

# Freight trip generation to buildings under construction: a comparative analysis with linear regression and generalised linear regression

*Geração de viagens de carga para edifícios em construção: uma análise comparativa entre regressão linear e regressão linear generalizada*

Leise Kelli de Oliveira<sup>1</sup>, Rafael Tagliatti Herédia<sup>2</sup>, Bruno Vieira Bertoncini<sup>3</sup>,  
Renata Lúcia Magalhães de Oliveira<sup>4</sup>

<sup>1</sup>Universidade Federal de Minas Gerais, Minas Gerais – Brasil, leise@etg.ufmg.br

<sup>2</sup>Universidade Federal de Minas Gerais, Minas Gerais – Brasil, rafaeltagliatti@hotmail.com

<sup>3</sup>Universidade Federal do Ceará, Ceará – Brasil, bruviber@det.ufc.br

<sup>4</sup>Centro Federal de Educação Tecnológica de Minas Gerais, Minas Gerais – Brasil, renataoliveira@cefetmg.br

## Recebido:

13 de novembro de 2018

## Aceito para publicação:

7 de junho de 2020

## Publicado:

15 de dezembro de 2020

## Editor de área:

Helena Beatriz Cybis

## Keywords:

Freight trip generation model.

Buildings under construction.

Linear regression.

Generalised linear regression.

Comparative analysis.

## Palavras-chaves:

Modelos de geração de viagens de carga.

Edifícios em construção.

Regressão linear.

Regressão linear generalizada.

Análise comparativa.

DOI:10.14295/transportes.v28i5.1885

## ABSTRACT

Estimating the number of trips generated by a company is an essential part of the process of freight demand modelling. In this context, the current study examines freight trip generation to buildings under construction (BUC) using generalised linear regression and linear regression, through a case study to Belo Horizonte. The main contributions of this paper are related to the models to estimate freight trips to BUC, the verification of the linearity assumptions of the linear models and the comparison of different modelling techniques for the freight trip generation models. Linearity assumptions verified the reliability of the results of the linear regression models. Results indicate that the models with the best accuracy in predicting freight trips to the BUC are the linear models that use the area as an explanatory variable.

## RESUMO

A estimativa do número de viagens geradas por uma empresa é parte essencial do processo de modelagem da demanda de carga. Nesse contexto, o presente estudo compara modelos de viagens de carga para edifícios em construção (BUC) estimados por regressão linear generalizada e regressão linear, em um estudo de caso para Belo Horizonte. As principais contribuições deste artigo estão relacionadas aos modelos para estimar viagens de carga para BUC, a verificação das premissas de linearidade dos modelos lineares e a comparação de diferentes técnicas de modelagem para a geração de viagens de carga. As suposições de linearidade permitiram avaliar a confiabilidade dos resultados dos modelos de regressão linear. Os resultados indicam que os modelos com melhor acurácia na previsão de viagens de carga ao BUC são os modelos lineares que utilizam a área como variável explicativa.



## 1. INTRODUCTION

Urban freight transport (UFT) supports the urban lifestyle because people need access to consumer goods (such as food, medicines, clothing and other products) and services (such as a

waste collection). However, UFT operation directly impacts both local and regional economies (Ferreira and Silva, 2016). Despite this, local authorities are not able to establish measures that minimise the effects of freight transport in the urban environment, since they do not know the dynamics of freight flow in their territory (Lawson et al., 2012).

In order to develop effective public policies that minimise the operational and environmental impacts of UFT, the number of trips generated by activities in a city must be known (Comi et al., 2012) and understood by transport planners (Gonçalves et al., 2012; Oliveira et al., 2016). This is particularly important given that the number of trips can cause harmful effects on road traffic in its immediate surroundings and, in some cases, hinder the accessibility of the region, aggravate vehicle and pedestrian safety conditions (Comi et al., 2012; Oliveira et al., 2016), and cause increased emissions of pollutants and fuel consumption. To address such a situation, freight trip generation models (FTGM) estimate the number of trips produced and attracted, based on variables that reflect characteristics of the region or of the phenomenon, considering the dynamics of the urban space (Lawson et al., 2012).

For Brazil, the literature presents FTGM to pubs and restaurants (Campos and Melo, 2004; Silva and Waisman, 2007; Oliveira et al., 2016), supermarkets (Campos and Melo, 2002; Gasparini et al., 2010; Oliveira et al., 2016; Reis et al., 2018), shopping centres (Gasparini et al., 2010), and retail sites, including those specialising in clothing, food, construction materials and fuel (Campos and Melo, 2004; Oliveira et al., 2018). Against such a backdrop, this paper presents freight trip generation models for buildings under construction (BUC), which have not been examined in the Brazilian context.

The construction sector is economically important to countries around the world. For example, in 2017, the share of civil construction in GDP was 4.4% in Brazil (CBIC, 2018), 5% in the European Union (EuroStat, 2018) and 4.3% in the United States (Statista, 2018). This sector generates a significant number of trips to guarantee production throughout a project's construction process, and it does not always consider the internalisation of loading and unloading operations. Also, BUC can occur in any location within a town, which adds an additional layer of complexity if it is compared with the supplying of other sectors. Thus, the development of FTGM for BUC was seen as a timely research opportunity.

Although BUC have transitory movements and the attractiveness of freight trips exists only during the construction time frame, the impact of these movements is significant, especially when the location of BUC takes place in dense and congested regions. Also, many cities offer temporary unloading areas in front of the BUC. The decision on the size of unloading areas is arbitrary due to the lack of knowledge of transportation analysts about the construction process and the generation of trips. Thus, this paper contributes to the literature about freight trip generation models in BUC.

