

Thaís Ventura Chibiaqui

**MÉTODO PARA ELABORAÇÃO DE MODELOS DE GERAÇÃO
DE VIAGENS PARA CAMINHÕES: APLICAÇÃO PARA
CONTEINERES NO PORTO DE ITAJAÍ (SC)**

Tese submetida ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Santa Catarina, para a obtenção do grau de Doutora em Engenharia Civil.

Orientadora: Prof^a. Dra. Lenise Grando Goldner

Florianópolis
2017

Ficha de identificação da obra elaborada pelo autor
através do Programa de Geração Automática da Biblioteca Universitária
da UFSC.

Chibiaqui, Thais Ventura

Método para elaboração de modelos de geração de viagens para caminhões: Aplicação para contêineres no porto de Itajaí (SC) / Thais Ventura Chibiaqui ; orientador, Lenise Grando Goldner, 2017.

228 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós Graduação em Engenharia Civil, Florianópolis, 2017.

Inclui referências.

1. Engenharia Civil. 2. 1. Portos marítimos 2. Viagens de caminhões 3. Modelos de geração de viagens. I. Goldner, Lenise Grando. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Civil. III. Título.

Thaís Ventura Chibiaqui

**MÉTODO PARA ELABORAÇÃO DE MODELOS DE GERAÇÃO
DE VIAGENS PARA CAMINHÕES: APLICAÇÃO PARA
CONTAINERES NO PORTO DE ITAJAÍ (SC)**

Esta Tese foi julgada adequada para obtenção do Título de “Doutor em Engenharia Civil” e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Santa Catarina – UFSC.

Florianópolis, 25 de agosto de 2017.

Prof. Glicério Trichês, Dr.

Coordenador do Curso

Prof. Lenise Grando Goldner, Dra.

Universidade Federal de Santa Catarina (Orientadora)

Banca examinadora:

Prof. Márcio de Almeida D’Agosto, Dr.

Universidade Federal do Rio de Janeiro (videoconferência)

Prof. Adelaida Pallavicini Fonseca, Dra.

Universidade Federal de Brasília

Prof. Liseane Padilha Thives, Dra.

Universidade Federal de Santa Catarina

Prof. João Victor Staub de Melo, Dr.

Universidade Federal de Santa Catarina

Prof. Amir Mattar Valente, Dr.

Universidade Federal de Santa Catarina

Dedico este trabalho à minha família e amigos como símbolo de todo o meu amor e amizade.

AGRADECIMENTOS

Aos meus pais, Nicolau Pedro Ventura e Rosana dos Santos Ventura, pelo exemplo de determinação, perseverança e dedicação, e, principalmente, por sempre me relembrem o significado da palavra “incondicional”.

Aos meus avós, Alfredo dos Santos e Dilza dos Santos, e ao meu irmão, Thiago dos Santos Ventura, por tudo que fizeram por mim até hoje e por compreender os momentos em que estive ausente ao longo dessa caminhada.

Ao meu marido Gustavo Chibiaqui Moreira, pela força dada, pelos abraços reconfortantes e pelas palavras de incentivo quando delas precisei.

À administração do porto de Itajaí (SC), por toda sua simpatia em me atender e por ceder informações essenciais para o desenvolvimento desse trabalho.

À Eng. Msc. Cintia Machado de Oliveira, pelo auxílio na revisão bibliográfica sistemática; ao Eng. Msc. Daniel Neves Schmitz Gonçalves, pelo auxílio na elaboração dos modelos de RNAs; e ao Eng. Msc. Gabriel Henrique de Almeida Pereira, pelo auxílio na elaboração dos modelos SVMs.

Um agradecimento especial: ao Prof. Dr. Marcio de Almeida D’Agosto, membro da banca de qualificação e de defesa, que me convidou para uma visita de uma semana ao Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia da Universidade Federal do Rio de Janeiro e me apresentou à Cintia e ao Daniel (supracitados), e a outras pessoas que agregaram conhecimento importante a esse trabalho; e ao Prof. Dr. Gueibi Peres Souza, membro da banca de qualificação, não só pelo auxílio na elaboração dos modelos de regressão, mas também pelo auxílio na revisão de toda a tese, com destaque para a parte geral de modelagem. Muito obrigada por todo o tempo de vocês e atenção!

A todos os membros da banca da defesa, Prof. Dr. Marcio de Almeida D’Agosto, Prof. Dr. Adelaida Pallavicini Fonseca, Prof. Dr. Liseane Padilha Thives, Prof. Dr. João Victor Staub de Melo e Prof. Dr. Amir Mattar Valente, e aos membros da banca de qualificação Prof. Dr. Gueibi Peres Souza e Prof. Dr. Marcos Aurélio Noronha, pelos comentários, dicas e sugestões que conferiram e irão conferir prestígio e valor a esse trabalho.

À minha orientadora, Prof. Dra. Lenise Grando Goldner, pela confiança na minha capacidade e pelo suporte prestado para que tudo ocorresse bem no transcorrer dessa produção.

Ao Programa de Pós-Graduação em Engenharia Civil da UFSC pela oportunidade concedida e aos funcionários do Departamento de Engenharia Civil da UFSC, pela disposição demonstrada sempre.

Aos colegas (e amigos) do Laboratório de Transportes e Logística, com destaque ao meu coordenador Eng. Msc. Rodolfo Carlos Nicolazzi Philippi, pelo apoio ao meu crescimento acadêmico.

A todos os meus amigos, especialmente aqueles que escutaram meus desabafos durante esse período.

A todos que contribuíram direta ou indiretamente para a realização deste trabalho.

Muito obrigada!

“É muito melhor lançar-se em busca de conquistas grandiosas, mesmo expondo-se ao fracasso, do que alinhar-se com os pobres de espírito, que nem gozam muito nem sofrem muito, porque vivem numa penumbra cinzenta, onde não conhecem nem vitória, nem derrota.”

(Theodore Roosevelt)

RESUMO

Apesar de gerar muitos benefícios, os portos marítimos também podem gerar impactos negativos na sua vizinhança como o congestionamento provocado pelo tráfego de caminhões. Este ocorre, principalmente, por causa do planejamento da rede e sistemas de gestão de tráfego inadequados; e, gera, além de incomodações para o entorno do porto, dificuldades para o desenvolvimento desse setor. Este trabalho tem como objetivo propor um método detalhando passo a passo o que deve ser feito para encontrar um modelo de geração de viagens para portos/ terminais marítimos e considerando as variáveis explicativas e as técnicas que podem ser utilizadas. Uma aplicação no porto de Itajaí, no Sul do Brasil, para elaboração de modelos de geração de viagens para caminhões foi realizada utilizando técnicas de regressão linear múltipla, regressão de Poisson, binomial negativa, Redes Neurais Artificiais (RNAs) e Máquinas de Vetores de Suporte (*Support Vector Machine* – SVM). Assim como foi elaborada uma comparação em termos de Erro Percentual Médio Absoluto (*Mean Absolute Percentage Error* – Mape) e da Mediana dos Erros Percentuais Absolutos (*Median Absolute Percentage Error* - Mdape) e análise crítica desses modelos. O método proposto foi validado com sucesso e, quanto aos modelos, a técnica SVM foi a que apresentou melhor resultado entre as técnicas utilizadas. A técnica de binomial negativa também se destacou em termos de Mape frente aos outros modelos, e é mais simples de aplicar e requer um tamanho de amostra menor que o SVM. Por ora, pode-se afirmar que o método proposto e o uso dessas técnicas, as quais não são comumente utilizadas para previsão de viagens, podem auxiliar na estimativa do fluxo de tráfego nas áreas portuárias de modo a contribuir para que se faça um planejamento da rede e gestão de tráfego.

Palavras-chave: Portos marítimos; Viagens de caminhões; Modelos de geração de viagens.

ABSTRACT

Despite of generating many benefits, ports generate highly negative impact to their neighborhood as traffic caused by truck. This occurs mostly due to the limited studies on traffic flow of the port area, consequently resulting on poor network planning and inadequate traffic management systems. Traffic port not only troubles port surroundings affects every logistic mode and function, delaying the development of the sector. The objective of this work is to propose a step-by-step method, which should be done to find a model of trip generation for ports/maritime terminals and considering the explanatory variables and the techniques that may be used. A case study in the port of Itajaí, southern Brazil, for the elaboration of models of generation of trips for trucks was performed using multiple linear regression, Poisson regression, negative binomial, Artificial Neural Networks (RNAs) and Support Vector Machine (SVM) techniques. The proposed method was successfully validated. Then, a comparison and critical analysis between these models was also performed. A comparison of their results was made using the Mean Absolute Percentage Error (MAPE) and the Median Absolute Percentage Error (MDAPE), in which the SVM technique presented the best result among the techniques used. The negative binomial technique also stood out in terms of Mape compared to other models, and is simpler to apply and requires a smaller sample size than the SVM. For the time being, it can be stated that the proposed method and the use of these techniques, which are not commonly used for trip forecasting, can help to estimate traffic flow in port areas in order to network and traffic management.

Keywords: Maritime Port; Truck Trip; Trip generation models.

LISTA DE FIGURAS

Figura 1.1 – Procedimento adotado no estudo	33
Figura 2.1 – Esquema dos procedimentos para análise do impacto viário	44
Figura 2.2 – Estrutura global de análise de impactos	48
Figura 2.3 – Processo do planejamento estratégico do transporte de carga	49
Figura 2.4 – Protocolo da revisão bibliográfica sistemática desse trabalho	53
Figura 2.5 – Distribuição dos trabalhos em relação à sua fonte de publicação.....	56
Figura 2.6 – Distribuição dos trabalhos em relação ao ano de publicação	57
Figura 2.7 – Distribuição dos trabalhos em relação aos países de origem	57
Figura 2.8 – Distribuição dos trabalhos em relação à abordagem temática	58
Figura 2.9 – Distribuição dos trabalhos em relação à abordagem metodológica	58
Figura 2.10 – Modelo de quatro etapas adaptado para carga: abordagem com base a) em viagens e b) em volume de carga.....	70
Figura 2.11 – Procedimento para elaboração de modelo de geração de viagens ou volume de carga	77
Figura 3.1 – Protocolo da revisão bibliográfica sistemática desse trabalho	84
Figura 3.2 – Distribuição dos trabalhos em relação à sua fonte de publicação.....	88
Figura 3.3 – Distribuição dos trabalhos em relação ao ano de publicação	88
Figura 3.4 – Distribuição dos trabalhos em relação aos países de origem	89
Figura 3.5 – Distribuição dos trabalhos em relação à abordagem temática	90
Figura 3.6 – Distribuição dos trabalhos em relação à abordagem metodológica	90
Figura 3.7 – Configuração do modelo de RNA para o porto de Miami	98
Figura 3.8 – Processo de modelagem de fretamento multimodal.....	115
Figura 4.1 – Fluxograma do método	146

Figura 5.1 – Mapa de localização da área objeto de estudo (porto de Itajaí/SC)	157
Figura 5.2 – TUPs do complexo portuário de Itajaí (SC)	160
Figura 5.3 – Rede de apoio logístico do complexo portuário de Itajaí (SC)	161
Figura 5.4 – Evolução da movimentação (em toneladas) no porto de Itajaí (SC)	163
Figura 5.5 – Evolução da movimentação (em toneladas) no complexo portuário de Itajaí (SC).....	164
Figura 5.6 – Zoneamento do porto de Itajaí (SC).....	165
Figura 5.7 – Infraestrutura de acostagem do porto de Itajaí (SC)	166
Figura 5.8 – Estruturas de armazenagem do porto de Itajaí (SC).....	167
Figura 5.9 – Canais de acesso e bacia de evolução do porto de Itajaí (SC)	168
Figura 5.10 – Acessos rodoviários do porto de Itajaí (SC)	169
Figura 5.11 – Acessos rodoviários do porto de Itajaí (SC)	171
Figura 5.12 – Trajeto de acesso ao portão principal do porto de Itajaí (SC)	171
Figura 5.13 – Acessos ao portão secundário do porto de Itajaí (SC) ..	171
Figura 5.14 – Gráfico com os melhores resultados de R^2 do subconjunto de validação segundo o número de neurônios na camada oculta para os modelos de exportação	193
Figura 5.15 – Gráfico com os melhores resultados de R^2 do subconjunto de validação segundo o número de neurônios na camada oculta para os modelos de importação.....	193
Figura 5.16 – Gráfico com os melhores resultados de R^2 do subconjunto de validação segundo o número de neurônios na camada oculta para os modelos de importação.....	200

LISTA DE TABELAS

Tabela 2.1 – Descrição dos procedimentos disponíveis	39
Tabela 2.2 – Resumo da revisão bibliográfica para estações metroviárias	63
Tabela 2.3 – Categorias e variáveis mais relevantes	64
Tabela 2.4 – Síntese dos modelos de geração de viagens selecionados para carga a) nacionais	72
Tabela 3.1 – Modelos de geração de viagens do ITE para portos marítimos	92
Tabela 3.2 – Descrição das variáveis testadas.....	121
Tabela 3.3 – Modelos para entrada diária de contêineres.....	122
Tabela 3.4 – Modelos para saída diária de contêineres	123
Tabela 3.5 – Fontes bibliográficas encontradas sobre modelos de geração de viagens para caminhões em portos	125
Tabela 5.1 – Resultado do modelo de exportação para o porto de Itajaí utilizando regressão linear múltipla.....	180
Tabela 5.2 – Resultados dos testes estatísticos do modelo de regressão linear múltipla de exportação para o porto de Itajaí	181
Tabela 5.3 – Resultado do modelo de importação para o porto de Itajaí utilizando regressão linear múltipla.....	183
Tabela 5.4 – Resultados dos testes estatísticos do modelo de regressão linear múltipla de importação para o porto de Itajaí.....	184
Tabela 5.5 – Resultado do modelo de exportação para o porto de Itajaí utilizando regressão de Poisson.....	187
Tabela 5.6 – Resultado do modelo de exportação para o porto de Itajaí utilizando distribuição binomial negativa.....	188
Tabela 5.7 – Resultado do modelo de importação para o porto de Itajaí utilizando regressão de Poisson.....	189
Tabela 5.8 – Resultado do modelo de importação para o porto de Itajaí utilizando distribuição binomial negativa.....	191
Tabela 5.9 – Resultado de Mape para os modelos desenvolvidos para o porto de Itajaí.....	197
Tabela 5.10 – Resultado de Mdape para os modelos desenvolvidos para o porto de Itajaí.....	197
Tabela 5.11 – Projeção da demanda de cargas no porto de Itajaí entre os anos 2013 (observado) e 2030 (projetado) em toneladas para cargas do tipo contêineres, navegação de longo curso e sentido de embarque	200

Tabela 5.12 – Volume de cargas estimado (segunda linha) e taxa de crescimento utilizada (terceira linha) para cada ano (primeira linha) entre o período de 2015 a 2020 201

LISTA DE ABREVIATURAS E SIGLAS

Anpet	Associação Nacional de Pesquisa e Ensino em Transportes
ARD	<i>Automatic Relevance Determination</i> (Determinação Automática da Relevância)
Arima	<i>Autoregressive Integrated Moving Average</i> (Auto Regressivo Integrado de Média Móvel)
BPN	<i>Backpropagation Neural Network</i> (Rede Neural de Retro-Propagação)
CET-SP	Companhia de Engenharia de Tráfego de São Paulo
CNT	Confederação Nacional de Transporte
Denatran	Departamento Nacional de Trânsito
DI	Despacho de Importação
DTC	Declaração de Trânsito de Contêiner
EUA	Estados Unidos da América
FCL	<i>Full Container Load</i> (Contêiner Totalmente Carregado)
FRNN	<i>Fully Recurrent Neural Network</i> (Rede Neural Totalmente Recorrente)
GP	<i>Gaussian Process</i> (Processo Gaussiano)
HCM	<i>Highway Capacity Manual</i> (Manual de Capacidade Rodoviária)
IDH	Índice de Desenvolvimento Humano
ITE	<i>Institute of Transportation Engineers</i> (Instituto de Engenheiros de Transporte)
LCL	<i>Less Container Load</i> (Contêiner Menos Carregado)
LM	<i>Lagrange Multiplier</i> (Multiplicador de Lagrange)
LMA	<i>Levenberg-Marquardt Algorithm</i> (Algoritmo de Levenberg–Marquardt)
MAD	<i>Mean Absolute Deviation</i> (Desvio Médio Absoluto)
Mape	<i>Mean Absolute Percentage Error</i> (Erro Percentual Médio Absoluto)
Mdape	<i>Median Absolute Percentage Error</i> (Mediana dos Erros Percentuais Absolutos)
MHC	

	<i>Mobile Harbor Crane</i> (Guindaste Móvel Portuário)
ML	<i>Maximum Likelihood</i> (Máxima Verossimilhança)
MLFNN	<i>Multilayer Feedforward Neural Network</i> (Rede Neural Acíclica Multicamada)
MLP	<i>Multilayer Perceptron</i> (Perceptron Multicamada)
MSD	<i>Mean Square Displacement</i> (Média do Desvio Quadrado)
MSE	<i>Mean of Squared Errors</i> (Média dos Quadrados dos Erros)
NB	<i>Negative Binomial</i> (Binomial Negativa)
Nittrans	Niterói Transporte e Trânsito
PAC	Programa de Aceleração do Crescimento
Panam	Pan-Americano de Engenharia de Transportes e Trânsito
PGM	Polo Gerador de Viagem de Uso Misto
PGV	Polo Gerador de Viagem
PIB	Produto Interno Bruto
Pluris	Planejamento Urbano, Regional, Integrado e Sustentável
PMGV	Polo Múltiplo Gerador de Viagem
Portobras	Empresa de Portos do Brasil
PPP	Parceria Público Privada
RAC	Recinto Alfandegado Contíguo
RBF	<i>Radial Basis Function</i> (Função Base Radial)
Reset	<i>Regression Specification Error Test</i> (Teste de Erro de Especificação da Regressão)
RFB	Receita Federal do Brasil
RMSE	<i>Root Mean Square Error</i> (Raiz da Média dos Erros Quadráticos)
RNA	Rede Neural Artificial
RS	Rio Grande do Sul
RTRL	<i>Real-Time Recurrent Learning</i> (Aprendizagem Recorrente em Tempo Real)
SC	Santa Catarina
SEP	Secretaria de Portos
SSE	<i>Sum of Squared Errors</i> (Soma dos Quadrados dos Erros)
SVM	<i>Support Vector Machine</i> (Máquina de Vetores de Suporte)
Tecon	Terminal de Contêineres

Teconvi	Terminal de Contêineres do Vale do Itajaí
TEU	<i>Twenty Feet Equivalent Unit</i> (Unidade Equivalente a Vinte Pés)
TRB	<i>Transportation Research Board</i> (Conselho de Pesquisa de Transporte)
TUP	Terminais de Uso Privado
VHP	Volume Hora de Pico
VMDH	Volume Médio Diário Horários
VTMS	<i>Vessel Tracking and Monitoring System</i> (Sistema de Acompanhamento e Monitoramento de Embarcações)

SUMÁRIO

1 INTRODUÇÃO	27
1.1 OBJETIVOS	30
1.1.1 Objetivo geral	30
1.1.2 Objetivos específicos	30
1.2 JUSTIFICATIVA E RELEVÂNCIA DA PESQUISA	30
1.3 PROCEDIMENTOS PARA ALCANÇAR O OBJETIVO.....	31
1.3.1 Revisão bibliográfica sistemática.....	32
1.4 OBJETO DE ESTUDO.....	34
1.5 ESTRUTURA DA TESE.....	35
2 MODELOS DE GERAÇÃO DE VIAGENS PARA VEÍCULOS EM POLOS GERADORES DE VIAGENS	37
2.1 ESTUDOS DE IMPACTO	38
2.2 GERAÇÃO DE VIAGENS	49
2.3 PLANEJAMENTO DA REVISÃO BIBLIOGRÁFICA	52
2.3.1 Identificação da necessidade da revisão	52
2.3.2 Elaboração da proposta da revisão	52
2.3.3 Desenvolvimento do protocolo de revisão	53
2.4 REALIZAÇÃO DA REVISÃO BIBLIOGRÁFICA	54
2.4.1 Identificação e seleção dos trabalhos.....	55
2.4.2 Avaliação dos trabalhos selecionados.....	55
2.4.3 Extrair dados e informações	55
2.4.4 Síntese dos estudos	59
2.5 COMUNICAÇÃO E DIVULGAÇÃO DOS RESULTADOS DA REVISÃO.....	59
2.6 DESCRIÇÃO DOS ESTUDOS ENCONTRADOS	59
2.6.1 Shopping centers	59
2.6.2 Equipamentos institucionais	60
2.6.3 Estabelecimentos residenciais	61
2.6.4 Hotéis.....	62
2.6.5 Estações metroferroviárias.....	62
2.6.6 Estabelecimentos de ensino	65
2.6.7 Hospitais e estabelecimentos de saúde.....	66
2.6.8 Hipermercados e supermercados.....	67
2.6.9 Polos de uso misto e polos de uso múltiplo.....	68
2.6.10 Polos geradores de viagens de cargas	69

2.7	ANÁLISE DA REVISÃO BIBLIOGRÁFICA.....	79
3	MODELOS DE GERAÇÃO DE VIAGENS PARA CAMINHÕES EM PORTOS MARÍTIMOS	82
3.1	PLANEJAMENTO DA REVISÃO BIBLIOGRÁFICA	83
3.1.1	Identificação da necessidade da revisão	83
3.1.2	Elaboração da proposta da revisão	83
3.1.3	Desenvolvimento do protocolo de revisão	83
3.2	REALIZAÇÃO DA REVISÃO BIBLIOGRÁFICA	85
3.2.1	Identificação e seleção dos trabalhos.....	86
3.2.2	Avaliação dos trabalhos selecionados.....	87
3.2.3	Extrair dados e informações	87
3.2.4	Síntese dos estudos	91
3.3	COMUNICAÇÃO E DIVULGAÇÃO DOS RESULTADOS DA REVISÃO.....	91
3.4	DESCRIÇÃO DOS ESTUDOS ENCONTRADOS	91
3.4.1	<i>Estudos desenvolvidos para todo os EUA.....</i>	<i>92</i>
3.4.2	<i>Estudos desenvolvidos para Baltimore (EUA)</i>	<i>93</i>
3.4.3	<i>Estudos desenvolvidos para a Florida (EUA)</i>	<i>93</i>
3.4.4	<i>Estudos desenvolvidos para o Texas (EUA)</i>	<i>105</i>
3.4.5	<i>Estudos desenvolvidos para toda a Singapura</i>	<i>112</i>
3.4.6	<i>Estudos desenvolvidos para Taiwan (República da China)</i>	<i>112</i>
3.4.7	<i>Estudos desenvolvidos para Montevideú (Uruguai)</i>	<i>119</i>
3.5	ANÁLISE DA REVISÃO BIBLIOGRÁFICA.....	124
3.5.1	Regressão linear múltipla	133
3.5.2	Regressão de Poisson e distribuição binomial negativa	134
3.5.3	Redes neurais artificiais.....	137
3.5.4	Aprendizado de máquina baseado em <i>kernel</i>	139
3.5.4.1	<i>Funções kernel</i>	141
4	MÉTODO DA PESQUISA	144
4.1	ETAPA 1.....	148
4.2	ETAPA 2.....	148
4.3	ETAPA 3.....	150
4.4	ETAPA 4.....	152
4.5	ETAPA 5.....	153
4.6	ETAPA 6.....	154
4.7	ETAPA 7.....	154

5	APLICAÇÃO DO MÉTODO.....	157
5.1.1	Limites do estudo	157
5.2	ETAPA 1.....	158
5.3	ETAPA 2.....	158
5.3.1	Caracterização do PGV e sua operação por meio de fontes secundárias.....	159
5.3.1.1	<i>Histórico</i>.....	161
5.3.1.2	<i>Dados gerais</i>.....	165
5.3.2	Caracterização do PGV e sua operação por meio de entrevista com funcionários do porto	173
5.3.3	Definição das variáveis a serem testadas	175
5.4	ETAPA 3.....	176
5.5	ETAPA 4.....	178
5.6	ETAPA 5.....	186
5.6.1	Regressão de Poisson e distribuição binomial negativa	186
5.6.2	Redes neurais artificiais.....	192
5.6.3	<i>Support vector machine</i>	194
5.7	ETAPA 6.....	196
5.8	ETAPA 7.....	199
6	CONCLUSÕES E RECOMENDAÇÕES	203
6.1	CONCLUSÕES	203
6.2	RECOMENDAÇÕES E SUGESTÕES PARA NOVAS PESQUISAS	207
	APÊNDICE.....	221
	APÊNDICE A - ALGORITMO SVM UTILIZADO NA ELABORAÇÃO DOS MODELOS.....	221

1 INTRODUÇÃO

Por ser responsável pelo fluxo de cargas, o transporte desempenha papel fundamental no desenvolvimento econômico de uma localidade (CRISTIAN, 2012). Nesse contexto, o transporte marítimo se distingue por cobrir a maior parte das quantidades de bens trocadas entre os diferentes países, respondendo por mais de 80% do comércio mundial de mercadorias (UNCTAD, 2012).

Definitivamente, quando bem localizados e administrados, os portos geram benefícios econômicos. Entre outros benefícios, os portos reduzem os custos do comércio, geram valor agregado e emprego, e fomentam inovação tecnológica no setor (MERK, 2013).

Entretanto, os portos marítimos produzem um contingente significativo de viagens por caminhão que necessitam de grandes espaços para estacionamento, carga e descarga, embarque e desembarque dos mesmos, promovendo, conseqüentemente, potenciais impactos no seu entorno e se caracterizando como um Polo Gerador de Viagens (PGV) (REDEPGV, 2012).

Logo, é preciso analisar os problemas que podem ser gerados pelo porto marítimo no seu entorno, principalmente se o mesmo se localizar em área urbana. De acordo com Facchini (2006), a movimentação de carga e descarga, pode agravar os problemas de tráfego já existentes na cidade, como elevados níveis de congestionamento e problemas econômicos, sociais e ambientais, impactando nas transportadoras e no cidadão.

Entre os impactos negativos gerados pelos portos marítimos, o congestionamento foi eleito uma das questões mais críticas pelo Comitê Executivo do Conselho de Pesquisa de Transporte (*Transportation Research Board – TRB*) (2006).

O tráfego intenso nas áreas portuárias afeta todos os modos e funções logísticas dos portos (FAN, WILSON & DAHL, 2012). Como prova disso, por motivos de congestionamento tanto nos acessos rodoviários aos portos quanto nos seus canais de acesso, observou-se que menos da metade de todos os navios porta-contêineres atracam no porto dentro do cronograma, causando, conseqüentemente, aumento nos custos de transporte (BLOOMBERG, 2011).

Além disso, o fluxo de tráfego de caminhões gerado pelo porto utiliza as vias da sua área de entorno para acessar o mesmo, incrementando o fluxo de tráfego já existente naquelas vias, e podendo

acarretar, então, em um aumento: no tempo de viagem para os outros usuários daquelas vias, na deterioração dos pavimentos, na taxa de acidentes, e na poluição atmosférica e sonora na região (AL-DEEK, 2001a; PAROLA & MAUGERI, 2013).

No Brasil, a partir de dados levantados pela Secretaria de Portos, o Laboratório de Transportes e Logística (Labtrans, 2015) analisou o desempenho dos acessos rodoviários a 27 portos brasileiros, calculando suas capacidades e comparando-as com o fluxo de veículos que circulavam nesses acessos.

De acordo com essa análise, em 2013, os acessos rodoviários de seis portos (dos 27 analisados) se encontravam com suas capacidades esgotadas, e dois portos se encontravam com uma alta utilização dos seus acessos rodoviários (de 70 a 100% de utilização da sua capacidade). Sendo que a previsão, segundo o Labtrans (2015), é que muitos dos acessos rodoviários que ainda não tinham esgotado sua capacidade, a esgotarão nos próximos anos, principalmente nos horários de pico, caso não haja maiores investimentos em infraestrutura viária.

Isso pode ser explicado pelo aumento no número de mercadorias transportadas pelos portos nos últimos anos, o tem ocasionando, como consequência, em um aumento no número de viagens de caminhões gerados por esses portos (FAN, WILSON & DAHL, 2012).

Além disso, o fato de os portos estarem geralmente inseridos em áreas urbanas (no Brasil, por exemplo, mais de 80% dos portos se situam em área urbana¹) contribui ainda mais para essa problemática, principalmente nos países latino-americanos. Isso porque os mesmos sofreram processos de urbanização intensos, acelerados, e concentrados no espaço, particularmente em algumas poucas cidades (como exemplo, as nove principais metrópoles brasileiras compreendem 30% da população do País) (IBGE, 2010).

Nesse âmbito, de acordo com Portugal et al. (2010), a infraestrutura de transportes se insere no contexto de escassez e uso inadequado dos recursos públicos. Como exemplo, segundo o CNT (2013), o transporte rodoviário é responsável por 61,1% da matriz de

¹ Considerou-se como situa-se em área urbana, os portos que se situam em municípios classificados desde grande metrópoles nacionais a centros locais pelo estudo Regiões de Influências das Cidades (IBGE, 2007).

transportes brasileira; o ferroviário, por 20,7%; o aquaviário, por 13,6%; o dutoviário, por 4,2%; e o aéreo, por 0,4%.

No Brasil, a infraestrutura de transporte ferroviário e transporte por cabotagem ligando os centros de produção de mercadorias aos terminais portuários é pequena ou, até mesmo, inexistente e a maioria dos acessos rodoviários aos portos já se encontram com alta utilização ou capacidade esgotada.

Como exemplo, cerca de 98% das mercadorias de exportação que chegam ao porto de Santos, o principal porto do País em termos de movimentação, são transportados até ele por meio de caminhões, gerando congestionamento e conflito na relação porto-cidade (MT, 2014).

No Brasil, o crescimento acelerado da motorização também é outro fator agravante, uma vez acarreta em aumento no número de veículos e muitas vezes, como consequência, em alta utilização das vias, principalmente as urbanas (CAF, 2010; PORTUGAL et al., 2010). Ou seja, boa parte das vias urbanas já se encontram com um alto volume de tráfego que somado ao volume de caminhões gerados pelos portos brasileiros em áreas urbanas, resulta em esgotamento de capacidade dos acessos rodoviários aos portos (MT, 2015).

Yang, Chen & Moodie (2010) citam que o problema de tráfego em áreas portuárias vem da interação das seguintes causas: um grande número de caminhões que entram e saem do porto; infraestrutura de transportes, incluindo a rodoviária, insuficiente no porto; e pouco conhecimento do fluxo de tráfego da área portuária, o que resulta, conseqüentemente, em planejamento da rede e sistemas de gestão de tráfego inadequados.

Portanto, para melhor definir soluções para o congestionamento, segundo Yang, Chen & Moodie (2010), é necessário conhecer o fluxo de tráfego da área portuária para que, assim, se faça um planejamento da rede e gestão de tráfego visando reduzir ou acomodar esse número de caminhões adequadamente, seja por meio da implantação de uma nova conexão ferroviária, da criação de plataformas logísticas no entorno ou de outras ações.

Muitas vezes, para se conhecer o fluxo de tráfego nas áreas portuárias é preciso estimar o volume de viagens geradas pelo porto para soma-lo ao fluxo de veículos já existentes naquelas vias. Mas como estimar esse volume de viagens para um porto marítimo e qual seria a(s) técnica(s) e as variáveis mais recomendada(s) para desenvolver um modelo de geração de viagens para esse tipo de PGV?

1.1 OBJETIVOS

1.1.1 Objetivo geral

Este trabalho tem como objetivo elaborar um método para construção de modelos de geração de viagens para caminhões de em portos/ terminais marítimos.

1.1.2 Objetivos específicos

- 1) identificar, por meio de revisão bibliográfica, os modelos existentes de geração de viagens em PGVs, de um modo geral, e para portos marítimos, visando verificar as técnicas e as variáveis explicativas utilizadas na modelagem de geração de viagens para os mesmos, verificando quais são as novas abordagens e quais são as vantagens e desvantagens de cada uma por meio de uma análise crítica;
- 2) a aplicação de diferentes técnicas para modelar a geração de viagens de caminhões para o porto de Itajaí (SC), de modo a compará-las entre si e identificar a técnica utilizada que estatisticamente melhor explicou o fenômeno de geração de viagens;
- 3) comparar entre si, os modelos elaborados com diferentes variáveis, de modo a analisar as variáveis independentes utilizadas para a construção dos modelos.

1.2 JUSTIFICATIVA E RELEVÂNCIA DA PESQUISA

Melhorias nas estradas podem ser implantadas por meio da compreensão do comportamento de movimentação de cargas e por meio da identificação dos fatores que afetam os modelos de geração de viagem e de divisão modal para esse tipo de movimento (AL-DEEK, 2001a).

No entanto, apesar dessa necessidade, não há um método consolidado detalhando passo a passo o que deve ser feito para encontrar um modelo parcimonioso de geração de viagens para portos/ terminais

marítimos considerando as variáveis explicativas e as técnicas que podem vir a ser utilizadas.

Essa última afirmação tem como embasamento uma revisão sistemática no endereço eletrônico online Science Direct, Scopus, Periódicos Capes e Trid Online por meio das palavras chaves “porto”, “caminhões”, e “modelos de geração de viagens”, e suas combinações por meio dos operadores booleanos “e”, conforme mostrado no Capítulo 3.

Diante desse cenário, entendeu-se como relevante um método detalhado para construção de modelos de geração de viagens para caminhões em portos/ terminais marítimos, considerando o exposto no parágrafo anterior.

Como para validar o método, este foi aplicado a um porto brasileiro. Isso possibilitou que fossem apresentados os critérios utilizados para a montagem dos modelos com diferentes técnicas, tais como regressão linear múltipla, regressão de Poisson e distribuição binomial negativa, redes neurais artificiais e aprendizado de máquina baseado em *kernel* (SVM).

Assim como, realizar uma comparação entre os resultados obtidos com as diferentes técnicas utilizadas, pois não foi constatado, pela revisão bibliográfica do Capítulo 3, nenhum estudo que comparasse, por exemplo, os resultados de uma regressão de Poisson e distribuição binomial negativa com redes neurais artificiais ou aprendizado de máquina baseado em *kernel*, de modo a identificar a técnica utilizada que estatisticamente melhor explique o fenômeno de geração de viagens de caminhões para o porto analisado.

Com isso, pretende-se contribuir para que as agências de transporte e as autoridades portuárias possuam ainda mais orientações sobre os métodos, as técnicas e as variáveis a serem utilizadas para estimar viagens geradas por portos marítimos, de modo que possam melhor planejar a infraestrutura para acomodar o tráfego de caminhões de um novo terminal portuário ou até mesmo de alguma expansão de um porto existente.

1.3 PROCEDIMENTOS PARA ALCANÇAR O OBJETIVO

Para alcançar o objetivo desse trabalho foi desenvolvida, como procedimento técnico, uma pesquisa bibliográfica sistemática no Capítulo 2 e 3, conforme etapas mostradas no subitem 1.3.1 a seguir, que se

estendeu por todo o período do trabalho, buscando-se assuntos relacionados, respectivamente, a modelos de geração de viagem para PGVs e, especificamente, para portos marítimos.

Os conhecimentos adquiridos por meio da revisão foram, então, utilizados na construção do método e o mesmo foi aplicado em uma área de estudo real para verificação de sua validade.

A aplicação do método permitiu testar diferentes técnicas para modelar a geração de viagens de caminhões para o porto de Itajaí (SC), de modo a compará-las entre si e identificar a técnica utilizada que estatisticamente melhor explicou o fenômeno de geração de viagens.

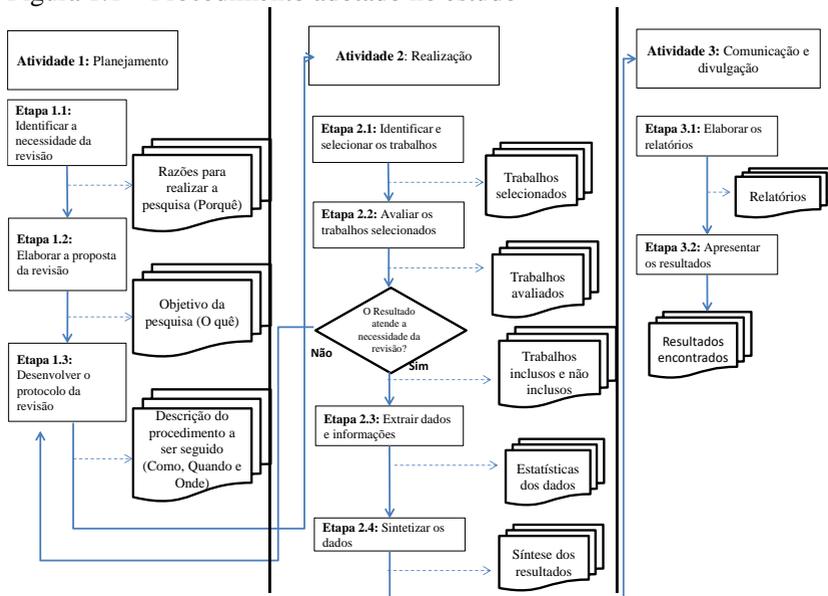
1.3.1 Revisão bibliográfica sistemática

As revisões bibliográficas sistemáticas são levantamentos de informações científicas por meio de métodos planejados para responder uma pergunta específica e estreita (COOK, MULROW & HAYNES, 1997).

Logo, a principal diferença entre esse tipo de revisão e a revisão bibliográfica narrativa é que a primeira segue um processo científico rigoroso (OLIVEIRA, 2015; BENCHEIKH, LANDRY & AMARA, 2006). Isso acarreta no aumento da confiabilidade dos resultados em relação à revisão narrativa (OLIVEIRA, 2015; BERETON et al., 2007; COOK, MULROW & HAYNES, 1997), bem como na possibilidade dessa revisão ser reproduzida por qualquer outro pesquisador (COOK, MULROW & HAYNES, 1997).

O procedimento adotado nesse estudo para a realização da revisão bibliográfica sistemática é apresentado na Figura 1.1

Figura 1.1 – Procedimento adotado no estudo



Fonte: Oliveira, 2015.

Ao se observar a Figura 1.1, percebe-se que o procedimento pode ser resumido em três atividades: o planejamento, a realização e a comunicação e divulgação (OLIVEIRA, 2015; BERETON et al., 2007; TRANFIELD, DENYER & SMART, 2003).

O planejamento da revisão é formado pela identificação da necessidade, elaboração da proposta e desenvolvimento do protocolo de revisão. Enquanto a realização da revisão em si consiste na identificação e seleção dos trabalhos, avaliação dos trabalhos selecionados, extração dos dados e informações e sintetização dos dados. Por fim, a comunicação e divulgação da pesquisa são compostas pelas etapas de elaboração dos relatórios e apresentação dos resultados (OLIVEIRA, 2015; BERETON et al., 2007; TRANFIELD, DENYER & SMART, 2003).

O desenvolvimento do protocolo de revisão pôde ser realizado por meio das seguintes etapas:

- definição da(s) base(s) de dados a ser(e) utilizada(s) no levantamento dos dados;

- b) definição das palavras chaves a serem usadas nesse levantamento (e suas combinações);
- c) definição de quais tipos de estudos serão considerados e em qual parte deles será feita a busca;
- d) definição da abrangência geográfica e temporal dessa busca;
- e) definição dos critérios para seleção dos artigos encontrados na pesquisa; e
- f) após selecionados os artigos, definição dos critérios de inclusão e exclusão deles no trabalho.

1.4 OBJETO DE ESTUDO

Para o desenvolvimento dos modelos de geração de viagens para caminhões foi considerado o porto de Itajaí, visto que o mesmo faz parte junto com o porto de Navegantes do complexo portuário de Itajaí, responsável pela segunda maior movimentação de contêineres do Brasil e pela primeira de Santa Catarina (SC).

O complexo portuário do Itajaí se localiza no litoral norte do Estado de Santa Catarina, na região sul do Brasil. É constituído pelo porto público de Itajaí e demais terminais portuários instalados nas margens direita e esquerda do Rio Itajaí, além de instalações de apoio logístico em operações nas cidades de Itajaí e Navegantes (PORTOITAJAÍ, 2014).

A estrutura do porto de Itajaí, situada na margem direita do Rio Itajaí, é composta por quatro berços de atracação, sendo dois arrendados e dois públicos, além de píer turístico de passageiros (PORTOITAJAÍ, 2014).

Em 2016, o complexo foi responsável pela movimentação de 12.049.551 toneladas de mercadorias (LABTRANS, 2015).

O complexo movimenta mercadorias provenientes e destinadas a 20 Estados nas cinco Regiões do País e ao Distrito Federal, sendo as cargas predominantes: produtos congelados, produtos cerâmicos, madeira e derivados, máquinas, motores e acessórios, açúcar, maçãs e tabacos (BERGER & BERGER, 2009).

Mais detalhes sobre o complexo e sobre o porto em si podem ser vistos no item 5.3, na página 158.

1.5 ESTRUTURA DA TESE

O trabalho foi estruturado em seis capítulos. O Capítulo 1, Introdução, situa o leitor quanto ao tema estudado, circunscreve a relevância do trabalho, fixa o objetivo geral do estudo e o modo de alcançá-lo por meio dos objetivos específicos delimitados.

Os Capítulos 2 e 3, Revisão Bibliográfica, objetivam por meio de revisão bibliográfica revisar a literatura, identificar os métodos empregados para estimar as viagens geradas pelos PGVs de um modo geral, isto no Capítulo 2, e, especificamente, pelos portos marítimos, no Capítulo 3, assim como as técnicas e as variáveis utilizadas nas construções dos modelos.

No Capítulo 4, Método da Pesquisa, é apresentado o método proposto para atingir os objetivos propostos pelo trabalho, estruturando todo a aplicação a ser desenvolvida; e o Capítulo 5 trata da aplicação do método e dos resultados encontrados com esse trabalho.

Por fim, o Capítulo 6, Conclusões, corresponde às conclusões alcançadas com esse estudo.

No final desse volume ainda se encontram as referências utilizadas e os apêndices do trabalho.

2 MODELOS DE GERAÇÃO DE VIAGENS PARA VEÍCULOS EM POLOS GERADORES DE VIAGENS

Os primeiros estudos realizados de modo mais sistemático e com respaldo técnico com foco em PGVs foram iniciados na década de 1970 pelo Instituto de Engenheiros de Transporte (*Institute of Transportation Engineers – ITE*), dos Estados Unidos da América (EUA). O ITE iniciou um programa para coletar as características da geração de viagens para variados usos do solo durante décadas. A acumulação desse banco de dados resultou na publicação *Trip Generation*, cuja versão mais atualizada é de 2012 (PORTUGAL, 2012).

No Brasil, as primeiras menções sobre o assunto datam da década de 1970, no qual esses grandes empreendimentos foram chamados de polos geradores de tráfego (Companhia de Engenharia de Tráfego de São Paulo – CET-SP, 1983; Departamento Nacional de Trânsito – DENATRAN, 2001), visto que os impactos mais diretos e visíveis na época estavam associados à circulação e ao estacionamento dos veículos, em particular dos automóveis (PORTUGAL, 2012).

Desde então, os conceitos iniciais de polos geradores de tráfego evoluíram e passaram a considerar que, além do tráfego, eles modificam o uso e ocupação do solo, a dinâmica de circulação e atraem e geram viagens de diferentes modos – como as não motorizadas, por transporte público e as de carga (KNEIB, 2004).

Assim, o termo polos geradores de viagens passou a substituir o termo polos geradores de tráfego, visto que procura contemplar os potenciais impactos no seu sentido mais amplo, ou seja, considerando impactos nos sistemas viários e de transportes (congestionamentos, acidentes e repercussões no ambiente), na estrutura urbana, bem como no desenvolvimento socioeconômico e na qualidade de vida da população (REDEPGV, 2012). Deste modo, ficou definido como PGVs:

Locais ou instalações de distintas naturezas que têm em comum o desenvolvimento de atividades em porte e escala capazes de exercer grandes atratividades sobre a população, produzir um contingente significativo de viagens, necessitarem de grandes espaços para estacionamento, carga e descarga, embarque e desembarque, promovendo consequentemente, potenciais impactos (REDEPGV, 2012, s/p.).

Ressalta-se que, de acordo com Portugal (2012), está em fase de desenvolvimento e consolidação, entre os pesquisadores brasileiros, o conceito relacionado a polos geradores de viagens sustentáveis e o de polos geradores de desenvolvimento e qualidade de vida. A ideia, segundo Santos (2011), é de incorporar uma definição que demonstre preocupação com a sustentabilidade e qualidade de vida dentro do seu espaço privado e nos espaços urbanos que os integram, por meio de promoção da vitalidade das ruas e dos espaços de circulação, estimulando as viagens a pé, de bicicleta, e por transporte público.

São exemplos de PGVs: shoppings centers, equipamentos institucionais, estabelecimentos residenciais, hotéis, estações metroferroviárias, estabelecimentos de ensino, hospitais, supermercados, entre outros como os portos marítimos, objetos desse trabalho.

Independentemente do tipo de PGV ou seu porte, os impactos precisam ser previstos e tratados. Uma das etapas fundamentais dos estudos de impacto é a de estimativa das viagens geradas.

2.1 ESTUDOS DE IMPACTO

O PGV se caracteriza por produzir um contingente significativo de viagens até ele ou a partir dele. Os impactos sentidos no seu entorno são função desse volume de deslocamentos (BASTOS, 2004).

Nesse contexto, é necessária a realização de estudos de impactos adequados a serem submetidos e apreciados num processo mais abrangente de licenciamento. Além da identificação dos impactos, devem-se identificar as medidas a serem utilizadas para maximização dos impactos positivos e eliminação ou mitigação dos impactos negativos, propiciando sua viabilidade financeira e garantindo seu compromisso com o interesse social (PORTUGAL, 2012).

Apesar disso, a maioria das cidades não utiliza um método específico sistematizado para exercer o controle efetivo sobre a implantação de um PGV, tratando essa questão com base apenas em restrições de lei de uso do solo ou lei de zoneamento (SOARES, 1990).

Na bibliografia disponível encontram-se diversas propostas de procedimentos para a realização dos estudos de impactos em PGVs, as quais 14 destas foram resumidas, por Portugal (2012), segundo o tipo de PGV, o tipo de impacto e as etapas de análise, e apresentada na Tabela 2.1.

Tabela 2.1 – Descrição dos procedimentos disponíveis

(continua)

Entidade/ Autor (Ano)	PGV abordado	Impacto considerado	Etapas adotadas
CET-SP (1983)	Diversos	Sistema viário	Modelos de geração de viagens; viagens na hora-pico; divisão modal; área de influência; sugere a análise de impacto avaliada em três níveis: impacto na área do entorno, impacto nas vias de acesso e impacto na área de influência.
Cox Consultores (1984)	Shopping centers	Sistema viário	Delimitação da área de influência; geração e distribuição de viagens; alocação de viagens; avaliação e índice de desempenho do sistema viário; dimensionamento do estacionamento.
<i>United States Department of Transportation/ ITE</i> (1985)	Diversos	Sistema viário	Estabelece 7 fases de estudo: definição da área de estudo; estimativa futura sem o PGV do pico horário do tráfego; estimativa do pico horário a partir do funcionamento do PGV; estabelece horizontes de estudo; taxas de geração, distribuição e alocação de viagens; e negociação entre órgãos locais e planejadores e implementação de melhorias.
Grando (1986)	Shopping centers	Sistema viário	Caracterização do problema; delimitação da área de influência; análise do sistema viário e de transportes; geração de viagens; demanda; oferta; escolha modal; distribuição de viagens; projeção da situação atual em diferentes horizontes; análise de desempenho; soluções alternativas e tomada de decisão.

Tabela 2.1 – Descrição dos procedimentos disponíveis

(continuação)

Entidade/ Autor (Ano)	PGV abordado	Impacto considerado	Etapas adotadas
Goldner & Portugal (1995)	Shopping centers	Sistema viário e de transportes	Delimitação da área de influência; geração de viagens; distribuição de viagens; alocação do tráfego; oferta viária; escolha modal; definição de cenários de análise; soluções alternativas, análise de desempenho, avaliação e tomada de decisões.
Espanhola (década de 1990)	Shopping centers	Sistema viário	Avaliação da situação atual; geração de viagens; demanda de veículos; distribuição de viagens geradas; estabelecimento da área de influência; análise dos fluxos atuais, taxas de crescimento do fluxo veicular; proposições.
Cybis et al. (1999)	Diversos	Sistema viário	Caracterização da área de influência; geração de viagens; distribuição das viagens geradas; prognóstico do crescimento do tráfego; alocação de viagens.
Menezes (2000)	Diversos	Ambiental e sistema viário	Localização e caracterização do PGT; determinação da área de influência; do fluxo de veículos; da capacidade e níveis de serviço da rede viária; determinação e verificação dos níveis ambientais e de serviço para a situação anterior e pós-implantação do PGT; determinação do número de viagens geradas; distribuição e alocação das viagens geradas pelo PGT; níveis de serviço das vias; determinação dos níveis ambientais da rede após o PGT.

