

ANÁLISE COMPARATIVA DE TÉCNICAS DE MODELAGEM DA GERAÇÃO DE VIAGENS DE CARGA

Leise Kelli de Oliveira

Ana Sara Spindola da Silva

André da Cunha Melo Tolentino

Arthur Henrique Farias de Souza Rachid

Universidade Federal de Minas Gerais

Bruno Vieira Bertoncini

Universidade Federal do Ceará

Jessica Helena de Lima

Universidade Federal de Alagoas

Lílian dos Santos Fontes Pereira Bracarense

Universidade Federal do Tocantins

Renata Lúcia Magalhães de Oliveira

Centro Federal de Educação Tecnológica de Minas Gerais

RESUMO

Neste é artigo apresentada uma análise a partir dos resultados de duas técnicas de modelagem da geração de viagens de carga (regressão linear e regressão linear generalizado), para os setores de alimentos e bebidas em quatro cidades brasileiras. O número médio de entregas por semana, o número de funcionários e a área do estabelecimento foram coletados diretamente em estabelecimentos comerciais de Belo Horizonte (MG), Quixadá (CE), Palmas (TO) e Recife (PE). Os resultados indicam que os modelos que empregam como variável independente o número de funcionários foram os que apresentaram menores erros quadráticos médios. Sobre a técnica de modelagem, apesar do modelo linear generalizado ser adequado para modelar dados de contagem, este não apresentou melhores resultados que o modelo linear. Por fim, modelos locais são imprescindíveis para análise dos impactos da geração de viagens de carga.

ABSTRACT

We present in this paper the results of two modeling techniques for the freight trip generation (linear regression and generalized linear regression) for the food and beverage sectors in four Brazilian cities. The average number of deliveries per week, the number of employees and the establishment area were collected directly from commercial establishments in Belo Horizonte (MG), Quixadá (CE), Palmas (TO) and Recife (PE). The results indicate that the models that use the number of employees as independent variable presented smaller root mean square error. Regarding the modeling technique, although the generalized linear model was adequate to model the counting data, it did not present better results than the linear model. Finally, local models are essential for analyzing the impacts of the freight trips generation in urban areas.

1. INTRODUÇÃO

O fenômeno gerador da demanda por transporte de carga é muito diferente e mais complexo do que o observado quando se analisa o transporte de passageiros, visto que é o resultado de interações econômicas entre múltiplos agentes que participam da cadeia de suprimentos, bem como do governo que regula vários aspectos do sistema (Holguín-Veras e González-Calderón, 2013). Assim, estimar a quantidade de viagens de veículos de carga geradas por uma zona ou um estabelecimento comercial é um dos mais importantes processos da modelagem da demanda por transporte de carga (Jaller *et al.*, 2014; Lim *et al.*, 2017; Kaszubowski, 2018). Nesses casos, os modelos de geração de viagens são importantes ferramentas analíticas para estimar as trocas envolvendo estabelecimentos comerciais urbanos (Alho e Silva, 2014) e respaldam a avaliação de políticas de curto prazo (Comi *et al.*, 2012).

O desenvolvimento de modelos de geração exige uma sequência de etapas desde a coleta de dados até as análises estatísticas de validação das equações. Ao longo desse desenvolvimento

é necessário escolher a abordagem do modelo em relação à variável dependente (número de veículos, volume de carga ou número de entregas), eleger as variáveis independentes, definir o tipo de coleta de dados (dados primários ou secundários) e os métodos empregados para obtê-los, e por fim a técnica de modelagem a ser utilizada. As escolhas feitas em cada uma dessas etapas podem influenciar nos resultados, sendo necessário um estudo prévio para embasar as decisões e contribuir para sejam as mais adequadas à realidade que se deseja analisar.

Face a essa problemática, busca-se, com este trabalho, comparar duas técnicas de modelagem de geração de viagens no transporte de cargas, regressão linear e regressão linear generalizada, em estabelecimentos comerciais que comercializam alimentos e bebidas, para a realidade de quatro cidades brasileiras, Belo Horizonte, Quixadá, Palmas e Recife. Para tanto, a seção 2 busca apresentar e discutir uma revisão sistemática da literatura sobre a modelagem da geração de viagens de carga. Na seção 3 tem-se uma proposta de método para construção dos modelos, com a consequente apresentação e discussão dos resultados obtidos ocorrendo na seção 4. Por fim, as considerações finais decorrentes do presente estudo são apresentadas na seção 5.

2. A MODELAGEM DA GERAÇÃO DE VIAGENS NO TRANSPORTE DE CARGA

Uma revisão sistemática da literatura foi conduzida para identificar as técnicas de coleta de dados, de modelagem da geração de viagens de carga, e as variáveis comumente utilizadas. O Google acadêmico foi a base de dados utilizada e buscaram-se artigos utilizando as palavras-chaves “*freight trip generation*” e “*modelling*”. Foram identificados 291 artigos e, através da análise do título e resumo, selecionados 106 artigos. Importante ressaltar que 80 registros não foram considerados por se tratar de relatório e/ou apresentação. Após a leitura de cada artigo, 39 referências foram consideradas nessa revisão. Na Figura 1 é apresentada a evolução temporal das publicações sobre modelagem da geração de viagens no transporte de carga, com destaque para o ano de 2017, com 8 artigos publicados, sendo a maior quantidade nos últimos anos. Os artigos foram publicados no *TRB Annual Meeting* (21%), *Transportation Research Record* (15%), *Transportation* (5%), *Journal of Transport Geography* (5%), *Procedia Social and Behavioral Sciences* (5%), e *Transportes* (5%). Os principais autores são J. Holguín-Veras (10%), A. R. Alho (10%), M. Jaller (8%), I. Sánchez-Días (8%).

