

PAULO JÚNIO MOURA ROSA

**Análise de modelos de escolha discreta espaciais para escolha de modo em viagens a pé
na cidade de São Paulo**

SÃO PAULO

2024

PAULO JÚNIO MOURA ROSA

**Análise de modelos de escolha discreta espaciais para escolha de modo em viagens a pé
na cidade de São Paulo**

Versão Corrigida

Dissertação apresentada à Escola Politécnica da
Universidade de São Paulo como requisito para obtenção
do título de Mestre em Ciências.

Área de Concentração: Engenharia de Transportes

Orientador: Prof. Dr. Cassiano Augusto Isler

SÃO PAULO

2024

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 29 de janeiro de 2024

Assinatura do autor: Paulo Júnio Moura Rosa

Assinatura do orientador: Cassiano Augusto Isler

Catálogo-na-publicação

Rosa, Paulo Júnio Moura

Análise de modelos de escolha discreta espaciais para escolha de modo em viagens a pé na cidade de São Paulo / P. J. M. Rosa -- versão corr. - São Paulo, 2024.

90 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Transportes.

1.Pedestres 2.Transportes I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Transportes II.t.

À minha família, que de todas as formas foi suporte, abrigo e força, em todos os momentos da
minha vida.

AGRADECIMENTOS

Toda a minha gratidão, que não pode ser expressa apenas em palavras, seja expressa com a minha vida, ao meu Deus, Ele sempre faz caminhos, abre portas e nos ensina que a perseverança é o trilho do caráter cristão!

Aos meus pais, Paulo César e Karla Beatriz, que sem sombra de dúvidas foram minha base, minha força, meu impulso, se hoje chego aqui, é porque estou sentado em ombros de gigantes, vocês sempre me impulsionaram e me ensinaram a ir além. As minhas irmãs, Maria Júlia e Ana Luiza, de tudo o que eu poderia ter vocês são o meu bem mais precioso, todo o meu amor e dedicação por vocês, obrigado por serem minhas melhores amigas. Amo vocês!

A minha amiga Luiza Beatriz, você sabe que eu não estaria aqui se não fosse você, em um momento de extrema dificuldade, você acreditou em mim e me fez romper aos meus limites, você tem uma participação tão especial nesse mestrado que com toda certeza essa conquista também é sua! A você gratidão pela amizade, pelo ensino de vida, e a certeza que sempre aqui você terá um amigo, toda minha admiração e carinho por você!

Ao meu orientador Professor Cassiano, com toda certeza eu não chegaria até aqui sem o senhor. O senhor definitivamente foi um grande conselheiro nessa jornada, me soube incentivar, me direcionar, me esticar, em cada momento. Eu não poderia desejar ou pensar em nenhuma outra pessoa para me ensinar tanto, mais do que um profissional o senhor é um grande ser humano, toda a minha admiração e respeito ao senhor, muito obrigado!

A toda a família Moura e Rosa, vocês são a minha alegria, em vocês tenho repouso e paz para os dias difíceis, eu amo ser um Moura Rosa. A minha amiga/irmã Amanda, que por toda a minha jornada sempre acreditou e me incentivou com palavras ou atitudes, sempre me ensinando que eu posso sonhar sim, porque sonhos existem para se realizar, que dividiu comigo minha segunda mãe, Tia Alessandra. A minha amiga Estella, que com palavras e orações, sempre foi um porto seguro para rasgar meu coração!

A minha família IBA Centro, em especial ao FIM, Artus e Rebeca, vocês me incentivaram e oraram sempre por mim, somos nós daqui até a eternidade. Sou grato às minhas irmãs, Élyda, Daniele e Gabriela, que sempre me ensinaram o valor de uma amizade, vocês são minha família em Goiás para sempre.

Meu muito obrigado à professora Cira Pitombo, ao professor André Chagas, e a professora Renata Magalhães por toda contribuição e direcionamentos na qualificação e na defesa, e que contribuíram significativamente para essa dissertação. Toda minha admiração aos senhores!

Aos amigos e companheiros de trabalho, Igor, Kaio, Ademar, Mariana, Ícaro, Cristiano, Olga e Fred, por tamanha paciência durante esse caminho e por nunca dificultarem o processo do mestrado, mas sempre me incentivarem, vocês são profissionais exímios.

À CAPES pelo apoio financeiro prestado nesta pesquisa. Obrigado à toda comunidade USP pelo esforço em permanecer sendo a melhor universidade brasileira e uma das melhores do mundo, por acreditar na Engenharia de Transportes como uma solução para melhorar a qualidade de vida das pessoas.

Minha gratidão a todos os professores e pessoas que de uma forma única e especial me incentivaram e acreditaram em mim, nem em meus maiores sonhos eu pensei que isso aconteceria, mestre pela USP, o que um dia era impossível se torna real! Perseverança e fé continuam sendo o melhor par de sapatos para o caminho árduo da conquista!

“Acima de tudo, guarde o seu coração, pois dele depende toda a sua vida.”

Provérbios 4:23

RESUMO

ROSA, P. J. M. **Análise de modelos de escolha discreta espaciais para escolha de modo em viagens a pé na cidade de São Paulo**. Dissertação (Mestrado em Engenharia de Transportes) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2023.

A análise do comportamento de viagens é uma das tradicionais áreas de estudos na literatura de Engenharia de Transportes. Apesar disso, ainda existe uma ausência de estudos na literatura quanto à modelagem de escolha de modos de viagens considerando o modo a pé como referência. Os modelos de escolha modal identificados na literatura são majoritariamente de base econométrica, e trabalhos incorporando os efeitos de dependência espacial nas modelagens têm sido utilizados para análise do fenômeno de viagens pelo modo a pé desde o início dos anos 2000. Nos estudos que incorporam estes efeitos, os resultados evidenciam uma melhora no desempenho dos modelos, além de viabilizar melhor compreensão sobre como as escolhas dos indivíduos e os aspectos que influenciam na escolha se relacionam no espaço. O objetivo principal deste estudo analisar a dependência espacial na escolha pelo modo a pé em viagens urbanas por motivo trabalho e estudo. A análise é feita utilizando modelos de escolha discreta espaciais e não espaciais, considerando os efeitos presentes na variável explicativa, por meio da abordagem *RUM*. Os resultados obtidos são relacionados a modelos *logit* tradicionais e *logit* espaciais nas variáveis explicativas. Ao se analisar os modelos estimados, o modelo espacial defasado nas variáveis explicativas (*Spatial Lag of X, SLX*) com matriz de vizinhança com critério de distância de 500 metros apresenta melhor desempenho. Os coeficientes dos atributos dos modos e do ambiente construído significativo é o tempo de viagem pelo modo a pé (*WALK*), os atributos sociodemográficos significativos são idade, grau de instrução e quantidade de automóveis, enquanto para os atributos espaciais apenas o critério de renda mais baixo é significativo. A análise do fenômeno permite a possibilidade de seu uso em políticas públicas, principalmente ao identificar os atributos de maior impacto na escolha, em especial o atributo que indica uma dependência espacial entre os indivíduos de maiores rendas quanto à escolha do modo a pé para viagens urbanas no contexto de São Paulo.

Palavras chaves – comportamento de viagens a pé; modelos de escolha discreta espaciais; dependência espacial; escolha de modo

ABSTRACT

ROSA, P. J. M. **Analysis of spatial discrete choice models for mode choice in pedestrian trips in the city of São Paulo**. Dissertação (Mestrado em Engenharia de Transportes) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2023.

The analysis of travel behavior is one of the traditional areas of study in Transportation Engineering literature. However, there is still a lack of studies in the literature regarding the modeling of travel mode choice considering walking mode as a reference. The modal choice models identified in the literature are predominantly econometric-based, and works incorporating the effects of spatial dependence in modeling have been used to analyze the phenomenon of walking mode trips since the early 2000s. In studies that incorporate these effects, the results demonstrate an improvement in model performance, as well as enabling a better understanding of how individuals' choices and influencing factors relate in space. The main objective of this study is to analyze spatial dependence in mode choice for walking in urban trips for work and study purposes. The analysis is conducted using spatial and non-spatial discrete choice models, considering the effects present in the explanatory variable through the Random Utility Model (RUM) approach. The obtained results are related to traditional logit models and spatial logit models in the explanatory variables. When analyzing the estimated models, the spatially lagged model in the explanatory variables (Spatial Lag of X, SLX) with a neighborhood matrix with a distance criterion of 500 meters shows better performance. The coefficients of mode attributes and significant built environment attributes are travel time by walking mode (WALK), significant sociodemographic attributes are age, level of education, and number of automobiles, while for spatial attributes only the lower-income criterion is significant. The analysis of the phenomenon allows for the possibility of its use in public policies, mainly by identifying the attributes with the greatest impact on choice, especially the attribute indicating spatial dependence among higher-income individuals regarding the choice of walking mode for urban trips in the context of São Paulo.

