

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E
AMBIENTAL**

**Desenvolvimento de um Modelo de Previsão da Demanda
de Passageiros do Transporte Rodoviário Interestadual
Utilizando Regressão com Efeitos Espaciais Locais**

ARLEY COUTO DE MENDONÇA

ORIENTADORA: YAEKO YAMASHITA, PhD

DISSERTAÇÃO DE MESTRADO EM TRANSPORTES

PUBLICAÇÃO: T.DM-014A/2008

BRASÍLIA/DF: SETEMBRO/2008

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E
AMBIENTAL

**Desenvolvimento de um Modelo de Previsão da Demanda
de Passageiros do Transporte Rodoviário Interestadual
Utilizando Regressão com Efeitos Espaciais Locais**

ARLEY COUTO DE MENDONÇA

DISSERTAÇÃO DE MESTRADO SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE.

APROVADA POR:

YAEKO YAMASHITA, PhD (UnB)
(ORIENTADORA)

PASTOR WILLY GONZALES TACO, Dr (UnB)
(EXAMINADOR INTERNA)

ANTONIO NELSON RODRIGUES DA SILVA, Dr (EESC/USP)
(EXAMINADOR EXTERNO)

Data: Brasília/DF, 08 de setembro de 2008

FICHA CATALOGRÁFICA

MENDONÇA, ARLEY COUTO DE

Desenvolvimento de um Modelo de Previsão da Demanda de Passageiros do Transporte Rodoviário Interestadual Utilizando Regressão com Efeitos Espaciais Locais, 2008.

xiii, 123p., 210X297mm (ENC/FT/UnB, Mestre, Transportes, 2008).

Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia. Departamento de Engenharia Civil e Ambiental.

1 – Transporte Rodoviário Interestadual de Passageiros

2 – Demanda de Passageiros 3 – Análise Espacial

4 – Regressão com Efeitos Espaciais Locais

I – ENC/FT/UnB

II – Título (série)

REFERÊNCIA BIBLIOGRÁFICA

MENDONÇA, A. C. de. Desenvolvimento de um Modelo de Previsão da Demanda de Passageiros do Transporte Rodoviário Interestadual Utilizando Regressão com Efeitos Espaciais Locais, 2008. Dissertação de Mestrado, Publicação T.DM-014A/2008, Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, DF, 2008.

CESSÃO DE DIREITOS

NOME DO AUTOR: Arley Couto de Mendonça

TÍTULO DA DISSERTAÇÃO DE MESTRADO: Desenvolvimento de um Modelo de Previsão da Demanda de Passageiros do Transporte Rodoviário Interestadual Utilizando Regressão com Efeitos Espaciais Locais.

GRAU/ANO: Mestre/2008

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito do autor.

Arley Couto de Mendonça
QE 28 Conjunto S casa 06 Guará II1
Brasília, DF – Brasil – CEP: 71060-192

Dedicatória

*A Deus, pelo o dom da vida, à meus pais, José e Marina e
minha avó, Maria, por todas as lições de vida e
à meus irmãos, Uiran e Uiara, pela paciência.*

Agradecimentos

Diversas pessoas participaram, direta e indiretamente, na conclusão desta nova etapa da minha vida. Relacionar todas é muito improvável, mas espero conseguir listar aquelas que lembro no momento e, caso esqueça alguém, que a pessoa compreenda com caridade, os vários fatores que influenciam a memória de um ser humano.

À professora Yaeko, por toda a orientação, pela paciência com este ser enrolado e por arrumar tempo, quando não tinha, para discussões sobre o desenvolvimento da pesquisa e claro, para dizer que tinha que deixar um pouco a estatística de lado e pensar mais como *transporteiro*.

Ao amigo e quase co-orientador Alan que, além do incentivo para o meu ingresso neste mestrado em transportes, auxiliou-me nas diversas análises de regressão com efeitos espaciais.

Aos amigos de mestrado e CEFTRU que me estimularam a reconhecer minhas virtudes e por dois anos toleraram minhas histórias, me viram chorar e principalmente, sorrir quase todos os dias: Alan, Luciany, Ângela, Miguel, Fernando, Rodriguinho, Gizelle Netto, Cristiano, André Paraíba, Raphael, Francesca, Lillian, Venina, Jocilene, Elis, Edcélio, Denise, Leandro Sertanejo, Lara, Carlos Eduardo, Andréia, Thania, Leonardo, Kátia, Sandra, Arthur, Sandro, Érika, Evandro, Heitor, Marne, Giovanna, Michelle, Luiz Gustavo, Bruninha, Ernesto, Daniel, Marcelo, Ricardo, Alexandre.

Aos amigos *Equipistas* que, por meio dos ensinamentos adquiridos em comunidades cristãs, me ajudaram a ser mais humano e pacientemente caminharam sempre ao meu lado. Como não é possível listar todos, vou citar da seguinte forma: aos exs e atuais membros das EJNS das Graças, a todos os membros de Brasília, Rio de Janeiro, São Paulo e Pernambuco e claro, aos amigos que conquistei fora do país.

A todos os professores pelos ensinamentos na área de transporte, principalmente ao Professor Pastor, pela disponibilidade de 24 horas por dia para todos os alunos e, aos funcionários em geral do PPGT, principalmente ao Júlio, pela atenção com os detalhes do departamento e pelas ajudas necessárias.

A todos os membros da minha família que sempre me apoiaram e auxiliaram com os meus estudos.

E ao CNPq, pelo apoio financeiro de dois anos, fundamental para a conclusão desse trabalho...

RESUMO

MODELO DE PREVISÃO DA DEMANDA DE PASSAGEIROS DO TRANSPORTE RODOVIÁRIO INTERESTADUAL UTILIZANDO REGRESSÃO COM EFEITOS ESPACIAIS LOCAIS

O transporte rodoviário interestadual de passageiros (TRIP) por ônibus é o meio de transporte predominante no Brasil, representando aproximadamente 95% dos deslocamentos de pessoas e um faturamento anual de mais de R\$ 2,5 bilhões, sendo um serviço essencial para economia do país. Desse modo, existe a necessidade de estudos específicos para adquirir informações operacionais, além de identificar fatores que influenciam na produção e na atração de passageiros do transporte rodoviário interestadual. Assim, estudos da demanda de passageiros são de fundamental importância na tomada de decisão. Devido à relação entre transporte e espaço (geográfico, distância, dinâmica regional, etc), as técnicas que podem ser as mais adequadas para a previsão da demanda de passageiros, são as que utilizam Análise Espacial (AE). Ou seja, modelos que não possuem alguma estrutura espacial, como por exemplo, a matriz de proximidade espacial W , não são os mais adequados para o trabalho. Dessa forma, pela existência de dados espaciais, provavelmente verifica-se a presença da dependência espacial, e assim, o modelo proposto é de regressão com efeitos espaciais. Mas se além dessa dependência, forem identificadas sub-regiões com os mesmos padrões espaciais (não-estacionariedade), é necessária a utilização de modelos de regressão com efeitos espaciais locais que podem ser trabalhados de duas formas: discreto e contínuo. O presente trabalho ostenta além de uma metodologia para a construção dos modelos de regressão com efeitos espaciais, uma análise exploratória e confirmatória por meio de um estudo de caso que engloba passageiros de municípios como São Paulo e Natal em viagens interestaduais e passageiros em viagens intermunicipais do estado do Ceará, além das variáveis socioeconômicas de ambos os lugares. Os resultados mostraram que, na existência de dependência espacial e da não-estacionariedade, os modelos de regressão com efeitos espaciais locais da forma discreta, que analisa cada sub-região com suas particularidades, apresentaram diagnósticos que permitem afirmar e compreender melhor de que o espaço com suas atividades e dinâmicas, influencia na demanda de passageiros do transporte rodoviário interestadual.

ABSTRACT

DEMAND FORECAST MODEL FOR PASSENGERS TRAVELS IN INTERSTATE TRANSPORTATION BUS SERVICES USING SPATIAL REGRESSION WITH LOCAL EFFECTS

Interstate transportation bus services (TRIP) are the principal transport mode used in Brazil, responsible for more than 95% of personal displacements, and an annual turnover of over R\$ 2.5 billion. In order to enable better decision making, it is important to develop studies focused on operational information, identifying demand performance. Due to the Space and transport relationship (geographic, distance and regional dynamics), only spatial structured models like spatial weights matrix (W) should be used for forecasting passenger demand. Dealing with Spatial data, it is expected to observe spatial dependence, represented by regression models with spatial effects. But beyond this dependence, it could be observed sub-regions with the same spatial patterns (non-stationary), where regression models with local spatial effects are properly recommended in both cases: discrete and continuous. This paper presents, beyond a suitable methodology for developing regression models with spatial effects, an exploratory and confirmatory analysis, considering socioeconomics variables, by a study case including interstate passenger travels in Brazil, from cities of São Paulo and Natal, and intercity passenger travels in Ceará State. The results have shown that, existing spatial dependence and non-stationarity, regression models with local spatial effects in discrete way can allow diagnoses that promote a better understanding of urban dynamic space, influencing transportation passenger demand on interstate travels, due to the evaluation of each sub-region in its merits given by the technique.

SUMÁRIO

1. INTRODUÇÃO	1
1.1. APRESENTAÇÃO	1
1.2. IDENTIFICAÇÃO DO PROBLEMA	2
1.3. JUSTIFICATIVA	3
1.4. HIPÓTESE	4
1.5. OBJETIVO	5
1.6. METODOLOGIA DA PESQUISA	5
1.7. ESTRUTURA METODOLÓGICA DA DISSERTAÇÃO	5
2. TRANSPORTE RODOVIÁRIO INTERESTADUAL DE PASSAGEIROS NO BRASIL	8
2.1. APRESENTAÇÃO	8
2.2. DEFINIÇÃO E CARACTERÍSTICAS	8
2.2.1.Histórico do Transporte Rodoviário de Passageiros	9
2.2.2.Importância do Transporte Rodoviário Interestadual de Passageiros.....	11
2.2.3.Regulação do Transporte Rodoviário Interestadual de Passageiros	13
2.2.4.Funcionamento do Transporte Rodoviário Interestadual de Passageiros	14
2.3.MODELOS DE DEMANDA DE TRANSPORTE	17
2.3.1.Demanda em Transporte	18
2.3.2.Tipos de Modelos de Demanda.....	20
2.3.2.1.Modelos Convencionais.....	20
2.3.2.1.1.Técnicas de Geração de Viagens	21
2.3.2.1.2.Técnicas de Distribuição de Viagens.....	22
2.3.2.2.Modelos Desagregados	24
2.3.2.2.1.Modelos Comportamentais.....	25
2.3.2.2.2.Modelos Atitudinais.....	26
2.3.2.1.Modelos com Base em Novas Tecnologias.....	28
2.3.3.Modelo de demanda utilizado pela ANTT	29
2.4.TÓPICOS CONCLUSIVOS	32
3. REGRESSÃO COM EFEITOS ESPACIAIS	34
3.1.APRESENTAÇÃO	34
3.2.CONCEITOS DE ANÁLISE ESPACIAL	35
3.3.ESTATÍSTICA ESPACIAL	37
3.3.1.Matriz de Proximidade Espacial (W)	38
3.3.2.Dependência Espacial	41
3.3.2.1.Índice Global I de Moran.....	42
3.3.2.2.Índice Local I_i de Moran	44
3.3.2.3.Análise Gráfica da Dependência Espacial.....	45

3.4.MODELOS DE REGRESSÃO COM EFEITOS ESPACIAS.....	49
3.4.1.Regressão com Efeitos Espaciais Globais.....	50
3.4.2.Regressão com Efeitos Espaciais Locais.....	52
3.5. TÓPICOS CONCLUSIVOS.....	56
4. METODOLOGIA PARA CONSTRUÇÃO DE MODELOS DE REGRESSÃO COM EFEITOS ESPACIAIS LOCAIS	59
4.1.APRESENTAÇÃO	59
4.2.ESTRUTURA DO MÉTODO PROPOSTO	59
4.3.ETAPAS DO MÉTODO PROPOSTO	62
Etapa 1: Descrição e definição da área de estudo.....	62
Etapa 2: Identificação das variáveis	62
Etapa 3: Análise da qualidade do banco de dados disponível	64
Etapa 4: Seleção das variáveis independentes	65
Etapa 5: Verificação da dependência espacial.....	66
Etapa 6: Construção do modelo de regressão com efeitos espaciais locais	67
4.4.TÓPICOS CONCLUSIVOS.....	69
5. ESTUDO DE CASO	70
5.1.APRESENTAÇÃO	70
5.2.APLICAÇÃO DA METODOLOGIA.....	70
Etapa 1: Descrição e definição da área de estudo.....	70
Etapa 2: Identificação das variáveis	77
Etapa 3: Análise da qualidade do banco de dados disponível	79
Etapa 4: Seleção das variáveis independentes	80
Etapa 5: Verificação da dependência espacial.....	83
Etapa 6: Construção do modelo de regressão com efeitos espaciais locais	97
5.3. ANÁLISE DA ADEQUABILIDADE DO MODELO E INTERPRETAÇÃO DOS PARÂMETROS.....	105
5.4.TÓPICOS CONCLUSIVOS.....	112
6. CONCLUSÕES E RECOMENDAÇÕES	113
6.1. APRESENTAÇÃO	113
6.2. AVALIAÇÃO DA CONTRIBUIÇÃO DO MODELO PROPOSTO.....	113
6.3. AVALIAÇÃO DAS FRAGILIDADES DO MODELO PROPOSTO.....	115
6.3. RECOMENDAÇÕES PARA TRABALHOS FUTUROS	116
BIBLIOGRAFIA	118

LISTA DE TABELAS

Tabela 2.1. Participação das 20 maiores empresas por passageiros/km transportados em 2006	12
Tabela 2.2. Tipos de modelos de previsão da demanda.....	20
Tabela 2.3. Evolução dos Métodos de Fator de Crescimento.....	23
Tabela 2.4. Evolução dos Modelos Gravitacionais	30
Tabela 3.1. Tipos de modelagem de regressão com efeitos espaciais locais.....	53
Tabela 5.1. Quantidade de passageiros transportados em produzidas por Região e UF	72
Tabela 5.2. Os sete maiores emissores de passageiros para o Ceará (Out./2004)	74
Tabela 5.3. Análise descritiva das possíveis variáveis independentes.	80
Tabela 5.4. Matriz de correlação das variáveis.....	82
Tabela 5.5. Indicadores de autocorrelação global para a variável <i>Pass</i>	83
Tabela 5.6. Indicadores de autocorrelação global para a variável <i>Pass_{munic}</i>	90
Tabela 5.7. Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre São Paulo a Crato.....	98
Tabela 5.8. Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre São Paulo a Crateús.	99
Tabela 5.9. Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre São Paulo a Canindé.	100
Tabela 5.10. Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre Natal e Juazeiro do Norte.	102
Tabela 5.11. Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre Natal e Aracati	103
Tabela 5.12. Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre Natal e Região Metropolitana de Fortaleza	104
Tabela 5.13. Números estimados, observados e erros relativos para a quantidade de passageiros de cada O/D no ano de 2005	106
Tabela 5.14. Números estimados, observados, erros relativos e erros médios para a quantidade de passageiros de cada O/D, com a variável <i>renda</i> estimada para 2005	106
Tabela 5.15. Relação da quantidade de passageiros (2004-2005).....	107
Tabela 5.16. Resultados do modelo de regressão espacial com efeitos locais da forma discreta para passageiros em viagens até Crato.....	107
Tabela 5.17. Números estimados, observados, erros relativos e erros médios para a quantidade de passageiros de cada O/D para 2005.....	109

LISTA DE FIGURAS

Figura 2.1: Ônibus do final do século XIX e início do século XX.....	9
Figura 2.2: Prefixo da linha base Rio de Janeiro (RJ) – Brasília (DF)	15
Figura 2.3: Relação modelo-realidade	18
Figura 3.1: Inter-relação entre os quatro grupos de ferramentas de análise espacial.....	36
Figura 3.2: Matriz de proximidade espacial.....	39
Figura 3.3: Representação de autocorrelação positiva (a) e (b) e negativa (c) e (d).....	42
Figura 3.4: Gráfico de espalhamento de Moran.....	45
Figura 3.5: Exemplo de <i>Box Map</i>	47
Figura 3.6: Exemplo de <i>Lisa Map</i>	48
Figura 3.6: Exemplo de <i>Moran Map</i>	49
Figura 4.1: Estrutura metodológica para o cálculo da quantidade de passageiros do TRIP utilizando regressão com efeitos espaciais locais.....	61
Figura 5.1: Quantidade de passageiros transportados em viagens produzidas por município (Brasil, Out./2004).....	73
Figura 5.2: Quantidade de passageiros transportados em viagens produzidas por estado (Brasil, Out./2004).....	73
Figura 5.3: Quantidade de Passageiros ($Pass_{ij}$) que viajaram do município de São Paulo/SP aos municípios do Ceará em Outubro de 2004.....	75
Figura 5.4: Quantidade de Passageiros ($Pass_{ij}$) que viajaram do município de São Paulo/SP aos municípios do Ceará em Dezembro de 2004.....	75
Figura 5.5: Quantidade de Passageiros ($Pass_{munic}$) em viagens intermunicipais no Estado do Ceará em Outubro de 2004.....	76
Figura 5.6: <i>Box Map</i> da quantidade de passageiros utilizando a matriz W de contigüidade (a) e das distâncias rodoviárias (b).	84
Figura 5.7: <i>Moran Map</i> da quantidade de passageiros utilizando a matriz W de contigüidade (a) e das distâncias (b).....	87
Figura 5.8: <i>Moran Scatterplot</i> da quantidade de passageiros utilizando a matriz W de contigüidade (a) e das distâncias (b), considerando a Região Metropolitana de Fortaleza (RMF)....	89
Figura 5.9: <i>Moran Scatterplot</i> da quantidade de passageiros utilizando a matriz W de contigüidade (a) e das distâncias (b), sem considerar a Região Metropolitana de Fortaleza (RMF)....	89
Figura 5.10: <i>Moran Map</i> da variável $Pass_{munic}$ no município de Crato – Matriz W contigüidade.....	92
Figura 5.11: <i>Moran Map</i> da variável $Pass_{munic}$ no município de Juazeiro do Norte – Matriz W de contigüidade.....	92
Figura 5.12: <i>Moran Map</i> da variável $Pass_{munic}$ no município de Canindé – Matriz W contigüidade.....	93
Figura 5.13: <i>Moran Map</i> da variável $Pass_{munic}$ no município de Aracati – Matriz W contigüidade.....	93
Figura 5.14: <i>Moran Map</i> da variável $Pass_{munic}$ no município de Crateús – Matriz W contigüidade.....	93
Figura 5.15: <i>Moran Map</i> da variável $Pass_{munic}$ na RMF – Matriz W de contigüidade.....	93
Figura 5.16: <i>Moran Map</i> da variável $Pass_{munic}$ no município de Crato – Matriz W das distâncias rodoviárias.	94
Figura 5.17: <i>Moran Map</i> da variável $Pass_{munic}$ no município de Juazeiro do Norte – Matriz W das distâncias rodoviárias	94

Figura 5.18: <i>Moran Map</i> da variável $Pass_{munic}$ no município de Canindé – Matriz W das distâncias rodoviárias	94
Figura 5.19: <i>Moran Map</i> da variável $Pass_{munic}$ no município de Aracati – Matriz W das distâncias rodoviárias.	94
Figura 5.20: <i>Moran Map</i> da variável $Pass_{munic}$ no município de Crateús – Matriz W das distâncias rodoviárias	95
Figura 5.21: <i>Moran Map</i> da variável $Pass_{munic}$ no município de Crateús – Matriz W das distâncias rodoviárias	95

1. INTRODUÇÃO

1.1. APRESENTAÇÃO

No início do século passado, muitas cidades brasileiras e povoados se encontravam isolados entre si, pois não existiam caminhos ou estradas que os ligavam, já que os transportes por estradas de ferro e por navegação a vapor estavam subordinados à economia agro-exportadora (Brasileiro *et al.*, 2001). Devido a esse fator, juntamente com a chegada dos primeiros veículos automotores no País, o Governo Federal decidiu incentivar a construção de rodovias no Brasil.

Em 1926, o presidente Washington Luís, que tinha a filosofia de que “governar é abrir estradas”, iniciou o trabalho de crescimento da malha rodoviária do País, que, segundo Brasileiro *et al.* (2001), passou de plano à diretriz de política pública em âmbito federal.

Dessa forma, inicia-se o deslocamento de pessoas de uma cidade à outra, para fins de trabalho, turismo, entre outras atividades. Com isso, o Transporte Rodoviário de Passageiros foi surgindo naturalmente e por iniciativa própria de poucos que tinham veículos importados ou adaptavam caminhões para transportar pessoas, de forma que as viagens costumavam ser curtas, devido às condições precárias das estradas (Wright, 1992).

Nessa perspectiva, com o aumento de usuários e o crescimento da malha rodoviária, o transporte rodoviário de passageiros passou a ser uma atividade altamente significativa na estrutura econômica e social do Brasil, o que se perpetua na atual década (Fernandes, 2003).

No Brasil, o ônibus é a principal modalidade de transporte de passageiros, composto por transporte urbano (dentro do município), intermunicipal (dentro do estado), interestadual (de um estado para outro) e internacional (de um estado do Brasil para outro País). Segundo a Agência Nacional de Transportes Terrestres (ANTT, 2007), o Transporte Rodoviário Interestadual de Passageiros (TRIP), juntamente com o internacional, “é um serviço público essencial, responsável por uma movimentação superior a 140 milhões de usuários/ano”.

O TRIP tem importância significativa no País, pois é responsável por quase 95% dos deslocamentos das pessoas, além da expressiva participação na economia brasileira. Atualmente, o TRIP assume um faturamento anual de mais de R\$ 2,5 bilhões na prestação de serviços regulares pelas empresas permissionárias, com aproximadamente 13.400 ônibus.

Para adquirir essas informações operacionais são necessários estudos relacionados à demanda do TRIP. Segundo Gonçalves *et al* (2006), o conhecimento da demanda, através do TRIP por ônibus no Brasil, é uma importante contribuição ao planejamento e à regulação desse serviço. Para estimar essa demanda de passageiros, o modelo que vem sendo utilizado no Brasil é o Modelo Gravitacional (MG) que envolve variáveis como população e distância rodoviária entre os pontos de origem e destino.

1.2. IDENTIFICAÇÃO DO PROBLEMA

A distribuição espacial da demanda geralmente produz problemas de coordenação que afetam o equilíbrio entre oferta e demanda (Ortúzar, 2000), como, por exemplo, ônibus realizando viagens com poucos passageiros entre cidades com mais de 100 mil habitantes, enquanto existem passageiros em potencial esperando em outras cidades com menos de 50 mil habitantes.

Os modelos gravitacionais tendem a não considerar a influência de variáveis espaciais e temporais, o que pode representar uma previsão de demanda de passageiros não significativa (Carneiro, 2005). Logo, esses modelos utilizados no Brasil apresentam erros quando comparados aos valores reais de operação dos serviços, podendo acarretar em sub ou superestimação da demanda real de passageiros (Aguiar Junior *apud* Aguiar Junior, 2006). Além de os modelos brasileiros serem anuais e não mensais, o que daria uma estimativa mais sensata, já que existem meses com mais passageiros do que outros.

Dessa forma, técnicas que se destacavam em diversas áreas, como Redes Neurais Artificiais, Minerador de Dados e Sistemas de Informação Geográfica, começaram a ser desenvolvidas na área de transporte. Nos últimos anos, essas técnicas foram utilizadas para estimar a demanda por transportes, e obtiveram resultados um pouco mais satisfatórios do que os obtidos por meio de modelos gravitacionais.

Nesse contexto, a fim de identificar padrões de demanda do TRIP, uma melhor estimativa para esta e de contribuir para o seu planejamento e regulação, o presente trabalho buscará responder a seguinte pergunta: “Como representar a quantidade de passageiros nos deslocamentos entre os pares de Origem e Destino (O/D) do transporte rodoviário interestadual em países de grande extensão e diversidades socioeconômicas, de forma a auxiliar e melhor entender o planejamento desse serviço?”.

1.3. JUSTIFICATIVA

Os deslocamentos realizados por pessoas são, em função das vontades e necessidades de sobrevivência, representados pelas diversas atividades realizadas, sendo que essas dependem das condições socioeconômicas e das disposições espaciais. Os deslocamentos são definidos por um par de origem e destino, em que essa relação pode ser representada pela dependência espacial entre eles, que é o objeto de análise deste estudo.

A análise espacial aplicada aos transportes é um fato recente (especialmente a regressão espacial), mas que aos poucos vem sendo difundido entre os profissionais ligados à área, sobretudo aqueles que necessitam de ferramentas que auxiliem nas tomadas de decisão. Henrique (2004) afirma que a maioria dos trabalhos envolvendo análise espacial aplicada aos transportes se limita às ferramentas de seleção e manipulação de dados espaciais, sendo que as ferramentas de análise exploratória e confirmatória ainda estão em um estágio inicial de suas aplicações.

Araújo (2005) utilizou estas ferramentas de análise para identificar a dependência espacial em variáveis que posteriormente fariam parte de modelos com efeitos espaciais. Esta etapa é essencial para a decisão em se utilizar modelos convencionais ou espaciais, em que a diferença básica entre eles é que os espaciais têm a presença de uma estrutura (espacial) diferenciada no fenômeno estudado. Assim, Araújo (2005), ao fim de seu trabalho, constatou a existência dessa dependência espacial, sugerindo a utilização de modelos de regressão com efeitos espaciais para a determinação da geração de viagens para o transporte público intermunicipal de passageiros no estado do Ceará.

Outro exemplo é o trabalho de Silva (2006), que desenvolveu um modelo de regressão com efeitos espaciais globais para a previsão de demanda no transporte rodoviário de carga.

Neste se concluiu que modelos de regressão com efeitos espaciais produzem melhores ajustes que os modelos de regressão convencional. Isto por incorporarem a estrutura espacial no modelo, sendo que essa estrutura, nesse trabalho, foi do tipo contigüidade e do tipo distância euclidiana entre os centróides de cada município.

Outro método, bem menos utilizado no Brasil e de fundamental importância, é o de regressão com efeitos espaciais locais. Existem dois modos de se trabalhar com esse tipo de modelagem: discreto e contínuo. Loureiro *et al.* (2006) desenvolveram uma metodologia de análise da regressão, com efeitos espaciais locais contínuos (regressão geograficamente ponderada), aplicada ao fenômeno de viagens intermunicipais, em que os objetivos principal e secundário, respectivamente, foram: “tentativa de incorporar de uma forma mais eficiente e eficaz as técnicas quantitativas e qualitativas nos processos de análise dos sistemas de transportes, dado o caráter explicitamente espacial destas análises” e “incentivar o estudo e a utilização mais intensa das ferramentas de análise exploratória e confirmatória na análise, planejamento e operação em transporte”.

Assim, a contribuição deste trabalho está em buscar um melhor entendimento do processo espacial que ocasiona a demanda de passageiros no transporte rodoviário interestadual nos diversos pares de O/D. Para isso, buscará aplicar ferramentas de estatística espacial, além de desenvolver os dois métodos de regressão com efeitos espaciais locais.

1.4. HIPÓTESE

A hipótese considerada nesta pesquisa é que municípios que apresentam uma relação espacial têm maiores níveis de deslocamento entre eles. Logo, uma modelagem da demanda de passageiros do transporte rodoviário interestadual considerando os efeitos espaciais, por meio de informações socioeconômicas dos municípios e da quantidade de passageiros nas viagens intermunicipais, representa, de forma mais realística, esse fenômeno de transporte.

As hipóteses secundárias são:

- Modelos de regressão com efeitos espaciais locais (REL) têm resultados melhores em relação à previsão da demanda do que modelos de regressão com efeitos espaciais globais (REG), devido à particularidade de cada localidade.

- Representa-se, na matriz de proximidade espacial ou de vizinhança (W), o grau de associação entre diferentes áreas, ou seja, o quanto o município X está associado ao município Y em relação à demanda de passageiros. Logo, quando se trabalha na área de transporte, a matriz composta pela distância tem melhores resultados do que uma matriz W de contigüidade (1 fronteira e 0 não fronteira).

1.5. OBJETIVO

O objetivo geral deste trabalho é o desenvolvimento de um modelo de demanda para o transporte rodoviário interestadual de passageiros para um par de origem e destino, considerando os efeitos espaciais – por meio de regressão com efeitos espaciais locais – prevendo de forma mais realística a quantidade de passageiros, a fim de subsidiar o planejamento estratégico desse serviço.

A aplicação do modelo será em dois estágios: (i) os passageiros apenas em viagens interestaduais e (ii) utilizando as viagens intermunicipais para chegar ao seu destino final.

Como objetivo específico tem-se:

- Comparar os resultados encontrados utilizando as diferentes estruturas para a matriz de vizinhança (distância via malha rodoviária e contigüidade).

1.6. METODOLOGIA DA PESQUISA

Esta pesquisa tem sua origem em um problema e, por meio de algumas tentativas e eliminação de erros, procura-se, ao longo de seu desenvolvimento, traçar uma possível solução. Com isso, o método de abordagem adotado neste trabalho é o *hipotético dedutivo*, que também apresenta uma melhor sistemática àqueles que nunca trabalharam com estudos científicos.

O método de procedimento utilizado neste trabalho é o que se fundamenta em teorias estatísticas (*método estatístico*) e na técnica de pesquisa descritiva, visto que se observa a realidade (os modelos de regressão na previsão de viagens) sem manipulá-las.

1.7. ESTRUTURA METODOLÓGICA DA DISSERTAÇÃO

Visando comprovar a hipótese considerada neste trabalho, assim como atingir o objetivo proposto, o estudo foi desenvolvido em duas diferentes etapas:

- **Revisão Bibliográfica:**

Nesta etapa buscou-se contextualizar o transporte rodoviário interestadual de passageiros em relação ao histórico, a importância, a regulação e o funcionamento desse serviço. Também apresentará alguns conceitos e tipos de modelos de demanda em transporte. Em seguida, conceitua a análise espacial e a estatística espacial. Complementando essa revisão bibliográfica, apresentam-se os principais modelos de regressão com efeitos espaciais.

- **Estudo de caso, conclusões e recomendações:**

Neste tópico desenvolvem-se os modelos propostos para o cálculo da demanda para o transporte rodoviário interestadual de passageiros, por meio do banco de dados fornecido pela ANTT com informação do número de passageiros que utilizam esse serviço, empregando técnicas de estatística espacial. Em seguida, finaliza-se a dissertação com as conclusões e as recomendações para futuros trabalhos.

O presente trabalho foi dividido em 6 capítulos, sendo que, o primeiro capítulo introduz o assunto a ser trabalhado, além da estrutura da dissertação. O segundo capítulo tem o objetivo de apresentar a evolução do TRIP, do primeiro ônibus que surgiu na França há mais de 200 anos até os dias de hoje, mostrando sua importância para o Brasil. Nesse mesmo capítulo são apresentadas a legislação e a forma como é regulado o TRIP, assim como alguns conceitos que envolvem o funcionamento desse serviço. Destaca-se também nesse capítulo os modelos de demanda em transporte, que são representações simplificadas de um fenômeno real, neste caso, transporte. O terceiro capítulo introduz a análise e a estatística espacial, explicando os conceitos de matriz de proximidade ou de vizinhança e dependência espacial, sendo finalizado com os principais modelos de regressão com efeitos espaciais globais e principalmente locais. O quarto capítulo apresenta a metodologia utilizada para a construção do modelo de regressão com efeitos espaciais locais. O quinto capítulo é a aplicação desta metodologia no estudo de caso, que foi elaborado com base no banco de dados fornecido pela ANTT e pelo Plano Diretor e Operacional do Transporte

Intermunicipal de Passageiros do Estado do Ceará (PDOTIP-CE). Por fim, o sexto capítulo contém as conclusões e as recomendações para trabalhos futuros.

2. TRANSPORTE RODOVIÁRIO INTERESTADUAL DE PASSAGEIROS NO BRASIL

2.1. APRESENTAÇÃO

O meio de transporte predominante em território nacional para viagens interestaduais é o transporte rodoviário por ônibus, o qual é responsável por quase 95% dos deslocamentos de pessoas no Brasil. Segundo Wright (1992), o principal concorrente era a aviação que, apesar das reduções tarifárias, ainda é uma modalidade financeiramente inacessível à grande maioria da população, ressalvando épocas do ano em que as empresas colocam tarifas mais baixas que o serviço de ônibus.

Nos últimos anos, as principais empresas de TRIP vêm se preocupando com diversos fatores para não perder passageiros para o transporte aéreo. Alguns desses fatores são a qualidade e a qualificação da mão-de-obra; a segurança em viagens (direção defensiva e controle de assalto em ônibus); os cuidados na condição de saúde, psicológicas e familiares do motorista; as condições de descanso e alimentação; além de tecnologias avançadas, conforto e rapidez (Revista ABRATI, 2000). Apesar desses avanços, não é possível afirmar com precisão, quantos passageiros o transporte rodoviário interestadual de passageiros (TRIP) perdeu, ou até mesmo ganhou, nos últimos anos, pois estudos em que o alvo é o comportamento do passageiro desse tipo de transporte são complicados de se executar.

O presente capítulo apresentará algumas estatísticas dos passageiros que utilizam esse tipo de serviço, ressaltando a importância do TRIP, assim como o histórico, o funcionamento, as características e os conceitos. Ainda neste capítulo serão apresentados os principais modelos de demanda de transporte, assim como o modelo utilizado pela ANTT, para obter um conhecimento maior sobre os métodos utilizados para o cálculo da demanda de passageiros.

2.2. DEFINIÇÃO E CARACTERÍSTICAS

Segundo a ANTT (2007a), o transporte rodoviário interestadual de passageiros é um serviço público essencial, que transpõe os limites de Estado, de Território ou do Distrito Federal e é responsável por uma ampla parcela de deslocamento de pessoas, por diversos motivos. O TRIP é o principal modo de deslocamento coletivo de pessoas no âmbito interestadual. Antes de apresentar alguns dados que comprovem essa afirmativa, é importante descrever a história e a evolução do ônibus e do transporte rodoviário de passageiros no Brasil.

2.2.1. Histórico do Transporte Rodoviário de Passageiros no Brasil

O primeiro ônibus surgiu a mais de 200 anos atrás, na França, onde as diligências faziam o transporte de até 24 pessoas, puxadas por seis cavalos. No final do século XIX, iniciou-se a produção do primeiro modelo em série de um carro sobre três rodas, impulsionado por um motor de dois tempos a gasolina e com uma carruagem motorizada (Revista ABRATI, 2000).

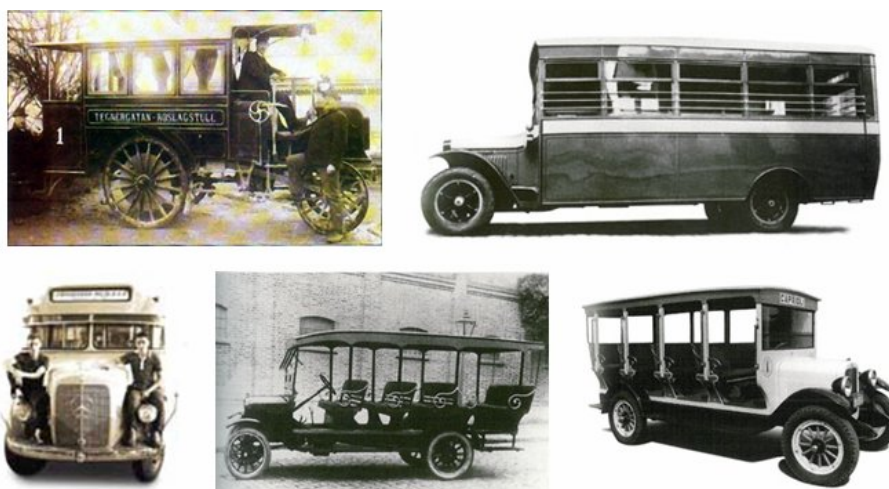


Figura 2.1: Ônibus do final do século XIX e início do século XX

Mas apenas em 1911 foi construída a primeira carroceria motorizada no Brasil. Provavelmente em consequência do incentivo do Governo Federal à construção de rodovias que ocorreu em outubro de 1910 com o decreto nº. 8.324, o qual previa “subvenções a serem pagas pelo governo por quilômetro de rodovia construída dentro das especificações de projeto e execução dele constantes que pretendiam introduzir no País padrões mínimos de tecnologia rodoviária” (Brasileiro *et al.*, 2001).

Tiveram vários fatores, no início do século XX, que impulsionaram o crescimento do transporte rodoviário no País, principalmente de passageiros. A seguir, um resumo da evolução do transporte rodoviário de passageiros no Brasil do começo da república até os dias de hoje:

▪ **Primeira fase republicana (1889–1930):**

Economia brasileira dependente do comércio exterior. Época em que poucos grupos privados e empresas estrangeiras exploravam estradas de ferro e bondes urbanos. Em 1919, a Ford foi a primeira empresa a conseguir uma autorização do governo brasileiro para instalar uma montadora no País (Brasileiro *et al.*, 2001). Mas apenas no final da década de 20, surge a primeira empresa registrada no setor de transporte rodoviário de passageiros (Auto Viação Catarinense, antiga Empresa Catharinense), fundada em Blumenau (Gômara *apud* Fernandes, 2003).