Tabela 2.1 – Continuação da descrição dos procedimentos disponíveis

(continuação)

Entidade/ Autor (Ano)	PGV abordado	Impacto considerado	Etapas adotadas
Sistema de Evaluación de Impactos sobre el Sistema de Transporte Urbanos/ Chile (2012)	Diversos	Sistema viário e de transportes	Estudo da estimativa de demanda; análise da oferta viária; estabelecimento de quatro tipos de procedimentos em função do porte do empreendimento, estudo tático complexo e estudo estratégico. Envolvem estabelecimento da área de influência, caracterização da situação atual, estimativa da demanda de transporte, definição da oferta viária, modelagem e simulação, proposição de medidas de mitigação, avaliação de alternativas, esquema físico e operacional, apresentação das informações para aprovação.
Sinay & Quadros (2002)	Diversos	Sistema viário e de transportes	Localização; determinação da área de influência e caracterização da rede; matriz origem/destino, distribuição e alocação de viagens; determinação do nível ambiental da rede; instalação do PGV.
ITE (2005)	Diversos	Sistema viário	Previsão do tráfego não local (de passagem) e a previsão do tráfego gerado pelo PGV (tráfego local); estudo de distribuição de viagens e alocação de tráfego.

Tabela 2.1 – Continuação da descrição dos procedimentos disponíveis

(conclusão)

Entidade/ Autor (Ano)	PGV abordado	Impacto considerado	Etapas adotadas
Vargas (2005)	Diversos	Sistema viário	Início (necessidade identificada e transformada em um problema a ser resolvido); planejamento (detalhamento do projeto); execução (materialização do planejamento); controle e monitoramento (ocorre paralelamente ao planejamento operacional e à execução do projeto); encerramento (avaliação para que erros similares não ocorram novamente)
<i>Department for Transport – Reino Unido Guidance on Transport Assessment (2007)</i>	Diversos	Sistema viário e de transportes	Incentivo à sustentabilidade ambiental; melhoria da acessibilidade ao local do empreendimento; gerenciamento da rede viária existente com melhoramentos na rede local de transporte público; mitigação de impactos residuais através de gerenciamento da operação e pequenas melhorias na rede viária; necessidade de estudo de impacto sempre que o empreendimento gerar 30 viagens veiculares no horário pico; análise do tipo e geração de viagens; horizontes de projeto (situação atual, cinco e 10 anos).

Fonte: adaptado de Portugal e Goldner (2003) e de Coelho (2011) apud Portugal (2012).

Uma análise comparativa desses procedimentos foi realizada por Cunha (2009) em Portugal (2012), que ressaltou que os critérios para o embasamento das análises de impacto têm, na sua maioria, o estabelecimento da previsão de viagens geradas pela implantação de novos empreendimentos.

Cunha (2009) ainda observa que em alguns dos procedimentos disponíveis há preocupação apenas com os impactos externos (fluidez e congestionamento do tráfego, análise do nível de serviço, enquanto outros se preocupam também com o dimensionamento das instalações internas, bem como com os acessos e interfaces.

De acordo com Cunha (2009), os procedimentos americanos são direcionados genericamente para novos PGVs e para estudos de geração de tráfego de automóveis, havendo necessidade de estendê-los aos pedestres e aos veículos de carga. Como ponto positivo dos procedimentos americanos, Cunha (2009) destaca a objetividade do processo de análise com regras técnicas e a transferência do ônus da implantação aos empreendedores, além da análise de diferentes horizontes ao longo dos anos, sem e com a implantação do PGV.

Nas metodologias brasileiras, segundo Cunha (2009), a preocupação maior é com as condições do tráfego no ano da abertura do empreendimento. Os procedimentos brasileiros são ainda: mais focados na implantação de certos tipos de PGVs, como os shopping centers, sem, contudo, perder sua abrangência e aplicabilidade aos PGVs em geral; e mais condizentes com a nossa realidade e necessidades (CUNHA, 2009).

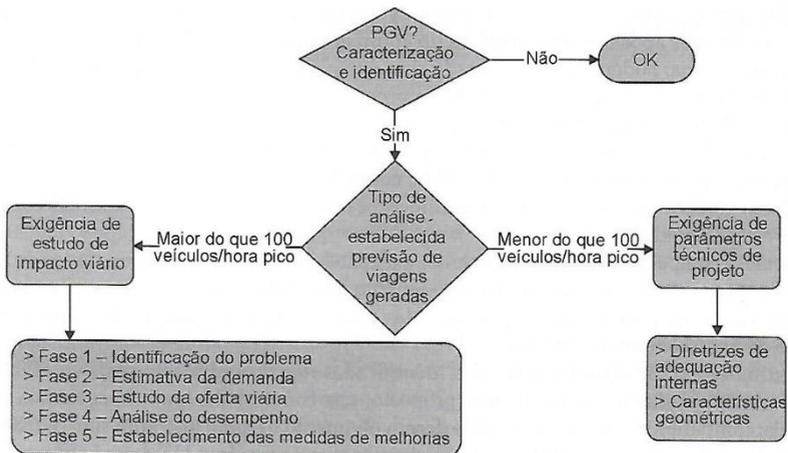
Cunha (2009) identificou algumas restrições nos procedimentos existentes: maior preocupação com os aspectos ambientais, compatibilização de aspectos de desenvolvimento urbano, interação com outros setores, a inserção dos estudos em um plano de desenvolvimento mais amplo, a inclusão do tráfego não motorizado, a contemplação do transporte público e carga, a utilização de modelos compatíveis com a realidade e o emprego de ferramentas de simulação.

Com isso, Cunha (2009) elaborou um procedimento para a análise dos estudos de impactos nos sistemas viários e de transportes de uma cidade, e que também contribua para o aperfeiçoamento da análise e licenciamento dos PGVs quanto ao estabelecimento de uma metodologia e normativas de apreciação, dentro das características que atendam às cidades brasileiras. O mesmo é apresentado a seguir.

No procedimento proposto por Cunha (2009), o primeiro passo é identificar a necessidade de realização do estudo de impacto, assim como

do seu escopo e nível de detalhe. Isso pode ser feito, conforme mostra a Figura 2.1, por meio da verificação da quantidade de adicional gerado de viagens produzidas: caso seja inferior a 100 veículos/hora-pico pode-se optar apenas pela exigência de parâmetros técnicos de projeto; e caso seja superior a 100 veículos/hora-pico, deverá ser exigida a realização do estudo de impacto viário.

Figura 2.1 – Esquema dos procedimentos para análise do impacto viário



Fonte: Cunha (2009).

São cinco fases para a realização do estudo de impacto viário (CUNHA, 2009).

A primeira fase, identificação do problema, consiste na caracterização do empreendimento (seu tipo, porte e sua localização) e dos potenciais problemas derivados, compreendendo as suas dimensões espaciais e temporais, assim estabelecidas (CUNHA, 2009):

- 1) levantamento da situação atual e proposta;
- 2) identificação dos empreendimentos relevantes quanto a impacto viário no entorno do local considerado;
- 3) caracterização do uso do solo;
- 4) delimitação da área crítica; delimitação da área de estudo do empreendimento (em função do tempo de viagem associado

- à identificação do tipo de atividade a ser implantada, bem como à população estimada);
- 5) identificação e análise das características temporais e sazonalidades – períodos críticos a serem considerados no estudo; estabelecimento do horizonte de estudo (em função do número de viagens geradas);
 - 6) identificação de tipos de impactos a serem investigados (aspectos ambientais quanto à população e ruídos); e
 - 7) tipo, etapas de estudo e ferramentas de análise a serem adotadas.

A segunda fase, estimativa de demanda, trata da geração de viagens e da alocação do tráfego, identificando as rotas de acesso, a origem/ destino das viagens e a previsão de distribuição de viagens e respectivas modalidades de transporte, assim definidas (CUNHA, 2009):

- 1) previsão de geração de viagens com o uso de modelos e taxas apropriadas;
- 2) distribuição e escolha modal de viagens;
- 3) previsão de crescimento da área; e
- 4) estimativa de demanda atual e futura.

Esses resultados devem convergir para o dimensionamento das necessidades internas (estacionamentos, carga/descarga, corredores e dispositivos de circulação de pessoas e veículos) e interfaces do empreendimento (dimensionamento de acessos, paradas de automóveis, táxis, pontos de ônibus) (CUNHA, 2009).

A terceira fase, estudo da oferta viária, abrange a análise dos componentes e da rede viária do entorno, a caracterização da oferta viária e de transportes nas áreas críticas e de influência, assim definidas (CUNHA, 2009):

- 1) dimensão das vias do entorno e suas características operacionais;
- 2) identificação da malha hierarquizada;
- 3) identificação de interseções e pontos críticos;
- 4) capacidade das vias; e
- 5) estabelecimento da rede viária atual e futura.

A quarta fase, análise do desempenho, envolve a resultante da alocação e interação entre a demanda e a oferta viária, atual e futura, possibilitando delimitar os impactos provenientes da implementação do empreendimento, assim estabelecidos:

- 1) alocação e simulação do tráfego com o objetivo de estabelecer o desempenho, expresso por indicadores concretos e representativos dos impactos;
- 2) análise de capacidade, nível de serviço e desempenho das componentes e rede viária para as condições atual e futura, considerando dois cenários, sem e com PGV; e
- 3) identificação dos trechos viários e interseções críticas e análise do desempenho para os dois cenários, sem e com o PGV, isolando os impactos provenientes da implantação do empreendimento (CUNHA, 2009).

Por último, a quinta fase, estabelecimento das medidas de melhoria, aborda a identificação, análise e seleção das medidas recomendadas quanto à abrangência das intervenções, compreendendo (CUNHA, 2009):

- 1) estabelecimento das melhorias viárias necessárias (obras viárias, adequações geométricas);
- 2) estabelecimento da sinalização viária necessária; adequações aos pedestres (calçadas, abrigos, passarelas), dentro das normas do desenho universal (acessibilidade para as pessoas com deficiência) e ciclistas; estabelecimento do ônus do empreendedor (custos inerentes às adequações);
- 3) estabelecimento de responsabilidades;
- 4) estabelecimento de prazos e cronogramas; implemento de medidas de mitigação para minimizar ou compensar os impactos causados;
- 5) implemento de medidas compensatórias para minimizar os impactos ambientais (poluição e ruídos); e
- 6) acompanhamento da execução do projeto, controle e monitoramento operacional.

Ressalta-se que as medidas devem ser derivadas do desempenho dos sistemas viários e de transportes, testadas, avaliadas e selecionadas, e

devem ser detalhadas, seus custos quantificados e estabelecidos os repasses necessários ao empreendedor (CUNHA, 2009).

No procedimento proposto por Portugal e Goldner (2003) apud Portugal (2012), Figura 2.2, são consideradas, além dos impactos no sistema viário (automóveis, ônibus e pedestres), as necessidades internas de armazenamento dos meios de transporte (automóveis, caminhões e táxis) e de circulação de veículos e pessoas no interior do PGV.

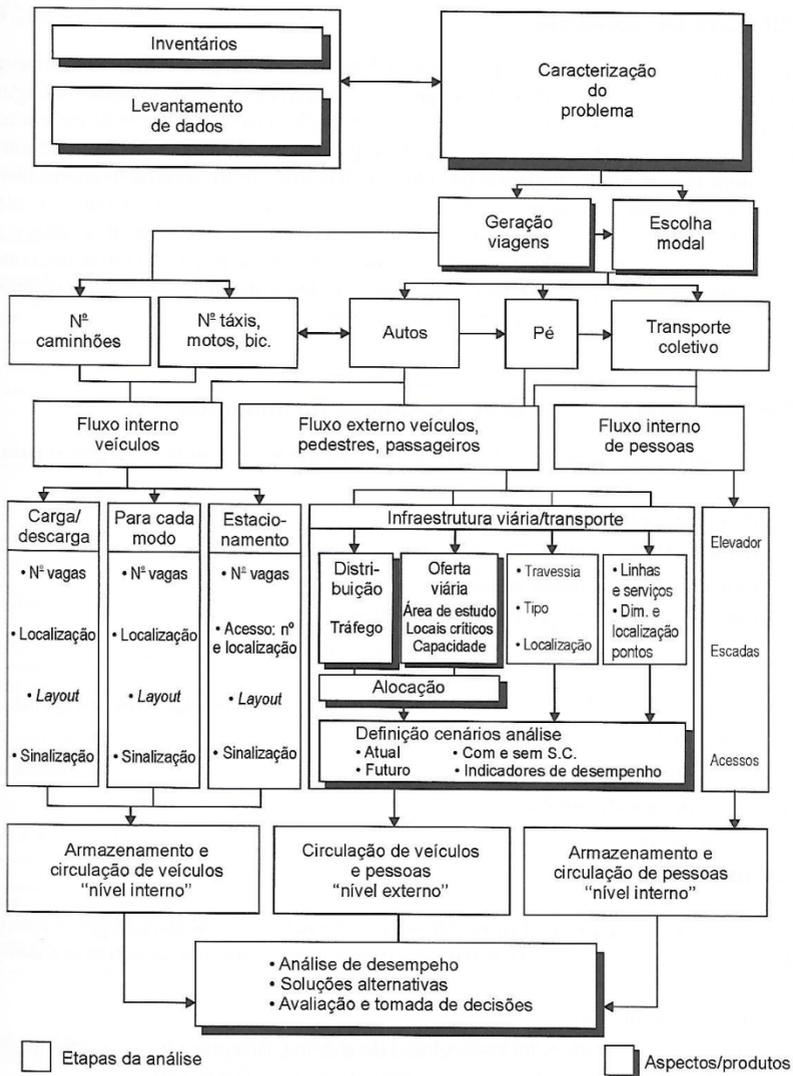
Em relação às técnicas de análise utilizadas nos estudos de impactos, as mesmas podem variar desde técnicas analíticas, como é o caso do Manual de Capacidade Rodoviária (*Highway Capacity Manual – HCM*), até técnicas de simulação, como é o caso do uso de Transplan, Transcad, Netsim, Paramics, etc. (PORTUGAL, 2012).

O critério de escolha da técnica pode depender: do tipo de intervenção e da dimensão espacial e do horizonte temporal do estudo; do número de viagens no horário de pico; ou do tipo de projeto e velocidade de alteração do mesmo (PORTUGAL, 2005; ITE, 2009; LOPES, 2010).

Entre esses, de acordo com Gonçalves (2012), o critério do número de viagens no horário de pico parece ser o mais apropriado, uma vez que o mesmo pode indicar o potencial de impactos do empreendimento. Ele recomenda ainda que a análise analítica seja utilizada principalmente para potenciais baixos, embora também possa ser empregada para potenciais médios, e as técnicas de simulação, para potenciais médio, alto e regional.

Como medidas mitigadoras, pode-se citar: pequenas intervenções físicas na oferta viária, na sua sinalização e controle de tráfego; a construção de alargamentos e mesmo de novas vias, ampliando a sua capacidade; o gerenciamento da demanda e investimentos na melhoria dos sistemas de transporte público e nas alternativas não motorizadas; melhorias de produtividade, ocupação e diversificação dos padrões de uso do solo (PORTUGAL, 2012).

Figura 2.2 – Estrutura global de análise de impactos

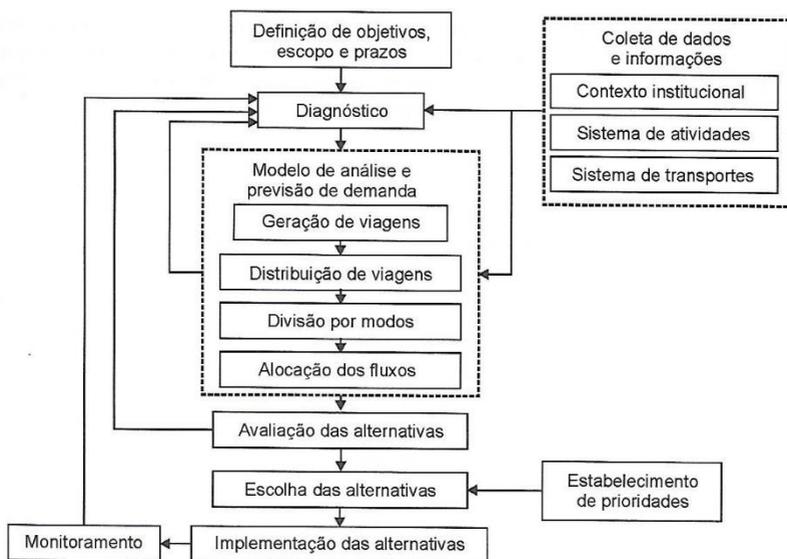


Fonte: Portugal e Goldner (2003) apud Portugal (2012).

2.2 GERAÇÃO DE VIAGENS

O modelo de geração de viagens é parte componente de um modelo de análise e previsão de demanda e este é parte central de um processo de planejamento mais amplo, o processo de planejamento estratégico de transportes, conforme mostra a Figura 2.3; e que possui duas abordagens: o modelo quatro etapas, que é formado pela geração de viagens, distribuição de viagens, escolha modal e alocação de viagens; ou, ainda, os modelos baseados em atividades, no qual a geração, a distribuição e a escolha modal são substituídas por uma única etapa (PORTUGAL, 2012).

Figura 2.3 – Processo do planejamento estratégico do transporte de carga



Fonte: elaborado a partir de Adler (1978); Anderson (1994) e FHA (1999); apud Portugal (2012).

No contexto de um PGV, segundo Pinto, Diógenes & Lindau (2003), a geração deve ser entendida como o processo de produção e atração de viagens.

Nos EUA, a publicação mais difundida e utilizada para os cálculos de previsão de viagens em PGVs é a *Trip Generation*, do ITE. Os modelos do ITE são elaborados a partir do levantamento do número de viagens de veículos nos acessos ao empreendimento em vários PGVs de mesmo tipo e, de posse desses dados, elaboram-se as taxas e modelos de geração de viagens, correlacionando o número de viagens com variáveis explicativas, tais como área do empreendimento, número de empregados, etc. (GOLDNER & SCHMITZ, 2010).

Ressalta-se que as taxas e modelos de geração de viagens propostos na publicação do ITE são utilizadas também em vários outros países, quando os mesmos não possuem estudos sobre a realidade local. No entanto, deve-se tomar cuidado ao se utilizar as taxas do ITE em condições distintas, visto que essas taxas tendem a superestimar a geração de viagens, principalmente se utilizadas no meio urbano, em áreas compactas ou em regiões ocupadas por população de baixo poder aquisitivo. Por causa disso, é importante o desenvolvimento sistemático de tais taxas que compreendam as especificidades locais e sua atualidade, para previsões mais compatíveis com a realidade (PORTUGAL, 2012).

Assim, como essa tese tem como um dos objetivos contribuir com as pesquisas no campo de PGVs, inicialmente se propôs fazer uma revisão bibliográfica sobre os modelos de geração de viagens utilizados para PGVs em geral.

Ressalta-se que optou-se por realizar uma análise baseada no livro “Polos Geradores de Viagens orientados a qualidade de vida e ambiental: modelos e taxas de geração de viagens”, escrito por diversos autores ibero-americanos especializados nessa área e organizado por Portugal (2012). Este é um trabalho reconhecido pelo mundo acadêmico e compila os principais resultados de estudos realizados sobre diferentes PGVs e, por isso, foi utilizado como principal fonte.

Entre os autores do livro se encontram:

- 1) Adriana Scovino da Rocha;
- 2) Ana Paula Borba Gonçalves;
- 3) Anabela Salgueiro Narciso Ribeiro;
- 4) Ângela Beatriz Souza Bertazzo;
- 5) Angela Rosas Meza;
- 6) Antônio Néelson Rodrigues da Silva;
- 7) Antonio Paulo de Hollanda Cavalcante;
- 8) Archimedes Azevedo Raia Jr.;

- 9) Bianca Côrtes Cardoso;
- 10) Cecília Rafaela dos Santos Pinto;
- 11) Ciro Jaramillo Molina;
- 12) Claudio Falavigna;
- 13) Cristiane Duarte Ribeiro de Souza;
- 14) Denise Beer Frenkel;
- 15) Diana Scabelo da Costa Pereira da Silva Lemos;
- 16) Diego Westphal;
- 17) Eduardo Pessoa de Andrade;
- 18) Elisabeth Frieco;
- 19) Eloisa Basto Amorim de Moraes;
- 20) Erika Cristine Kneib;
- 21) Eymard de Albuquerque Pinheiro;
- 22) Fábio dos Santos Gonçalves;
- 23) Gabriel Vendrúsculo de Freitas;
- 24) Geisa Aparecida da Silva Gontijo;
- 25) Giovanna Megumi Isida Tedesco;
- 26) Gonçalo Homem de Almeida Rodrigues Correia;
- 27) Ilce Marília Dantas Pinto de Freitas;
- 28) Inês Frade;
- 29) Isolina Cruz;
- 30) Jackeline Murillo Hoyos;
- 31) Jorge Augusto Martins Gonçalves;
- 32) Jorge Galarraga, Josefina Flórez Díaz;
- 33) Juan Carlos Sanáñez;
- 34) Juan Pedro Moreno Delgado;
- 35) Laura Sinay;
- 36) Lenise Grando Goldner;
- 37) Licínio da Silva Portugal;
- 38) Luciana Santos Pires;
- 39) Marcela Rubert;
- 40) Marcello Marinho Berenger Vianna;
- 41) Marcelo Herz;
- 42) Marcio de Almeida D'Agosto;
- 43) Maria Alice Prudêncio Jacques;
- 44) Maria Cristina Fogliatti;
- 45) Maria Laura Albrieu;
- 46) Maria Leonor Alves Maia;
- 47) Mariana de Paiva, Osmar Soares;

- 48) Patricia Ingrid de Souza Coelho;
- 49) Regina Fátima de Faria Cunha;
- 50) Sergio Marcolini;
- 51) Thais de Cássia Martinelli Guerreiro; e
- 52) Vânia Gouvêa Campos.

Além disso, como o livro foi publicado em 2012, foi proposta a realização de uma revisão bibliográfica sistemática, conforme procedimento apresentado na Figura 1.1, de modo a identificar estudos publicados entre 2013 e a data presente.

2.3 PLANEJAMENTO DA REVISÃO BIBLIOGRÁFICA

O primeiro passo do procedimento da revisão bibliográfica, que é o planejamento da revisão, é descrito a seguir.

2.3.1 Identificação da necessidade da revisão

Conforme citado no Capítulo 1, os portos e terminais marítimos são um tipo de PGV e, assim como os outros PGVs, sua implantação e operação causam impactos que precisam ser previstos e tratados, sendo para isso, necessário estimar as viagens geradas por esses PGVs. De modo a comparar as técnicas e variáveis utilizadas nos modelos de geração de viagens para portos e terminais marítimos com as técnicas e variáveis utilizadas nos modelos de geração de viagens para os outros PGVs, foi proposta essa revisão bibliográfica sistemática sobre modelos de geração de viagens para PGVs, de modo a complementar a referência base Portugal (2012) supracitada.

2.3.2 Elaboração da proposta da revisão

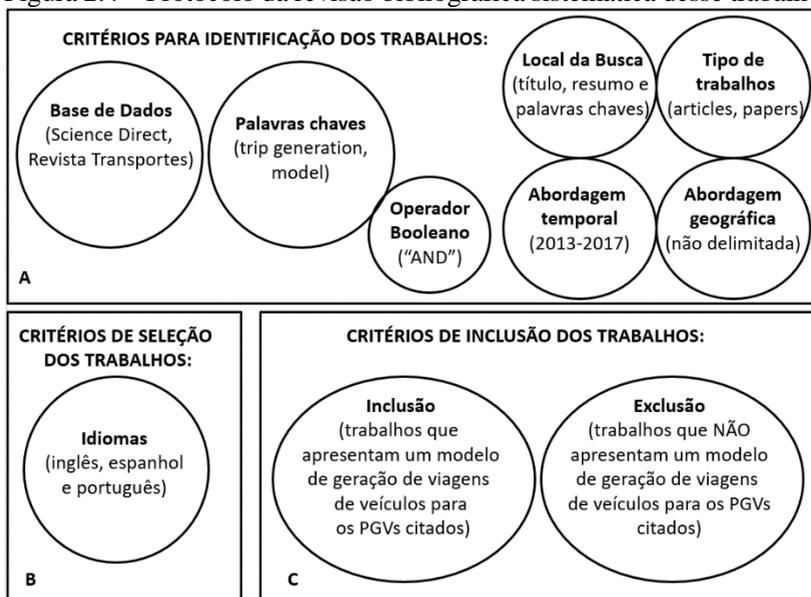
A finalidade dessa revisão bibliográfica sistemática foi identificar estudos de modelos de geração de viagens para veículos em PGVs, tais como: shopping centers, equipamentos institucionais, estabelecimentos residenciais, hotéis, estações metroferroviárias, estabelecimentos de ensino, hospitais e estabelecimentos de saúde, hipermercados e

supermercados, polos de uso misto e polos de uso múltiplo, e polos geradores de viagens de carga (exceto para portos ou terminais marítimos, o qual foi abordado no Capítulo 3); destacando a técnica utilizada e as variáveis utilizadas.

2.3.3 Desenvolvimento do protocolo de revisão

Nesse trabalho, o protocolo de revisão empregado foi o exposto na Figura 2.4.

Figura 2.4 – Protocolo da revisão bibliográfica sistemática desse trabalho



Fonte: Elaboração própria com base em Oliveira, 2015.

Em resumo:

- foram utilizadas as bases de dados Science Direct e Revista Transportes;
- foram usadas as palavras-chaves “*trip generation*” e “*models*”, ou seja, “geração de viagens” e “modelos” e suas

- combinações por meio do operador booleano “*and*”, ou seja, “*e*”;
- c) foram considerados os estudos do tipo “*papers*”, ou seja, “artigos” disponíveis na base de dados e a investigação das palavras chaves foi feita no título, resumo e palavras-chaves desses estudos;
 - d) não se limitou a busca a alguma abrangência geográfica ou temporal;
 - e) foram selecionados apenas os trabalhos que se encontravam em língua inglesa, espanhola ou portuguesa; e
 - f) após selecionados os artigos, para a inclusão e exclusão deles no trabalho foi considerado se eles atingiam o objetivo principal ou secundário dessa tese, ou seja, mostravam a construção de um modelo de geração de viagens para PGVs.

Os resultados dessa coleta podem ser vistos no item 2.4 a seguir.

Cabe ressaltar que não foi utilizado o operador booleano “*ou*”, visto que resultava em um número grande de trabalhos que não atingiam o objetivo dessa revisão sistemática. Ainda, consideraram-se apenas artigos, visto que, de acordo com Nord et al. (1995) e Ngai & Wat (2002), profissionais e pesquisadores utilizam os periódicos científicos para adquirir informações e divulgar novos resultados, o que justifica essa escolha. Também se optou por não restringir a busca a alguma geográfica, de modo que se pretendia observar quais países estavam estudando o tema. Em relação à abordagem temporal, como a revisão bibliográfica sistemática visou complementar Portugal (2012), foi restringida para o período entre 2013 e 2017. Por fim, para a seleção dos artigos apenas se levou em conta o idioma em que se encontravam, uma vez que os periódicos constantes nessas bases de dados são de boa qualidade e optou-se por não limitar a nenhuma classificação de Qualis dos mesmos.

2.4 REALIZAÇÃO DA REVISÃO BIBLIOGRÁFICA

Com o planejamento concluído, deu-se sequência ao procedimento com a etapa de realização do mesmo conforme apresentado a seguir.

2.4.1 Identificação e seleção dos trabalhos

Os trabalhos foram identificados e selecionados segundo os critérios descritos no subitem 2.3.3, na página 53. Como todos estavam no idioma inglês ou espanhol, todos foram selecionados para a avaliação.

2.4.2 Avaliação dos trabalhos selecionados

Os resumos dos trabalhos encontrados no subitem anterior do presente item foram lidos com o objetivo de verificar se atendiam os critérios de inclusão, ou seja, se eles mostravam a construção de um modelo de geração de viagens de veículos para os tipos de PGVs citados. Caso não tenha ficado claro pela leitura do resumo se eles atingiam ou não esse objetivo, esses trabalhos foram lidos na íntegra para que pudessem ser avaliados.

Dos documentos selecionados, 8 foram incluídos na pesquisa: (i) Díaz (2017); (ii) Guzman, Gomez & Rivera (2017); (iii) Oliveira et al. (2017); (iv) Ducret & Gonzalez-Feliu (2016); (v) Mommensa, van Lier & Macharis (2016); (vi) Arliansyah & Hartono (2015) (vii) Herz & Galarraga (2014); (viii) Torquato & Raia Jr. (2014); e, portanto, tiveram dados e informações extraídos no subitem 2.4.3 desse item a seguir de modo a subsidiar o restante dessa pesquisa.

O restante não foi incluso, pois não atingia o objetivo da revisão bibliográfica sistemática, tratando sobre temas como: modelos de geração de viagens para motociclistas, ciclistas, pedestres ou caronas; modelos para identificar zonas de colisões de veículos; modelos para estimar a emissão de gás; modelos de escolha e distribuição de viagens; técnicas de simulação; modelos para planos de evacuação; coleta de dados com o uso de celulares; entre outros.

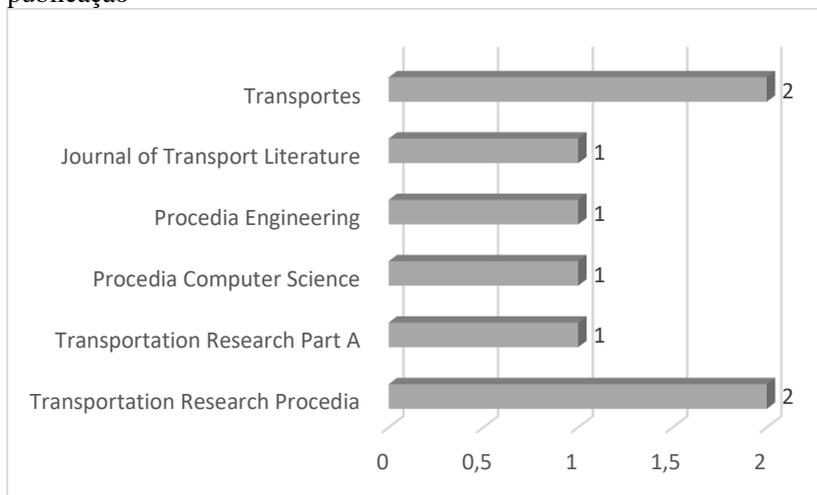
2.4.3 Extrair dados e informações

Nessa fase, os trabalhos incluídos na revisão bibliográfica sistemática foram classificados de acordo com os periódicos científicos do qual fazem parte, o ano de sua publicação, a sua abordagem regional, temática, e metodológica.

No caso dos termos utilizados para identificação dos trabalhos, como no levantamento foi utilizado três palavras-chaves e o operador booleano “e”, todos os documentos identificados e selecionados possuíam essas duas palavras-chaves: “geração de viagens” e “modelos”.

Em relação à distribuição dos trabalhos segundo suas fontes de publicação, dois (25%) deles foram publicados pela Revista Transportes, dois (25%) pela *Transportation Research Procedia*, um (12,5%) pelo *Journal of Transport Literature*, um (12,5%) pelo *Procedia Computer Science* e um (12,5%) pelo *Transportation Research Part A* (Figura 2.5).

Figura 2.5 – Distribuição dos trabalhos em relação à sua fonte de publicação

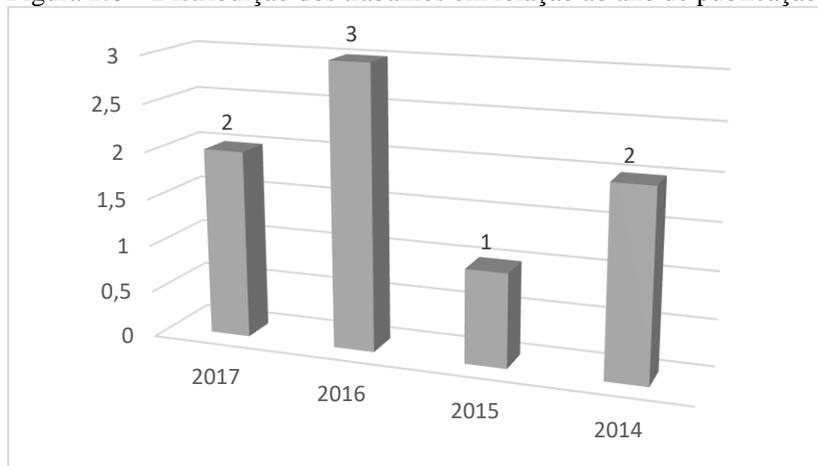


Fonte: Elaboração própria com base em Oliveira, 2015.

Quanto à abordagem temporal, percebe-se que, segundo essa revisão, a partir de 2013, o tema foi abordado duas vezes em 2014, uma vez em 2015, três vezes em 2016 e duas vezes em 2017 (Figura 2.6).

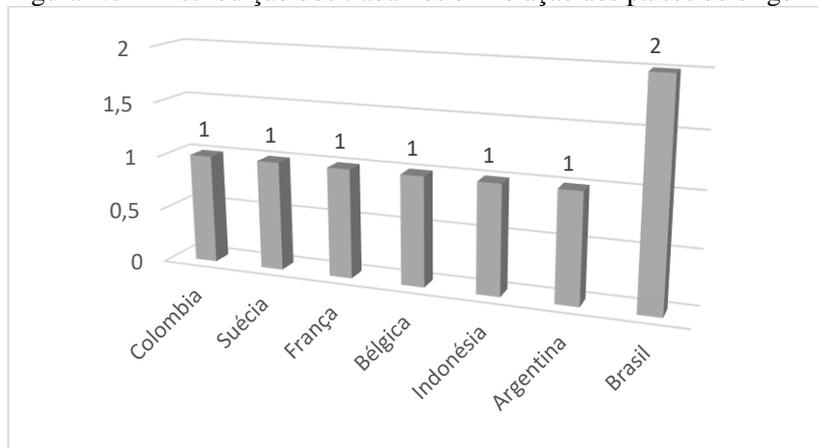
No que se concerne à distribuição dos trabalhos em relação aos países de procedência, dois se originaram no Brasil, e um se originou na Argentina, um na Colômbia, um na França, um na Bélgica, um na Suécia e um na Indonésia (Figura 2.7).

Figura 2.6 – Distribuição dos trabalhos em relação ao ano de publicação



Fonte: Elaboração própria com base em Oliveira, 2015.

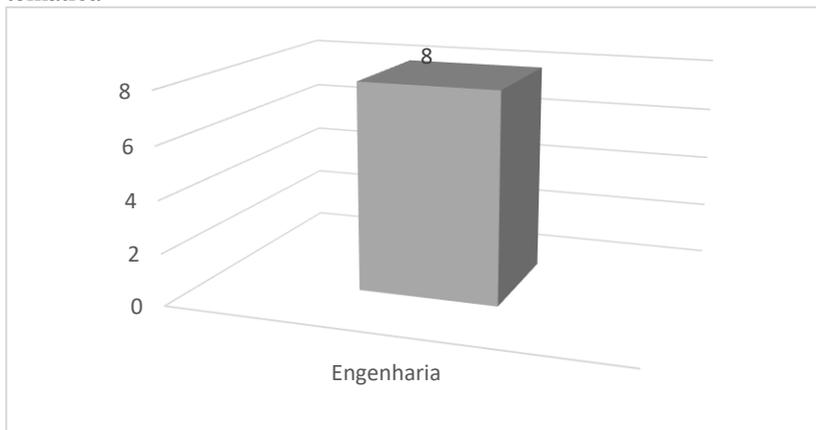
Figura 2.7 – Distribuição dos trabalhos em relação aos países de origem



Fonte: Elaboração própria com base em Oliveira, 2015.

Sobre a abordagem temática, todos os artigos (100%) se encontram relacionados ao tema Engenharia (Figura 2.8).

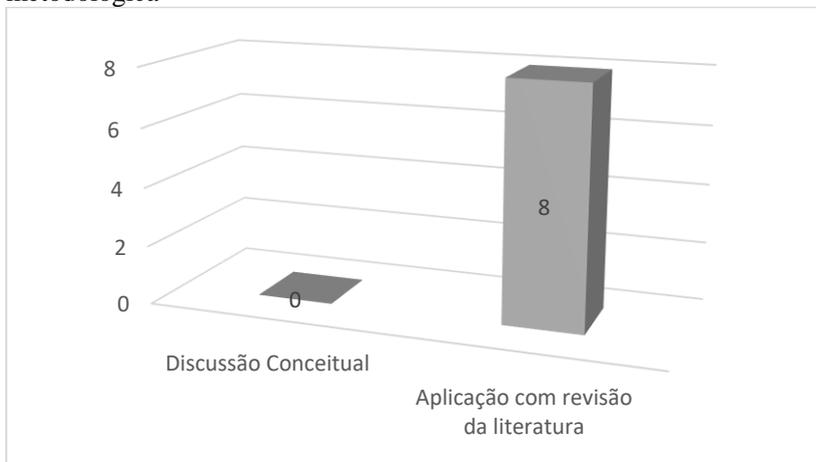
Figura 2.8 – Distribuição dos trabalhos em relação à abordagem temática



Fonte: Elaboração própria com base em Oliveira, 2015.

No que tange à abordagem metodológica, todos os trabalhos (100%) realizam uma aplicação dos conceitos pesquisados e apresentaram uma seção de revisão bibliográfica (Figura 2.9).

Figura 2.9 – Distribuição dos trabalhos em relação à abordagem metodológica



Fonte: Elaboração própria com base em Oliveira, 2015.

Cabe ressaltar que a revisão bibliográfica para o período anterior à 2013 foi baseada no livro “Polos Geradores de Viagens orientados a qualidade de vida e ambiental: modelos e taxas de geração de viagens”, escrito por diversos autores ibero-americanos especializados nessa área e organizado por Portugal (2012).

2.4.4 Síntese dos estudos

A síntese dos estudos é apresentada no item 2.7, referente à análise da revisão bibliográfica.

2.5 COMUNICAÇÃO E DIVULGAÇÃO DOS RESULTADOS DA REVISÃO

Por fim, com os resultados das etapas anteriores em mãos, esse documento final foi redigido e submetido à banca examinadora com o objetivo de obter o grau de Doutora em Engenharia Civil; resultando, portanto, na publicação de uma tese de doutorado e em um artigo científico baseado nela.

2.6 DESCRIÇÃO DOS ESTUDOS ENCONTRADOS

Detalhes sobre Portugal (2012) e os oito trabalhos incluídos na pesquisa por meio da revisão bibliográfica sistemática, podem ser visualizados a seguir de acordo com o tipo de PGV para o qual foram desenvolvidos seus modelos de geração de viagens.

2.6.1 Shopping centers

De acordo com Portugal (2012), nos estudos sobre PGVs, os shopping centers são comumente contemplados, seja pela sua capacidade de gerar viagens, seja pela proliferação desse tipo de empreendimento.

Nesse subitem é apresentada uma análise de Portugal (2012) sobre os modelos de geração de viagem encontrados no contexto ibero-americano. Os trabalhos que foram analisados por Portugal (2012) são os

seguintes: Goldner (1994); Martins (1996); Espejo (2001); Rosa (2003); Cárdenas (2003); Andrade (2005); De Andrade (2005); e CET-SP (2011); todos apud Portugal (2012).

A análise realizada confirmou a influência exercida pelo ITE, onde a maioria dos trabalhos se divide no que diz respeito à abordagem em: regressão linear e taxas de geração de viagens. Destes, todos utilizaram a técnica de regressão linear simples, exceto um deles que utilizou uma regressão linear bivariada.

Das variáveis independentes utilizadas nos modelos, se destacou a área bruta locável.

Na revisão bibliográfica sistemática, foi identificado o trabalho de Oliveira et al. (2017), o qual elaborou modelos de geração de viagens para a cidade de Belo Horizonte, incluindo modelos para shoppings centers. Os modelos foram desenvolvidos com a técnica de regressão linear simples e múltipla e as variáveis independentes utilizadas foram: área construída, número de vagas de estacionamento, número de lojas e capacidade do cinema.

2.6.2 Equipamentos institucionais

Nesse subitem são tratados estabelecimentos de diferentes usos, embora todos eles com caráter institucional, tais como: prédios da administração pública, agências de correios, igrejas, bibliotecas, prisões e prédios militares.

Da mesma forma que o subitem 2.6.1, é apresentada uma análise de Portugal (2012) sobre os modelos de geração de viagem para esses tipos de PGVs, a qual incluiu os seguintes estudos: de escritórios governamentais por Mobley (2006) e ITE (2008) código 730; de estabelecimentos religiosos por ITE (2008) código 560, ITE (2008) código 561, Ratrou (2009) e Murillo et al. (2011); de bibliotecas por ITE (2008) código 590; de agências dos correios por Azar & Pant (1985) e ITE (2008) código 732; de prisões por French et al. (2000) e ITE (2008) código 571; e de bases militares por Hazarvatian (1998) e ITE (2008) código 501; todos apud Portugal (2012).

Alguns desses estudos consideraram somente taxas de viagens, que derivaram: de equações de análise de regressão linear simples e múltipla e de regressão logarítmica. Outros elaboraram modelos de regressão linear simples, regressão linear bivariada e de regressão logarítmica.

Das variáveis independentes utilizadas nos modelos, se destacou a área total construída (ou área bruta de piso). As exceções foram a utilização do número de leitos para o caso de prisões; a utilização do número de civis e de militares ou do número de empregados, e do número de veículos autorizados para o caso das bases militares; e a utilização do número de membros e do número de assentos disponíveis para o caso de sinagogas e igrejas.

Não foram encontrados estudos a partir de 2013 para equipamentos institucionais na revisão bibliográfica sistemática.

2.6.3 Estabelecimentos residenciais

Usualmente as viagens de passageiros têm como origem ou destino a residência dos mesmos – cerca de 80%, de acordo com Bruton (1979) apud Portugal (2012). Entretanto, apesar dessa relevância, Portugal (2012) afirma que há relativamente poucos estudos sobre modelos e taxas de geração de viagens para esses empreendimentos.

Logo, na análise de Portugal (2012) foram incorporados os trabalhos: do ITE (2008), de Arrington & Cervero (2008), Leániz et al. (2008), e de Niterói Transporte e Trânsito (Nittrans, 2011), todos apud Portugal (2012).

Destes, todos produziram suas estimativas através de índices ou taxas, exceto por ITE (2008) apud Portugal (2012) que apresentou para cada uso do solo estudado, as equações de regressão.

As variáveis independentes utilizadas nas diferentes tipologias de estabelecimentos residenciais foram: o número de unidades residenciais, o número de pessoas, o número de veículos, o número de vagas de estacionamento e a área urbanizada em acres.

Na revisão bibliográfica sistemática foram identificados os seguintes trabalhos com foco em modelos de geração de viagens para estabelecimentos residenciais a partir de 2013: Guzman, Gomez & Rivera (2017), Oliveira et al. (2017) e Arliansyah & Yusuf Hartono (2015), que elaboraram modelos de geração de viagens para cidades específicas, mas que incluíam residências; e Torquato & Raia Jr. (2014), que elaboraram modelos de geração de viagens para condomínios residenciais horizontais.

O primeiro utilizou a técnica de regressão linear múltipla; o segundo, regressão linear simples e múltipla; o terceiro, redes neurais

artificiais; e o último, regressão linear simples com base logarítmica. Em relação às variáveis independentes, foram utilizadas: quantidade de residências, área construída, quantidade de residentes, quantidade de dormitórios por residência, quantidade de estudantes e trabalhadores por residência, nível de renda e quantidade de automóveis por residência.

2.6.4 Hotéis

Esse tipo de PGV está presente em grande número nas médias e grandes cidades e, por causa disso, se faz necessário conhecer os estudos existentes sobre os mesmos, bem como estimular a realização de novos estudos. Todavia, Portugal (2012) acredita que estes não são estudados ainda com frequência.

Portugal (2012) destacou as seguintes fontes bibliográficas sobre esse tema: os estudos do ITE (2008), desenvolvidos para diferentes categorias de hotéis; os trabalhos de Ackret & Hosea (1992) e Rowe et al. (2002), para hotéis cassino de Las Vegas; o estudo de Feitosa & Balassiano (2003), desenvolvido para os padrões de viagens de hotéis, residência da cidade do Rio de Janeiro; e os estudos de Goldner et al. (2006) e Goldner & Inocêncio (2007), em relação aos padrões de viagens e à geração de viagens em hotéis da cidade de Florianópolis; todos apud Portugal (2012).

A análise dos mesmos evidencia a influência exercida pelos trabalhos do ITE na produção de pesquisas sobre o tema. Todos os trabalhos utilizaram regressão linear simples ou múltipla.

Todos os modelos utilizaram como variáveis independentes o número de quartos ou o número de empregados.

Não foram encontrados estudos a partir de 2013 para hotéis na revisão bibliográfica sistemática.

2.6.5 Estações metroferroviárias

Com o crescente problema de espraiamento quanto à ocupação do solo e de congestionamento nas grandes e médias cidades, o sistema metroferroviário voltou a ganhar destaque como opção de transporte. Consequentemente, os mesmos passaram a ser objeto de estudo de geração de viagens para suas estações (PORTUGAL, 2012).

Na análise de Portugal (2012) sobre esse tipo de PGV, foram incluídas 14 fontes bibliográficas, as quais encontram-se na Tabela 2.2.

Tabela 2.2 – Resumo da revisão bibliográfica para estações metroviárias

Autor	Local	Método de análise	Categorias das variáveis intervenientes
Bertolini (1999)	Holanda	Método nó-lugar	Uso do solo e integração modal
Reusser et al. (2008)	Suíça	Regressão linear e gráfico nó-lugar	Uso do solo e integração modal
Kuby et al. (2004)	EUA	Regressão linear	Uso do solo, integração modal, centralidade e renda
Kim et al. (2007)	EUA	Modelo Multinomial logit	Empregos, integração modal e socioeconômicos
Lara et al. (2007)	Rio de Janeiro	Regressão linear	Empregos, renda, acesso à estação e desenvolvimento social
Sohn & Shin (2010)	Seul	Regressão linear	Ambiente construído (uso do solo), integração modal e centralidade
Gutiérrez et al. (2011)	Madrid	SIG e regressão linear	Uso do solo, integração modal, região metropolitana, estrutura da rede e socioeconômicos
Miller & Thorn (2009)	EUA	Modelo gravitacional	Viagens a negócios, população e distância
ITE (2008)	EUA	Regressão linear	Park and ride
Cervero & Arrington (2008)	EUA	Regressão linear	Uso do solo
O'Sullivan & Morral (2008)	Calgary	Gráficos estatísticos	Distância
Silva (2008)	Brasília	Modelo prima-espaco-tempo	Distância e tempo
Maghelal (2011)	Dallas	Pesquisa de campo	Distância
Gonçalves (2006)	Rio de Janeiro	Centralidade com base na Teoria das Grafos	-

Fonte: Portugal (2012).

Quanto às variáveis independentes utilizadas, pode-se citar as mostradas na Tabela 2.3.

Tabela 2.3 – Categorias e variáveis mais relevantes

Categoria	Variável	Descrição
Uso do solo	População próxima à estação	Número de empregos próximos à estação
	Empregos próximos à estação	População próxima à estação (raio = 500m)
	Polos Geradores de Viagens	Influência dos PGVs atraindo usuários
	Área Central de Negócios	Atraem grande contingente de pessoas e viagens
Integração modal	Vagas em estacionamentos	Acesso, por carro, de todas as áreas facilitado
	Integrações	Atratividade financeira e comodidade
	Congestionamentos	Horários regulares e conforto do sistema ferroviário
	Espaço público	Atrai usuários de outros sistemas e aumenta a atratividade para os que estão no sistema
Ambiente urbano: aspectos socioeconômicos	Renda	Falta de recursos para compra de automóvel/combustível; maior uso do transporte público
	Índice de Desenvolvimento Humano (IDH)	Índice de Desenvolvimento Humano que considera, além da dimensão econômica, outras características sociais, culturais e políticas que influenciam na qualidade de vida humana
Sistema ferroviário: qualidade e custo de serviço	Número de usuários	Vagões lotadas reduzem o número de passageiros
	Estrutura tarifária	Tarifa alta afasta as classes de menor poder aquisitivo
Sistema ferroviário: estrutura da rede	Estação terminal	A estação terminal possui área de influência mais abrangente
	Estações de transferências	Maior número de acesso devido ao desembarque de diversas linhas
	Centralidade	Ligada diretamente à atratividade, aumento o número de viagens na área

Fonte: Portugal (2012).

Percebe-se que a regressão linear é popular entre os estudos analisados.

Não foram encontrados estudos a partir de 2013 para estações metroferroviárias na revisão bibliográfica sistemática.

2.6.6 Estabelecimentos de ensino

As viagens geradas pelos estabelecimentos de ensino ocorrem de modo regular e em horários coincidentes com os picos de tráfego no sistema viário. Por causa disso, é importante delimitar os impactos que a implantação de um PGV desse tipo pode causar no seu entorno (PORTUGAL, 2012).

Nesse subitem foram consideradas as instituições de ensino de Educação Básica e Superior analisadas por Portugal (2012): CET-SP (2000), ITE (2008) e Betarzzo (2008); e para a Educação Superior: CET-SP (1983), TECTRAN (2003, 2004), Nunes (2005), Souza (2007), ITE (2008) e Herz et al. (2007, 2009); todos apud Portugal (2012).

Todos os estudos utilizaram como variáveis explicativas o número de alunos ou de empregados; e apresentaram modelos de regressão ou apenas estimativas através de taxas.

Na revisão bibliográfica sistemática foram identificados os seguintes trabalhos com foco em modelos de geração de viagens para estabelecimentos de ensino a partir de 2013: Guzman, Gomez & Rivera (2017), Oliveira et al. (2017) e Arliansyah & Yusuf Hartono (2015), que elaboraram modelos de geração de viagens para cidades específicas, mas que incluíam estabelecimentos de ensino.