Keywords – walking travel behavior; spatial discrete choice models; spatial dependence; mode choice.

LISTA DE FIGURAS

Figura 1 - Fluxograma esquemático do método	40
Figura 2 - Zonas de Tráfego da pesquisa OD 2017 na RMSP	41
Figura 3 - Procedimento para tratamento do banco de dados	42
Figura 4 - Etapas da caracterização dos atributos das alternativas	47
Figura 5 - Escolha das alternativas da amostra analisada por origem	57
Figura 6 - Resultado teste I de Moran variável propensão de escolha do modo	58
Figura 7 - Origens dos indivíduos classificados por critério de renda mais baixa (CR2) e destinos por motivo de viagem.....	68
Figura 8 - Origens dos indivíduos classificados por critério de renda mais baixa (CR2) e destinos por escolha de modo (WALK/NOWALK).....	69

LISTA DE TABELAS

Tabela 1 - Relação da sigla do atributo e sua descrição.....	50
Tabela 2 - Relação de bibliotecas e funções utilizadas.....	54
Tabela 3 - Descrição das variáveis analisadas estatisticamente	56
Tabela 4 - Topologia das matrizes de vizinhança estimadas.....	59
Tabela 5 - Teste de dependência espacial I de Moran com matriz de vizinhança de 500m nas variáveis explicativas modeladas.....	60
Tabela 6 - Resultados gerais dos modelos estimados não espacial e espacial de 200m, 300m, 400m, 500m e 600m.....	61
Tabela 7 - Teste de multicolinearidade VIF	63
Tabela 8 - Resultados gerais dos modelos espaciais estimados de 500m, com cutoffs de 0,1; 0,2, 0,3 e 0,5	63
Tabela 9 - Resultados dos parâmetros estimados para os modelos não espacial e espacial com defasagem da 500 metros e cutoff 0,3.....	65

SUMÁRIO

1 INTRODUÇÃO.....	13
1.1 OBJETIVOS.....	16
1.2 HIPÓTESE.....	16
1.3 JUSTIFICATIVA.....	16
1.4 ESTRUTURA DO TEXTO.....	18
2 REVISÃO DA LITERATURA.....	19
2.1 COMPORTAMENTO DE VIAGENS PELO MODO A PÉ.....	19
2.1.1 Modelos de escolha de rota em viagens a pé.....	19
2.1.2 Modelos de escolha de destinos a pé.....	21
2.1.3 Modelos de escolha de modos de viagens a pé.....	21
2.1.4 Conclusões da seção.....	24
2.2 DEPENDÊNCIA ESPACIAL EM MODELOS DE ESCOLHA DISCRETA.....	24
2.2.1 Modelos de escolha discreta.....	25
2.2.2 Especificações de efeitos espaciais.....	30
2.2.3 Matriz de vizinhança e pesos espaciais.....	33
2.2.4 Estimadores.....	35
2.2.5 Aplicações.....	36
3. MÉTODO.....	40
3.1 BASE DE DADOS.....	41
3.1.1 Caracterização da pesquisa OD 2017.....	42
3.1.2 Seleção de viagens no município de São Paulo.....	43
3.1.3 Filtro por motivo de viagem.....	43
3.1.4. Definição das alternativas.....	43
3.1.5 Obtenção de dados sociodemográficos.....	44
3.1.6 Eliminação de observações duplicadas.....	44
3.1.7 Obtenção da distância mínima para matriz de vizinhança.....	45
3.2 CARACTERIZAÇÃO DOS ATRIBUTOS DAS ALTERNATIVAS.....	46
3.2.1 Níveis dos atributos das alternativas.....	48
3.3 RESUMO DOS ATRIBUTOS PARA MODELAGEM.....	49
3.4 ESTATÍSTICAS ESPACIAIS.....	50
3.4.1 Matriz de vizinhança.....	51
3.5 MODELOS DE ESCOLHA DISCRETA.....	52
3.5.1 Modelo <i>RUM</i> não espacial.....	52
3.5.2 Modelo <i>RUM</i> espacial <i>SLX</i>	52

3.6 SOFTWARE.....	54
4. RESULTADOS	55
4.1 RESULTADOS DAS ESTATÍSTICAS DESCRITIVAS	55
4.2 RESULTADOS DAS ESTATÍSTICAS ESPACIAIS.....	57
4.3 RESULTADOS GERAIS	61
4.4 RESULTADOS DOS PARÂMETROS ESTIMADOS.....	64
4.4.1 Parâmetros dos atributos dos modos e ambiente construído.....	65
4.4.2 Parâmetros dos atributos sociodemográficos	66
4.4.3 Parâmetros dos atributos espaciais	67
4.5 CONSIDERAÇÕES FINAIS	69
5 CONCLUSÕES.....	72
5.1 LIMITAÇÕES E RECOMENDAÇÕES PARA TRABALHOS FUTUROS.....	75
REFERÊNCIAS BIBLIOGRÁFICAS	77
APÊNDICE A- RESULTADOS DOS MODELOS	87

1 INTRODUÇÃO

A localização das atividades realizadas no contexto urbano e as condições dos sistemas de transportes para acessá-las afetam os deslocamentos e, conseqüentemente, a qualidade de vida das pessoas. No contexto dos países em desenvolvimento, observa-se ainda a dificuldade de adequação das cidades quanto ao ordenamento do uso do solo e a promoção da qualidade de deslocamentos pelos modos coletivos e ativos (DAVIS, 2006). Essas dificuldades geram diminuição da qualidade de vida dos indivíduos, além de efeitos diretos e indiretos na economia, saúde e desenvolvimento social (VASCONCELLOS, 2018).

Os modos de transporte ativos (a pé ou não motorizados) e coletivos são priorizados na legislação brasileira estabelecida pela Política Nacional de Mobilidade Urbana por meio da Lei nº 12.587/2012. Ao considerar esse instrumento para melhoria da mobilidade urbana, os modos ativos possuem prioridade sobre os demais (BRASIL, 2012). Concomitantemente à vigência dessa legislação, o número de pesquisas relacionadas aos modos ativos tem aumentado pela relevância de sistemas como *bikesharing*, pela implementação de infraestruturas que incentivam o uso desses modos, além do modo a pé sempre estar relacionado à última milha (ABREU; ALMEIDA; TURINI, 2021).

Estudos relacionados ao comportamento de viagens pelo modo a pé indicam dois principais tipos de análise. A primeira abordagem se relaciona ao comportamento do pedestre, por exemplo para explicar como as pessoas se comportam em situações de emergência. Esses estudos surgiram para propor planos de evacuação que zelassem pela segurança, diminuindo as possibilidades de conseqüências graves em caso de acidentes. Em 2001, a *Pedestrian and Evacuation Dynamics*, primeira conferência internacional debatendo o assunto, foi realizada em Duisburg, Alemanha (SCHRECKENBERG; SHARMA, 2002). A partir dessa conferência aumentou o interesse pela simulação do comportamento de pedestres além do estabelecimento de conceitos sobre o tema (BIERLAIRE; ANTONINI; WEBER, 2003).

Neste tipo de análise, os agentes e o espaço são cruciais para modelagem. O agente é o líder, sendo analisado para capturar o padrão de comportamento do pedestre pelas interações com os demais indivíduos a partir de uma seqüência de escolhas. Esse líder tem uma função semelhante à observada em análises de Engenharia de Tráfego quanto ao veículo líder em um fluxo. O espaço é entendido como o ambiente onde ocorrem as relações entre os agentes, sendo definido de acordo com a escala da análise a ser realizada (BIERLAIRE; ANTONINI; WEBER,

2003). As três principais interações que surgem a partir desses elementos são: agentes-agentes, agentes-indivíduos e agentes-espaço.