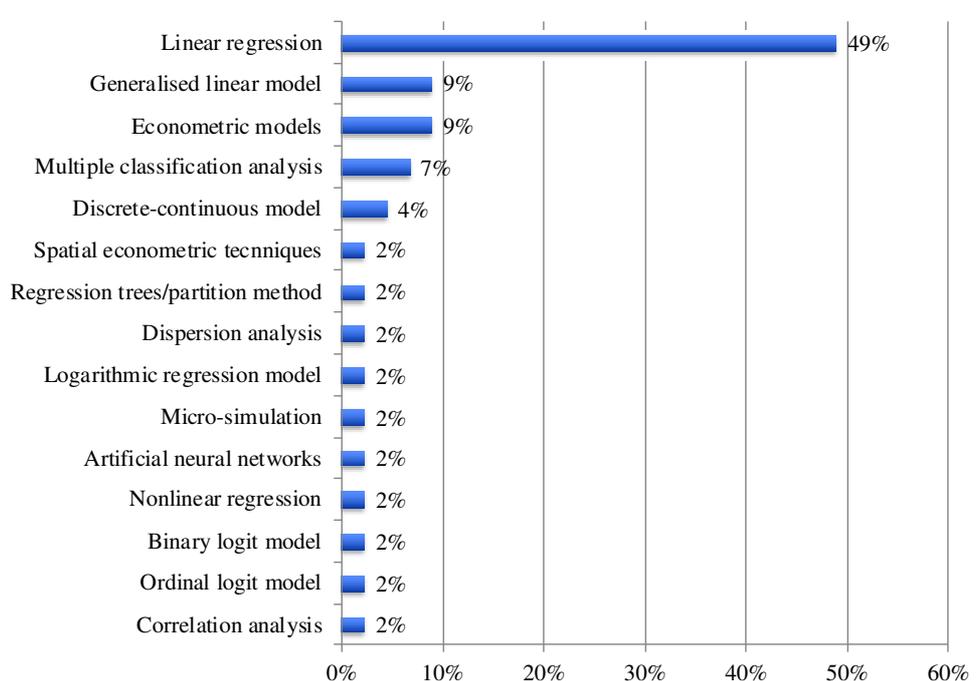
In this way, the paper presents freight trip generation models in BUC through a case study to Belo Horizonte. The following hypotheses were considered: (i) FTGM in BUC is a linear phenomenon; (ii) the linearity assumptions are fundamental to evaluating the accuracy of the model; (iii) cross-validation allows us to identify the predictive capacity of the model; and (iv) although generalized linear regression is the technique usually applied, linear regression provides models with better predictive capability. Thus, from this set of hypotheses, the paper seeks to contribute to the field in three ways: (i) the development of FTGMs in BUC; (ii) the use of linearity assumptions and a cross-validation technique to evaluate the accuracy

and predictive capacity of the model, respectively; and (iii) the comparison of different trip generation modelling techniques.

## 2. FREIGHT TRIP GENERATION MODELLING

A systematic literature review (SLR) was carried out to identify the state of the art of FTGMs based on the review protocol proposed by Wee and Banister (2016). The SLR consisted of three steps. In the first step, data sources and keywords were identified. Science Direct, Web of Knowledge, Scopus (Elsevier) and TRID (Transport Research International Documentation) were used as database sources. Freight trip and freight trip generation models were used as keywords (note: keywords in Portuguese and English were considered). In the second step, abstracts were read, and papers were selected according to the following approaches: FTGM (the specific theme of this study), freight demand management (related to the freight trip generation) and urban goods distribution (due to the relationship between urban development patterns and freight trip). In the third step, papers not related to the theme were excluded, and those identified by the snowball strategy were included.

After this procedure, 63 papers were identified by keywords, with 12 redundancies noted (i.e. the same paper in a different database). After analyses of the abstracts, 25 papers were included in the SLR. Also, 17 articles were identified by the snowball strategy, yielding a total of 42 papers for the SLR. Notably, the literature review did not identify freight trip generation models for BUC, as proposed in this paper. The focus of this SLR comprises the modelling techniques, explanatory variables and sector considered to estimate freight trips.



**Figure 1.** Techniques used in freight trip generation models

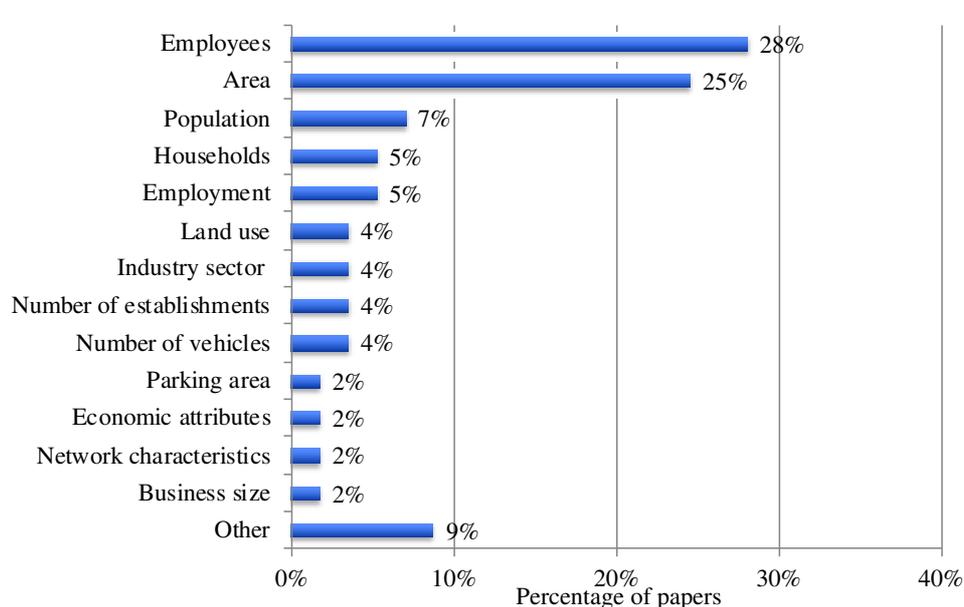
Figure 1 presents the modelling techniques used to develop FTGMs, of which linear regression is the most popular (Iding et al., 2002; Silva and Waisman, 2007; Gasparini et al., 2010; Holguín-Veras et al., 2011; Campbel et al., 2012; Lawson et al., 2012; Alho and Silva, 2014a; Jaller et al., 2014; Kulpa, 2014; Alho and Silva, 2015; Gonzalez-Feliu et al., 2015; Jaller et al., 2015a,

2015b; Ferreira and Silva, 2016; Gonzalez-Feliu and Peris-Plab, 2017; Günay et al., 2016; Oliveira et al., 2016; Lim et al., 2017; Boarnet et al., 2017; Rudani et al., 2017; Oliveira et al., 2017, 2018). There is also an increasing number of papers that use generalised linear models (Alho and Silva, 2014a, 2017; Oliveira et al., 2018; Reis et al., 2018) and other econometric techniques (Holguín-Veras et al., 2013; Jaller et al., 2015a, 2015b; Sánchez -Días et al., 2016). In addition, some analyses used linear regression and generalised linear regression (Alho and Silva, 2014a; 2015b, 2017; Oliveira et al., 2018).

The accuracy of the models was generally assessed using the coefficient of determination ( $R^2$ , 32%), root mean square error (RMSE, 24%);  $R^2$  and RMSE (8%); t-value,  $R^2$  and RMSE (4%); mean squared error, t-value and  $R^2$  (4%) and other combinations of measures (28%), such as Akaike information criteria and the spatial error model. Only Alho and Silva (2015b) evaluated the accuracy of the model analysing linearity assumptions, as proposed in this paper.