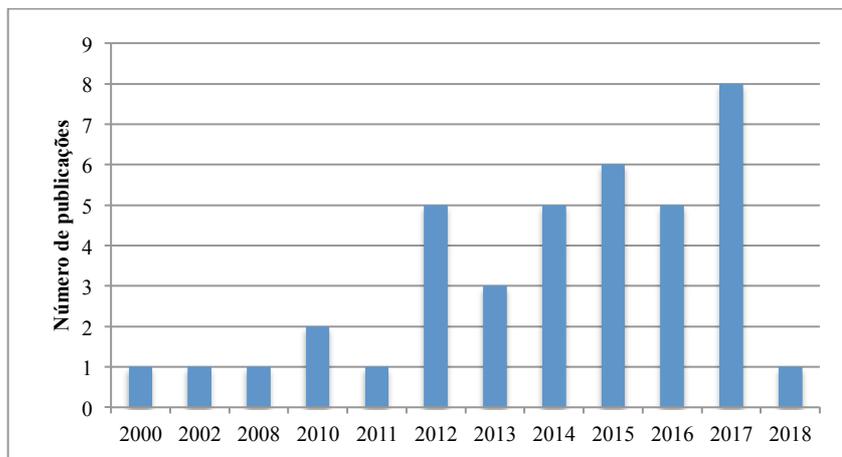


Figura 1: Evolução das publicações referentes à modelagem da geração de viagens de carga.

A modelagem da geração de viagens no transporte de carga pode ser baseada em veículo,

mercadoria e entrega (Comi *et al.*, 2012; Kaszubowski, 2018). Os modelos baseados em veículos têm como unidade de referência a viagem de veículo de carga e permite estimar o número de veículos atraídos ou produzidos por zonas individuais, possibilitando fornecer, em etapa subsequente, a distribuição espacial (matrizes O-D) dos fluxos estimados anteriormente, em etapa subsequente (Comi *et al.*, 2012). Uma vantagem dos modelos baseados em veículos é a utilização de dados de volume de tráfego, que são facilmente coletados com baixo custo (Comi *et al.*, 2012; Cantilho *et al.*, 2012). Os modelos baseados em mercadorias consideram a quantidade de mercadoria (toneladas, peso, volume, carga) como unidade de referência e permitem estimar os fluxos de carga a partir de dados socioeconômicos (Cantilho *et al.*, 2012). Segundo Doustmohammadi *et al.* (2016), os modelos baseados em veículos não modelam os comportamentos econômicos dos quais derivam a demanda, enquanto os modelos baseados em mercadorias deixam de ser realisticamente responsáveis pelas atividades dos veículos, especialmente em ambientes urbanos, em que a avaliação de impacto é crucial. Por fim, os modelos baseados na entrega se concentram nas coletas e entregas de mercadorias realizadas nos estabelecimentos. O uso da entrega como unidade de referência permite relacionar diretamente geradores e operadores de transporte, por meio do uso da mesma unidade de referência. Os modelos propostos dentro da abordagem baseada na entrega consistem em uma sequência de modelos estatístico-descritivos (Comi *et al.*, 2012).

A obtenção dos dados referentes à movimentação do transporte urbano de carga, costumeiramente, ocorre por quatro tipos de pesquisa (Allen *et al.*, 2014; Aditjandra *et al.*, 2016): (i) **gerais** baseados no fluxo de veículos ou volume de mercadorias; (ii) baseado nos **stakeholders**, por meio de entrevistas e/ou questionários com transportadores/embarcadores/varejistas; (iii) baseado nos **veículos**, usando dados de movimentação de veículos; (iv) em uma **área específica**, por meio de uma abordagem geográfica. Esses dados são utilizados para obter conhecimento sobre as operações da carga e, para isto, os dados são classificados em viagens de veículos, fluxos de mercadorias e atividade do veículo (González-Feliu e Routhier, 2012; Aditjandra *et al.*, 2016). Contudo, os modelos dependem da quantidade e qualidade dos dados e a construção de categorias pode ter impacto na qualidade do modelo (González-Feliu *et al.*, 2015; Günay *et al.*, 2016).

No Quadro 1 são apresentadas as pesquisas utilizadas para obtenção de dados para modelagem da geração de viagens de carga, sendo que a pesquisa em estabelecimentos comerciais é mais comumente utilizada. Allen *et al.* (2012) e Kaszubowski (2018) afirmam que a pesquisa em estabelecimentos comerciais é a técnica comumente utilizada para obter dados, geralmente realizada através de pesquisas face-a-face, por telefone ou questionário. Essa técnica tem como pontos positivos o fornecimento de dados de fluxo de mercadorias e atividade do veículo e vincula o fluxo de mercadorias e a atividade dos veículos ao setor econômico e ao uso do solo. Como pontos negativos, Allen *et al.* (2012) e Kaszubowski (2018) destacam que os entrevistados geralmente não têm muito conhecimento sobre o tipo de veículo de mercadorias, local de carga/descarga e horários. A pesquisa depende do conhecimento dos entrevistados, sendo que as taxas de resposta podem ser afetadas se a pesquisa for muito longa. Esse tipo de pesquisa não fornece informações sobre a rota dos veículos e é uma pesquisa que pode ser relativamente cara. Para conhecer detalhes sobre as técnicas de coleta de dados para modelos de geração de viagens, consultar Allen *et al.* (2012), Holguín-Veras e Jaller (2012) e Tavasszy e Jong (2014).

