

UNIVERSIDADE FEDERAL DE MINAS GERAIS

CURSO DE MESTRADO EM GEOTECNIA E TRANSPORTES

**ESTIMATIVA DE GERAÇÃO DE VIAGENS DE
VEÍCULOS DE CARGA EM ÁREAS URBANAS
UTILIZANDO
MODELAGEM GEOGRÁFICA E MODELO
LINEAR GENERALIZADO**

Victor Lima Migliorini

Belo Horizonte

2018

Victor Lima Migliorini

**ESTIMATIVA DE GERAÇÃO DE VIAGENS DE
VEÍCULOS DE CARGA EM ÁREAS URBANAS
UTILIZANDO
MODELAGEM GEOGRÁFICA E MODELO LINEAR
GENERALIZADO**

Dissertação apresentada ao Curso de Mestrado em Geotecnia e Transportes da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Geotecnia e Transportes.

Área de concentração: Transportes

Orientador: Rodrigo Affonso de Albuquerque Nóbrega

Coorientadora: Leise Kelli de Oliveira

Belo Horizonte

Escola de Engenharia da UFMG

2018

M634e

Migliorini, Victor Lima.

Estimativa de geração de viagens de veículos de carga em áreas urbanas utilizando modelagem geográfica e modelo linear generalizado [manuscrito] / Victor Lima Migliorini. – 2018.
xii, 110 f., enc.: il.

Orientadora: Rodrigo Affonso de Albuquerque Nóbrega.
Coorientadora: Leise Kelli de Oliveira.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Apêndices: f. 85-110.

Bibliografia: f. 77-84.

1. Transportes - Teses. 2. Logística empresarial - Teses.
3. Modelos lineares (Estatística) - Teses. 4. Transporte de mercadorias - Teses. I. Nóbrega, Rodrigo Affonso de Albuquerque. I. Oliveira, Leise Kelli de. III. Universidade Federal de Minas Gerais. Escola de Engenharia.
IV. Título.

CDU: 656(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOTECNIA E TRANSPORTES

UFMG

FOLHA DE APROVAÇÃO

ESTIMATIVA DE GERAÇÃO DE VIAGENS DE VEÍCULOS DE CARGA EM ÁREAS URBANAS UTILIZANDO MODELAGEM GEOGRÁFICA E MODELO LINEAR GENERALIZADO

VICTOR LIMA MIGLIORINI

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em GEOTECNIA E TRANSPORTES, como requisito para obtenção do grau de Mestre em GEOTECNIA E TRANSPORTES, área de concentração TRANSPORTES.


Aprovada em 08 de fevereiro de 2018, pela banca constituída pelos membros:


Prof. Rodrigo Affonso de Albuquerque Nobrega - Orientador

UFMG


Prof(a). Leise Kelli de Oliveira

UFMG


Prof(a). Cira Souza Pitombo

USP


Prof(a). Renata Lucia Magalhães de Oliveira

CEFET-MG

Belo Horizonte, 8 de fevereiro de 2018.

AGRADECIMENTOS

Primeiramente, agradeço a Deus, por ter me guiado por toda a minha vida.

Aos meus pais Iva e José Carlos, por todo apoio, carinho, dedicação e por não pouparem esforços para tornar os meus sonhos possíveis. Sem vocês nada disso seria possível.

Agradeço a toda a minha família, em especial aos meus irmãos Fabrício e Mariana por sempre estarem ao meu lado me apoiando. Agradeço também aos meus amigos pelo companheirismo.

Agradeço à minha amada Jéssica, por estar sempre ao meu lado durante todo este período, me incentivando, me animando e por tornar os meus dias mais leves.

Ao meu orientador Professor Rodrigo Affonso de Albuquerque Nóbrega e à minha coorientadora Leise Kelli de Oliveira, pela confiança, pelos ensinamentos e contribuições dadas ao longo deste trabalho.

À Prefeitura de Belo Horizonte, por ter cedido uma parte dos dados utilizados nesta pesquisa e aos meus colegas do Programa de Mestrado em Engenharia de Transporte, por terem desenvolvido pesquisas sem as quais este trabalho não poderia ter sido realizado.

À UFMG e todos os mestres e colaboradores que contribuíram na minha formação e por terem transmitido os seus conhecimentos durante todo o período da minha graduação e do meu mestrado.

À Locale Trânsito e Transporte, por ter aberto as portas da área de transportes para mim, pela oportunidade de aprendizado e crescimento na área e por ter possibilitado a realização deste trabalho.

*"Nunca ande pelo caminho traçado, pois ele conduz somente até onde os outros
já foram."*

(Alexander Graham Bell)

RESUMO

MIGLIORINI, Victor Lima. **Estimativa de geração de viagens de veículos de carga em áreas urbanas utilizando modelagem geográfica e modelo linear generalizado**. 110f. Dissertação (Mestrado em Geotecnia e Transportes) – Universidade Federal de Minas Gerais, Belo Horizonte, 2018.

O tema logística urbana tem ganhado importância na gestão de transportes e na definição de políticas públicas nas médias e grandes cidades no Brasil e no mundo. Estimar com precisão o número de viagens de veículos de carga que atendem os estabelecimentos comerciais é uma das necessidades apontadas pelos planejadores de transporte. O processo usual para estimar o número de caminhões gerados emprega a técnica de regressão linear que, embora adequado em alguns casos, apresenta uma limitação quanto à heterogeneidade do contexto urbano, bem como o comportamento de não-linearidade das variáveis envolvidas no problema. Além disso, pesquisas recentes sobre modelos lineares generalizados, aplicadas em outras áreas de transporte, apresentaram resultados promissores, porém, ainda não foram avaliadas no contexto das viagens de veículos de carga. Neste sentido, este estudo apresenta modelos para estimar a geração de viagens de veículos de carga de supermercados em áreas urbanas, comparando os resultados obtidos por meio de modelos lineares, lineares generalizados e geograficamente ponderados. Por meio de dados de geração de viagens de veículos de carga para supermercados de Belo Horizonte e de dados socioeconômicos, foram desenvolvidos cenários utilizando os três métodos de modelagem. A comparação dos resultados mostrou que o modelo linear generalizado e o modelo geograficamente ponderado apresentam ganhos relativos quando comparados ao modelo linear tradicional. Estatisticamente, o modelo linear generalizado apresentou resultados ligeiramente melhores do que o modelo geograficamente ponderado. Porém, ao contrário do modelo linear generalizado que fornece um número único de geração, o modelo geograficamente ponderado produz uma superfície contínua com estimativa local de viagens de veículos de carga por *pixel*. Isso pode mudar o paradigma e inovar a forma como a geração de veículos de carga é modelada. Por fim, por meio da análise das variáveis independentes utilizadas nos modelos, constatou-se que quanto maior a renda e a densidade populacional, de empregos e de pessoas fazem compras, menor a taxa de geração de viagens de veículos de carga na região.

Palavras-chave: geração de viagens de veículos de carga, regressão linear generalizada, regressão linear generalizada geograficamente ponderada

ABSTRACT

MIGLIORINI, Victor Lima. **Estimation of freight trip generation in urban areas using geographic modeling and generalized linear model.** 110f. Dissertation (Master in Geotechnics and Transportation) – Federal University of Minas Gerais, Belo Horizonte, 2018.

The urban freight transportation has gained importance in transport management and public policies in medium and large cities in Brazil and the world. Accurately estimating the number freight trips to supply retail stores is one of the needs highlighted by transport planners. The traditional process for estimating the number of truck-trips employs linear regression. Although suitable, they present limitation regarding the heterogeneity of the urban context as well as the nonlinearity of the variables involved in the problem. Recent investigations on generalized linear models applied other areas of transportation have delivered promising results, therefore not yet evaluated for urban freight. This dissertation introduces models to estimate freight-trip generation to supermarket in urban areas, comparing the results obtained through linear model, generalized linear model and geographically weighted model. Using data from freight-trip generation to markets and supermarkets in Belo Horizonte and socioeconomic data, scenarios were developed using the three models. Findings showed generalized linear models presenting relative gains when compared to the traditional linear models. The geographically weighted model also presented better results than the linear regression. Statistically, the generalized linear model presented slightly better results than the geographically weighted model. However, instead of a static number, the geographically weighted model outputs a continuous surface with local estimative of freight-trip per pixel. This can change the paradigm and innovate the way freight-trip generation is modeled. Finally, through the analysis of the independent variables used in the models, it was found that the higher the average income and the population density, density of jobs and density of customers, lower is freight-trip generation in the region.

Keywords: trip generation model, generalized linear regression, geographic weighted generalized linear model

SUMÁRIO

1	INTRODUÇÃO.....	1
1.1	PROBLEMA DE PESQUISA E HIPÓTESES DO ESTUDO	2
1.2	OBJETIVOS DO TRABALHO.....	3
1.3	JUSTIFICATIVA DO ESTUDO	3
1.4	ESTRUTURA DA DISSERTAÇÃO	4
2	REVISÃO BIBLIOGRÁFICA.....	6
2.1	MODELOS DE GERAÇÃO DE VIAGENS DE VEÍCULOS DE CARGA URBANA	6
2.1.1	<i>Literatura Internacional</i>	9
2.1.2	<i>Literatura Nacional</i>	14
2.2	ANÁLISE ESPACIAL PARA A GERAÇÃO DE VIAGENS	17
2.3	SÍNTESE DA REVISÃO DA LITERATURA	19
3	ÁREA DE ESTUDO, SETOR ECONÔMICO E OBTENÇÃO DE DADOS	23
3.1	DEFINIÇÃO DA ÁREA DE ESTUDO.....	23
3.2	DEFINIÇÃO DO SETOR ECONÔMICO.....	26
3.3	OBTENÇÃO DOS DADOS.....	26
4	ABORDAGEM METODOLÓGICA	30
4.1	VARIÁVEIS DO MODELO	31
4.1.1	<i>Análise de correlação entre variáveis</i>	32
4.1.2	<i>Análise da autocorrelação espacial</i>	32
4.2	MODELAGEM DA GERAÇÃO DE VIAGENS DE VEÍCULOS DE CARGA URBANA	34
4.2.1	<i>Modelo de regressão linear</i>	34
4.2.2	<i>Modelo linear generalizado</i>	36
4.2.3	<i>Modelo linear generalizado geograficamente ponderado</i>	38
4.3	VALIDAÇÃO CRUZADA DOS MODELOS.....	41

4.4	COMPARAÇÃO DOS MODELOS	41
5	RESULTADOS E DISCUSSÕES	42
5.1	ANÁLISE DE CORRELAÇÃO ENTRE VARIÁVEIS	42
5.2	ANÁLISE DA AUTOCORRELAÇÃO ESPACIAL.....	43
5.3	DADOS DO MODELO	44
5.4	MODELO LINEAR.....	45
5.5	MODELO LINEAR GENERALIZADO	54
5.6	MODELO LINEAR GENERALIZADO GEOGRAFICAMENTE PONDERADO.....	62
5.7	ANÁLISE COMPARATIVA DOS MODELOS	69
5.8	ANÁLISE DAS VARIÁVEIS INDEPENDENTES EM RELAÇÃO AOS MODELOS OBTIDOS.....	72
6	CONSIDERAÇÕES FINAIS	74
6.1	RECOMENDAÇÕES.....	75
	REFERÊNCIAS	77
	APÊNDICE A – ANÁLISE DE RESÍDUOS DOS MODELOS LINEARES	85
	APÊNDICE B – ANÁLISE DE RESÍDUOS DOS MODELOS LINEARES GENERALIZADOS.....	98

LISTA DE FIGURAS

Figura 1: Área de Estudo.....	24
Figura 2: Principais corredores de acesso à região central (BHTRANS, 2013).....	25
Figura 3: Supermercados e hipermercados no município de Belo Horizonte.	28
Figura 4: Fluxograma da abordagem metodológica do trabalho.	30
Figura 5: Teste da pseudo-significância – Índice I de Moran (ESRI, 2017).....	34
Figura 6: Estabelecimentos com respostas validadas.	45
Figura 7: Resíduos versus valores ajustados – LM 01.	52
Figura 8: Resíduos versus valores ajustados – LM 18.	52
Figura 9: Gráfico normal de probabilidades – LM 01.....	52
Figura 10: Gráfico normal de probabilidades – LM 18.....	52
Figura 11: Resíduos versus valores ajustados – GLM 01.	61
Figura 12: Resíduos versus valores ajustados – GLM 18.	61
Figura 13: Gráfico normal de probabilidades – GLM 01.....	61
Figura 14: Gráfico normal de probabilidades – GLM 18.....	61
Figura 15: <i>Raster</i> do valor do intercepto – GWGLM 11.	66
Figura 16: <i>Raster</i> do coeficiente relativo à variável Área do estabelecimento – GWGLM 11.	66
Figura 17: <i>Raster</i> do coeficiente relativo à variável Densidade emprego – GWGLM 11.	66
Figura 18: <i>Raster</i> do coeficiente relativo à variável Renda média – GWGLM 11.	66
Figura 19: Geração de viagens de veículos de carga – Supermercados com 500m ²	67
Figura 20: Geração de viagens de veículos de carga – Supermercados com 1000m ²	67

Figura 21: Geração de viagens de veículos de carga – Supermercados com 1500m²..... 68

Figura 22: Geração de viagens de veículos de carga – Supermercados com 2000m²..... 68

LISTA DE TABELAS

Tabela 1: Variáveis independentes utilizadas nos modelos na literatura (JALLER <i>et al.</i> , 2013).	21
Tabela 2: Variáveis independentes e sua categoria de classificação utilizados nos modelos identificados na literatura (FERREIRA E SILVA, 2015; JALLER <i>et al.</i> , 2013).....	22
Tabela 3: Restrição de tráfego de veículos pesados (BHTRANS, 2013).....	25
Tabela 4: Resumo das informações solicitadas aos entrevistados (RAMOS, 2015).....	29
Tabela 5: Variáveis independentes selecionadas.....	31
Tabela 6: Matriz de correlação entre variáveis dependente e independentes.....	42
Tabela 7: Índices de autocorrelação espacial (I de Moran Global) das variáveis dependente e independentes.	44
Tabela 8: Modelos de geração de viagens de veículos de carga utilizando regressão linear. ..	46
Tabela 9: Modelos de geração de viagens de veículos de carga utilizando modelo linear generalizado.....	55
Tabela 10: Modelos de geração de viagens de veículos de carga utilizando GWGLM.....	63
Tabela 11: Análise comparativa do RMSD.....	70

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

ABRAS	Associação Brasileira de Supermercados
AH	Área Homogênea
AICc	<i>Akaike Information Criterion Correction</i> – Critério de Informação Akaike Corrigido
ANN	<i>Artificial neural networks</i> – Rede Neural Artificial
<i>Bandwidth</i>	Largura de banda
BHTRANS	Empresa de Transportes e Trânsito de Belo Horizonte
<i>Tractor-trailer</i>	Carreta
<i>Checkouts</i>	Caixas de supermercados
CMC	Cadastro Municipal de Contribuintes de Tributos Mobiliários
CNAE	Classificação Nacional de Atividades Econômicas
CNPJ	Cadastro Nacional de Pessoa Jurídica
<i>Cross validation</i>	Validação cruzada
ESDA	<i>Exploratory Spatial Data Analysis</i> – Análise Exploratória de Dados Espaciais
ESRI	<i>Environmental Systems Research Institute</i> – Instituto de Pesquisa de Sistemas Ambientais
GIS	<i>Geographic Information System</i> – Sistema de Informação Geográfico
GWGLM	<i>Geographically Weighted Generalised Linear Model</i> – Modelo Linear Generalizado Ponderado Geograficamente
GWR	<i>Geographically Weighted Regression</i> – Regressão geograficamente ponderada

hab./km ²	Habitantes por quilômetro quadrado
IBGE	Instituto Brasileiro de Geografia e Estatística
IDW	<i>Inverse Distance Weighting</i> – Ponderação pelo Inverso da Distância
ISSQN	Imposto sobre Serviços de Qualquer Natureza
ITE	<i>Institute of Transportation Engineers</i> – Instituto dos Engenheiros de Transportes
km	Quilômetros
km ²	Quilômetros quadrados
LM	<i>Linear model</i> – Modelo linear
MCA	<i>Multiple classification analysis</i> – Análise de classificação múltipla
MMQ	Método dos Mínimos Quadrados
NCFRP	<i>Freight Trip Generation and Land Use Report</i> – Relatório de Geração de Viagens de Frete e Uso da Terra
OD	Origem e Destino
OLS	<i>Ordinary Least Squares</i> – Método dos Mínimos Quadrados
PCA	<i>Principal Componentes Analysys</i> – Análise de Componentes Principais
PGV Carga	Polos Geradores de Viagens de veículos de carga
PIB	Produto Interno Bruto
R ²	Coefficiente de determinação
RMBH	Região Metropolitana de Belo Horizonte
RMRJ	Região Metropolitana de São Paulo
RMSD	<i>Root-mean-square deviation</i> – Média dos quadrados dos erros

RMSP	Região Metropolitana do Rio de Janeiro
SEGEM	Secretaria de Gestão Metropolitana
SETOP	Secretaria de Transportes e Obras Públicas
SIC	<i>Standard Industrial Classification</i> – Classificação Industrial Padrão
TFEP	Taxa de Fiscalização de Engenhos de Publicidade
TFLF	Taxa de Fiscalização, Localização e Funcionamento
TFS	Taxa de Fiscalização Sanitária
UFMG	Universidade Federal de Minas Gerais
VI	Variável independente
VUC	Veículo urbano de carga

1 INTRODUÇÃO

Um dos maiores problemas atualmente enfrentado pelas grandes cidades é o excesso de veículos que circulam nas vias, principalmente nos horários de pico. Os efeitos negativos do excesso de veículos, crescente na grande maioria das cidades, têm impulsionado nos últimos anos a necessidade de mudanças em políticas públicas, em especial políticas relacionadas a transportes. A gestão pública municipal tem incluído diretrizes voltadas para a luta contra o congestionamento, a gestão do espaço público e também a redução dos impactos ambientais, visando economia de energia e redução de emissões de gases de efeito estufa (PATIER e ROUTHIER, 2008).

Contudo, o congestionamento não é apenas um problema para os gestores públicos, os condutores e usuários dos veículos, mas também para embarcadores e transportadores de carga, cuja receita depende, dentre outros fatores, da eficiência logística do sistema de transportes. Para Iding *et al.* (2002), a malha viária em torno de instalações industriais, em geral, não tem capacidade suficiente para o tráfego de caminhões, resultando em congestionamento nessas áreas. Se, por um lado, a maior parte da infraestrutura viária urbana e interurbana não tem sido adequada ao elevado volume de veículos em horários críticos, por outro lado a solução do problema do congestionamento apenas pela perspectiva da reestruturação viária resulta em excessivos gastos com infraestrutura, demandando alto investimento e incentivando ainda mais o aumento da frota.

Para os embarcadores e transportadores de carga, bem como para os estabelecimentos receptores dessas cargas, a ineficiência logística consequente do excesso de veículos e saturação das vias afeta diretamente a estabilidade logística e econômica do sistema, comprometendo fortemente a distribuição urbana de mercadorias.

Conceitualmente, segundo Dablanc (2007), a distribuição urbana de mercadorias pode ser entendida como o conjunto de deslocamentos de produtos que entram, atravessam e/ou saem das áreas urbanas. Tais produtos podem ser de diversas categorias, como, por exemplo, mercadorias para o varejo, serviços de correios, resíduos urbanos, transporte de valores, entre outros. Além disso, não obstante aos fluxos de mercadorias, as atividades de carga e descarga e armazenamento de produtos devem ser consideradas no planejamento de circulação dessas cargas (DABLANC, 2007).

Um aspecto importante da distribuição urbana de mercadorias é a geração de viagens de veículos de carga, que ocorrem em pontos denominados Polos Geradores de Viagens de veículos de carga – PGV Carga (GASPARINI *et al.*, 2010, OLIVEIRA *et al.*, 2016). Os impactos gerados no sistema viário decorrentes da instalação destes polos devem ser analisados. Essa análise, usualmente, é inicialmente realizada por meio de modelos que estimam a movimentação local de mercadorias por meio de dados de caracterização do empreendimento (GASPARINI *et al.*, 2010). Todavia, esses modelos mais usuais na literatura são analíticos, sem ou com limitado apoio de informações de contexto geográfico.

Apesar da importância da distribuição urbana de mercadorias, responsável pela geração de empregos e serviços para economia das cidades, essa temática tem sido historicamente negligenciada pelas autoridades, pelas pesquisas de transporte e pelo planejamento estratégico das cidades (DABLANC, 2009). Neste contexto, encontrar métodos eficientes para apoiar a formação de políticas públicas e a tomada de decisão no planejamento e na gestão dos sistemas de transportes é um dos papéis e desafios da academia. Tal papel é exercitado no presente trabalho, que emprega modelos lineares, lineares generalizados e de geoprocessamento para investigar a estimativa da geração de viagens de veículos de carga urbana.

Além da abordagem quantitativa baseada em modelos lineares generalizados, este estudo também leva em consideração a utilização de modelos espaciais. Segundo Costa *et al.* (2013), quando são incorporadas variáveis relacionadas ao espaço no modelo a ser desenvolvido, ampliam-se as possibilidades de análise melhorando a precisão do mesmo por meio da adequação a características locais.

1.1 Problema de Pesquisa e Hipóteses do Estudo

Um dos objetivos do planejamento de transportes, segundo Ribeiro (2012), é modelar a demanda, em viagens, para que o gestor de transportes identifique os movimentos atuais e possa estimar os futuros, que subsidiarão a tomada de decisão em relação às questões referentes à mobilidade urbana.

Desse modo, o problema de pesquisa desta dissertação é: “Como inferir a quantidade de movimentação de veículos de carga em áreas urbanas a partir de polos geradores de viagens de veículos de carga?”.

Para responder a esse problema de pesquisa, foram consideradas duas hipóteses. A primeira hipótese considera que é possível estimar a geração de viagens de veículos de carga em áreas urbanas por meio da utilização de um modelo de análise geográfica, baseado em dados secundários. Esses dados podem ser referentes a dados socioeconômicos, cadastro imobiliário, uso e ocupação do solo e índice de empregos existente na área de estudo. A segunda hipótese considera que modelos geograficamente ponderados apresentam resultados melhores que os modelos lineares generalizados que, por sua vez, apresentam melhores resultados que os modelos lineares.

1.2 Objetivos do trabalho

O objetivo principal deste estudo é desenvolver modelos para estimar a geração de viagens de veículos de carga em áreas urbanas utilizando modelagem geográfica e modelo linear generalizado, por meio de dados secundários.

Para tanto, são objetivos específicos:

- adquirir e tratar dados secundários para representar geograficamente um estudo quantitativo da geração de viagens urbanas de mercadorias;
- construir e analisar modelos de geração de viagens de veículos de carga utilizando dados amostrais;
- avaliar o potencial do uso de ferramentas geoespaciais para estimar a geração de viagens de veículos de carga considerando que o fenômeno não é independente no espaço;
- comparar os resultados da modelagem geográfica e do modelo linear generalizado com a regressão linear, descrevendo os benefícios e restrições do uso de modelos lineares generalizados e do uso da modelagem geográfica para a estimativa da geração de viagens de veículos de carga em áreas urbanas;

1.3 Justificativa do estudo

O desenvolvimento deste trabalho justifica-se primeiramente pela necessidade de elaboração de modelos que trabalhem com dados secundários, visando amenizar a escassez de dados primários para o planejamento do transporte urbano de mercadorias, bem como prover

subsídios para políticas públicas e ambientais eficientes em relação à Logística Urbana, conforme apontado por Stich *et al.* (2011) e Pereira e Oliveira (2014).

Em relação à modelagem da demanda da carga urbana, Ortúzar e Willumsen (1994), Iding *et al.*, (2012) e Novak *et al.* (2011) afirmam que, apesar do impacto substancial que o transporte de carga tem sobre a competitividade econômica, bem como as externalidades causadas pelos veículos de carga, os modelos para o planejamento do transporte de mercadorias são normalmente menos investigados que a modelagem da demanda por passageiros. Os modelos de previsão de demanda de carga urbana surgiram posteriormente a partir da adaptação dos modelos às necessidades particulares do contexto da carga urbana (IDING *et al.*, 2012).

Segundo Souza *et al.* (2010), é observada uma quantidade muito maior de modelos de geração de viagens internacionais do que nacionais, indicando que esta é uma área de pesquisa a ser explorada no Brasil, considerando as características específicas de cada cidade.

Em relação aos modelos de geração de viagens de veículos de carga, a metodologia mais utilizada é a regressão linear simples, que resultam em modelos que podem retornar bons resultados de predição e fornecer boas estimações (SOUZA *et al.* 2010, NOVAK *et al.*, 2011). Porém, segundo Turkman e Silva (2000), em casos nos quais a variável dependente do modelo a ser construído é uma contagem, formando um conjunto de dados discretos positivos, a utilização dos modelos lineares generalizados pode ser indicada, com possibilidades de melhoria nos ajustes dos modelos construídos.

Além disso, os modelos lineares largamente utilizados para modelagem de viagens de veículos de carga podem apresentar resíduos espacialmente correlacionados. Sendo assim, modelos de regressão espacial podem apresentar melhor ajuste do que modelos não espaciais e, portanto, eliminar problemas associados à autocorrelação espacial na estimativa de viagens.

Portanto, o método que será desenvolvido neste trabalho visa suprir, de forma indireta, a carência por informações sobre transporte de cargas, demanda, rotas e viagens, auxiliando o planejamento do transporte urbano de mercadorias.

1.4 Estrutura da dissertação

Este trabalho está assim organizado: o Capítulo 2 é apresentada uma revisão bibliográfica abordando modelos de geração de viagens de veículos de carga urbana publicados em

literatura nacional e internacional e os estudos que abordam a temática da análise espacial na previsão de viagens.

No Capítulo 3 é apresentada a área de estudo, a definição do setor econômico e as fontes de dados utilizadas ao longo do estudo. O Capítulo 4 discorre sobre os procedimentos adotados e etapas realizadas para o desenvolvimento do modelo de estimativa de viagens. No Capítulo 5 são apresentados os resultados encontrados pelos métodos empregados no decorrer do trabalho e a comparação entre eles, com intuito de encontrar o método que apresenta a melhor acurácia na estimativa de geração de viagens de carga.

No Capítulo 6 são apresentadas as considerações finais da dissertação e, em seguida, têm-se as referências bibliográficas citadas no decorrer do trabalho.

Por fim, nos Apêndices A e B são apresentadas as análises dos resíduos dos modelos desenvolvidos ao longo do estudo.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo é apresentada a revisão bibliográfica que buscou abordar os principais modelos de geração de viagens de veículos de carga urbana existentes na literatura, o emprego de análise espacial na geração de viagens, bem como identificar as principais lacunas referentes à modelagem de viagens de veículos de carga.

2.1 Modelos de Geração de Viagens de Veículos de Carga Urbana

Na distribuição urbana de mercadorias, foco do presente trabalho, os movimentos das mercadorias resultam de um conjunto de escolhas feitas por habitantes e clientes, que decidem onde fazer as compras e dos varejistas, atacadistas e distribuidores, que decidem a localização das lojas e escolhem a localização dos centros de distribuição e qual a forma de abastecimento dos clientes varejistas. Além disso, as autoridades públicas também exercem papel fundamental na movimentação de mercadorias, pois tomam medidas no intuito de minimizar os custos de transporte, de congestionamento e os custos externos – poluição e segurança viária (TANIGUCHI *et al.*, 2001)

Para tanto, o desenvolvimento de modelos de previsão da movimentação de cargas, em áreas urbanas, é de extrema importância para o planejamento de transporte, uma vez que permitem prever a quantidade ou o valor associado à atividade de transporte de mercadorias (NOVAK *et al.*, 2011). No entanto, a criação de um modelo é uma tarefa difícil, pois a relação entre a atividade econômica e os fluxos de transporte não é estática. Ela é influenciada por fatores internos e externos, como as estratégias da empresa, desenvolvimentos tecnológicos, regras governamentais, etc. (IDING *et al.*, 2002).

Segundo Patier e Routhier (2008), o transporte de mercadorias urbanas tem características específicas ligadas:

- aos produtos: numerosos receptores de mercadorias, rápidas mudanças na logística da distribuidora e do receptor;
- à logística urbana: o domínio dos transportes por caminhões, a grande diversidade dos tipos de veículos de entrega, diferentes padrões de viagem (viagens de ida e volta), existência do transporte por conta própria (isto é, produtos e transportador),

movimento complexo da última milha¹, congestionamento, fortes restrições ao compartilhamento rodoviário e a intervenção das autoridades locais para regulação e planejamento.

Apesar disso, os modelos desempenham um papel fundamental, pois permitem estimar e avaliar o desempenho e os impactos gerados por diferentes políticas (COMI *et al.*, 2012). Segundo Gonzalez-Feliu e Routhier (2012), um modelo consiste em reduzir um objeto estudado para compreender os mecanismos envolvidos naquele fenômeno. Para tanto, é necessário identificar claramente o objeto de estudo, indicando objetivos que, em seguida, tornam possível a determinação de variáveis para o desenvolvimento do modelo. Com base nessa identificação, é possível aplicar métodos matemáticos e de processamento de computador para a construção de um modelo adaptado para o objeto de estudo.

Para a construção de um bom modelo, Holguín-Veras *et al.* (2011) e González-Feliu e Routhier (2012) afirmam ser necessário um conhecimento detalhado da situação existente, principalmente daquelas características que estão sujeitas a alterações ou que são invariáveis. A melhor unidade de observação deve ser tão rica quanto possível em termos de qualidade da informação.

Ao serem observadas as diferentes abordagens para a modelagem de carga, percebe-se que a escolha das variáveis depende da disponibilidade de dados e da aplicação do modelo. Mesmo o modelo mais sofisticado não irá fornecer bons resultados quando os dados de entrada forem incertos ou ausentes (KULPA, 2014).

González-Feliu e Routhier (2012) definem quatro modelos principais relacionados a movimentação de mercadorias em áreas urbanas:

- modelos de estimativa de demanda têm a função de estimar o número de viagens de veículos de carga em determinada região e relacionar essa estimativa com variáveis socioeconômicas, sociais ou intrínsecas como o potencial turístico;
- modelos de otimização de demanda fixa estão relacionados com métodos de programação linear, estes que por sua vez, objetivam-se a otimizar os processos de oferta de transporte e outras operações associadas, cuja demanda já é conhecida ou foi estimada a partir de outros modelos;

¹ A última milha ou *last mile* é um termo comumente utilizado para indicar o transporte de mercadorias que saem do centro de distribuição para seguir ao seu destino final.

- modelos de simulação multi-ator têm a função de simular o comportamento das partes interessadas envolvidas;
- modelos de apoio a decisões macroeconômicas e públicas avaliam as ações e soluções para melhoria de distribuição urbana de pessoas e mercadorias e são dependentes de outros modelos propostos.

O foco principal desta pesquisa é a primeira categoria de modelos (modelos de estimativa de demanda), pois constitui a base do diagnóstico dos movimentos de mercadorias em áreas urbanas e fornece dados de entrada para os outros modelos.

Comi *et al.* (2012) apresentam uma revisão da literatura e identificam diferentes modelos de demanda para o transporte de carga para áreas urbanas e metropolitanas. Os autores concluem que vários modelos e métodos têm sido propostos pela comunidade científica, geralmente com modelagem multi-estágio e podem ser classificados em termos de unidade de volume de caminhões gerados pelo empreendimento, volume de mercadorias geradas, volume de entregas e mistos (número de caminhões gerados, volume de mercadorias e de entregas).

Os modelos baseados em número de caminhões, segundo Comi *et al.* (2012), usualmente são estruturados em duas etapas. A primeira permite estimar o número de veículos de carga atraídos ou produzidos por zonas individuais. O passo seguinte fornece a distribuição espacial (matriz Origem Destino – OD) dos fluxos estimados. Além disso, os modelos são especificados e calibrados para diferentes tipos de veículos. Os modelos baseados em número de caminhões têm a vantagem da facilidade de coleta de dados (por exemplo, utilizando contagens de tráfego), facilitando as etapas de calibração e validação.

Os modelos baseados em volume de mercadorias têm como objetivo estimar os fluxos de cargas entre diferentes áreas a partir de dados socioeconômicos para converter esta quantidade de carga em número de veículos circulantes. Uma dificuldade do uso deste modelo é a obtenção de dados em termos de fluxos de mercadorias, pois as empresas não os tornam públicos por questões comerciais (COMI *et al.*, 2012).

Os modelos baseados em volume de entregas permitem uma ligação direta entre os geradores e os operadores de transporte, por meio da utilização da mesma unidade de referência (COMI *et al.*, 2012). Dentre os modelos, pode-se citar o FRETURB e o *CityGoods*.

Por fim, o quarto grupo de modelos, identificado por Comi *et al.* (2012), são os modelos mistos, que combinam as três unidades de referência anteriores (número de caminhões

gerados, volume de mercadorias e de entregas) e criam um vínculo direto entre os comportamentos interativos entre consumidores e fornecedores / varejistas. Segundo os autores, este tipo de modelo pode tornar-se uma ferramenta útil para a avaliação dos diferentes cenários logísticos de um município.

As subseções a seguir apresentam alguns trabalhos identificados na literatura internacional e nacional por meio dos quais foram desenvolvidos modelos de geração de viagens de veículos de carga em áreas urbanas. A forma de apresentação destes trabalhos seguiu a ordem cronológica de publicação.

2.1.1 Literatura Internacional

Brogan (1980) estudou o efeito da estratificação adicional das viagens de caminhões com base no tamanho do veículo, no motivo da viagem do veículo carga (entrega de mercadorias e serviço) e no tipo do uso do solo do destino da viagem nos modelos de geração de viagens de veículos de carga (uso industrial e não-industrial). Por meio de dados de origem e destinos de quatro cidades norte-americanas (Gastonia, Carolina do Norte; Flint, Michigan; Saginaw, Michigan; e Columbus, Ohio), foram desenvolvidos os modelos com regressão linear múltipla para os veículos estratificados e os resultados foram comparados com os modelos sem a divisão dos veículos por categorias de tamanho. Por meio de indicadores estatísticos (teste F e R^2) e do erro médio, o autor concluiu que a estratificação por tipo de veículo não indicou melhorias no modelo, ou seja, os modelos gerais foram mais estatisticamente adequados e as variáveis independentes mais significativas nas equações estratificadas foram comparáveis às que aparecem no modelo total de caminhão. Já a estratificação por motivo da viagem e por tipo do uso do solo do destino da viagem resultaram em melhorias estatísticas nos modelos de geração de viagens, indicando que estes dois tipos de estratificação devem ser refinados, pois têm potencial de retornarem resultados ainda melhores.

Ogden (1992) propôs modelos de geração de viagens de veículos de carga para as cidades de Chicago, Vancouver e Melbourne. Na cidade de Chicago, a geração de veículos de carga foi estimada em função de variáveis específicas das atividades comerciais e foram correlacionadas com viagens de passageiros. Já em Vancouver e Melbourne, foram utilizadas a relação entre os veículos e as atividades, bem como outras variáveis socioeconômicas e de acessibilidade. Todos os modelos são específicos para a área onde eles foram calibrados e têm mostrado sérias deficiências na adequação para outros contextos urbanos. Os modelos foram

desenvolvidos utilizando regressão linear. Também com regressão linear Tadi e Baldach (1994) desenvolveram modelos baseados em viagens de veículos de carga e dividiram os modelos de acordo com o tamanho do veículo de carga.

Iding *et al.* (2002) construíram modelos que consideraram diferentes tipos de indústrias. Para tanto, utilizaram dados reais de geração de viagens de veículos de carga, obtidos em pesquisa de campo realizada nas indústrias e empreendimentos comerciais da Holanda, totalizando uma amostra de 1529 empresas. Os modelos foram desenvolvidos por meio de regressão linear e os resultados indicaram que as duas principais variáveis independentes que explicam a geração de viagens de caminhões são a área do estabelecimento e o número de empregados. Porém, a importância de uma ou outra varia de acordo com o tipo do estabelecimento ou indústria. Além disso, os autores concluíram que o tipo de estabelecimento também é uma variável muito importante na previsão de viagens de veículos de carga.

O estudo de Kawamura *et al.* (2008) identificou as variáveis que podem ser utilizadas para o desenvolvimento de modelos de previsão de geração de viagens de veículos de carga. Os dados utilizados neste estudo foram obtidos de três formas: por meio de questionário de pesquisa, de visitas a lojas e por contatos telefônicos. Foram obtidas informações detalhadas sobre a quantidade de mercadorias entregues nas lojas e o número de viagens de entrega por semana por loja. No total, foram pesquisadas 426 lojas de 9 redes nacionais de varejo, 5 redes de móveis e 4 redes de calçados. Foi utilizada a técnica de regressão logit binomial para o desenvolvimento dos modelos. Os autores concluíram que a geração de viagens de veículos de carga pode variar consideravelmente dentro do mesmo setor de varejo dependendo da estratégia da cadeia de suprimentos adotada. Além disso, ao contrário do que é comumente utilizado na maioria de estudos do tipo, o número de empregados e a área total do estabelecimento como variáveis independentes não retornaram bons resultados quanto aos modelos desenvolvidos.

Bastida e Holguín-Veras (2009) analisaram comparativamente duas abordagens para a estimativa de geração de viagens de veículos de carga: a metodologia dos mínimos quadrados (*ordinary least squares* – OLS) e a classificação cruzada. Os modelos OLS foram utilizados para estimar a relação entre o número de veículos de carga gerados e as variáveis independentes. Por sua vez, a utilização da técnica de classificação cruzada teve como objetivo identificar uma estrutura de classificação com uma boa representação do processo de geração de viagens. Para tanto, utilizou-se a análise de classificação múltipla (*multiple classification analysis* –

MCA) para identificar os grupos de variáveis independentes que explicassem a geração de viagens. Esta técnica é descrita mais detalhadamente por Ortúzar e Willumsen (1994). A base de dados utilizada foi obtida por meio de pesquisas em estabelecimentos comerciais localizados nas regiões de Manhattan e Brooklyn, em Nova York. Os resultados mostraram que as técnicas de OLS e a MCA retornam os mesmos resultados de quais são as variáveis mais importantes que explicam a geração de viagens de carga. O segmento da indústria, o número de vendas totais por ano e o tipo de mercadoria do estabelecimento se mostraram importantes fatores para a estimativa de geração de viagens de veículos de carga.

Russo e Comi (2010) desenvolveram uma modelagem para simular os movimentos de mercadorias em uma área urbana por meio de um modelo multinível, considerando uma abordagem desagregada para cada nível de decisão. Os movimentos foram examinados em dois níveis: análise dos fluxos gerados advindos do consumo das mercadorias considerando a quantidade e o número de veículos. No primeiro nível do modelo é calculada a quantidade de mercadorias movimentadas devido ao consumo ou reabastecimento, no segundo nível, determinam-se o tipo de estabelecimento, os veículos utilizados e o tempo de entrega, assim como a via escolhida para o reabastecimento dos pontos de venda, a fim de estimar o fluxo de veículos na área urbana.

Holguín-Veras *et al.* (2011) analisaram a relação entre o tamanho do estabelecimento e as taxas de geração de viagens de veículos de carga. Para tanto, utilizaram dados de 400 empresas transportadoras e 400 empresas receptoras de mercadorias, localizadas em Nova York e Nova Jersey. O processo de modelagem utilizado foi a modelagem linear simples considerando o número de empregados como única variável independente. Os mínimos quadrados ordinários foram usados para estimar os modelos e a produção de viagens (viagens de caminhão feitas por dia) e a atração de viagens (entregas recebidas por dia), que foram as variáveis dependentes dos modelos. Os estabelecimentos foram classificados por códigos de Classificação Industrial Padrão (SIC) e agrupados em oito setores. Os resultados dos modelos indicaram que em apenas 18% dos setores da indústria, a partir dos quais os dados foram coletados, a suposição de proporcionalidade entre a geração de viagens de veículos de carga e número de empregados estava ajustada. Ou seja, o uso desse modelo não foi ajustado em 82% dos setores.

Para justificar este resultado, Holguín-Veras *et al.* (2011) afirmam que, ao contrário do caso de transporte de passageiros, no qual há uma correspondência bastante estreita entre o número

de viagens de passageiros geradas e as viagens correspondentes do veículo, no transporte de carga, a correspondência entre demanda e tráfego é muito mais fraca, uma vez que as empresas podem mudar drasticamente os tamanhos de embarque. Na geração de carga, o tamanho do empreendimento está diretamente relacionado com a quantidade de mercadoria recebida ou produzida. Já a geração de viagens de veículos de carga é resultado de decisões logísticas internas ao empreendimento. Por conta disso, os autores afirmam que a geração de carga e a geração de viagens de veículos de carga são fenômenos diferentes e devem ser estudados por modelos diferentes.