Figure 2 presents all explanatory variables identified in the SLR. In general, the explanatory variables commonly used to model the phenomenon are the area of the company and number of employees. Silva and Waisman (2007), Holguín-Veras et al. (2011), Alho and Silva (2014a, 2017), Aditjandra et al. (2016), Oliveira et al. (2016, 2017, 2018) and Sánchez-Días et al. (2017) present models using the area of the company and number of employees as explanatory variables. Other explanatory variable highlights are as follows:

- socioeconomic variables were used to estimate freight trip generation in the Metropolitan Region of Rio de Janeiro (Ferreira and Silva, 2016b);
- population, number of households, employees and area were used to identify shopping trip attraction rate (Gonzalez-Feliu and Peris-Plab, 2017); and
- land-use variables were used to estimate freight trips (Lawson et al., 2012).



**Figure 2.** Explanatory variables used in freight trip generation models

FTGM were estimated to logistics sites or warehouses (Holguín-Veras et al., 2011; Jaller et al., 2014, 2015; Gonzalez-Feliu et al., 2015; Günay et al., 2016), wholesale trade/industry (Iding et al., 2002; Lawson et al., 2012; Jaller et al., 2014; Gonzalez-Feliu et al., 2015; Jaller et al. 2015a;

Lim et al., 2017; Rudani et al., 2017), trade sector (Gasparini et al., 2010; Holguín-Veras et al., 2011; Lawson et al., 2012; Holguín-Veras et al., 2013; Alho and Silva, 2014a, 2014b, 2015; Jaller et al., 2014; Gonzalez-Feliu et al., 2015; Jaller et al. 2015a; Gonzalez-Feliu and Peris-Plab, 2016; Sánchez-Días, 2017; Oliveira et al., 2018) and food service (Gasparini et al., 2010; Silva and Waisman, 2007; Holguín-Veras et al., 2011; Lawson et al., 2012; Holguín-Veras et al., 2013; Jaller et al., 2014; Jaller et al., 2015a; Gonzalez-Feliu and Peris-Plab, 2016; Oliveira et al., 2016, 2017; Sánchez-Días, 2017; Reis et al., 2018). Similar to the sector proposed for analysis in this paper, Jaller et al. (2015a) presented FTGM for the construction sector that consider the North American Classification System (code 23), which includes establishments involved in the construction of buildings and engineering projects.

The results of the SLR highlight the contribution of this paper, as FTGM in BUC are not considered in the literature. Also, it is unusual to use linear assumption and cross-validation to identify the predictive capacity of the model, as is proposed in the current study. Accordingly, this paper contributes to the literature in two ways: concerning the method, it used techniques to assess the accuracy of the prediction of the model, and, concerning the phenomenon, it addressed a topic that had not yet been studied in the literature.

### 3. RESEARCH APPROACH

The research approach used to determine the freight trip generation models for buildings under construction was based on Oliveira et al. (2016) and Campos et al. (2012), as described in the next sections.

#### 3.1. Definition of objective, scope and area of study

FTGM contributes to urban freight transportation planning (Oliveira et al., 2016) by creating estimate models for BUC (residential or commercial).

The study scope is based on trips generated to BUC in the study area of Belo Horizonte, MG, Brazil. A building under construction passes through five main stages during the construction process: foundation, structure construction, brickwork and interior rough-in, coating and finishing. Depending on the type of constructive structure of the building (concrete, structural masonry or mixed structure), a different number of trips can be generated using many types of freight vehicles. Dump trucks are typically used in the foundation stage, cement mixer trucks are used in the structure construction stage and flatbed and box trucks are used in other phases of construction.

Regardless of the stage and type of construction, BUC attracts freight vehicles in order to supply goods for the construction project. Moreover, BUC also produces trips to remove construction waste. Despite it being possible to analyse these trips together, in this paper, the produced and attracted trips were analysed separately, since the BUC generated a different number of trips in each phase of construction. Another motivation for considering these phenomena individually is the scarcity of models to explain them. It is important to mention that regardless of whether the trip is attracted or produced, at some point, an empty trip (where the vehicle is empty) will be produced by the BUC, but these empty trips were not analysed in this study.

#### 3.2. Definition of dependent and explanatory variables

In freight transportation, an estimate of the trips generated can be made in terms of the number of trips or the amount of cargo transported (Nuzzolo et al., 2012). In this paper, the dependent

variable is the number of freight vehicles attracted (i.e. deliveries of goods to BUC) and produced (i.e. collection of waste of the BUC) by week. The time unit 'week' was chosen because the planning of a BUC is on a weekly basis. In this sense, it is logical to use the same time unit in the freight trip generation modelling.

Regarding the exploratory variables, area and number of employees were considered as explanatory variables since they are the classical explanatory variables in FTGM (Alho and Silva, 2017; Oliveira et al., 2018). Additionally, the number of floors and units were alternative explanatory variables considered in the modelling.

### 3.3. Identification of buildings under construction in the area of study and data collection

The identification of BUC was obtained through the Approved Building Projects Report (Belo Horizonte, 2018), which identifies BUC projects approved by the Belo Horizonte municipality. This report provides the location of the BUC in Belo Horizonte.

A questionnaire was designed to obtain the data. The structure of the questionnaire is presented in Table 1. Data were collected considering the stage of construction, since the number of employees is related to the services performed and varies according to the construction phase.

**Table 1** – Structure of the questionnaire

Block	Theme	Variable	Type of response
Block 1	BUC information	Address	Text
		Type of building	Categorical (4 class)
		Type of constructive structure	Categorical (4 class)
		Total area	Continuous
		Number of floors	Continuous
		Construction time	Continuous
		Stage of construction	Categorical (5 class)
		Number of employees	Continuous
Block 2	Earthmoving process	Earthmoving process	Binary
		Volume (m <sup>3</sup> )	Continuous
		Number of trucks	Continuous
Block 3	Freight collection and delivery process	Number of deliveries per week	Continuous
		Number of collects per week	Continuous
		Freight vehicle type	Categorical (8 class)
		Peak weekday	Categorical (6 class)
		Day period	Categorical (4 class)
		Parking place	Categorical (3 class)
		Loading time	Continuous
		Unloading time	Continuous

Engineers responsible for new projects in a construction company in Belo Horizonte were interviewed to validate the questionnaire. The questionnaire was answered by engineers or construction managers involved in the routine of the building under construction.

#### 3.3.1. Sample

Data from the Project Report approved by the Belo Horizonte municipality was used to define the sample. This report presents information about the projects that requested a building permit from January 2017 to December 2017. We identified 604 BUC in Belo Horizonte in 2017, and the data collection was planned considering the number of BUC in nine administrative

regions in Belo Horizonte. The sample size estimated was 83 BUC, with a 95% confidence level and 10% margin of error.

