

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

**ANÁLISE DA RELAÇÃO ENTRE O AMBIENTE
CONSTRUÍDO E A ESCOLHA PELO MODO A PÉ POR MEIO
DE MODELO ESPACIAL DE ESCOLHA DISCRETA**

CARLOS EDUARDO TEIXEIRA NEVES

ORIENTADOR: ALAN RICARDO DA SILVA

DISSERTAÇÃO DE MESTRADO EM TRANSPORTES

**PUBLICAÇÃO: T.DM-005/2019
BRASÍLIA/DF: FEVEREIRO/2019**

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

ANÁLISE DA RELAÇÃO ENTRE O AMBIENTE CONSTRUÍDO E A
ESCOLHA PELO MODO A PÉ POR MEIO DE MODELO ESPACIAL
DE ESCOLHA DISCRETA

CARLOS EDUARDO TEIXEIRA NEVES

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 EM TRANSPORTES.

APROVADA POR:



PROF. ALAN RICARDO DA SILVA, Dr. (UnB)
(ORIENTADOR)



PROF. PASTOR WILLY GONZALES TACO, Dr. (UnB)
(EXAMINADOR INTERNO)



PROFª. ANA MARGARITA LARRAÑAGA URIARTE, Drª. (UFRGS)
(EXAMINADOR EXTERNO)

BRASÍLIA/DF, 28 DE FEVEREIRO DE 2019.

FICHA CATALOGRÁFICA

NEVES, CARLOS EDUARDO TEIXEIRA

Análise da relação entre o ambiente construído e a escolha pelo modo a pé por meio de modelo espacial de escolha discreta [Distrito Federal] 2019.

xiii, 83p., 210 x 297 mm (ENC/FT/UnB, Mestre, Transportes, 2019).

Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia.

Departamento de Engenharia Civil e Ambiental.

1. Modo a pé

2. Ambiente Construído

3. Escolha modal

4. Modelos espaciais

I. ENC/FT/UnB

II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

NEVES, C. E. T. (2019). Análise da relação entre o ambiente construído e a escolha pelo modo a pé por meio de modelo espacial de escolha discreta, Publicação T.DM-005/2019, Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 83p.

CESSÃO DE DIREITOS

AUTOR: Carlos Eduardo Teixeira Neves.

TÍTULO DA DISSERTAÇÃO: Análise da relação entre o ambiente construído e a escolha pelo modo a pé por meio de modelo espacial de escolha discreta.

GRAU: Mestre

ANO: 2019

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 dessa dissertação de mestrado pode ser reproduzida sem autorização por escrito do autor.

Carlos Eduardo Teixeira Neves

carlos_etn@outlook.com

Dedicatória

Aos meus pais, Francisco e Ângela, por me darem todo o suporte para chegar até aqui e à Paula, pelo incansável companheirismo nesses dois anos de esforço.

Agradecimentos

Acredito que escrever agradecimentos é sempre uma tarefa ingrata devido a grande quantidade de pessoas e instituições que me acompanharam e inspiraram durante essa jornada. Entretanto, prometo fazer o meu melhor.

À Prof^ª. Dr^ª. Fabiana, que desde o início me apoiou e orientou nos altos e baixos ao longo do trabalho e compartilha do meu entusiasmo pelo planejamento urbano e pelos modos ativos.

Ao Prof. Dr. Alan, que me aceitou como seu orientando, me iniciou nessa ciência incrível que é a Estatística e cujo conhecimento foi essencial para o sucesso dessa pesquisa.

Ao Prof. Dr. Pastor, que me introduziu aos Transportes no curso de Engenharia Civil e colaborou imensamente para minha formação ao estar presente em todas as bancas que me avaliaram desde a graduação.

Aos meus antigos colegas de EGL Engenharia, por sempre incentivarem o estudo e a capacitação e me proporcionarem grande aprendizado em Transportes.

Aos meus colegas de PPGT, em especial Maiara, Elayne, Helen, Lorena e Guadiana, pelo companheirismo ao longo de todas as disciplinas, trabalhos e eventos.

À minha família e aos meus amigos de infância e de graduação, que sempre me apoiaram.

À CAPES, pela bolsa de Mestrado em parte desses dois anos.

Finalmente, à Universidade de Brasília, minha eterna *alma mater*, por me proporcionar crescimento intelectual, amizades eternas e memórias incríveis.

RESUMO

ANÁLISE DA RELAÇÃO ENTRE O AMBIENTE CONSTRUÍDO E A ESCOLHA PELO MODO A PÉ POR MEIO DE MODELO ESPACIAL DE ESCOLHA DISCRETA

A pesquisa sobre como as características urbanas, comumente categorizadas nos 5 D's, influenciam a escolha modal nos deslocamentos cotidianos é recorrente, principalmente em países desenvolvidos. O objetivo principal é que o entendimento dessas relações consiga subsidiar a criação de políticas urbanas que promovam o desenvolvimento sustentável. Nesse sentido, a caminhada representa um dos meios de locomoção a serem incentivados, pois é o modo com menor gasto energético e exigência de espaço de circulação. Este estudo buscou então analisar a relação entre a escolha por esse modo e as características do ambiente construído por meio de modelo de regressão espacial. Portanto, foi estabelecido como objetivo secundário a avaliação se a utilização de modelo espacial melhora a representatividade dos resultados quando comparado a um modelo de regressão tradicional. A pesquisa utilizou como base de dados de viagens a pesquisa Origem-Destino de 2007 do município de São Paulo e as características do ambiente construído foram obtidas em ambiente GIS. Após o tratamento e filtragem da amostra, foram consideradas 7.329 viagens. De forma preliminar, foram aplicados testes para diagnosticar as diferentes relações espaciais existentes entre as observações das variáveis. Estimaram-se então quatro modelos Probit, sendo um não espacial e três que consideram os efeitos espaciais de formas distintas. Os resultados indicaram que a inclusão de dependência espacial na escolha (variável dependente) não adiciona ajuste em comparação ao modelo não espacial, mas ocorre um pequeno aumento de ajuste quando são considerados os efeitos espaciais das variáveis independentes. Ademais, os resultados indicam que, para o município de São Paulo, as medidas do ambiente construído mais relacionadas à escolha pelo modo a pé são a entropia e a distância para o centro, variáveis representantes da diversidade de usos e da disponibilidade de destinos, respectivamente.

ABSTRACT

ANALYSIS OF THE RELATIONSHIP BETWEEN THE BUILT ENVIRONMENT AND THE CHOICE FOR WALKING USING A SPATIAL DISCRETE CHOICE MODEL

The research on how the urban characteristics, commonly categorized in the 5 D's, influences the modal choice in daily trips is recurrent, mainly in developed countries. The main objective of this work is that the understanding of these relations subsidize the creation of urban policies that promote sustainable urban development. In this sense, walking represents one of the ways of transportation to be encouraged, since it is the mode with lower energy expenditure and requirement of circulation space. This study aimed the analysis of the relationship between the choice for this mode and the characteristics of the built environment using a spatial regression model, type of regression analysis that includes spatial effects in its design. Therefore, it was established as a secondary objective the evaluation if the use of a spatial perspective improves the model goodness of fit when compared to a traditional regression model. The research used the travel database of the 2007 Origin – Destination survey from São Paulo and the characteristics of the built environment were obtained in a GIS system. After treatment and filtering of the sample, 7,329 trips were considered. Firstly, tests were applied to diagnose the different spatial relationships between the observations of the variables. Four Probit models were applied, one non-spatial and three considering spatial effects in different ways. The results indicated that the inclusion of the spatial dependence in the choice (dependent variable) does not increase the goodness of fit of the model in comparison to the non-spatial model, but the goodness of fit increases when the spatial effects of the independent variables are considered. In addition, the results indicate that, for São Paulo, the measurements of the built environment most related to the choice for walking are the entropy and the distance to downtown, variables representing the diversity of land use and the destination accessibility, respectively.

ÍNDICE

1. INTRODUÇÃO	1
1.1. PROBLEMA	3
1.2. PERGUNTA DE PESQUISA	4
1.3. OBJETIVOS.....	4
1.3.1. Objetivo Geral	4
1.3.2. Objetivos Específicos	4
1.4. JUSTIFICATIVA.....	4
1.5. ESTRUTURA DO TRABALHO.....	5
2. ESCOLHA PELO MODO A PÉ E O AMBIENTE CONSTRUÍDO	6
2.1. OS CINCO D'S	8
2.2. ASPECTOS DA LITERATURA	10
2.3. VARIÁVEIS DE CONTROLE.....	16
2.4. DENSIDADE	19
2.5. DIVERSIDADE	21
2.6. DESENHO	22
2.7. DISPONIBILIDADE DE DESTINOS.....	24
2.8. DISTÂNCIA PARA O TRANSPORTE PÚBLICO	25
2.9. CONSIDERAÇÕES FINAIS	26
3. MODELOS DE ESCOLHA DISCRETA E ESPACIAIS.....	27
3.1. MATRIZ DE PROXIMIDADE ESPACIAL	29
3.2. ÍNDICES DE AUTOCORRELAÇÃO ESPACIAL	31
3.2.1. <i>I</i> de Moran	31
3.2.2. <i>C</i> de Geary.....	32
3.2.3. Índices Locais.....	33
3.3. MODELOS DE REGRESSÃO ESPACIAL.....	34
3.3.1. Modelo Espacial Autoregressivo Misto	35
3.3.2. Modelo com Erro Espacial Autoregressivo	36
3.3.3. Modelo Regressivo Cruzado Espacial.....	37
3.3.4. Modelo Espacial Durbin.....	38
3.3.5. Modelo Probit Espacial	39
3.4. AJUSTE DE MODELOS DE REGRESSÃO ESPACIAL	41
3.5. CONSIDERAÇÕES FINAIS	42
4. MÉTODO PARA DESENVOLVIMENTO DO TRABALHO	43
4.1. ESCOLHA DO MODELO.....	43
4.2. SELEÇÃO E TRATAMENTO DA BASE DE DADOS DE VIAGENS E DA VARIÁVEL DEPENDENTE	44

4.3. ESPECIFICAÇÃO DA MATRIZ DE PROXIMIDADE ESPACIAL	48
4.4. ESCOLHA E ESPECIFICAÇÃO DAS VARIÁVEIS INDEPENDENTES	49
4.5. VERIFICAÇÃO DA EXISTÊNCIA DE DEPENDÊNCIA ESPACIAL NAS VARIÁVEIS EXPLICATIVAS.....	56
4.6. ESTIMATIVA DO MODELO ESPACIAL	56
4.7. ESTIMATIVA DO MODELO NÃO ESPACIAL.....	57
4.8. COMPARAÇÃO DO AJUSTE DOS MODELOS	57
4.9. AVALIAÇÃO DOS RESULTADOS	57
4.10. CONSIDERAÇÕES FINAIS	58
5. ANÁLISE DOS RESULTADOS.....	59
5.1. MATRIZ DE PROXIMIDADE	59
5.2. ESCOLHA DE DISTÂNCIA DE BUFFER E ESTATÍSTICA DESCRITIVA.....	60
5.3. DEPENDÊNCIA ESPACIAL NAS VARIÁVEIS INDEPENDENTES.....	62
5.4. PARÂMETROS E AJUSTES DOS MODELOS.....	64
5.5. EFEITOS MARGINAIS	66
5.6. CONSIDERAÇÕES FINAIS	68
6. CONCLUSÕES.....	71
6.1. LIMITAÇÕES DO TRABALHO	71
6.2. SUGESTÕES PARA TRABALHOS FUTUROS	72
REFERÊNCIAS BIBLIOGRÁFICAS	74
APÊNDICE A – RESULTADOS DOS MODELOS.....	80

LISTA DE TABELAS

Tabela 4.1 – Perfil de escolaridade.	47
Tabela 4.2 – Perfil socioeconômico (Critério de Classificação Econômica Brasil).	47
Tabela 4.3 – Perfil etário.	48
Tabela 5.1 – Teste de distâncias para a Matriz de Proximidade.	59
Tabela 5.2 – Teste de distâncias para o <i>buffer</i>	60
Tabela 5.3 – Estatística descritiva das variáveis (<i>buffer</i> de 400 metros e $n = 7.329$).	61
Tabela 5.4 – Teste de dependência espacial nas variáveis independentes.	63
Tabela 5.5 – Coeficientes (β) e ajuste dos modelos Probit.	65
Tabela 5.6 – Efeitos marginais médios.	66
Tabela 5.7 – Relação entre distância para o centro do destino e distância de viagem.	67

LISTA DE FIGURAS

Figura 2.1 – Relação entre os 5 D's.	10
Figura 2.2 – Cidades estudadas na literatura.	14
Figura 3.1 – Exemplo de Matriz de Proximidade Espacial.	30
Figura 3.2 – Comparação de modelos espaciais.	35
Figura 3.3 – <i>Spatial Auto Regressive</i> (SAR).	36
Figura 3.4 – <i>Spatial Error Model</i> (SEM).	37
Figura 3.5 – <i>Spatial lag of X</i> (SLX).	38
Figura 3.6 – <i>Spatial Durbin Model</i> (SDM).	39
Figura 4.1 – Método aplicado.	43
Figura 4.2 – Comparação entre diferentes tipos de <i>buffer</i> (500 metros).	51
Figura 5.1 – Índice <i>I</i> de Moran por distância de corte em metros.	60
Figura A.1 – Resultados do modelo Probit não espacial.	80
Figura A.2 – Resultados do modelo Probit SAR.	81
Figura A.3 – Resultados do modelo Probit SLX (primeira parte).	81
Figura A.4 – Resultados do modelo Probit SLX (segunda parte).	82
Figura A.5 – Resultados do modelo Probit SDM.	82
Figura A.6 – Efeitos marginais médios do modelo Probit não espacial.	83
Figura A.7 – Efeitos marginais médios do modelo Probit SLX.	83

LISTA DE QUADROS

Quadro 2.1 – Características do referencial.	12
Quadro 2.2 – Características dos dados utilizados nas referências.	13
Quadro 2.3 – Tipos de variáveis de controle.	17
Quadro 2.4 – Conclusões com relação às variáveis de controle.	18
Quadro 2.5 – Resumo das variáveis relativas à densidade.	20
Quadro 2.6 – Resumo das variáveis relativas à diversidade.	22
Quadro 2.7 – Resumo das variáveis relativas ao desenho.	23
Quadro 2.8 – Resumo das variáveis relativas à disponibilidade de destinos.	24
Quadro 2.9 – Resumo das variáveis relativas à distância para o transporte público.	25
Quadro 4.1 – Variáveis independentes selecionadas.	54
Quadro 4.2 – Relação entre as etapas do método e os resultados.	58

LISTA DE SIGLAS

AIC – *Akaike Information Criterion*

BIC – *Bayesian Information Criterion*

CBD – *Central Business District*

CEM/Cepid – Centro de Estudos da Metrópole

CET – Companhia de Engenharia de Tráfego

DECAR – Departamento de Arrecadação e Cobrança

GIS – *Geographic Information System*

IBGE – Instituto Brasileiro de Geografia e Estatística

IPTU – Imposto Predial e Territorial Urbano

LIK – Log-verossimilhança

LISA – *Local Indicators of Spatial Association*

MAUP – *Modifiable Areal Unit Problem*

Metrô-SP – Companhia do Metropolitano de São Paulo

SAR – *Spatial Auto Regressive*

SDM – *Spatial Durbin Model*

SEM – *Spatial Error Model*

SLX – *Spatial Lag of X*

TP – Transporte Público

TPCL – Cadastro Territorial e Predial, de Conservação e Limpeza

1. INTRODUÇÃO

O incentivo à mobilidade ativa (não-motorizada) é tido como iniciativa fundamental para lidar com os atuais problemas de mobilidade nos grandes centros urbanos. O transporte individual motorizado, antes amplamente encarado como prioritário, possui problemas críticos quando utilizado em excesso e se apresenta contrário aos conceitos de sustentabilidade urbana e ambiental. Congestionamentos e poluição acarretam custos sociais ou econômicos elevados e estão presentes no modelo de desenvolvimento urbano das grandes cidades, principalmente em países em desenvolvimento como o Brasil.

Além da sustentabilidade, a mobilidade ativa também é importante instrumento na busca por melhores indicadores na saúde pública. Segundo dados do IBGE (2015), 62,1% dos brasileiros acima de 15 anos não praticam quaisquer atividades físicas. Vários estudos (WANG *et al.*, 2004; WEN & RISSEL, 2008; CERVERO *et al.*, 2009; EVENSON *et al.*, 2012; SIQUEIRA REIS *et al.*, 2013; FEUILLET *et al.*, 2016; LU *et al.*, 2017) relacionam a utilização dos modos não motorizados à realização de atividade física, ou seja, a mudança de hábitos diários básicos, como a decisão de qual modo utilizar para percorrer pequenas distâncias no cotidiano, pode acarretar em melhores condições de saúde para a população. Dentre os dois modos ativos mais comuns, a bicicleta e a caminhada, o segundo merece destaque por ser a forma mais acessível e barata de locomoção. Além disso, a caminhada é o modo de menor gasto energético, o menos poluente e o que ocupa menos espaço urbano.

Um dos focos existentes nos estudos sobre os pedestres é quais fatores incentivam as pessoas a optarem por esse modo em detrimento de outros. Essa escolha envolve diversos aspectos, como as características socioeconômicas (HEARST *et al.*, 2013; MARQUET & MIRALLES-GUASCH, 2014), os padrões comportamentais (BIRD *et al.*, 2013) e as características urbanísticas (EWING & CERVERO, 2010; KOH & WONG, 2013; SUN *et al.*, 2017; LINDELÖW *et al.*, 2017; GUERRA *et al.*, 2018). Logo, a busca pela troca modal se mostra bastante complexa, pois deve agir em várias frentes distintas, que muitas vezes já possuem valores consolidados no cotidiano individual. O estudo dessas influências já foi analisado por diversos autores, mas não existe consenso em relação aos impactos existentes, visto a complexidade do comportamento humano e a diversidade de contextos sociais e urbanos.

Dentre os fatores citados, merece destaque as características urbanísticas, que compõem o ambiente no qual o deslocamento é realizado. Dentre elas é possível citar o uso do solo, a infraestrutura para pedestres, as características das fachadas, a forma urbana, entre outras. Segundo Cervero & Kockelman (1997), as características do ambiente construído tem influência direta na forma de deslocamento das pessoas. Ademais, Litman (2018) afirma que, embora as características de uso do solo tenham impactos individuais pequenos, geralmente afetando uma pequena proporção do número de viagens total, eles são cumulativos e sinérgicos, ou seja, quando considerados em conjunto possuem relevância significativa. Como essas características devem ser combinadas para promover a mobilidade do pedestre ainda é tema de discussões. O conhecimento amplo desse modo de transporte requer, além da constante coleta de dados, a utilização de modelos de análise capazes de avaliar, de forma confiável, o seu comportamento, principalmente por meio da previsão da demanda e da análise da oferta.

Nesse contexto, os modelos de escolha discreta são usados para estudar essas relações, tanto para pedestres quanto para outros modos (HANDY, 1996). Essa análise é utilizada para examinar a relação entre uma escolha entre duas ou mais opções e uma ou mais variáveis independentes. Depois de realizada a análise, as estatísticas de regressão podem ser usadas para prever a probabilidade de escolha por determinada opção quando as variáveis independentes são conhecidas. Logo, a análise vai além da correlação, adicionando recursos de previsão, dependendo da qualidade da modelagem realizada. Ademais, como análise adicional, é possível computar os efeitos marginais da variação das variáveis na probabilidade de escolha. Dessa forma, são obtidas conclusões quanto à importância relativa de determinadas características no processo decisório estudado.

É notável que as características do ambiente construído possuem caráter essencialmente espacial e geográfico, o que pode prejudicar a eficiência dos modelos estatísticos tradicionais pela existência de correlação entre as variáveis (ANSELIN *et al.*, 2004). Esse efeito de dependência espacial é definido por Almeida (2012) como “um caso especial de *cross sectional dependence*, que aflora quando as unidades de corte transversal – indivíduos, domicílios, empresa ou regiões – não são mais independentes entre si”. Segundo LeSage & Pace (2009), a dependência espacial reflete uma situação na qual uma observação em determinado ponto é influenciada pelas observações próximas. Essa dependência, tema de estudo da Econometria Espacial, pode ocorrer em relação à variável dependente, às variáveis independentes ou aos aspectos não considerados no modelo (erro), portanto existem uma série de diferentes

especificações de modelos de regressão espacial que buscam lidar com esse fenômeno (GOLGHER, 2015).

Logo, além da discussão sobre o efeito das características do ambiente construído na escolha pelo modo a pé, é importante discutir quais modelos podem ser utilizados e se são adequados para esse tipo de estudo, de forma a melhorar as conclusões e previsões obtidas a partir deles. Este trabalho busca então investigar, no município de São Paulo, a relação entre o ambiente construído na escolha pelo modo a pé em conjunto com a aplicabilidade de modelo de regressão espacial nesse tipo de pesquisa.

1.1. PROBLEMA

A modelagem de como os aspectos urbanísticos influenciam na escolha pela caminhada no cotidiano urbano tem como objetivo analisar os pesos e a relevância do ambiente construído nessa decisão. A meta final desse tipo de análise é subsidiar o desenvolvimento de diretrizes para a promoção da mobilidade sustentável no contexto urbano, de forma a promover melhor qualidade de vida. Porém, grande parte da pesquisa já realizada sobre o assunto foi realizada em países ocidentais desenvolvidos e pouco foi feito em países em desenvolvimento (CERVERO *et al.*, 2009; REIS *et al.*, 2013; LARRAÑAGA *et al.*, 2015; GUERRA *et al.*, 2018).

Ademais, a validade dos modelos estatísticos de escolha discreta, comumente utilizados em pesquisas na área, é intrinsecamente conectada às suposições adotadas em sua concepção, seja com relação às condições das variáveis independentes e dependentes, seja com relação aos procedimentos de amostragem. Em grande parte da pesquisa na área, as relações espaciais entre os dados não foram consideradas (LEE *et al.*, 2017) e, como mostrado mais adiante nesse trabalho, a existência desse tipo de relação pode comprometer a premissa básica de independência entre as variáveis presente nas técnicas mais comuns de análise. Logo, surge a questão se, caso a visão espacial seja incorporada à análise do problema da relação entre o ambiente construído e a escolha pelo modo a pé, as relações já encontradas na literatura se manteriam ou se os resultados seriam divergentes.

1.2. PERGUNTA DE PESQUISA

Frente ao problema exposto, surge a principal pergunta desta pesquisa: como as características do ambiente construído influenciam na escolha pelo modo a pé quando são consideradas as relações espaciais entre as suas características e entre os indivíduos no processo de escolha? Em segundo plano, destaca-se o questionamento se a aplicação dos efeitos espaciais melhora ou não o ajuste de modelos de regressão comumente utilizados nesse tipo de estudo.

1.3. OBJETIVOS

1.3.1. Objetivo Geral

Este trabalho tem como objetivo principal analisar a relação existente entre o ambiente construído e a escolha pelo modo a pé por meio de modelo espacial de escolha discreta.

1.3.2. Objetivos Específicos

Os objetivos específicos são:

- Analisar os fatores de influência (variáveis independentes) do ambiente construído para a escolha do modo a pé;
- Avaliar a existência de relação espacial entre as variáveis;
- Analisar a aderência dos modelos espaciais e não espaciais.

1.4. JUSTIFICATIVA

Entender como o espaço urbano influencia a escolha pelo modo a pé é um caminho para a definição de diretrizes de planejamento que busquem incentivar a caminhada e reduzir o uso dos modos motorizados, de forma a tratar os atuais problemas de mobilidade urbana. Guerra *et al.* (2018) destacam que mesmo após várias pesquisas na área, ainda restam dúvidas sobre o impacto quantificável do ambiente construído sobre o padrão de viagens, principalmente na América Latina. Logo, além de o problema posto ter importância significativa no atual contexto urbano, principalmente de países em desenvolvimento, ainda existe a necessidade de novas contribuições para melhor conhecimento do assunto.

Com relação à técnica escolhida, os modelos espaciais de regressão possuem algumas aplicações em estudos de transporte no geral (GOETZKE, 2008; BOUSSAUW *et al.*, 2012; WANG *et al.*, 2015; SOLTANI & ASKARI, 2017; CHEN *et al.*, 2017), entre elas pesquisas que tem foco nos pedestres (SANNI & ABRANTES, 2010; FEUILLET *et al.*, 2016; YANG *et al.*, 2017; LEE *et al.*, 2017). No Brasil, a Econometria Espacial aplicada à área de transportes foi encontrada em teses e dissertações (QUEIROZ, 2003; TEIXEIRA, 2003; KREMPI, 2004; LOPES, 2005; SILVA, 2006; LOPES, 2010; RIBEIRO, 2012; SOUZA FILHO, 2013). Os modelos de regressão espacial, devido à consideração das relações espaciais em sua concepção, se apresentam como potenciais ferramentas de análise em problemas e investigações no planejamento de transportes.

1.5. ESTRUTURA DO TRABALHO

Esta dissertação é composta por 6 Capítulos, além das Referências Bibliográficas, organizados na seguinte forma:

1. **Introdução:** contextualização sobre o tema da pesquisa e apresentação de seus objetivos e justificativas.
2. **Escolha pelo modo a pé e o ambiente construído:** levantamento na literatura dos fatores de influência do ambiente construído estudados, principais técnicas de análise adotadas nas pesquisas e resultados encontrados.
3. **Modelos espaciais:** exposição da teoria que embasa os modelos espaciais de regressão.
4. **Método para desenvolvimento do trabalho:** detalhamento do método com base no referencial teórico dos Capítulos 2 e 3.
5. **Resultados obtidos:** análise dos resultados frente aos objetivos determinados no Capítulo 1.
6. **Conclusão:** considerações finais e recomendações para estudos posteriores.

2. ESCOLHA PELO MODO A PÉ E O AMBIENTE CONSTRUÍDO

A utilização da caminhada como meio de transporte oferece uma ampla gama de benefícios para os indivíduos e para a sociedade. Do ponto de vista do transporte e do planejamento urbano, mais viagens a pé podem ser acompanhadas por menos viagens de veículos individuais motorizados e, portanto, menos tráfego, poluição do ar e outros impactos ambientais. Do ponto de vista da saúde pública, as viagens a pé significam maior atividade física e, assim, melhoram o bem-estar e reduzem os custos associados a sistemas públicos de saúde (LU *et al.*, 2017).