Quadro 1: Técnicas para modelagem da geração de viagens de carga e respectivas referências

Fonte dos dados	Referências
Contagens de tráfego	Aditjandra <i>et al.</i> (2016), Günay <i>et al.</i> (2016), Holguín-Veras <i>et al.</i> (2013)
Dados de planos diretores	Günay <i>et al.</i> (2016)
Matriz origem destino de carga	Kulpa (2014)
Pesquisa de fluxo de carga	Jaller <i>et al.</i> (2014, 2015a), Lim <i>et al.</i> (2017)
Pesquisa baseada em rotas	Ambrosini <i>et al.</i> (2010)
Pesquisa com empresas de transporte	Campbell <i>et al.</i> (2012), Rudani <i>et al.</i> (2017)
Pesquisa em estabelecimentos comerciais	Ambrosini <i>et al.</i> (2010), Campbell <i>et al.</i> (2012), Holguín-Veras <i>et al.</i> (2013), Alho e Silva (2014a, b, 2015, 2017), Priya <i>et al.</i> (2015), Oliveira <i>et al.</i> (2016, 2017), Sánchez-Días (2017), Kaszubowski (2018)

No Quadro 2 são apresentadas as variáveis comumente utilizadas nos modelos de geração de viagens de carga. O número de empregados, que caracteriza o tamanho da empresa, é uma variável comumente utilizada nos modelos pela maioria dos estudos. A área do estabelecimento tem sido também utilizada, recentemente, como uma variável explicativa.

Quadro 2: Variáveis independentes utilizadas nos modelos de geração de viagens.

Variável	Referências
Área do estabelecimento	Iding <i>et al.</i> (2002), Jaller <i>et al.</i> (2014), Priya <i>et al.</i> (2015), Aditjandra <i>et al.</i> (2016), Oliveira <i>et al.</i> (2016, 2017), Alho e Silva (2017), Rudani <i>et al.</i> (2017), Sánchez-Días (2017)
Número de empregados	Iding <i>et al.</i> (2002), Holguín-Veras <i>et al.</i> (2011; 2013), Campbell <i>et al.</i> (2012), Lawson <i>et al.</i> (2012), Alho e Silva (2014a, b, 2015, 2017), Gonzalez-Feliu <i>et al.</i> (2015), Jaller <i>et al.</i> (2015a, b), Priya <i>et al.</i> (2015), Ferreira e Silva (2016), Günay <i>et al.</i> (2016), Oliveira <i>et al.</i> (2016, 2017) Boarnet <i>et al.</i> (2017), Lim <i>et al.</i> (2017), Sánchez-Días (2017)
Número de habitantes	Kulpa (2014), Ferreira e Silva (2016), Lim <i>et al.</i> (2017)
Número de domicílios	Ferreira e Silva (2016)
Número de empresas	Kulpa (2014), Lim <i>et al.</i> (2017)
Variáveis relacionadas ao uso da terra	Lawson <i>et al.</i> (2012), Sánchez-Días <i>et al.</i> (2013; 2016b)
Número de veículos	Priya <i>et al.</i> (2015), Lim <i>et al.</i> (2017)
Número de veículos de carga	Lim <i>et al.</i> (2017)
Valor da folha de pagamento anual	Lim <i>et al.</i> (2017)
Atributos econômicos	Sánchez-Días <i>et al.</i> (2013; 2016b)
Área de estacionamento	Priya <i>et al.</i> (2015)
Características da rede	Sánchez-Días <i>et al.</i> (2013; 2016b)

No Quadro 3 são apresentadas as principais técnicas de modelagens utilizadas nas publicações identificadas na revisão sistemática da literatura. Observa-se que a regressão linear é o método clássico para definição de modelos de geração de viagens de carga, seguido do método dos mínimos quadrados. Pendyala *et al.* (2000), Souza *et al.* (2010), Ambrosini *et al.* (2008; 2010), Comi *et al.* (2012), Tavasszy e Jong (2014), Doustmohammadi *et al.* (2016), Ferreira e Silva (2016), Kaszubowski (2018) apresentam revisão da literatura sobre os modelos de geração de viagens de carga e a tendência das variáveis comumente utilizadas.

Quadro 3: Técnica utilizada para modelagem da geração de viagens de carga

Tipo de pesquisa	Referências
Análise de classificação múltipla	Campbell <i>et al.</i> (2012), Lawson <i>et al.</i> (2012)
Análise de dispersão	Gonzalez-Feliu <i>et al.</i> (2015)
Método dos mínimos quadrados (ou regressão linear)	Alho e Silva (2015a,b), Boarnet <i>et al.</i> (2017), Campbell <i>et al.</i> (2012), Ferreira e Silva (2016), Gonzalez-Feliu e Peris-Plab (2016), Gonzalez-Feliu <i>et al.</i> (2015), Günay <i>et al.</i> (2016), Holguín-Veras <i>et al.</i> (2011), Iding <i>et al.</i> (2002), Jaller <i>et al.</i> (2014; 2014a), Kulpa (2014), Lawson <i>et al.</i> (2012), Lim <i>et al.</i> (2017), Oliveira <i>et al.</i> (2016; 2017), Priya <i>et al.</i> (2015), Rudani <i>et al.</i> (2017)
Modelo contínuo-discreto	Jaller <i>et al.</i> (2015b), Sánchez-Días (2017)
Modelos econométricos	Jaller <i>et al.</i> (2015b), Sánchez-Días <i>et al.</i> (2013, 2016)
Modelos espaciais	Boarnet <i>et al.</i> (2017), Oliveira <i>et al.</i> (2016; 2017), Sánchez-Días <i>et al.</i> (2013, 2016)
Redes neurais artificiais	Kulpa (2014), Sánchez-Días <i>et al.</i> (2013; 2016)
Regressão linear generalizada	Alho e Silva (2014a)
Regressão logística	Jaller <i>et al.</i> (2015b), Lim <i>et al.</i> (2017)
Regressão não linear	Kulpa (2014)

A análise espacial para a geração de viagens de carga foi utilizada por poucos autores (Boarnet *et al.*, 2017; Oliveira *et al.*, 2016; 2017; Sánchez-Días *et al.*, 2013, 2016). Sánchez-Días *et al.* (2013; 2016b) utilizaram técnicas econométricas espaciais. Oliveira *et al.* (2016) utilizaram análise de cluster para identificar concentração de entregas. Boarnet *et al.* (2017) utilizaram mapa de calor para relacionar empregos com atividades de carga. Por fim, Oliveira *et al.* (2017) utilizaram análise da dispersão geográfica para analisar a concentração de entregas em bares e restaurantes.