### 3.4. Analysis of the correlation

The analysis of the correlation between variables determines the strength of the relationship between two paired observations (Stevenson, 1981). Pearson's correlation coefficient was used to analyse the correlation between the dependent variable and the exploratory variables. If there is a correlation between the independent and dependent variable, it is more probable to obtain a model with statistical significance.

### 3.5. Modelling of freight trip generation

Linear regression and generalised linear regression (GLM) were used as modelling techniques. According to Maia (2017), regression analysis consists of obtaining an equation that tries to explain the variation of the dependent variable by the variation of the independent variable. Washington et al. (2010) and Maia (2017) present the details of the linear regression technique.

GLM is used when linear regression is inadequate (i.e. the dependent variable is not asymmetric or represents data from counts, or the data is binary). McCullagh and Nelder (1989) developed the GLM, incorporating exponential family distributions to the regression adjustment. The Poisson distribution is indicated for the regression adjustment (Washington et al., 2010) considering FTGM. The log-likelihood function is used to estimate the parameters.

The non-bias and minimum variance in the estimation of the coefficient were verified through the t-test (Maia, 2017). The null hypothesis is that the estimation of the coefficient is significant for the model at  $t\text{-test} < 0.05$ . The coefficient of determination ( $R^2$ ) was used to identify the proportion of the variability of the dependent variable that is explained by the independent variable of the model. Analysis of variance (ANOVA - test f) allowed for verification as to whether the model contributes to explain the dependent variable. For this, the model contributes to explain the dependent variable if the  $p\text{-value} \leq \alpha$  (null hypothesis).

Linearity assumptions – linearity, mean of errors equal to zero, homoscedasticity, autocorrelation between errors and normality of residues – were tested to assess the accuracy of the model (Washington et al., 2010). Residuals vs fitted plot was used to evaluate linearity; homoscedasticity was verified by the Breusch-Pagan test (Breusch; Pagan, 1979). The null hypothesis considers the model as homoscedasticity, i.e. there is constant variance in the residuals at the 5% significance level,  $P[\chi^2] > 0.05$  (Maia, 2017). The normality of the residues was verified by the Shapiro-Wilk test, in which the sample comes from a normal distribution (null hypothesis), rejecting it if the result of the test is less than  $P [W_{\text{calculado}}] < P [W_{\alpha}]$ , where  $P [W_{\alpha}]$  is from p-value.

The model has accuracy if the model meets all linearity assumptions. The predictive capacity of the model was evaluated by the chi-square test, the square root of the mean square error (RMSE) and the cross-validation test (Hyndman, 2006; Arlot and Celisse, 2010). Leave-one-out cross-validation (LOOCV) process was used; for more details, including the procedure associated with LOOCV, see Arlot and Celisse (2010).

Concerning the GLM, the hat-value identifies the leverage values in the model. Estimation of the GLM coefficients was evaluated by the Pearson chi-square test ( $\chi^2$ ), which is used to determine if there is a significant difference between the expected values in a predictive model and the values observed in the data set. Akaike information criterion (AIC) was used to classify the

models (Akaike, 1974). Alternatively, Hurvich and Tsai (1989) proposed the corrected AIC (AICc), which is indicated for small samples with a normal distribution, as being a more suitable criterion to select the model (Davison, 2001). Burnham and Anderson (2002) recommend using AIC to select models when the number of observations is at least 40 times greater than the number of parameters.

The models were estimated using the software R version 3.4.4, launched on March 15, 2018, through the following package: RSQ (Zhang, 2018), metrics (Frasco, 2018), lmtest (Hothorn et al., 2018), ISLR (James et al., 2017) and boot (Canty and Ripley, 2017).

## 4. RESULTS

Ninety BUC were randomly selected from the Report of Belo Horizonte Municipality. In the field survey, 36 BUC are located and were replaced by another BUC in the same region. Data from 105 BUC located in Belo Horizonte were obtained. The data collection occurred between August 2018 until November 2018.

### 4.1. Characterization of the buildings under construction

Of the 105 BUC, 79.0% of the projects are of concrete frame network, 16.2% of structural masonry and 4.8% use mixed structure, i.e. part in reinforced concrete and part in structural masonry. Residential projects are the majority (87.6%), 3.8% are buildings exclusively for commercial use and 8.6% are mixed buildings (commercial use on the ground floor and residential units on other floors). Also, 17% were in the foundation phase, 19% in the structure construction stage, 21.9% in the brickwork and interior rough-in stage, 23.8% in the coating phase and 18.1% in the finishing phase.

Regarding deliveries, 15.2% occur on Monday, 18.1% on Tuesday, 23.8% on Wednesday and 4.8% on Thursday. Some of the interviewees (38.1%) were not able to specify one day with more frequency of deliveries. Deliveries mainly occur between 7-10 hours (51.4%), 16.2% occur between 10-12 hours, 3.8% between 12-14 hours and 4.8% between 14-17 hours, and 23.8% of respondents did not know when the deliveries occur.

For unloading operations, 78.1% of the vehicles park on the street, in front of the BUC, and 10.5% park in a temporary unloading area. Also, in 11.4% of the BUC, the vehicles park at the construction site, mainly in the foundation phase. According to the Transport and Traffic Company of Belo Horizonte (BHTRANS), only five BUC requested a temporary unloading area in 2107.

Unloading time ranges from 30 minutes (21.9% of respondents) to 45 minutes (28.5%), 60 minutes (39.0%), 90 minutes (4.9%) and 120 minutes (5.7%). The capacity of vehicles range from 3.5-6 tonnes (1.9%) to 6-10 tonnes (36.2%), 10-15 tonnes (48.6%) and 15-40 tonnes (13.3%). Fixed deliveries occur in 22% of BUC, as one delivery per week (20%), two deliveries per week (0.95%) or four deliveries per week (3.81%). Waste construction collection occurred in 45.7% of the BUC. In general, construction time is around 2 years.

### 4.2. Freight trip generation models

Data were segregated according to the type of the constructive structure of the building (concrete, structural masonry and mixed) and the construction stage of the construction for the development of the models. Table 2 presents the sample of BUC according to the constructive

structure and stage of construction. A minimum sample of five BUC was considered for modelling. In this way, it was possible to obtain equations for all stages considering the BUC in concrete. Also, it is possible to obtain equations for structural masonry constructive structure for brickwork and interior rough-in stage, and coating stage.