As interações do tipo agentes-agentes buscam identificar os aspectos do comportamento de um agente que influenciam nas escolhas de outros agentes. As interações agentes-indivíduos indicam padrões de comportamento quanto ao estilo de caminhada e tomada de decisão que os indivíduos realizam em detrimento de copiar o agente. A relação agentes-espaço ocorre pelo conjunto de escolhas que o agente faz de acordo com as características do espaço onde ocorre sua caminhada (BIERLAIRE; ANTONINI; WEBER, 2003). Essas relações são discutidas por Bierlaire, Antonini e Weber (2003), Antonini (2005), Antonini, Bierlaire e Weber (2006) e Robin *et al.* (2009), em que estruturam um padrão de comportamento dos pedestres.

O segundo tipo de análise que aborda o comportamento de viagens pelo modo a pé busca explicar o comportamento dos pedestres e estabelecer padrões coletivos. Nesse sentido, existem estudos que se dedicam a analisar a escolha do modo, os destinos de viagens e as rotas escolhidas em viagens a pé. Nas análises de escolha de modo identificam-se as características socioeconômicas das pessoas que decidem viajar a pé, assim como os atributos que condicionam a escolha. Nas análises de escolha de destino, tem-se o objetivo de determinar os destinos dos indivíduos ou atividades que preferem realizar pelo modo a pé. Por fim, os estudos de escolha de rotas se dedicam a identificar as diferentes rotas que as pessoas selecionam para realizar seus deslocamentos utilizando o modo a pé. Nesses três formatos, o pressuposto é que o tomador de decisão faz escolhas para maximizar a sua utilidade.

Em resumo, é possível organizar os temas analisados em duas principais vertentes, a que se dedica ao comportamento dos pedestres para explicar como os indivíduos se deslocam no espaço, e a que se dedica ao comportamento de viagens pelo modo a pé, buscando explicar o motivo da escolha pela viagem pelo modo a pé, incluindo estudos de escolha de rota, de destino e de modo. Esta dissertação tem enfoque no comportamento em viagens pelo modo a pé.

Modelos de escolha discreta têm sido tradicionalmente utilizados para explicar as motivações das escolhas individuais, sendo baseados em teorias econômicas para representar as escolhas dos indivíduos perante um conjunto de alternativas. Para definir as alternativas utilizadas, considera-se em geral uma série de atributos que caracterizam aquelas disponíveis aos indivíduos (MCFADDEN, 1974; BEN-AKIVA; LERMAN, 1985).

A Teoria da Utilidade é a mais utilizada para a modelagem do fenômeno de escolha de modos de transportes. A partir da abordagem da maximização da utilidade aleatória (*Random*

Utility Maximization - RUM) define-se que um indivíduo escolherá a alternativa de maior utilidade entre aquelas disponíveis. Os atributos dessas alternativas e as características específicas do indivíduo são fatores determinantes para caracterizar a utilidade, existindo também uma parcela não observável relacionada a um termo de erro (MCFADDEN, 1974; BEN-AKIVA e LERMAN, 1985; ORTÚZAR; WILLUMSEN, 2011).

Tradicionalmente, os coeficientes dos atributos referentes ao custo monetário e tempo de viagem associados aos modos de transporte apresentam significância estatística para explicar as escolhas dos indivíduos (ORTÚZAR; WILLUMSEN, 2011). Entretanto, ao considerar viagens pelo modo a pé, o custo monetário é igual a zero e outros atributos podem ser utilizados para caracterizar a utilidade desse modo de transporte.

A abordagem *RUM* tem sido utilizada para explicar as complexas escolhas dos indivíduos, destacando-se especialmente em aplicações relacionadas a problemas de escolha de modos de viagens (CHORUS, 2012). A partir da análise do fenômeno da escolha de modo, aspectos espaciais e sociais foram incorporados para representar as relações que atuam na tomada de decisão como apresentado por Zhou *et al.* (2016) e Carrión-Flores *et al.* (2018).

A partir da Econometria Espacial são derivados os modelos que proporcionam a análise dos efeitos espaciais. Segundo Anselin (2003), existem dois tipos de efeitos considerados em modelagens espaciais, os efeitos de impacto global e local. Por um modelo geral linear, são especificados três principais modelos com o intuito de controlar a dependência espacial: o *Spatial Autoregressive Model (SAR)*; o *Spatial Lag of X (SLX)*; e o *Spatial Error Model (SEM)*.

De forma resumida, o modelo *SAR* é de efeito global e sintetiza as estruturas de dependência entre as preferências dos indivíduos ou interdependência nas escolhas. O *SLX* é um modelo de efeito local e sintetiza as estruturas de dependência entre os atributos não observados ou alternativas selecionadas (efeito de transbordamento das características dos vizinhos denominado *spillover*). Por fim, o modelo *SEM* é de efeito global e integra os efeitos *spillover* locais nas especificidades dos indivíduos, tratando a heterogeneidade a partir de choques comuns que são propagados nas variáveis explicativas, caracterizando a observação com um efeito direto e influenciando a variável dependente de forma indireta. O modelo *SAR* é mais aplicado nos estudos de escolha de modo para viagem (SIDHARTHAN *et al.*, 2011; WANG *et al.*, 2015; SHENG; SHARP, 2019) e, em contrapartida, são poucas as aplicações do modelo *SLX* nesse contexto.

A aplicação da modelagem espacial proporciona um entendimento do comportamento de viagens que relaciona as escolhas dos indivíduos de acordo com sua localização no espaço, viabilizando a interpretação dos efeitos dos atributos do modo e sociodemográficos pela componente espacial. Por meio desses modelos é possível obter informações relevantes sobre o fenômeno e indicações de políticas públicas em transportes que se relacionam aos efeitos de dependência espacial.

Apesar da existência de estudos sobre escolha de modo para viagens a pé, poucos são aqueles com enfoque na dependência espacial; pelo conhecimento deste autor somente o trabalho de Neves (2019) e Neves, Silva e Arruda (2021), esta condição pode estar relacionada dentre outros fatores pela dificuldade de modelagem deste modo, pelo quantitativo de viagens que são realizadas exclusivamente pelo modo a pé, e principalmente por ser um modo menos considerado na formulação de políticas públicas em detrimento dos demais. Neste contexto, considera-se a seguinte pergunta de pesquisa: Qual a influência do efeito espacial local nas escolhas individuais na realização de viagens a pé?

1.1 OBJETIVOS

O objetivo principal da pesquisa é analisar a dependência espacial na escolha pelo modo a pé em viagens urbanas por motivo trabalho e estudo. Para isso propõe-se o objetivo específico de identificar os atributos que afetam a escolha de modo em viagens urbanas considerando o modo a pé como uma alternativa.

1.2 HIPÓTESE

A hipótese deste estudo é de que existem evidências de dependência espacial na escolha do modo a pé para viagens urbanas por motivo trabalho e estudo.

1.3 JUSTIFICATIVA

A compreensão do comportamento relacionado a viagens tem sido realizada por diferentes estratégias e abordagens, a fim de obter uma melhor modelagem do comportamento

dos indivíduos. Recentemente, a estratégia de incluir efeitos exógenos de diferentes formas tem sido adotada. Hess *et al.* (2012) consideraram efeitos exógenos nas estratégias de decisão, Chorus, Rose e Hensher (2013) utilizaram-nos no processamento dos atributos, enquanto Greene e Hensher (2003) e Hensher e Greene (2003) utilizaram efeitos relacionados aos tomadores de decisão. Outro formato de abordagem se baseia nas atividades, observando as escolhas das pessoas quanto aos motivos que levam a realizar atividades no ambiente urbano, atribuindo a esses modelos um componente que captura os efeitos sociais e espaciais (PINJARI; BHAT, 2011).

A partir da abordagem *RUM*, existe uma série de possibilidades para analisar o comportamento de viagens. Mesmo com o desenvolvimento de estudos na área, ainda foram identificadas duas vertentes com poucos trabalhos no contexto da literatura da Engenharia de Transportes. A primeira se relaciona à incorporação de efeitos de interação espacial em análises de viagens, sendo esta ausência relacionada à dificuldade de estimativa dos modelos, e a segunda sobre modelos de escolha modal que tenham viagens a pé como base, dada a dificuldade de modelagem do fenômeno.