Os resultados desta revisão permitem concluir que as pesquisas tendem em focar estabelecimentos comerciais para obtenção de dados de geração de viagens de carga. As variáveis comumente utilizadas são área do estabelecimento e número de funcionários. Para o desenvolvimento dos modelos, a regressão linear é a técnica comumente utilizada. Contudo, Handy (2015) destaca a necessidade de melhores técnicas para estimar a geração de viagens de carga. Alho e Silva (2014) propõem os modelos de regressão lineares generalizados como uma alternativa para evitar transformações em variáveis devido a heterocedasticidade. De acordo com Manning e Mullahy (2001), não existe uma única técnica de modelagem que seja melhor em todas as condições analisadas. Desta forma, este artigo apresenta a comparação de duas técnicas de modelagem em quatro diferentes condições, no intuito de identificar a potencialidade de outras abordagens de modelagem para a geração de viagens de carga baseada em dados coletados em estabelecimentos comerciais.

3. METODOLOGIA

Esta seção descreve a metodologia utilizada para obtenção de dados e o desenvolvimento dos modelos de regressão linear e regressão linear generalizada. A primeira etapa da metodologia descreve a coleta de dados, a segunda parte descreve as variáveis utilizadas nos modelos e a terceira parte descreve o desenvolvimento dos modelos.

3.1 Framework da coleta de dados

Neste estudo, os dados foram coletados nos estabelecimentos comerciais localizados na região central das cidades que compõem esta análise: Belo Horizonte, Quixadá, Palmas e Recife. Esta estratégia justifica-se pelos estudos que têm utilizado essa técnica de coleta de dados, bem como pela facilidade na obtenção de dados diretamente nos estabelecimentos comerciais (Sánchez-Días, 2017).

Os dados foram obtidos por meio de um formulário aplicado por um entrevistador, que solicitava informações da área do estabelecimento, número de empregados, número de entregas semanais, tipo predominante de produto comercializado e endereço completo. Os dados foram estratificados de acordo com as atividades econômicas, considerando que o tipo de produto comercializado nos estabelecimentos comerciais pode influenciar no padrão de geração de viagens de carga, conforme exposto por Campbel *et al.* (2012) e Cherrett *et al.* (2012). Na Tabela 1 é apresentada a amostra para o desenvolvimento dos modelos de geração de viagens.

Tabela 1: Número de estabelecimentos comerciais por cidade.

Atividade econômica	Tamanho da amostra
Belo Horizonte	31
Quixadá	11
Palmas	31
Recife	142

González-Feliu *et al.* (2015) consideram 30 observações como uma amostra mínima para estabelecer modelos de geração de viagens de carga. Oliveira *et al.* (2016) apresentam a amostra para os modelos nacionais que, em geral, é composto por poucas observações. Desta forma, apesar do número de estabelecimentos comerciais entrevistados que comercializam alimentos não ser suficiente para elaborar modelos de geração de viagens de carga em Quixadá (CE), de acordo com González-Feliu *et al.* (2015), foram elaborados modelos por ser uma cidade de pequeno porte, tornando um caso de análise interessante, em termos de comparação, daí ela configurar uma das cidades estudadas no trabalho. As demais cidades, apresentam amostra considerada suficiente para o desenvolvimento dos modelos.

3.2 Variáveis dos modelos

De acordo com Sánchez-Días (2017), as principais premissas para o desenvolvimento do modelo são que o tamanho de um estabelecimento (medido pela área ou número de funcionários) é um bom indicador da intensidade de atividade comercial e que estabelecimentos de diferentes setores comerciais têm diferentes necessidades logísticas. Cherrett *et al.* (2012) assumem que grandes estabelecimentos comerciais (em área) são responsáveis por um número maior de entregas e, segundo Alho e Silva (2014b), o uso desta variável se difundiu devido à falta de dados de outras variáveis explicativas. Neste estudo utilizou-se a área do estabelecimento e o número de empregados como variáveis explicativas do fenômeno modelado, cujo resumo das estatísticas descritivas encontra-se na Tabela 4.

Tabela 2: Estatísticas descritivas.

Atividade econômica	Estatística	Belo Horizonte	Quixadá	Palmas	Recife
Número de entregas	Mínimo	1,00	1,00	1,00	1,00
	Média	6,85	2,73	4,81	6,10
	Máximo	30,00	7,00	21,0	60,00
Área do estabelecimento (m ²)	Mínimo	16,00	25,00	20,00	3,00
	Média	85,74	229,90	165,70	309,82
	Máximo	380,0	1138,0	1100,0	9200,0
Número de funcionários	Mínimo	1,00	2,00	1,00	1,00
	Média	5,65	9,18	8,81	16,11
	Máximo	18,00	23,00	46,00	200,00

Na Tabela 3 é apresentada a correlação de Pearson entre as variáveis utilizadas nos modelos. Segundo Holguín-Veras *et al.* (2011), a maioria dos modelos apresentados na literatura foi obtida sem testar a significância da variável independente ou a validade de sua forma funcional. Recentemente, González-Feliu *et al.* (2015), Günay *et al.* (2016) e Oliveira *et al.* (2016) utilizaram a correlação de Pearson para verificar essa correlação. Neste estudo, observa-se que os dados para Belo Horizonte não apresentam correlação entre o número de entregas e as variáveis independentes (a correlação entre a área do estabelecimento e número de funcionários é menor que 0,5, sendo considerada fraca). Palmas apresenta correlação moderada. Neste estudo, serão desenvolvidos modelos mesmo para aqueles que não tem correlação para demonstrar a importância da análise das variáveis no desenvolvimento dos

modelos.