Segundo Tian & Ewing (2017), essas duas áreas de estudo distintas fornecem evidências da ligação entre o ambiente construído e a caminhada. A pesquisa de comportamento de viagens, baseada nas áreas de engenharia de transportes e planejamento urbano, focou na caminhada como um meio de transporte, ou seja, na viagem a pé com o objetivo de chegar até um destino. Já as pesquisas na área da saúde pública têm se concentrado na caminhada como uma forma de exercício físico, com ou sem consideração do motivo pelo qual ela está sendo realizada.

Apesar de um grande número de estudos contribuir para o debate sobre a relação entre o ambiente construído e a escolha do modo de transporte, a avaliação de como diferentes características do ambiente urbano influenciam na escolha pelo modo a pé ainda é objeto de pesquisas no meio científico devido à variedade de técnicas de pesquisa e contextos sociais e urbanos existentes. Nesse sentido, Tian & Ewing (2017) apontam que existem diversas revisões da literatura e metanálises sobre o assunto. Apenas no estudo de Ewing & Cervero (2010) foram inspecionados 200 estudos que relacionam de forma quantitativa o ambiente construído com medidas relacionadas a viagens.

Esse conjunto de pesquisas fornece evidências da existência de uma correlação entre o ambiente construído e a opção pelo modo a pé. Por exemplo, Cervero & Duncan (2003) concluíram que a diversidade do uso do solo está positivamente relacionada à escolha por esse modo e Vale & Pereira (2016) apontaram que a acessibilidade do local de trabalho tem uma associação positiva com a frequência com que o indivíduo caminha. Kamruzzaman *et al.* (2016) afirmaram que aspectos do ambiente construído afetam a caminhada como meio de transporte, mas que também deve ser dada a devida importância às atitudes e percepções dos indivíduos no processo de escolha do modo.

No geral, as pesquisas focam em testar se as diretrizes do movimento chamado *New Urbanism* (Novo Urbanismo) realmente influenciam na divisão modal. Segundo Mumford (2018), o Novo Urbanismo possui a premissa central de que comunidades desenhadas de forma compacta tendem a promover a redução do uso do automóvel individual e a aumentar o uso da caminhada e do transporte público. Larrañaga *et al.* (2015) citam ainda que o tema é de grande interesse de correntes como o *Smart Growth* (Crescimento Inteligente) e o *Transit Oriented Development* (Desenvolvimento Orientado ao Transporte Sustentável).

Nesse contexto, o objetivo principal deste Capítulo é levantar na literatura existente estudos que buscaram de alguma forma investigar a relação entre o ambiente construído e a caminhada. Essa revisão visou à busca pelo entendimento dessa relação e ao fornecimento de subsídios para a escolha das variáveis de análise desta pesquisa, além de aspectos como a escala e a forma de levantamento dessas variáveis. É importante destacar que não foram encontradas somente pesquisas com foco exclusivo nos pedestres, já que existem estudos que avaliam o tema deste trabalho em conjunto com a escolha por outros modos de transporte.

Primeiramente, é importante definir exatamente qual o significado do termo “ambiente construído” utilizado nesta pesquisa. Vale ressaltar que a intenção não é criar novas definições ou conceitos, mas apenas trazer como a expressão é utilizada na literatura e estabelecer em qual sentido ela foi aqui utilizada.

No contexto dos estudos levantados, o termo “ambiente construído”, ou “*built enviroment*”, é empregado de diversas formas, mas em nenhum artigo procurou-se estabelecer um conceito de forma explícita. Entretanto, é possível analisar o posicionamento dos autores quanto à questão pelo emprego do termo em suas pesquisas. Targa & Clifton (2004) utilizam os termos *land use* e *urban form* inseridos dentro do ambiente construído, ou seja, esse engloba os conceitos de uso do solo e forma urbana. Já Ewing & Cervero (2010) vão além e enquadram os 5 D’s (densidade, diversidade, desenho, destino e distância para o transporte público), que são tema da próxima seção, dentro da concepção de ambiente construído. A mesma diretriz é adotada por McKibbin (2011). Logo, nos exemplos aqui colocados, o ambiente construído é formado tanto pelo aspecto físico e construtivo do meio urbano, quanto pelas atividades que são realizadas em seu espaço.

Dessa forma, optou-se nesse trabalho por utilizar o termo “ambiente construído” na forma utilizada por Ewing & Cervero (2010). O ambiente construído abrange então as características do conjunto de intervenções físicas construídas pelo homem no contexto urbano que permitem o desenvolvimento de suas atividades, além de abranger a distribuição e configuração dessas. Já o termo “uso do solo” será adotado apenas em tópicos referentes às atividades desenvolvidas no meio urbano.

Este capítulo está dividido na seguinte forma: a primeira seção trata sobre a conceituação teórica aplicada na análise das variáveis e a segunda traz uma visão geral da literatura encontrada. Já da terceira à oitava seções são apresentadas as variáveis independentes utilizadas na literatura e as conclusões gerais dos autores em relação a elas. Por fim, a nona seção traz as considerações finais sobre a temática.

2.1. OS CINCO D’S

Ewing & Cervero (2010) destacam que as variáveis utilizadas para analisar a influência do ambiente construído na demanda de viagens são frequentemente categorizadas em palavras ou expressões iniciadas, na língua inglesa, com a letra “D”. Cervero & Kockelman (1997) começaram essa especificação pela criação dos três primeiros “D’s”: densidade (*density*), desenho (*design*) e diversidade (*diversity*). Essas três categorias abrangem as principais métricas encontradas na literatura. Posteriormente, acrescentou-se a disponibilidade de destinos (*destination accessibility*) e a distância para o transporte coletivo (*distance to transit*) às categorias de variáveis (EWING & CERVERO, 2010).

A densidade é medida como a variável de estudo pela unidade de área. A variável pode representar diversos aspectos, como população, empregos, habitações e área construída. Em determinados casos, duas medidas podem ser combinadas. Ewing & Cervero (2010) dão o exemplo da combinação de população e empregos para gerar uma “densidade de atividades”.

A diversidade diz respeito aos diferentes usos do solo em uma determinada área e a sua importância relativa em termos de área ocupada ou de emprego. A entropia, em que valores baixos indicam ambientes de uso único e valores altos indicam diversidade no uso do espaço, é amplamente utilizada em estudos de viagens.

Já o desenho inclui características da rede de vias. A malha viária varia entre densas redes urbanas com alta conectividade e vias diretas para redes suburbanas esparsas com baixa conectividade e muitos “becos sem saída”. As medidas incluem o tamanho médio dos quarteirões, a proporção de interseções em “cruz” (quatro aproximações), a quantidade de interseções por unidade de área e a densidade de vias. O desenho também é medido em função de variáveis físicas que diferenciam ambientes orientados para pedestres de ambientes orientados para veículos, como a quantidade de árvores, a largura média das calçadas e a presença de travessias para pedestres. Larrañaga *et al.* (2015) comentam que “o desenho urbano significa o desenho da cidade e os elementos físicos dentro dele, incluindo tanto sua organização como sua aparência.”

A disponibilidade de destinos mede o acesso a potenciais destinos de interesse, como regiões com alta concentração de empregos e de estabelecimentos comerciais. São exemplos de variáveis nessa categoria a distância ao *Central Business District (CBD)*, a quantidade de empregos acessíveis dentro de uma determinada quantidade de tempo e a distância da residência ao supermercado mais próximo.

Por fim, a distância para o transporte coletivo diz respeito à acessibilidade de estações de transporte ferroviário ou rodoviário, que pode ser representada pela distância de acesso, pela densidade de linhas na área de estudo ou pela densidade de pontos de acesso. É importante destacar que também podem ser utilizadas outras medidas que considerem a oferta e a qualidade do transporte coletivo na região.

Nota-se que essas variáveis não são completamente independentes, pois, por exemplo, áreas urbanas densas tendem a ser diversas, acessíveis e tem elevado nível de serviços de transporte coletivo (CERVERO *et al.*, 2009). Logo, segundo Ewing & Cervero (2010), os domínios dessas categorias se sobrepõem (Figura 2.1) e a classificação de determinadas medidas pode variar na literatura, o que afetaria os modelos de análise apenas se esse processo considerasse a categorização de forma explícita.

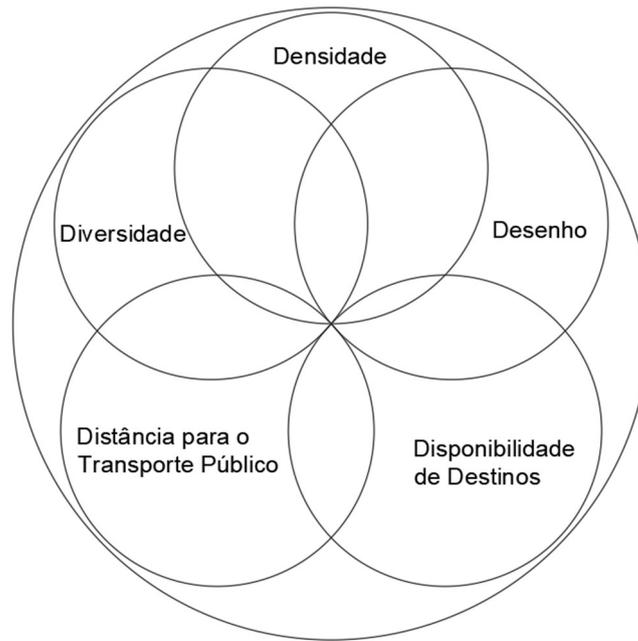


Figura 2.1 – Relação entre os 5 D's.

Fonte: Adaptado de Cervero *et al.* (2009).

Esse modelo classificatório é então utilizado nesse estudo, de forma a estruturar a revisão teórica e a organizar a análise dos resultados. Abordagem semelhante foi aplicada por diversos estudos encontrados na literatura (AMÂNCIO & SANCHES, 2008; MCKIBBIN, 2011; LARRAÑAGA *et al.*, 2015; SUN *et al.*, 2017; TIAN & EWING, 2017; LU *et al.*, 2017).

2.2. ASPECTOS DA LITERATURA

Siqueira Reis *et al.* (2013) ressaltam que o ambiente construído na América Latina historicamente é diferente do encontrado em países desenvolvidos e Guerra *et al.* (2018) destacam a falta de evidências empíricas sobre o assunto nos países latinos não europeus. Sun *et al.* (2017) comentam que, embora muitos estudos tenham mostrado significância estatística entre o ambiente construído e a caminhada em países ocidentais, pouca pesquisa foi realizada nas megacidades da Ásia. Portanto, nesse contexto, buscou-se também na literatura estudos que utilizaram como objeto cidades fora de países ocidentais desenvolvidos. O Quadro 2.1 e o Quadro 2.2 resumem as características dos estudos encontrados e a Figura 2.2 mostra a localização e os países de origem dos dados. Foram encontrados cinco estudos realizados na América Latina, sendo três em cidades brasileiras (AMÂNCIO & SANCHES, 2008; REIS *et al.*, 2013; LARRAÑAGA *et al.*, 2015), um na Colômbia e um no México.

Os estudos foram classificados quanto a área de pesquisa a que pertencem. Como colocado, existem dois campos de estudo com pesquisas relacionadas à conexão entre o ambiente construído e o modo a pé: o planejamento urbano ou de transportes e a saúde pública. Entretanto, os estudos focados na caminhada como um exercício físico também fornecem uma base de medidas do ambiente construído que pode ser utilizada em estudos sobre planejamento de transportes.

É notável que o campo de pesquisa ao qual o trabalho pertence afeta diretamente na variável de análise escolhida (variável dependente). Todos os trabalhos que avaliam a caminhada como uma atividade física utilizam variáveis associadas ou com o tempo de caminhada em determinado intervalo de tempo, ou com a frequência de viagens realizadas a pé. Já os trabalhos da área de transportes lidam principalmente com a escolha discreta entre diferentes modos ou a realização ou não de viagens a pé. Esse tipo de variável dependente é o mais adequado para a análise de processos de escolha (HANDY, 1996) e é o tipo escolhido para este trabalho. Mais sobre a especificação da variável dependente está presente no Capítulo 4. Ademais, nos estudos da área de transportes, também foi utilizada a proporção de viagens realizadas a pé como variável de análise (SANNI & ABRANTES, 2010; LAMÍQUIZ & LÓPEZ-DOMÍNGUEZ, 2015).

Quadro 2.1 – Características do referencial.

Referência	Área	Local	Variável analisada (dependente)	Unidade espacial	Principais técnicas de análise
Guerra <i>et al.</i> , 2018	Transporte/ Planej. Urbano	100 cidades (México)	Escolha discreta entre modos (Carro, TP e Modos Ativos)	Cidade	Regressão Logística Multinomial
Boulangue <i>et al.</i> , 2017	Transporte/ Planej. Urbano	Melbourne (Austrália)	Binária (realizou ou não uma viagem em determinado modo)	Buffer de rede (1.600m)	Modelo Multinível (logístico)
Lu <i>et al.</i> , 2017	Saúde Pública	Hong Kong (China)	Minutos de caminhada por semana	Buffer (400m)	Modelo Multinível (linear)
Ding <i>et al.</i> , 2017	Transporte/ Planej. Urbano	Baltimore (EUA)	Escolha discreta entre modos (Carro, TP e Modos Ativos)	Variada (<i>buffers</i> , quadras e zonas de tráfego)	Equações Estruturais Regressão Logística Multinomial
Sun <i>et al.</i> , 2017	Transporte/ Planej. Urbano	Xangai (China)	Escolha discreta entre modos (Carro, TP, Bicicleta e Caminhada) Distância de Viagem	Subdistrito	Cópula (Regressão Logística Multinomial e Regressão Linear)
Tian & Ewing, 2017	Transporte/ Planej. Urbano	Portland (EUA)	1 - Binária (se ocorreu viagem a pé ou não) 2 - Quantas viagens a pé ocorreram	Buffer de rede (400m, 800m e 1.600m)	1 - Regressão Logística 2 - Regressão Binomial Negativa
Feuillet <i>et al.</i> , 2016	Saúde Pública	Paris (França)	Tempo de caminhada (horas por semana)	Buffer (500m)	Regressão de Poisson Regressão de Poisson Geograficamente Ponderada
Koohsari <i>et al.</i> , 2016	Saúde Pública	Adelaide (Austrália)	Viagens a pé por semana	Setor censitário	Equações de Estimativas Generalizadas (GEE)
Lamíquiz & López-Domínguez, 2015	Transporte/ Planej. Urbano	Madri (Espanha)	Proporção de viagens a pé	Variada (Zonas de Transporte e Bairros)	Regressão Linear Múltipla
Larrañaga <i>et al.</i> , 2015	Transporte/ Planej. Urbano	Porto Alegre (Brasil)	Binária (realizou ou não uma viagem a pé)	Buffer (500m)	Regressão Logística Multinomial
Reis <i>et al.</i> , 2013	Saúde Pública	Curitiba (Brasil)	Binária (andou ou não mais de 150 minutos por semana)	Setor censitário	Regressão Logística Multinomial
Sanni & Abrantes, 2010	Transporte/ Planej. Urbano	Leeds (Reino Unido)	Proporção de viagens a pé	Setor censitário	Regressão Linear Múltipla Modelo de Durbin Espacial
Cervero <i>et al.</i> , 2009	Saúde Pública	Bogotá (Colômbia)	Binária (mais de 30 minutos de caminhada durante 5 dias ou não)	Buffer (500m e 1.000m)	Modelo Multinível (não-linear)
Amâncio & Sanches, 2008	Transporte/ Planej. Urbano	São Carlos (São Paulo)	Escolha discreta entre modos (Carro e Caminhada)	Setor censitário	Regressão Logística

Quadro 2.2 – Características dos dados utilizados nas referências.

Referência	Tipo de fonte de dados	Tipo de viagem	Perfil de idade
Guerra <i>et al.</i> , 2018	Secundária (Mexico's 2015 Intercensus)	Utilitária	-
Boulangé <i>et al.</i> , 2017	Secundária (Victorian Integrated Survey of Travel Activity 2009–2010)	Utilitária	+18
Lu <i>et al.</i> , 2017	Primária (questionário residencial) - 2014	Utilitária e Lazer (separadamente)	18-65
Ding <i>et al.</i> , 2017	Secundária (National Household Travel Survey 2001)	Iniciada na residência	+16
Sun <i>et al.</i> , 2017	Primária (questionário em parques públicos) - 2009	Trabalho	-
Tian & Ewing, 2017	Secundária (Oregon Household Travel and Activity Survey of 2011)	Iniciada na residência	-
Feuillet <i>et al.</i> , 2016	Primária (questionário eletrônico) - 2013	Tarefas e Lazer (separadamente)	+18
Koohsari <i>et al.</i> , 2016	Secundária (Physical Activity in Localities and Community Environments 2004)	Utilitária	20-65
Lamíquiz & López-Domínguez, 2015	Secundária (Regional Mobility Home-Based Survey 2004)	-	-
Larrañaga <i>et al.</i> , 2015	Primária (questionário residencial) - 2012	-	-
Reis <i>et al.</i> , 2013	Primária (questionário residencial) - 2102	Utilitária	18-65
Sanni & Abrantes, 2010	Secundária (2001 Census Travel to Work)	Trabalho	-
Cervero <i>et al.</i> , 2009	Primária (questionário residencial) - 2009	Utilitária	+18
Amâncio & Sanches, 2008	Secundária (Arruda, 2005)	-	-

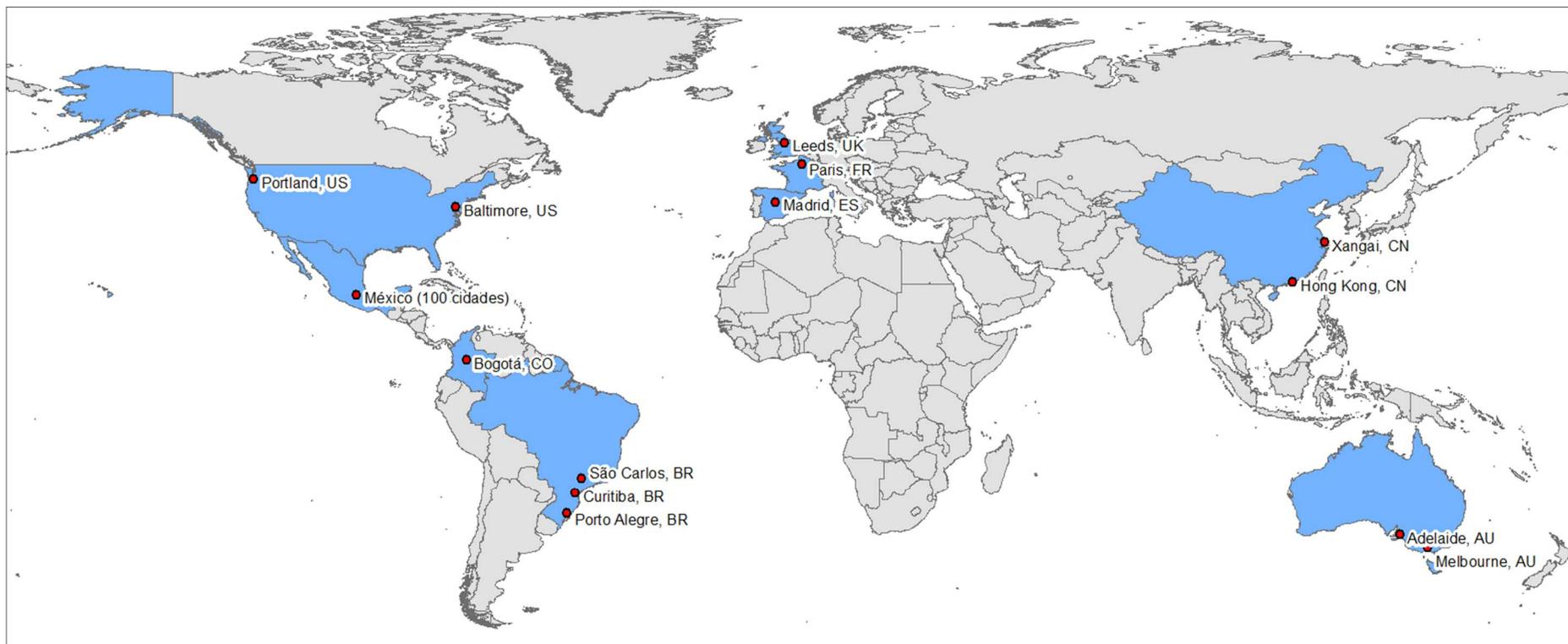


Figura 2.2 – Cidades estudadas na literatura.

Na literatura, as variáveis que representam o ambiente construído são computadas dentro de uma determinada área, que depende da abordagem adotada na pesquisa. Quando são analisados setores censitários ou zonas de transporte, a medida é relativa à área da zona ou setor que engloba o ponto de análise. Essa abordagem, apesar de facilitar o cálculo das variáveis, agrega os dados em áreas e pode gerar erros caso a escala dessas áreas não esteja adequada ao problema em questão ou se o ponto de referência estiver próximo ao perímetro. Esse problema está relacionado ao Problema da Unidade de Área Modificável (*Modifiable Areal Unit Problem – MAUP*). De forma resumida, esse problema ocorre devido à sensibilidade dos resultados quanto à forma e à escala de agregação de dados espaciais (ALMEIDA, 2012). Logo, algumas pesquisas utilizam como área de análise um *buffer* (área circular) centrado na localização em estudo para quantificação dos fatores, o que impõe um caráter local as variáveis do ambiente construído. Dentre as referências avaliadas, o raio da área circular varia entre 400 e 1.600 metros. Entretanto, essa abordagem não é completamente livre do MAUP, pois os resultados provavelmente dependem do tamanho de *buffer* adotado. Em um problema como a escolha do modo a pé, a escala deve ser adequada a esse tipo de modo, já que ele apresenta limitações quanto à distância viável de viagem (LAMÍQUIZ & LÓPEZ-DOMÍNGUEZ, 2015).

Quanto às técnicas de análise, os estudos utilizam a modelagem por regressão como ferramenta básica. Ela aparece em suas diferentes formas dependendo da variável de análise (dependente) e da base conceitual estabelecida em cada estudo. Entre as técnicas de análise estão formas comuns de regressão como as regressões linear, logística, de Poisson e binomial negativa. Também são aplicadas técnicas de combinação como Modelos Multinível e Cópula. Entretanto, merece destaque a utilização de modelos de regressão espacial por dois trabalhos (SANNI & ABRANTES, 2010; FEUILLET *et al.*, 2016). Esse tipo de regressão considera a correlação espacial que existe entre seus elementos e foi o adotado nesse trabalho. Mais detalhes sobre esse assunto podem ser vistos no Capítulo 3.

Os dados utilizados (Quadro 2.2) provêm principalmente de fontes secundárias, ou seja, eles não foram levantados especificamente para cada estudo. No geral são utilizados levantamentos realizados por alguma instituição de planejamento de transportes ou censos realizados por institutos de pesquisa e estatística, quando esses incluem pesquisas relacionadas aos padrões de transporte do indivíduo. Essa abordagem é adotada devido ao alto custo e à dificuldade de levantamento dos dados necessários (HANDY, 1996).

Os estudos costumam aplicar restrições ao tipo de viagem analisada e, em determinados casos, à faixa etária dos indivíduos que a realizam. A segunda restrição busca basicamente manter apenas indivíduos adultos. Já a primeira se justifica devido às diferenças que existem no processo de escolha do modo para diversos motivos de viagem. No geral, os estudos analisam viagens utilitárias, ou seja, viagens com o objetivo de se realizar alguma atividade no destino (trabalho, estudos ou outras tarefas). Além disso, alguns estudos avaliam apenas viagens iniciadas na residência.

A próxima seção inicia a discussão sobre as variáveis independentes analisadas na literatura. As variáveis foram categorizadas em variáveis de controle e nas 5 classes comumente utilizadas: densidade, diversidade, desenho, disponibilidade de destinos e distância para o transporte coletivo.

2.3. VARIÁVEIS DE CONTROLE

Em modelos de regressão, as variáveis de controle são variáveis que provavelmente afetam a variável dependente objeto da análise e que devem ser levadas em consideração, mesmo que não sejam o foco do estudo. Dessa forma, é possível isolar o efeito das variáveis independentes de interesse, que, neste trabalho, são as características do ambiente construído. Geralmente, essas variáveis são características socioeconômicas dos indivíduos, como idade, gênero, renda e escolaridade. Entretanto, outras podem ser utilizadas dependendo da análise realizada.

Essas variáveis foram classificadas em dez categorias e o Quadro 2.3 traz as diferentes formas em que essas categorias são medidas na literatura. Pelo caráter qualitativo de algumas das características utilizadas como variável de controle, essas são inseridas nos modelos como variáveis do tipo *dummy* ou binárias, ou seja, que assumem apenas os valores 0 ou 1.

Idade e gênero, características demográficas dos indivíduos, estão entre o tipo de variável mais presente nos estudos. Além disso, questões socioeconômicas como renda, emprego e escolaridade também representam um número significativo. Outro aspecto que é utilizado em alguns estudos é a presença ou não de crianças nos domicílios. Essa inclusão provavelmente tem base no fato de que ocupantes de residências com crianças podem tender a utilizar determinados modos de transporte devido às restrições impostas pelas características dessas.

Quadro 2.3 – Tipos de variáveis de controle.

Tipo	Especificações
Idade	<ul style="list-style-type: none"> • Idade em anos • Faixas de idade
Gênero	<ul style="list-style-type: none"> • Masculino ou Feminino
Disponibilidade de Veículo	<ul style="list-style-type: none"> • Quantidade de veículos • Nenhum veículo ou pelo menos um • Faixas de quantidade de veículos • Veículos/residentes
Renda	<ul style="list-style-type: none"> • Renda mensal ou anual familiar • Faixas de renda
Emprego	<ul style="list-style-type: none"> • Empregado ou não • Número de empregados na residência • Local de trabalho (casa ou não)
Crianças na Família	<ul style="list-style-type: none"> • Quantidade de crianças na residência • Nenhuma criança ou pelo menos uma
Aspectos da Residência	<ul style="list-style-type: none"> • Tamanho da residência • Tipo de residência • Tempo de moradia naquela residência • Número de moradores
Escolaridade	<ul style="list-style-type: none"> • Classes de escolaridade • Classes de anos de educação
Estado Civil	<ul style="list-style-type: none"> • Solteiro, casado ou outro
Atitudes e percepções	<ul style="list-style-type: none"> • Percepção do estado de saúde • Percepção sobre diferentes modos de transporte
Outros	<ul style="list-style-type: none"> • Etnia • Status socioeconômico • Fumante ou não fumante • Comprimento da viagem

A disponibilidade de veículos também está entre as categorias mais presentes nos estudos, o que é condizente com o objeto desses. É razoável considerar que a possibilidade de uso de automóvel influencia as decisões dos indivíduos. Logo, esse tipo de variável de controle pode representar um dos mais importantes fatores em estudos de escolha modal. Merece destaque a abordagem utilizada por Amâncio & Sanches (2008), que dividiram a quantidade de veículos pela quantidade de residentes, gerando assim uma medida de disponibilidade do automóvel.