▪ **Segunda fase republicana (1930-1960):**

Os bens que anteriormente eram importados passaram a ser manufaturados no Brasil. Com isso, houve a necessidade de uma rede de transportes por rodovias para a distribuição desses no País (Prado, 2006). O número de empresas de transporte de passageiros continuava aumentando, apesar das dificuldades (segundo Brasileiro *et al.* (2001), menos de 0,5% dos 276.700 km de estradas de rodagem eram pavimentadas em 1942). A implantação da indústria automobilística, no final da década de 50, foi um dos fatores responsáveis pelo desenvolvimento do sistema rodoviário no País. Em 1960, o Brasil tinha uma frota de 27.645 ônibus, uma extensão de 32.402 km de rede rodoviária pavimentada e 60,5% de passageiros transportados através das rodovias.

▪ **Anos 1960 a 2000:**

Continua a expansão do sistema de transporte rodoviário de passageiros no País, pois se acreditou que as rodovias possuíam o acesso mais democrático e seriam mais adequadas à ocupação do território nacional (Dantas Filho, 1995). Desse modo, ao longo dos anos, houve uma degradação do sistema ferroviário brasileiro e com isso um aumento do setor rodoviário. Em 1976, o movimento total de passageiros pelo sistema de transporte rodoviário interestadual e internacional era de aproximadamente 19 bilhões de passageiros quilômetros transportados (Wright, 1992). Segundo a ANTT

(2007b), no ano de 2000, o movimento total de passageiros por esse sistema era de 31 bilhões de passageiros quilômetros transportados.

▪ **Anos 2001 aos dias de hoje:**

No dia 5 de junho de 2001 é instituída a Agência Nacional de Transportes Terrestres (ANTT), com o objetivo de regular e fiscalizar as atividades de exploração da infraestrutura de transportes, exercidas por terceiros; a prestação do serviço de transporte rodoviário interestadual e internacional de passageiros; além do transporte nacional e internacional de cargas. Após a criação da Agência, houve melhorias dos serviços e redução de custos aos usuários, mas, mesmo assim, em 2006, o movimento total de passageiros por esse sistema era inferior ao ano de 2000, de aproximadamente 28 bilhões de passageiros quilômetros transportados (ANTT, 2007c).

2.2.2. Importância do Transporte Rodoviário Interestadual de Passageiros no Brasil

A existência de um sistema de transporte rodoviário de passageiros é vital para o País, já que o Brasil possui uma malha rodoviária de aproximadamente 1,8 milhões de quilômetros. Responsável, em 2006, por aproximadamente quatro milhões de viagens realizadas, quase 1,5 bilhões de quilômetros percorridos e mais de 130 milhões de passageiros, o TRIP é o serviço mais importante nessa esfera (ANTT, 2006). Segundo a Revista ABRATI (2005), todas as cidades brasileiras estão interligadas por ônibus, e as responsáveis por esse tipo de serviço são centenas de empresas permissionárias de diversos tipos de tamanho (empresas com frotas de dois ou três veículos e outras com mais de 1.000 ônibus).

O TRIP, juntamente com o transporte rodoviário internacional de passageiros, engloba mais de 220 empresas permissionárias e em torno de 16 mil ônibus em circulação, sendo que 64% dos passageiros quilômetros transportados em 2006 estão nas 20 maiores empresas, como apresentado na Tabela 2.1 (Revista ANTT, 2006). Esses dois serviços juntos geram aproximadamente 80.000 empregos diretos e o quadro de motoristas é de 30.000 profissionais (Revista ABRATI, 2005).

Tabela 2.1. Participação das 20 maiores empresas por passageiros.km transportados em 2006.

Nº.	Empresa	Passageiros.km	Participação (%)
1	Viação Itapemirim S/A	3.485.091.922	12,24
2	Empresa Gontijo de Transportes Ltda.	1.791.683.669	6,30
3	Cia. São Geraldo de Viação	1.708.827.485	6,00
4	Viação Anapolina Ltda.	1.380.863.831	4,85
5	Expresso Guanabara S/A	1.317.170.798	4,63
6	Viação Cometa S/A	806.437.984	2,83
7	Viação Garcia Ltda.	756.616.969	2,66
8	Transbrasiliana – Transportes e Turismo Ltda.	640.243.672	2,25
9	Taguatur – Taguatinga Transporte e Turismo Ltda.	620.985.511	2,18
10	Viação Águia Branca S/A	569.121.965	2,00
11	Pluma Conforto e Turismo S/A	561.185.925	1,97
12	Auto Viação Catarinense Ltda.	560.998.922	1,97
13	Real Expresso Ltda.	539.710.543	1,90
14	Viação Motta Ltda.	538.443.726	1,89
15	Empresa de Ônibus Nossa Senhora da Penha S/A	527.956.577	1,85
16	Reunidas S/A – Transportes Coletivos	527.471.539	1,85
17	Unesul de Transportes Ltda.	513.322.215	1,80
18	Empresa de Transportes Andorinha S/A	487.579.932	1,71
19	Rápido Planaltina Ltda.	484.445.820	1,70
20	Empresa Santo Antônio Transporte e Turismo Ltda.	467.062.802	1,64
	Total das 20 maiores empresas	18.285.221.807	64,25
	Total Geral	28.461.511.025	

Fonte: Empresas Permissionárias – Anuário Estatístico da ANTT (Ano Base 2006).

O transporte rodoviário de passageiros também tem um valor expressivo para a economia brasileira. Segundo Prado (2006), é uma das bases para a integração e o desenvolvimento da economia do País. O faturamento anual do TRIP é superior a R\$ 2,5 bilhões na prestação dos serviços regulares.

Apesar desses números alentadores, o transporte rodoviário de passageiros convive com dificuldades para exercer suas funções, devido a alguns fatores como: a precariedade das estradas e rodovias nacionais; a falta de segurança e fiscalização; situação deficiente de alguns terminais rodoviários e a ação predatória das empresas ilegais que, segundo Aragão *et al.* (2000), estão tirando uma parte ativa do mercado brasileiro de transporte de passageiros. Brasileiro *et al.* (2001) e Prado (2006) mencionam algumas razões que buscam explicar o crescimento do transporte rodoviário interestadual ilegal de passageiros:

- rigidez do regulamento brasileiro de transporte, que impede a oferta de serviço mais flexível;
- taxas crescentes de desemprego que levam parcelas da população a atuar neste mercado, pois custos afundados (que não podem ser recuperados quando uma empresa decide sair do mercado) são relativamente baixos, ocorrendo à presença de contestabilidade do setor;
- agressiva política de *marketing* dos fabricantes internacionais de chassis, veículos de pequeno porte e das fábricas de carrocerias nacionais.

Ainda segundo Brasileiro *et al.* (2001), uma razão essencial para o crescimento dos operadores ilegais são as falhas regulatórias no serviço do TRIP. Uma dessas falhas é o regulamento federal que nunca submeteu as empresas, que funcionam há bastante tempo, a um processo licitatório (de verdadeira competição), em que as empresas que apresentassem as propostas de menor tarifa venceriam a licitação e entrariam no mercado.

2.2.3. Regulação do Transporte Rodoviário Interestadual de Passageiros no Brasil

As origens do processo regulatório do transporte rodoviário de passageiros iniciam-se, verdadeiramente, a partir do Decreto-lei nº. 8.463, de 27 de dezembro de 1945 (Lei Joppert), que concedeu autonomia financeira ao Departamento Nacional de Estradas de Rodagem (DNER) e criou o Fundo Rodoviário Nacional (FRN). Todavia, o serviço de TRIP teve sua primeira regulamentação apenas com o Decreto nº. 68.961, de 20 de julho de 1971, que atribui ao DNER a competência para sua execução (Castro Junior, 2003). A partir deste documento a administração e o gerenciamento do TRIP foram exercidos pelo DNER até o ano de 1990. Segundo Brasileiro *et al.* (2001), até essa década, as características indispensáveis dos regulamentos federal e estadual estavam alinhadas com a longa tradição brasileira de práticas de concessões no setor de transporte. Alguns aspectos devem ser considerados sobre esse assunto:

- Uma única rota é concedida a uma única empresa.
- Com fundamento em uma planilha de custo padrão e de uma estimativa da ocupação média do veículo, o poder público define os preços das passagens.
- Contratos de concessão devem ser concedidos por meio de um processo de licitação que, no geral, não ocorre. Ou seja, os contratos são renovados em tempos

determinados sem nenhuma negociação ou exigência adicional da parte do órgão responsável.

Passando por diversas transformações, apenas em 20 de março de 1998, com o Decreto nº. 2.521, a exploração, diretamente ou mediante permissão ou autorização, do Transporte Rodoviário Interestadual de Passageiros se torna competência da União, e sempre precedida por licitação (Brasil, 1998).

Em 2001 é instituída a Agência Nacional de Transportes Terrestres (ANTT), pela Lei nº. 10.233, com objetivo de atuar como fiscalizadora e reguladora das atividades de prestação de serviço e de exploração da infra-estrutura do TRIP, além de outros serviços do transporte terrestre. A ela cabe ainda determinar os preços das passagens, controlar a execução dos serviços, além de arbitrar conflitos e impedir situações que configurem como competição imperfeita ou infração da ordem econômica. Assim, em âmbito interestadual, somente empresas autorizadas pela ANTT, mediante homologação, permissão ou concessão, podem transportar passageiros (Brasil, 2001).

Outra atribuição da ANTT é estudar o mercado dessas linhas do TRIP (Brasil, 2001). Segundo Martins (2006), as condições nos mercados do TRIP podem ser caracterizadas em função dos seguintes parâmetros: frota, serviço, demanda, política pública, tipos de insumo, concentração da oferta de serviços e barreiras de entrada.

2.2.4. Funcionamento do Transporte Rodoviário Interestadual de Passageiros no Brasil

Sendo um meio de transporte público predominante no âmbito interestadual do Brasil, sua operação será realizada por meio de linhas de transporte que transpõem os limites de Estados, Distrito Federal e Territórios.

Como citado anteriormente, o transporte público interestadual de passageiros é feito, quase que exclusivamente nas estradas e sua principal função é o deslocamento de pessoas de um Estado para outro, através da malha rodoviária brasileira. Esse deslocamento é realizado em um itinerário da linha, que é o percurso a ser utilizado na execução do serviço.

Na resolução nº 16 (ANTT, 2002) foi aprovado o glossário de termos e conceitos utilizados pela ANTT na regulamentação da prestação dos serviços do TRIP, com o seguinte conceito de linha (conceito alterado conforme a resolução nº 2850 de 13 de agosto de 2008): “serviço de transporte rodoviário coletivo de passageiros executado em uma ligação de dois pontos terminais, nela incluída os seccionamentos e as alterações operacionais efetivadas, inclusive o serviço diferenciado, aberto ao público em geral, de natureza regular e permanente, com itinerário definido no ato de sua delegação mediante permissão”. Ou seja, as linhas onde as empresas operam são definidas como o próprio serviço.

Dessa forma, as linhas de transporte interurbano de passageiros podem ser classificadas como:

- *intermunicipal*: transporte de passageiros que transpõe os limites dos municípios, ligando-os dentro do mesmo Estado (Carneiro, 2005);
- *intermunicipal semi-urbano (metropolitano)*: transporte realizado entre municípios de regiões metropolitanas ou entre municípios vizinhos quaisquer (Ceará, 2001);
- *interestadual (convencional)*: transporte de passageiros que transpõe os limites de Estados, do Distrito Federal ou de Territórios (Brasil, 1998);
- *interestadual (semi-urbano)*: transporte de passageiros que, com extensão igual ou inferior a setenta e cinco quilômetros, transpõe os limites de Estados, do Distrito Federal ou de Territórios (Brasil, 1998);
- *internacional*: transporte de passageiros que transpõe as fronteiras nacionais (Brasil, 1998).

Para o funcionamento de uma linha do TRIP, é essencial definir o prefixo da linha, pois segundo Martins (2007), é um parâmetro importante para caracterizar a diferenciação dos serviços prestados em termos de operação, fiscalização, gestão e planejamento. O prefixo da linha é definido por uma seqüência de oito algarismos, conforme o exemplo da Figura 2.2:

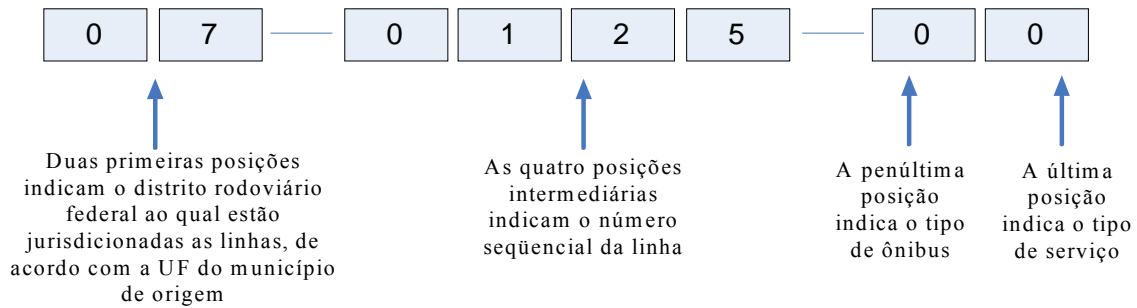


Figura 2.2: Prefixo da linha base Rio de Janeiro (RJ) – Brasília (DF)

Cada empresa possui um prefixo exclusivo para cada tipo de ônibus, de cada linha operada. Duas empresas que operem em um mesmo itinerário, com o mesmo tipo de serviço e de ônibus, possuem dois prefixos distintos, porque o número seqüencial da linha (as quatro posições intermediárias) é diferente para cada operadora.

O serviço de TRIP abrange outros conceitos que merecem ser considerados para uma melhor compreensão de seu funcionamento e para algumas análises que serão realizadas no presente trabalho, tais como:

- *pontos terminais*: definidos como os extremos do itinerário onde se dará o início e término da viagem (EBTU, 1988);
- *ligação*: denominada como uma conexão entre um par de origem e destino. Neste estudo serão consideradas apenas municípios que estão situadas em diferentes UFs e que são chamadas de *ligação inter-regional* (ANTT, 2002);
- *seção ou seccionamento*: designa um serviço realizado em trecho de itinerário de linha ou de sua área de influência, com fracionamento do preço de passagem (Ceará, 2001);
- *ponto de seção*: denomina a localidade, entre os municípios de origem e destino, onde é permitida à empresa a venda de passagens para embarque de passageiros na linha (Carneiro, 2005);
- *ponto de apoio*: refere-se ao local destinado a socorro, reparos e manutenção de veículos em viagem e atendimento da tripulação (Brasil, 1998);
- *ponto de parada*: relativo ao local de parada obrigatória, ao longo do itinerário, de forma a assegurar a alimentação, o conforto e o descanso aos passageiros e à tripulação dos ônibus, sempre respeitando o curso da viagem e o tempo devido (Brasil, 1998);

- *serviço regular e permanente*: concernente ao serviço operados em caráter contínuo, obedecendo a horários ou intervalos pré-estabelecidos, para o atendimento das necessidades básicas de transporte da população (EBTU, 1988);
- *mercado*: designa o núcleo de população em que há potencial de passageiros capaz de gerar demanda suficiente para exploração econômica de uma linha (Brasil, 1998);
- *estudo de mercado*: consiste na análise dos fatores que influenciam a caracterização da demanda de um mercado, para efeito de dimensionamento e avaliação da viabilidade de ligação de transporte rodoviário de passageiro, por meio de levantamento de dados e informações e de aplicação de modelos de previsão da demanda (Brasil, 1998);
- *demanda*: designa o movimento de passageiros, entre pares de localidades em um período de tempo determinado (Brasil, 1998).

A demanda é a responsável pela existência do serviço de TRIP, ou seja, quando não há uma demanda de passageiros em um determinado local, não é necessário uma linha para esta localidade. Este assunto será abordado no próximo tópico deste capítulo.

2.3. MODELOS DE DEMANDA DE TRANSPORTE

Demanda é um conceito histórico e específico a um determinado tipo de economia. Segundo Castro e Lessa (1967), é o volume de bens ou serviços que a sociedade está disposta a adquirir a determinados preços. Mas quando o preço de um determinado bem ou serviço se eleva e todas as outras coisas permanecem inalteradas, a quantidade demandada diminui ou vice-versa. Essa relação é observada com tanta frequência nos mercados e os economistas a chamam de *lei da demanda* (Hall e Lieberman, 2005).

No âmbito do transporte interestadual de passageiros, as empresas aéreas são as principais concorrentes das operadoras de ônibus. Assim, a quantidade demandada do transporte aéreo aumenta quando existem as reduções tarifárias, portanto espera-se que a demanda do TRIP diminua. Mas nem sempre a tarifa é o fator determinante para mudança do modo de transporte público, como exemplo pode-se citar os diversos atrasos nos vôos, deslocando passageiros do modo aéreo para o modo rodoviário. O presente trabalho se preocupará em

apenas em desenvolver um modelo de previsão da demanda para o TRIP e não entrará no contexto aéreo.

Precedentemente a apresentação do modelo utilizado pela ANTT para a previsão de demanda anual, vê-se a necessidade de conceituar o fenômeno demanda em transporte e apresentar os principais modelos utilizados no transporte de passageiros.

2.3.1. Demanda em Transporte

A necessidade ou o desejo de movimentação de uma pessoa, de um grupo ou até mesmo de cargas entre diferentes locais, geram uma demanda por transporte. Essa necessidade por transporte é o tronco para a interação entre atividades sociais e econômicas dispersas no espaço.

Dessa forma, Kanafani (1983) afirma que “a diversidade dos padrões da interação socioeconômica e a complexidade resultante na evolução da necessidade de transporte indicam que as análises formais e sistemáticas são essenciais para entender as relações entre a distribuição espacial das atividades e transporte”. Esse é o principal objetivo da análise de demanda por transporte. Segundo Ferreira (1999), a análise dessa demanda é um processo em que se procura identificar os determinantes da demanda e a maneira como eles interagem e afetam a evolução do volume de tráfego (ou viagens).

O processo de análise da demanda de transporte de passageiros é desenvolvido através de modelos. Segundo Ortúzar (2000), modelo é uma representação simplificada das características mais relevantes de um fenômeno ou de uma situação real, que busca abstrair maior clareza conceitual destes, reduzindo suas variedades e complexidades a níveis que permitam compreendê-los e especificá-los de forma adequada (Figura 2.3). A maioria dos resultados desta análise é realizada empregando modelos estatísticos e matemáticos. Esses modelos de análise da demanda de transporte, segundo Novaes (1986), podem ser utilizados para previsões de curto prazo com análise da situação presente, de médio e longo prazo com projeções detalhadas de variáveis socioeconômicas e ainda, de longo prazo envolvendo planejamento regional e de uso do solo.



Figura 2.3: Relação modelo-realidade

Segundo Molinero e Arellano (1998), é possível fazer modelos para previsão da demanda do transporte urbano de passageiros utilizando-se dois métodos:

- projetar a demanda de maneira proporcional à evolução do crescimento populacional ou em função do aumento na mobilidade individual.
- projetar a demanda por meio da comparação com outras cidades, onde o padrão de vida e a mobilidade são similares ou um pouco maiores.

Os modelos de previsão da demanda do transporte urbano não são muito práticos para a adaptação no TRIP, devido às longas distâncias e ao tamanho dos sistemas interurbanos comparados aos sistemas urbanos, além de variáveis que podem ser bem representativas para o transporte urbano e nem tanto para o TRIP.

No transporte interurbano, como no transporte urbano, a finalidade, a distância e o comportamento da viagem são os fatores mais importantes para a definição das variáveis a serem estudadas na análise da demanda. Vale ressaltar que esses fatores têm visões diferentes quando analisados em cada um dos sistemas, principalmente a finalidade da viagem (Kanafani, 1983).

Segundo Kanafani (1983), as finalidades ou os motivos de viagem do transporte urbano são trabalho e compras, diferentemente do TRIP, que são passeios, negócios pessoais e trabalho. Segundo EMBRATUR/FIPE (*apud* TCU, 2005), mais de 75% das viagens interestadual têm por fim lazer (passeios, visitas a parentes e amigos e férias). A distância é outro fator que influencia muito a construção desse tipo de modelo para o TRIP, pois

quando a distância é menor que 100 km entre uma cidade a outra, os comportamentos e os motivos são bastante parecidos com os utilizados em transporte urbano (Kanafani, 1983).

Em relação ao comportamento de indivíduos nas viagens interurbanas, são poucos os estudos nessa área, pois são caros e difíceis de conduzir. Portanto, como o transporte urbano, esses comportamentos permanecem sem muita compreensão, mas se sabe que esse tipo de serviço se diferencia das viagens interurbanas acima de 100 km, principalmente, por terem viagens rotineiras (repetidas).

2.3.2. Tipos de Modelos de Demanda

Os modelos de previsão da demanda de transporte foram classificados de diversas formas, conforme a literatura. O presente trabalho distinguirá os modelos de acordo com a Tabela 2.2 (Carneiro, 2005):

Tabela 2.2: Tipos de modelos de previsão da demanda

Classificação	Tipos de modelos
Modelos Convencionais	Geração de Viagens
	Distribuição de Viagens
	Divisão Modal
	Alocação de viagens
Modelos Desagregados	Modelos Comportamentais
	Modelos Atitudinais
Modelos baseados em novas tecnologias	Sistema de Informação Geográfica
	Sistemas Inteligentes

Fonte: Carneiro (2005)

2.3.2.1. Modelos Convencionais

Os modelos convencionais visam simplificar o problema dividindo-o em partes menores. Este processo é tradicional na área de planejamento em transporte, e a divisão do problema é realizada por meio do que se imaginava ser a seqüência de decisão do usuário. Conhecido como *Urban Transportation Planning System (UTPS)*. Este método seqüencial se divide em quatro etapas:

- Modelo de Geração de Viagens.
- Modelo de Distribuição de Viagens.
- Modelo de Divisão Modal.
- Modelo de Alocação de Viagens.

Uma das principais críticas a esse processo é em relação à seqüência das etapas, visto que elas podem não representar fielmente as decisões do usuário. Segundo Marshment (2000), os modelos convencionais tiveram sua utilidade após sua criação, na década de 1950. Entretanto, nos dias de hoje, eles têm se mostrado incapazes de atender a diversos problemas no que tange o planejamento em transporte. Contudo, apesar dessas críticas, estas etapas ainda são utilizadas, não necessariamente contíguas ou seguindo o método seqüencial.

Os modelos convencionais também são conhecidos como modelos baseados em zonas, posto que, em sua grande maioria, utilizam dados agregados das variáveis associadas às características socioeconômicas e demográficas das zonas de tráfego de áreas urbanas. O nível de agregação dos dados nestas áreas depende do estudo realizado, ou seja, além da divisão por zonas, há também domiciliar ou individual (Carneiro, 2005). Geralmente é possível a aplicação desses modelos para o planejamento do TRIP, bastando que se mude o nível de agregação das variáveis. Todavia, essa tarefa não é muito simples quando se trata da obtenção de dados para todos os municípios do País.

A presente dissertação abordará apenas os principais conceitos e modelos vinculados às duas primeiras etapas, pois a pesquisa será realizada com base na geração e distribuição de viagens do TRIP.

2.3.2.1.1. Técnicas de Geração de Viagens

Os estudos relativos à geração de viagem constituem uma fase vital do processo de planejamento dos transportes. Por isso é essencial que os atuais fatores que determinam a produção e a atração de viagens sejam claramente entendidos antes que se avalie a natureza das futuras demandas de viagens (Brutton, 1979). Segundo Gonzales Taco (1997), é possível definir a geração de viagens como “a determinação do número total de viagens

associada a uma zona de tráfego, consistindo em viagens produzidas ou atraídas pela área de estudo”.

A previsão do número de viagens produzidas e atraídas é determinada por meio de relações matemáticas estabelecidas entre o uso do solo e os padrões de deslocamentos para as condições presentes. Essas informações são adquiridas através da coleta de dados. Antes da coleta é necessário, porém, que se defina a área de estudo. Como a pesquisa é no âmbito de transportes, define-se a área de estudo em zonas de tráfego. Nesse sentido, experiências têm mostrado que o nível de renda que afeta o número, a frequência e o modo de viagens, e a população total (fator significativo para estimar a geração de viagem) são variáveis necessárias para a previsão de viagens, assim como os padrões do uso do solo e a capacidade do sistema de transporte existente.

Diversas são as técnicas capazes de prever o número de viagens produzidas e atraídas (modelos de regressão linear convencional, fatores de crescimento, relação tráfego e uso do solo e análise de categorias). Esta previsão é obtida, em geral, por meio da regressão linear convencional (simples ou múltipla), em que o objetivo principal é analisar a influência de um ou mais fatores independentes agindo simultaneamente no total das viagens. O propósito desta análise é desenvolver uma equação para a produção e outra para atração de viagens da seguinte forma:

$$Y = \alpha + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \dots + \beta_n \cdot X_n \quad (2.1)$$

onde,

Y = é a variável dependente (ou seja, número de viagens produzidas ou atraídas por modo e propósito);

$X_1 \dots X_n$ = são as variáveis independentes relacionadas, por exemplo, o uso do solo, ou características socioeconômicas e/ou capacidade do sistema;

$\beta_1 \dots \beta_n$ = são os coeficientes das respectivas variáveis independentes;

α = é uma constante cuja finalidade é representar a parcela de Y que não foi explicada pelas variáveis independentes.

2.3.2.1.2. Técnicas de Distribuição de Viagens

O modelo de geração de viagens desenvolve o somatório de linhas e colunas da matriz de origem/destino (O/D). Assim, o modelo de distribuição de viagens tem o objetivo de relacionar as origens com os destinos (somatório das linhas com somatório das colunas), determinando o número de viagens entre um par de zonas de tráfego (i,j) (Novaes, 1982). As origens e os destinos citados podem ser os municípios brasileiros onde se encontram as linhas do TRIP.

Neste contexto, com as técnicas de distribuição de viagens, é possível conhecer os movimentos entre as linhas interestaduais e estabelecer a frequência de viagens realizadas entre os municípios. Dentre as técnicas, os modelos de fator de crescimento e o modelo de distribuição gravitacional (ou apenas, modelo gravitacional) são os mais utilizados e conhecidos.

Os métodos de fator de crescimento utilizam um procedimento de previsão da demanda futura baseado em uma matriz O/D histórica anterior (Ortúzar, 2000). De acordo com a literatura, vários métodos foram apresentados para o fator de crescimento, dependendo da evolução destes, conforme apresentado na Tabela 2.3 (Abreu, 1999):

Tabela 2.3: Evolução dos Métodos de Fator de Crescimento

Tipos de Métodos	Representação Matemática	Considerações sobre o Fator de Crescimento (E)
Fator Uniforme	$T_{ij} = t_{ij}E$	É uniforme para toda área de estudo.
Fator Médio	$T_{ij} = t_{ij} \left(\frac{E_i + E_j}{2} \right)$	É obtido pelo valor médio entre fatores de crescimento da zona de origem (E_i) e da zona de destino (E_j).
Fratar	$T_{ij} = \frac{T_{i(G)} \times t_{ij} E_j}{t_{ij} E_j + t_{ik} E_k + \dots + t_{in} E_n}$	Segundo Brutton (1979), este método parte das seguintes suposições: 1) A distribuição de viagens futuras de uma dada zona de origem é proporcional à sua distribuição atual. 2) A distribuição destas viagens futuras é modificada pelo fator de crescimento da zona para a qual estas viagens são atraídas.
Detroit	$T_{ij} = t_{ij} \frac{E_i + E_j}{E}$	Composto pelo fator de crescimento da zona de origem (E_i) na proporção de crescimento da zona de destino (E_j) dividido pelo crescimento da área de Estudo (E).

Onde:

T_{ij} = número de viagens previstas da zona i para zona j ;
 $T_{i(G)}$ = número de viagens futuras esperadas, geradas na zona i ;
 $t_{ij} \dots t_{in}$ = número existente de viagens entre a zona i e todas as outras zonas $j \dots n$;
 $E_i \dots em$ = fator de crescimento da zona $i \dots n$;
 E = fator de crescimento.

Fonte: Adaptado de Abreu (1999)

Segundo Ortúzar (2000), métodos de fator de crescimento são fáceis de entender e aplicar, e quando se faz uma previsão de curto prazo ou para áreas que tem uma estrutura de desenvolvimento estável, obtêm-se valores significativos e aceitáveis. Entretanto, estes modelos precisam dos mesmos dados de modelos mais sofisticados para produzirem resultados menos pontuais, além de, não indicarem as principais variáveis ou as causas dos deslocamentos de pessoas dentro de uma área de estudo antes da previsão dos padrões de viagens.

Partindo destas considerações, buscou-se desenvolver melhorias nestes métodos de fator de crescimento para a distribuição de viagens, ou seja, procurou-se explicar os motivos para os deslocamentos de passageiros. Dentre os métodos desenvolvidos se destaca o modelo gravitacional (MG). A concepção deste modelo será exposta na apresentação do método utilizado pela ANTT.

É importante ressaltar que esses modelos costumam calcular o número de viagens, pois, muitas vezes, a construção da matriz O/D é desenvolvida com base no número de deslocamento de pessoas. Como as matrizes O/D do presente trabalho são constituídas pelo número de passageiros em viagens interestaduais e intermunicipais (apenas do estado do Ceará) e como já existem modelos de regressão que utilizam passageiros substituindo viagens (por exemplo, ANTT), o fenômeno a ser estudado não será a quantidade de viagens, mas quantidade de passageiros.

2.3.2.2. Modelos Desagregados

São modelos baseados em teorias comportamentais e construídos através de informações de comportamentos e atitudes de indivíduos como dados de entrada. Segundo Ortúzar (2000), os modelos desagregados podem ser mais eficientes no uso de informações do que os modelos convencionais, já que além de permitir uma representação mais flexível de

variáveis que são relevantes para estudos em transporte (população, renda, etc), eles têm uma maior probabilidade de serem estáveis no espaço e no tempo.

Os modelos desagregados são divididos em modelos comportamentais e atitudinais. Os comportamentais determinam, através de atributos, a probabilidade de o usuário selecionar um destino qualquer e/ou modo de transporte que utilizará. Os atitudinais fundamentam-se nas escolhas dos usuários (Carneiro, 2005).

2.3.2.2.1 Modelos Comportamentais

Os Modelos Comportamentais consistem em métodos que buscam a relação entre as necessidades individuais dos usuários com os seus deslocamentos dentro do sistema de transporte. Estes modelos, afirma Novaes (1986), partem dos conceitos da Teoria do Consumidor, em que o usuário de transporte é um consumidor deste serviço.

De acordo com Abreu (1999) e Novaes (1986), a regra que serve de base para a construção dos modelos comportamentais é a de que as decisões adotadas por um indivíduo em relação à preferência envolvem fatores subjetivos e/ou racionais, lembrando que, embora as decisões desses indivíduos incorporem variáveis subjetivas, elas se mantêm com o mesmo padrão comportamental ao longo do tempo. Em outras palavras, supõe-se que os padrões comportamentais, mesmo subjetivos, em parte, não são incertos ou totalmente aleatórios, mas se conservam dentro de determinadas circunstâncias (Novaes, 1986).

Com isso, para modelar os comportamentos individuais, devem-se medir quantitativamente as preferências desses usuários, o que, de acordo com Novaes (1986), se formaliza por meio de uma função utilidade. Essa, por sua vez, está vinculada aos fatores ou atributos que representam o nível de serviço. As variáveis, que em geral compõem este nível de serviço, são variáveis ligadas ao custo, ao tempo, à segurança, ao conforto e à conveniência.

Os conceitos estatísticos e matemáticos mais utilizados para a construção de modelos comportamentais são os modelos de regressão de escolha qualitativa e a análise de

discriminante, sendo que com o primeiro é possível analisar três abordagens: *modelo de probabilidade linear, logit e probit*. O modelo *logit* é o mais utilizado na área de transporte e a expressão geral é dada pela seguinte Equação 2.2:

$$P_k = \frac{e^{U_k}}{\sum_{L=1}^n e^{U_L}} \quad (2.2)$$

onde,

P_k = probabilidade de o usuário decidir viajar pelo modo k ;

U_k = utilidade da alternativa k ;

U_L = utilidade da alternativa L , para $L = 1, \dots, n$;

N = número de alternativas do modo de transporte.

Segundo Abreu (1999), os modelos comportamentais podem realizar as etapas de distribuição de viagens (ou de passageiros) e divisão modal simultaneamente. Dessa forma, no momento em que é definida a probabilidade de um determinado usuário escolher o modo de transporte, pode-se aplicar a Equação (2.2) para estimar o fluxo de passageiros entre i (origem) e j (destino) utilizando o modo k , como apresentado na seguinte equação:

$$F_{ij}^k = F_{ij} \times P(k \setminus i, j) \quad (2.3)$$

onde,

F_{ij}^k = número de passageiros entre os municípios i e j utilizando o modo k ;

F_{ij} = número de passageiros entre os municípios i e j ;

$P(k \setminus i, j)$ = probabilidade de um passageiro escolher o modo de transporte k dado que a viagem é realizada da zona i para a zona j .

Os modelos comportamentais contribuem significativamente para a reformulação teórica dos principais modelos de distribuição de viagens. Entretanto, são métodos complexos para estudos do TRIP, visto que a obtenção dos dados referentes à necessidade e à preferência dos usuários tem um custo elevado e exigiria tempo para uma pesquisa em âmbito nacional.

2.3.2.2.2 Modelos Atitudinais

O comportamento de um usuário em relação às opções de transporte tem diferenças, não muito significativas, se comparado à sua atitude frente a situações específicas. Isso ocorre porque a atitude do usuário em relação às opções de transporte a ele sugeridas é resultado de uma interpretação pessoal e que pode se concretizar de forma bastante diferente do previsto (Novaes, 1986). Ou seja, supõe-se que se alguma vantagem for oferecida dentro do serviço de transporte para os usuários, através dos modelos comportamentais, é possível fazer uma análise da demanda dentro desta vantagem. Mas algumas vezes estes usuários não aceitam estas vantagens por diversas percepções, como a falta de esclarecimentos e até mesmo por motivos culturais.

Os modelos atitudinais (ou modelos com preferência declarada) visam identificar reações dos usuários não captadas por meio dos modelos convencionais e comportamentais. Os modelos atitudinais são métodos que necessitam de informações mais detalhadas, colhidas em situações específicas e concretas. Isso faz com que estes modelos sejam aplicados em análises da operação de sistemas de transporte já implantados, visando sua melhoria.

Os dados necessários ao desenvolvimento e ajuste de modelos atitudinais são obtidos através de entrevista direta com cada indivíduo usuário do serviço, ou até mesmo indiretamente por meio de envio de questionário (Novaes, 1986). Estes dados coletados, segundo Carneiro (2005), são quantificados em função da diferença dos atributos (tarifa, tempo de viagem, distância, segurança, entre outros) em cada uma das alternativas e de acordo com uma constante que represente o peso de cada um dos atributos. Logo, a formulação geral para os modelos atitudinais é:

$$F_k = f(\alpha_1 - \alpha_2) \quad (2.4)$$

onde,

F_k = número de passageiros no modo k ;

$(\alpha_1 - \alpha_2)$ = diferença entre os modos 1 e 2 no atributo α .

Como os modelos comportamentais, os atitudinais são métodos complexos para estudos do TRIP devido à dificuldade na obtenção de dados dos usuários em todos os municípios do Brasil.

2.3.2.3. Modelos com Base em Novas Tecnologias

Nas últimas décadas, em virtude das limitações dos modelos convencionais e desagregados, da evolução populacional sujeita a fatores exógenos e de um mundo onde a informação é uma das principais matérias-primas para o planejamento; viu-se a necessidade de técnicas mais específicas para coleta de dados relacionados a transporte e para um melhor entendimento das inter-relações entre as variáveis intervenientes na geração e na distribuição de viagens (Gonzales Taco, 1997; Silva, 2003). Desse modo, os modelos, com base em novas tecnologias, vêm surgindo na última década, pois além de possibilitarem um trabalho com maior número de variáveis, eles proporcionam uma melhor compreensão e representação do fenômeno estudado.

Entre as diversas técnicas dos modelos com base em novas tecnologias, o Sistema de Informações Geográficas (SIG) é uma técnica bastante utilizada na área de transportes que, segundo Câmara (1999), “consiste de um conjunto de ferramentas capazes de adquirir, armazenar, recuperar, transformar e emitir informações espaciais”.

Gonzales Taco (1997) desenvolveu uma metodologia para a geração de viagens baseada em fundamentos de Sistemas de Informação Geográfica e técnicas de Sensoriamento Remoto.

Outra técnica também muito utilizada reside nos Sistemas Inteligentes (composto por Lógica *Fuzzy*, Redes Neurais, Mineradores de Dados, *Data Warehouse*, entre outros) que

têm por objetivo prover as organizações sistematicamente de coleta de dados e transformá-los em informações de caráter estratégico (Silva, 2003).

Em 1999, Abreu desenvolveu uma metodologia para distribuição de viagens utilizando, de forma integrada, a Lógica *Fuzzy*, o SIG e o Método de Análise Hierárquica. Com esta metodologia é possível identificar as variáveis que intervêm na distribuição de viagens. Silva (2003) explanou um sistema inteligente, utilizando tecnologia *Data Warehouse* e *OLAP*, de apoio à tomada de decisão para as empresas operadoras do transporte rodoviário de passageiros.