Tabela 3: Correlação entre as variáveis área e número de funcionários (variáveis independentes) com o número de entregas (variável dependente).

Variável	Belo Horizonte	Quixadá	Palmas	Recife
Área	0,16	0,21	0,67	0,39
Funcionários	0,29	0,59	0,63	0,62

A regressão linear (ML) e a regressão linear generalizada (MLG) foram utilizadas como técnicas de modelagem neste estudo. Detalhes sobre as técnicas podem ser obtidos em McCullagh e Nelder (1972), Alho e Silva (2014) e Sánchez-Días (2017).

Os modelos foram considerados estatisticamente válidos por meio da análise do teste-t, teste-F e p-valor, medidas utilizadas na literatura (Holguín-Veras *et al.*, 2011; Jaller *et al.*, 2014). O coeficiente de determinação (R^2) e critério de informação de Akaike (AIC) foram utilizados para definir os melhores modelos, considerando o equilíbrio entre a complexidade e o ajuste para as duas técnicas analisadas. Além disso, a raiz quadrada do erro médio (*Root Mean Square Error*, RMSE, em inglês) foi utilizada como medida para avaliar o modelo mais adequado para estimar a geração de viagens. A RMSE é comumente utilizada para comparar modelos e fornece uma boa indicação da acurácia preditiva dos modelos de regressão (Campbell *et al.*, 2012; Lawson *et al.*, 2012; González-Feliu *et al.*, 2015; Alho e Silva, 2017). As análises foram desenvolvidas utilizando o software R.

4. RESULTADOS

Na Tabela 4 são apresentados os resultados dos modelos de geração de viagens para estabelecimentos que comercializam alimentos e bebidas por meio da técnica de regressão linear. Os modelos para os estabelecimentos de Belo Horizonte (MG) e Quixadá (CE) não foram considerados estatisticamente válidos, estando diretamente relacionados ao fato das variáveis não serem correlacionadas. Para Palmas, o modelo ML6 teve pior ajuste ($R^2 = 0,39$) contudo apresenta adequado poder de predição (RMSE = 1,79). Para os dados de Recife, o modelo ML8 tem o melhor ajuste ($R^2 = 0,38$) e melhor poder de predição (RMSE = 3,07). Os resultados permitem concluir que, para o desenvolvimento de modelos de geração de viagens de carga:

- o tamanho da amostra e a homogeneidade dos dados são fundamentais;
- a confiabilidade da informação fornecida pelos entrevistados é fundamental;
- é fundamental avaliar a correlação entre as variáveis utilizadas;
- para a escolha do melhor modelo, recomenda-se a utilização do RMSE em vez do AIC ou R^2 .

Na Tabela 5 são apresentados os modelos de geração de viagens de carga por meio da técnica de regressão linear generalizada. Para Quixadá, o modelo MLG3 não é válido estatisticamente (valor de z inferior a 1,96). Dentre os modelos desenvolvidos para Belo Horizonte, o modelo MLG 1 tem menor RMSE, apesar do AIC ser maior quando comparado com o modelo MLG2. Para Palmas, o modelo MLG06 apresenta menores RMSE e AIC.

Tabela 4: Modelos de geração de viagens através da técnica de regressão linear.

Cidade	Variável independente	Modelo		Teste-t	p-valor	R ²	AIC	RMSE	
Belo Horizonte	Área	ML1	Coefficiente	0,025	0,85	0,40	0,02	209,28	nc
			Intercepto	5,64	2,31				
	Funcionários	ML2	Coefficiente	0,666	1,57	0,13	0,04	207,49	nc
			Intercepto	4,10	1,67				
Quixadá	Área	ML3	Coefficiente	0,001	0,65	0,53	0,04	50,95	nc
			Intercepto	2,42	3,12				
	Funcionários	ML4	Coefficiente	0,18	1,23	0,05	0,35	46,63	nc
			Intercepto	1,10	2,23				
Palmas	Área	ML5	Coefficiente	0,008	4,65	0,00	0,46	129,26	1,80
			Intercepto	2,956	6,02				
	Funcionários	ML06	Coefficiente	0,172	4,16	0,00	0,39	131,95	1,79
			Intercepto	2,877	5,33				
Recife	Área	ML07	Coefficiente	0,003	4,99	0,00	0,15	1001,8	4,35
			Intercepto	5,07	7,13				
	Funcionários	ML08	Coefficiente	0,22	3,52	0,00	0,38	957,34	3,07
			Intercepto	2,47	9,26				

(nc = não calculado)

Tabela 5: Modelos de geração de viagens através da técnica de regressão linear generalizado.

Cidade	Variável independente	Modelo		z-valor	AIC	RMSE	
Belo Horizonte	Área	MLG1	Coefficiente	0,009	3,73	248,73	9,58
			Intercepto	2,21	12,26		
	Funcionários	MLG2	Coefficiente	0,14	4,49	243,33	10,07
			Intercepto	2,08	11,54		
Quixadá	Área	MLG3	Coefficiente	0,0004	0,76	46,69	nc
			Intercepto	1,56	8,29		
	Funcionários	MLG4	Coefficiente	0,05	2,26	42,21	1,58
			Intercepto	1,13	4,29		
Palmas	Área	MLG5	Coefficiente	0,002	4,057	129,96	5,29
			Intercepto	1,762	15,07		
	Funcionários	MLG06	Coefficiente	0,036	3,93	128,36	4,19
			Intercepto	1,737	14,27		
Recife	Área	MLG07	Coefficiente	0,0005	11,45	1277,10	INF
			Intercepto	2,27	52,00		
	Funcionários	MLG08	Coefficiente	0,03	19,20	1007,60	INF
			Intercepto	1,79	35,45		

(nc = não calculado).