O primeiro utilizou a técnica de regressão linear múltipla; o segundo, regressão linear simples e múltipla; o terceiro, redes neurais artificiais. Em relação às variáveis independentes, foram utilizadas: quantidade de escolas e universidades, área construída de escolas e universidades, quantidade de estudantes registrados nas escolas e universidades, quantidade de professores ou funcionários registrados nas escolas e universidades, quantidade de salas de aulas nas escolas e universidades, quantidade de vagas de estacionamento, quantidade de estudantes por residência e quantidade de automóveis por residência.

2.6.7 Hospitais e estabelecimentos de saúde

Existem diferentes concepções de estabelecimentos que têm como finalidade a assistência de enfermos nos diversos países. Essa diferença de características desses empreendimentos dificulta a comparação dos elementos de análise desses PGVs para países distintos e faz com que os mesmos precisem ser considerados nas suas especificidades e geografia espacial (PORTUGAL, 2012).

Portugal (2012) afirma que o número de casos estudados no campo voltado aos estabelecimentos de saúde, principalmente no que diz respeito aos modelos de geração de viagens, é relativamente pequeno quando comparado com outros PGVs tradicionais, como os shopping centers; e que grande parte dos trabalhos encontrados faz análises de impactos utilizando os modelos e taxas encontradas em ITE (2004 e 2008).

Os estudos analisados por Portugal (2012) foram os seguintes: CET-SP (1983), ARPC (2001), Macêdo, Filizola & Souza (2002), Carqueja (2006), Latinopoulou, Tsohos & Basbas (2010), Fehr & Pers (2008), Gontijo & Raia Jr. (2009, 2010a, 2010b), Martínez, Hueso & Sánchez (2010), Albrieu, Pastor & Galarraga (2009, 2011), Rosas & Sanánez (2010, 2012) e Nishimori & Raia Jr. (2011); todos apud Portugal (2012).

Destes, a maioria está enfocada na geração de taxas de viagens, e só cinco deles desenvolvem modelos de regressão para estimar as viagens geradas.

As variáveis independentes mais frequentes nos modelos foram a quantidade de leitos instalados e a área construída. A quantidade de médicos e de funcionários nos hospitais e prontos-socorros foram outras variáveis também consideradas em alguns modelos.

Na revisão bibliográfica sistemática foram identificados os seguintes trabalhos com foco em modelos de geração de viagens para hospitais e estabelecimentos de saúde a partir de 2013: Oliveira et al. (2017), que elaboraram modelos de geração de viagens para a cidade de Belo Horizonte, no Brasil, mas que incluíam hospitais e estabelecimentos de saúde.

Oliveira et al. (2017) utilizou a técnica de regressão linear simples e múltipla para elaborar os modelos de geração de viagens, e as variáveis independentes utilizadas foram: área construída, número de leitos e número de funcionários.

2.6.8 Hipermercados e supermercados

De acordo com Portugal (2012), na América Latina o conceito de supermercado está relacionado aos estabelecimentos comerciais que oferecem bens de consumo em sistemas de autoatendimento, com oferta diversa de alimentos e outros produtos; enquanto que o de hipermercado, a aqueles que tenham, além das características supracitadas, superfícies extensas e possuam classificação dos produtos por setores ou departamentos.

Na análise de Portugal (2012) foram incorporados os seguintes estudos: ITE (2008) para os códigos 813, 850 e 854; Brehmer & Butorac (2003); Pearson et al. (2009); CET-SP (1983); Goldner (1994); Barbosa & Gonçalves (2000); Silva (2006); Galarraga et al. (2007); Freitas & Raia Jr. (2008, 2011); e Freitas (2009); todos apud Portugal (2012).

ITE (2008) apud Portugal (2012) estimou valores das taxas e modelos utilizando regressão linear simples. A maioria dos outros estudos utilizou o de ITE como base e também fez uso de regressão linear simples para estimar seus modelos, exceto por um dos estudos que se destacou por utilizar regressão linear múltipla.

A variável independente frequentemente utilizada pelos estudos foi a área total construída do empreendimento.

Na revisão bibliográfica sistemática foram identificados os seguintes trabalhos com foco em modelos de geração de viagens para hipermercados e supermercados a partir de 2013: Oliveira et al. (2017), que elaboraram modelos de geração de viagens para a cidade de Belo Horizonte, no Brasil, mas que incluíam hipermercados e supermercados; e Herz & Galarraga (2014), que elaboraram modelos de geração de viagens para hipermercados e supermercados na Argentina.

Oliveira et al. (2017) utilizou a técnica de regressão linear simples e múltipla para elaborar os modelos de geração de viagens, e as variáveis independentes utilizadas foram: área construída, número de vagas de estacionamento e número de *checkout*. Herz & Galarraga (2014) elaboraram modelos de regressão linear e regressão linear com base logarítmica, e com as variáveis independentes área total construída, quantidade de funcionários e volume horário de tráfego da rua adjacente.

2.6.9 Polos de uso misto e polos de uso múltiplo

Existem divergências entre as definições de desenvolvimento Multiuso e Polos Múltiplos Geradores de Viagens (PMGVs), principalmente entre as usadas nos estudos americanos e as brasileiras.

Conforme o ITE (2004), um desenvolvimento Multiuso é tipicamente um projeto imobiliário único que consiste em dois ou mais usos do solo segundo classificação do próprio ITE, no qual as viagens entre os vários usos do solo são realizadas internamente, ou seja, nas vias dentro do próprio local.

Enquanto que no Brasil, surgiu o conceito defendido por Goldner et al. (2010) de que PMGVs são os conjuntos de estabelecimentos comerciais ou de serviços próximos entre si que formam um agrupamento de atividades, sendo que as viagens são facilitadas pela proximidade entre os diferentes usos do solo, podendo utilizar tanto o sistema viário interno quanto o externo.

Nesse subitem, cabe ressaltar ainda que desenvolvimento Multiuso/ PMGVs diferem de desenvolvimento de Polos Geradores de Viagens de Uso Misto (PGMs). Portugal (2012) acredita que: a interpretação de PGMs está mais direcionada à composição, em um único edifício, enquanto que a interpretação de Multiuso/ PMGVs está ligada a composição, dentro de um mesmo espaço delimitado, podendo ser uma quadra, quarteirão ou subcentro, de diversos edifícios com diversas atividades integrantes (PORTUGAL, 2012).

Apesar de não haver um consenso entre os diferentes pesquisadores a respeito desses conceitos, o ITE (2004) publicou um método de análise para os mesmos.

Por outro lado, no Brasil estudos na área ainda estão sendo desenvolvidos, como é o caso: do estudo de Silva & Freitas (2011) que fez um método para determinar a área de influência desses PGVs; de Cavalcante (2002) que fez um modelo de previsão de viagens para esses PGVs através de regressão linear múltipla e utilizando as variáveis independentes o número de grupos de usos considerados para cada iteração segundo os critérios de classificação especificados e variável *dummy* de ajuste do modelo, no caso, renda média da população fixa e flutuante; e de Goldner et al. (2010) que fez entrevistas para obter características das viagens para esses PGVs.

Quanto à estimação de viagens, em resumo, tem-se que o número de viagens em um PMGV não é simplesmente a soma de viagens do empreendimento A com o empreendimento B, mas sim o somatório das viagens de cada empreendimento menos as viagens feitas no seu interior, que atenuam o número de viagens geradas no sistema viário externo de acesso ao desenvolvimento como um todo (SILVA & FREITAS, 2011). Essa taxa interna da captação pode ser definida como o percentual de redução aplicada às estimativas da geração de viagens para uso de solos individuais e explicam as viagens internas ao local (PORTUGAL, 2012).

Não foram encontrados estudos a partir de 2013 para polos de uso misto e polos de uso múltiplo na revisão bibliográfica sistemática.

2.6.10 Polos geradores de viagens de cargas

Segundo Marra (1999), a movimentação de carga envolve uma grande variedade de atores e, de acordo com Gasparini (2008), de fatores – uso do solo, variedade de mercadorias, de diferentes tipos e tamanhos, locais para carga e descarga, tipos de veículos, seus movimentos, fiscalização de tais movimentos pelas autoridades responsáveis e características da rede viária – que influenciam em seu aspecto operacional. Quando essa movimentação ocorre em área urbana, Marra (1999) ainda afirma que a mesma se torna mais complexa.

Conforme Gasparini (2008), a movimentação de carga urbana pode estar concentrada nas áreas do entorno de PGVs, como, por exemplo, em lojas de departamento, supermercados, restaurantes, terminais de carga, etc.

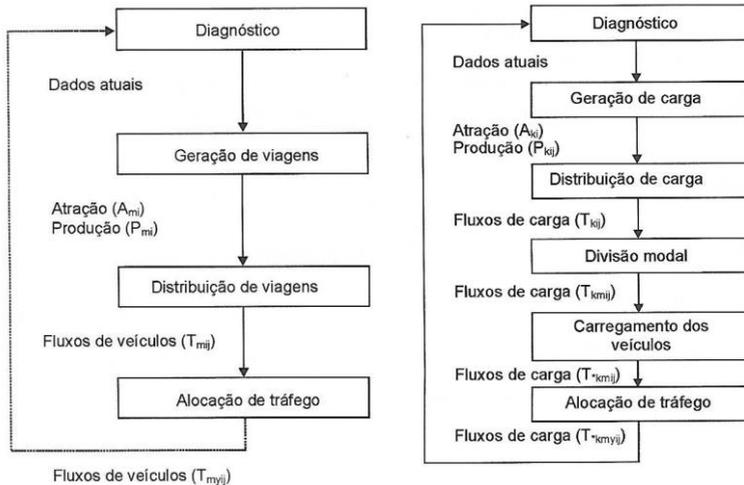
De acordo com Marra (1999), para PGVs urbanos de carga, os modelos mais utilizados para previsão de demanda podem ser baseados em deslocamento de carga (Figura 2.10a) ou em viagens de veículos (Figura 2.10b).

A diferença principal é que o modelo baseado em viagens pressupõe que a seleção do modo de transporte (m) já foi realizada e, portanto, não se faz necessário efetuar a etapa de divisão modal, ao contrário do modelo baseado em volume de carga (PORTUGAL, 2012).

Ambos os modelos iniciam com o diagnóstico, o qual fornece as informações do ano-base. O procedimento segue para a etapa de geração de viagens ou de cargas, no qual é necessário definir: as taxas de geração de viagens ou o modelo de regressão específico para cada zona (i) por

modo de transporte (m), para o caso do modelo baseado em viagens; ou as taxas de geração de carga ou o modelo de regressão específico para cada zona (i), desagregado por tipo de carga (k), para o caso do modelo baseado em volume de carga. O resultado dessa etapa alimenta a distribuição de viagens, de modo a obter o fluxo de viagens ou de cargas (PORTUGAL, 2012).

Figura 2.10 – Modelo de quatro etapas adaptado para carga: abordagem com base a) em viagens e b) em volume de carga



Fonte: Hoeguin-Veras e Thorson (2000) apud Portugal (2012).

Na etapa de divisão modal, a qual, conforme supracitado, só ocorre para o modelo baseado em volume de carga, é distribuído o volume de carga obtido na matriz origem-destino entre os modos de transporte contemplados no estudo, obtendo o fluxo de carga entre zonas, por tipo de carga e por modo de transporte. Após a divisão modal, é necessário, ainda, converter os volumes de carga em viagens de veículos, de acordo com a densidade da carga e a capacidade do veículo (PORTUGAL, 2012).

A última etapa para ambos os modelos trata da alocação de tráfego, que aloca, a cada segmento da rede de transporte, os fluxos verificados na matriz origem-destino (PORTUGAL, 2012).

Em relação às vantagens de cada modelo, dos baseados em viagens destacam-se: os dados de tráfego são relativamente fáceis de obter; um

crescente número de sistemas inteligentes de transporte é capaz de seguir os movimentos dos veículos, constituindo uma fonte de dados de tráfego; e as viagens vazias já são consideradas no modelo. Já os modelos baseados em volume de carga têm como vantagens: permitir captar com mais precisão os mecanismos fundamentais que direcionam a movimentação das cargas, que são determinados por seus atributos, como forma e peso por unidade (PORTUGAL, 2012).

Quanto, especificamente, aos modelos de geração de viagens para PGVs de carga, a Tabela 2.4 (a) e (b) apresenta uma síntese de uma seleção de estudos, nacionais e internacionais, respectivamente, sobre o tema, feita por Portugal (2012), destacada pelo seu enfoque: se para gerenciamento de tráfego ou para planejamento estratégico.

Os modelos desenvolvidos com a finalidade de fornecer subsídio para o gerenciamento de tráfego geralmente buscam verificar o número de viagens de veículos de carga geradas pelo PGV por meio de variáveis independentes como a área construída do empreendimento analisado ou de outras variáveis que forneçam indícios sobre o tamanho do empreendimento, como número de funcionários, número de clientes ou, ainda, número de vagas de estacionamento (PORTUGAL, 2012).

Nos modelos desenvolvidos com a finalidade de fornecer subsídio ao planejamento estratégico, deve-se elaborar modelos que forneçam informações que possam ser utilizadas como entrada para o modelo quatro etapas. Nesse caso, o enfoque é mais amplo, buscando-se analisar não apenas o PGV e seu entorno, mas a região em que este se insere (PORTUGAL, 2012).

Desse modo, é comum que tais modelos busquem informações mais desagregadas, utilizando como variável dependente o volume de carga transportada total por tipo de produto ou o volume de viagens totais, por tipo de veículo ou por tipo de atividade, variando conforme o objetivo do estudo; e como variável independente as características relacionadas à infraestrutura do empreendimento, como as citadas anteriormente, ou ainda variáveis que possam estimar o potencial de demanda da região, como, por exemplo, rendimento médio per capita, número de residências, população por zona, etc. (PORTUGAL, 2012).

Tabela 2.4 – Síntese dos modelos de geração de viagens selecionados para carga a) nacionais

(continua)

Referência	Enfoque	Tipo de modelo	Variável dependente		Variável independente		Modelo matemático utilizado	
			Grupo	Variável	Grupo	Variável	Tipo	Modelo
Gasparini (2008) e Silva & Waisman (2007)	Gerenciamento de tráfego	Baseado em viagens	Volume de viagens	Totais	Espaço ocupado	Área construída	Regressão linear simples	$Y = Ax_1 + B$
				Na hora de pico		Área bruta locável		
						Vagas no estacionamento		
						Área de vendas		
				Fora da hora de pico	Usuários do sistema	Volume de clientes	Regressão linear múltipla	$Y = Ax_1 + Bx_2 + C$
						Número de funcionários		

Tabela 2.4 – Síntese dos modelos de geração de viagens selecionados para carga a) nacionais

(continuação)

Referência	Enfoque	Tipo de modelo	Variável dependente		Variável independente		Modelo matemático utilizado	
			Grupo	Variável	Grupo	Variável	Tipo	Modelo
Melo (2002)	Estratégico	Baseado em viagens	Volume de viagens	Por atividades	Espaço ocupado	Área construída	Regressão linear simples	$Y = Ax + B$
					Usuários do sistema	Carga horária trabalhada		
					Distância percorrida			
					Outros	Carga própria da empresa	Regressão linear múltipla	$Y = Ax1 + Bx2 + Cx3 + Dx4 + E$
						Número de veículos utilizados		
Marra (1999)		Baseado em volume de carga	Volume de carga	Total de carga no período	Espaço ocupado	Área construída	Regressão linear simples	$Y = Ax1 + B$
				Usuários do sistema	Número de funcionários			
				Total de carga no período por bairro	Outros	Rendimento médio (per capita)	Ajuste exponencial	$Y = A^{x1}$

Tabela 2.4 – Síntese dos modelos de geração de viagens selecionados para carga b) internacionais

(conclusão)

Referência	Enfoque	Tipo de modelo	Variável dependente		Variável independente		Modelo matemático utilizado	
			Grupo	Variável	Grupo	Variável	Tipo	Modelo
Allen (2002)				Por tipo de caminhão	Espaço ocupado	Área construída	Regressão linear simples	$Y = Ax1 + B$
Iding et al. (2002)		Baseado em viagens	Volume de viagens	Por setor	Usuários do sistema	Área total	Regressão linear múltipla	$Y = Ax1 + Bx2 + C$
Ogden (1977 e 1992)				Por categoria		Número de funcionários		$Y = Ax1 + Bx2 + Cx3 + Dx4 + E$
Tadi & Baldach (1994)	Estratégico			Por tipo de uso do solo		Número de residências/população	Taxa	$Y = Axi$
Black (1999)		Baseado em volume de carga	Volume de carga	Por tipo de produto	Usuários do sistema	Número de funcionários	Regressão linear simples	$Y = mAx1 + B$
Ogden (1977)						Número de residências/população	Regressão linear múltipla	$Y = Ax1 + Bx2 + C$
-						Outros	Receita bruta por setor	

Fonte: Portugal (2012).

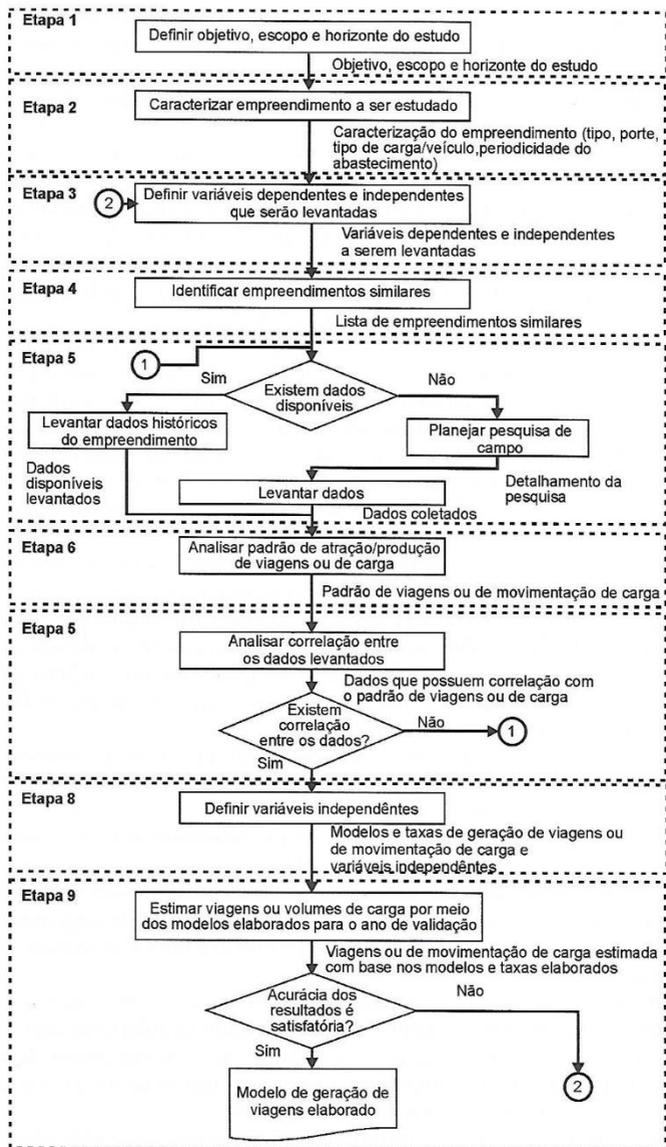
Percebe-se também, pela Tabela 2.4, que uma característica da modelagem da geração de viagens para o PGV de carga está na necessidade de estabelecer diferentes equações ou taxas para diferentes tipos de carga e classes de veículos. Consequentemente, pode-se ter diversas equações, até mesmo utilizando técnicas diferentes, para um mesmo par de regiões dependendo do número de tipos de carga e classes de veículos. Essa situação demonstra a complexidade de se modelar a geração de viagens de carga (PORTUGAL, 2012).

Entre todas as técnicas utilizadas para a construção dos diversos modelos analisados, destacou-se a regressão linear simples ou múltipla. Portugal (2012) acredita que essa possa ser uma indicação de que é possível obter resultados satisfatórios com modelos simples, que utilizam poucas variáveis independentes, uma vez que a compreensão dos mecanismos que levam à escolha das variáveis muitas vezes é mais importante do que a complexidade da função matemática associada ao modelo (PORTUGAL, 2012).

Por fim, cabe ressaltar que os modelos de geração de viagens são muitos, e variam de acordo com o tipo do empreendimento, porte, atividade, localização, etc.; sendo que muitas vezes os modelos existentes não se mostram adequados ao PGV em estudo, sendo necessário elaborar um modelo específico que permita uma avaliação mais adequada. Desse modo, Portugal (2012) apresenta um procedimento (Figura 2.11) que tem por objetivo auxiliar profissionais que necessitam desenvolver um modelo com potencial de apresentar maior aderência ao PGV de carga que se pretende estudar.

O procedimento (Figura 2.11) se inicia, na primeira etapa, com a avaliação do objetivo do estudo a ser realizado – se o mesmo tem como enfoque o gerenciamento do tráfego de entorno ou o planejamento estratégico da região – de maneira a subsidiar a delimitação da área de estudo e a identificação das variáveis que poderão ser utilizadas na construção do modelo; assim como com a delimitação do escopo e do horizonte do estudo, de maneira a identificar: se apenas o PGV em questão será avaliado ou se também é necessário caracterizar o seu entorno, qual o período para levantamento de dados e qual o horizonte para a projeção dos dados (PORTUGAL, 2012).

Figura 2.11 – Procedimento para elaboração de modelo de geração de viagens ou volume de carga



Fonte: Portugal (2012).

Na segunda etapa, deve-se caracterizar o empreendimento a ser estudado, buscando conhecer qual o tipo do PGV, seu porte, os tipos de carga e/ou veículos de carga que atrai/produz e a periodicidade de seu abastecimento. O resultado dessa etapa subsidia a seleção dos PGVs similares a serem utilizados para a elaboração do modelo (PORTUGAL, 2012).

Na terceira etapa, deve-se definir variáveis dependentes e independentes que serão levantadas em empreendimentos já existentes; e na quarta etapa, deve-se identificar PGVs existentes parecidos ao que se pretende analisar (PORTUGAL, 2012).

Na quinta etapa, a partir da seleção dos PGVs similares ao que se deseja estudar, deve-se levantar os dados para elaboração de um modelo de geração de viagens. Se existirem dados disponíveis, os mesmos podem ser levantados via consulta a documentos públicos ou cedidos pela administração dos PGVs, caso não existam dados disponíveis, os mesmos devem ser levantados a partir de pesquisa de campo. Para a realização da pesquisa de campo, deve-se definir qual o tipo de instrumento de coleta que será utilizado (questionário, entrevista, observação, etc.), o tamanho da amostra, o período para a coleta de dados e os dados que serão levantados (PORTUGAL, 2012).

Na sexta etapa, de posse dos dados referentes ao conjunto de PGVs selecionados, deve-se analisá-los, buscando conhecer o padrão de atração/ produção de viagens ou de movimentação de carga, conforme objetivo definido na primeira etapa, observando-se a ocorrência de períodos de sazonalidade ou tendências (PORTUGAL, 2012).

Na sétima etapa, deve-se analisar a correlação existente entre as variáveis independentes e dependentes. Caso seja verificado que essa correlação é baixa, sugere-se retornar à quinta etapa – levantamento de dados (PORTUGAL, 2012).

Na oitava etapa, deve-se definir as variáveis que serão utilizadas. Recomenda-se que a verificação da aderência dos modelos deva ser iniciada pelos modelos mais simples, como regressão linear simples, para que se possa selecionar as variáveis que melhor explicam o comportamento das viagens ou da movimentação de cargas. Em seguida, pode-se desenvolver modelos que relacionem duas variáveis bem correlacionadas (PORTUGAL, 2012).

Por último, na nona etapa, deve-se testar o modelo por meio de comparação entre os resultados estimados para um período conhecido e os dados reais. Para fazer essa verificação, pode-se utilizar os mesmos

PGVs selecionados para a elaboração do modelo ou outro PGV, desde que existam dados para isso. Quanto maior for o número de PGVs em que for feita essa verificação, melhor. Caso os resultados sejam validados estatisticamente, tem-se um modelo de geração de viagens ou volume de carga que pode ser utilizado para a realização de projeção de viagens ou volume de carga para PGVs similares. Caso contrário, deve-se voltar à terceira etapa – definição das variáveis –, de modo a verificar se são necessárias outras informações sobre o PGV, se os dados levantados foram suficientes, e se a escolha das variáveis foi adequada (PORTUGAL, 2012).

Em relação à revisão bibliográfica sistemática foram identificados os seguintes trabalhos com foco em modelos de geração de viagens para PGVs de carga, exceto portos ou terminais marítimos, a partir de 2013: Díaz (2017), Ducret & Gonzalez-Feliub (2016) e Mommens, van Lier & Macharis (2016).

Os dois primeiros trabalhos utilizaram a técnica de regressão linear múltipla para elaborarem seus modelos e o último trabalho utilizou a técnica de regressão linear com base logarítmica. As variáveis independentes utilizadas foram: área do estabelecimento, número de empregados por estabelecimento, variáveis binárias para o tipo do estabelecimento (perecíveis, alimentação, saúde e serviços, serviços públicos e educação), número de entregas e volume de mercadorias.

2.7 ANÁLISE DA REVISÃO BIBLIOGRÁFICA

Na maioria dos estudos analisados por Portugal (2012) para os diversos tipos de PGVs – shoppings centers, equipamentos institucionais, estabelecimentos residenciais, hotéis, estações metroferroviárias, estabelecimentos de ensino, hospitais, supermercados, polos multiuso e PGVs de carga – notou-se uma forte influência dos estudos do ITE (2004; 2008), que foram os percussores em estudos nessa área. Isso porque, todos os estudos desenvolveram modelos ou taxas de geração de viagens, utilizando, principalmente a técnica de regressão linear simples ou múltipla.

O que pode ter ocorrido é que, como foram embasados no ITE, esses estudos optaram inicialmente por essa técnica, e como a mesma se mostrou estatisticamente válida, com resultados satisfatórios, além de ser

simples de aplicar e interpretar, não foi necessário o desenvolvimento de modelos matemáticos mais complexos para esses tipos de PGVs.

Outro ponto a ser destacado é que a definição das variáveis a serem utilizadas nos modelos foi de suma importância nesses estudos e podem ter colaborado no desenvolvimento de modelos com boa aderência. As variáveis independentes mais utilizadas para explicar o volume de viagens gerados por um PGV se relacionaram à área construída do empreendimento ou à demanda do mesmo – nº de unidades residenciais para estabelecimentos residenciais, nº de quartos/empregados para hotéis, nº de usuários do sistema metroferroviário, nº de alunos/empregados para estabelecimentos de ensino e nº de leitos para hospitais.

Os modelos de geração de viagens existentes para PGVs são muitos, e variam de acordo com o tipo do empreendimento, porte, atividade, localização, etc.; por isso taxas e modelos de geração de viagens já propostos para PGVs devem ser utilizados com cuidado em PGVs em condições distintas quando os mesmos não possuem estudos sobre a realidade local, visto que os mesmos podem superestimar ou subestimar a geração de viagens para o PGV a ser analisado. Ressalta-se, no entanto, que muitas vezes não há modelos existentes que possuam características similares ao PGV que se deseja estudar, sendo necessário elaborar um modelo específico, que compreendam as especificidades locais e sua atualidade, para uma avaliação mais adequada de seus impactos.

Especificamente em relação aos PGVs de carga, foi destacada a complexidade de se estudar tais PGVs, uma vez que são muitos fatores que influenciam no seu aspecto operacional. Apesar disso, embora possua algumas especificidades no seu procedimento para construção de modelos de geração de viagens, a maioria dos estudos supracitados utilizaram a técnica de regressão linear e variáveis independentes relacionadas à área construída do empreendimento ou à demanda, como o número de funcionários.

Ressalta-se, no entanto, que nenhum dos estudos analisados abordou o PGV portos marítimos, sendo que a maioria abordou movimento de caminhões em área urbana de forma geral ou, ainda, abordou como PGVs de carga bares e restaurantes, e firmas. Acredita-se que, entre os PGVs de carga, os portos marítimos inseridos em áreas urbanas são ainda mais complexos no que diz respeito à construção de um modelo de geração de viagens, uma vez que os mesmos são responsáveis por um volume muito maior que os outros PGVs de carga de transportes

de mercadorias e há maiores variações dos fatores que influenciam em seu aspecto operacional.

3 MODELOS DE GERAÇÃO DE VIAGENS PARA CAMINHÕES EM PORTOS MARÍTIMOS

Esse Capítulo trata sobre a aplicação do procedimento de revisão bibliográfica sistemática apresentado na Figura 1.1, assim como os resultados encontrados nesse levantamento, no que diz respeito aos modelos de geração de viagens para caminhões em portos marítimos.

3.1 PLANEJAMENTO DA REVISÃO BIBLIOGRÁFICA

O primeiro passo do procedimento da revisão bibliográfica, que é o planejamento da revisão, é descrito a seguir.

3.1.1 Identificação da necessidade da revisão

Conforme citado no Capítulo 1, para solucionar problemas de congestionamentos terrestres no acesso dos portos marítimos, é necessário conhecer o fluxo de tráfego da área portuária para que, assim, se faça um planejamento da rede e gestão de tráfego eficiente. Nesse contexto, estudar modelos de geração de viagens para caminhões em portos marítimos pode contribuir na solução dessa problemática.

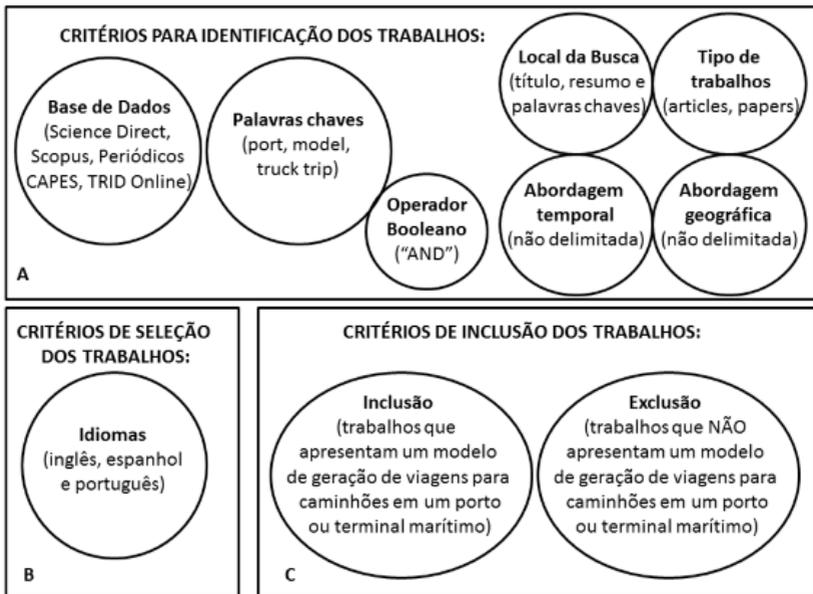
3.1.2 Elaboração da proposta da revisão

A finalidade dessa revisão bibliográfica sistemática foi identificar estudos de modelos de geração de viagens para caminhões em portos marítimos, destacando a técnica utilizada e as variáveis utilizadas, e os resultados encontrados, além das vantagens e desvantagens da utilização dessas técnicas.

3.1.3 Desenvolvimento do protocolo de revisão

Nesse trabalho, o protocolo de revisão empregado foi o exposto na Figura 3.1.

Figura 3.1 – Protocolo da revisão bibliográfica sistemática desse trabalho



Fonte: Elaboração própria com base em Oliveira, 2015.

Em resumo:

- g) foram utilizadas as bases de dados Science Direct, Scopus, Periódicos Capes e Trid Online;
- h) foram usadas as palavras-chaves “*port*”, “*truck*” e “*trip generation models*”, ou seja, “porto”, “modelos de geração de viagens” e “caminhões” e suas combinações por meio do operador booleano “*and*”, ou seja, “e”;
- i) foram considerados os estudos do tipo “*papers*”, ou seja, “artigos” disponíveis na base de dados e a investigação das palavras chaves foi feita no título, resumo e palavras-chaves desses estudos;
- j) não se limitou a busca a alguma abrangência geográfica ou temporal;
- k) foram selecionados apenas os trabalhos que se encontravam em língua inglesa, espanhola ou portuguesa; e

- l) após selecionados os artigos, para a inclusão e exclusão deles no trabalho foi considerado se eles atingiam o objetivo principal ou secundário dessa tese, ou seja, mostravam a construção de um modelo de geração de viagens para caminhões em algum porto ou terminal marítimo.

Os resultados dessa coleta podem ser vistos no item 3.2 a seguir.

Cabe ressaltar que não foi utilizado o operador booleano “ou”, visto que resultava em um número grande de trabalhos que não atingiam o objetivo dessa revisão sistemática. Ainda, consideraram-se apenas artigos, visto que, de acordo com Nord et al. (1995) e Ngai & Wat (2002), profissionais e pesquisadores utilizam os periódicos científicos para adquirir informações e divulgar novos resultados, o que justifica essa escolha. Também se optou por não restringir a busca a alguma abordagem temporal ou geográfica, de modo que se pretendia observar quais países estavam estudando o tema e desde que ano o mesmo vem sendo investigado. Por fim, para a seleção dos artigos apenas se levou em conta o idioma em que se encontravam, uma vez que os periódicos constantes nessas bases de dados são de boa qualidade e optou-se por não limitar a nenhuma classificação de Qualis dos mesmos.

Para se chegar nesse protocolo, se realizaram buscas preliminares nessas mesmas bases de dados considerando a combinação por meio do operador booleano “e” de duas ou três das palavras chaves utilizadas, além de outras como: “*marine terminal*”, “*trip generation*”, “*truck traffic*” e “*freight modeling*”, ou seja, terminal marítimo, gerador de viagens, tráfego de caminhões e modelagem de transporte de mercadorias. Apesar de algumas dessas buscas derivarem um maior número de trabalhos selecionados, ao se fazer a avaliação preliminar dos mesmos se chegava a um número igual ou até menor de trabalhos incluídos na tese pelo protocolo escolhido.

3.2 REALIZAÇÃO DA REVISÃO BIBLIOGRÁFICA

Com o planejamento concluído, deu-se sequência ao procedimento com a etapa de realização do mesmo conforme apresentado a seguir.

3.2.1 Identificação e seleção dos trabalhos

Os trabalhos foram identificados e selecionados segundo os critérios descritos no subitem 3.1.3, na página 83. Ao todo, foram identificados:

- a) dois trabalhos na base de dados Science Direct: (i) Roberts et al. (2014); e, (ii) Hartman & Clott (2015);
- b) 25 na base de dados Scopus: (i) Hartman & Clott (2015); (ii) Qian, Zhu & Xie (2014); (iii) Roberts et al. (2014) (iv) Islam, Olsen & Ahmed (2013); (v) Chen, Langevin & Lu (2013); (vi) Kappanna et al. (2012); (vii) Protopapas, Kruse & Olson (2011); (viii) Chu (2011); (ix) Xie & Huynh (2010a); (x) Xie & Huynh (2010b); (xi) Lee et al. (2009); (xii) Rey, Muñoz & Weintraub (2009); (xiii) Rahimi, Asef-Vaziri & Harrison (2008); (xiv) Stephens (2006); (xv) Sarvareddy et al. (2005); (xvi) Zietsman & Perkinson (2005); (xvii) Klodzinski et al. (2004); (xviii) Klodzinski & Al-Deek (2004); (xix) Lazic & Morian (2004); (xx) Klodzinski & Al-Deek (2003); (xxi) Holguín-Veras, López-Genao & Salam (2002); (xxii) Al-Deek (2002); (xxiii) Al-Deek (2001a); (xxiv) Al-Deek (2001b); (xxv) Al-Deek et al. (2000);
- c) dois trabalhos na base de dados Periódicos Capes: (i) Chu (2011); e, (ii) Al-Deek (2001a); e
- d) 15 trabalhos na base de dados Trid Online: (i) Hartman & Clott (2015); (ii) You & Ritchie (2012); (iii) Protopapas, Kruse & Olson (2012); (iv) Chu (2011); (v) Protopapas, Kruse & Olson (2011); (vi) Rahimi, Asef-Vaziri & Harrison (2008); (vii) Sarvareddy et al. (2005); (viii) Agnello, Jonnes & Allen (2004); (ix) Klodzinski & Al-Deek (2004); (x) Klodzinski et al. (2004); (xi) Klodzinski & Al-Deek (2003); (xii) Al-Deek (2002); (xiii) Luk et al. (2002); (xiv) Al-Deek (2001a); e, (xv) Al-Deek (2001b).

No entanto, desconsiderando os trabalhos que foram encontrados em mais de uma base, chegou-se ao total de 29 trabalhos identificados.

Como todos estavam no idioma inglês ou espanhol, todos foram selecionados para a avaliação.

3.2.2 Avaliação dos trabalhos selecionados

Os resumos dos 29 trabalhos encontrados no subitem anterior do presente item foram lidos com o objetivo de verificar se atendiam os critérios de inclusão, ou seja, se eles mostravam a construção de um modelo de geração de viagens para caminhões em algum porto ou terminal marítimo. Caso não tenha ficado claro pela leitura do resumo se eles atingiam ou não esse objetivo, esses trabalhos foram lidos na íntegra para que pudessem ser avaliados.

Dos 29 documentos selecionados, 14 foram incluídos na pesquisa: (i) Chu (2011); (ii) Xie & Huynh (2010a); (iii) Xie & Huynh (2010b); (iv) Sarvareddy et al. (2005); (v) Klodzinski et al. (2004); (vi) Klodzinski & Al-Deek (2004); (vii) Klodzinski & Al-Deek (2003); (viii) Holguín-Veras, López-Genao & Salam (2002); (ix) Al-Deek (2002); (x) Al-Deek (2001a); (xi) Al-Deek (2001b); (xii) Al-Deek et al. (2000); (xiii) Agnello, Jonnes & Allen (2004); e, (xiv) Luk et al. (2002); e, portanto, tiveram dados e informações extraídos no subitem 3.2.3 desse item a seguir de modo a subsidiar o restante dessa pesquisa.

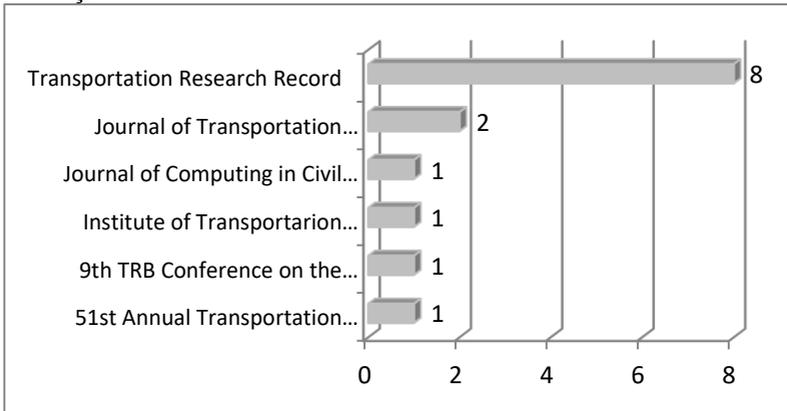
3.2.3 Extrair dados e informações

Nessa fase, os trabalhos incluídos na revisão bibliográfica foram classificados de acordo com os periódicos científicos do qual fazem parte, o ano de sua publicação, a sua abordagem regional, temática, e metodológica.

No caso dos termos utilizados para identificação dos trabalhos, como no levantamento foi utilizado três palavras-chaves e o operador booleano “e”, todos os documentos identificados e selecionados possuíam essas palavras: “porto”, “modelo” e “viagens de caminhões”.

Em relação à distribuição dos trabalhos segundo suas fontes de publicação, 8 (67%) deles foram publicados pela *Transportation Research Record*, dois (17%) pelo *Journal of Transportation Engineering*, um (8%) pelo *Journal of Computing in Civil Engineering* e um (8%) pelo *51st Annual Transportation Research Forum* (Figura 3.2).

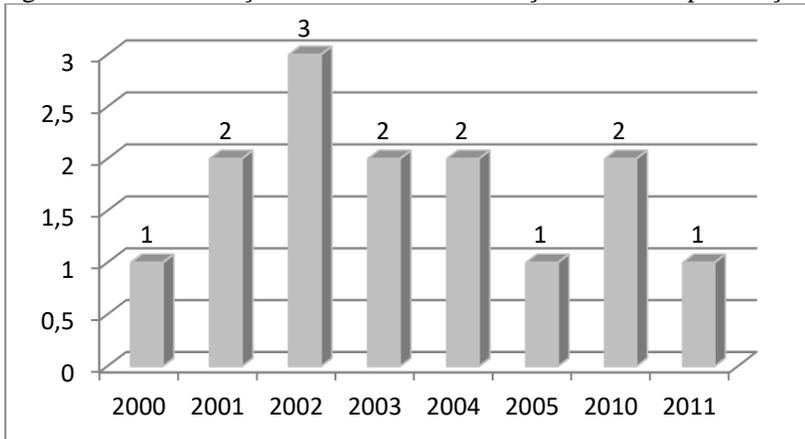
Figura 3.2 – Distribuição dos trabalhos em relação à sua fonte de publicação



Fonte: Elaboração própria com base em Oliveira, 2015.

Quanto à abordagem temporal, percebe-se que, segundo essa revisão, o tema foi abordado pela primeira vez com esse foco em um artigo em 2000, com pico em 2002 (Figura 3.3).

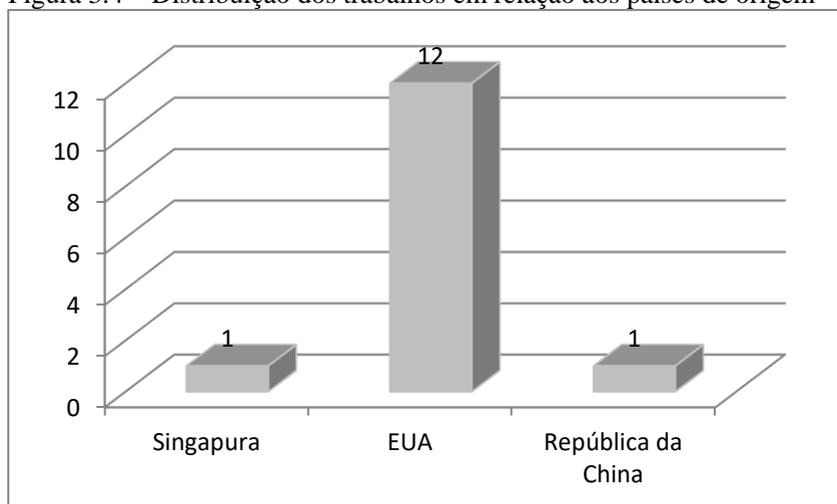
Figura 3.3 – Distribuição dos trabalhos em relação ao ano de publicação



Fonte: Elaboração própria com base em Oliveira, 2015.

No que se concerne à distribuição dos trabalhos em relação aos países de procedência, todos se originaram dos EUA, com exceção de um dos trabalhos que é teve como origem a República da China, e de outro que é de origem de Singapura (Figura 3.4). Destaca-se que não foram identificados trabalhos com origem em outros continentes segundo os critérios da revisão sistemática apresentados no item 3.1.3, na página 83. Logo, pressupõe-se que estudos com esse foco ainda não foram concluídos em outros continentes e/ou que ainda não foram publicados artigos no nível de qualidade especificado nessa revisão bibliográfica sistemática.

Figura 3.4 – Distribuição dos trabalhos em relação aos países de origem

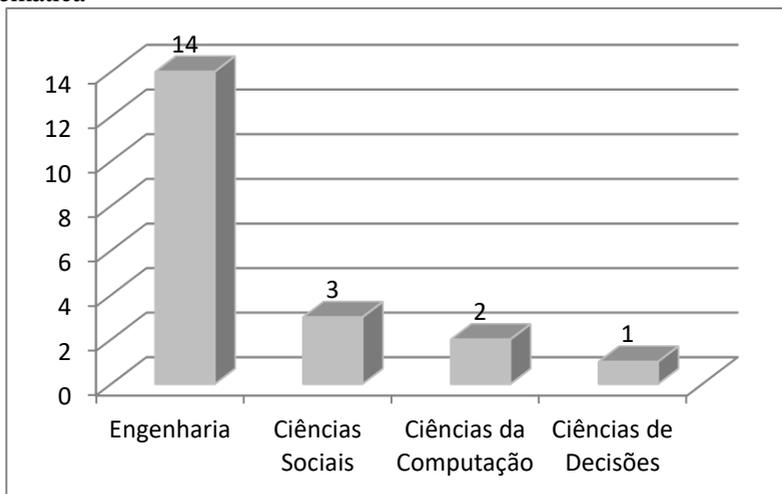


Fonte: Elaboração própria com base em Oliveira, 2015.

Sobre a abordagem temática, todos os artigos (100%) se encontram relacionados ao tema Engenharia. Como os artigos podem ser classificados em mais de uma temática, 21% dos artigos foram classificados, além da temática de Engenharia, também no tema Ciências Sociais, 14%, no tema Ciências da Computação e 7%, no tema Ciências da Decisão (Figura 3.5).

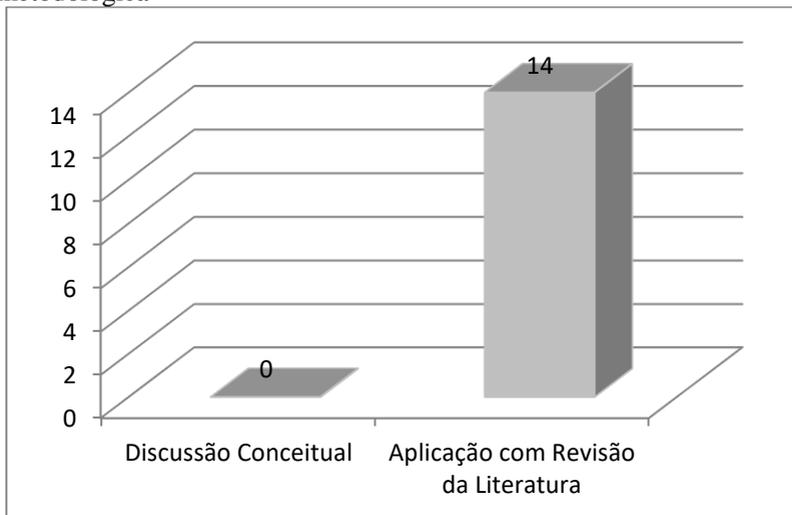
No que tange à abordagem metodológica, todos os trabalhos (100%) realizam uma aplicação dos conceitos pesquisados e apresentaram uma seção de revisão bibliográfica (Figura 3.6).

Figura 3.5 – Distribuição dos trabalhos em relação à abordagem temática



Fonte: Elaboração própria com base em Oliveira, 2015.

Figura 3.6 – Distribuição dos trabalhos em relação à abordagem metodológica



Fonte: Elaboração própria com base em Oliveira, 2015.

Por fim, cabe ressaltar que como a publicação mais difundida e utilizada para os cálculos de previsão de viagens em PGVs é a Trip Generation do ITE, e as taxas e modelos de geração de viagens propostos nesta publicação são utilizadas, não apenas nos EUA, origem da publicação, como também em vários outros países, quando os mesmos não possuem estudos sobre a realidade local; esse documento também foi incluído na pesquisa.

Além disso, de modo a ampliar o escopo da pesquisa se fez uma nova identificação de trabalhos com os critérios utilizados nessa revisão bibliográfica, só que utilizando como base de dados anais de congressos dos quais a autora teve acesso: Associação Nacional de Pesquisa e Ensino em Transportes (Anpet) de 2004 a 2012, Planejamento Urbano, Regional, Integrado e Sustentável (Pluris) 2008, Pan-Americano de Engenharia de Transportes e Trânsito (Panam) 2008, 2012 e 2014 e TRB 2011. Com isso, identificou-se e incluiu-se na pesquisa mais um trabalho, o de Pereyra, Rubinstein & Pérez (2008), de interesse por utilizar como área de estudo um porto na América do Sul, mais especificamente no Uruguai.

3.2.4 Síntese dos estudos

A síntese dos estudos é apresentada no item 3.5, referente à análise da revisão bibliográfica.

3.3 COMUNICAÇÃO E DIVULGAÇÃO DOS RESULTADOS DA REVISÃO

Por fim, com os resultados das etapas anteriores em mãos, esse documento final foi redigido e submetido à banca examinadora com o objetivo de obter o grau de Doutora em Engenharia Civil; resultando, portanto, na publicação de uma tese de doutorado e em um artigo científico baseado nela.

3.4 DESCRIÇÃO DOS ESTUDOS ENCONTRADOS

Detalhes sobre os trabalhos 14 incluídos na pesquisa por meio da revisão bibliográfica sistemática, mais a publicação do ITE (2012) e o

artigo de Pereyra, Rubinstein & Pérez (2008), podem ser visualizados a seguir de acordo com a localização para o qual foram desenvolvidos seus modelos de geração de viagens para caminhões.

3.4.1 Estudos desenvolvidos para todo os EUA

O *Trip Generation*, versão mais atualizada – 2012 – do ITE, apresenta para o uso do solo número 010, correspondente ao porto/terminal marítimo, um modelo de regressão linear simples utilizando como variável dependente o número de viagens de veículos nos acessos aos portos/terminais marítimos, e como variável independente o número de berços do terminal ou a área em acres, resultando, respectivamente, em um modelo com taxa média de viagens de 171,52 por berço de atracação e 11,93 por acre, ambos por dia útil da semana, e com R^2 de 0,58 e 0,93 conforme mostra a Tabela 3.1.