Utilizando a mesma base de dados de Holguín-Veras *et al.* (2011), Jaller *et al.* (2014) avaliaram o desempenho de modelos de geração de viagens de veículos de carga, baseados na área do estabelecimento comercial. Esses modelos foram comparados com os modelos de geração de viagens de veículos de carga, baseados no emprego desenvolvidos pelos autores em Holguín-Veras *et al.* (2011). Os modelos foram desenvolvidos utilizando a regressão linear simples. A análise comparativa dos erros nos dados de validação entre os modelos utilizando essas duas variáveis independentes indicaram claramente que os modelos baseados em emprego funcionam melhor que a base da área. Os resultados também indicam que, embora os modelos baseados em áreas tenham desempenho inferior, é apenas uma pequena diferença para a maioria das categorias de estabelecimentos analisadas.

Comi e Nuzzolo (2014) desenvolveram um modelo para estimar a geração de viagens de veículos de carga em áreas urbanas a partir dos resultados de um estudo realizado para investigar o comportamento do consumidor final na cidade de Roma (Itália). Partindo do resultado dessa pesquisa, os autores utilizaram o modelo de regressão logit multinomial e, como variáveis correlacionadas, as características físicas do comércio e dados socioeconômicos da população. Com este modelo, foram estimados os valores de geração de viagens de mercadorias baseados nas escolhas do consumidor final em relação ao tipo de ponto de venda (por exemplo, pequeno, médio ou grande porte), pois, segundo os autores, tais escolhas impactam os fluxos de distribuição de mercadorias.

Kulpa (2014) desenvolveu equações de geração de viagens de veículos de carga com base em diferentes fontes de dados, dados secundários e abordagens indiretas. Regressão linear, regressão não linear e Redes Neurais Artificiais (*Artificial Neural Networks – ANN*) foram utilizadas para determinação das equações de geração de viagens. Como resultado do estudo, na regressão linear, a variável com melhores resultados foi o número de habitantes na área e

com os piores resultados foi o número de empregados na agricultura. Já na regressão múltipla, as variáveis mais correlacionadas foram número de empregados no setor de serviços, número de empregados na indústria e número de habitantes. O autor concluiu que, apesar da regressão linear ser o método mais simples, foi a que obteve os melhores resultados, se comparado com a regressão múltipla e as Redes Neurais Artificiais.

O estudo de Gunay *et al.* (2016) teve como objetivo melhorar a compreensão dos processos subjacentes que geram viagens de veículos de carga e, conseqüentemente, melhorar a modelagem de geração de viagens de veículos de carga. Para tanto, utilizando dados obtidos do Plano Diretor de Logística da cidade de Kocaeli (Turquia), os autores inicialmente criaram grupos de estabelecimentos logísticos que possuíam características homogêneas de geração de viagens de veículos de carga. Após a criação e análise dos grupos criados, foi identificado um grupo que possuía muitas viagens zero para a categoria carreta (*tractor-trailer*). Por conta disso, foi utilizado o modelo de logit binomial, que é indicado para este tipo de configuração de dados, e posteriormente, foi comparado com a abordagem tradicional que utiliza regressão linear.

Os resultados do estudo indicaram que a criação de grupos homogêneos de estabelecimentos logísticos foi válido para o refinamento do modelo e a nova abordagem de modelagem aplicada utilizando o modelo logit binomial obteve melhores resultados que a modelagem de regressão regular.

Sánchez-Díaz (2016) estudou a geração de viagens de veículos de carga em estabelecimentos comerciais urbanos e os impactos no tráfego usando dados coletados dos estabelecimentos da cidade de Gotemburgo, na Suécia. Os dados abrangem diferentes zonas da cidade e incluem diversos setores comerciais, como varejistas, serviços de alimentação, de saúde, escritórios do setor público e educação. No total, foram coletadas informações de 195 estabelecimentos. O autor fez uso de um conjunto de modelos estatísticos (regressão logit binomial e probit) que foram desenvolvidos com base em análises de regressão e modelos de escolha discreta para estimar o número de viagens de veículos de carga produzidas e atraídas por semana. Além disso, os modelos foram capazes de estimar os volumes e pesos das mercadorias movimentadas pelo estabelecimento. Os resultados indicaram que os estabelecimentos que trabalham com produtos perecíveis possuem uma maior taxa de atração de viagens de veículos de carga, seguido dos escritórios do setor público e varejistas. O estudo constatou

também que não existia no local uma política de racionalização das entregas e consequente redução dos impactos que as entregas geram no tráfego urbano.

Uma das fontes mais importantes para o estudo de geração de viagens na literatura internacional é o Manual de Geração de Viagens do ITE (*Institute of Transportation Engineers*). Ele consiste em uma compilação de estudos de geração de viagens, desenvolvidos por agências públicas, empresas de consultoria e universidades. Apesar do foco do Manual ITE não seja especificamente transporte de mercadorias, alguns resultados podem ser aplicados diretamente à geração de viagens de veículos de carga (ITE, 2008 *apud* SÁNCHEZ-DÍAZ, 2016).

Outra importante fonte é o NCFRP 25 *Freight Trip Generation and Land Use Report* (Holguín-Veras *et al.*, 2012). Voltado exclusivamente para a geração de viagens de veículos de carga, possui uma discussão abrangente sobre a modelagem destas viagens e um conjunto de estudos de caso como modelos de geração baseados em dados de estabelecimentos da cidade de Nova York. Neste documento, são apresentadas estimativas de geração de viagens para diferentes tipos de estabelecimentos, diferentes usos do solo e são utilizadas diferentes técnicas para estimativa destas viagens, como taxas de viagens, OLS e MCA.

2.1.2 Literatura Nacional

No Brasil, um dos primeiros estudos sobre modelos de geração de viagens de veículos de carga foi desenvolvido por Marra (1999) para identificar os padrões da demanda por mercadorias em áreas residenciais e comerciais do município de Campinas (São Paulo) em função de variáveis socioeconômicas e operacionais, como, por exemplo, áreas de imóveis, número de moradores e empregados e nível de renda. Para o desenvolvimento do trabalho, foram utilizados dados de pesquisa de campo e utilização do *software TransCAD*. Os modelos tiveram coeficiente de determinação (R^2) acima de 0,88, constatando que há correlação entre a geração de viagens de veículos de carga e as variáveis socioeconômicas em questão. No entanto, Marra (1999) enfatizou que esta correlação apenas existe quando analisada de uma forma agregada, como um bairro, por exemplo, e não de forma local, em um domicílio em específico. O autor também analisou a geração de viagens de veículos de carga considerando o uso do solo, cujos resultados indicaram que o porte do imóvel e o tipo de atividade econômica praticada no estabelecimento influenciam diretamente na geração de mercadorias.

Melo (2002) desenvolveu um modelo para avaliação da demanda por transporte de carga para o município do Rio de Janeiro. Foram realizadas pesquisas em transportadoras e em estabelecimentos comerciais como, por exemplo, supermercados, comércio varejista, bares e restaurantes, material de construção e combustível, para identificar a quantidade de carga movimentada. A autora utilizou, para o caso de lojas comerciais, a regressão linear simples, correlacionando volume de viagens geradas com área construída. Para modelar o fluxo de viagens de empresas de transporte, foi utilizada a regressão múltipla.

Silva e Waisman (2007) desenvolveram um modelo para bares e restaurantes da cidade de São Paulo, com dados de pesquisa de campo. Com isso, foi desenvolvido um modelo correlacionando o volume de viagens de veículos de carga com a área do estabelecimento e número de funcionários.

O estudo de Gasparini *et al.* (2010) determinou modelos de geração de viagens de veículos de carga para *shopping centers* e supermercado no Rio de Janeiro. Para tanto, foi elaborado e aplicado um questionário em 21 supermercados e 7 *shopping centers* da cidade. Contatou-se que, para os supermercados e *shopping centers*, o dia de maior volume (pico) de viagens é a sexta-feira, o caminhão é o tipo de veículo de carga mais utilizado, sendo do tipo pequeno e médio para entregas em *shopping centers*. Domingo é um dia atípico para entregas, sendo estas mínimas e eventuais e o maior número de viagens de veículos de carga ocorre no período da manhã.

Além disso, Gasparini *et al.* (2010) observaram que, para os *shopping centers*, no período próximo ao Natal, há um aumento de aproximadamente 45% das viagens. Para os supermercados que não possuem depósito, o estudo mostrou que o comportamento relativo a viagens ocorre com maior volume no início da semana, geralmente nas segundas e terças-feiras durante o período da manhã. Quando o supermercado possui depósito, o maior volume de viagens ocorre nas sextas-feiras. Os modelos que obtiveram melhores coeficientes de determinação consideraram, como variável explicativa, o número de clientes e/ou número de funcionários para o caso de supermercados e, número de clientes para *shopping centers*.

Barroso *et al.* (2013) desenvolveram um estudo para avaliar a correlação entre a demanda por transporte de carga e informações socioeconômicas da população em Belo Horizonte. Foram utilizadas as equações de geração propostas por Melo (2002) e dados de cadastro imobiliário para estimar o número de viagens de mercadorias. O estudo concluiu que a geração de

viagens de veículos de carga, principalmente para o setor de bar e restaurante, tem correlação com as características socioeconômicas da população na cidade, sendo que os maiores valores de coeficientes de correlação foram obtidos para o número de viagens geradas, a renda média e o nível de instrução.

Pereira e Oliveira (2014) propuseram uma metodologia para estimar fluxos de carga utilizando dados secundários, a fim de obter uma matriz OD de carga para Belo Horizonte. A metodologia desenvolvida consistiu em sete etapas: definição da área de estudo, zoneamento, escolha dos segmentos econômicos, coleta e tratamento de dados secundários, geração de viagens, modelagem da distribuição de viagens e calibração. Foi aplicado, neste estudo, o modelo integrado baseado em atividades, denominado TRANUS e foram utilizadas as equações de geração de carga presentes no estudo de Melo (2002) e Tadi e Baldach (1994). Como resultados, constatou-se que a maior parte das viagens com destino ao município de Belo Horizonte se originam em outros municípios e que as vias mais carregadas com o tráfego de cargas segundo o modelo são as avenidas Amazonas, Antônio Carlos e Cristiano Machado, Via expressa e Anel Rodoviário. A matriz OD obtida por Pereira e Oliveira (2014) obteve uma correlação com os dados reais de 0,81 e um R^2 de 0,65.

Ferreira e Silva (2015) utilizaram variáveis territoriais, como população, número de domicílios e empregos na região, para verificar a relação destas com as viagens de veículos de carga em áreas urbanas na Região Metropolitana do Rio de Janeiro (RMRJ). Como resultado, verificou-se que as variáveis domicílio e população possuem elevado potencial para explicar viagens de veículos de carga. Como desdobramento deste trabalho, Ferreira e Silva (2016) analisaram a relação entre variáveis socioeconômicas e geração de viagens de caminhões na RMRJ usando regressão linear. Os autores concluem que existe relação entre o fluxo de carga e população, número de domicílios e emprego, sendo que estas variáveis podem auxiliar o planejamento de transportes e o planejamento estratégicos de empresas.

Oliveira *et al.* (2016) desenvolveram modelos de geração de viagens de veículos de carga para bares, restaurantes, mercados e supermercados em Belo Horizonte. Para a obtenção dos dados, foi realizada uma pesquisa de campo através de entrevistas com funcionários do setor administrativo que possuíam conhecimento sobre o funcionamento das operações de carga e descarga do empreendimento. Como resultado, concluiu-se que a área do empreendimento comercial pode ser associada à geração de viagens de veículos de carga. Foi realizada ainda uma comparação com os modelos obtidos por Melo (2002), Silva e Waisman (2007) e

Gasparini (2008) que indicaram semelhanças e divergências entre os modelos, reforçando a necessidade de elaboração de um modelo de geração de viagens local. Além disso, os autores concluem que não é necessária uma amostra expressiva, e sim consistente para a validação dos resultados.

2.2 Análise Espacial para a Geração de Viagens

A análise espacial tem como princípio básico o entendimento da ocorrência de fenômenos observados no espaço e a relação existente entre eles (HENRIQUE, 2004). A análise do comportamento espacial das variáveis ganhou destaque a partir da década de 1950, quando se constatou que apenas a estatística básica não era capaz de explicar os fenômenos espaciais (ROCHA *et al.*, 2016).

Pela vasta gama de procedimentos existentes, as operações de análise espacial podem ser sintetizadas em ferramentas de análises topológicas (operações de adjacência e conectividade geográfica, por exemplo), ferramentas de análise multicritério (baseadas em álgebra de mapas), ferramentas de inferência espacial (interpolação e análise geoestatística), e ferramentas de modelagem preditiva (elaboração de cenários futuros baseados em eventos passados e presentes).

No entanto, para Câmara *et al.* (2002), grande parte dos usuários limita o uso de ferramentas espaciais operações de visualização, tirando conclusões intuitivas. Assunção (2001) salienta que a visualização é um fator importante da estatística espacial, mas complementar à descrição, comparação e interpretação dos dados.

Miller (1999) aponta várias questões analítico-espaciais que devem ser consideradas na análise de transportes e, dentre elas, está o caso da dependência espacial, que é o foco deste trabalho, mais precisamente na análise de demanda por transportes. Segundo Ribeiro (2012), as técnicas que utilizam análise espacial para a previsão de geração de viagens são as mais adequadas devido à forte relação entre transporte e espaço. Assim, é adequado inferir que as áreas são diferentes entre si e possuem características específicas, afetando diretamente a demanda por viagens (RIBEIRO, 2012).

Em relação aos trabalhos relacionados ao planejamento de transportes, que utilizaram técnicas de análise espacial, tem-se Bhat e Zhao (2002) que desenvolveram um estudo cujo objetivo

foi identificar os problemas espaciais que precisam ser reconhecidos na modelagem da demanda e propor uma modelagem multi-nível para previsão de viagens.

Lopes (2005), que confrontou a análise tradicional e modelos de análise espacial na temática do planejamento de transportes e concluiu que com a análise espacial foi possível obter melhores resultados que o método tradicional, retornando menores erros de estimativa.

Pitombo *et al.* (2010) comparam duas técnicas de análise de dados espaciais para a previsão de viagens urbanas na Região Metropolitana de São Paulo (RMSP). Os autores utilizaram as técnicas de krigagem (técnica de análise geoestatística) e a regressão geograficamente ponderada (*Geographically Weighted Regression – GWR*). Além disso, foi proposta uma metodologia para aplicação conjunta da Análise de Componentes Principais (PCA) para prevenção de problemas com multicolinearidade. Foram encontrados bons resultados estatísticos para as duas técnicas e verificou-se que o GWR se destaca pela possibilidade de visualização dos padrões de deslocamentos espaciais para cada uma das variáveis envolvidas no modelo.

Novak *et al.* (2011) investigaram a aplicação da regressão linear e técnicas de regressão espacial para a previsão de viagens de veículos de carga no âmbito nacional. Os autores constataram que, devido à natureza zonal ou espacial da demanda por viagens e da geração de viagens de veículos de carga, há questões espaciais que precisam ser reconhecidas e abordadas na construção do modelo de geração. Eles demonstraram no artigo que o uso de modelos de regressão linear multivariada pode resultar em boas estimativas para a geração de viagens de veículos de carga em nível nacional utilizando variáveis socioeconômicas. Entretanto, foram observados resíduos espacialmente correlacionados nestes modelos. Por fim, os autores concluíram que a utilização de técnicas de modelagem de regressão espacial pode melhorar o ajuste dos modelos e eliminar os problemas associados à autocorrelação espacial dos resíduos.

Ribeiro (2012) realizou um estudo cujo objetivo foi construir um modelo para previsão de demanda por viagens por meio de regressão espacial e de regressão geograficamente ponderada. A autora aplicou a modelagem nos dados da Pesquisa OD de Vitória, Espírito Santo, realizadas nos anos de 1998 e 2007. Após os ajustes dos modelos, concluiu-se que o modelo de regressão geograficamente ponderada pelas distâncias obteve melhores resultados, se comparados com os modelos convencionais, que utilizam regressão linear.

Rocha *et al.* (2016) afirma que os modelos de geração de viagens tradicionais não correspondem à atual dinâmica socioespacial das cidades, pois muitos desses modelos não consideram fatores geográficos, que são componentes básicos na estrutura urbana. Ainda segundo a autora, a geração de viagens está intimamente relacionada a fatores socioeconômicos e, devido a isto, a incorporação da variável espacial permite enriquecer o modelo.

Por fim, Sánchez-Díaz *et al.* (2016) realizaram uma análise exploratória da atração de viagens de veículos de carga urbana a fim de identificar a correlação com as características do ambiente urbano para a cidade de Nova York. Os autores constataram que a localização do empreendimento desempenha um papel importante no número final de viagens atraídas, pois a localização está diretamente relacionada ao valor da terra e à distância percorrida pelos veículos até o cliente final, por exemplo. Além disso, estabelecimentos de varejo, localizados em áreas de alto emprego, tendem a ter uma taxa de atração por empregado maior que a média. Por fim, os autores concluíram que o valor de atração desse tipo de viagens é melhor estimado com o uso de técnicas espaciais, variáveis de localização e especificações não lineares, por conta da autocorrelação espacial das variáveis explicativas dos modelos.

2.3 Síntese da revisão da literatura

De acordo com a revisão da literatura realizada, pode-se constatar que a regressão linear é uma das abordagens de modelagem de geração de viagens de veículos de carga urbanas mais utilizadas na literatura. Segundo Novak *et al.* (2011), este fato pode ser explicado de modo que os modelos de regressão linear podem ser utilizados sem investir em soluções de *software* de alto custo e personalizadas. Souza *et al.* (2010) obtiveram conclusão semelhante após analisarem 25 artigos da literatura nacional e internacional sobre modelos de geração de viagens de veículos de carga, quando identificaram uma predominância do uso de métodos de regressões lineares simples e múltiplas para o desenvolvimento dos modelos. Esses modelos geralmente são estatisticamente validados através de R^2 , t-test e p-value.

Apesar da grande e larga utilização desta técnica para modelagem de viagens de veículos de carga, as características da variável dependente (número de viagens ou volume de mercadorias) fazem com que a utilização de modelos generalizados seja mais indicada, pois forma um conjunto de dados discretos (TURKMAN e SILVA, 2000). Foram identificados poucos estudos na literatura nacional e internacional (KAWAMURA *et al.*, 2008; GUNAY *et*

al., 2016; SÁNCHEZ-DÍAZ, 2016) fazem uso deste tipo de abordagem. Por conta disso, foi utilizada a técnica de regressão linear generalizada neste trabalho.

Ainda em relação aos modelos de regressão linear, Novak *et al.* (2011) demonstraram que estes podem retornar bons resultados de estimativas de geração de viagens de veículos de carga urbana. Porém, podem apresentar resíduos espacialmente correlacionados. Portanto, segundo os autores, modelos de regressão espacial podem melhorar o ajuste do modelo e eliminar problemas associados à autocorrelação espacial.

Como os ambientes urbanos são dinâmicos e espacialmente complexos, as soluções em planejamento de transporte e políticas públicas devem ser sensíveis ao contexto geográfico. No entanto, há uma lacuna na literatura sobre o uso de análises geográficas abrangentes para melhorar o cálculo da geração de viagens de veículos de carga dentro das cidades. Inúmeros estudos empregam análise espacial apenas como ilustração dos resultados e não como parte integrante do processo do processo de criação do modelo.

Diante disso, este trabalho vem preencher esta lacuna na literatura, pois utiliza técnicas de análise espacial para a construção de modelos de geração de viagens de carga.

Em relação às variáveis dependentes utilizadas, a revisão da literatura identificou que os modelos de geração de número de caminhões gerados pelo empreendimento e em termos de volume de cargas são os mais recorrentes tanto na literatura nacional, quanto na internacional (SOUZA *et al.*, 2010). Ferreira e Silva (2015) realizaram uma análise dos modelos de geração de carga, presentes na literatura, a fim de identificar quais variáveis poderiam ser utilizadas para explicar a geração de viagens de veículos de carga em áreas urbanas. Como resultado da revisão bibliográfica realizada, observou-se que a variável dependente mais utilizada é o número de viagens, com 84% dos trabalhos, e a variável volume de carga foi observada em apenas 16% dos trabalhos consultados.

Segundo Ogden (1992), a abordagem de modelagem baseada em veículos é a mais adequada para áreas urbanas. Porém, é importante ressaltar que, como o número de viagens geradas por um estabelecimento depende do volume de carga e das decisões logísticas do empreendimento, o número de viagens não é necessariamente diretamente proporcional ao tamanho do negócio (HOLGUÍN-VERAS *et al.*, 2011; SÁNCHEZ-DÍAZ, 2016). Além disso, Holguín-Veras *et al.* (2011) afirmam que a geração de carga e a geração de viagens de veículos de carga são fenômenos diferentes e devem ser estudados por modelos diferentes.

Já com relação às variáveis independentes utilizadas, Jaller *et al.* (2013) fizeram uma extensa revisão da literatura de modelos de geração de viagens de veículos de carga e a Tabela 1 apresenta um resumo dos resultados encontrados.

Tabela 1: Variáveis independentes utilizadas nos modelos na literatura (JALLER *et al.*, 2013).

Variável Independente	Quantidade	Variável Independente	Quantidade
Área	786 41,59%	Individuais	15 0,79%
Empregados	565 29,89%	Tipo de carga	13 0,69%
Tipo de estabelecimento	278 14,71%	Total de vendas	5 0,26%
Uso do solo	211 11,16%	Segmento da indústria	2 0,11%
Características físicas dos domicílios	47 2,49%	Volume de tráfego	2 0,11%
Outras	41 2,17%	Renda	1 0,05%
Frota	36 1,90%	Estacionamento	1 0,05%

Ainda segundo Jaller *et al.* (2013), a revisão encontrou ao todo 1.890 modelos, que variavam desde taxas de viagem até modelos de regressão tanto para a produção de viagens de veículos de carga quanto para a atração. Esses modelos utilizam uma variedade de variáveis independentes, nas quais o emprego e a área são as mais comuns. Dos modelos que utilizam apenas uma variável, 53% são taxas constantes por empregado e cerca de 70% são taxas constantes por unidade de área. Além disso, foram encontrados 52 modelos de regressão que incluíam emprego e área como variáveis independentes.

Na mesma linha de raciocínio, Sánchez-Díaz (2016) concluiu que o tamanho de um estabelecimento (medido em área ou emprego) é um bom indicador de sua intensidade de atividade comercial e estabelecimentos de diferentes setores comerciais têm diferentes necessidades de logística e, portanto, têm diferentes índices de geração de viagens.

Ferreira e Silva (2015) e Jaller *et al.* (2013) classificaram e categorizaram as variáveis independentes mais utilizadas na literatura nacional e internacional, conforme a Tabela 2.

Tabela 2: Variáveis independentes e sua categoria de classificação utilizados nos modelos identificados na literatura (FERREIRA E SILVA, 2015; JALLER *et al.*, 2013).

Tipo	Categoria	Variáveis utilizadas
Estabelecimento / Atividade	Área	Área construída, área de vendas
	Características físicas	Número de estabelecimentos / lojas / empresas / escritórios
		Vagas de estacionamento
	Características operacionais	Carga horária trabalhada / dias de funcionamento
		Carga própria da empresa
		Distância percorrida
		Número de fornecedores
		Número / volume de unidades entregues
	Emprego / atividade	Emprego por setor de atividade
		Número de empregados por estabelecimento
Frota	Tamanho da frota própria	
Territoriais	Domicílios	Número de domicílios na zona
	Emprego	Total de empregos por zona
	População	População residente na zona
		Renda média per capita
	Uso do solo	Tipo de uso do solo

Segundo Souza *et al.* (2010), a correta escolha de uma variável independente é mais importante para o resultado final do que a utilização de modelagem complexa no desenvolvimento de um modelo de geração de viagens. Esta afirmação corrobora com Marra (1999), que concluiu que modelos que consideram ajustes de curva para função potência ou polinomial, não apresentaram resultados tão bons quanto aqueles que utilizaram regressão linear simples ou ajuste exponencial. Portanto, entende-se que a correta escolha de variáveis independentes contribui mais para um bom resultado do modelo do que o uso de modelos com inúmeras variáveis independentes ou excessivamente elaborados (MARRA, 1999).

3 ÁREA DE ESTUDO, SETOR ECONÔMICO E OBTENÇÃO DE DADOS

Neste capítulo será contextualizada a área de estudo deste trabalho, o setor econômico escolhido para a construção dos modelos e as fontes de informações utilizadas para obtenção dos dados.

3.1 Definição da Área de Estudo

A definição da área de estudo é uma etapa essencial para o desenvolvimento de modelos de geração de viagens e, para fins de pesquisa acadêmica, fundamental para realizar comparações com outros modelos, visto que, o modelo deve ser compatível com as necessidades locais (GRIECO e MARCOLINI, 2012; OLIVEIRA *et al.*, 2017). Desta forma, a primeira etapa foi a definição da área de estudo e a obtenção das bases de dados para tal área. É importante certificar também que existem fontes de dados suficientes e consistentes para a área em questão.

A abordagem metodológica proposta foi desenvolvida e aplicada no município de Belo Horizonte (Figura 1) em função da maior facilidade de compreensão territorial, da existência de diversas bases de dados potencialmente úteis para o estudo e, principalmente, devido à existência de uma pesquisa de campo na área central do município desenvolvida no ano de 2011 pela UFMG em parceria com a BHTRANS (OLIVEIRA, 2016). Esta pesquisa foi de grande importância na etapa de validação, em que se comparam os resultados da abordagem metodológica proposta com os dados já existentes.

Belo Horizonte possui atualmente cerca de 2,5 milhões de habitantes, área de 331,40 km², densidade demográfica de 7.584 hab./km² e um produto interno bruto *per capita* (PIB *per capita*) de R\$35.187,85 (IBGE, 2017). O município foi planejado e construído no local do antigo povoado Curral del Rey para ser a nova capital do estado de Minas Gerais. Foi inaugurado em 12 de dezembro de 1897, porém sua construção não estava totalmente finalizada por conta da crise econômica que acometia o País e que limitou os recursos para finalizar o projeto.

Ao longo dos 100 primeiros anos de existência de Belo Horizonte, a cidade originalmente planejada para o entorno da Avenida do Contorno sofreu uma expansão urbana sem o devido planejamento urbano. A expansão desordenada do espaço urbano por décadas gerou reflexos negativos na estrutura viária, quase sempre construída de forma a remediar a crescente

demanda da população e da indústria. Por conta disso, a cidade enfrenta problemas de mobilidade de seus habitantes e de naturezas logística e operacionais comuns a outras cidades brasileiras de grande porte. Tais problemas impactam diretamente o setor de transporte de cargas, que é primordial para o funcionamento da cidade, porém frequentemente negligenciado no planejamento da mesma.

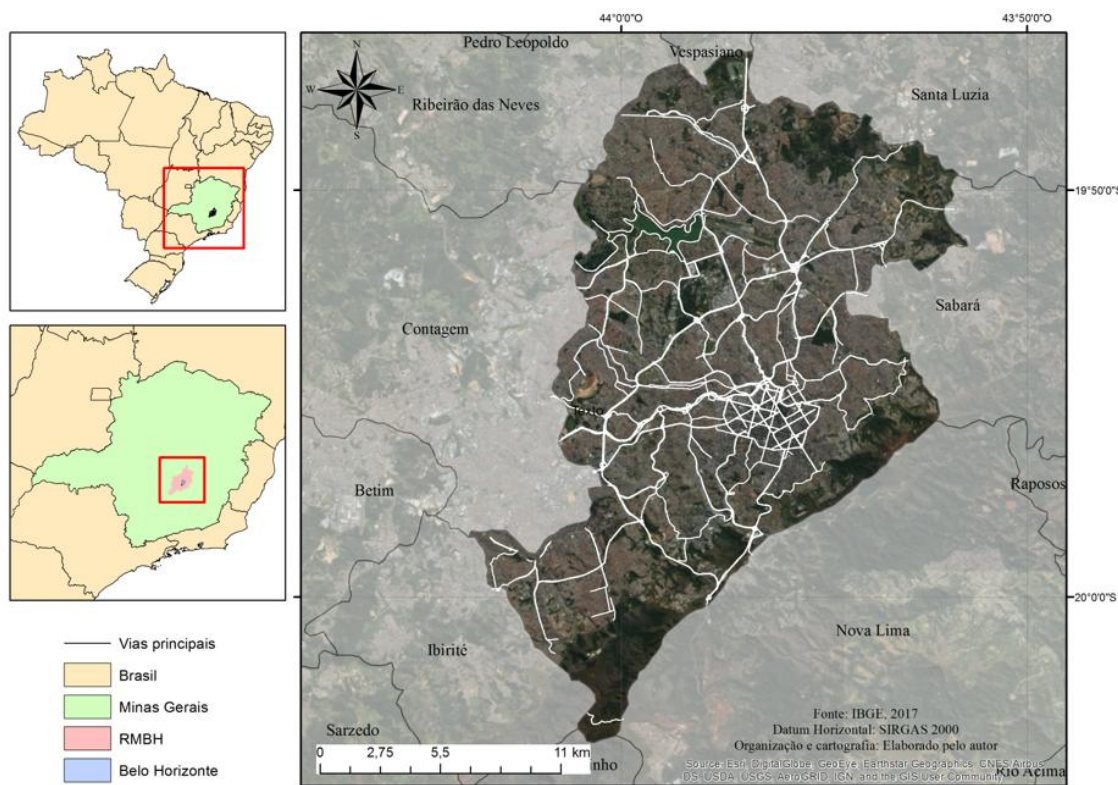


Figura 1: Área de Estudo.

Em 2007, a Prefeitura de Belo Horizonte implantou restrição ao tráfego de veículos de carga na área central e nos principais corredores de acesso à mesma. Esta medida restringiu o tráfego de veículos acima de determinado tamanho e comprimentos, priorizando a utilização de veículos de carga menores com o intuito de promover maior agilidade no trânsito. De forma direta, a medida impôs maior rotatividade nas vagas de carga e descarga, visou garantir maior segurança no trânsito, bem como visou preservar o pavimento das vias do município (BHTRANS, 2013). Todavia, o êxito da medida depende de ações paralelas, como a implementação de políticas públicas e mudanças comportamentais de motoristas e empreendedores quanto à logística de entregas.

Pela regra vigente, o tráfego de veículos com lotação de até 5 toneladas e comprimento de até 6,50 metros é livre em todas as vias no município. Já veículos maiores, seguem as seguintes regras apresentadas na Tabela 3.

Tabela 3: Restrição de tráfego de veículos pesados (BHTRANS, 2013).

CAPACIDADE DO VEÍCULO	HORÁRIOS
Até 5 toneladas e comprimento de até 6,50 metros	Livre
Acima de 5 toneladas ou comprimento acima de 6,50 metros.	<p>Área Central (Hipercentro, Savassi, Lourdes, Assembleia, Barro Preto, Região Hospitalar) e Avenida Nossa Senhora do Carmo:</p> <ul style="list-style-type: none"> • Proibido de segunda a sexta: das 7h às 20h; • Sábados: das 7h às 15h <p>Corredores de tráfego:</p> <ul style="list-style-type: none"> • Proibido de segunda a sexta: das 7h às 9h e das 17h às 20h; • Sábados: das 7h às 9h.
Carretas e cavalos mecânicos	<p>Área Central (Hipercentro, Savassi, Lourdes, Assembleia, Barro Preto, Região Hospitalar) e Avenida Nossa Senhora do Carmo:</p> <p>Expressamente proibido em qualquer horário</p>

Os corredores de acesso à Área Central com restrição de tráfego de caminhões de grande porte, incluídos no disciplinamento, são ilustrados na Figura 2.



Figura 2: Principais corredores de acesso à região central (BHTRANS, 2013).

3.2 Definição do setor econômico

A definição do setor econômico objeto de estudo é importante, pois os modelos de geração de viagens de veículos de carga atendem apenas ao setor econômico para o qual foram validados. Portanto, este estudo será baseado na geração de viagens de veículos de cargas em supermercados.

A escolha dos supermercados se deve ao fato de ser uma atividade econômica com presença em todas as regiões do município. Além disso, tem como característica a elevada geração de viagens de veículos de carga, que pode afetar todos os níveis da rede viária, desde vias locais a vias arteriais. Outro ponto importante a ser destacado é que uma técnica de geração de viagens desenvolvido para o setor de supermercados pode ser replicada para outras localidades, posto que este é um setor econômico presente em todas as cidades.

No ano de 2016, o setor do varejo supermercadista correspondeu a 5,4% das riquezas (bens e serviços) produzidas no Brasil. Em todo o País, há cerca de 89 mil estabelecimentos, empregando 1,8 milhões de pessoas. Estes números corroboram a importância deste setor na economia brasileira (ABRAS, 2017).

Ainda segundo a Associação Brasileira de Supermercados (ABRAS), em Minas Gerais localiza-se 10% do total dos estabelecimentos brasileiros, sendo o terceiro estado com maior número de lojas. Em relação ao faturamento, o setor gerou R\$ 29 bilhões de receita, representando 10,8% do faturamento nacional em 2016.

3.3 Obtenção dos dados

Segundo Ogden (1992) e Dablanc (2009), apesar dos avanços observados nos Estados Unidos, Japão e na Europa, o transporte por mercadorias ainda é menos monitorado do que o transporte de passageiros. Essa situação é ainda mais discrepante em países em desenvolvimento, onde não se encontram bases comparativas relativas ao transporte de carga e cujas metodologias de pesquisas são geralmente locais e diferentes entre si. Por conta dessa escassez de dados, foram utilizados dados secundários para a formulação do modelo de geração de viagens de veículos de carga urbana.

A primeira fonte de dados utilizada foi o Cadastro Municipal de Contribuintes de Tributos Mobiliários (CMC) do município de Belo Horizonte do ano de 2015 (BELO HORIZONTE, 2015). O CMC consiste no cadastro de pessoas jurídicas que exercem atividades que estão

sujeitas ao pagamento do Imposto sobre Serviços de Qualquer Natureza (ISSQN) ou das Taxas Mobiliárias (Taxa de Fiscalização, Localização e Funcionamento – TFLF, Taxa de Fiscalização Sanitária – TFS e Taxa de Fiscalização de Engenhos de Publicidade – TFEP) no município de Belo Horizonte (BELO HORIZONTE, 2017).

O banco de dados do CMC de Belo Horizonte traz, entre outros dados, o nome da empresa, CNPJ (Cadastro Nacional de Pessoa Jurídica), o endereço de funcionamento, a data do início da atividade, a área utilizada e o código CNAE (Classificação Nacional de Atividades Econômicas) da atividade. O CNAE é a classificação oficialmente adotada pelo Sistema Estatístico Nacional e pelos órgãos federais gestores de registros administrativos (IBGE, 2016).

Definido o setor econômico de análise, foi necessário filtrar os empreendimentos do setor na base de dados do CMC Belo Horizonte (BELO HORIZONTE, 2015). Para a identificação destes estabelecimentos, utilizou-se o CNAE. Os códigos utilizados são apresentados a seguir:

- **4711-3/01:** Compreende as seguintes categorias:
 - Hipermercado;
 - Hipermercado – área de venda superior a 5000 metros quadrados; comércio varejista;
 - Mercadorias em geral, com predominância de produtos alimentícios, com área de venda superior a 5000 metros quadrados; comércio varejista;
- **4711-3/02:** Compreende as seguintes categorias:
 - Mercadorias em geral, com predominância de produtos alimentícios, com área de venda entre 300 e 5000 metros quadrados; comércio varejista;
 - Supermercado;
 - Supermercado – área de venda de 300 a 5000 metros quadrados; comércio varejista.

Com base nos dados do CMC do ano de 2015 (BELO HORIZONTE, 2015), foram identificados 310 estabelecimentos no município de Belo Horizonte. Na Figura 3 é apresentada a distribuição espacial destes estabelecimentos em Belo Horizonte, sendo possível observar uma concentração de estabelecimentos ao longo dos principais corredores de tráfego e na área central de Belo Horizonte.

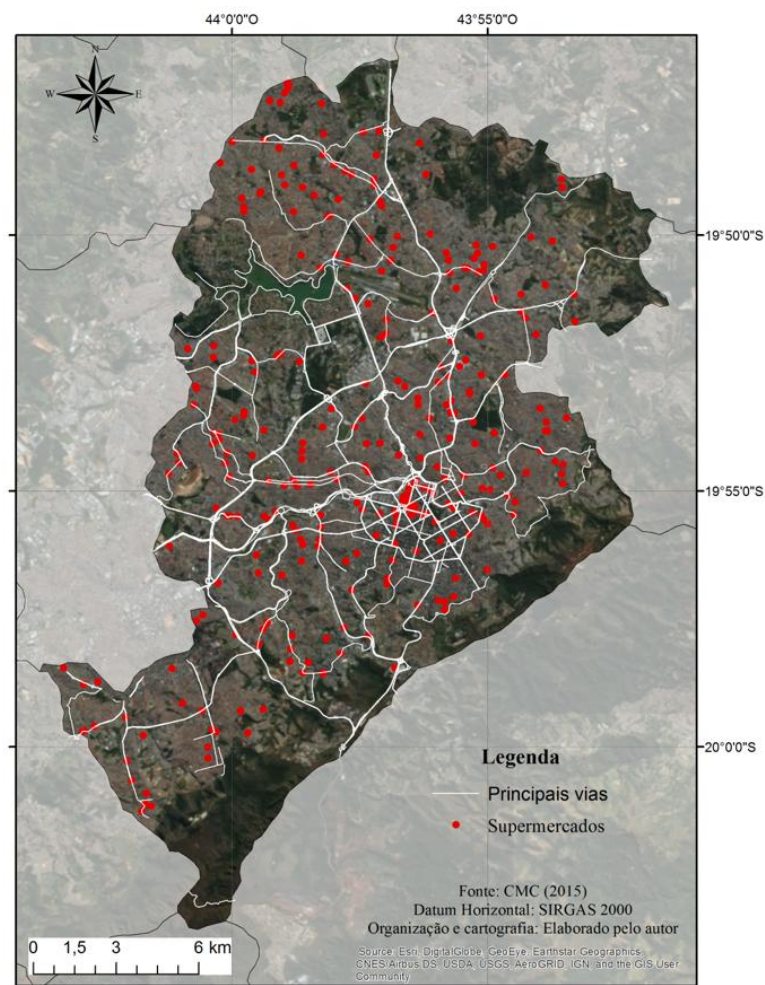


Figura 3: Supermercados e hipermercados no município de Belo Horizonte.

Além dos dados do CMC, foi utilizada também a base de dados proveniente da Pesquisa Origem e Destino da Região Metropolitana de Belo Horizonte de 2012 (Pesquisa OD RMBH de 2012). A Pesquisa OD foi o resultado do trabalho de parceria entre as Secretarias de Gestão Metropolitana (SEGEM) e de Transportes e Obras Públicas (SETOP) e os resultados constituem uma relevante ferramenta para o Sistema de Planejamento e Gestão Metropolitana. O objetivo principal da pesquisa foi coletar dados para conhecer a dinâmica dos deslocamentos das pessoas na RMBH, associando as características das viagens realizadas a variáveis socioeconômicas, permitindo identificar padrões nas viagens urbanas (MINAS GERAIS, 2013).

Por meio do tratamento dos resultados da Pesquisa OD de 2012, foram obtidos dados para a definição de variáveis independentes para o modelo, como o número de empregos em uma determinada região e o número de pessoas com destino à determinada região por motivo compras. Esta contabilização foi realizada para as **áreas homogêneas (AH)**, uma unidade espacial criada e utilizada especificamente para a Pesquisa Origem e Destino de Belo Horizonte e definida como

um conjunto de setores censitários agregados de acordo com critérios socioeconômicos e de ocupação, em unidades espaciais definidas para serem unidades de coleta adequada para estabelecer amostras mínimas em pesquisas socioeconômicas.

Outra fonte de dados utilizada nesta pesquisa foi o Censo 2010, realizado pelo IBGE (Instituto Brasileiro de Geografia e Estatística) com objetivo de retratar as características socioeconômicas da população brasileira, bem como as características dos domicílios (IBGE, 2017). Com esta base de dados, foi possível obter a população residente e a renda média *per capita* por setor censitário (unidade territorial de coleta das operações censitárias, definido pelo IBGE). Apesar da menor unidade de agregação dos dados do IBGE ser o setor censitário, foram também utilizadas as AH, para padronizar a unidade de análise com os dados da Pesquisa OD.

