técnicas de visualização dos dados e de análise espacial, as quais têm potencial de trazer uma visão mais rica sobre o comportamento dos acidentes que não somente a oriunda da interpretação de modelos e testes estatísticos. Desta maneira, a visualização de mapas mostrando variáveis demográficas se tornou uma grande ferramenta descritiva, capaz de revelar padrões impossíveis de serem percebidos em tabelas.

O objetivo da pesquisa é de se utilizar dos recursos da geoinformação associado à metodologia de inferência estatística. Para isso, a hipótese levantada é de relacionar os dados dos acidentes de trânsito no trecho urbano da BR 101 do município do Recife-PE, com as informações socioeconômicas de cada setor censitário de cada trecho da região de estudo, bem como outros aspectos informados pela Polícia Rodoviária Federal relacionados ao registro da ocorrência, a exemplo de informações temporais e das características do evento. Para verificar a probabilidade associada a esse tipo de acidente, se utiliza da variável dependente na forma de saída dicotômica, no entanto, em análises de regressão clássica linear, se verifica a inadequabilidade desse tipo de saída para o modelo linear, devido a incompatibilidade ao atendimento dos pressupostos básicos do modelo de regressão, a

exemplo da esperança nula, normalidade, homoscedasticidade, não autocorrelação e aleatoriedade. (PARDOE, 2006).

Portanto, para se obter uma correlação mais forte entre os vetores da variável dependente ajustado versus calculado, bem como se ajustar um modelo que atenda aos pressupostos básicos, o método adotado é de regressão logística, no caso, uma regressão não-linear, onde se possibilita se verificar a probabilidade associada ao evento de acidente por atropelamento com as variáveis explicativas do comportamento deste fenômeno.

Com as informações obtidas do modelo gerado, agregadas aos trechos da rodovia, se possibilita a geração de produtos cartográficos que possibilitem a visualização e interpretação do fenômeno estudado. Portanto, um produto útil a tomada de decisão por parte das autoridades competentes no sentido de desenvolver políticas públicas e de intervenções físicas, a exemplo de implantação de passarelas, para minimizar os prejuízos econômico-sociais relativos aos acidentes de trânsito.

3- METODOLOGIA E DESENVOLVIMENTO

Para a fase de coleta das informações necessárias para o estudo, as informações relativas aos acidentes foram obtidas através do sistema VGEO do DNIT - Departamento Nacional de Infraestrutura (2017), e também de dados de setores censitários, IDH dos bairros, renda média mensal dos chefes de família para o bairro, população residente, quantidade de famílias obtidos no Censo Demográfico 2010 (2017).

A proposta ao se utilizar de um modelo logístico, é de se utilizar de informações de um conjunto de variáveis independentes para prever a ocorrência de acidente por atropelamento, o que esta seria a variável dependente. O que se espera é que todas as variáveis que estejam consideradas no modelo tenham relação entre si, na forma de um problema multivariável. A associação calculada pelo modelo logístico é obtida por meio da comparação de eventos que diferem apenas na característica de ocorrência para o ajuste estatístico. Para a situação, se torna aplicável para o estudo de ocorrência de atropelamento visto que a função que varia pode assumir valores de 0 (zero) ou 1 (um), neste a ocorrência e naquela a não-ocorrência. Pelo fato do modelo linear não se ajustar a uma função S, o modelo logístico adotado de resposta probabilística é dado pela Equação (1), onde os parâmetros deste modelo deverão ser estimados a partir da amostra obtida pelo método da máxima verossimilhança por métodos iterativos.

$$y = \frac{1}{b_0 + x_1 \cdot b_1 + x_2 \cdot b_2 + \dots + x_n \cdot b_n} \quad (1)$$

Entretanto, após as primeiras simulações com a utilização de todas as variáveis citadas, e com os indicadores dos testes de ponderação, as variáveis escolhidas para o modelo foram Sentido da Via, Renda média das famílias para o setor censitário da posição, Número de domicílios para o setor censitário da posição, Hora do dia, sendo este último utilizado em uma forma senoidal para ajustar das oscilações dos horários de maior tráfego das vias. Sendo que para todas as variáveis a significância individual dos parâmetros dado pelo teste bicaudal, a probabilidade foi a nível de 20%. Ainda para os testes de Omnibus de coeficientes do modelo, se obteve uma significância do modelo a um nível de 0,5%, e a sumarização do modelo com uma estimação finalizada no número de seis iterações, e com o R quadrado Nageikerke a

48,5%, que embora minimizado, é esperado para modelos probabilísticos lineares.

Após a adoção do modelo, foi realizada a integração com o Sistema de Informação Geográfica para a concepção do produto descrito pela Figura 1 e a Figura 2. Através do produto gerado, se pode verificar visualmente que há uma preponderância do aparecimento da probabilidade de ocorrência de acidente por atropelamento no setor censitário do bairro de Jardim São Paulo, à Sul, tanto para os sentidos crescente e decrescente, bem como para os variados horários do dia. Dessa forma, a metodologia proposta se coloca como uma importante diretriz para estudos de verificação de implantação de travessias urbanas do tipo passarela, para as regiões que apresentem índices mais críticos. Ou inclusive, para ações de conscientização do trânsito, direcionado à populações mais vulneráveis aos atropelamentos.