Merece destaque ainda a utilização de medidas atitudinais em relação aos modos de transporte por Feuillet *et al.* (2016) e Larrañaga *et al.* (2015), ou seja, os modelos consideram as percepções e impressões dos indivíduos sobre os diferentes modos de transporte. Além disso, Amâncio & Sanches (2008) utilizaram a distância de viagem como medida de controle, já que é lógico assumir que quanto mais longa a viagem menor a probabilidade de uso do modo a pé.

Ainda, de forma objetiva e quando aplicável, o Quadro 2.4 apresenta as conclusões dos estudos selecionados sobre as variáveis de controle. É importante destacar que alguns estudos (FEUILLET *et al.*, 2016; KOOHSARI *et al.*, 2016; BOULANGE *et al.*, 2017; LU *et al.*, 2017) não buscaram tecer conclusões sobre os resultados encontrados para as variáveis de controle e focaram a análise apenas nas variáveis relativas às características do ambiente construído.

Quadro 2.4 – Conclusões com relação às variáveis de controle.

Referência	Conclusões
Guerra <i>et al.</i> , 2018	<ul style="list-style-type: none"> • Rendas altas levam a maior probabilidade de uso do automóvel e redução considerável da probabilidade de uso do transporte público e dos modos ativos; • O grau de instrução não apresentou resultado significativo na escolha do modo; • Pessoas até 44 anos de idade tendem a usar mais o automóvel do que modos não-motorizados; • Trabalhadores do setor informal são mais propensos aos modos ativos que aos modos motorizados.
Ding <i>et al.</i> , 2017	<ul style="list-style-type: none"> • Renda familiar e número de ocupantes na residência estão associados positivamente com a escolha pelo automóvel; • Número de crianças na residência está positivamente associado com a escolha por transporte público ou pelos modos ativos; • Pessoas mais novas (16-39) e mais velhas (+60) são mais prováveis de escolher o transporte público ou os modos ativos; • Pessoas com maior grau de instrução tendem a escolher o automóvel.
Sun <i>et al.</i> , 2017	<ul style="list-style-type: none"> • Existe associação positiva entre renda e a escolha pelo automóvel; • Homens tendem a escolher o automóvel; • Pessoas com 31-40 anos de idade são menos prováveis de andar até o trabalho; • Pessoas pertencentes a famílias maiores tendem a escolher o automóvel; • Pessoas com casa própria tendem a escolher o automóvel e inquilinos tem probabilidade maior de escolher os modos ativos.
Tian & Ewing, 2017	<ul style="list-style-type: none"> • No geral, características socioeconômicas são fortes preditoras das viagens a pé; • Tamanho da residência e quantidade de empregados influenciam positivamente a realização de viagens a pé; • A renda média domiciliar influencia negativamente a realização ou não de viagens a pé.

Referência	Conclusões
Larrañaga <i>et al.</i> , 2015	<ul style="list-style-type: none"> • Idade e renda médias e altas estão negativamente associadas com a decisão de realizar viagens a pé; • Atitude positiva em relação à caminhada está positivamente associada à decisão de viajar a pé; • Atitude positiva em relação ao veículo individual está negativamente associada à decisão de viajar a pé.
Reis <i>et al.</i> , 2013	<ul style="list-style-type: none"> • Não foi encontrada relação entre renda e tempo de caminhada.
Sanni & Abrantes, 2010	<ul style="list-style-type: none"> • Posse de veículo negativamente associada com a proporção de viagens realizadas a pé.
Cervero <i>et al.</i> , 2009	<ul style="list-style-type: none"> • Posse de veículo negativamente associada com o tempo de caminhada por dia.
Amâncio & Sanches, 2008	<ul style="list-style-type: none"> • Disponibilidade do automóvel positivamente associada com a escolha pelo veículo individual; • Distância da viagem negativamente associada com a escolha pelo modo a pé.

2.4. DENSIDADE

Segundo Sun *et al.* (2017), estudos mostram que a densidade, que é medida alternativamente por população, unidades habitacionais e emprego, está associada negativamente à distância da viagem, à duração da viagem e ao uso do carro. Segundo os autores, a literatura no geral estabelece que a alta densidade encurta as viagens, permite caminhar, andar de bicicleta e utilizar o transporte público, além de reduzir a dependência de carros particulares. No entanto, Cervero *et al.* (2009) chegaram a uma conclusão diferente para a cidade de Bogotá, já que a densidade não apresentou influência na utilização da caminhada como transporte.

Larrañaga *et al.* (2015) acrescentam que, pela natureza do cálculo, geralmente a densidade na origem e a densidade no destino são calculadas em separado. Entretanto, seria a combinação entre as duas que influenciaria no comportamento de viagem. Áreas com elevadas densidades habitacionais e baixa presença de comércio e serviços não promoveriam a redução de distâncias e a opção pela caminhada. O Quadro 2.5 resume as variáveis relativas à densidade encontradas na literatura.

Dentre as opções da tabela, a variável mais aplicada é a densidade populacional, que é substituída pela densidade de domicílios/habitações ou de empregos em alguns casos (CERVERO *et al.*, 2009; REIS *et al.*, 2013; BOULANGE *et al.*, 2017). Nota-se ainda a existência de variáveis que tentam medir o aproveitamento da área urbana para a construção (AMÂNCIO & SANCHES, 2008; CERVERO *et al.*, 2009).

Guerra *et al.* (2018) chegaram à conclusão que, para um residente médio mexicano, um aumento de 35% na densidade populacional gera um aumento de 38% na probabilidade de escolha pelo modo a pé. Sun *et al.* (2017) chegaram à conclusão semelhante para Xangai, mas quando se trata da densidade de empregos, a relação é invertida ao geralmente encontrado na literatura. Os autores afirmam que a densidade de trabalho em áreas residenciais está associada a uma maior probabilidade de dirigir, provavelmente resultante do fato de que chineses com maior renda geralmente vivem no centro da cidade (próximo a muitos empregos) e tem preferência por dirigir por questões de *status* social. Ademais, em outro estudo na China, Lu *et al.* (2017) obtiveram resultados que sugerem que a densidade populacional afeta a decisão de caminhar apenas em regiões com densidades baixas para o padrão de Hong Kong.

Quadro 2.5 – Resumo das variáveis relativas à densidade.

Tipo	Especificações
População	<ul style="list-style-type: none"> • Pessoas por unidade de área
Domicílios/habitações	<ul style="list-style-type: none"> • Domicílios ou habitações por unidade de área
Empregos	<ul style="list-style-type: none"> • Empregos por unidade de área
Atividade	<ul style="list-style-type: none"> • Empregos + População por unidade de área • Empregos + População + Estudantes por unidade de área
Comércio	<ul style="list-style-type: none"> • Estabelecimentos comerciais por unidade de área • Estabelecimentos comerciais de alimentação por unidade de área
Outros	<ul style="list-style-type: none"> • Altura média dos edifícios • Proporção de área ocupada por construções • Área construída pela área do terreno • Área construída pela área total

Boulangue *et al.*, 2017 constataram que ambientes com densidade residencial bruta superior a 20 habitações por hectare incentivam o uso de caminhadas, ciclismo e transporte público, ao mesmo tempo que desencorajam o uso do automóvel em Melbourne. Já Lamíquiz & López-Domínguez (2015) concluíram que a associação entre densidade e a mobilidade de pedestres em Madri é sutil e que seu comportamento é bastante assimétrico. Uma divisão modal de pedestres alta não foi estabelecida como limitante pela baixa densidade do bairro e, na verdade, ocorreu o contrário: sempre que o bairro tinha uma densidade mínima, a proporção de viagens de pedestres nunca caía abaixo de um certo limite. Esse aspecto pode ser consequência do contexto europeu onde a pesquisa foi desenvolvida, no qual existe um consenso de que as pessoas possuem comportamento mais voltado para a caminhada e para o transporte público.

Em contraponto, Cervero *et al.* (2009) encontraram que a densidade, diferente dos países desenvolvidos, com destaque para os Estados Unidos, não influencia no tempo semanal de caminhada em Bogotá. Os autores sugerem que isso se deve às condições de urbanização da cidade. As densidades em Bogotá são relativamente uniformes entre suas diferentes regiões devido, segundo os autores, ao processo característico de evolução urbana de países subdesenvolvidos. Esse aspecto tira o poder explicativo da variável na modelagem estatística. Entretanto, Larrañaga *et al.* (2015) obtiveram resultados em Porto Alegre que indicam que valores altos de densidade populacional e de comércio e serviços estimulam os deslocamentos a pé, resultado esse condizente com o geralmente encontrado para países desenvolvidos. Nessa mesma linha, Tian & Ewing (2017) afirmam que a densidade de atividades (população e emprego) está positivamente associada à decisão de caminhar como modo de transporte.

2.5. DIVERSIDADE

A diversidade é conceitualmente uma medida de relatividade entre diferentes tipos de uso do solo. O Quadro 2.6 traz alguns exemplos de métricas de diversidade de uso do solo encontrados na literatura. A entropia é a medida mais utilizada (GUERRA *et al.*, 2018) e assume valores entre 0 e 1. O primeiro representa apenas um uso e o segundo o perfeito balanceamento entre os diferentes usos considerados. É obtida pela seguinte expressão:

$$E = - \frac{\sum_j^n (p_j \times \ln(p_j))}{\ln(n)} \quad (2.1)$$

onde,

E = Entropia;

p_j = proporção do uso j ;

n = número total de usos considerados.

Segundo Boulange *et al.* (2017), indivíduos que utilizaram a caminhada como meio de transporte vivem em áreas com valores mais altos de entropia. Além disso, Ding *et al.* (2017) e Amâncio & Sanches (2008) constataram que o uso misto do solo, representado por entropia próxima a 1, está associado com probabilidades maiores de escolha pela caminhada. Ademais, Tian & Ewing (2017) indicam que, quanto maior a entropia, maior é o número de viagens a pé. Merece destaque que os estudos de Boulange *et al.* (2017), Ding *et al.* (2017) e Tian & Ewing (2017) utilizaram como objeto cidades de países ocidentais desenvolvidos (Baltimore, Portland e Melbourne, respectivamente). Por outro lado, Lu *et al.* (2017) e Cervero *et al.* (2009) não

encontraram relação significativa entre a entropia e o modo a pé para Hong Kong e Bogotá, respectivamente.

Ainda, com menos frequência, também é utilizado o balanço entre empregos ou estabelecimentos comerciais pelo número de residentes. Segundo Ewing & Cervero (2010), o balanço empregos/população possui relação positiva mais forte com a caminhada que a própria entropia. Já Lamíquiz & López-Domínguez (2015) encontraram baixa associação entre esse tipo de balanço e a proporção de viagens realizadas a pé em zonas de transporte.

Quadro 2.6 – Resumo das variáveis relativas à diversidade.

Tipo	Especificações
Entropia	<ul style="list-style-type: none"> • Índice de Entropia
Balanço entre empregos/comércio e população	<ul style="list-style-type: none"> • Empregos por residente • Estabelecimentos comerciais por residente
Outros	<ul style="list-style-type: none"> • Coeficiente de Gini entre setores censitários • Presença de edifícios de uso misto • Área construída com uso misto

2.6. DESENHO

O desenho está diretamente associado à configuração urbana, ou seja, à maneira como as vias se articulam para formar os caminhos e os quarteirões no espaço urbano. Além disso, também são consideradas nessa classificação determinadas amenidades comumente relacionadas ao aumento da atratividade da caminhada, como arborização, priorização frente aos modos motorizados e iluminação pública. O Quadro 2.7 traz as principais medidas encontradas na literatura.

Medidas relacionadas à quantidade ou densidade de interseções ou vias são predominantes na literatura. Segundo Guerra *et al.* (2018), um incremento de 10% na densidade de vias nas cidades mexicanas gera um aumento de 1% na probabilidade de escolha pelo carro como meio de transporte. Ademais, Lu *et al.* (2017) encontraram que a densidade de interseções não possui associação significativa com a caminhada para fins utilitários ou de lazer.

Em contraponto, Sun *et al.* (2017) colocam que residentes de áreas de Xangai com alta densidade de vias tendem a dirigir menos e a escolher outros modos de transporte. Os autores também citam que o aumento da proporção de interseções com quatro aproximações reduz a

probabilidade de escolha pelo automóvel, resultando semelhante ao obtido por Koohsari *et al.* (2016), que afirmam que a densidade de interseções está positivamente associada com a caminhada para o trabalho, e por Larrañaga *et al.* (2015), que indicam que configurações viárias com muitas interseções incentivam o modo a pé, visto que aumentam a conectividade. Resultado semelhante foi encontrado por Cervero *et al.* (2009).

Quadro 2.7 – Resumo das variáveis relativas ao desenho.

Tipo	Especificações
Interseções	<ul style="list-style-type: none"> • Densidade ou quantidade de interseções com 3 ou mais aproximações • Quantidade de interseções com 4 aproximações • Densidade ou quantidade de becos sem saída • Quantidade de interseções por número de vias
Vias	<ul style="list-style-type: none"> • Densidade de vias • Extensão total das vias
Amenidades	<ul style="list-style-type: none"> • Número de árvores por metro de via • Quantidade de semáforos para pedestres • Quantidade de dispositivos de moderação de tráfego
Sintaxe Espacial / Análise de Rede	<ul style="list-style-type: none"> • Integração • Inteligibilidade • Permeabilidade
Outros	<ul style="list-style-type: none"> • Declividade média • Tamanho médio dos lotes • Tamanho médio das quadras

Merece destaque a utilização da Sintaxe Espacial por Lamíquiz & López-Domínguez (2015) para a geração de medidas de desenho que pudessem ter sua correlação testada com a escolha pelo modo a pé. Segundo os autores, a Sintaxe Espacial foi concebida como uma teoria descritiva sobre formas espaciais e métricas oriundas desse campo mostraram associação com a caminhada em áreas urbanas.

Por fim, também merece destaque a escolha de Larrañaga *et al.* (2015) e de Cervero *et al.* (2009) em utilizar a quantidade de acidentes envolvendo pedestres como uma medida do desenho em suas análises. Essa abordagem tem a mesma base que a consideração da presença de amenidades nos estudos, ou seja, indivíduos cercados por ambientes mais hostis à caminhada tendem a escolher outros modos. Entretanto, sua utilização pode ser dificultada pela ausência de dados georreferenciados de acidentes.

2.7. DISPONIBILIDADE DE DESTINOS

A disponibilidade de destinos diz respeito à presença de localizações de interesse próximas ao ponto de origem da viagem. De certa forma, é possível afirmar que a disponibilidade é consequência da combinação da densidade, que aumenta o uso do solo, e da diversidade, que aproxima diferentes usos. O Quadro 2.8 traz as principais medidas encontradas na literatura.

Quadro 2.8 – Resumo das variáveis relativas à disponibilidade de destinos.

Tipo	Especificações
Comércio e Serviços	<ul style="list-style-type: none">• Distância para o centro comercial da cidade (<i>CBD</i>)• Densidade de comércio e serviços• Distância para o supermercado mais próximo• Quantidade de destinos (lojas, supermercados, farmácias, entre outros)• Pontuação de vida local (ponderação sobre a presença de destinos próximos)
Empregos	<ul style="list-style-type: none">• Total de empregos na área• Acessibilidade relativa a empregos

No geral, os estudos tratam a disponibilidade de destinos por meio da quantidade de comércio e serviços presente nos arredores da origem da viagem, geralmente representada pela residência. Os resultados de Boulange *et al.* (2017) indicam que o aumento de destinos dentro dos bairros aumenta a chance de utilização da caminhada como meio de transporte dentro do próprio bairro.

Nesse sentido, Larrañaga *et al.* (2015) encontraram elasticidade positiva e significativa entre a escolha pelo modo a pé e a densidade de comércio e serviços. Ademais, a distância da residência o centro comercial da cidade, outra variável utilizada para medir a presença de destinos de interesse, possui caráter regional (EWING & CERVERO, 2010) e costuma apresentar relação inversa à escolha pelo modo a pé e direta à escolha pelo veículo individual (SUN *et al.*, 2017). Entretanto, Koohsari *et al.* (2016) afirmam que o efeito positivo da presença de destinos próximos está associado à integração das vias (medida de desenho), que gera maior movimentação de pedestres, o que por sua vez atrai o comércio de caráter local.

Outra forma de quantificação da disponibilidade de destinos é a presença de empregos. Além de os próprios postos de trabalhos como destinos de interesse, áreas com mais empregos tendem a possuir mais serviços disponíveis. Segundo os resultados de Tian & Ewing (2017) para

Portland, quanto maior a quantidade de empregos acessíveis nos arredores, maior a probabilidade de ocorrerem viagens a pé.

2.8. DISTÂNCIA PARA O TRANSPORTE PÚBLICO

Várias pesquisas colocam que a distância para o transporte público é aspecto importante na escolha modal nos deslocamentos diários no meio urbano. A ideia geral dessa categoria é que um melhor acesso (menor distância) ao transporte público promoveria o uso desse modo frente aos demais. Logo, alguns estudos utilizam a distância ao ponto de acesso mais próximo, ao passo que outros utilizam medidas mais refinadas de disponibilidade, como a frequência de viagens, a capacidade ofertada ou a presença de serviços de alta capacidade. Merece destaque também a abordagem de Guerra *et al.* (2018), que definiram a quantidade de trabalhadores dos serviços de transporte urbano como uma medida de aproximação do tamanho da oferta. O Quadro 2.9 traz os parâmetros utilizados na literatura.

Quadro 2.9 – Resumo das variáveis relativas à distância para o transporte público.

Categoria	Especificações
Distância de acesso	<ul style="list-style-type: none">• Distância para os pontos de ônibus mais próximo• Distância para a estação de trem ou metrô mais próxima
Disponibilidade	<ul style="list-style-type: none">• Densidade de pontos de transporte público• Número de viagens (oferta) de transporte público• Extensão da rede de transporte de alta capacidade por tamanho da população
Outros	<ul style="list-style-type: none">• Trabalhadores no setor de transportes

Certo cuidado é necessário na avaliação da influência da distância para o transporte público e o modo a pé, já que essa relação depende da definição de viagem adotada. Caso o deslocamento até o ponto de acesso ao sistema seja considerado como uma viagem por si só, os dois modos são complementares e ganha força a hipótese de que existe associação positiva entre eles. Entretanto, se a viagem é definida como um deslocamento completo entre origem e destino, os modos passariam a ser concorrentes no processo de escolha, visto que o deslocamento a pé para acesso ao sistema seria desconsiderado.

Todavia, alguns estudos mostram que, apesar dessa “concorrência”, existe proporcionalidade entre as medidas de acesso ao transporte público e a caminhada como modo principal no deslocamento. Guerra *et al.* (2018) encontraram que há relação positiva da extensão da rede de

transporte de alta capacidade pelo tamanho da população com a escolha pelo modo a pé. Resultado semelhante foi obtido por Boulange *et al.* (2017) e Ding *et al.* (2017), que concluíram que quanto maior a distância para o ponto de transporte público mais próximo menos são realizadas viagens a pé.

Esse aspecto pode ter duas explicações. A primeira é que a presença de bons sistemas de transporte público é comumente associada a outras características urbanas que, segundo a literatura, promovem a escolha pelo modo a pé, como a densidade e a diversidade. A segunda possibilidade é que a presença de mais pessoas nas ruas se deslocando para acessar o transporte público melhora a sensação de segurança, o que incentiva o uso da caminhada, fenômeno esse descrito por Jacobs (1961).

2.9. CONSIDERAÇÕES FINAIS

No geral, os resultados das pesquisas na área indicam relação positiva entre a densidade, a diversidade, a disponibilidade de destinos, a acessibilidade ao transporte público e a conectividade das vias urbanas (desenho) com a escolha pelo modo a pé. Esse tipo de resultado, já apontado por Ewing & Cervero (2010), é condizente com o atual estado da teoria na área, que embasa tendências como o Desenvolvimento Orientado ao Transporte Sustentável.

Entretanto, é notável que nem sempre todas essas características se mostram importantes no processo de escolha. Segundo Cervero *et al.* (2009), somente as medidas de desenho possuem relação com a caminhada em Bogotá. Lu *et al.* (2017) não encontraram relação entre a diversidade, representada pela entropia, e o modo a pé em Hong Kong. Além disso, em diversas situações nos estudos encontrados, não foram feitos comentários pelos autores sobre determinadas variáveis do ambiente construído que não tiveram resultado estatisticamente significativo. A falta de significância indica que não é possível afirmar que existe relação entre a variável independente e a variável dependente.

Portanto, fica evidente a importância de aprofundamento dos estudos no assunto, principalmente com a utilização de técnicas diferenciadas, como os modelos espaciais, e de dados de diferentes contextos urbanos.

3. MODELOS DE ESCOLHA DISCRETA E ESPACIAIS

Segundo Ortúzar & Willumsen (2011), os modelos de escolha discreta postulam que a probabilidade de escolha por determinada opção depende da atratividade dessa opção e das características de quem faz a escolha. Esses modelos são baseados na existência de uma variável latente, comumente chamada de utilidade, que determina a escolha adotada ou a ocorrência ou não de determinado evento. Para determinar se uma alternativa será escolhida ou não, o valor da utilidade deve ser comparado com a utilidade de outras opções e transformado em um valor entre 0 e 1 (ORTÚZAR & WILLUMSEN, 2011). Normalmente, são utilizados os modelos Logit, que possui o erro seguindo uma distribuição Gumbel, e Probit, que possui erro seguindo uma distribuição Normal. Esse segundo pode ser expressado pela seguinte expressão:

$$\begin{cases} y = 1, & \text{se } y^* \geq 0 \\ y = 0, & \text{se } y^* < 0 \end{cases}$$
$$y^* = \alpha + X\beta + \varepsilon \quad (3.1)$$
$$\varepsilon \sim N(0, \sigma^2)$$

onde,

y = variável dependente binária relativa;

y^* = variável latente;

X = matriz de variáveis independentes;

α, β = coeficientes da regressão;

ε = erro aleatório seguindo uma distribuição normal com média zero e variância σ^2 .

Esse tipo de modelo é muito utilizado na pesquisa e no planejamento de transportes, principalmente em análises de escolha modal. Segundo Meyer (2016), o uso de modelos de escolha discreta revolucionou a modelagem da escolha dos indivíduos, principalmente pela evolução na medição e especificação das variáveis explicativas utilizadas.

Entretanto, o espaço tem grande importância em estudos relacionados ao desenvolvimento urbano, como os de transporte, já que os dados coletados e analisados nesse tipo de trabalho possuem a localização como fator fundamental a ser considerado. Esse aspecto pode resultar na inadequação dos modelos tradicionais de escolha discreta a problemas espaciais, visto que pressupostos básicos, como a independência entre as observações da amostra e os erros

independentes e identicamente distribuídos, podem não se mostrar verdadeiros. Segundo Anselin (1988) e LeSage (1999), dados espaciais são diferenciados pela existência de dois problemas: a dependência espacial (autocorrelação espacial) e a heterogeneidade (estrutura espacial).

O primeiro diz respeito à dependência que uma amostra em determinada região tem com as amostras de outras regiões da área de estudo. Druck *et al.* (2004) citam que esse conceito parte da chamada Primeira Lei da Geografia, enunciada por Tobler (1970): “tudo está relacionado a todo o resto, mas aquilo que está mais próximo se parece mais que aquilo que está mais distante”. Como exemplo prático no contexto desse trabalho, é possível argumentar que a diversidade do uso do solo de determinado setor censitário está relacionada à densidade dos setores vizinhos, assim como é lógico considerar que essa dependência diminui com a distância. LeSage (1999) cita duas razões para a dependência espacial ocorrer:

- A coleta de dados baseada em divisões de área (setor censitário, bairro, zona de transporte) pode apresentar erros se o zoneamento adotado não reflete adequadamente o fenômeno de estudo;
- A dimensão espacial de problemas regionais e urbanos geralmente representa um fator importante nos modelos, já que a localização importa nesses fenômenos.

Já o segundo problema representa a variação das relações estudadas no espaço. Esse problema considera que, por exemplo, determinada relação encontrada entre variáveis de diferentes setores censitários da zona sul de uma cidade pode ser diferente da encontrada na zona norte. De forma simples, a heterogeneidade é o fenômeno no qual a dependência espacial citada anteriormente não é homogênea ao longo de toda a área de estudo.

Nesse sentido, alguns estudos da área de transportes buscaram considerar os efeitos espaciais em suas análises. Sanni & Abrantes (2010) utilizaram o *Spatial Durbin Model* (SDM) para estudar a proporção de viagens realizadas a pé nos setores censitários de Leeds, Reino Unido. Os autores concluíram que o uso do modelo espacial melhora o ajuste da modelagem, logo os resultados são mais consistentes. Além disso, Feuillet *et al.* (2016) utilizaram modelos geograficamente ponderados para determinar a relação entre a quantidade horas de caminhada por semana e as características do ambiente construído em Paris. A utilização do modelo espacial local permitiu aos autores a avaliação da existência de *clusters* no espaço urbano do comportamento estudado.

Este Capítulo busca principalmente apresentar as características dos modelos espaciais e suas possibilidades de aplicação. A primeira seção trata sobre a matriz de proximidade espacial, ou de ponderação espacial, elemento essencial para esses modelos. A segunda trata dos indicadores mais utilizados para verificação da dependência espacial. Já a terceira seção descreve alguns modelos de regressão espacial existentes e a quarta trata de formas de quantificar o ajuste desses modelos aos dados. Por fim, a quinta e última seção traz as considerações finais sobre a temática.

3.1. MATRIZ DE PROXIMIDADE ESPACIAL

LeSage (1999) afirma que a primeira tarefa que deve ser realizada antes da análise da dependência espacial e da heterogeneidade é a quantificação dos aspectos locacionais do conjunto de dados que estão sendo analisados. O autor cita duas fontes de informação para determinação da estrutura espacial: a localização cartesiana em latitude e longitude e a contiguidade.