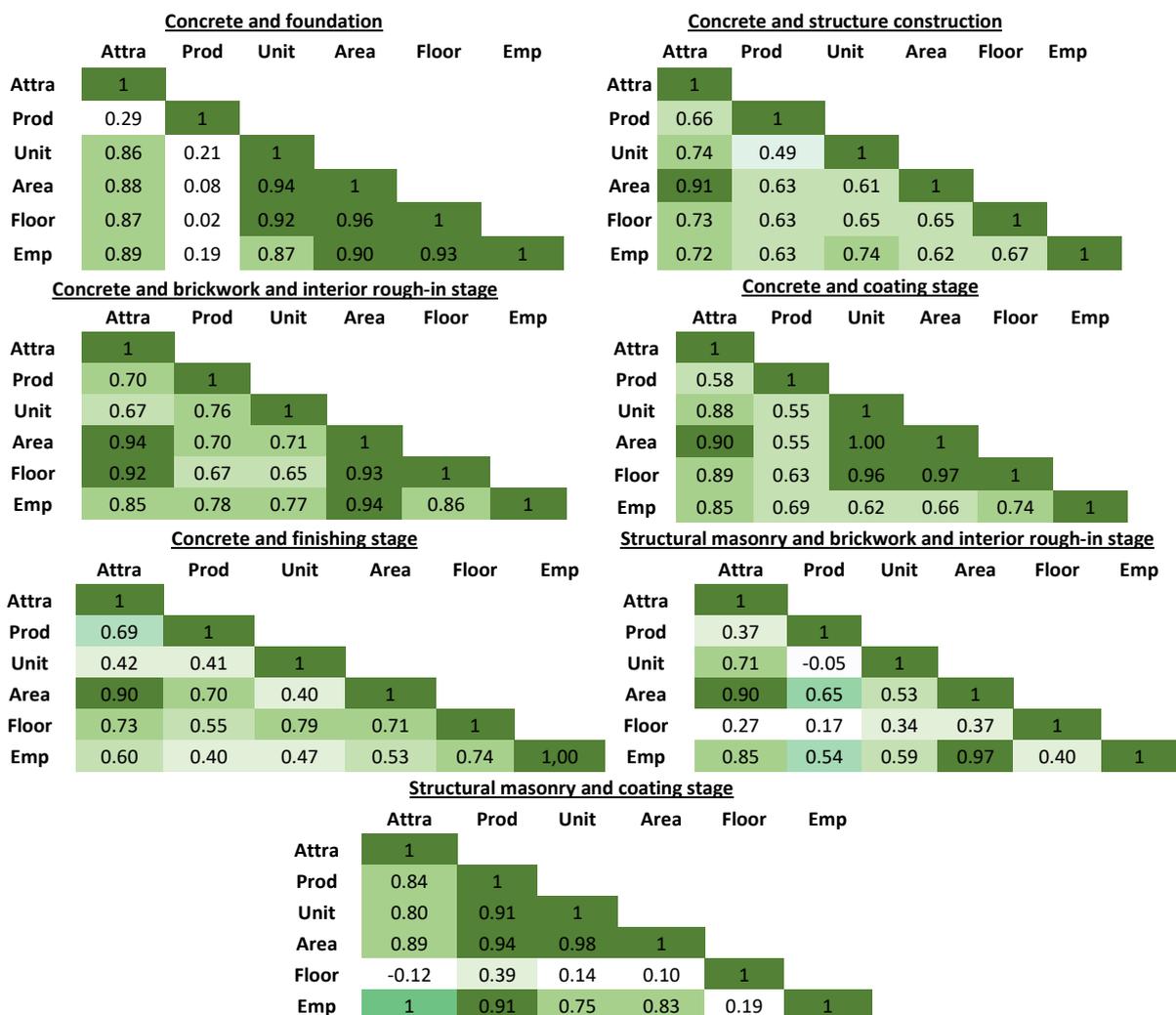
**Table 2** – Number of BUC considering the constructive structure of the building and stage of the construction

Constructive structure of the building	Stage of the construction				
	Foundation construction	Structure construction	Brickwork and interior rough-in stage	Coating stage	Finishing stage
Concrete	18	20	12	18	15
Structural masonry	-	-	8	5	4
Mixed	-	-	3	2	-

**Table 3** – Descriptive statistics of dependent and exploratory variables

Variable	Short name	Minimum	1 <sup>st</sup> Quartile	Average	Median	3 <sup>rd</sup> Quartile	Maximum
FV attracted	Attra	1.00	3.00	4.11	4.00	5.00	12.00
FV produced	Prod	0.00	1.00	1.86	2.00	3.00	5.00
Units	Unit	1.00	6.00	13.75	8.00	16,00	117.00
Area	Area	284.80	674.20	1,769.70	997.20	2,160.50	13,403.20
Floors	Floor	2.00	4.00	6.71	6.00	8.00	27.00
Employees	Emp	5.00	8.00	12.86	12.00	17.00	27.00

**Table 4** – Pearson correlation between the variables by constructive structure of the building and stage of construction



#### 4.2.1. FTGM by linear regression

We developed 110 models using linear regression – 55 from attraction, 55 from production. From that, only nine models were validated by statistical tests and linearity assumptions. These models are presented in Table 5. Regarding the attraction models, freight vehicle attraction models for BUC in concrete, independent of the construction stage, were obtained. Also, model LM6 is predictive due to the cross-validation result. This model considers a linear freight vehicle attraction model for any constructive structure or stages of BUC.

**Table 5** – Linear freight vehicle generation models (p-v = p-value)

Type/stage and model ID	Variable and intercept	Value estimated	t-test	p-value	ANOVA	R <sup>2</sup>	AIC	Breusch-Pagan test (BP)	Shapiro-Wilk test (W)	LOOCV	RSME
<b>Attraction (*99% confidence level)</b>											
Concrete and Foundation stage LM1	Area	0.0005	7.44	1.40E-06*	55.37	0.78	41.5	BP=0.002 p-v=0.96	W=0.93 p-v=0.20	2.04	0.65
	Intercept	2.08	10.63								
Concrete and Structure stage LM2	Area	0.0007	9.01	4.30E-08*	81.25	0.82	53.4	BP=0.00 p-v=0.99	W=0.97 p-v=0.72	1.25	0.79
	Intercept	3.40	12.40								
Concrete and Brickwork and interior rough-in stage LM3	Area	0.002	8.78	5.17E-06*	77.08	0.89	30.3	BP=0.95 p-v=0.33	W=0.92 p-v=0.26	0.60	0.67
	Intercept	2.16	5.97								
Concrete and coating stage LM4	Area	0.0007	8.41	2.70E-07*	71.42	0.82	57.1	BP=0.17 p-v=0.68	W=0.93 p-v=0.17	12.95	1.00
	Intercept	3.10	10.40								
Concrete and finishing stage LM5	Area	0.0008	7.42	5.05E-06*	55.06	0.81	35.8	BP=3.27 p-v=0.07	W=0.94 p-v=0.33	0.51	0.65
	Intercept	2.39	9.80								
GENERAL LM6	Area	0.0004	9.16	2.20E-16*	209.5	0.80	266.4	BP=5.78 p-v=0.06	W=0.99 p-v=0.81	0.76	0.83
	Emp	0.17	9.39								
	Intercept	1.16	5.42								
<b>Production (*99% confidence level)</b>											
Concrete and structure stage LM7	Area	0.0003	3.408	3.14E-03*	11.61	0.39	55.6	BP=0.001 p-v= 0.97	W=0.93 p-v=0.13	0.87	0.84
	Intercept	1.33	4.611								
Concrete and coating stage LM8	Area	0.0003	3.51	3.82E-03*	12.34	0.49	26.9	BP=0.08 p-v= 0.78	W=0.88 p-v=0.05	0.60	0.49
	Intercept	1.51	8.30								
Structural masonry GENERAL LM9	Area	0.001	4.970	1.68E-04*	24.7	0.62	42.0	BP=2.80 p-v=0.09	W=0.99 p-v=0.58	1.05	1.01
	Intercept	1.02	3.169								