No que tange a primeira vertente, a incorporação de efeitos espaciais em modelos de escolha discreta na área de Engenharia de Transportes ainda é pouco explorada. Em geral, o estado da arte na Econometria Espacial tem enfoque nas ciências regionais, com destaque aos estudos que analisam os efeitos espaciais do tipo global pelos modelos *SAR* e *SEM*.

Outra forma de analisar a dependência espacial nos modelos de escolha discreta é observar se existem efeitos espaciais do tipo local pelo modelo *SLX*, cuja interdependência espacial é resultado da reflexão sobre as relações das pessoas com seu círculo social, em que a escolha do indivíduo é resultado das suas relações em seu convívio social. Essas relações podem ser observadas tanto no espaço social como no espaço geográfico (PÁEZ; SCOTT, 2007; PÁEZ; SCOTT; VOLZ, 2008).

Análises de influência social na literatura sobre Engenharia de Transportes aumentaram no contexto do comportamento relacionado a viagens (DUGUNDJI *et al.*, 2011). Em relação à influência espacial, as pesquisas têm investigado esses efeitos em escolhas de modo de transportes, inicialmente por Goetzke (2003) e em seguida por Sidharthan *et al.* (2011), Wang *et al.* (2015) e Sheng e Sharp (2019).

A segunda vertente na literatura se estabelece pela ausência de estudos que investigam modelos de escolha de modo para viagens a pé com efeitos espaciais. Apesar de existirem

estudos sobre escolhas de modo sob efeitos espaciais como citados anteriormente, são raros aqueles que analisam as viagens realizadas a pé especificamente. Os estudos que se propõem a investigar o comportamento dos pedestres têm enfoque nas relações com o espaço, com os outros pedestres, e suas decisões quanto às rotas para realizar suas viagens (BIERLAIRE; ANTONINI; WEBER, 2003; ANTONINI, 2005; ANTONINI; BIERLAIRE; WEBER, 2006; ROBIN *et al.*, 2009). Assim, não são considerados os fatores que levam os indivíduos a realizarem suas viagens a pé ao invés de outros modos de transporte.

Esta pesquisa pretende analisar especificamente as escolhas de viagem pelo modo a pé considerando os efeitos espaciais que podem influenciá-las. Além disso, busca contribuir para a literatura sobre modelagem de escolha discreta considerando viagens a pé e efeitos de interação espacial, com aplicação no contexto urbano do município de São Paulo.

1.4 ESTRUTURA DO TEXTO

Este texto é composto por 5 capítulos. O capítulo 2 apresenta o estado da arte sobre análise quanto ao comportamento de viagens a pé, além dos modelos de escolha discreta espaciais e suas aplicações na análise de comportamento relacionado a viagens. O capítulo 3 apresenta o processo de construção do banco de dados e o método proposto para incorporar a dependência espacial em modelos de escolha discreta na aplicação à cidade de São Paulo. O capítulo 4 descreve a base de dados utilizada na pesquisa e resultados da modelagem aplicada. Finalmente, o capítulo 5 apresenta as conclusões da pesquisa e recomendações para trabalhos futuros.

2 REVISÃO DA LITERATURA

Esta revisão da literatura busca aprofundar a compreensão dos estudos voltados para o comportamento de viagens a pé, com foco particular na abordagem de modelos de escolha discreta. Além dessa análise, também são explorados os modelos de escolha discreta espaciais, com destaque para suas características específicas, tipologias, métodos de estimação e aplicações empíricas.

2.1 COMPORTAMENTO DE VIAGENS PELO MODO A PÉ

A literatura aborda as decisões tomadas pelos indivíduos durante suas viagens a pé, englobando escolhas de rotas, destino e comparação com outros modos de deslocamento. Essas escolhas são frequentemente estudadas através de modelos de escolha discreta e revelam como os atributos influenciam a seleção de determinada alternativa e os padrões de escolha das pessoas quando são apresentados cenários semelhantes à vida real.

Os dados empregados nos modelos de escolha discreta provêm, em sua maioria, da caracterização de atributos que abrangem características sociodemográficas dos indivíduos amostrados, além daqueles que descrevem as características das alternativas analisadas. Nas próximas seções serão apresentados detalhadamente esses formatos de análise, bem como as aplicações específicas relacionadas ao contexto das viagens a pé.

2.1.1 Modelos de escolha de rota em viagens a pé

Os estudos sobre escolhas de rotas durante viagens a pé abordam as decisões diante das diferentes opções de trajetos disponíveis aos indivíduos. Hill (1982) identificou que a seleção de rotas pelos pedestres ocorre de maneira subconsciente. Um fator chave que influencia a escolha de uma rota específica é sua direção, não apenas em relação à distância, mas também considerando as mudanças ao longo do percurso. Senevarante e Morall (1986) indicaram que, em geral, os pedestres tendem a optar pela rota mais curta, mesmo que essa escolha seja tomada inconscientemente.

Além da distância, outros estudos destacam a importância da atratividade de uma rota na determinação da escolha (HOOGENDOORN; BOVY, 2004; PAPADIMITRIOU; YANNIS; GOLIAS, 2009). Outros fatores relevantes incluem hábito, número de travessias, níveis de poluição e ruído, segurança, seguridade, proteção contra intempéries e estímulo ao ambiente. Esses fatores têm diferentes influências na escolha dependendo dos propósitos ou motivações da viagem (BOVY e STERN, 1990).

Hoogendoorn e Bovy (2004) exploraram a teoria do comportamento do pedestre sob incerteza baseada na maximização da utilidade. Discutiram fatores influentes na escolha de rotas, bem como abordagens de Teoria de Filas, modelos de escolha discreta, modelos de escolha racional subjetiva e modelos *Space Syntax* que indicam que a forma e a organização do espaço físico têm impactos positivos significativos na maneira como as pessoas se movem e interagem em um determinado ambiente. Atributos relevantes incluem custo generalizado de caminhada, utilidade da atividade, tempo de viagem, velocidade do pedestre e desconforto ao caminhar próximo a obstáculos e paredes.

Papadimitriou, Yannis e Golias (2009) realizaram uma revisão crítica dos modelos de escolha de rotas e do comportamento de travessia em áreas urbanas e indicaram a complementaridade e possível interdependência desses dois aspectos, embora muitas vezes examinados separadamente. No entanto, ressaltaram a falta de uma abordagem abrangente do comportamento do pedestre em uma viagem urbana.

Borst *et al.* (2009) investigaram como o ambiente influencia a caminhada de idosos utilizando um modelo de escolha de rota que considera as características das ruas com aplicação de ferramentas de análise espacial para modelar a escolha das rotas. Os resultados sugerem que o ambiente local impacta o comportamento de caminhada dos idosos, com características de acesso às residências desempenhando um papel importante.

Finalmente, Tribby *et al.* (2017) desenvolveram uma abordagem combinando seleção de dados orientada e modelagem de escolha de rota para analisar a caminhabilidade. A partir de dados de Salt Lake City, EUA, aplicaram a técnica de floresta aleatória para selecionar variáveis nos modelos de escolha de rota. Essa técnica melhorou consideravelmente a qualidade de ajuste dos modelos em relação aos baseados em dimensões predefinidas de caminhabilidade.

2.1.2 Modelos de escolha de destinos a pé

Os modelos de escolha de destinos estão relacionados com as distribuições de viagens, sendo exemplificado pelo modelo gravitacional, embora frequentemente sejam abordados por meio de modelos desagregados como os de escolha discreta. Normalmente, as especificações incluem dois condicionantes, os termos de impedância (distância, custos generalizados), as informações relacionadas ao propósito da viagem e a estrutura espacial (BEN-AKIVA e LERMAN, 1985; BHAT *et al.*, 1998).

Na literatura sobre escolha de destinos a pé, vários fatores relacionados ao ambiente construído que influenciam a caminhada foram amplamente investigados, incluindo: distância aos destinos; densidade residencial; densidade de empregos; uso do solo misto; conectividade; acessibilidade ao transporte público; e características das calçadas (SAELENS e HANDY, 2008; EWING e CERVERO, 2010; SAELENS *et al.*, 2003; SCHNEIDER *et al.*, 2009; RODRIGUEZ e JOO, 2004).