Estas ferramentas também tiveram suas utilidades no TRIP. Carneiro (2005) apresentou uma metodologia para a previsão da demanda do TRIP nas ligações entre municípios utilizando mecanismos de Rede Neurais Artificiais. Aguiar *et al.* (2006) construíram um modelo para estimativa de demanda do TRIP aplicando técnicas de Mineradores de Dados e de RNA, com sua utilização em diversas tarefas, como planejamento estratégico, avaliação de corredores de transporte e estudos tarifários.

2.3.3. Modelo de demanda utilizado pela Agência Nacional de Transportes Terrestres

Os modelos convencionais, citados anteriormente, são métodos tradicionais na área de planejamento em transporte. Dentre as técnicas de distribuição de viagens do tipo convencional se destacam os modelos gravitacionais. O primeiro modelo gravitacional tem origem, por analogia, na lei gravitacional de Newton, que propõe que a atração entre dois corpos é diretamente proporcional à massa deles e inversamente proporcional ao quadrado da distância que os separa. Segundo Kanafani (1983), as primeiras aplicações destes conceitos gravitacionais foram voltadas ao transporte interestadual.

Diferentemente dos modelos de fatores de crescimento, os modelos gravitacionais têm o objetivo de estimar o número de viagens ou passageiros para cada célula da matriz O/D sem a necessidade de obter padrões de viagem anteriores ao ano de estudo (Ortúzar, 1994 *apud* Carneiro, 2005). De acordo com Abreu (1999), ao longo dos anos, vários modelos têm sido propostos a partir de alterações realizadas no primeiro modelo gravitacional, conforme a Tabela 2.4:

Tabela 2.4: Evolução dos Modelos Gravitacionais

Evolução	Representação	Considerações
Modelo Original	$T_{ij} = K \frac{P_i P_j}{d_{ij}^2}$ <p>P_i, P_j = população das cidades ou zonas de tráfego i e j; d_{ij} = distância entre i e j; K : fator de ajustamento ou proporcionalidade.</p>	Falhas ao considerar o fluxo proporcional à população das localidades e inversamente proporcional ao quadrado da distância.
Primeira Evolução	$T_{ij} = K \frac{O_i D_j}{f(d_{ij})}$ <p>O_i = número de viagens com origem na zona i; D_j = número de viagens atraídas por j; d_{ij} = distância entre i e j; $f(d_{ij})$ = função distância (Ex: $f(d_{ij}) = d_{ij}^n$), em que n será um parâmetro, a ser estimado, de calibração do modelo.</p>	Busca minimizar as falhas apresentadas na abordagem anterior através da melhoria da função da impedância.
Segunda Evolução	$T_{ij} = \frac{O_i D_j F_{ij} K_{ij}}{\sum_{j=1}^n D_j F_{ij} K_{ij}}$ <p>O_i = número de viagens com origem na zona i; D_j = número de viagens atraídas por j; F_{ij} = fator de tempo de viagem obtido empiricamente para expressar o efeito da separação espacial (Brutton, 1979); K_{ij} = fator de ajustamento, função dos padrões socioeconômicos das zonas de tráfego; n: número de zonas de tráfego.</p>	Melhora a etapa anterior com a introdução de variáveis mais complexas e relevantes, como o tempo de viagem.
Custo Generalizado	$T_{ij} = A_i O_i B_j D_j f(C_{ij})$ <p>A_i = fator de equilíbrio que assegura que o número total de viagens que deixa a zona é igual ao número de viagens produzidas pela zona; B_j = fator de equilíbrio que assegura que o número de viagens que entra na zona é igual ao número de</p>	

	viagens atraídas pela zona; C_{ij} = custo generalizado da viagem entre i e j por um determinado modo de transporte; $f(C_{ij})$ = função custo generalizado, cujas formulações podem ser vistas no Brutton (1979).	
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

Fonte: Adaptado de Abreu (1999)

Como citado anteriormente, a ANTT estuda o mercado das linhas do TRIP, sendo que, uma das etapas deste estudo é a previsão da demanda anual de passageiros. O método utilizado pela Agência, com a finalidade de estimar esta previsão, fundamenta-se em uma variação dos modelos gravitacionais para o cálculo da demanda, modificando-a de forma a incorporar diversas variáveis socioeconômicas. Então, segundo a ANTT (2007), a previsão dessa demanda anual para cada seção de uma linha do TRIP, após a divisão do espaço territorial por distância e por agrupamentos das ligações regionais e inter-regionais, é dada pela seguinte equação:

$$D_{ij} = e^{\left[\begin{array}{l} \beta_0 + \beta_1 \ln\left(\frac{P_i}{1000}\right) + \beta_2 \ln\left(\frac{P_j}{1000}\right) + \beta_3 \ln\left(\frac{r_i r_j}{10000}\right) + \beta_4 \ln\left(\frac{M_i}{P_i}\right) + \beta_5 \ln\left(\frac{M_j}{P_j}\right) + \\ + \beta_6 \ln(d_{ij}) + \beta_7 \ln(dummy) + \beta_8 \ln(mig_i) + \beta_9 \ln(mig_j) \end{array} \right]} \quad (2.5)$$

onde,

D_{ij} = fluxo anual total de passageiros entre os municípios i e j ;

P_i, P_j = população dos municípios i e j ;

r_i, r_j = renda média per capita dos municípios i e j ;

M_i, M_j = movimento total anual de embarques de passageiros nos terminais dos municípios i e j , considerando tão somente fluxos interestaduais;

d_{ij} = distância rodoviária entre os municípios i e j ;

$dummy$ = variável que determina se o município é ou não um pólo turístico;

mig_i, mig_j = taxa de habitantes que não são naturais dos Estados dos municípios de origem e destino;

$B_0, B_1, B_2, B_3, B_4, B_5, B_6, B_7, B_8$ e B_9 = parâmetros da equação em função da extensão.

Entretanto, como pode haver linhas com várias seções, a demanda corrigida para estas seções intermediárias é apresentada conforme a Equação 2.6:

$$D_c = D_{ij} \times \frac{d_s}{d_l} \quad (2.6)$$

onde,

D_c = emanda corrigida para seção em estudo;

D_{ij} = Demanda Anual entre os municípios i e j ;

d_s = extensão da seção em estudo;

d_l = extensão da linha;

Importante lembrar que, como o número de seções em cada município é ilimitado, a cada nova linha com seção em um município qualquer, uma nova demanda pode ser prevista.

2.4. TÓPICOS CONCLUSIVOS

- O sistema rodoviário no País teve sua ascendência a partir do aumento dos manufaturados fabricados no Brasil e da implantação da indústria automobilística no final da década de 50.
- Com uma malha rodoviária de 1,8 milhões de quilômetros e a maioria dos deslocamentos de pessoas em viagens interurbanas através do transporte rodoviário por ônibus, o TRIP torna-se um serviço vital ao Brasil.
- O serviço de transporte rodoviário de passageiros convive com dificuldades para exercer suas funções, como precariedade das rodovias brasileiras, falta de segurança e fiscalização, ruim estado de alguns terminais rodoviários e principalmente a ação predatória das empresas ilegais.
- Em 2001, é instituída a Agência Nacional de Transportes Terrestres (ANTT) com objetivo de atuar como fiscalizadora e reguladora das atividades de prestação de serviço e de exploração da infra-estrutura do TRIP. Ou seja, no âmbito

interestadual, somente empresas autorizadas pela ANTT, mediante homologação, permissão ou concessão podem transportar passageiros.

- Além destes objetivos, a ANTT também tem a competência de estudar o mercado das linhas do TRIP, de modo que uma das etapas deste estudo é realizar a previsão da demanda anual. A identificação dessa demanda, independente do tempo (anual, mensal, entre outros), é fundamental para o planejamento do serviço.
- O processo de análise da demanda de transporte é construído por meio de modelos, que são uma representação simplificada das características mais relevantes de um fenômeno.
- Um único modelo, ou mesmo um conjunto de modelos, pode não representar as relações socioeconômicas entre um par de O/D em um país continental como o Brasil. Então, existe a necessidade de se desenvolver novos modelos que substituam os anteriores e possam atender, com mais representatividade, estas diversidades socioeconômicas entre os estados e municípios.

3. REGRESSÃO COM EFEITOS ESPACIAIS

3.1. APRESENTAÇÃO

Como mencionado anteriormente, diversas são as técnicas que vem sendo testadas e são capazes de prever a demanda de passageiros. Dentre estas, se destaca a principal ferramenta da econometria, o modelo de regressão linear (simples ou múltipla). É uma ferramenta para analisar e descrever a relação existente em duas ou mais variáveis, de modo que, uma destas seja explicada pelas demais. Essa ferramenta segue três importantes pressupostos necessários para sua validação:

- erros normais com média zero – $E(e_i) = 0$;
- erros com variância constante – $\text{Var}(e_i) = \sigma^2$;
- erros não correlacionados – $\text{Cov}(e_i, e_j) = 0$ para $i \neq j$.

O fato é que, constatada a dependência espacial, por meio de ferramentas da estatística espacial (seção 3.4), a utilização de modelos de regressão linear convencional torna-se inválida, pois é pouco provável que o pressuposto de os erros serem não correlacionados seja verdadeiro. Dessa forma, recorre-se a modelos alternativos que incorporem estruturas espaciais, sendo estes denominados de modelos de regressão com efeitos espaciais.

Câmara *et al.* (2004) afirmam que, quando o processo é não-estacionário, necessariamente utiliza-se modelos de regressão com efeitos espaciais locais, os quais permitem um melhor entendimento do processo espacial.

Anteriormente à apresentação das técnicas de regressão com efeitos espaciais locais, o presente capítulo indicará alguns conceitos de análise espacial, além de mostrar algumas ferramentas da estatística espacial (matriz de proximidade espacial e indicadores para identificação da dependência espacial) e técnicas de regressão com efeitos espaciais globais.

3.2. CONCEITOS DE ANÁLISE ESPACIAL

Diversas são as ferramentas e técnicas utilizadas pelos planejadores em transporte para estudar o deslocamento de pessoas e as viagens realizadas dentro de um determinado espaço. Esta movimentação humana no espaço, Tarrus (*apud* Vasconcellos, 2001) define em três estágios: a migração internacional e regional, para a permanência durante um período; a mobilidade residencial no município, relacionada aos ciclos familiar e aos condicionantes socioeconômicos e a mobilidade diária, referente às atividades cotidianas.

Dessa forma, devido a esta relação entre espaço e transporte, em que o último é influenciado continuamente pela dinâmica urbana e regional, as técnicas que podem ser as mais adequadas para estudos relacionados aos três estágios citados são as que utilizam a Análise Espacial (AE).

Então, dada a importância de se analisar espacialmente atributos ligados ao sistema de transportes, bem como as incipientes análises que consideram os aspectos espaciais desses atributos na modelagem da previsão da demanda, destaca-se a necessidade da utilização de ferramentas de AE. Essas têm a capacidade de permitir uma compreensão mais clara do comportamento espacial dos vários fenômenos em transportes por trabalhar desde a coleta de dados até a análise confirmatória.

Os princípios básicos da Análise Espacial consistem em entender de que forma os dados procedentes dos fenômenos ocorridos no espaço se organizam e qual a relação existente entre eles (Henrique, 2004). Segundo Teixeira (2003), pode-se definir análise espacial como qualquer sistema que torna possível a apresentação (ou seleção), manipulação, análise, inferência e estimação de dados espaciais.

Dados espaciais, por sua vez, são caracterizados no espaço em função de um determinado sistema de coordenadas absolutas ou relativas. Chou (1997) define dados espaciais como aqueles que possam ser representados em forma de mapa, enquanto que para Anselin (1992) os dados espaciais são os que possuem uma posição absoluta ou relativa no espaço. Para a análise das relações espaciais entre esse tipo de dados utilizam-se algumas ferramentas, com destaque para as propostas feitas por Anselin (1997) que as classificam em quatro tipos diferentes, como apresentado na Figura 3.1.

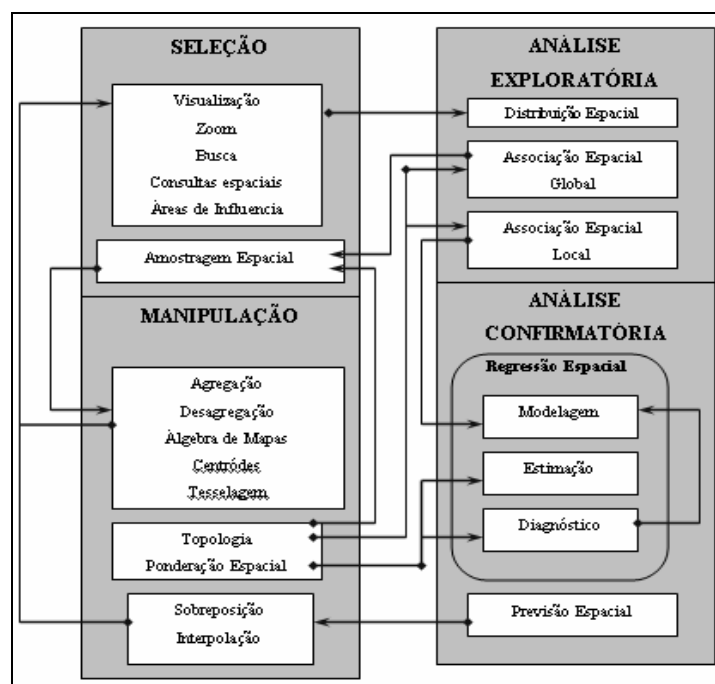


Figura 3.1 : Inter-relação entre os quatro grupos de ferramentas de análise espacial
 Fonte: Anselin (1997)

Na fase de *seleção* dos dados espaciais, apenas navega-se em um banco de dados geográficos, por meio de consultas e mapas simples. A fase de *manipulação* envolve diversificados tipos de função que criam dados espaciais, que vão da simples função de agregação ou de desagregação de dados, para se chegar à área de estudo, até a utilização de mapas para visualização das variáveis no espaço. Cabe à fase de *análise exploratória* verificar como os dados estão distribuídos no espaço, descobrir padrões de associação espacial, sugerir a existência de não-estacionariedade (conceito mais adiante) e identificar observações atípicas. E, por último, a *análise confirmatória*, que se encarrega da parte de obtenção dos modelos, das estimações e da realização de diagnósticos por meio dos resultados obtidos (Câmara, 2000; Krempi, 2004).

Muitas vezes, planejadores em transporte utilizam apenas os dois primeiros grupos de ferramenta de análise espacial. Mas, de forma a aumentar a robustez nas análises que utilizam ferramentas de AE, utiliza-se a Estatística Espacial, que engloba basicamente o ferramental matemático da análise exploratória e da análise confirmatória. A Estatística Espacial auxilia o planejador a estabelecer critérios quantitativos de agrupamento ou dispersão de dados espaciais, a obter o grau de dependência espacial entre as observações, entre outros aspectos (Teixeira, 2003). Na seqüência, tratar-se-á da Estatística Espacial

com foco na identificação da dependência espacial da variável dependente e da não-estacionariedade.

3.3. ESTATÍSTICA ESPACIAL

Estatística é a ciência que busca explicar fenômenos, experimentos e determinados estudos por meio de teorias probabilísticas e de amostragem. Esta teoria de amostragem, segundo Griffith (1984), implica articulação de percepções estatísticas de população, amostra, espaço amostral e probabilidade. Entretanto, a Estatística Espacial (EE) é conceituada como a aplicação desta teoria para a situação geográfica, ou seja, envolve a interpretação destas quatro percepções em situações geográficas.

O objetivo principal da EE é caracterizar padrões espaciais (similaridade de áreas que fazem fronteira). Esses padrões causam problemas de mensuração, chamados de efeitos espaciais, como, por exemplo, dependência espacial e heterogeneidade espacial, afetando a validade dos métodos estatísticos convencionais, bem como de modelos desagregados.

Um dos principais conceitos relacionados à EE é o de dependência espacial ou autocorrelação espacial, derivado do conceito estatístico de correlação, utilizado para mensurar a relação entre duas variáveis aleatórias. No caso, a correlação é medida por meio da mesma variável (por isso o nome autocorrelação), indicando o grau de influência que uma unidade espacial tem sobre seus vizinhos. Caso a autocorrelação seja positiva diz-se que a ocorrência de evento influencia o acontecimento de outro semelhante ao seu redor, implicando uma distribuição aglomerada. Se um mesmo evento influencia ou impede a ocorrência de um outro ao seu redor, diz-se que existe autocorrelação negativa e uma distribuição aproximadamente equidistante dos eventos (Teixeira, 2003). No tópico 3.3.2, sobre dependência espacial, discutir-se-á mais a fundo os conceitos e o modo de encontrar a autocorrelação positiva e negativa.

O termo heterogeneidade espacial é outro conceito relacionado à EE, e se refere à variação no relacionamento excedente ao espaço geográfico, ou seja, a uma diferenciação espacial ou regional que segue da intrínseca singularidade de cada local. Segundo Anselin (1992) e Peña (2006), as diferenças sistemáticas na ocorrência de determinados fenômenos, em diferentes regiões, podem ser comprovadas em vários aspectos da análise estatística, como:

podem ocorrer sob a forma de diferentes distribuições influenciadas pelo espaço ou pelos subconjuntos de dados, ou mesmo sob a forma de diferentes médias aritméticas, variâncias ou outros parâmetros entre os subconjuntos.

Com isso, se uma determinada região tem o número médio de passageiros transportados pelo sistema de TRIP diferente de outra, ou se até mesmo as variáveis estudadas têm um efeito distinto em certas regiões, este será um típico caso de heterogeneidade espacial.

Percebe-se, assim, que as técnicas de análise espacial, e especialmente as de estatística espacial, podem ampliar significativamente a capacidade de entendimento dos padrões espaciais associados aos dados de área e aos diferentes casos de heterogeneidade espacial, sobretudo quando se referem aos indicadores sociais que possuem autocorrelação.

3.3.1. Matriz de Proximidade Espacial (W)

Anteriormente ao cálculo de indicadores de autocorrelação espacial, é necessário estudar a matriz de proximidade espacial ou de vizinhança (ou apenas matriz W), pois ela é a responsável por representar a estrutura espacial. A matriz W , segundo Krempi (2004), estima a variabilidade espacial de dados de áreas (envolve elementos que associam o mapa geográfico a uma base de dados), sendo uma ferramenta muito útil para descrever o arranjo espacial dos objetos.

A matriz W consiste em uma matriz quadrada ($n \times n$), não estocástica, em que cada elemento w_{ij} desta matriz representa uma medida de proximidade espacial, um peso ou um grau de conectividade entre diferentes áreas (zonas, municípios, entre outros) i e j . Uma determinada área i não sofre influência dela mesma, ou está distante zero quilômetro ou zero minuto (ou qualquer outra unidade de medida) de si própria. Então, a diagonal da matriz W é nula, ou seja, $w_{ii} = 0$ para todo $i = 1, 2, \dots, n$ (Ripley, 1981; Krempi, 2004; Silva, 2006).

Diversas são as maneiras de construção da matriz W , sendo que a forma mais utilizada é a de contigüidade (tipo binária), ou seja, o indicador binário recebe o valor 1 se duas áreas têm fronteiras em comum e o valor 0 caso contrário, como mostra a Figura 3.2:

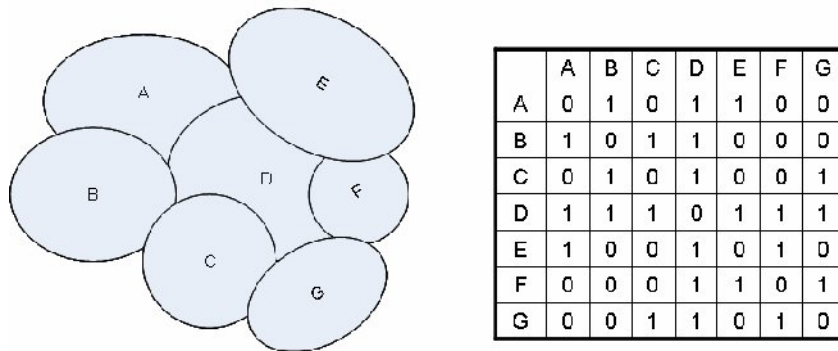


Figura 3.2: Matriz de proximidade espacial

Assunção *apud* Silva (2006) e Silva (2006) apresentam algumas escolhas comuns para a construção da matriz W , além das que compartilham fronteiras, tais como:

1. $w_{ij} = 1$, se o centróide da área i está distante a menos de “ x ” quilômetros da área j (a distância é escolhida de acordo com o objetivo do trabalho), e $w_{ij} = 0$, caso contrário. Como na maneira apresentada anteriormente da construção da matriz W , esta também é uma matriz simétrica.
2. $w_{ij} = 1$, se o centróide da área j é um dos k centróides mais próximos daquela área i e $w_{ij} = 0$, caso contrário. É importante salientar que k pode assumir qualquer valor inteiro e que a matriz W não é simétrica, pois não necessariamente os k vizinhos mais próximos de i serão o mesmo para j .
3. Seja d_{ij} a distância entre os centróides de i e j . Fazendo $w_{ij} = 1/(1 + d_{ij})$, as áreas extremamente próximas, com $d_{ij} \approx 0$, teriam assim $w_{ij} \approx 1$, e à medida que estas áreas se afastam, teriam $w_{ij} \approx 0$. Como a distância entre i e i é zero, o valor de w_{ii} seria igual a 1. Mas, neste caso, deve-se forçar o resultado, ou seja, $w_{ii} = 0$.
4. $w_{ij} = l_{ij}/l_i$ para $i \neq j$ e $w_{ij} = 0$ para $i = j$, em que l_{ij} é o comprimento da fronteira comum entre as áreas i e j e l_i o perímetro de i . Este modelo leva em consideração o tamanho da fronteira entre as áreas i e j .
5. A matriz W pode ser utilizada de uma forma que não considere os limites e as posições geográficas das áreas, como a mesma forma utilizada no item 3, mas substituindo o d_{ij} pelo tempo de viagem, por um modo determinado entre as áreas i e j .
6. Outro modo de construir a matriz W seria utilizando w_{ij} como o valor ou a quantidade de trocas comerciais entre as áreas i e j . Neste contexto, áreas distantes

geograficamente, mas com trocas comerciais significativas, teriam mais peso na matriz do que áreas próximas geograficamente, mas com poucas trocas comerciais.

As linhas da matriz W podem ser padronizadas, em alguns casos, a fim de auxiliar a derivação de fórmulas e as propriedades estatísticas vinculadas ao estudo. Entretanto, esta padronização é realizada de forma que, as medidas de proximidades espaciais, ou os pesos ou os graus de conectividade (w_{ij}) associados a uma determinada área i , somem 1.

Apesar das diversas formas de construção da matriz de vizinhança, a binária ainda é a mais utilizada. Portanto, alguns trabalhos científicos mostraram que o uso da matriz W com indicadores binários é, na maioria das vezes, inadequado para estudos que envolvam transportes. Silva (2006) construiu um modelo de regressão com efeitos espaciais a fim de identificar e analisar os fatores medianeiros na distribuição espacial da frota de veículos rodoviários de carga.

Anteriormente à construção do modelo, tem-se a necessidade de verificar a dependência espacial através dos indicadores, que são calculados fazendo uso da matriz W . Silva (2006) encontrou sinais de dependência espacial em relação à quantidade de carrocerias de certo tipo. Como é necessário quantificar e validar tais indícios, utilizou-se um dos índices de dependência espacial (com uma matriz W com indicadores binários). O valor calculado foi significativo, mas considerado baixo (22%), o que levou a pensar que esta estrutura espacial com indicadores binários não seria a mais adequada para o estudo e que, provavelmente, a matriz W , com indicadores e com a distância rodoviária entre os municípios, teria um resultado mais relevante.

Outro trabalho, realizado por Silva *et al.* (2007), foi referente à atração de viagens no modo ônibus, por zona de tráfego na cidade de Manaus. A matriz W , considerada neste estudo, foi o tempo médio de percurso de uma zona de tráfego a outra. Este tempo médio, além de ser o fator que determina a “vizinhança” de i e de encontrar um valor significativo para dependência espacial, leva em consideração as condições da via, possíveis congestionamentos, entre outras variáveis que resultam em um valor bastante significativo para este tipo de estudo. É importante lembrar que foram feitos testes, anteriormente à utilização da matriz de vizinhança, considerando o tempo médio, e que não se encontrou

dependência espacial utilizando-se a matriz de vizinhança com indicadores binários para o cálculo dos índices de autocorrelação.

Por fim, é essencial ter o conhecimento que apenas a matriz W não dispõe de informações que comprovem a dependência espacial. Para isso, é necessária a utilização de índices e testes relativos para este tipo de método, como será apresentado a seguir.

3.3.2. Dependência Espacial

O conceito mais importante para a análise e a compreensão de fenômenos espaciais é de dependência espacial. Peña (2006) define dependência espacial como a primeira idéia deste método, em que uma variável qualquer tende a assumir valores similares em unidades geograficamente próximas dando lugar ao surgimento de *clusters*, como, por exemplo, áreas urbanas pobres tenderiam a estar cercadas de outras áreas pobres.

Outra idéia de dependência espacial é que uma determinada variável assume valores similares não somente em unidade geograficamente próximas, mas também em relação à conectividade (trocas comerciais e tempo médio de viagem) das áreas. Então, o preceito básico da dependência espacial é a Primeira Lei da Geografia, proposta por Waldo Tobler (1979), que diz “*tudo está relacionado a tudo, mas as coisas próximas estão mais relacionadas que as distantes*”, tal que a palavra “próxima” significa o padrão espacial geográfico ou conectivo.

A formulação matemática da dependência espacial é dada pela autocorrelação espacial, termo originado do conceito estatístico de correlação (medida que descreve, por meio de um número, o grau de associação entre duas variáveis). Segundo Lopes (2005), a preposição “auto” aponta para o fato de que a medida de correlação é executada com a mesma variável, encontradas em locais distintos do espaço.

Assim como a medida de correlação, a autocorrelação pode ser positiva ou negativa, em que a autocorrelação positiva está relacionada com a Primeira Lei da Geografia. Silva (2006) ilustra a autocorrelação conforme as figuras a seguir:

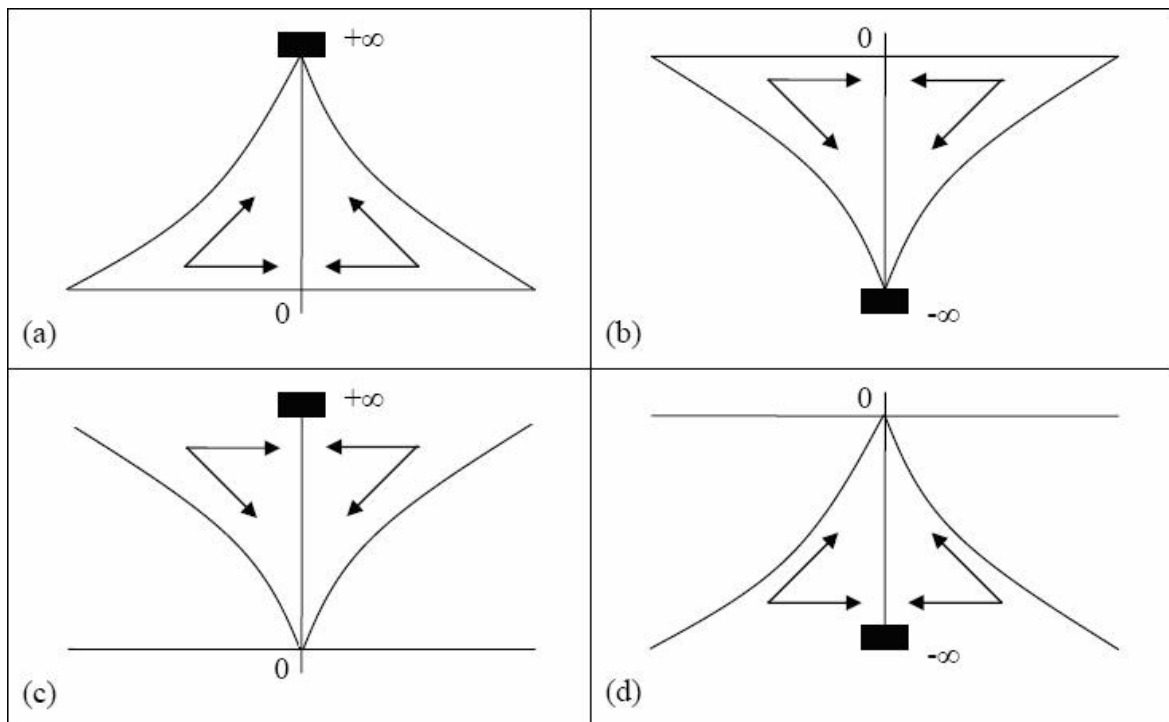


Figura 3.3: Representação de autocorrelação positiva (a) e (b) e negativa (c) e (d)
 Fonte: Silva (2006)

De acordo com as Figuras 3.3(a) (retângulo em negrito representa um alto valor) e 3.3(b) (retângulo em negrito representa um baixo valor), verifica-se que quanto mais próximo do retângulo, mais parecido com ele será. A recíproca é verdadeira para o caso contrário. Quando se visualiza as Figuras 3.3(c) e 3.3(d), considerando o mesmo retângulo, observa-se que quanto mais próximo do retângulo, mais diferente dele será. Lembrando que a recíproca é verdadeira para o caso contrário.

Existem índices para se detectar a presença da dependência espacial em um conjunto de dados. O mais conhecido, e o que será utilizado neste trabalho, é o índice *I* de Moran. De acordo com Silva (2006), existem duas formas para se medir a dependência espacial: através de índices globais (detêm tal dependência em apenas um fator) e índices locais (verificam a dependência espacial em cada área estudada).

3.3.2.1. Índice global *I* de Moran

Os índices de autocorrelação global objetivam estimar o quanto um valor observado em uma área depende dessa mesma variável nas localizações vizinhas (próximas), resultando um valor único de associação espacial para todo o conjunto de dados (Henrique, 2004). O

índice global I de Moran, introduzido na década de 50 pelo estatístico australiano Patrick Moran, é o mais antigo teste para determinação da autocorrelação espacial global. Como citado anteriormente, ele foi desenvolvido a partir da idéia do coeficiente de Pearson, com a diferença de: medir a correlação entre os pares y_i e y_j e por ser ponderado pela proximidade espacial.

A forma de calcular o índice global I de Moran é dada pela Equação 3.1, cujos valores podem variar entre -1 e +1. Quando os valores estão próximos de zero indicam a inexistência de autocorrelação espacial.

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\left(\sum_{i=1}^n (y_i - \bar{y})^2 \right) \sum_{i \neq j} w_{ij}} \quad (3.1)$$

onde,

I = índice Global de Moran;

n = quantidade de áreas;

w_{ij} = pesos ou graus de conectividade atribuídos conforme a relação topológica entre as áreas i e j ;

y_i = valor do atributo considerado na área i ;

y_j = valor do atributo considerado na área j ;

\bar{y} = valor médio do atributo em todas as áreas em estudo.

Padronizando-se a matriz W , de forma que a soma dos elementos de cada linha seja igual a 1, como citado anteriormente, e definindo-se uma variável auxiliar que represente os desvios em relação à média ($z_i = y_i - \bar{y}$ e $z_j = y_j - \bar{y}$), têm-se a seguinte equação (Câmara, 2000):

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} z_i z_j}{\sum_{i=1}^n z_i^2} \quad (3.2)$$

onde,

I = índice Global de Moran;

n = número de áreas;

z_i = diferença entre o valor do atributo no local i e a média de todos os atributos;

z_j = diferença entre o valor do atributo no local j e a média de todos os atributos;

w_{ij} = pesos ou graus de conectividade atribuídos conforme a relação topológica entre as áreas i e j .

3.3.2.2. Índice local I_i de Moran

Na caracterização de uma região como um todo, é necessário apenas um único valor de medida de associação espacial para o conjunto de dados. Mas quando se lida com um grande número de áreas, como os municípios do País que possuem o serviço de TRIP, é bastante provável que ocorram diferentes regimes de associação espacial e que apareçam máximos locais de autocorrelação espacial, em que a dependência espacial é ainda mais pronunciada.

Logo, para analisar os valores de associação espacial e examinar padrões com mais detalhes, utilizam-se os índices de autocorrelação local (LISA – *Local Indicators of Spatial Association*) que, segundo Câmara *et al.* (2004), produzem um valor específico para cada área, permitindo assim a identificação de agrupamentos, cuja formulação é apresentada na Equação 3.3:

$$I_i = \frac{\sum_{j=1}^n w_{ij} z_i z_j}{\sum_{i=1}^n z_i^2} \quad (3.3)$$

onde,

I_i = índice local de Moran;

z_i = diferença entre o valor do atributo no local i e a média de todos os atributos;

z_j = diferença entre o valor do atributo no local j e a média de todos os atributos;

w_{ij} = pesos ou graus de conectividade atribuídos conforme a relação topológica entre as áreas i e j .

Segundo Anselin (1995), LISA é qualquer estatística que satisfaz a duas condições: a primeira é que índices de autocorrelação local, para cada observação, fornecem uma indicação de agregações significativas em torno de certa observação. Já a segunda condição é que a soma destes índices de autocorrelação local, para todas as observações, é proporcional ao índice de autocorrelação global.

3.3.2.3. Análise Gráfica da Dependência Espacial

Como visto anteriormente, a quantificação da dependência espacial ocorre por meio de estatísticas globais e locais, com o uso de indicadores de autocorrelação espacial (Índice Global de Moran e Índice Local de Moran, por exemplo). Mas existem maneiras adicionais em que é possível visualizar esta dependência espacial. Um exemplo importante é o diagrama de espalhamento de Moran, que é a forma gráfica do índice global I de Moran e tem como principal objetivo analisar o comportamento da variabilidade espacial.

Esse diagrama de espalhamento de Moran é um gráfico bidimensional, construído com base nos valores normalizados (Z) e comparados com a média dos vizinhos (W_z), dividido em quatro quadrantes, conforme a Figura 3.4. Esses quadrantes indicam tanto os pontos de associação espacial positiva, em que os valores dos vizinhos são semelhantes ao valor do local (Q1 e Q3), como os pontos de associação espacial negativa, cujos vizinhos possuem valores distintos do valor local (Q2 e Q4).

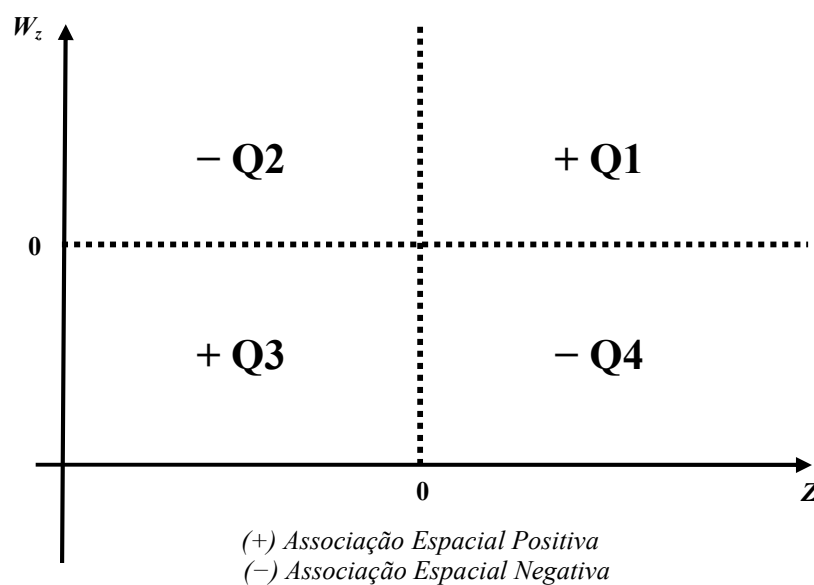


Figura 3.4: Gráfico de espalhamento de Moran

De acordo com a Figura 3.4, verifica-se que os pontos encontrados no quadrante Q1 indicarão que, para altos valores de Z , na média, existem altos valores de W_z , e serão chamados de alto-alto (*High-High*). Já os pontos do quadrante Q3 indicarão que, para baixos valores de Z , na média, existem baixos valores de W_z , e serão chamados de baixo-baixo (*Low-Low*). Os pontos localizados em Q2 e Q4 apontarão que, para baixos (ou altos) valores de Z , na média, existem, respectivamente, altos (ou baixos) valores de W_z . Os pontos em Q2 são chamados de baixo-alto (*Low-High*) e os pontos em Q4 de alto-baixo (*High-Low*) (Silva, 2006).

A análise do diagrama de Espalhamento de Moran permite a identificação de valores extremos, ou *outliers*, os quais não seguem o mesmo processo de dependência espacial que a maioria das outras observações. Quanto maior a quantidade de observações nos quadrantes Q1 e Q3, mais caracterizada será a dependência espacial na variável em estudo. Diferentemente, quando os quadrantes Q2 e Q4 têm a maioria dos pontos, significará que a estrutura espacial não influencia no valor da variável dependente. A definição exata das observações em relação ao quadrante será o ponto decisivo para a utilização da regressão com efeitos espaciais locais, pois esse tipo de modelo é recomendado apenas para as áreas (no caso da dissertação, serão os municípios brasileiros) definidas como *High-High* e *Low-Low*.