Tabela 3.1 – Modelos de geração de viagens do ITE para portos marítimos

Uso do solo	Nº de estudos	X médio	Unidade ou variável explicativa (X)	Taxa média de viagens geradas por unidade de tempo (intervalo)	Equação (R^2)
010	7	3	Nº de berços	171,52/dia útil da semana (38,60-338,57)	298,56X - 417,40 (0,58)
010	7	47	Nº de acres	11,93/dia útil da semana (4,95-19,47)	18,01X - 287,06 (0,93)

Fonte: adaptado de ITE, 2012.

Holguín-Veras, López-Genao & Salam (2002) também desenvolveram uma pesquisa nacional nos EUA com o intuito de construir modelos de geração de viagens de caminhões para terminais de contêineres. Nessa pesquisa, os autores levantaram dados de 21 terminais de contêineres no País e elaboraram um modelo de regressão linear múltipla considerando como variável dependente o número de viagens de caminhões em um sentido de viagem e, como variáveis independentes a área do terminal, o número de TEUs e o número de caixas de contêineres manuseadas por terminal em um determinado ano.

3.4.2 Estudos desenvolvidos para Baltimore (EUA)

Segundo Agnello, Jonnes & Allen (2004), as entrevistas para obtenção de informações sobre o movimento de caminhões que é utilizada na abordagem tradicional para a elaboração de modelos com taxas de viagens de caminhões, padrões de distribuição e rotas, são de difícil execução e custosas. Considerando isso, eles elaboraram uma abordagem inovadora, rápida e menos custosa para alcançar esse mesmo objetivo por meio de técnicas chamadas de “atribuição adaptável”.

Em Baltimore já existiam estudos sobre modelos de geração de viagens de veículos em geral pelo método de quatro etapas, só que os mesmos estavam defasados e era preciso atualizá-los, levando em consideração o aumento no percentual de caminhões no número total de veículos que circulam pela região.

Ao invés de se gastar recursos fazendo uma nova pesquisa de campo, esses autores propuseram esse novo método chamado de “atribuição adaptável”. Nesse método, um modelo inicial foi criado adaptando parâmetros de outras áreas urbanas, resultando em uma tabela inicial de viagens. Por meio de processos interativos, essa tabela foi adaptada até se chegar o mais próximo possível da tabela de dados reais de contagem. Essa última tabela foi comparada com a tabela inicial de modo a tentar entender as causas de suas diferenças. Então, ajustes foram feitos no modelo inicial para que ele refletisse essas diferenças. Ressalta-se que um dos ajustes feitos diz respeito à identificação de zonas que atraem mais caminhões, como por exemplo, portos marítimos.

3.4.3 Estudos desenvolvidos para a Florida (EUA)

Na Florida, agentes públicos e líderes de empresas demonstravam uma grande preocupação sobre a capacidade da infraestrutura existente nas regiões portuárias, uma vez que a mesma poderia não ser suficiente para acomodar o crescimento do transporte de fretamento em um futuro próximo. Era preciso estimar precisamente a quantidade de movimentos de caminhões nos terminais portuários, no entanto os estudos floridianos existentes não supriam essa demanda.

Tendo isso em vista, Al-Deek passou a desenvolver, sozinho ou em conjunto com outros autores, alguns estudos de modelos de geração de viagens para caminhões aplicados aos portos da Florida.

Em 1998, Al-Deek apud Sarvareddy et al. (2005) desenvolveu um modelo de geração de viagens para caminhões para o porto de Miami. Esse modelo foi baseado na análise de regressão linear simples. O modelo foi restrito a configurações de caminhões reboque e contêineres, porque eles eram o único frete disponível no porto de Miami para o estudo. Também foram desenvolvidos modelos de séries temporais para examinar tendências de variações sazonais e de curto prazo (cinco anos ou menos).

Continuamente, Al-Deek et al. (2000) estudaram o tráfego de frete nos principais portos da Florida. O porto de Miami foi selecionado para a primeira fase do estudo e seus resultados foram publicados pelos autores em um artigo na revista *Transportation Research Record*.

De acordo com esses autores, as abordagens tradicionais de modelagem de fretamento utilizando dados agregados de demografia e economia regional (população, emprego, renda, etc.) não são as mais adequadas para um estudo microscópico, como é o caso de regiões portuárias. Então, eles desenvolveram um novo método para analisar o porto de Miami que consistia nos seguintes passos:

- 1) examinar a malha rodoviária na vizinhança do porto, fazendo observações de campo e revisando informações gerais de tráfego;
- 2) registrar os volumes de tráfego de caminhões e classificações (tipo, número de eixos e configurações) durante diferentes períodos de pico e fora de pico em locais específicos;
- 3) entrevistar os prepostos do porto familiarizados com a operação – isso inclui pessoal da administração e operações de campo, e agentes de transporte, terminais privados, companhias de caminhões, segurança, contabilidade e marketing;
- 4) coletar amostra de dados para elaborar análises preliminares por meio de processo interativo – diferentes fontes de dados foram analisadas para determinar as variáveis independentes a serem utilizadas nos modelos;
- 5) formular modelos para correlacionar o volume de caminhões de fretamento na única e principal rodovia de acesso ao porto

com a atividade portuária – um modelo separado foi desenvolvido para cada direção (entrada e saída do volume de caminhões);

- 6) validar esses modelos para assegurar que ele simula o sistema atual – isso foi realizado por meio da introdução de dados que não foram utilizados durante o processo de formulação do modelo, sendo que os valores estimados do volume de caminhões foram comparados com os volumes medidos;
- 7) examinar tendências sazonais e de longo prazo por meio de análises de séries temporais – análises de séries temporais foram aplicadas no total mensal da principal variável independente (volume de unidades de carga marítimas), baseadas nos dados históricos obtidos para essa variável para o período de 1978-1998;
- 8) determinar a distribuição horária dos movimentos dos caminhões no portão do porto marítimo; e
- 9) interpretar os resultados para estabelecer conclusões e fazer recomendações para análises futuras, assim como fazer sugestões para melhorar as operações de tráfego baseadas na aplicação dos modelos.

Portanto, o modelo inicial para essa pesquisa envolveu certas variáveis baseadas nas informações específicas aprendidas sobre o porto e na disponibilidade de dados. Essas variáveis puderam ser incorporadas no modelo ou retiradas dele durante o processo de análise interativa, de acordo com sua significância com o tráfego de caminhões.

Na análise de regressão linear simples foram consideradas como variáveis dependentes o volume de entrada e saída de caminhões pesados e como variáveis independentes o número de contêineres importados e exportados. Para a construção do modelo ainda foi utilizado um conceito de “grupo de dias”, criado para representar a característica do porto de que os caminhões operam somente de segunda a sexta-feira enquanto os contêineres operam todos os dias da semana, resultando em: o número total de caminhões circulando no porto na segunda-feira pode ser explicado pelo volume total de contêineres na segunda (grupo 1); de terça a quinta, pelo de terça a sexta (grupo 2); e, de sexta-feira, pelo de sábado e domingo (grupo 3).

Vários cenários compuseram as análises de regressões para o desenvolvimento de modelos de entrada e saída de caminhões de sistemas

portuários. Os modelos com maior precisão de acordo com os valores de R^2 , menor percentagem de observações discrepantes, menor valor da razão da Soma dos Quadrados dos Erros (*Sum of Squared Errors* - SSE) pela proporção média, e tamanho da amostra para o modelo de entrada e de saída de caminhões, foram os expostos, respectivamente, nas Equações 1 e 2:

$$\text{INTK}_g = 1,20 (\text{EXPFU}_g) \quad (1)$$

Onde:

INTK_g = volume de caminhões carregados entrando no porto para o grupo g; e

EXPFU_g = número total de unidades de mercadorias exportadas para o grupo g.

$$\text{OUTTK}_g = 310 + 0,7 (\text{IMPFU}_g) \quad (2)$$

Onde:

OUTTK_g = volume de caminhões carregados saindo do porto para o grupo g; e

IMPFU_g = número total de unidades de mercadorias importadas para o grupo g.

Os modelos resultaram em um R^2 de 0,686 para o modelo de entrada de caminhões e 0,784 para o de saída de caminhões.

Um total de 28 pontos estava disponível para cada modelo. Somente vinte pontos foram utilizados para construir o modelo e os restantes (oito pontos) foram utilizados para validar o modelo desenvolvido.

Uma limitação potencial de modelos de regressão linear é exceder a faixa de dados. Portanto, alguns passos foram conduzidos para formular modelos de séries temporais do tipo Auto-Regressivo Integrado de Média Móvel (*Autoregressive Integrated Moving Average* – Arima) para examinar tendências sazonais e de curto/ médio prazo. Modelos de séries temporais baseadas em funções logaritmos naturais das variáveis foram desenvolvidos para cada direção de movimento de fretamento (importação e exportação de unidades de frete) conforme é apresentado, respectivamente, nas Equações 3 e 4:

$$\ln (\text{IMPFU}_m) = 0,0135 + \ln (\text{IMPFU}_{m-1}) - 0,218 [\ln (\text{IMPFU}_{m-9}) - \ln (\text{IMPFU}_{m-10})] \quad (3)$$

Onde:

m = índice do mês;

IMPFU_m = número total de unidades de mercadorias importadas para qualquer mês m ;

IMPFU_{m-1} = número total de unidades de mercadorias importadas para o mês precedendo o mês m ;

IMPFU_{m-9} = número total de unidades de mercadorias importadas para os nove meses precedendo o mês m ; e

IMPFU_{m-10} = número total de unidades de mercadorias importadas para os dez meses precedendo o mês m .

$$\ln (\text{EXPFU}_m) = 0,0128 + \ln (\text{EXPFU}_{m-1}) - 0,18 [\ln (\text{EXPFU}_{m-9}) - \ln (\text{EXPFU}_{m-10})] \quad (4)$$

Onde:

m = índice do mês;

EXPFU_m = número total de unidades de mercadorias exportadas para qualquer mês m ;

EXPFU_{m-1} = número total de unidades de mercadorias exportadas para o mês precedendo o mês m ;

EXPFU_{m-9} = número total de unidades de mercadorias exportadas para os nove meses precedendo o mês m ; e

EXPFU_{m-10} = número total de unidades de mercadorias exportadas para os dez meses precedendo o mês m .

Ressalta-se que como o resultado desses modelos de regressão linear é o volume total de caminhões por grupo de dias é preciso convertê-lo em volume diário de caminhões. Para obter o volume diário de caminhões é preciso multiplicar a esse resultado a percentagem média de movimento de caminhões obtida de distribuições empíricas. Essas distribuições são construídas por meio de dados de volume de caminhões entrando e saindo (para um grupo de dias dentro de uma semana e para cada dia dentro de um grupo específico) levantados no porto em questão.

Al-Deek (2001a) dá sequência a esse trabalho publicando um artigo que compara o modelo de regressão linear simples supracitado com

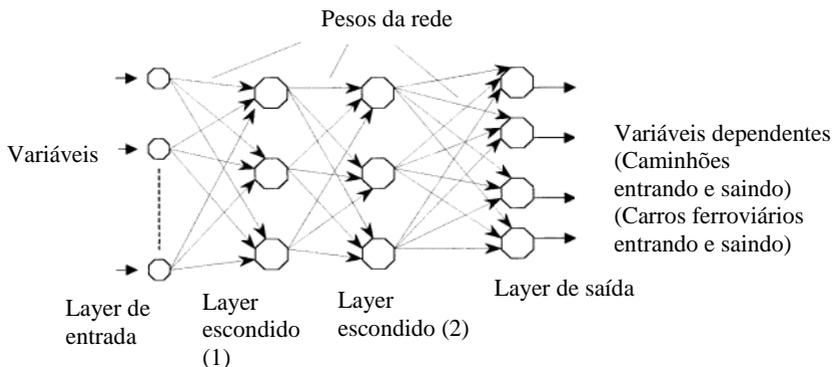
um novo modelo elaborado por meio da técnica de Rede Neural Acíclica Multicamada (*Multilayer Feedforward Neural Network – MLFNN*), com o uso de Algoritmo de Retro-propagação (*Backpropagation Neural Network – BPN*). O objetivo principal desse trabalho foi descobrir qual das duas abordagens seria mais adequada para ser aplicada no porto de Miami, nos EUA.

Na nova abordagem foi utilizada uma rede neural com uma camada de entrada e uma de saída. A camada de entrada contém vários neurônios, que representam as variáveis independentes, enquanto que a camada de saída possui dois neurônios, que são as variáveis dependentes (tráfego diário de caminhões saindo e entrando do porto).

Foram consideradas como variáveis independentes: E = unidades de contêineres exportados; e I = unidades de contêineres importados. Um dos desafios foi correlacionar os sete dias de atividades de carregamento de navios com os cinco dias de tráfego de caminhões no período de uma semana. Como os contêineres podem ser armazenados durante um máximo de três dias no porto de Miami, o tráfego de caminhões em certo dia nesse porto pode ser influenciado pelas atividades marítimas desenvolvidas nos três dias anteriores e posteriores ao analisado. Devido à sazonalidade diária do volume de tráfego, também foi utilizada como variável independente o dia da semana analisado: MO = segunda; TU = terça; WE = quarta; TH = quinta; e, FR = sexta-feira.

Após alguns ensaios, o autor chegou à seguinte configuração de modelo de RNA (Figura 3.7):

Figura 3.7 – Configuração do modelo de RNA para o porto de Miami



A camada de entrada possui 19 neurônios, equivalente as variáveis independentes supracitadas: E-3, E-2, E-1, E0, E1, E2, E3, I-3, I-2, I-1, I0, I1, I2, I3, MO, TU, WE, TH, e, FR; e a camada de saída possui dois neurônios: o volume diário de caminhões saindo e entrando do porto estudado. Foi arbitrado que a primeira e a segunda camada oculta poderiam ter até cinco funções de transferência dentre os três tipos de funções de transferência que podem ser utilizadas: a função sigmóide-logística (log-sig), a função tangente hiperbólica sigmóide (tansing), e a função linear (purelin).

Uma série de experimentos, usando diferentes combinações na quantidade e tipo das funções nessas camadas, foi desenvolvida com a finalidade de encontrar o modelo com menor valor de Média dos Quadrados dos Erros (*Mean of Squared Errors* – MSE). O modelo com menor valor de MSE foi o com duas camadas ocultas, três funções do tipo tansing na primeira camada oculta, duas funções do tipo tansing segunda camada oculta e duas funções lineares na camada de saída.

Testes estatísticos (teste t, teste de normalidade e teste de Kolmogorov-Smirnov) foram realizados acerca da diferença entre os valores estimados pelos modelos e os valores reais encontrados tanto para o modelo de regressão linear quanto para o modelo de RNA. Com base nos resultados, o autor concluiu que a RNA se mostrou mais precisa e, conseqüentemente, mais adequada para ser aplicada no porto de Miami, porém ressaltou que essa técnica necessita de uma amostra grande de dados para obter bons resultados. O autor ainda faz uma aplicação de MLFNN no porto de Jacksonville, nos EUA.

No mesmo ano, Al-Deek (2001b) desenvolveu um método de modelagem de âmbito estadual para o movimento de caminhões pesados de carga em cinco portos da Flórida. Foram desenvolvidos modelos específicos para todos os portos utilizando a técnica MLFNN. Os modelos desenvolvidos foram utilizados para previsões de curto prazo.

Em todos os modelos, as variáveis dependentes consistiam no volume de caminhões entrando e saindo do porto e as variáveis independentes consistiam em atividade de carga no porto (mercadorias importadas e exportadas de navios).

Pesquisas e testes de abordagens de modelagem foram elaborados comparando-se os modelos desenvolvidos com a técnica de RNA com modelos desenvolvidos com a técnica de regressão linear, sendo que, nesse estudo, os modelos de RNAs apresentaram resultados mais próximos aos reais do que modelos de regressão linear caracterizando-se,

portanto, na técnica mais adequada para ser aplicada aos portos da Florida.

Posteriormente, Al-Deek (2002) apresenta um novo método que combina os modelos de MLFNN com séries temporais para prever o volume de caminhões entrando e saindo de portos marítimos. Ressalta-se que modelos foram desenvolvidos em separado para o tráfego de caminhões que entram e saem dos portos.

O método é aplicado com sucesso em cinco portos na Florida: Miami, Tampa, Palm Beach, Jacksonville e Everglades. No artigo é demonstrado com mais detalhes a aplicação realizada para o porto de Everglades.

Informações sobre o fretamento marítimo foram utilizadas para identificar quais parâmetros eram relevantes para o uso como dados de entrada do modelo de geração de viagens para caminhões. Entre os parâmetros que se destacaram e foram utilizados como variáveis independentes estão: contêineres importados, tonelagem importada, barris importados, contêineres exportados, tonelagem exportada, e o dia da semana particular de operação.

Como resultado, foi estimado um aumento de 33% e de 30% na média de entrada e saída, respectivamente, de caminhões por dia para o porto de Everglades considerando um período de cinco anos a partir do ano 2000. A precisão de ambos os modelos foi acima de 90%.

No ano seguinte, Klodzinski e Al-Deek (2003) aplicaram o modelo de RNA desenvolvido anteriormente em Al-Deek (2001b) no porto de Tampa e no porto Canaveral, na Florida, para testar a possibilidade de transferência da técnica de modelagem de RNA.

A partir de dados diários de navios de carga, modelos para ambos os portos foram desenvolvidos com sucesso e validados no nível de confiança de 95%, com dados coletados em campo. Os modelos validados foram executados para previsões de curto prazo de viagens de caminhões em ambos os portos. Foi previsto uma diminuição, para o porto de Tampa, de 0,31% na média anual de caminhões pesados, atribuída à tendência de queda nos embarques de mercadorias a granel; e um aumento médio anual de 5,07%, para o porto Canaveral, que se correlacionou com as tendências históricas e estimativas futuras para a atividade de carga.

Continuando esses estudos, Klodzinski et al. (2004) desenvolveram um aplicativo Java para executar modelos de geração de viagens para caminhões em portos marítimos baseados nos modelos

desenvolvidos pelo Instituto de Sistemas de Transporte da Universidade da Florida Central.

O aplicativo permite executar dois modelos: um modelo selecionado MLFNN para o porto de Everglades e um híbrido do modelo desse porto, que permite ao usuário treinar o modelo de RNA antes da sua execução para os dados desejados. Embora o modelo treinável tenha maior complexidade, possui o benefício de produzir uma rede com um maior grau de precisão.

Ainda em relação à Florida, um método inovador foi desenvolvido por Klodzinski e Al-Deek (2004) para a análise de movimentação de mercadorias portuárias em redes rodoviárias locais, mesclando modelos de geração de viagens para caminhões anteriormente desenvolvidos utilizando RNA com um modelo de simulação microscópica de rede.

Por meio da simulação de computador, esse método abrangente analisa um porto considerado um PGV de caminhões pesados e uma rede viária adjacente que inclui rotas intermodais identificadas que se conectam ao porto.

O tráfego de caminhões a partir do porto é inicialmente modelado com RNA utilizando as atividades de navios de carga no porto. Esses modelos de RNAs foram incorporados ao método para fornecer precisamente os volumes de caminhões e os volumes totais de tráfego para a modelagem das redes. Esse método foi testado com sucesso com dois modelos de simulação microscópica de rede. A transferibilidade foi testada com sucesso com dois portos com características diferentes. Três meses de histórico de dados de campo de cada porto e de locais selecionados nas redes foram utilizados na calibração e validação dos modelos. Ambos os modelos foram validados com sucesso e não apresentaram diferença estatisticamente significativa entre os dados de saída de campo e dos modelos.

Por fim, Sarvareddy et al. (2005) avaliou duas técnicas de modelagem de geração de viagens para caminhões para instalações intermodais.

Uma vez que os estudos anteriores mostraram que a RNA pode ser usado para modelar viagens de caminhões para portos, outras arquiteturas de RNA foram testadas que são mais fáceis de aplicar e treinar, como é o caso da Rede Neural Totalmente Recorrente (*Fully Recurrent Neural Network* – FRNN), que pode ser formada por meio da Aprendizagem Recorrente em Tempo Real (*Real-Time Recurrent Learning* – RTRL). A natureza recorrente faz com que RTRL seja dinâmico, resultando em uma

melhor análise de relações complexas. Devido ao processo mais simples de modelagem e robustez, os autores trabalharam com a hipótese de a técnica de FRNN ser possivelmente a melhor ferramenta de modelagem para a geração de viagens de caminhões intermodais nos portos marítimos utilizando dados de embarcações que transportam cargas, quando comparada à técnica de MLFNN.

Desse modo, esse estudo testou a adaptação do método desenvolvido por Al-Deek et al. (2001) para o porto de Canaveral e comparou os resultados com a performance de um modelo FRNN para o mesmo porto.

Dois locais foram selecionados no porto de Canaveral para a coleta de dados de volume de caminhões. Os dados foram coletados a partir de 20 de setembro de 2001 a 25 de maio de 2002, para as direções de entrada e de saída em cada terminal. Dados com erro ou coletados em dias atípicos foram desconsiderados, resultando em 99 dias de contagens de entrada e 107 dias de contagem de saída.

Entrevistas com as autoridades portuárias, fornecedores e empresas de caminhões concluíram que caminhões não pernoitam nos terminais portuários, dando a entender que a contagem de caminhões de entrada deve ser semelhante à contagem de saída. Além disso, padrões de viagem de caminhões sem carga e com carga se mostraram consistentes e com pouca variação, com base em entrevistas com inquilinos do porto que manipulam a maior parte da atividade de transporte intermodal. Um teste-t de amostras pareadas também foi conduzido e nenhuma diferença significativa entre os caminhões entrando e saindo de cada terminal foi encontrada.

Somente um modelo foi considerado necessário para determinar os volumes de caminhões no porto de Canaveral, uma vez que as contagens são as mesmas em ambas as direções. Como havia mais dados de contagem de saída de caminhões do que de entrada, o modelo foi desenvolvido para esses dados. As contagens de entrada de caminhões foram utilizadas como um conjunto de dados adicional de teste.

A autoridade portuária forneceu dados diários de frete de 30 de agosto de 2001 a 12 de julho de 2002, totalizando 422 registros individuais de embarcações. Os dados relacionavam mercadorias por unidade, tipo, montante unitário, data de chegada e saída, se a mercadoria era importada ou exportada, nome do navio e agente de transporte. Entrevistas realizadas mostraram que certas mercadorias importadas eram utilizadas pelo próprio porto, não gerando viagens de caminhões. Essas

mercadorias foram desconsideradas na modelagem. A movimentação de contêineres também era insignificante e, por isso, não foi incluída como uma variável independente.

Para a construção do modelo MLFNN foi utilizado o software Matlab. Como variáveis independentes foram considerados os dados de frete de embarcações (barris importados, madeira importada, tonelagem importada e tonelagem exportada) e indicador de fim de semana), e como variáveis dependentes os dados de contagem de caminhões (contagem de caminhões saindo diariamente).

Ressalta-se que variáveis de adiantamento e atraso não podem ser consideradas para essas mercadorias porque algumas delas têm um tempo de armazenamento tão longo quanto noventa dias. O grande desafio se trata da distribuição dos dados de frete de forma a relacionar as atividades de remessa com as atividades de caminhões. Assim, os barris importados mensalmente foram investigados para uso como um valor distribuído uniformemente por todo o mês, assim como a madeira importada. As médias mensais de barris importados e madeira importada foram distribuídas para os valores diários e todas as tonelagens de mercadorias pelo seu tempo de resposta.

Como um teste preliminar, um modelo de regressão linear, com as mesmas variáveis, foi comparado com a estrutura do modelo MLFNN preliminar. Os resultados da regressão mostraram coeficientes positivos para as quatro categorias de mercadorias e um coeficiente negativo para o indicador de fim de semana, o que implica volumes mais baixos de caminhões nos fins de semana. Além disso, a constante foi um valor baixo, indicando que o modelo era bom em relação a técnicas anteriores de modelagem adotadas para outro porto na Florida.

Para o desenvolvimento do modelo MLFNN, os dados foram divididos em conjuntos para treinamento e teste, sendo que o conjunto de treinamento foi subdividido em conjunto de dados de calibração e conjunto de dados de validação cruzada. Diferentes configurações foram usadas na modelagem, variando o número de camadas ocultas, número de neurônios em cada camada oculta, número de fases de treinamento, tipo e número de variáveis de entrada. Foi selecionado o modelo MLFNN que resultou no menor valor de MSE. Esse modelo foi composto por uma camada oculta com um nó oculto e produziu uma precisão de 84,60% conforme mostraram as análises. Somado a isso, o modelo MLFNN também capturou com sucesso a tendência sazonal de tráfego de caminhões no porto.

Depois de construído o modelo MLFNN para o porto de Canaveral, um modelo de séries temporais do tipo Arima foi desenvolvido para produzir uma projeção dessa estimativa para um período de cinco anos. Foi estimado que o porto de Canaveral teria um crescimento médio anual de 5,07%.

No desenvolvimento do modelo FRNN também foi utilizada a ferramenta Matlab, assim como as mesmas variáveis utilizadas no modelo MLFNN. O período de influência ou período de armazenamento é automaticamente considerado pelo modelo por meio de sua arquitetura e capacidade de calibração interna. Assim, a distribuição de dados não é necessária para a formação do FRNN. Embora o modelo FRNN não exija que os dados sejam distribuídos, a modelagem preliminar identificou um problema para as toneladas exportadas, que foi resolvido com a associação de uma data anterior às mercadorias exportadas em função do seu tempo de resposta.

Para o desenvolvimento da FRNN, duas variações da rede foram avaliadas para aplicabilidade. A primeira variação foi desenvolvida com os dados divididos em: calibração, validação cruzada e teste. O melhor modelo encontrado, segundo os resultados de MSE, possuía um nó oculto e uma precisão de apenas 71,85%. Esse desempenho baixo pode ser justificado pela quantidade limitada de dados disponíveis com variação sazonal. O modelo foi treinado usando dados fora da alta temporada.

Para fornecer mais dados para o treinamento da FRNN, a segunda variação dividiu os dados somente em dois grupos: calibração e teste. O modelo que produziu o melhor resultado de MSE não possuía nenhum nó oculto. Sua precisão foi de 80,20%, no entanto o modelo produziu resultados insatisfatórios na análise de sensibilidade. Com isso, concluiu-se que o modelo foi insensível às mudanças significativas nos dados de entrada e, por conseguinte, não seria útil na previsão.

A técnica FRNN falhou em produzir um bom modelo para o caso estudado devido à insuficiência de dados, o que torna, nesse quesito, a técnica MLFNN mais vantajosa por necessitar de menos dados. Além disso, no desenvolvimento do modelo MLFNN, os dados foram distribuídos e as contagens de caminhão faltantes não foram relevantes no desempenho do modelo. Contudo para o modelo FRNN, as contagens de caminhões faltantes privaram o modelo de aprender com sucesso as tendências nas contagens de caminhões.

Embora a técnica MLFNN tenha suas vantagens e tenha sido aplicada com sucesso em cinco portos na Florida, uma das desvantagens

é que o modelo se desenvolve considerando cada registro como um evento independente. Mercadorias com extenso tempo de resposta podem causar um aumento nas variáveis de entrada e, assim, os dados precisam ser distribuídos ao contrário da técnica de FRNN. Se os dados são pobremente distribuídos, a técnica pode não produzir um bom modelo.

Portanto, nesse caso específico, o modelo MLFNN mostrou melhores resultados para previsões de curto prazo do que o modelo FRNN. No entanto, ressalta-se que o modelo MLFNN nem sempre reconhece a relação de caminhos para períodos de armazenamento de mercadorias. Já o modelo FRNN pode determinar o efeito do armazenamento na contagem de caminhos desde que seja alimentado por grande quantidade de dados.

3.4.4 Estudos desenvolvidos para o Texas (EUA)

Xie & Huynh (2010a e 2010b) publicaram dois artigos derivados de um estudo sobre modelos de geração de viagens para caminhões em um porto no Texas, nos EUA.

Segundo os autores, uma série de estudos desenvolveram modelos para prever o tráfego de caminhões em portos com base em dados de exploração diária de terminal, sendo os mais recentes baseados em modelos de redes neurais artificiais.

Uma das principais razões para a popularidade das RNAs é que elas têm uma capacidade muito forte de aproximação da função (HORNIK et al., 1989 apud XIE & HUYN, 2010) e podem modelar melhor a relação não linear potencial entre os volumes de tráfego de caminhões e os dados das operações portuárias. Somado a isso, a aplicação de RNAs não exige a especificação explícita de uma formulação de modelo como é exigida por muitas outras técnicas.

Embora as RNAs tenham muitas características atraentes, aplicar essa técnica não é uma tarefa fácil. Muitas decisões difíceis têm de ser feitas corretamente no que diz respeito à formação e seleção do modelo a fim de obter um modelo válido, como por exemplo: arquitetura de rede, tipo de funções de ativação de transferência, taxa de aprendizagem e número de neurônios ocultos (RASMUSSEN & WILLIAMS, 2006).

Além disso, no caso um modelo de RNA com arquitetura do tipo MLFNN, o ajuste precisa superar o excesso de ajuste (“*overfitting*”) e problemas de mínimos locais. O que torna essa tarefa mais difícil é que

as soluções para esses dois problemas são contraditórias. Por um lado, o processo de formação do modelo tenta minimizar o erro de formação e se afastar de eventuais pontos de mínimos locais. Por outro lado, é desejável evitar o excesso de treinamento (“*overtraining*”) do modelo e parar o processo de treinamento prematuramente antes do erro de validação aumentar. Tal contradição é especialmente verdadeira para aplicações com tamanhos de dados de formação limitados, ou seja, quando as informações contidas nas amostras são limitadas é fácil para o modelo MLFNN ser superajustado.

Quando o superajuste acontece, o modelo ajustado pode ter um erro muito pequeno de previsão sobre os dados de formação. No entanto, quando o modelo é aplicado ao conjunto de dados de teste, o erro de previsão pode ser muito grande. Várias estratégias têm sido desenvolvidas para abordar o problema de superajuste.

Uma estratégia de “parar cedo” é adotada nesse estudo. Isto porque na fase de treinamento inicial, tanto os erros de treinamento quanto de validação diminuem drasticamente conforme o número de iterações de treinamento aumenta. Depois de certo número de iterações de treinamento, o erro de treinamento continua a diminuir, enquanto que o erro de validação aumenta. Portanto, o erro de treinamento sempre apresentará uma tendência decrescente. Por outro lado, o erro de validação inicialmente diminui devido à aprendizagem da tendência subjacente, mas, inevitavelmente, começa a aumentar devido ao acréscimo da quantidade de informações sobre o ruído conhecida a partir dos dados de treinamento.

Logo, o seguinte critério foi utilizado nesse artigo para interromper o processo de treinamento: ao começar a partir de qualquer iteração de treinamento, se o erro de validação não melhorar após cem vezes de iterações, o processo de treinamento é finalizado e a rede correspondente ao menor erro de validação é selecionada. Para todos os modelos MLFNN, a taxa de aprendizagem foi definida como 0,01 e uma camada oculta com seis neurônios ocultos foi considerada.

Quanto ao problema de mínimos locais, nesse artigo, o mesmo foi resolvido com o uso de uma estratégia que consiste em executar o processo de treinamento do modelo vinte vezes e escolher a rede formada com o menor erro de validação (não o erro de treinamento).

A principal contribuição desse trabalho é apresentar, além do modelo MLFNN, duas técnicas de aprendizado de máquina baseado em *kernel* para a previsão de viagens de caminhões geradas em terminais

portuários. Essas duas técnicas, *support vector machine* e Processo Gaussiano (Gaussian Process – GP), também provaram ter forte capacidade de aproximação da função com a vantagem de exigir muito menos esforço para o ajuste do modelo do que o modelo MLFNN. Segundo o conhecimento dos autores, essas técnicas não tinham sido estudadas no contexto de geração de viagens de caminhões até então. As duas técnicas são avaliadas com base em dados reais de operação do porto de Houston, nos EUA, e seus resultados são comparados com o modelo MLFNN.

Os modelos SVM são desenvolvidos com base no princípio da minimização do risco estrutural (SUYKENS et al., 2002 apud XIE e HUYN, 2010), em oposição ao princípio da minimização do risco empírico utilizado em redes neurais convencionais. Por isso, os modelos SVM teoricamente podem melhor resolver o problema de superajuste e têm melhor capacidade de generalização do que as redes neurais convencionais. Outra característica importante desse modelo é que os mesmos podem garantir uma solução global ótima para um dado conjunto de dados de treinamento (ZHANG & XIE, 2008, WU et al. 2004 apud XIE & HUYN, 2010).

Assumindo-se um problema de previsão de viagem de caminhões para um porto marítimo com N dados de entradas $\{x(i)\}_{i=1}^N$ e N dados de saída $\{y(i)\}_{i=1}^N$, onde $x(i) \in \mathbb{R}^n$ e $y(i) \in \mathbb{R}^1$, o modelo SVM mapeia primeiro os dados de entrada para um espaço n -dimensional em um espaço maior h -dimensional usando uma função $\phi[x(i)]$ de tal modo que a relação não linear potencial entre $x(i)$ e $y(i)$ pode ser linearizada. Nesta nova e maior dimensão de entrada, a função de estimativa de saída $y(i)$ é a Equação 5:

$$\hat{y}(i) = f[x(i)] = w^T \phi[x(i)] + b \quad (5)$$

Onde: $w \in \mathbb{R}^h$ e $b \in \mathbb{R}^1$ = coeficientes a serem estimados por meio da resolução de um problema duplo de otimização (SUYKENS et al., 2002; SMOLA & SCHOLKOPF, 2004 apud XIE & HUYN, 2010) detalhado em Xie & Huyn (2010).

Uma vez que o duplo problema é resolvido, a função de previsão pode ser escrita como a Equação 6 (SMOLA & SCHOLKOPF, 2004 apud XIE & HUYN, 2010):

$$\hat{y} = f(x) = \sum_{y=1}^N (\alpha_i - \partial_i^*) \phi[x(i)]^T \phi(x) + b = \sum_{y=1}^N (\alpha_i - \partial_i^*) K[x(i), x] + b \quad (6)$$

Onde: $K[x(i), x(j)] = \phi[x(i)]^T \phi[x(j)]$ = função de *kernel* a ser especificada.

Existem vários tipos de funções de *kernel*, incluindo linear, polinomial, base radial, sigmóide e funções de *kernel* de Determinação Automática da Relevância (*Automatic Relevance Determination* - ARD). Nesse estudo, a função de *kernel* sigmóide foi selecionada para o modelo SVM devido ao seu melhor desempenho.

Uma vez que o modelo SVM pode garantir uma solução global ideal, o seu treinamento é relativamente fácil em comparação com o modelo MLFNN. Não são necessárias várias rodadas para superar o problema de mínimos locais.

Para o modelo SVM utilizado nesse estudo, três parâmetros precisaram ser determinados: C, ϵ , e um parâmetro da função de *kernel* (γ). Com base no valor recomendado por Chang & Lin (2001) apud Xie & Huyn (2010), um valor de ϵ igual a 0,05 foi selecionado para todos os modelos SVM. Para cada modelo SVM individual, um teste foi realizado para encontrar os melhores valores para C e γ : os valores entre 1 e 100, com incremento de 0,1 foram testados para C, e os valores entre 0,0001 e 0,01 foram testados para γ com incrementos de 0,0005. Ao todo, 991 valores para C e vinte para γ foram avaliados para cada modelo SVM. Os valores de C e γ que resultaram no menor erro de validação foram selecionados.

Outra técnica apresentada nesse trabalho para estimar viagens de caminhões para portos marítimos é o modelo GP. Modelos GPs recentemente atraíram considerável atenção na comunidade de aprendizagem de máquina. Eles têm sido amplamente utilizados para aplicações de regressão e classificação devido à sua forte capacidade de aproximação de funções (RASMUSSEN & WILLIAMS, 2006; ZHAO et

al., 2008 apud XIE & HUYN, 2010). A formulação do modelo GP é baseada em uma estrutura Bayesiana e consiste em simplesmente maximizar a função de verossimilhança do tipo II em relação a um vetor de parâmetros da função de *kernel*. Essa estrutura Bayesiana possibilita excelente capacidade de generalização aos modelos GPs e permite que os mesmos possam gerar previsões estatisticamente interpretáveis.

Assim como os modelos SVMs, os modelos GPs também exigem um mapeamento do espaço de entrada original em um novo espaço de dimensão superior. A função de regressão nova é mostrada na Equação (7):

$$\hat{y} = f(x) = w^T \phi(x) + \varepsilon \quad (7)$$

Onde:

w = parâmetros de regressão;

$\phi(\cdot)$ = função de mapeamento; e

ε = um termo de ruído que segue uma distribuição independente e idêntica à distribuição Gaussiana $N(0, \sigma^2)$.

É assumido que w também segue uma distribuição Gaussiana com uma média de zero e matriz covariância Σ [ou seja, $N(0, \Sigma)$]. Dadas as distribuições anteriores assumidas e o dado observado de entrada (x) e de saída (y), a distribuição posterior dos parâmetros de regressão w pode ser calculada pela regra de Bayes na Equação (8):

$$p(w|X, y) = \frac{p(y|X, w)p(w)}{p(y|X)} \quad (8)$$

Para um teste de entrada x^* , o resultado previsto $f^* = f(x^*)$ é a média de saída sobre todos os parâmetros possíveis de regressão. A distribuição de f^* é mostrada na Equação (9):

$$p(f^*|x^*, X, y) = \int p(f^*|x^*, w)p(w|X, y)dw \quad (9)$$

A média e a variância de f^* podem ser derivadas como mostrado, respectivamente, nas Equações (10) e (11):

$$\mu(f^*) = K(x^*, X) [K(X, X) + \sigma^2 I]^{-1} y \quad (10)$$

$$\text{var}(f^*) = K(x^*, x^*) - K(x^*, X) [K(X, X) + \sigma^2 I]^{-1} K(X, x^*) \quad (11)$$

Semelhante ao SVMs, as funções de *kernel* são introduzidas na função de previsão de modelos de GPs. Para o modelo de GP utilizado no presente estudo, uma função de *kernel* ARD foi usada.

Entre os três modelos, o modelo GP exige o menor esforço para ajustar. Não há necessidade de usar um conjunto de dados de validação em separado para seleção de um modelo. A formação do modelo GP simplesmente envolve encontrar os melhores parâmetros de *kernel* para a função de *kernel* ARD usando o conjunto de dados de treinamento. Isto é equivalente a maximizar a função de máxima verossimilhança tipo II (XIE et al., 2010 apud XIE e HUYN, 2010) em relação a um vetor de parâmetros de *kernel*. Os detalhes desse processo de treinamento podem ser encontrados em Rasmussen & Williams (2006) e Xie et al. (2010) apud Xie e Huyn (2010) e não foram repetidos aqui.

Teoricamente, esse processo de maximização também pode ficar preso em pontos mínimos locais e gerar diferentes resultados de várias rodadas de treinamento. No entanto, os testes realizados nesse estudo mostraram que várias rodadas de treinamento do mesmo modelo GP produzem resultados consistentes.

Para testar os modelos propostos, dados de dois terminais de contêineres no porto de Houston, Bayport e Barbours Cut, foram usados. Os conjuntos de dados incluem dados operacionais diários detalhados abrangendo sete meses, de 1º de junho de 2008 a 30 de janeiro de 2009, como: a contagem de contêineres descarregados, contagem de contêineres com carga, número de caminhões que descarregam cargas, e número de caminhões que coletam cargas. Pontos de dados faltantes e anormais causados por feriados e eventos incomuns, como o furacão Ike, foram removidos dos dados originais. Isso resulta em 150, 154, 139, e 148 dias de dados para o conjunto de dados, respectivamente, de exportação do Bayport, importação do Bayport, exportação do Barbours Cut e importação do Barbours Cut.

As viagens de caminhão geradas pelas atividades de importação e exportação foram investigadas separadamente, gerando quatro conjuntos

de dados para testes e comparação de modelos. A análise dos dados resumidos sugere que a maioria dos contêineres de importação e exportação não permanece mais do que 12 dias nos dois terminais. Portanto, viagens de caminhão de exportação são modeladas com base nos números de contêineres de exportação nos próximos 12 dias, enquanto viagens de caminhão de importação são modeladas com base nos números de contêineres de importação nos últimos 12 dias.

Cada conjunto de dados é separado aleatoriamente em três subgrupos de treinamento, validação e teste. O conjunto de dados de treinamento é usado para formar os modelos. Para os modelos MLFNN e SVM, o conjunto de dados de validação é usado para encontrar os melhores parâmetros, tais como C e para evitar superajuste dos dados de treinamento. Para o modelo de GP, o conjunto de dados de validação não é necessário e pode realmente ser combinado com os dados de treinamento estabelecidos para fins de formação. No entanto, para fornecer uma base para comparação de modelos, o modelo GP é apenas formado usando o conjunto de dados de treinamento.

Dois cenários de teste foram considerados nesse estudo para comparar a eficácia entre os modelos SVM, GP, e MLFNN para estimar viagens de caminhões para terminais marítimos. O primeiro cenário usou setenta pontos de dados para treinamento, trinta pontos de dados para validação, e os pontos de dados restantes para os testes. O segundo cenário usou oitenta pontos de dados para treinamento, quarenta pontos de dados para validação, e os pontos de dados restantes para os testes. Para cada cenário de teste, 12 modelos foram montados e testados. Por fim, essas três técnicas foram comparadas com base no desempenho da média do erro percentual absoluto dos modelos.

Para ambos os cenários de teste, os modelos GP e SVM produziram desempenho de previsão semelhante e, em ambos os casos de teste seu desempenho de previsão comparou-se favoravelmente com a do modelo MLFNN. Para comparar a sua transferência, os três modelos formados usando os dados do terminal Barbours Cut foram utilizados para prever o conjunto inteiro de dados do terminal Bayport. Uma comparação dos resultados sugeriu que as duas técnicas de aprendizado de máquina baseado em *kernel* demonstraram melhor desempenho de transferência do que a técnica MLFNN para ambos os cenários I e II.

Cabe ressaltar que entre as duas técnicas baseadas em *kernel*, o modelo GP se destaca, pois, além de ser mais simples de ajustar, ele pode estimar os desvios-padrão dos dados de previsão de volume de

caminhões, sendo muito útil para a construção de intervalos de confiança para os valores previstos. Uma vez que geralmente são inevitáveis erros nas previsões de tráfego de caminhões em portos, proporcionar um intervalo de confiança seria muito mais informativo e significativo do que simplesmente apresentar um valor único.

Em resumo, o forte desempenho global dos modelos GP e SVM indica que eles são abordagens viáveis para a previsão de viagens de caminhões para portos marítimos. Pode ser possível que o desempenho de previsão do modelo MLFNN pode ser melhorado por meio da utilização de outras técnicas que lidem melhor com o superajuste e problemas de mínimos locais. No entanto, dado que os modelos GP e SVM produzem um desempenho de previsão satisfatório e sua aplicação é relativamente simples, eles oferecem aos pesquisadores uma boa alternativa frente ao modelo MLFNN.

3.4.5 Estudos desenvolvidos para toda a Singapura

Luk et al. (2002) desenvolveram um artigo com uma modelagem para a circulação de veículos que fazem fretamento de mercadorias (caminhões) em Singapura.

O estudo considerou várias zonas com diferentes usos de solo, no total de 32 internas e quatro zonas externas. Dentre as quatro zonas externas, se encontra o porto de Singapura, motivo pelo qual esse estudo foi incluído na revisão bibliográfica.

Um modelo do tipo gravitacional duplamente vinculado (restringido) com uma função exponencial de separação, no qual o tempo de viagem é a medida de separação, foi usado para distribuir viagens de caminhões entre as várias zonas internas e externas. Ressalta-se que o modelo foi calibrado parcialmente devido à falta de informações completas de origem-destino dos caminhões que trafegavam pelas zonas consideradas.

3.4.6 Estudos desenvolvidos para Taiwan (República da China)

Chu (2011) publicou um método empírico para estimar viagens internas e externas de caminhões para um porto principal em Taiwan, na República da China.

De acordo com o autor, devido ao grande crescimento do volume de fretamento nos EUA, alguns estados e organizações de planejamento metropolitano tem incorporado o transporte de mercadorias nos modelos de geração de viagens regionais e estaduais com o objetivo de identificar as prováveis deficiências da rede viária.

Nesse sentido, Chu cita Al-Deek (2000), Holguín-Veras, López-Genao & Salam (2002) e Sarvareddy et al. (2005). Além desses estudos, ele também destaca os de:

- 1) Shin & Kawamura (2006), que desenvolveram um modelo logit binário para estimar o número de entregas realizadas por caminhões por semana com base numa série de decisões estratégicas e operacionais de cadeias de varejo de móveis, sendo que foi descoberto que variáveis explicativas das características da localização e tipos de loja tem efeito significativo na estimativa do número de caminhões;
- 2) Novak et al. (2008), que usa um modelo de regressão linear para estimar demanda de frete baseada em variáveis independentes econômicas, tais como empregos totais por zona, e variáveis espaciais, tais como extensão da rodovia, extensão da ferrovia e zonas de influência do porto;
- 3) Fischer et al. (2000), que desenvolveu um modelo com base em dados de emprego, uso da terra, e facilidade comercial para estimar o fluxo de mercadorias, posteriormente convertido para viagens de caminhões usando fatores de carregamento de caminhões por classe de peso do caminhão;
- 4) Harris et al. (2009), que desenvolveu um método para desagregar dados nacionais de origem/ destino de fretes com base em pesquisa de fluxo de mercadorias para o nível local, sendo que os fatores econômicos usados para estimar a demanda de frete foram o valor do envio de frete, renda pessoal, população e emprego.

No quesito estado da prática, Chu incluiu o uso de um modelo de estimativa demanda de frete e o uso do método de três ou quatro etapas. Ele ainda ressaltou que, usualmente, modelos de demanda de frete são caracterizados pelo modelo baseado em viagens de veículos (caminhões) ou pelo modelo baseado no fluxo de mercadorias.

As vantagens dos modelos baseados em viagens de veículos (caminhões) incluem: coleção de dados e integração com rastreamento de veículos por sistemas inteligentes de transportes relativamente fáceis; fornecimento direto como dado de saída do número de viagens atribuído na rede; e modelagem das viagens com caminhões cheios e vazios. As desvantagens incluem: dificuldade de capturar mudanças políticas e características econômicas das cargas; e aplicação inadequada do modelo no sistema de transporte de frete multimodal na escolha modal e seleção de veículo.

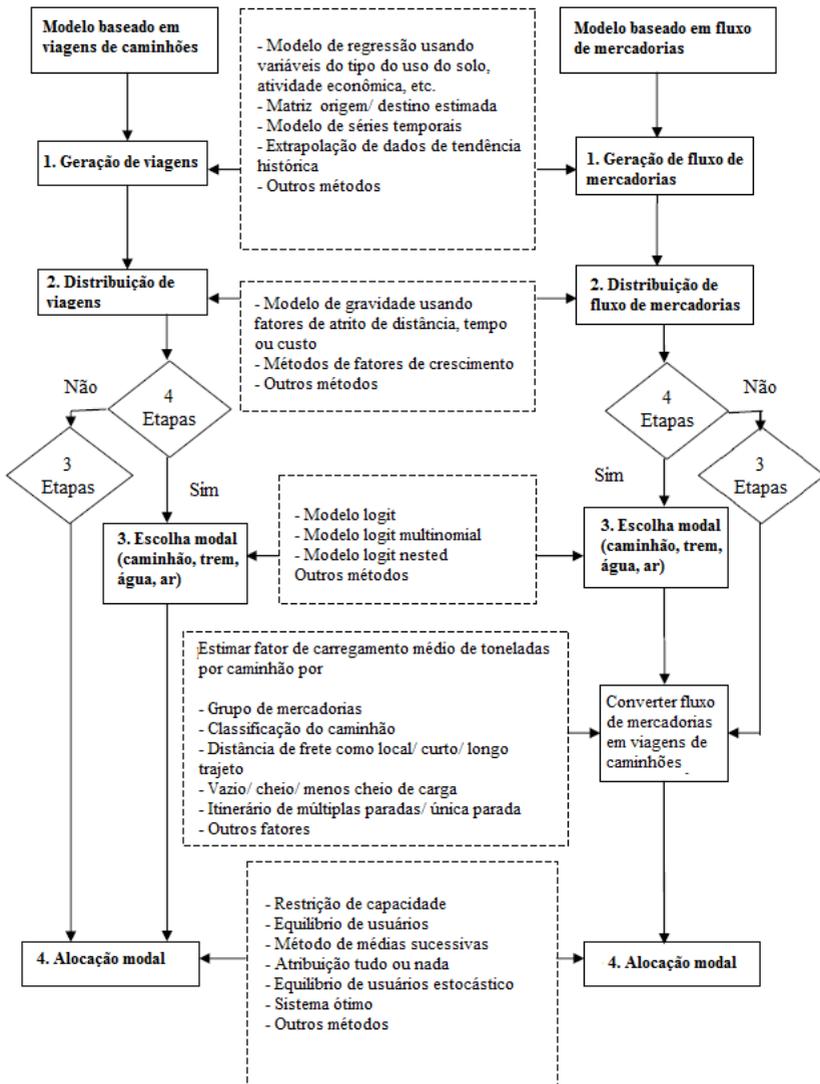
As vantagens dos modelos baseados no fluxo de mercadorias incluem: capturar mudanças políticas e características econômicas das cargas. As desvantagens incluem: dificuldade de coletar dados suficientes sobre mercadorias; requer um fator de carregamento de caminhões para superar o problema de converter quantidades de mercadorias em número necessário de viagens de caminhões; deficiência no modelo de viagens de caminhões vazios; e dificuldade de calibração e validação do modelo.