A localização nos permite o cálculo da distância entre todos os pontos da amostra, que é a base para o cálculo dos efeitos espaciais esperados, ou seja, pontos mais próximos possuem maior dependência espacial e se comportam de forma semelhante. Já a contiguidade está relacionada à posição relativa entre os pontos da amostra, como, por exemplo, a existência ou não de fronteira entre diferentes bairros.

É importante destacar que a distância não necessariamente precisa ser entendida como um parâmetro físico. Ela pode ser tomada de modo amplo, não apenas geográfico, mas também social, econômico ou político (ALMEIDA, 2012; ANSELIN, 1988). Por exemplo, a distância pode ser entendida como a quantidade de contatos que separam duas pessoas em determinada rede social em estudos de comportamento (GOLGHER, 2015). Ou então, o tempo de deslocamento entre dois pontos em estudos de transporte (SILVA, 2006).

Similar à abordagem utilizada para grafos ou redes, a estrutura espacial é formalmente expressa por uma matriz de proximidade espacial (ANSELIN, 1988). Nela, cada unidade espacial é representada nas linhas e nas colunas e os campos indicam qual a relação espacial existente entre as unidades. Logo, o elemento W_{ij} de uma matriz de proximidade espacial indica a relação

existente entre as unidades espaciais i e j , de acordo com a concepção adotada para a estrutura espacial.

Anselin (1988) explica que a noção original de dependência espacial era baseada na contiguidade entre unidades espaciais. Nessa concepção, os elementos da matriz recebem valor 1 quando existe fronteira entre as unidades espaciais e valor 0 quando essa fronteira não existe. Um exemplo dessa abordagem é mostrado na Figura 3.1. Nota-se que, por definição, a diagonal dessa matriz é igual a zero ($W_{ii} = 0$).

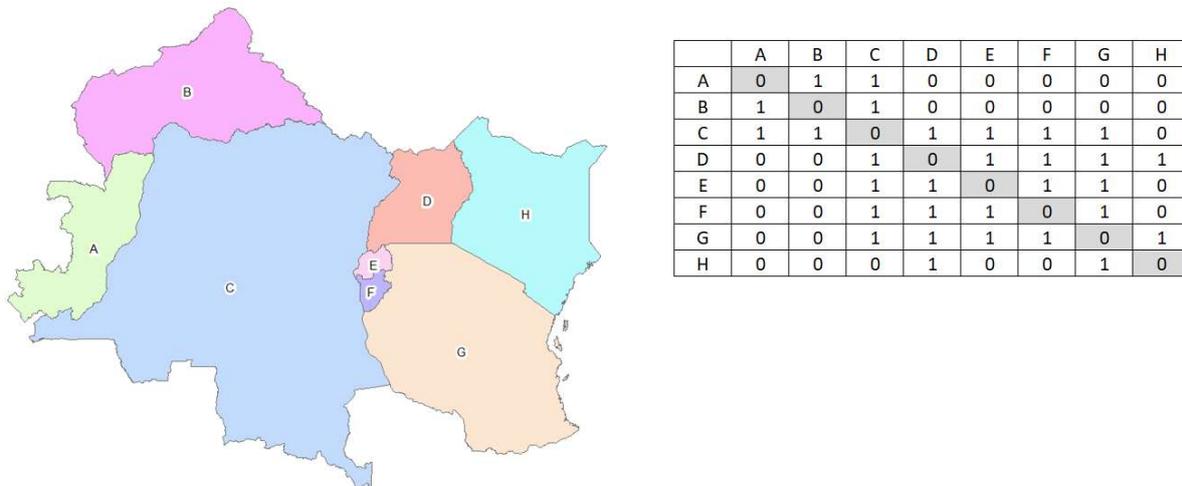


Figura 3.1 – Exemplo de Matriz de Proximidade Espacial.

Fonte: Druck *et al.* (2004) com adaptações.

Existem ainda diversas maneiras de construção de matrizes de proximidade com base na contiguidade, como a consideração de fronteiras apenas em determinadas direções ou de unidades que façam fronteiras com unidades que circundam a unidade de referência (contiguidade de segunda ordem). Segundo LeSage (1999), definições mais elaboradas às vezes são utilizadas pela consideração do tamanho das fronteiras entre as unidades. Dessa forma, são relativizados os casos nos quais a fronteira existe, mas tem dimensões reduzidas.

Além disso, as matrizes de proximidade podem se basear nas distâncias geográficas entre os elementos de estudo. Os elementos aqui considerados podem ser, por exemplo, a distância entre centroides de setores censitários ou a distância entre as residências onde foram coletados os dados. Um exemplo apresentado por Almeida (2012) é a matriz baseada no inverso da distância entre dois pontos, na qual pode ser aplicado um ponto de corte. Ou seja, a partir de determinada

distância limite, a matriz recebe o valor 0. Existem outros exemplos, como a função de distância exponencial (SANNI & ABRANTES, 2010; ALMEIDA, 2012), a função inversa do quadrado da distância (GOLGHER, 2015) e forma binária na qual é atribuído o valor 1 para distâncias menores que determinado ponto de corte e 0 para distâncias maiores.

Druck *et al.* (2004) destacam ainda que, como a matriz é aplicada no cálculo de indicadores de dependência espacial, é útil a normalização de suas linhas para que a soma de seus elementos seja igual a 1. Entretanto, é preciso cuidado nesse tratamento, já que a resolução de determinados modelos espaciais exige que a matriz de proximidade espacial seja invertível.

LeSage (1999) destaca que o princípio básico para definição da matriz deve ser a natureza do problema em si. Logo, sua especificação deve estar alinhada à base conceitual do problema que está sendo estudado. Um fenômeno que conceitualmente possui sua estrutura espacial baseada na distância entre os elementos não deve então ser especificado pela existência ou não de fronteiras.

Por fim, a falta de critérios objetivos para a especificação da matriz de proximidades é uma das críticas em relação ao uso dos modelos espaciais (GOLGHER, 2015), problema esse também apontado por Anselin (2010). Na prática, a abordagem mais comum é o teste dos modelos espaciais com diferentes especificações para a matriz e a escolha daquela que apresenta maior aderência em relação aos dados.

3.2. ÍNDICES DE AUTOCORRELAÇÃO ESPACIAL

Segundo Almeida (2012), os índices de autocorrelação espacial buscam testar se os dados espaciais estão aleatoriamente distribuídos, ou seja, se não existe dependência espacial nos seus valores. Druck *et al.*, (2004) destacam que a caracterização da dependência espacial, objetivo desses coeficientes, é uma etapa essencial da análise exploratória de dados espaciais. A seguir são apresentados os principais índices utilizados na literatura.

3.2.1. *I* de Moran

O índice *I* de Moran, desenvolvido por Moran (1950), é o principal indicador utilizado para a quantificação de dependência espacial. Silva (2006) coloca que sua concepção é baseada no

coeficiente de correlação de Pearson, mas a introdução da ponderação pela proximidade espacial é determinada na matriz de proximidades. Seu valor é dado por:

$$I = \frac{n}{\sum_i^n \sum_j^n w_{ij}} \times \frac{\sum_i^n \sum_j^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i^n (x_i - \bar{x})^2} \quad (3.2)$$

onde,

$I = I$ de Moran;

x_i = valor da variável x na região i ;

x_j = valor da variável x na região j ;

\bar{x} = média dos valores de x ;

w_{ij} = elemento da linha i e coluna j da matriz de proximidade;

n = tamanho da amostra.

Segundo Druck *et al.* (2004), o I de Moran adota como hipótese nula a independência espacial, o que é caracterizado quando seu valor é igual a zero. O índice varia de -1 (correlação negativa) a 1 (correlação positiva). Os autores destacam ainda que importante determinar se o índice calculado possui validade estatística, sendo a abordagem mais comum a associação do índice a uma distribuição normal. Entretanto, outra possibilidade é realização de um teste de pseudo-significância. Logo, o I de Moran fornece três informações: se existe autocorrelação espacial ou não (significância estatística), se essa dependência é positiva ou negativa (sinal do resultado) e sua dimensão (distância do valor nulo).

3.2.2. C de Geary

Outro índice de dependência espacial foi o proposto por Geary (1954). Segundo Almeida (2012), a diferença em relação ao índice I de Moran reside na medida de covariância adotada. Semelhante ao índice anterior, a hipótese nula é a inexistência de dependência espacial. Entretanto, essa hipótese no índice C de Geary é representada pelo valor 1, sendo que o índice varia entre 0 (dependência espacial positiva) e 2 (dependência espacial negativa). Ele é dado pela expressão:

$$C = \frac{n - 1}{2 \sum_i^n \sum_j^n w_{ij}} \times \frac{\sum_i^n \sum_j^n w_{ij} (x_i - x_j)^2}{\sum_i^n (x_i - \bar{x})^2} \quad (3.3)$$

onde,

$C = C$ de Geary;

x_i = valor da variável x na região i ;

x_j = valor da variável x na região j ;

\bar{x} = média dos valores de x ;

w_{ij} = elemento da linha i e coluna j da matriz de proximidade;

n = tamanho da amostra.

3.2.3. Índices Locais

Os dois indicadores apresentados são globais, ou seja, eles consideram que as relações espaciais são únicas e constantes em toda a área de estudo. Entretanto, como explicado anteriormente, a estrutura espacial pode variar, o que prejudica o teste dos indicadores apresentados. Em casos nos quais são analisadas muitas áreas, é provável que ocorram diferenças na estrutura espacial (DRUCK *et al.*, 2004). Os Indicadores Locais de Associação Espacial ou *Local Indicators of Spatial Association* (LISA), termo designado por Anselin (1995), buscam analisar, unidade espacial por unidade espacial, a dependência espacial com as unidades próximas. Logo, é obtido um resultado específico para cada área, o que permite a identificação de áreas que possuem dependência espacial muito diferente das outras, evidenciando assim a existência de *clusters*.

Um dos índices disponíveis na literatura é o índice local de Moran (Anselin, 1995). Almeida (2012) destaca que a soma proporcional do índice local de Moran para diversas áreas representa o I de Moran global. Logo, é possível afirmar que o primeiro é uma decomposição do segundo. O índice local é dado pela expressão:

$$I_i = \frac{x_i \sum_j^n w_{ij} x_j}{\sum_j^n x_j^2} \quad (3.4)$$

onde,

$I_i = I$ de Moran local na região de interesse i ;

x_i = valor da variável x na região de interesse i ;

x_j = valor da variável x nas regiões periféricas j ;

w_{ij} = elemento da linha i e coluna j da matriz de proximidade;

n = tamanho da amostra.

3.3. MODELOS DE REGRESSÃO ESPACIAL

Segundo Fávero & Belfiore (2017), as técnicas de regressão são as mais utilizadas em análise de dados que buscam o entendimento da relação entre variáveis preditoras (independentes) e um fenômeno específico (dependente), mesmo que haja apenas uma relação de correlação e não causalidade. Tais modelos servem tanto para a análise do efeito que determinadas variáveis exercem em um fenômeno, quanto para modelagem e predição. Existem diversos modelos de regressão muito utilizados na pesquisa científica, como a Regressão Linear, a Regressão Logística e a Regressão de Poisson.

Os modelos de regressão espacial são variações dos modelos clássicos de regressão que de alguma forma incorporam as relações espaciais em sua concepção. Logo, em teoria, sua utilização melhora a aderência da regressão em estudos com característica espacial. Anselin (2010) explica que variáveis espacialmente ajustadas podem ser incluídas para a variável dependente, para as variáveis explicativas e termos de erro, bem como combinações destes, gerando um conjunto extenso de modelos espaciais disponíveis.

A Figura 3.2 mostra a variedade de modelos lineares disponíveis pela inclusão da dependência espacial em seus diferentes elementos. A seguir são apresentados cinco diferentes modelos de regressão espacial. É importante destacar que todos os modelos aqui apresentados são globais, ou seja, consideram que a estrutura espacial é uniforme em toda a área de estudo. Semelhante aos índices da seção anterior, a utilização de modelos globais de regressão para a análise de fenômenos com variações locais pode gerar distorções nos resultados. Como exemplo de modelo local de regressão é possível citar a Regressão Geograficamente Ponderada.

Segundo LeSage & Pace (2009), a estimação de modelos espaciais pelo Método dos Mínimos Quadrados (MMQ), ou Mínimos Quadrados Ordinários (MQO), pode resultar em estimativas inconsistentes dos parâmetros da regressão, dos parâmetros espaciais e dos erros. Os autores afirmam que o Método da Máxima Verossimilhança é consistente para a estimação desses modelos.

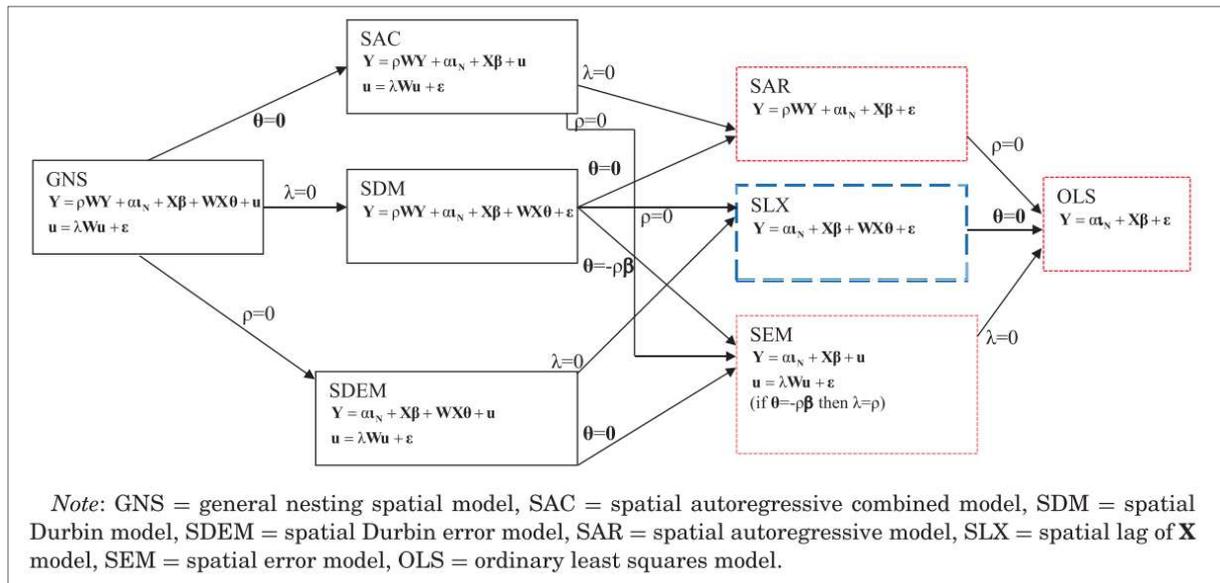


Figura 3.2 – Comparação de modelos espaciais.

Fonte: Vega & Elhorst (2015).

3.3.1. Modelo Espacial Autoregressivo Misto

O Modelo Espacial Autoregressivo Misto, ou *Spatial Auto Regressive* (SAR), é resultado da combinação entre um modelo tradicional de regressão e uma estrutura espacial autoregressiva (LESAGE & PACE, 2009). Nesse modelo, a interação espacial ocorre entre as variáveis dependentes, como ilustrado na Figura 3.3. Tal modelo é denominado misto devido a existência de variáveis explicativas exógenas. Segundo LeSage & Pace (2009), o modelo é expresso por:

$$\begin{aligned}
 y &= \rho W y + \alpha \iota_n + X \beta + \varepsilon \\
 y &= (I_n - \rho W)^{-1} (\alpha \iota_n + X \beta) + (I_n - \rho W)^{-1} \varepsilon \\
 \varepsilon &\sim N(0, \sigma^2 I_n)
 \end{aligned}
 \tag{3.5}$$

onde,

y = variável dependente;

X = variáveis independentes;

ι_n = vetor $n \times 1$ com todos os valores unitários;

W = matriz de proximidade espacial;

α, β = coeficientes da regressão;

ρ = coeficiente autoregressivo espacial;

I_n = matriz identidade $n \times n$;

ε = erro aleatório seguindo uma distribuição normal com média zero e variância σ^2 .

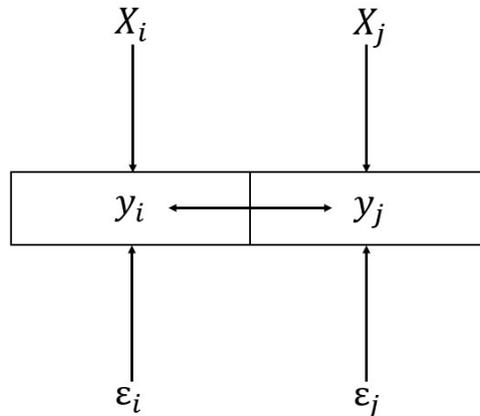


Figura 3.3 – *Spatial Auto Regressive* (SAR).

Fonte: Almeida (2012).

O coeficiente autoregressivo espacial (ρ) é o elemento que traz a dependência espacial para o modelo. Um valor positivo indica autocorrelação espacial positiva, ou seja, valores altos ou baixos em determinado local estão relacionados, respectivamente, a existência de valores altos ou baixos em locais próximos (ALMEIDA, 2012). Já valores negativos indicam a relação inversa. A hipótese nula desse modelo é que ρ é igual a zero, logo, se ele não for estatisticamente significativo, considera-se que não existe a dependência espacial.

3.3.2. Modelo com Erro Espacial Autoregressivo

O Modelo com Erro Espacial Autoregressivo, ou *Spatial Error Model* (SEM), considera a dependência espacial nos erros, ou seja, nos termos não modelados que tem influência na variável dependente. A Figura 3.4 ilustra as relações nesse tipo de modelo. O modelo é expresso por (LESAGE & PACE, 2009):

$$\begin{aligned} y &= \alpha I_n + X\beta + u \\ u &= \rho W u + \varepsilon \\ \varepsilon &\sim N(0, \sigma^2 I_n) \end{aligned} \quad (3.6)$$

onde,

y = variável dependente;

X = variáveis independentes;

ι_n = vetor $n \times 1$ com todos os valores unitários;

W = matriz de proximidade espacial;

α, β = coeficientes da regressão;

ρ = coeficiente autoregressivo espacial;

u = erro das observações;

ε = erro aleatório seguindo uma distribuição normal com média zero e variância σ^2 ;

I_n = matriz identidade $n \times n$.

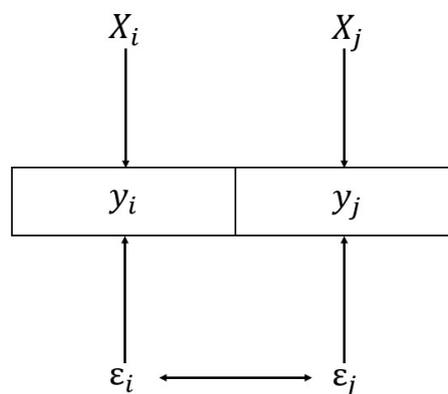


Figura 3.4 – *Spatial Error Model* (SEM).

Fonte: Almeida (2012).

Segundo Almeida (2012), os erros das observações (u) são uma média dos erros das regiões de influência mais um erro aleatório (ε). Semelhante ao modelo SAR, a hipótese nula é que o coeficiente autoregressivo espacial é igual a zero, o que indicaria que a dependência espacial não existe. Logo, a estimativa desse coeficiente precisa ser estatisticamente significativa para que se possa considerar a existência da dependência espacial.

3.3.3. Modelo Regressivo Cruzado Espacial

O Modelo Regressivo Cruzado Espacial, ou *Spatial Lag of X* (SLX), considera que a dependência espacial ocorre nas variáveis independentes, ou seja, os efeitos dessas variáveis transbordam espacialmente. Logo, o modelo considera que determinada unidade espacial não é somente influenciada pelos valores de suas respectivas variáveis independentes, mas também pelos valores das variáveis das áreas que a cercam. A Figura 3.5 esquematiza as relações desse modelo.

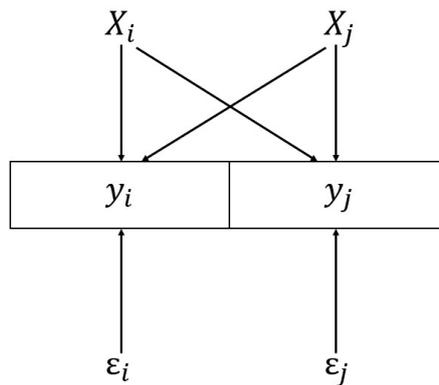


Figura 3.5 – *Spatial lag of X (SLX)*.

Fonte: Almeida (2012).

Segundo Almeida (2012), a defasagem ocorre pela inclusão de componentes de transbordamento espacial como variáveis independentes do modelo. Esses componentes nada mais são que os valores das variáveis independentes ponderados pela matriz de proximidades, que define a estrutura espacial existente. É importante destacar que nem todas as variáveis independentes precisam ter o transbordamento espacial considerado no modelo. O modelo SLX é dado pela seguinte expressão:

$$y = X\beta + WX\gamma + \varepsilon \quad (3.7)$$

onde,

y = variável dependente;

X = variáveis independentes;

W = matriz de proximidade espacial;

γ = vetor de coeficientes autoregressivos espaciais;

ε = erro aleatório seguindo uma distribuição normal com média zero e variância σ^2 .

3.3.4. Modelo Espacial Durbin

O Modelo Espacial Durbin, ou *Spatial Durbin Model (SDM)*, é uma variação do modelo SAR com a inclusão da dependência espacial nas variáveis explicativas, além das variáveis dependentes. De certa forma, é possível afirmar que ele é uma combinação dos modelos SAR

e SLX. A Figura 3.6 ilustra as relações desse modelo. Segundo LeSage & Pace (2009), ele é dado por:

$$\begin{aligned}
 y &= \rho W y + \alpha \iota_n + X \beta + W X \gamma + \varepsilon \\
 y &= (I_n - \rho W)^{-1} (\alpha \iota_n + X \beta + W X \gamma + \varepsilon) \\
 \varepsilon &\sim N(0, \sigma^2 I_n)
 \end{aligned}
 \tag{3.8}$$

onde,

y = variável dependente;

X = variáveis independentes;

ι_n = vetor $n \times 1$ com todos os valores unitários;

W = matriz de proximidade espacial;

α, β = coeficientes da regressão;

ρ, γ = coeficientes autoregressivos espaciais;

I_n = matriz identidade $n \times n$;

ε = erro aleatório seguindo uma distribuição normal com média zero e variância σ^2 .

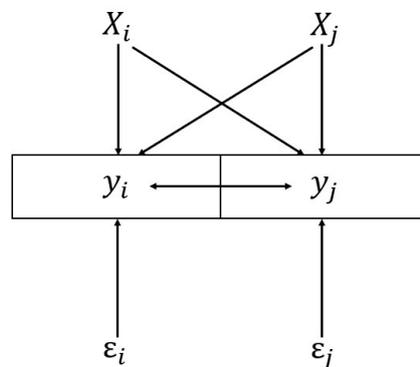


Figura 3.6 – *Spatial Durbin Model (SDM)*.

Fonte: Almeida (2012).

3.3.5. Modelo Probit Espacial

Diante do exposto sobre modelos de escolha discreta e sobre os modelos espaciais, é possível ainda a aplicação de efeitos espaciais em modelos com variável dependente categórica. LeSage & Pace (2009) indicam que a inclusão da dependência espacial na variável dependente em modelos de escolha discreta resulta na situação em que as escolhas realizadas em um

determinado local são similares às escolhas realizadas nas regiões próximas. Os autores utilizam como exemplo exatamente a escolha modal em fluxos pendulares diários, já que viajantes com origens próximas enfrentam as mesmas condições de acesso ao transporte público e a outros modos de transporte. Já Smirnov (2010) entende a dependência espacial em modelos de escolha discreta como a dependência entre as preferências dos indivíduos que estão próximos. Logo, surge o fenômeno da decisão influenciada pelo contexto social, representado nesse caso pelo espaço, visto que existe dependência entre as decisões.

LeSage & Pace (2009) descrevem o modelo Probit espacial em sua forma binária da seguinte forma:

$$\begin{cases} y_i = 1, & \text{se } y_i^* \geq 0 \\ y_i = 0, & \text{se } y_i^* < 0 \end{cases}$$

$$y_i^* = (I_n - \rho W)^{-1} X \beta + (I_n - \rho W)^{-1} \varepsilon \quad (3.9)$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

onde,

y_i = variável dependente binária relativa ao ponto i ;

y_i^* = variável latente relativa ao ponto i ;

X = variáveis independentes;

I_n = vetor $n \times 1$ com todos os valores unitários;

W = matriz de proximidade espacial;

β = coeficiente da regressão;

ρ = coeficiente autoregressivo espacial;

I_n = matriz identidade $n \times n$;

ε = erro aleatório seguindo uma distribuição normal com média zero e variância σ^2 .

As equações indicam que caso a utilidade de ocorrência do evento (1) supere a utilidade da não ocorrência (0), ele ocorrerá. Caso contrário, o evento não ocorrerá. Nota-se que a expressão utilizada para definição da utilidade é igual à expressão do modelo SAR. LeSage & Pace (2009) indicam ainda que, da mesma forma que o modelo SAR pode ser expandido para o modelo SDM, o modelo Probit aqui apresentado pode incorporar a dependência espacial em suas variáveis explicativas.

Visto que o objeto do presente estudo, a escolha pelo modo a pé, representa uma variável binária, o modelo Probit espacial foi escolhido para a análise da relação entre as medidas do ambiente construído, estudadas no Capítulo anterior, e a escolha pela caminhada. Mais considerações sobre o método adotado serão apresentadas no Capítulo 4.

3.4. AJUSTE DE MODELOS DE REGRESSÃO ESPACIAL

A análise do ajuste de modelos espaciais é um importante aspecto dos estudos que os aplicam. Anselin (1988) afirma que essa é uma tarefa mais complexa do que no caso de modelos não espaciais, já que não existe medida de ajuste padronizada como o coeficiente de ajuste (R^2). Em casos de erros não esféricos, o R^2 perde sua função de indicador da qualidade da regressão (ALMEIDA, 2012).

Anselin (1988) coloca a utilização dos critérios de informação de Akaike (AIC) e de Schwarz ou bayesiano (BIC) como possíveis medidas de ajuste de modelos espaciais. Entretanto, o autor recomenda cautela na utilização de ambos os critérios, visto que não é claro qual dos dois deve ser utilizado em cada caso. Os critérios são baseados no valor da função de log-verossimilhança (LIK) e quanto menor o valor do critério, melhor o ajuste do modelo.