O diagrama pode ser visto também na forma de mapa temático, denominado *Box Map* (*Moran Scatterplot*), conforme a Figura 3.5, em que o estudo busca, através de ferramentas de estatística espacial, estimar o número de viagens atraídas por zona de tráfego na cidade de Manaus.

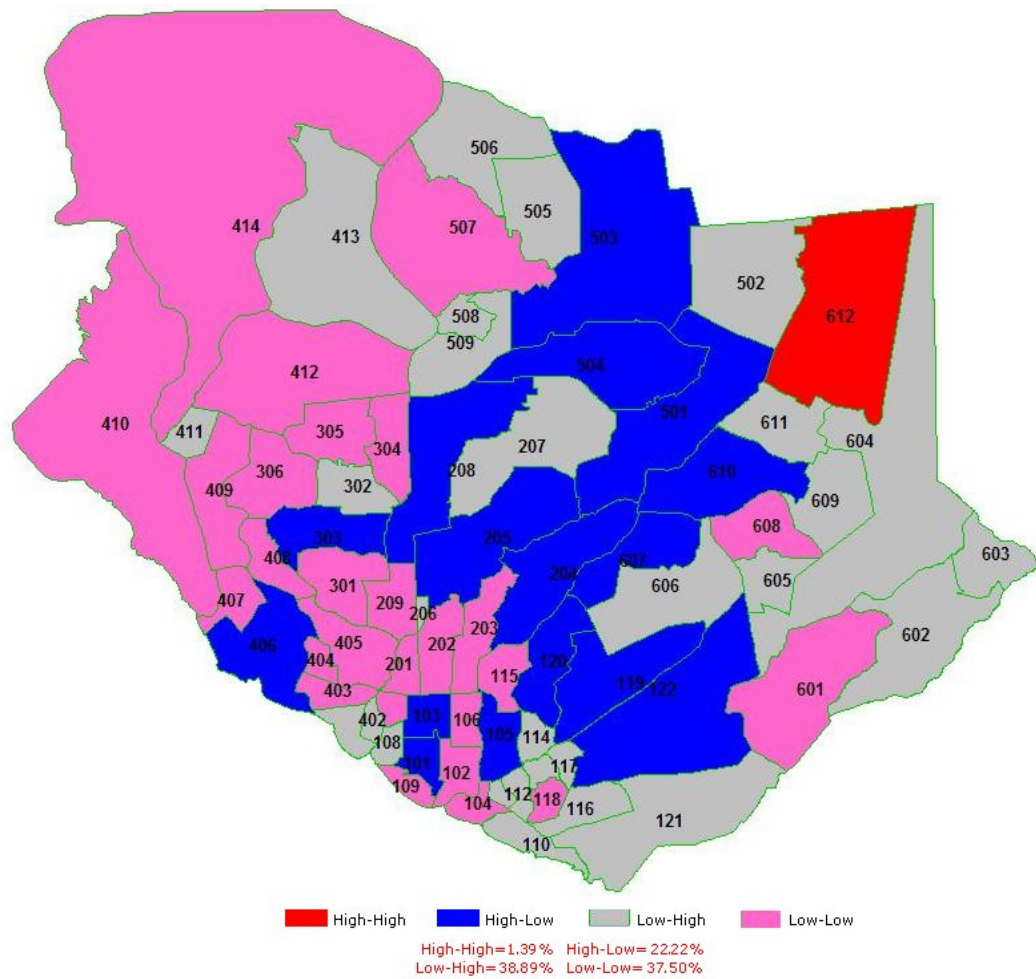


Figura 3.5: Exemplo de *Box Map*

No caso da visualização e indicação das regiões que apresentam correlação local significativamente diferente do resto dos dados, o mapa temático utilizado é o *LISA Map*. Segundo Anselin (1995), na geração deste mapa, os valores do índice de autocorrelação local (I_i de *Moran*) são classificados em três níveis de significância (quando existem os três níveis no resultado): 95%, 99% e 99,9%, além dos não significantes, conforme a Figura 3.6.

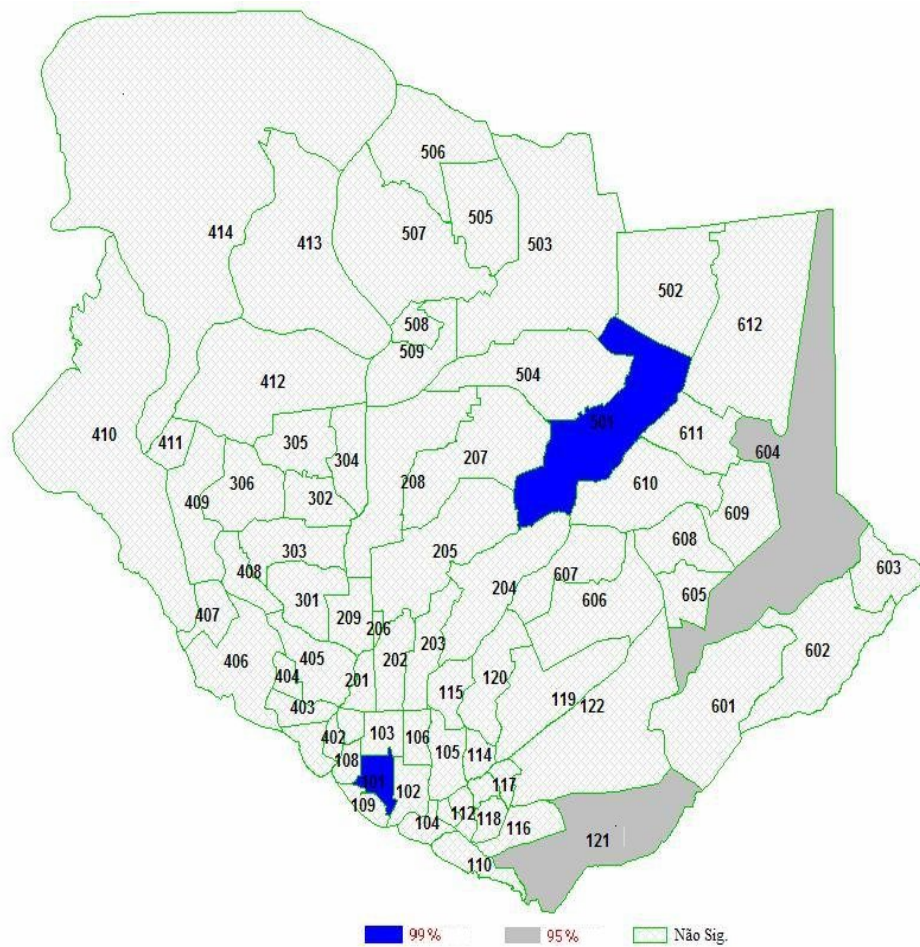


Figura 3.6: Exemplo de Lisa Map

Os resultados do *LISA Map* fornecem, conjuntamente com o *Box Map*, outro tipo de mapa, denominado *Moran Map*. Ele é construído apenas com os valores significativos do I_i de *Moran*, mas classificados conforme os quadrantes do diagrama de espalhamento de Moran, como mostra a Figura 3.7.

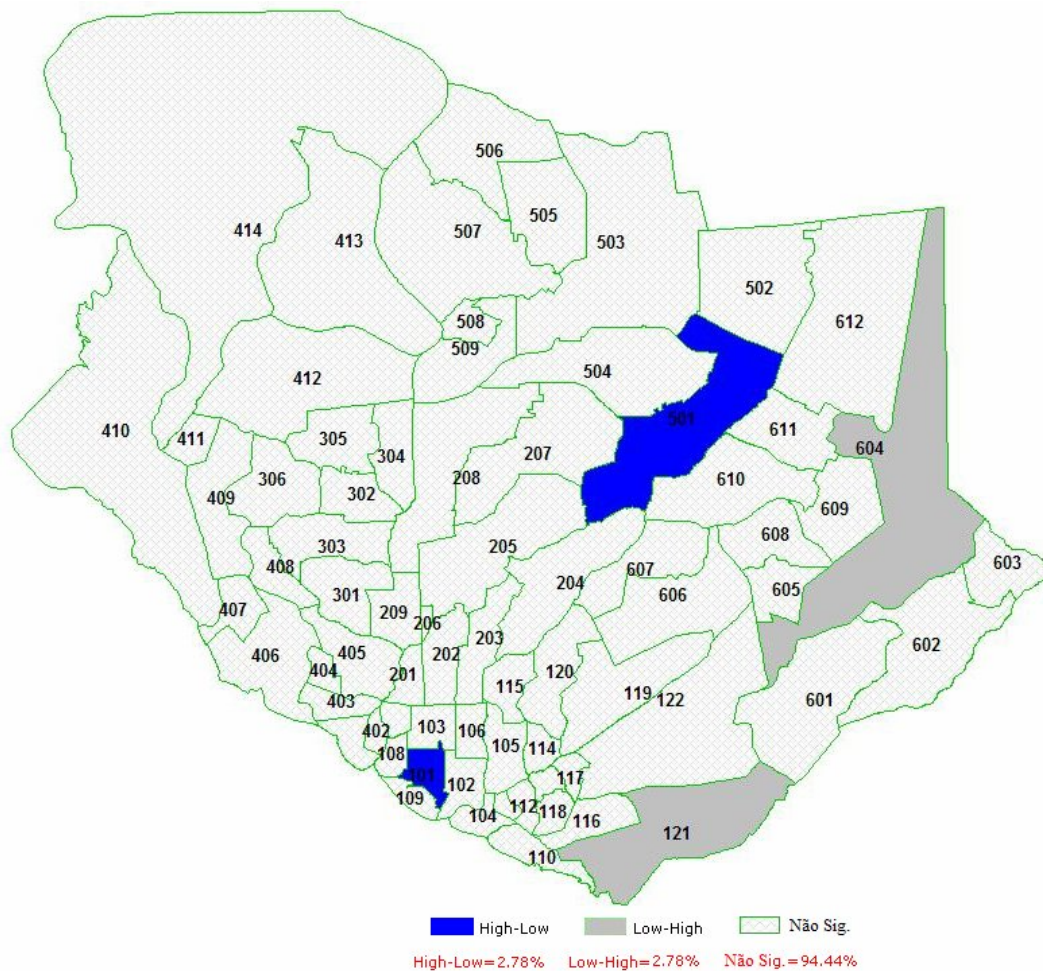


Figura 3.7: Exemplo de *Moran Map*

Segundo Lopes (2005), todos esses índices e mapas apresentados até agora fazem parte das técnicas *ESDA* (*Exploratory Spatial Data Analysis*), cuja função é auxiliar a identificação de objetos com valores altos e baixos, de áreas de transição e casos atípicos. Logo, após a identificação da dependência espacial utilizando técnicas *ESDA*, o passo seguinte é a utilização de modelos de regressão com efeitos espaciais.

3.4. MODELOS DE REGRESSÃO COM EFEITOS ESPACIAIS

O fato de se constatar a dependência espacial, por meio da utilização das técnicas *ESDA*, faz com que os modelos convencionais de regressão sejam invalidados por violarem pressupostos como independência espacial das variáveis e a homocedasticidade (Silva, 2006), além de superestimar alguns parâmetros destes modelos. Dessa forma, recorre-se a modelos alternativos que incorporem esses fenômenos à estrutura da função, os quais são

conhecidos como modelos de regressão com efeitos espaciais (ou apenas regressão espacial).

Esses modelos, assim como os modelos de regressão linear convencional, devem atender os três pressupostos básicos citados no início deste capítulo. Por incorporarem uma estrutura espacial, os modelos de regressão espacial eliminam, na maioria das vezes, os problemas de heterocedasticidade, e a hipótese de observações não-correlacionadas não é verdadeira (Silva, 2006).

Neste contexto, esta seção apresentará os diferentes tipos de modelos de regressão espacial, desde um breve resumo dos modelos que possuem efeitos espaciais globais até um estudo mais aprofundado dos modelos que possuem efeitos espaciais locais.

3.4.1. Regressão com Efeitos Espaciais Globais

Os modelos de regressão com efeitos espaciais globais (ou apenas regressão espacial global) são os mais simples dos modelos que envolvem efeitos espaciais, pois segundo Câmara *et al.* (2004), nesses existem a possibilidade de capturar a estrutura de correlação espacial em um único parâmetro, que é adicionado ao modelo de regressão linear convencional.

Existem diversas formas de tratar a autocorrelação espacial global em um modelo de regressão, mas as mais utilizadas são: modelo espacial auto-regressivo misto (*Spatial Autoregressive – SAR*) e modelo do erro espacial (*Spatial Error Model – SEM*).

No SAR, a autocorrelação espacial é atribuída à variável dependente. Sendo assim, a idéia básica desse modelo é que ele possua, em sua composição, uma estrutura espacial. Dessa forma, esse modelo pode ser expresso conforme a Equação 3.4:

$$y_i = \rho \left(\sum_j w_{ij} y_j \right) + \sum_i x_i \beta_i + \varepsilon_i \quad (3.4)$$

onde,

y_i = variável dependente na área i ;

y_j = variável dependente na área j ;

ρ = coeficiente espacial auto-regressivo;

x_i = variáveis independentes;

β_i = parâmetros da regressão para serem estimados;

ε_i = erros aleatórios;

w_{ij} = pesos ou graus de conectividade atribuídos conforme a relação topológica entre as áreas i e j .

É importante salientar que um processo auto-regressivo é aquele em que a variável é explicada por ela mesma, ou seja, para esse caso de modelo espacial auto-regressivo, a variável dependente y é explicada por seus vizinhos (Silva, 2006). Isso significa que a demanda do TRIP (variável dependente) de um determinado município pode ser explicada pela demanda de outros municípios “próximos”. Esse modelo pode se apresentar, também, na forma matricial (formalmente), de acordo com a expressão a seguir:

$$Y = \rho WY + X\beta + \varepsilon \quad (3.5)$$

Quando a dependência espacial está presente no termo *erro aleatório* da regressão e os efeitos espaciais são importunos, ou seja, um fator que precisa ser eliminado, o modelo com parâmetros globais a ser utilizado será o SEM (Anselin, 1992). Então, usando a forma matricial para expressar um modelo de regressão, a dependência espacial nos erros é especificada da seguinte forma:

$$\begin{aligned} Y &= X\beta + \varepsilon, \\ \varepsilon &= \lambda W\varepsilon + \xi \end{aligned} \quad (3.6)$$

onde,

Y = vetor de dimensão $n \times 1$ (variável dependente);

X = matriz de dimensão $n \times k$ (variáveis independentes);

β = vetor de dimensão $k \times 1$ a ser estimado (parâmetros da regressão);

ε, ξ = vetores de dimensão $n \times 1$ (erros aleatórios);

λ = coeficiente espacial auto-regressivo;

$W\varepsilon$ = este produto expressa a dependência espacial em ε .

Existem outros modelos que utilizam os efeitos espaciais globais, como o modelo regressão espacial Durbin, os *probit* e *tobit*, e os baseados em um panorama Bayesiano. Esses modelos não serão aqui apresentados, mas podem ser estudados em LeSage (1999a), LeSage (1999b) e Silva (2006).

Em todos esses modelos de regressão espacial com efeitos globais, como citado anteriormente, os padrões de autocorrelação espacial existentes nos dados podem ser capturados em um único parâmetro. Ou seja, o princípio da utilização desses tipos de modelos é que o processo espacial subentendido aos dados analisados é estacionário. Câmara *et al.* (2004) afirmam que a estacionariedade nem sempre ocorre. Nesse contexto, desenvolveram-se modelos de regressão espacial com efeitos locais, que serão estudados no tópico a seguir.

3.4.2. Regressão com Efeitos Espaciais Locais

O processo espacial é estacionário quando as particularidades estatísticas da variável independem de sua localização absoluta. Em outras palavras, o valor esperado e a variância são constantes em qualquer sub-região ou pontos do espaço, e a covariância depende somente da localização relativa entre duas localidades quaisquer (Câmara *et al.*, 2004).

Então, quando as propriedades estatísticas citadas dependem de sua localização absoluta (não-estacionariedade), os coeficientes de regressão precisam refletir a heterogeneidade espacial. Com isso, um ajuste mais específico para cada sub-região determina uma melhor representatividade do processo como um todo. Dessa forma, são utilizados modelos de regressão com efeitos espaciais locais (Fotheringham *et al.*, 2002; Câmara *et al.*, 2004).

O primeiro modelo desse tipo, chamado de modelo espacial de expansão, foi introduzido em 1972 por Casetti e expresso da seguinte forma (LeSage, 1999a):

$$\begin{aligned} Y &= X\beta + \varepsilon, \\ \beta &= ZJ\beta_0 \end{aligned} \tag{3.7}$$

onde,

Y = vetor de dimensão $n \times 1$ (variável dependente);

X = matriz de dimensão $n \times nk$ (variáveis independentes);

β = vetor de dimensão $nk \times 1$ (parâmetros da regressão);
 β_0 = vetor com $2k$ parâmetros para serem estimados: β_y e β_x ;
 Z = matriz das coordenadas latitude-longitude das observações;
 J = matriz de dimensão $n \times k$;
 ε = vetor de dimensão $nk \times 1$ (erros aleatórios).

Esse modelo atrairá para si a heterogeneidade espacial, permitindo a variação da dependência relacionada, de modo que os conjuntos de observações “próximas”, medidos pelas coordenadas latitude-longitude, assumem semelhantes valores de parâmetro. Como existe uma variação nas localizações, a regressão relacionada é mudada para se adaptar a um ajuste linear localmente, através dos conjuntos de observações na vizinhança (LeSage, 1999a).

Mas, atualmente, segundo Câmara *et al.* (2004), existem duas possibilidades de se trabalhar com modelos de regressão com efeitos espaciais locais: modelar a tendência espacial de forma contínua, com parâmetros variantes no espaço, ou modelar a variação espacial de forma discreta, ao dividir o espaço em sub-regiões estacionárias, chamadas de regimes espaciais, conforme a Tabela 3.1, que apresenta um resumo destas duas formas:

Tabela 3.1: Tipos de modelagem de regressão com efeitos espaciais locais

	Modelos	Onde	Observações
Modelagem discreta	$Y_1 = X_1\beta_1 + \varepsilon_1, ind = 1$ $Y_2 = X_2\beta_2 + \varepsilon_2, ind = 2$	Y_1 e Y_2 = variáveis dependentes; X_1 e X_2 = variáveis independentes; β_1 e β_2 = parâmetros do modelo; ε_1 e ε_2 = erros aleatórios.	<ul style="list-style-type: none"> • Identifica regimes espaciais; • Divide em sub-regiões criando indicadores; • Estima parâmetros do modelo conjuntamente.
Modelagem contínua	$Y(\pi) = \beta(\pi)X + \varepsilon$	$Y(\pi)$ = variável dependente representando o processo em π ; $\beta(\pi)$ = indica que os parâmetros são estimados em π ; X = variável independente; ε = erro aleatório; π = vetor que indica as coordenadas espaciais (ponto ou sub-região considerada).	<ul style="list-style-type: none"> • Ajuste de um modelo a cada ponto observado; • Pondera outras observações em função da distância.

A modelagem da variação espacial de forma discreta também é chamada de modelo de regressão com regimes espaciais. Segundo Lopes (2005), a idéia dos regimes espaciais é dividir uma determinada região de estudo em sub-regiões, cada uma com seu padrão espacial próprio, e realizar regressões para cada uma delas. A Tabela 3.1 mostra as observações classificadas em dois subconjuntos (ou dois modelos). Mas dependendo de como foram determinados os regimes espaciais, por meio de técnicas para a verificação da autocorrelação local, será possível a classificação em três ou mais subconjuntos.

Lopes (2005) afirma que, apesar de cada regime espacial conter os seus próprios valores de coeficientes, estes valores são estimados contiguamente, isto é, todo o conjunto de observações disponível é utilizado no modelo.

Modelos de regressão com regimes espaciais tendem a apresentar melhores resultados que modelos de regressão linear convencional ou de regressão espacial global, quando se trabalha com dados socioeconômicos característicos dos municípios brasileiros. Isso ocorre, segundo Câmara *et al.* (2004), em função da alta desigualdade social no Brasil, que ocasiona descontinuidades repentinas nos fenômenos estudados.

Outra forma de modelar a tendência espacial é de forma contínua, em que a idéia é ajustar um modelo de regressão linear a cada ponto observado, ponderando todas as demais observações como função da distância a esse ponto (Câmara *et al.*, 2004). Assim, com esta técnica, é possível explorar se a associação entre demanda de transporte rodoviário e suas variáveis explicativas é constante em toda região ou se é possível identificar variações entre sub-regiões. Essa técnica é denominada na literatura como regressão geograficamente ponderada - RGP (*Geographically Weighed Regression* – GWR).

De acordo com Foody (2003), a característica essencial da RGP é que permite que os parâmetros estimados variem localmente. Assim, cada ponto ou sub-região considerada terá seus próprios coeficientes de ajuste. A estimação desses parâmetros, segundo Brunson *et al.* (1998), para o modelo da Tabela 3.1 (modelagem contínua), é dada por:

$$\hat{\beta}(\pi) = (X^T W(\pi) X)^{-1} X^T W(\pi) Y \quad (3.8)$$

onde,

$\hat{\beta}(\pi)$ = estimador para $\beta(\pi)$;

X = matriz das variáveis independentes; e

$$W(\pi) = \begin{pmatrix} \alpha_{\pi 1} & 0 & \dots & 0 \\ 0 & \alpha_{\pi 2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_{\pi N} \end{pmatrix} \quad (3.9)$$

é uma matriz cujos elementos diagonais correspondem aos pesos (α) ao se calibrar uma RGP em torno de uma sub-região ou um ponto π ; e N é o número de observações.

Devido à variação dos pesos na RGP, de acordo com a localização de π , uma consideração inicial e importante: o *círculo de inclusão* de observações em torno de π (Brunsdon *et al.*, 1998). Esse círculo (ou quadrado) de inclusão será uma área onde se utilizará um determinado modelo de regressão. Se a área for muito grande, então os dados inclusos em cada parâmetro estimado $\beta(\pi)$ compreenderão a maioria da área em estudo. Em outras palavras, as estimativas, para cada caso, tornar-se-ão muito semelhantes aos resultados obtidos utilizando técnicas de regressão espacial global. Quando é muito pequena, poucas observações estarão inclusas na calibração do modelo, o que conduzirá às estimativas do $\beta(\pi)$ com erros padrões muito grandes.

Segundo Fotheringham *et al.* (2002), para minimizar esses tipos de erros no modelo, além de se combater determinados problemas de descontinuidades dos pesos, um caminho é especificar $\alpha_{\pi j}$ como uma função contínua de $d_{\pi j}$ (distância entre a localização das observações π e j), em que a ponderação é dada por:

$$\alpha_{\pi j} = \exp\left[\frac{-d_{\pi j}^2}{2b^2}\right] \quad (3.10)$$

onde,

b = parâmetro que regula o grau de suavidade da função e estimado conforme as técnicas de mínimos quadrados.

A modelagem de regressão com efeitos espaciais locais é uma abordagem recente, mas que aos poucos vem sendo utilizada nas mais distintas áreas. Brunson *et al.* (1998) utilizaram-na nos dados sobre uma determinada variável atribuída a doenças, do censo de 1991 no Reino Unido. Neste trabalho, os autores supunham que a aplicação desta técnica permitiria investigar se quaisquer relações existentes entre essa variável, relacionada ao tipo de doença e variáveis quantitativas (sociais e econômicas), seriam estáveis em todo o espaço, ou ainda, se mudariam para refletir as características de diferentes localidades na área estudada.

Huang e Leung (2002) mostraram que o modelo de RGP que estima parâmetros “locais” é significativamente melhor do que modelo de regressão convencional, que apenas produz “média” e estima parâmetros globais. Os autores aplicaram, em uma província chinesa chamada Jiangsu, as duas técnicas para analisar a relação entre o nível de industrialização regional e vários fatores, como PIB per capita, porcentagem de trabalhadores urbanos na população total, saída em valor bruto da produção da indústria e da agricultura, entre outros. Nesse tipo de trabalho é possível verificar o local e quanto que uma determinada variável independente (fatores) poderá influenciar na variável dependente (industrialização regional).

Além desses trabalhos, Foody (2003) mostrou o potencial da ferramenta no estudo da relação do índice de vegetação por diferença normalizada e dos alagamentos ao longo do norte da África e o Oriente Médio, durante 8 anos. Enquanto Wang *et al.* (2005) aplicaram a regressão com efeitos espaciais locais para obtenção de uma produção primária líquida, que inclui a não-estacionariedade espacial na estimação dos parâmetros para os ecossistemas florestais da China.

3.5. TÓPICOS CONCLUSIVOS

- Diversas são as técnicas capazes de prever as viagens pelos diferentes modos e a quantidade (demanda) de passageiros. Devido à relação entre espaço e transporte, as técnicas mais empregadas são as que utilizam a Análise Espacial (AE).
- A Estatística Espacial, que engloba o ferramental matemático da análise exploratória e confirmatória, auxilia o planejador e o operador em transporte a

estabelecer critérios quantitativos de agrupamentos ou dispersão de dados espaciais, a obter o grau de dependência espacial entre as observações, entre outros.

- A dependência espacial é a correlação da mesma variável, que indica o grau de influência que uma unidade espacial tem sobre seus vizinhos. Já a heterogeneidade espacial se refere à variação no relacionamento excedente ao espaço geográfico.
- A matriz de proximidade espacial (W) é responsável por representar a estrutura espacial. Apesar de não dispor de informações que provem a dependência espacial, a matriz W é essencial para o cálculo dos indicadores de autocorrelação espacial.
- Alguns trabalhos científicos mostraram que o uso da matriz de proximidade espacial com indicadores binários é, na maioria das vezes, inadequado para estudos que envolvam transportes. Isso mostra que uma definição adequada para a matriz W conduz a resultados mais confiáveis.
- Para estimar o quanto um valor observado em uma área depende dessa mesma variável nas localizações vizinhas, são utilizados índices de autocorrelação. O índice de autocorrelação global é aplicado na caracterização da área como um todo, e o índice de autocorrelação local, para examinar padrões em grande número de áreas com mais detalhes.
- Os modelos de regressão com efeitos espaciais globais são os mais simples dos modelos que envolvem efeitos espaciais e têm a possibilidade de capturar a estrutura espacial em um único parâmetro.
- Quando o processo espacial é não-estacionário, os coeficientes de regressão precisam refletir a heterogeneidade espacial e, necessariamente, deve-se utilizar modelos de regressão espacial com efeitos locais.
- A modelagem, utilizando regressão com efeitos espaciais locais, poderá representar melhor a quantidade de passageiros que se deslocam pelo transporte rodoviário, pois, além do poder de correção de modelos mal especificados, tem um ajuste mais específico para cada sub-região.

- Vale ressaltar que o modelo que será aplicado nesta dissertação utilizará dois dos três estágios citados por Tarrius (*apud* Vasconcellos, 2001): a migração regional, em que se observa os passageiros de viagens interestaduais, e a mobilidade residencial no município, na qual os passageiros de viagens interestaduais utilizam o serviço intermunicipal para chegar ao seu destino final.

- Uma forma de modelagem da regressão com efeitos espaciais locais é a forma discreta, em que se divide uma determinada região de estudo em sub-regiões, de acordo com o padrão espacial de cada uma, e diferentes modelos de regressão são desenvolvidos para cada sub-região.

- Quando se ajusta um modelo de regressão a cada ponto observado, ponderando todas as demais observações como função da distância a esse ponto, utiliza-se a forma contínua de modelagem, chamada regressão geograficamente ponderada (RGP).

4. METODOLOGIA PARA A CONSTRUÇÃO DE MODELOS DE REGRESSÃO COM EFEITOS ESPACIAIS LOCAIS

4.1. APRESENTAÇÃO

A necessidade ou o desejo de movimentação de pessoas e cargas, entre diferentes locais, geram uma demanda por transporte, conforme visto no Capítulo 2. A análise da demanda por transporte – que estuda a diversidade dos padrões da interação socioeconômica e a complexidade resultante na evolução da necessidade por esse serviço – é ponto fundamental para a existência do transporte rodoviário interestadual de passageiros (TRIP).

Dessa forma, modelos que visam o cálculo da quantidade de passageiros são de suma importância ao planejamento e à operação de linhas interestaduais, pois, segundo Ramos *et al.*(2004), medidas que descrevam a mobilidade de indivíduos permitem avaliar o nível de atividade e a dinâmica de um determinado território. Mas, em muitos casos, estes modelos representam uma previsão não significativa, por não considerarem influências de variáveis espaciais.

Assim, devido à necessidade de modelos de demanda regionais para adaptação da oferta do TRIP, o presente capítulo busca desenvolver um método para a construção de modelos de regressão com efeitos espaciais locais, com intuito de estimar a demanda de passageiros deste serviço.

4.2. ESTRUTURA DO MÉTODO PROPOSTO

Diversas são as metodologias para a construção de modelos de regressão convencional. Não existe muita diferença na construção de modelos de regressão com efeitos espaciais, pois esses seguem os mesmos pressupostos e aproximadamente os mesmos conceitos.

O diferencial da regressão com efeitos espaciais é a etapa em que se verifica a dependência espacial do fenômeno em estudo, na qual também se verifica a não-estacionariedade (para modelos locais). Dessa forma, será discriminada uma estrutura metodológica (Figura 4.1) por meio de 6 etapas:

1. descrição e definição da área de estudo;
2. identificação das variáveis;
3. análise da qualidade do banco de dados disponível;
4. seleção das variáveis independentes;
5. análise espacial e verificação da dependência e do processo espacial e
6. construção do modelo de regressão com efeitos espaciais locais.

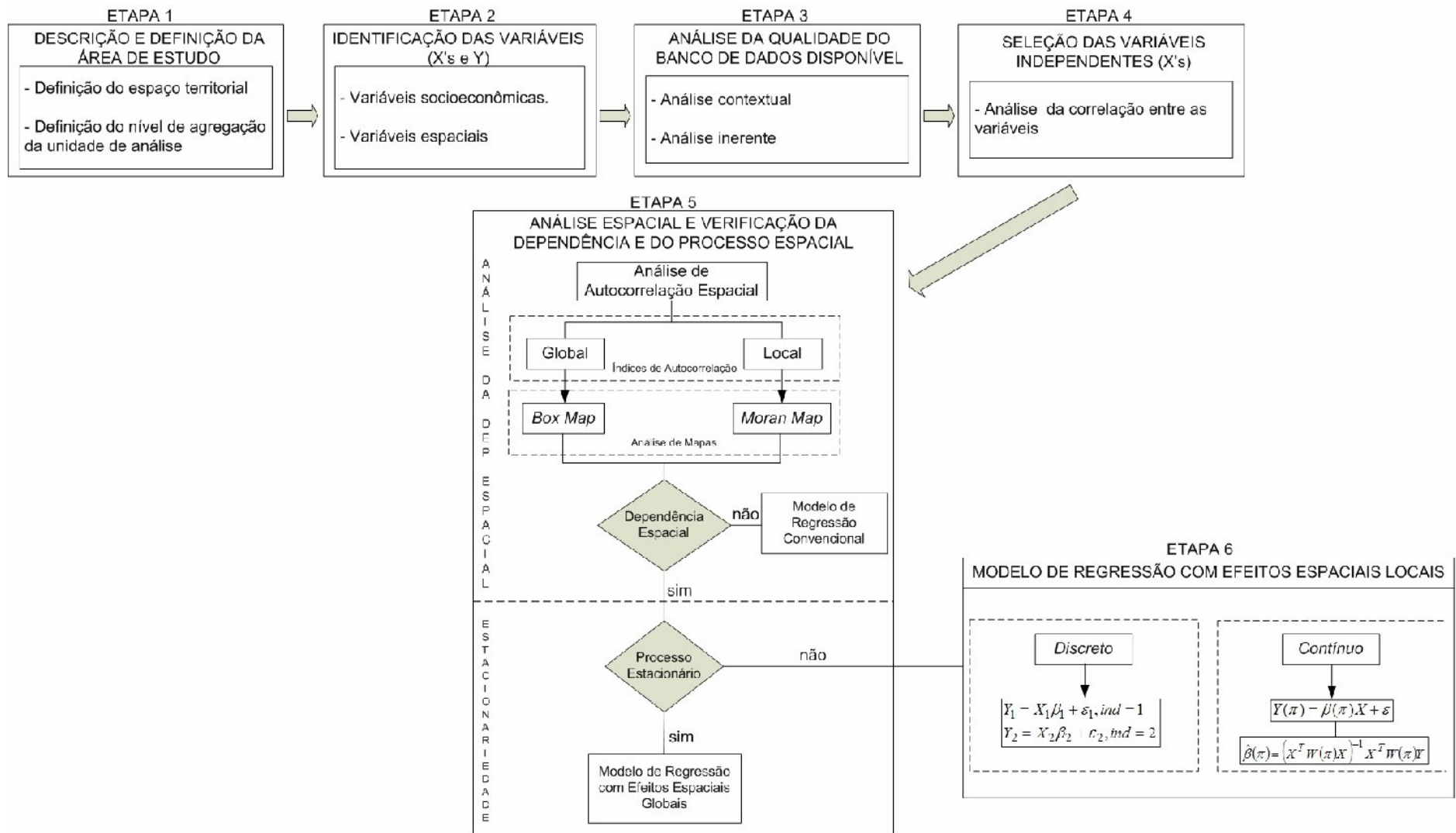


Figura 4.1: Estrutura metodológica para o cálculo da quantidade de passageiros do TRIP utilizando regressão com efeitos espaciais locais.

4.3. ETAPAS DO MÉTODO PROPOSTO

São apresentadas a seguir as sete etapas de forma mais detalhada.

Etapa 1: Descrição e definição da área de estudo

A identificação e a demarcação da área de estudo é um dos pontos importantes para o início do desenvolvimento de diversos modelos. Devido à falta de recursos e de tempo, muitos trabalhos são executados em uma pequena parte para resolver um problema como o todo.

Visto que a dissertação tem como objetivo desenvolver um modelo para o cálculo da variável dependente *demandas de passageiros nas linhas do transporte rodoviário interestadual*, o nível de agregação da área de estudo será municipal, atentando às variáveis que possam influenciar nesta demanda.

Esta etapa ainda terá como objetivo realizar uma análise descritiva da variável dependente e verificar a distribuição desta demanda dentro da área de estudo. Além de informações numéricas sobre o comportamento da variável dependente na área de estudo, utilizará mapas temáticos para a visualização desse fenômeno. A construção desses, através de tecnologias SIG, é uma técnica que vem sendo explorada por diversas áreas para análise, avaliação e seleção de alternativas à tomada de decisões. Este tipo de ferramenta será útil para compreender, de forma mais simples, a distribuição de viagens e a movimentação de passageiros dentro das rodovias brasileiras.

Apesar de ser símbolo das ferramentas de análise espacial, é importante acentuar que, mapas temáticos não são ferramentas de análise confirmatória, mas apenas exploratória, pois a construção destes apenas sugere algo (por exemplo, a existência de não-estacionariedade), sem nenhum poder de estimar parâmetros e de realizar testes estatísticos.

Etapa 2: Identificação das variáveis

Para o desenvolvimento de modelos econométricos, existe a necessidade de identificar variáveis que possibilitam sua representação. No caso da modelagem da demanda do TRIP,

segundo Carneiro (2005), as variáveis podem ser classificadas em socioeconômicas e espaciais, de acordo com a Equação 4.1:

$$D_{ij} = f(VS_{ij}; VE_{ij}) \quad (4.1)$$

onde,

D_{ij} = demanda prevista entre os municípios de origem i e destino j ;

VS_{ij} = variáveis socioeconômicas dos municípios i e j ;

VE_{ij} = variáveis espaciais dos municípios i e j .

As variáveis socioeconômicas são atributos que buscam definir uma população em relação a sua qualidade de vida. Já as variáveis espaciais (ou dados espaciais), de acordo com o tópico 3.2, são caracterizadas no espaço em função de um determinado sistema de coordenadas absolutas ou relativas. Com auxílio desses dois conjuntos de variáveis, pode-se compreender diversos fenômenos de uma determinada região, inclusive os relacionados a transportes.

Essas variáveis serão chamadas de independentes ou explicativas e têm como principal objetivo buscar explicação lógica para a variável dependente (quantidade de passageiros dos pares de O/D do transporte rodoviário interestadual). Elas podem ser obtidas por meio de pesquisas com aplicação de questionário ou através de instituições confiáveis que já tenham feito esse tipo de trabalho. É importante salientar que a aplicação de questionários é um método eficiente, mas são necessários tempo e recursos. Por isso, muitas vezes deixa-se de lado esse meio e procura-se instituições para adquirir estas informações.

Existem casos em que as variáveis independentes relevantes para o desenvolvimento do modelo não são do ano desejado. Assim, se utilizadas, acarretarão problemas na estimação dos parâmetros da regressão. Dessa forma, se possível, deve-se disponibilizá-las por meio de estimação utilizando anos anteriores.