Com base nessa revisão bibliográfica, o autor desenvolveu um método para modelagem de fretamento multimodal com ambos os modelos baseados em viagens de caminhões e em fluxo de mercadorias, a qual é mostrada na Figura 3.8.

Esse método se diferenciou das desenvolvidos até então por incorporar fatores importantes nos mecanismos de estimativas de viagens de caminhões, especialmente para fatores associados com fatores de carregamento de caminhões. A maioria dos modelos baseados em fluxo de mercadorias converte quantidades de caminhões em viagens de caminhões utilizando fatores de carregamento “toneladas-por-caminhão” ao invés de converter mercadoria somente em quantidade de caminhões. Fatores de carregamento “toneladas-por-caminhão” são estimados como um valor médio que varia de acordo com o grupo de mercadorias, classificação do caminhão, distância de envio de frete, ou número de caminhões vazios.

Para melhor entender o método usado para prever a geração de viagens de caminhões e examinar a diferença associada entre a conversão do modelo baseado em viagens de caminhões e em fluxo de mercadorias, um estudo de caso no porto de Kaohsiung, em Taiwan, foi desenvolvido como segue.

Figura 3.8 – Processo de modelagem de fretamento multimodal



Fonte: Chu, 2011.

Uma pesquisa nas rodovias foi conduzida de agosto a outubro de 2009 para coletar dados sobre a movimentação de caminhões na área metropolitana de Kaohsiung. Uma amostra de seis mil motoristas de caminhões pesados carregando TEUs completaram o questionário. As entrevistas foram realizadas nos terminais do porto, nos centros de distribuição e nos pontos de parada de caminhões na rodovia.

O questionário aplicado possuía as seguintes questões: padrão de motorista (motoristas contratados ou privados), configuração do caminhão, número de eixos, peso bruto do veículo, carregamentos do caminhão (cheio, menos que cheio ou vazio), peso de carga útil, tipo de mercadoria transportada, distância de transporte, origem da viagem, destino da viagem, rotas viajadas, padrão de viagens de caminhão (viagem externa-externa, interna-externa ou interna-interna), tipo de viagens (número de paradas), conexão com outros modos (trem, navio ou avião) e solicitação de entrega *just-in-time*. Por causa do método face a face se alcançou taxas de resposta de mais de 90% em todas as questões.

Foram selecionadas nove potenciais variáveis explicativas para o modelo: movimentação de carga em toneladas; número de TEUs; eficiência no manuseio da carga (tonelada de carga manuseada por hora trabalhável); taxa de utilização de armazéns (armazenagem em toneladas por dia dividida pela capacidade do armazém); realização de expansão do porto ou projetos de desenvolvimento; nova operação ou melhoria de infraestrutura em portos vizinhos concorrentes; Produto Interno Bruto (PIB); percentagem de alteração no preço da gasolina no ano anterior; e número de tufões que atravessaram o porto. Dados mensais associados com essas potenciais variáveis explicativas foram coletados de janeiro de 2000 a dezembro de 2008.

O artigo analisou as flutuações das séries históricas desses dados, bem como suas correlações. O tráfego de caminhões no porto teve correlação positiva com movimentação de carga, TEUs, eficiência no manuseio da carga, taxa de utilização de armazéns e PIB. Somente a correlação com o número de TEUs e com o PIB deu acima de 0,50, o que indica uma forte correlação com o volume de tráfego no porto. Correlações negativas apareceram para nova operação ou melhoria de infraestrutura em portos vizinhos concorrentes e para percentagem de alteração no preço da gasolina no ano anterior, com valor de correlação -0,45 e -0,47, respectivamente. O número de tufões apresentou baixa correlação positiva, com coeficiente igual a 0,3.

Diferentes métodos – regressão linear múltipla, séries-históricas, e redes neurais artificiais – foram utilizados para desenvolver um modelo de geração de viagens internas e externas de caminhões em um terminal portuário.

Por causa da variação sazonal, a equação de geração de viagens de caminhões que utilizou uma regressão não-linear múltipla tomando como base uma transformação logaritmo natural apresentou um nível de significância melhor que uma regressão linear. O modelo final resultou na Equação 12:

$$\ln Y_{im} = \beta_0 + \beta_{1im} \times \ln X_{1im} + \beta_{2im} \times X_{2im} + \beta_{3im} \times X_{3im} + \beta_{4im} \times X_{4im} + \varepsilon_{im} \quad (12)$$

Onde:

Y_{im} = viagens de caminhões (média do volume diário anual) durante o ano i e o mês m;

X_{1im} = TEUs durante o ano i e o mês m;

X_{2im} = nova operação ou melhoria de infraestrutura em portos vizinhos concorrentes durante o ano i e o mês m (variáveis *dummy*, 0 = não e 1 = sim);

X_{3im} = PIB durante o ano i e o mês m;

X_{4im} = alteração no preço da gasolina no ano anterior i-1 no mês m; e

ε_{im} = erro randômico durante ano i no mês m.

Com base nesse modelo de regressão múltipla, uma análise de sensibilidade das variáveis explicativas considerando a precisão de converter a variável resposta do fluxo de mercadorias em viagens de caminhões é conduzida. Para essa conversão, a tonelada total de carga manuseada no porto intermodal é multiplicada pela percentagem de divisão modal da tonelada de carga por caminhão e então dividida pelo fator de carregamento de caminhão, definido segundo a Equação 13:

$$T = CF \times \alpha \div \lambda \quad (13)$$

Onde:

T = viagens de caminhões (média do volume diário anual de caminhões);

CF = fluxo de mercadorias (em toneladas);
 α = divisão modal do total de toneladas por caminhão (em porcentagem); e
 λ = fator de carregamento de caminhão (em tonelagem de mercadoria por caminhão).

O fator de carregamento médio para caminhões que transportam TEUs é derivado da pesquisa nas rodovias. Os resultados mostraram que a diferença entre viagens de caminhões e mercadorias convertidas em viagens de caminhões está na faixa de 3-17%. No entanto, uma vez a diferença sendo maior que 30%, as variáveis de nova operação ou melhoria de infraestrutura em portos vizinhos concorrentes e de alteração no preço da gasolina no ano anterior apresentam insignificância estatística, o que indica que viagens de caminhões convertidas por meio do fluxo de mercadorias são muito sensíveis a essas duas variáveis. Não obstante, os resultados mostraram que as variáveis TEUs e PIB ainda possuem forte correlação com o volume de caminhões.

Além disso, ambos os valores de R^2 e de R^2 ajustado no modelo de fluxo de mercadorias são menores que os valores no modelo de viagens de caminhões, principalmente porque o fator de carregamento médio de caminhões não pode exatamente explicar a variação no nível de carregamento durante a cadeia de viagens com múltiplas paradas e mudanças de comportamento do remetente ou de motoristas de caminhões na logística de fretamento.

Modelos de séries históricas também foram analisados, entre eles: modelo de tendência linear, modelo de tendência quadrática, modelo de tendência de crescimento exponencial, modelo de suavização exponencial, método de Winters e modelo Arima. O detalhamento de cada um desses modelos pode ser visto em Chu (2011).

O modelo Arima pode ser usado para modelar diferentes séries históricas, com ou sem tendência ou componentes sazonais, característica que o torna mais flexível no ajuste de dados para previsões de curto prazo (BOX e JENKINS, 1994). Visto que a série histórica de dados de volume de caminhões indica padrões não estacionários, um modelo Arima sazonal foi desenvolvido com base no nono ano dos dados históricos. Os resultados indicam que o modelo Arima (3, 1, 0) tem o menor valor de MSE e parece se ajustar melhor nas estimativas de volume de caminhões. Sua configuração é expressa por meio da Equação 14:

$$Y_t = 8,074 - 1.074Y_{t-1} - 1.071Y_{t-2} - 0.954Y_{t-3} + \varepsilon_t \quad (14)$$

Onde:

Y_t = volume mensal de caminhões no mês t ; e

ε_t = termo de erro randômico durante o mês t .

Por fim, um modelo de RNA com algoritmo BPN e com três camadas – uma camada de entrada com quatro variáveis, uma camada oculta com nove neurônios, e uma camada de saída com a variável volume de caminhões – foi desenvolvido com base na Raiz da Média dos Erros Quadráticos (*Root Mean Square Error* – RMSE). As quatro variáveis na camada de entrada são as mesmas variáveis explicativas utilizadas no modelo de regressão múltipla.

Para validar todos os modelos construídos com base nos dados levantados de janeiro 2000 a dezembro de 2008, uma observação adicional de janeiro de 2009 a dezembro de 2009 foi utilizada. Os resultados do Mape, do Desvio Médio Absoluto (*Mean Absolute Deviation* –MAD), e da Média do Desvio Quadrado (*Mean Square Displacement* – MSD) indicam que o modelo de RNA tem melhor ajuste aos dados, seguido pelo modelo de regressão múltipla e pelo modelo Arima, no entanto as diferenças entre eles são pequenas. O modelo de RNA demonstrou que é o melhor modelo de previsão devido à capacidade de identificar relações não-lineares entre o volume de caminhão e os fatores influentes. Além disso, o desempenho da RNA poderia ser muito melhor caso fosse utilizada uma amostra maior.

3.4.7 Estudos desenvolvidos para Montevideu (Uruguai)

Pereyra, Rubinstein & Pérez (2008) publicaram um artigo que trata de taxas geradoras de viagens de caminhões para o porto de Montevideu, no Uruguai.

De acordo com os autores, até a data de publicação desse artigo, a bibliografia a respeito de modelagem de entrada/saída de caminhões transportando contêineres em um porto era pequena, bem como não era eficiente para todos os casos. No levantamento de revisões bibliográficas para elaboração do trabalho, é citado o artigo escrito por Al-Deek et al.

(2000) e o artigo de autoria de Al-Deek (2001a), ambos já descritos na presente tese.

Para elaboração do modelo foram considerados os seguintes dados do porto de Montevideu no período do ano de 2006: quantidade de escalas diárias de navios; volume diário de contêineres operados no cais, classificados como carregados ou descarregados; e o volume diário de caminhões registrados segundo entrada e saída em todos os acessos do porto marítimo.

Cabe ressaltar que devido a sua condição de “porto livre”, a entrada e saída de caminhões nem sempre está ligada a operação do cais. Isto se deve ao fato de que alguns dos contêineres descarregados de navios são armazenados no porto antes de serem carregados para outros navios, nunca deixando a zona portuária pelos portões. O mesmo acontece com alguns contêineres descarregados de caminhões, que nunca deixam a zona portuária pelo mar.

Esses dados – quantidade de navios e quantidade de contêineres – possuem as seguintes características: somente valores positivos, grande variabilidade dentro de alguns dias da semana, zeros observados com alta frequência, e se comportam como uma distribuição de Poisson. Por causa disso, os autores optaram por desenvolver um modelo de contagem para estimar a geração de viagens de caminhões para o porto.

A entrada e saída de caminhões foram modeladas de forma independente. Variáveis binárias representando os dias da semana e meses foram testadas, bem como variáveis *dummy* para tratar dados atípicos observados. Além da distribuição de Poisson, foram estudados modelos de distribuição binomial negativa e modelos com componentes autoregressivos para compreender fenômenos de atraso operacional. Tabelas resumo com as variáveis consideradas e os modelos desenvolvidos podem ser vistas a seguir (Tabelas 3.2, 3.3 e 3.4).

Tabela 3.2 – Descrição das variáveis testadas

Nome da variável	Tipo	Descrição
ing	discreta	Quantidade de contêineres que entram no terminal portuário por dia
sal	discreta	Quantidade de contêineres que saem do terminal portuário por dia
buques	discreta	Quantidade de navios que transportam contêineres que fazem escalas diariamente
cga	discreta	Quantidade de contêineres carregados para o navio diariamente
dga	discreta	Quantidade de contêineres descarregados do navio diariamente
Lun, ..., sab	<i>dummy</i>	Variáveis binárias segundo o dia da semana
Mesdum1, ..., mesdum11	<i>dummy</i>	Variáveis binárias segundo o mês do ano
Ing _j	autoregressivo	Variável autoregressiva de entradas para o dia j anterior a data de chegada estimada
Sal _j	autoregressivo	Variável autoregressiva de entradas para o dia j anterior a data de partida estimada
Dum _k	<i>dummy</i>	Variável binária para tratar eventos fora de classificação

Fonte: Pereyra, Rubinstein e Pérez, 2008.

Tabela 3.3 – Modelos para entrada diária de contêineres

Modelo	Variáveis explicativas	Variáveis <i>dummy</i>	Atrasos	Dumk	Pseudo-R²
1A Poisson	Buques	diadumX, mesdum Y	-	Dum3	0,7679
1B Poisson	Cga, dga	“	-	Dum3	0,7691
2A Binomial negativa	Buques	“	-	Dum3	0,1581
2B Binomial negativa	Cga, dga	diadumX, mesdum Y	-	Dum3	0,1582
3A Poisson	Buques	“	Ing_1	Dum3	0,7741
3B Poisson	Cga, dga	“	Ing_1	Dum3	0,7752
4A Poisson	Buques	“	Ing_1,..., ing_6	Dum3	0,7820
4B Poisson	Cga, dga	diadumX, mesdum Y	“	Dum3	0,7829
5A Poisson	Buques	“	Ing_1,..., ing_7,ing_14	Dum3	0,7823
5B Poisson	Cga, dga	“	“	Dum3	0,7828
6A Binomial negativa	Buques	diadumX, mesdum Y	Ing_1	Dum3	0,1564
6B Binomial negativa	Cga, dga	“	Ing_1	Dum3	0,1565

Fonte: Pereyra, Rubinstein e Pérez, 2008.

Tabela 3.4 – Modelos para saída diária de contêineres

Modelo	Variáveis explicativas	Variáveis <i>dummy</i>	Atrasos	Dumk	Pseudo-R²
1A Poisson	Buques	diadumX, mesdum Y	-	Dum3	0,7222
1B Poisson	Cga, dga	“	-	Dum3	0,7224
2A	Buques	“	-	Dum3	0,1094
Binomial negativa					
2B	Cga, dga	diadumX, mesdum Y	-	Dum3	0,1092
Binomial negativa					
3A Poisson	Buques	“	Sal_1	Dum3	0,7347
3B Poisson	Cga, dga	“	Sal_1	Dum3	0,7350
4A Poisson	Buques	diadumX, mesdum Y	Sal_1,..., sal_7	Dum3	0,7426
4B Poisson	Cga, dga	“	“	Dum3	0,7427
5A Poisson	Buques	“	Sal_1,..., sal_7,sal _14	Dum3	0,7412
5B Poisson	Cga, dga	diadumX, mesdum Y	“	Dum3	0,7413
6A	Buques	“	Sal_1,..., sal_6	Dum3	0,1156
Binomial negativa					
6B	Cga, dga	“	“	Dum3	0,1148
Binomial negativa					

Fonte: Pereyra, Rubinstein e Pérez, 2008.

Os modelos e suas respectivas variáveis foram comparados entre si utilizando parâmetros estatísticos, como o teste t, R^2 , e menor MSE. Os resultados revelaram que os modelos que mais se ajustaram aos dados foram os de regressão de Poisson, mais especificamente os modelos 5Bs tanto para entrada quanto para saída de caminhões. Cabe ressaltar que os autores acreditam que esse modelo seja simples de ser elaborado sem perda de especificidade, o que se torna uma alternativa interessante para os modelos existentes – regressão linear e RNA –, embora nenhuma comparação prática tenha sido feita.

Os ensaios realizados também mostraram que as variáveis para dia da semana e mês foram significativas em todos os modelos. Por sua vez, a incorporação de atraso melhorou substancialmente a qualidade da estimativa, indicando ainda a dependência do tempo da mesma. Em todos os casos foi preferível utilizar a informação de operação no cais, discriminada entre carga e descarga, do que a chegada de navios como variável explicativa.

3.5 ANÁLISE DA REVISÃO BIBLIOGRÁFICA

Em resumo, quanto aos estudos de modelos de geração de viagens de caminhões que tratam especificamente de portos marítimos, foram levantadas as fontes mostradas na Tabela 3.5.

Essa tabela apresenta as fontes em colunas segundo referência bibliográfica e país de origem, a técnica, as variáveis dependentes e independentes utilizadas na construção dos modelos, assim como as conclusões obtidas para cada estudo.

Tabela 3.5 – Fontes bibliográficas encontradas sobre modelos de geração de viagens para caminhões em portos
(continua)

Documento / País	Modelo	Variável dependente	Variáveis independentes	Conclusões
ITE (2012)/ EUA	Elaboração de modelo de regressão linear simples.	Número de viagens de veículos nos acessos a portos/terminais marítimos.	Número de berços ou número de acres dos portos/terminais marítimos.	Taxa média de viagens de 171,52 por berço de atracação ou 11,93 por acre dia útil da semana R ² de 0,58 ou 0,93.
Al-Deek et al. (2000) / EUA	Elaboração de modelo de regressão linear simples.	Volume de caminhões carregados entrando/saindo do porto para o grupo de dia g.	Número total de unidades de mercadorias exportadas/importadas para o grupo de dia g.	R ² de 0,686 para importação e 0,784 para exportação.
Al-Deek (2001a) / EUA	Comparação entre modelo de regressão linear de Al-Deek, Jonson & Mohamed (2000) e RNA (MLFNN).	Volume de caminhões entrando/ saindo do porto.	Unidades de contêineres importados/exportados no dia analisado e nos três dias anteriores e posteriores ao analisado e o dia da semana.	RNA: 2 camadas ocultas, 3 funções tansing na 1ª camada oculta, 2 funções tansing na 2ª camada oculta e 2 funções lineares na camada de saída. RNA: mais precisa, porém necessita de uma amostra grande de dados.

Tabela 3.5 – Fontes bibliográficas encontradas sobre modelos de geração de viagens para caminhões em portos
(continuação)

Documento / País	Modelo	Variável dependente	Variáveis independentes	Conclusões
Al-Deek (2001b) / EUA	Comparação entre regressão linear múltipla e RNA (MLFNN).	Volume de caminhões entrando/ saindo do porto.	Atividade de carga no porto (mercadorias importadas e exportadas de navios).	RNA: resultados mais próximos aos reais do que modelos de regressão linear múltipla.
Holguín-Veras, López-Genao & Salam (2002) / EUA	Elaboração de modelo de regressão linear múltipla.	Número de viagens de caminhões em um sentido de viagem.	Área do terminal, número de TEUs e o número de caixas de contêineres manuseadas por terminal em um determinado ano.	O modelo mostrou-se válido.
Luk et al. (2002)/ Singapura	Elaboração de modelo gravitacional.	Viagens de caminhão.	Tempo de viagem e zonas com diferentes usos de solo (incluindo zonas portuárias).	O modelo mostrou-se válido.

Tabela 3.5 – Fontes bibliográficas encontradas sobre modelos de geração de viagens para caminhões em portos
(continuação)

Documento / País	Técnica empregada	Variável dependente	Variáveis independentes	Conclusões
.Al-Deek (2002) / EUA	Combinação de RNA (MLFNN) com séries históricas.	Volume de caminhões entrando/ saindo do porto.	Contêineres importados, tonelagem importada, barris importados, contêineres exportados, tonelagem exportada e o dia da semana particular de operação.	Estimado um aumento de 33% e de 30% na média de, respectivamente, entrada e saída de caminhões por dia para o porto de Everglades (EUA) considerando o período de 2000-2005. Modelos: precisão acima de 90%.
Klodzinski & Al-Deek (2003) / EUA	Transferabilidade do modelo RNA (MLFNN) de Al-Deek (2001b).	Volume de caminhões entrando/ saindo do porto.	Atividade de carga no porto (mercadorias importadas e exportadas de navios).	Estimado uma diminuição de 0,31% na média anual de caminhões para o porto de Tampa e um aumento de 5,07% para o porto Canaveral Modelos: desenvolvidos com sucesso e validados no nível de confiança de 95% com dados coletados em campo.

Tabela 3.5 – Fontes bibliográficas encontradas sobre modelos de geração de viagens para caminhões em portos
(continuação)

Documento / País	Técnica empregada	Variável dependente	Variáveis independentes	Conclusões
Agnello, Jonnes & Allen (2004) / EUA	Adaptação de modelo quatro etapas.	Número de caminhões.	Fatores em relação às zonas que atraem caminhões ou em que há proibição de circulação de caminhões, equivalência de caminhões em relação a carros de passageiros, uso do solo e jurisdição.	O modelo mostrou-se válido.
Klodzinski et al. (2004) / EUA	Elaboração de aplicativo de aplicativo Java para executar modelos de RNA (MLFNN) de Al-Deek (2001b).	Volume de caminhões entrando/ saindo do porto.	Atividade de carga no porto (mercadorias importadas e exportadas de navios).	O modelo mostrou-se válido.

Tabela 3.5 – Fontes bibliográficas encontradas sobre modelos de geração de viagens para caminhões em portos
(continuação)

Documento / País	Técnica empregada	Variável dependente	Variáveis independentes	Conclusões
Klodzinski & Al-Deek (2004) / EUA	Combinação de modelo RNA (MLFNN) de Al-Deek (2001b) com simulação microscópica de rede.	Volume de caminhões entrando/saindo do porto.	Atividade de carga no porto (mercadorias importadas e exportadas de navios).	O modelo mostrou-se válido.
Sarvaderry et al. (2005) / EUA	Comparação entre modelo RNA (MLFNN) de Al-Deek (2001b) e RNA (FRNN).	Volume de caminhões entrando/saindo do porto.	Atividade de carga no porto (mercadorias importadas e exportadas de navios).	MLFNN: 1 camada oculta com 1 nó oculto, e precisão de 84,6%. FRNN: 1 nó oculto, e precisão de 71,85%. MLFNN: melhores resultados para previsões de curto prazo do que o modelo FRNN. FRNN: poderia até fornecer estimativas mais precisas do que o MLFNN se a amostra de dados fosse maior.

Tabela 3.5 – Fontes bibliográficas encontradas sobre modelos de geração de viagens para caminhões em portos
(continuação)

Documento / País	Técnica empregada	Variável dependente	Variáveis independentes	Conclusões
Pereyra, Rubinstein & Pérez (2008)/ Uruguai	Comparação entre binomial negativa e regressão de Poisson.	Entrada e saída de caminhões do porto.	Quantidade de contêineres que entram/saem do terminal portuário por dia, quantidade navios que transportam contêineres que fazem escalas diariamente, quantidade de contêineres carregados/ descarregados para o navio diariamente, dia da semana, mês do ano, variável autoregressiva de entradas para o dia j anterior à data de chegada/partida estimada, variável binária para tratar eventos fora de classificação.	Regressão de Poisson: se ajustou melhor aos dados (R^2 de 0,7413), com as variáveis quantidade de contêineres, dia da semana, mês do ano e eventos fora de classificação Modelos: simples de ser elaborado e sem perda de especificidade, alternativa interessante para os modelos existentes – regressão linear e RNA –, embora nenhuma comparação prática tenha sido feita.

Tabela 3.5 – Fontes bibliográficas encontradas sobre modelos de geração de viagens para caminhões em portos (conclusão)

Documento / País	Técnica empregada	Variável dependente	Variáveis independentes	Conclusões
Xie & Huynh (2010a) e Xie & Huynh (2010b) / EUA	RNA (MLFNN) e modelos baseados no <i>kernel</i> (SVM e GP).	Viagens de caminhão de exportação/importação .	Números de contêineres de exportação nos próximos/últimos 12 dias.	SVM e GP: desempenho de previsão semelhante com a RNA, com a vantagem de requerer menos esforço para o ajuste do modelo, e desempenho na transferabilidade melhor que a RNA.
Chu (2011)/ China	Comparação entre modelo de regressão linear múltipla, séries-históricas e RNA (BPN).	Viagens de caminhões durante o ano i e o mês m .	TEUs, nova operação ou melhoria de infraestrutura em portos vizinhos concorrentes e PIB durante o ano i e o mês m , alteração no preço da gasolina no ano anterior $i-1$ no mês m , e erro randômico.	RNA: mais precisa em relação ao restante das abordagens, seguido pela regressão linear e pela Arima, mas as diferenças entre eles foram pequenas.

Fonte: Autoria Própria.

Percebe-se, pela Tabela 3.5, que, ao contrário da tendência de utilizar basicamente a técnica de regressão linear simples ou múltipla para construir modelos de geração de viagens, verificada no item 2.7, na página 80, para os outros PGVs, no caso dos portos marítimos a maioria dos estudos analisados optou por fazer uso de modelos matemáticos mais complexos como redes neurais artificiais ou aprendizado de máquina baseado em *kernel*.

No caso desses estudos, entre as técnicas utilizadas, os modelos de regressão linear múltipla se mostraram mais precisos do que os modelos de regressão linear simples, que por sua vez apresentam precisão inferior aos dos modelos de redes neurais artificiais. Como os modelos de redes neurais artificiais são de difícil aplicação, alguns autores estudaram alternativas a esse modelo, como a distribuição de Poisson e distribuição binomial negativa, e aprendizado de máquina baseados em *kernel*. Esses modelos não são tradicionalmente usados para esse tipo de análise, mas apresentaram bons resultados – os de *kernel* apresentaram resultados semelhantes aos dos modelos de RNAs com a vantagem de exigirem menos esforço para o seu desenvolvimento. Infelizmente não foram realizados estudos comparando o desempenho do modelo de regressão de Poisson e distribuição binomial negativa em relação aos outros.

Quanto às variáveis, a maioria dos autores utilizou como variável dependente o fluxo de caminhões, visto que os mesmos tinham como intuito apenas o gerenciamento de tráfego de entorno do PGV. Os modelos foram formados em separado para caminhões entrando e saindo do porto ou terminal portuário, pois a diferença de volume nos dois sentidos era significativa.

Das variáveis independentes, notou-se que grande parte dos estudos utilizou como variável “principal” o volume de contêineres que entram/saem do porto ou terminal portuário durante certo período de tempo, podendo considerar o fluxo de mercadorias ocorrido em um dia ou até doze dias conforme a operação do PGV estudado; ou seja, a principal variável independente utilizada se relaciona à área ou demanda do empreendimento, assim como ocorreu nos outros PGVs (item 2.7, página 80).

Além do número de contêineres, muitos autores incluíram variáveis *dummy* para especificar o dia da semana e o mês do ano analisado nos seus modelos, visando captar a sazonalidade semanal e mensal. Menos comumente, também foram consideradas como variáveis independentes: a média de armazenamento de TEUs por caminhão, a

quantidade de navios que chegam/saem do porto, melhoria de infraestrutura em portos vizinhos concorrentes, PIB, alteração no preço da gasolina, e variável binária para tratar eventos fora de classificação.

Os estudos realizados revelaram que as variáveis de tempo, como dia da semana, mês e atrasos, são significativas em praticamente todos os modelos desenvolvidos.

Cabe ressaltar que todos os modelos analisados foram aplicados em países do exterior, sendo que não se encontrou modelos de geração de viagens desenvolvidos para portos marítimos no Brasil, e tampouco uma comparação entre o uso de uma regressão de Poisson ou distribuição binomial negativa com uma regressão linear ou redes neurais para esta realidade.

Baseada nessa análise, neste trabalho será proposto o uso dos seguintes modelos:

3.5.1 Regressão linear múltipla

A análise de regressão é o estudo de uma variável dependente em função de uma variável independente, no caso da regressão linear simples, ou mais variáveis independentes, no caso da regressão linear múltipla, com o objetivo de estimar e/ou prever a média populacional ou valor médio da variável dependente, utilizando valores observados por amostragem das variáveis independentes (GUJARATI, 2000); ou seja, o resultado de ambas as regressões é uma equação linear que pode ser utilizada para estimar a variável dependente como uma combinação linear das variáveis independentes.

Na Equação 15 é apresentada a função de regressão linear simples e, na Equação 16, de regressão linear múltipla.

$$y = \beta_0 + \beta_1x + \varepsilon \quad (15)$$

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p + \varepsilon \quad (16)$$

Onde:

y = variável dependente, variável explicada ou variável resposta;

x_j = variáveis independentes, variáveis explicativas, variáveis regressoras ou covariáveis (j = 1, 2, ..., p);

β_j = parâmetros da população; e
 ε = erros aleatórios.

Observa-se que ambas as funções são lineares devido ao fato de que nenhum de seus parâmetros aparece com expoente ou é multiplicado ou dividido por outros parâmetros.

Os parâmetros da população – β_j – são os coeficientes de regressão, no qual β_0 é o intercepto da reta no eixo da ordenada ou superfície de resposta (regressão) e $\beta_1, \beta_2, \dots, \beta_p$ é a inclinação da reta. Isso significa, que se a abrangência do modelo inclui $(0, 0, \dots, 0)$ então β_0 representa a resposta média $E(Y)$ neste ponto (em outras situações, β_0 não tem qualquer outro significado como um termo separado no modelo de regressão); e que podemos interpretar β_j como a mudança esperada em y devido ao aumento de uma unidade em x_j , estando as outras variáveis x_k , $k \neq j$, fixas.

Para estimar esses coeficientes de regressão, uma das técnicas mais usadas é a de mínimos quadrados. Essa técnica estima os parâmetros de regressão de forma a minimizar as somas quadráticas dos resíduos (MONTGOMERY et al., 2001; NETER et al., 2005; WEISBERG, 2005). Os estimadores determinados por essa técnica são não viesados e consistentes (NETER et al., 2005).

O termo ε é o erro aleatório e reflete a possível interferência de inúmeros fatores nas observações da variável dependente y (CHARNET et al., 1999). Assume-se que esse termo como normal e independentemente distribuído, com média zero e variância (σ^2) constante e desconhecida.

3.5.2 Regressão de Poisson e distribuição binomial negativa

Modelos de contagem são utilizados quando o resultado de interesse é um número inteiro não negativo, ou uma contagem, denotada por y , e $y \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$. O objetivo é analisar y em uma configuração de regressão, dando um vetor de K covariáveis, x . Uma vez que a variável resposta é discreta, esta distribuição coloca a densidade de probabilidade somente em valores inteiros não negativos. Formulações totalmente paramétricas de modelos de contagem acomoda essa propriedade de distribuição (CAMERON & TRIVEDI, 2009).

Regressões contáveis são não lineares; $E(y | x)$ é usualmente uma função não linear, mais comumente uma função de índice único, como $\exp(x'\beta)$. Vários recursos especiais de modelos de regressão de contagem são intimamente conectados com valores discretos e não linearidade (CAMERON & TRIVEDI, 2009).

Algumas das complicações padrões na análise de dados de contagem incluem as seguintes: presença de heterogeneidade não observada semelhante a variáveis omitidas; excesso de zeros, o que frequentemente resulta em superdispersão dos dados (a variância é maior que a média); truncamento na distribuição observada y ; e, regressores endógenos (CAMERON & TRIVEDI, 2009).

O ponto natural de início para análises de contagem é a distribuição de Poisson e o modelo de Poisson, que é frequentemente usado para modelar o número de ocorrências de um evento por um certo período de tempo ou por um certo volume ou por uma certa área. A distribuição de Poisson univariada, denotada por Poisson ($y | \mu$), para número de ocorrências do evento y sobre um período de exposição fixo tem a função de densidade de probabilidade, segundo Cameron & Trivedi (2009):

$$\Pr(Y = y) = \frac{e^{-\mu} \mu^y}{y!}, y = 0, 1, 2, \dots \quad (17)$$

Onde μ é a intensidade ou taxa do parâmetro. Os dois primeiros momentos são de acordo com Cameron & Trivedi (2009):

$$E(Y) = \mu \quad (18)$$

$$\text{Var}(Y) = \mu \quad (19)$$

Isso mostra a bem conhecida igualdade de propriedade da média e variância, também chamada de propriedade de equidispersão da distribuição de Poisson (CAMERON & TRIVEDI, 2009).

A parametrização média padrão é $\mu = \exp(x'\beta)$ para garantir que $\mu > 0$. Isso implica que o modelo é intrinsecamente heteroscedástico (CAMERON e TRIVEDI, 2009).

A propriedade de equidispersão (igualdade da média e variância) é comumente violada em trabalhos práticos, porque a superdispersão (a variância é maior que a média) é frequente. Assim a variância

(condicional) excede a média (condicional). Tal dispersão adicional pode ser considerada de várias maneiras, sendo a presença de heterogeneidade não observada uma das mais habituais (CAMERON & TRIVEDI, 2009).

Heterogeneidade não observada, que produz variabilidade adicional em y , pode ser gerada ao introduzir aleatoriedade multiplicativa. Substitui-se μ por μv , onde v é uma variável aleatória, consequentemente, $y \sim \text{Poisson}(y \mid \mu v)$. Supondo que se especifique v como $E(v) = 1$ e $\text{Var}(v) = \sigma^2$. Em seguida, é fácil demonstrar que v preserva a média, mas aumenta a dispersão. Especificamente, $E(y) = \mu$ e $\text{Var}(y) = \mu(1 + \mu\sigma^2) > E(y) = \mu$. O termo “sobredispersão” descreve a característica $\text{Var}(y) > E(y)$, ou mais precisamente $\text{Var}(y \mid x) > E(y \mid x)$, em um modelo de regressão (CAMERON & TRIVEDI, 2009).

No caso especial bem conhecido que $v \sim \text{Gama}(1, \alpha)$, em que α é o parâmetro da variância de distribuição gama, a distribuição marginal de y é uma mistura de Poisson-gama com uma forma fechada – a distribuição Binomial Negativa (*Negative Binomial* – NB) denotado por NB (μ, α) - cuja função massa de probabilidade é, segundo Cameron & Trivedi (2009):

$$\Pr(Y = y \mid \mu, \alpha) = \frac{\Gamma(\alpha^{-1} + y)}{\Gamma(\alpha^{-1})\Gamma(y + 1)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu} \right)^{\alpha^{-1}} \left(\frac{\mu}{\mu + \alpha^{-1}} \right)^y \quad (20)$$

Onde $\Gamma(\cdot)$ denota a integral de gama que se especializa em um fatorial de um argumento inteiro. O modelo NB é mais geral do que o modelo de Poisson, porque acomoda sobredispersão e reduz ao modelo de Poisson como $\alpha \rightarrow 0$ (CAMERON & TRIVEDI, 2009).

O modelo de regressão NB permite $\mu = \exp(x'\beta)$ e o deixa como uma constante. A opção padrão para a regressão NB na maioria dos softwares estatístico é a versão com uma variação quadrática (NB2). Os momentos do NB2 são $E(y \mid \mu, \alpha) = \mu$ e $\text{Var}(y \mid \mu, \alpha) = \mu(1 + \alpha\mu)$. Empiricamente, a função de variância quadrática é uma aproximação versátil em uma ampla variedade de processos de dados com sobredispersão (CAMERON & TRIVEDI, 2009). Outra variante do NB na literatura tem uma função linear da variância, $\text{Var}(y \mid \mu, \alpha) = (1 + \alpha)\mu$, e é chamado de modelo NB1 (CAMERON & TRIVEDI, 2009).

Dada a $\mu = \exp(x'\beta)$, e a suposição de que as observações ($y_i \mid x_i$) são independentes, a Máxima Verossimilhança (*Maximum Likelihood – ML*) de Poisson é muitas vezes o ponto de partida de um exercício de modelagem, especialmente se toda a distribuição e não apenas a média condicional é o objeto de interesse (CAMERON & TRIVEDI, 2009).

Dados de contagem são muitas vezes sobredispersos. Uma abordagem é manter a premissa da média condicional $E(y \mid x) = \exp(x'\beta)$. Em seguida, pode-se continuar a usar o máximo estimador de verossimilhança do Poisson, que mantém a sua consistência, mas relaxa a suposição de equivariância para obter uma estimativa robusta da matriz de variância-covariância do estimador. Alternativamente, o modelo de NB, que explicita a sobredispersão de modelos, pode ser usado com a estimativa de ML (CAMERON & TRIVEDI, 2009).

3.5.3 Redes neurais artificiais

As RNAs são utilizadas para resolver problemas práticos por meio de uma espécie de simulação do funcionamento do cérebro humano. Essa simulação do cérebro humano é possível graças à arquitetura das RNAs, que consiste em sistemas paralelos distribuídos, compostos por unidades de processamento simples (neurônios) que calculam determinadas funções matemáticas, que normalmente não são lineares (ALVES; SILVA; WAERDEN, 2012).

Conforme já citado no Capítulo 3, uma das principais razões para a popularidade das RNAs é que elas têm uma capacidade muito forte de aproximação da função. Em contrapartida, a dificuldade dessa técnica diz respeito às muitas decisões que tem de ser tomadas quanto à formação e seleção do modelo a fim de que o modelo obtido seja válido.

Com o intuito de auxiliar na tomada dessas decisões, Gonçalves (2015) desenvolveu um procedimento para aplicar RNAs para problemas na área de planejamento de transportes, especificamente quanto à previsão de demanda/viagens, a partir da identificação de softwares, objetivos, número de neurônios na camada intermediária, função de normalização, variáveis de entrada, número de camadas ocultas, algoritmos de aprendizagem utilizados e a proporção dos dados utilizados no treinamento, validação e teste utilizados em diversos trabalhos que utilizaram RNAs nessa área.

A primeira etapa do procedimento proposto por Gonçalves consistiu na definição das variáveis independentes que serão utilizadas como entrada para a RNA.

Na segunda etapa deve-se determinar a quantidade de camadas ocultas na RNA, cuja topologia é composta por uma camada de saída, uma de entrada e uma ou mais ocultas. Além disso, nessa etapa também se deve definir o algoritmo de aprendizado e a função de ativação que serão utilizados.

Para as aplicações em previsões de viagens, o autor sugere ainda avaliar o comportamento da RNA para uma camada oculta, utilizando como função de ativação a tangente hiperbólica. Isso porque, segundo o autor: testes preliminares não mostraram ganho de desempenho com mais de uma camada oculta e com outras funções de ativação, além do que Batista (2012) ressalta que o aumento do número de camadas acarreta o aumento da complexidade e do tempo de processamento da rede; e Aguiar Júnior (2004) afirma que a melhor opção para função de ativação dos neurônios, tanto na camada intermediária quanto na camada de saída, é a função da tangente hiperbólica.

Dentre diversas variações do algoritmo *backpropagation*, Gonçalves sugere a utilização do Algoritmo de Levenberg-Marquardt (*Levenberg-Marquardt Algorithm* – LMA), que segundo Araújo & Melo (2006) e Yuehua et al. (2008) caracteriza-se como um dos métodos mais rápidos para o treinamento de RNA Acíclicas de tamanho moderado, e é eficaz na minimização de funções com um pequeno número de parâmetros e o tempo de convergência da rede tende a ser menor quando comparado aos outros modelos e possui elevada precisão.

Depois disso, na etapa seguinte do procedimento deve-se determinar a proporção dos dados a serem utilizados no treinamento, validação e teste. Nesse quesito, Gonçalves recomenda que os dados sejam divididos aleatoriamente em 70% para treinamento, 15% para teste e 15% para validação.

Por fim, a última etapa consiste na determinação da quantidade de neurônios que a camada oculta deve possuir para o conjunto de treinamento. Para isso, Gonçalves recomenda que sejam testados um intervalo de 1 a 20 neurônios por diversas vezes cada um até gerar seu melhor resultado, que deve ser avaliado por meio dos respectivos R^2 do subconjunto de validação; e, após isso, deve ser identificado o número de neurônios em que o R^2 se estabilizou e, no qual, a partir desse ponto, os ganhos de precisão foram pequenos.

Como software para desenvolver as RNAs, Gonçalves propõe o uso do Matlab, visto que além de oferecer um ambiente completamente amigável, este admite a utilização de ferramentas úteis ao desenvolvimento intuitivo de RNAs (ARAÚJO & MELO, 2006).

3.5.4 Aprendizado de máquina baseado em *kernel*

Em aprendizado de máquina, o modelo resultante é definido por meio de alguns parâmetros, sendo que é utilizada a programação de computadores para otimizar esses parâmetros por meio de um conjunto de dados de treinamento (RODRIGUES, 2009).

A aprendizagem da máquina utiliza teorias estatísticas para a construção de modelos matemáticos, uma vez que a principal finalidade é fazer inferências a partir de uma amostra. A computação se faz presente, primeiramente, no treino, no qual se precisa de algoritmos eficientes para a resolução de problemas de otimização, assim como para lidar com uma quantidade considerável de dados; e, em segundo lugar, uma vez criado o modelo, a sua representação e solução algorítmica para a inferência também precisa ser eficiente (RODRIGUES, 2009).

Particularmente, o *support vector machine* é um conjunto de métodos supervisionados de aprendizagem utilizado para regressão e previsão (RODRIGUES, 2009). De acordo com Rodrigues (2009), o algoritmo original foi proposto por Vapnik & Chervonenkis (1963), e a técnica se baseia na procura de uma fronteira ou um hiperplano que separe um conjunto de dados em classes discretas, utilizando-se de processo iterativo e de exemplos de treinamento para ajustar este hiperplano (PEREIRA, 2013). Para a representação deste hiperplano podem ser usadas funções lineares ou até mesmo funções *kernel*, sendo as mais utilizadas as de base polinomial, Radial (*Radial Basis Function* – RBF) e *kernel* sigmóide (PETROPOULOS, ARVANITIS & SIGRIMIS, 2012).

De modo geral, segundo Foody & Mathur (2004), o SVM é fundamentado na noção de se formar um hiperplano ótimo de separação entre classes, focando-se nas amostras de treino que se encontram na extremidade das distribuições de classe, os vetores suporte. Tendo isso em vista, chega-se à conclusão de que a classificação de dados pode ser feita a partir de uma amostra menor do que a amostra necessária para redes neurais artificiais, uma vez que apenas as amostras de treino que se

encontram nas fronteiras entre as classes são necessárias para a discriminação (PEREIRA, 2013).

Outra vantagem do método é que o mesmo converge para uma única solução ótima, ao contrário dos modelos de RNA. Além disso, o SVM é um método não paramétrico e, portanto, não assume uma distribuição estatística conhecida a priori dos dados a serem analisados. Dessa maneira, pode-se superar técnicas com base em máxima verossimilhança, pois a normalidade nem sempre dá uma suposição correta da distribuição real dos dados, que muitas vezes podem ter distribuições desconhecidas (PEREIRA, 2013).

O desafio de sua aplicação diz respeito ao fato de que o SVM é um classificador binário, ou seja, permite classificar os dados em apenas duas categorias; e, isso muitas vezes não condiz com a realidade de muitos conjuntos de dados que necessitam de mais divisões de classificação. No entanto, abordagens utilizando a redução do problema multiclasse a um conjunto binário têm sido utilizadas para resolver essa questão (PEREIRA, 2013).

Outro método aprendizado de máquina muito utilizado se baseia em GP. Considerando a entrada como x e a saída como y e um conjunto de dados D de n observações, ou seja, $D = \{(x_i, y_i) \mid i = 1, \dots, N\}$, tem-se como objetivo fazer previsões para novas entradas x^* que não tenham sido visualizadas no conjunto de treinamento. Para isso, é necessário fazer suposições sobre as características da função subjacente, pois senão qualquer função que fosse consistente com os dados de treinamento seria igualmente válida. Em outras palavras, há um conjunto enorme de funções possíveis e, para solucionar esse problema, é utilizado o processo de Gauss (RASMUSSEN & WILLIAMS, 2006).

Um processo de Gauss é uma generalização da distribuição de probabilidade Gaussiana. A distribuição de probabilidade descreve variáveis aleatórias que são escalares ou vetores (para distribuições multivariadas), no qual um processo estocástico rege as propriedades das funções. De modo simplificado, pode-se imaginar uma função como um vetor muito longo, onde cada entrada no vetor especifica o valor da função $f(x)$ a uma entrada especial x . Neste caso, se forem solicitadas somente as propriedades da função para um número finito de pontos, a inferência do processo de Gaussian dá como resposta a mesma que daria caso tivesse levado em conta os pontos infinitos (RASMUSSEN & WILLIAMS, 2006).

Sob o ponto de vista processo de Gauss, os modelos podem ser mais fáceis de manusear e interpretar do que seus homólogos, tais como, por exemplo, redes neurais. Em grosso modo, um processo estocástico é uma generalização de uma distribuição de probabilidade (que descreve uma variável aleatória finito-dimensional) para funções. Ao concentrar-se em processos que são Gaussinos, verifica-se que os cálculos necessários para inferência e aprendizado tornam-se relativamente fáceis (RASMUSSEN & WILLIAMS, 2006).

3.5.4.1 Funções kernel

Considere o modelo de regressão $y = m(x) + u$, onde x é um escalar e a função média condicional $m(\cdot)$ não é especificada. A estimativa de regressão local $m(x)$ em $x = x_0$ é uma média local ponderada de y_i , $i = 1, \dots, N$, que dá grande peso em observações para as quais x_i está perto de x_0 e pouco ou nenhum peso sobre observações para as quais x_i está longe de x_0 , formalmente, segundo Cameron & Trivedi (2009):

$$\hat{m}(x_0) = \sum_{i=1}^N w(x_i, x_0, h) y_i \quad (21)$$

Em que os pesos $w(x_i, x_0, h)$ somam sobre i para um e diminuem à medida que a distância entre o x_i e x_0 aumenta. Como a largura de banda de parâmetros h aumenta, mais peso é colocado nas observações para as quais x_i está perto de x_0 (CAMERON & TRIVEDI, 2009).

A plotagem é obtida pela escolha de uma função de ponderação, $w(x_i, x_0, h)$; a escolha de uma largura de banda, h ; avaliando $m(x_0)$ num intervalo de valores de x_0 ; e traçando $m(x_0)$ contra estes valores x_0 (CAMERON & TRIVEDI, 2009).

Por exemplo, os métodos de aprendizado de máquina baseado em *kernel* são utilizados para estimar curvas de densidades onde cada observação é ponderada pela distância em relação a um valor central, o núcleo.

Neste caso, regressões de *kernel* utilizam o peso $w(x_i, x_0, h) = K\{(x_i - x_0)/h\} / \sum_{i=1}^N K\{(x_i - x_0)/h\}$, onde $K(\cdot)$ é uma função de *kernel* definida posteriormente (CAMERON & TRIVEDI, 2009). Uma

gama de funções do *kernel* são comumente usadas dependendo do objetivo:

- a) função linear: é utilizada como um teste da não-linearidade no conjunto de treino, assim como uma base para a eventual melhoria da classificação obtida com *kernels* não lineares (RODRIGUES, 2009);

$$\Phi = x_i \cdot x_j \quad (22)$$

Onde:

x = variáveis (x1, x2, ..., xn)

- b) função polinomial: é usada para a modelação de relações não lineares. Sua vantagem é que é um método simples e eficiente. No entanto, sua desvantagem é o sobreajuste que pode aparecer quando o grau da equação aumenta, visto que a superfície de classificação torna-se mais complexa (RODRIGUES, 2009);

$$\Phi = (x_i \cdot x_j + 1)^d \quad (23)$$

Onde:

d = grau do polinômio

x = variáveis (x1, x2, ..., xn)

- c) função básica radial: neste caso, os parâmetros são otimizados utilizando o método da validação cruzada (RODRIGUES, 2009); e

$$\Phi = \exp(-\gamma \|x_i \cdot x_j\|^2) \quad (24)$$

Onde:

γ = coeficiente

x = variáveis (x1, x2, ..., xn)

- d) função sigmóide: é a mais utilizada função de transferência para redes neurais e consta em uma função tangente hiperbólica (\tanh) com uma forma sigmoidal, conforme mostra a Equação 25 (RODRIGUES, 2009).

$$\Phi = \tanh(\gamma x_i \cdot x_j + c) \quad (25)$$

Onde:

c = constante

γ = coeficiente

$x = (x_1, x_2, \dots, x_n)$

4 MÉTODO DA PESQUISA

O desenvolvimento do método proposto nesta tese foi embasado na revisão da literatura referente aos métodos empregados para estimar as viagens geradas pelos PGVs de um modo geral e, especificamente, de carga, bem como nas técnicas e nas variáveis utilizadas nas construções dos modelos para portos marítimos.

Como diferencial, o método proposto buscou alcançar um maior detalhamento do processo de desenvolvimento de um modelo de geração de viagens para um porto, assim como englobar todas as opções encontradas tanto de técnicas de modelagem quanto das variáveis utilizadas com o intuito de facilitar sua utilização.

Assim sendo, propõe-se um método a ser desenvolvido em sete etapas, como apresentado no fluxograma da Figura 4.1. As etapas do fluxograma propostas são descritas nos itens a seguir.