Clifton *et al.* (2016) analisaram o comportamento de escolha de destinos de pedestres a partir da pesquisa de viagens familiares de Portland, EUA, em 2011, para identificação das escolhas de destinos em viagens a pé por meio de um modelo *logit* multinomial para seis propósitos de viagem. Como resultados, indicaram que a distância foi um impedimento significativo para a escolha do destino de pedestres, e as pessoas em domicílios sem carro ou sem filhos eram menos sensíveis à distância para alguns propósitos de viagem.

2.1.3 Modelos de escolha de modos de viagens a pé

É importante destacar certas características relacionadas a viagens pelo modo a pé, uma vez que este modo de deslocamento apresenta particularidades. A primeira a ser considerada é que ocorre sem o uso de um veículo, não motorizado ou motorizado, e a própria pessoa realiza o deslocamento. Por essa condição, o custo monetário não existe e o tempo é o principal atributo que caracteriza o custo generalizado.

Como nos demais modos, existem outras características que explicam a escolha do indivíduo. Como observado por Olojede, Yoade e Olufemi (2017) e Ton *et al.* (2019), as estruturas do ambiente onde ocorre a caminhada, ou seja, a distância, o uso do solo misto na região, e a densidade de acesso ao serviço de transporte coletivo são alguns atributos que

influenciam na escolha do modo a pé. Além dessas, características socioeconômicas como a renda e a posse de veículos são atributos que contribuem para a escolha desse modo na realização de viagens. Os estudos conduzidos por Kamargianni e Polydoropoulou (2013) e Ding *et al.* (2017) abordam a integração dos modos a pé e bicicleta, referindo-se a eles como modos ativos ou modos não motorizados. Na literatura, são encontrados diversos estudos que tratam do modo a pé como uma opção de deslocamento, seja como um único modo ou como um modo ativo (com a bicicleta).

Kamargianni e Polydoropoulou (2013) aplicaram uma pesquisa de preferência declarada para capturar o comportamento de adolescentes e estimar modelos de escolha de modos híbridos, nos quais as utilidades dependem tanto dos atributos do modo quanto da variável latente “propensão à caminhar ou pedalar”. Um modelo de equações estruturais foi estimado com dados extraídos de uma pesquisa realizada em todas as escolas de ensino médio de Chipre em 2012. A amostra foi composta por 4.174 adolescentes de 12 a 18 anos e abrangeu 8,7% da população total do ensino médio. Como principais resultados foi apontado que a existência de ciclovias e calçadas largas afetam significativamente a escolha do transporte ativo. Disposição para caminhar e pedalar têm efeito positivo na escolha dessas alternativas e efeito negativo na escolha pelo automóvel.

Ding *et al.* (2017) concentraram suas análises no impacto das atitudes para caminhar e andar de bicicleta na escolha do modo de deslocamento, pela estimativa de um modelo integrado de escolha discreta e de equações estruturais para examinar os dados de uma pesquisa na cidade de Zhenjiang, China, em 2015. Os autores indicam que, pela comparação dos modelos, o de escolha híbrida supera o modelo de escolha discreta tradicional.

Olojede, Yoade e Olufemi (2017) examinaram os fatores que influenciam a escolha da caminhada como modo de viagem ativo na cidade de Ilesa, Nigéria. Os autores estimaram um modelo de escolha de modo utilizando a técnica de amostragem em multiestágio com 524 respostas de três principais zonas residenciais da cidade quanto à escolha do modo a pé para viagens. Como resultado foi indicado que as variáveis que mais influenciam a escolha do modo são: não possuir veículos; renda; comprimento da viagem; custo da viagem; e benefícios à saúde.

Ton *et al.* (2019) buscaram identificar os determinantes que influenciam a escolha por um modo de transporte ativo quando se considera um conjunto mais abrangente de modos na Holanda, onde há alto uso de modos ativos. Foi utilizado um modelo de escolha de modo *logit*

misto multinomial com enfoque nos modos ativos. O modelo foi estimado com os dados do Painel de Mobilidade da Holanda e de uma pesquisa adicional focada em modos ativos. Os autores indicaram que andar de bicicleta e caminhar devem ser consideradas alternativas distintas nas análises. Além disso, o uso do modo ativo é mais sensível às mudanças nas características da viagem e no ambiente construído que o modo motorizado.

Caldeira *et al.* (2021) analisaram os padrões de deslocamento urbano a pé de indivíduos maiores de 50 anos e suas variações entre macrorregiões e cidades de diferentes estratos populacionais no Brasil. Os autores utilizaram modelos mistos e não-mistos de regressão binomial negativo inflacionado de zeros, utilizando-se o banco de dados do Estudo Longitudinal da Saúde dos Idosos Brasileiros (ELSI) para identificar a existência de heterogeneidades no contexto nacional e analisar os fatores que afetam a escolha pela caminhada, bem como a intensidade daqueles que caminham. Resultados evidenciaram a dificuldade de se locomover como principal fator relacionado à escolha pela caminhada, enquanto a segurança pública, qualidade das calçadas e risco de atravessar a rua são fatores que diminuem o tempo semanal de caminhada.

Um último estudo na temática foi desenvolvido por Neves, Silva e Arruda (2021). O objetivo do trabalho foi investigar a conexão entre as variáveis do ambiente construído e a escolha de viajar caminhando em São Paulo, e a comparação dos resultados com a literatura para países em desenvolvimento. Os autores empregaram modelos de escolha de modo de transporte em suas pesquisas para examinar a seleção do modo a pé considerando variáveis do ambiente construído e socioeconômicas. Mais especificamente no que tange às características do ambiente construído, os grupos de variáveis utilizadas foram densidade (populacional e de empregos), diversidade (entropia quanto ao uso misto do solo), desenho (densidade de ruas por km²), acessibilidade do destino (distância ao centro de comércio e negócios) e distância ao transporte coletivo.

Os autores estimaram um modelo de escolha discreta espacial, mas o incremento marginal em relação ao modelo *logit* tradicional foi pequeno conforme relatado pelos autores. Como principais resultados indicaram que, para a cidade de São Paulo, as variáveis do ambiente construído são mais relevantes na origem e a dimensão mais relacionada à escolha da caminhada é a diversidade, provavelmente por motivos socioeconômicos. As características dos indivíduos também tiveram efeitos significativos, como a idade, o sexo e a renda.

Como evidenciado pelo estudo, as variáveis do ambiente construído são um dos fatores que mais se relacionam à escolha do modo a pé em viagens urbanas. A presente pesquisa faz a conexão direta entre os elementos dos dois grupos de estudo evidenciados na literatura, comportamento do pedestre (ambiente construído) e comportamento de viagens (escolha de modo para viagem).

Conforme observado, os principais grupos de variáveis explicativas para o comportamento de viagens a pé dizem respeito ao ambiente construído, características da viagem (tempo de viagem, distância de viagem), atitudinais (propensão de viajar pelo modo por benefícios da saúde, propensão de viajar pelo modo por benefícios ao meio ambiente) e sociodemográficas (renda, posse de veículo, idade).

2.1.4 Conclusões da seção

De forma geral, os estudos que utilizam a modelagem de escolha discreta buscam descrever o comportamento de viagens a pé perante um conjunto de alternativas. Para cada tipo de problema (escolha de rota, escolha de destino ou escolha de modo), um conjunto de atributos é elencado para descrever a escolha do indivíduo. Os atributos aqui discutidos são fundamentais para a comparação quanto aos resultados obtidos nesta pesquisa.

Todos os tópicos apresentados se relacionam para o entendimento geral do comportamento de viagens a pé. A compreensão do padrão de caminhada, os aspectos do agente, assim como do espaço, além das variáveis explicativas da escolha da rota, do destino e do modo, são relevantes para subsidiar discussões sobre certas preferências e os comportamentos que são capturadas pelos modelos de escolha discreta.

2.2 DEPENDÊNCIA ESPACIAL EM MODELOS DE ESCOLHA DISCRETA

Esta seção descreve técnicas de inclusão de dependência espacial nos modelos de escolha discreta, além da especificação de matrizes de vizinhança e de pesos espaciais, e aplicações encontradas na literatura de comportamento de viagens na Engenharia de Transportes.