Vale ressaltar que a função de selecionar as variáveis independentes será de responsabilidade da etapa 4, pois na etapa 2 apenas serão identificadas as variáveis que possam representar a quantidade de passageiros em viagens interestaduais pelo modo ônibus.

Etapa 3: Análise da qualidade do banco de dados disponível

Nesta fase, além de se verificar a disponibilidade dos dados que serão necessários, confere-se como esses estão organizados, realiza-se uma análise da qualidade e, se necessário, algumas possíveis estimativas ou imputações.

Como visto na etapa 2, diversas variáveis ou indicadores para o uso em trabalhos acadêmicos são adquiridos através de fontes conceituadas, visto que efetuar uma pesquisa para adquirir informação é um método com custos altos e que demanda bastante tempo. Com isso, esta dissertação buscará fontes confiáveis para o desenvolvimento do modelo, averiguando a veracidade destas informações.

Uma das formas mais eficiente de se verificar a veracidade destas informações é analisando a qualidade dos dados e das informações. Existem diversos indicadores que são utilizados para analisar a qualidade, em que Lee *et al* (2001) classificam-nos em quatro grupos distintos: Qualidade Inerente, Contextual, de Representação e Acessibilidade. Segundo Correia (2004), os indicadores mais eficientes são os de qualidade contextual e inerente, considerando que estão completamente relacionados à informação e aos dados, e não com o seu modo de armazenamento.

De forma geral, a qualidade contextual é relativa ao que deve ser considerado dentro do contexto da utilização da informação, que deve ser pertinente, adequada, oportuna e completa. Já a qualidade inerente refere-se aos valores que os dados ou as informações possuem ou podem assumir.

A qualidade contextual é dividida em diferentes indicadores para análise dos atributos dos dados ou das informações. Um dos indicadores é chamado de *utilidade*, que tem como objetivo localizar, em um banco de dados, variáveis ou atributos que não serão utilizados para um determinado estudo ou até mesmo dados que podem prejudicar a análise.

Outro indicador é o de *integralidade* que, segundo Correia (2004), é o mais importante indicador de qualidade contextual, pois ele basicamente analisa a existência ou a omissão de dados ou informação em um banco de dados.

O indicador de *atualidade* ou *temporalidade* é um indicador relacionado à periodicidade da informação. Esse tipo de indicador será importante para o trabalho realizado nesta dissertação, já que a ferramenta envolve um modelo estatístico. Como visto na etapa 2, variáveis podem ser estimadas, caso não seja encontrado o ano apropriado para o desenvolvimento do modelo.

Como a qualidade contextual, a inerente também é dividida em diferentes indicadores para a análise da informação. O indicador *exatidão* é o que pode ser quantificado como a diferença entre o valor apresentado no banco de dados e o valor considerado aceitável ou verdadeiro. Quase sinônimo de exatidão, a *precisão* não se refere ao valor do dado e sim ao seu detalhamento, ou seja, é, essencialmente, uma medida que especifica o domínio de uma variável.

Outro indicador é a *unicidade*, por meio da qual se busca obter apenas uma informação para a variável em estudo, apesar de ocorrer a existência de variáveis semelhantes em fontes distintas possuindo valores diferentes.

Consistência é um indicador que busca nas informações a concordância entre a armazenagem dos dados e a estrutura do banco; a concordância conceitual da informação e a exatidão das características topológicas dos dados. Este indicador compreende a análise exploratória do banco de dados (média, mediana, máximo, mínimo e dados faltantes). Mais detalhes sobre métodos para análise da qualidade da informação ou de dados podem ser vistos em Correia (2004).

Etapa 4: Seleção das variáveis independentes

Depois da identificação das variáveis, da montagem e da análise do banco de dados, o próximo passo é a seleção das variáveis explicativas, ou independentes, do modelo. Nesta etapa, empregam-se ferramentas estatísticas para observar a relação entre estas variáveis independentes e a variável dependente da demanda de passageiros nas ligações interestaduais pelo modo ônibus, ou até mesmo nas relações lineares aproximadamente exatas entre si (multicolinearidade).

A multicolinearidade é um problema típico dos modelos de regressão que pode levar a consequências irremediáveis, como erros-padrão elevados e parâmetros impossíveis de serem estimados. Com isso, é necessário identificar a existência de uma possível relação linear entre as variáveis independentes, construindo uma matriz de correlação. O índice mais utilizado para o cálculo da correlação entre duas variáveis é o *coeficiente de correlação de Pearson*, definido conforme a equação 4.2:

$$r = \text{Corr}(X, Z) = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{z_i - \bar{z}}{s_z} \right) \quad (4.2)$$

onde,

x e z = variáveis independentes;

\bar{x} e \bar{z} = médias de x e de z , respectivamente;

s_x e s_z = desvios padrão de x e de z , respectivamente.

Antes de se identificar a relação linear entre as variáveis independentes, deve-se verificar a relação linear entre a variável dependente e as variáveis independentes, que podem ser examinadas através de um método simples, a construção de gráficos de dispersão ou também a matriz de correlação. Lembrando que a estatística R^2 (medida que calcula o quanto a variação total em uma determinada variável é explicada pela variabilidade da outra) também é um meio para se verificar a relação linear.

Etapa 5: Verificação da dependência espacial

Essa etapa é fundamental ao desenvolvimento do objetivo desta dissertação, pois é a fase em que se verifica a necessidade ou não da construção de um modelo de regressão com efeitos espaciais. Se a dependência espacial for constatada, então o melhor modelo a ser utilizado será o espacial; caso contrário, aconselha-se trabalhar com o convencional, pois além de apresentar os mesmos resultados, o processo na calibração e na construção do modelo será mais simples.

Para se verificar a hipótese de que o fenômeno estudado incorpora atributo espacial, tem-se a necessidade de calcular os indicadores de autocorrelação espacial. Para isso, devem-se

definir as matrizes de proximidade espacial (W), de acordo com as formas já apresentadas no capítulo 3.

No passo seguinte, serão calculados os índices de autocorrelação global que, além de detectar a presença de dependência espacial em toda área de estudo, quantifica o nível de autocorrelação espacial da variável dependente, demanda do TRIP. Pode-se também construir o diagrama de espalhamento de Moran ou *Box Map*, que são formas gráficas e de mapa, respectivamente, do índice global de Moran.

Por fim, trabalha-se com técnicas *LISA* (índice local de Moran e *Moran Map*), que identificam as unidades geográficas que possuem comportamento semelhante ou inverso, permitindo assim verificar se o processo espacial é estacionário ou não. Então, caso seja constatada a dependência espacial por meio de índices de autocorrelação global ou local e, além disso, o processo seja não-estacionário, é necessária a utilização de modelos de regressão espacial com efeitos locais.

Etapa 6: Construção do modelo de regressão com efeitos espaciais locais

Depois de verificar-se a dependência espacial e a não-estacionariedade do processo espacial, os próximos passos são a construção do modelo de regressão com efeitos espaciais locais e a verificação de sua adequabilidade. Como visto no tópico 3.4.2, existem duas possibilidades para se trabalhar com esse tipo de modelo: de forma discreta e de forma contínua.

Em conjunto com a construção do modelo de regressão espacial, devem-se verificar os pressupostos básicos para sua validação, como na análise de regressão convencional. Esses são erros normais, com média zero; erros com variância constante e erros não correlacionados. Este não poderá ser comprovado, dado que o único teste para demonstrá-los só é possível em séries temporais.

Esses pressupostos podem ser averiguados por meio de gráficos ou testes estatísticos específicos. Para confirmar a normalidade dos erros utiliza-se o *qq-plot* (gráfico que compara os *quantis* da distribuição normal com os *quantis* da distribuição empírica dos

dados) ou os testes como Shapiro-Wilk, em que o tamanho da amostras deve ser entre 3 e 50, ou como Kolmogorov-Smirnov, para amostras acima de 50. A heterocedasticidade (erros com variância não constante) também pode ser verificada através dos gráficos de resíduos entre as variáveis independentes, ou até mesmo, com maior confiabilidade, através do teste de White. Existem outros testes menos utilizados, como mostra Gujarati (2006).

Testes estatísticos também são utilizados para confirmar a existência do modelo (*teste F*), ou seja, se os parâmetros são significativos (parâmetros em conjunto são diferentes de zero) e, para indicar a significância de cada parâmetro do modelo (*teste t*). Todos esses testes só terão sentido se os pressupostos citados acima forem comprovados. Caso leve em conta, por exemplo, a heterocedasticidade, os testes *F* e *t* darão resultados inexatos e, com isso, os parâmetros que pareciam ser estatisticamente não significativos, na verdade, são significativos. Ou seja, uma variável que deveria estar no modelo não aparecerá.

O coeficiente de determinação (R^2), como visto anteriormente, é medida que indica o quanto os dados se ajustam à reta de regressão. Quanto mais próximo de 1 estiver o valor do R^2 , melhor o ajuste do modelo; e quanto mais próximo de zero, pior esse ajuste. É importante salientar que, na prática, é quase impossível uma correlação perfeita (em que o R^2 é igual a 1), pois existem diversos fatores que definem as relações entre variáveis em um fenômeno real.

Segundo Câmara (2004), as medidas R^2 são índices limitados e devem ser encaradas com certos cuidados quando se estuda modelos com efeitos espaciais. Então, para verificar o melhor modelo para a previsão e sua adequabilidade, pode-se utilizar o cálculo do Erro Relativo Médio (*ERM*), expresso pela Equação 4.3:

$$ERM = \frac{\sum_{i=1}^n \frac{|YE_i - YO_i|}{YO_i}}{n} \quad (4.3)$$

onde,

YE_i = número estimado pelo modelo;

YO_i = número observado (real);

n = número de observações.

No caso de modelos de regressão com efeitos espaciais locais, tanto no caso discreto como contínuo (RGP), será necessário o cálculo dos R^2 de cada sub-região. O *ERM* será utilizado para a comparação dos diferentes modelos construídos para o cálculo da demanda do TRIP.

4.4. TÓPICOS CONCLUSIVOS

- Esta proposta procura sistematizar o estudo da demanda de passageiros do transporte rodoviário interestadual levando em consideração as relações espaciais existentes entre cada par de O/D.
- Avalia se existe uma relação ou dependência espacial em cada par de O/D, e até mesmo a existência de sub-regiões na unidade federativa que foi destino da viagem (não-estacionariedade).
- Devido à estrutura contida na metodologia proposta e ao fato de a matriz O/D das viagens interestaduais está composta por diversos zeros, a regressão com efeitos espaciais locais da forma contínua não será o modelo mais adequado para a pesquisa.
- Por fim, a metodologia proposta não sugere como deve ser considerado os pares de O/D em que não exista dependência espacial ou até mesmo onde o fenômeno é estacionário.

5. ESTUDO DE CASO

5.1. APRESENTAÇÃO

O principal papel do Transporte Rodoviário Interestadual de Passageiros no Brasil é o deslocamento de pessoas, entre os Estados, através da sua malha rodoviária. De acordo com o capítulo 2, este deslocamento é realizado em um itinerário da linha, que é a rota (ligação de dois pontos terminais) a ser utilizada na execução do serviço. Nesse contexto, o estudo de caso se preocupará em desenvolver um modelo de demanda para determinar a demanda de passageiros entre as ligações do transporte rodoviário interestadual regulado e fiscalizado pela ANTT. Para isso, neste capítulo verificar-se-á a aplicabilidade da metodologia proposta no capítulo anterior, tendo como objetivo principal a aplicação da metodologia para alguns estados do Brasil. Dessa forma, o capítulo é apresentado em duas seções, passando etapa por etapa da metodologia proposta, e finalizando com alguns tópicos conclusivos.

5.2. APLICAÇÃO DA METODOLOGIA

Esta seção, por meio das etapas da metodologia proposta no capítulo 4, buscará desenvolver o estudo de caso, a fim de verificar a sua viabilidade. Nesse sentido, serão apresentadas todas as etapas de forma detalhada.

Etapa 1: Descrição e definição da área de estudo

O estudo da previsão da demanda de passageiros do transporte rodoviário interestadual pode ser desenvolvido tanto em linhas isoladas como em todas as linhas do território brasileiro. A intenção deste estudo era, a princípio, testar a viabilidade da proposta em diversas linhas, mas decidiu-se por estudar o par São Paulo/algumas cidades do Ceará e Natal/algumas cidades do Ceará, em função da disponibilidade dos dados para a realização do estudo e pelos motivos descritos nesta etapa.

Segundo ANTT (2007), em razão da extensão do país e de suas diferenças regionais, o universo da análise deve ser dividido por grupos (distâncias, movimentação total, renda, população, entre outros), conforme a decisão do pesquisador.

Desse modo, dois fatores devem ser levados em consideração: o primeiro é que o presente trabalho foi segmentado pelos passageiros que fizeram viagens interestaduais acima de 100 km entre os municípios (migração regional), onde o destino não é apenas uma cidade, e sim uma área de influência. Esta área significa uma sub-região com mesmo padrão espacial, conforme visto na seção 3.3.2. O segundo ponto é que o modelo não será aplicado para todo território brasileiro, mas para alguns estudos de caso específico em que existam uma área de influência dos municípios (as variáveis socioeconômicas e os passageiros das viagens intermunicipais).

Para a definição dessa sub-região para o cálculo da demanda de passageiros do transporte rodoviário interestadual, buscou-se a preponderância das possíveis variáveis independentes dentro desta área de influência (destino), observando a demanda de passageiros do transporte rodoviário intermunicipal (mobilidade residencial).

É importante apresentar um contexto dos deslocamentos de passageiros em viagens interestaduais realizadas por meio de ônibus regular, para melhor compreender o presente estudo de caso.

Este estudo tem como base a demanda de passageiros do mês de outubro de 2004. O total de passageiros transportados, em viagens produzidas neste período, no Brasil, utilizando o transporte rodoviário interestadual para viagens acima de 100 km, foi de 1.564.051. A região com maior demanda por este tipo de serviço foi a região sudeste, com quase 50% dos passageiros, como se pode observar na Tabela 5.1 e nas Figuras 5.1 e 5.2. Outro ponto a se destacar é que aproximadamente 23% do número total de passageiros saiu do Estado de São Paulo (por volta de 364 mil). Destaca-se o estado do Amapá que, segundo fontes da ANTT, nesse período, não apresentou passageiros transportados em viagens produzidas.

Tabela 5.1: Quantidade de passageiros transportados em viagens produzidas por Região e UF em outubro de 2004

	Quantidade de Passageiros	Percentual (%)
Brasil	1.564.051	100,00
Norte	32.059	2,05
Rondônia	3.649	0,23
Acre	1.325	0,08
Amazonas	1.598	0,10
Roraima	2.682	0,17
Pará	12.253	0,78
Amapá	0	0,00
Tocantins	10.552	0,67
Nordeste	235.596	15,06
Maranhão	19.500	1,25
Piauí	13.190	0,84
Ceará	30.990	1,98
Rio Grande do Norte	20.919	1,34
Paraíba	25.601	1,64
Pernambuco	37.931	2,43
Alagoas	16.503	1,06
Sergipe	20.669	1,32
Bahia	50.293	3,22
Sudeste	768.268	49,12
Minas Gerais	230.652	14,75
Espírito Santo	82.811	5,29
Rio de Janeiro	90.739	5,80
São Paulo	364.066	23,28
Sul	322.490	20,62
Paraná	184.739	11,81
Santa Catarina	84.607	5,41
Rio Grande do Sul	53.144	3,40
Centro-Oeste	205.638	13,15
Mato Grosso do Sul	21.324	1,36
Mato Grosso	26.717	1,71
Goiás	102.079	6,53
Distrito Federal	55.518	3,55

Fonte: SISDAP/ANTT, 2004

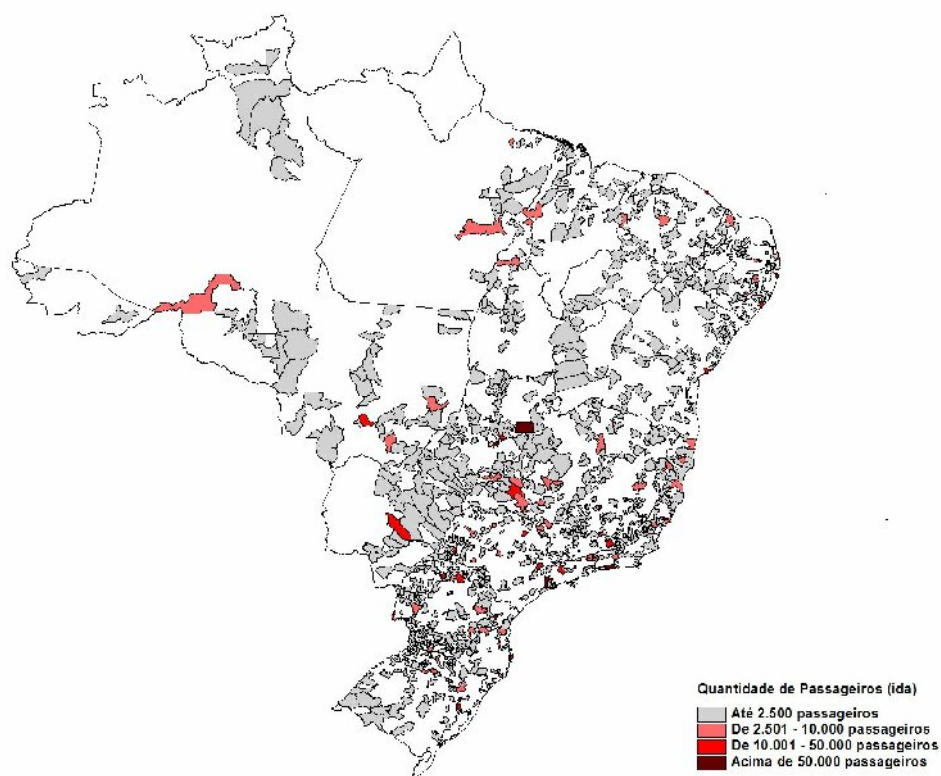


Figura 5.1: Quantidade de passageiros transportados em viagens produzidas por município (Brasil, Out./2004).

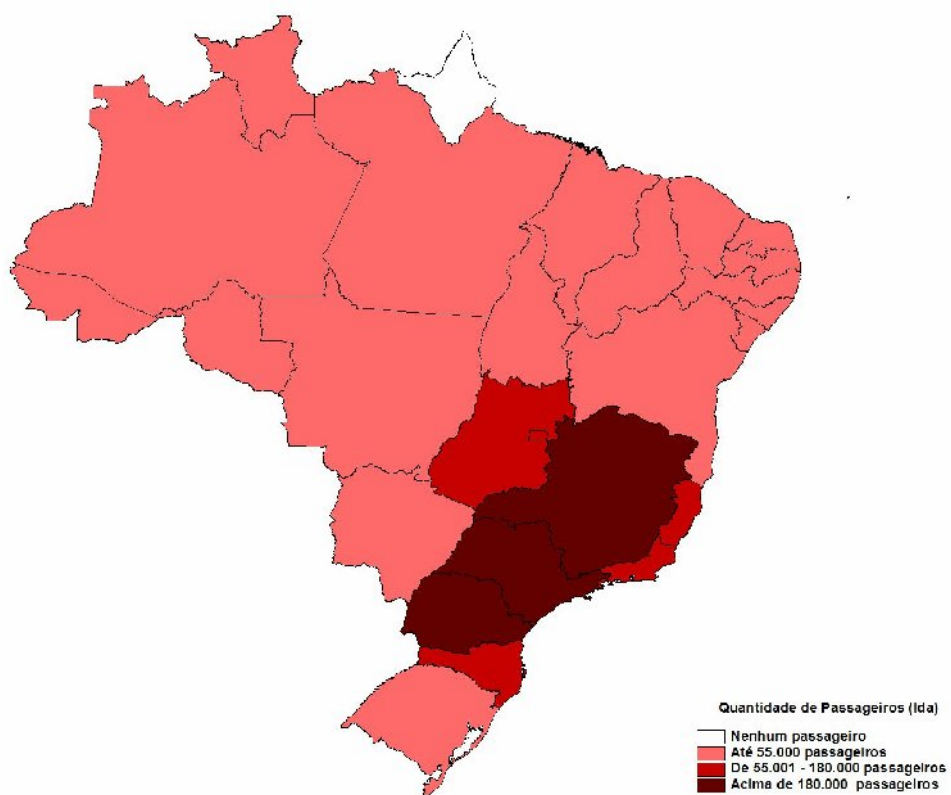


Figura 5.2: Quantidade de passageiros transportados em viagens produzidas por estado (Brasil, Out./2004).

A cidade de São Paulo foi o município com o maior deslocamento de pessoas em todo país utilizando o transporte rodoviário interestadual, com 226 mil passageiros (aproximadamente 15% do total de brasileiros). Destes, mais de 155 mil passageiros com destino a municípios localizados na própria região sudeste. A outra parte de passageiros, com origem a cidade de São Paulo, se dividiu entre as outras regiões, sendo que os municípios do sul e do nordeste foram os mais visitados.

Da cidade de São Paulo aos municípios do Estado do Ceará, no ano de 2004, aconteceram mais de 2.300 viagens (ida), totalizando em aproximadamente 50 mil passageiros. No mês de outubro, mais de 3.500 passageiros saíram da cidade de São Paulo com destino ao Ceará. Uma quantidade consideravelmente pequena, mas foi o terceiro município com maior número de passageiros no período, perdendo apenas para Teresina/PI e Natal/RN, conforme a Tabela 5.2 dos sete maiores emissores de passageiros para o estado do Ceará:

Tabela 5.2: Os sete maiores emissores de passageiros para o Ceará (Out./2004)

Municípios/Estados	Quantidade de Passageiros
Natal/RN	6.254
Teresina/PI	5.443
São Paulo/SP	3.607
Parnaíba/PI	2.306
Rio de Janeiro/RJ	2.188
Recife/PE	1.622
Belém/PA	1.476

Fonte: SISDAP/ANTT, 2004

Em determinadas épocas do ano, nos chamados meses atípicos (janeiro, fevereiro, julho e dezembro), existe uma concentração maior de passageiros do transporte rodoviário chegando aos municípios do Ceará. Mas como neste estudo é essencial ter a informação dos passageiros das viagens intermunicipais e esta só está disponível para o mês de outubro, obrigatoriamente trabalhou-se com esse mês nas viagens interestaduais.

Nas Figuras 5.3 e 5.4, pode-se observar a quantidade de passageiros em viagens interestaduais com origem à cidade de São Paulo em duas épocas do ano, outubro e dezembro. As figuras mostram que provavelmente o estudo seria mais satisfatório se estivessem disponíveis os dados das viagens intermunicipais para o mês de dezembro, já que a quantidade de passageiros nesse mês é muito superior a outubro.

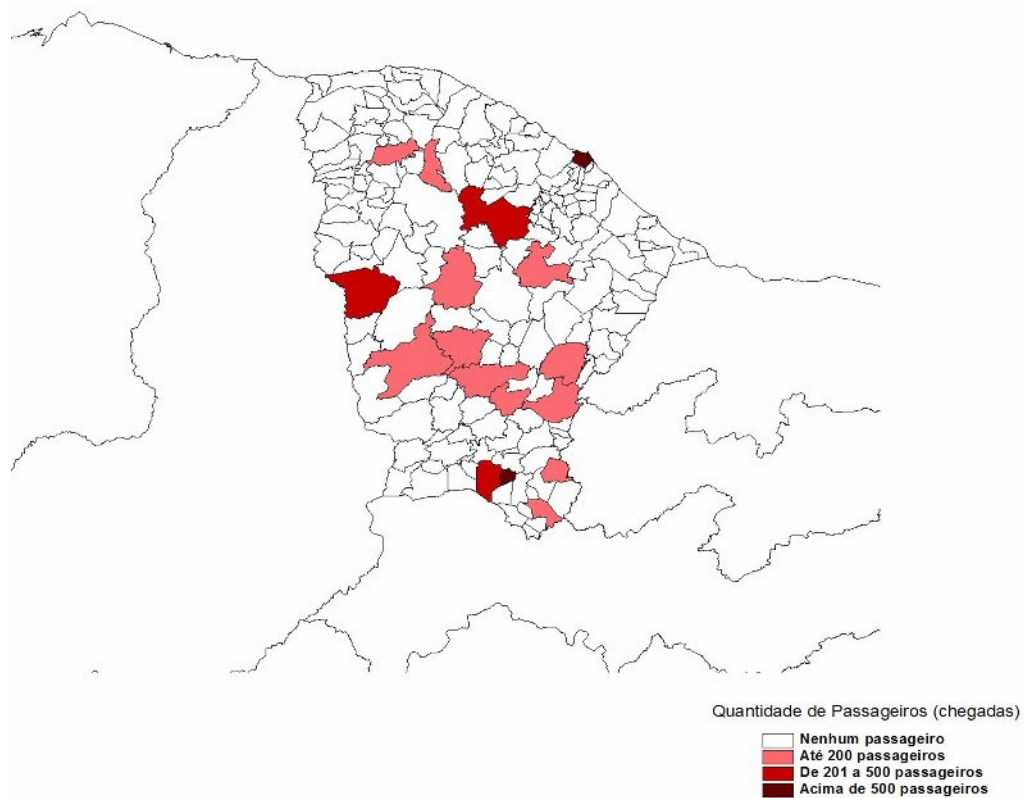


Figura 5.3: Quantidade de Passageiros ($Pass_{ij}$) que viajaram do município de São Paulo/SP aos municípios do Ceará em Outubro de 2004.

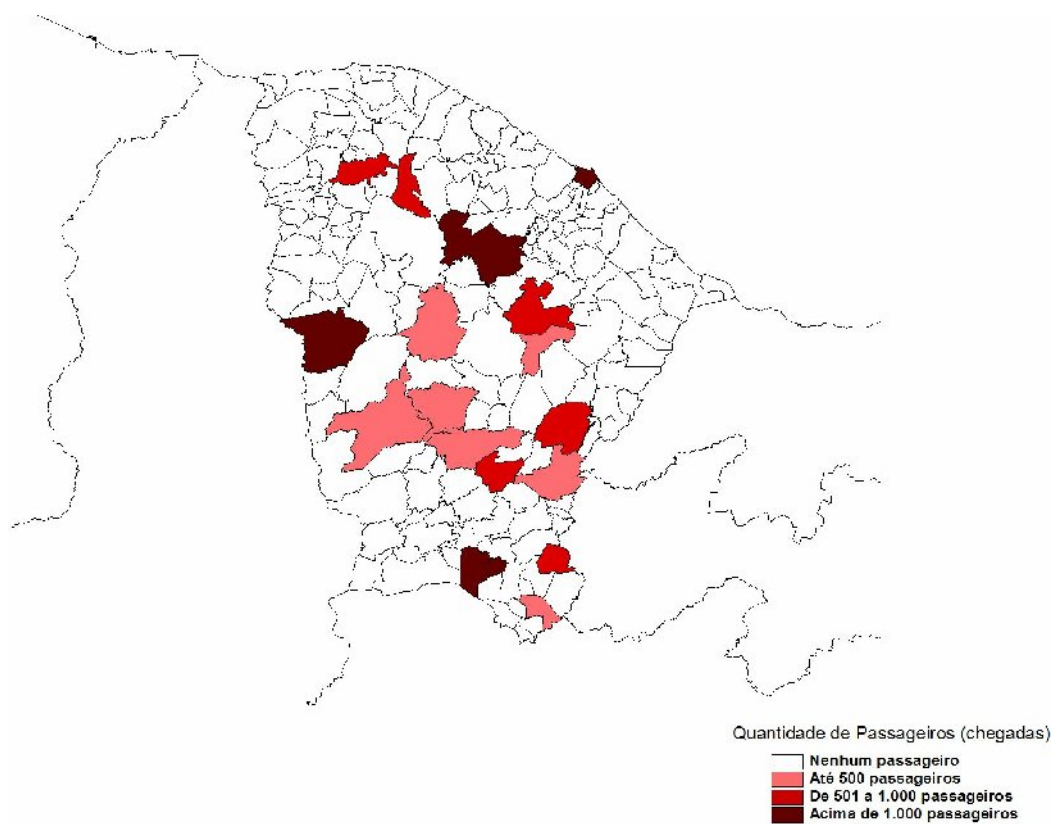


Figura 5.4: Quantidade de Passageiros ($Pass_{ij}$) que viajaram do município de São Paulo/SP aos municípios do Ceará em Dezembro de 2004.

Como visto anteriormente, Natal/RN é o município com maior quantidade de passageiros com destino aos municípios do Ceará (mais de 6 mil passageiros). Vale à pena ressaltar que apenas 8 municípios do Estado do Ceará foram pontos terminais dos passageiros que saíram de Natal no mês de outubro (Fortaleza, Crato, Aracati, Milagres, Sobral, Tianguá, Juazeiro do Norte, Canindé). Então, além do trabalho realizado com base nas viagens que saem de São Paulo para municípios cearenses, desenvolve-se, em conjunto, um estudo complementar para um município mais próximo (Natal/RN), a fim de mostrar a adequabilidade do modelo para as diferentes distâncias.

Com relação ao estado do Ceará, as viagens intermunicipais produzidas se distribuem (com relação à quantidade de passageiros) conforme a Figura 5.5. Observa-se que a maioria dos passageiros que se deslocam por meio do transporte rodoviário intermunicipal está na Região Metropolitana de Fortaleza e Itapipoca, ao nordeste do estado; em Sobral, mais ao noroeste, e Iguatu, Crato, Juazeiro do Norte e Barbalha, ao sul do Ceará. Com isso é possível que essas regiões sejam pontos terminais ou até mesmo pontos de integração de uma viagem interestadual.

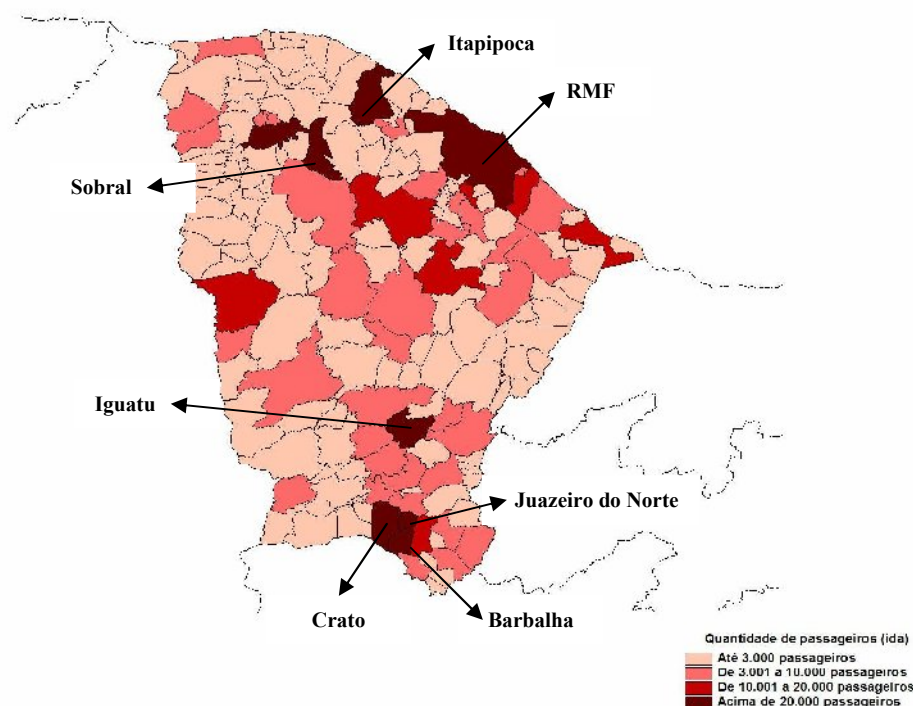


Figura 5.5: Quantidade de Passageiros ($Pass_{munic}$) em viagens intermunicipais no Estado do Ceará em Outubro de 2004.

Nesse contexto, a área de estudo para aplicação da metodologia é o par origem-destino (O/D) de São Paulo/SP e algumas cidades do Ceará, em função da disponibilidade dos dados de viagens intermunicipais. Como uma análise complementar, será realizado um estudo no município de Natal/RN.

Etapa 2: Identificação das variáveis

Esta etapa tem o propósito de identificar as principais variáveis independentes que possibilitam explicar a variável dependente (demanda de passageiros dos pares de O/D do transporte rodoviário interestadual). Esta quantidade de passageiros é informada pela ANTT através do Sistema de Controle dos Dados dos Serviços de Transporte Rodoviário de Passageiros (SISDAP). O ano de referência consultado no SISDAP para esta dissertação foi o de 2004.

De acordo com a metodologia apresentada no Capítulo 4, as variáveis independentes foram classificadas como socioeconômicas e espaciais. As variáveis como renda e população são importantes para explicar os deslocamentos de pessoas no âmbito interestadual utilizando o modo ônibus e são classificadas como variáveis socioeconômicas. Dessa forma, as disponíveis para o estudo, obtidas no IBGE e em outras instituições conceituadas, que de alguma forma apresentam relação com a produção de viagens e com a demanda de passageiros que se deslocam na área de estudo são:

- **População (*Pop*)** – Estimativa de habitantes do município em 2004 (IBGE).
- **Renda per capita (*Renda*)** – Renda média per capita por município em 2000 (IBGE).
- **PIB per capita (*PIB*)** – Produto Interno Bruto por município em 2004 (IBGE).
- **Pólo Turístico (*TUR*)** – Municípios que são roteiros turísticos no país em 2004/05 (Embratur/Ministério do Turismo).
- **Migração (*MIG*)** – Taxa de habitantes que não são naturais do Estado em 2004 (IBGE).
- **Empregos Formais (*EMP*)** – Número de empregos formais no município em 2004 (Ministério do trabalho).
- **Leitos Hospitalares (*LH*)** – Número de leitos hospitalares no município em 2004 (Ministério da Saúde).

- **Famílias Pobres (FP)** – Número de famílias abaixo da linha de pobreza (até R\$ 120,00 reais per capita) no município em 2004 (IPEA e PNAD/IBGE).

As variáveis espaciais, obtidas em trabalhos científicos, disponíveis para a possível construção do modelo proposto, são as que, de alguma forma, influenciam espacialmente o fenômeno, e estão definidas a seguir:

- **Distância (Dist)** – Distância rodoviária entre os municípios de origem e de destino nas ligações (ANTT, 2007).
- **Tempo de viagem (Tempo)** – Tempo de viagem entre os municípios de origem e de destino nas ligações, pela malha rodoviária utilizando o modo ônibus (ANTT, 2007).
- **Índice de cobertura espacial rodoviária (ICER)** – Índice de cobertura espacial rodoviária para o município que considera a área de influência da rodovia (Magalhães, 2004).

As variáveis *Dist* ou *Tempo* são elementos que compõem a matriz de proximidade espacial (W) no cálculo dos índices de autocorrelação e no desenvolvimento do modelo de regressão com efeitos espaciais.

A variável *ICER*, proposta por Magalhães (2004), é um índice que representa o poder de influência ou de atração de uma rodovia em relação a determinado município e que indiretamente representa a oferta de transporte nas ligações interestaduais. Ela pode ser calculada da seguinte forma:

$$ICER_x = \frac{\sum_{i=1}^n \gamma(i)A_i}{A_x} \quad (5.1)$$

onde,

ICER_x = índice de cobertura espacial rodoviário para o município *x*;

γ(i) = função que determina o peso atribuído à *i*-ésima faixa cobertura, tal que, *γ(i)* está entre zero (0) e um (1);

A_i = área da *i*-ésima faixa pertencente a *x*;

A_x = área do município de estudo *x*;

n = quantidade de faixas (*buffers*).

Etapa 3: Análise da qualidade do banco de dados disponível

A primeira ação realizada nesta etapa é a análise da demanda de passageiros no âmbito interestadual no banco de dados chamado SISDAP, referente ao ano de 2004, fornecido pela ANTT. Diversas tabelas estão compondo esse banco, mas apenas duas são de interesse para o trabalho, a saber: *movimento_secao* e *cidade*.

A tabela *movimento_secao* é utilizada para coletar informação da variável dependente. Nessa tabela encontram-se os códigos dos municípios de origem e de destino, com informações do número de passageiros na viagem de ida e de volta em todos os meses do ano de 2004. A tabela *cidade* é utilizada para separar as viagens realizadas no Brasil (interestaduais e intermunicipais) e para buscar os nomes das localidades onde se situam os pontos de seção ou os pontos terminais.

A fim de analisar a qualidade dos dados, essas duas tabelas foram transformadas em apenas uma. Depois dessa junção, o próximo passo foi extrair apenas a movimentação de passageiros na esfera interestadual e acima de 100 km. Foram retirados desta pesquisa origens e destinos com sede em distritos, subdistritos, restaurantes, postos, trevos, entre outros que não sejam municípios.

Na segunda parte dessa etapa, trabalhar-se-á com uma tabela da quantidade de passageiros nas viagens intermunicipais fornecida pelo Plano Diretor e Operacional do Transporte Intermunicipal de Passageiros do Estado do Ceará (PDOTIP-CE), referente a uma pesquisa realizada em uma semana do mês de outubro de 2004.

Após verificar a qualidade dos dois principais bancos de dados, também é necessário analisar esta qualidade nas variáveis identificadas na etapa anterior. Vale ressaltar que esta análise não será tão aprofundada, pois objetiva-se apenas a eliminar dados incoerentes.