Figura 4.1 – Fluxograma do método

(continua)

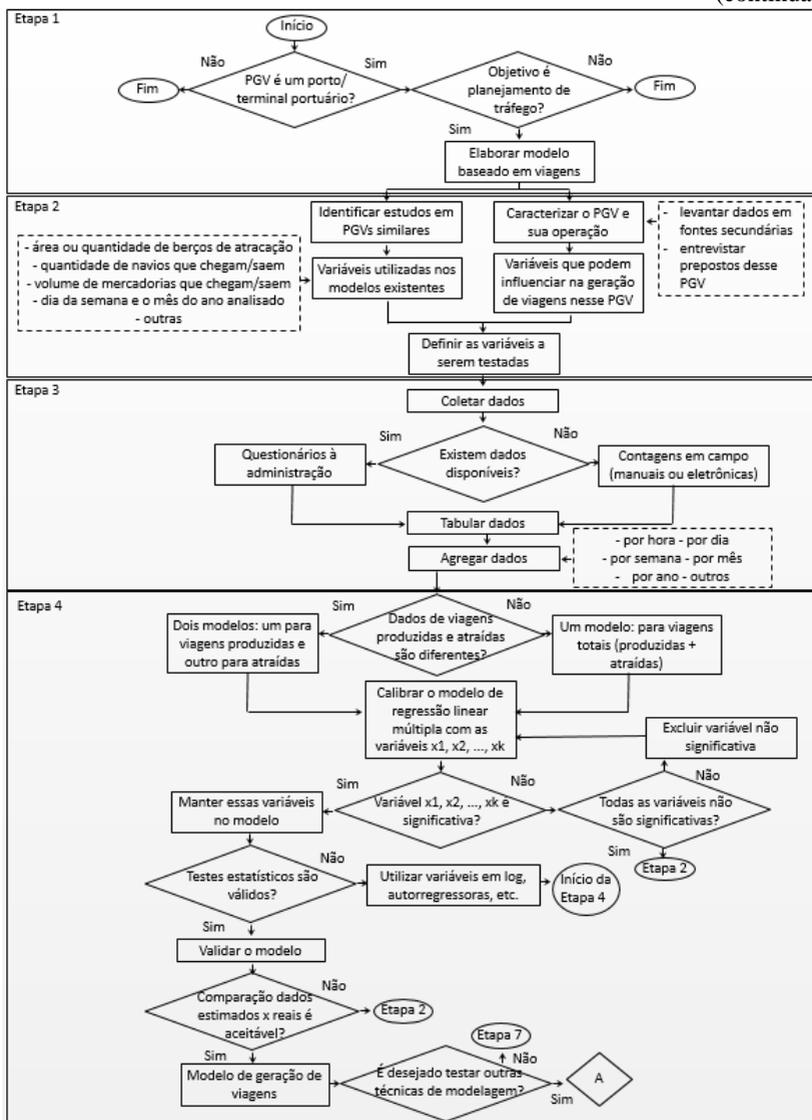
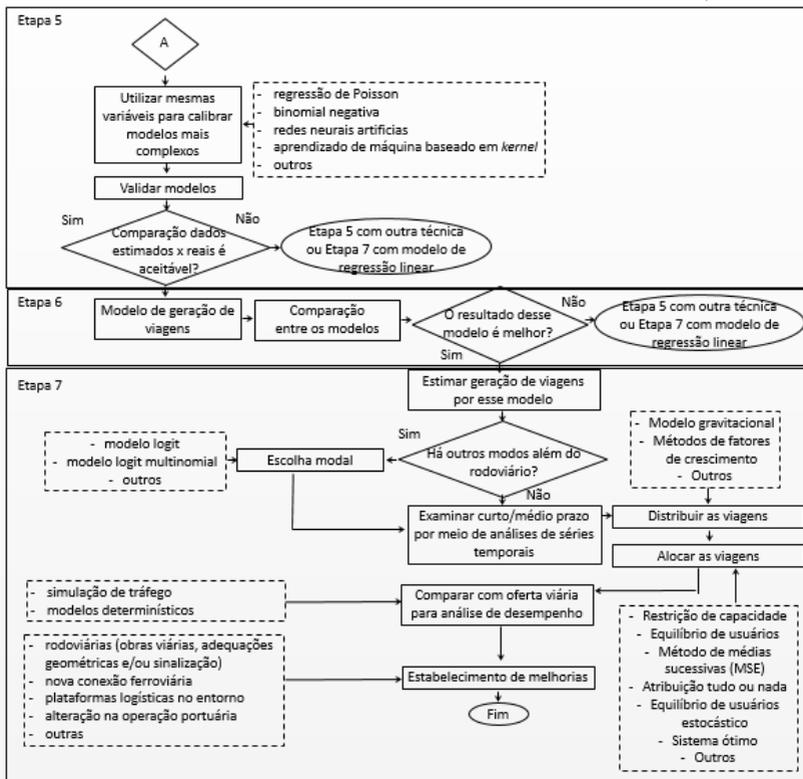


Figura 4.1 – Continuação do fluxograma do método

(conclusão)



Fonte: Autoria Própria.

4.1 ETAPA 1

A etapa 1 tem início com uma pergunta: o PGV que se pretende analisar é um porto² ou terminal portuário³? Caso a resposta seja positiva, se dá prosseguimento ao método, se não, o método é finalizado nesse ponto, visto que seria necessário buscar um método para estimar modelos de geração de viagens para especificamente o tipo de PGV em questão.

No passo seguinte é feita mais uma pergunta: o objetivo do estudo é planejamento de tráfego de entorno do PGV? Se sim, segue-se com o método apresentado; se não, a utilização desse método finaliza nesse ponto, uma vez que seria necessário buscar métodos de elaboração de modelos baseados em fluxo de carga, conforme citado no subitem 2.6.10, na página 70. Nesse caso, pode-se utilizar séries históricas para projetar o volume de contêineres e, posteriormente, transformar em volumes de caminhões, sabendo-se a média de armazenamento de TEUs por caminhão, como foi realizado no Plano Mestre do Porto de Itajaí (SC), apresentado no item 5.3, página 158.

Se o objetivo é o planejamento de tráfego de entorno de um porto ou terminal marítimo, há a necessidade, então, de elaboração de um modelo de geração baseado em viagens e pode-se seguir para a etapa 2 do método.

4.2 ETAPA 2

Para elaborar o modelo de geração baseado em viagens, a etapa 2 prevê a identificação de estudos existentes sobre modelos de geração de

² É a denominação geral dada ao complexo composto por vários terminais, equipamentos portuários, fundeadouro, canal de acesso ao porto, vias de acesso ferroviárias, rodoviárias e fluviais, etc. É sempre um lugar de uma costa marítima ou de um rio ou lago, que oferece às embarcações as condições para entrada e fundeio.

³ É a unidade menor em que é dividido um porto, e que é administrada independentemente por um operador portuário privado ou, quando a administração é estatal, pela União, Estado ou Município. Pode ser composto por um ou mais berços de atracação. São normalmente especializados em tipos de cargas ou de navios.

viagens para portos/ terminais marítimos, com a finalidade de se verificar quais são as variáveis independentes utilizadas nesses estudos.

Na revisão bibliográfica realizada no Capítulo 3, verificou-se que os estudos existentes acerca dessa temática fazem uso, principalmente, do volume de mercadorias que chegam/ saem do porto/ terminal marítimo e do dia da semana e mês do ano analisado. Também foram utilizadas, menos frequentemente, variáveis relacionadas à área do empreendimento, ao cronograma de chegada e saída de navios, entre outras conforme mostrado no item 3.5, na página 125. Essas variáveis encontradas na revisão bibliográfica estão apresentadas no fluxograma do método (Figura 4.1).

Além da identificação de estudos existentes sobre o assunto, deve-se buscar encontrar informações acerca do PGMV objeto de estudo e de sua operação – tais como: localização, área, acessos rodoviários, infraestrutura de acostagem, estruturas de armazenagem, movimentação portuária, especificidades de operação, etc. – por meio de consulta, por exemplo, nos endereços eletrônicos dos mesmos e de órgãos relacionados aos mesmos; ou ainda por meio de entrevistas com os prepostos do porto familiarizados com a operação, incluindo, se julgado necessário, pessoal da administração e operações de campo, e agentes de transporte, terminais privados, companhias de caminhões, segurança, contabilidade e marketing.

O objetivo dessa caracterização é adquirir informações específicas sobre como se dá a operação na prática que possam auxiliar na identificação de variáveis, não detectadas nas consultas aos estudos existentes, que influenciam no volume de caminhões atraído/ produzido para, especificamente, esse porto/ terminal marítimo analisado, bem como adquirir subsídios para posteriores interpretações dos resultados dos parâmetros dos modelos.

Então, com base nos subsídios conseguidos por meio da revisão bibliográfica e das entrevistas aplicadas aos prepostos do porto podem ser definidas as variáveis que serão testadas, inicialmente, na construção do modelo de geração de viagens e que deverão ser levantadas na etapa 3. Ressalta-se que, se no decorrer do desenvolvimento do modelo, durante a etapa 4, caso se faça necessário definir mais variáveis a serem testadas, as mesmas podem ser levantadas e acrescentadas posteriormente ao modelo.

Aqui, ainda cabe destacar as recomendações do ITE (2004) quanto à definição das variáveis independentes do modelo: aquela variável

parece ser a causa da variação da quantidade de viagens produzidas por uso do solo, o que não significa apenas haver correlação estatística; é obtida por medição primária e não ser derivada de um dado secundário; produz taxa ou equação com os melhores índices de acerto; é de fácil acesso; e é relacionada à construção e operação e não somente às características do terreno.

4.3 ETAPA 3

A etapa 3 é a etapa de coleta de dados para elaboração dos modelos de geração de viagens. A partir da definição das variáveis independentes a serem inicialmente testadas, as quais devem ser definidas na etapa anterior (etapa 2), as mesmas devem ser coletadas por meio de questionários enviados à administradores do empreendimento, caso os dados estejam disponíveis, ou de pesquisa de campo – contagens *in loco*, que podem ser manuais, pneumáticas ou por vídeo –, caso os mesmos não estejam disponíveis.

Para a realização da pesquisa de campo, deve-se definir qual o tipo de instrumento de coleta (questionário, entrevista, observação, etc.), o tamanho da amostra e o período para a coleta de dados. Em relação ao tamanho da amostra, pode-se utilizar a fórmula para cálculo do tamanho da amostra (Equação 26), de modo a garantir o grau de confiança escolhido e também a margem de erro tolerada na estimativa dos parâmetros.

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E^2} \right)^2 \quad (26)$$

Onde:

n = número de indivíduos da amostra;

$z_{\alpha/2}$ = valor crítico que corresponde ao grau de confiança desejado⁴;

⁴ $z_{\alpha/2} = 1,645$ para um grau de confiança de 90%, o qual geralmente é recomendado.

σ = desvio-padrão populacional da variável estudada; e
E= margem de erro ou erro máximo da estimativa.

A Equação 26 exige que se substitua por algum valor o desvio-padrão populacional σ , mas se este for desconhecido, podemos utilizar a aproximação $\sigma \approx \text{amplitude}/4^5$ ou ainda realizar um estudo piloto, iniciando o processo de amostragem, calcular o desvio-padrão dessa amostra S e utilizá-lo em lugar de σ .

Outra opção é buscar coletar o tamanho máximo de amostra possível, recomenda-se pelo menos seis meses de dados, e, posteriormente, analisar a margem de erro, verificando se a amostra estava de magnitude satisfatória (geralmente é utilizado valores entre 1% e 10% como margem de erro, mas podem-se aceitar valores maiores desde que esteja destacada a existência dessa margem escolhida como ressalva).

Se a resposta for negativa, pode ser usada a estimativa do desvio padrão dos resíduos de modo a verificar o tamanho da amostra necessário para que resulte em magnitudes aceitáveis. Quando possível, se faz novo levantamento de dados de modo a complementar o anterior. Ressalta-se ainda que há técnicas que necessitam de amostras maiores para obterem precisão adequada e superaram erros de *overfitting* ou *overtraining*, como é o caso de determinados tipos de RNAs e de aprendizado de máquina baseado em *kernel*.

Após coletados os dados, estes devem ser tabulados em planilhas de Excel, com cada variável sendo alocada em uma coluna, caso já não se preparados para serem usados nas análises preliminares de modelos de regressão linear múltipla por meio de processo iterativo. Essa preparação pode incluir o agrupamento dos dados segundo o período para o qual se pretende obter as estimativas de viagens, seja por hora, dia, semana, mês ou ano. O mais comum é que sejam agrupados por dia, conforme mostra os resultados da revisão bibliográfica sistemática sobre o assunto (item 3.5, na página 125).

⁵ Esse valor corresponde ao valor encontrado empiricamente na estatística no qual resultou em maior ou igual ao real σ pelo menos 95% das vezes.

4.4 ETAPA 4

A etapa 4 corresponde a formulação do modelo de regressão linear múltipla, com o objetivo de mensurar o impacto das variáveis independentes definidas no volume de viagens geradas pelo porto/terminal marítimo. Caso não haja nenhuma diferença significativa entre o volume de viagens atraídas e produzidas pelo empreendimento, o que pode ser verificado, como por exemplo, por meio de um teste-t de student de amostras pareadas, pode-se fazer apenas um modelo contemplando o volume total de viagens. No entanto, caso haja diferença nesse volume de viagens atraídas e produzidas, deve-se elaborar modelos separados para cada direção de geração (atração e produção).

Assim, deve ser feita a inserção desses dados em um software estatístico que irá processar os cálculos para calibração do modelo, ou seja, o ajuste dos parâmetros do modelo aos dados.

Por meio do resultado do teste-t de student das variáveis independentes incluídas no modelo pode-se identificar se as variáveis independentes são estatisticamente significantes. Caso sejam identificadas variáveis independentes estatisticamente não significantes, deve-se excluí-las e calibrar o modelo novamente. Caso nenhuma variável independente tenha resultado em estatisticamente significativa, deve-se voltar, então, à etapa 2 para definição de novas variáveis.

Estando o modelo calibrado com as variáveis que se mostraram estatisticamente significativas, deve-se fazer a verificação se o mesmo cumpre as hipóteses da regressão linear, por meio dos seguintes testes estatísticos: teste de não-linearidade, teste de erro de especificação da regressão (*Regression Specification Error Test* – Reset), teste de White para a heteroscedasticidade, teste da normalidade dos resíduos, teste de Chow para a falha estrutural e teste Multiplicador de Lagrange (*Lagrange Multiplier* – LM) para autocorrelação.

Se o modelo não “passar” em um ou mais de um teste, deve-se ser feita alguma alteração, como por exemplo, utilizar as variáveis em logaritmo, acrescentar variável dependente autoregressora como variável independente, entre outras. Caso mesmo assim, o modelo continue não cumprindo as premissas de uma regressão linear, deve-se voltar à etapa 2 para verificação se as variáveis independentes escolhidas são adequadas para explicar o modelo ou se há outras variáveis importantes que não foram consideradas.

Caso o modelo cumpra todas as premissas, segue-se para a validação do mesmo, no qual, por meio da equação de regressão linear múltipla encontrada, deve ser estimado o volume de viagens e este deve ser comparado aos volumes reais medidos. Outra maneira de fazer essa verificação é por meio de parâmetros estatísticos como R^2 , Mape, Mdape, entre outros.

Se a diferença for aceitável (pode-se utilizar como critério, por exemplo, R^2 acima de 0,58 que foi o menor valor apresentado pelo ITE (2012) para portos/ terminais marítimos), tem-se, assim, um modelo aceitável de geração de viagens por meio da técnica de regressão linear múltipla. Caso contrário, volta-se a etapa de especificação do modelo de modo a identificar novas variáveis que consigam melhor explicar a geração de viagens de caminhões em portos/ terminais marítimos.

Nesse momento, pode-se simplesmente pular para a etapa 7, caso não se deseje testar outras técnicas de modelagem com o intuito de encontrar um modelo parcimonioso, ou pode-se continuar com a sequência normal do método, caso se opte por testar outra ou outras técnicas.

4.5 ETAPA 5

De acordo com a revisão bibliográfica sistemática sobre o assunto (Capítulo 3), além da regressão linear, há outras técnicas que estão sendo atualmente utilizadas para a elaboração de modelos de geração de viagens para portos/ terminais marítimos, tais como: regressão de Poisson e distribuição binomial negativa, RNAs e SVMs.

Portanto, caso se opte por desenvolver o modelo de geração de viagens com outra técnica, o modelo com a técnica escolhida deve ser, então, calibrado considerando as mesmas variáveis que resultaram estaticamente significativas durante a construção do modelo de regressão linear múltipla. Após calibração do mesmo, devem-se ser feitos, caso necessário, os testes estatísticos para estes, assim como a validação do mesmo como é proposto na etapa 4 no desenvolvimento do modelo de regressão linear múltipla (item 4.4, na página 151).

Se o resultado for aceitável, pode-se passar para a etapa de comparação dos modelos ou, ainda, caso se deseje, pode-se testar outras técnicas. Todavia, se o resultado não for aceitável, pode-se iniciar novamente a etapa 5, testando outras técnicas, ou ainda continuar com a

seqüência do método – etapa 6 – por meio apenas do modelo de regressão linear múltipla encontrado na etapa 4.

4.6 ETAPA 6

Caso se o desenvolvimento das etapas anteriores tenha resultado em mais de um modelo válido, pode ser feita uma comparação entre eles do desempenho de suas previsões e utilizar o modelo com melhor resultado para elaborar, assim, as estimativas de viagens.

A comparação entre si dos modelos pode ser realizada com base em parâmetros estatísticos, como por exemplo, maior R^2 , menor MSE, menor Mape, menor Mdape, menor U de Theil, ou outros.

4.7 ETAPA 7

Na etapa 7, deve-se estimar o volume de viagens de caminhões geradas pelo porto por meio do modelo que obteve melhor precisão. Tendo-se essas estimativas, pode-se fazer uma projeção delas para o período que se pretende analisar, assim como distribuí-las e aloca-las na rede.

Ressalta-se que caso haja mais de um modo de transporte envolvido no deslocamento de carga, deve-se ser realizada a divisão modal das viagens antes da etapa de projeção, distribuição e alocação das viagens. A divisão modal procura determinar qual modo de transporte será utilizado para um determinado número de viagens, para isso são utilizados modelos determinísticos ou probabilísticos, sendo no último caso muito usado o modelo logit.

Com as viagens estimadas para o modo rodoviário e, se for o caso, também para os outros modos, pode-se examinar tendências de curto, médio e longo prazo por meio de análises de séries temporais. Essas análises podem ser aplicadas, por exemplo, no total mensal da principal variável independente (como, volume de carga), baseadas nos dados históricos obtidos para essa variável; e insere-se, então, esses novos valores da variável independente no modelo para estimar a variável dependente “volume de viagens geradas pelo PGV”.

Com a estimativa dos futuros volumes de viagens, deve-se distribuí-los, fazendo uso de modelos gravitacionais ou métodos de

fatores de crescimento, e aloca-los nos diversos segmentos da rede, fazendo uso de métodos de restrição de capacidade, equilíbrio de usuários, método de médias sucessivas, atribuição tudo ou nada, equilíbrio de usuários estocástico ou sistema ótimo.

Posteriormente, deve-se fazer o levantamento da oferta dessa rede (capacidade do sistema viário) para seja possível fazer uma comparação entre o nível de desempenho da situação atual (sem porto ou sem ampliação do porto, por exemplo) com o nível de desempenho resultante da soma de viagens da situação atual mais a previsão de viagens obtida com os modelos encontrados na etapa anterior referente a situação futura (novo porto ou ampliação do porto, por exemplo). Isso pode ser feito através de técnica de simulação de tráfego ou com o uso de modelos determinísticos conhecidos na literatura.

Assim, é possível avaliar o impacto que o PGV pode causar na rede viária. Conforme o nível de desempenho encontrando, sugerem-se alternativas para mitigar ou reduzir esse impacto, como por exemplo: implementação de obras viárias, adequações geométricas e/ou melhorias na sinalização; implementação de nova conexão ferroviária; implementação de plataformas logísticas no entorno do PGV; alteração na operação portuária; entre outras.

5 APLICAÇÃO DO MÉTODO

Neste Capítulo é realizada a aplicação do método proposto no porto de Itajaí (SC), com localização apresentada na Figura 5.1, seguindo as mesmas etapas mostradas no Capítulo 4.

Figura 5.1 – Mapa de localização da área objeto de estudo (porto de Itajaí/SC)



Fonte: PortoItajaí, 2014.

5.1.1 Limites do estudo

Essa aplicação do método teve como limites:

- 1) o porto foi estudado isoladamente assim como os outros estudos com esse foco, uma vez os portos são complexos de serem analisados por possuírem características distintas, e elaborar um modelo para vários portos poderia resultar em menor precisão das estimativas de geração de viagens;
- 2) uma vez que não se fez necessário estudar a concorrência interportos, visto que o volume de mercadorias que é uma das variáveis do modelo;

- 3) o levantamento das variáveis que foram testadas no modelo correspondeu ao período de 20 meses entre 01/01/2013 e 31/08/2015, sendo que o registro do período entre 01/01/2013 e 31/12/2014 foi realizado manualmente (nesse período o porto de Itajaí ainda não tinha recursos disponíveis para fazer o registro eletronicamente) e o registro do período entre 01/01/2015 e 31/08/2015 foi realizado eletronicamente;
- 4) a unidade da variável “volume de mercadorias” considerada foi contêineres, pois corresponde a 90% do transporte de mercadorias do porto de Itajaí, para qual foi construído o modelo;
- 5) o tipo de caminhões considerado no “volume de caminhões” semirreboques porta-contêineres, pois praticamente todas as mercadorias que saem e entram do porto via acesso rodoviário utilizam esse tipo de caminhão;
- 6) a sazonalidade foi considerada no estudo por meio da inclusão no modelo de variáveis *dummy* para os dias da semana e meses do ano; e
- 7) as técnicas utilizadas para a construção dos modelos foram: regressão linear múltipla, regressão de Poisson, distribuição binomial negativa, redes neurais artificiais e aprendizado de máquina baseado em *kernel* (SVM).

5.2 ETAPA 1

Respondendo às duas perguntas iniciais do método: o PGV que se pretende analisar é um porto ou terminal portuário? Sim. O objetivo do estudo é planejamento de tráfego de entorno do PGV? Sim. Portanto, como essa aplicação tem como finalidade obter um modelo de geração de viagens para o porto de Itajaí (SC), se dá prosseguimento ao método por meio da etapa 2 do mesmo.

5.3 ETAPA 2

Além da identificação de estudos existentes sobre o assunto, a qual já foi realizada no Capítulo 3, na etapa 2 deve-se buscar encontrar informações acerca do PGV objeto de estudo e de sua operação por meio

de consulta em fontes secundárias ou ainda por meio de entrevistas com os prepostos do porto familiarizados com a operação.

Para a aplicação dessa tese no porto de Itajaí (SC), para o qual foram construídos os modelos de geração de viagens para caminhões, as informações obtidas para esse porto por meio de pesquisa em fontes secundárias podem ser visualizadas no subitem 5.3.1 que se segue; assim como as informações obtidas por meio de entrevistas com os prepostos do porto de Itajaí (SC) familiarizados com a operação são apresentadas no subitem 5.3.2.

O objetivo dessa caracterização era adquirir subsídios para identificação das variáveis independentes a serem utilizadas na construção dos modelos de geração de viagens para esse PGV. Então, com base nesses subsídios somado ao conhecimento adquirido com a revisão bibliográfica sobre o assunto mostrado no item 3.5 foram definidas as variáveis independentes que foram testadas, inicialmente, na construção do modelo de geração de viagens, conforme mostra o subitem 5.3.3.

5.3.1 Caracterização do PGV e sua operação por meio de fontes secundárias

O complexo portuário de Itajaí está localizado as margens do Rio Itajaí-Açu, no litoral norte do Estado de Santa Catarina. A posição geográfica o coloca no centro da Região Sul, englobando, no raio de 600 quilômetros, as capitais de Santa Catarina, Rio Grande do Sul, Paraná e São Paulo, além de importantes municípios desses quatro estados (PORTOITAJAÍ, 2014).

Fazem parte desse complexo tanto o porto público de Itajaí quanto seus Terminais de Uso Privado (TUP) – Teporti Terminal Portuário Itajaí S.A., Poly Terminais S.A., Trocadeiro Terminal Portuário, Terminal Portuário Braskarne e Portonave S.A. (Terminais Portuários de Navegantes). Ressalta-se que também está projetada a instalação de mais um TUP – o Barra do Rio Terminal Portuário. A Figura 5.2 ilustra a localização do porto público de Itajaí e TUPs que fazem parte desse complexo portuário (LABTRANS, 2015).

Figura 5.2 – TUPs do complexo portuário de Itajaí (SC)



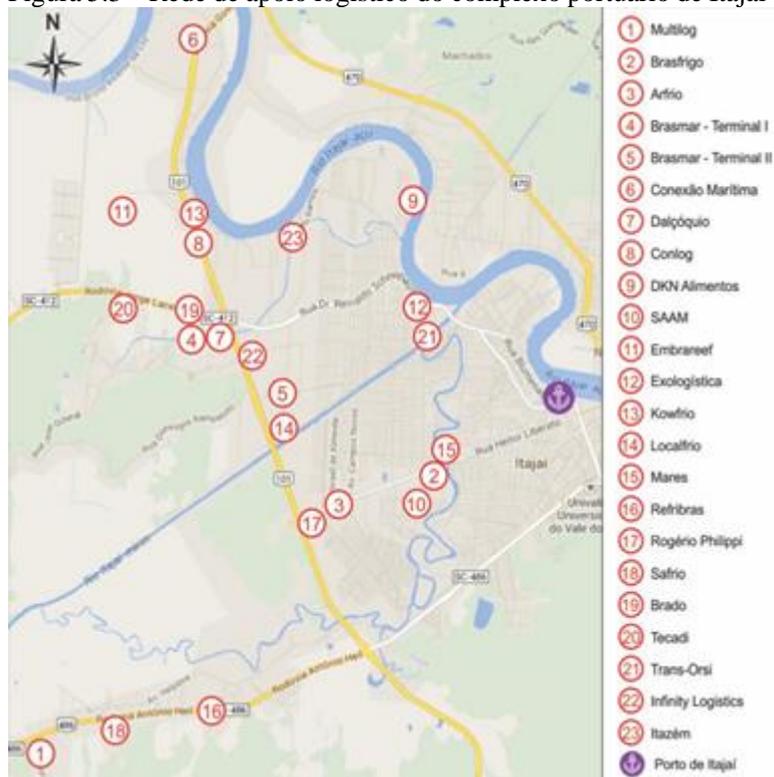
Fonte: LabTrans, 2015.

O complexo portuário de Itajaí possui uma rede de apoio logístico para suas operações, mostrada na Figura 5.3 (LABTRANS, 2015).

No total, o complexo do porto de Itajaí tem capacidade estática total para armazenagem de cargas congeladas para 160 mil toneladas ou 150 mil posições *pallets*; capacidade total de armazenagem de contêineres cheios e vazios para 140 mil TEUs; 9,6 mil tomadas para contêineres *reefer*; e, área total de 465 mil m² para armazenagem coberta, dividida em 28 armazéns (LABTRANS, 2015).

O complexo movimenta mercadorias provenientes e destinadas a 20 Estados nas cinco regiões do País e ao Distrito Federal. Ao longo dos anos recentes, as principais cargas que passaram por ele foram: madeira e derivados; frangos congelados (maior porto exportador do Brasil); cerâmicos; papel *kraft*; máquinas e acessórios; tabacos; veículos, têxteis; açúcar e carne congelada. Destaca-se a movimentação de contêineres, uma vez que o complexo portuário de Itajaí é o segundo do Brasil nesse tipo de transporte, atrás somente do porto de Santos. Com relação ao tipo de operação, as exportações, que na década passada chegaram a representar 90% das operações do complexo, hoje respondem por aproximadamente 50%, enquanto os outros 50% estão a cargo das importações (PORTOITAJAÍ, 2014).

Figura 5.3 – Rede de apoio logístico do complexo portuário de Itajaí (SC)



Fonte: LabTrans, 2015.

5.3.1.1 Histórico

Os registros históricos indicam que em 1905 foram realizados os primeiros estudos técnicos sobre o porto de Itajaí, os quais foram desenvolvidos pela “Comissão de Melhoramentos dos portos e Rios”. No entanto, foi somente após quase 10 anos, em 1914, que começaram as obras, constituídas de 700 metros lineares do molhe sul. Posteriormente, foram realizadas outras obras, incluindo as do molhe norte (PORTOITAJAÍ, 2014).

O porto, propriamente dito, teve sua construção iniciada em 1938, com um cais de 233 metros de extensão em estrutura de concreto armado,

pátios pavimentados em paralelepípedos e um armazém. Durante a década de 50, foi feita a complementação desse cais, com mais de 570 metros, resultando em um total de 803 metros. Em 1956, no ano que o prolongamento do cais foi concluído, teve início a edificação do primeiro armazém frigorífico do porto de Itajaí (PORTOITAJAÍ, 2014).

Por meio do decreto 58.780/66, em 1966, o porto em questão foi considerado porto organizado ao se criar a Junta Administrativa do porto de Itajaí, subordinada ao Departamento Nacional de Portos e Vias Navegáveis e responsável pela gestão e operação das atividades portuárias. Ressalta-se, contudo, que o gerenciamento passou a ser exercido pela administração do porto, vinculada a estatal, após a instituição da Empresa de Portos do Brasil S.A. (Portobras), em 1975 (PORTOITAJAÍ, 2014).

Na mesma época, foi erguido o prédio administrativo do porto e seu terceiro armazém e, devido à necessidade de manutenção, nos anos 80, foi realizada uma reforma e recuperação completa do cais acostável (PORTOITAJAÍ, 2014).

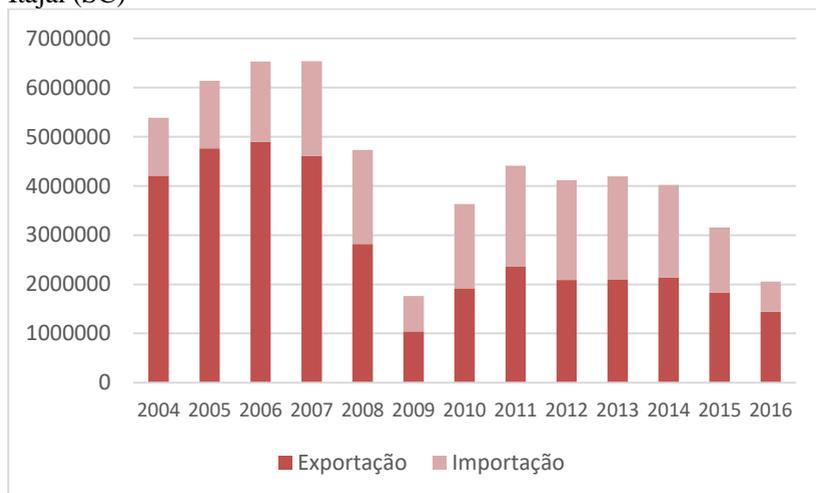
Em relação às operações, até o final da década de 60, eram alavancadas pelas cargas de madeira, trazidas das regiões Oeste, Meio Oeste e Planalto do Estado. Entretanto, a partir dessa data o ritmo de extração da madeira começou a cair, contrariando a longa tendência de crescimento desde o início do século (PORTOITAJAÍ, 2014).

Tendo isso em vista e aproveitando do bom momento econômico no País, no início dos anos 70, o porto de Itajaí se adaptou à nova realidade e diversificou suas mercadorias, inserindo a cana de açúcar, produtos congelados e, logo em seguida, os contêineres. Hoje, estes últimos são o principal motivo de operações no porto. Assim que a movimentação de cargas containerizadas aumentou, em 1990, a Portobras foi extinta, gerando uma nova crise nos portos brasileiros (PORTOITAJAÍ, 2014).

Na busca de uma solução, a administração do porto foi atrelada, então, à Companhia Docas de São Paulo; mas os problemas só foram realmente solucionados quando a gestão administrativa do porto de Itajaí foi delegada ao município de Itajaí pelo período de 25 anos, renovável por mais 25. A partir daí o município assumiu a gestão do porto e, segundo a legislação em vigência, foi promovido o arrendamento do Terminal de Contêineres (Tecon) a empresa de Terminal de Contêineres do Vale do Itajaí (Teconvi), hoje APM Terminals Itajaí. Com isso, o porto pode receber, enfim, investimentos em equipamentos e modernização (PORTOITAJAÍ, 2014).

A partir de então, o porto só cresceu: dados (Figura 5.4) mostram que a movimentação portuária aumentou expressivamente a partir do ano de 2004, até atingir o pico de cerca de 6,5 milhões de toneladas em 2007. Em 2008, a movimentação no porto de Itajaí deu uma reduzida principalmente pelo início de operação de sua concorrente, o TUP Portonave (LABTRANS, 2015).

Figura 5.4 – Evolução da movimentação (em toneladas) no porto de Itajaí (SC)



Fonte: LabTrans, 2015.

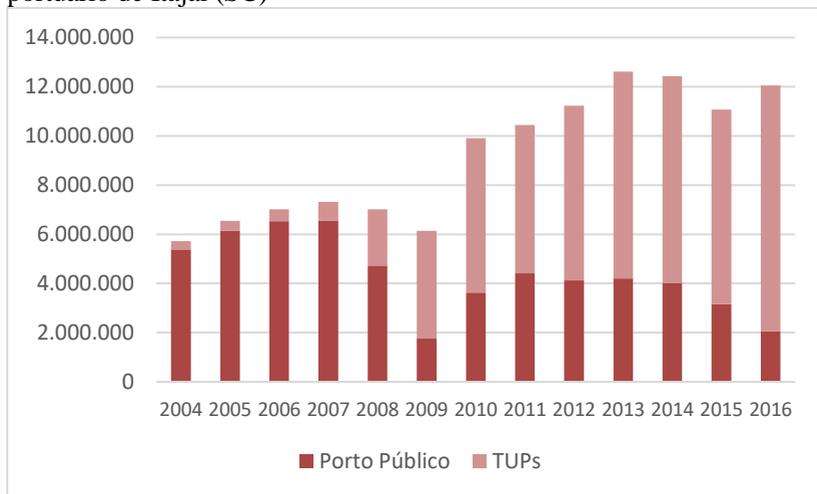
Cabe ressaltar que, além dos dilemas administrativos, o porto enfrentou duas grandes crises devido às danificações em sua estrutura causadas por duas grandes enchentes, uma em julho de 1983 e outra em novembro de 2008. Na primeira enchente, o porto foi reconstruído no mesmo ano, mas, na segunda, as obras só foram concluídas em novembro de 2010, resultando em uma redução significativa na movimentação portuária em 2009 – cerca de 37% daquela do ano anterior e a 27% do pico verificado em 2007 (LABTRANS, 2015).

Em 2010, com as obras finalizadas, a movimentação portuária alcançou um patamar superior a 4 milhões de toneladas ao ano, e desde então vem se mantendo aproximadamente no mesmo nível (apesar da paralização das operações no berço 1 de setembro de 2011 a junho de

2014, devido às implicações de uma enchente) – por exemplo, no ano de 2013 foram movimentadas 4.196.105 toneladas de carga nas instalações de uso público, quase exclusivamente de contêineres, distribuídas por 244,644 mil unidades ou 411,504 mil TEUs. Nesse mesmo ano, o TUP Portonave movimentou 426,193 mil unidades (693,236 mil TEUs).

Já no complexo portuário de Itajaí, em 2010, foi movimentado um total de 954,38 mil TEUs, superando as expectativas em 154,38 mil TEUs para esse mesmo ano e avançando 61% em relação a 2009, conforme mostra a Figura 5.5.

Figura 5.5 – Evolução da movimentação (em toneladas) no complexo portuário de Itajaí (SC)



Fonte: LabTrans, 2015.

Por sua vez, em 2016, o complexo foi responsável pela movimentação de 12.049.551 toneladas de mercadorias (LABTRANS, 2015).

Ainda considerando todo o complexo mostrado na Figura 5.5, observa-se que, com exceção do ano de 2009, no qual ocorreu a enchente e a crise mundial, a evolução da movimentação foi quase sempre positiva. Concluindo-se, então, que a redução de movimentação no porto público se deveu principalmente à transferência de parte da mesma para o TUP Portonave (LABTRANS, 2015).

Atualmente o porto público possui um número maior de contêineres importados que exportados: os desembarques de contêineres cheios, em 2013, totalizaram 151,577 mil TEU, enquanto os embarques atingiram 130,113 mil TEU. No TUP Portonave, ocorreu o contrário: 275,308 mil TEU cheios foram embarcadas e 221,412 mil, desembarcadas (LABTRANS, 2015).

Em 2014, os principais projetos de expansão e desenvolvimento diziam respeito à: reconstrução dos berços 2 e 3; implantação de linha de transmissão de 138 KVA para suprir a demanda de contêineres *reefer*; dragagem de aprofundamento dos canais interno e externo e bacia de evolução para 14 m; ampliação do píer turístico em 1,6 mil m²; construção da Marina do Saco da Fazenda, integrando o porto de Lazer, por meio de Parceria Público Privada (PPP); elaboração e lançamento de edital de arrendamento dos dois berços do porto Público; e, implantação do Sistema de Acompanhamento e Monitoramento de Embarcações (*Vessel Tracking and Monitoring System – VTMS*) para o gerenciamento e controle de navios (PORTOITAJAÍ, 2014).

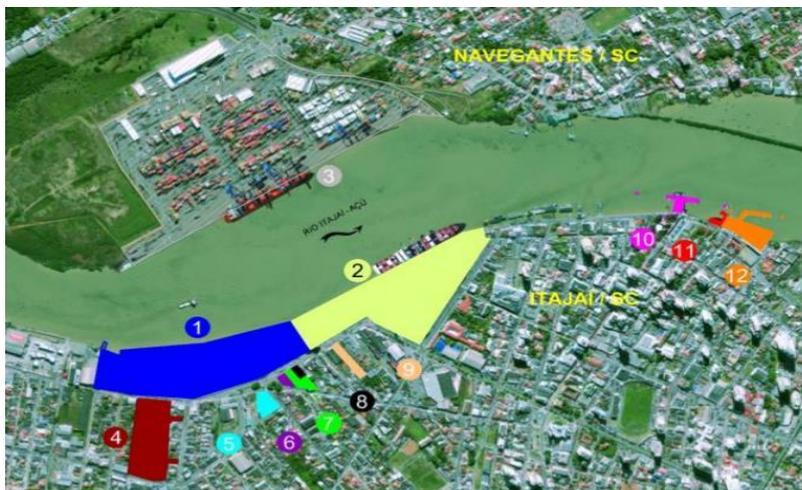
5.3.1.2 *Dados gerais*

O zoneamento geral do porto de Itajaí é mostrado na Figura 5.6 e detalhado a seguir.

O porto de Itajaí conta com obras de abrigo para proporcionar a tranquilidade das águas na bacia de evolução e junto ao berço de atracação. Existem sete espigões de larguras variando entre 120 m e 140 m situados na margem norte do rio; e dois molhes – molhe norte e sul, com respectivamente, 1070 m de comprimento e 800 m de comprimento. Destaca-se que as obras da nova bacia de evolução no Saco da Fazenda promoverão alterações nas obras de abrigo atuais (LABTRANS, 2015).

Quanto à infraestrutura de acostagem, o porto de Itajaí possui quatro berços, sendo dois arrendados e dois públicos, e um Píer Turístico. A Figura 5.7 identifica a localização dos berços, assim como o Recinto Alfandegado Contíguo (RAC) (LABTRANS, 2015).

Figura 5.6 – Zoneamento do porto de Itajaí (SC)

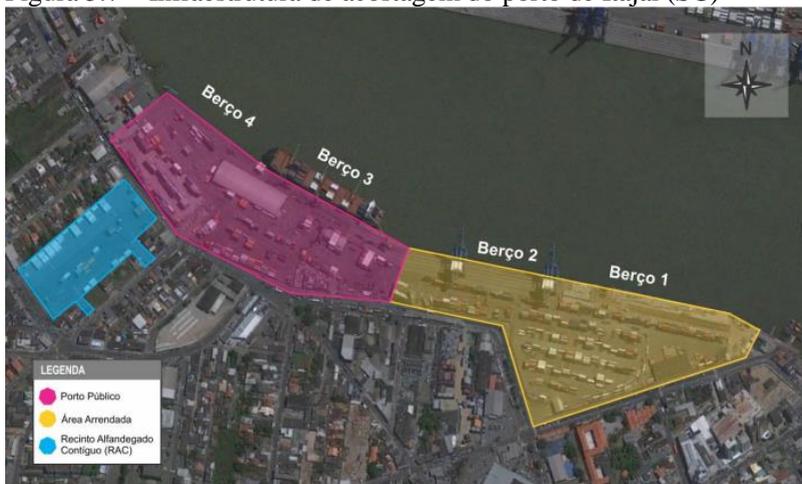


LEGENDA

- | | | |
|------------------------------------|-----------------------------------|------------------------------------|
| 1 Porto de Itajaí - Área Primária | 5 Sede Administrativa | 10 Centro Integrado de Atendimento |
| 2 Porto de Itajaí - Área Arrendada | 6 Centro de Treinamento Portuário | 11 Terminal de Passageiros |
| 3 TUP Portonave | 8 Oficina Arrendada | 12 Serviços de Praticagem |
| 4 Recinto Alfandegado Contíguo | 9 Base de Emergência | 13 Autoridade Marítima |

Fonte: LabTrans, 2015.

Figura 5.7 – Infraestrutura de acostagem do porto de Itajaí (SC)

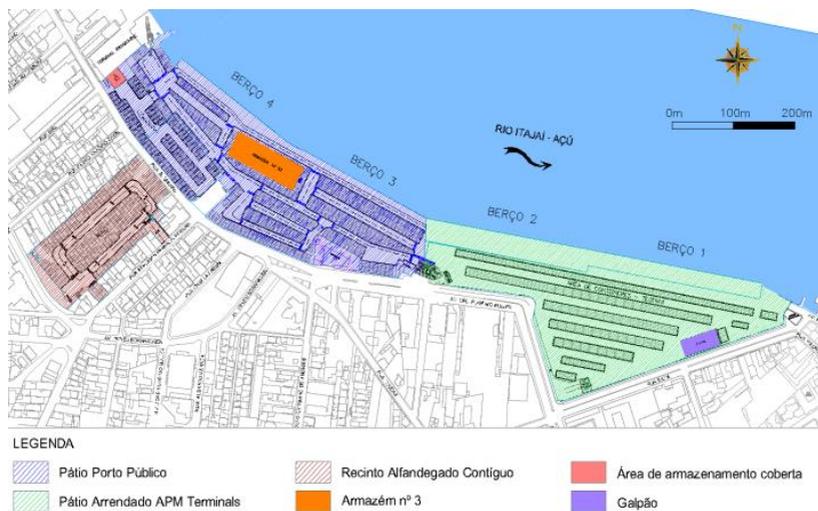


Fonte: LabTrans, 2015.

Os berços 1 e 2, arrendados à empresa APM Terminals, são especializados em contêineres e possuem, respectivamente 285 metros e 272,3 metros de comprimento, e 14,5 metros e 15 metros de profundidade. Os berços 3 e 4, que compõem o cais público, onde podem operar todos os operadores portuários credenciados junto ao porto de Itajaí, possuem, respectivamente, 248 metros e 215 metros de comprimento, e 11,6 metros e 8,8 metros de profundidade. Ambos os berços estão sendo reforçados e realinhados (LABTRANS, 2015).

Em relação às estruturas de armazenagem, não se pode deixar de citar as principais, mostradas na Figura 5.8 e que são as seguintes: pátio da área arrendada, de 60 mil m² em pavimento intertravado; pátio do porto público, de aproximadamente 73 mil m² em pavimentação asfáltica; recinto alfandegado contíguo, em área de aproximadamente 27,1 mil m²; armazém 03, de 4,75 mil m², localizado na retroárea do porto público; área de armazenamento coberta de 566 m² no porto público; e galpão de 1,5 mil m², localizado na área arrendada do terminal. Destaca-se que há projetos para aumentar essa área de armazenagem (LABTRANS, 2015).

Figura 5.8 – Estruturas de armazenagem do porto de Itajaí (SC)

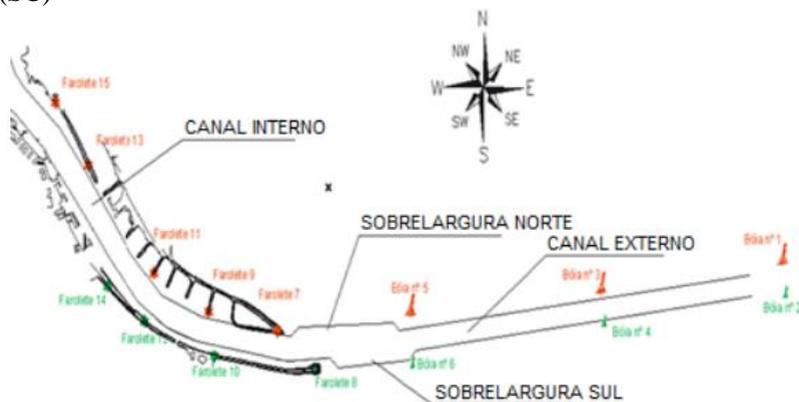


Fonte: LabTrans, 2015.

No que diz respeito aos equipamentos portuários, o porto de Itajaí possui três guindastes sobre rodas do tipo Guindaste Móvel Portuário (*Mobile Harbor Crane – MHC*), com capacidade para 100 toneladas brutas, e dois portêineres pós-Panamax do tipo *twin-lift* para movimentação de cargas no terminal portuário, e empilhadeiras do tipo *reach stacker* e do tipo *empty container handlers*, além de caminhões *terminal tractor* para movimentação de cargas no terminal retroportuário. O porto ainda dispõe de cinco balanças para caminhões (LABTRANS, 2015).

O acesso aquaviário do porto de Itajaí (Figura 5.9) permite que as embarcações em operações possuam, no máximo, as seguintes medidas: 306 metros de comprimento; 48,2 metros de boca; e 95 mil toneladas de porte bruto. Sua bacia de evolução é localizada em frente ao TUP da Portonave e possui 400 metros de diâmetro. Seu canal está dividido em dois trechos, um externo e um interno com, respectivamente, 120 metros de largura, 3,3 mil metros de extensão e 13,2 metros de profundidade, sem curvas, e com 120 e 150 metros (devido a variação de largura entre nas curvas) de largura, 3,2 mil metros de extensão e 12,7 metros de profundidade. Em ambos o trânsito autorizado é de mão única sem ultrapassagens e a velocidade máxima permitida é de sete nós (LABTRANS, 2015).

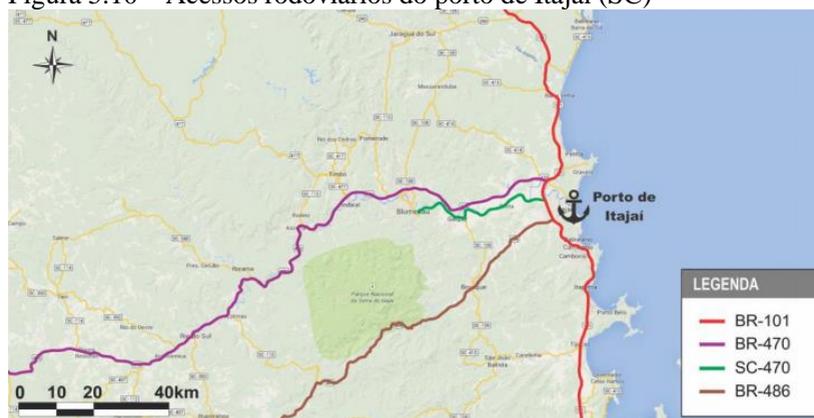
Figura 5.9 – Canais de acesso e bacia de evolução do porto de Itajaí (SC)



Fonte: LabTrans, 2015.

Não há acesso ferroviário ao porto de Itajaí, há apenas acesso rodoviário. Esse acesso rodoviário se encontra a poucos quilômetros das rodovias BR-101 e BR-470, que compõem um dos principais entroncamentos rodoviários do Sul do Brasil. Além dessas rodovias, conectam também o porto com sua hiterlândia, embora com menor importância, a BR-486 e a SC-470, conforme mostra a Figura 5.10 (LABTRANS, 2015).

Figura 5.10 – Acessos rodoviários do porto de Itajaí (SC)



Fonte: LabTrans, 2015.

Por ser a principal ligação entre Santa Catarina e os demais estados do Sul do Brasil, e entre todas as cidades do litoral catarinense, a BR-101 é utilizada por um grande fluxo de veículos, que aumenta, devido ao turismo, no período de verão. O trecho entre a divisa com o Paraná até a cidade de Palhoça é totalmente duplicado (velocidade máxima permitida é de 100 km/h para veículos leves) e é concessionado à Auto Pista Litoral Sul, empresa pertencente ao grupo Arteris, desde 2008 por um período de 25 anos. Tanto as sinalizações horizontal e vertical quanto o pavimento encontram-se em bom estado de conservação, de acordo com o Relatório da Pesquisa da Confederação Nacional de Transporte (CNT) de Rodovias 2013 (LABTRANS, 2015).

O trecho mais crítico da BR-101 é o trecho entre a interseção da BR-101 com a Avenida Governador Adolfo Konder e a interseção da BR-101 com a BR-486, que atinge nível de serviço F nos horários de pico, seguido pelo trecho entre a entrada para SC-412 e a SC-414 que apresenta

nível de serviço E nos horários de pico. Tendo isso em vista, novas obras de ampliação de capacidade podem vir a ser necessárias em breve (LABTRANS, 2015).

Já a BR-470 é a principal ligação entre as cidades do trecho Navegantes (SC) a Camaquã (Rio Grande do Sul – RS). Em Santa Catarina, onde se encontram 359,1 dos 472 quilômetros dessa rodovia, a mesma se destaca por fazer ligação entre o planalto e o oeste ao litoral. O trecho catarinense possui pista simples pavimentada em bom estado de conservação, acostamento na maioria do trajeto, e terceira pista ao longo da rodovia, notadamente em trechos de aclive. É detectada a presença de sinalizações verticais e horizontais na via, esta última bastante desgastada pelo tráfego de veículos em alguns trechos, de acordo com o Relatório da Pesquisa CNT de Rodovias 2013 (LABTRANS, 2015).