Outra forma de comparação de ajuste baseada na log-verossimilhança é o pseudo- R^2 de McFadden, que possui valor entre 0 e 1. A ideia desse indicador é que, se o modelo não tem potencial para predição, a sua verossimilhança não será muito maior que a do modelo considerando apenas o intercepto. Ele é calculado pela seguinte expressão:

$$R_{McFadden}^2 = 1 - \frac{LIK}{LIK_{null}} \quad (3.9)$$

onde,

$R_{McFadden}^2$ = pseudo- R^2 de McFadden;

LIK_{null} = log-verossimilhança do modelo sem variáveis explicativas;

LIK = log-verossimilhança do modelo;

3.5. CONSIDERAÇÕES FINAIS

Diante do exposto neste Capítulo, é evidente o potencial que a econometria espacial e seus modelos tem para a pesquisa e o planejamento nas áreas de transportes e de desenvolvimento urbano. A consideração explícita do espaço, fator inerente a qualquer estudo de transportes, pode melhorar o resultado de pesquisas e de modelos de simulação. Ademais, a dependência espacial pode ocorrer de formas diferentes, logo existem vários modelos e indicadores que buscam considerá-la ou quantificá-la. Nesse sentido, o método desse trabalho, apresentado no próximo Capítulo, buscou incorporar esses conceitos. Portanto, espera-se comparar os resultados de modelos espaciais aos de não espaciais, de forma a avaliar, complementarmente ao objetivo principal do trabalho, se a consideração dos efeitos espaciais melhora o ajuste das técnicas de análise de regressão.

4. MÉTODO PARA DESENVOLVIMENTO DO TRABALHO

O presente Capítulo apresenta o método utilizado neste trabalho. Adotou-se uma abordagem semelhante aos estudos levantados na literatura, que utilizam a modelagem por regressão para a investigação das relações entre a o ambiente construído e o modo a pé. A Figura 4.1 ilustra as etapas do método.

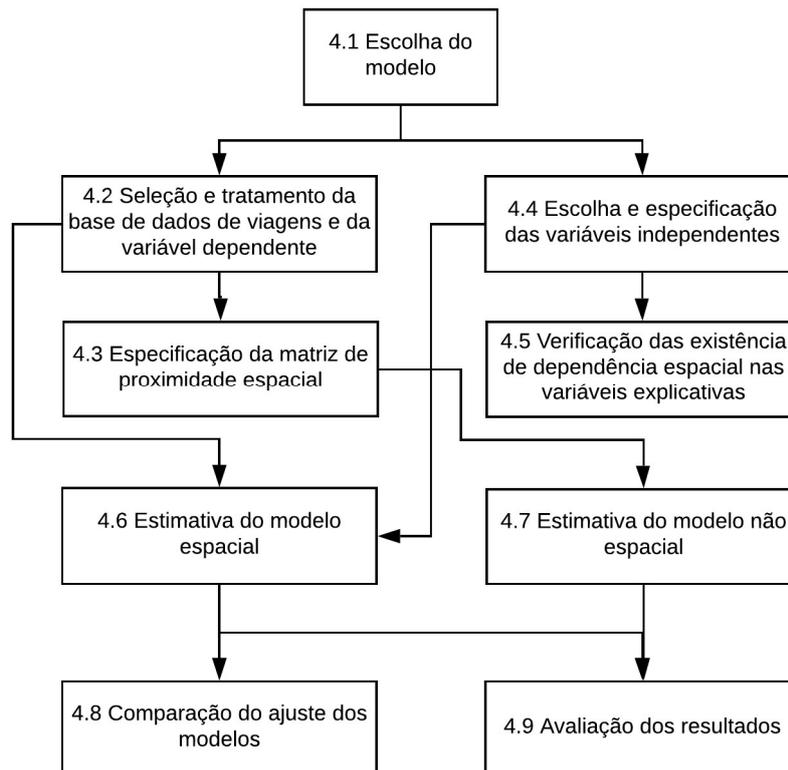


Figura 4.1 – Método aplicado.

4.1. ESCOLHA DO MODELO

Segundo Ortúzar & Willumsen (2011), modelos de escolha aplicados a viagens preveem a probabilidade de um indivíduo escolher alguma alternativa específica baseados em uma variável latente, normalmente chamada de utilidade, que dá uma medida de comparação entre as opções disponíveis. Segundo a autora, essa abordagem tem forte básica teórica e chega mais perto de testar relações causais que outros métodos, já que a análise é desagregada e baseada no indivíduo e no seu processo decisório.

Além disso, Páez & Scott (2004) citam que a única forma de realizar análises urbanas de forma independente e evitando o Problema da Unidade de Área Modificável (MAUP), citado no Capítulo 2, é utilizando dados ao nível do indivíduo. Nesses casos, o dado é representado por coordenadas geográficas e não por áreas de análise.

Portanto, para a investigação proposta nesse trabalho, foi escolhido como modelo base o Probit com defasagem espacial na variável dependente (endógena) descrito por LeSage & Pace (2009) e apresentado no Capítulo anterior. Segundo Smirnov (2010), o Probit espacial é mais utilizado na literatura científica do que o modelo Logit espacial, que é o mais popularmente aplicado no planejamento de transportes em sua versão não espacial. Anselin (2002) indica que o Probit espacial possui vantagem em relação ao Logit espacial devido à distribuição normal do erro. Além disso, Hahn & Soyer (2005) encontraram diferença prática de ajuste entre os modelos Probit e Logit apenas em situações específicas que não se encaixam no contexto desse trabalho.

Além do modelo com dependência espacial na variável dependente, também foi avaliada a aplicação da dependência nas variáveis independentes, quando o teste descrito mais adiante indicar autocorrelação espacial significativa. Logo, a abordagem aqui adotada se assemelha ao modelo SDM (*Spatial Durbin Model*) apresentado no Capítulo anterior. É importante destacar que, pelo caráter exógeno das variáveis independentes, o processo de estimação do modelo não é afetado pela inclusão da dependência espacial relacionada a essas variáveis, semelhante ao que ocorre quando se parte do modelo SAR para o modelo SDM.

4.2. SELEÇÃO E TRATAMENTO DA BASE DE DADOS DE VIAGENS E DA VARIÁVEL DEPENDENTE

Segundo Handy (1996), umas das abordagens mais promissoras para o estudo da influência do ambiente construído no comportamento de viagens é a utilização de dados de agências de planejamento urbano ou de outras instituições públicas, como institutos de pesquisa e estatística. Isso ocorre porque tais instituições possuem recursos e meios para a realização de grandes pesquisas periódicas focadas em padrões de transporte. Entretanto, merece destaque a utilização do *big data* na pesquisa e no planejamento de transportes, pois recentemente passou o cruzamento de diversas bases de dados ficou mais prático (MILNE & WATLING, 2018).

Logo, para esse estudo optou-se pela utilização da base de dados da Pesquisa Origem e Destino de São Paulo/SP, realizada a cada dez anos desde 1967 para a Região Metropolitana pela Companhia do Metropolitano de São Paulo (Metrô-SP). Diferente da maioria das pesquisas desse tipo, as informações de ponto de origem e ponto de destino foram tabuladas a nível de coordenadas geográficas, o que permite análises desagregadas. Além disso, a pesquisa levantou diversas características socioeconômicas e há disponibilidade de outras informações do Município, como uso do solo e distribuição de empregos.

A pesquisa foi realizada por meio de entrevistas domiciliares, que buscam levantar todas as viagens realizadas por todos os ocupantes da residência no dia útil típico anterior, com informações de origem, destino, horário, modo e motivo. Logo, é possível estimar um panorama da mobilidade local. Segundo o Metrô/SP, foi obtido um total aproximado de trinta mil domicílios com entrevistas válidas.

É importante destacar que foi utilizada a base da pesquisa realizada em 2007, visto que a mais atual, de 2017, ainda está na fase de tabulação e consolidação dos dados, logo suas informações ainda não estavam disponíveis até o momento de conclusão deste trabalho. Entretanto, apesar de não ser a abordagem ideal, como visto no Quadro 2.2, diversos estudos da literatura utilizaram bases de dados defasadas em relação ao ano do trabalho. Essa deficiência ocorre devido à dificuldade para a realização de levantamentos desse tipo.

A base de dados da OD de 2007 contém um total de 196.698 viagens registradas. Entretanto, nem todas foram consideradas para esse trabalho. Os seguintes filtros foram aplicados para seleção da amostra:

- Retirada das entradas com dados inconsistentes ou ausentes: em alguns casos faltam dados necessários para a estimação do modelo, provavelmente devido a resíduos no banco de dados da pesquisa, e em outros os dados são inconsistentes, como nos casos em que a origem e o destino estão com as coordenadas iguais. Logo, essas viagens foram excluídas;
- Seleção apenas de viagens internas ao Município de São Paulo, já que os dados de uso e ocupação de solo e empregos só estão disponíveis para essa cidade;
- Seleção apenas de viagens iniciadas na residência (coordenadas da origem igual às coordenadas da residência). Essa abordagem é a mesma adotada por Tian & Ewing (2017). Neste trabalho ela é justificada pela dependência que viagens iniciadas em

outros locais tem com a primeira viagem iniciada na residência. É lógico assumir que o modo escolhido para sair da residência afetará a escolha do modo para voltar ou para atingir alguma outra localidade a partir do primeiro destino;

- Apenas pessoas com mais de 18 anos de idade, conforme utilizado na literatura (CERVERO *et al.*, 2009; REIS *et al.*, 2013; BOULANGE *et al.*, 2017)
- Apenas viagens por motivo de trabalho ou estudo. Essa consideração busca a utilização apenas de viagens utilitárias, como observado na literatura (SANNI & ABRANTES, 2010; SUN *et al.*, 2017; BOULANGE *et al.*, 2017; GUERRA *et al.*, 2018), e que tenham um caráter cotidiano e essencial, já que essas viagens representam grande parte dos deslocamentos no meio urbano;
- Apenas viagens cuja distância de deslocamento seja no máximo igual a 4 quilômetros, visto que, a partir dessa distância, pode-se considerar que a caminhada deixa de ser uma opção de modo (VALE *et al.*, 2018). Logo, a inclusão de viagens acima dessa distância enviesaria os resultados, pois indicaria uma representação menor do modo a pé. Essa distância foi medida em linha reta (distância euclidiana) entre a origem e o destino e está presente na própria base de dados da pesquisa.

Todos esses filtros estão relacionados ou à consistência do banco de dados ou à base teórica do problema de pesquisa. Entretanto outras considerações foram realizadas devido ao método adotado. Primeiro, para que não aparecessem linhas repetidas na matriz de proximidades, optou-se pela utilização de apenas pontos de origem e destino não sobrepostos. O seguinte procedimento foi adotado para evitar o enviesamento da seleção da amostra:

- Sorteio aleatório para cada origem (conjunto latitude e longitude) existente na base de apenas uma entrada da amostra (viagem);
- Do conjunto de viagens selecionadas na etapa anterior, faz-se um sorteio aleatório para cada destino (conjunto latitude e longitude) existente na base de apenas uma entrada da amostra (viagem).

Segundo, como descrito mais adiante, foram testadas diferentes distâncias de corte com base nos pontos de origem para a construção da matriz de proximidades. Entretanto, para distâncias menores, alguns pontos da amostra resultante dos cortes anteriores ficam sem “vizinhos” dentro daquele raio (linha e coluna relativas àquele ponto zeradas na matriz), o que invalida os testes

e a modelagem. Logo, foram excluídas da amostra todas as entradas em que essa questão surgia para a menor distância testada (400 metros).

Portanto, a amostra final considerada é igual a 7.329 viagens, das quais 42% foram realizadas a pé (3.068). Além disso, 74% das viagens tiveram como motivo o trabalho em diferentes setores (6% na indústria, 16% no comércio e 52% setor de serviços). Ainda, 54% dos indivíduos são mulheres e 84% possuem trabalho. A Tabela 4.1, a Tabela 4.2 e a Tabela 4.3 trazem a os perfis de escolaridade, socioeconômico e etário dos indivíduos presentes na amostra.

Tabela 4.1 – Perfil de escolaridade.

Escolaridade	Quantidade	%
Não-Alfabetizado/Primário Incompleto	293	4,00%
Primário Completo/Ginásio Incompleto	775	10,57%
Ginásio Completo/Colegial Incompleto	1.118	15,25%
Colegial Completo/Superior Incompleto	2.687	36,66%
Superior Completo	2.456	33,51%
Total	7.329	100,00%

Tabela 4.2 – Perfil socioeconômico (Critério de Classificação Econômica Brasil).

Perfil	Quantidade	%
A1	284	3,88%
A2	1.004	13,70%
B1	1.348	18,39%
B2	1.674	22,84%
C1	1.604	21,89%
C2	983	13,41%
D	397	5,42%
E	35	0,48%
Total	7.329	100,00%

Tabela 4.3 – Perfil etário.

Faixa de Idade	Quantidade	%	% Acumulada
18 - 25	1.058	14,44%	14,44%
25 - 30	974	13,29%	27,73%
30 - 35	945	12,89%	40,62%
35 - 40	925	12,62%	53,24%
40 - 45	867	11,83%	65,07%
45 - 50	806	11,00%	76,07%
50 - 55	687	9,37%	85,44%
55 - 60	421	5,74%	91,19%
60 - 65	281	3,83%	95,02%
65 - 70	196	2,67%	97,69%
70 - 100	169	2,31%	100,00%
Total	7.329	100,00%	

Tendo em mente os objetivos desse trabalho, a variável dependente binária do modelo adotado será a escolha ou não do modo a pé como modo de transporte nas especificações descritas anteriormente. Logo, ela assumiu valor igual a 1 quando esse modo foi escolhido e 0 quando qualquer outro modo foi escolhido. Essa mesma abordagem e modelo espacial foram aplicados por Wang *et al.* (2015), mas em relação à escolha da bicicleta como meio de transporte. É importante destacar que a tabulação da Pesquisa OD considerou como viagem a pé apenas viagens em que esse modo foi utilizado do início ao fim, logo não são considerados deslocamentos para acesso a outros modos. Além disso, a pesquisa desconsiderou as viagens a pé com distância inferior a 500 metros que não fossem por motivo de trabalho ou estudo, o que não afeta este estudo, já que foram selecionadas apenas viagens para esses dois motivos.

4.3. ESPECIFICAÇÃO DA MATRIZ DE PROXIMIDADE ESPACIAL

Como explicado no Capítulo anterior, a matriz de proximidade espacial (W) precisa ser baseada na estrutura espacial do fenômeno estudado. Dentre as duas principais categorias de formas de construção da matriz, contiguidade e distância geográfica, a primeira não se aplica a esse trabalho, já que foram analisados diferentes pontos no espaço (origens e destinos) e não áreas que possuem fronteiras entre si.

Portanto, a Matriz foi construída com base na distância entre os pontos da amostra. Para um determinado ponto, consideram-se seus vizinhos, recebendo o valor 1 na matriz, aqueles que estão até determinada distância de corte. Pontos com distância maior a essa distância receberam

valor zero na matriz. Essa abordagem torna o cálculo computacionalmente mais leve por zerar muitos elementos da matriz.

Essa distância de corte foi investigada pela aplicação do *I* de Moran descrito no Capítulo anterior na variável dependente considerando matrizes de proximidade construídas com base em diferentes pontos de corte (de 400 a 4.000 metros, de 200 em 200 metros). O teste foi executado utilizando o pacote “lctools”, publicado por Stamatis Kalogirou, na versão gratuita do *software* RStudio, que é um ambiente de desenvolvimento integrado para R, linguagem muito utilizada em análises estatísticas. A distância que indicou maior dependência espacial (valor mais próximo de 1) foi a adotada.

É importante destacar que foram construídas duas matrizes nesta pesquisa: uma com base nos pontos de origem das viagens, que foi utilizada para a variável dependente, para as variáveis de controle e para as variáveis do ambiente construído levantadas na origem; e outra com base nos pontos de destino das viagens, que foi utilizada para as variáveis do ambiente construído levantadas no destino.

4.4. ESCOLHA E ESPECIFICAÇÃO DAS VARIÁVEIS INDEPENDENTES

Com base na literatura analisada no Capítulo 2, foram selecionadas as variáveis apresentadas no Quadro 4.1. A tabela indica a unidade da variável, a forma de levantamento e a base de dados utilizada como referência. Em alguns casos reportados na literatura, a análise das características do ambiente construído é realizada com base na origem, que, como adotado nesse trabalho, geralmente representa a residência do indivíduo. Entretanto, os resultados encontrados por Sun *et al.* (2017) para Xangai indicam que em determinadas situações as características do destino são mais relevantes que as da origem. A importância do destino também é apontada por Vale *et al.* (2018). Logo, optou-se neste estudo pelo levantamento dessas variáveis em ambos os pontos, de forma a promover essa investigação no contexto urbano de São Paulo.

Na literatura, as variáveis que representam o ambiente construído comumente são levantadas com *buffers* ao redor do ponto de análise e essa abordagem foi a adotada neste trabalho. Entretanto, Tian & Ewing (2017) colocam que a utilização de *buffers* de rede representam de forma mais acurada o ambiente ao redor de determinado ponto que *buffers* circulares comuns. Essa conclusão tem sentido, principalmente em estudos de transporte, visto que a distância

realmente importante para o deslocamento não é a euclidiana, mas sim a distância dentro da rede de vias. Logo, para o levantamento das variáveis, serão aplicados *buffers* às vias acessíveis em determinada distância de caminhada a partir do ponto de análise (*sausage network buffer*).

Essa abordagem foi analisada por Forsyth *et al.* (2012) e Oliver *et al.* (2007), que comparam diferentes métodos de análise por *buffer*. Os autores colocam que *buffers* de rede gerados diretamente por *softwares* como o ArcGIS, também chamados de áreas de serviço, carecem de metodologia clara de execução, logo a aplicação de *buffers* a rede de vias acessível (*sausage network buffer*) é uma boa alternativa para levantamento de características espaciais. A Figura 4.2 compara um *buffer* circular tradicional, uma área de serviço gerada pelo *Network Analyst Tools* do ArcGIS, e um *sausage network buffer*.

Já a distância aplicada nos *buffers* variou bastante na literatura. Dentre os estudos levantados que utilizam *buffers* de rede, Boulange *et al.* (2017) aplicaram a distância de 1.600 metros e Tian & Ewing (2017) testaram a utilização de raios de 400, 800 e 1.600 metros. Nesse segundo estudo não foram observadas diferenças significativas entre os resultados dos diferentes *buffers*, logo é incerto qual distância é a mais adequada. Portanto, optou-se nesse estudo pela modelagem com diferentes distâncias dentro da amplitude encontrada na literatura (400, 800, 1.200 e 1.600 metros). O modelo que indicou o melhor ajuste foi o considerado para a análise dos resultados. A execução dos *buffers* foi realizada em um Sistema de informação Geográfica (*Geographic Information System*) em duas etapas: levantamento, para cada ponto de origem e destino, da rede de vias acessível em 300, 700, 1.100 e 1.500 metros de distância e aplicação de *buffer* de 100 metros à rede de vias obtida na primeira etapa. Essa mesma abordagem é utilizada por Oliver *et al.* (2007).

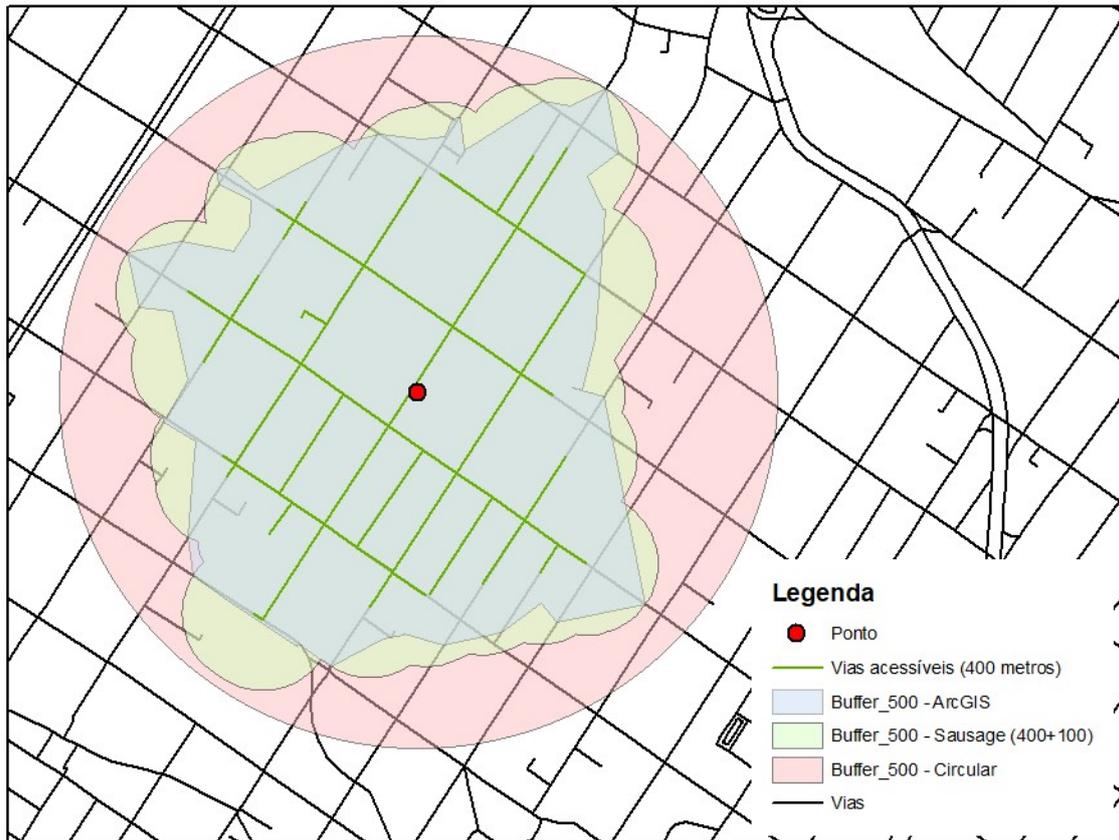


Figura 4.2 – Comparação entre diferentes tipos de *buffer* (500 metros).

É importante destacar que a utilização de *buffers*, devido à personalização da área de análise para cada ponto da amostra, diminui os problemas de agregação de dados em áreas gerais, como setores censitários e zonas de transporte. Entretanto, os resultados ainda podem ser sensíveis ao tamanho escolhido para o *buffer*.

Ainda, são necessárias algumas observações sobre o levantamento das variáveis independentes. O uso do solo teve como base o Cadastro Territorial e Predial, de Conservação e Limpeza (TPCL), mantido atualmente pelo Departamento de Arrecadação e Cobrança (DECAR) da Secretaria Municipal de Fazenda da Prefeitura de São Paulo. Foi utilizado o cadastro com referência 2008, que considera os dados atualizados pelo Poder Público no ano anterior (2007). É um cadastro de base fiscal, com a finalidade de permitir o lançamento e a cobrança do Imposto Predial e Territorial Urbano (IPTU) sobre a propriedade imobiliária. Nele são indicadas por lote informações como área construída, valor do terreno, tipo de uso, entre outros. Logo, buscou-se o agrupamento dos usos presentes na base em 4 diferentes tipos: residencial, comercial, industrial e outros (hospitais, escolas, templos, clubes, entre outros). Além disso, foram desconsiderados os terrenos vazios. Entretanto, essas informações não estão georreferenciadas

a nível de lote, sendo possível apenas a sua associação espacial a nível de quadra fiscal, que possui base georreferenciada disponível. O seguinte procedimento foi então adotado:

- Consolidação, para cada quadra fiscal, da quantidade total de área construída de cada tipo de uso do solo (A_{ji} , onde j é o tipo de uso e i a identificação da quadra). A utilização da área construída considera toda a área da construção e não somente a área de sua projeção horizontal. Logo, o cálculo da proporcionalidade de ocupação entre os diferentes tipos fica mais refinada;
- Determinação da proporção de área construída de cada tipo para cada quadra ($p_{ji}=A_{ji}/A_i$, onde A_i é área construída total da quadra fiscal);
- Determinação do aproveitamento da área da quadra fiscal ($q_i=A_i/A_{Hi}$, onde A_{Hi} é a área da projeção horizontal da quadra i);
- Levantamento, para as quadras fiscais que possuem interseção com o respectivo *buffer* da amostra, das áreas dessas quadras abrangidas pelo *buffer* (A_{bi}), procedimento semelhante ao adotado para computação da densidade populacional com os setores censitários;
- Cálculo da área construída total do *buffer* (A) pelo somatório das áreas das quadras fiscais abrangidas pelo *buffer* (A_{bi}) multiplicadas pelos respectivos aproveitamentos (q_i);
- Cálculo da área construída total do *buffer* ocupada por determinado tipo de uso do solo (A_j) pelo somatório das áreas das quadras abrangidas pelo *buffer* multiplicadas pelos respectivos aproveitamentos (q_i) e proporção do uso (p_{ji});
- Obtenção da proporção de área construída para determinado tipo de uso (p_j) pela divisão da área total construída desse uso (A_j) pela área total construída (A).

Já para a variável “Distância para o centro”, optou-se pela utilização da referência do centro expandido da cidade. O centro expandido foi estabelecido pela Companhia de Engenharia de Tráfego (CET) de São Paulo como a área onde se aplica o rodízio de veículos de acordo com a placa e o dia da semana. O centro expandido é visto como a área que concentra grande parte dos serviços e empregos da capital paulista, logo foi adotado como referência para esse trabalho. A variável é representada pela distância em quilômetros ao centro expandido e recebeu valor igual a 0 quando o ponto estava localizado dentro de seu perímetro, sendo esta a mesma abordagem adotada por Sun *et al.* (2017).

A densidade de empregos foi computada a partir de dados georreferenciados e disponibilizados pelo Centro de Estudos da Metrópole (CEM/Cepid). Essa base foi elaborada a partir do cadastro central de empresas do Instituto Brasileiro de Geografia e Estatística (IBGE). Ela indica a localização das empresas no município e outras informações como tipo de atividade e faixa de quantidade de empregos da unidade local (0-4, 5-9, 10-49, 50-99, 100-499, maior ou igual a 500). Para o cálculo da densidade de empregos considerou-se o ponto médio das faixas utilizadas e o valor de 500 empregos para a última categoria.