Utilizando o indicador de *temporalidade*, identifica-se que existe uma variável, renda média, com ano diferente ao ano que se deseja estudar (2004). Esta se refere ao ano de 2000, e poderia acarretar problemas na estimação dos parâmetros do modelo. Mesmo assim, devido à dificuldade de aquisição, a variável renda média fará parte do modelo.

A Tabela 5.3 apresenta uma análise exploratória da planilha que compõe as possíveis variáveis independentes, utilizando os indicadores de *precisão* e *consistência*.

Tabela 5.3: Análise exploratória das possíveis variáveis independentes

Variáveis	Válidos	Missing	Média	Mediana	Desv Padrão	Mínimo	Máximo	Domínio
<i>Município</i>	5.564	0	-	-	-	-	-	-
<i>Pop</i>	5.564	0	32.634,980	10.562,00	194.959,618	818	10.838.581	0 a ∞
<i>Renda</i>	5.507	57	170,814	159,100	96,425	28,38	954,65	0 a ∞
<i>PIB</i>	5.560	4	7.683,061	5.061,470	11.244,015	763,36	315.208,07	0 a ∞
<i>FP</i>	5.564	0	1.995,460	860,530	6.776,804	23	285,06	0 a ∞
<i>TUR</i>	5.564	0	,080	0	,267	0	1	0 a 1
<i>MIG</i>	5.564	0	39,893	38,700	8,681	25,50	61,30	0 a ∞
<i>ICER</i>	5.507	57	,9317	,9843	,1576	0	1	0 a 1
<i>LH</i>	5.507	57	80,050	20,000	474,591	0	19.114	0 a ∞
<i>EMP</i>	5.564	0	4.438,460	365,000	44.968,328	0	2.632.019	0 a ∞

De acordo com a Tabela 5.3, todas as variáveis estão dentro dos respectivos domínios, mas alguns problemas foram encontrados. Em relação aos valores faltantes (*missing*) para *Renda*, *ICER* e *LH*, estes dizem respeito aos 57 municípios não existentes no ano 2000 na base de dados do IBGE e do Ministério da Saúde para o cálculo dessas variáveis. A variável *PIB* utilizou apenas o número de municípios brasileiros que existiam até o ano de 2002. Apesar dessas questões, os municípios faltantes não alterarão o resultado final.

Com referência à análise exploratória dos dados, observa-se a existência de valores muito discordantes para a média e a mediana de algumas variáveis. Esse problema indica a presença de valores extremos, como mostra, por exemplo, a variável *Pop*. Esse fato pode interferir nos pressupostos de erros normais e de homogeneidade nas variâncias, já que existe a possibilidade da não normalidade dos dados.

Etapa 4: Seleção das variáveis independentes

Todas essas possíveis variáveis independentes serão agregadas aos dois bancos de dados (passageiros interestaduais e passageiros intermunicipais) para, em seguida, verificar-se quais farão parte do modelo. Anteriormente a essa junção, foi necessário excluir a variável que identifica o tempo de viagem (*Tempo*), pois não se obteve essa informação para as viagens intermunicipais. A variável que identifica a distância entre os municípios não será trabalhada na matriz de correlação, pois será o elemento de proximidade espacial (matriz W) na construção do modelo de regressão com efeitos espaciais locais.

Para selecionar as variáveis independentes, será utilizada a matriz de correlação de Pearson com as variáveis identificadas na Etapa 2, juntamente com a demanda de passageiros (*Pass*) que se deslocaram nas viagens interestaduais acima de 100 km, no mês de outubro de 2004, conforme a Tabela 5.4. Vale ressaltar que estas variáveis serão atribuídas tanto para os municípios de origem (*V_1*) como os de destino (*V_2*), e que a utilização do mês de outubro de 2004 tem a justificativa de que é o mesmo mês e ano da pesquisa realizada para as viagens intermunicipais do Estado do Ceará.

Tabela 5.4: Matriz de correlação das variáveis

	<i>Pass</i>	<i>Pop</i> ₁	<i>Renda</i> ₁	<i>PIB</i> ₁	<i>FP</i> ₁	<i>TUR</i> ₁	<i>MIG</i> ₁	<i>ICER</i> ₁	<i>LH</i> ₁	<i>EMP</i> ₁	<i>Pop</i> ₂	<i>Renda</i> ₂	<i>PIB</i> ₂	<i>FP</i> ₂	<i>TUR</i> ₂	<i>MIG</i> ₂	<i>ICER</i> ₂	<i>LH</i> ₂	<i>EMP</i> ₂
<i>Pass</i>	1	,440**	,421**	,021**	,439**	,396**	,001	,216**	,247**	,241**	,421**	,424**	,031**	,423**	,391**	-,051**	,237**	,244**	,221**
<i>Pop</i> ₁		1	,495**	,121**	,972**	,412**	,015*	,077**	,959**	,997**	-,049**	-,088**	-,061**	-,046**	-,026**	-,148**	,021**	-,050**	-,050**
<i>Renda</i> ₁			1	,414**	,456**	,413**	,247**	,223**	,557**	,504**	-,093**	,059**	,055**	-,119**	-,066**	,108**	-,027**	-,100**	-,087**
<i>PIB</i> ₁				1	,091**	,087**	,137**	,117**	,100**	,130**	-,052**	,080**	,087**	-,077**	-,061**	,125**	,004	-,060**	-,046**
<i>FP</i> ₁					1	,454**	-,051**	,074**	,949**	,954**	-,045**	-,110**	-,079**	-,036**	-,009	-,188**	,028**	-,044**	-,047**
<i>TUR</i> ₁						1	-,025**	,074**	,479**	,401**	-,028**	-,036**	-,033**	-,021**	,010	-,091**	-,023**	-,024**	-,029**
<i>MIG</i> ₁							1	-,079**	-,026**	,027**	-,102**	,008	,040**	-,136**	-,113**	,365**	-,207**	-,117**	-,095**
<i>ICER</i> ₁								1	,101**	,080**	-,017*	,024**	,046**	-,029**	-,042**	,013	,099**	-,018*	-,013
<i>LH</i> ₁									1	,956**	-,054**	-,091**	-,065**	-,048**	-,017*	-,156**	,023**	-,054**	-,055**
<i>EMP</i> ₁										1	-,050**	-,082**	-,056**	-,048**	-,029**	-,137**	,019*	-,052**	-,050**
<i>Pop</i> ₂											1	,543**	,138**	,980**	,440**	-,041**	,124**	,968**	,998**
<i>Renda</i> ₂												1	,400**	,521**	,424**	,157**	,202**	,588**	,544**
<i>PIB</i> ₂													1	,120**	,072**	,090**	,088**	,133**	,142**
<i>FP</i> ₂														1	,485**	-,097**	,131**	,964**	,969**
<i>TUR</i> ₂															1	-,007	,051**	,496**	,425**
<i>MIG</i> ₂																1	-,324**	-,090**	-,030**
<i>ICER</i> ₂																	1	,150**	,124**
<i>LH</i> ₂																		1	,964**
<i>EMP</i> ₂																			1

*. Correlação significativa ao nível de 5%.

**.. Correlação significativa ao nível de 10%.

Verifica-se que todas variáveis escolhidas têm correlação significativa com a variável *Demanda de Passageiros*, porém variáveis como *Migração* e *PIB per capita* com correlações muito próximas de zero. Contudo, a variável relacionada à população do município (tanto origem como destino) apresenta uma alta correlação com as variáveis *Famílias Pobres*, *Leitos Hospitalares* e *Empregos Formais*. Como visto no capítulo 4, essa correlação indica um possível problema de multicolinearidade, o que já era esperado, visto que municípios de grande porte (acima de 100 mil habitantes) têm mais famílias abaixo da linha de pobreza, mais leitos hospitalares, mais empregos formais, e vice-versa.

Nesse contexto, as variáveis utilizadas no modelo de previsão da demanda de passageiros do transporte rodoviário interestadual são *População*, *Renda*, *Pólo Turístico* e *Índice de Cobertura Espacial Rodoviária*.

Etapa 5: Verificação da dependência espacial

Uma análise descritiva e exploratória da variável dependente *Pass* indica uma possível dependência espacial. Assim, há necessidade de verificá-la, por meio de técnicas específicas de estatística espacial, para o desenvolvimento de modelos espaciais. Primeiramente, será verificada a presença da dependência espacial e da não-estacionariedade na quantidade de passageiros em viagens interestadual (*Pass*) e em seguida, na quantidade de passageiros nas viagens intermunicipais no estado do Ceará (*Pass_{munic}*).

- *Verificação da dependência espacial nos passageiros de viagens interestaduais*

Utilizando-se o índice de autocorrelação global com a matriz W de contigüidade e das distâncias rodoviárias, onde no ultimo caso utilizar-se-á um dos métodos proposto por Silva (2006), em que $w_{ij} = 1/(1 + d_{ij})$, têm-se os seguintes resultados (Tabela 5.5):

Tabela 5.5: Indicadores de autocorrelação global para a variável *Pass*

Matriz W de contigüidade	Matriz W das distâncias
$I = - 0,0084$	$I = 0,0016$
$(p=0,375)$	$(p=0,376)$

O valor entre parênteses é o p-valor calculado (a hipótese nula não será rejeitada se p-valor > 0,1)

Considerando os valores encontrados, pode-se afirmar que a hipótese nula de não dependência espacial na variável *Pass* não será rejeitada, ou seja, a variável *demanda de passageiros (Pass)* não assume valores similares em unidades geograficamente próximas (tipo binária) e na relação de conectividade entre os municípios (distâncias).

Como visto anteriormente, existem maneiras adicionais em que é possível verificar a presença da dependência espacial, além do índice global de Moran. Essas maneiras serão apresentadas a seguir por meio das Figuras 5.6 (a) e (b):

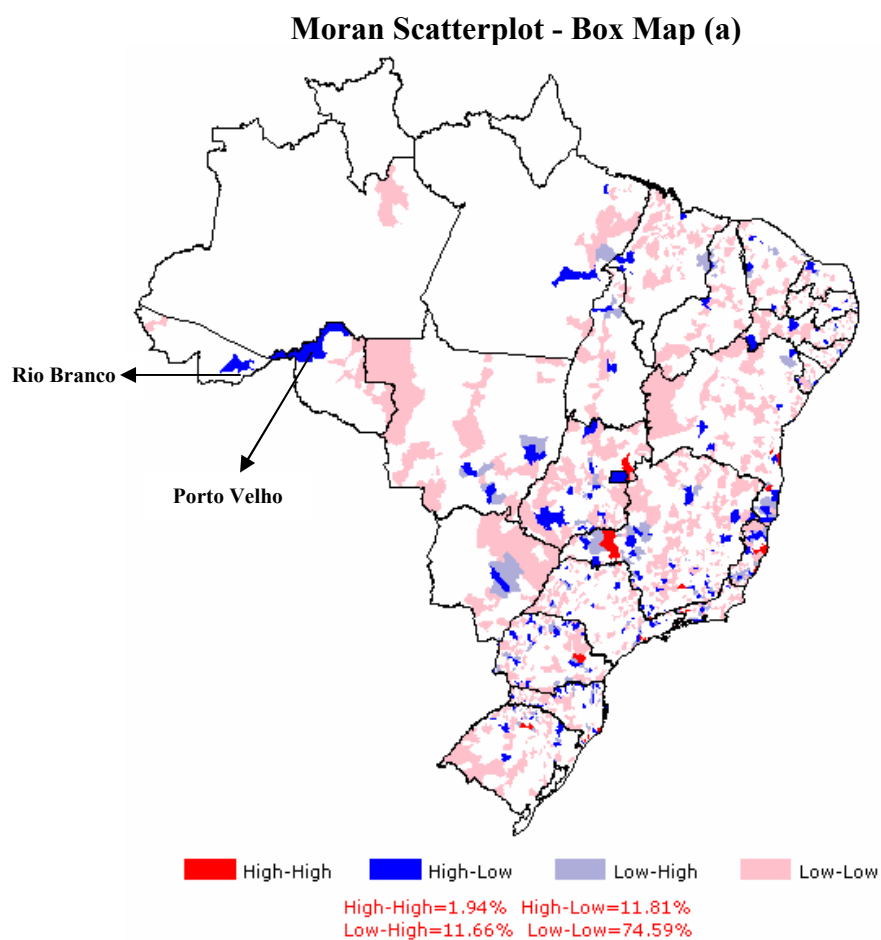


Figura 5.6 (a): *Box Map* da quantidade de passageiros utilizando a matriz W de contigüidade.

Moran Scatterplot - Box Map (a)

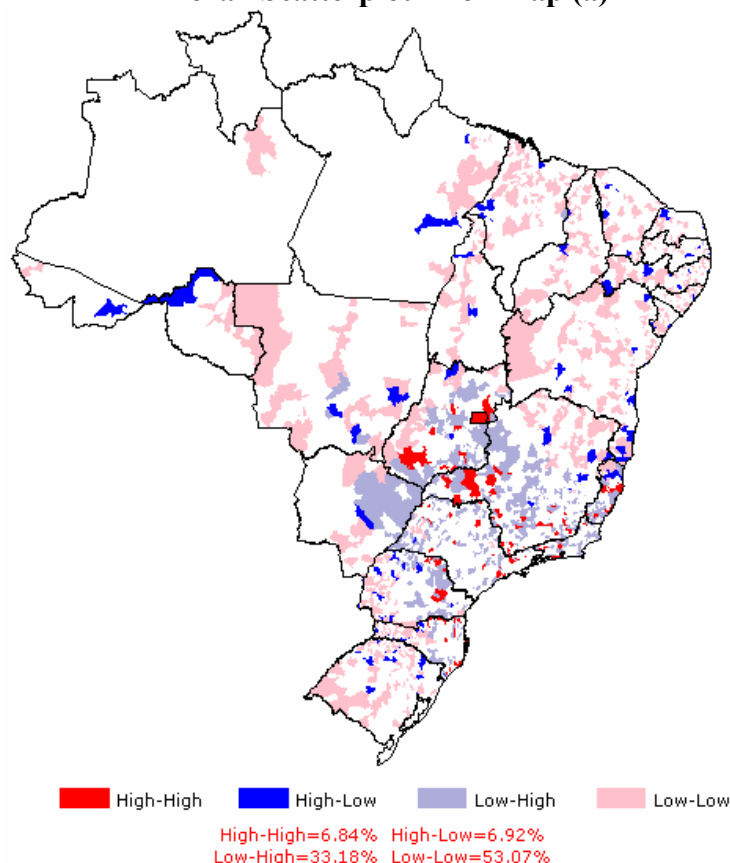


Figura 5.6 (b): *Box Map* da quantidade de passageiros utilizando a matriz W das distâncias rodoviárias.

Observa-se nas Figuras 5.6 (a) e (b), que mesmo com a quantidade de áreas classificadas como definidoras de dependência espacial (*High-High* e *Low-Low*) somarem mais de 60% da informação, o índice global não foi significativo, conforme a Tabela 5.5. Mesmo sem a definição da dependência espacial, verifica-se muitos municípios classificados como *Low-Low* no Nordeste, tanto na Figura 5.6 (a) quanto na Figura 5.6 (b). Isso significa que existia uma pequena quantidade de passageiros que saíram desses municípios e dos seus vizinhos no mês de Outubro de 2004.

Quando a matriz W de contiguidade é utilizada, observam-se apenas alguns municípios *High-High*, ou seja, os municípios com números elevados de passageiros que viajam no modo interestadual e seus vizinhos idem (apenas os que fazem fronteira). Já quanto à matriz W das distâncias rodoviárias entre os municípios, verifica-se uma concentração maior, principalmente no Centro-Oeste e Sudeste. Isso demonstra, nesse caso, que a utilização da matriz W com as distâncias rodoviárias entre os municípios fornece

resultados mais relevantes quando se analisa a dependência espacial do que a de contigüidade. Porém, existe a possibilidade deste fato não ocorrer em todas as análises.

Vale lembrar que diversos municípios brasileiros não atraem e não produzem viagens interestaduais diretamente, ou seja, nestas localidades não existem linhas para todos os municípios e os passageiros costumam terminar suas viagens utilizando o transporte intermunicipal. Com isso, observa-se nas Figuras 5.6 (a) e (b) que existem numerosas áreas brancas, onde não houve passageiros em viagens produzidas em outubro de 2004.

Outro aspecto importante que vale ressaltar é que, com a realidade desses municípios onde não possuem linhas diretas regulamentadas do transporte rodoviário interestadual de passageiros, existem municípios, em alguns casos, que são considerados *High-Low* por seus vizinhos não gerarem viagens, como Rio Branco e Porto Velho.

Depois da análise da caracterização da demanda de passageiros que se deslocam em todo Brasil (análise global do fenômeno), será necessária uma análise local. Como não se encontrou dependência espacial no todo, é bastante provável a existência de sub-regiões com seus padrões espaciais próprios.

Dessa forma, utilizando o índice local de *Moran*, observa-se significância em alguns municípios, conforme as Figuras 5.7 (a) e (b):

Moran Map (a)



Figura 5.7 (a): *Moran Map* da quantidade de passageiros utilizando a matriz W de contigüidade.

Moran Map (b)



Figura 5.7 (b): *Moran Map* da quantidade de passageiros utilizando a matriz W das distâncias rodoviárias.

Verifica-se que tanto na matriz W de contigüidade, como das distâncias rodoviárias, apresentam algumas áreas significativas (heterogeneidade espacial), indicando padrões espaciais diferenciados para algumas sub-regiões. Tanto I_i local de *Moran* de Natal/RN (*High-Low*) como de São Paulo/SP (*High-High*) foram significativos. Visto isso, no caso do cálculo da demanda de passageiros, principalmente os que têm como origem São Paulo, espera-se que o modelo de regressão com efeitos espaciais locais, estime de forma mais realista do que modelos de regressão convencional e com efeitos espaciais globais, conforme citado na seção 3.3.2.

- *Verificação da dependência espacial nos passageiros de viagens intermunicipais*

Após a verificação da não-estacionaridade (*Moran Map*) na variável *Pass*, em alguns municípios do país, é indispensável confirmar a não-estacionaridade na variável *Pass_{munic}* (passageiros das viagens intermunicipais) nos municípios destinos das viagens interestaduais. Conseqüentemente, realizou-se a construção de mapas e testes para os índices de autocorrelação espacial (Figuras 5.8 (a) e (b), 5.9 (a) e (b), e Tabela 5.6) para verificar os efeitos espaciais nos municípios cearenses.

Vale ressaltar que as informações de Fortaleza, nas viagens intermunicipais, englobam os passageiros da Região Metropolitana (além de Fortaleza, Maranguape, São Gonçalo do Amarante, Caucaia, Maracanaú, Pacatuba, Guaiúba, Itaitinga, Eusébio, Aquiraz, Horizonte, Pacajús e Chorozinho). Ou seja, passageiros que viajam entre estes municípios são considerados passageiros em viagens semi-urbanas (metropolitano) e não são contabilizados em viagens intermunicipais.

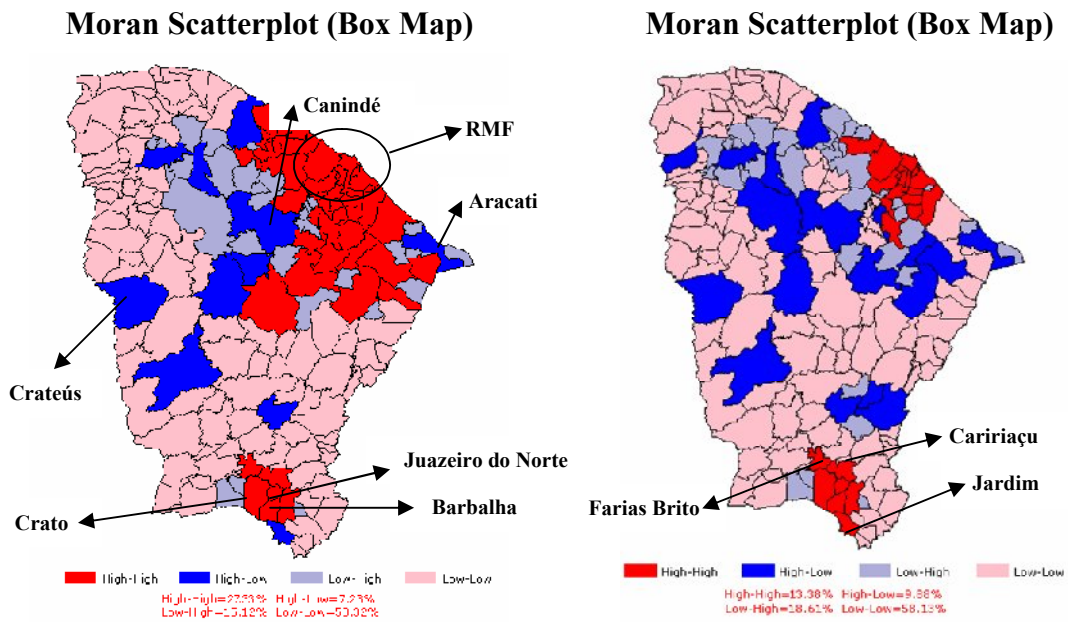


Figura 5.8: *Moran Scatterplot* da quantidade de passageiros utilizando a matriz W de contigüidade (a) e das distâncias (b), considerando a Região Metropolitana de Fortaleza (RMF).

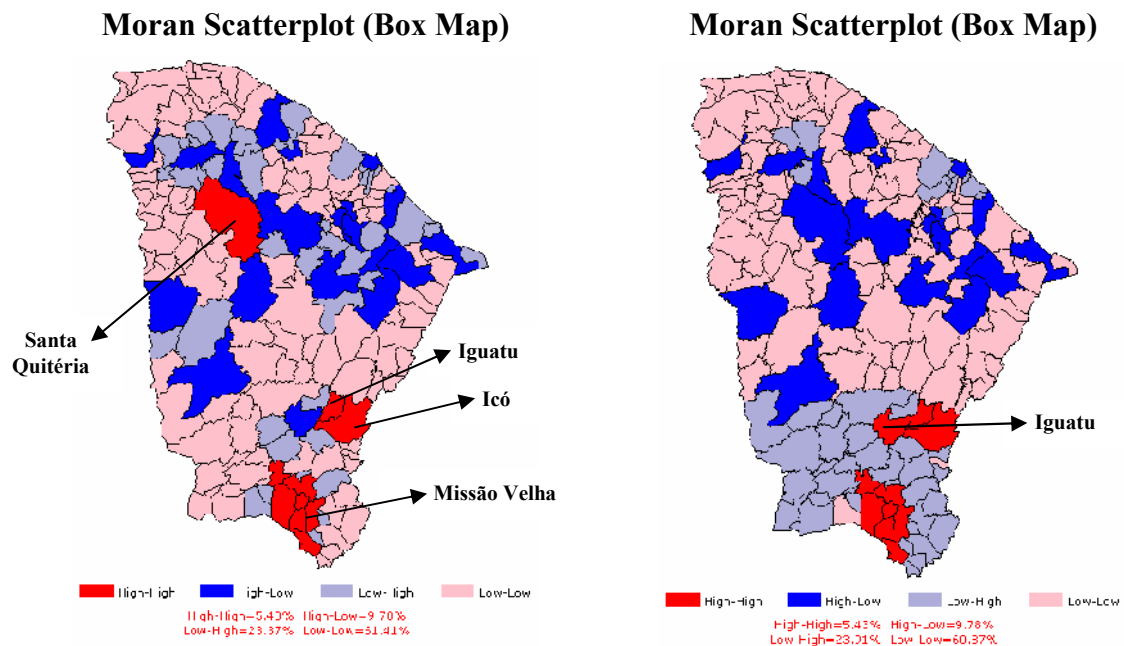


Figura 5.9: *Moran Scatterplot* da quantidade de passageiros utilizando a matriz W de contigüidade (a) e das distâncias (b), sem considerar a Região Metropolitana de Fortaleza (RMF).

Analisando a quantidade de passageiros que utilizam o transporte intermunicipal, verifica-se diferentes tipos de comportamentos, tanto na matriz W de contigüidade como na matriz W das distâncias rodoviárias. Mas, na maioria das localidades, a quantidade de passageiros que se desloca no modo intermunicipal é baixa, inserida em sub-regiões formadas por municípios também abaixo da média (*Low-Low*).

Utilizando-se o índice de autocorrelação global com a matriz W de contigüidade e das distâncias rodoviárias entre os municípios, com e sem a definição da Região Metropolitana de Fortaleza (RMF), tem-se os resultados na Tabela 5.6:

Tabela 5.6: Indicadores de autocorrelação global para a variável $Pass_{munic}$

Matriz W de contigüidade (RMF)	Matriz W das distâncias (RMF)
$I = 0,4460$ ($p=0,001$)	$I = 0,3500$ ($p=0,001$)
Matriz W de contigüidade	Matriz W das distâncias
$I = 0,2539$ ($p=0,000$)	$I = 0,0819$ ($p=0,000$)

O p entre parênteses é o p-valor calculado (a hipótese nula não será rejeitada se $p\text{-valor} > 0,1$)

De acordo com a Tabela 5.6, verifica-se que a hipótese nula de não dependência espacial na variável $Pass_{munic}$ foi rejeitada em todos os casos, ou seja, a variável *demandas de passageiros em viagens intermunicipais* assumiu valores similares em unidades geograficamente próximas e na relação de conectividade entre os municípios. Por conseguinte, há a evidência estatística de que, se tratando de passageiros em viagens intermunicipais, foi constatada a dependência espacial no processo.

Com relação aos municípios destinos para o estudo de caso, optou-se pelos seis municípios que mais receberam passageiros de São Paulo e Natal. Para São Paulo escolheu-se os destinos Crato, Crateús e Canindé, e para Natal escolheu-se os destinos Fortaleza, Juazeiro do Norte e Aracati.

Dos seis municípios analisados, três são *High-Low*, ou seja, municípios que podem significar influência socioeconômica sobre seus vizinhos, representando um pólo da região. Dependendo dos casos, são municípios auto-suficientes, diferentes dos seus vizinhos, que realizam viagens intermunicipais com esses e mantêm uma ligação relativamente sólida com os municípios mais ricos e mais distantes dentro do estado ou até mesmo em municípios de outros estados vizinhos.

Nas Figuras 5.8 e 5.9, observa-se que desses três municípios *High-Low*, dois deles têm algo em comum, pois estão “cercados” ou próximos de municípios *Low-Low* (Crateús e Aracati). Com isso, de acordo com Loureiro *et al.* (2006), talvez o tipo de modelo aplicado

nesse trabalho não seja o mais expressivo para o cálculo da demanda de passageiros para Crateús e Aracati, visto que existe a possibilidade de subestimação dos parâmetros da regressão desses dois municípios e da superestimação dos parâmetros da regressão dos municípios vizinhos a esses.

No município de Canindé, apesar de ser definido como *High-Low*, é possível que esses efeitos não ocorram, pois em volta desse, além de municípios classificados como *Low-Low*, existem outros classificados como *High-High*, *High-Low* e *Low-High*. Ou seja, existem municípios próximos a Canindé que também exercem uma influência socioeconômica sobre os vizinhos.

Em Crato e Juazeiro do Norte, observa-se em todos os mapas que a demanda de passageiros nas linhas intermunicipais que saíram destes municípios é elevada, e o mesmo acontece com seus vizinhos (*High-High*). Com esse resultado, é bastante provável que a modelagem proposta por essa dissertação será válida para esse caso, o que será demonstrado no decorrer do capítulo. Além disso, é possível sugerir com os resultados apresentados pelos *Box Maps* que o ponto terminal da viagem interestadual para Crato ou para Juazeiro do Norte poderá se encontrar em qualquer município vizinho *High-High*.

No caso de Fortaleza, só é exequível fazer uma análise das Figuras 5.8 (a) e (b), já que essas são relativas à RMF, e as informações das viagens realizadas dentro dessa região, como citado anteriormente, não são consideradas linhas do transporte rodoviário intermunicipal de passageiros e sim do transporte semi-urbano de passageiros. É possível que os resultados da regressão com efeitos espaciais locais encontrados nos municípios de Crato e de Juazeiro do Norte sejam diferentes dos resultados encontrados na RMF, devido à quantidade de vizinhos em torno da cada área e da distância entre eles. Ainda nesta seção, será realizada uma análise mais aprofundada destes dois tipos de casos.

Na Figura 5.9 (b) observa-se um ponto interessante. Quando se analisa os passageiros em viagens intermunicipais, os municípios definidos como *High-High* estão “cercados” por municípios *Low-High*. Isso significa que esses municípios podem ter uma dependência econômica com Crato, Juazeiro do Norte e seus vizinhos *High-High*. Logo, é possível que os passageiros que chegaram de viagens interestaduais tenham se deslocado, em seguida,

para os municípios definidos como *Low-High*, ou seja, Crato e/ou Juazeiro do Norte provavelmente foram os principais “pontos de integração” da viagem.

Depois da análise de todo o estado, viu-se a necessidade de verificar o comportamento de cada município utilizando técnicas do *LISA*, por meio do qual será possível localizar áreas de influência (sub-regiões semelhantes). As Figuras 5.10 a 5.15 mostram os resultados deste comportamento utilizando a matriz W de contigüidade, e as Figuras 5.16 a 5.21 com a matriz W das distâncias.

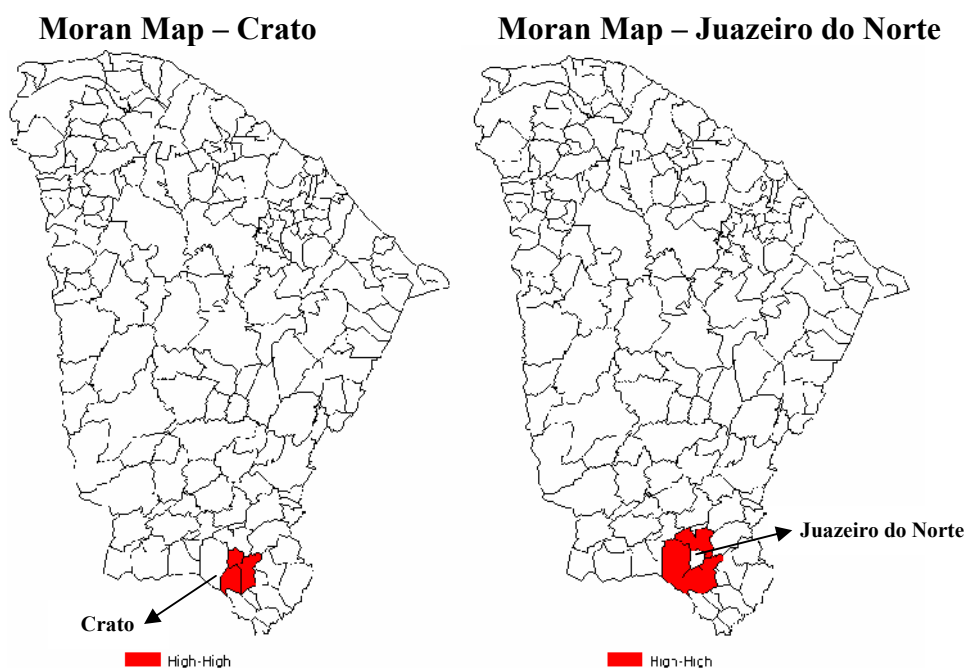


Figura 5.10: Moran Map da variável $Pass_{munic}$ no município de Crato – Matriz W de contigüidade.

Figura 5.11: Moran Map da variável $Pass_{munic}$ no município de Juazeiro do Norte – Matriz W de contigüidade.

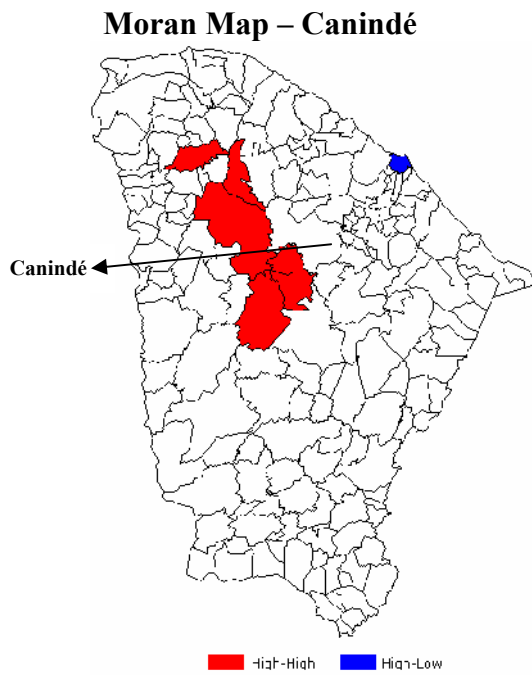


Figura 5.12: Moran Map da variável $Pass_{munic}$ no município de Canindé – Matriz W de contigüidade.

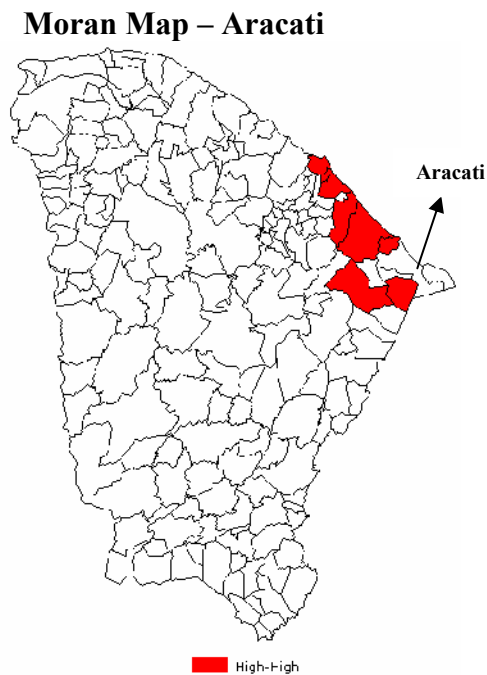


Figura 5.13: Moran Map da variável $Pass_{munic}$ no município de Aracati – Matriz W de contigüidade.

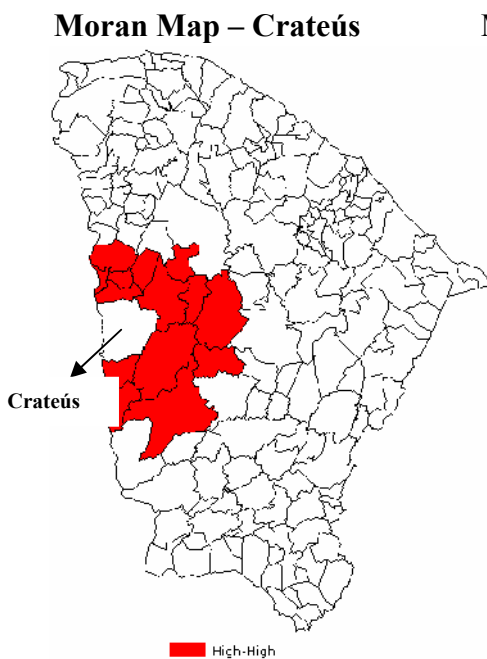


Figura 5.14: Moran Map da variável $Pass_{munic}$ no município de Crateús – Matriz W de contigüidade.

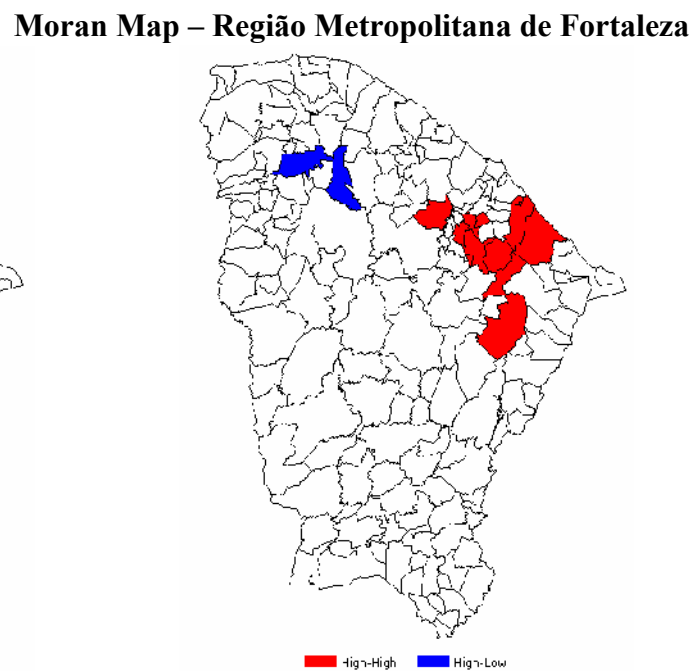


Figura 5.15: Moran Map da variável $Pass_{munic}$ na RMF – Matriz W de contigüidade.

Moran Map – Crato

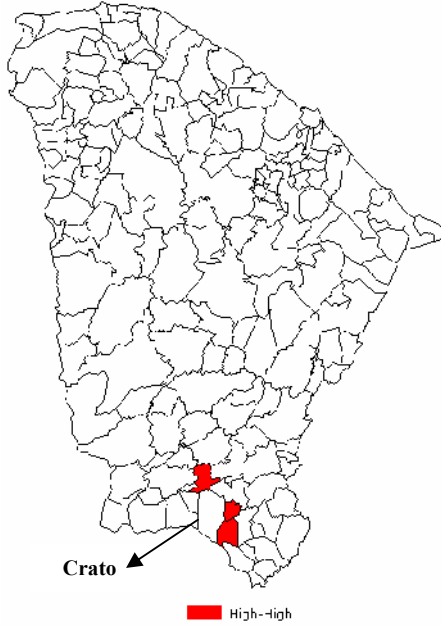


Figura 5.16: Moran Map da variável $Pass_{munic}$ no município de Crato – Matriz W das distâncias rodoviárias.