4.1. Discussão dos Resultados

Os modelos baseados em entrega têm sido considerados como os mais promissores em termos de possíveis resultados e capacidade de representar os diferentes aspectos das cadeias de suprimentos urbanas de maneira coerente (Kaszubowski, 2018). Contudo, deve-se ter atenção à fidedignidade na coleta de dados. Um procedimento sistematizado pode garantir uma melhor qualidade do dado, visto que o entrevistado pode não dizer a verdade sobre os dados coletados. Além disso, os resultados permitem concluir que as técnicas estatísticas devem ser melhor empregadas para análise dos dados. Os modelos que não foram considerados estatisticamente válidos não apresentaram correlação. Contudo, isto não é um teste comumente verificado na literatura, pois poucos estudos utilizam o teste de Pearson para verificar a correlação entre as variáveis (González-Feliu *et al.*, 2015; Günay *et al.*, 2016; Oliveira *et al.*, 2016).

Ainda, a agregação de dados tem impacto na qualidade e significância do modelo, conforme afirmam González-Feliu *et al.* (2015). Neste estudo, todos os estabelecimentos que comercializam alimentos e bebidas foram agrupados. Contudo, os resultados deste artigo permitem concluir que é necessária cautela na agregação dos dados, pois observou-se baixa correlação em alguns casos.

Neste estudo, os modelos que empregam como variável independente o número de funcionários foram os que apresentaram melhores RMSE. Contudo, conforme afirmam González-Feliu *et al.* (2015), deve-se ter cautela na utilização desta variável, visto que as empresas podem declarar números maiores/menores de funcionários, produzindo estimativas enganosas. Além disso, a utilização dessa variável como preditora para geração de viagens de futuros estabelecimentos pode ser de mais difícil aplicação, dependendo das características de operação e até mesmo da política interna da empresa. Em uma aplicação com fins de planejamento, a área se mostra uma variável de mais fácil obtenção e confiabilidade.

Sobre a técnica de modelagem, apesar do modelo linear generalizado ser adequado para modelar dados de contagem (McCullagh e Nelder, 1972; Manning e Mullahy, 2001), que é caso do número de entregas, ele não apresentou melhores resultados que o modelo linear. Os resultados obtidos neste estudo são similares aos obtidos por Alho e Silva (2014), que concluíram na comparação das mesmas técnicas, que o modelo linear generalizado não fornece predição superior ao modelo linear.

5. CONSIDERAÇÕES FINAIS

Este artigo apresenta uma comparação da modelagem da geração de viagens de carga utilizando regressão linear e regressão linear generalizada. Foram utilizados dados de quatro cidades brasileiras (Belo Horizonte, Palmas, Quixadá e Recife), coletados diretamente nos estabelecimentos comerciais. Os modelos foram considerados estatisticamente válidos analisando o resultado do teste-t e p-valor. A raiz quadrada do erro quadrático médio (RMSE) foi utilizada para identificar os modelos de melhor acurácia na predição.

Os dados de Belo Horizonte e Quixadá apresentaram correlação baixa e modelos foram desenvolvidos para demonstrar que a falta de correlação acarreta na não validação estatística dos modelos.

Em relação aos modelos desenvolvidos para Palmas e Recife considerando o modelo linear, o número de funcionários é a variável que melhor explica o fenômeno de entrega de alimentos e bebidas em estabelecimentos comerciais, sendo o RMSE igual a 4,35 para Palmas e 3,07 para Recife.

Em relação ao modelo linear generalizado, Quixadá não apresentou modelo estatisticamente válido para a variável número de funcionários. Palmas obteve um MLG com melhor poder de predição se comparado com o ML, cujo RMSE é de 4,19 (correlacionando o número de entregas com o número de funcionários).

Nas análises feitas, para um mesmo setor foram gerados modelos distintos para cada cidade, o que pode estar relacionado às diferenças de características urbanísticas e socioeconômicas das cidades brasileiras. Portanto, os modelos precisam ser locais, confirmando o estudo de

Oliveira et al. (2016) que demonstrou que cada regional de Belo Horizonte gera diferentes taxas de geração de viagens de carga para um mesmo setor. Desta forma, os modelos são válidos apenas para a área em que os dados foram coletados.

Para futuros trabalhos sugere-se padronizar a agregação dos dados utilizando a classificação nacional de atividades econômicas (CNAE), conforme utilizado por Oliveira *et al.* (2016) e explorado por González-Feliu et al. (2015) e Günay et al. (2016). Por fim, a comparação das técnicas requer ainda testes para identificar problemas de normalidade, assimetria, curtose e de heterocedasticidade, conforme apontado por Alho e Silva (2014) para avaliar a validade dos modelos lineares, além da análise dos resíduos do modelo, conforme proposto por Manning e Mullahy (2001). Tais fatos não foram explorados neste artigo, sendo sugestões para futuros estudos.

Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Fundação de Amparo a Pesquisa de Minas Gerais (FAPEMIG) pelo suporte à pesquisa.