Além disso, para a Escolaridade optou-se pela não utilização de variáveis *dummy*, mas pela consideração de anos de estudo de acordo com as classes presentes no banco de dados da pesquisa. Nesse sentido, foram considerados apenas os ciclos de estudo completo e a duração padrão de cada um: não-alfabetizado (nenhum ano de estudo), primário completo (4 anos), ginásio completo (8 anos), colegial completo (11 anos) e superior completo (16 anos).

Quadro 4.1 – Variáveis independentes selecionadas.

Tipo	Nome	Código	Unidade/descrição	Fonte	Levantamento
Controle	Idade	Idade	Anos	OD São Paulo 2007	Idade indicada no banco de dados
	Gênero	Gen	1: Masculino 0: Feminino		Gênero indicado no banco de dados
	Crianças no domicílio	Crian	Número de pessoas com menos de 16 anos na família		Para a família correspondente, contabilizadas quantas pessoas tem menos de 16 anos
	Disponibilidade de veículos/motos	Disp_veic	Veículos e motos por membro da família		Para a família correspondente, soma do número de veículos mais o número de motos, dividido pelo número de membros da família com mais de 18 anos
	Renda familiar	Ren	Reais por mês		Renda familiar indicada no banco de dados
	Escolaridade	Esc	Anos de estudo: • Não-Alfabetizado/Primário Incompleto: 0 • Primário Completo/Ginásio Incompleto: 4 • Ginásio Completo/Colegial Incompleto: 8 • Colegial Completo/Superior Incompleto: 11 • Superior Completo: 16	Grau de instrução indicado no banco de dados	
Densidade	Densidade populacional	Den_pop_o Den_pop_d	Pessoas por quilômetro quadrado	Censo IBGE 2010	Ponderação da densidade populacional dos setores censitários que interceptam o <i>buffer</i> pelas áreas desses setores abrangidas pelo <i>buffer</i>
	Densidade de empregos	Den_emp_o Den_emp_d	Empregos por quilômetro quadrado	Cadastro Central de Empresas do IBGE (2000)	Quantidade de empregos dentro da área de <i>buffer</i> dividida pela área total
Diversidade	Entropia	Ent_o Ent_d	Intervalo de 0 (apenas um uso) a 1 (equilíbrio perfeito entre os usos)	TPCL (2008)	$E = - \frac{\sum_j^n (p_j \times \ln(p_j))}{\ln(n)}$ <p>p_j = proporção do uso j n = número total de usos considerados</p>

Tipo	Nome	Código	Unidade/descrição	Fonte	Levantamento
Desenho	Densidade de vias	Den_via_o Den_via_d	Quilômetro de via por quilômetro quadrado	<i>Open Street Maps</i>	Extensão total de vias dentro do <i>buffer</i> dividida pela área desse
	Densidade de interseções	Den_int_o Den_int_d	Interseções com 3 ou mais aproximações por quilômetro quadrado	<i>Open Street Maps</i>	Quantidade de nós da rede ligados a 3 ou mais <i>links</i> dentro do <i>buffer</i> dividida pela área desse
	Densidade de vias sem saída	Den_said_o Den_said_d	Vias sem saída por quilômetro quadrado	<i>Open Street Maps</i>	Quantidade de nós da rede ligados a apenas 1 <i>link</i> dentro do <i>buffer</i> dividida pela área desse
Disponibilidade de destinos	Distância para o centro	Dist_CE_o Dist_CE_d	Quilômetros	CET (2018)	Distância do ponto ao perímetro do centro expandido. Foi considerada igual a zero para pontos dentro dessa área
Distância para o Transporte Público	Densidade de pontos de parada e estações de trem ou metrô	Den_tp_o Den_tp_d	Pontos/estação por quilômetro quadrado	Portal Geosampa	Quantidade de pontos dentro do <i>buffer</i> dividida pela área desse

4.5. VERIFICAÇÃO DA EXISTÊNCIA DE DEPENDÊNCIA ESPACIAL NAS VARIÁVEIS EXPLICATIVAS

Nessa etapa foi calculado o índice I de Moran global descrito no Capítulo 3 em todas as variáveis explicativas (independentes) consideradas nesse trabalho e levantadas pela aplicação da distância de *buffer* encontrada na etapa anterior. O objetivo era determinar em quais variáveis ocorre a dependência espacial entre suas observações. Logo, a inclusão de efeitos espaciais pela matriz de proximidade espacial ocorreu apenas nas variáveis independentes que apresentaram I de Moran acima de 0,5 e significativo estatisticamente.

É importante destacar que o efeito espacial medido pelo I de Moran não é exatamente o mesmo presente em modelos como o SDM e SLX. Esses modelos consideram os efeitos das variáveis explicativas relativas aos vizinhos na variável dependente no ponto da observação. Já o I de Moran quantifica a dependência espacial entre as observações de cada uma das variáveis explicativas. Entretanto, a existência dessa dependência de forma elevada pode indicar alta interação espacial nos processos que envolvem essas variáveis.

Novamente, a matriz de proximidades aplicada para as variáveis do ambiente construído na origem não teve os mesmos valores da utilizada para as variáveis do destino, já que as localizações geográficas desses pontos são diferentes. Além disso, para análise da dependência espacial das variáveis de controle, relacionadas as características socioeconômicas, foi considerada a matriz construída com base nos pontos de origem.

4.6. ESTIMATIVA DO MODELO ESPACIAL

Quanto ao método de estimação dos parâmetros, Smirnov (2010) cita que abordagens por máxima verossimilhança para modelos espaciais de escolha discreta demandam mais capacidade computacional e que normalmente o processo de calibração dos parâmetros é realizado por aproximações ou simulações. Uma alternativa é a abordagem bayesiana desenvolvida por LeSage & Pace (2009). Entretanto, nessa pesquisa, foi utilizado o pacote “ProbitSpatial”, publicado por Davide Martinetti, também no *software* RStudio, que utiliza a máxima verossimilhança para a estimação do modelo. Dessa forma, foi possível utilizar o valor da log-verossimilhança para posteriores comparações de ajuste entre modelos.

4.7. ESTIMATIVA DO MODELO NÃO ESPACIAL

Para fins de comparação, também foi estimado um modelo Probit não espacial com base nas mesmas variáveis, de forma a comparar seu ajuste ao do modelo espacial. O modelo foi estimado utilizando o pacote “stats”, publicado pelo R-Project, no *software* RStudio.

4.8. COMPARAÇÃO DO AJUSTE DOS MODELOS

Tanto para comparação do ajuste entre as diferentes distâncias de *buffer* da Seção 4.4, quanto para comparação entre o modelo espacial e o não espacial, foi utilizado o pseudo-R² de McFadden apresentado no Capítulo anterior. Segundo Fávero & Belfiore (2017), esse critério é muito utilizado para a comparação de ajuste e seleção de modelos, na qual o modelo com maior pseudo-R² é o de melhor ajuste.

4.9. AVALIAÇÃO DOS RESULTADOS

Os resultados dos modelos Probit foram avaliados de acordo com os seguintes critérios:

- Significância estatística dos parâmetros;
- Sinal dos parâmetros;
- Efeitos marginais das variáveis independentes.

O primeiro buscou indicar quais parâmetros possuem resultados estatisticamente significantes. Já o segundo faz parte da investigação se a direção da relação entre a escolha pelo modo a pé e as variáveis está coerente ou não com os resultados encontrados na literatura para os diversos contextos urbanos.

Já os efeitos marginais, medida muito utilizada para análise de modelos de escolha discreta, é a medida de aumento ou redução de probabilidade da escolha a partir da redução ou aumento do valor de determinada variável. Logo, eles indicam a importância que as variáveis têm para a escolha e o efeito que sua alteração gera, caso todo o resto se mantenha constante. Como seu valor é diferente para cada ponto da amostra, a abordagem mais comum, que foi adotada nesse trabalho, é o cálculo da média dos efeitos nos diferentes pontos (*Average Marginal Effects*).

Os efeitos marginais médios são uma boa forma de análise para avaliação de políticas, visto que eles proveem informação do tipo “caso tal medida seja adotada, esse será o impacto”. Eles

foram calculados por meio do pacote “margins”, publicado por Thomas Leeper, no *software* RStudio. No tema tratado neste trabalho, eles podem ajudar a responder a seguinte pergunta: no caso de alteração das características do ambiente construído por políticas de desenvolvimento urbano, qual será o impacto na escolha pelo modo a pé para transporte?

4.10. CONSIDERAÇÕES FINAIS

As etapas do método buscam então alcançar os objetivos elencados no Capítulo 1. No próximo Capítulo foram apresentados os resultados obtidos. O Quadro 4.2 relaciona as etapas do método às Seções do Capítulo 5.

Quadro 4.2 – Relação entre as etapas do método e os resultados.

Método	Resultados	Descrição
4.3	5.1 Matriz de proximidade	Resultados dos testes para escolha da distância de corte
4.4	5.2 Escolha de distância de <i>buffer</i> e estatística descritiva	Modelagem inicial para escolha da distância de <i>buffer</i> e estatística descritiva das variáveis utilizadas no modelo final
4.5	5.3 Dependência espacial nas variáveis independentes	Seleção das variáveis independentes para aplicação da dependência espacial no modelo final
4.6 a 4.9	5.4 Parâmetros e ajustes dos modelos	Valores estimados dos parâmetros de ambas as regressões. Significância estatística das estimativas. Comparação do ajuste dos modelos
	5.5 Efeitos marginais	Análise dos sinais dos parâmetros e comparação com a literatura. Análise dos efeitos marginais.

5. ANÁLISE DOS RESULTADOS

5.1. MATRIZ DE PROXIMIDADE

Utilizou-se o índice I de Moran, calculado conforme a Equação (3.1), para testar para diferentes distâncias de corte e segundo a definição da Matriz de Proximidade escolhida no método, qual resultaria em maior valor de dependência espacial para a variável dependente (escolha pelo modo a pé).

Os resultados apresentados na Figura 5.1 e na Tabela 5.1, mostram que a dependência espacial na escolha pelo modo a pé diminui à medida que a distância de corte para seleção dos vizinhos aumenta. Logo, a correlação entre as escolhas tende a ser maior a um nível local. Todos os resultados foram significativos ao nível de 99% de confiança.

Tabela 5.1 – Teste de distâncias para a Matriz de Proximidade.

Distância (metros)	I de Moran	Distância (metros)	I de Moran
400	0,116	2.400	0,071
600	0,108	2.600	0,069
800	0,103	2.800	0,066
1.000	0,097	3.000	0,064
1.200	0,093	3.200	0,062
1.400	0,088	3.400	0,060
1.600	0,082	3.600	0,058
1.800	0,078	3.800	0,056
2.000	0,075	4.000	0,054
2.200	0,073		

Esse resultado é coerente ao se considerar que, em um contexto urbano, a escolha pelo modo de transporte é relacionada a escolha dos vizinhos mais próximos e não a escolha de vizinhos mais distantes que pouco interagem ou que possuem características socioeconômicas e do ambiente construído diferenciadas. Portanto, todos os modelos e testes das próximas etapas do método adotaram a matriz de proximidade com corte de distância igual a 400 metros, na qual pontos com distância menor a esse limite recebem o valor 1 e pontos com distância maior recebem o valor 0, ou seja, não são considerados vizinhos.

Por fim, o maior valor encontrado do indicador não é elevado, o que indica que, apesar de existir dependência espacial na escolha pelo modo a pé, essa não é alta. Mais sobre esse aspecto foi tratado nos próximos resultados desse estudo.

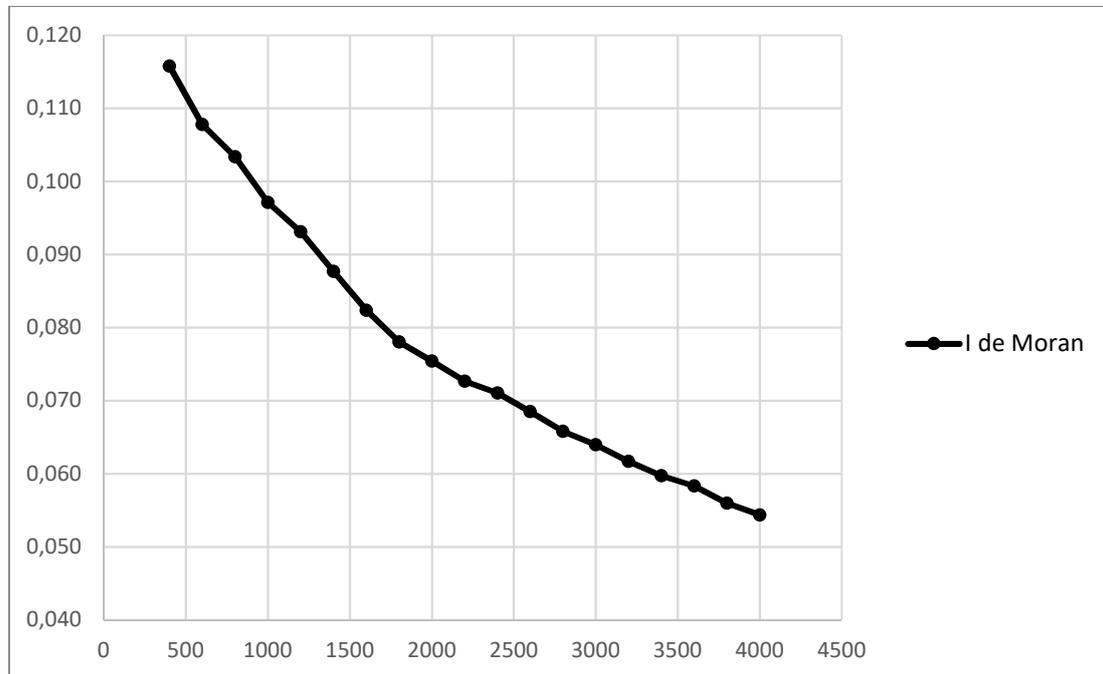


Figura 5.1 – Índice *I* de Moran por distância de corte em metros.

5.2. ESCOLHA DE DISTÂNCIA DE BUFFER E ESTATÍSTICA DESCRITIVA

Definida a distância de corte para a matriz de proximidades, a próxima etapa foi o teste de escolha de distância de *buffer* para levantamento das variáveis do ambiente construído. Foram estimados então quatro modelos Probit do tipo SAR (Seção 3.3.5), ou seja, os modelos consideraram a dependência espacial apenas na variável dependente (escolha pelo modo a pé). Cada modelo considerou todas as variáveis do Quadro 4.1 e as do ambiente construído foram levantadas em ambiente GIS para as respectivas distâncias de *buffer*. A Tabela 5.2 traz os resultados de ajuste dos modelos.

Tabela 5.2 – Teste de distâncias para o *buffer*.

Distância de buffer (metros)	ρ	LIK	LIK _{null}	pseudo-R ²
400	0,131	4.095,09	4.868,27	0,1588
800	0,114	4.090,25	4.868,27	0,1598
1.200	0,113	4.090,51	4.868,27	0,1597
1.600	0,115	4.095,57	4.868,27	0,1587

Para todos os modelos o resultado de ρ (coeficiente autoregressivo espacial) é significativo ao nível de 99% de confiança, logo é possível afirmar que existe dependência espacial na escolha, resultado alinhado ao teste aplicado na Seção anterior. Os modelos apresentaram bons valores de ajuste (pseudo-R²), semelhantes aos encontrados na literatura (LARRAÑAGA *et al.*, 2015; TIAN & EWING, 2017; GUERRA *et al.*, 2018).

Entretanto, os valores encontrados foram praticamente iguais entre si. Esse resultado indica que não há interferência no ajuste do modelo ao se aumentar ou diminuir a distância de *buffer* para levantamento das variáveis. Logo, ainda é incerto qual distância deve ser a adotada. Essa mesma conclusão foi obtida por Tian & Ewing (2017), que testaram distâncias de 400, 800 e 1.600 metros. Uma possível explicação é que 1.600 metros, maior distância adotada, não é suficiente para ocorrer alteração significativa das características do meio urbano. Portanto, a progressão do tamanho de *buffer* não altera significativamente o resultado, já que ela está ocorrendo dentro de uma mesma região homogênea.

Para o prosseguimento do trabalho foi considerada então distância igual a 400 metros. Essa abordagem busca reduzir a sobreposição entre os *buffers* de cada ponto da amostra, visto que quanto maior a sobreposição das áreas de levantamento, maior será a dependência espacial entre os valores das variáveis independentes. Dessa forma, a Tabela 5.3 traz a estatística descritiva das variáveis utilizadas nos modelos das próximas etapas.

Tabela 5.3 – Estatística descritiva das variáveis (*buffer* de 400 metros e $n = 7.329$).

Tipo	Variável	Unidade	Mínimo	Máximo	Média/Proporção	Desvio Padrão
Dependente	Escolha pelo modo a pé	-			42% (viagens a pé)	
	Idade	Anos	18	93	40	14
Controle	Gen	-			46% (homens)	
	Crian	Pessoas	0	9	0,68	0,93
	Disp_veic	Veic. por morador	0	15	0,43	0,45
	Ren	R\$	0	46.000	4.006	3.884
	Esc	Anos	0	16	11	4
Densidade	Den_pop_o	Pess./km ²	0	116.748	14.188	8.621
	Den_pop_d		0	94.433	12.312	7.670
	Den_emp_o	Emp./km ²	0	154.592	8.774	11.541
	Den_emp_d		0	205.102	14.495	22.309
Diversidade	Ent_o	-	0	1	0,47	0,21
	Ent_d	-	0	0,99	0,51	0,20

Tipo	Variável	Unidade	Mínimo	Máximo	Média/Proporção	Desvio Padrão
Desenho	Den_via_o	Km/km ²	8	46	20	4
	Den_via_d		7	45	21	5
	Den_int_o	Int./km ²	19	566	146	62
	Den_int_d		20	590	161	74
	Den_saíd_o	Vias sem saída/km ²	0	301	16	16
	Den_saíd_d		0	316	16	16
Disponibilidade de destinos	Dist_CE_o	Km	0	30	3	4
	Dist_CE_d		0	29	3	4
Distância para o transporte público	Den_tp_o	Pont./km ²	0	230	24	17
	Den_tp_d		0	240	30	21

5.3. DEPENDÊNCIA ESPACIAL NAS VARIÁVEIS INDEPENDENTES

A próxima etapa consistiu no teste de existência de dependência espacial nas variáveis independentes da Tabela 5.3. Como colocado, o objetivo foi avaliar a inclusão ou não dos efeitos espaciais dessas variáveis nos modelos Probit. O teste foi realizado com o indicador *I* de Moran e seus resultados são apresentados na Tabela 5.4.

Com exceção do Gênero, todos os resultados foram significativos ao nível de 99% de confiança. O resultado encontrado para o Gênero é adequado, já que não existe razão para existir agrupamento de pessoas do mesmo gênero em regiões da cidade. Ademais, os outros valores indicam que, apesar de existir certa dependência espacial nas variáveis de controle, essa não é elevada. Dentre essas variáveis, os maiores valores foram obtidos, em ordem decrescente, para a Renda, a Escolaridade e a Disponibilidade de veículos. Esse resultado é esperado, visto que, no meio urbano, pessoas com renda semelhantes tendem a se agrupar em bairros considerados condizentes com a sua situação social. Já que a escolaridade e a disponibilidade de veículos podem ser diretamente associadas à renda, essas duas variáveis também apresentaram certa dependência espacial.

Além disso, todas as variáveis do ambiente construído, tanto em relação à origem, quanto em relação ao destino, possuem *I* de Moran acima de 0,5, o que indica dependência espacial considerável entre as observações. Merece destaque o valor obtido para a Distância para o centro, que é praticamente igual a 1. Esse resultado tem base na concepção da variável, que consiste no levantamento de uma distância em relação a uma área fixa no espaço para todos os pontos da amostra. Logo, seu valor é muito semelhante para pontos próximos uns dos outros.

Tabela 5.4 – Teste de dependência espacial nas variáveis independentes.

Tipo	Variável		I de Moran
Controle	Idade	Anos	0,028
	Gen*	-	0,001
	Crian	Pessoas	0,073
	Disp_veic	Veíc. por morador	0,175
	Ren	R\$	0,313
	Esc	Anos	0,242
Densidade	Den_pop_o	Pess./km ²	0,824
	Den_pop_d		0,802
	Den_emp_o	Emp./km ²	0,791
	Den_emp_d		0,904
Diversidade	Ent_o	-	0,778
	Ent_d	-	0,755
Desenho	Den_via_o	Km/km ²	0,702
	Den_via_d		0,741
	Den_int_o	Int./km ²	0,685
	Den_int_d		0,757
	Den_said_o	Vias sem saída/km ²	0,702
	Den_said_d		0,669
Disponibilidade de destinos	Dist_CE_o	Km	0,999
	Dist_CE_d		0,999
Distância para o transporte público	Den_tp_o	Pont./km ²	0,564
	Den_tp_d		0,628

*p-valor = 0,83

Pelo método proposto, os resultados indicaram que os efeitos espaciais deveriam ser considerados para todas as variáveis do ambiente construído. A princípio, dois modelos seriam estimados: um Probit não espacial e um Probit espacial do tipo SDM, com efeito espacial na variável dependente e nas variáveis independentes referentes ao ambiente construído.

Entretanto, a magnitude dos indicadores e seu tamanho relativo aos encontrados para a variável dependente (Seção 5.1) indicam que os efeitos espaciais estão muito presentes nas variáveis independentes e em maior dimensão que na variável dependente. Logo, optou-se por aumentar a quantidade de modelos estimados inicialmente prevista, com o objetivo de avaliar como a inclusão ou não da dependência espacial nas variáveis independentes afeta o ajuste dos modelos. Estimou-se quatro modelos no total:

- Probit não espacial;
- Probit do tipo SAR (defasagem espacial na variável dependente);

- Probit do tipo SLX (efeitos espaciais nas variáveis independentes do ambiente construído);
- Probit do tipo SDM (combinação dos dois modelos anteriores).

5.4. PARÂMETROS E AJUSTES DOS MODELOS

A Tabela 5.5 e o Apêndice A trazem os resultados obtidos para as quatro regressões. Na tabela são indicados quais os coeficientes possuem significância estatística ao nível de 95% de confiança. Conclui-se pelos valores obtidos de pseudo-R² que a inclusão da defasagem espacial na variável dependente (SAR e SDM) não acrescenta ajuste à regressão, pois os modelos com maior ajuste foram, em ordem decrescente, SLX, não espacial, SDM e SAR. Esse aspecto pode estar baseado na baixa dependência espacial verificada nos testes da Seção 5.1. Logo, mesmo que o coeficiente autoregressivo (ρ) seja estatisticamente significativo, os modelos sem os efeitos espaciais na variável dependente possuem melhor ajuste.

Outra possível conclusão dos resultados dos ajustes diz respeito à origem da correlação espacial encontrada na Seção 5.1. Ela tanto pode indicar que as escolhas dos vizinhos afetam a escolha do indivíduo, quanto que os mesmos fatores que afetam a escolha dos vizinhos afetam a escolha do indivíduo, sejam qualidades do ambiente construído, sejam quaisquer outras características atitudinais ou socioeconômicas. Portanto, os resultados indicam que a correlação está associada à segunda opção, ou seja, que a correlação tem origem nas variáveis explicativas dos vizinhos e não na influência entre as escolhas.

Para prosseguimento da análise e discussão sobre os resultados obtidos nas regressões, foram considerados apenas o Probit não espacial e o espacial com melhor ajuste, ou seja, o modelo SLX. Além disso, não é necessária a análise de coeficientes sem significância estatística, logo somente os parâmetros significantes ao nível de 95% foram considerados. Os resultados estão na Tabela 5.5.

Tabela 5.5 – Coeficientes (β) e ajuste dos modelos Probit.

	Resultados	Unidade	Não espacial	SAR	SLX	SDM	
Intercepto			7,78E-01*	9,05E-01*	6,33E-01*	7,76E-01*	
Controle	Idade	Anos	-5,15E-03*	-5,12E-03*	-5,27E-03*	-5,24E-03*	
	Gen	-	-2,72E-01*	-2,71E-01*	-2,76E-01*	-2,75E-01*	
	Crian	Pessoas	6,26E-02*	5,90E-02*	6,26E-02*	5,95E-02*	
	Disp_veic	Veíc. por morador	-8,06E-01*	-7,95E-01*	-7,99E-01*	-7,91E-01*	
	Ren	R\$	-4,44E-05*	-4,17E-05*	-4,33E-05*	-4,12E-05*	
	Esc	Anos	-5,48E-02*	-5,36E-02*	-5,53E-02*	-5,43E-02*	
Densidade	Den_pop_o	Pess./km ²	2,22E-06	8,68E-07	3,89E-06	4,43E-06	
	Den_pop_d		1,06E-05*	1,04E-05*	1,80E-05*	1,81E-05*	
	Den_emp_o	Emp./km ²	8,23E-06*	8,15E-06*	-2,07E-06	-2,67E-06	
	Den_emp_d		-2,29E-06*	-2,28E-06*	-1,64E-06	-1,83E-06	
Diversidade	Ent_o	-	4,24E-01*	3,46E-01*	-3,15E-01	-3,02E-01	
	Ent_d	-	7,05E-02	6,15E-02	-1,48E-01	-1,50E-01	
Desenho	Den_via_o	Km/km ²	-8,40E-03	-9,37E-03	2,43E-03	2,13E-03	
	Den_via_d		7,36E-03	6,03E-03	-1,41E-02	-1,42E-02	
	Den_int_o	Int./km ²	-2,24E-05	-3,67E-06	-5,34E-04	-4,70E-04	
	Den_int_d		-4,29E-04	-3,73E-04	-2,16E-04	-1,75E-04	
	Den_said_o	Vias sem saída/km ²	-1,12E-03	-1,08E-03	-3,35E-03	-3,43E-03	
	Den_said_d		-2,69E-04	-3,80E-04	1,59E-03	1,43E-03	
Disponibilidade de destinos	Dist_CE_o	Km	-9,33E-02*	-9,78E-02*	-5,05E-01*	-5,05E-01*	
	Dist_CE_d		1,14E-01*	1,14E-01*	3,88E-01	3,82E-01	
Distância para o transporte público	Den_tp_o	Pont./km ²	1,43E-03	1,06E-03	9,91E-04	9,20E-04	
	Den_tp_d		-1,58E-03	-1,58E-03	8,79E-04	9,70E-04	
Efeitos Espaciais	ρ			1,31E-01*		1,15E-01*	
	Den_pop_o (γ)	Pess./km ²			-2,69E-06	-4,73E-06	
	Den_pop_d (γ)				-8,89E-06	-9,10E-06	
	Den_emp_o (γ)	Emp./km ²			1,46E-05*	1,51E-05*	
	Den_emp_d (γ)				-1,34E-06	-1,09E-06	
	Ent_o (γ)	-			9,32E-01*	7,99E-01*	
	Ent_d (γ)				1,89E-01	1,91E-01	
	Den_via_o (γ)	Km/km ²			-1,59E-02	-1,59E-02	
	Den_via_d (γ)				3,08E-02	2,96E-02	
	Den_int_o (γ)	Int./km ²			5,63E-04	4,82E-04	
	Den_int_d (γ)				-3,99E-04	-3,99E-04	
	Den_said_o (γ)	Vias sem saída/km ²			3,42E-03	3,47E-03	
	Den_said_d (γ)				-2,73E-03	-2,62E-03	
	Dist_CE_o (γ)	Km			4,19E-01	4,15E-01	
	Dist_CE_d (γ)				-2,77E-01	-2,71E-01	
	Den_tp_o (γ)	Pont./km ²			1,24E-03	8,49E-04	
	Den_tp_d (γ)				-3,52E-03	-3,64E-03	
	Ajuste	LIK		-4104	4095	-4082	4076
		LIK _{null}		-4983	4868	-4983	4868
		Pseudo-R²		0,1763	0,1588	0,1806	0,1627

*p-valor < 0,05

5.5. EFEITOS MARGINAIS

A Tabela 5.6 traz os resultados dos efeitos marginais médios das variáveis com resultados estatisticamente significativos e com efeito marginal diferente de zero com arredondamento em duas casa decimais, o que engloba quase todas as variáveis de controle de ambos os modelos, que tiveram resultados muito semelhantes. O sinal negativo dos efeitos marginais da Idade e da Escolaridade (Esc) é o esperado de acordo com a literatura. Ademais, os resultados indicam que homens (Gen), em média, tendem a escolher menos a caminhada como meio de transporte.