Na BR-470, a saturação é natural, visto que a rodovia apresenta pista simples e detém grande volume de tráfego (o volume horário de pico em 2013 era de 2.037 veículos por hora no trecho de Blumenau a Indaial, segundo o Departamento Nacional de Infraestrutura de Transportes). O trecho da BR-470 de 73 quilômetros entre a cidade de Gaspar (SC) e Navegantes (SC) é o mais movimentado, sendo responsável por um alto índice de acidentes. Visando atenuar isso, foi projetada e está sendo realizada a obra de duplicação do trecho Navegantes a Blumenau com verbas do Programa de Aceleração do Crescimento (PAC) do Governo Federal. Além da duplicação da pista, o projeto inclui a recuperação e reabilitação do pavimento, viadutos, pontes, passarelas, ciclovias e calçadas. No entanto, a obra se encontra atrasada (LABTRANS, 2015).

O trajeto para acessar os portões de entrada do porto a partir dessas duas rodovias – BR-101 ou BR 470 – é mostrado na Figura 5.11 (linha em laranja). Ao todo, são dois portões de entrada/saída do porto – um portão principal compartilhado entre área pública e arrendada e um portão secundário que dá acesso à área pública –, um portão apenas de saída do porto, e um portão de acesso e outro de saída ao RAC (LABTRANS, 2015).

Esse trajeto considerando à chegada no porto pelo portão principal se dá passando pelas ruas apontadas na Figura 5.12. O trajeto considerando à chegada pelo portão secundário é igual ao trajeto supracitado até a Avenida Irineu Bornhausen, se diferenciando a partir daí de acordo com a Figura 5.13.

Destaca-se que esse trajeto atravessa a cidade de Itajaí, sendo que o grande número de intersecções somada à forte influência do tráfego

urbano torna a viagem dos caminhões ao porto mais lenta (LABTRANS, 2015).

Figura 5.11 – Acessos rodoviários do porto de Itajaí (SC)



Fonte: LabTrans, 2015.

Figura 5.12 – Trajeto de acesso ao portão principal do porto de Itajaí (SC)



Fonte: LabTrans, 2015.

Figura 5.13 – Acessos ao portão secundário do porto de Itajaí (SC)



Fonte: LabTrans, 2015.

No Plano Mestre desenvolvido para o porto pelo LabTrans (2015) foram realizadas a projeção de tráfego para os trechos das rodovias BR-101 e BR-470 utilizados para acessar o empreendimento, seguindo os seguintes passos:

- 1) presumiu-se que o número de veículos nas rodovias que não têm relação direta com o porto continuará crescendo segundo a média histórica de crescimento do PIB de Santa Catarina nos últimos doze anos, que é de 3,3% ao ano e, com isso, calculou-se a projeção dessa parcela do tráfego rodoviário em Volume Médio Diário horários (VMDH) total e Volume Hora de Pico (VHP) total para cada ano;
- 2) já a parcela do tráfego rodoviário que tem relação direta com o porto foi calculada, então, a partir do histórico de movimentação de cargas do porto e de sua projeção por meio de modelos painéis de dados, no qual realizou-se a alocação das cargas nas rodovias, calculando as quantidades de caminhões necessários para transportar aquelas cargas (também em VMDH total e VHP total para cada ano), e levando em conta a origem e o destino das mesmas;
- 3) por fim, a soma dessas duas parcelas de tráfego (sem e com relação direta com o porto) resultou no volume de tráfego total (em VMDH total e VHP total) estimado para essas rodovias para os próximos anos.

Por meio da metodologia HCM foram calculadas as capacidades dessas rodovias e estas foram comparadas à demanda rodoviária futura

obtida. Essa comparação demonstrou que é nítida a saturação dos trechos rodoviários analisados, refletindo no porto de Itajaí em uma perda de competitividade do mesmo.

Nesse sentido, Labtrans (2015) ainda destaca: que as obras de duplicação da BR-470 são fundamentais a fim de que o porto continue sendo uma opção viável para o escoamento da produção da região oeste; e que a BR-101 necessitará de ampliação nos trechos analisados, visto que a demanda urbana em horários de pico esgota a capacidade da via e que nos anos futuros essa situação tende apenas a piorar.

Por fim, cabe ressaltar ainda que os Planos Mestres trazem em seu portfólio de investimentos e expansões necessárias visando o crescimento e eficiência do porto não só orientações quanto aos acessos terrestres, mas também uma série de medidas considerando a infraestrutura, operações e gestão portuária.

5.3.2 Caracterização do PGV e sua operação por meio de entrevista com funcionários do porto

Inicialmente, entrou-se em contato com o diretor do porto de Itajaí e nesse primeiro contato foi repassado todo o histórico sobre o porto. Em seguida, foram recebidos materiais, abrangendo informações gerais e mapas de cada elemento do empreendimento.

Quanto ao funcionamento, foi repassado que o porto possui um cronograma de chegada e saída de navios que se repete semanalmente ao longo do ano. Assim que as mercadorias chegam ao porto para exportação ou importação, são transferidas para o navio (no caso de exportação) ou caminhão (no caso de importação). No caso da importação, essas mercadorias são transportadas para os terminais portuários e portos secos e, nesses locais, esse material fica armazenado até que as transportadoras de cada terminal portuário ou porto seco levem o mesmo para seu destino de importação.

Em relação à sazonalidade mensal, segundo o diretor do porto, pode-se afirmar que a movimentação do segundo semestre geralmente é maior em função de duas condições específicas: os contratos comerciais devem ser cumpridos até o final do exercício (mês de dezembro); e tem-se um aumento das importações por conta das festas de final de ano. Já nos primeiros meses geralmente ocorre uma redução histórica relacionada com o início dos contratos comerciais

e férias das empresas. No entanto, segundo o diretor do porto, essa diferença não é significativa e não causa mudanças na escala dos navios que se mantêm praticamente constantes semanalmente ao longo do ano.

Sobre a sazonalidade semanal, a chegada de boa parte dos navios nos sábados e domingos reflete em um fluxo maior de caminhões chegando e saindo do entorno desse PGV durante o início/ meio da semana.

Na visão do entrevistado, as variáveis que deveriam ser consideradas no estudo, pois influenciam o movimento de caminhões, são: o volume de contêineres transportados; a economia do País, que impacta diretamente no volume de mercadorias transportadas; e as chuvas acima de 40 milímetros, que atrasam o cronograma dos navios devido à dificuldade de operação ou até assoreamento do canal de acesso. Também foi detectado pela entrevista que o dia da semana influencia na demanda de caminhões.

Quanto aos dados, ele disse haver disponibilidade para entrega de registro de movimentação de contêineres, segundo:

- 1) código do contêiner; tamanho do contêiner (20 ou 40 pés);
- 2) categoria (importação ou exportação);
- 3) indicação da condição física do contêiner, ou seja, se o mesmo está Totalmente Carregado (Full Container Load - FCL), Menos Carregado (Less Container Load - LCL) ou sem observação (célula vazia);
- 4) se a carga é classificada como perigosa ou não;
- 5) se o contêiner é refrigerado ou não;
- 6) indicação de uma condição/característica especial de retirada ou entrega do contêiner, conforme explicado a seguir;
- 7) tipo do modo no qual o contêiner chegou ao terminal, ou seja, navio para importação ou caminhão para exportação;
- 8) placa do caminhão, no caso da importação, ou número do navio, no caso da exportação, que descarregou o contêiner;
- 9) data em que o contêiner foi descarregado;
- 10) tipo do modo no qual o contêiner saiu do terminal, ou seja, caminhão para importação ou navio para exportação, exceto no caso de dray-off; e
- 11) data em que o contêiner foi carregado.

Quanto à indicação de uma condição/característica especial de retirada ou entrega do contêiner, a mesmo representa:

- 1) *dray-off* = condição de contêiner de exportação, que por algum motivo, o exportador decide por retirar o contêiner do porto antes do seu embarque no navio. Em suma, é um cancelamento de embarque no navio e retirada via caminhão;
- 2) retirado via DTC = condição do contêiner de importação, no qual a entrega de contêineres descarregados dos navios ocorre via Declaração de Trânsito de Contêiner (DTC), ou seja, quando o cliente importador, opta por transferir seu contêiner logo após a descarga do navio para outro recinto alfandegado de sua escolha, devidamente autorizado pela Receita Federal do Brasil (RFB) a receber cargas oriundas do porto de Itajaí, e dar continuidade ao processo de nacionalização destes;
- 3) liberado em área primária = condição do contêiner de exportação e importação, no qual a entrega de contêineres descarregados dos navios ocorre via Despacho de Importação (DI). Nesse caso, o cliente não solicita transferência para nenhum outro terminal e, desta maneira, o contêiner é presenciado junto a RFB dentro da área primária, e passa pelo processo de liberação dentro do porto, podendo ser retirado diretamente pelo cliente importador, após os trâmites junto à alfandega; e
- 4) célula vazia = nulo, ou seja, não existe nenhuma característica especial.

5.3.3 Definição das variáveis a serem testadas

Em relação às variáveis utilizadas nos modelos existentes, segundo a revisão bibliográfica sistemática sobre o assunto (item 3.5, na página 125), a caracterização da área de estudo via fontes secundárias (subitem 5.3.1, na página 159) e via entrevistas com prepostos do porto familiarizados com a operação (subitem 5.3.2, na página 173), podem ser considerados como dados de entrada, inicialmente, nos modelos de geração de viagem para o porto:

- 1) o volume de mercadorias (em TEUs) que chegam/ saem do porto durante certo período de tempo, podendo considerar o fluxo de mercadorias ocorrido em um dia

- ou até o número máximo de dias que as cargas geralmente permanecem armazenadas, conforme a operação do PGV estudado;
- 2) condição física do contêiner (vazio ou cheio);
 - 3) se o contêiner é de carga perigosa;
 - 4) se o contêiner é refrigerado;
 - 5) indicação de uma condição/característica especial de retirada ou entrega do contêiner, conforme explicado na página 175;
 - 6) o dia da semana e o mês do ano analisado nos modelos;
 - 7) PIB da região ou, caso não exista essa informação, do País; e
 - 8) se houve precipitações pluviométricas acima de 40 milímetros no dia analisado.

5.4 ETAPA 3

Seguindo os devidos objetivos, o estudo prosseguiu com o levantamento das variáveis definidas para o porto de Itajaí (SC).

O PIB do município pode ser levantado no site do Instituto Brasileiro de Geografia e Estatística e o volume de chuvas pode ser levantado junto à Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina; este foi considerado como variável *dummy* identificando os dias em que o volume de chuvas foi igual ou superior a 40 milímetros, o que poderia ter causado atrasos na operação portuária.

No que diz respeito ao restante das informações, foi encaminhado um pedido de solicitação das mesmas à administração do porto de Itajaí (SC). A administração do terminal de contêineres forneceu os dados sobre o volume de contêineres movimentados e o volume de caminhões que entraram e saíram do porto no período de 20 meses entre 01/01/2013 e 31/08/2015. Esses dados foram encaminhados em planilhas de Excel, e possuíam as informações citadas no subitem 5.3.3, na página 175. Ressalta-se que apenas o registro dos dados no período de 01/01/2015 a 31/08/2015 foram realizados eletronicamente, visto que anterior a esse período o porto não possuía recursos técnicos para realizar o registro de modo eletrônico, apenas de modo manual.

Além desses dados, foram adicionadas variáveis qualitativas para representar o dia da semana ao qual se estava fazendo referência. Para

isso, foram utilizadas variáveis do tipo *dummy* (0 = não; 1 = sim). Foi criada uma variável *dummy* para a ocorrência de feriado e seis variáveis *dummy* para o dia da semana: segunda, terça, quarta, quinta, sexta-feira e sábado. O domingo foi escolhido como grupo de referência e, por isso, não foi incluído como variáveis *dummy*. Isto porque é recomendada a utilização de $N - 1$ variáveis *dummy* com a finalidade de se evitar o caso de perfeita colinearidade, problema esse que é chamado de armadilha das variáveis *dummy*.

Dessa forma, para o dia 01/01/2015, por exemplo, a variável *dummy* feriado assumiu o valor 1, por se tratar do Ano Novo; as variáveis *dummy* segunda-feira, terça-feira, quarta-feira, sexta-feira e sábado assumiram o valor 0 enquanto que a variável *dummy* quinta-feira assumiu o valor 1. Seguindo o mesmo raciocínio, para o dia 04/04/2015, por exemplo, a variável *dummy* feriado assumiu o valor 0; e as variáveis *dummy* segunda-feira, terça-feira, quarta-feira, quinta-feira, sexta-feira e sábado assumiram o valor 0, caracterizando, portanto, e por exclusão, o dia em questão como sendo um domingo.

Além disso, foram consideradas, ainda, variáveis autorregressivas para os contêineres, ou seja, foram considerados os valores defasados delas próprias. Isto porque tanto os contêineres de importação quanto os de exportação permanecem um tempo armazenado na área portuária antes de seguir seu destino.

Dessa forma, foi adicionada à lista de variáveis para teste as variáveis autorregressivas do número de contêineres considerando os 11 dias anteriores ao dia avaliado para contêineres de exportação e 13 dias anteriores para contêineres de importação. Esses números foram escolhidos considerando que os contêineres para exportação não permanecem mais do que 11 dias armazenados antes de seguir para seu destino e os contêineres para importação não permanecem mais que 13 dias armazenados antes de seguir para seu destino.

Com isso, iniciou-se a preparação dos dados para que os modelos pudessem ser elaborados. A primeira etapa consistiu em fazer um somatório dos dados quantitativos – como, por exemplo, número de caminhões, número de contêineres, etc. – por dia, durante o período analisado.

5.5 ETAPA 4

Com esse conjunto de dados em mãos, iniciaram-se os testes de modelos de regressão linear múltipla pelo software Gretl. Ressalta-se que se optou por elaborar modelos em separado para a exportação de TEUs (caminhões entrando) e importação de TEUs (caminhão saindo), visto que essas duas operações possuíam características muito diferentes no porto de Itajaí (SC).

No primeiro caso, foram verificadas equações de análise de regressão múltipla passo a passo para identificar as variáveis independentes com melhor correlação e numa segunda etapa, foram estimadas regressões lineares múltiplas somente com as variáveis estatisticamente significativas. Informações como PIB do município e volume de chuvas não foram significativas, além de a inserção das mesmas resultava em problemas de colinearidade nos modelos, visto que estas refletem diretamente em um aumento ou diminuição do número de contêineres. Assim, essas variáveis foram desconsideradas.

Além disso, cabe destacar que após realizar os primeiros testes estatísticos, considerando o modelo com todas as variáveis que eram estatisticamente significativas, foi necessário a inserção de defasagens da variável y (volume de caminhões) do lado direito da equação (como variável explicativa) para sanar problemas de autocorrelação.

Ressalta-se que a inclusão da variável defasada do volume de caminhões como dado de entrada não impossibilita porventura o uso do modelo desenvolvido para um porto novo. Como nos estudos que antecedem a implementação de novos portos geralmente há a necessidade de uma previsão de movimentação de caminhões. A partir destas previsões se poderia inserir os impactos das demais variáveis do modelo.

Apesar de ter incluído a defasagens da variável y do lado direito da equação, o modelo ainda não se encontrava válido quanto aos testes estatísticos e, ao analisar o registro de dados, perceberam algumas inconsistências nos dados em que havia sido feito o registro manual. Dessa forma, optou-se por utilizar apenas os dados em que houve registro eletrônico dos mesmos, ou seja, os dados do período de 01/01/2015 a 31/08/2015.

Com isso, o modelo de exportação (caminhões entrando) resultou no modelo final apresentado na Equação 27.

$$CAMENT = -37,500 + 194,364SEG + 246,217TER + 229,017QUA + 198,045QUI + 141,133SEX + 87,730SÁB + 0,032TEUEXP_{-3} + 0,496CAMENT_{-1} \quad (27)$$

Onde:

CAMENT = volume de caminhões entrando no porto no dia analisado;

SEG = variável *dummy* para segunda-feira;

TER = variável *dummy* para terça-feira;

QUA = variável *dummy* para quarta-feira;

QUI = variável *dummy* para quinta-feira;

SEX = variável *dummy* para sexta-feira;

SAB = variável *dummy* para sábado;

TEUEXP₋₃ = volume de contêineres (TEU) exportados há três dias;

e

CAMENT₋₁ = volume de caminhões entrando no porto há um dia.

Na Tabela 5.1 são apresentados os valores do teste-t, assim como do R² ajustado, que resultou em 52,68%.

Os sinais positivos das variáveis indicam que as mesmas tendem a crescer na mesma direção que a variável dependente, sendo que variações de 1 ponto percentual (p.p.) no volume de caminhões entrando no dia anterior provocaram variação de 0,50 p.p. no volume de caminhões entrando no dia analisado; e que variações de 1 p.p. no volume de TEUs exportados há três dias provocaram variação de 0,03 p.p. sobre o volume de caminhões entrando no dia analisado. A terça-feira por si só afetou 246,22 p.p. o volume de caminhões entrando no dia analisado; a quarta-feira, 229,02 p.p.; a quinta-feira, 198,04 p.p.; a segunda-feira, 194,36 p.p.; a sexta-feira, 141,13 p.p.; o sábado, 87,73 p.p.; e o domingo, 0,00 p.p.

As variáveis *dummy* para dia da semana mostraram que o volume de caminhões entrando no porto de Itajaí no início/meio da semana – terça, quarta, quinta e segunda-feira – é maior que o volume de caminhões entrando no porto no final da semana – sexta-feira, sábado e domingo. Isso pode ser explicado pelo fato de muitos navios vazios atracam no porto no final de semana para serem descarregados e carregados com TEUs e como estes permanecem alguns poucos dias armazenados, eles geram maior movimentação de caminhões ao longo do início/meio da semana.

Tabela 5.1 – Resultado do modelo de exportação para o porto de Itajaí utilizando regressão linear múltipla

Variável Independente	Coefficiente	Erro Padrão	teste-t	p-valor	
const	-37,5004	22,6026	-1,6591	0,0985	*
SEG	194,364	28,3269	6,8615	<0,0001	***
TER	246,217	27,3412	9,0053	<0,0001	***
QUA	229,017	27,8644	8,2190	<0,0001	***
QUI	198,045	28,4765	6,9547	<0,0001	***
SEX	141,133	28,6847	4,9201	<0,0001	***
SAB	87,7298	27,5265	3,1871	0,0016	***
TEUEXP_3	0,0323793	0,0167307	1,9353	0,0542	*
CAMENT_1	0,496091	0,0566521	8,7568	<0,0001	***
Média var. dependente	264,5774		D.P. var. dependente	161,8628	
Soma resíd. quadrados	2851593		E.P. da regressão	111,3473	
R ²	0,542684		R ² ajustado	0,526778	
F(8, 230)	34,11685		P-valor(F)	3,67e-35	
Log da verossimilhança	-1460,864		Critério de Akaike	2939,728	
Critério de Schwarz	2971,016		Critério Hannan-Quinn	2952,336	
rô	-0,006207		h de Durbin	-	
				0,198804	

Fonte: Autoria Própria.

Para mais detalhes acerca da significância estatística do modelo ver Tabela 5.2.

Tabela 5.2 – Resultados dos testes estatísticos do modelo de regressão linear múltipla de exportação para o porto de Itajaí

(continua)

Teste de não-linearidade (quadrados) -

Hipótese nula: a relação é linear

Estatística de teste: LM = 0,140028

com p-valor = $P(\text{Qui-quadrado}(2) > 0,140028) = 0,932381$

Teste Reset para especificação (apenas quadrados) -

Hipótese nula: a especificação é adequada

Estatística de teste: $F(1, 229) = 3,3034$

com p-valor = $P(F(1, 229) > 3,3034) = 0,0704434$

Teste de White para a heteroscedasticidade -

Hipótese nula: sem heteroscedasticidade

Estatística de teste: LM = 28,4801

com p-valor = $P(\text{Qui-quadrado}(23) > 28,4801) = 0,198169$

Teste da normalidade dos resíduos -

Hipótese nula: o erro tem distribuição Normal

Estatística de teste: $\text{Qui-quadrado}(2) = 3,99024$

com p-valor = 0,135997

Teste de Chow para a falha estrutural na observação 2015-04-30

-

Hipótese nula: sem falha estrutural

Estatística de teste: $F(9, 221) = 1,79756$

com p-valor = $P(F(9, 221) > 1,79756) = 0,0699043$

Teste LM para autocorrelação até a ordem 7 -

Hipótese nula: sem autocorrelação

Estatística de teste: LMF = 0,528372

com p-valor = $P(F(7, 223) > 0,528372) = 0,812597$

Tabela 5.2 – Resultados dos testes estatísticos do modelo de regressão linear múltipla de exportação para o porto de Itajaí

(conclusão)

Fatores de Inflacionamento da Variância (VIF)

Valor mínimo possível = 1,0

Valores > 10,0 podem indicar um problema de colinearidade

SEG	1,887
TER	1,758
QUA	1,826
QUI	1,907
SEX	1,935
SAB	1,782
TEUEXP_3	1,166
CAMENT_1	1,605

$VIF(j) = 1/(1 - R(j)^2)$, onde $R(j)$ é o coeficiente de correlação múltipla entre a variável j e a outra variável independente

Fonte: Autoria Própria.

Uma regressão linear múltipla também foi elaborada para o modelo de importação (caminhões saindo) e, obteve-se, o modelo final apresentado na Equação 28.

$$CAMSAI = -50,236 + 207,674SEG + 168,3857TER + 260,382QUA + 210,758QUI + 143,884SEX + 150,086SÁB + 0,161TEUIMP_{-1} + 0,076TEUIMP_{-4} + 0,233CAMSAI_{-1} \quad (28)$$

Onde:

CAMSAI = volume de caminhões saindo do porto no dia analisado;

SEG = variável *dummy* para segunda-feira;

TER = variável *dummy* para terça-feira;

QUA = variável *dummy* para quarta-feira;

QUI = variável *dummy* para quinta-feira;

SEX = variável *dummy* para sexta-feira;

SAB = variável *dummy* para sábado;

- e TEUEXP₋₁ = volume de contêineres (TEU) exportados há um dia;
- e TEUEXP₋₄ = volume de contêineres (TEU) exportados há quatro dias; e
- CAMSAI₋₁ = volume de caminhões saindo do porto há um dia.

Na Tabela 5.3 são apresentados os valores do teste-t, assim como do R^2 ajustado, que resultou em 51,14%.

Tabela 5.3 – Resultado do modelo de importação para o porto de Itajaí utilizando regressão linear múltipla

Variável Independente	Coefficiente	Erro Padrão	teste-t	p-valor	
const	-50,2359	26,6081	-1,8880	0,0603	*
SEG	207,674	31,922	6,5057	<0,0001	***
TER	168,385	29,3778	5,7317	<0,0001	***
QUA	260,382	28,7086	9,0698	<0,0001	***
QUI	210,758	29,1644	7,2265	<0,0001	***
SEX	143,884	30,7416	4,6804	<0,0001	***
SAB	150,086	27,7731	5,4040	<0,0001	***
TEUIMP_1	0,160852	0,0174986	9,1923	<0,0001	***
TEUIMP_4	0,0758047	0,0184833	4,1013	<0,0001	***
CAMSAI_1	0,23278	0,0561566	4,1452	<0,0001	***
Média var. dependente	307,9289		D.P. var. dependente	162,5641	
Soma resíd. quadrados	2956926		E.P. da regressão	113,6325	
R^2	0,529874		R^2 ajustado	0,511398	
F(8, 230)	28,67817		P-valor(F)	4,84e-33	
Log da verossimilhança	-1465,198		Critério de Akaike	2950,397	
Critério de Schwarz	2985,161		Critério Hannan-Quinn	2964,406	
rô	-0,066913		h de Durbin	-	
				2,084380	

Fonte: Autoria Própria.

Para mais detalhes acerca da significância estatística do modelo ver Tabela 5.4.

Tabela 5.4 – Resultados dos testes estatísticos do modelo de regressão linear múltipla de importação para o porto de Itajaí

(continuação)

Teste de não-linearidade (quadrados) -

Hipótese nula: a relação é linear

Estatística de teste: LM = 6,48913

com p-valor = $P(\text{Qui-quadrado}(3) > 6,48913) = 0,0900922$

Teste Reset para especificação (apenas quadrados) -

Hipótese nula: a especificação é adequada

Estatística de teste: $F(1, 228) = 1,00354$

com p-valor = $P(F(1, 228) > 1,00354) = 0,317516$

Teste de White para a heteroscedasticidade -

Hipótese nula: sem heteroscedasticidade

Estatística de teste: LM = 34,2019

com p-valor = $P(\text{Qui-quadrado}(33) > 34,2019) = 0,409775$

Teste da normalidade dos resíduos -

Hipótese nula: o erro tem distribuição Normal

Estatística de teste: Qui-quadrado(2) = 4,95751

com p-valor = 0,0838474

Teste de Chow para a falha estrutural na observação 2015-04-30

-
Hipótese nula: sem falha estrutural

Estatística de teste: $F(10, 219) = 1,64133$

com p-valor = $P(F(10, 219) > 1,64133) = 0,0964335$

Teste LM para autocorrelação até a ordem 7 -

Hipótese nula: sem autocorrelação

Estatística de teste: LMF = 1,11906

com p-valor = $P(F(7, 222) > 1,11906) = 0,351978$

Tabela 5.4 – Resultados dos testes estatísticos do modelo de regressão linear múltipla de importação para o porto de Itajaí

(conclusão)

Fatores de Inflacionamento da Variância (VIF)

Valor mínimo possível = 1,0

Valores > 10,0 podem indicar um problema de colinearidade

SEG	2,358
TER	1,949
QUA	1,861
QUI	1,921
SEX	2,134
SAB	1,742
TEUIMP_1	1,320
TEUIMP_4	1,459
CAMSAI_1	1,548

$VIF(j) = 1/(1 - R(j)^2)$, onde $R(j)$ é o coeficiente de correlação múltipla entre a variável j e a outra variável independente

Fonte: Autoria Própria.

Os sinais positivos das variáveis indicam que as mesmas tendem a se moverem na mesma direção que a variável dependente, sendo que variações de 1 p.p. no volume de caminhões saindo no dia anterior provocaram variação de 0,23 p.p. no volume de caminhões saindo no dia analisado; e que variações de 1 p.p. no volume de TEUs exportados há um e a quatro dias provocaram variação, respectivamente, de 0,16 e 0,08 p.p. sobre o volume de caminhões saindo no dia analisado. A quarta-feira por si só afetou 260,38 p.p. o volume de caminhões entrando no dia analisado; a quinta-feira, 210,76 p.p.; a segunda-feira, 207,67 p.p.; a terça-feira, 168,38 p.p.; o sábado, 150,09 p.p.; a sexta-feira, 143,88 p.p.; e o domingo, 0,00 p.p.

As variáveis *dummy* para dia da semana mostraram que o volume de caminhões saindo do porto de Itajaí no início/meio da semana – quarta, quinta, segunda e terça-feira – é maior que o volume de caminhões saindo do porto no final da semana – sábado, sexta-feira e domingo; assim como ocorreu para o modelo de exportação.

5.6 ETAPA 5

Nessa etapa, iniciaram-se os testes de modelos de regressão de Poisson e distribuição binomial negativa pelo software Gretl e, posteriormente, utilizando as mesmas variáveis, os modelos de RNAs e SVM pelo software Matlab.

5.6.1 Regressão de Poisson e distribuição binomial negativa

Durante o processo de desenvolvimento de modelos de geração de viagens para o Porto de Itajaí (SC), optou-se também pela estimação do modelo de Poisson, pois este é utilizado quando as variáveis são dados discretos com unidade de medida contínua.

Foram utilizadas as mesmas variáveis que as utilizadas para o modelo de regressão linear, resultando em um modelo de Poisson, para o caso da exportação, com os parâmetros mostrados na Tabela 5.5.

Observa-se nessa tabela que o valor da estatística qui-quadrado (79,3271) para o teste de excesso de dispersão do modelo de Poisson proporciona um p-valor menor do que 5%, indicando que a variância é maior do que a média e o índice de dispersão é maior do que um, ou seja, o modelo não é válido, visto que infringe uma das restrições do modelo de Poisson – média e variância serem iguais.

Tabela 5.5 – Resultado do modelo de exportação para o porto de Itajaí utilizando regressão de Poisson

Variável Independente	Coefficiente	Erro Padrão	z	p-valor	
const	3,88542	0,0210765	184,3479	<0,0001	***
SEG	1,27601	0,0235661	54,1461	<0,0001	***
TER	1,48424	0,0220617	67,2770	<0,0001	***
QUA	1,36981	0,0221296	61,8992	<0,0001	***
QUI	1,26315	0,0225001	56,1398	<0,0001	***
SEX	1,1076	0,0227757	48,6310	<0,0001	***
SAB	0,864263	0,0233928	36,9457	<0,0001	***
TEUEXP_3	9,48974e-05	8,64314e-06	10,9795	<0,0001	***
CAMENT_1	0,00175258	3,02446e-05	57,9471	<0,0001	***
Média var. dependente	264,5774		D.P. var. dependente	161,8628	
Soma resíd. quadrados	2841727		E.P. da regressão	111,1546	
R ² de McFadden	0,517470		R ² ajustado	0,516815	
Log da verossimilhança	-6631,588		Critério de Akaike	13281,18	
Critério de Schwarz	13312,46		Critério Hannan-Quinn	13293,78	
rô	-0,006207		h de Durbin	2,012281	
Teste de excesso de dispersão: Qui-quadrado(1) = 79,3271 [0,0000]					

Fonte: Autoria Própria.

Nesse caso, uma alternativa à utilização do modelo de Poisson é a utilização do modelo de distribuição binomial negativa, também utilizado para dados de contagem. O resultado da aplicação desse modelo com os dados do porto de Itajaí é apresentado na Tabela 5.6 para o caso da exportação.

Tabela 5.6 – Resultado do modelo de exportação para o porto de Itajaí utilizando distribuição binomial negativa

Variável Independente	Coefficiente	Erro Padrão	z	p-valor	
const	3,84691	0,137183	28,0421	<0,0001	***
SEG	1,26322	0,154792	8,1608	<0,0001	***
TER	1,5077	0,145964	10,3293	<0,0001	***
QUA	1,45775	0,146104	9,9775	<0,0001	** *
QUI	1,34162	0,150412	8,9196	<0,0001	***
SEX	1,17827	0,151401	7,7825	<0,0001	***
SAB	0,963648	0,153698	6,2697	<0,0001	***
TEUEXP_3	0,00013128 6	6,10952e- 05	2,1489	<0,0001	***
CAMENT_1	0,0016322	0,00020764 8	7,8604	<0,0001	***
alpha	55,0288	5,35714	10,2720	<0,0001	***
Média var. dependente	264,5774		D.P. var. dependent e	161,862 8	
Log da verossimilhança	-1443,210		Critério de Akaike	2906,41 9	
Critério de Schwarz	2941,184		Critério Hannan-Quinn	2920,42 8	
rô	-0,006207		h de Durbin	2,01228 1	

Fonte: Autoria Própria.

Observa-se nessa tabela que o parâmetro alpha foi significativamente diferente de zero ao nível de 5%, o que quer dizer que o parâmetro de dispersão afeta os valores esperados, reforçando a hipótese de que a distribuição de Poisson não é adequada aos dados ao contrário da distribuição binomial negativa.

As variáveis dependentes e independentes tendem a se moverem na mesma direção, sendo que variações de 1 p.p. no volume de caminhões entrando no dia anterior, bem como no volume de TEUs exportados há três dias, não provocaram variações no volume de caminhões entrando no dia analisado. A terça-feira por si só afetou 1,50 p.p. o volume de caminhões entrando no dia analisado; a quarta-feira, 1,46 p.p.; a quinta-

feira, 1,34 p.p.; a segunda-feira, 1,26 p.p.; a sexta-feira, 1,17 p.p.; o sábado, 0,96 p.p.; e o domingo, não afetou o volume de caminhões. Assim como ocorreu no modelo de regressão linear, as variáveis *dummy* para dia da semana mostraram que o volume de caminhões entrando no porto de Itajaí no início/meio da semana é maior que o volume de caminhões entrando no porto no final da semana.

O mesmo procedimento foi feito para o caso da importação e seus resultados podem ser observados nas Tabelas 5.7 e 5.8.

O valor da estatística qui-quadrado (85,536) para o teste de excesso de dispersão do modelo Poisson proporciona um p-valor menor do que 5%, indicando que a variância é maior do que a média e o índice de dispersão é maior do que u_i ; e o parâmetro alpha do modelo de distribuição binomial negativa foi significativamente diferente de zero ao nível de 5%, o que quer dizer que o parâmetro de dispersão afeta os valores esperados. Tendo isso em vista, pode-se supor que a distribuição de Poisson não é adequada aos dados ao contrário da distribuição binomial negativa.

As variáveis dependentes e independentes tendem a se moverem na mesma direção, sendo que variações de 1 p.p. no volume de caminhões saindo no dia anterior, bem como no volume de TEUs exportados há um e três dias, não provocaram variações no volume de caminhões saindo no dia analisado. A quarta-feira por si só afetou 1,41 p.p. o volume de caminhões entrando no dia analisado; a segunda-feira, 1,29 p.p.; a quinta-feira, 1,27 p.p.; a terça-feira, 1,15 p.p.; a sexta-feira, 1,06 p.p.; o sábado, 1,03 p.p.; e o domingo, não afetou o volume de caminhões. Assim como ocorreu no modelo de regressão linear e de Poisson para a exportação, as variáveis *dummy* para dia da semana mostraram que o volume de caminhões saindo do porto de Itajaí no início/meio da semana é maior que o volume saindo do porto no final da semana.

Tabela 5.7 – Resultado do modelo de importação para o porto de Itajaí utilizando regressão de Poisson

Variável Independente	Coefficiente	Erro Padrão	z	p-valor	
const	4,16336	0,0191078	217,8882	<0,0001	** *
SEG	1,06031	0,0206321	51,3910	<0,0001	** *
TER	0,897172	0,0193838	46,2845	<0,0001	** *
QUA	1,19921	0,0189258	63,3641	<0,0001	** *
QUI	1,05807	0,0195515	54,1174	<0,0001	** *
SEX	0,847028	0,0199914	42,3696	<0,0001	** *
SAB	0,856099	0,0192524	44,4672	<0,0001	** *
TEUIIMP_1	0,000490802	8,31128e-06	59,0525	<0,0001	** *
TEUIIMP_4	0,000222978	8,79796e-06	25,3442	<0,0001	** *
CAMSAI_1	0,000822827	2,87709e-05	28,5992	<0,0001	** *
Média var. dependente	307,9289		D.P. var. dependent e	162,5641	
Soma resíd. quadrados	3154897		E.P. da regressão	117,3748	
R ² de McFadden	0,465817		R ² ajustado	0,465037	
Log da verossimilhança	-6848,911		Critério de Akaike	13717,82	
Critério de Schwarz	13752,59		Critério Hannan-Quinn	2952,336	
rô	-0,066913		h de Durbin	2,127814	
Teste de excesso de dispersão: Qui-quadrado(1) = 85,536 [0,0000]					

Fonte: Autoria Própria.

Tabela 5.8 – Resultado do modelo de importação para o porto de Itajaí utilizando distribuição binomial negativa

Variável Independente	Coefficiente	Erro Padrão	z	p-valor	
const	3,89555	0,144905	26,8835	<0,0001	** *
SEG	1,29018	0,162673	7,9311	<0,0001	** *
TER	1,15316	0,152893	7,5423	<0,0001	** *
QUA	1,41273	0,14576	9,6922	<0,0001	** *
QUI	1,26744	0,148964	8,5083	<0,0001	** *
SEX	1,06101	0,155916	6,8050	<0,0001	** *
SAB	1,02863	0,150013	6,8569	<0,0001	** *
TEUIMP_1	0,00044995 5	6,37548e- 05	7,0576	<0,0001	** *
TEUIMP_4	0,00023988 1	6,59436e- 05	3,6377	<0,0001	** *
CAMSAI_1	0,00104677	0,00021271 6	4,9210	<0,0001	** *
alpha	63,918	6,29328	10,1565	<0,0001	** *
Média var. dependente	307,9289		D.P. var. dependente	162,564 1	
Log da verossimilhança	-1481,382		Critério de Akaike	2984,76 5	
Critério de Schwarz	3023,006		Critério Hannan-Quinn	3000,17 5	
rô	-0,066913		h de Durbin	2,12781 4	

Fonte: Autoria Própria.

5.6.2 Redes neurais artificiais

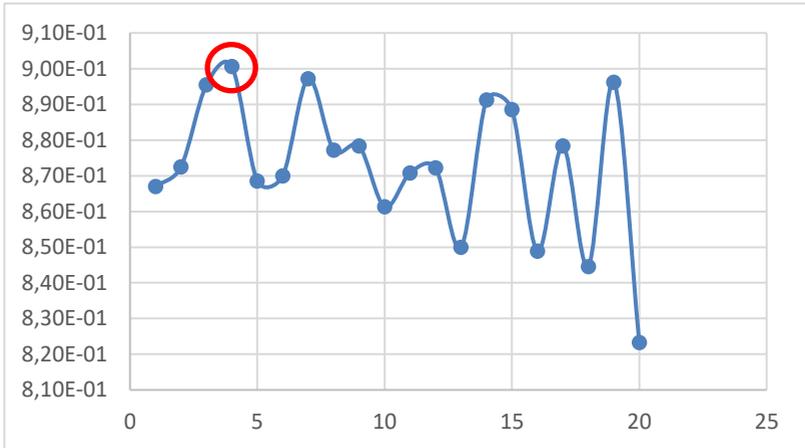
Outra técnica testada na busca de modelos de geração de viagens para o Porto de Itajaí (SC), foi a de RNAs, uma vez que a mesma possui como diferencial a capacidade de capturar a não linearidade por detrás dos dados, resultando, muitas vezes, em um desempenho superior quando comparada a tradicionais técnicas de regressão.

Cabe ressaltar, no entanto, que apesar de ter essa vantagem, a técnica de RNA pode apresentar *overfitting* ou *overtraining* se não tomado os devidos cuidados, o que pode levar a rede a memorizar os dados de treinamento. Além disso, é muito difícil definir a arquitetura ideal da RNA, uma vez que não existem regras claras para se definir quantas unidades devem existir nas camadas intermediárias, quantas camadas, ou como devem ser as conexões entre essas unidades. E há, ainda, o fato de as RNAs serem vistas como "caixas pretas", na qual quase não se sabe porque a rede chega a um determinado resultado, uma vez que os modelos não apresentam justificativas para suas respostas.

Na definição da arquitetura da RNA, foi utilizado o default do programa Matlab – Perceptron Multicamada (*Multilayer Perceptron* – MLP) com uma camada oculta, algoritmo de aprendizado Levenberg-Marquardt e como função de ativação a tangente hiperbólica – por ser um dos mais precisos e pela simplicidade da sua implementação, sendo um dos algoritmos mais recentemente utilizados.

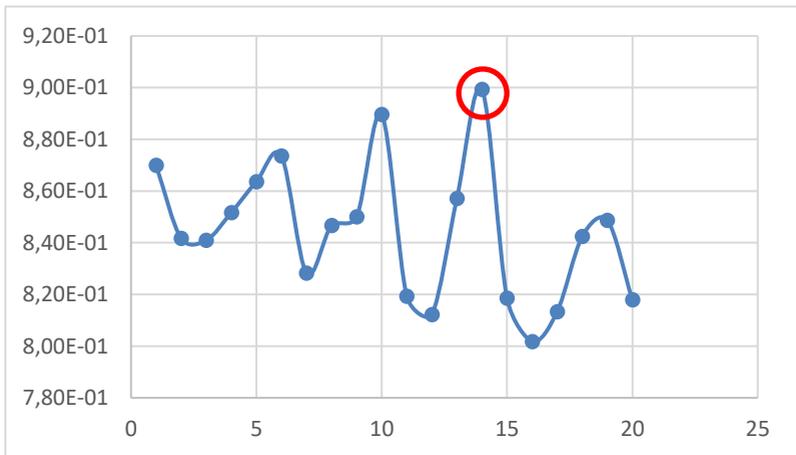
Para definição do número de neurônios na camada oculta, visando atenuar efeitos de *overfitting* ou *overtraining*, foi adotada a seguinte regra: foi testado um intervalo de um a 20 neurônios, por diversas vezes cada um – cem vezes – até gerar seu melhor resultado, avaliado por meio dos respectivos R^2 do subconjunto de validação; e, após isso, foi identificado o número de neurônios em que o R^2 se estabilizou e, no qual, a partir desse ponto, os ganhos de precisão foram pequenos. Dessa forma, o maior valor de R^2 (0,81 para exportação e 0,81 para importação) obtido foi com, respectivamente, 4 e 14 neurônios na camada oculta conforme mostram as Figuras 5.14 e 5.15.

Figura 5.14 – Gráfico com os melhores resultados de R^2 do subconjunto de validação segundo o número de neurônios na camada oculta para os modelos de exportação



Fonte: Autoria Própria.

Figura 5.15 – Gráfico com os melhores resultados de R^2 do subconjunto de validação segundo o número de neurônios na camada oculta para os modelos de importação



Fonte: Autoria Própria.

A rede originada para prever o volume de caminhões saindo (modelo de exportação) possui como variáveis de entrada: segunda, terça, quarta, quinta, sexta-feira e sábado, volume de TEUs exportados há três dias, volume de caminhões entrando no porto no dia anterior; assim como os modelos de regressão linear e binomial negativo resultantes.

Já a rede originada para prever o volume de caminhões entrando (modelo de importação) possui como variáveis de entrada: segunda, terça, quarta, quinta, sexta-feira, sábado, as variáveis regressivas volume de TEUs exportados há quatro e um dia e volume de caminhões saindo do porto há um dia; assim como os modelos de regressão linear e binomial negativo resultantes.

Apesar de ambas as redes terem resultado em um R^2 de 0,81, as RNAs tem uma estrutura razoavelmente complexa para que um pesquisador possa interpretar os resultados do modelo de previsão de inadimplência. Mesmo se tendo os pesos de cada neurônio, não há como mensurar as variações no volume de caminhões resultantes em variações, por exemplo, de 1 p.p. em cada uma das variáveis independentes de forma similar como ocorre em uma regressão linear.

5.6.3 *Support vector machine*

Por fim, elaborou-se também um modelo com a técnica SVM, que recentemente vem sendo utilizada como uma alternativa as RNA. Os SVMs são classificadores binários, ou seja, aceitam a classificação dos dados em apenas duas classes.

Como a variação de caminhões no porto de Itajaí pode ser de 0 a 1.000 caminhões por dia, existiriam 1.000 possibilidades de classes, o que não é compatível com a abordagem do SVM. Com isso, uma forma de contornar essa limitação e aplicar SVM para problemas como abordado nesse trabalho, é a geração de SVM especialistas, ou seja, para esse caso, criar 1.000 SVM especialistas, um deles em identificar se as variáveis de entrada se enquadram ou não (classificação binária) na saída de um caminhão, outro para identificação de dois caminhões, outro para identificação de três caminhões, e assim por diante.

Entretanto, devido à escassez de dados, não foi possível apresentar grande variabilidade de dados para o treinamento de 1.000 SVM especialistas. Caso fossem apresentados poucos dados para treinamento de cada SVM especialista, poderiam ocorrer problemas de *overfit* ou

overtraining, onde os SVM não identificariam condições diferentes das entradas com as quais foram treinados. Com isto, houve a necessidade de agregar alguns dados e criar intervalos de treinamento – SVM especialistas na identificação de intervalos de 20 em 20, 30 em 30, 50 em 50 caminhões. O modelo resultaria em uma previsão mais genérica que os anteriores, mas desse modo se teria mais condições e dados para treinar cada algoritmo do SVM.

Levando isso em consideração, foram feitos alguns testes para separação das classes, usando polinomial e RBF, as funções mais utilizadas. Em todos os testes foram utilizadas as mesmas variáveis independentes que compõem os modelos de regressão linear, binomial negativo e RNAs resultantes. O algoritmo desenvolvido para elaboração dos modelos é apresentado no Apêndice A.

Para o intervalo de 20 a 20 caminhões, a polinomial de quinto grau foi a que gerou maiores quantidades de acertos para cada teste – 80,99% de acertos para o modelo de exportação e 99,60% de acertos para o modelo de importação - frente ao polinomial de terceiro, quarto e segundo grau, e ao RBF 1 e 2 sigma.

Aumentando o intervalo para formação das classes, se teria mais dados para treinar o algoritmo, o que evitaria problemas de *overtrainig*. Dessa forma, os testes para um intervalo de 30 a 30 caminhões e para o polinomial de quinto grau, resultaram em uma quantidade de acertos de 78,19% para o modelo de exportação e 99,59% para o modelo de importação.

Aumentando ainda para o intervalo de 50 a 50 caminhões, a percentagem de acertos resultou, para o polinomial de quinto grau, em 78,51% para o modelo de exportação e 99,18% para o modelo de importação.

Para o intervalo de 30 a 30 e de 50 a 50 caminhões também foram feitos testes com outras funções, mas a polinomial de quinto grau foi a que apresentou maior quantidade de acertos.

Conforme pode-se observar, o SVM com o intervalo de 20 a 20 caminhões foi o que apresentou a maior percentagem de acertos tanto para o modelo de importação quanto de exportação, seguido pelo SVM com o intervalo de 30 a 30 caminhões e, por último, o SVM com o intervalo de 50 a 50 caminhões.

Esse fato poderia ser explicado pelo *overtraining*, ou seja, uma possível superespecialização da rede devido à pouca disponibilidade de dados no caso do SVM com o intervalo de 20 a 20 caminhões. No entanto,

nota-se que mesmo aumentando-se para o intervalo de 30 a 30 caminhões, ou ainda, para o intervalo de 50 a 50 caminhões, no qual a quantidade de dados passou a ser mais que o dobro, a quantidade de acertos reduziu relativamente pouco, o que leva a crer que essa precisão dos resultados se deve a aderência satisfatória do SVM aos dados.

Por fim, ressalta-se que assim como as RNAs, o SVM também possui uma estrutura que torna não trivial a interpretação dos resultados do modelo pelo pesquisador.

5.7 ETAPA 6

Uma comparação entre os resultados dos mesmos foi feita utilizando o Mape – calculado por meio da Equação 29 – e o Mdape – Equação 30 –, e uma comparação também foi feita com os resultados de estudos anteriores, de modo a subsidiar uma análise crítica sobre os modelos estudados.

$$MAPE = \frac{1}{N} \sum_{k=1}^N \left| \frac{v\hat{o}l(k) - vol(k)}{vol(k)} \right| \times 100\% \quad (29)$$

Onde:

$v\hat{o}l(k)$ = valor previsto para o período;

$vol(k)$ = valor real no período; e,

N = número de períodos de previsão.

$MdAPE = \text{median}(p_1, p_2, \dots, p_N)$, sendo que (30)

$$p_t = \left| \frac{y_t - f_t}{y_t} \right|$$

Onde:

y_t = é o valor real do resultado no período t ; e

f_t = é o valor da previsão no período t .

A diferença entre o Mape e o Mdape é que é baseado no cálculo da média e o segundo na da mediana. Devido a sua fragilidade (esconder discrepantes). O Mdape aparece como uma alternativa para reduzir os efeitos dos “outliers”, já que o mesmo se utiliza de uma metodologia de corte que descarta os altos e baixos valores de erros.

Para efeitos de comparação, foi elaborada uma tabela contendo o Mape e o Mdape calculado de todos os modelos desenvolvidos para o porto de Itajaí – respectivamente, Tabela 5.9 e 5.10 –, exceto para o modelo de regressão de Poisson que não se mostrou válido.

Tabela 5.9 – Resultado de Mape para os modelos desenvolvidos para o porto de Itajaí

Modelo/Mape	Exportação	Importação
Regressão linear	225,74	143,34
Binomial negativa	162,65	145,42
MLP de RNA	334,05	131,89
SVM p/ intervalo de 20 e polinomial de quinto grau	63,67	16,87
SVM p/ intervalo de 30 e polinomial de quinto grau	63,36	27,08
SVM p/ intervalo de 50 e polinomial de quinto grau	80,61	43,32

Fonte: autoria própria

Tabela 5.10 – Resultado de Mdape para os modelos desenvolvidos para o porto de Itajaí

Modelo/Mdape	Exportação	Importação
Regressão linear	297,46	262,21
Binomial negativa	291,37	274,45
MLP de RNA	331,42	220,46
SVM p/ intervalo de 20 e polinomial de quinto grau	4,14	3,26
SVM p/ intervalo de 30 e polinomial de quinto grau	6,86	5,39
SVM p/ intervalo de 50 e polinomial de quinto grau	11,29	7,76

Fonte: autoria própria

Tanto em termos de Mape quanto de Mdape, para os modelos de exportação, o modelo SVM foi o que pareceu mais preciso, seguido pelo modelo de distribuição binomial negativa, de regressão linear e pelo RNA. Enquanto que, para o modelo de importação, o modelo SVM foi o que apresentou melhor resultado de Mape e Mdape, seguido pela RNA, pela regressão e pelo binomial, ressaltando que esses três últimos apresentaram resultados muito similares.