Moran Map – Juazeiro do Norte

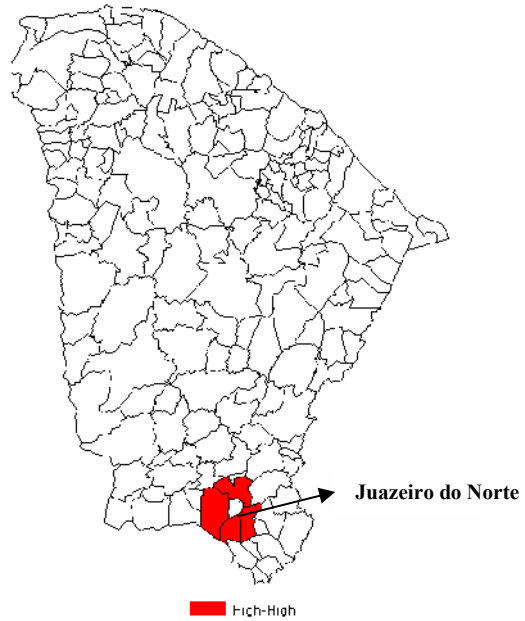


Figura 5.17: Moran Map da variável $Pass_{munic}$ no município de Juazeiro do Norte – Matriz W das distâncias rodoviárias.

Moran Map – Canindé

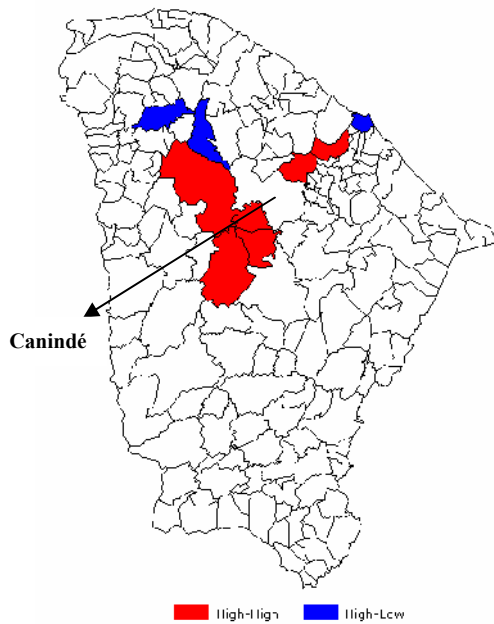


Figura 5.18: Moran Map da variável $Pass_{munic}$ no município de Canindé – Matriz W das distâncias rodoviárias.

Moran Map – Aracati

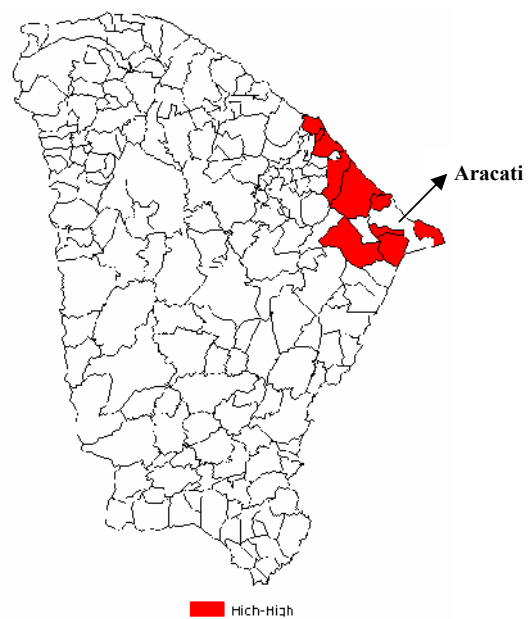


Figura 5.19: Moran Map da variável $Pass_{munic}$ no município de Aracati – Matriz W das distâncias rodoviárias.

Moran Map – Crateús

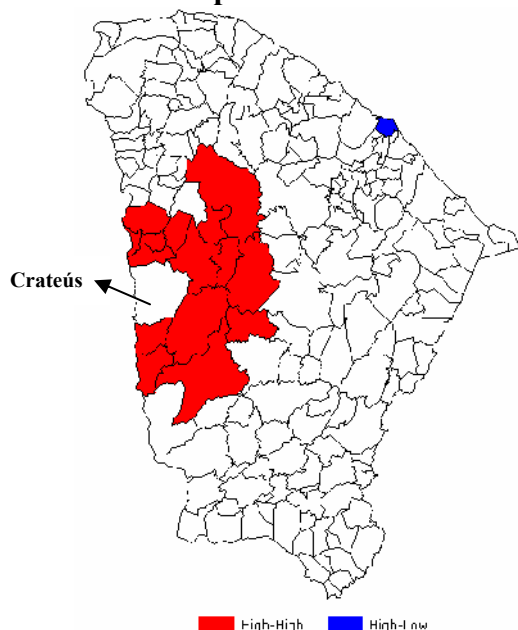


Figura 5.20: Moran Map da variável $Pass_{munic}$ no município de Crateús – Matriz W das distâncias rodoviárias.

Moran Map – Região Metropolitana de Fortaleza

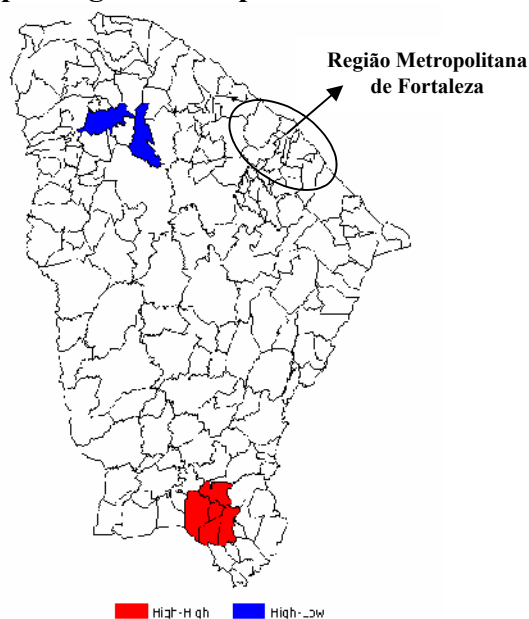


Figura 5.21: Moran Map da variável $Pass_{munic}$ no município de Crateús – Matriz W das distâncias rodoviárias.

Como já citado anteriormente, a análise utilizando técnicas LISA fornece apenas uma indicação de agregações significativas de um local ou uma observação, ou seja, todos os resultados estarão relacionados a um determinado município. Dessa forma, para cada um dos cinco municípios foram definidas 184 hipóteses nulas de não dependência espacial em relação à variável *demanda de passageiros em viagens intermunicipais* que saíram desses (184 municípios restantes). Apenas a RMF terá 171 hipóteses nulas, já que a quantidade de municípios que recebem passageiros da região diminui.

Exemplificando o parágrafo anterior, supõe-se que a hipótese nula de Crato relacionado a Barbalha é rejeitada. Então, pode-se afirmar que existe evidência estatística da *demanda de passageiros em viagens intermunicipais* assumir valores similares entre estes dois municípios.

Fazendo uma breve análise das Figuras 5.10 a 5.21, juntamente com as figuras anteriores, tem-se os seguintes resultados:

- As Figuras 5.10 e 5.21 mostram a não-estacionariedade do processo. Ou seja, as propriedades estatísticas da demanda de passageiros em viagens intermunicipais dependem da sub-região.

- Como já citado anteriormente, dois municípios possuem características em comum, Crateús e Aracati. Nestas figuras, observa-se que nas distintas matrizes de vizinhança, os resultados são bastante parecidos: os municípios significativos são os mais próximos desses, ou seja, existe uma quantidade alta de passageiros que se deslocam para Crateús e Aracati. Dessa forma, além de municípios considerados *High-Low*, provavelmente as viagens que ocorram em suas sub-regiões sejam rotineiras (trabalho e/ou estudo).
- Regressão com efeitos espaciais locais provavelmente não será o modelo mais expressivo para Crateús e Aracati.
- Existe a possibilidade de que a regressão com efeitos espaciais locais seja um modelo significativo para Canindé, pois além de ser um dos municípios que mais receberam passageiros da cidade de São Paulo e de possuir uma das maiores populações do estado, está relativamente próximo das regiões mais ricas e das mais populosas (RMF e Sobral). Nesse contexto, é possível que os passageiros que desembarcaram no terminal de Canindé tenham se deslocado para municípios mais próximos.
- Com relação a Crato e a Juazeiro do Norte, não se obteve informações mais relevantes. Espera-se que o modelo proposto seja o mais expressivo para esses, já que foram considerados municípios *High-High*.
- No caso da Região Metropolitana de Fortaleza, observa-se na Figura 5.21 que apesar da distância de mais de 200 quilômetros de Sobral e de aproximadamente 480 quilômetros da região composta por municípios do sul do estado (Crato, Juazeiro do Norte, Caririçu, Missão Velha e Barbalha), existem propriedades estatísticas similares entre estes locais e a RMF.
- No caso da Região Metropolitana de Fortaleza, é provável que o modelo proposto não expresse realmente a quantidade de passageiros de viagens interestaduais que chegam à sub-região, apesar de ser uma região *High-High* e possuir propriedades estatísticas parecidas com municípios *High-High* do estado. Isso poderá ocorrer, visto que a RMF é uma sub-região com padrões espaciais como as outras localizadas no estado, mas com uma diferença: não acontecerão viagens intermunicipais na região metropolitana, mas nas viagens semi-urbanas (mais viagens rotineiras).
- Apesar disso, desenvolveu-se um modelo de regressão com efeitos espaciais locais para os passageiros que chegaram a Fortaleza, supondo que a chegada foi para qualquer município da RMF.

Dessa forma, apesar de não existência neste estudo de um teste estatístico para verificar se o processo é não-estacionário, constatou-se a dependência espacial em sub-regiões separadas por meio dos índices de autocorrelação, como aconteceu em São Paulo e Natal nas viagens interestaduais e em alguns municípios do Ceará em viagens intermunicipais. Assim, com a possível não-estacionariedade do processo e a dependência espacial em diversificadas sub-regiões, os coeficientes de regressão provavelmente precisarão representar a heterogeneidade espacial e, indispensavelmente, utiliza-se de modelos de regressão espacial com efeitos locais.

Etapa 6: Construção do modelo de regressão com efeitos espaciais locais

Após a etapa de verificação da dependência espacial e da possibilidade de o fenômeno ser não-estacionário, o passo seguinte é a construção do modelo de regressão com efeitos espaciais locais das duas formas: discreta e contínua, comparando-as após a construção. O modelo será desenvolvido com base na demanda de passageiros que se deslocaram da cidade de São Paulo e Natal (onde os valores do índice de autocorrelação local foram significativos) para determinadas sub-regiões no Ceará.

Para a construção do modelo, as variáveis independentes (X) escolhidas na etapa 4 foram *População* (Pop_1 e Pop_2), *Renda* ($Renda_1$ e $Renda_2$), *Pólo Turístico* (TUR_1 e TUR_2) e *Índice de Cobertura Espacial Rodoviária* ($ICER_1$ e $ICER_2$). Todas as variáveis X_1 estão relacionadas às origens e todas as variáveis X_2 estão relacionadas aos destinos.

Primeiramente, o estudo de caso buscou calibrar modelos de regressão com efeitos espaciais locais de forma discreta. Dessa forma, calculou-se a demanda de passageiros dos municípios de São Paulo e Natal até os seis municípios (um de cada vez) por meio das informações da área de influência do destino. Essa influência será dada pela demanda de passageiros nas viagens intermunicipais e pelas variáveis independentes de cada município vizinho do ponto terminal da viagem interestadual.

Municípios vizinhos nesta pesquisa foram definidos de duas formas: a primeira é que o vizinho de um determinado município são os municípios que fazem fronteiras com esse. A segunda forma é determinada pelas menores distâncias rodoviárias entre esse determinado município e os outros.

O estudo iniciou-se com os passageiros que saíram de São Paulo em outubro de 2004 com destino a Crato, Crateús e Canindé. As áreas de influência desses municípios destinos foram classificadas em diferentes regimes espaciais ou ordens (1ª ordem – vizinhos do destino, 2ª ordem – vizinhos dos vizinhos do destino, e 3ª ordem – vizinhos dos vizinhos dos vizinhos do destino). Os resultados desses modelos se encontram nas Tabelas 5.7, 5.8 e 5.9.

Tabela 5.7: Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre São Paulo a Crato

		São Paulo – Crato					
Variáveis	1ª ordem		2ª ordem		3ª ordem		
	Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>	
Modelos iniciais	Intercepto	-	-	-	-	-	-
	<i>Pop_{div}</i>	-0,467* (0,008)	-4,83	-0,45* (0,000)	-5,42	0,401* (0,001)	3,69
	<i>Pop₂</i>	0,685* (0,007)	11,35	0,746* (0,000)	11,28	1,174* (0,000)	11,60
	<i>Renda_{div}</i>	-0,050 (0,768)	-0,323	0,171 (0,362)	0,947	-0,313* (0,028)	-2,37
	<i>Renda₂</i>	-0,190 (0,403)	-0,972	-0,516* (0,006)	-3,21	-0,608* (0,001)	-3,84
	<i>TUR_{div}</i>	-	-	-	-	-	-
	<i>TUR₂</i>	-	-	0,006 (0,952)	0,054	0,096 (0,294)	1,09
	<i>ICER_{div}</i>	-	-	0,391 (0,202)	1,349	0,323* (0,039)	2,21
	<i>ICER₂</i>	-	-	0,404 (0,203)	1,347	-3,53 (0,254)	-1,18
Modelos finais	<i>Pop_{div}</i>	-970,742* (0,009)	-4,76	-728,09* (0,001)	-4,18	1.096,489** (0,062)	1,99
	<i>Pop₂</i>	0,843* (0,000)	11,35	0,931* (0,000)	11,28	1,174* (0,000)	11,60
	<i>Renda_{div}</i>	-	-	-	-	-46.354,770* (0,009)	-2,92
	<i>Renda₂</i>	-	-	-142,66* (0,044)	-2,11	-1.540,435* (0,001)	-4,16
	<i>TUR_{div}</i>	-	-	-	-	-	-
	<i>ICER_{div}</i>	-	-	-	-	194.766,906* (0,004)	3,3
	R ²	0,970		0,941		0,931	
	R ² Ajustado	0,955		0,928		0,911	
	F	64,40* (0,001)		69,42* (0,000)		46,05* (0,000)	
	White (χ^2)	0,32 ^{NH} (2gl – 5,99)		1,92 ^{NH} (3gl – 7,81)		1,80 ^{NH} (5gl – 11,07)	
	Normalidade (SW)	<i>p</i> -valor = 0,004		<i>p</i> -valor = 0,465		<i>p</i> -valor = 0,418	
n	6		16		22		

(**) Significativos para *p*-valor < 0,10 e (*) Significativos para *p*-valor < 0,05. (NH) – Não há heterocedasticidade.

Tabela 5.8: Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre São Paulo a Crateús

São Paulo – Crateús							
Variáveis	1ª ordem		2ª ordem		3ª ordem		
	Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>	
Modelos iniciais	Intercepto	-	-	-	-	-	-
	<i>Pop_{div}</i>	-0,329 (0,437)	-0,86	-0,218 (0,250)	-1,20	-0,152 (0,346)	-0,97
	<i>Pop₂</i>	-	-	-0,102 (0,770)	-0,30	-0,123 (0,721)	-0,36
	<i>Renda_{div}</i>	-0,329 (0,437)	-0,86	-0,571 (0,165)	-1,47	-0,307 (0,415)	-0,83
	<i>Renda₂</i>	-	-	-0,267 (0,723)	-0,36	-0,371 (0,571)	-0,58
	<i>TUR_{div}</i>	-	-	-	-	-	-
	<i>TUR₂</i>	-	-	-	-	-0,206 (0,195)	-1,343
	<i>ICER_{div}</i>	4.226,86* (0,045)	2,05	3.834,57* (0,001)	4,60	3.253,012* (0,000)	4,34
	<i>ICER₂</i>	-8,731 (0,301)	-1,19	-3,710 (0,148)	-1,54	-4,916** (0,058)	-2,03
Modelos finais	<i>ICER_{div}</i>	4.226,86* (0,045)	2,05	3.834,57* (0,001)	4,60	25.001,22* (0,027)	2,38
	<i>ICER₂</i>	-	-	-	-	-23.315,59** (0,047)	-2,11
	R ²	0,458		0,585		0,607	
	R ² Ajustado	0,349		0,555		0,560	
	F	4,221** (0,095)		19,731* (0,001)		13,112* (0,000)	
	White (χ^2)	1,20 ^{NH} (1gl – 3,84)		2,01 ^{NH} (1gl – 3,84)		1,10 ^{NH} (2gl – 5,99)	
	Normalidade (SW)	<i>p</i> -valor = 0,011		<i>p</i> -valor = 0,001		<i>p</i> -valor = 0,001	
n	6		15		22		

(**) Significativos para *p*-valor < 0,10 e (*) Significativos para *p*-valor < 0,05. (NH) – Não há heterocedasticidade.

Tabela 5.9: Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre São Paulo a Canindé

São Paulo – Canindé							
	Variáveis	1ª ordem		2ª ordem		3ª ordem	
		Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>
Modelos iniciais	Intercepto	-	-	-	-	-	-
	<i>Pop_{div}</i>	-0,304 (0,317)	-1,08	-0,263* (0,050)	-2,12	-0,235** (0,085)	-1,81
	<i>Pop₂</i>	-	-	0,058* (0,000)	6,62	0,053* (0,000)	6,24
	<i>Renda_{div}</i>	-0,366 (0,325)	-1,06	-0,240 (0,108)	-1,70	0,055 (0,891)	0,14
	<i>Renda₂</i>	27,262* (0,024)	2,77	0,087 (0,694)	0,401	-0,063 (0,789)	-0,27
	<i>TUR_{div}</i>	-	-	-	-	-	-
	<i>TUR₂</i>	-0,322 (0,273)	-1,19	-0,215** (0,099)	-1,76	-0,159 (0,248)	-1,19
	<i>ICER_{div}</i>	-	-	0,009 (0,960)	0,05	-0,015 (0,937)	-0,08
	<i>ICER₂</i>	-	-	0,014 (0,942)	0,07	-0,016 (0,929)	-0,09
Modelos finais	<i>Pop_{div}</i>	-	-	-29,619* (0,050)	-2,12	-26,39** (0,085)	-1,81
	<i>Pop₂</i>	-	-	0,065* (0,000)	7,56	0,057* (0,000)	6,80
	<i>Renda₂</i>	27,262* (0,024)	2,77	0,087 (0,694)	0,401	-0,063 (0,789)	-0,27
	<i>TUR₂</i>	-	-	-1,469,93** (0,09)	-1,76	-0,159 (0,248)	-1,19
	R ²	0,490		0,819		0,699	
	R ² Ajustado	0,426		0,783		0,669	
	F	7,682* (0,024)		22,654* (0,000)		23,189* (0,000)	
	White (χ^2)	0,09 ^{NH} (1gl – 3,84)		1,55 ^{NH} (3gl – 7,81)		0,004 ^{NH} (2gl – 5,99)	
	Normalidade (SW)	<i>p</i> -valor = 0,134		<i>p</i> -valor = 0,555		<i>p</i> -valor = 0,496	
	n	9		18		22	

(**) Significativos para p-valor < 0,10 e (*) Significativos para p-valor < 0,05. (NH) – Não há heterocedasticidade.

Alguns fatores devem ser observados nestas tabelas. O primeiro é que todas as variáveis referentes à origem (X_1) receberam um tratamento (Equação 5.2) em relação ao valor original, já que essas variáveis poderiam acarretar graves problemas ao modelo. Apenas a variável TUR_1 não recebeu esse tratamento, já que não fazia sentido, por ser uma variável *dummy*. Com isso, essa foi excluída do modelo.

$$X_{div} = \frac{X_1}{X_2} \quad (5.2)$$

Outro ponto que merece consideração é o fato de o intercepto (constante) não fazer parte do modelo, já que apesar de a maioria dos estudos considerá-lo como a média do fenômeno observado, neste caso não teria sentido algum quando variáveis do tipo “população” e “rendimento” não fizessem parte do modelo. Buscou-se fazer alguns testes para o intercepto, mas nenhum deles foi significativo aos níveis de confiança 95% e 99%.

O último passo diz respeito à construção dos modelos quando se considera a área de influência de 1ª ordem, devido à quantidade mínima de informações. Deve-se utilizar menos variáveis independentes do que a quantidade de pontos (informações). Então, antes de se executar um modelo utilizando 1ª ordem, é importante verificar as variáveis mais significativas para aquela região e excluir algumas.

Nas simulações para as diferentes áreas de influência, há a necessidade de se verificar os dois principais pressupostos para a validade dos modelos: erros normalmente distribuídos (Teste de Shapiro-Wilk – *SW*) e com variância constante (Teste de White). O primeiro, sob a hipótese nula de *erros normalmente distribuídos*, e o segundo, sob a hipótese nula de *erros com variância não constante (heterocedasticidade)*.

Feito isso, o modelo não conseguiu obter uma explicação lógica no momento de estimar a demanda de passageiros que saem de São Paulo e vão para Crateús. Além de ter um dos pressupostos inválidos para as três áreas de influência (não-normalidade dos erros amostrais), a única variável significativa foi a variável *Índice de Cobertura Espacial Rodoviária* (ICER) que, isolada, não explica o fenômeno.

Em relação aos outros dois municípios, obtiveram-se melhores ajustes. Segundo os pressupostos, apenas para a sub-região formada pela 1ª ordem do município Crato não há normalidade nos erros amostrais. Além disso, o modelo está bem ajustado para os trechos entre *São Paulo-Crato* e *São Paulo-Canindé*, considerando o R^2 (**0,941** e **0,819**, respectivamente) da melhor área de influência segundo os resultados (regime espacial de 2ª ordem). As variáveis estudadas, até o momento, para a demanda de passageiros do transporte rodoviário interestadual e para a verificação da adequabilidade do modelo, serão vistas na seção 5.3 deste capítulo.

Dado os resultados para a origem em um município distante do Ceará, foi realizado o mesmo processo para Natal, um município mais próximo do estado. Os resultados estão apresentados nas tabelas 5.10, 5.11 e 5.12

Tabela 5.10: Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre Natal e Juazeiro do Norte

Natal – Juazeiro do Norte							
Variáveis	1ª ordem		2ª ordem		3ª ordem		
	Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>	
Modelos iniciais	Intercepto	-	-	-	-	-	-
	<i>Pop_{div}</i>	-0,663 (0,164)	-2,15	-1,06** (0,088)	-1,89	-0,142 (0,629)	-0,50
	<i>Pop₂</i>	-1,136 (0,111)	-2,75	-2,981* (0,033)	-2,47	-0,094 (0,712)	-0,37
	<i>Renda_{div}</i>	-	-	9,30** (0,067)	2,05	-0,987 (0,017)	-2,54
	<i>Renda₂</i>	1.866,31* (0,050)	4,27	12,39* (0,026)	2,61	1,52* (0,001)	3,83
	<i>TUR_{div}</i>	-	-	-	-	-	-
	<i>TUR₂</i>	-	-	-0,051 (0,85)	-0,19	0,183 (0,319)	1,02
	<i>ICER_{div}</i>	-	-	-13,88** (0,073)	-1,98	-6,514 (0,267)	-1,14
	<i>ICER₂</i>	-	-	-2,46 (0,681)	-0,42	-2,15 (0,648)	-0,46
Modelos finais	<i>Pop_{div}</i>	-	-	-3.679,31** (0,060)	-2,07	-	-
	<i>Pop₂</i>	-	-	-4,087* (0,014)	-2,88	-	-
	<i>Renda_{div}</i>	-	-	237.476,893* (0,034)	2,39	-21.955,42* (0,001)	-3,84
	<i>Renda₂</i>	1.866,31* (0,050)	4,27	6.738,22* (0,007)	3,22	672,718* (0,000)	5,08
	<i>ICER_{div}</i>	-	-	-862.706,33* (0,018)	-2,75	-	-
	R ²	0,928		0,772		0,481	
	R ² Ajustado	0,821		0,676		0,446	
	F	8,643 (0,10)		8,108 (0,002)		13,89 (0,000)	
	White (χ^2)	0,45 ^{NH} (1gl – 3,84)		0,24 ^{NH} (5gl – 11,07)		0,003 ^{NH} (4gl – 9,49)	
	Normalidade (SW)	<i>p</i> -valor = 0,751		<i>p</i> -valor = 0,869		<i>p</i> -valor = 0,476	
n	5		17		32		

(**) Significativos para *p*-valor < 0,10 e (*) Significativos para *p*-valor < 0,05. (NH) – Não há heterocedasticidade.

Tabela 5.11: Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre Natal e Aracati

Natal – Aracati							
	Variáveis	1ª ordem		2ª ordem		3ª ordem	
		Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>
Modelos iniciais	Intercepto	-	-	-	-	-	-
	Pop_{div}	-0,814** (0,059)	-2,44	-0,068 (0,781)	-0,29	0,203 (0,309)	1,04
	Pop_2	0,254 (0,472)	0,79	0,051* (0,002)	4,03	0,048* (0,000)	5,71
	$Renda_{div}$	0,294 (0,710)	0,40	-0,137 (0,689)	-0,41	0,110 (0,654)	0,45
	$Renda_2$	0,830* (0,011)	3,64	0,397 (0,385)	0,91	-0,403** (0,092)	-1,76
	TUR_{div}	-	-	-	-	-	-
	TUR_2	-	-	0,301 (0,238)	1,25	0,447* (0,004)	3,59
	$ICER_{div}$	-	-	0,249 (0,506)	0,69	0,559 (0,330)	0,99
	$ICER_2$	-	-	0,230 (0,538)	0,64	0,370 (0,554)	0,60
Modelos finais	Pop_{div}	-270,606** (0,059)	-2,44	-	-	-	-
	Pop_2	-	-	0,051* (0,002)	4,03	0,053* (0,002)	3,60
	$Renda_2$	39,607* (0,006)	4,59	-	-	-9,918** (0,092)	-1,76
	TUR_2	-	-	-	-	2.155,76* (0,002)	3,59
	R^2	0,858		0,596		0,744	
	$R^2_{Ajustado}$	0,801		0,559		0,709	
	F	15,057 (0,008)		16,209 (0,002)		21,321 (0,000)	
	White (χ^2)	4,03 ^{NH} (2gl – 5,99)		3,72 ^{NH} (1gl – 3,84)		0,725 ^{NH} (3gl – 7,81)	
	Normalidade (SW)	<i>p</i> -valor = 0,044		<i>p</i> -valor = 0,327		<i>p</i> -valor = 0,021	
	n	7		12		25	

(**) Significativos para p -valor < 0,10 e (*) Significativos para p -valor < 0,05. (NH) – Não há heterocedasticidade.

Tabela 5.12: Resultados do modelo de regressão espacial com efeitos locais da forma discreta para a viagem entre Natal e Região Metropolitana de Fortaleza

Natal – RM de Fortaleza							
	Variáveis	1ª ordem		2ª ordem		3ª ordem	
		Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>
Modelos iniciais	Intercepto	-	-	-	-	-	-
	<i>Pop_{div}</i>	-0,843* (0,029)	-2,50	-0,620* (0,046)	-2,10	-0,346* (0,004)	-2,99
	<i>Pop₂</i>	-0,008* (0,000)	-3,53	-0,321 (0,199)	-1,32	0,393 (0,168)	1,40
	<i>Renda_{div}</i>	-0,233 (0,707)	-0,39	-0,076 (0,855)	-0,18	-2,07* (0,042)	-2,09
	<i>Renda₂</i>	2,02* (0,005)	3,62	1,484* (0,000)	4,13	-1,494 (0,182)	-1,36
	<i>TUR_{div}</i>	-	-	-	-	-	-
	<i>TUR₂</i>	-0,272 (0,506)	-0,70	0,096 (0,615)	0,51	0,238** (0,061)	2,19
	<i>ICER_{div}</i>	-7,19 (0,405)	-0,89	-1,523 (0,722)	-0,36	-1,197 (0,666)	-0,435
	<i>ICER₂</i>	14,56 (0,178)	1,50	3,028 (0,557)	0,60	3,946* (0,032)	2,21
Modelos finais	<i>Pop_{div}</i>	-901,887* (0,033)	-2,44	-251,796* (0,034)	-2,23	-230,937* (0,004)	-2,99
	<i>Pop₂</i>	-0,0057* (0,049)	-3,53	-	-	-	-
	<i>Renda_{div}</i>	-	-	-	-	-3.837,857** (0,054)	-1,97
	<i>Renda₂</i>	172,665* (0,003)	3,88	121,848* (0,000)	5,57	-	-
	<i>TUR₂</i>	-	-	-	-	3.830,49* (0,033)	2,19
	<i>ICER₂</i>	-	-	-	-	13.597,873* (0,000)	3,81
	R ²	0,699		0,592		0,583	
	R ² _{Ajustado}	0,617		0,563		0,548	
	F	8,52 (0,003)		20,35 (0,000)		16,76 (0,000)	
	White (χ^2)	1,29 ^{NH} (2gl – 5,99)		0,012 ^{NH} (3gl – 7,81)		6,63 ^{NH} (4gl – 9,49)	
	Normalidade (SW)	<i>p</i> -valor = 0,23		<i>p</i> -valor = 0,000		<i>p</i> -valor = 0,000	
n	14		29		51		

(**) Significativos para *p*-valor < 0,10 e (*) Significativos para *p*-valor < 0,05. (NH) – Não há heterocedasticidade.

Observa-se que a exemplo de Crato, o modelo proposto não é o mais adequado para Juazeiro do Norte quando a área de influência é de 1ª ordem, provavelmente pela quantidade de pontos. Já segundo os testes realizados, a partir da 2ª ordem, os resultados foram significativos. Nesta área de influência, encontrou-se um R² que apresenta dados relativamente bem ajustados à reta de regressão (0,772).

O município de Aracati apresentou erros não-normais nas 1ª e 3ª ordens, e na 2ª ordem apenas uma variável independente foi significativa. Assim, pode-se deduzir que o modelo

proposto não será o mais indicado para estimar a demanda de passageiros nas viagens *Natal-Aracati*.

No caso da RMF, quando se analisa as 2ª e 3ª ordens, tem-se erros não-normais, o que pode acarretar a superestimação e a subestimação dos parâmetros da regressão. Isso significará que o modelo só terá sentido quando trabalhado na 1ª ordem (apenas os vizinhos da RMF). Um resultado esperado, pois, pela área territorial da RMF, “os vizinhos dos vizinhos” já serão municípios mais próximos das 1ª e 2ª ordens de Canindé e Sobral (municípios com terminais que receberam passageiros).

Os resultados para RMF foram diferentes de Crato e Juazeiro do Norte por dois motivos visíveis. O primeiro pelo tamanho da área, que é composta por 13 municípios, e o segundo pelo banco de dados não contabilizar os passageiros que fazem viagens dentro da RMF.

5.3. ANÁLISE DA ADEQUABILIDADE DO MODELO E INTERPRETAÇÃO DOS PARÂMETROS

Além do coeficiente de determinação (R^2), que verificou o ajuste dos modelos de regressão com efeitos locais discretos, é necessário averiguar a confiabilidade e a consistência. Dessa forma, aplicou-se a Equação 5.2 (Erro Relativo – *ER*) e a Equação 4.3, do capítulo 4 (Erro Relativo Médio - *ERM*), em que o número real será a demanda de passageiros de cada O/D do ano seguinte (2005).

$$ER = |YE - YO| \quad (5.2)$$

onde,

YE = número estimado pelo modelo;

YO = número observado (real);

Essa técnica foi aplicada em todos os municípios que obtiveram resultados significativos (Crato, Canindé, Juazeiro do Norte e Região Metropolitana de Fortaleza) e áreas de influência de 2ª ordem, exceto a RMF, para o qual utilizou-se de 1ª ordem, conforme a Tabela 5.13.

Tabela 5.13: Números estimados, observados, e erros relativos e médios para a demanda de passageiros de cada O/D no ano de 2005.

Município Origem	Município Destino	Número Estimado (YE)	Número Observado (YO)	Erro Relativo (ER)	Erro Relativo Médio (ERM)
São Paulo	Crato	1.534	283	1.251	2,20
São Paulo	Canindé	494	218	276	7,13
Natal	Juazeiro do Norte	- 310	168	478	5,86
Natal	Reg. Metrop. de Fortaleza	964	3.759	2.795	1,67

Analisando a Tabela 5.13, observa-se elevados Erros Relativos para cada O/D significativa, o que sustenta a não adequabilidade do modelo proposto. Porém, como mencionado no início do capítulo, a variável *renda* poderia influenciar os valores estimados, visto que é uma informação do ano de 2000. Desse modo, buscou-se estimá-la para o ano de 2005, após a calibração do modelo por meio da variação percentual, entre os anos de 1991 e 2000, de cada município envolvido no processo.

Exemplificando o método realizado: Crato obteve uma renda média per capita em 1991 de R\$ 114,67 e em 2000 de R\$ 168,05 (uma variação de 78%). Assim, utilizou-se a mesma variação na estimação da renda do ano de 2005 e a Tabela 5.14 apresenta os resultados. É importante salientar que a variável *renda* não foi significativa na O/D *São Paulo-Canindé*.

Tabela 5.14: Números estimados, observados, e erros relativos e médios para a demanda de passageiros de cada O/D, com a variável *renda* estimada para 2005.

Município Origem	Município Destino	Número Estimado (YE)	Número Observado (YO)	Erro Relativo (ER)	Erro Relativo Médio (ERM)
São Paulo	Crato	428	283	145	0,75
São Paulo	Canindé	494	218	276	7,13
Natal	Juazeiro do Norte	269	168	101	5,16
Natal	Reg. Metrop. de Fortaleza	3.917	3.759	158	1,61

De acordo com a Tabela 5.14, pode-se afirmar para *São Paulo-Canindé* obteve um elevado Erro Relativo (aproximadamente 56% a mais do que a demanda real de passageiros). Mas nos casos de *São Paulo-Crato* e *Natal-Juazeiro do Norte*, já se observa erros relativamente moderados (33% e 37%, respectivamente). E, por último, na O/D *Natal-RMF*, com erro suficientemente baixo (4%). Vale ressaltar que os casos *São Paulo-Crato* e *Natal-RMF* obtiveram *ERM* relativamente baixos (de um a dois passageiros em média).

Apesar do elevado erro para *São Paulo-Canindé* (em média, 7 passageiros), mesmo compreendendo que o município de Canindé foi considerado *High-Low* e, com isso, o tipo

de modelagem não é a mais apropriada; vale à pena relatar outro fato: em todas as O/D houve um decréscimo, em 2005, na demanda de passageiros com relação a 2004, conforme a Tabela 5.15. Dessa forma, o Erro Relativo seria menor se houvesse um crescimento, constante ou não, da demanda de passageiros para todas as O/D, inclusive para O/D *São Paulo-Canindé*, já que variáveis relevantes, como renda e população, se alteram ano a ano.

Tabela 5.15: Relação da demanda de passageiros (2004-2005).

Município Origem	Município Destino	Número Observado (YO) - 2004	Número Observado (YO) - 2005
São Paulo	Crato	403	283
São Paulo	Canindé	274	218
Natal	Juazeiro do Norte	205	168
Natal	Reg. Metrop. de Fortaleza	3.981	3.759

Outra forma de se verificar a adequabilidade do modelo é selecionar um dos municípios *High-High* do estado do Ceará e realizar a simulação com outros municípios. Nesse intuito, optou-se, como destino, por Crato, com área de influência de 2ª ordem, e como origem quatro municípios de características distintas e distâncias rodoviárias variadas (Juazeiro/BA, Ouricuri/PE, Rio de Janeiro/RJ e Salgueiro/PE). Os resultados estão apresentados na Tabela 5.16.