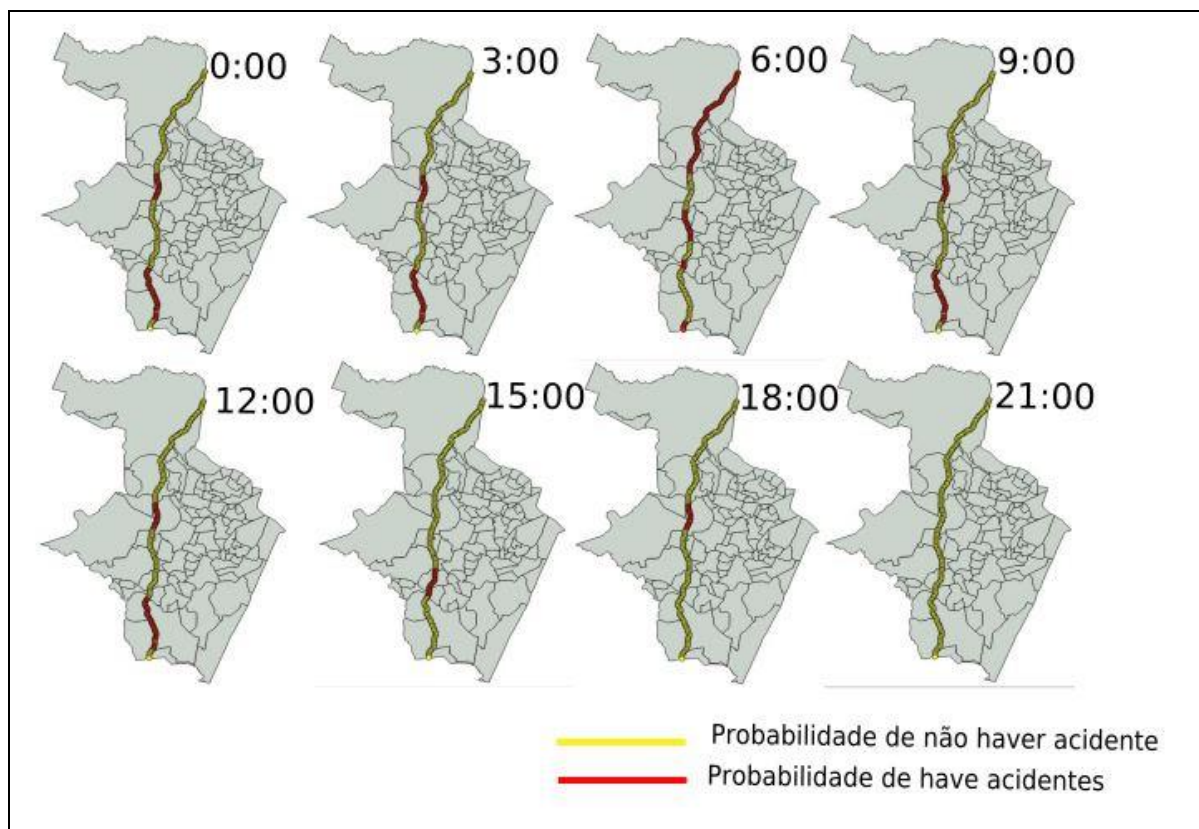


Fig. 1 – Representação das probabilidades de ocorrência de acidente por atropelamento no trecho da BR 101 (Recife), para o sentido crescente (Norte à Sul), para intervalos de três horas, dado pelo modelo logístico.