Ressalta-se que em todos os modelos encontrados, como esperado, as variáveis independentes que se mostraram significativas foram as mesmas, sendo que tanto na regressão linear quanto na binomial negativa mostraram que os dias da semana, principalmente o início e meio da semana, produzem mais variações no volume de caminhões entrando e saindo do porto, seguido pelo volume de caminhões registrados nos dias anteriores e, por fim, no volume de contêineres transportados. A RNA e o SVM, devido à sua estrutura relativamente complexa não possibilitaram esse tipo de interpretação dos resultados.

Em relação as diferenças de aderência dos dados encontradas nos modelos, principalmente do SVM e RNA em relação à regressão linear e binomial negativa, reiteram-se o fato de que tanto o SVM quanto a RNA não exigem a especificação explícita de uma formulação de modelo, possui uma capacidade forte de aproximação da função, e podem modelar relações não lineares. Como portos marítimos geralmente possuem uma complexidade maior no que diz respeito ao volume de viagens atraídas e geradas por ele em relação a outros polos, essas técnicas conseguiram obter distribuições que se ajustaram melhor aos dados amostrais.

No entanto, apesar do modelo SVM ter apresentado melhor precisão tanto para a exportação quanto para a importação, considerando o cálculo do Mape e Mdape, deve-se fazer a ressalva de que, no caso estudado, esse modelo resultou em uma previsão mais genérica, visto que seus dados de saída são intervalos de valores. Isso ocorreu, conforme supracitado, devido ao tamanho e característica da amostra, sendo que para previsões menos genéricas, seria necessária uma amostra maior.

Logo, a desvantagem da utilização desse modelo, além da necessidade de uma amostra grande, é que nenhum dos dois softwares utilizados possui um default pronto para desenvolver o modelo e, portanto, há a necessidade de conhecimento de preparação de algoritmos. Além disso, é importante destacar que o SVM tem limitações similares as RNAs no que diz respeito a serem classificadores do tipo caixa-preta, ou seja, não permitem interpretação da estratégia de decisão como as

árvores; e o mesmo também pode vir a ter problemas de *overfitting* ou *overtraining*, caso não seja evitado o excesso de treinamento.

O modelo de distribuição binomial negativa também se mostrou uma alternativa interessante para o porto de Itajaí, visto que parece ter alcançado o segundo melhor resultado de Mape e Mdape para o modelo de exportação e, apesar de ter ficado na última posição quanto ao Mape e Mdape para o modelo de importação, este foi bastante similar aos Mapes e Mdapes calculados para os outros modelos – regressão linear e RNA. O modelo é simples de executar, não necessita de amostras tão grandes quanto o SVM e RNA, não tem os problemas de *overfitting* e *overtraining*, e já havia sido aplicado com sucesso para outro porto na América do Sul.

5.8 ETAPA 7

Na etapa 7, a título de exemplo, foi realizada as estimativas de viagens de caminhões entrando no porto a partir do modelo de regressão linear de exportação (Equação 31) e sua projeção para o ano de 2020.

$$CAMENT = -37,500 + 194,364SEG + 246,217TER + 229,017QUA + 198,045QUI + 141,133SEX + 87,730SÁB + 0,032TEUEXP_{-3} + 0,496CAMENT_{-1} \quad (27)$$

Onde:

CAMENT = volume de caminhões entrando no porto no dia analisado;

SEG = variável *dummy* para segunda-feira;

TER = variável *dummy* para terça-feira;

QUA = variável *dummy* para quarta-feira;

QUI = variável *dummy* para quinta-feira;

SEX = variável *dummy* para sexta-feira;

SAB = variável *dummy* para sábado;

TEUEXP₋₃ = volume de contêineres (TEU) exportados há três dias;

e

CAMENT₋₁ = volume de caminhões entrando no porto há um dia.

Cabe destacar que os dados encaminhados pela APM Terminals do porto de Itajaí somente contemplavam o volume de caminhões entrando

no porto e o volume de contêineres (TEU) exportados, ambos por dia, até o dia 31 de agosto de 2015.

Desse modo, para realizar esse exemplo, inicialmente foi feita uma projeção do volume de contêineres (TEU) exportados por dia para o período de 01 de setembro de 2015 a 31 de dezembro de 2020.

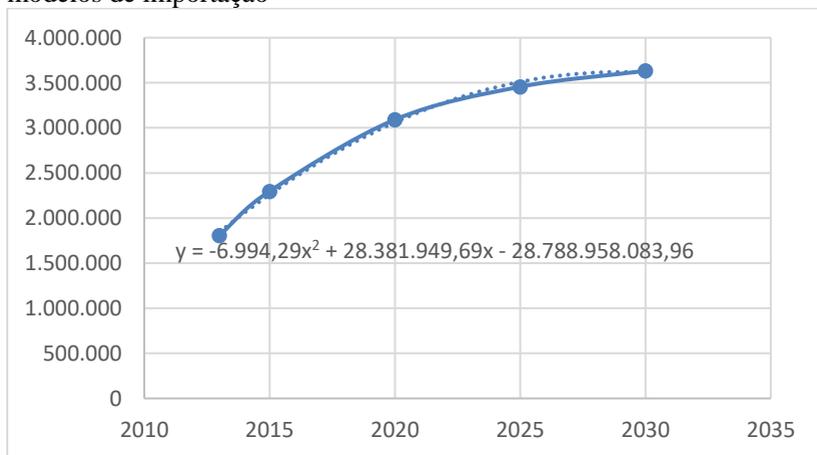
Para essa projeção, se considerou o crescimento do estudo do Labtrans (2015), apresentado na Tabela 5.11. Esses valores foram então plotados em um gráfico (Figura 5.17). Como as estimativas de crescimento não seguiam uma tendência linear, precisava-se identificar qual a equação utilizada para o crescimento do volume de cargas, de modo que se pudesse estimar também qual seria o crescimento para os anos de 2016, 2017, 2018 e 2019, o qual não foram apresentados no estudo do Labtrans (2015).

Tabela 5.11 – Projeção da demanda de cargas no porto de Itajaí entre os anos 2013 (observado) e 2030 (projetado) em toneladas para cargas do tipo contêineres, navegação de longo curso e sentido de embarque

2013	2015	2020	2025	2030
1.807.607	2.296.946	3.090.989	3.455.547	3.631.973

Fonte: Labtrans (2015).

Figura 5.16 – Gráfico com os melhores resultados de R² do subconjunto de validação segundo o número de neurônios na camada oculta para os modelos de importação



Fonte: Autoria Própria.

A linha de tendência que mais se aproximou da curva do gráfico foi a de uma polinomial de segundo grau, cuja equação está apresentada no gráfico. A partir dessa equação, pode-se estimar o volume de cargas e a taxa de crescimento para todos os anos de 2015 a 2020 (Tabela 5.12).

Tabela 5.12 – Volume de cargas estimado (segunda linha) e taxa de crescimento utilizada (terceira linha) para cada ano (primeira linha) entre o período de 2015 a 2020

2015	2016	2017	2018	2019	2020
2.279.426	2.467.393	2.641.371	2.801.361	2.947.361	3.079.374
1,10	1,08	1,07	1,06	1,05	1,04

Fonte: Labtrans (2015).

Aplicou-se então a taxa de crescimento de 2015 ao volume de contêineres (TEU) exportados por dia para o período de 01 de setembro a 31 de dezembro de 2014, para se estimar o volume de contêineres (TEU) exportados por dia para esse período de 2015. Em seguida, aplicou-se a taxa de crescimento de 2016 ao volume de contêineres (TEU) exportados por dia para o período de 01 de janeiro a 31 de dezembro de 2015, para se estimar o volume de contêineres (TEU) exportados por dia para esse período de 2016. E assim por diante até se obter a estimativa do volume de contêineres (TEU) exportados por dia para o período de 01 de janeiro a 31 de dezembro de 2020.

Quanto ao dado volume de caminhões entrando no porto há um dia, o mesmo foi estimado através da Equação 27. Como se tinha o dado do volume de caminhões entrando no porto para o dia 31 de agosto de 2015, pode-se estimar o volume de caminhões entrando no porto para o dia 01 de setembro de 2015. Com a estimativa do volume de caminhões entrando no porto para o dia 01 de setembro de 2015, pode-se estimar o volume de caminhões entrando no porto para o dia 02 de setembro de 2015, e assim por diante até o dia 31 de dezembro de 2020.

Foram obtidas então, as estimativas do volume de caminhões entrando no porto para cada dia do ano de 2020. Somando-se o volume de caminhões entrando no porto por dia para esse ano de 2020, resultou em uma estimativa de 100.791 caminhões entrando no porto de Itajaí (APM Terminals) para descarregar contêineres para exportação.

A distribuição e alocação modal dessas viagens não foi realizada, visto que os dados sobre a movimentação de contêineres fornecidos pela administração do porto de Itajaí (SC) não possuíam registros de origem/

destino dos mesmos. No entanto, deixa-se como sugestão para estudos posteriores.

De qualquer forma, como exemplificação dessas outras partes do modelo quatro etapas, cita-se os resultados do Plano Mestre para o porto de Itajaí (SC), realizados pelo Labtrans (2015), e apresentados no subitem 5.3.1, na página 160. A distribuição e alocação do volume de viagens na rede rodoviária nesse estudo e a posterior avaliação do desempenho dessa rede demonstrou que as obras de duplicação da BR-470 são fundamentais a fim de que o porto continue sendo uma opção viável para o escoamento da produção da região oeste; e que a BR-101 também necessitará de ampliação nos trechos analisados, visto que a demanda urbana em horários de pico esgota a capacidade da via e que nos anos futuros essa situação tende apenas a piorar.

Além disso, os coeficientes das variáveis *dummy* para dia da semana nos modelos encontrados na etapa 4 e 5 mostraram que o volume de caminhões saindo do porto de Itajaí no início/ meio da semana – segunda, terça, quarta e quinta-feira – é maior que o volume de caminhões saindo do porto no final da semana – sexta-feira, sábado e domingo.

Em conversa com o diretor do Porto de Itajaí ele afirmou que um número significativo de embarcações atraca no porto durante o sábado e o domingo. Devido a isso, esses dias apresentam menos movimento de caminhões, visto que, no caso da exportação, as cargas devem chegar ao porto anteriormente a esses dias para passarem pelo processo de alfândega e posterior carregamento nos navios; e, no caso da importação, as cargas devem passar por todo o processo de descarregamento e alfândega antes de serem liberadas para o seu destino através dos caminhões, o que requer algum tempo.

Portanto, uma melhor programação da distribuição do número de navios atracando no porto durante todos os dias da semana, poderia vir a ocasionar conseqüentemente uma melhor distribuição do volume de caminhões circulando no entorno do porto durante todos os dias da semana. Ou seja, seria reduzido o volume de caminhões no início/ meio da semana e aumentado o volume de caminhões ao menos na sexta-feira e no sábado (visto que domingo é um dia de trabalho que envolve maiores custos aos empresários do ramo de fretamento).

6 CONCLUSÕES E RECOMENDAÇÕES

6.1 CONCLUSÕES

Conforme visto no Capítulo 2, a maioria dos estudos sobre PGVs utiliza técnicas de regressão linear ou múltipla e como variável dependente o número de viagens de veículos e independente alguma que seja relacionada à área construída do empreendimento ou à demanda do mesmo como variáveis. Uma explicação para essa utilização poderia ser de que muitos estudos foram baseados no ITE, que foi o precursor do desenvolvimento de modelos de geração de viagens para PGVs, e utiliza essa técnica. Acredita-se que como a mesma é fácil de aplicar e se mostrou estatisticamente válida, não foi necessário o desenvolvimento de modelos matemáticos mais complexos para esses tipos de PGVs.

No entanto, a própria literatura destaca a complexidade de se estudar PGVs de carga em comparação aos outros PGVs, visto que são muitos fatores que influenciam no seu aspecto operacional. No caso especificamente dos portos marítimos, devido aos mesmos geralmente estarem inseridos em áreas urbanas, possuírem um tamanho significativo e movimentarem uma grande quantidade de cargas, e sofrerem maiores variações dos fatores que influenciam em seu aspecto operacional, pode-se dizer que os mesmos são ainda mais complexos de serem estudados.

Essa observação pode ser constatada através do Capítulo 3, no qual percebe-se que, ao contrário da tendência de utilizar basicamente a técnica de regressão linear simples ou múltipla para construir modelos de geração de viagens, no caso dos portos marítimos a maioria dos estudos analisados optou por fazer uso de modelos matemáticos mais complexos como redes neurais artificiais ou aprendizado de máquina baseado em *kernel*.

Além do que, a maioria dos estudos encontrados analisou apenas um porto marítimo, e preparou um modelo de geração de viagens específico para o mesmo, provavelmente porque não havia segurança de que o porto estudado possuía características similares a dos portos para os quais já existiam taxas e modelos de geração de viagens, e seu uso poderia superestimar ou subestimar a geração de viagens para o PGV a ser analisado. Essa é outra grande dificuldade, visto que os portos marítimos possuem grandes variações de localização, porte, atividade ou até mesmo operação, sendo muitas vezes necessário elaborar um modelo

específico, que compreenda as especificidades locais e sua atualidade, para uma avaliação mais adequada de seus impactos.

Logo, o desenvolvimento do método para construção de modelos de geração de viagens para caminhões em portos/ terminais marítimos tem como objetivo contribuir para que o planejador de transportes possua uma ferramenta para desenvolver um modelo de geração de viagens para o porto marítimo que deseja estudar, e assim possa melhor avaliar e tratar os impactos de tráfego do entorno desse PGV.

A evolução das técnicas de modelagem, as quais estão sendo aplicadas nas análises de impacto no entorno de portos/ terminais marítimos, e seus resultados positivos mostraram que as mesmas devem ser, de algum modo, incorporadas durante o desenvolvimento de modelos de geração de viagens para esses empreendimentos.

O método proposto buscou detalhar em passo a passo como encontrar modelos de geração de viagens, de modo a obter uma estimativa das viagens geradas por esses tipos de PGVs, considerando essa evolução das técnicas de modelagem, assim como as variáveis explicativas comumente utilizadas na construção dos modelos.

Após desenvolvido o método, buscou-se aplica-lo a um porto no Brasil, visto que não foi encontrado modelos de geração de viagens desenvolvidos para portos marítimos no País, e essa aplicação, além de possibilitar verificar se o método cumpre seu objetivo, ainda resultaria em um modelo que poderia vir a ser utilizado para portos brasileiros similares ao de Itajaí.

A aplicação do método ainda foi importante para apresentar critérios utilizados para a montagem dos modelos com diferentes técnicas, tais como regressão linear múltipla, regressão de Poisson e distribuição binomial negativa, redes neurais artificiais e aprendizado de máquina baseado em *kernel* (SVM); de modo a auxiliar outras pessoas que desejem aplicar uma dessas técnicas a outro porto/ terminal marítimo. Além de ter possibilitado fazer uma comparação entre o desempenho do modelo de regressão de Poisson e de distribuição binomial negativa com outros modelos, inclusive com RNAs, a qual não havia sido feita nos estudos encontrados.

Sobre os modelos desenvolvidos, as variáveis independentes que foram significantes no modelo foram o volume de TEUs autorregressivos (valores defasados do volume de TEUs) e os dias da semana, similar aos estudos anteriores sobre o assunto. Ressalta-se ainda que as variáveis *dummy* para dia da semana mostraram que o volume de caminhões

entrando no porto de Itajaí no início/meio da semana é maior que o volume de caminhões entrando no porto no final da semana. Isso pode ser explicado pelo fato de muitos navios atracam no porto no final de semana para serem descarregados e carregados com TEUs e como estes permanecem alguns poucos dias armazenados, eles geram maior movimentação de caminhões ao longo do início/meio da semana.

Entretanto, cabe ressaltar que uma variável independente que não havia sido levada em consideração nos estudos anteriores e que foi essencial para o desenvolvimento do modelo para o porto de Itajaí é o volume de caminhões no dia anterior ao analisado, que teve coeficiente angular maior até que o volume de contêineres transportados, ou seja, essa variável tem mais influência no volume de caminhões saindo no dia analisado do que as outras variáveis independentes.

Destaca-se ainda que se optou por elaborar modelos em separado para a exportação de TEUs (caminhões entrando) e importação de TEUs (caminhão saindo), visto que essas duas operações possuíam características muito diferentes no porto de Itajaí (SC). No geral, todos os modelos tiveram resultados menores de Mape e Mdape para os modelos de importação do que para os modelos de exportação, ou seja, os dados previstos se aproximaram mais dos dados reais no caso do modelo de importação do que no caso do modelo de exportação. Isso pode ser explicado pela variação menor dos dados de volume de caminhões por dia no caso da importação quando comparada ao caso de exportação. O mesmo não foi identificado nos estudos realizados em outros países.

É fato que todas as técnicas utilizadas têm suas vantagens e desvantagens. No entanto, para o caso do porto de Itajaí, a técnica que pareceu mais vantajosa, considerando apenas o valor de Mape e Mdape, foi o SVM. Nesse caso, cabe destacar ainda que o SVM apresentou resultados de Mape e Mdape relativamente superiores aos outros modelos, inclusive aos de RNA, diferente do estudo de Xie & Huynh (2010a e 2010b) para o porto de Texas, no qual o SVM apresentou resultados de Mape e Mdape similares e/ou pouco maiores do que os de RNA.

Apesar do modelo SVM ter apresentado valores vantajosos de Mape e Mdape, deve-se fazer a ressalva de que, no caso estudado, esse modelo resultou em uma previsão mais genérica, visto que seus dados de saída são intervalos de valores. Isso se deve ao tamanho e característica da amostra que não possibilitou fazer uma previsão menos genérica. Logo, essa técnica necessita de amostras grandes para que possa fazer

uma previsão mais específica. Além disso, outra desvantagem dessa técnica é que nenhum dos dois softwares utilizados possui um default pronto para desenvolver o modelo e, portanto, há a necessidade de conhecimento de preparação de algoritmos, e o mesmo também pode vir a ter problemas de *overfitting* ou *overtraining* similares aos que podem ocorrer nas RNAs, caso não seja evitado o excesso de treinamento.

Certamente mais pesquisas precisam ser desenvolvidas para que se possa obter uma resposta conclusiva acerca do uso desta técnica para essa finalidade, mas esse estudo e o de Xie & Huynh já mostraram que a técnica pode ser empregada para previsão de viagens de caminhões em portos.

Apesar da técnica de binomial negativa ter resultado em valores piores de Mape e Mdape do que o SVM, a mesma é mais simples de aplicar que o SVM, pois este requer conhecimento em algoritmos e, nesse caso específico, em elaborar algoritmos especialistas para identificar cada dado de saída do modelo; como não sofre com problemas de *overtraining*, requerem um tamanho de amostra menor que o SVM; e seu layout de saída permite uma interpretação de maneira direta os coeficientes angulares das variáveis. Mais pesquisas também precisam ser desenvolvidas com essa técnica, a qual se mostrou pertinente em outras análises anteriormente realizadas.

Logo, técnicas como binomial negativa e SVM, as quais não são comumente utilizadas para previsão de tráfego, podem auxiliar na estimativa do fluxo de tráfego de caminhões nas áreas portuárias de modo a contribuir para que se faça um melhor planejamento da rede viária como um todo (caminhões, automóveis, ônibus, etc.) e melhoramentos na gestão de tráfego visando reduzir ou acomodar esse número de caminhões adequadamente, seja por meio da implantação de uma nova conexão ferroviária, da criação de plataformas logísticas no entorno ou de outras medidas cabíveis.

Por fim, cabe ressaltar que a etapa 7 do procedimento, que diz respeito à comparação entre o nível de desempenho da situação atual com o nível de desempenho da situação futura e a sugestão de alternativas para mitigar ou reduzir os impactos, não foi realizada, visto que os dados sobre a movimentação de contêineres fornecidos pela administração do porto de Itajaí (SC) não possuíam registros de origem/ destino dos mesmos.

No entanto, para exemplificar a aplicação da etapa 7, pode-se considerar o estudo realizado pelo Labtrans (2015) no Plano Mestre do porto de Itajaí (SC), o qual fez essa comparação entre as suas situações –

atual e futura – e identificou como medidas mitigadoras: duplicação da BR-470; e ampliação da BR-101 nos trechos analisados.

Somado a isso, a aplicação do método dessa tese também identificou, através da interpretação dos coeficientes das variáveis *dummy* para dia da semana nos modelos, que uma distribuição mais equitativa do número de navios atracando no porto durante todos os dias da semana, poderia resultar em uma melhor distribuição do fluxo de caminhões ao longo da semana.

6.2 RECOMENDAÇÕES E SUGESTÕES PARA NOVAS PESQUISAS

Com base nos resultados obtidos neste estudo, assim como o conhecimento adquirido no seu desenvolvimento, recomenda-se o seu prosseguimento mediante a realização de seguintes estudos para:

- a) aplicação dos modelos de viagens encontrados para o porto de Itajaí (SC) em outro porto similar com finalidade de testar sua transferibilidade;
- b) constante aprimoramento e atualização do método proposto para construção de modelos de geração de viagens para caminhões em portos marítimos;
- c) pesquisar mais sobre o resultado do uso da técnica de regressão de Poisson e distribuição binomial negativa para construção de modelos de geração de viagens para portos marítimos;
- d) pesquisar mais sobre o resultado do uso da técnica de com aprendizado de máquina baseada em *kernel* para construção de modelos de geração de viagens para portos marítimos; e
- e) verificar a possibilidade da utilização da técnica de lógica *fuzzy* para construção de modelos de geração de viagens para portos marítimos.

REFERÊNCIAS

AGNELLO, P.; JONNES, J.; ALLEN JR., W. G. **An innovative approach to truck modeling**. Proceedings of 9th TRB Conference on the Application of Transportation Planning Methods, Baton Rouge, Luisiana, v. 16, p. 362-371. 2004.

AGUIAR JR., S. R. **Modelo RAPIDE: uma aplicação de mineração de dados e redes neurais artificiais para a estimativa da demanda por transporte rodoviário interestadual de passageiros no Brasil**. Dissertação (Mestrado). Programa de Pós-Graduação em Gestão do Conhecimento e da Tecnologia da Informação, Universidade Católica de Brasília, Brasília. 2004.

AL-DEEK, H. M.; JOHNSON, G.; MOHAMED, A.; EL-MAGHRABY, A. **Truck trip generation models for seaports with container and trailer operation**. Transportation Research Record, v. 1719, p. 1-9. 2000.

AL-DEEK, H. M. **Comparison of two approaches for modeling freight movement at seaports**. Journal of Computing in Civil Engineering, v. 15, n. 4, p. 284-291. 2001a.

_____. **Which method is better for developing freight planning models at seaports - neural networks or multiple regression?** Transportation Research Record, v. 1763, p. 90-97. 2001b.

_____. **Use of vessel freight data to forecast heavy truck movements at seaports**. Transportation Research Record, v. 1804, p. 217-224. 2002.

ALVES, V. F. B.; SILVA, A. N. R.; WAERDEN, P. **Técnicas exploratórias para localizar potenciais usuários de transporte público urbano**. Journal of Transport Literature, v. 6, n. 3, p. 180-203. 2012.

ARLIANSYAH, J.; HARTONO, Y. **Trip attraction model using radial basis function neural networks**. Procedia Engineering, v. 125, p. 445-451. 2015.

BASTOS, M. A. B. **Uma reflexão sobre as necessidades de mudança na gestão dos polos geradores de tráfego**. Dissertação (Mestrado). Programa de Pós-Graduação em Engenharia de Transportes, Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2004.

BATISTA, B. C. F. **Soluções de equações diferenciais usando redes neurais de múltiplas camadas com os métodos da descida mais íngreme e Levenberg-Marquardt**. Dissertação (Mestrado). Programa de Pós-Graduação em Matemática, Universidade Federal do Pará, Bélem. 2012.

BENCHEIKH, N.; LANDRY, R.; AMARA, N. **Lessons from innovation empirical studies in the manufacturing sector: A systematic review of the literature from 1993-2003**. *Technovation*, v. 26, p. 644-664. 2006.

BERETON, P.; KITCHENHAM, B.; BUDGEN, D.; TURNER, M.; KHALIL, M. **Lessons from applying the sistematic literature review process within the software engineering domain**. *The Journal of System and Software*, v. 80, p. 571-583. 2007.

BERGER, A.; BERGER, F. R. **Portos e Terminais Marítimos do Brasil**. Editora Bela Catarina, 2ª edição. 2009.

BLOOMBERG. **How to fix world transportation**. 2011. Disponível em: <<https://www.bloomberg.com/news/articles/2011-12-01/how-to-fix-world-transportation>> Acesso em: 28 nov. 2015.

CAF – Corporación Andina de Fomento. **Observatorio de Movilidad Urbana para América Latina**. Corporación Andina de Fomento, Bogotá, Colombia. 2010.

CAMERON, A. C.; TRIVEDI, P. K. **Microeconometrics using Stata**. Stata Press, Original edition, Texas. 2009.

CAVALCANTE, A. P. H. **Metodologia de previsão de viagens a polo gerador de tráfego de uso misto: estudo de caso para a cidade de Fortaleza**. Dissertação (Mestrado), Programa de Pós-Graduação em Engenharia de Transportes, Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2002.

CET-SP – Companhia de Engenharia de Tráfego de São Paulo. **Polos geradores de tráfego**. Boletim Técnico 32. Companhia de Engenharia de Tráfego, São Paulo. 1983.

CHARNET, R.; BONVINO, H; FREIRE C. A. L; CHARNET, E. M. R. **Análise de modelos de regressão linear com aplicações**. Unicamp, São Paulo, 354 f. 1999.

CHEN, L.; LANGEVIN, A.; LU, Z. **Integrated scheduling of crane handling and truck transportation in a maritime container terminal**. European Journal of Operational Research, v. 225, n. 1, p. 142-152. 2013.

CHU, H. **Empirical method for predicting internal-external truck trips at a major port**. Journal of Transportation Engineering, v. 137, n. 7, p. 496-508. 2011.

CNT - Confederação Nacional de Transporte. **Matriz de Transportes**. 2013. Disponível em: <<http://www.cnt.org.br/>> Acesso em: ago. 20, 2017.

COOK, D. J.; MULROW, C. D.; HAYNES, R. B. **Systematic reviews: synthesis of best evidence for clinical decisions**. Annals of Internal Medicine, v. 126, n. 5, p. 376-380. 1997.

CRISTIAN, D. **Developments of Maritime Transport Economy in Europe**. Annals of 17th Constanta Maritime University, p. 257-260. 2012.

CUNHA, R. F. F. **Uma sistemática de avaliação e aprovação de projetos de Polos Geradores de Viagens (PGVs)**. Dissertação (Mestrado), Programa de Pós-Graduação em Engenharia de Transportes, Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2009.

DENATRAN - Departamento Nacional de Trânsito. **Manual de procedimentos para o tratamento de polos geradores de tráfego**. Departamento Nacional de Trânsito, Ministério da Justiça, Brasília. 2001.

DÍAZ, S. **Modeling urban freight generation: A study of commercial establishments' freight needs**. Transportation Research Part A: Policy and Practice, v. 102, p. 3–17. 2017.

DUCRET, R.; GONZALEZ-FELIU, J. **Connecting demand estimation and spatial category models for urban freight: First attempt and**

research implications. Transportation Research Procedia, v. 12, p. 142–156. 2016.

FACCHINI, D. **Análise dos “GAPS” de percepção dos atores envolvidos no transporte urbano de carga em Porto Alegre.** Dissertação (Mestrado). Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal do Rio Grande do Sul, Rio Grande do Sul. 2006.

FAN, L.; WILSON, W. W.; DAHL, B. **Congestion, port expansion and spatial competition for US container imports.** Transportation Research Part E: Logistics and Transportation Review, v. 48, n. 6, p. 1121-1136. 2012.

FOODY, G. M.; MATHUR, A. **A Relative evaluation of multiclass image classification by support vector machines.** Transactions on Geoscience and Remote Sensing, v. 42, n. 6. 2004.

GASPARINI, A. **Atratividade do transporte de carga para polos geradores de viagem em áreas urbanas.** Dissertação (Mestrado), Programa de Pós-Graduação em Engenharia de Transportes, Instituto Militar de Engenharia, Rio de Janeiro. 2008.

GOLDNER, L. G.; SCHMITZ, A. **Estudo de polos múltiplos geradores de viagens: o caso do parque Montigalá em Barcelona.** Anais do XVI Congresso Pan-Americano de Engenharia de Tráfego e Transportes e Logística, Lisboa, Portugal. 2010.

GOLDNER, L. G.; WESTPHAL, D.; De FREITAS, I. M. D. P.; SANTOS, D. V. C. **Polos Múltiplos Geradores de Viagens (PMGV).** Revista Transportes, v. 18, n. 1, p. 113-121. 2010.

GONÇALVES, F. S. **Classificação dos PGVs e sua relação com as técnicas de análise de impactos viários.** Dissertação (Mestrado), Programa de Pós-Graduação em Engenharia de Transportes, Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2012.

GONÇALVES, D. N. S.; SILVA, M. A.; D’AGOSTO, M. A. **Procedimento para uso de redes neurais artificiais no planejamento estratégico de fluxo de carga no Brasil.** Journal of Transport Literature, v. 9, n. 1, p. 45-49. 2015.

GUJARATI, D. N. **Econometria básica**. Makron Books, São Paulo. 2000.

GUZMAN, L. A.; GOMEZ, A. M.; RIVERA, C. **A Strategic Tour Generation Modeling within a Dynamic Land-Use and Transport Framework: A Case Study of Bogotá, Colombia**. *Transportation Research Procedia*, v. 25C, pp. 2540-2555. 2017.

HARTMAN, B. C.; CLOTT, C. **Intermodal chassis supply in the US - A bayesian game model**. *Research in Transportation Business and Management*, v. 14, p. 66-71. 2015.

HERZ, M.; GALARRAGA, J. **Análisis de tasas y modelos para generación de viajes en hipermercados y supermercados**. *Journal of Transportation Literature*, v. 8, n. 3, p. 172-198. 2014.

HOLGUÍN-VERAS, J.; LÓPEZ-GENAO, Y.; SALAM, A. **Truck-trip generation at container terminals: results from a nationwide survey**. *Transportation Research Record*, v. 1790, p. 89-96. 2002.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Região de Influência das Cidades**. 2007. Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv40677.pdf>> Acesso em: set. 20, 2017.

_____. **Censo Demográfico**. 2010. Disponível em: <<https://censo2010.ibge.gov.br/>> Acesso em: set. 20, 2017.

ISLAM, S.; OLSEN, T.; AHMED, M. D. **Reengineering the seaport container truck hauling process: reducing empty slot trips for transport capacity improvement**. *Business Process Management Journal*, v. 19, n. 5, p. 752-782. 2013.

ITE - Institute of Transportation Engineers. **Trip Generation Handbook: an ITE recommended practice**. Institute of Transportation Engineers, Washington, EUA. 2004.

_____. **Trip Generation Handbook: an ITE recommended practice**. Institute of Transportation Engineers, Washington, EUA. 2008.

_____. **Transportation Planning Handbook**. Institute of Transportation Engineers, Washington, EUA. 2009.

_____. **Trip Generation Handbook: an ITE recommended practice.** Institute of Transportation Engineers, Washington, EUA. 2012.

KAPPANNA, H. K.; BESCH, M. C.; THIRUVENGADAM, A.; THIRUVENGADAM, P.; BONSACK, P.; CARDER, D. K.; OSHINUGA, A. **Evaluation of drayage truck chassis dynamometer test cycles and emissions measurement.** Proceedings of Internal Combustion Engine Division Fall Technical Conference, Vancouver, Canada, p. 563-575. 2012.

KLODZINSKI, J.; AI-DARASEH, A.; GEORGIPOULOS, M.; AL-DEEK, H. M. **Development of a java applet for generating truck trips from freight data.** Transportation Research Record, v. 1870, p. 10-17. 2004.

KLODZINSKI, J.; AL-DEEK, H. M. **Transferability of an intermodal freight transportation forecasting model to major Florida seaports.** Transportation Research Record, v. 1820, p. 36-45. 2003.

_____. **Methodology for modeling a road network with high truck volumes generated by vessel freight activity from an intermodal facility.** Transportation Research Record, v. 1873, p. 35-44. 2004.

KNEIB, E. C. **Caracterização de empreendimentos geradores de viagens: contribuição conceitual à análise de seus impactos no uso, ocupação e valorização do solo urbano.** Dissertação (Mestrado). Programa de Pós-Graduação em Transportes, Universidade de Brasília, Brasília. 2004.

LABTRANS - Laboratório de Transportes e Logística. **Plano Mestre, Porto Público de Itajaí.** 2015. Disponível em: <<http://www.MT.gov.br/assuntos-1/pnpl/arquivos/planos-mestres-sumarios-executivos/se14.pdf>> Acesso em: jul. 7, 2016.

LAZIC, M. B.; MORIAN, D. **Advanced approach in designing and building heavy duty pavements for high-density container terminal operation.** Proceedings of Port Conference, Houston, Texas, p. 1201-1210. 2004.

LEE, G.; YOU, S.; RITCHIE, S. G.; SAPHORES, J.; SANGKAPICHAI, M.; JAYAKRISHNAN, R. **Environmental impacts of a major freight corridor.** Transportation Research Record, v. 2123, p. 119-128. 2009.

LOPES, S. B. **Uma ferramenta para planejamento da mobilidade sustentável com base em modelos de uso do solo e transportes**. Tese (Doutorado), Programa de Pós-Graduação em Engenharia de Transportes, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Paulo. 2010.

LUK, J. K.; OLSZEWSKI, P.; WONG, Y-D.; LAM, S-H. **Road freight modelling in Singapore. Operations, transport and safety: Outside the square**. Proceedings of Institute Of Transportation Engineers International Conference, Melbourne, Victoria, Australia, 10 f. 2002.

MARRA, C. **Caracterização de demanda de movimentações urbanas de carga**. Dissertação (Mestrado), Programa de Pós-Graduação em Engenharia Civil, Universidade Estadual de Campinas, São Paulo. 1999.

MERK, O. **Competitiveness of Global Port Cities Synthesis Report**. 2013. Disponível em: <<https://www.oecd.org/cfe/regional-policy/Competitiveness-of-Global-Port-Cities-Synthesis-Report.pdf>> Acesso em: ago 20, 2015.

MT - Ministério dos Transportes. **Relações internacionais**. Disponível em: <<http://www.MT.gov.br/assuntos-1/relacoes-internacionais>> Acesso em: abr. 21, 2014.

_____. **Plano Nacional de Logística Portuária**. 2015. Disponível em: <<http://www.MT.gov.br/assuntos-1/pnpl/plano-nacional-de-logistica-portuaria>> Acesso em: set. 17, 2016.

MOMMENS, K.; VAN LIER, T.; MACHARIS, C. **Loading Unit in Freight Transport Modelling**. Procedia Computer Science, v. 83, p. 921-927. 2016.

MONTGOMERY, D. C.; VINING, G. G.; PECK, E. A. **Introduction to linear regression analysis**. John Wiley & Sons, 3ª edição, Nova Iorque, EUA, 641 f. 2001.

NETER, J.; KUTNER, M H.; NACHTSHEIM, C. J.; LI, W. **Applied linear statistical models**. Mc Graw-Hill/Irwin, 5ª edição, Nova Iorque, EUA, 1396 f. 2005.

NGAI, E. W. T.; WAT, F. K. T. **A literature review and classification of electronic commerce research, Information & Management**. Journal of Electronic Commerce Research, v. 39, p. 415-429. 2002.

NORD, J. H.; NORD, G. D. **MIS research: journal status assessment and analysis**. Information & Management, v. 29, n. 1, p. 29-42. 1995.

OLIVEIRA, C. M. **Aprimorando a sustentabilidade da logística por meio da gestão da operação do transporte de carga**. Qualificação (Doutorado). Programa de Pós-Graduação em Engenharia de Transportes, Programa de Pós Graduação de Engenharia de Transportes, Universidade Federal do Rio de Janeiro, Rio de Janeiro. 2015.

OLIVEIRA, L. K.; STUBBS, L. C. M.; GONTIJO, N. T.; OLIVEIRA, R. L. M. **Proposição de modelos de geração de viagens para Belo Horizonte**. Revisa Transportes, v. 25, n. 2, p. 137-155. 2017.

PAROLA, F.; MAUGERI, S. **Origin and taxonomy of conflicts in seaports: towards a research agenda**. Research in Transportation Business and Management, v. 8, p. 114-122. 2013.

PEREIRA, G. H. A. **Avaliação do uso combinado de métodos de modelagem da Direção do Escoamento Superficial e Modelos Digitais de Superfície Derivados de Laser Scanner em Regiões Urbanas**. Dissertação (Mestrado). Programa de Pós-Graduação em Engenharia de Recursos Hídricos e Ambiental, Setor de Tecnologia da Universidade Federal do Paraná, Curitiba. 2013.

PEREYRA, A.; RUBINSTEIN, E.; PÉREZ, M. **Tasa generadora de viajes para el Puerto de Montevideo – Una propuesta metodológica**. Anais do XV Congresso Panamericano de Engenharia de Trânsito e Transporte e Logística, Colombia. 2008.

PETROPOULOS, G. P.; ARVANITIS, K.; SIGRIMIS, N. **Hyperion hyperspectral imagery analysis combined with machine learning classifiers for land use/cover mapping**. Expert Systems with Applications. v. 39., n. 3., p. 3800-3809. 2012.

PINTO, A. B.; DIÓGENES, M. C.; LINDAU, L. A. **Quantificação dos impactos de polos geradores de tráfego**. 2003. Disponível em: <http://www.producao.ufrgs.br/arquivos/disciplinas/412_impactos_polo_gerador_versao_lindau_rev.doc> Acesso em: ago 13, 2012.

PORTOITAJAI **Perfil do porto**. Disponível em: <<http://www.portoitajai.com.br/novo/l/perfil-do-porto>> Acesso em: abr. 21, 2014.

- PORTUGAL, L. S. **Simulação de Tráfego – Conceitos e técnicas de modelagem**. Editora Interciência, Rio de Janeiro. 2005.
- PORTUGAL, L. S.; FLOREZ, J.; SILVA, A.N.R. **Rede de pesquisa em transportes: um instrumento de transformação e melhora da qualidade de vida**. Revista Transportes, v. XVIII, n. 1, p. 6-16. 2010.
- PORTUGAL, L. S. (Org.) **Polos geradores de viagens orientados à qualidade de vida e ambiental: modelos e taxas de geração de viagens**. Editora Interciência, Rio de Janeiro. 2012.
- PROTOPAPAS, A.; KRUSE, C.; OLSON, L. **Transportation rate analysis of waterborne trip movements on gulf intracoastal waterway west**. Transportation Research Record, v. 2222, p. 98-102. 2011.
- PROTOPAPAS, A.; KRUSE, C.; OLSON, L. **Transportation Rate Analysis of Waterborne Trip Movements Through the Calcasieu Lock**. Proceedings of 91st Annual Meeting of Transportation Research Board, Washington, DC, 11 f. 2012.
- QIAN, J.; ZHU, X.; XIE, X. **"Quays, trailers and yard" two way operations plan coordinated model**. Journal of Transportation Systems Engineering and Information Technology, v. 14, n. 2, p. 138-143. 2014.
- RAHIMI, M.; ASEF-VAZIRI, A.; HARRISON, R. **An inland port location-allocation model for a regional intermodal goods movement system**. Maritime Economics and Logistics, v. 10, n. 4, p. 362-379. 2008.
- RASMUSSEN, C. E.; WILLIAMS, C. K. I. (2006). **Gaussian Processes for Aprendizado de máquina**. The MIT Press, Cambridge, Massachusetts. 2006.
- REDEPGV - Rede Ibero-americana de Estudo em Polos Geradores de Viagens. **Conceitos**. Disponível em: <<http://redpgv.coppe.ufrj.br>> Acesso em: Ago 8, 2012.
- REY, P. A.; MUÑOZ, J. A.; WEINTRAUB, A. **A column generation model for truck routing in the chilean forest industry**. INFOR, v. 47, n. 3, p. 215-221. 2009.
- ROBERTS, B.; ROSE, A.; HEATWOLE, N.; WEI, D.; AVETISYAN, M.; CHAN, O.; MAYA, I. **The impact on the US economy of changes in wait times at ports of entry**. Transport Policy, v. 35, p. 162-175. 2014.

RODRIGUES, F. J. J. **Avaliação de compostos com atividade inibidora da glicoproteína-p**. Dissertação (Mestrado). Programa de Pós-Graduação em Bioquímica Aplicada, Universidade da Madeira, Funchal. 2009.

SANTOS, D. V. C. **Polos Geradores de Viagens Sustentáveis: uma proposta para o licenciamento e análise de Projetos**. Dissertação (Mestrado). Programa de Pós-Graduação em Engenharia Civil, Universidade Federal da Bahia, Salvador. 2011.

SARVAREDDY, P.; AL-DEEK, H.; KLODZINSKI, J.; ANAGNOSTOPOULOS, G. **Evaluation of two modeling methods for generating heavy-truck trips at an intermodal facility by using vessel freight data**. Transportation Research Record, v. 1906, p. 113-120. 2005.

SILVA, G. L.; FREITAS, I. M. D. P. **Um estudo sobre delimitação da área de influência em Polos Múltiplos Geradores de Viagens (PMGV)**. Anais do XXV Congresso de Pesquisa e Ensino em Transportes, Belo Horizonte. 2011.

SOARES, M. V. **Método para estabelecimento da capacidade de uma rede viária: análise dos efeitos da implantação de polos geradores de tráfego**. Dissertação (Mestrado). Programa de Pós-Graduação em Engenharia de Transportes, Universidade Federal do Rio de Janeiro, Rio de Janeiro. 1990.

STEPHENS, S. **PLM and steam generator replacement at SONGS**. Nuclear News, v. 49, n. 11, p. 47-50. 2006.

TORQUATO, T. L. L.; RAIÁ JÚNIOR; A. A. **Modelos de geração de viagens para condomínios residenciais horizontais**. Revista Transportes, v. 22, n. 1, p.56-64. 2014.

TRANFIELD, D.; DENYER, D.; SMART, P. **Towards a methodology for developing evidence-informed management knowledge by means of systematic review**. British Journal of Management, v. 14, p. 207-222. 2003.

TRB - Transportation Research Board. **Critical issues in transportation**. 2006. Disponível em:
<<http://onlinepubs.trb.org/Onlinepubs/general/criticalissues06.pdf>>
Acesso em: abr. 21, 2014.

UNCTAD – United Nations Conference on Trade and Development. **Review of Maritime transport**. 2012. Disponível em: <http://unctad.org/en/publicationslibrary/rmt2012_en.pdf> Acesso em: mar. 14, 2015.

WEISBERG, S. **Applied linear regression**. John Wiley & Sons, 3ª edição, Nova Iorque, EUA, 310 f. 2005.

XIE, Y.; HUYNH, N. **Kernel-based machine learning methods for modeling daily truck volume at seaport terminals**. Proceedings of 51st Annual Transportation Research Forum, Arlington, Virginia, v. 1, p. 409-427. 2010a.

_____. **Kernel-based machine learning models for predicting daily truck volume at seaport terminals**. Journal of Transportation Engineering, v. 136, n. 12, p. 1145-1152. 2010b.

YANG, Z.; CHEN, G.; MOODIE, D. R. **Modeling road traffic demand of container consolidation in a chinese port terminal**. Journal of Transportation Engineering, v. 136, n. 10, p. 881-886. 2010.

YOU, S. I.; RITCHIE, S. G. **Tour Behavior of Clean Drayage Trucks in Southern California**. Proceedings of 91st Annual Meeting of the Transportation Research Board, Washington, DC, 18 f. 2012.

YUEHUA, Z. H. U.; CHEN, Y.; GENG, X.; LIU, L. **Transport modal split of commercial sites based on artificial neural network**. Journal of Transportation, v. 8, n. 1. 2008.

ZIETSMAN, J.; PERKINSON, D. G. **Estimating extended idling emissions of heavy-duty diesel trucks in Texas**. Transportation Research Record, v. 1941, p. 34-42. 2005.

APÊNDICE

APÊNDICE A - Algoritmo SVM utilizado na elaboração dos modelos


```

%% Iniciando Processamento
clc
clear

display('===== Iniciando Leitura dos dados
=====')
display(['Iniciado às '
datestr(now, 'HH:MM:SS')])

%% Leitura dos dados
display ('Leitura dos dados...')

caminho = 'D:\Dropbox\SVM - Thais\';
% caminho para a pasta onde os dados se
encontram
arquivo_excel = 'Dados por dia.xlsx';
% nome da tabela do excel com extensão onde
estão os dados

[num_Dados_IMP,txt_Dados_IMP,raw_Dados_IMP] =
xlsread([caminho arquivo_excel],1);
[num_Dados_IMPIn,txt_Dados_IMPIn,raw_Dados_IMPIn
] = xlsread([caminho arquivo_excel],2);
[num_Dados_IMPOut,txt_Dados_IMPOut,raw_Dados_IMP
Out] = xlsread([caminho arquivo_excel],3);
[num_Dados_EXP,txt_Dados_EXP,raw_Dados_EXP] =
xlsread([caminho arquivo_excel],4);
[num_Dados_EXPIn,txt_Dados_EXPIn,raw_Dados_EXPIn
] = xlsread([caminho arquivo_excel],5);
[num_Dados_EXPOut,txt_Dados_EXPOut,raw_Dados_EXP
Out] = xlsread([caminho arquivo_excel],6);

%% Preparando os dados
display ('Preparando os dados...')

ModeloIMPDadosIn = num_Dados_IMPIn;
ModeloIMPDadosOut = num_Dados_IMPOut;

ModeloEXPDadosIn = num_Dados_EXPIn;
ModeloEXPDadosOut = num_Dados_EXPOut;

```

```

[mIMP,nIMP] = size(ModeloIMPdadoIn);
[mEXP,nEXP] = size(ModeloEXPdadoIn);

classesModeloIMP = unique(ModeloIMPdadoOut);
classesModeloEXP = unique(ModeloEXPdadoOut);

N_classesModeloIMP = max(max(classesModeloIMP));
N_classesModeloEXP = max(max(classesModeloEXP));

intervalo = 50; % quantidade
de caminhões utilizada para separação das
classes

N_IMP_intervalo =
ceil(N_classesModeloIMP/intervalo);
N_EXP_intervalo =
ceil(N_classesModeloEXP/intervalo);

ModeloIMPTreinoOut =
zeros(N_IMP_intervalo,nIMP);
ModeloEXPTreinoOut =
zeros(N_EXP_intervalo,nEXP);

for i = 1:N_IMP_intervalo
    for j = 1:nIMP
        if (ModeloIMPdadoOut(1,j) > intervalo*(i-
1)) && (ModeloIMPdadoOut(1,j) <= intervalo*i)
            ModeloIMPTreinoOut(i,j) = 1;
        end
    end
end

for i = 1:N_EXP_intervalo
    for j = 1:nEXP
        if (ModeloEXPdadoOut(1,j) > intervalo*(i-
1)) && (ModeloEXPdadoOut(1,j) <= intervalo*i)
            ModeloEXPTreinoOut(i,j) = 1;
        end
    end
end

```

```

end
end

%% Treinando os SVM
display ('Treinando SVM...')

ModeloSVM_IMP = cell(N_IMP_intervalo,1);
ModeloSVM_EXP = cell(N_EXP_intervalo,1);

% ===== parâmetros a serem alterados para
% modificar os SVM gerados
kernel_function = 'polynomial';
metodo = 'SMO';
rbf_sigma = 1;
polyorder = 6;
options = optimset('maxiter',1500000);
%
=====
=====

% ===== gera SVMs especialistas para cada uma
% das classes, definidas pela quantidade de
% caminhões que existirão em cada classe
for i=1:N_IMP_intervalo
    ModeloSVM_IMP{i} =
    svmtrain(ModeloIMPdadoIn(:,:),'ModeloIMPTreinoOu
    t(i,:),'Kernel_Function',kernel_function,'polyor
    der',polyorder,'rbf_sigma',rbf_sigma,'Method',me
    todo,'options',options);
end

for i=1:N_EXP_intervalo
    ModeloSVM_EXP{i} =
    svmtrain(ModeloEXPdadoIn(:,:),'ModeloEXPTreinoOu
    t(i,:),'Kernel_Function',kernel_function,'polyor
    der',polyorder,'rbf_sigma',rbf_sigma,'Method',me
    todo,'options',options);
end

%% Classificar novos dados

```

```

display ('Classificando novos dados com o
SVM...')

% classificação dos dados utilizando cada um dos
SVMs especialistas gerados
for i=1:N_IMP_intervalo
    for j=1:nIMP
        classificadoIMP(i,j) =
svmclassify(ModeloSVM_IMP{i},ModeloIMPDadosIn(:,j)
');
        classificadoIMPfinal(i,j) =
classificadoIMP(i,j)*i*intervalo;
    end
end

for i=1:N_EXP_intervalo
    for j=1:nEXP
        classificadoEXP(i,j) =
svmclassify(ModeloSVM_EXP{i},ModeloEXPDadosIn(:,j)
');
        classificadoEXPfinal(i,j) =
classificadoEXP(i,j)*i*intervalo;
    end
end

classificadoIMPfinal(end+1,:) =
max(classificadoIMPfinal);
classificadoEXPfinal(end+1,:) =
max(classificadoEXPfinal);

%% Verificacao dos resultados
% verificação inicial dos resultados. A planilha
do excel que criei
% quantifica melhor os resultados

% acertosIMP = 0;
% errosIMP = 0;
%
% for i = 1:N_IMP_intervalo

```



```
display(['Finalizado às '  
datestr(now, 'HH:MM:SS')])  
display('Executado com sucesso!')  
display('===== Fim =====')
```