2.2.1 Modelos de escolha discreta

Os modelos de escolha discreta em Engenharia de Transportes têm sido extensivamente utilizados para representar o comportamento dos indivíduos perante um conjunto de alternativas para escolha mediante suas características socioeconômicas e as condições do meio em que ocorre a tomada de decisão. A partir de teorias probabilísticas, os modelos de escolha discreta capturam as escolhas mais prováveis de um indivíduo perante as alternativas disponíveis.

Entre um conjunto de alternativas estabelecidas pelo modelador podem surgir discrepâncias nas escolhas feitas pelo tomador de decisões que não podem ser diretamente observadas. Segundo a abordagem *RUM*, os indivíduos sempre buscam a maximização da utilidade na comparação entre as alternativas disponíveis (BEN-AKIVA e LERMAN, 1985).

Por esse conceito, a utilidade de uma alternativa é medida por uma parcela mensurável (determinística) e uma parcela aleatória (erro aleatório) conforme a equação (1). Assim, entende-se que o termo de erro aleatório não pode ser medido ou não é conhecido, e pode representar idiosincrasias dos indivíduos quanto às suas escolhas.

$$U_{ni} = V_{ni} + \varepsilon_{ni} = \sum_{k=1}^K \beta_k x_{nk} + \varepsilon_{ni} \quad (1)$$

em que U_{ni} é a utilidade da alternativa i para indivíduo n ;

V é a parcela determinística da utilidade da alternativa i para o indivíduo n ;

ε é o termo de erro aleatório da alternativa i para o indivíduo n ;

x_{nk} são os k atributos que influenciam a utilidade do indivíduo n ; e

β_k são os k parâmetros estimados sob a abordagem *RUM*.

Diferentes modelos probabilísticos podem ser derivados ao se atribuir uma distribuição de densidade de probabilidade ao termo de erro. As distribuições mais consideradas na literatura são Valor Extremo do Tipo I e Normal. Ao considerar uma distribuição de Valor Extremo do Tipo I, o modelo probabilístico resultante é conhecido por *logit*, e no caso da distribuição Normal o modelo resultante é o *probit*. A probabilidade de escolha de uma alternativa por um indivíduo pode ser escrita conforme a equação (2) (TRAIN, 2009).

$$P_{ni} = Prob(U_{ni} > U_{nj} \quad \forall j \neq i) \quad (2)$$

$$\begin{aligned}
&= Prob(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \quad \forall j \neq i) \\
&= Prob(\varepsilon_{nj} - \varepsilon_{ni} < V_{nj} - V_{ni} \quad \forall j \neq i)
\end{aligned}$$

Essa probabilidade representa uma função de distribuição acumulada, em que a probabilidade de um componente aleatório $\varepsilon_{nj} - \varepsilon_{ni}$ está abaixo de um valor observado $V_{nj} - V_{ni}$. Assim, a partir da função de densidade $f(\varepsilon_n)$, a função de probabilidade pode ser rescrita conforme a equação (3) (TRAIN, 2009).

$$P_{ni} = Prob(\varepsilon_{nj} - \varepsilon_{ni} < V_{nj} - V_{ni} \quad \forall j \neq i) \quad (3)$$

$$= \int_{\varepsilon} I(\varepsilon_{nj} - \varepsilon_{ni} < V_{nj} - V_{ni} \quad \forall j \neq i) f(\varepsilon_n) d\varepsilon_n$$

A função $I(\cdot)$ é indicadora, assumindo valor 1 quando o termo é verdadeiro, e 0 caso contrário. Assim, diferentes modelos de escolha discreta podem ser derivados. No caso do modelo *logit* os termos de erro são Independentes e Identicamente Distribuídos (*IID*) e assumem distribuição Valor Extremo do Tipo I, tal que a integral da equação (3) assume uma forma fechada e a probabilidade de escolha da alternativa pelo indivíduo é calculada conforme a equação (4) (BEN-AKIVA; LERMAN, 1985; TRAIN, 2009).

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} \quad (4)$$

Essa expressão representa a probabilidade de um indivíduo n escolher uma alternativa i entre um conjunto de alternativas $j \geq 2$. Outras especificações para a função de densidade dos termos de erro não apresentam formas fechadas de cálculo, como é o caso da distribuição Normal. Essa condição faz com que os modelos com outras especificações de distribuição sejam menos utilizados na literatura, tendo em vista principalmente a necessidade de simulação para a estimação das probabilidades.

Aos modelos derivados da Teoria de Maximização da Utilidade têm sido incorporados os efeitos de aspectos espaciais na tomada de decisão. A especificação de modelos de escolha discreta espaciais não é uma tarefa simples, já que existem poucos testes na literatura disponíveis para determinar a presença de dependência espacial em modelos não lineares, além da aplicação não ocorrer de forma direta como nos modelos lineares. As primeiras aplicações deste tipo de modelagem ocorreram nas ciências regionais, principalmente para analisar fenômenos de problemas de localização de firmas e escolha de uso/parcelamento do solo. Nesse

contexto, os modelos *SAR* e *SEM* se destacam na literatura. Na área de análise do comportamento individual em transportes isso se repete pela especificação recorrente do modelo *SAR* (CALDEIRA, 2022).

De acordo com Billé e Arbia (2019), a formulação geral da função utilidade de um modelo binário considerando os efeitos espaciais pelo modelo aninhado não linear geral (*GNNM – General nonlinear nested model*) tem a estrutura das equações (5), (6) e (7).

$$\mathbf{y}^* = \rho W_{1,n} \mathbf{y}_n^* + \mathbf{X}_n \boldsymbol{\beta} + W_{2,n} \tilde{\mathbf{X}}_n \boldsymbol{\theta} + \mathbf{u}_n \quad (5)$$

$$\mathbf{u}_n = \lambda \mathbf{M}_n \mathbf{u}_n + \varepsilon_n, \quad \varepsilon_n \sim \varepsilon_n(\mathbf{0}_n, \Sigma_\varepsilon) \quad (6)$$

$$\mathbf{y}_n = \mathbb{I}_n(\mathbf{y}_n^* > \mathbf{0}_n) \quad (7)$$

em que \mathbf{y}^* é o vetor coluna de variáveis dependentes contínuas e latentes de dimensão n ;

\mathbf{y}_n é o vetor coluna de dimensão n de variáveis dependentes binárias definidas pela função indicadora $\mathbb{I}_n(\mathbf{y}_n^* > \mathbf{0}_n) = I(y_{n1}^* > 0), \dots, I(y_{nn}^* > 0)'$ de dimensão n ;

\mathbf{X}_n é a matriz ($n \times k$) de variáveis exógenas, incluindo o termo constante; $\tilde{\mathbf{X}}_n$ é a mesma matriz de variáveis exógenas, excluindo-se apenas o termo constante;

$W_{1,n}$ e $W_{2,n}$ e \mathbf{M}_n são matrizes espaciais ($n \times n$) de constantes conhecidas;

$\boldsymbol{\theta}$ é um vetor coluna de dimensão k que capta os efeitos de correlação espacial local;

$\boldsymbol{\gamma} = (\boldsymbol{\beta}', \rho, \lambda)'$ é um vetor de dimensão $k + 2$ de parâmetros com coeficientes autorregressivos ρ e λ que capturam efeitos de *spillovers* globais e com $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k)'$;

ε_n é um vetor normal multivariado de inovações com média zero e variância $\sigma_\varepsilon^2 < \infty$, tal que $\Sigma_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_n$.

A partir desse modelo genérico é possível definir modelos específicos capazes de representar a dependência espacial presente nos diferentes fenômenos analisados. O modelo com $\theta=0$ e $\lambda=0$ é conhecido como *Spatial Autoregressive Model (SAR)*, indicando a dependência espacial presente na variável dependente. O modelo com $\rho=0$ e $\lambda=0$ é denominado *Spatial Lag of X (SLX)*, que indica a dependência espacial nas variáveis independentes. Quando $\rho=0$ e $\theta=0$ o modelo é conhecido por *Spatial Error Model (SEM)* e representa dependência

espacial nos erros. Cabe destacar novamente que o modelo *SAR* e *SEM* são de efeitos globais, enquanto o modelo *SLX* é de efeito local.