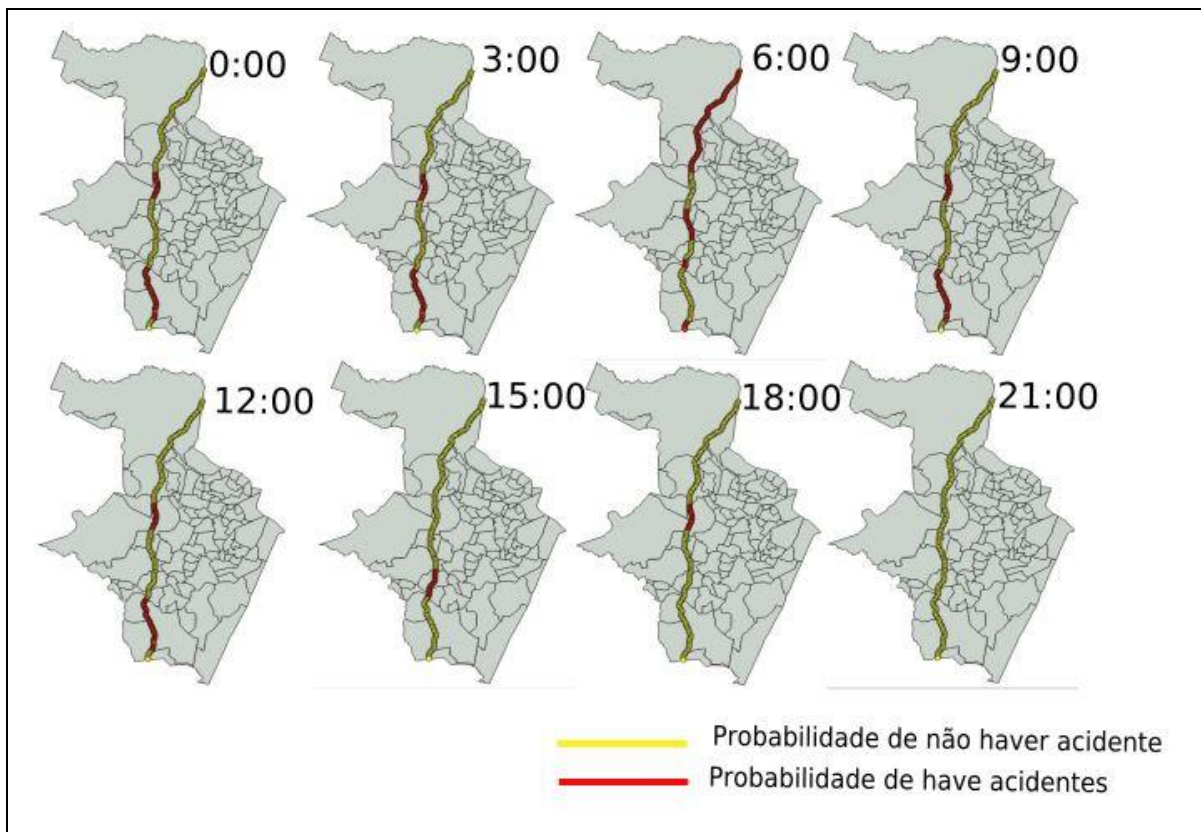


Fig. 2 – Representação das probabilidades de ocorrência de acidente por atropelamento no trecho da BR 101 (Recife – Pernambuco), para o sentido decrescente (Sul à Norte), para intervalos de três horas, dado pelo modelo logístico.

REFERÊNCIAS BIBLIOGRÁFICAS

ARIAS, Santa. Rethinking space: an outsider's view of the spatial turn. *GeoJournal*, vol. 75, n.1, p. 29-41, 2010.

Cardoso, G; Modelos para previsão de acidentes de trânsito em vias arteriais urbanas Tese (Doutorado em Engenharia de produção, sistemas de transportes); Programa de Pós-Graduação em Engenharia de Produção; Universidade Federal do Rio Grande do Sul; 2006.

Cardoso, G; Utilização de um sistema de informações geográficas visando o gerenciamento da segurança viária no município de São José-SC; Dissertação (Mestrado em Engenharia Civil); Programa de Pós-Graduação em Engenharia de Civil; Universidade Federal de Santa Catarina; 1999.

CENSO DEMOGRÁFICO 2010. Características da População e dos domicílios: resultados do universo. Rio de Janeiro: IBGE, 2011. Acompanha 1 CD-ROM. Disponível em: <http://www.ibge.gov.br/home/estatistica/populacao/censo2010/caracteristicas_da_populacao/resultados_do_universo.pdf>. Acesso em: julho. 2017.

DEPARTAMENTO NACIONAL DE INFRAESTRUTURA (DNIT). VGEO. Disponível em

<http://servicos.dnit.gov.br/vgeo/#> Acesso em: julho de 2017.

MOUW, Ted. Visions of The Future: New Directions in Population Research. In: *Population Association of Annual Meeting*. Los Angeles, USA, 2000.

Nodari, C. T; Lindau, L. A; *Auditoria da segurança viária*; Universidade Federal do Rio Grande do Sul; 2001.

RINDFUSS, Ronald R.; STERN, Paul C. Linking Remote Sensing and Social Science: The Need and the Challenges. In: LIVERMAN, Diana; MORAN, Emilio F.; RINDFUSS, Ronald R.; STERN, Paul C. (Ed.). *People and Pixels: Linking Remote Sensing and Social Science*. Washington, D.C.: National Academy Press, 1998

WARF, Barney; ARIAS, Santa. Introduction: the reinsertion of space into the social sciences and humanities. In: WARF, Barney; ARIAS, Santa (Ed.). *The Spatial Turn: Interdisciplinary Perspectives*. Londres e Nova York: Routledge, 2009.

WARF, Barney; ARIAS, Santa. Introduction: the reinsertion of space into the social sciences and humanities. In: WARF, Barney; ARIAS, Santa (Ed.). *The Spatial Turn: Interdisciplinary Perspectives*. Londres e Nova York: Routledge, 2009.