Por fim, para a obtenção de dados relativos à movimentação de cargas nos empreendimentos, utilizou-se uma pesquisa de campo realizada por Ramos (2015). Um funcionário do setor administrativo com conhecimento das operações de recebimento de mercadorias respondeu o questionário estruturado para a pesquisa (RAMOS, 2015; OLIVEIRA *et al.*, 2016). Na Tabela 4 são apresentadas as principais informações obtidas por Ramos (2015).

Tabela 4: Resumo das informações solicitadas aos entrevistados (RAMOS, 2015).

Categoria	Informação solicitada
Informações do Empreendimento	Nome e endereço do empreendimento
	Nome e função do entrevistado
	Telefone de contato e e-mail
Características do Empreendimento	Tempo de funcionamento
	Dias e horários de funcionamento
	Áreas dos empreendimentos
	Número de funcionários
	Número de <i>checkouts</i> (caixas), para supermercados
Recebimento de Mercadorias	Número de docas
	Vagas externas regulamentadas na via
	Tipos de mercadoria mais recebidas
	Tipos de veículos de carga atraídos
	Número médio de veículos de carga recebidos por dia da semana
	Horário das entregas
	Tempo de descarga
	Problemas verificados na operação

4 ABORDAGEM METODOLÓGICA

A abordagem metodológica desenvolvida para esse estudo consiste em nove etapas, conforme ilustrado na Figura 4. Nas subseções a seguir são detalhados os procedimentos de cada etapa.

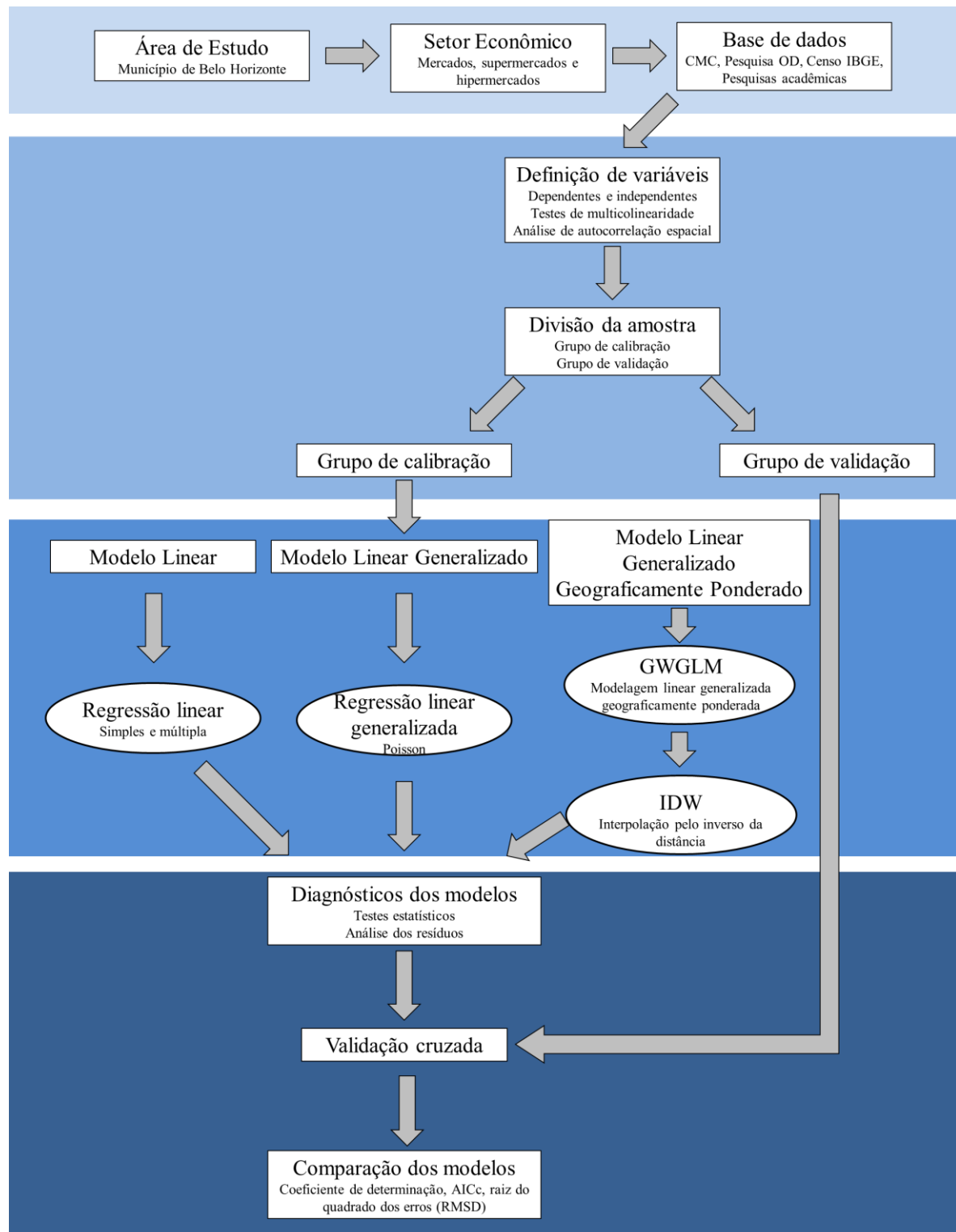


Figura 4: Fluxograma da abordagem metodológica do trabalho.

4.1 Variáveis do modelo

Para a construção de um modelo, é imprescindível a definição das variáveis independentes e dependentes. Conforme o exposto no Capítulo 2, o número de viagens de veículos de carga geradas pelo empreendimento foi considerado como variável dependente, devido sua maior utilização nos trabalhos nacionais e internacionais e pela maior facilidade de obtenção dos dados, se comparado com a variável volume de carga. Mais especificamente, foi utilizado neste estudo o número de viagens por dia de veículos de carga na sexta-feira, dia com maior movimentação de mercadorias (OLIVEIRA *et al.*, 2016).

Na Tabela 5 são apresentadas as variáveis independentes utilizadas neste estudo.

Tabela 5: Variáveis independentes selecionadas.

	Variável	Fonte	Unidade
VI 1	Área do empreendimento	CMC (2015)	m ²
VI 2	Número de <i>checkouts</i> (caixas)	RAMOS (2015)	Unidades
VI 3	Número de funcionários no empreendimento	RAMOS (2015)	Unidades
VI 4	Densidade populacional	IBGE (2017)	Hab/m ²
VI 5	Densidade de empregos	Pesquisa OD RMBH 2012	Empregos/m ²
VI 6	Densidade de pessoas que fazem compra na região	Pesquisa OD RMBH 2012	Pessoas/m ²
VI 7	Renda média <i>per capita</i>	IBGE (2017)	Salários mínimos

As três primeiras variáveis (VI 1, 2 e 3) referem-se a características físicas do estabelecimento que independem da localização do mesmo. Já as outras quatro variáveis independentes (VI 4, 5, 6 e 7) referem-se às características socioeconômicas da região ao entorno do empreendimento. Esta região ao entorno foi definida por meio dos limites da AH na qual o estabelecimento está localizado. A densidade populacional e a renda média *per capita* da população foram obtidas por meio dos dados fornecidos pelo IBGE. Já a densidade de empregos e de pessoas que fazem compras na AH foram obtidas por meio dos dados da Pesquisa OD da RMBH de 2012, conforme já explicado anteriormente no item 3.3 – Obtenção dos dados.

Optou-se pela utilização de densidade populacional, de empregos e de pessoas que fazem compra em detrimento ao número absoluto de cada uma das variáveis para atenuar as diferenças de valores entre AHs vizinhas. Como as áreas de análise, no caso as áreas

homogêneas, apresentam tamanhos variados, a adoção do número absoluto poderia prejudicar a calibração do modelo.

Após a definição das variáveis independentes, é necessária a realização de algumas análises para atestar a viabilidade de utilização destas variáveis nos modelos. Os itens a seguir abordam estas análises.

4.1.1 Análise de correlação entre variáveis

Segundo Farrar e Glauber (1967), é desejável que variáveis independentes tenham uma forte correlação linear com a variável dependente. Já entre variáveis independentes, não é desejável esta correlação, pois além de não agregar informação explicativa ao modelo, duas variáveis com estas características podem interferir no resultado final estimado.

4.1.2 Análise da autocorrelação espacial

A estatística espacial tem como principal objetivo a caracterização espacial entre diferentes dados (BAILEY e GATRELL, 1995). Quando em um conjunto de variáveis se observa um padrão de distribuição espacial, diz-se que tal conjunto possui uma dependência espacial. Uma consequência importante da dependência em uma amostra espacial é que a inferência estatística não será tão eficiente como a obtida em uma amostra independente do mesmo tamanho. Em outras palavras, a dependência leva a uma perda de informação. A perda de eficiência pode ser superada de três maneiras: aumentando o tamanho da amostra, desenvolvendo um esquema de amostragem que acumule observações de modo que sua interação seja insignificante ou pode ser levado em consideração esta dependência por meio de métodos estatísticos espacializados (ANSELIN, 1992).

A dependência espacial de uma variável é definida como autocorrelação espacial, cujo termo deriva da análise estatística que mensura o relacionamento entre duas variáveis (CÂMARA *et al.*, 2002). No entanto, ao contrário do conceito estatístico, a autocorrelação espacial consiste na comparação de uma variável com ela mesmo, porém em um lugar diferente no espaço, como indica o prefixo “auto” presente no termo.

O conceito de uma variável ser influenciada pela localização no espaço confirma o que Tobler (1970) enunciou e que hoje é conhecida como Primeira Lei da Geografia:

“No mundo, todas as coisas se parecem, mas coisas mais próximas são mais parecidas que aquelas mais distantes”

Segundo Câmara *et al.* (2002), a autocorrelação pode ser positiva ou negativa. Na autocorrelação positiva, a ocorrência de um evento influencia a ocorrência de um outro evento semelhante em suas proximidades. Portanto, a Primeira Lei da Geografia é respeitada. Já na autocorrelação negativa, a ocorrência de determinado evento dificulta a repetição de outro evento semelhante em seu entorno. Espacialmente, é possível concluir que uma autocorrelação positiva gera uma aglomeração de eventos no espaço e a negativa gera uma distribuição que tende a ser equidistante (TEIXEIRA, 2003).

Anselin (1995) define as técnicas para avaliar o grau de autocorrelação espacial como Análise Exploratória de Dados Espaciais (ESDA – *Exploratory Spatial Data Analysis*), sendo de suma importância para compreensão e desenvolvimento das diferentes etapas de modelagem espacial. As ferramentas de ESDA mais utilizadas são o Índice I Global de Moran e o Índice C de Geary. Ambos índices indicam o grau de associação espacial refletido no conjunto de dados como um todo. Porém, enquanto o I de Moran é baseado em produtos cruzados para medir a associação de valor, o C da Geary emprega diferenças ao quadrado (ANSELIN, 1992).

Neste trabalho, será utilizado o Índice I Global de Moran para a análise da autocorrelação espacial das variáveis. O Índice I Global de Moran varia de -1 a +1, sendo que valores positivos indicam que a variável analisada tende a ter uma correlação direta e os negativos, uma correlação inversa. Assim que calculado, deve-se estabelecer a validade estatística do modelo, pois trata-se de uma estatística inferencial, o que significa que os resultados da análise são sempre interpretados no contexto de sua hipótese nula que, neste caso, afirma que a variável que está sendo analisada é distribuída aleatoriamente na área de estudo (ESRI, 2017).

A abordagem mais comum para testar a hipótese é através do teste da pseudo-significância (CÂMARA *et al.*, 2002). O teste consiste em gerar inúmeras permutações de valores associadas no espaço, sendo que cada combinação os valores são redistribuídos no espaço, gerando um diferente arranjo espacial. Assim, é construída uma situação empírica do índice I de Moran Global, como apresentado na Figura 5, considerando um grau de confiança de 95%. A hipótese nula indica que os dados são distribuídos aleatoriamente em toda a área de estudo.

Caso o índice da situação real esteja nos extremos desta distribuição, é possível rejeitar a hipótese nula e considerar que o grupo testado possui autocorrelação espacial.

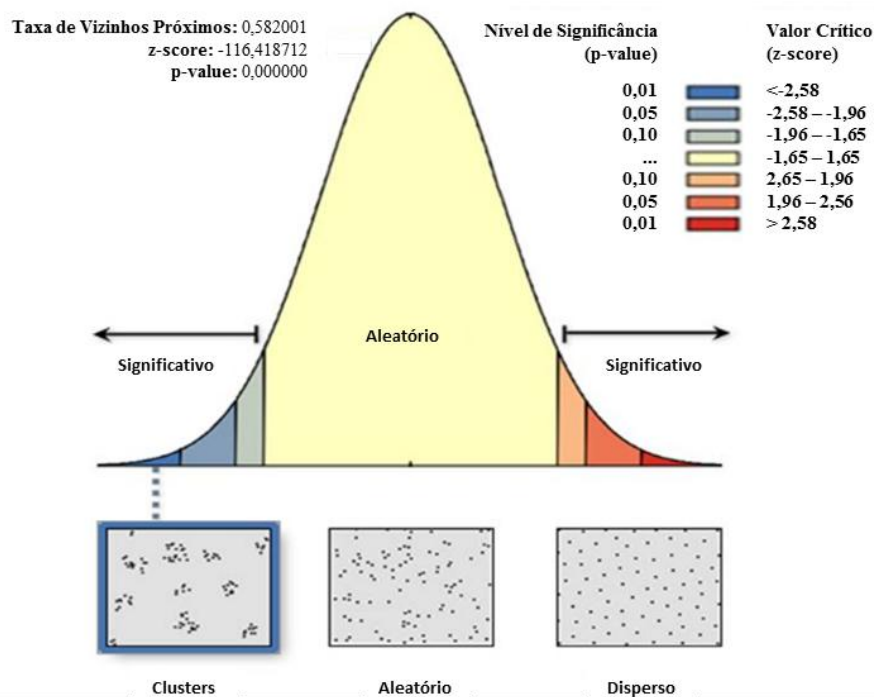


Figura 5: Teste da pseudo-significância – Índice I de Moran (ESRI, 2017).

4.2 Modelagem da geração de viagens de veículos de carga urbana

Para o desenvolvimento dos modelos de geração de viagens de veículos de carga urbana foram utilizadas técnicas de regressão. A regressão é utilizada para avaliar a relação entre duas ou mais variáveis para que seja possível compreender melhor o que acontece em determinado espaço e prever comportamentos futuros (ESRI, 2017). Neste estudo foram utilizadas três categorias diferentes de regressão: a regressão linear, modelos lineares generalizados e regressão geograficamente ponderada. Nos itens a seguir são apresentadas as descrições de cada um destes métodos.

4.2.1 Modelo de regressão linear

Por meio da revisão da literatura identificou-se que a regressão linear é o método mais utilizado para a modelagem da geração de viagens de veículos de carga. Portanto, neste trabalho também foi utilizada, a fim de se comparar os resultados obtidos com esta técnica com os outros métodos.

O Método dos Mínimos Quadrados (MMQ), ou ainda *Ordinary Least Squares* (OLS) é utilizado para a obtenção do modelo de regressão linear. Segundo Elian (1988), a regressão linear consiste em um método global e se baseia no uso de apenas uma equação para explicar a relação entre as variáveis estudadas (dependente e independentes). Neste caso, parte-se do pressuposto que apenas uma equação é capaz de explicar a relação entre as variáveis em toda área de estudo, ou seja, não é considerada a possibilidade de heterogeneidade no espaço. O MMQ pode ser considerado um ponto de partida adequado para a análise de regressão espacial (ESRI, 2017), que será também abordada neste estudo.

A regressão linear pode ser do tipo simples ou múltipla, sendo que a diferença está no número de variáveis explicativas. Na regressão simples, é utilizada apenas uma variável explicativa. O modelo de regressão linear simples é dado pela Equação 1.

$$Y_i = \beta_0 + \beta_1 X_1 + \varepsilon_1 \quad (1)$$

em que:

- Y_i é o i -ésimo valor da variável resposta;
- β_0 e β_1 são os parâmetros (coeficientes de regressão);
- X_1 é o i -ésimo valor da variável preditora (é uma constante conhecida, fixo);
- ε_i é o termo do erro aleatório com $E(\varepsilon_i) = 0$ e $\sigma^2(\varepsilon_i) = \sigma^2$, sendo σ^2 a variância;

O modelo de regressão linear múltipla é dado pela Equação 2.

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon_i \quad (2)$$

em que:

- $\beta_0, \beta_1, \dots, \beta_n$ são os parâmetros (coeficientes de regressão);
- X_1, X_2, \dots, X_n são constantes conhecidas;
- ε_i são independentes com distribuição $N(0, \sigma^2)$;
- $i = 1, 2, \dots, n$.

O método dos mínimos quadrados é utilizado para o ajuste da equação, sendo uma técnica matemática capaz de determinar o melhor ajuste de uma equação segundo um conjunto de dados. Neste método, a soma dos desvios verticais dos pontos em relação a reta gerada é zero e a soma dos quadrados dessas distâncias (ou desvios) é a mínima possível (CHARNET *et al.*, 1999).

Para identificar os melhores modelos, obtidos a partir da combinação das variáveis independentes e dependente, e verificar a validade do modelo, foram calculados os valores de R^2 , teste-t, p-value e AICc resultantes das regressões lineares. Tais combinações serão chamadas de cenários neste estudo. Ressalta-se que um modelo é considerado válido quando for confirmado estatisticamente. Deste modo, para um nível de confiança de 95%, o teste-t deve ser maior que 1,96 e o p-value menor que 0,05. Para a estatística R^2 , não há um valor mínimo para ser considerado válido, porém quanto maior o valor, melhor é o nível de ajuste do modelo. O Critério de Informação Akaike Corrigido (AICc) consiste em uma medida de desempenho do modelo. O AICc não consiste em uma medida absoluta de ajuste de um modelo, no entanto, é útil para comparar diferentes modelos de regressão com diferentes variáveis explicativas, mas com a mesma variável dependente.

Além disso, foi realizada a análise dos resíduos encontrados em cada modelo para avaliar a adequabilidade na estimativa. Esta análise foi realizada por meio de dois tipos de gráficos. O primeiro foi o gráfico de resíduos versus valores ajustados, cujo padrão, para um modelo bem ajustado, é a distribuição aleatória dos resíduos, centrada em zero e com variância constante.

Já o segundo foi o gráfico normal de probabilidade que representa em um dos eixos os quantis amostrais dos resíduos e no outro eixo, os quantis teóricos da distribuição normal. Caso o modelo seja bem ajustado, o gráfico segue um padrão de alinhamento dos pontos na reta que representa a identidade dos quantis amostrais e teóricos.

Foi utilizada a linguagem e o ambiente de computação estatística R (R CORE TEAM, 2017) para o desenvolvimento do modelo. O R fornece uma grande variedade de estatísticas e técnicas gráficas e está disponível como *software* livre. A biblioteca de dados utilizada neste estudo foi a Biblioteca Car (FOX e WEISBERG, 2011).

4.2.2 Modelo linear generalizado

Quando se deseja associar uma variável dependente com variáveis independentes, é comum se utilizar a modelagem linear. No entanto, uma limitação dos modelos lineares é a necessidade da variável dependente que compõe o modelo siga uma distribuição normal (DOBSON, 1997). Portanto, é necessário buscar uma alternativa para associar as variáveis dependentes e independentes de forma satisfatória.

Uma classe de modelos capaz de realizar esta associação entre variáveis independentes e uma variável dependente de comportamento não-normal são os modelos lineares generalizados. Segundo Dobson (1997), o modelo linear generalizado permite ajustar modelos de regressão para dados de resposta univariada que seguem uma distribuição da chamada família exponencial. A família exponencial inclui as distribuições normal, binomial, Poisson, geométrico, binomial negativa, exponencial, gamma e distribuição normal inversa.

Neste estudo, foi utilizada a distribuição de Poisson, largamente considerada para a análise de processos de chegada e números de acidentes. Essas variáveis, assim como o número de viagens geradas em um estabelecimento, são contagens, formando um conjunto de dados discretos (TURKMAN; SILVA, 2000).

O modelo linear generalizado com distribuição de Poisson é dado pela Equação 3.

$$\ln(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon_i \quad (3)$$

em que:

- $\beta_0, \beta_1, \dots, \beta_n$ são os parâmetros (coeficientes de regressão);
- X_1, X_2, \dots, X_n são constantes conhecidas;
- ε_i são independentes com distribuição $N(0, \sigma^2)$;
- $i = 1, 2, \dots, n$.

Para identificar os melhores cenários (combinações entre variáveis) e verificar a sua validade estatística, foram obtidos o AICc e o teste-z de cada cenário. Ressalta-se que para um modelo ser considerado válido, é preciso que haja validade estatística. Desse modo, para um nível de confiança de 95%, o teste-z deve ser maior que 1,96.

Realizou-se também a análise dos resíduos das regressões. A exemplo das regressões lineares, os modelos lineares generalizados também foram desenvolvidos com a utilização do *software* estatístico R (R CORE TEAM, 2017) e a Biblioteca Car (FOX e WEISBERG, 2011).

4.2.3 Modelo linear generalizado geograficamente ponderado

O modelo linear generalizado geograficamente ponderado pode ser dividido em duas etapas. A primeira consiste na utilização do Modelo Linear Generalizado Geograficamente Ponderado (ou *Geographically Weighted Generalised Linear Model* – GWGLM). Já a segunda etapa consiste na interpolação dos coeficientes encontrados na etapa anterior através do Método de Interpolação pelo Inverso da Distância (ou *Inverse Distance Weighting* – IDW).

4.2.3.1 Modelo linear generalizado geograficamente ponderado (GWGLM)

A técnica GWGLM tem como base a Regressão Geograficamente Ponderada (ou *Geographically Weighted Regression* – GWR), que é uma técnica de regressão espacial para construção de modelos e explicações de fenômenos muito utilizada em diferentes áreas do conhecimento. Ela é capaz de explorar características heterogêneas na área de estudo, pois consiste em um modelo de regressão local que possibilita a criação de uma equação diferente para cada elemento do conjunto de dados, incorporando variáveis dependentes e explicativas dos elementos vizinhos que se encontram dentro da distância limite de busca, chamada de largura de banda ou *bandwidth*, sendo assim capaz de capturar possíveis heterogeneidades no espaço (CHARLTON e FOTHERINGHAM, 2009, BASCUÑÁN e QUEZADA, 2016 e ESRI, 2017). Cabe ao usuário do modelo determinar a forma e o tamanho da largura de banda que será considerada no modelo.

Ao contrário dos modelos de regressão globais, como é o caso da MMQ, o GWR estima um modelo para cada elemento do banco de dados, permitindo identificar variações locais das variáveis. Em outras palavras, o GWR pode ser considerado um modelo de regressão com parâmetros geograficamente variáveis (NAKAYA, 2012). O GWR convencional é descrito pela Equação 4.

$$y_i = \sum_k \beta_k(u_i, v_i) x_{k,i} + \varepsilon_i \quad (4)$$

em que:

- y_i : variável independente
- $x_{k,i}$: k-ésima variável independente;
- ε_i : erro gaussiano para a localização i
- (u_i, v_i) : coordenada x-y para a i -ésima localização;
- $\beta_k(u_i, v_i)$: coeficientes que variam de acordo com a localização.

Ao contrário do modelo global, o GWR é capaz de retornar diferentes valores de $\beta_0, \beta_1, \dots, \beta_k$ considerando as características do entorno daquele ponto (CARVALHO *et al.*, 2006; CHARLTON e FOTHERINGHAM, 2009). Por premissa, assume-se que os resíduos do modelo devem ser independentes e normalmente distribuídos com média igual a zero. Utilizando a estrutura de modelagem linear generalizada geograficamente ponderada (GWGLM), é possível fazer uso de regressão logística ou de Poisson com coeficientes geograficamente ponderados, úteis para modelagem de dados de contagem (NAKAYA, 2012), como é o caso da geração de viagens. Neste trabalho, foi utilizada a regressão Poisson. O GWGLM utilizando a regressão de Poisson é descrito pela equação:

$$\ln(y_i) = \sum_k \beta_k(u_i, v_i) x_{k,i} + \varepsilon_i \quad (5)$$

em que:

- y_i : variável independente
- $x_{k,i}$: k-ésima variável independente;
- ε_i : erro gaussiano para a localização i
- (u_i, v_i) : coordenada x-y para a y-ésima localização;
- $\beta_k(u_i, v_i)$: coeficientes que variam de acordo com a localização.

Nesta etapa foram realizadas regressões lineares generalizadas geograficamente ponderadas utilizando o mesmo grupo de calibração e as mesmas combinações de variáveis dos modelos anteriores, compondo assim os cenários. Para identificar os melhores cenários e verificar a validade dos mesmos, foram calculados os valores da estatística AICc para cada combinação de variáveis. Segundo Câmara *et al.* (2002), como os tradicionais estimadores de regressão não considera os efeitos espaciais, as indicações de qualidade de ajuste baseadas em R^2 são incorretas.

No presente estudo, foi utilizado o *software* GWR4² para construção dos modelos. O GWR4 é um *software* livre e permite o desenvolvimento dos modelos GWR e GWGLM com testes estatísticos associados e seleções de modelo por dados definidos pelo usuário e configurações do modelo (NAKAYA, 2012).

² O *software* GWR4 foi desenvolvido e programado pelo professor Tomoki Nakaya do Departamento de Geografia da Universidade Ritsumeikan de Kyoto, Japão.

Foi utilizado nos cenários desenvolvidos o *kernel* do tipo “*Fixed Gaussian (distance)*”. Já para como método de seleção “*Bandwidth*” ou “largura de banda”, foi utilizada a opção “*Golden section search*”

Após a construção dos cenários utilizando o modelo GWGLM, passou-se para a segunda etapa do método espacial, que é a interpolação dos resultados obtidos nesta etapa.

4.2.3.2 Interpolação IDW

A interpolação dos dados foi realizada para obter os coeficientes de regressão para qualquer ponto da área de estudo e, conseqüentemente, estimar viagens de veículos carga urbana para outros empreendimentos que não pertençam ao conjunto de dados utilizados para a execução do GWGLM.

A técnica de interpolação consiste na estimativa de valores de atributos em um espaço contínuo a partir de um conjunto de pontos amostrados em uma área. Existem dois grupos de técnicas de interpolação: determinísticas e geoestatísticas. As técnicas de interpolação determinísticas são capazes de criar superfícies baseando-se em pontos medidos e considerando a extensão de similaridade (por exemplo, a interpolação ponderada pelo inverso da distância) ou no grau de suavização (por exemplo, as funções base radiais). Por sua vez, as técnicas de interpolação geoestatística (por exemplo, a krigagem) utilizam as propriedades estatísticas dos pontos medidos, quantificando a autocorrelação espacial entre pontos medidos e representando a configuração espacial dos pontos de amostra em torno do local de predição (ESRI, 2017).

A técnica de interpolação escolhida para este estudo foi o inverso da distância, ou IDW (*Inverse Distance Weighting*), que consiste em uma técnica de interpolação determinística que cria uma superfície a partir dos pontos medidos, com base na extensão da similaridade, colocando os pontos mais próximos do local de previsão maior do que aqueles mais distantes, por isso o nome da distância inversa ponderada (JOHNSTON *et al.*, 2001). Neste modelo, a superfície que está sendo interpolada deve ser a de uma variável localmente dependente (ESRI, 2017).

Como resultado da interpolação, tem-se uma superfície *raster* para os coeficientes de cada variável independente empregada no modelo GWGLM e uma superfície *raster* para o valor do intercepto, parte integrante da equação. A partir destas superfícies, foi possível definir uma equação de geração de viagens de veículos de carga para cada localidade da área de estudo e, através de álgebra de mapas, gerar mapas de geração de viagens. A álgebra de mapas consiste

em um desdobramento da álgebra tradicional, porém, utilizando campos geográficos como variáveis a serem manipuladas (BERRY, 1993).

Na etapa de interpolação, foi utilizado o *software* ArcGIS para a execução desta etapa de interpolação. Desenvolvido pela ESRI (*Environmental Systems Research Institute*), o ArcGIS é um Sistema de Informação Geográfica (GIS) muito utilizado em todo o mundo para a criação de mapas, compilação, análise e gestão de dados geográficos em geral.

4.3 Validação cruzada dos modelos

A validação cruzada, também conhecida como *cross validation*, é uma técnica utilizada para avaliar a capacidade explicativa de um modelo de predição construído a partir de um banco de dados (KOHAVI, 1995). Portanto, o objetivo principal da validação cruzada neste trabalho foi de estimar a precisão do modelo.

Segundo Kokavi (1995), a validação cruzada pode ser utilizada para a escolha do melhor modelo a ser utilizado para a estimativa de determinado fenômeno. Portanto, neste estudo, a validação cruzada foi utilizada para testar a adequabilidade dos modelos de geração de viagens de veículos de carga utilizando novos dados e assim evitar o sobreajuste dos dados por parte dos modelos.

Dentre os métodos de validação cruzada, foi utilizado o de *holdout*, que divide a amostra em um grupo de calibração e um grupo de validação (KOKAVI, 1995) para estimar a precisão do modelo. O grupo de calibração é um conjunto de dados utilizado para o desenvolvimento do modelo de predição. O grupo de validação é um conjunto utilizado para a validação do modelo elaborado a partir dos dados de calibração, a fim de medir a acurácia do modelo desenvolvido na predição ou considerando novos dados. Com o grupo de calibração e de validação, foi possível comparar os valores reais com os valores preditos pelo modelo. A escolha dos estabelecimentos para cada um dos grupos foi realizada de modo aleatório.

4.4 Comparação dos modelos

Após a realização da validação cruzada dos resultados dos três métodos de estimativa de viagens utilizados (regressão linear, GLM e GWGLM), os resultados foram comparados para identificar a técnica que melhor explica o fenômeno de geração de viagens de veículos de carga.

5 RESULTADOS E DISCUSSÕES

Neste capítulo serão apresentados os passos e resultados obtidos no processo de construção dos modelos de geração de viagens de veículos de carga urbana. Por meio da aplicação da abordagem metodológica proposta para este trabalho, primeiramente faz-se necessária a análise e tratamento dos dados, a fim de que possam ser utilizadas como variáveis dependentes e independentes no modelo a ser elaborado, além da definição da amostra de pesquisa. Após estas etapas, são apresentados os passos realizados e resultados obtidos na elaboração dos cenários segundo as três técnicas de modelagem descritas anteriormente: modelo de regressão linear, modelo GLM e GWGLM. Por fim, são apresentados os resultados da validação cruzada dos modelos e análise do método que apresenta a melhor acurácia na predição.

5.1 Análise de correlação entre variáveis

A matriz de correlação linear (Tabela 6) foi gerada calculando os coeficientes de correlação de Pearson entre as variáveis dependente e independentes utilizadas no estudo.

Tabela 6: Matriz de correlação entre variáveis dependente e independentes.

	VD1 – Nº de entregas na sexta-feira	VI1 – Área do empreendimento	VI2 – Nº de <i>checkouts</i>	VI3 – Número de funcionários	VI4 – Densidade populacional	VI5 – Densidade de compras	VI6 – Densidade de empregos	VI7 – Renda média <i>per capita</i> da população do entorno
VD1 – Número de viagens geradas na sexta-feira	1,00	0,77	0,79	0,78	0,05	0,04	-0,06	0,18
VI1 – Área do empreendimento	-	1,00	0,67	0,77	0,37	0,24	0,18	0,51
VI2 – Número de <i>checkouts</i>	-	-	1,00	0,90	-0,05	0,19	0,21	0,28
VI3 – Número de funcionários	-	-	-	1,00	0,17	0,10	0,11	0,44
VI4 – Densidade populacional da área ao entorno	-	-	-	-	1,00	-0,01	-0,23	0,14
VI5 – Densidade de compras da área ao entorno	-	-	-	-	-	1,00	0,74	0,13
VI6 – Densidade de empregos da área ao entorno	-	-	-	-	-	-	1,00	0,32
VI7 – Renda média <i>per capita</i> da população do entorno	-	-	-	-	-	-	-	1,00

Na Tabela 6 foram destacados os valores acima de 0,50, critério adotado para auxiliar nas análises no momento da especificação dos modelos tradicionais. Pode-se observar através das duas primeiras linhas da matriz que as variáveis independentes “área do empreendimento”, “número de *checkouts*” e “número de funcionários” apresentam altos índices de correlação com a variável dependente “número de entregas na sexta-feira”. Observa-se que há correlação linear entre as três variáveis relacionadas à estrutura física do empreendimento (VI1, VI2 e VI3). Além disso, constatou-se correlação entre a variável “densidade de pessoas que fazem compra na região” e a variável “densidade de emprego na região”. Esta análise de multicolinearidade entre variáveis será utilizada na elaboração das combinações de variáveis que comporão os diferentes cenários modelados.

Destaca-se que as variáveis independentes territoriais tiveram uma pequena correlação linear com a variável dependente, porém serão importantes na construção do modelo GWGLM e com isso podem se tornar explicativas.

5.2 Análise da autocorrelação espacial

A análise de autocorrelação espacial das variáveis permite identificar a dependência espacial das variáveis que pode prejudicar a acurácia do modelo não espacial. Porém, uma das formas de contornar o problema é considerar esta dependência por meio de métodos estatísticos espacializados. Portanto, caso seja identificado a autocorrelação espacial nas variáveis utilizadas, a utilização do modelo GWGLM desenvolvido neste trabalho torna-se justificável.

São apresentados na Tabela 7 os resultados obtidos após a execução do teste de autocorrelação I de Moran Global. Como apresentado anteriormente na Figura 5, é possível afirmar com um nível de confiança de 95%, se uma variável possui autocorrelação caso o resultado do z-score for menor que 1,96 (autocorrelação inversa) ou maior que 1,96 (autocorrelação direta) e o p-value for menor que 0,05.

Tabela 7: Índices de autocorrelação espacial (I de Moran Global) das variáveis dependente e independentes.

Variáveis	I de Moran Global	z score	p-value
VD1 – Número de entregas na sexta-feira	0,343	0,484	0,628
VI1 – Área do empreendimento	0,166	0,286	0,774
VI2 – Número de <i>checkouts</i>	0,173	0,276	0,782
VI3 – Número de funcionários	0,274	0,385	0,7
VI4 – Densidade populacional da área ao entorno	0,0717	3,307	0,001
VI5 – Densidade de compras da área ao entorno	1,2239	61,23	0,001
VI6 – Densidade de empregos da área ao entorno	0,317	15,46	0,001
VI7 – Renda média <i>per capita</i> da população do entorno	0,204	5,401	0,001

Nos índices apresentados na Tabela 7 observa-se que as variáveis independentes VI 4, VI 5, VI 6 e VI 7 apresentam autocorrelação espacial, indicando que tais variáveis são espacialmente dependentes. Este resultado já era esperado por se tratar de variáveis de domínio geográfico (dados demográficos e socioeconômicos). Tal fato justifica a utilização de um modelo que considere as características locais das variáveis. Já as variáveis correlacionadas às características físicas do estabelecimento não apresentaram esta correlação espacial.

5.3 Dados do modelo

Neste estudo, os dados brutos obtidos por Ramos (2015) foram utilizados para o desenvolvimento dos modelos. Os dados foram analisados e excluindo os *outlayers* obteve-se uma amostra de 27 estabelecimentos com respostas consideradas válidas. Esta amostra foi dividida em dois grupos: um grupo de calibração com 22 estabelecimentos e um grupo de validação com 5 estabelecimentos. Conforme apresentado na Figura 7, os pontos do grupo de validação se encontram distribuídos ao longo de toda a área de estudo.

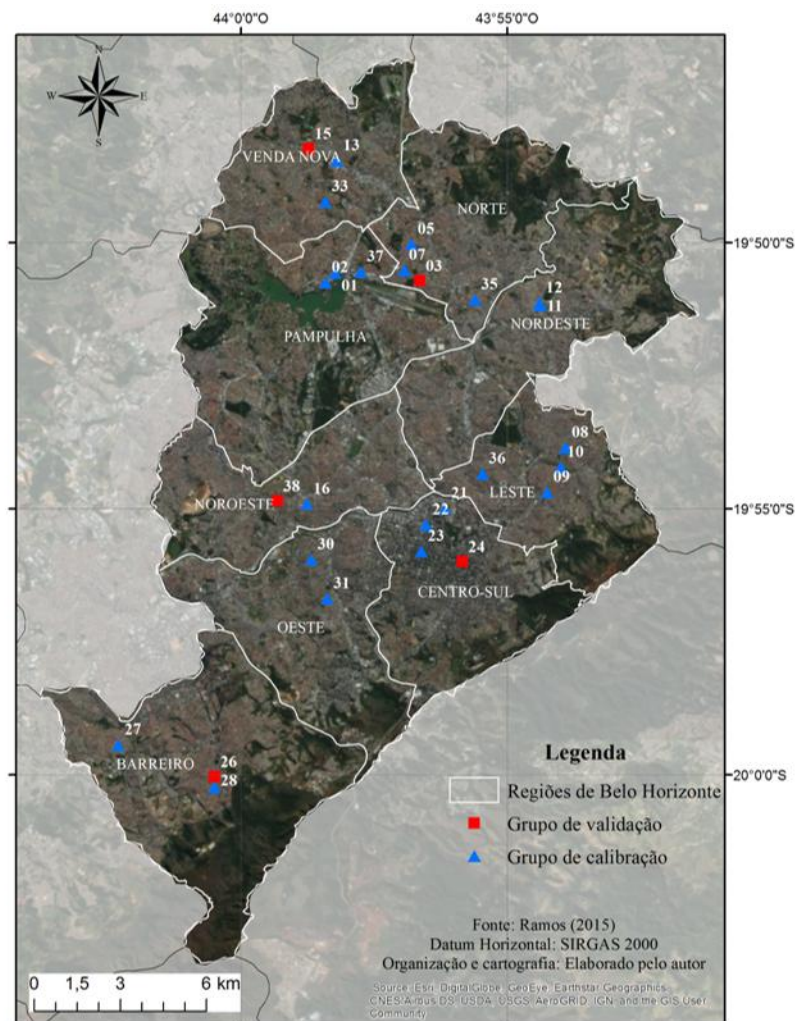


Figura 6: Estabelecimentos com respostas validadas.

5.4 Modelo linear

Nesta etapa do estudo foi utilizada a técnica de modelagem linear para a elaboração de diversos cenários utilizando o grupo de calibração com diferentes combinações entre variáveis dependente e independentes. Estas combinações foram elaboradas evitando que aquelas variáveis independentes que apresentam uma alta colinearidade entre si estejam no mesmo modelo (multicolinearidade). A multicolinearidade deve ser evitada, pois pode prejudicar a calibração do modelo e aumentar a variância dos coeficientes da regressão. Os resultados das análises de correlação linear entre as variáveis serão apresentados no item 5.1.1 – Análise de correlação linear entre variáveis.

São apresentados na Tabela 8 os cenários elaborados com os respectivos valores dos coeficientes que compõe a equação de previsão de viagens e os resultados de R^2 , teste-t, p-value e AICc resultantes das regressões lineares utilizando o grupo de calibração.

Tabela 8: Modelos de geração de viagens de veículos de carga utilizando regressão linear.

Cenário	Variável Dependente	Variáveis Independentes	Valor do intercepto	Valor dos coeficientes	R2	Teste t		p value	AICc	RMSD
						Intercepto	Var. Independente			
LM 01	Entrega sexta-feira	Área	4,528	0,007	0,68	2,52	6,48	0,000	144,62	7,11
LM 02	Entrega sexta-feira	Funcionários	1,026	0,191	0,66	0,44	6,17	0,000	146,05	6,52
LM 03	Entrega sexta-feira	<i>Checkout</i>	-7,111	2,610	0,74	-2,45	7,47	0,000	140,19	8,50
LM 04	Entrega sexta-feira	Área	10,475	0,008	0,74	3,36	7,42	0,000	141,46	6,90
		Densidade populacional		-675,044			-2,24	0,000		
LM 05	Entrega sexta-feira	Área	5,070	0,007	0,70	2,79	6,70	0,000	144,80	6,66
		Densidade emprego		-55,359			-1,28	0,000		
LM 06	Entrega sexta-feira	Área	4,722	0,007	0,70	2,63	6,61	0,000	145,17	7,14
		Densidade compras		-132,857			-1,14	0,000		
LM 07	Entrega sexta-feira	Área	5,649	0,007	0,70	2,80	6,17	0,000	145,06	5,96
		Renda		-0,702			-1,19	0,000		
LM 08	Entrega sexta-feira	Área	13,820	0,008	0,82	4,64	8,97	0,000	136,02	6,22
		Densidade populacional		-943,490			-3,37	0,000		
		Densidade emprego		-100,094			-2,69	0,000		
LM 09	Entrega sexta-feira	Área	11,082	0,008	0,77	3,64	7,80	0,000	140,83	6,90
		Densidade populacional		-717,757			-2,45	0,000		
		Densidade compras		-157,483			-1,51	0,000		
LM 10	Entrega sexta-feira	Área	12,267	0,008	0,78	3,83	7,39	0,000	140,54	5,68
		Densidade populacional		-725,351			-2,49	0,000		
		Renda		-0,844			-1,60	0,000		

Tabela 8: Modelos de geração de viagens de veículos de carga utilizando regressão linear – continuação.