Tabela 5.16: Resultados do modelo de regressão espacial com efeitos locais da forma discreta para passageiros em viagens até Crato

		Destino Crato			
		Juazeiro/BA		Ouricuri/PE	
Variáveis		Parâmetro	Teste <i>t</i>	Parâmetro	Teste <i>t</i>
Modelos iniciais	Intercepto	-	-	-	-
	Pop_{div}	0,344* (0,002)	4,02	0,405* (0,005)	3,35
	Pop_2	1,288* (0,000)	19,66	0,873* (0,000)	6,92
	$Renda_{div}$	-1,898* (0,000)	-5,11	0,185 (0,360)	0,95
	$Renda_2$	-2,961* (0,000)	-8,87	-0,516* (0,006)	-3,212
	TUR_{div}	-	-	-	-
	TUR_2	0,126 (0,187)	1,69	0,024 (0,845)	0,20
	$ICER_{div}$	-1,272* (0,153)	-1,74	0,493 (0,142)	1,57
	$ICER_2$	-1,796 (0,382)	-0,91	0,494 (0,170)	1,46
Modelos finais	Pop_{div}	2.305,619* (0,004)	3,62	2.713,492* (0,005)	3,35
	Pop_2	1,288* (0,000)	17,70	1,174* (0,000)	10,39

	$Renda_{div}$	-40.564,431* (0,001)	-4,32	-	-
	$Renda_2$	-531,215* (0,000)	-7,90	-544,210* (0,000)	-5,27
	R^2	0,977		0,926	
	$R^2_{Ajustado}$	0,967		0,909	
	F	95,114* (0,000)		54,397* (0,000)	
	White (χ^2)	1,920 ^{NH} (5gl – 11,07)		2,320 ^{NH} (3gl – 7,81)	
	Normalidade (SW)	p -valor = 0,755		p -valor = 0,168	
	n	16		16	

Destino Crato

	Variáveis	Rio de Janeiro/RJ		Salgueiro/PE	
		Parâmetro	Teste t	Parâmetro	Teste t
Modelos iniciais	Intercepto	-	-	-	-
	Pop_{div}	-0,435* (0,000)	-4,86	2.724,211** (0,066)	2,02
	Pop_2	0,646* (0,000)	6,90	1,612* (0,000)	10,17
	$Renda_{div}$	-0,420 (0,734)	-0,35	0,164 (0,431)	0,82
	$Renda_2$	-0,194 (0,263)	-1,17	-1,114* (0,000)	-4,83
	TUR_{div}	-	-	-	-
	TUR_2	-0,491 (0,690)	-0,41	0,024 (0,846)	0,20
	$ICER_{div}$	-0,550 (0,644)	-0,47	0,494 (0,145)	1,56
	$ICER_2$	-0,570 (0,633)	-0,49	-1,764 (0,375)	-0,93
Modelos finais	Pop_{div}	-1.669,410* (0,000)	-4,86	2.724,211* (0,005)	3,35
	Pop_2	0,792* (0,000)	11,94	1,174* (0,000)	10,36
	$Renda_2$	-	-	-545,722* (0,000)	-5,27
	R^2	0,911		0,926	
	$R^2_{Ajustado}$	0,898		0,908	
	F	71,43* (0,000)		53,923* (0,000)	
	White (χ^2)	0,134 ^{NH} (2gl – 5,99)		2,320 ^{NH} (3gl – 7,81)	
	Normalidade (SW)	p -valor = 0,762		p -valor = 0,166	
	n	16		16	

(**) Significativos para p -valor < 0,10 e (*) Significativos para p -valor < 0,05. (NH) – Não há heterocedasticidade.

De acordo com a Tabela 5.16, constata-se que as hipóteses de não heterocedasticidade e de normalidade dos erros não foram rejeitadas e houve um bom ajuste dos modelos (97,7% para o modelo com Juazeiro/BA, 92,6% para o modelo com Ouricuri/PE, 91,1% para o modelo com Rio de Janeiro/RJ e 92,6% para o modelo com Salgueiro/PE).

A Tabela 5.17 apresenta os números estimados, observados, e os erros relativos e médios para a demanda de passageiros que se deslocaram dos quatro municípios para Crato, com a variável *renda* estimada para 2005. Verifica-se erros relativos moderados e erros relativos médios de, no máximo, um passageiro, como a O/D *São Paulo-Crato*.

Tabela 5.17: Números estimados, observados, e erros relativos e médios para a demanda de passageiros de cada O/D para 2005.

Município Origem	Município Destino	Número Estimado (YE)	Número Observado (YO)	Erro Relativo (ER)	Erro Relativo Médio (ERM)
Juazeiro	Crato	186	108	78	0,97
Ouricuri	Crato	629	502	127	0,25
Rio de Janeiro	Crato	251	173	78	0,95
Salgueiro	Crato	148	60	88	1,02

Após a análise da adequabilidade do modelo de regressão com efeitos espaciais locais, o passo seguinte é interpretar cada parâmetro com suas particularidades. Para isso, utilizou-se as Tabelas 5.7, 5.9, 5.10, 5.12 e 5.16 e, exclusivamente, a área de influência de 2ª ordem (somente em *Natal-RMF* utilizou-se 1ª ordem).

Analisando os parâmetros, quando o destino é Crato, verifica-se uma variação no sinal do parâmetro da variável Pop_{div} . Conforme mostrado nas tabelas 5.7 e 5.16, quando um determinado município é intensamente populoso, o sinal do parâmetro é negativo. Isso significa que se a população do município origem for excessivamente maior em relação à população de Crato e seus vizinhos (São Paulo – 98 vezes maior e Rio de Janeiro – 55 vezes maior), menor é a demanda de passageiros em viagens interestaduais com destino a essa sub-região. A possível explicação para este fato é que, apesar do crescimento populacional dessas metrópoles, há menos originários de Crato, de Canindé (onde a variável Pop_{div} também foi negativa) e de suas respectivas áreas de influência, residindo em esses municípios pólos do país.

Exemplificando o que foi mencionado no parágrafo anterior: segundo o IBGE, em 2005, São Paulo tinha 10,93 milhões, e Crato, 113,5 mil habitantes. Calculando a variável Pop_{div} , encontrou-se aproximadamente um valor absoluto de 97 e isso quer dizer que a demanda de passageiros diminuiria em 70 mil passageiros entre *São Paulo-Crato*. Este resultado tem o objetivo de mostrar somente o quantitativo da variável, visto que existem outras variáveis significativas no modelo. Ou seja, enquanto o parâmetro da variável Pop_{div} influencia de

modo negativamente o modelo, existirá outro parâmetro que influenciará positivamente, já que é improvável o deslocamento de 70 mil passageiros de São Paulo com destino a Crato em um mês.

Análogo ao caso do parâmetro da variável Pop_{div} , os parâmetros relacionados à renda ($Renda_{div}$ e $Renda_2$) – em que todos os municípios foram confrontados com Crato – obtiveram sinais negativos. Dessa forma, quanto maior a renda, principalmente de Crato e da área de influência, menor será a demanda de passageiros em viagens interestaduais com destino a essa sub-região. Este resultado é relativamente esperado, porque as reduções tarifárias do transporte aéreo, juntamente com o crescimento da renda média per capita da população, fazem presumir uma queda na demanda de passageiros do transporte rodoviário interestadual.

No caso do parâmetro da variável Pop_2 , no município Crato, não houve qualquer variação nos sinais dos cinco municípios estudados. Ou seja, quanto maior a população de Crato e da área de influência (2ª ordem), maior a demanda de passageiros em viagens interestaduais com destino a essa sub-região.

Quando analisadas as variáveis $Renda_{div}$, $Renda_2$ e Pop_2 , em Juazeiro do Norte e na Região Metropolitana de Fortaleza, com suas respectivas áreas de influência, observam-se sinais diferenciados dos parâmetros analisados em Crato, conforme mostrado nas Tabelas 5.10 e 5.12.

Em relação às variáveis $Renda_{div}$ e $Renda_2$, quanto maior a renda de Juazeiro do Norte e da RMF, maior a demanda de passageiros em viagens interestaduais com destino a essas sub-regiões. Nesse caso, o resultado faz algum sentido, pois a distância rodoviária é relativamente pequena entre a origem (Natal) e os destinos comparados, por exemplo, a *São Paulo-Crato* e *Rio de Janeiro-Crato*. Com isso, aparentemente, as tarifas do transporte aéreo não influenciam a demanda de passageiros que se deslocam entre *Natal-Juazeiro do Norte* e *Natal-RMF*.

As variáveis Pop_{div} e Pop_2 , tanto para Juazeiro do Norte quanto para RMF, obtiveram sinais negativos. Dessa forma, no caso do parâmetro da Pop_{div} , se a população do município origem for maior que a população de Juazeiro do Norte e da RMF, menor será a

demanda de passageiros em viagens interestaduais para as sub-regiões. Caso contrário, a demanda de passageiros em viagens interestaduais continuará menor para essas sub-regiões, mas com menos intensidade.

Uma explicação para este fator é que a população destino influencia em uma maior demanda de passageiros se deslocando em viagens interestaduais, conforme indicado nas duas variáveis, e, desse modo, é possível afirmar que a população de Juazeiro do Norte e da RMF, com as suas respectivas áreas de influência, têm viajado mais de transporte rodoviário interestadual.

Analisando os parâmetros das variáveis $ICER_{div}$, $ICER_2$ e TUR_2 não se obteve respostas muito concretas. A variável TUR_2 (municípios pólos turísticos) é significativa somente quando analisada nos dados relacionados à Canindé e à sua área de influência, e, mesmo assim, quando aplicada exclusivamente ao município não faz sentido algum, já que Canindé não é classificado como pólo turístico (EMBRATUR/FIPE *apud* TCU, 2005).

No caso do índice de cobertura espacial rodoviária, somente o parâmetro da variável $ICER_{div}$ obteve, negativamente, influência no resultado da O/D Natal-Juazeiro do Norte, conforme a Tabela 5.10. Assim, para um aumento no $ICER$ do município de origem maior do que no $ICER$ de Juazeiro do Norte, é necessária uma menor demanda de passageiros em viagens interestaduais para as sub-regiões. Caso contrário, a demanda de passageiros em viagens interestaduais continuará menor para essa sub-região, mas com menos intensidade.

Nesse contexto, verifica-se que a influência de uma rodovia em relação a uma determinada sub-região é mais expressiva nos municípios destinos do que nos municípios origens, em relação ao acréscimo de passageiros em viagens interestaduais. Supondo um $ICER = 0,8$ para Juazeiro do Norte e um $ICER = 1,0$ para Natal, isso significaria que o $ICER_{div} = 1,25$ e a demanda de passageiros diminuiriam em 1,7 milhão. Ao contrário, com $ICER = 1,0$ para Juazeiro do Norte e um $ICER = 0,8$ para Natal, tem-se que o $ICER_{div} = 0,8$ e a demanda de passageiros diminuiriam em 690 mil, uma diferença de aproximadamente um milhão de passageiros. Vale ressaltar que somente este resultado não faz sentido para o modelo, já que existem outras variáveis.

5.4. TÓPICOS CONCLUSIVOS

- O presente capítulo teve como objetivo verificar a aplicabilidade do modelo de regressão com efeitos espaciais locais utilizando-se de um banco de dados com informação de passageiros em viagens interestaduais (ANTT) e um banco de dados com informações de passageiros em viagens intermunicipais (PDOTIP-CE).
- A demanda de passageiros do transporte rodoviário interestadual não apresentou dependência espacial, mas, sub-regiões com padrões espaciais (não-estacionaridade). Já a demanda de passageiros do transporte rodoviário intermunicipal apresentou dependência espacial, além de não-estacionaridade.
- Utilizando o *Box Map* e o *Lisa Map*, foi possível identificar sub-regiões pólos, dentro do estado, capazes de receber passageiros de viagens interestaduais.
- Modelos de regressão com efeitos espaciais locais discretos apresentaram bons ajustes aos dados e erros razoáveis, além de que conseguiram refletir a demanda de passageiros do transporte rodoviário interestadual quando a origem apresenta regime espacial *High-High* ou *High-Low* e o destino apresenta regime espacial *High-High*. Nos outros casos, o modelo não apresentou resultados satisfatórios.
- Diferentemente do discreto, os modelos de regressão com efeitos espaciais locais contínuos não obtiveram resultados satisfatórios em nenhum dos casos, provavelmente pela quantidade de zeros na matriz O/D.
- Por meio desses modelos, observou-se que as variáveis *população* e *renda* realmente influenciam o deslocamento de pessoas no âmbito interestadual pelo modo ônibus e que, dependendo da situação, essa influência pode ocorrer de forma positiva ou negativa.

6. CONCLUSÕES E RECOMENDAÇÕES

6.1. APRESENTAÇÃO

A pesquisa apresentada neste trabalho permite enunciar algumas conclusões e sugestões futuras para a modelagem da demanda de passageiros do transporte rodoviário interestadual por meio de ferramentas de estatística espacial mais avançadas (análise exploratória e confirmatória).

Estas conclusões foram elaboradas a fim de representar os principais resultados do modelo proposto, além de propiciar a discussão a cerca da melhor forma de modelagem da demanda do transporte rodoviário interestadual e intermunicipal de passageiros e a importância da utilização de técnicas de análise espacial para esse fenômeno.

6.2. AVALIAÇÃO DA CONTRIBUIÇÃO DO MODELO PROPOSTO

O transporte rodoviário de passageiros é uma atividade altamente expressiva na estrutura econômica e social de uma região. Dessa forma, estudos que busquem conhecer características de um determinado lugar e a demanda de passageiros do transporte público têm uma importante contribuição ao planejamento e à regulação do serviço.

Diversas são as ferramentas utilizadas para o planejamento em transporte público, principalmente no cálculo da demanda de passageiros. Portanto, a grande maioria das ferramentas é utilizada para o transporte urbano, já que os fatores como finalidade e comportamento da viagem são conhecidos.

Nos últimos anos, pesquisadores têm buscado novas técnicas para estimar a demanda de passageiros do transporte rodoviário interestadual, já que a utilizada pela ANTT (modelos gravitacionais) não tem produzido resultados satisfatórios.

Nesse contexto, ferramentas de análise espacial se apresentam como técnicas avançadas à compreensão de determinados fenômenos de exploração e modelagem de dados com caráter espacial, garantindo ao planejador em transporte uma análise mais consistente.

A metodologia apresentada neste trabalho procurou sistematizar o estudo da demanda de passageiros do transporte rodoviário, levando em consideração as relações espaciais existentes entre cada par de O/D. Portanto, buscou-se verificar a dependência espacial nessa variável e a não-estacionariedade do processo, pois somente desse modo pode-se utilizar modelos de regressão com efeitos espaciais locais.

Juntamente com as informações de sub-regiões, com o mesmo padrão espacial, e da dependência espacial em relação aos passageiros em viagens interestaduais e intermunicipais, a hipótese da pesquisa foi corroborada, uma vez que esta técnica obteve resultados que representaram de forma realista o fenômeno do TRIP. Isso porque os modelos desenvolvidos nesta pesquisa utilizaram sub-regiões criadas por meio das variáveis socioeconômicas e dos passageiros em viagens intermunicipais. Portanto, modelos que estudam cada sub-região com suas particularidades permitem entender melhor os deslocamentos, e, sobretudo, oferecem subsídio às tomadas de decisão estratégicas de se criar ou retirar linhas do Trip.

Exemplificando esta afirmação, observa-se o município de Crato, que constitui uma das maiores economias do sul cearense, que originou a cidade de Juazeiro do Norte e constituiu também um entroncamento rodoviário para três estados do Nordeste, além da Região Metropolitana de Fortaleza. Dessa forma, como se esperava, o tipo de modelo proposto foi satisfatório e obteve ajustes mais favoráveis quando a área de influência foi considerada de 2ª e 3ª ordem. Pode-se concluir então que a influência do município de Crato ultrapassa suas fronteiras e valida, portanto, a hipótese de que o espaço, juntamente com suas atividades e dinâmicas, influencia a demanda de passageiros do transporte rodoviário interestadual.

Pode-se concluir também que a matriz de proximidade espacial (W), seja de contigüidade ou de distâncias rodoviárias entre os municípios, obteve resultados interessantes. Se houver possibilidade, deve-se utilizá-las sempre em conjunto com ferramentas de análise espacial. Porém, vale ressaltar, as condições das rodovias ou qualquer outro fator externo pode atrair ou repelir a demanda de passageiros e somente matrizes de proximidade espacial, diferentemente da forma binária, têm a possibilidade de identificar esse fenômeno.

Por fim, a interpretação dos parâmetros é essencial para o entendimento do fenômeno. Modelos de regressão com efeitos espaciais locais têm a mesma estrutura do modelo de regressão convencional no que diz respeito à interpretação desses, ou seja, de simples compreensão. Assim, mostrou-se que variáveis como população e renda sempre preponderam na demanda de passageiros em transporte público, seja no transporte urbano, intermunicipal ou interestadual.

6.3. AVALIAÇÃO DAS FRAGILIDADES DO MODELO PROPOSTO

Apesar dos resultados satisfatórios do método proposto por este trabalho, é necessário ressaltar algumas restrições relacionadas ao processo de construção e interpretação do modelo de regressão com efeitos espaciais locais:

- A confiabilidade e a qualidade dos dados são essenciais a construção de qualquer tipo de modelo. Os bancos de dados relacionados à ANTT e ao PDOTIP-CE foram as principais informações utilizadas na construção do modelo proposto. Portanto, a análise da confiabilidade e da qualidade desses bancos foi superficial, já que essas informações são as divulgadas à sociedade.
- Um ponto importante, e provavelmente a maior fragilidade na utilização desse modelo proposto, é que grande maioria dos estados não possuem a cultura de construir uma matriz O/D para a quantidade de passageiros em viagens intermunicipais no modo ônibus.
- O modelo apresentou resultados satisfatórios para o cálculo da demanda de passageiros do TRIP somente quando o município origem apresentou regime espacial *High-High* ou *High-Low* e o município destino um regime espacial *High-High*. Portanto, o município Canindé, mesmo com regime espacial *High-Low*, conseguiu refletir resultados significativos. Ou seja, existe a possibilidade de municípios considerados *High-Low* obterem resultados representativos, mas não é recomendável a sua utilização.
- Segundo as informações obtidas no SISDAP (ANTT), a quantidade de passageiros que se deslocaram em 2004, na área de estudo, foi maior do que em 2005,

fenômeno que o modelo proposto não consegue identificar com as variáveis utilizadas.

6.4. RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Como visto, este tipo de modelagem é de suma importância ao planejamento em transporte, o que supõe a sua utilização em outras atividades relacionadas à área. Nesse contexto, seguem algumas recomendações e sugestões para trabalhos futuros:

- O tempo de viagem é um fator que influencia a demanda do transporte rodoviário de passageiros. Dessa forma, além de trabalhar com as matrizes W de contigüidade e das distâncias rodoviárias, recomenda-se, se possível, a utilização do tempo médio de viagem.
- Para todo tipo de modelagem regressão linear, as variáveis independentes devem ser do mesmo período (mês, ano, outros) do fenômeno estudado. Assim, se alguma variável estiver em um período diferenciado, vale a pena estimá-la para o uso no modelo, tanto antes como depois da calibração.
- A modelagem da demanda de passageiros do transporte rodoviário interestadual, por meio da regressão com efeitos espaciais locais de forma discreta, apresentou bons ajustes quando se utilizou a influência dos passageiros em viagens intermunicipais do Ceará, além das variáveis socioeconômicas dos seus municípios. Contudo, existe a necessidade de se verificar o mesmo processo em viagens intermunicipais de outros estados destinos de viagens interestaduais e até mesmo, utilizando-se as informações das viagens intermunicipais do estado origem.
- A modelagem da demanda, utilizando regressão com efeitos espaciais locais da forma contínua (RGP), não apresentou resultados satisfatórios. Todavia, esse tipo de modelo provavelmente terá resultados melhores no cálculo da demanda do transporte urbano ou metropolitano, devido à formação da matriz O/D , que certamente estará com uma quantidade mínima de zeros em sua estrutura.

- Conhecer a demanda de passageiros do transporte rodoviário interestadual é de fundamental importância ao planejamento na área. Todavia, estudar a oferta do transporte, seja ela aérea ou terrestre, é útil para melhor entender os comportamentos e as atitudes preponderantes das pessoas em relação às suas necessidades e para conhecer, de forma mais expressiva, a procura destes passageiros pelo serviço de transporte rodoviário interestadual.

BIBLIOGRAFIA

- ABREU, V. C. de. Desenvolvimento de uma Metodologia de distribuição de Viagens com Aplicação da Lógica *Fuzzy*. Brasília: UnB/DF, Publicação TU.DM-08A/99, Dissertação de Mestrado, Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, 1999.
- AGUIAR JUNIOR, S. R.; BRASIL, L. M.; FIGUEIREDO, A. S. Aplicando Mineração de Dados para Estimar Demanda por Transporte Rodoviário Interestadual de Passageiros. Panorama Nacional da Pesquisa em Transportes, Congresso da ANPET, Brasília, 2006.
- ANSELIN, L. Spatial Data Analysis with GIS: An Introduction to Application in the Social Sciences. National Center for Geographic Information and Analysis. University of California, Santa Barbara, USA, 1992.
- ANSELIN, L. Local Indicators of Spatial Association – Lisa. Geographical Analysis. USA, 1995. Disponível em: <<http://66.102.1.104/scholar?hl=pt-BR&lr=&q=cache:rNs0Vzm6q3YJ:www.drs.wisc.edu/personnel/faculty/voss/Courses/RS977/readings/week%2520/Anselin%25201995%2520LISA.pdf+%22Local+indicators+of+spatial+association%22>>. Acesso em: 01/07/2007.
- ANSELIN, L. Exploratory Spatial Data Analysis Linking SpaceStat and ArcView. In: FISCHER, M. M. and GETIS, A. Recent Developments in Spatial Analysis – Spatial Statistics, Behavioural Modelling and Computational Intelligence. Berlin, Germany: Springer, 1997.
- ANTT, Agência Nacional de Transportes Terrestres. Brasília, ANTT. Glossário dos termos e conceitos utilizados pela Agência Nacional de Transportes Terrestres Resolução nº 16, 2002. Disponível em: <http://www.antt.gov.br/resolucoes/00100/resolucao016_2002.htm>. Acesso em: 04/02/2007.
- ANTT, Agência Nacional de Transportes Terrestres. Brasília, ANTT. Disponível em: <<http://www.antt.gov.br/passageiro/apresentacaopas.asp>>. Acesso em: 04/02/2007a.
- ANTT, Agência Nacional de Transportes Terrestres. Brasília, ANTT. Anuário Estatístico 2001 - Ano Base 2000. Quadros Estatísticos Relativos aos Serviços de Transporte Rodoviário Interestadual e Internacional de Passageiros Disponível em: <<http://www.antt.gov.br/passageiro/anuarios/anuario2001/411a415.htm>>. Acesso em: 11/07/2007b.
- ANTT, Agência Nacional de Transportes Terrestres. Brasília, ANTT. Anuário Estatístico 2001 - Ano Base 2000. Resumo geral dos dados operacionais – 2006. Disponível em: <<http://www.antt.gov.br/passageiro/anuarios/anuario2007/411.htm>>. Acesso em: 11/07/2007c.
- ANTT, Agência Nacional de Transportes Terrestres. Brasília, ANTT. Disponível em: <http://appeantt.antt.gov.br/acpublicas/apublica2007_72/AnexoVIII.pdf>. Acesso em: 04/01/2008.
- ANTT, Agência Nacional de Transportes Terrestres. Brasília, ANTT. Definição de linha no glossário dos termos e conceitos utilizados pela Agência Nacional de Transportes Terrestres, 2008. Disponível em: <http://www.antt.gov.br/resolucoes/04000/resolucao2850_2008.htm>. Acesso em: 11/09/2008.
- ARAÚJO, C. E. F. Análise Espacial da Geração de Viagens do Transporte Intermunicipal de Passageiros do Estado Ceará. Grupo de Pesquisa em Transporte, Trânsito e Meio Ambiente, Departamento de Engenharia de Transporte da Universidade Federal do Ceará, 2005.

- BRASIL. Decreto nº 2.521, de 20 de março de 1998. Dispõe sobre a exploração mediante permissão e autorização de serviços de transporte rodoviário interestadual e internacional de passageiros e dá outras providências, 1998.
- BRASIL. Lei nº 10.233, de 5 de Junho de 2001. Dispõe sobre a reestruturação dos transportes aquaviário e terrestre, cria o Conselho Nacional de Integração de Políticas de Transporte, a Agência Nacional de Transportes Terrestres, a Agência Nacional de Transportes Aquaviários e o Departamento Nacional de Infra-Estrutura de Transportes, e dá outras providências, 2001.
- BRASILEIRO, A.; SANTOS, E. M.; ARAGÃO, J. J. G.; SENNA, J. M.; NETO, O. L.; FILHO, R. D. O. Transporte no Brasil: História e Reflexões. Empresa Brasileira de Planejamento de Transportes/GEIPOT, Recife: Ed. Universitária da UFPE, 2001.
- BRASILEIRO, A.; SANTOS, E.M.; ARAGÃO, J.J.G.; SILVA, A.J.T. Agências Regulatórias e Organismos de Tutela Nacionais: O Transporte Interestadual de Passageiros no Brasil, Congresso Latinoamericano de Transporte Público y Urbano, CLATPU, La Habana, 2001.
- BRUNSDON, C.A.; FOTHERINGHAM, A.S.; CHARLTON, M.E. Geographically weighted regression - Modeling Spatial Nonstationary. *The Statistician* 47,43 1-443. NY, USA: JSTOR, 1998.
- BRUTTON, M. J. Introdução ao Planejamento dos Transportes. 2.ed. São Paulo: Universidade de São Paulo, 1979.
- CÂMARA, G. Sistemas de Informação Geográfica para aplicação ambientais e cadastrais: Uma visão Geral. Instituto Nacional de Pesquisas Espaciais – INPE, Divisão de Processamento de Imagens. São José dos Campos, 1999.
- CÂMARA, G.; MONTEIRO, A. M. V.; CARVALHO, M. S. Análise Espacial e Geoprocessamento. In: DRUCK, S.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. M. V. Análise Espacial de Dados Geográficos. Instituto Nacional de Pesquisas Espaciais – Divisão de Processamento de Imagens, São José dos Campos, São Paulo, 2000.
- CÂMARA, G.; CARVALHO, M. S.; CRUZ, O. G.; CORREIA, V. Análise Espacial de Áreas. In: DRUCK, S.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. M. V. Análise Espacial de Dados Geográficos. Planaltina, Brasília, DF: Ed. Embrapa Cerrado, 2004.
- CÂMARA, G.; MONTEIRO, A. M.; DRUCK, S.; CARVALHO, M. S. Análise Espacial e Geoprocessamento. In: DRUCK, S.; CARVALHO, M. S.; CÂMARA, G.; MONTEIRO, A. M. V. Análise Espacial de Dados Geográficos. Planaltina, Brasília, DF: Ed. Embrapa Cerrado, 2004.
- CARNEIRO, L. G. P. L. Desenvolvimento de uma Metodologia para a Previsão de Demanda de Passageiros para o Transporte Rodoviário Interestadual por ônibus. Brasília: UnB/DF, Publicação T.DM-002A/2005, Dissertação de Mestrado, Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, 2005.
- CASTRO, A. B.; LESSA, C. F. Introdução à Economia. Rio de Janeiro: Ed. Forense, 1967.
- CASTRO JUNIOR, N. R. Transporte Rodoviário de Passageiros: Estrutura, Desempenho e Desafios Regulatórios. Núcleo de Estudos e Modelos Espaciais Sistêmicos/NEMESIS. Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro/FAPERJ, Rio de Janeiro, 2003. <<http://www.nemesis.org.br>>.

- CEARÁ. Decreto nº 26.103, de 12 de janeiro de 2001. Aprova o regulamento dos serviços de transporte rodoviário intermunicipal de passageiros do Estado do Ceará e dá outras providências, 2001.
- CHOU, Y. Exploring Spatial Analysis in Geographic Information Systems. Santa Fé, New Mexico, USA: On Word Press, 1997.
- CORREIA, D. E. R. Metodologia para Identificação da Qualidade da Informação: Uma Aplicação para o Planejamento de Transportes. Brasília: UnB/DF, Publicação T.DM-005/2004, Dissertação de Mestrado, Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, 2004.
- DANTAS FILHO, N. M. Rotação de Tripulação em Transporte rodoviário por ônibus. Dissertação de Mestrado. Universidade Federal do Rio de Janeiro. COPPE, Rio de Janeiro, 1995.
- EBTU, Empresa Brasileira dos Transportes Urbanos. STPP Gerencia do Sistema de Transporte Público de Passageiros - Módulos de Treinamento. Brasília, Distrito Federal, 1988.
- FERNANDES, L. A. Engenharia Simultânea e Qualidade no Transporte Rodoviário de Passageiros: Um Estudo de Caso. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis, 2003.
- FERREIRA, E. A. Um Método de Utilização de Dados de Pesquisa Embarque/Desembarque na Calibração de Modelos de Distribuição do Tipo Gravitacional. Dissertação de Mestrado, Escola de Engenharia de São Carlos da Universidade de São Paulo. São Carlos, 1999.
- FOODY, G. M. Geographical weighting as a further refinement to regression modeling: An example focused on the NDVI – rainfall relationship. Remote Sensing of Environment 88, 283-293. Netherlands: Elsevier Science, 2003.
- FOTHERINGHAM, A.S; BRUNSDON, C.A.; CHARLTON, M.E. Geographically Weighted Regression – The analysis of spatially varying relationships. Chichester, England: Wiley, 2002.
- GONÇALVES, M. B.; BEZ, E. T.; MEDEIROS, H. C.; LUZ, G.; PHILIPPI, R. C. N. Um estudo sobre as preferências dos usuários do transporte rodoviário interestadual de passageiros. Panorama Nacional da Pesquisa em Transportes, Congresso da ANPET, Brasília, 2006.
- GONZALES TACO, P. W. Modelo de Geração de Viagens com Aplicação dos Sistemas de Informação Geográfica e Sensoriamento Remoto. Brasília: UnB/DF, Publicação TU.DM-02A/97, Dissertação de Mestrado, Departamento de Engenharia Civil, Faculdade de Tecnologia, Universidade de Brasília, 1997.
- GRIFFITH, D. A. Theory of Spatial Statistics. In: GAILE, G. L.; WILLMOTT, C. J. Spatial Statistics and Models. USA: D. Reidel Publishing Company. Springer, 1984.
- GUJARATI, D. N. Econometria Básica. Traduzido por Maria José Cyhlar Monteiro, 4th, ed. Rio de Janeiro: Editora Elsevier, 2006.
- HALL, R. E.; LIEBERMAN, M. Microeconomics: Principles and Applications. 3rd Edition. Thomson South-Western, Mason, OH. USA, 2005.
- HENRIQUE, C. S. Diagnóstico Espacial da Mobilidade e da Acessibilidade dos Usuários do Sistema Integrado de Transporte de Fortaleza. Dissertação de Mestrado, Programa de Mestrado em Engenharia de Transportes, Universidade Federal. Fortaleza, 2004.

- HUANG, Y.; LEUNG Y. Analysing Regional Industrialisation in Jiangsu Province using Geographically Weighted Regression. *Journal of Geographical Systems* 4: 233-249. Berlin/Heidelberg, Germany: Springer, 2002.
- KANAFANI, A. K. *Transportation Demand Analysis*. University of California, Berkeley, USA: McGraw-Hill Book Company, 1983.
- KREMPI, A. P. *Explorando Recursos de Estatística Espacial para Análise da Acessibilidade da Cidade de Bauru*. São Carlos: UFSCar/SP, Dissertação de Mestrado, Universidade Federal de São Carlos, 2004.
- LEE, Y. W.; STRONG, D. M.; KAHN, B. K.; WANG, R. Y. *AIMQ: A Methodology for Information Quality Assessment*. Information & Management. Netherlands: Noth Holland, Elsevier Science, 2001.
- LESAGE, J. P. *Spatial Econometrics*. Department of Economics, University of Toledo, OH. USA, 1999a.
- LESAGE, J. P. *Applied Econometrics using MATLAB*. Department of Economics, University of Toledo, OH. USA, 1999b.
- LOPES, S. B. *Efeitos da Dependência Espacial em Modelos de Previsão de Demanda por Transporte*. São Carlos: UFSCar/SP, Dissertação de Mestrado, Universidade Federal de São Carlos, 2005.
- LOUREIRO, C. F. G.; SILVA, H. N.; CARVALHO, L. E. X. *Metodologia de Análise de Regressão Geograficamente Ponderada Aplicada ao Fenômeno das Viagens Intermunicipais*. Panorama Nacional da Pesquisa em Transportes, Congresso da ANPET, Brasília, 2006.
- MAGALHÃES, M. T. Q. *Metodologia para Desenvolvimento de Sistemas de Indicadores: Uma Aplicação no Planejamento e Gestão da Política Nacional de Transportes*. Brasília: UnB/DF, Publicação T.DM-015A/2004, Dissertação de Mestrado, Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, 2004.
- MARTINS, F. G. D.; SILVA, F. G. F.; ROCHA, C. H. *Uma Análise de Fatores Determinantes das Condições de Concorrência nos Mercados de Transporte Rodoviário de Passageiros*. Panorama Nacional da Pesquisa em Transportes, Congresso da ANPET, Brasília, 2006.
- MARTINS, F. G. D. *Mercados de Transporte Rodoviário Interestadual de Passageiros: Uma Análise sob Enfoque Dinâmico de Elementos Determinantes das Condições de Concorrência*. Brasília: UnB/DF, Publicação T.DM-001A/2007, Dissertação de Mestrado, Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, 2007.
- MARSHMENT, R. *Transportation Planning Challenges and Opportunities*. Committee on Transportation Planning Applications, University of Oklahoma. USA, 2000. Disponível em: <<http://pubsindex.trb.org/document/view/default.asp?lbid=639419>>. Acesso em: 10/09/2007.
- MOLINERO, A. M.; ARELLANO, I. S. *Transporte Público: Planeación, Diseño, Operación y Administración*. 3ª. Ed., Colônia Nápoles, México: Fundación ICA, A. C., 1998.

- NOVAES, A. G. Modelos em Planejamento Urbano, Regional e de Transportes. São Paulo: Ed. Edgard Blücher Ltda, 1982.
- NOVAES, A. G. Sistemas de Transportes Volume1: Análise da Demanda. São Paulo: Ed. Edgard Blücher Ltda, 1986.
- ORTÚZAR, J. de D. Modelos de Demanda de Transporte. 2ª. Ed., Pontificia Universidad Catolica do Chile, México: Alfaomega Grupo Editor, 2000.
- PEÑA, L. L. S. Métodos para el Análisis Espacial. Una Aplicación al Estudio de la Geografía de la Pobreza. Congreso de la Asociación Latinoamericana de Población – ALAP, Guadalajara, México, 2006.
- PRADO, M. V. Data Warehouse para Apoio a Gestão da Operação em Empresas do Transporte Rodoviário Interestadual de Passageiros. Brasília: UnB/DF, Publicação T.DM-009A/2006, Dissertação de Mestrado, Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, 2006.
- QUEIROZ, M.P. Análise Espacial dos Acidentes de Trânsito do Município de Fortaleza. Fortaleza: Dissertação de Mestrado, Programa de Mestrado em Engenharia de Transportes, Universidade Federal do Ceará, 2003.
- RAMOS, R. A. R.; SILVA, A. N. R.; MIRANDA, V. P. Comparação de Metodologias de Análise Espacial para Definição de Zonas Urbanas Homogêneas. VIII Encontro de Utilizadores de Informação Geográfica – USIG. Tagus Park, Lisboa, Portugal, 2004.
- REVISTA ABRATI. A fantástica viagem do ônibus. nº 23, dez/2000.
- REVISTA ABRATI. Transporte democrático e desconcentrado. nº 41, jun/2005.
- RIPLEY, B. D. Spatial Statistics. USA: John Wiley & Sons Inc, 1981.
- SILVA, E. J. Sistema de Informação para Empresas Operadoras do Transporte Rodoviário de Passageiros com Utilização de Tecnologia de *Data Warehouse* e Ferramenta *OLAP*. Brasília: UnB/DF, Publicação TU.DM-008A/03, Dissertação de Mestrado, Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, 2003.
- SILVA, A. R. da. Avaliação de Modelos de Regressão Espacial para a Análise de Cenários do Transporte Rodoviário de Cargas. Brasília: UnB/DF, Publicação T.DM-013A/2006, Dissertação de Mestrado, Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, 2006.
- SILVA, A. R. da; MENDONÇA, A. C. de; ARAÚJO, C. E. F.; TACO, P. W. G.; YAMASHITA, Y. Nova Abordagem de Modelo de Geração de Viagens para o Transporte Público Urbano: Regressão Espacial. Congresso Brasileiro de Transporte e Trânsito – ANTP, Maceió, 2007.
- TEIXEIRA, G. L. Uso de Dados Censitários para Identificação de Zonas Homogêneas para Planejamento de Transportes Utilizando Estatística Espacial. Brasília: UnB/DF, Publicação T.DM-010A/03, Dissertação de Mestrado, Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, 2003.
- TCU, Tribunal de Contas da União. Transporte Rodoviário Interestadual e Internacional de Passageiros no Brasil. Brasília: Secretaria de Fiscalização de Desestatização, Apresentação Adylson Motta, 2005. Disponível em: <<http://www2.tcu.gov.br/pls/portal/docs/PAGE/TCU/>>

PUBLICACOES/LISTAPUBLICACOES/AUDITORIA_TRANSPORTE_RODOVIARIO_PO
RTUGUES.PDF>. Acesso em: 27/08/2007.

TOBLER, W. R. Cellular Geography, Philosophy in Geography. Edited by S. Gale and G. Olsson. Eds, Amsterdam, 1979.

VASCONCELLOS, E. A. Transporte Urbano, Espaço e Equidade – Análise das políticas públicas. 3ª. Ed., São Paulo: Annablume Editora, 2001.

WANG, Q.; NI, J.; TENHUNEN, J. Application of a geographically-weighted regression analysis to estimate net primary production of Chinese forest ecosystems. *Global Ecology & Biogeography* 14: 379-393. Boston, USA: Blackwell Publishing, 2005.

WRIGHT, C. L. Transporte Rodoviário de Ônibus. Brasília: IPEA, 1992.