Regarding freight vehicle production, area is the exploratory variable in all validated models. Also, it was possible to obtain a general model for structural masonry buildings under construction (LM9).

#### 4.2.2. Linear generalised freight vehicle generation models

We developed 80 models using GLM – 40 for attraction, 40 for production. From that, only five were validated by the statistical tests and are presented in Table 6.

Considering the Akaike information criterion (AIC), the generalised linear models do not present better predictive power when compared to linear models. For example, model GLM1 has the same variables as LM1 and the AIC (GLM1) is 78.07, while the AIC (LM1) is 53.4. Comparing these results, we concluded that LM1 is more suitable to estimate the number of freight vehicles attracted by buildings under construction. Similar results were obtained comparing the models GLM2 and LM4. Also, three models (GLM3, GLM4 and GLM5) were validated using the

independent variable *employee*, while we did not obtain a linear regression model using this variable. Also, the values of  $c$ -hat do not indicate the dispersion of the data, being smaller than one for all the models.

**Table 6** – Linear generalised freight vehicle generation models ( $p$ -v =  $p$ -value)

Type/stage and model ID	Variable and intercept	Value estimated	z-value	AIC	AICc	c-hat	Chi-squared test	LOOCV	RMSE
<b>Attraction (*99% confidence level; **95% confidence level)</b>									
Concrete and structure stage GLM1	Area	0.0001	3.25*	77.36	78.07	0.22	$\chi^2=3.96$ $p$ -v=1.0	7.71	1.01
	Intercept	1.35	9.40*						
Concrete and coating stage GLM2	Area	0.0002	2.68*	52.87	53.87	0.18	$\chi^2=2.34$ $p$ -v=1.0	1.18	0.70
	Intercept	0.99	5.30*						
GENERAL GLM3	Emp	0.06	7.51*	367.68	367.80	0.27	$\chi^2=27.57$ $p$ -v=1.0	1.24	1.08
	Intercept	0.56	4.24*						
Concrete GLM4	Emp	0.06	7.14*	291.46	291.61	0.27	$\chi^2=22.10$ $p$ -v=1.0	1.28	1.09
	Intercept	0.53	3.61*						
<b>Production (*99% confidence level; **95% confidence level)</b>									
Concrete GLM5	Emp	0.08	5.62*	239.62	239.77	0.60	$\chi^2=48.69$ $p$ -v=1.0	0.99	0.97
	Intercept	-0.56	2.40**						

### 4.3. Discussion of results

We obtained 15 valid models with the data obtained for Belo Horizonte, with accuracy and predictive capacity, including a freight vehicle attracted model (LM6) regardless of the construction type and the construction stage. Thus, it was possible to estimate FTGM in BUC.

In general, the models using number of units or floors as explanatory variables obtained the worst adjustment. These variables could influence the number of trips because they are related to the volume of goods necessary to build each floor. However, they did not contribute to explaining the phenomenon analysed in this paper.

Regarding the number of employees, three (from 15) models were estimated using this explanatory variable (GLM3, GLM4 and GLM5), all using the GLM technique. This result brings a reflection about the location of BUC and the technology used in this construction. Analysing the data, on average, the BUC have 13.16 employees (standard deviation = 3.05; minimum = 4; maximum = 38). Thus, it is possible to assume that with the advent of technology, fewer employees are necessary. Thus, this variable needs to be used with parsimony to explain the FTGM in BUC.

Regarding the techniques, more models were obtained using linear regression than GLM. Thus, although the literature suggests the use of GLM for counting data, the results demonstrate that it is possible to obtain linear models that meet linearity assumptions with accuracy and with predictive capacity. Also, the results prove that the FTGM in BUC is a linear phenomenon.

Also, the results show that an analysis focused on the coefficient of determination ( $R^2$ ) is not enough to make a conclusion about the efficiency, accuracy and predictive capacity of the model. As an example, if the accuracy of the models presented in this paper was evaluated looking

only at  $R^2$ , 30% (equivalent to 3 models) of them were discarded. However, despite the low coefficient of determination, the models explain the phenomenon under study, taking into account the linearity assumptions. Consequently, other tests are essential to making a conclusion about the accuracy and predictive capacity of the model, such as those used in this study. Therefore, linearity assumptions are fundamental to evaluating the accuracy of the model. In the same manner, cross-validation allows for the identification of the predictive capacity of the model.

## 5. CONCLUSION

This paper presented freight vehicle generation models for buildings under construction. Data from BUC in Belo Horizonte were considered for the modelling. Models were estimated using linear regression and GLM.

Data were obtained from interviews in 105 BUC. We estimated 190 equations (110 using linear regression and 80 using GLM) and obtained 15 models that were statistically valid. Considering the Akaike information criterion (AIC), the models obtained by linear regression were more suitable to estimate the freight vehicle generations. Also, FTGM to BUC is a linear phenomenon in the Belo Horizonte case.

Among the variables used to explain the phenomenon in BUC, the best results were obtained using the area and employees. The number of floors and units (related to the characteristics of the sector) did not present statistical significance in the estimations.

Finally, linearity assumptions and cross-validation reduce the number of models obtained. However, the models present accuracy and predictive capacity. Thus, it is possible to conclude that the linearity assumptions are fundamental to evaluating the accuracy of the model. Also, it is important to verify the predictive capacity of the model using the cross-validation test.