Dois resultados das variáveis de controle merecem destaque. O efeito marginal da Renda (Ren), que possui coeficiente estatisticamente significante, é praticamente igual a zero, o que não condiz com o encontrado na literatura (DING *et al.*, 2017; TIAN & EWING, 2017; GUERRA *et al.*, 2018), que indica associação negativa entre a realização de viagens a pé e a renda. O segundo resultado diz respeito à Disponibilidade de veículos (Disp_veic), que possui efeito marginal médio com o sinal esperado, ou seja, quanto mais veículos disponíveis, menor a probabilidade de escolher o modo a pé. O destaque está na dimensão desse efeito, que no geral é muito maior que a maioria dos outros efeitos marginais. Esse aspecto mostra a importância que a presença de veículos na residência tem para a escolha do modo.

Tabela 5.6 – Efeitos marginais médios.

Resultados			Probit Não espacial	SLX
Controle	Idade	Anos	-0,16%	-0,17%
	Gen	-	-8,65%	-8,74%
	Crian	Pessoas	1,99%	1,98%
	Disp_veic	Veíc. por morador	-25,69%	-25,30%
	Esc	Anos	-1,75%	-1,75%
Diversidade	Ent_o	-	13,50%	
Disponibilidade de destinos	Dist_CE_o	Km	-2,97%	-15,99%
	Dist_CE_d		3,64%	
Distância para o transporte público	Den_tp_d	Pont./km ²	-0,05%	
	Ent_o (γ)	-		29,53%

Além disso, poucas variáveis do ambiente construído obtiveram resultados estatisticamente significativos e, dentre as que são significativas, várias possuem efeito marginal próximo de zero. Basicamente, para São Paulo, apenas medidas de diversidade e disponibilidade de destinos

foram significativas e exercem influência considerável na escolha pelo modo a pé. Nem mesmo a densidade populacional, medida com resultados mais consolidados na literatura, obteve efeito marginal médio significativo. Vale ressaltar que o significado da não significância dos coeficientes é que não é possível rejeitar a hipótese nula de que ele é igual a zero, o que significa que a variável explicativa não tem influência no fenômeno estudado.

A Distância para o centro na origem (Dist_CE_o) obteve resultados com sinal negativo em ambos os modelos, o que era esperado (SUN *et al.*, 2017), já que o centro é referência para a concentração de comércio e serviços. Logo, quanto maior a distância dessa área, menor a probabilidade de escolher a caminhada como meio de transporte, resultado esse já presente na literatura. Entretanto, o modelo espacial, que possui melhor ajuste, indica efeito marginal cinco vezes maior que o modelo não espacial.

Além disso, o modelo não espacial indicou efeito marginal positivo para a Distância para o centro no destino (Dist_CE_d). Esse resultado vem da característica da amostra de que, quanto maior a distância do destino da viagem ao centro expandido, menor, em média, é a distância total do deslocamento, como mostrado na Tabela 5.7. Essa conclusão, apesar de, a princípio parecer destoante com o esperado, mostra que destinos distantes do centro expandido tendem a ser escolhidos apenas quando a origem é próxima, o que reforça que a ideia de atração que o centro e sua disponibilidade de comércio e serviços exerce.

Tabela 5.7 – Relação entre distância para o centro do destino e distância de viagem.

Intervalo de distância do destino para o centro (quilômetros)	Quantidade de observações da amostra	Média das distâncias de viagem (metros)
0 - 5	5.660	1.620
5-10	1.021	1.325
10-15	440	1.135
15-20	179	1070
20-25	26	853
25-30	3	212
Total / Média	7.329	1.533

O resultado mais expressivo nas variáveis do ambiente construído foi o da Entropia na origem (Ent_o), dimensão da diversidade de uso do solo. Em ambos os modelos sua relação com a escolha pelo modo a pé foi positiva, mas houve diferença entre os dois. No modelo espacial, o resultado significativo ocorreu não na Entropia referente ao próprio ponto de origem da viagem,

mas na ponderação da Entropia de seus vizinhos (γ). Ademais, o efeito marginal médio é maior que o da própria Disponibilidade de veículos, o que pode indicar que o efeito espacial das Entropias dos vizinhos influencia mais na decisão que a posse de veículos.

5.6. CONSIDERAÇÕES FINAIS

Os resultados apresentados neste capítulo buscaram atender não só às perguntas e aos objetivos específicos e geral colocados no Capítulo 1, mas também a todas as questões que foram levantadas na consolidação da revisão da literatura e no método utilizado neste trabalho.

Para desenvolvimento do estudo houve a necessidade de delimitar a distância de corte ideal dos vizinhos na Matriz de Proximidade. Os testes aplicados na Seção 5.1 mostraram que, quanto menor a distância de corte dos vizinhos, maior a dependência espacial na escolha. Logo, essa dependência ocorre em um nível local. De toda forma, ela não possui valor elevado, pois o indicador I de Moran foi igual a 0,116. Além disso, as regressões da Seção 5.4 mostraram que a consideração dos efeitos espaciais das características do ambiente construído dos vizinhos (SLX e SDM) aumenta o ajuste do modelo. Logo, é provável que a origem da dependência espacial entre as escolhas seja essas características e não a influência das escolhas entre si.

Além disso, surgiu a questão de qual seria a distância ideal de *buffer* para levantamento das variáveis do ambiente construído. Os testes da Seção 5.2 mostraram que não há aumento de ajuste quando é alterada a distância de *buffer*, resultado semelhante ao obtido por Tian & Ewing (2017). Logo, não foi obtida conclusão clara para esse questionamento. Esse resultado pode ser explicado pela escala da heterogeneidade do ambiente urbano. A maior distância de *buffer* utilizada de 1.600 metros, não é suficiente, em média, para ocorrerem alterações significativas das características urbanas. Ademais, a consideração de distâncias maiores provavelmente não é adequada para o problema desta pesquisa, pois o modo a pé não é viável para distâncias muito elevadas.

A análise da literatura trouxe a discussão de qual características são mais importantes para a escolha do modo a pé, as da origem ou as do destino. A princípio, pelos resultados da Seção 5.5, as características da origem são mais relevantes que as do destino, mesma conclusão obtida por Sun *et al.* (2017). Entretanto, certo cuidado é necessário nessa conclusão, pois poucas variáveis do ambiente construído tiveram resultados estatisticamente significativos. Logo, é

incerto se, de maneira geral, quais características são mais relevantes: as da origem ou do destino.

A literatura sobre a Econometria Espacial indica que dados espaciais tendem a possuir dependência entre si. Logo, essa possibilidade foi testada. A análise da Seção 5.3 mostrou que, nas variáveis de controle, a dependência é inexistente para o Gênero e reduzida para o restante. A Renda foi a variável com maior valor para o índice I de Moran, o que faz sentido, visto que famílias com perfis semelhantes de renda tendem a morar nos mesmos bairros. Já as variáveis do ambiente construído possuem alta dependência espacial entre suas observações, o que comprova a suposição de que as características urbanísticas de determinada área estão relacionadas às características das áreas vizinhas.

Dentre os objetivos desta pesquisa, está a investigação se a inclusão de efeitos espaciais melhora o ajuste das regressões utilizadas para analisar a escolha pelo modo a pé. Os resultados da Seção 5.4 indicaram que a consideração dos efeitos espaciais das variáveis explicativas do ambiente construído melhora o ajuste da regressão, ao passo que a inclusão da dependência espacial na variável dependente (escolha pelo modo a pé) o reduz. Entretanto, é importante destacar que esse aumento não é elevado, o que gera um questionamento quanto à real necessidade de consideração dos efeitos espaciais, principalmente em estudos técnicos de planejamento urbano. De toda forma, é possível concluir que, mesmo que o resultado de determinada modelagem mostre que os efeitos espaciais são praticamente irrelevantes, esse teste deve ser realizado em estudos com características espaciais para se chegar a essa conclusão, da mesma forma que outros aspectos, como a multicolinearidade e a aderência à distribuição normal, são comumente analisados na modelagem estatística.

Finalmente, a principal pergunta de pesquisa deste trabalho é como as características do ambiente construído influenciam a escolha pelo modo a pé. Considerando os efeitos marginais médios da Seção 5.5, apenas duas medidas apresentaram resultados significantes: a Distância para o centro e a Entropia. A primeira teve o sinal esperado e efeito não tão elevado no modelo não espacial (cerca de 3%) e mais elevado no modelo espacial (cerca de 16%), o que significa que uma redução de uma unidade (sinal negativo) na distância para o centro aumenta em 16% a chance de escolha pelo modo a pé. Logo, fica evidente que a disponibilidade de destinos é importante para a escolha.

Entretanto, o maior efeito foi o positivo encontrado para a Entropia na origem, semelhante ao encontrado na literatura para países desenvolvidos (BOULANGE *et al.*, 2017; DING *et al.*, 2017; TIAN & EWING, 2017), sendo os efeitos espaciais da Entropia dos vizinhos a variável relevante no modelo espacial. Nesse modelo, o efeito marginal é inclusive maior que o da Disponibilidade de veículos, variável de controle importante no processo de escolha.

Por fim, a relevância de ambas as dimensões do ambiente construído, diversidade e disponibilidade de destinos, aponta para a importância, na cidade de São Paulo, de se aproximar diferentes possibilidades de destinos das residências, o que aumenta a chance de escolha pela caminhada como meio de transporte principal.

6. CONCLUSÕES

Esta pesquisa buscou, com base na diversa literatura existente, investigar a relação entre o ambiente construído e a escolha pelo modo a pé no contexto urbano de uma cidade brasileira. Procurou-se também inovar na técnica de pesquisa pela utilização de modelos espaciais de regressão. Devido a questões de disponibilidade e formato de dados, a cidade de São Paulo foi a escolhida como exemplo para aplicação do método proposto.

Os resultados mostraram que, dentre os 5D's propostos na literatura, a disponibilidade de destinos e a diversidade são relevantes para incentivar a utilização desse modo em São Paulo. Logo, a elaboração de políticas públicas deve incentivar o uso misto do solo como base para a promoção da caminhada como meio de transporte. Ademais, os modelos espaciais se mostraram como uma nova ferramenta para análises de planejamento urbano e de transportes e foi possível concluir que os efeitos espaciais devem ser ao menos testados antes de estudos desse tipo.

6.1. LIMITAÇÕES DO TRABALHO

É importante destacar que as conclusões são baseadas na análise de dados de apenas uma cidade brasileira. Portanto, é necessário que planejadores realizem estudos semelhantes para os diferentes contextos urbanos com o objetivo de avaliar como os padrões de transporte desses se comportam. Além disso, um fenômeno apontado na literatura não considerado é o viés de seleção (*self-selection*). Segundo Sun *et al.* (2017), a escolha pelo local de residência está muito associado às características do ambiente construído. Logo, indivíduos que preferem a caminhada como meio de transporte tendem a viver em regiões propensas para isso, ou seja, que possuem as características presentes na literatura como incentivadoras do modo a pé.

Já em relação ao método adotado, algumas limitações são pertinentes de serem colocadas. Como citado no Capítulo 3, a escolha da matriz de proximidade é um processo muito baseado no conhecimento do pesquisador sobre o objeto da pesquisa, pois não existem métodos objetivos para a sua especificação. Além disso, os modelos espaciais e indicadores utilizados neste trabalho são globais e não lidam com a heterocedasticidade das relações espaciais no contexto urbano. Ademais, este estudo, assim como a maioria dos estudos que aplicam a regressão estatística, tem caráter puramente transversal. Logo, formalmente, as conclusões implicam apenas em relações de correlação e não de causalidade.

Por fim, apesar das limitações citadas, este estudo contribui ao aplicar técnica de regressão espacial para a análise de fenômeno ainda pouco estudado no contexto urbano de países não desenvolvidos. Espera-se então que os resultados encontrados promovam novas investigações e inspirem os planejadores no desenvolvimento de estudos análogos e na elaboração de políticas públicas.

6.2. SUGESTÕES PARA TRABALHOS FUTUROS

A partir das limitações citadas e dos resultados obtidos surgem novas possibilidades de investigação. Dois questionamentos propostos não tiveram conclusões definitivas: a distância adequada de *buffer* e a importância relativa entre as características da origem e do destino da viagem. Sugere-se então, para a primeira questão, que sejam avaliadas mais distâncias de forma separada para diferentes variáveis, visando a uma melhor análise do efeito da área de abrangência nos resultados. Para a segunda, uma abordagem possível seria a modelagem de modelos exclusivos para características da origem e do destino e posterior comparação entre os resultados e ajustes.

Segundo Cao *et al.* (2009), existem nove diferentes abordagens para tratar do problema do viés de seleção, que vão desde a consideração de variáveis de controle relativas a medidas atitudinais até a modelagem conjunta entre escolha residencial e escolha do modo. Portanto, a utilização de modelos que o considerem pode ser um novo caminho de pesquisa nesse tema.

Além disso, cabe a investigação sobre a utilização de outras formas de Matriz de Proximidade, como o inverso da distância ao quadrado. Com relação à heterocedasticidade em estudos espaciais, um exemplo de modelo alternativo para lidar com essa questão é o proposto por Páez (2006), que apresenta um modelo Probit cujo erro segue uma função baseada na Regressão Geograficamente Ponderada (RGP). Quanto à transversalidade do método adotado, a utilização de modelos longitudinais em dados coletados em diferentes momentos pode trazer novas percepções sobre o assunto. Esse tipo de técnica permitiria o estudo da famosa relação sugerida por Jacobs (1961) de que quanto mais pessoas caminham mais há incentivo para outras realizarem essa escolha.

Portanto, fica evidente que ainda existem diversos caminhos de pesquisa nesse assunto que podem ser explorados futuramente. Além das possibilidades já citadas, existe a necessidade de aplicação de análises semelhantes a outros contextos urbanos brasileiros que não o do município de São Paulo. Dessa forma, será possível consolidar o conhecimento já existente ou promover novas percepções sobre as relações aqui exploradas.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, E. (2012) *Econometria Espacial Aplicada*. Alínea, Campinas, SP, 498 p.
- AMÂNCIO, M. A. & SANCHES, S. da P. (2008) A forma urbana e as viagens a pé - estudo de caso em uma cidade brasileira de porte médio. *Acta Scientiarum - Technology*, v. 30, n. 2, p. 147-154.
- ANSELIN, L. (1988) *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Santa Barbara, EUA, 284 p.
- ANSELIN, L. (1995) Local indicators of spatial association - LISA. *Geographical Analysis*, v. 27, n. 2, p. 93-115.
- ANSELIN, L. (2002) Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, v. 27, n. 3, p. 247-267.
- ANSELIN, L., FLORAX, R. J. G. M. & REY, S. J. (2004) *Advances in Spatial Econometrics: Methodology Tools and Applications*. Springer-Verlag Berlin Heidelberg, Nova Iorque, EUA, 513 p.
- ANSELIN, L. (2010) Thirty years of spatial econometrics. *Papers in Regional Science*, v. 89, n. 1, p. 3-25.
- ARRUDA, F. S. (2005) *Aplicação de um modelo baseado em atividades para análise da relação uso do uso e transportes no contexto brasileiro*. Tese de Doutorado, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, SP, 145 p.
- BIRD, E. L., BAKER, G., MUTRIE, N., OGILVIE, D., SAHLQVIST, S. & POWELL, J. (2013) Behavior change techniques used to promote walking and cycling: A systematic review. *Health Psychology*, v. 32, n. 8, p. 829-838.
- BOULANGE, C., GUNN, L., GILES-CORTI, B., MAVOA, S., PETTIT, C. & BADLAND, H. (2017) Examining associations between urban design attributes and transport mode choice for walking, cycling, public transport and private motor vehicle trips. *Journal of Transport and Health*, v. 6, p. 155-166.
- BOUSSAUW, K., NEUTENS, T. & WITLOX, F. (2012) Relationship between Spatial Proximity and Travel-to-Work Distance: The Effect of the Compact City. *Regional Studies*, v. 46, n. 6, p. 687-706.
- CAO, X., MOKHTARIAN, P. & HANDY, S. (2009) Examining the Impacts of Residential Self-Selection on Travel Behaviour: A Focus on Empirical Findings. *Transport Reviews: A Transnational Transdisciplinary Journal*, v. 29, n. 3, p. 359-395.
- CERVERO, R. & DUNCAN, M. (2003) *Walking, Bicycling, and Urban Landscapes: Evidence from the San Francisco Bay Area*. Working Paper, University of California Transportation Center, Berkeley, EUA, 14 p.
- CERVERO, R. & KOCKELMAN, K. (1997) Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, v. 2, n. 3, p. 199-219.

- CERVERO, R., SARMIENTO, O. L., JACOBY, E., GOMEZ, F. & NEIMAN, A. (2009) Influences of Built Environments on Walking and Cycling: Lessons from Bogotá. *International Journal of Sustainable Transportation*, v. 3, p. 203-226.
- CHEN, J., NI, J., XI, C., LI, S. & WANG, J. (2017) Determining intra-urban spatial accessibility disparities in multimodal public transport networks. *Journal of Transport Geography*, v. 65, p. 123-133.
- DING, C., WANG, D., LIU, C., ZHANG, Y. & YANG, J. (2017) Exploring the influence of built environment on travel mode choice considering the mediating effects of car ownership and travel distance. *Transportation Research Part A: Policy and Practice*, v. 100, p. 65-80.
- DRUCK, S., CARVALHO, M. S., CÂMARA, G., MONTEIRO, A. M. V., CAMARGO, E. C. G., FELGUEIRAS, C. A., CRUZ, O. G. & CORREA, V. (2004) Análise Espacial de Dados Geográficos. Embrapa Cerrados, Brasília, DF, 209 p.
- EVENSON, K. R., SATINSKY, S. B., RODRÍGUEZ, D. A. & AYTUR, S. A. (2012) Exploring a public health perspective on pedestrian planning. *Health Promotion Practice*, v. 13, n. 2, p. 204-213.
- EWING, R., & CERVERO, R. (2010) Travel and the built environment. *Journal of the American Planning Association*, v. 76, n. 3, p. 265-294.
- FÁVERO, L. P. & BELFIORE, P. (2017) Manual de Análise de Dados: Estatística e Modelagem Multivariada com Excel, SPSS e Stata. Elsevier, Rio de Janeiro, RJ, 1187p.
- FEUILLET, T., SALZE, P., CHARREIRE, H., MENAI, M., ENAUX, C., PERCHOUX, C., HESS, F., KESSE-GUYOT, E., HERCBERG, S., SIMON, C., WEBER, C. & OPPERT, J. M. (2016) Built environment in local relation with walking: Why here and not there? *Journal of Transport and Health*, v. 3, 500-512.
- FORSYTH, A., VAN RIPER, D., LARSON, N., WALL, M. & NEUMARK-SZTAINER, D. (2012) Creating a replicable, valid cross-platform buffering technique: The sausage network buffer for measuring food and physical activity built environments. *International Journal of Health Geographics*, v. 11, p. 1-9.
- GEARY, R. C. (1954) The Contiguity Ratio and Statistical Mapping. *The Incorporated Statistician*, v. 5, n. 3, p. 115-145.
- GOETZKE, F. (2008) Network effects in public transit use: Evidence from a spatially autoregressive mode choice model for New York. *Urban Studies*, v. 45, n. 2, p. 407-417.
- GOLGHER, A. B. (2015) Introdução à Econometria Espacial. Paco Editorial, Jundiaí, SP, 348 p.
- GUERRA, E., CAUDILLO, C., MONKKONEN, P. & MONTEJANO, J. (2018) Urban form, transit supply, and travel behavior in Latin America: Evidence from Mexico's 100 largest urban areas. *Transport Policy*, v. 69, p. 98-105.

- HAHN, E. D., SOYER, R. (2005) *Probit and Logit models: differences in the multivariate realm*. Working Paper, EUA.
- HANDY, S. (1996) Methodologies for Exploring the Link Between Urban Form and Travel Behavior. *Transportation Research Part D: Transport and Environment*, v. 1, n. 2, p. 151-165.
- HEARST, M. O., SIRARD, J. R., FORSYTH, A., PARKER, E. D., KLEIN, E. G., GREEN, C. G. & LYTLE, L. A. (2013) The relationship of area-level sociodemographic characteristics, household composition and individual-level socioeconomic status on walking behavior among adults. *Transportation Research Part A: Policy and Practice*, v. 50, p. 149-157.
- IBGE (2015) Instituto Brasileiro de Geografia e Estatística. *Práticas de esporte e atividade física*. Rio de Janeiro, RJ, 81 p.
- JACOBS, J. (1961) *The Death and Life of Great American Cities*. Random House, Nova Iorque, EUA, 458 p.
- KAMRUZZAMAN, M., WASHINGTON, S., BAKER, D., BROWN, W., GILES-CORTI, B. & TURRELL, G. (2016) Built environment impacts on walking for transport in Brisbane, Australia. *Transportation*, v. 43, n. 1, p. 53-77.
- KOH, P. P. & WONG, Y. D. (2013) Comparing pedestrians' needs and behaviours in different land use environments. *Journal of Transport Geography*, v. 26, p. 43-50.
- KOOHSARI, M. J., SUGIYAMA, T., MAVOA, S., VILLANUEVA, K., BADLAND, H., GILES-CORTI, B. & OWEN, N. (2016) Street network measures and adults' walking for transport: Application of space syntax. *Health and Place*, v. 38, n. 89-95.
- KREMPI, A. P. (2004) *Explorando Recursos De Estatística Espacial Para Análise Da Acessibilidade Da Cidade De Bauru*. Dissertação de Mestrado. Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, SP, 82 p.
- LAMÍQUIZ, P. J. & LÓPEZ-DOMÍNGUEZ, J. (2015) Effects of built environment on walking at the neighbourhood scale. A new role for street networks by modelling their configurational accessibility? *Transportation Research Part A: Policy and Practice*, v. 74, p. 148-163.
- LARRAÑAGA, A. M., CYBIS, H. B. B. & TORRES, T. B. (2015) Influência da Estrutura Urbana na Decisão de Realizar Viagens a pé em Porto Alegre. *Transportes*, v. 23, n. 4, p. 89-97.
- LEE, S., SUNG, H. & WOO, A. (2017) The Spatial Variations of Relationship between Built Environment and Pedestrian Volume: Focused on the 2009 Seoul Pedestrian Flow Survey in Korea. *Journal of Asian Architecture and Building Engineering*, v. 16, n. 1, 147-154.
- LESAGE, J. P. (1999) *The Theory and Practice of Spatial Econometrics*. Department of Economics, University of Toledo, EUA, 296 p.

- LESAGE, J. P. & PACE, R. K. (2009) *Introduction to Spatial Econometrics*. Chapman & Hall, EUA, 321 p.
- LINDELÖW, D., SVENSSON, Å., BRUNDELL-FREIJ, K. & WINSLOTT HISELIUS, L. (2017) Satisfaction or compensation? The interaction between walking preferences and neighbourhood design. *Transportation Research Part D: Transport and Environment*, v. 50, p. 520-532.
- LITMAN, T. (2018) *Land Use Impacts on Transport: How Land Use Factors Affect Travel Behavior*. Victoria Transport Policy Institute, Austrália, 88 p.
- LOPES, S. B. (2005) *Efeitos da dependência espacial em modelos de previsão de demanda por transporte*. Dissertação de Mestrado, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, SP, 137 p.
- LOPES, S. B. (2010) *Uma ferramenta para planejamento da mobilidade sustentável com base em modelo de uso do solo e transportes*. Tese de Doutorado, Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, SP, 184 p.
- LU, Y., XIAO, Y. & YE, Y. (2017) Urban density, diversity and design: Is more always better for walking? A study from Hong Kong. *Preventive Medicine*, n. 103, p. S99-S103.
- MARQUET, O. & MIRALLES-GUASCH, C. (2014) Walking short distances. The socioeconomic drivers for the use of proximity in everyday mobility in Barcelona. *Transportation Research Part A: Policy and Practice*, v. 70, p. 210-222.
- MCKIBBIN, M. (2011) The influence of the built environment on mode choice - evidence from the journey to work in Sydney. *Proceedings of the Australasian Transport Research Forum*, Adelaide, Australia.
- MEYER, M. D. (2016) *Transportation Planning Handbook*. John Wiley & Sons, EUA, 4ª Edição, 1203 p.
- MILNE, D. & WATLING, D. (2018) Big data and understanding change in the context of planning transport systems. *Journal of Transport Geography*. No prelo.
- MORAN, P. A. P. (1950) Notes on Continuous Stochastic Phenomena. *Biometrika*, v. 37, n. 1, p. 17-23.
- MUMFORD, E. (2018) *Designing the Modern City: Urbanism since 1850*. Yale University Press, Londres, Reino Unido, 499 p.
- OLIVER, L. N., SCHUURMAN, N. & HALL, A. W. (2007) Comparing circular and network buffers to examine the influence of land use on walking for leisure and errands. *International Journal of Health Geographics*, v. 6, p. 1-11.
- ORTÚZAR, J. de D., WILLUMSEN, L. G. (2011) *Modelling Transport*. John Wiley & Sons, Reino Unido, 4ª Edição, 586 p.
- PÁEZ, A. (2006) Exploring contextual variations in land use and transport analysis using a probit model with geographical weights. *Journal of Transport Geography*, v. 14, n. 3, 167-176.