Cenário	Variável Dependente	Variáveis Independentes	Valor do intercepto	Valor dos coeficientes	R2	Teste t		p value	AICc	RMSD
						Intercepto	Var. Independente			
LM 11	Entrega sexta-feira	Área	5,937	0,007	0,72	2,93	6,25	0,000	145,65	5,87
		Densidade emprego		-47,862			-1,09	0,000		
		Renda		-0,588			-0,98	0,000		
LM 12	Entrega sexta-feira	Área	5,824	0,008	0,72	2,90	6,32	0,000	145,54	6,02
		Densidade compras		-130,741			-1,13	0,000		
		Renda		-0,691			-1,18	0,000		
LM 13	Entrega sexta-feira	Área	14,991	0,009	0,84	4,99	8,75	0,000	135,51	5,65
		Densidade populacional		-963,068			-3,54	0,000		
		Densidade emprego		-92,467			-2,53	0,000		
		Renda		-0,671			-1,44	0,000		
LM 14	Entrega sexta-feira	Área	12,861	0,009	0,80	4,15	7,84	0,000	139,55	5,70
		Densidade populacional		-767,582			-2,73	0,000		
		Densidade compras		-156,623			-1,57	0,000		
		Renda		-0,839			-1,65	0,000		
LM 15	Entrega sexta-feira	Funcionários	2,747	0,194	0,66	0,72	6,07	0,000	147,68	6,44
		Densidade populacional		-184,854			-0,57	0,000		
LM 16	Entrega sexta-feira	Funcionários	1,339	0,192	0,66	0,56	6,09	0,000	147,65	6,09
		Densidade emprego		-26,721			-0,59	0,000		
LM 17	Entrega sexta-feira	Funcionários	1,086	0,192	0,66	0,45	6,01	0,000	148,01	6,51
		Densidade compras		-23,683			-0,20	0,000		
LM 18	Entrega sexta-feira	Funcionários	1,495	0,202	0,66	0,61	5,65	0,000	147,57	5,81
		Renda		-0,393			-0,65	0,000		

Tabela 8: Modelos de geração de viagens de veículos de carga utilizando regressão linear – continuação.

Cenário	Variável Dependente	Variáveis Independentes	Valor do intercepto	Valor dos coeficientes	R2	Teste t		p value	AICc	RMSD
						Intercepto	Var. Independente			
LM 19	Entrega sexta-feira	Funcionários	3,870	0,197	0,67	0,93	6,05	0,000	148,97	5,81
		Densidade populacional		-259,015			-0,75	0,000		
		Densidade emprego		-36,882			-0,77	0,000		
LM 20	Entrega sexta-feira	Funcionários	2,821	0,195	0,66	0,71	5,91	0,000	149,63	6,43
		Densidade populacional		-185,986			-0,56	0,000		
		Densidade compras		-24,840			-0,20	0,000		
LM 21	Entrega sexta-feira	Funcionários	3,134	0,205	0,67	0,79	5,56	0,000	149,22	5,77
		Densidade populacional		-177,542			-0,54	0,000		
		Renda		-0,381			-0,62	0,000		
LM 22	Entrega sexta-feira	Funcionários	1,666	0,201	0,67	0,66	5,50	0,000	149,33	5,59
		Densidade emprego		-21,080			-0,44	0,000		
		Renda		-0,329			-0,52	0,000		
LM 23	Entrega sexta-feira	Funcionários	1,525	0,202	0,66	0,60	5,50	0,000	149,55	5,82
		Densidade compras		-15,751			-0,13	0,000		
		Renda		-0,384			-0,62	0,000		
LM 24	Entrega sexta-feira	Funcionários	3,989	0,205	0,68	0,94	5,46	0,000	150,73	5,42
		Densidade populacional		-242,614			-0,69	0,000		
		Densidade emprego		-31,411			-0,62	0,000		
		Renda		-0,281			-0,43	0,000		
LM 25	Entrega sexta-feira	Funcionários	3,175	0,205	0,67	0,78	5,41	0,000	151,20	5,78
		Densidade populacional		-178,495			-0,52	0,000		
		Densidade compras		-17,119			-0,14	0,000		
		Renda		-0,372			-0,58	0,000		

Tabela 8: Modelos de geração de viagens de veículos de carga utilizando regressão linear – continuação.

Cenário	Variável Dependente	Variáveis Independentes	Valor do intercepto	Valor dos coeficientes	R2	Teste t		p value	AICc	RMSD
						Intercepto	Var. Independente			
LM 26	Entrega sexta-feira	<i>Checkout</i>	-10,085	2,627	0,75	-2,41	7,51	0,000	141,09	8,64
		Densidade populacional		273,380			0,99	0,000		
LM 27	Entrega sexta-feira	<i>Checkout</i>	-6,890	2,685	0,76	-2,43	7,77	0,000	140,01	7,73
		Densidade emprego		-54,361			-1,41	0,000		
LM 28	Entrega sexta-feira	<i>Checkout</i>	-7,194	2,675	0,75	-2,48	7,51	0,000	141,10	8,47
		Densidade compras		-103,287			-0,98	0,000		
LM 29	Entrega sexta-feira	<i>Checkout</i>	-7,547	2,541	0,74	-2,51	6,90	0,000	141,67	9,05
		Renda		0,326			0,68	0,000		
LM 30	Entrega sexta-feira	<i>Checkout</i>	-8,960	2,688	0,77	-2,11	7,66	0,000	141,49	7,92
		Densidade populacional		187,618			0,66	0,000		
		Densidade emprego		-47,402			-1,17	0,000		
LM 31	Entrega sexta-feira	<i>Checkout</i>	-10,199	2,694	0,76	-2,44	7,55	0,000	141,92	8,62
		Densidade populacional		276,131			1,00	0,000		
		Densidade compras		-104,338			-0,99	0,000		
LM 32	Entrega sexta-feira	<i>Checkout</i>	-10,214	2,570	0,75	-2,39	6,91	0,000	142,73	9,08
		Densidade populacional		252,360			0,89	0,000		
		Renda		0,267			0,55	0,000		
LM 33	Entrega sexta-feira	<i>Checkout</i>	-7,495	2,596	0,77	-2,58	7,28	0,000	140,82	8,39
		Densidade emprego		-62,276			-1,58	0,000		
		Renda		0,476			1,00	0,000		
LM 34	Entrega sexta-feira	<i>Checkout</i>	-7,696	2,602	0,76	-2,56	6,99	0,000	142,40	9,09
		Densidade compras		-110,575			-1,04	0,000		
		Renda		0,370			0,77	0,000		

Tabela 8: Modelos de geração de viagens de veículos de carga utilizando regressão linear – continuação.

Cenário	Variável Dependente	Variáveis Independentes	Valor do intercepto	Valor dos coeficientes	R2	Teste t		p value	AICc	RMSD
						Intercepto	Var. Independente			
LM 35	Entrega sexta-feira	<i>Checkout</i>	-8,953	2,607	0,78	-2,09	7,14	0,000	142,54	8,47
		Densidade populacional		137,514			0,47	0,000		
		Densidade emprego		-56,407			-1,34	0,000		
		Renda		0,429			0,87	0,000		
LM 36	Entrega sexta-feira	<i>Checkout</i>	-10,355	2,631	0,77	-2,43	6,99	0,000	143,41	9,13
		Densidade populacional		251,750			0,89	0,000		
		Densidade compras		-110,376			-1,03	0,000		
		Renda		0,312			0,63	0,000		

É apresentado na Tabela 8, para cada cenário, o valor do intercepto e dos coeficientes de cada variável independente que compõe o modelo. Por meio dela, é possível compor as equações de geração de viagens de veículos de carga e estimar a geração para qualquer estabelecimento do mesmo setor econômico de estudo desde que os modelos sejam válidos estatisticamente.

Percebe-se que, para aqueles cenários com a presença da variável independente número de *checkout*, o valor do intercepto é negativo. A estimativa de valores tanto positivos, quanto negativos, é uma das características do modelo linear, pois ele assume uma distribuição normal. No entanto, cabe ressaltar que o número de viagens de veículos de carga geradas deve ser sempre positivo ou, no mínimo, nulo. Por conta disso, é primordial respeitar o domínio de cada variável do modelo. Estes domínios são os seguintes:

- Área do estabelecimento (m²): [171, 5300];
- Número de funcionários: [3, 180];
- Número de *checkouts*: [3, 14].

Considerando os cenários utilizando os modelos lineares desenvolvidos e apresentados na Tabela 8, foi possível constatar que todos os cenários analisados apresentaram coeficientes de determinação (R²) satisfatórios, variando de 0,66 até o 0,84 (observado no modelo 13). As equações associadas à área do estabelecimento e ao número de *checkouts* apresentaram os melhores valores de R², resultado semelhante ao encontrado nos trabalhos de Gasparini (2008) e Ramos (2015). Já as equações associadas ao número de funcionários apresentaram menores coeficientes de determinação.

Foi necessário também observar a validade estatística dos modelos construídos em cada cenário através do teste-t e do p-value. Na Tabela 8, os cenários destacados (linhas sombreadas) foram considerados estatisticamente válidos seguindo os critérios apresentados no item 4.2.1. Observa-se que apenas quatro deles (LM 01, LM 03, LM 04 e LM 08) foram validados. Nos outros cenários, algumas variáveis apresentaram resultados de teste-t estatisticamente válidos, porém as outras variáveis do modelo não.

Esse resultado já era esperado, pois a análise de correlação entre as variáveis dependente e as independentes realizada anteriormente (item 5.1 – Análise de correlação entre variáveis) apontou uma correlação apenas entre o número de entregas na sexta-feira com a área do empreendimento, número de funcionários e número de *checkouts*.

Com relação ao AICc, o índice variou de 135,51 a 151,20, sendo que quanto menor o valor, melhor é o modelo. Os melhores resultados foram observados naqueles modelos com a presença da variável independente número de *checkouts* e aos piores resultados foram observados nos modelos com a presença da variável independente número de funcionários.

Foi realizada também a análise gráfica dos resíduos das regressões lineares de cada modelo. As figuras a seguir apresentam os gráficos gerados a partir dos resultados obtidos nos modelos LM 01 e LM 18. No Apêndice A são apresentados os gráficos referentes aos outros cenários.

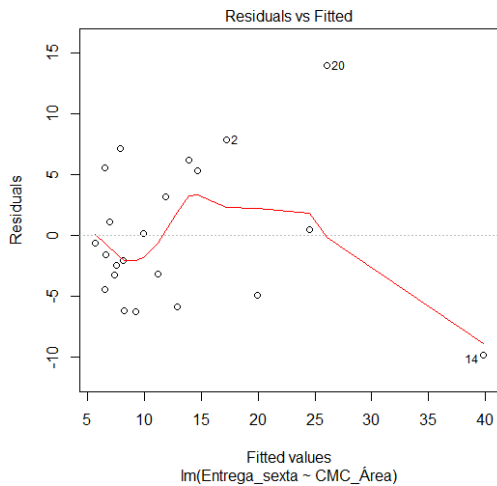


Figura 7: Resíduos versus valores ajustados – LM 01.

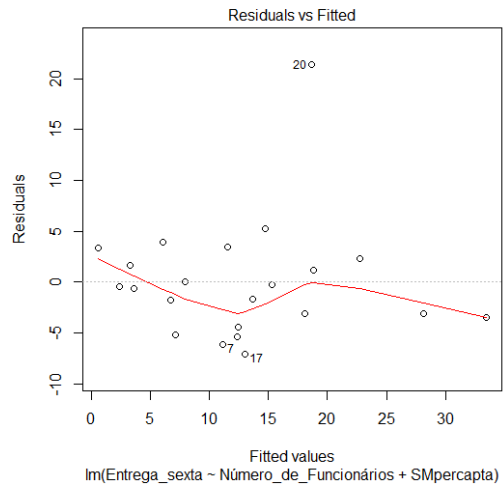


Figura 8: Resíduos versus valores ajustados – LM 18.

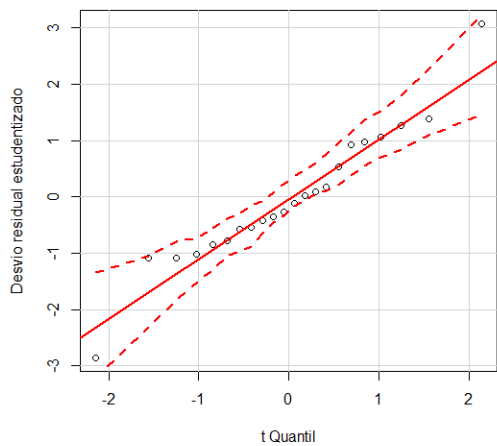


Figura 9: Gráfico normal de probabilidades – LM 01.

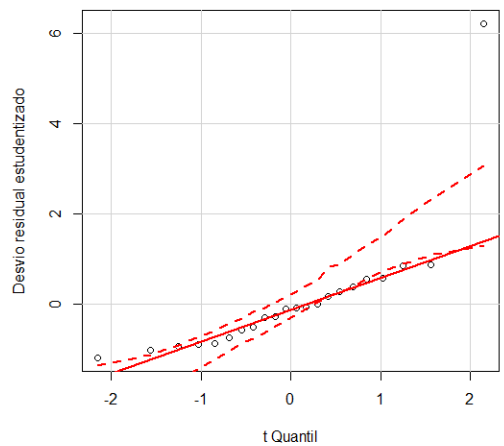


Figura 10: Gráfico normal de probabilidades – LM 18.

São apresentados na Figura 7 e na Figura 8 os gráficos de resíduos versus valores ajustados e percebe-se que para o cenário LM 01, a distribuição dos resíduos foi próximo à normal, apesar de possuir dois *outlayers* (pontos 14 e 20). O mesmo pode ser percebido para o cenário LM 18. Portanto, ambos os modelos não apresentam variância constante.

Já na Figura 9 e na Figura 10 são apresentados os gráficos normais de probabilidades e percebe-se que para o cenário LM 01, os pontos seguem próximos à reta, indicando um bom ajuste. No entanto, no cenário LM 18, apesar da maioria dos pontos seguirem o alinhamento da reta, o resíduo do ponto 20 se distancia muito da mesma. Desta forma, realizando a análise visual dos resíduos dos outros modelos lineares gerados, pode-se concluir que os cenários LM 01, 05, 06, 07, 09, 11 e 12 apresentaram resíduos dentro da normalidade.

Outra forma de avaliar o desempenho da predição é através da validação cruzada. Foi realizada, em cada um dos modelos lineares gerados, esta validação utilizando o grupo de validação. Através da equação gerada em cada modelo, foram estimados os valores de geração de viagens para cada um dos cinco estabelecimentos que formam o grupo de validação.

Feito isso, foi calculada a média dos quadrados dos erros (*root-mean-square deviation* – RMSD), que consiste no desvio padrão dos residuais (diferença entre valor observado e estimado). O RMSD é um importante meio de comparar erros de previsão de diferentes cenários. Os resultados obtidos para cada cenário também se encontram na Tabela 8.

Os resultados indicam que o valor médio do RMSD foi de 6,97. O menor valor observado foi de 5,42, no cenário LM 24 e o maior valor foi de 9,13, no cenário LM 36. Os cenários com a presença da variável independente número de funcionários apresentaram os menores índices e, conseqüentemente, a melhor qualidade na predição. Aqueles cenários com a variável número de *checkouts* apresentaram os piores resultados. Considerando apenas os quatro cenários validados estatisticamente por meio do teste-t e do p-value, a média foi de 7,18. Já considerando aqueles modelos com os resíduos dentro da normalidade, a média foi de 6,52.

Destaca-se ainda que, conforme apresentado na Tabela 8, os cenários que não foram considerados válidos estatisticamente através dos resultados do teste-t e da análise visual dos resíduos apresentaram resultados na validação cruzada iguais ou melhores que aqueles considerados válidos por meio dos testes estatísticos realizados.

5.5 Modelo linear generalizado

A segunda técnica de modelagem utilizada neste estudo foi o modelo linear generalizado. Conforme já mencionado no capítulo de Abordagem Metodológica, devido ao fato de que o número de viagens de veículos de carga em áreas urbanas (variável dependente do modelo) forma um conjunto de dados discretos, um modelo linear generalizado tenderia a apresentar melhores resultados de calibração se comparados com modelos convencionais de regressão linear.

São apresentados na Tabela 9 os cenários elaborados com os respectivos valores dos coeficientes que compõe a equação de previsão de viagens e os resultados do teste-z e AICc dos modelos GLM desenvolvidos a partir do grupo de calibração. A metodologia da escolha das combinações testadas foi similar à utilizada para o modelo anterior, evitando a multicolinearidade entre as variáveis.

Tabela 9: Modelos de geração de viagens de veículos de carga utilizando modelo linear generalizado.

Cenário	Variável Dependente	Variáveis Independentes	Valor do intercepto	Valor dos coeficientes	Teste z		AICc	RMSD
					Intercepto	Var. independente		
GLM 01	Entrega sexta-feira	Área	1,9908	0,0004	21,76	10,28	165,49	6,70
GLM 02	Entrega sexta-feira	Funcionários	1,6758	0,0118	14,28	10,31	157,84	6,51
GLM 03	Entrega sexta-feira	Checkout	0,7696	0,2039	3,95	10,78	128,73	8,52
GLM 04	Entrega sexta-feira	Área	2,5824	0,0005	18,33	10,057	142,88	6,59
		Densidade populacional		-83,8644		-4,964		
GLM 05	Entrega sexta-feira	Área	2,0341	0,0004	21,97	10,336	161,83	6,18
		Densidade emprego		-6,0710		-2,191		
GLM 06	Entrega sexta-feira	Área	2,0064	0,0004	21,76	10,373	166,4	6,65
		Densidade compras		-5,5761		-1,004		
GLM 07	Entrega sexta-feira	Área	2,0494	0,0005	21,30	8,946	155,24	5,29
		Renda		-0,0950		-3,425		
GLM 08	Entrega sexta-feira	Área	2,7591	0,0006	-18,16	10,697	132,93	6,11
		Densidade populacional		-97,7618		-5,536		
		Densidade emprego		-7,8586		-3,149		
GLM 09	Entrega sexta-feira	Área	2,6136	0,0005	18,32	10,223	143,07	6,49
		Densidade populacional		-85,0189		-5,037		
		Densidade compras		-7,2524		-1,277		
GLM 10	Entrega sexta-feira	Área	2,6358	0,0006	17,57	10,574	135,67	5,67
		Densidade populacional		-81,6193		-4,684		
		Renda		-0,0779		-2,948		
GLM 11	Entrega sexta-feira	Área	2,0737	0,0005	21,50	9,038	154,55	5,26
		Densidade emprego		-4,1755		-1,553		
		Renda		-0,0836		-2,972		

Tabela 9: Modelos de geração de viagens de veículos de carga utilizando modelo linear generalizado – continuação.

Cenário	Variável Dependente	Variáveis Independentes	Valor do intercepto	Valor dos coeficientes	Teste z		AICc	RMSD
					Intercepto	Var. independente		
GLM 12	Entrega sexta-feira	Área	2,0635	0,0005	21,36	9,057	156,19	5,24
		Densidade compras		-5,2988		-0,988		
		Renda		-0,0941		-3,42		
GLM 13	Entrega sexta-feira	Área	2,7677	0,0006	17,54	10,43	130,24	5,64
		Densidade populacional		-93,3118		-5,16		
		Densidade emprego		-6,4114		-2,541		
		Renda		-0,0558		-2,103		
GLM 14	Entrega sexta-feira	Área	2,6591	0,0006	17,63	10,19	136,18	5,58
		Densidade populacional		-82,2900		-4,732		
		Densidade compras		-6,3610		-1,163		
		Renda		-0,0759		-2,894		
GLM 15	Entrega sexta-feira	Funcionários	1,9032	0,0131	12,45	10,068	154,42	6,41
		Densidade populacional		-31,5609		-2,352		
GLM 16	Entrega sexta-feira	Funcionários	1,7022	0,0121	14,33	10,264	157,85	5,84
		Densidade emprego		-3,5382		1,348		
GLM 17	Entrega sexta-feira	Funcionários	1,6745	0,0118	14,06	10,302	159,84	6,51
		Densidade compras		0,3408		0,063		
GLM 18	Entrega sexta-feira	Funcionários	1,6452	0,0155	13,27	9,011	150,62	4,91
		Renda		-0,0789		-2,958		
GLM 19	Entrega sexta-feira	Funcionários	1,9395	0,0133	12,28	10,226	154,43	5,80
		Densidade populacional		-31,7865		-2,348		
		Densidade emprego		-3,3209		-1,348		

Tabela 9: Modelos de geração de viagens de veículos de carga utilizando modelo linear generalizado – continuação.

Cenário	Variável Dependente	Variáveis Independentes	Valor do intercepto	Valor dos coeficientes	Teste z		AICc	RMSD
					Intercepto	Var. independente		
GLM 20	Entrega sexta-feira	Funcionários	1,8995	0,0132	12,38	10,052	156,35	6,43
		Densidade populacional		-31,8789		-2,364		
		Densidade compras		1,3972		0,261		
GLM 21	Entrega sexta-feira	Funcionários	1,7816	0,0154	10,10	9,149	151,52	5,15
		Densidade populacional		-16,5005		-1,06		
		Renda		-0,0631		-2,119		
GLM 22	Entrega sexta-feira	Funcionários	1,6557	0,0154	13,09	8,858	152,47	4,84
		Densidade emprego		-1,0262		-0,386		
		Renda		-0,0749		-2,634		
GLM 23	Entrega sexta-feira	Funcionários	1,6330	0,0156	12,89	8,975	152,35	4,93
		Densidade compras		2,7764		0,53		
		Renda		-0,0812		-2,99		
GLM 24	Entrega sexta-feira	Funcionários	1,8130	0,0152	9,87	8,956	153,17	5,05
		Densidade populacional		-18,2909		-1,153		
		Densidade emprego		-1,5446		-0,586		
		Renda		-0,0554		-1,732		
GLM 25	Entrega sexta-feira	Funcionários	1,7705	0,0154	9,96	9,116	153,22	5,17
		Densidade populacional		-16,7356		-1,073		
		Densidade compras		2,9192		0,556		
		Renda		-0,0652		-2,164		
GLM 26	Entrega sexta-feira	<i>Checkout</i>	0,7201	0,2030	3,11	10,685	130,57	8,54
		Densidade populacional		5,4834		0,396		
GLM 27	Entrega sexta-feira	<i>Checkout</i>	0,8074	0,2052	4,15	10,913	128,61	7,91
		Densidade emprego		-3,3444		-1,395		

Tabela 9: Modelos de geração de viagens de veículos de carga utilizando modelo linear generalizado – continuação.

Cenário	Variável Dependente	Variáveis Independentes	Valor do intercepto	Valor dos coeficientes	Teste z		AICc	RMSD
					Intercepto	Var. independente		
GLM 28	Entrega sexta-feira	<i>Checkout</i>	0,7802	0,2046	4,01	10,856	130,18	8,48
		Densidade compras		-3,7792		-0,718		
GLM 29	Entrega sexta-feira	<i>Checkout</i>	0,7383	0,1999	3,73	10,29	129,81	9,05
		Renda		0,0202		0,976		
GLM 30	Entrega sexta-feira	<i>Checkout</i>	0,7550	0,2046	3,20	10,855	130,46	7,92
		Densidade populacional		5,5713		0,392		
		Densidade emprego		-3,3743		-1,393		
GLM 31	Entrega sexta-feira	<i>Checkout</i>	0,7228	0,2037	3,13	10,77	131,97	8,50
		Densidade populacional		6,3760		0,459		
		Densidade compras		-3,9782		-0,752		
GLM 32	Entrega sexta-feira	<i>Checkout</i>	0,7458	0,2000	3,25	10,284	131,8	9,06
		Densidade populacional		-0,9942		-0,065		
		Renda		0,0209		0,887		
GLM 33	Entrega sexta-feira	<i>Checkout</i>	0,7672	0,1999	3,90	10,385	128,6	8,43
		Densidade emprego		-4,3750		-1,692		
		Renda		0,0311		1,455		
GLM 34	Entrega sexta-feira	<i>Checkout</i>	0,7478	0,2005	3,79	10,366	131,1	9,06
		Densidade compras		-4,3574		-0,813		
		Renda		0,0219		1,06		
GLM 35	Entrega sexta-feira	<i>Checkout</i>	0,8082	0,2001	3,47	10,387	130,49	8,49
		Densidade populacional		-5,1579		-0,329		
		Densidade emprego		-4,4563		-1,727		
		Renda		0,0350		1,433		

Tabela 9: Modelos de geração de viagens de veículos de carga utilizando modelo linear generalizado – continuação.

Cenário	Variável Dependente	Variáveis Independentes	Valor do intercepto	Valor dos coeficientes	Teste z		AICc	RMSD
					Intercepto	Var. independente		
GLM 36	Entrega sexta-feira	<i>Checkout</i>	0,7516	0,2005	3,28	10,361	133,1	9,07
		Densidade populacional		-0,4996		-0,032		
		Densidade compras		-4,3513		-0,812		
		Renda		0,0223		0,945		

São apresentados na Tabela 9, para cada cenário, os respectivos valores do intercepto e dos coeficientes de cada variável independente do modelo. Com isso é possível compor as equações de geração de viagens de veículos de carga e estimar a geração para qualquer estabelecimento do mesmo setor econômico de estudo.

Foram realizados testes para observar a validade estatística dos cenários utilizando o teste-z, o qual testa a hipótese baseado na estatística z, que segue a distribuição normal padrão sob a hipótese nula. Na Tabela 9, os cenários destacados (linhas sombreadas) foram considerados estatisticamente válidos, seguindo os critérios apresentados no item 4.2.2. Observa-se que apenas 11 deles (GLM 01, 02, 03, 04, 05, 07, 08, 10, 13, 15 e 18) foram validados. Nos outros cenários, algumas variáveis apresentaram resultados de teste-z estatisticamente válidos, porém as outras variáveis do modelo não.

Observa-se, portanto, que a utilização do modelo GLM possibilitou um número maior de cenários estatisticamente válidos se comparado com o modelo linear. Na prática, isso possibilita uma amplitude maior de alternativas a serem consideradas, que, se por um lado aumenta a sensibilidade do modelo, por outro lado demanda maior conhecimento por parte do analista / decisor.

Com relação ao AICc, o índice variou de 128,61 a 159,84. A variação do AICc entre os diferentes cenários seguiu o mesmo padrão do modelo linear. Os melhores resultados foram observados naqueles cenários com a presença da variável independente número de *checkouts* e aos piores resultados foram observados nos cenários com a presença da variável independente número de funcionários.

De forma análoga à técnica utilizando a regressão linear apresentada na seção 6.4, foi realizada também a análise gráfica dos resíduos das regressões lineares de cada cenário. As figuras a seguir apresentam os gráficos gerados a partir dos resultados obtidos nos modelos GLM 01 e GLM 18. No Apêndice B é apresentado os gráficos referentes aos outros modelos.

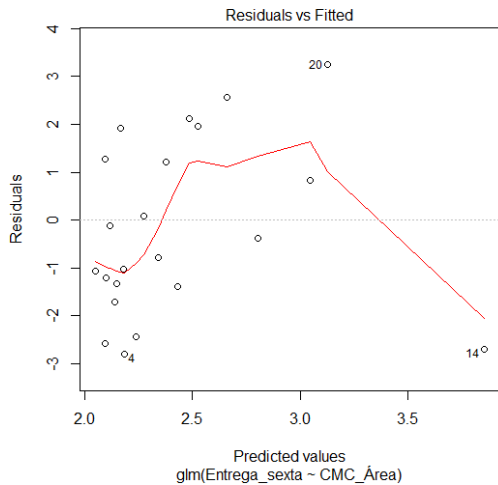


Figura 11: Resíduos versus valores ajustados
– GLM 01.

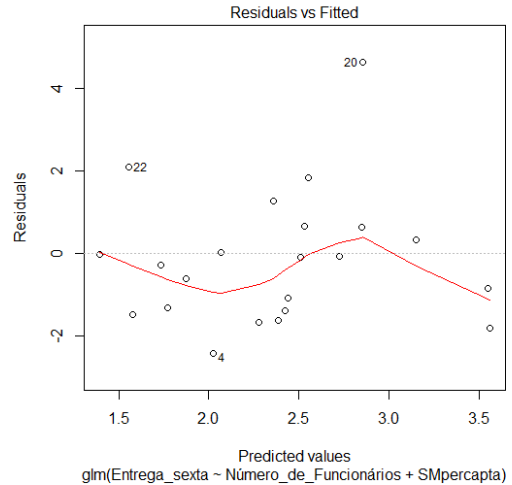


Figura 12: Resíduos versus valores ajustados
– GLM 18.

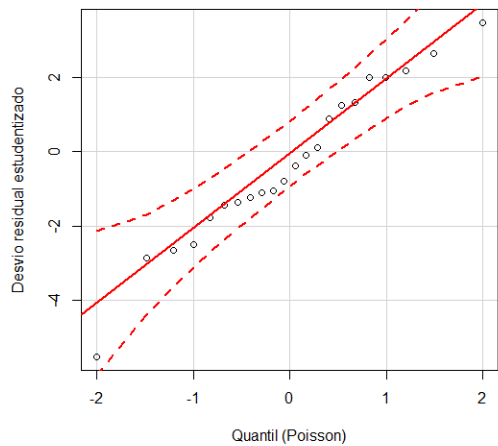


Figura 13: Gráfico normal de probabilidades
– GLM 01.

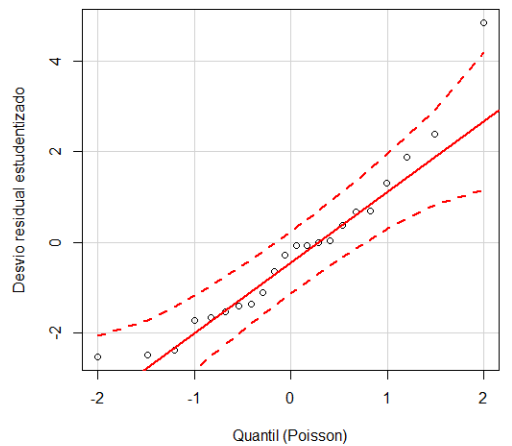


Figura 14: Gráfico normal de probabilidades
– GLM 18.

Por meio da análise da Figura 11 e da Figura 12, é possível indicar que no cenário GLM 01, a distribuição dos resíduos foi próxima à normal, apesar de possuir dois *outlayers* (pontos 14 e 20). O mesmo pode ser percebido para o modelo GLM 18. Portanto, ambos os modelos não apresentam variância constante. Já a Figura 13 e a Figura 14 mostram que no cenário GLM 01, os pontos seguem próximos à reta, indicando um bom ajuste. Já o cenário GLM 18 apresenta um ponto fora do intervalo de confiança considerado, não sendo, portanto, estatisticamente válido.

Realizando a análise visual dos resíduos dos outros cenários, pode-se concluir que os cenários GLM 01, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 14, 26, 27, 30 e 31 apresentaram resíduos dentro da normalidade, totalizando 16 modelos com resíduos dentro da normalidade, número significativamente superior ao obtido nos modelos lineares.

A exemplo do modelo linear, foi realizada uma validação cruzada em cada um dos modelos utilizando o grupo de validação. Através da equação gerada em cada cenário, foram estimados os valores de geração de viagens para cada um dos cinco estabelecimentos que formam o grupo de validação. Feito isso, foi calculado o indicador RMSD e os resultados obtidos para cada cenário também se encontram na Tabela 9.

Observa-se que o valor médio do RMSD foi de 6,72. O menor valor observado foi de 5,03, no cenário GLM 22 e o maior valor foi de 9,24, no cenário GLM 36. Os cenários com a presença da variável independente número de funcionários apresentaram os menores índices e, conseqüentemente, os melhores ajustes. Já aqueles cenários com a variável número de *checkouts* apresentaram os piores resultados. Considerando apenas os onze cenários validados estatisticamente por meio do teste-z, a média foi de 6,39. Já considerando aqueles cenários com os resíduos dentro da normalidade, a média foi de 6,69.

Como também observado anteriormente nos modelos lineares, os cenários GLM que não foram considerados válidos estatisticamente através dos resultados do teste-z e da análise visual dos resíduos apresentaram resultados na validação cruzada iguais ou melhores que aqueles considerados válidos.

5.6 Modelo linear generalizado geograficamente ponderado

O modelo GWGLM foi desenvolvido em duas etapas: utilização da modelagem linear generalizada geograficamente ponderada (GWGLM) e em seguida a interpolação dos dados com a utilização da técnica IDW. São apresentados na Tabela 10 os cenários desenvolvidos utilizando o grupo de calibração com os respectivos valores da estatística AICc. Ao contrário dos modelos anteriores, no GWGLM foi avaliado apenas o AICc e este não consiste em uma medida absoluta de ajuste de um modelo e sim em uma medida comparativa de desempenho do modelo. Por conta disso, não foi possível atestar a validade estatística destes cenários. A metodologia da escolha das combinações testadas foi a mesma utilizada para os modelos anteriores, evitando a multicolinearidade entre as variáveis.

Tabela 10: Modelos de geração de viagens de veículos de carga utilizando GWGLM.

Cenário	Variável Dependente	Variáveis Independentes				AICc	RMSD
		Variável 01	Variável 02	Variável 03	Variável 04		
GWGLM 01	Entrega sexta-feira	Área	-	-	-	72,11	6,77
GWGLM 02	Entrega sexta-feira	Funcionários	-	-	-	66,07	6,38
GWGLM 03	Entrega sexta-feira	<i>Checkout</i>	-	-	-	37,02	8,50
GWGLM 04	Entrega sexta-feira	Área	Densidade populacional	-	-	53,64	6,62
GWGLM 05	Entrega sexta-feira	Área	Densidade emprego	-	-	69,68	6,34
GWGLM 06	Entrega sexta-feira	Área	Densidade compras	-	-	77,71	6,75
GWGLM 07	Entrega sexta-feira	Área	Renda	-	-	65,13	5,44
GWGLM 08	Entrega sexta-feira	Área	Densidade populacional	Densidade emprego	-	45,26	6,11
GWGLM 09	Entrega sexta-feira	Área	Densidade populacional	Densidade compras	-	55,39	6,49
GWGLM 10	Entrega sexta-feira	Área	Densidade populacional	Renda	-	48,00	5,67
GWGLM 11	Entrega sexta-feira	Área	Densidade emprego	Renda	-	65,68	5,49
GWGLM 12	Entrega sexta-feira	Área	Densidade compras	Renda	-	67,29	5,42
GWGLM 13	Entrega sexta-feira	Área	Densidade populacional	Densidade emprego	Renda	-	-
GWGLM 14	Entrega sexta-feira	Área	Densidade populacional	Densidade compras	Renda	-	-
GWGLM 15	Entrega sexta-feira	Funcionários	Densidade populacional	-	-	65,6	6,43
GWGLM 16	Entrega sexta-feira	Funcionários	Densidade emprego	-	-	67,94	5,84
GWGLM 17	Entrega sexta-feira	Funcionários	Densidade compras	-	-	69,38	6,45
GWGLM 18	Entrega sexta-feira	Funcionários	Renda	-	-	60,82	4,82

Tabela 10: Modelos de geração de viagens de veículos de carga utilizando GWGLM – continuação.

Cenário	Variável Dependente	Variáveis Independentes				AICc	RMSD
		Variável 01	Variável 02	Variável 03	Variável 04		
GWGLM 19	Entrega sexta-feira	Funcionários	Densidade populacional	Densidade emprego	-	66,63	5,85
GWGLM 20	Entrega sexta-feira	Funcionários	Densidade populacional	Densidade compras	-	68,59	6,45
GWGLM 21	Entrega sexta-feira	Funcionários	Densidade populacional	Renda	-	63,47	5,08
GWGLM 22	Entrega sexta-feira	Funcionários	Densidade emprego	Renda	-	63,51	4,79
GWGLM 23	Entrega sexta-feira	Funcionários	Densidade compras	Renda	-	63,53	4,86
GWGLM 24	Entrega sexta-feira	Funcionários	Densidade populacional	Densidade emprego	Renda	-	-
GWGLM 25	Entrega sexta-feira	Funcionários	Densidade populacional	Densidade compras	Renda	-	-
GWGLM 26	Entrega sexta-feira	<i>Checkout</i>	Densidade populacional	-	-	39,72	8,52
GWGLM 27	Entrega sexta-feira	<i>Checkout</i>	Densidade emprego	-	-	38,05	7,95
GWGLM 28	Entrega sexta-feira	<i>Checkout</i>	Densidade compras	-	-	39,4	8,48
GWGLM 29	Entrega sexta-feira	<i>Checkout</i>	Renda	-	-	38,67	9,06
GWGLM 30	Entrega sexta-feira	<i>Checkout</i>	Densidade populacional	Densidade emprego	-	41,33	7,95
GWGLM 31	Entrega sexta-feira	<i>Checkout</i>	Densidade populacional	Densidade compras	-	42,51	8,50
GWGLM 32	Entrega sexta-feira	<i>Checkout</i>	Densidade populacional	Renda	-	42,24	9,08
GWGLM 33	Entrega sexta-feira	<i>Checkout</i>	Densidade emprego	Renda	-	39,55	8,50
GWGLM 34	Entrega sexta-feira	<i>Checkout</i>	Densidade compras	Renda	-	41,45	9,08
GWGLM 35	Entrega sexta-feira	<i>Checkout</i>	Densidade populacional	Densidade emprego	Renda	-	-
GWGLM 36	Entrega sexta-feira	<i>Checkout</i>	Densidade populacional	Densidade compras	Renda	-	-

Os cenários número 13, 14, 24, 25, 35 e 36 não apresentam resultados, pois estes cenários resultaram em erro no programa no momento da definição da largura de banda ideal.

Analisando os resultados apresentados na Tabela 10, constata-se que a média do AICc calculado foi de 55,84, tendo variado entre 37,02 a 77,71. Este valor é significativamente menor se comparados aos mesmos índices calculados para os modelos lineares e lineares generalizados. Portanto, pressupõe-se que o ajuste dos modelos GWGLM são melhores que os testados anteriormente neste trabalho.

Além disso, observa-se também que os melhores resultados foram observados naqueles cenários com a presença da variável independente número de *checkouts* e aos piores resultados foram observados nos cenários com a presença da variável independente número de funcionários.

Após a etapa da regressão espacial, realizou-se, para cada cenário desenvolvido, a interpolação dos coeficientes encontrados para cada uma das variáveis independentes. Como já apresentado no capítulo Abordagem Metodológica, a interpolação espacial destes coeficientes e dos interceptos possibilitará que o valor de geração de viagens de veículos de carga urbana seja estimado para supermercados localizados em qualquer local da área de estudo.

Para exemplificar o método, as Figuras 15 a 18 apresentam as superfícies matriciais (*raster*) geradas através da interpolação dos coeficientes encontrados para cada uma das variáveis do modelo GWGLM 11. Ressalta-se que a interpolação espacial e a posterior validação cruzada foram realizadas para todos os cenários. Ou seja, apesar de serem apresentados a seguir apenas os *rasters* gerados através do modelo GWGLM 11, foi gerado para cada modelo um conjunto de superfícies *raster* com os respectivos coeficientes.

Através da análise destas superfícies não é possível realizar uma análise do comportamento da geração de viagens de veículos de carga ao longo do território, pois trata-se apenas dos valores do intercepto e dos coeficientes da equação de geração.

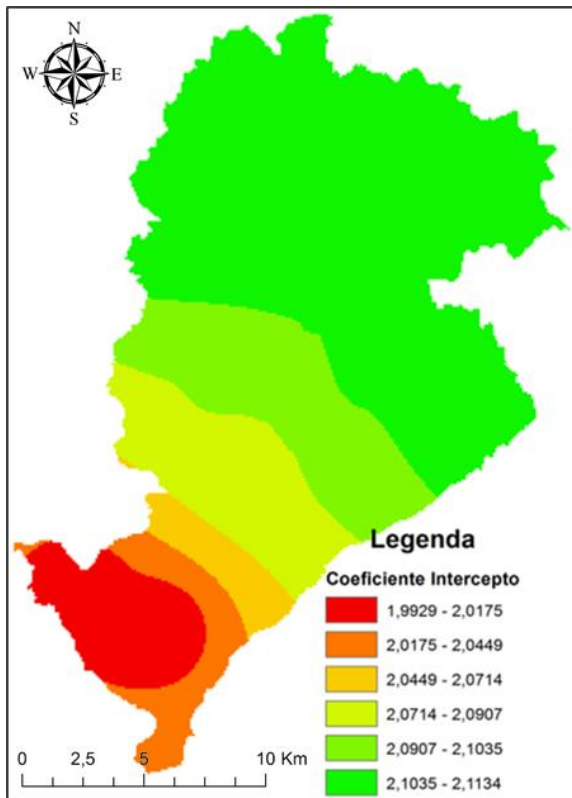


Figura 15: *Raster* do valor do intercepto – GWGLM 11.