REFERÊNCIAS BIBLIOGRÁFICAS

- Aditjandra, P.T., Galatioto, F, Bell, M. C. e T. H. Zunder (2016). Evaluating the impacts of urban freight traffic: application of micro-simulation at a large establishment. *European Journal of Transport and Infrastructure Research*, 16(1), 4-22. Disponível em <http://worldcat.org/issn/15677133>
- Al-Sobky, A.-S. A. e I. M. I. Ramadan (2015) A Generalized Piecewise Regression for Transportation Models. *International Journal of Computer Applications*, 129 (17), 16-22.
- Alho, A. R. e Silva, J. A. (2014a) Freight-Trip Generation Model: Predicting Urban Freight Weekly Parking Demand from Retail Establishment Characteristics. *Transportation Research Record: Journal of the Transportation Research Board*, 2411, 45-54. DOI 10.3141/2411-06
- Alho, A. R. e Silva, J. A. (2014b). The Development and Application of an Establishment-based Freight Survey: revealing retail establishments' characteristics, goods ordering and delivery processes for the city of Lisbon. In TRB 2014 Annual Meeting, Washington, D.C.
- Alho, A. R. e Silva, J. A. (2015) Lisbon's Establishment-based Freight Survey: revealing retail establishments' characteristics, goods ordering and delivery processes. *Eur. Transp. Res. Rev.*, 7(16), 1-17. DOI 10.1007/s12544-015-0163-7
- Alho, A. R., e Silva, J. A. (2017) Modeling retail establishments' freight trip generation: a comparison of methodologies to predict total weekly deliveries. *Transportation*, 44, 1195-1212. DOI 10.1007/s11116-016-9700-z
- Allen, J., Ambrosini, C., Browne, M., Patier, D., Routhier, J. e A. Woodburn (2014). Data Collection for Understanding Urban Goods Movement. In González-Feliu, J., Semet, F. and Routhier, J.-L. (Eds.) *Sustainable Urban Logistics Concepts Methods and Information Systems*, EcoProduction, Springer, Berlin Heidelberg, pp. 71–89.
- Allen, J., Browne, M. e T. Cherrett (2012). Survey Techniques in Urban Freight Transport Studies, *Transport Reviews*, 32(3), 287–311. DOI 10.1080/01441647.2012.665949
- Ambrosini, C., Meimbresse, B., Routhier, J. L., e Sonntag, H. (2008). Urban freight policy-oriented modeling in Europe. In E. Taniguchi, & R. G. Thompson (Eds.), *Innovations in City Logistics*. Nova Science Publishers, New York, US.
- Ambrosini, C., Patier, D. e Routhier, J.-L. (2010) Urban freight establishment and tour based surveys for policy oriented modeling. *Procedia Social and Behavioral Sciences*, 2, 6013-6026. DOI 10.1016/j.sbspro.2010.04.015
- Boarnet, M. G., Hong, A., Santiago-Bartolomei, R. (2017) Urban spatial structure, employment subcenters, and freight travel. *Journal of Transport Geography*, 60, 267-276. DOI 10.1016/j.jtrangeo.2017.03.007
- Campbel, S., Jaller, M., Sánchez-Díaz, I., Holguín-Veras, J., e Lawson, C. T. (2012) comparison between industrial classification systems in freight trip generation (FTG) modeling. In 91st TRB Annual Meeting, Washington, D.C.
- Cantilho, V., Jaller, M., e Holguín-Veras, J. (2012) Development of a national freight demand model with limited data and resources. In 91st TRB Annual Meeting, Washington, D.C.
- Cherrett, T., Allen, J., McLeod, F., Maynard, S., Hickford, A., Browne, M. (2012) Understanding urban freight