For future studies, it is suggested that construction time be included as an explanatory variable in the modelling. This variable is influenced by the stage of construction and influences the storage area and the number of trips. Also, it is suggested to include the number of units and floors as explanatory variables to confirm that they did not contribute to freight trip generation in BUC. Still, it is suggested to develop similar analyses in other cities to compare the results, including more extensive data collection efforts. Another suggestion is to carry out temporal data collection to analyse the influence of time on freight trip demand. This analysis could be interesting when associated with the explanatory variables: while the area of business is one variable that does not change over time, it is possible that the number of employees could change over the time and thus could influence the results of the modelling.

## ACKNOWLEDGEMENT

The authors would like to thank the Brazilian National Council for Scientific and Technological Development (CNPq) for financial support and encouragement in this research.

## REFERENCES

- Aditjandra, P.T., Galatioto, F., Bell, M.C., and 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*, v.16, n.1, p.4-22. <https://doi.org/10.18757/ejtr.2016.16.1.3110>
- Akaike, H. (1974) *A new look at the statistical model identification*. *IEEE Transactions on Automatic Control*, v.19, n.6, p.716-723. <https://doi.org/10.1109/TAC.1974.1100705>
- Alho, A.R. and J.A. Silva (2014a) Freight-Trip Generation Model: Predicting Urban Freight Weekly Parking Demand from Retail Establishment Characteristics. *Transportation Research Record: Journal of the Transportation Research Board*, v.2411, p.45-54. <https://doi.org/10.3141/2411-06>

- Alho, A.R. and J.A. Silva (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, US.
- Alho, A.R. and J.A. Silva (2015) Lisbon's Establishment-based Freight Survey: revealing retail establishments' characteristics, goods ordering and delivery processes. *Eur. Transp. Res. Rev.*, v.7, n.16, p. 1-17. <http://dx.doi.org/10.1007/s12544-015-0163-7>
- Alho, A.R. and J.A. Silva (2017) Modeling retail establishments' freight trip generation: a comparison of methodologies to predict total weekly deliveries. *Transportation*, v.44, p.1195-1212. <https://doi.org/10.1007/s11116-015-9670-6>
- Arlot, S. and A. Celise (2010) A survey of cross-validation procedures for model selection. *Statistics Survey*, v.10, p.40-79. <http://dx.doi.org/10.1214/09-ss054>
- Belo Horizonte (2018) Relatório de projetos de edificações aprovados. <https://prefeitura.pbh.gov.br/politica-urbana/regulacao-urbana/relatorios/edificacoes>
- Boarnet, M.G., Hong, A., and R. Santiago-Bartolomei (2017) Urban spatial structure, employment subcenters, and freight travel. *Journal of Transport Geography*, v.60, p.267-276. <https://doi.org/10.1016/j.jtrangeo.2017.03.007>
- Breusch, T.S. and A.R. Pagan (1979) A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, v.47, n.5, p.1287-1294. <http://dx.doi.org/10.2307/1911963>
- Burnham, K.P. and D.R. Anderson (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach*. Springer: New York.
- Campbell, S., Jaller, M., Sánchez-Díaz, I., Holguín-Veras, J., and C.T. Lawson (2012) Comparison between industrial classification systems in freight trip generation (FTG) modeling. In 91<sup>st</sup> TRB Annual Meeting, Washington, US.
- Campos, V.B.G. and I.C.B. Melo (2004) Análisis de la demanda de transporte de carga en area urbana bajo el punto de vista de la production y de la atracción de viajes. In VI Congreso de Ingenieria del Transporte, Zaragoza, Spain.
- Campos, V.B.G., D'Agosto, M.A., Cruz, I., and C.D. Souza (2012) Polos Geradores de Viagens de Carga. In: Portugal, L. S. (Org.) *Polos Geradores de Viagens Orientados a Qualidade de Vida e Ambiental: Modelos e Taxas de Geração de Viagens*. Interciência: Rio de Janeiro.
- Canty, A. and B. Ripley (2017) boot: Bootstrap Functions. Disponível em <https://CRAN.R-project.org/package=boot>
- CBIC (2018) PIB Brasil e Construção Civil. Câmara Brasileira da Indústria da Construção. Data from August of 2018. <http://www.cbicdados.com.br/menu/pib-e-investimento/pib-brasil-e-construcao-civil>
- Comi, A., Site, P.D., Fillipi, F., and A. Nuzzolo (2012) Urban Freight Transport Demand Modelling: a State of the Art. *European Transport \ Trasporti Europei*, v.51, n.7.
- Davison, A.C. (2001) Biometrika Centenary: Theory and General Methodology. *Biometrika*, v.88, p.13-52. <https://doi.org/10.1093/biomet/88.1.13>
- EuroStat (2018) Construction production (volume) index overview. Data from March 2018. [https://ec.europa.eu/eurostat/statistics-explained/index.php/Construction\\_production\\_\(volume\)\\_index\\_overview](https://ec.europa.eu/eurostat/statistics-explained/index.php/Construction_production_(volume)_index_overview)
- Ferreira, B.L.G. and M.A.V. Silva (2016) Truck trips in urban areas and its relation to socioeconomic variables. *Revista Gestão da Produção Operações e Sistemas*, v.11, n.4, p.197-212. <https://doi.org/10.15675/gepros.v11i4.1566>
- Frasco, M. (2018) Package 'Metrics': Evaluation Metrics for Machine Learning. <https://cran.r-project.org/web/packages/Metrics/Metrics.pdf>
- Gasparini, A., Campos, V.B.G., and M.A. D'Agosto (2010) Modelos para estimativa da demanda de viagens de veículos de carga para supermercados e shopping-centers *Transportes*, v.18, n.1, p.58-65. <https://doi.org/10.14295/transportes.v18i1.388>
- Gonçalves, F.S., Lemos, D.S.C.P.S., Kneib, E.C., and L.S. Portugal (2012) Caracterização dos Polos Geradores de Viagens. In: Portugal, L. S. (Org.) *Polos Geradores de Viagens Orientados a Qualidade de Vida e Ambiental: Modelos e Taxas de Geração de Viagens*. Interciência: Rio de Janeiro, p. 3-41.
- González-Feliu, J. and C. Peris-Plab (2017) Impacts of retailing attractiveness on freight and shopping trip attraction rates. *Research in Transportation Business & Management*, v.24, p.49-58. <https://doi.org/10.1016/j.rtbm.2017.07.004>
- González-Feliu, J., Sánchez-Díaz, I., and C. Ambrosini (2015) Aggregation level, variability and linear hypotheses for urban delivery generation models. In Transportation Research Board 95th Annual Meeting, Washington.
- Günay, G., Ergün, G., and I. Gökaşar (2016) Conditional Freight Trip Generation modeling. *Journal of Transport Geography*, v.54, p.102-111. <https://doi.org/10.1016/j.jtrangeo.2016.05.013>
- Holguín-Veras, J., Sánchez-Díaz, I., Lawson, C.T., Jaller, M., Campbell, S., Levinson, H.S., and H.-S. Shin (2013) Transferability of Freight Trip Generation Models. *Transportation Research Record: Journal of the Transportation Research Board*, v.2379, p.1-8. <https://doi.org/10.3141/2379-01>
- Holguín-Veras, J., Jaller, M., Destro, L., Ban, S., Lawson, C., and 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*, v.2224, p.68-81. <https://doi.org/10.3141/2224-09>
- Hothorn, T., Zeileis, A., Farebrother, R.W., Cummins, C., Millo, G., and M. Mitchell (2018) lmtest: Testing Linear Regression Models. <https://CRAN.R-project.org/package=lmtest>
- Hurvich, C.M. and C.-L. Tsai (1995) Model Selection for Extended Quasi-Likelihood Models in Small Samples. *Biometrics*, v. 51, p. 1077-1084. <https://doi.org/10.2307/2533006>
- Hyndman, R.J. and A.B. Koehler (2006) Another look at measures of forecast accuracy. *International Journal of Forecasting*, v.22, n.4, p.679-688. <http://dx.doi.org/10.1016/j.ijforecast.2006.03.001>