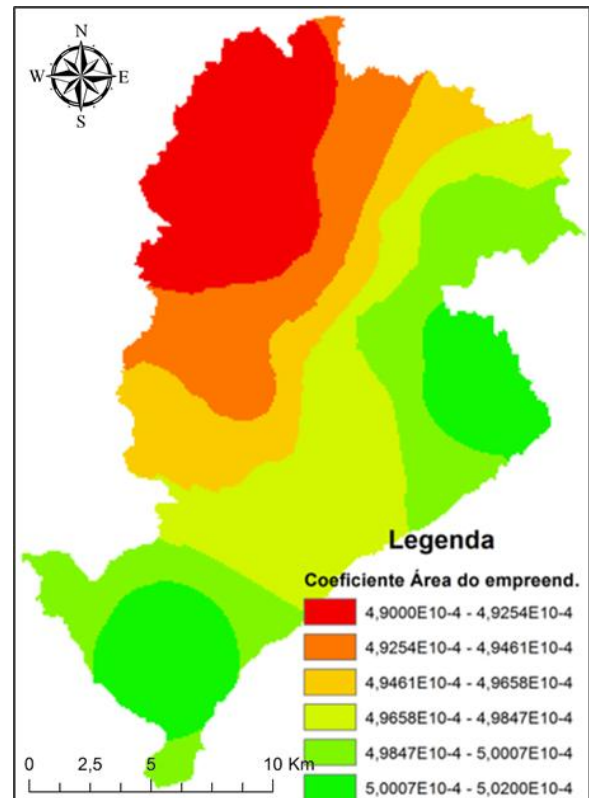


Figura 16: *Raster* do coeficiente relativo à variável Área do estabelecimento – GWGLM 11.

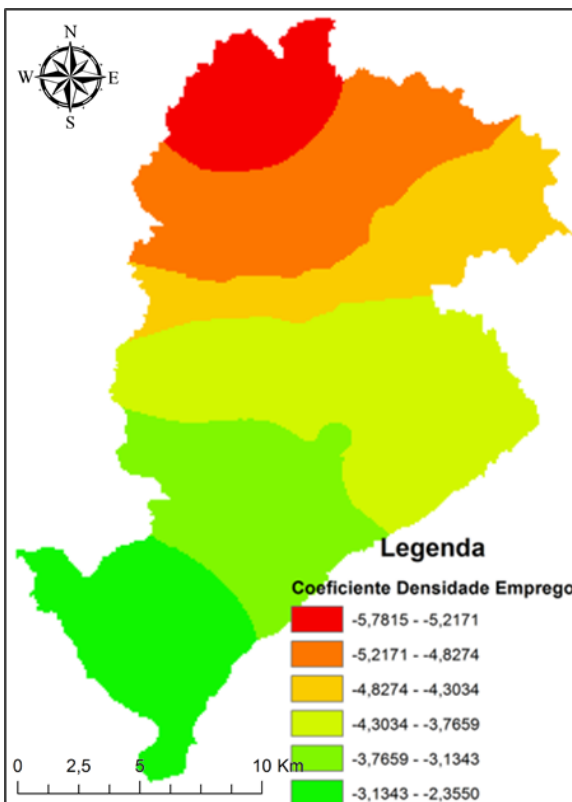


Figura 17: *Raster* do coeficiente relativo à variável Densidade emprego – GWGLM 11.

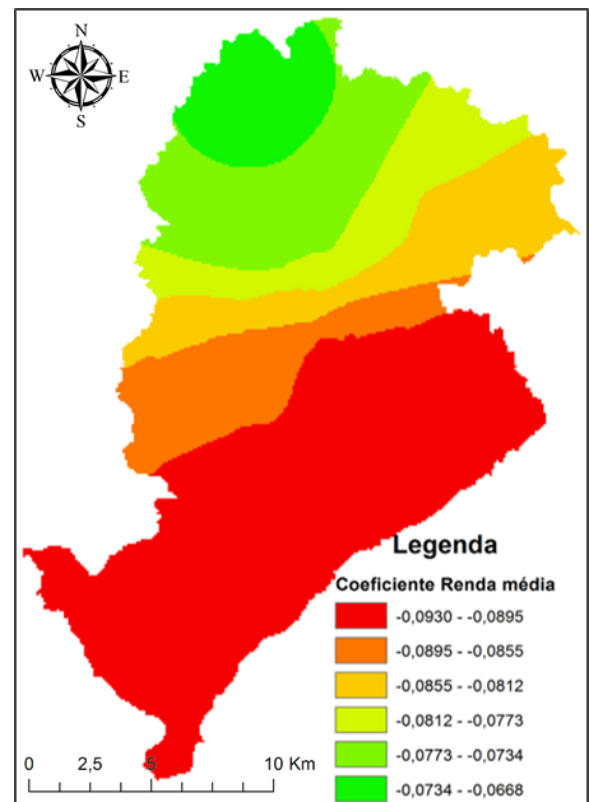


Figura 18: *Raster* do coeficiente relativo à variável Renda média – GWGLM 11.

No entanto, através da álgebra de mapas foi possível visualizar a representação espacial da geração de viagens de veículos de carga para qualquer estabelecimento. É apresentado nas Figuras 19 a 22 um exemplo hipotético, ilustrando o número de viagens de veículos de carga geradas em diferentes regiões do município de Belo Horizonte por estabelecimentos hipotéticos³ com 500m² de área (Figura 19), 1000m² (Figura 20), 1500m² (Figura 21) e 2000m² (Figura 22). Neste exemplo, foi utilizado o cenário GWGLM que possui como variáveis independentes a área do estabelecimento, a densidade populacional do entorno do empreendimento e a renda média per capita da população residente no entorno do estabelecimento. Em suma, as áreas homogêneas não mostram um cálculo consolidado de número de viagens e sim um potencial de geração de viagens de um estabelecimento localizado nesta região.

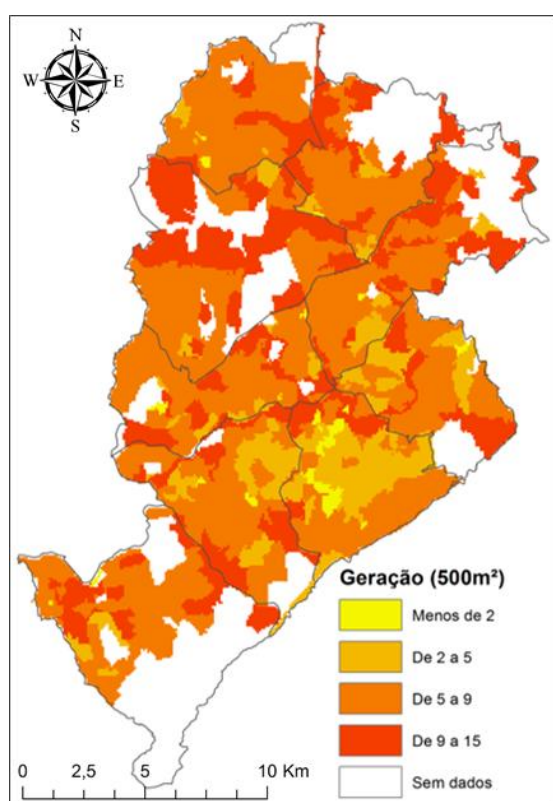


Figura 19: Geração de viagens de veículos de carga – Supermercados com 500m².

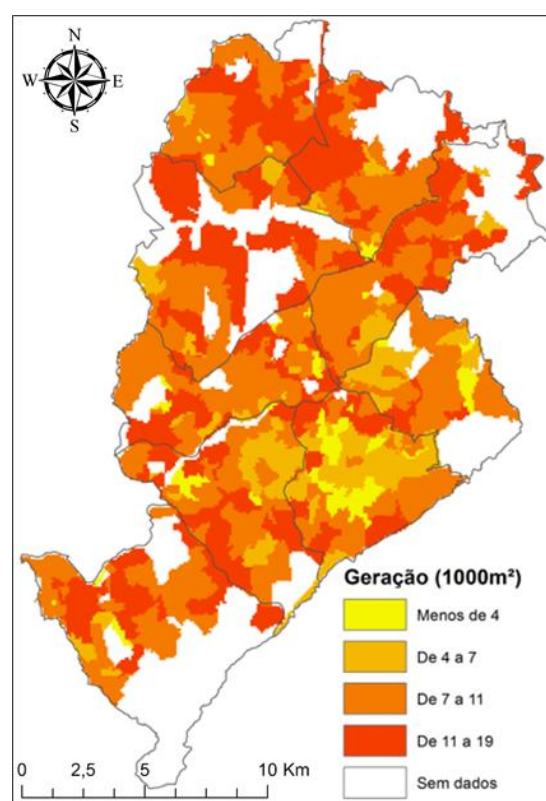


Figura 20: Geração de viagens de veículos de carga – Supermercados com 1000m².

³ Foi realizada esta categorização de empreendimentos pelo fato de que por se tratar de um modelo GWGLM que utiliza a distribuição de Poisson, a variável área do estabelecimento (variável relacionada à estrutura física do empreendimento utilizada neste exemplo) não varia de forma linear. Desta forma, não seria possível a elaboração de um mapa em função da área do empreendimento (por exemplo, um mapa cuja variável seria geração de viagens por m² de área do empreendimento).

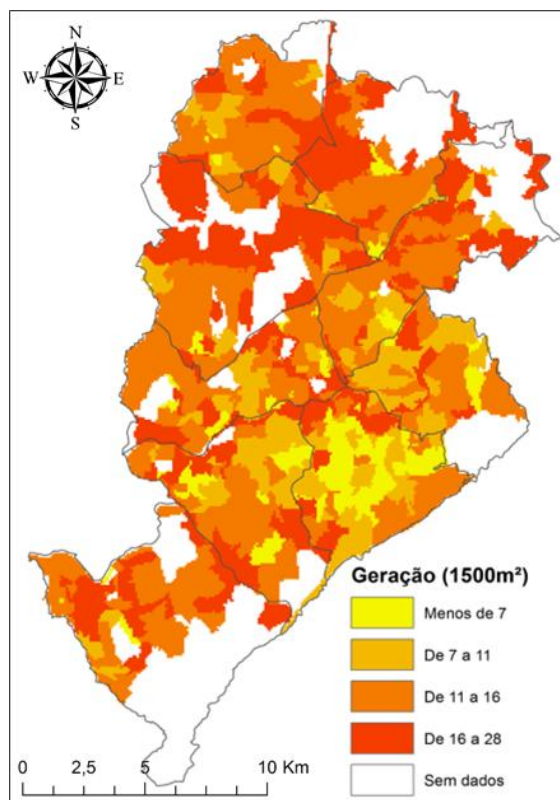


Figura 21: Geração de viagens de veículos de carga – Supermercados com 1500m².

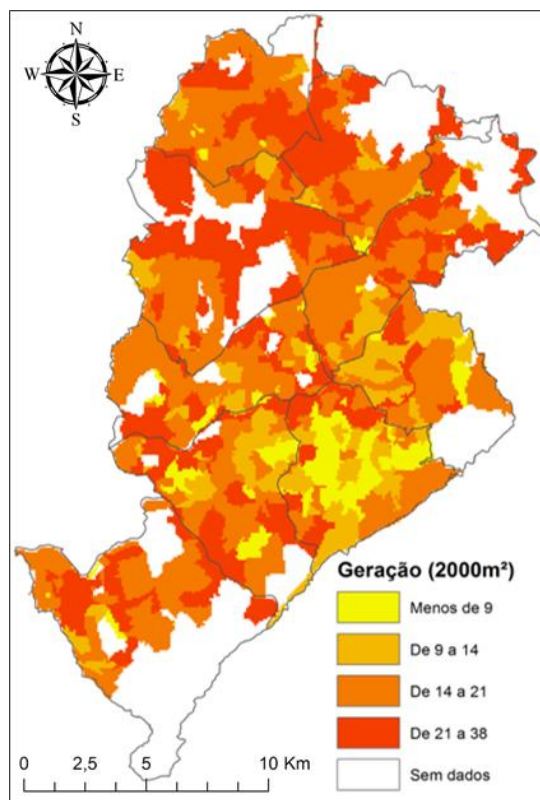


Figura 22: Geração de viagens de veículos de carga – Supermercados com 2000m².

Analisando as figuras anteriores, observa-se que possuem algumas áreas em branco pelo fato de não serem ocupadas por residências ou comércio/serviços. Além disso, foram identificados e descartados *outlayers* em algumas áreas da figura. Estas áreas foram identificadas pela legenda “sem dados”.

Portanto, realizando esta álgebra de mapas para cada um dos modelos, foi possível realizar a validação cruzada em cada um dos cenários utilizando o grupo de validação. Os resultados também se encontram na Tabela 10 apresentada anteriormente.

Observa-se que o valor médio do RMSD foi de 6,93. O menor valor observado foi de 4,79, no cenário GWGLM 22 e o maior valor foi de 9,08, no cenário GWGLM 34. Os cenários com a presença da variável independente número de funcionários apresentaram os menores índices e, conseqüentemente, os melhores ajustes. Já aqueles com a variável Número de *checkouts* apresentaram os piores resultados.

5.7 Análise comparativa dos modelos

Uma primeira análise a se fazer dos modelos elaborados é em relação aos modelos validados estatisticamente. Conforme apresentado, foram realizados testes estatísticos nos cenários que utilizam a técnica de regressão linear e nos modelos lineares generalizados. Ressalta-se que não foram realizados testes estatísticos para os modelos espaciais. No caso dos modelos lineares, foram validadas, através do teste-t, apenas 4 cenários, em um total de 36. Já nos modelos GLM, foram validadas através do teste-z 11 cenários.

Já em relação à análise dos resíduos, no caso dos modelos lineares 7 cenários apresentaram resíduos dentro da normalidade. Já no caso dos modelos GLM, foram 16 cenários, o que reforça o indicativo os modelos generalizados são mais eficientes que os modelos lineares.

Outra importante análise realizada para comparar as diferentes técnicas de regressão foi o comparativo dos valores de AICc encontrados após cada regressão. Conforme já mencionando anteriormente, não há um valor ideal para o AICc, pois este não consiste em uma medida absoluta de ajuste de um modelo, porém é útil na comparação entre modelos com a mesma variável dependente. No caso das regressões lineares, o valor médio do AICc foi de 144,2. Para os modelos GLM, o valor foi de 144,4. Já para os modelos GWGLM, o valor médio foi de 55,9. Observa-se que o índice foi praticamente igual para as duas primeiras técnicas de regressão e muito inferior quando utilizado o GWGLM. Conforme dito anteriormente, o AICc não é uma medida absoluta de ajuste de um modelo e sim um bom comparativo de ajuste do modelo. Portanto, os resultados indicam que o modelo GWGLM se ajustou melhor do que os modelos lineares e GLM.

Foram realizadas também para cada técnica de regressão a validação cruzada com dados do grupo de calibração e calculou-se o RMSD para cada um dos cenários desenvolvidos por meio das três técnicas. São apresentados na Tabela 11 os resultados obtidos.

Tabela 11: Análise comparativa do RMSD.

Cenário	LM	GLM	GWGLM	GLM/LM	GWGLM/LM
1	7,11*	6,70*	6,77	94,2%	95,2%
2	6,52	6,51*	6,38	99,8%	97,8%
3	8,50*	8,52*	8,50	100,1%	99,9%
4	6,90*	6,59*	6,62	95,5%	95,9%
5	6,66	6,18*	6,34	92,7%	95,1%
6	7,14	6,65	6,75	93,2%	94,6%
7	5,96	5,29*	5,44	88,6%	91,2%
8	6,22*	6,11*	6,11	98,3%	98,3%
9	6,90	6,49	6,49	94,1%	94,1%
10	5,68	5,67*	5,67	99,8%	99,8%
11	5,87	5,26	5,49	89,7%	93,5%
12	6,02	5,24	5,42	86,9%	90,0%
13	5,65	5,64*	-	99,9%	-
14	5,70	5,58	-	97,9%	-
15	6,44	6,41*	6,43	99,5%	99,8%
16	6,09	5,84	5,84	95,9%	95,9%
17	6,51	6,51	6,45	100,0%	99,0%
18	5,81	4,91*	4,82	84,4%	83,0%
19	5,81	5,80	5,85	99,9%	100,6%
20	6,43	6,43	6,45	100,0%	100,4%
21	5,77	5,15	5,08	89,2%	88,0%
22	5,59	4,84	4,79	86,5%	85,7%
23	5,82	4,93	4,86	84,7%	83,5%
24	5,42	5,05	-	93,3%	-
25	5,78	5,17	-	89,5%	-
26	8,64	8,54	8,52	98,8%	98,5%
27	7,73	7,91	7,95	102,3%	102,7%
28	8,47	8,48	8,48	100,1%	100,1%
29	9,05	9,05	9,06	100,0%	100,2%
30	7,92	7,92	7,95	100,0%	100,4%
31	8,62	8,50	8,50	98,7%	98,6%
32	9,08	9,06	9,08	99,8%	100,0%
33	8,39	8,43	8,50	100,6%	101,4%
34	9,09	9,06	9,08	99,7%	99,9%
35	8,47	8,49	-	100,3%	-
36	9,13	9,07	-	99,3%	-
Média	6,97	6,72	6,79	96,4%	97,4%

(*) Modelos validados estatisticamente pelo teste-t ou teste-z

Observa-se que, em média, os modelos GLM apresentaram um valor de RMSD 4% melhor em relação aos modelos lineares. Já os modelos GWGLM foram 3% melhores que os lineares. No entanto, se forem considerados apenas aqueles cenários estatisticamente validados, estas diferenças aumentam significativamente. Considerando os quatro modelos lineares válidos, o RMSD médio é de 6,69. Já o valor médio do RMSD dos 11 modelos GLM válidos é 6,23, 7% menor que o linear. Para completar a comparação, como não foram realizados testes estatísticos para os modelos GWGLM, foram consideradas as mesmas 11 combinações de

variáveis do modelo GLM. Com isso, a média do RMSD dos cenários GWGLM validados foi de 6,31, ou seja, 6% menor que o modelo linear.

Apesar do ganho relativamente pequeno (4% de melhoria do modelo LM para o GLM), o custo computacional para construção dos modelos utilizando as duas técnicas é semelhante, pois utilizam o mesmo software e os mesmos dados de entrada. Portanto, o modelo GLM apresenta um ganho de acurácia na predição se comparado com o modelo LM.

Comparando agora apenas o modelo GLM e o modelo GWGLM, observa-se que pela perspectiva do resultado da validação cruzada, o GLM se mostrou melhor, significando uma melhor capacidade de predição. Além disso, o GLM tem o menor custo computacional e é mais simples de implementar que o modelo GWGLM. No entanto, esse modelo tem uma grande vantagem sobre o GLM: a consideração da localização geográfica do empreendimento.

Como a geração de viagens não é constante ao longo do espaço, empreendimentos com a mesma área comercial, mesmo número de funcionários e/ou mesmo número de *checkouts* podem gerar diferentes valores de viagens de veículos de carga em diferentes pontos do município. Apesar de não apresentar um ganho significativo em termos de qualidade de predição, o método espacial apresenta estas informações detalhadas ao longo do espaço.

Por meio das Figuras 19 a 22, observa-se que dois estabelecimentos com a mesma área, sendo um localizado na região Centro Sul de Belo Horizonte e outro localizado na região Norte possuem valores de geração de viagens de carga diferentes. O estabelecimento da região Centro Sul tende a gerar um menor número de viagens por conta das restrições de tráfego existentes no local. Tal diferença pode ser explicada devido às restrições de tráfego e de horário existentes no local e a amplitude de estoque. Com isso, os estabelecimentos tendem a otimizar sua logística de abastecimento / entrega, realizando menos viagens de caminhões.

Já a área periférica do município, o número de geração tende a ser maior, pois não há esta restrição de tráfego de caminhões, possibilitando a livre circulação de veículos e a realização de pequenas entregas em diferentes estabelecimentos comerciais. Cabe ressaltar que essas diferentes configurações de entregas podem refletir nos custos de abastecimento dos estabelecimentos, e respectivamente nos custos dos produtos repassados aos consumidores.

Esta análise espacial da geração de viagens pode ser muito útil para os planejadores municipais, pois a reprodutibilidade da abordagem metodológica, considerando atributos locais, pode resultar em modelos de predição adequados à identificação de impactos de empreendimentos. Além disso, será possível saber em qual local o impacto de um empreendimento pode ser mais significativo. Desta forma, é possível a proposição de medidas mitigatórias específicas para aquele local.

Os resultados da análise comparativa permitem concluir que os modelos GLM e os modelos GWGLM resultaram em melhores resultados na validação cruzada, com uma pequena vantagem dos modelos GLM. Porém, a utilização da técnica espacial permite uma maior possibilidade de análise do impacto em diferentes regiões do município. Desta forma, a hipótese de que os modelos geograficamente ponderados apresentam melhores resultados que os modelos lineares generalizados que por sua vez apresentam melhores resultados que os modelos lineares é verdadeira.

5.8 Análise das variáveis independentes em relação aos modelos obtidos

Neste trabalho foram realizadas correlações entre a geração de viagens de veículos de carga e as variáveis independentes. Do conjunto de variáveis utilizadas, a área do estabelecimento, o número de funcionários e número de *checkouts* apresentaram uma grande correlação linear com a variável geração de viagens de veículos de carga.

A variável área do empreendimento foi a que apresentou maiores modelos válidos estatisticamente através do teste-t e teste-z, seguidas pela variável número de funcionários e, por último, pela variável número de *checkouts*.

Com relação às variáveis socioeconômicas utilizadas nos modelos, a densidade populacional, de empregos, de pessoas fazendo compras e a renda média da população no entorno do empreendimento se mostraram inversamente proporcionais à geração de viagens de veículos de carga. Este resultado vai ao encontro do senso comum de que uma região com maiores densidades de emprego, comércio e compras tendem a possuir um maior número de pessoas circulando e, conseqüentemente uma maior movimentação nos supermercados da região, gerando maiores taxas de geração de veículos de carga que abastecem o estabelecimento.

Uma explicação possível para este resultado é que regiões mais adensadas tendem a possuir um fluxo de veículos maior e, conseqüentemente, um maior índice de congestionamentos.

Com isso, os estabelecimentos, para diminuir custos com transportes, otimizam o máximo as entregas, resultando em um menor valor de geração de viagens.

Além disso, no caso específico de Belo Horizonte, as regiões mais adensadas e com a maior renda se encontram na área central, onde há uma restrição de tráfego de caminhões em determinados períodos do dia, conforme apresentado anteriormente. Este fator reforça a preocupação dos empreendimentos em aperfeiçoar suas entregas.

É importante ressaltar que os modelos desenvolvidos neste trabalho têm como base o número de viagens de veículos de carga geradas no estabelecimento e não o volume de carga gerado. Com isso, uma mesma viagem de carga pode ser realizada por um veículo urbano de carga (VUC) que transporta cerca de três toneladas ou um caminhão de maior porte que chega a 10 toneladas. Por conta disso, caso os modelos tivessem como variável dependente volume de entregas, a tendência seria que as variáveis territoriais influenciariam positivamente no resultado final do modelo.

A análise das variáveis permite concluir que é viável a utilização de dados secundários na estimativa de geração de viagens de veículos de carga. Desta forma, a hipótese de que é possível estimar a geração de viagens de veículos de carga em áreas urbanas por meio da utilização de um modelo de análise geográfica baseado em dados secundários, como dados socioeconômicos, cadastro imobiliário e índice de empregos existente na área de estudo é verdadeira.

6 CONSIDERAÇÕES FINAIS

Conforme o objetivo proposto inicialmente, neste trabalho foram desenvolvidos modelos para estimar a geração de viagens de veículos de carga em supermercados do município de Belo Horizonte utilizando modelo linear generalizado (GLM) e modelagem linear geográfica geograficamente ponderada (GWGLM). Os modelos elaborados foram validados através de testes estatísticos e da análise gráfica dos resíduos. Além disso, realizou-se uma validação cruzada através de dados provenientes do grupo de validação.

Para a construção destes modelos, foram utilizados dados secundários e dados amostrais de geração de viagens de veículos de carga obtidos por meio de pesquisas de campo. A utilização destes dados retornaram bons resultados na previsão de viagens.

Quanto aos modelos estudados, constatou-se que os modelos GLM apresentaram ganhos relativos se comparados com os modelos lineares que são usualmente utilizados para previsão de geração de viagens. Foi identificado um maior número de modelos estatisticamente válidos e constatou-se que os resultados provenientes da validação cruzada foram também melhores. Já os resultados obtidos na utilização do modelo GWGLM também foram melhores que os resultados da regressão linear.

Comparando o modelo linear generalizado e o modelo GWGLM, constatou-se que o primeiro obteve resultados na validação cruzada ligeiramente melhores que o segundo. Este fato, aliado ao menor custo computacional necessário para desenvolver os modelos colocam o modelo GLM em vantagem em termos de exatidão na geração de viagens.

No entanto, como o modelo GWGLM utiliza técnicas de modelagem geográfica em seu desenvolvimento, ele é capaz de variar as taxas de geração de viagens de acordo com a localização geográfica. Com isso é possível realizar não só a estimativa de geração de viagens, mas sim outras análises, como, por exemplo, identificar as áreas mais sensíveis à instalação de novos empreendimentos, do ponto de vista de impacto gerado. Por conta disso, tornasse viável a utilização deste modelo, pois além de retornar estimativas tão boas quanto os outros modelos, é acrescida a variável espacial na análise dos resultados.

Os modelos propostos neste trabalho, tanto o modelo GLM, quanto o modelo GWGLM apresentam uma importante contribuição técnico-científica, pois o mesmo pode ser utilizado como uma ferramenta útil para o planejamento de transportes no que tange a análise dos

impactos provocados pela movimentação de cargas no sistema viário ao entorno de um novo estabelecimento. Além disso, o modelo GWGLM apresenta um ganho maior no detalhamento destes impactos, pois apresenta como resultados diferentes taxas de geração de viagens de acordo com a localização geográfica do estabelecimento.

Em relação às variáveis independentes relacionadas às características físicas do empreendimento, a variável área do empreendimento foi a que apresentou um maior número de modelos validados estatisticamente, seja por testes estatísticos, seja através da análise dos resíduos. Já em relação às variáveis independentes relacionadas às características socioeconômicas do território, todas apresentaram uma influência negativa na geração de viagens, ou seja, quanto maior a renda e as densidades populacional, de empregos e de pessoas que fazem compras, menor a geração de viagens de veículos de carga. Este fato pode ser explicado pelas restrições na circulação de veículos nestas regiões com grandes densidades e renda e do excesso de tráfego nas vias, obrigando os estabelecimentos a desenvolverem um processo de otimização das entregas, resultando em um menor volume de caminhões gerados.

É importante ressaltar que as variáveis territoriais são muito sensíveis à regulação urbana, à operação viária e à estrutura morfológico e funcional das cidades. Portanto, estudos complementares devem ser realizados em cada localidade e em cada setor econômico para atestar a influência dessas variáveis na geração de viagens de veículos de carga.

6.1 Recomendações

Ao longo do desenvolvimento do trabalho foram percebidas algumas melhorias que se tornaram recomendações para pesquisas posteriores:

- Testar os modelos GLM e GWGLM utilizando amostra maior;
- Testar os modelos GLM e GWGLM em outros setores econômicos, como, por exemplo, shopping centers, armazéns, bares, restaurantes, etc., que são grandes geradores de demanda por mercadorias e que não foram abordados neste trabalho;
- Utilizar as variáveis socioeconômicas de acordo com a média da densidade de população compras e renda e um raio de influência do empreendimento e não pela área homogênea, pois o empreendimento pode se encontrar em uma área homogênea que não reflita as características locais;

- Utilizar as variáveis número de clientes, faturamento, volume de tráfego nas vias ou um indicador de mobilidade urbana como variáveis independentes dos modelos;
- Desenvolver modelos que utilizam como variável dependente o volume de mercadoria gerada pelo empreendimento e o número de veículos de carga gerados por tipo;
- Utilizar os cenários desenvolvidos por meio dos modelos GWGLM e GLM e os dados do CMC para estimar o número de viagens de veículos de carga em todo município de Belo Horizonte e analisar os impactos gerados no sistema viário;
- Investigar se os modelos gerados são válidos para outras cidades;
- Investigar se o comportamento das variáveis independentes socioeconômicas observado para o caso de Belo Horizonte de repete para em outras localidades;
- Avaliar a relação entre a validade estatística de um modelo através do teste-t e do teste-z e o resultado da validação cruzada, ou seja, entre ajuste do modelo e acurácia na previsão.

REFERÊNCIAS

- ABRAS. Associação Brasileira de Supermercados. Disponível em: <<http://www.abrasnet.com.br/economia-e-pesquisa/ranking-abras/os-numeros-do-setor/>>. Acesso em: 01 de junho de 2017.
- ANSELIN, L. *Spatial data analysis with GIS: An introduction to application in the Social Sciences*. National Center for Geographic Information and Analysis – University of California – Califórnia, EUA, 1992.
- ASSUNÇÃO, R. M. *Estatística Espacial com aplicações em Epidemiologia, Economia, Sociologia*. 7ª Escola de Modelos de Regressão, São Carlos, São Paulo, 2001.
- BAILEY, T. C.; GATRELL, A. C. *Interactive spatial data analysis*. Essex: Longman Scientific & Technical, Vol. 413, 1995.
- BARROSO, L. F.; OLIVEIRA, L. K.; CARDOSO, L.; LOBO, C. F. F. Avaliação da utilização de dados secundários para o planejamento de transporte de carga urbana. In: *XXVII Congresso de Pesquisa e Ensino em Transportes – ANPET*, 2013, Belém. Anais do XXVII Congresso de Pesquisa e Ensino em Transportes, 2013.
- BASCUÑÁN, M. M.; QUEZADA, C. R. Geographically weighted regression for modelling the accessibility to the public hospital network in Concepción Metropolitan Area, Chile. *Geospatial Health*, v. 11, n. 3, 2016.
- BASTIDA, C.; HOLGUIN-VERAS, J. Freight generation models: Comparative analysis of regression models and multiple classification analysis. *Transportation Research Record: Journal of the Transportation Research Board*, (2097), p. 51-61, 2009.
- BELO HORIZONTE. Cadastro municipal de contribuintes de tributos. Base de Dados. Prefeitura de Belo Horizonte, 2015.
- BELO HORIZONTE. Prefeitura Municipal Belo Horizonte. Disponível em: <<http://portalpbh.pbh.gov.br/pbh/ecp/comunidade.do?app=cmc>>. Acesso em: 29 de maio de 2017.

BERRY, J. K. Cartographic Modeling: The Analytical Capabilities of GIS. In: *Environmental Modeling with GIS*. Goodchild, M.; Parks, B.O., Steyaert, L.T., Oxford, Oxford University Press, 1993, pp. 58-74

BHAT, C.; ZHAO, H. The spatial analysis of activity stop generation. *Transportation Research Part B: Methodological*, v. 36, n. 6, p. 557-575, 2002.

BHTRANS. Empresa de Transportes e Trânsito de Belo Horizonte. Restrição de veículos Pesados, 2013. Disponível em: <<http://www.bhtrans.pbh.gov.br/portal/page/portal/portalpublico/Temas/Noticias/restricao-carga-descarga-contorno-050913>>. Acesso em: 06 de agosto de 2017.

BROGAN, J. D. Improving Truck Trip-Generation Techniques Through Trip-End Stratification. *Transportation Research Record 771*, TRB, National Research Council, Washington, D.C., pp. 1–6, 1980.

CÂMARA, G.; MONTEIRO, A. M.; CARVALHO, M. S. Análise Espacial e Geoprocessamento, In: *Análise Espacial de Dados Geográficos*, eds. Fuks, S.D.; Carvalho, M.S.; Câmara, G.; Monteiro, A.M.V. – Divisão de Processamento de Imagens – Instituto Nacional de Pesquisas Espaciais – São José dos Campos, São Paulo, 2002.

CARVALHO, L. E. X.; SILVA, H. N.; LOUREIRO, C. F. G.; MENEZES H. B. Regressão Linear Geograficamente Ponderada em ambiente SIG. *Transportes*, Volume XIV, nº 2: 18-26, 2006.

CHARLTON, M.E., FOTHERINGHAM, A.S. *Geographically Weighted Regression – A Tutorial on using GWR in ArcGIS 9.3*. National Centre for Geocomputation, National University of Ireland Maynooth. Maynooth, County Kildare, Ireland, 2009.

CHARNET, R.; FREIRE, C. D. L.; CHARNET, E. M.; BONVINO, H. Análise de modelos de regressão linear com aplicações. Campinas, São Paulo, Unicamp, 356p, 1999.

COMI, A.; DELLE SITE, P.; FILIPPI, F.; NUZZOLO, A. Urban Freight Transport Demand Modelling: a State of the Art. In: *European transport / trasporti europei*, ISTIEE, 2012, Trieste, Itália.

COMI, A.; NUZZOLO, A. Simulating urban freight flows with combined shopping and restocking demand models. *Procedia-Social and Behavioral Sciences*, 125, 49-61, 2014.

COSTA, A.; PITOMBO, C. S.; SALGUEIRO, A. R. Estimação de escolha modal através da geoestatística. In: *XXVII Congresso de Pesquisa e Ensino em Transportes – ANPET*, 2013, Belém.

DABLANC, L. Goods transport in large European cities: difficult to organize, difficult to modernize. *Transportation Research Part A*. Davis, v. 41, p. 280 – 285, 2007.

DABLANC, L. *Freight transport for development toolkit: Urban Freight*. Washington: Transport Research Support: The World Bank, 2009. 50 p. Relatório.

DOBSON, A. J. *An introduction to generalized linear models*. London: Chapman Hall, 1997.

ELIAN, S.N. *Análise de regressão*. São Paulo: USP / Dep. Estatística, 1988. 232 p.

ESRI. Environmental Systems Research Institute. A quick tour of geoprocessing tool references. Disponível em: <<http://desktop.arcgis.com/en/arcmap/latest/tools> >. Acesso em: 27 de maio de 2017.

FARRAR, D. E.; GLAUBER, R. R. Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics*, p. 92-107, 1967.

FERREIRA, B. L. G.; SILVA, M. A. V. Análise de geração de viagens de veículos de carga em área urbana. In: *XXIX Congresso de Pesquisa e Ensino em Transportes – ANPET*, 2015, Ouro Preto.

FERREIRA, B. L. G.; SILVA, M. A. V. Análise da relação entre variáveis socioeconômicas e o número de viagens de veículos de carga na região metropolitana do Rio de Janeiro. *Revista Produção e Desenvolvimento*, 2(1), 77-89, 2016.

FOX, J.; WEISBERG S. *An {R} Companion to Applied Regression*, Second Edition. Thousand Oaks CA: Sage, 2011. URL: <<http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>>.

GASPARINI, A.; CAMPOS, V. B. G.; D'AGOSTO M. A. Modelos para estimativa da demanda de viagens de veículos de carga para supermercados e shopping-centers. *Transportes*, 18(1), 2010.

GONZALEZ-FELIU, J.; ROUTHIER, J. Modeling urban goods movement: How to be oriented with so many approaches?. *Procedia-Social and Behavioral Sciences*, v. 39, p. 89-100, 2012.

GRIECO, E.; MARCOLINI S.; PORTUGAL L. S.; SOARES O. Estabelecimentos Residenciais. In: L. S. Portugal. *Polos Geradores de Viagens orientados a Qualidade de Vida e Ambiental: Modelos e Taxas de Geração de Viagens*. Editora Interciência Ltda., Rio de Janeiro, 2012. cap. 7, p. 207-236.

GÜNAY, G.; ERGÜN, G.; GÖKAŞAR, I. Conditional Freight Trip Generation modelling. *Journal of Transport Geography*, 54, p. 102-111, 2016.

HENRIQUE, C. S. *Diagnóstico Espacial da Mobilidade e da Acessibilidade dos Usuários do Sistema Integrado de Transportes de Fortaleza*. Dissertação (Mestrado), Programa de Mestrado de Transportes, Universidade Federal do Ceará, Fortaleza, CE, 165 fl. 2004.

HOLGUÍN-VERAS, J., JALLER M., DESTRO L., BAN X., LAWSON C. E LEVINSON H. Freight generation, freight trip generation, and the perils of using constant trip rates. *Transportation Research Record: Journal of the Transportation Research Board*, p. 68-81, 2011.

HOLGUÍN-VERAS, J., JALLER, M., SANCHEZ-DÍAZ, I., WOJTOWICZ, J.M., CAMPBELL, S., LEVINSON, H.S., *et al.* *NCHRP Report 739/NCFRP Report 19: Freight Trip Generation and Land Use*. Transportation Research Board of the National Academies, Washington D.C., NCHRP Report 739/NCFRP Report 19, 2012.

IBGE. Instituto Brasileiro de Geografia e Estatística. Cidades. Disponível em: <<http://cod.ibge.gov.br/5N7>>. Acesso em: 06 de agosto de 2017.

IBGE. Instituto Brasileiro de Geografia e Estatística . CNAE. Disponível em: <<http://cnae.ibge.gov.br/classificacoes/por-tema/atividades-economicas/classificacao-nacional-de-atividades-economicas>>. Acesso em: 05 de novembro de 2016.

IDING, M., MEESTER, W., TAVASSZY, L. Freight trip generation by firms. In: *XLII European congress of the regional science association*, 2002, Dortmund.

- ITE. Institute of Transportation Engineers. Trip Generation: An *ITE Informational Report*, eighth ed. Institute of Transportation Engineers, Washington, D.C., 2008.
- JALLER, M.; SANCHEZ-DIAZ, I.; HOLGUIN-VERAS, J.; LAWSON, C.T. Area based freight trip generation models. *93rd Transportation Research Board Annual Meeting. Transportation Research Board of the National Academies*, Washington, D.C., 2014.
- JOHNSTON, K.; VER HOEF, J. M.; KRIVORUCHKO, K.; LUCAS, N. *Using ArcGIS geostatistical analyst* (Vol. 380). Redlands – ESRI, 2001.
- KAWAMURA, K.; SHIN H. S.; MCNEIL S. *Business and Site Specific Trip Generation Methodology For Truck Trips*. Midwest Regional University Transportation Center, University of Wisconsin–Madison, 2008.
- KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *IJCAI* - Vol. 14, No. 2, pp. 1137-1145., 1995.
- KULPA, T. Freight truck trip generation modelling at regional level. *Procedia-Social and Behavioral Sciences*, 111, 197-202, 2014.
- LOPES, S. B. *Efeitos da dependência espacial em modelos de previsão de demanda por transporte*. 153 f. Dissertação (Mestrado). Universidade de São Paulo, São Carlos, 2005.
- MARRA, C. *Caracterização de demanda de movimentações urbanas de carga*. Dissertação (Mestrado em Engenharia Civil). Universidade Estadual de Campinas, São Paulo, 1999.
- MELO, I. C. B. *Avaliação da demanda por transporte de carga em áreas urbanas*. Dissertação (Mestrado em Engenharia de Transportes). Instituto Militar de Engenharia, Rio de Janeiro, 2002.
- MILLER, H. J. Potential contribution of spatial analysis to Geographic Information Systems for Transportation (GIS-T), *Geographical Analysis*, 31, 373-399, 1999.
- MINAS GERAIS. Agência de Desenvolvimento da Região Metropolitana de Belo Horizonte. *Pesquisa Origem e Destino 2011-2012*. Belo Horizonte, 2013.

MYERS, R. H., MONTGOMERY, D. C., VINING, G. G., ROBINSON, T. J. *Generalized linear models: with applications in engineering and the sciences*. Vol. 791. John Wiley & Sons, 2012.

NAKAYA, T. *GWR4 User Manual - Windows Application for Geographically Weighted Regression Modelling*, 2012.

NOVAK, D. C.; HODGDON, C.; GUO, F.; AULTMAN-HALL, L. *Nationwide freight generation models: A spatial regression approach*. *Networks and Spatial Economics*, v. 11, n. 1, p. 23-41, 2011.

OGDEN K. *Urban Goods Movement: A Guide to Policy and Planning*. Ashgate, Cambridge University Press, 1992.