- activity – key issues for freight planning. *Journal of Transport Geography*, 24, 22-32. DOI 10.1016/j.jtrangeo.2012.05.008
- Comi, A., Site, P.D., Filippi, F. e A. Nuzzolo (2012). Urban freight transport demand modelling: A state of the art. *European Transport - Trasporti Europei*, 51, 1–17.
- Development of Trip Generation Model for Goods Movement of Gozaria and Visnagar G.I.D.C. *International Journal of Advance Engineering and Research Development*, 4(4), 923-926.
- Doustmohammadi, E., Sisiopiky, V. P., Anderson, M. D., Doustmohammadi, M., Sullivan, A. (2016) Comparison of Freight Demand Forecasting Models. *International Journal of Traffic and Transportation Engineering*, 5(1), 19-26. DOI 10.5923/j.ijtte.20160501.03
- Ferreira, B. F. G., Silva, M. A. V. (2016) Viagens de caminhões em áreas urbanas e sua relação com variáveis socioeconômicas. *GEPROS. Gestão da Produção, Operações e Sistemas*, 11(4), 197-212. DOI 0.15675/gepros.v11i4.1566
- González-Feliu, J. e J.-L. Routhier (2012). Modeling Urban Goods Movement: How to be Oriented with so Many Approaches? *Procedia - Social and Behavioral Sciences*, 39, 89–100. DOI 10.1016/j.sbspro.2012.03.093
- Gonzalez-Feliu, J. e Peris-Plab, C. (2017) Impacts of retailing attractiveness on freight and shopping trip attraction rates. *Research in Transportation Business & Management*, 24, 49-58. DOI 10.1016/j.rtbm.2017.07.004
- González-Feliu, J., I. Sánchez-Díaz e C. Ambrosini (2015). Aggregation level, variability and linear hypotheses for urban delivery generation models. *Transportation Research Board 95th Annual Meeting*, Washington, D.C.
- Günay, G., G. Ergün e I. Gökaşar (2016). Conditional Freight Trip Generation modeling. *Journal of Transport Geography*, 54, 102-111. DOI 10.1016/j.jtrangeo.2016.05.013
- Handy, S. (2015). Trip generation: Introduction to the special section. *The Journal of Transport and Land Use*, 8(1), 1-4. DOI 10.5198/jtlu.2015.822
- Holguín-Veras, J. e González-Calderón, C. (2013) Urban Freight Tour Models: State of the Art and Practice. In M. Ben-Akiva, H. Meersman e E. V. de Voorde. *Freight transport modelling*. Emerald: Bingley, UK. Chapter 17, 335-351.
- Holguín-Veras, J. e M. Jaller (2012). Comprehensive Freight demand Data Collection Framework for Large Urban Areas. *Transportation Research Board (TRB) 91st Annual Meeting*, Washington, D.C.
- Holguín-Veras, J. e M. Jaller (2014). Comprehensive Freight demand Data Collection Framework for Large Urban Areas. In González-Feliu, J., Semet, F. and Routhier, J.-L. (Eds.) *Sustainable Urban Logistics: Concepts, Methods and Information Systems*. EcoProduction, Springer, Berlin, Heidelberg, pp. 91–112. DOI 10.1007/978-3-642-31788-0
- Holguín-Veras, J., I. Sánchez-Díaz, C. T. Lawson, M. Jaller, S. Campbell, H. S. Levinson, e H.-S. Shin (2013). Transferability of Freight Trip Generation Models. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2379, pp. 1–8. DOI 10.3141/2379-01
- Holguín-Veras, J., M. Jaller, L. Destro, S. Ban, C. Lawson, e H. S. Levinson (2011). Freight Generation, Freight Trip Generation, and Perils of Using Constant Trip Rates. *Transportation Research Record: Journal of the Transportation Research Board*, 2224, 68–81. DOI 10.3141/2224-09
- Iding, M. H. E., W. J. Meester e L. A. Tavasszy (2002). Freight trip generation by firms. 42nd European Congress of the Regional Science Association, Dortmund, Germany.
- Jaller, M., I. I. Sánchez-Díaz, J. Holguín-Veras, C. T. Lawson. (2014). Area based freight trip generation models. *Transportation Research Board (TRB) 93rd Annual Meeting*, Washington, D.C.
- Jaller, M., I. Sánchez-Díaz e J. Holgín-Veras (2015b). Identifying Freight Intermediaries: Implications for Modeling of Freight Trip Generation. *Transportation Research Record: Journal of the Transportation Research Board*, 2478, 48-56. DOI 10.3141/2478-06
- Jaller, M., X. Wang, e J. Holguín-Veras (2015a). Large urban freight traffic generators: Opportunities for city logistics initiatives. *Journal of Transport and Land Use*, 8(1), 1-17. DOI 10.5198/jtlu.2015.406
- Kaszubowski D. (2018) Urban Freight Transport Demand Modelling and Data Availability Constraints. In: Sierpiński G. (eds) *Advanced Solutions of Transport Systems for Growing Mobility*. TSTP 2017. *Advances in Intelligent Systems and Computing*, vol 631. Springer, Cham. DOI 10.1007/978-3-319-62316-0_14
- Kulpa, T. (2014) Freight truck trip generation modeling at regional level. *Procedia – Social and Behavioral Sciences*, 111, 197-202. DOI 10.1016/j.sbspro.2014.01.052
- Lawson, C. T., J. Holguín-Veras, Sánchez-Díaz, I., Jaller, M., Campbell, S. e E. Powers (2012) Estimated Generation of Freight Trips Based on Land Use. *Transportation Research Record: Journal of the Transportation Research Board*, 2269, 65–72. DOI 10.3141/2269-08

- Lim, H., S. M. Chin, H-L Hwang, L. D. Han (2017) Incorporating Information Complexity into Regression-Based Freight Generation Model Selection. *Transportation Research Board (TRB) 97rd Annual Meeting, Washington, D.C.*
- Manning, W. G. e J. Mullahy (2001) Estimating log models: to transform or not to transform? *Journal of Health Economics*, 20, 461-494. DOI
- McCullagh, P. e J. A. Nelder (1972). *Generalized Linear Models*. Chapman and Hall: London.
- Oliveira, L. K., R. A. A. Nóbrega, D. G. Ebias, B. G. S. Corrêa (2017) Analysis of freight trip generation model for food and beverage in Belo Horizonte (Brazil). *Region*, 4(1), 17-30. DOI 10.18335/region.v4i1.102
- Oliveira, L. K., R. L. M. Oliveira, C. M. F. Ramos, D. G. Ebias (2016) Modelo de geração de viagens de carga em áreas urbanas: um estudo para bares, restaurantes e supermercados. *Transportes*, 24 (2), 53-67. DOI 10.4237/transportes.v24i2.1058
- Pendyala, R. M., V. N. Shankar, R. G. McCullough (2000) Freight Travel Demand Modeling: Synthesis of Approaches and Development of a Framework. *Transportation Research Record: Journal of the Transportation Research Board*, 1725, 9–16. DOI 10.3141/1725-02
- Priya, C. D., G. Ramadurai e G. Devi (2015) Freight trip generation models for Chennai, India. *Transportation Research Board (TRB) 95rd Annual Meeting, Washington, D.C.*
- Rudani, H. R., V. F. Patel e H. K. Dave (2017)
- Sánchez-Días, I. (2017) Modeling urban freight generation: A study of commercial establishments' freight needs. *Transportation Research Part A*, 102, 3-17. DOI 10.1016/j.tra.2016.06.035
- Sánchez-Días, I., J. Holguín-Veras e C. Wang (2013) Assessing the Role of Land-Use, Network Characteristics and Spatial Effects on Freight Trip Attraction. *Transportation Research Board (TRB) 93rd Annual Meeting, Washington, D.C.*
- Sánchez-Días, I., J. Holguín-Veras e X. Wang (2016) An Exploratory Analysis of Spatial Effects of Freight Trip Attraction. *Transportation*, 43, 177-196. DOI 10.1007/s11116-014-9570-1
- Souza, C. D. R., S. D. Silva, M. A. D'agosto (2010) Modelos de geração de viagem para pólos geradores de viagens de carga. *Transportes*, 18 (1) 46-57. DOI 10.14295/transportes.v18i1.396
- Tavasszt, L. e G. Jong (2014) *Modelling freight transport*. Elsevier, London; Waltham.