- Iding, M.H.E., Meester, W.J., and L.A. Tavasszy (2002) Freight trip generation by firms. In 42nd European Congress of the Regional Science Association, Dortmund, Germany.
- Jaller, M., Sánchez-Díaz, I., and J. Holguín-Veras (2015b) Identifying Freight Intermediaries: Implications for Modeling of Freight Trip Generation. *Transportation Research Record: Journal of the Transportation Research Board*, v.2478, p.48-56. <https://doi.org/10.3141/2478-06>
- Jaller, M., Sánchez-Díaz, I., Holguín-Veras, J., and C.T. Lawson (2014) Area based freight trip generation models. In Transportation Research Board (TRB) 93rd Annual Meeting, Washington, US.
- Jaller, M., Wang, X., and J. Holguín-Veras (2015a) Large urban freight traffic generators: Opportunities for city logistics initiatives. *Journal of Transport and Land Use*, v.8, n.1, p.1-17. <http://dx.doi.org/10.5198/jtlu.2015.406>
- James, G., Witten, D., Hastie, T., and R. Tibshirani (2017) ISLR: Data for an Introduction to Statistical Learning with Applications in R. <https://CRAN.R-project.org/package=ISLR>
- Kulpa, T. (2014) Freight truck trip generation modeling at regional level. *Procedia – Social and Behavioral Sciences*, v. 111, p. 197-202. <https://doi.org/10.1016/j.sbspro.2014.01.052>
- Lawson, C.T., Holguín-Veras, J., Sánchez-Díaz, I., Jaller, M., Campbell, S., and E.L. Powers (2012) Powers Estimated Generation of Freight Trips Based on Land Use. *Transportation Research Record: Journal of the Transportation Research Board*, v.2269, p.65-72. <http://dx.doi.org/10.3141/2269-08>
- Lim, H., Chin, S.M., Hwang, H.-L., and L.D. Han (2017) Incorporating Information Complexity into Regression-Based Freight Generation Model Selection. In Transportation Research Board (TRB) 97th Annual Meeting, Washington, US.
- Maia, A.G. (2017) *Econometria: conceitos e aplicações*. Editora Saint Paul: São Paulo.
- McCullagh, P. and J.A. Nelder (1989) *Generalized Linear Models*. Chapman and Hall: London; New York.
- Nuzzolo, A., Crisalli, U., and A. Comi. (2012) A delivery approach modelling for urban freight restocking. *Journal of Civil Engineering and Architecture*, v.6, n.3, p.251-267. <http://dx.doi.org/10.17265/1934-7359/2012.03.001>
- Oliveira, L.K., Silva, A.S.S., Tolentino, A.C.M., Rachid, A.H.F.S., Bertoncini, B.V., Lima, J.H., Bracarense, L.S.F.P., and R.L.M. Oliveira (2018) Análise comparativa de técnicas de modelagem da geração de viagens de carga. Anais do XXXII Congresso Nacional de Ensino e Pesquisa em Transportes, Gramado, Brazil.
- Oliveira, L.K., Nóbrega, R.A.A., Ebias, D.G., and B.G.S. Corrêa (2017) Analysis of Freight Trip Generation Model for Food and Beverage in Belo Horizonte (Brazil). *Region*, v.4, n.1, p.17-30. <http://dx.doi.org/10.18335/region.v4i1.102>
- Oliveira, L.K., Oliveira, R.L.M., Ramos, C.M.F., and D.G. Ebias (2016) Modelo de geração de viagens de carga em áreas urbanas: um estudo para bares, restaurantes e supermercados. *Transportes*, v.2, n.24, p.53-67. <https://doi.org/10.14295/transportes.v24i2.1058>
- Reis, P.I.A.M., Bracarense, L.S.F.P., and L.K. Oliveira (2018) Geração de viagens de veículos de carga por supermercados: Estudo aplicado a Palmas-TO. Anais do XXXII Congresso Nacional de Pesquisa e Ensino em Transportes. ANPET, Gramado, Brazil.
- Rudani, H.R., Patel, V.F., and H.K. Dave (2017) Development of trip generation model for goods movement of Gozaria and Visnagar G.I.D.C. *International Journal of Advance Engineering and Research Development*, v.4, n.4, p.923-926
- Sánchez-Díaz, I. (2017) Modeling urban freight generation: A study of commercial establishments' freight needs. *Transportation Research Part A*, v.102, p.3-17. <https://doi.org/10.1016/j.tra.2016.06.035>
- Silva, M.R. and J. Waisman (2007) Cargas urbanas: estudo exploratório sobre a geração de viagens de caminhões em bares e restaurantes. Anais do 16º Congresso Brasileiro de Transporte e Trânsito, Maceió, Brazil.
- Statista (2018) Value added of the construction industry as a share of gross domestic product in the U.S. from 2007 to 2017. <https://www.statista.com/topics/974/construction/>
- Stevenson, W. J. (1981) *Estatística aplicada à administração*. Editora Harbra: São Paulo.
- Washington, S., Karlaftis, F., and M. Mannering (2011) *Statistical and econometric methods for transportation data analysis*. Chapman & Hall/CRC: Boca Raton.
- Wee, B.V. and D. Banister (2016) How to Write a Literature Review Paper? *Transport Reviews*, v.36, n.2, p.278-288. <https://dx.doi.org/10.1080/01441647.2015.1065456>
- Zhang, D. (2018) Package 'rsq': R-squared and related measures. <https://cran.rproject.org/web/packages/rsq/rsq.pdf>