OLIVEIRA, L. K.; OLIVEIRA, R. L. M.; RAMOS, C. M. F.; EBIAS, D. G. Modelo de geração de viagens de veículos de carga em áreas urbanas: um estudo para bares, restaurantes e supermercados. *Transportes*, 24(2), 53-67, 2016.

ORTÚZAR, J. D.; WILLUMSEN, L. G. *Modelling transport*. New Jersey: Wiley, 1994.

PATIER, D.; ROUTHIER J. How to Improve the Capture of Urban Goods Movement Data? Proc., *8th International Conference on Survey Methods in Transport*, Annecy, France, 2008.

PEREIRA, L. D. S. F.; OLIVEIRA, L. K. D. Methodology for estimating freight flows from secondary data: an application in Belo Horizonte. *Journal of Transport Literature*, 8(4), 279-315, 2014.

PITOMBO, C. S.; SOUSA A. J.; BIRKIN, M. Comparing Different Spatial Data Analysis to Forecast Trip Generation. *XII World Conference on Transport Research Society – WCTR*, p. 25, 2010.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL: <<https://www.R-project.org/>>.

RAMOS C. M. F. *Características operacionais e geração de viagens para supermercados*. Dissertação (Mestrado), Universidade Federal de Minas Gerais, Belo Horizonte, 2015.

RIBEIRO, V. C. *Análise de demanda por transportes de passageiros via modelos de regressão georeferenciados*. Tese (Doutorado), Universidade Federal do Espírito Santo, Vitória, ES, 2012.

ROCHA, S. S. *Análise de geração de viagens urbanas por transporte coletivo através de técnicas de geoestatística*. Dissertação (Mestrado), Universidade Federal da Bahia, Salvador, 2014.

ROCHA, S. S.; SALGUEIRO, A. R.; PITOMBO, C. S.. Interpolação Espacial de Viagens Urbanas por Transporte Coletivo Através de Amostra Sistemática Desagregada Artificial. *RBC. Revista Brasileira de Cartografia (Online)*, v. 68, p. 705-715, 2016.

RUSSO, F.; COMI, A. A modelling system to simulate goods movements at an urban scale. *Transportation*, 1-23, 2010.

SÁNCHEZ-DÍAZ, I. Modeling urban freight generation: A study of commercial establishments' freight needs. *Transportation Research Part A: Policy and Practice*, 2016.

SÁNCHEZ-DÍAZ, I.; HOLGUÍN-VERAS, J.; WANG, X. An exploratory analysis of spatial effects on freight trip attraction. *Transportation*, v. 43, n. 1, p. 177-196, 2016.

SILVA, J. E.; WAISMAN M. R. Cargas Urbanas: Estudo Exploratório sobre a Geração de Viagens de Caminhões em Bares e Restaurantes, In: *XVI Congresso da ANTP*, 2007, Maceió.

SOUZA, C. D. R.; SILVA, S. D.; ALMEIDA, M. (2010). Modelos de geração de viagem para pólos geradores de viagens de veículos de cargas. *Transportes*, 18(1), 2010.

STICH, B.; HOLLAND, J. H.; NOBREGA, R. A. A. ; O'HARA, C. G.. . Using multi-criteria decision making to highlight stakeholders values in the corridor planning process. *Journal of Transport and Land Use. JCR*, v. 4, p. 105-118, 2011.

TADI, R. R.; BALDACH, P. Truck trip generation characteristics of nonresidential land uses. *ITE Journal*, vol. 64, n. 7, pp. 43-47, 1994.

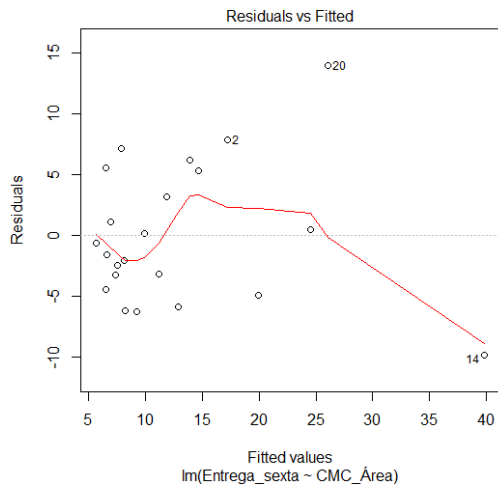
TANIGUCHI, E., THOMPSON, R.G., YAMADA, T.; DUIN, R. van. *City logistics – network modelling and intelligent transport systems*, Pergamon, 2001.

TEIXEIRA, G. L. *Uso de Dados Censitários para Identificação de Zonas Homogêneas para Planejamento de Transportes Utilizando Estatística Espacial*. 156 f. Dissertação (Mestrado). Universidade de Brasília. Faculdade de Tecnologia. Departamento de Engenharia Civil e Ambiental, Brasília, 2003.

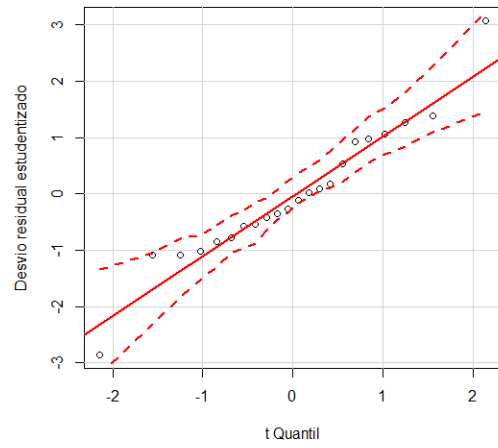
TOBLER, W. R. A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46:234-40, 1970.

TURKMAN, M. A. A., SILVA, G. L. Modelos Lineares Generalizados: da teoria à prática. In: *VIII Congresso Anual da Sociedade Portuguesa de Estatística*, Lisboa, 2000.

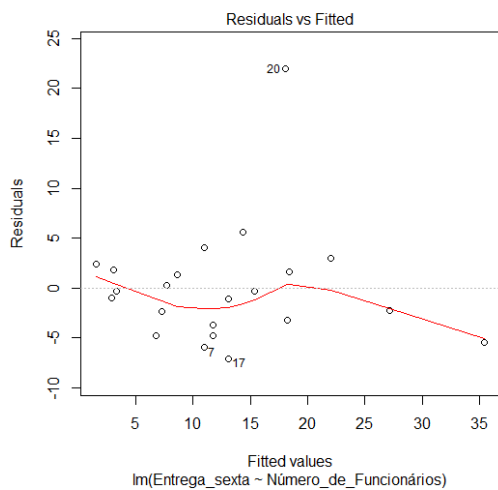
APÊNDICE A – ANÁLISE DE RESÍDUOS DOS MODELOS LINEARES



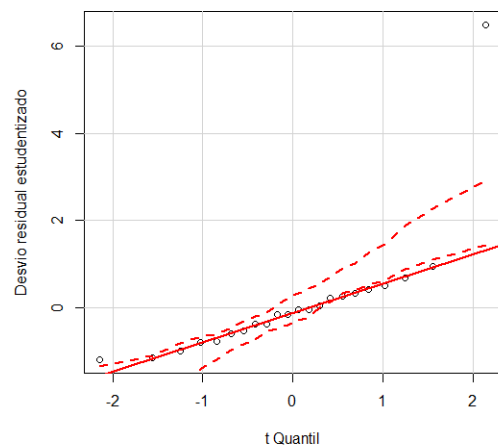
LM 01 – Resíduos versus valores ajustados.



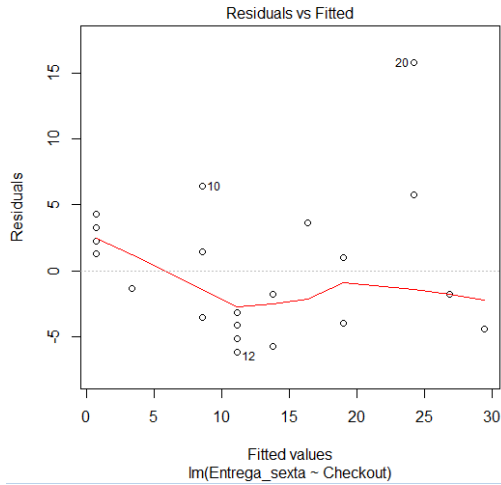
LM 01 – Gráfico normal de probabilidade.



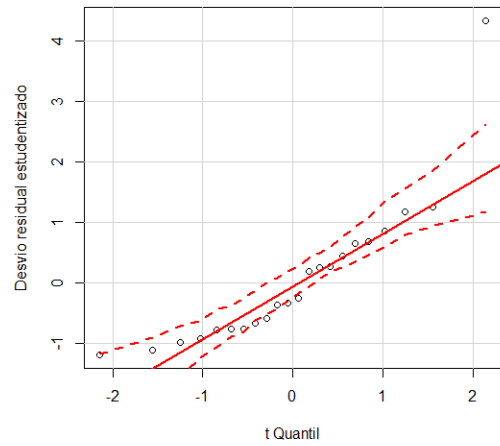
LM 02 – Resíduos versus valores ajustados.



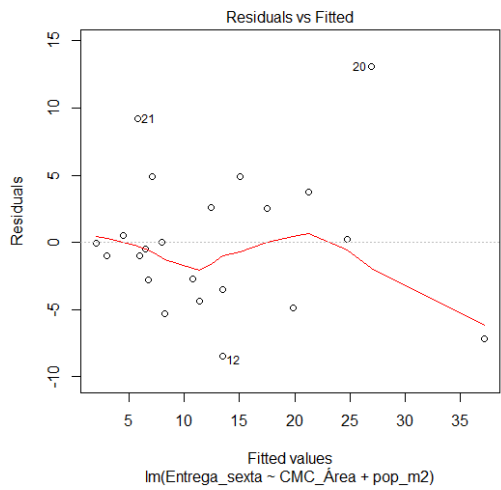
LM 02 – Gráfico normal de probabilidade.



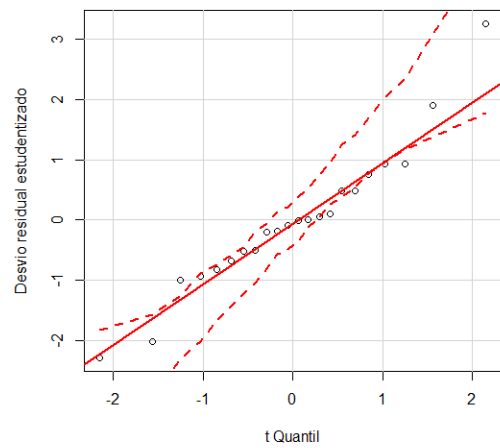
LM 03 – Resíduos versus valores ajustados.



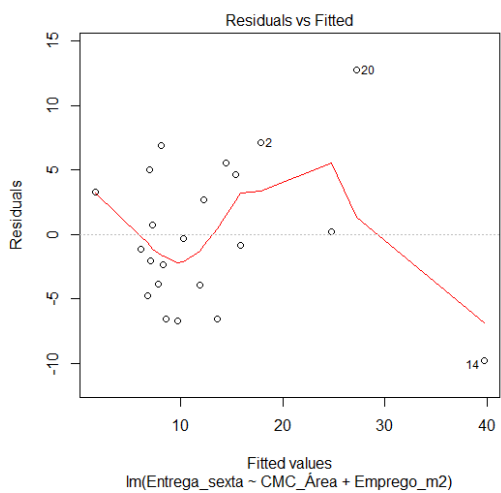
LM 03 – Gráfico normal de probabilidade.



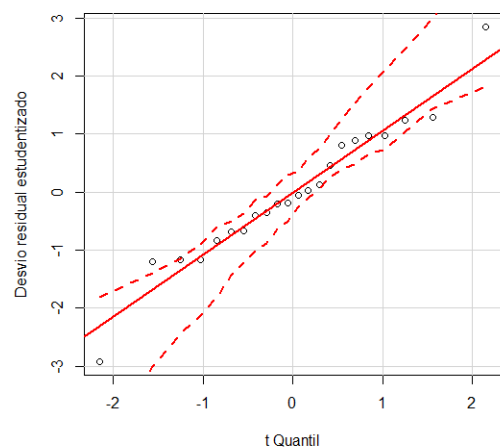
LM 04 – Resíduos versus valores ajustados.



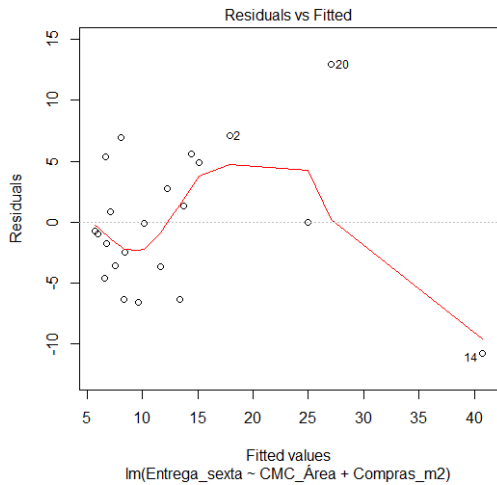
LM 04 – Gráfico normal de probabilidade.



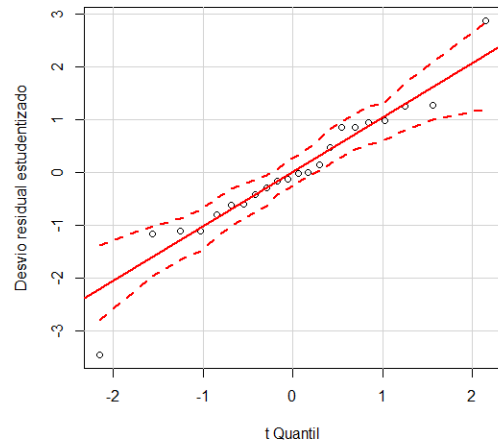
LM 05 – Resíduos versus valores ajustados.



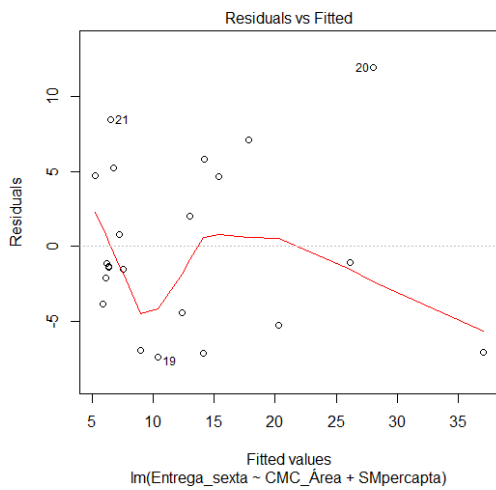
LM 05 – Gráfico normal de probabilidade.



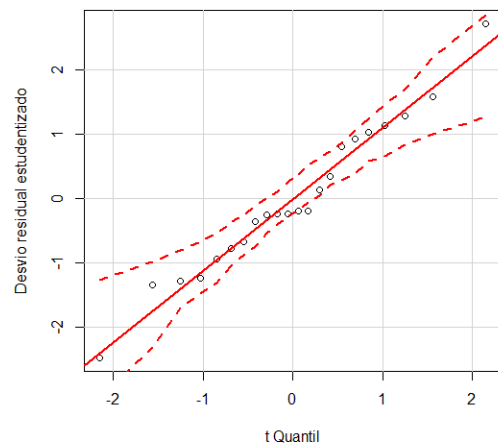
LM 06 – Resíduos versus valores ajustados.



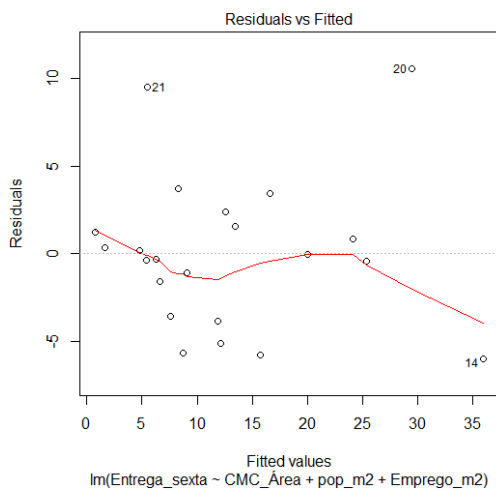
LM 06 – Gráfico normal de probabilidade.



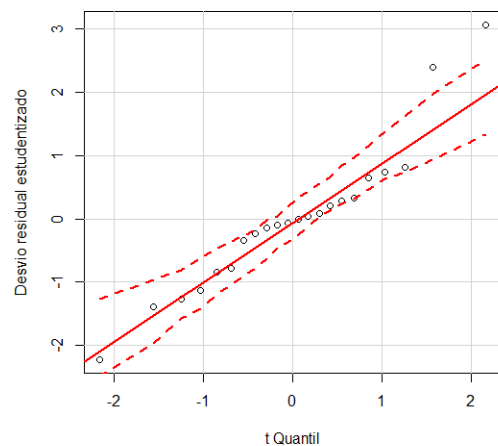
LM 07 – Resíduos versus valores ajustados.



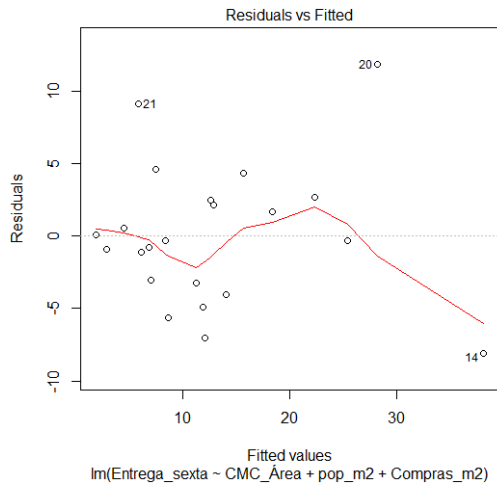
LM 07 – Gráfico normal de probabilidade.



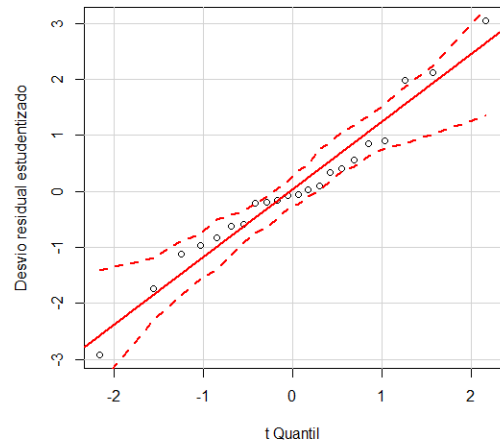
LM 08 – Resíduos versus valores ajustados.



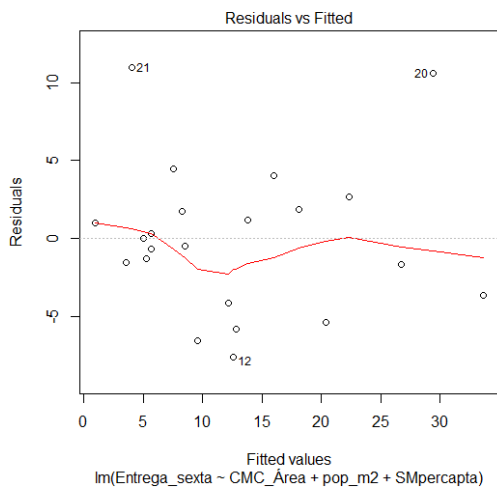
LM 08 – Gráfico normal de probabilidade.



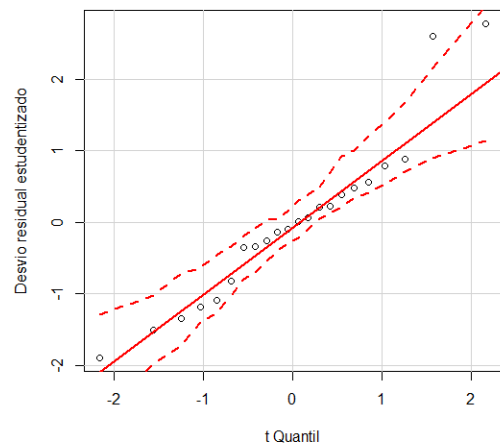
LM 09 – Resíduos versus valores ajustados.



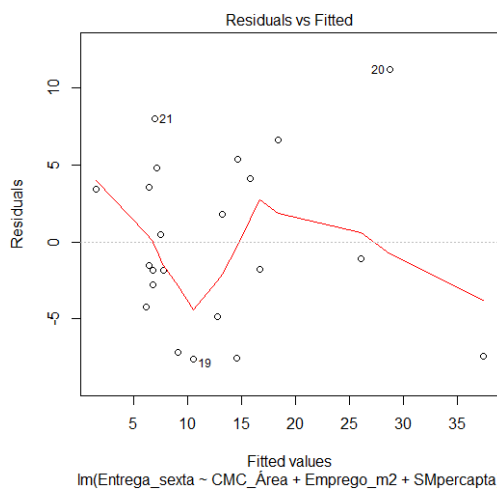
LM 09 – Gráfico normal de probabilidade.



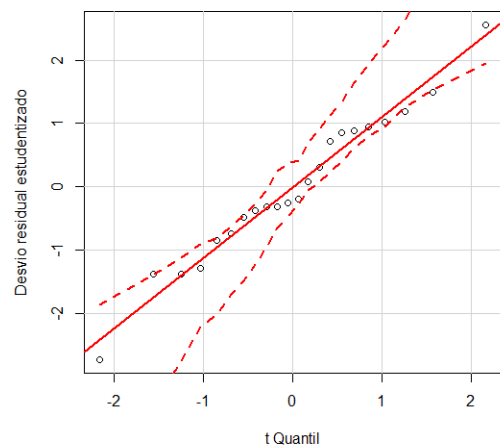
LM 10 – Resíduos versus valores ajustados.



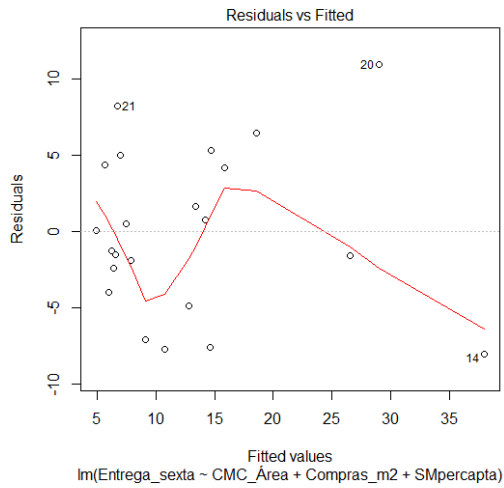
LM 10 – Gráfico normal de probabilidade.



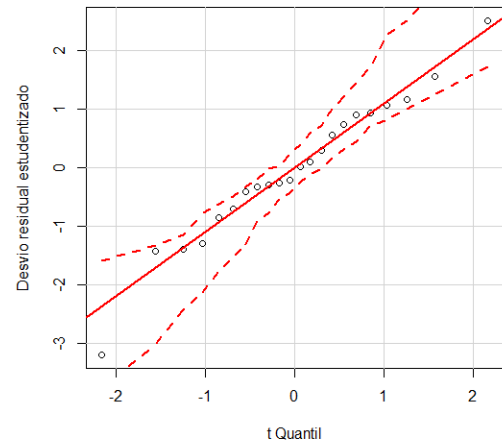
LM 11 – Resíduos versus valores ajustados.



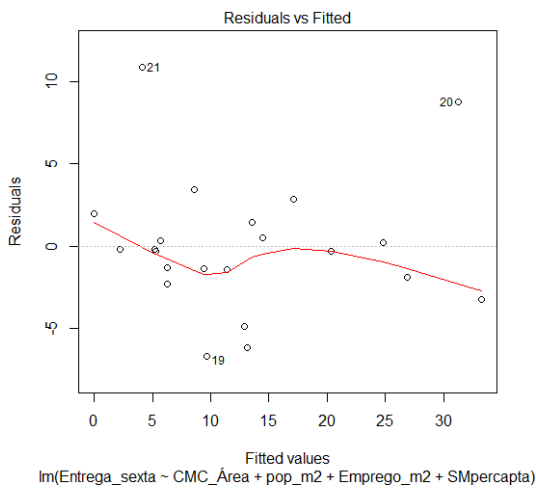
LM 11 – Gráfico normal de probabilidade.



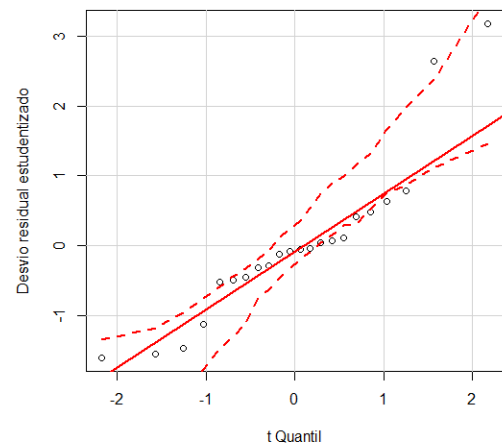
LM 12 – Resíduos versus valores ajustados.



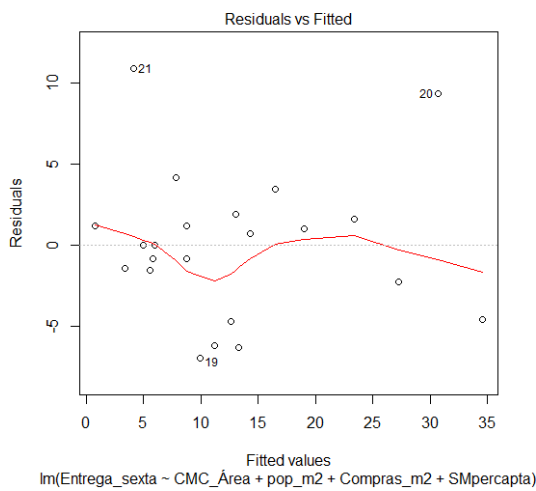
LM 12 – Gráfico normal de probabilidade.



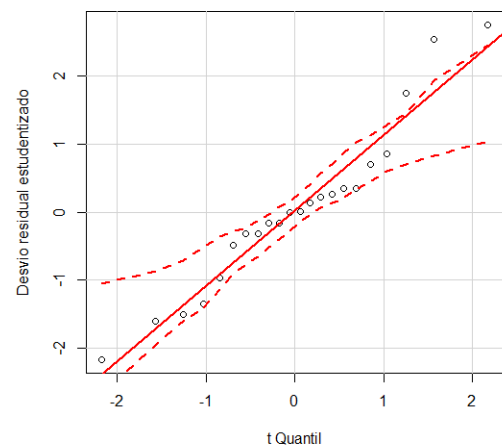
LM 13 – Resíduos versus valores ajustados.



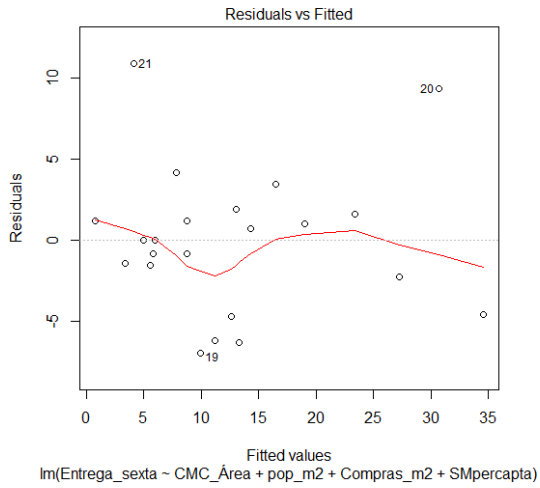
LM 13 – Gráfico normal de probabilidade.



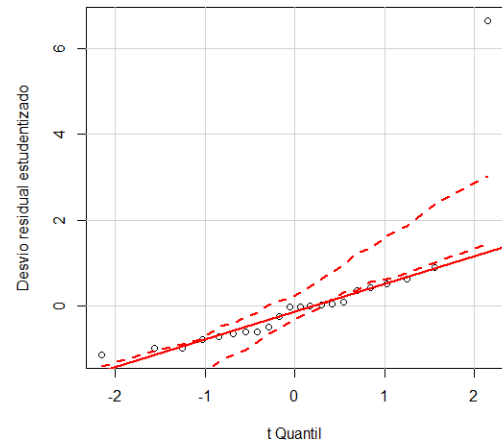
LM 14 – Resíduos versus valores ajustados.



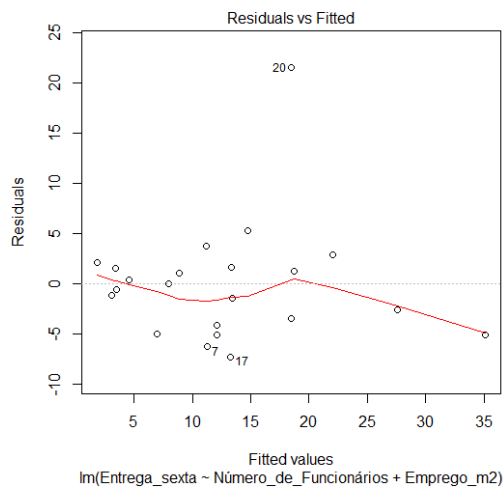
LM 14 – Gráfico normal de probabilidade.



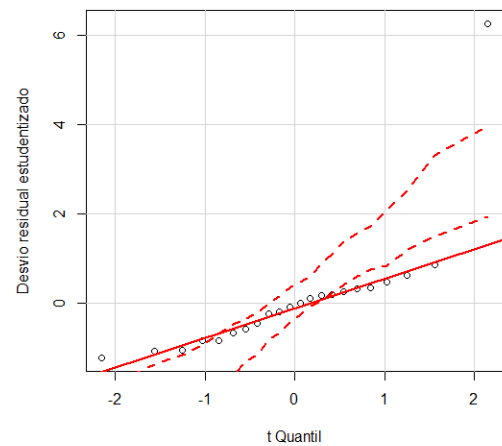
LM 15 – Resíduos versus valores ajustados.



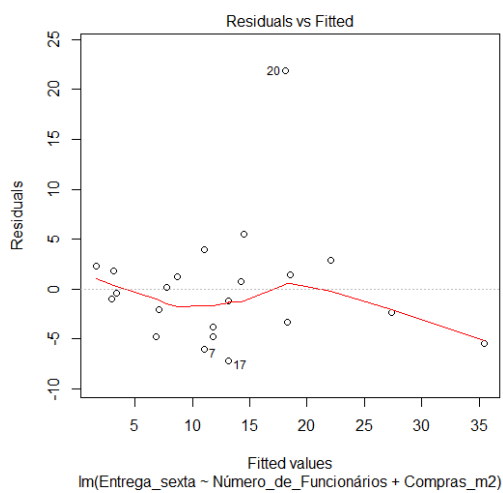
LM 15 – Gráfico normal de probabilidade.



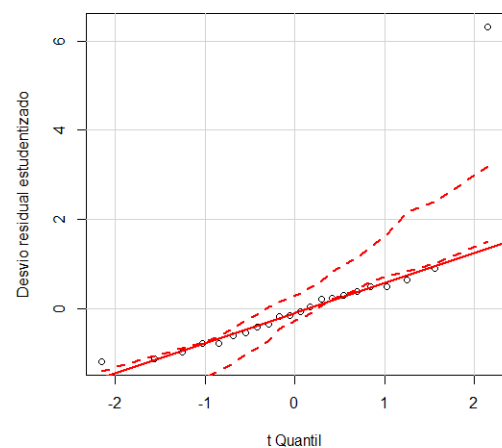
LM 16 – Resíduos versus valores ajustados.



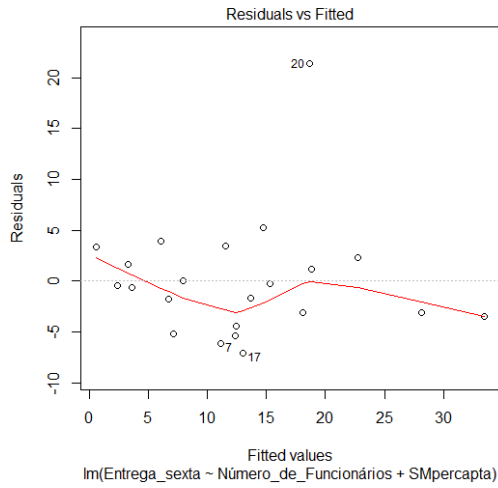
LM 16 – Gráfico normal de probabilidade.



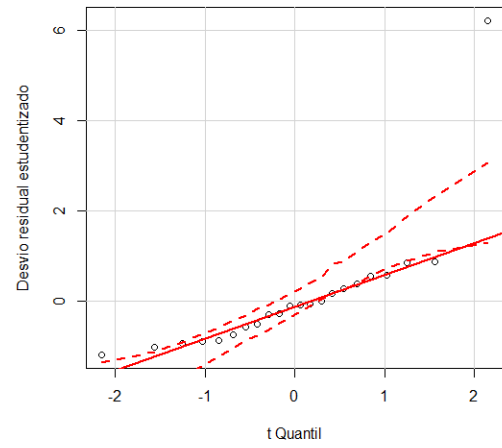
LM 17 – Resíduos versus valores ajustados.



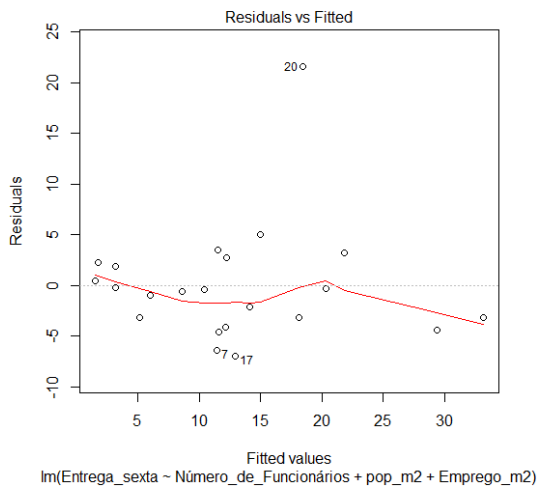
LM 17 – Gráfico normal de probabilidade.



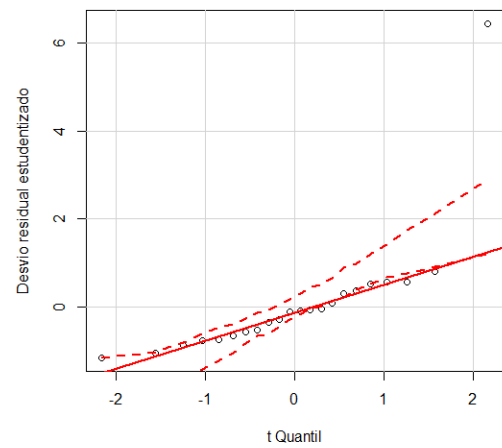
LM 18 – Resíduos versus valores ajustados.



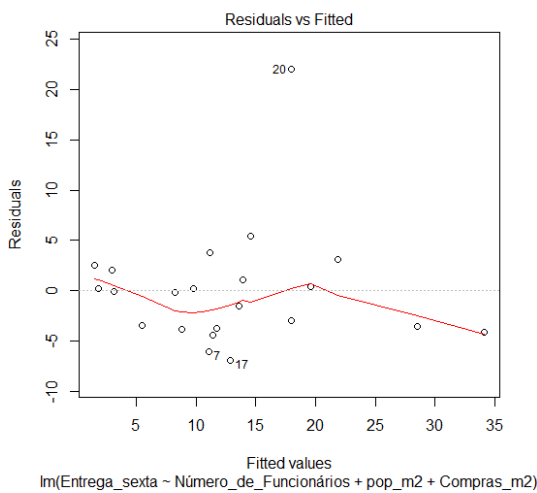
LM 18 – Gráfico normal de probabilidade.



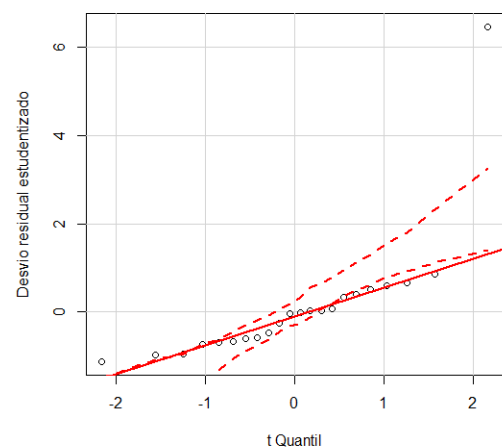
LM 19 – Resíduos versus valores ajustados.



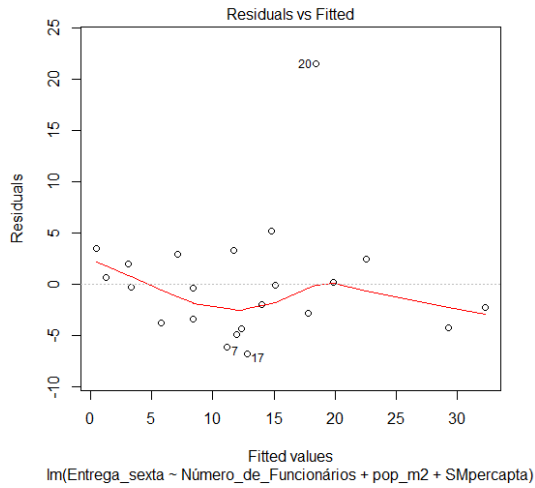
LM 19 – Gráfico normal de probabilidade.



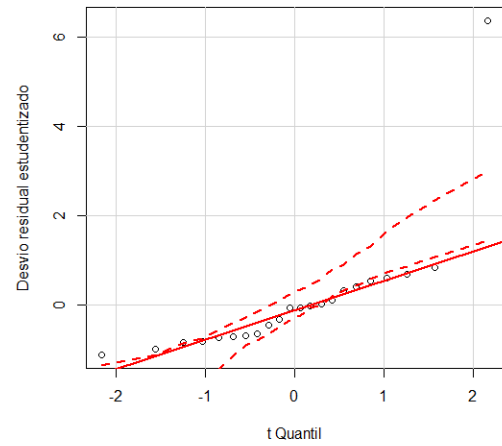
LM 20 – Resíduos versus valores ajustados.



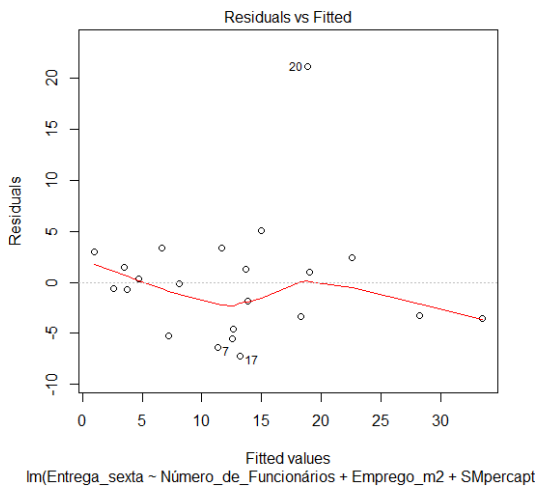
LM 20 – Gráfico normal de probabilidade.



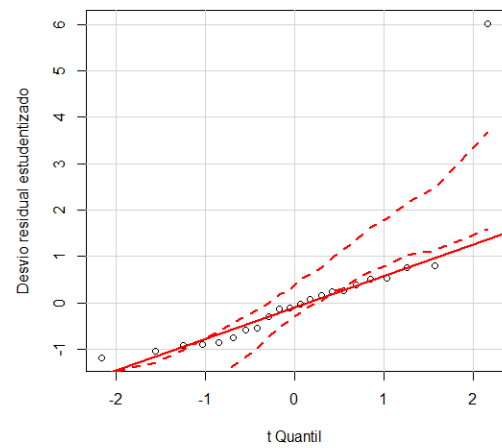
LM 21 – Resíduos versus valores ajustados.



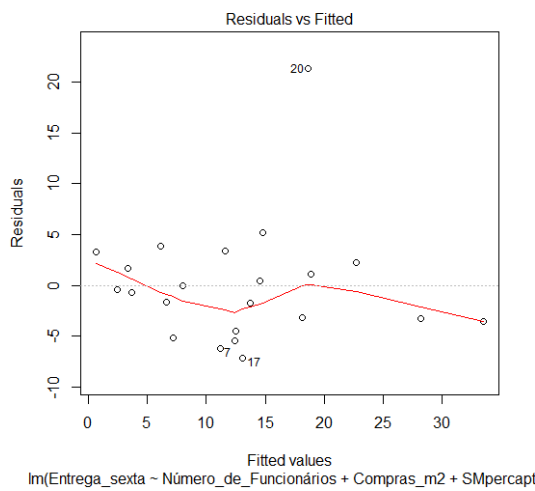
LM 21 – Gráfico normal de probabilidade.