A distribuição de probabilidade que melhor se adequa aos estimadores disponíveis na literatura e implementados em pacotes computacionais é referente ao termo de erro dado por uma distribuição Normal, resultando em modelos *probit*. É importante ressaltar que o uso de modelos *logit* na literatura de econometria espacial é restrito em aplicações empíricas, tendo em vista que a derivação dos modelos de escolha discreta espaciais autorregressivos nesses casos é mais complexa (ANSELIN, 2002; CALDEIRA, 2022).

Além dos modelos já apresentados, outros formatos surgem pela combinação dos parâmetros espaciais como o *Spatial Autoregressive Autoregressive (SARAR) probit* quando $\theta = 0$ (Billé e Arbia, 2019). Os autores afirmam que para a estimação dos modelos, o vetor θ pode ser utilizado em conjunto com cada um dos parâmetros λ e ρ , bem como em uma combinação de ambos.

Assumindo distribuição Normal para os termos de erro, os modelos de escolha discreta *probit* multinomiais são especificados pelas equações (8), (9) e (10) (BILLÉ e ARBIA, 2019).

$$U^*_{ij} = \rho \sum_{h=1}^n w_{1,ih} U^*_{hj} + \mathbf{x}'_{ij} \boldsymbol{\beta} + \sum_{h=1}^n w_{2,ih} \tilde{\mathbf{x}}'_{hj} \boldsymbol{\theta} + u_{ij} \quad (8)$$

$$u_{ij} = \lambda \sum_{k=1}^J m_{jk} u_{ik} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2_{\varepsilon_j}) \quad (9)$$

$$y_{ij} = 1 \text{ se } U^*_{ij} > U^*_{ik} \text{ para } k \neq j \quad (10)$$

$$y_{ij} = 0 \text{ caso contrário}$$

em que $i, h = 1, \dots, n$ são as observações;

$j, k = 1, \dots, J$ são as alternativas do conjunto de escolha em que a j -ésima alternativa é escolhida se sua utilidade é a maior em relação a todas as outras alternativas;

U^*_{ij} é a utilidade (ou preferência) do indivíduo i para a alternativa j ;

\mathbf{x}_{ij} são os regressores exógenos específicos do indivíduo com parâmetros $\boldsymbol{\beta}$;

y_{ij} é a variável multinomial observada correspondente;

$\rho \sum_{h=1}^n w_{1,ih} U^*_{hj}$ sumariza a estrutura de dependência entre a preferência de indivíduos;

$\lambda \sum_{k=1}^J m_{jk} u_{ik}$ sumariza a estrutura de dependência entre atributos não observados e alternativas selecionadas;

$\sum_{h=1}^n w_{2,ih} \tilde{\mathbf{x}}'_{hj} \boldsymbol{\theta}$ sumariza os efeitos de *spillovers* locais específicos do indivíduo; e

$\sigma^2_{\varepsilon_j} = 1$ para identificação do modelo.

Billé e Arbia (2019) indicam que a avaliação deste tipo de modelo é interessante quando existem indícios de correlação espacial entre as alternativas, além de permitir identificar se os atributos não observados relacionados aos indivíduos podem estar correlacionados espacialmente. Os modelos multinomiais do tipo *probit* não possuem uma forma fechada para estimação da função de probabilidade, dificultando ainda mais a aplicação dos modelos *SAR* e *SEM* (TRAIN, 2009).

Apesar dos autores considerarem os termos \mathbf{x}_{ij} como específicos do indivíduo, quando se analisa o termo $\sum_{h=1}^n w_{2,ih} \tilde{\mathbf{x}}'_{hj} \boldsymbol{\theta}$ é possível obter efeitos de *spillovers* locais que são provenientes da dependência espacial intrínseca das variáveis relativas às alternativas, como o custo de viagem e o tempo de viagem, por exemplo. Já ao se analisar o parâmetro ρ , observa-se que ele é genérico às alternativas. Essa condição não permite que as alternativas modeladas sejam analisadas de forma separada quanto ao efeito espacial capturado por este parâmetro.

Existem evidências de uso menos recorrente do modelo *SLX* relacionado ao fato de sua especificação não gerar problemas comuns aos autorregressivos, como é o caso da heterocedasticidade e a necessidade da inversão de matrizes de grande porte múltiplas vezes (CALABRESE; ELKINK, 2014; SMIRNOV, 2010; VEGA; ELHORST, 2015). Como o modelo *SLX* não se enquadra nessas estimações mais complexas, o estimador de máxima verossimilhança de modelos de escolha discreta tradicionais é suficiente para a estimação dos parâmetros utilizando os softwares existentes.

Outras evidências indicadas por Billé e Arbia (2019) é que os modelos *SAR* e *SEM* são pioneiros e, por isso, existem testes de especificação como o de Multiplicador de Lagrange, justificando a menor utilização do *SLX* (ELHORST; VEGA, 2017; VEGA; ELHORST, 2015). Em contrapartida, Ruttenauer (2019) demonstrou por meio de simulação *Markov Chain Monte Carlo* (MCMC) que os modelos *SLX* e *SDM* – *Spatial Durbin Model* podem ter desempenho

melhor que os autorregressivos no contexto linear. Além disso indica que existe um padrão indiscriminado de uso dos modelos *SAR* e *SEM* em trabalhos empíricos, sem justificativas teóricas que indiquem o uso desses modelos. No caso dos modelos lineares, a não consideração da dependência nas variáveis independentes (*SLX*) pode ser problemática pela omissão de variáveis relevantes para explicação dos modelos, o que torna os estimadores ineficientes e inconstantes.

O trabalho de Gibbons e Overman (2012) indica que os estudos com aplicações empíricas devem considerar primeiro especificações do tipo *SLX*, condizente com a formulação teórica apresentada na inversão da matriz, em que o maior efeito espacial está no primeiro termo que justamente é a parcela referente ao modelo *SLX*. Se o efeito espacial não for grande o suficiente ele não se propagará para as demais parcelas, que são justamente as que indicam a vizinhança de demais ordens e, por consequência, os modelos autorregressivos.

2.2.2 Especificações de efeitos espaciais

A econometria espacial surgiu pela intenção de identificar a presença de processos espaciais nas diferentes áreas das ciências regionais. Alguns dos pressupostos assumidos pelos modelos econométricos utilizados nas ciências sociais é a independência entre regiões e indivíduos. Quando o espaço é uma das dimensões de análise do modelo, é difícil assumir a validade da premissa de independência das observações. Assim, se a dependência espacial não é controlada, os parâmetros estimados tendem a ser viesados.

Dessa forma, de acordo com Anselin (1988), a dependência espacial pode ser entendida como “a existência do relacionamento funcional entre o que acontece em um ponto no espaço e o que acontece em qualquer outro”. Ainda, problemas de mensuração como a delimitação arbitrária, ou problemas de agregação de unidades espaciais, podem gerar dependência espacial.

A dependência espacial também pode ser resultado de diferentes fenômenos de interações espaciais. Um tipo específico dessas interações é o efeito do tipo *spillover*, que decorre do transbordamento do efeito de uma variável no espaço afetando direta ou indiretamente os vizinhos. Esse elemento é fundamental para a atribuição do espaço como elemento para a explicação do comportamento humano (ANSELIN, 1988).

Assim, entender as relações que se estabelecem nas unidades de análise por meio da dependência espacial é fundamental. Anselin (1988) indica que é possível identificar esses

processos por meio de noções topológicas de vizinhança ou de vizinho mais próximo. Após definir os vizinhos de uma observação, é possível obter o conjunto de observações e as relações entre eles através de uma matriz de vizinhança ou matriz de conectividade.

Anselin (2003) também pontua que a dependência espacial nos modelos econométricos é diretamente observada na matriz de variância-covariância, em que os resultados observados para a covariância são diferentes de zero, além das variâncias não serem constantes nos denominados modelos autorregressivos. A solução para este problema, segundo o autor, está em controlar a dependência espacial por meio da matriz de variância-covariância utilizando uma função de decaimento da distância, ou uma função de passo sob forma não paramétrica a exemplo dos procedimentos executados na Geoestatística.

Outra forma de indicar a estrutura espacial é por meio da matriz de pesos espaciais utilizada na Econometria Espacial. Neste formato a estrutura é capturada de forma indireta e um processo estocástico espacial é especificado de forma a relacionar o valor de uma variável aleatória em uma localização aos valores da variável utilizada em suas localizações vizinhas (CALDEIRA, 2022).