- PÁEZ, A. & SCOTT, D. M. (2004) Spatial statistics for urban analysis: A review of techniques with examples. *GeoJournal*, v. 61, n. 1, 53-67.
- QUEIROZ, M. P. (2003) *Análise Espacial Dos Acidentes De Trânsito Do Município De Fortaleza*. Dissertação de Mestrado, Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, CE, 124 p.
- REIS, R. S., HINO, A. A. F., RECH, C. R., KERR, J. & HALLAL, P. C. (2013) Walkability and physical activity: Findings from Curitiba, Brazil. *American Journal of Preventive Medicine*, v. 45, n. 3, p. 269-275.
- RIBEIRO, V. DA C. (2012) *Análise de demanda por transportes de passageiros via modelos de regressão georreferenciados*. Dissertação de Mestrado, Centro Tecnológico, Universidade Federal do Espírito Santo, Vitória, ES, 81 p.
- SANNI, T. & ABRANTES, P. A. L. (2010) Estimating walking modal share: a novel approach based on spatial regression models and GIS. *Journal of Maps*, v. 6, n. 1, 192-198.
- SILVA, A. R. (2006) *Avaliação de modelos de regressão espacial para análise de cenários do transporte rodoviário de carga*. Dissertação de Mestrado, Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 125 p.
- SMIRNOV, O. A. (2010) Modeling spatial discrete choice. *Regional Science and Urban Economics*, v. 40, n. 5, p. 292-298.
- SOLTANI, A., & ASKARI, S. (2017) Exploring spatial autocorrelation of traffic crashes based on severity. *Injury*, v. 48, n. 3, p. 637-647.
- SOUZA FILHO, L. A. (2013) *Análise espacial em modelos de geração de viagens*. Dissertação de Mestrado, Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, 90 p.
- SUN, B., ERMAGUN, A. & DAN, B. (2017) Built environmental impacts on commuting mode choice and distance: Evidence from Shanghai. *Transportation Research Part D: Transport and Environment*, v. 52, p. 441-453.
- TARGA, F. & CLIFTON, K. J. (2004) Built environment and trip generation for non-motorized travel. *Proceedings of the National Household Travel Survey Conference: Data for Understanding Our Nation's Travel*. Washington D.C., EUA.
- TEIXEIRA, G. L. (2003) *Uso de dados censitários para identificação de zonas homogêneas para planejamento de transportes utilizando estatística espacial*. Dissertação de Mestrado, Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 156 p.
- TIAN, G. & EWING, R. (2017) A walk trip generation model for Portland, OR. *Transportation Research Part D: Transport and Environment*, v. 52, p. 340-353.
- TOBLER, W. R. (1970) A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, v. 46, p. 234-240.

- VALE, D. S. & PEREIRA, M. (2016) Influence on pedestrian commuting behavior of the built environment surrounding destinations: A structural equations modeling approach. *International Journal of Sustainable Transportation*, v. 10, n. 8, p. 730-741.
- VALE, D. S., PEREIRA, M. & VIANA, C. M. (2018) Different destination, different commuting pattern? Analyzing the influence of the campus location on commuting. *Journal of Transport and Land Use*, v. 11, n. 1, p. 1-18.
- VEGA, S. H. & ELHORST, J. P. (2015) The SLX model. *Journal of Regional Science*, v. 55, n. 3, p. 339-363.
- WANG, C. H., AKAR, G., & GULDMANN, J. M. (2015) Do your neighbors affect your bicycling choice? A spatial probit model for bicycling to The Ohio State University. *Journal of Transport Geography*, v. 42, p. 122-130.
- WANG, G., MACERA, C. A., SCUDDER-SOUCIE, B., SCHMID, T., PRATT, M. & BUCHNER, D. (2004) Cost effectiveness of a bicycle/pedestrian trail development in health promotion. *Preventive Medicine*, v. 38, n. 2, p. 237-242.
- WEN, L. M. & RISSEL, C. (2008) Inverse associations between cycling to work, public transport, and overweight and obesity: Findings from a population based study in Australia. *Preventive Medicine*, v. 46, p. 29-32.
- YANG, H., LU, X., CHERRY, C., LIU, X. & LI, Y. (2017) Spatial variations in active mode trip volume at intersections: a local analysis utilizing geographically weighted regression. *Journal of Transport Geography*, v. 64, 184-194.

APÊNDICE A – RESULTADOS DOS MODELOS

```

Call:
glm(formula = Y ~ Idade + Gen + Crian + Disp_Veic + Ren + Esc +
     Den_pop_o + Den_emp_o + Ent_o + Den_via_o + Den_int_o + Den_said_o +
     Dist_CE_o + Den_tp_o + Den_pop_d + Den_emp_d + Ent_d + Den_via_d +
     Den_int_d + Den_said_d + Dist_CE_d + Den_tp_d, family = binomial(link = "probit"),
     data = Banco_400)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.2777  -0.9171  -0.4545   0.9786   3.1836

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  7.776e-01  1.653e-01  4.703 2.57e-06 ***
Idade        -5.149e-03  1.216e-03  -4.236 2.28e-05 ***
Gen          -2.716e-01  3.254e-02  -8.346 < 2e-16 ***
Crian         6.255e-02  1.782e-02   3.511 0.000447 ***
Disp_Veic    -8.062e-01  5.232e-02 -15.410 < 2e-16 ***
Ren          -4.442e-05  6.513e-06  -6.821 9.06e-12 ***
Esc          -5.479e-02  4.372e-03 -12.534 < 2e-16 ***
Den_pop_o    2.219e-06  2.205e-06   1.006 0.314210
Den_emp_o     8.231e-06  1.768e-06   4.655 3.25e-06 ***
Ent_o         4.238e-01  9.097e-02   4.658 3.19e-06 ***
Den_via_o    -8.401e-03  8.127e-03  -1.034 0.301241
Den_int_o    -2.242e-05  5.880e-04  -0.038 0.969583
Den_said_o   -1.115e-03  1.034e-03  -1.079 0.280645
Dist_CE_o    -9.332e-02  1.805e-02  -5.171 2.33e-07 ***
Den_tp_o     1.429e-03  1.038e-03   1.377 0.168511
Den_pop_d    1.062e-05  2.374e-06   4.473 7.70e-06 ***
Den_emp_d    -2.292e-06  9.871e-07  -2.322 0.020218 *
Ent_d        7.048e-02  9.204e-02   0.766 0.443828
Den_via_d    7.363e-03  7.552e-03   0.975 0.329621
Den_int_d   -4.292e-04  5.206e-04  -0.824 0.409765
Den_said_d   -2.687e-04  1.094e-03  -0.246 0.805999
Dist_CE_d    1.142e-01  1.810e-02   6.309 2.81e-10 ***
Den_tp_d    -1.576e-03  8.486e-04  -1.857 0.063347 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 9965.1  on 7328  degrees of freedom
Residual deviance: 8207.8  on 7306  degrees of freedom
AIC: 8253.8

Number of Fisher Scoring iterations: 4

> logLik(Est_400)
'log Lik.' -4103.88 (df=23)
> logLik(Est_400_0)
'log Lik.' -4982.545 (df=1)

```

Figura A.1 – Resultados do modelo Probit não espacial.

```

> Est_esp_400 <- SpatialProbitFit(
+ Y ~ Idade+Gen+Criar+Disp_Veic+Ren+Esc+Den_pop_o+Den_emp_o+Ent_o+Den_via_o+Den_int_o+Den_said_o+Dist_CE_o+Den_tp_o
+ +Den_pop_d+Den_emp_d+Ent_d+Den_via_d+Den_int_d+Den_said_d+Dist_CE_d+Den_tp_d, Banco_400, W_orig, DGP='SAR')
St. dev. of beta conditional on rho and Lik-ratio of rho
      Estimate Std. Error   Z-value   Pr(>|z|)
(Intercept)  9.046398e-01 1.564279e-01  5.783111831 7.333128e-09
Idade        -5.117733e-03 1.204292e-03 -4.249577121 2.141745e-05
Gen         -2.708035e-01 3.223291e-02 -8.401460065 0.000000e+00
Criar       5.896767e-02 1.758767e-02  3.352784184 8.000306e-04
Disp_Veic   -7.953573e-01 5.162149e-02 -15.407485875 0.000000e+00
Ren        -4.174735e-05 6.383536e-06 -6.539847774 6.158141e-11
Esc        -5.355739e-02 4.307704e-03 -12.432932009 0.000000e+00
Den_pop_o   8.676174e-07 2.039971e-06  0.425308787 6.706116e-01
Den_emp_o   8.150456e-06 1.640348e-06  4.968734758 6.739118e-07
Ent_o       3.461146e-01 8.425065e-02  4.108154039 3.988341e-05
Den_via_o  -9.370598e-03 7.555127e-03 -1.240296587 2.148657e-01
Den_int_o  -3.665498e-06 5.480597e-04 -0.006688137 9.946637e-01
Den_said_o -1.077367e-03 9.521791e-04 -1.131475045 2.578552e-01
Dist_CE_o  -9.777702e-02 1.755427e-02 -5.569984948 2.547613e-08
Den_tp_o   1.058343e-03 9.682785e-04  1.093014862 2.743873e-01
Den_pop_d   1.041188e-05 2.327711e-06  4.473013543 7.712489e-06
Den_emp_d  -2.275413e-06 9.720520e-07 -2.340834616 1.924069e-02
Ent_d       6.153537e-02 9.019029e-02  0.682283821 4.950595e-01
Den_via_d   6.033725e-03 7.442379e-03  0.810725255 4.175235e-01
Den_int_d  -3.726323e-04 5.136263e-04 -0.725493018 4.681496e-01
Den_said_d -3.801618e-04 1.072282e-03 -0.354535203 7.229378e-01
Dist_CE_d   1.139294e-01 1.767129e-02  6.447143860 1.139775e-10
Den_tp_d   -1.578914e-03 8.367822e-04 -1.886887316 5.917548e-02
lambda      1.307360e-01      NA      17.590271309 2.739864e-05
warning message:
glm.fit: fitted probabilities numerically 0 or 1 occurred
> Est_esp_400@loglik
[1] 4095.085
> Est_esp_400_0@loglik
[1] 4868.274

```

Figura A.2 – Resultados do modelo Probit SAR.

```

call:
glm(formula = Y ~ Idade + Gen + Criar + Disp_Veic + Ren + Esc +
  Den_pop_o + Den_emp_o + Ent_o + Den_via_o + Den_int_o + Den_said_o +
  Dist_CE_o + Den_tp_o + Den_pop_d + Den_emp_d + Ent_d + Den_via_d +
  Den_int_d + Den_said_d + Dist_CE_d + Den_tp_d + wDen_pop_o +
  wDen_emp_o + wEnt_o + wDen_via_o + wDen_int_o + wDen_said_o +
  wDist_CE_o + wDen_tp_o + wDen_pop_d + wDen_emp_d + wEnt_d +
  wDen_via_d + wDen_int_d + wDen_said_d + wDist_CE_d + wDen_tp_d,
  family = binomial(link = "probit"), data = Banco_w2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.1911 -0.9179 -0.4443  0.9746  3.2310

Coefficients:
      Estimate Std. Error z value Pr(>|z|)
(Intercept)  6.332e-01 1.942e-01  3.261 0.001110 **
Idade       -5.265e-03 1.222e-03 -4.309 1.64e-05 ***
Gen        -2.758e-01 3.265e-02 -8.447 < 2e-16 ***
Criar       6.263e-02 1.787e-02  3.504 0.000458 ***
Disp_Veic  -7.985e-01 5.252e-02 -15.204 < 2e-16 ***
Ren        -4.331e-05 6.554e-06 -6.608 3.89e-11 ***
Esc        -5.526e-02 4.392e-03 -12.582 < 2e-16 ***
Den_pop_o   3.889e-06 5.679e-06  0.685 0.493463
Den_emp_o  -2.066e-06 4.074e-06 -0.507 0.612038
Ent_o      -3.147e-01 2.022e-01 -1.556 0.119627
Den_via_o   2.427e-03 1.561e-02  0.155 0.876444
Den_int_o  -5.344e-04 1.059e-03 -0.504 0.613916
Den_said_o -3.346e-03 2.309e-03 -1.449 0.147341
Dist_CE_o  -5.045e-01 2.563e-01 -1.969 0.048999 *
Den_tp_o    9.908e-04 1.827e-03  0.542 0.587717
Den_pop_d   1.801e-05 6.706e-06  2.685 0.007254 **
Den_emp_d  -1.638e-06 2.906e-06 -0.563 0.573105
Ent_d      -1.479e-01 2.102e-01 -0.704 0.481632
Den_via_d  -1.408e-02 1.551e-02 -0.908 0.363988
Den_int_d  -2.164e-04 1.008e-03 -0.215 0.829947
Den_said_d  1.585e-03 2.251e-03  0.704 0.481303
Dist_CE_d   3.882e-01 2.605e-01  1.490 0.136102
Den_tp_d    8.790e-04 1.613e-03  0.545 0.585767
wDen_pop_o  -2.686e-06 6.528e-06 -0.411 0.680746
wDen_emp_o  1.461e-05 5.198e-06  2.811 0.004935 **
wEnt_o      9.318e-01 2.413e-01  3.862 0.000113 ***
wDen_via_o -1.594e-02 2.039e-02 -0.782 0.434379
wDen_int_o  5.634e-04 1.430e-03  0.394 0.693541

```

Figura A.3 – Resultados do modelo Probit SLX (primeira parte).

```

WDen_pop_o -2.686e-06 6.528e-06 -0.411 0.680746
WDen_emp_o 1.461e-05 5.198e-06 2.811 0.004935 **
WEnt_o 9.318e-01 2.413e-01 3.862 0.000113 ***
WDen_via_o -1.594e-02 2.039e-02 -0.782 0.434379
WDen_int_o 5.634e-04 1.430e-03 0.394 0.693541
WDen_said_o 3.419e-03 2.801e-03 1.221 0.222198
WDist_CE_o 4.193e-01 2.566e-01 1.634 0.102340
WDen_tp_o 1.238e-03 2.480e-03 0.499 0.617777
WDen_pop_d -8.887e-06 7.706e-06 -1.153 0.248794
WDen_emp_d -1.344e-06 3.266e-06 -0.412 0.680626
WEnt_d 1.889e-01 2.504e-01 0.754 0.450770
WDen_via_d 3.084e-02 1.937e-02 1.593 0.111272
WDen_int_d -3.991e-04 1.287e-03 -0.310 0.756414
WDen_said_d -2.734e-03 2.800e-03 -0.976 0.328892
WDist_CE_d -2.772e-01 2.608e-01 -1.063 0.287972
WDen_tp_d -3.521e-03 2.152e-03 -1.637 0.101729
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 9965.1 on 7328 degrees of freedom
Residual deviance: 8164.9 on 7290 degrees of freedom
AIC: 8242.9

Number of Fisher Scoring iterations: 5

> logLik(Est_400)
'log Lik.' -4082.471 (df=39)
> logLik(Est_400_0)
'log Lik.' -4982.545 (df=1)

```

Figura A.4 – Resultados do modelo Probit SLX (segunda parte).

```

> Est_400 <- SpatialProbitFit(Y ~ Idade+Gen+Cri+Disp_Veic+Ren+Esc+Den_pop_o+Den_emp_o+Ent_o+Den_via_o+Den_int_o+Den_said_o+Dist_CE_o+Den_tp_o
+ Den_pop_d+Den_emp_d+Ent_d+Den_via_d+Den_int_d+Den_said_d+Dist_CE_d+Den_tp_d+WDen_pop_o+WDen_emp_o+WEnt_o+WDen_via_o+
+ WDen_int_o+WDen_said_o+WDist_CE_o+WDen_tp_o
+ WDen_pop_d+WDen_emp_d+WEnt_d+WDen_via_d+WDen_int_d+WDen_said_d+WDist_CE_d+WDen_tp_d,
+ Banco_w2, W_orig, DGP='SAR')
St. dev. of beta conditional on rho and Lik-ratio of rho
Estimate Std. Error z-value Pr(>|z|)
(Intercept) 7.756557e-01 1.824702e-01 4.2508617 2.129497e-05
Idade -5.238790e-03 1.212198e-03 -4.3217281 1.548119e-05
Gen -2.752786e-01 3.238589e-02 -8.4999547 0.000000e+00
Cri 5.949133e-02 1.767709e-02 3.3654472 7.641968e-04
Disp_Veic -7.912853e-01 5.191933e-02 -15.2406705 0.000000e+00
Ren -4.120291e-05 6.443017e-06 -6.3949712 1.605776e-10
Esc -5.428946e-02 4.336501e-03 -12.5191851 0.000000e+00
Den_pop_o 4.427998e-06 5.690788e-06 0.7780993 4.365105e-01
Den_emp_o -2.667932e-06 4.042462e-06 -0.6599770 5.092686e-01
Ent_o -3.019114e-01 2.030469e-01 -1.4869043 1.370401e-01
Den_via_o 2.133402e-03 1.564859e-02 0.1363319 8.915589e-01
Den_int_o -4.696476e-04 1.062555e-03 -0.4419986 6.584903e-01
Den_said_o -3.431743e-03 2.314799e-03 -1.4825232 1.382011e-01
Dist_CE_o -5.047364e-01 2.529042e-01 -1.9957615 4.595989e-02
Den_tp_o 9.198528e-04 1.831125e-03 0.5023429 6.154263e-01
Den_pop_d 1.812728e-05 6.650320e-06 2.7257754 6.415064e-03
Den_emp_d -1.828342e-06 2.896378e-06 -0.6312513 5.278762e-01
Ent_d -1.495125e-01 2.084632e-01 -0.7172130 4.732427e-01
Den_via_d -1.416750e-02 1.540185e-02 -0.9198574 3.576473e-01
Den_int_d -1.747034e-04 1.001039e-03 -0.1745220 8.614552e-01
Den_said_d 1.432046e-03 2.234751e-03 0.6408079 5.216475e-01
Dist_CE_d 3.821497e-01 2.572481e-01 1.4855296 1.374036e-01
Den_tp_d 9.696176e-04 1.602037e-03 0.6052405 5.450192e-01
WDen_pop_o -4.732171e-06 6.452689e-06 -0.7333641 4.633364e-01
WDen_emp_o 1.506645e-05 5.072254e-06 2.9703650 2.974461e-03
WEnt_o 7.987235e-01 2.381627e-01 3.3536886 7.974205e-04
WDen_via_o -1.586995e-02 1.990503e-02 -0.7972831 4.252866e-01
WDen_int_o 4.818890e-04 1.392706e-03 0.3460092 7.293358e-01
WDen_said_o 3.465626e-03 2.753713e-03 1.2585287 2.082006e-01
WDist_CE_o 4.149201e-01 2.533533e-01 1.6377132 1.014815e-01
WDen_tp_o 8.488101e-04 2.409923e-03 0.3522146 7.246773e-01
WDen_pop_d -9.098964e-06 7.615093e-06 -1.1948592 2.321420e-01
WDen_emp_d -1.089513e-06 3.241071e-06 -0.3361584 7.367515e-01
WEnt_d 1.910674e-01 2.469064e-01 0.7738452 4.390224e-01
WDen_via_d 2.956115e-02 1.914229e-02 1.5442850 1.225193e-01
WDen_int_d -3.986658e-04 1.271252e-03 -0.3136010 7.538241e-01
WDen_said_d -2.624905e-03 2.763274e-03 -0.9499259 3.421499e-01
WDist_CE_d -2.713424e-01 2.575563e-01 -1.0535266 2.920997e-01
WDen_tp_d -3.635249e-03 2.127516e-03 -1.7086825 8.750977e-02
lambda 1.149022e-01 NA 12.7862234 3.491813e-04
warning message:
glm.fit: fitted probabilities numerically 0 or 1 occurred
> Est_400@loglik
[1] 4076.077
> Est_400_0@loglik
[1] 4868.274

```

Figura A.5 – Resultados do modelo Probit SDM.

```

> E_marg=margins(Est_400)
> summary(E_marg)
      factor    AME    SE      z      p    lower    upper
  Crian  0.0199 0.0057  3.5195 0.0004  0.0088  0.0310
Den_emp_d -0.0000 0.0000 -2.3235 0.0202 -0.0000 -0.0000
Den_emp_o  0.0000 0.0000  4.6737 0.0000  0.0000  0.0000
Den_int_d -0.0001 0.0002 -0.8244 0.4097 -0.0005  0.0002
Den_int_o -0.0000 0.0002 -0.0381 0.9696 -0.0004  0.0004
Den_pop_d  0.0000 0.0000  4.4907 0.0000  0.0000  0.0000
Den_pop_o  0.0000 0.0000  1.0065 0.3142 -0.0000  0.0000
Den_said_d -0.0001 0.0003 -0.2456 0.8060 -0.0008  0.0006
Den_said_o -0.0004 0.0003 -1.0791 0.2805 -0.0010  0.0003
Den_tp_d -0.0005 0.0003 -1.8581 0.0632 -0.0010  0.0000
Den_tp_o  0.0005 0.0003  1.3775 0.1684 -0.0002  0.0011
Den_via_d  0.0023 0.0024  0.9751 0.3295 -0.0024  0.0071
Den_via_o -0.0027 0.0026 -1.0340 0.3011 -0.0078  0.0024
Disp_Veic -0.2569 0.0159 -16.1785 0.0000 -0.2881 -0.2258
Dist_CE_d  0.0364 0.0057  6.3586 0.0000  0.0252  0.0476
Dist_CE_o -0.0297 0.0057 -5.1978 0.0000 -0.0410 -0.0185
Ent_d  0.0225 0.0293  0.7658 0.4438 -0.0350  0.0799
Ent_o  0.1350 0.0289  4.6775 0.0000  0.0785  0.1916
Esc -0.0175 0.0013 -12.9594 0.0000 -0.0201 -0.0148
Gen -0.0865 0.0102 -8.4617 0.0000 -0.1066 -0.0665
Idade -0.0016 0.0004 -4.2502 0.0000 -0.0024 -0.0009
Ren -0.0000 0.0000 -6.8719 0.0000 -0.0000 -0.0000

```

Figura A.6 – Efeitos marginais médios do modelo Probit não espacial.

```

> E_marg=margins(Est_400)
> summary(E_marg)
      factor    AME    SE      z      p    lower    upper
  Crian  0.0198 0.0056  3.5129 0.0004  0.0088  0.0309
Den_emp_d -0.0000 0.0000 -0.5633 0.5732 -0.0000  0.0000
Den_emp_o  0.0000 0.0000 -0.5071 0.6121 -0.0000  0.0000
Den_int_d -0.0001 0.0003 -0.2148 0.8299 -0.0007  0.0006
Den_int_o -0.0002 0.0003 -0.5045 0.6139 -0.0008  0.0005
Den_pop_d  0.0000 0.0000  2.6884 0.0072  0.0000  0.0000
Den_pop_o  0.0000 0.0000  0.6848 0.4935 -0.0000  0.0000
Den_said_d  0.0005 0.0007  0.7043 0.4813 -0.0009  0.0019
Den_said_o -0.0011 0.0007 -1.4495 0.1472 -0.0025  0.0004
Den_tp_d  0.0003 0.0005  0.5450 0.5857 -0.0007  0.0013
Den_tp_o  0.0003 0.0006  0.5422 0.5877 -0.0008  0.0014
Den_via_d -0.0045 0.0049 -0.9079 0.3639 -0.0141  0.0052
Den_via_o  0.0008 0.0049  0.1555 0.8764 -0.0089  0.0105
Disp_Veic -0.2530 0.0159 -15.9405 0.0000 -0.2841 -0.2219
Dist_CE_d  0.1230 0.0825  1.4911 0.1359 -0.0387  0.2847
Dist_CE_o -0.1599 0.0811 -1.9700 0.0488 -0.3189 -0.0008
Ent_d -0.0469 0.0666 -0.7037 0.4816 -0.1774  0.0837
Ent_o -0.0997 0.0640 -1.5571 0.1194 -0.2252  0.0258
Esc -0.0175 0.0013 -13.0117 0.0000 -0.0201 -0.0149
Gen -0.0874 0.0102 -8.5672 0.0000 -0.1074 -0.0674
Idade -0.0017 0.0004 -4.3241 0.0000 -0.0024 -0.0009
Ren -0.0000 0.0000 -6.6547 0.0000 -0.0000 -0.0000
wDen_emp_d -0.0000 0.0000 -0.4115 0.6807 -0.0000  0.0000
wDen_emp_o  0.0000 0.0000  2.8151 0.0049  0.0000  0.0000
wDen_int_d -0.0001 0.0004 -0.3102 0.7564 -0.0009  0.0007
wDen_int_o  0.0002 0.0005  0.3941 0.6935 -0.0007  0.0011
wDen_pop_d -0.0000 0.0000 -1.1534 0.2487 -0.0000  0.0000
wDen_pop_o -0.0000 0.0000 -0.4114 0.6808 -0.0000  0.0000
wDen_said_d -0.0009 0.0009 -0.9765 0.3288 -0.0026  0.0009
wDen_said_o  0.0011 0.0009  1.2210 0.2221 -0.0007  0.0028
wDen_tp_d -0.0011 0.0007 -1.6374 0.1015 -0.0025  0.0002
wDen_tp_o  0.0004 0.0008  0.4990 0.6178 -0.0011  0.0019
wDen_via_d  0.0098 0.0061  1.5933 0.1111 -0.0022  0.0218
wDen_via_o -0.0051 0.0065 -0.7818 0.4343 -0.0177  0.0076
wDist_CE_d -0.0878 0.0826 -1.0628 0.2879 -0.2498  0.0741
wDist_CE_o  0.1329 0.0813  1.6344 0.1022 -0.0265  0.2922
wEnt_d  0.0598 0.0794  0.7542 0.4507 -0.0957  0.2154
wEnt_o  0.2953 0.0762  3.8730 0.0001  0.1458  0.4447

```

Figura A.7 – Efeitos marginais médios do modelo Probit SLX.