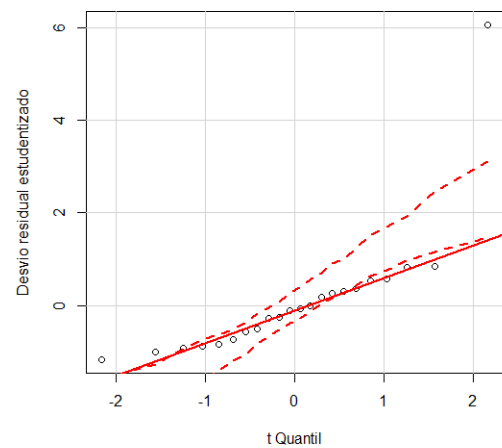
LM 22 – Resíduos versus valores ajustados.



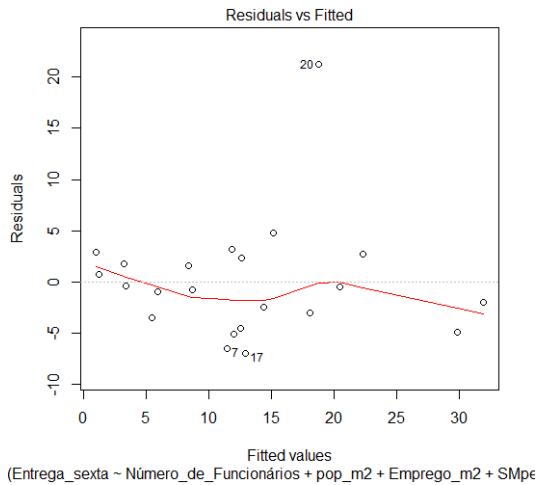
LM 22 – Gráfico normal de probabilidade.



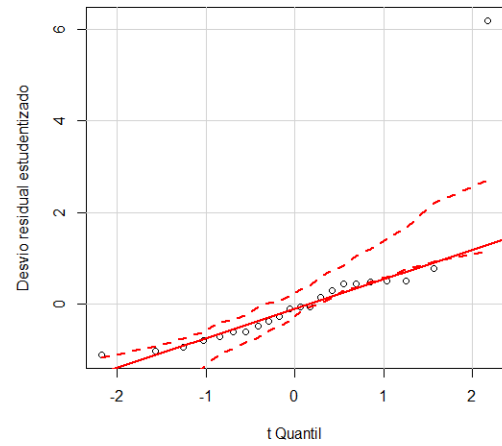
LM 23 – Resíduos versus valores ajustados.



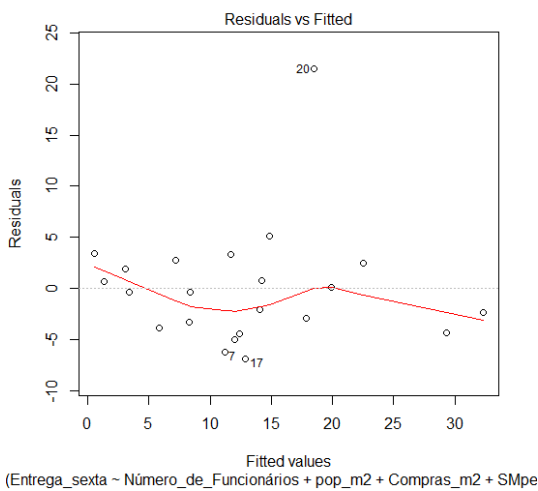
LM 23 – Gráfico normal de probabilidade.



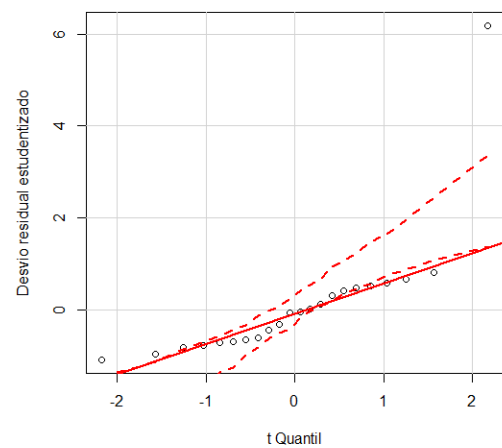
LM 24 – Resíduos versus valores ajustados.



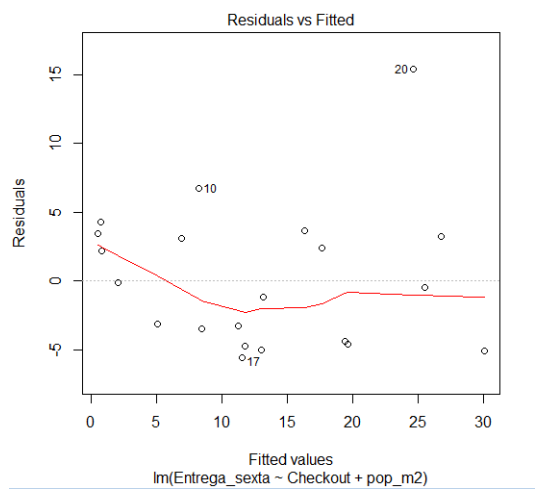
LM 24 – Gráfico normal de probabilidade.



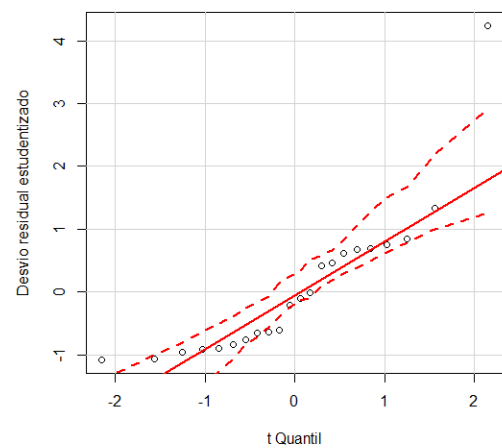
LM 25 – Resíduos versus valores ajustados.



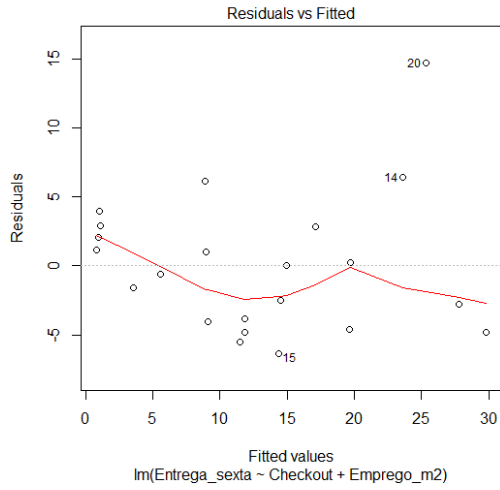
LM 25 – Gráfico normal de probabilidade.



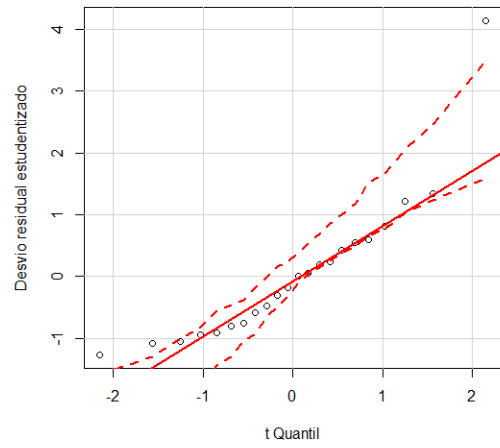
LM 26 – Resíduos versus valores ajustados.



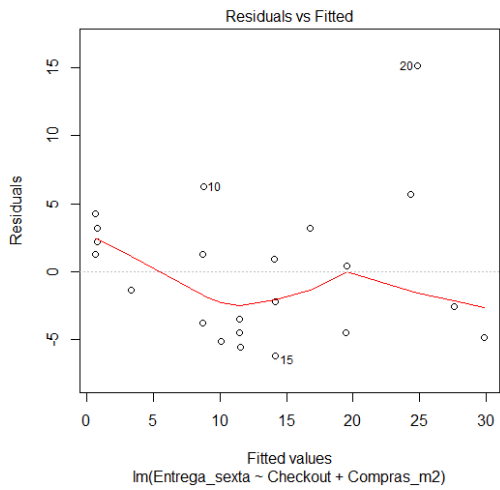
LM 26 – Gráfico normal de probabilidade.



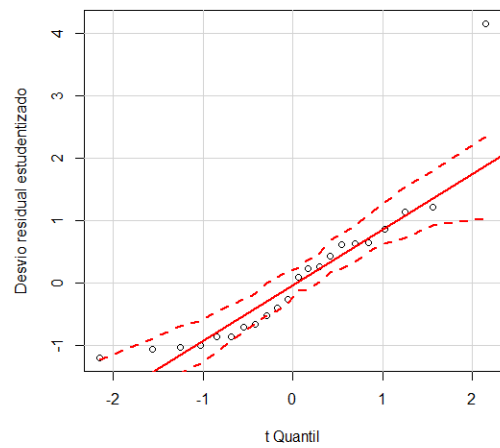
LM 27 – Resíduos versus valores ajustados.



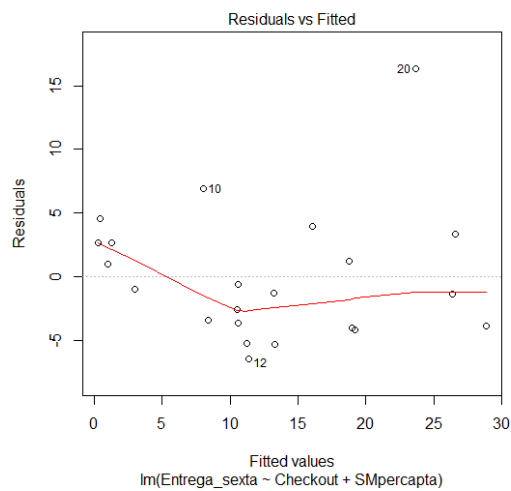
LM 27 – Gráfico normal de probabilidade.



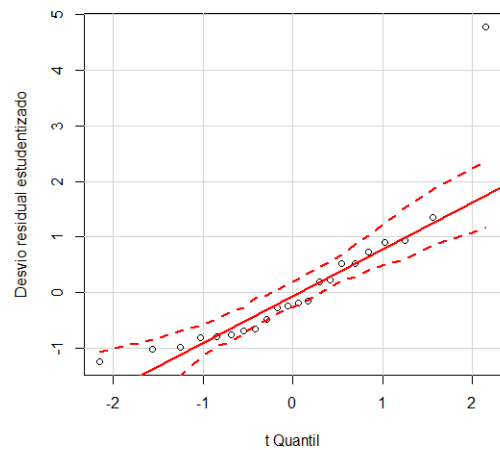
LM 28 – Resíduos versus valores ajustados.



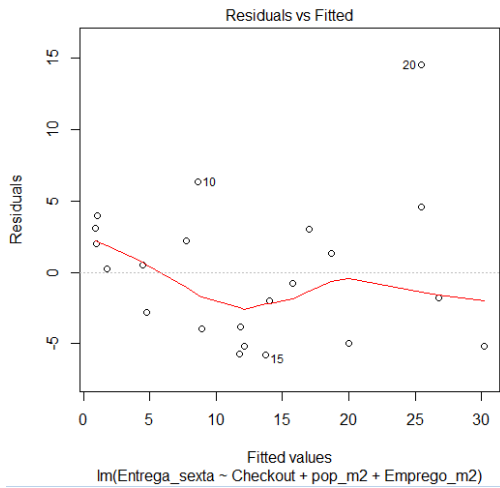
LM 28 – Gráfico normal de probabilidade.



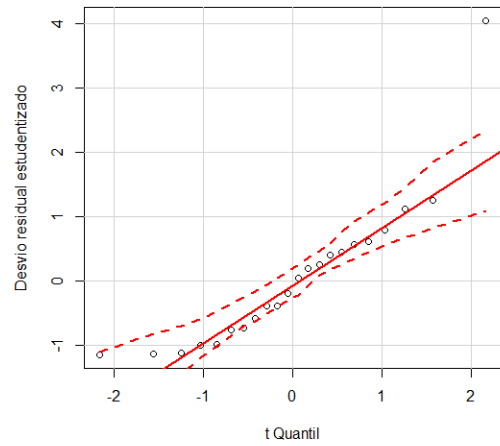
LM 29 – Resíduos versus valores ajustados.



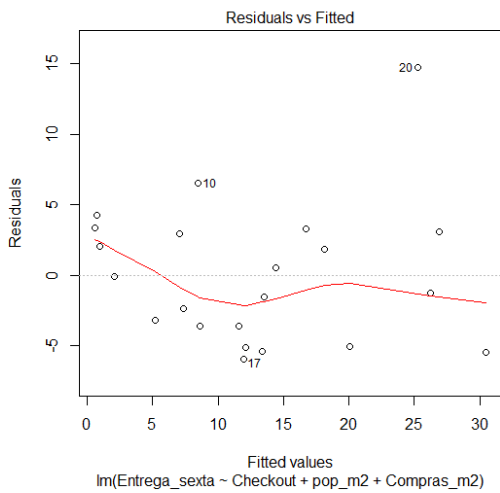
LM 29 – Gráfico normal de probabilidade.



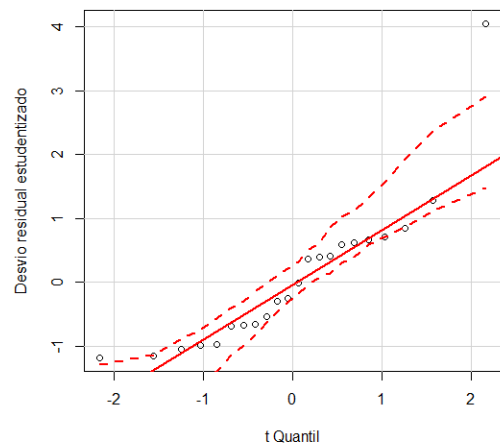
LM 30 – Resíduos versus valores ajustados.



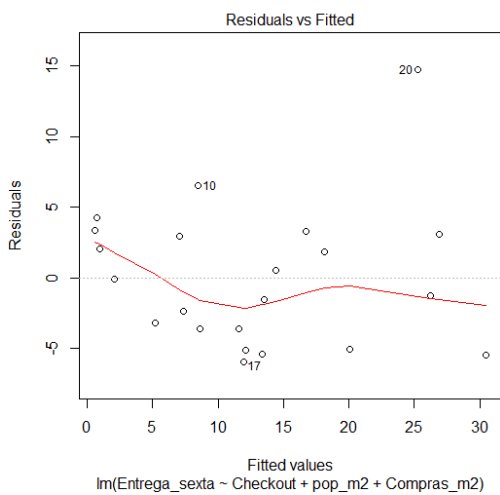
LM 30 – Gráfico normal de probabilidade.



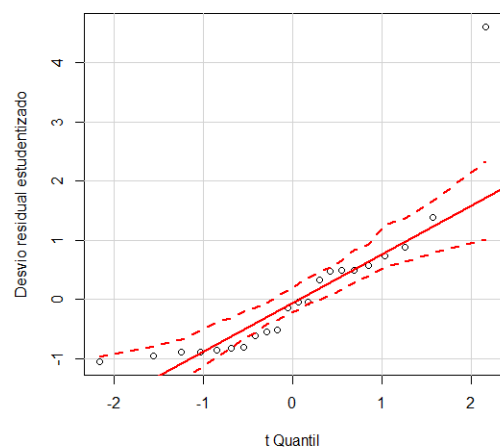
LM 31 – Resíduos versus valores ajustados.



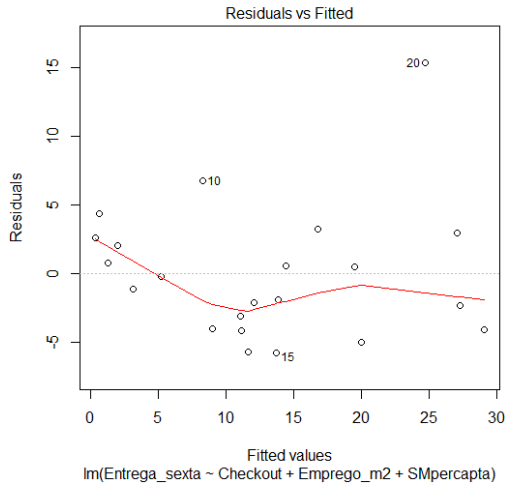
LM 31 – Gráfico normal de probabilidade.



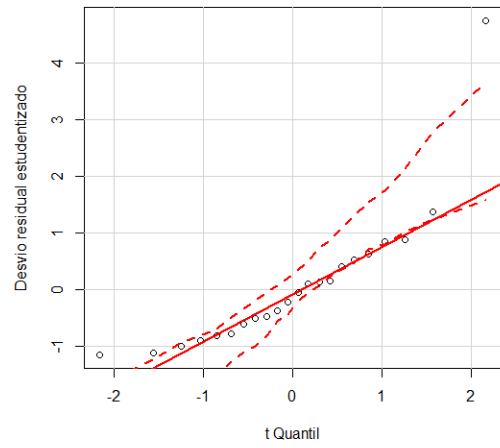
LM 32 – Resíduos versus valores ajustados.



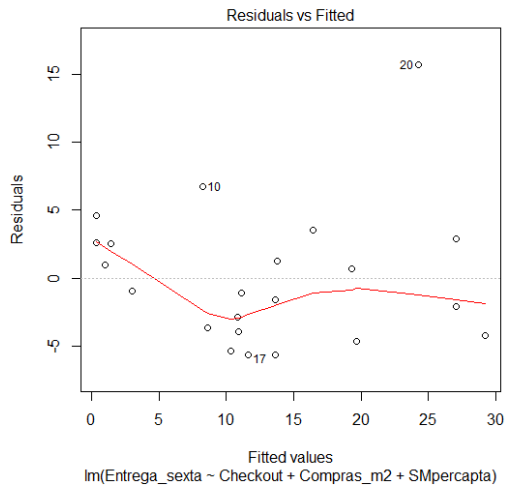
LM 32 – Gráfico normal de probabilidade.



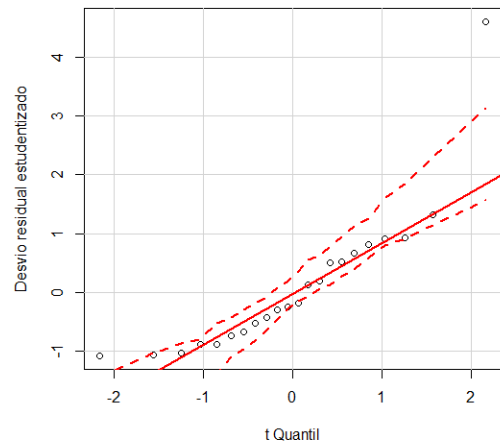
LM 33 – Resíduos versus valores ajustados.



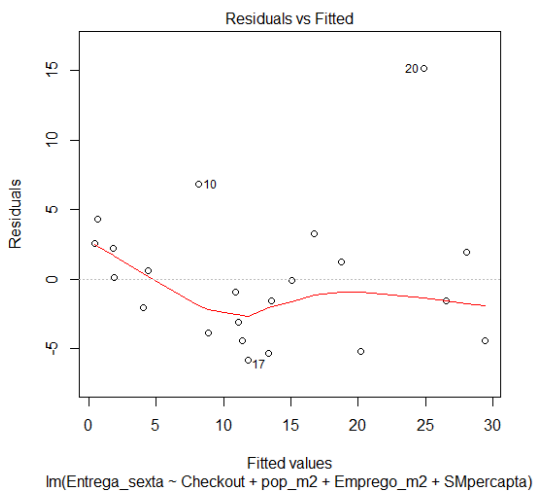
LM 33 – Gráfico normal de probabilidade.



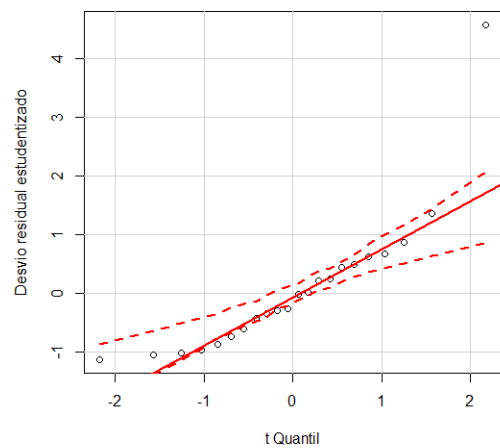
LM 34 – Resíduos versus valores ajustados.



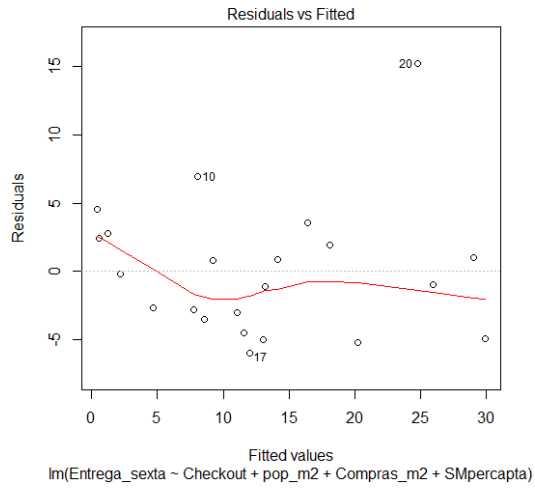
LM 34 – Gráfico normal de probabilidade.



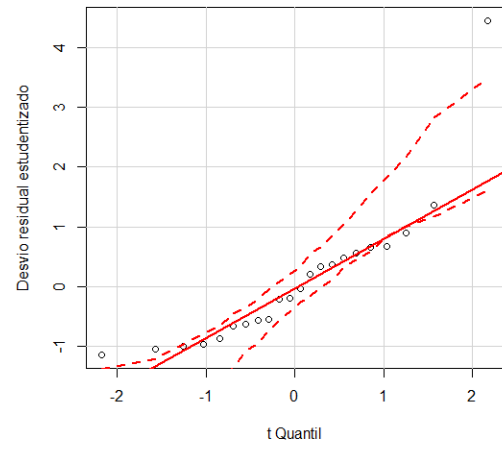
LM 35 – Resíduos versus valores ajustados.



LM 35 – Gráfico normal de probabilidade.

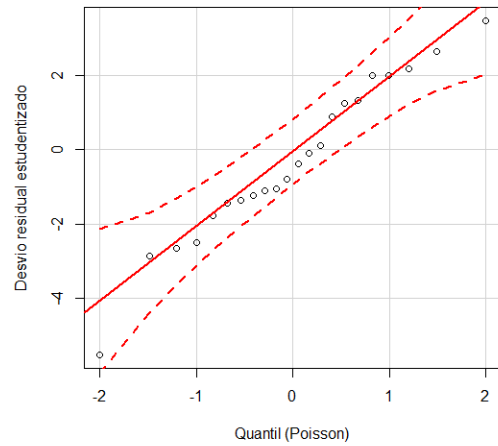
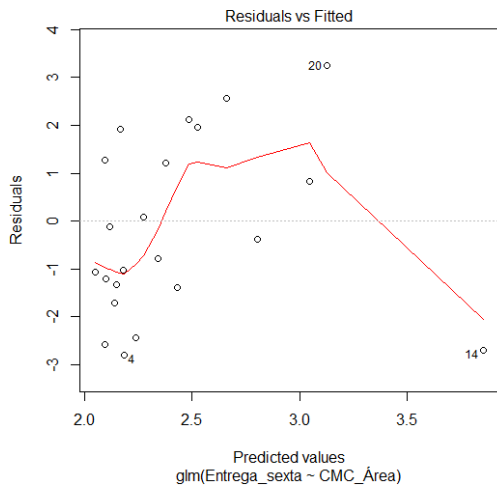


LM 36 – Resúdos versus valores ajustados.

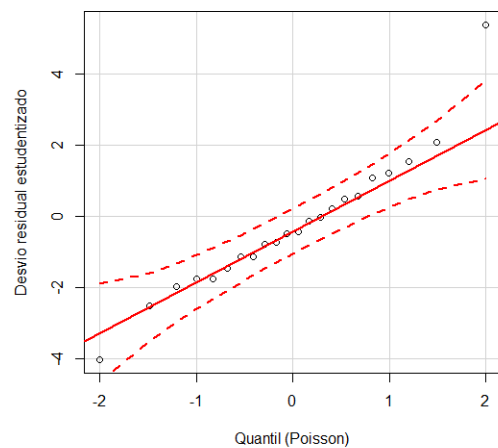
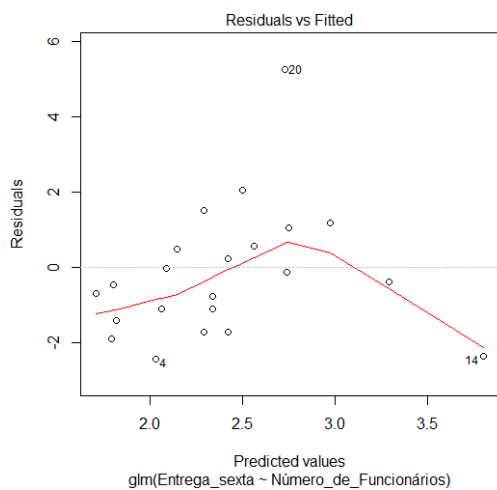


LM 36 – Gráfico normal de probabilidade.

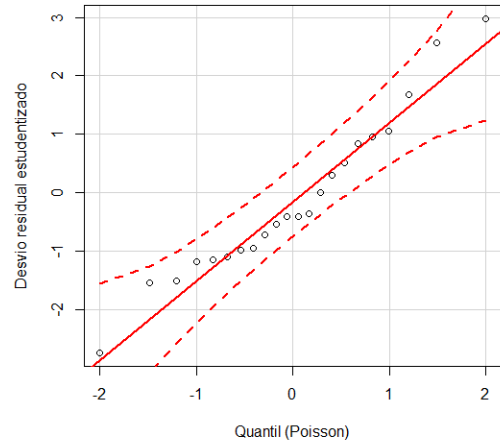
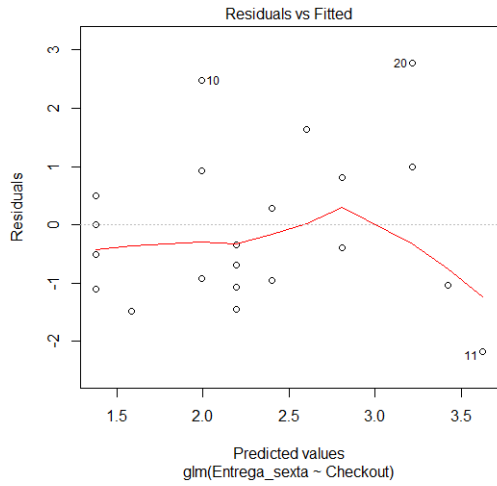
APÊNDICE B – ANÁLISE DE RESÍDUOS DOS MODELOS LINEARES GENERALIZADOS



GLM 01 – Resíduos versus valores ajustados. GLM 01 – Gráfico normal de probabilidade.

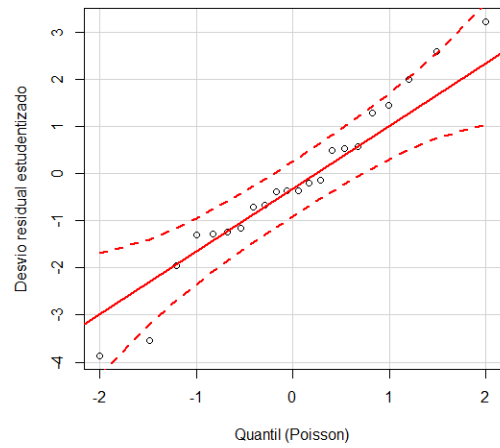
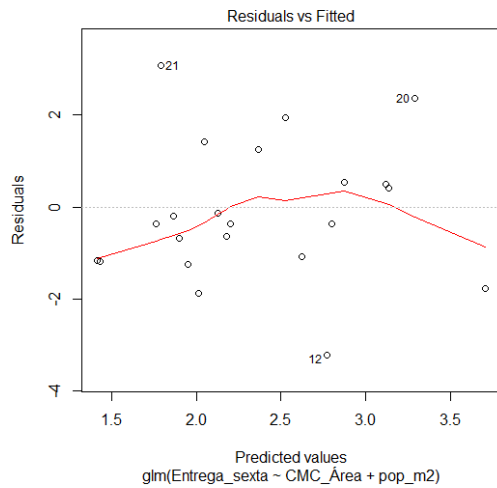


GLM 02 – Resíduos versus valores ajustados. GLM 02 – Gráfico normal de probabilidade.



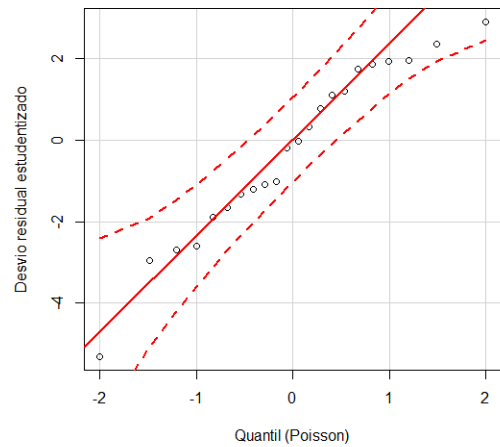
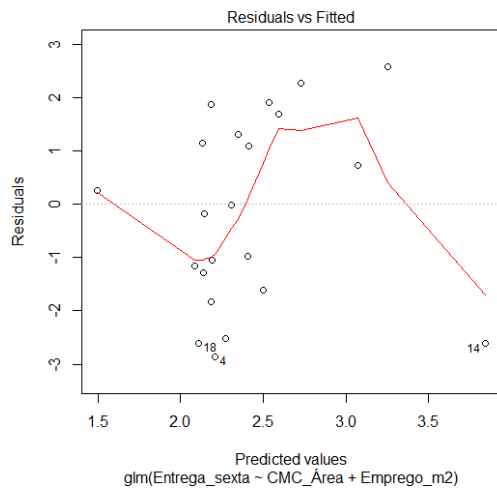
GLM 03 – Resíduos versus valores ajustados.

GLM 03 – Gráfico normal de probabilidade.



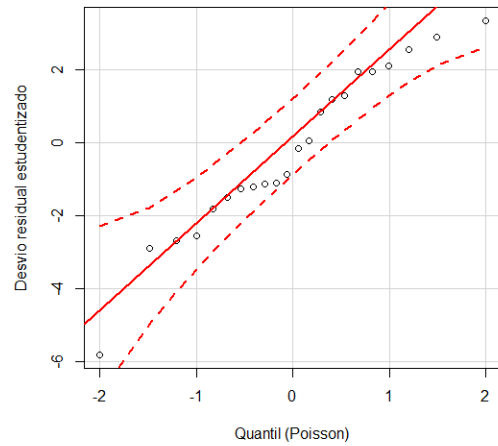
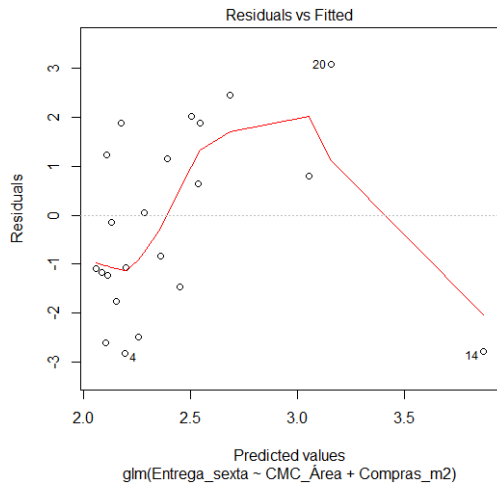
GLM 04 – Resíduos versus valores ajustados.

GLM 04 – Gráfico normal de probabilidade.



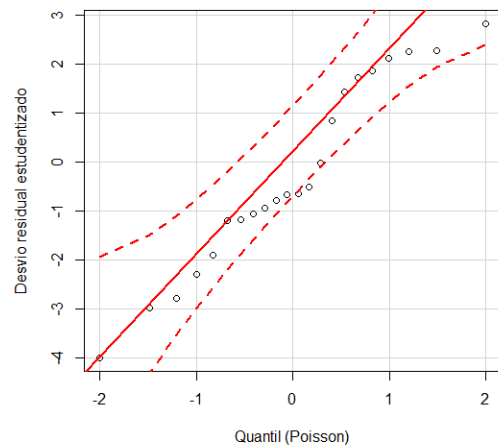
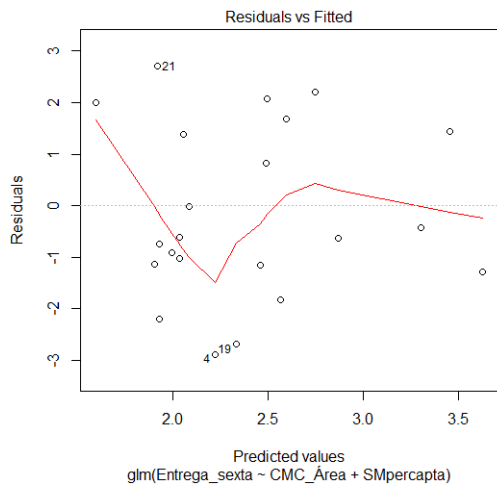
GLM 05 – Resíduos versus valores ajustados.

GLM 05 – Gráfico normal de probabilidade.



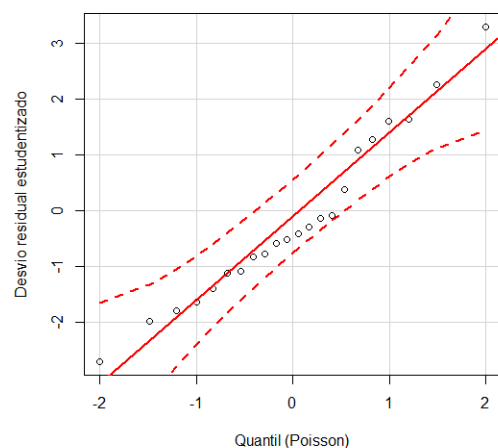
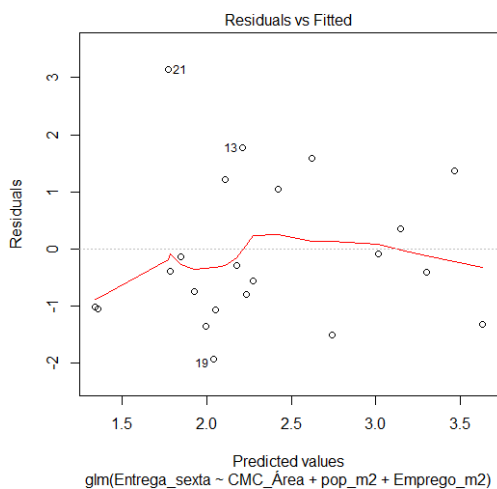
GLM 06 – Resíduos versus valores ajustados.

GLM 06 – Gráfico normal de probabilidade.



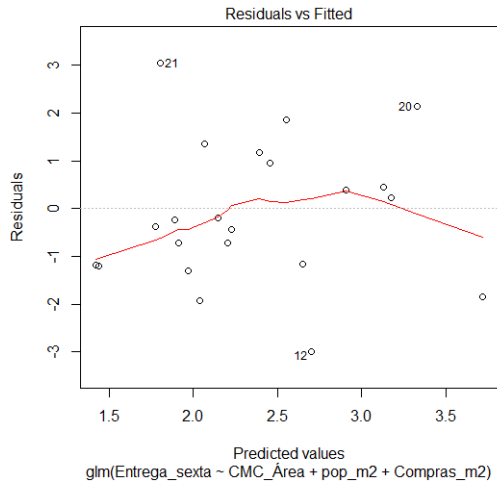
GLM 07 – Resíduos versus valores ajustados.

GLM 07 – Gráfico normal de probabilidade.

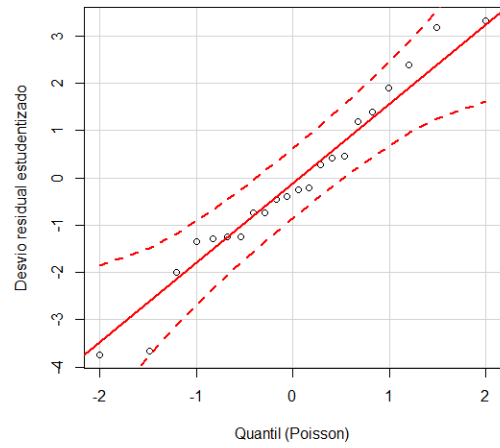


GLM 08 – Resíduos versus valores ajustados.

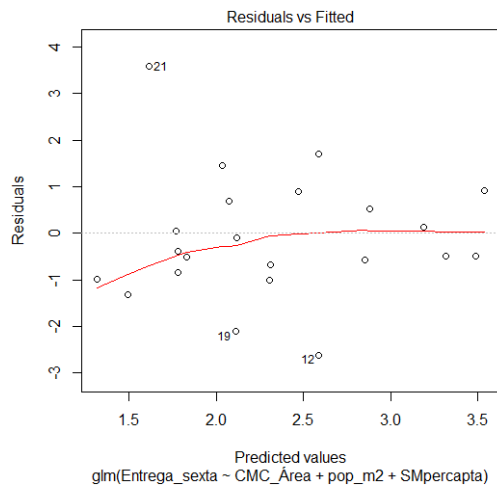
GLM 08 – Gráfico normal de probabilidade.



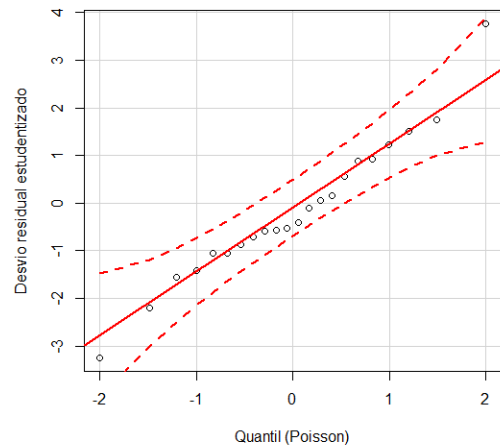
GLM 09 – Resíduos versus valores ajustados.



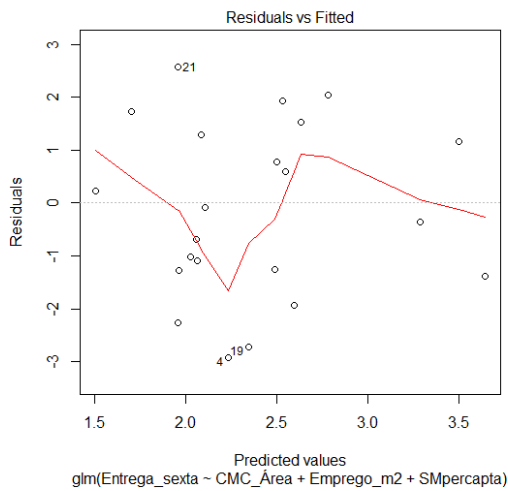
GLM 09 – Gráfico normal de probabilidade.



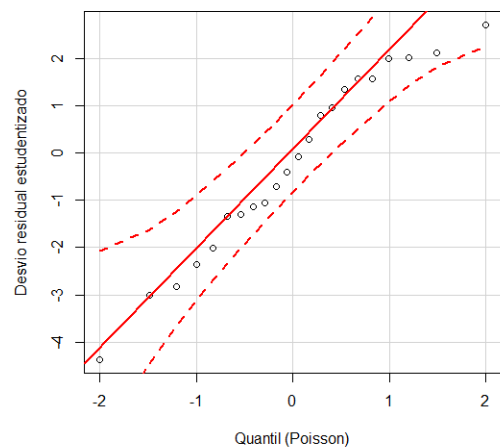
GLM 10 – Resíduos versus valores ajustados.



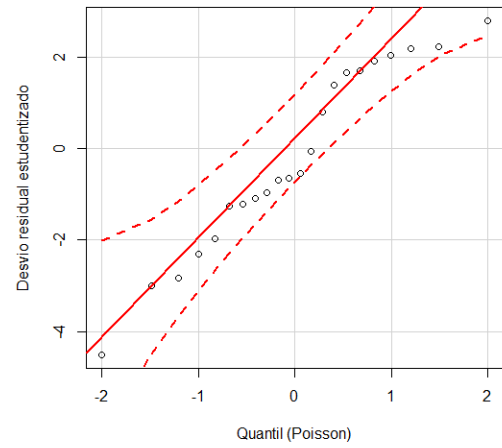
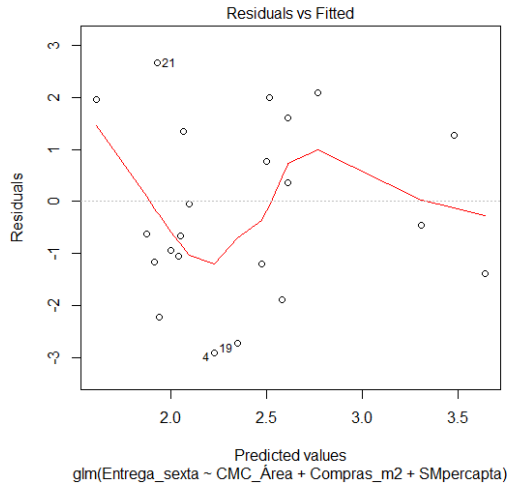
GLM 10 – Gráfico normal de probabilidade.



GLM 11 – Resíduos versus valores ajustados.

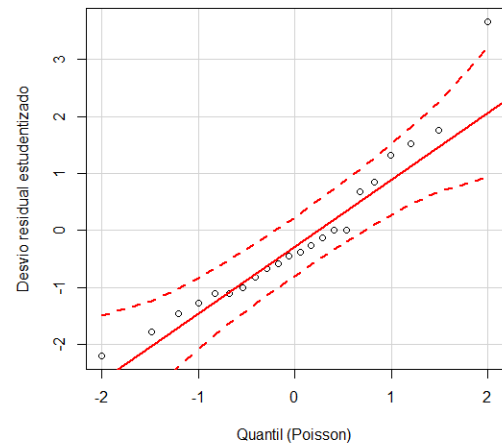
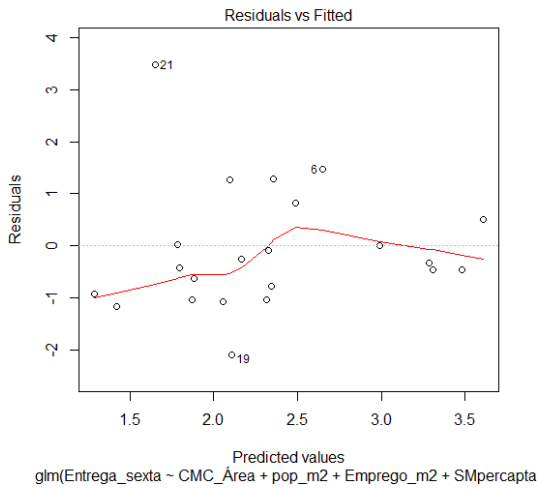


GLM 11 – Gráfico normal de probabilidade.



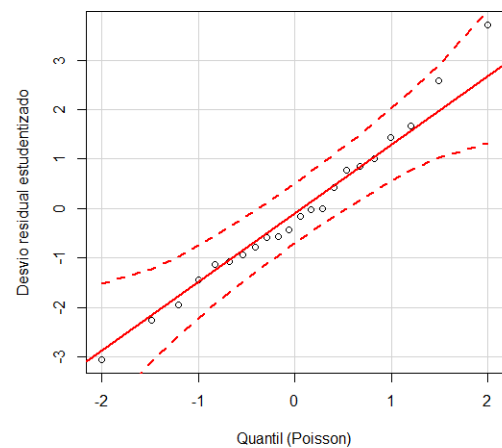
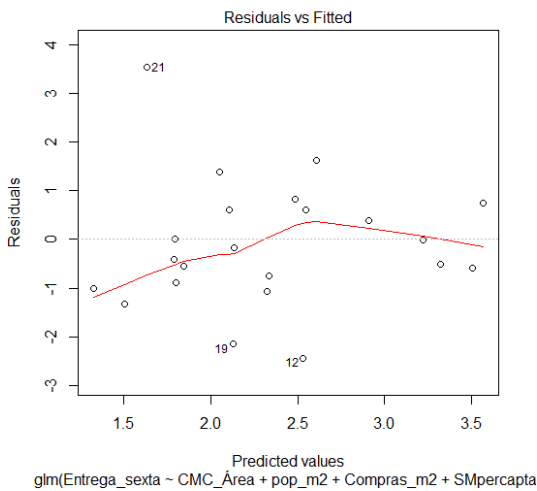
GLM 12 – Resíduos versus valores ajustados.

GLM 12 – Gráfico normal de probabilidade.



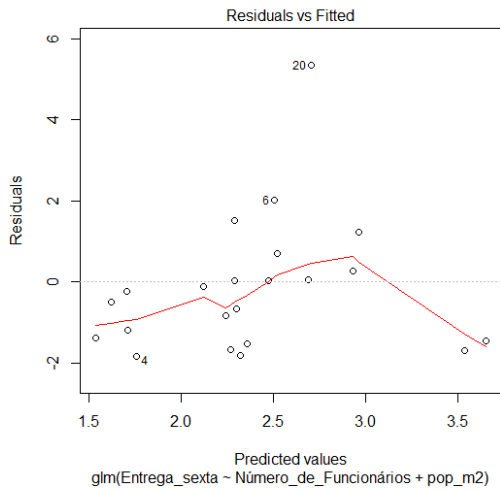
GLM 13 – Resíduos versus valores ajustados.

GLM 13 – Gráfico normal de probabilidade.

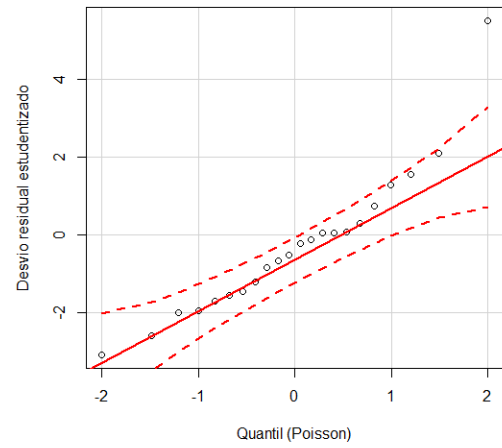


GLM 14 – Resíduos versus valores ajustados.

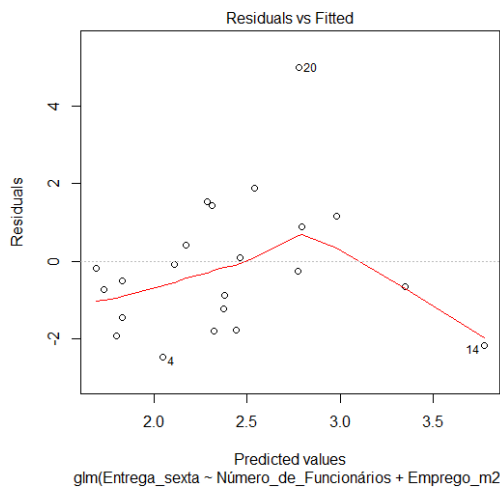
GLM 14 – Gráfico normal de probabilidade.



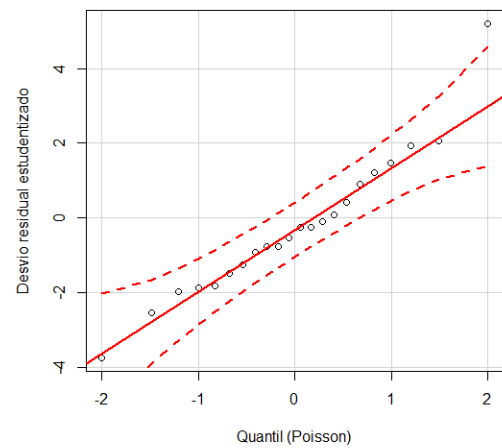
GLM 15 – Resíduos versus valores ajustados.



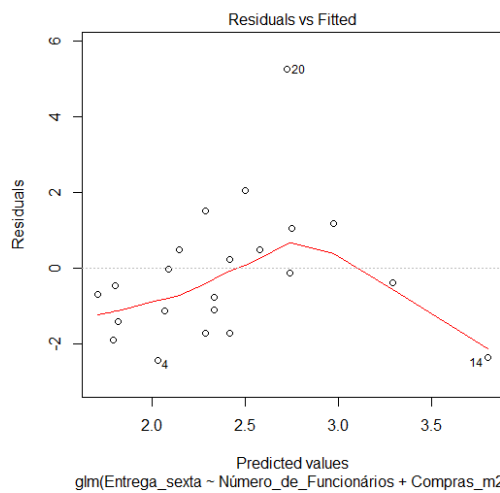
GLM 15 – Gráfico normal de probabilidade.



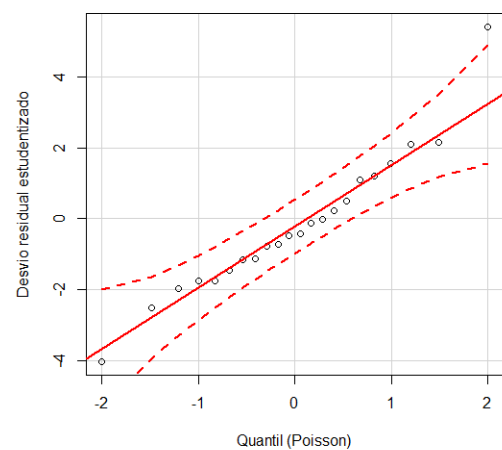
GLM 16 – Resíduos versus valores ajustados.



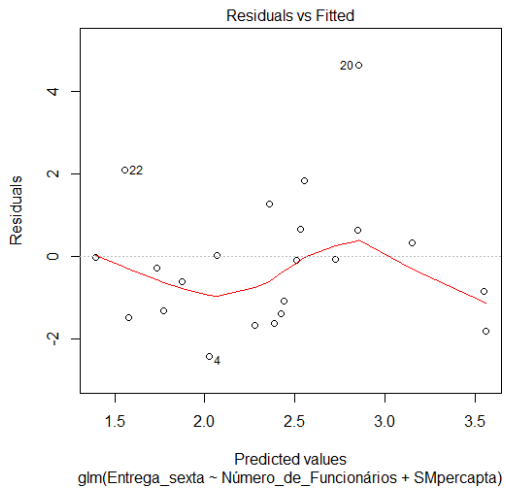
GLM 16 – Gráfico normal de probabilidade.



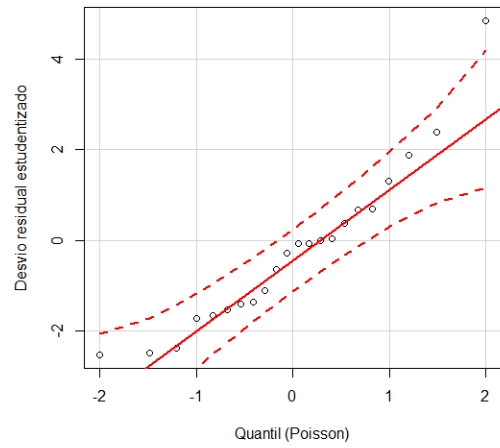
GLM 17 – Resíduos versus valores ajustados.



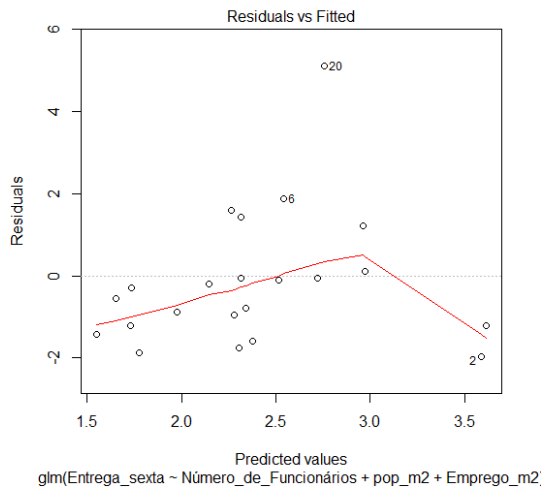
GLM 17 – Gráfico normal de probabilidade.



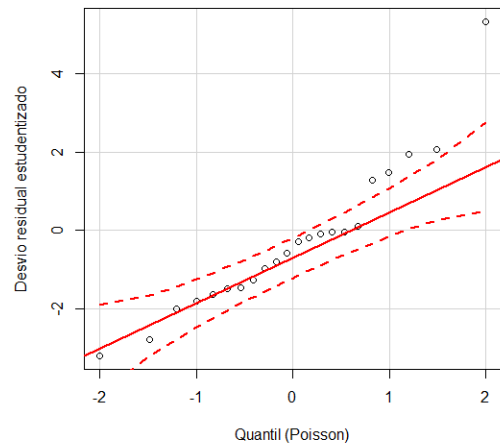
GLM 18 – Resíduos versus valores ajustados.



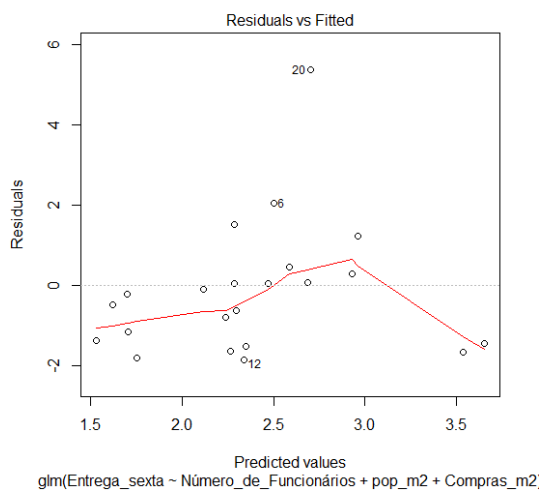
GLM 18 – Gráfico normal de probabilidade.



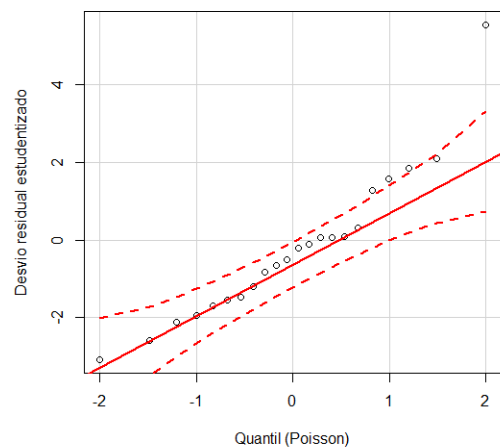
GLM 19 – Resíduos versus valores ajustados.



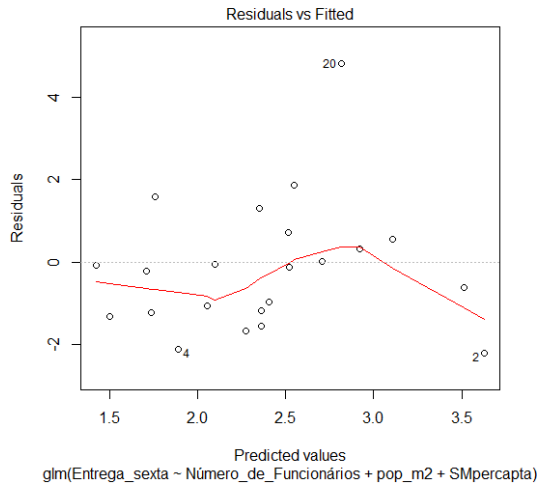
GLM 19 – Gráfico normal de probabilidade.



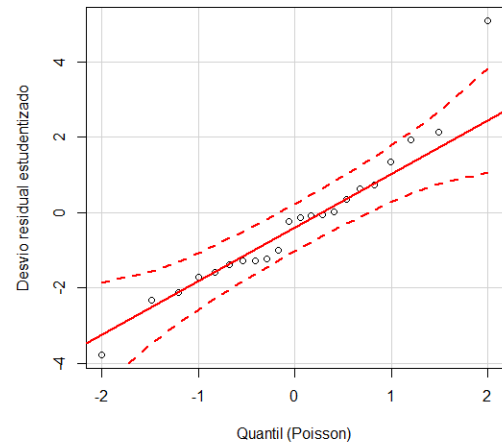
GLM 20 – Resíduos versus valores ajustados.



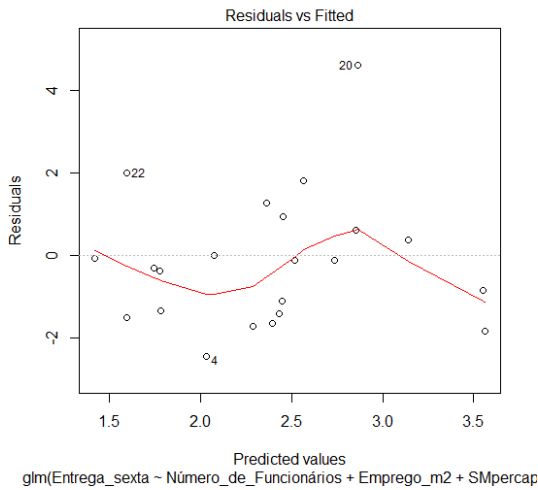
GLM 20 – Gráfico normal de probabilidade.



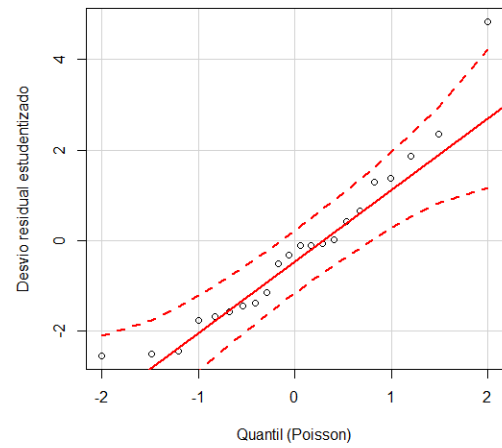
GLM 21 – Resíduos versus valores ajustados.



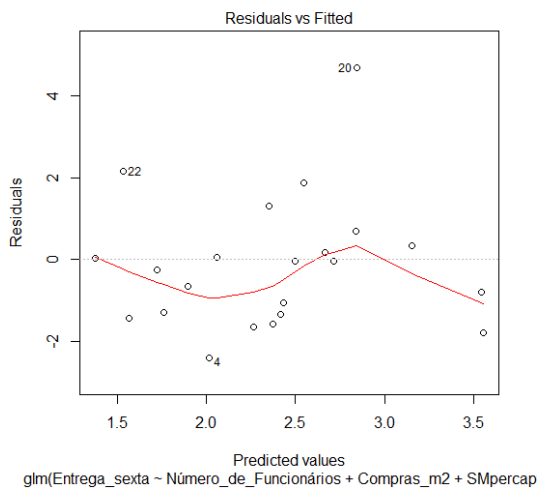
GLM 21 – Gráfico normal de probabilidade.



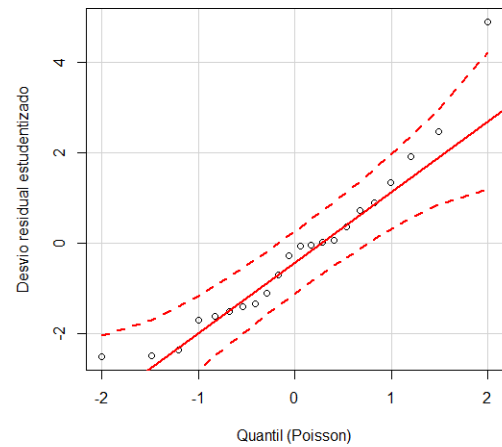
GLM 22 – Resíduos versus valores ajustados.



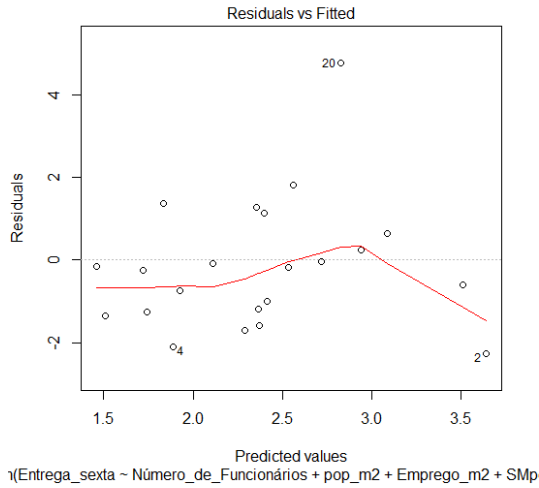
GLM 22 – Gráfico normal de probabilidade.



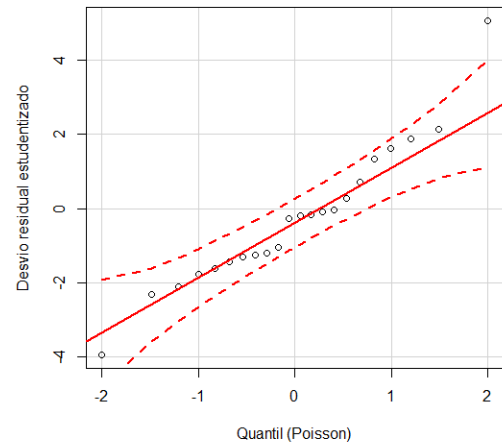
GLM 23 – Resíduos versus valores ajustados.



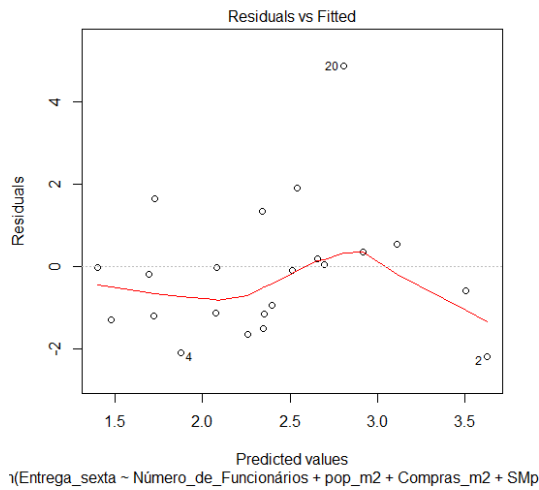
GLM 23 – Gráfico normal de probabilidade.



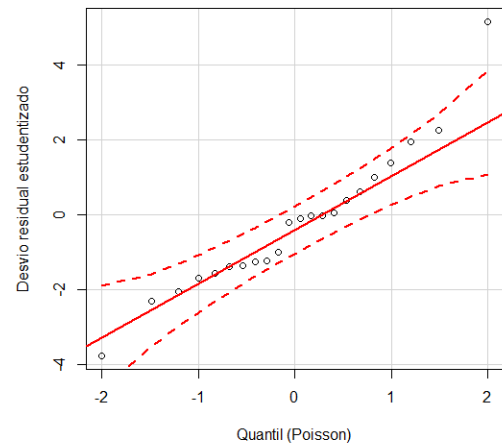
GLM 24 – Resíduos versus valores ajustados.



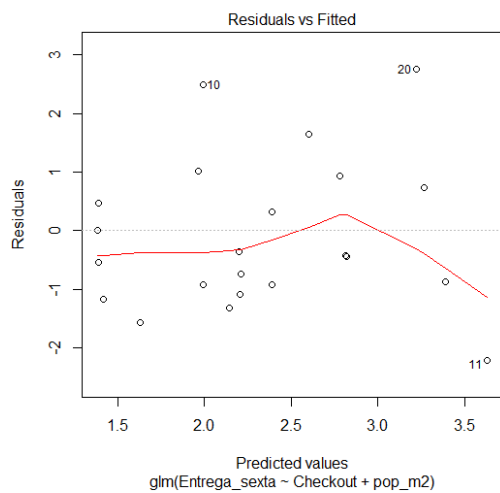
GLM 24 – Gráfico normal de probabilidade.



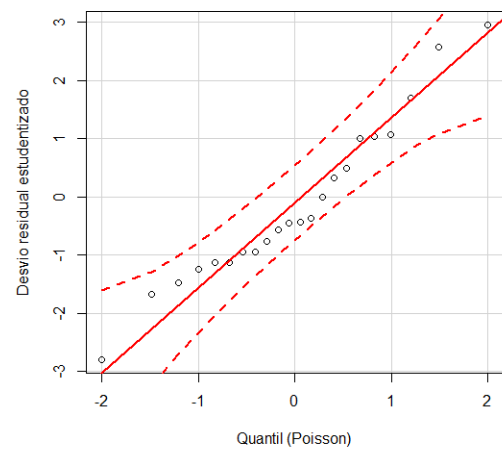
GLM 25 – Resíduos versus valores ajustados.



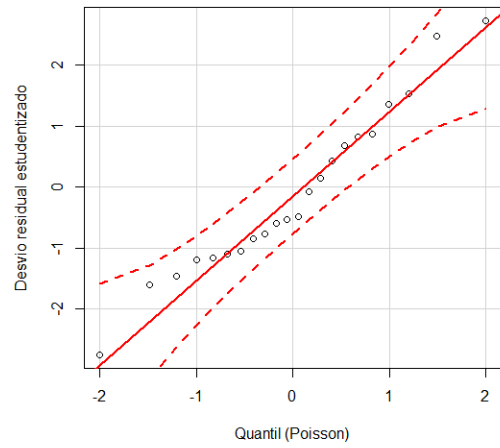
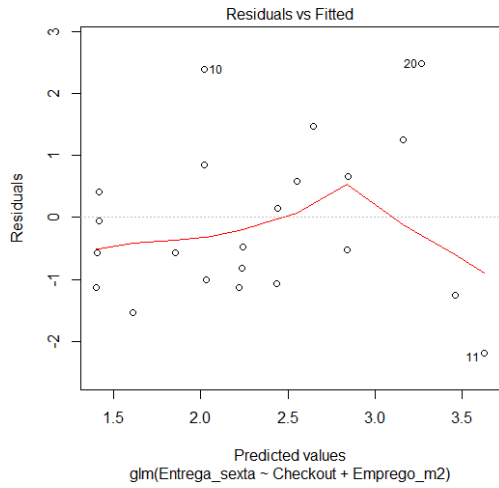
GLM 25 – Gráfico normal de probabilidade.



GLM 26 – Resíduos versus valores ajustados.

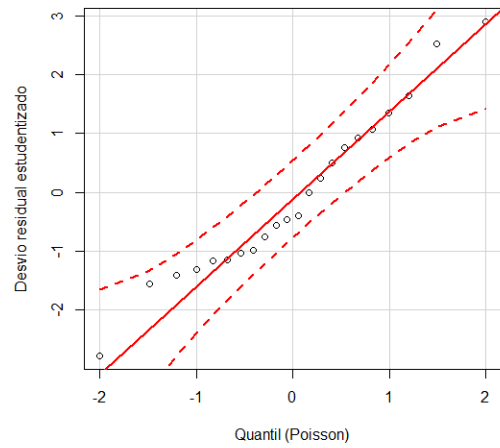
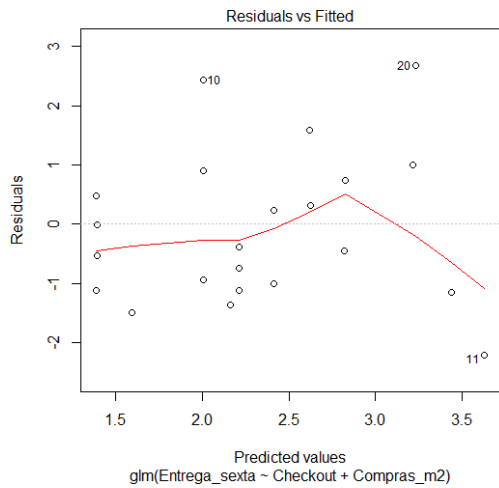


GLM 26 – Gráfico normal de probabilidade.



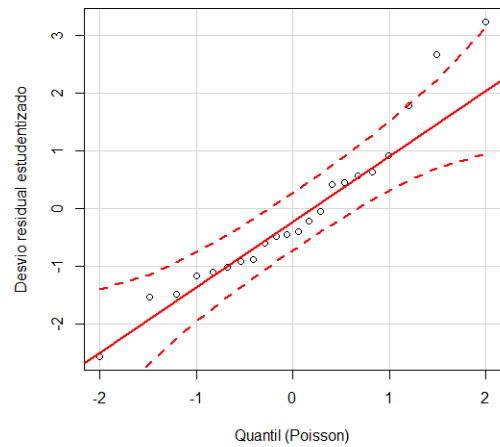
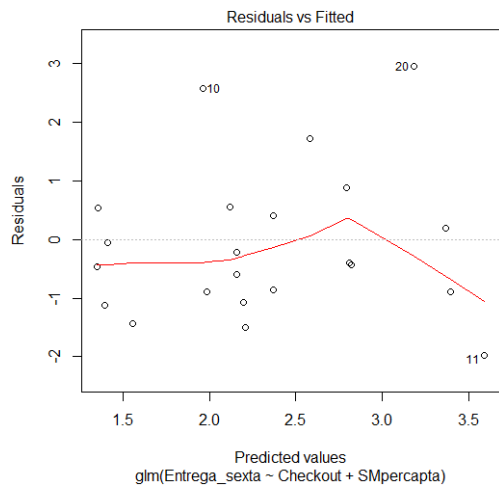
GLM 27 – Resíduos versus valores ajustados.

GLM 27 – Gráfico normal de probabilidade.



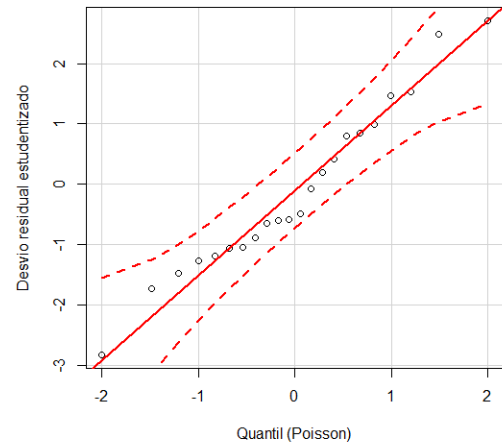
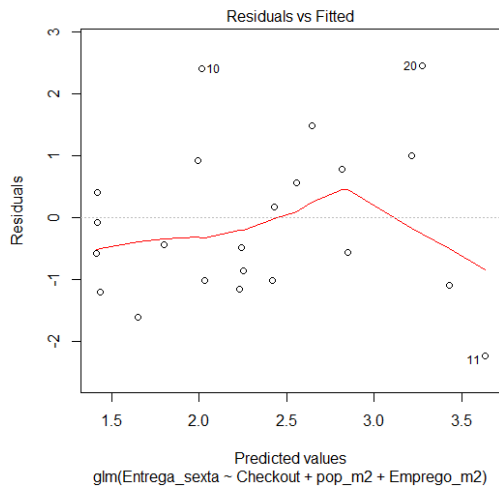
GLM 28 – Resíduos versus valores ajustados.

GLM 28 – Gráfico normal de probabilidade.



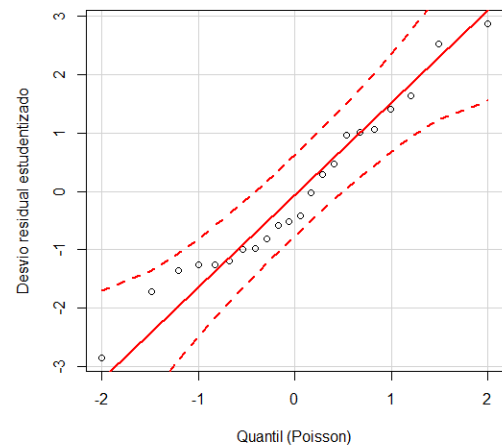
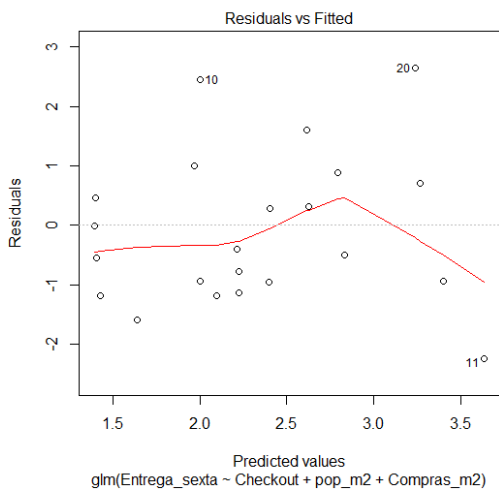
GLM 29 – Resíduos versus valores ajustados.

GLM 29 – Gráfico normal de probabilidade.



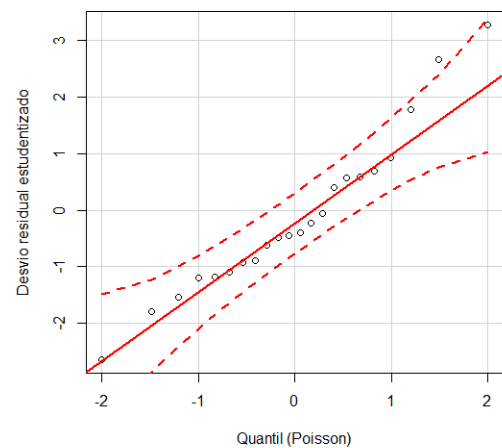
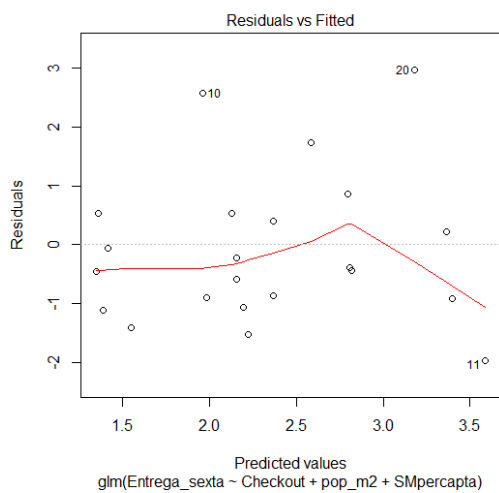
GLM 30 – Resíduos versus valores ajustados.

GLM 30 – Gráfico normal de probabilidade.



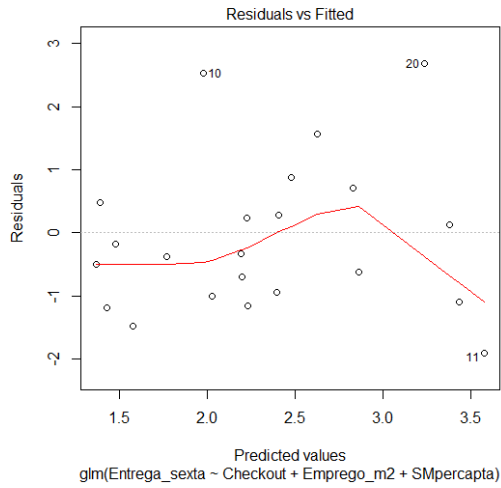
GLM 31 – Resíduos versus valores ajustados.

GLM 31 – Gráfico normal de probabilidade.

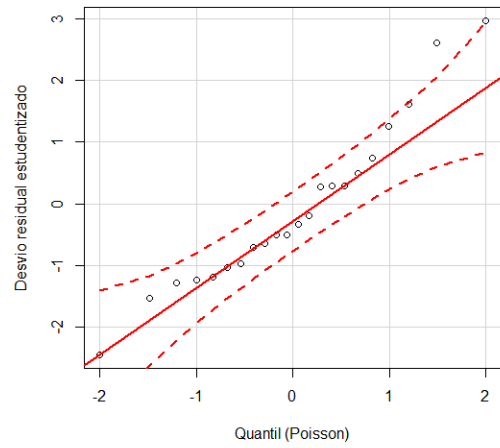


GLM 32 – Resíduos versus valores ajustados.

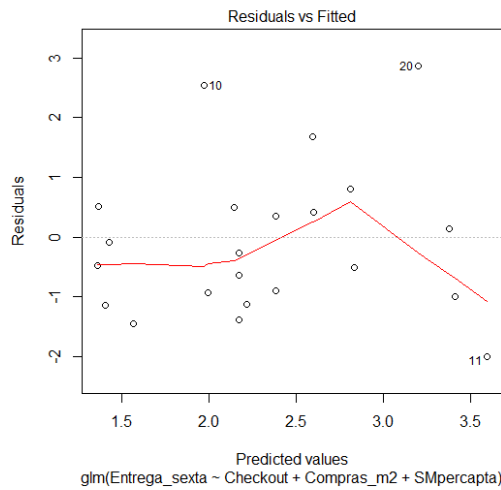
GLM 32 – Gráfico normal de probabilidade.



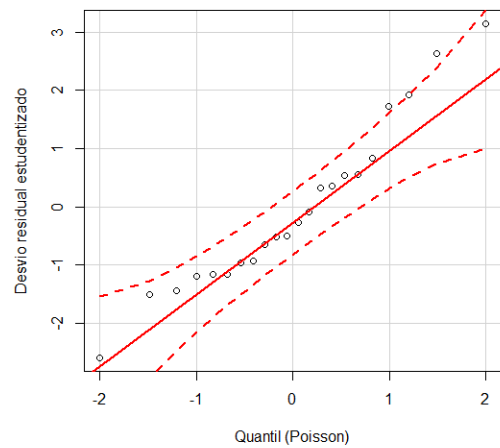
GLM 33 – Resúduos versus valores ajustados.



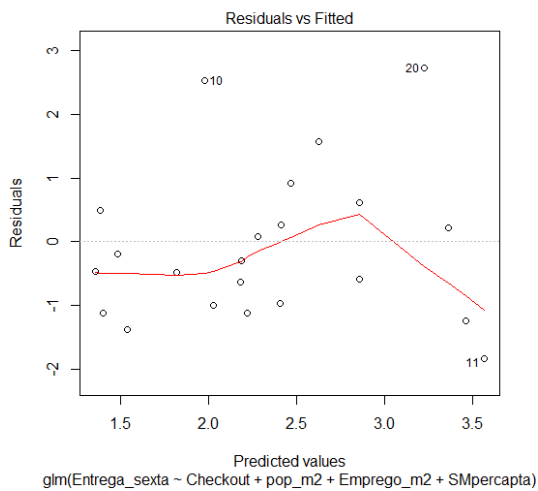
GLM 33 – Gráfico normal de probabilidade.



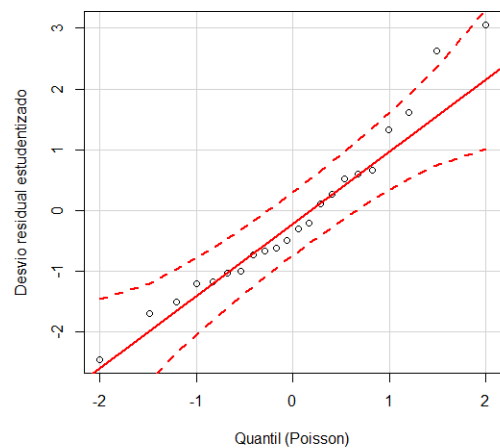
GLM 34 – Resúduos versus valores ajustados.



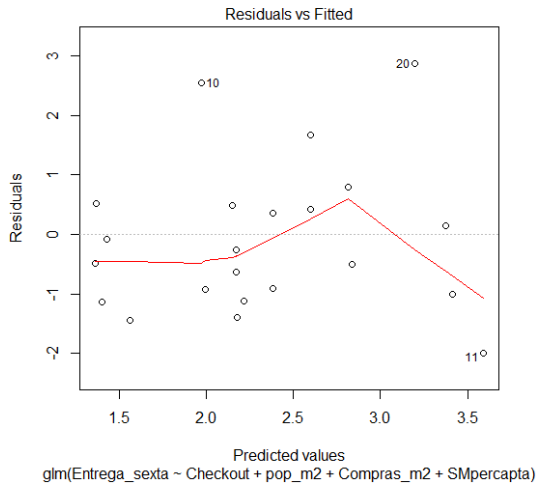
GLM 34 – Gráfico normal de probabilidade.



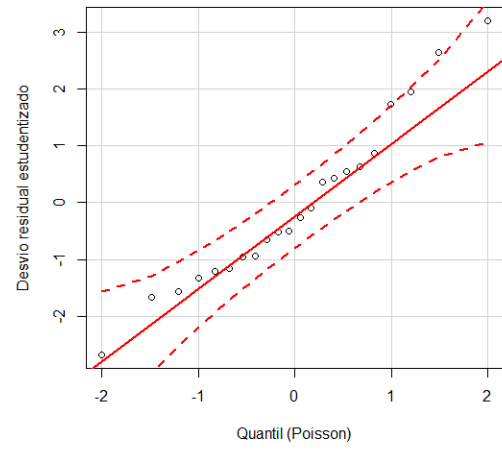
GLM 35 – Resúduos versus valores ajustados.



GLM 35 – Gráfico normal de probabilidade.



GLM 36 – Resíduos versus valores ajustados.



GLM 36 – Gráfico normal de probabilidade.