Dois tipos de efeitos espaciais podem surgir pela especificação de um modelo espacial, os de impacto global e os de impacto local. Os efeitos globais são provenientes de um processo espacial que gera um efeito de interação que se distribui sobre todas as localidades da região de análise, enquanto os efeitos locais se estabelecem apenas sobre as localidades vizinhas a observação. A especificação do modelo linear com dependência espacial nos termos de erro (*SEM*) representado pela equação (6) pode ser simplificada pela formulação da equação (11), considerando $\mathbf{M}_n = W$.

$$u = \lambda W u + \varepsilon \quad (11)$$

Assim, isolando a equação para o termo de erro espacialmente correlacionado, obtém-se a equação (12):

$$u = [I - \lambda W]^{-1} \varepsilon \quad (12)$$

Anselin (2003) discute que a matriz inversa expressa na equação anterior pode ser obtida por uma série de expansão $[I - \lambda W]^{-1} = I + \lambda W + \lambda^2 W^2 + \dots$. O termo W elevado à potência de segunda ordem indica matrizes de vizinhança de ordem dois (W^2), ou seja, consideram que os vizinhos da observação são os vizinhos dos vizinhos. A partir da análise da formulação, os

efeitos espaciais se propagam nas vizinhanças mais próximas de uma dada observação, e vai diminuindo para as demais localidades componentes da região de análise (com λ sendo uma constante que varia entre 0 e 1). Essa formulação se relaciona com a primeira lei da geografia de Tobler (1970) em que “todas as coisas estão relacionadas, mas coisas mais próximas estão mais relacionadas entre si que as mais distantes”.

Outra evidência destacada por Anselin (2003) é sobre o comportamento dos termos de variância representados pelos valores da diagonal principal na matriz de variância-covariância $[I - \lambda W]^{-1}$. Em situações em que a estrutura de vizinhança não é constante em todas as observações, como ocorre na maior parte dos casos, os termos de erro são heterocedásticos.

Ao se observar os efeitos locais, a percepção do efeito nas localidades diretamente vizinhas é mais simples pelo impacto limitado do fenômeno. Pelo processo de especificação é possível caracterizar esse efeito de acordo com a equação (13), em que γ é o parâmetro espacial (ANSELIN, 2003).

$$u = \gamma W \varepsilon + \varepsilon \quad (13)$$

Assim, o efeito espacial se aplica diretamente nos vizinhos de primeira ordem garantido pelo termo W . Essa forma reduzida da equação não exige a inversão da matriz de vizinhança, tal que a condição a matriz de variância-covariância é dada pela equação (14).

$$E[\varepsilon \varepsilon'] = \sigma^2 [(I + \gamma W)(I + \gamma W)'] = \sigma^2 [I + \gamma(W + W') + \gamma^2 W W'] \quad (14)$$

Por meio dessa equação é possível indicar que os elementos diferentes de zero fora da diagonal principal da matriz de variância-covariância são os que se correlacionam com os elementos diferentes de zero na matriz de vizinhança W (de forma equivalente W') e $W W'$. Assim, o efeito espacial se propaga até os vizinhos de segunda ordem, não se estabelecendo sobre as demais observações da área de estudo como ocorre no efeito global (ANSELIN, 2003).

A partir destas discussões, o campo da Econometria Espacial tem se estabelecido e apresentado análises empíricas nas diversas áreas das ciências sociais. Na Engenharia de Transportes, as primeiras discussões que consideraram o enviesamento dos parâmetros pela não consideração dos efeitos de dependência espacial em modelos de escolha discreta foi realizada por Goetzke (2003).

2.2.3 Matriz de vizinhança e pesos espaciais

De acordo com Anselin (2002), os efeitos de dependência espacial estão totalmente relacionados à definição das relações entre as observações e a intensidade de cada uma dessas relações. A matriz de vizinhança é uma matriz com dimensão $n \times n$, sendo n o tamanho da amostra em termos de número de indivíduos. Por meio das relações estabelecidas pela matriz de vizinhança são definidos os vizinhos de uma observação e os pesos espaciais que cada observação tem sobre os seus vizinhos.

O uso de matrizes de vizinhança tem recebido muitas críticas quanto à determinação porque na maioria dos estudos é considerada como conhecida (CORRADO; FINGLETON, 2012; GIBBONS; OVERMAN, 2012). Assim, assume-se que as relações e suas intensidades são exógenas, ou seja, não influenciam o fenômeno simultaneamente. Pela característica autorregressiva dos modelos, a estimação dos parâmetros de relação entre os elementos no espaço se torna complexa, de forma que a maioria dos estudos assumem a caráter exógeno das relações espaciais. O modelo *SLX* se destaca pela possibilidade de estimação dos pesos da matriz espacial pelos atributos modelados (VEGA; ELHORST, 2015).

Os principais critérios de definição de vizinhança são os de contiguidade e os baseados em distância. Os critérios de contiguidade são muito utilizados nas ciências regionais. Pela própria área de análise estar relacionada às regiões, a contiguidade é mais simples. Por esse critério, uma região é considerada vizinha a outra se compartilha a mesma fronteira ou tem fronteiras em comum. Uma outra forma de definição de contiguidade é pelos critérios *queen* (rainha) e *rook* (torre), que se relacionam ao formato de movimentação das peças no jogo de xadrez. Da mesma forma são determinados os vizinhos por esses critérios. Aos elementos que são considerados vizinhos são atribuídos um valor igual a 1, e os que não são vizinhos recebem o valor igual a 0. A equação (15) apresenta esse formato de atribuição dos pesos espaciais.

$$w_{ij} = \begin{cases} 1, & \text{se } i \text{ e } j \text{ são contíguos} \\ 0, & \text{caso contrário} \end{cases} \quad (15)$$

em que w_{ij} é o peso do indivíduo j sobre o indivíduo i .

Outra forma de definição é pelo critério de distância, em que é estabelecida a relação pela distância inversa entre as observações dentro de uma distância de corte conforme a

Equação (16). Assim, quanto maior a distância, menor é o peso espacial atribuído à observação de acordo com a lei de Tobler.

$$w_{ij}(d_{ij}) = \begin{cases} 1/d_{ij}^\alpha, & \text{se } d_{ij} \leq \bar{d} \\ 0, & \text{se } d_{ij} > \bar{d} \end{cases} \quad (16)$$

em que $w_{ij}(d_{ij})$ é o $\langle i, j \rangle$ -ésimo elemento da matriz de vizinhança W ;

d_{ij} é a distância entre as observações i e j ;

\bar{d} é a distância de corte na qual um indivíduo é considerado vizinho; e

α é um parâmetro que define a intensidade da relação entre as observações, assim quanto maior seu valor, menor a intensidade da relação entre as observações.

O processo de definição da matriz de vizinhança pelo critério de distância pode ainda ser realizado de duas outras formas, a primeira estabelece que todas as observações estão relacionadas e todos os indivíduos na amostra são definidos como vizinhos, enquanto a segunda estabelece um valor limite de corte (*cutoff*) para que observações que estiverem dentro do intervalo estabelecido sejam consideradas vizinhas. As equações (17) e (18) indicam esses critérios.

$$w_{ij} = 1/d \quad (17)$$

$$w_{ij} = \begin{cases} 1, & d \leq d_{m\acute{a}x} \\ 0, & \text{caso contr\`ario} \end{cases} \quad (18)$$

em que w_{ij} é o $\langle i, j \rangle$ -ésimo elemento da matriz de vizinhança W ;

d é a distância entre as observações; e

$d_{m\acute{a}x}$ é a distância máxima que um indivíduo é considerado vizinho.

Existem outros formatos de critérios para a definição da relação de vizinhança, entre eles o de distâncias sociais relacionado a estudos de influência social (MANESS; CIRILLO; DUGUNDJI, 2015; SIDHARTHAN *et al.*, 2011). Quanto à definição dos pesos espaciais, um procedimento comum é a padronização dos pesos por linhas, em que as matrizes binárias são obtidas por matrizes de pesos espaciais que indicam o efeito de cada vizinho na observação pelo peso médio ponderado de todas as observações vizinhas.

A definição do critério de construção da matriz de vizinhança é uma das etapas para a obtenção de uma estimação viável pela econometria. Algumas propriedades são elementares

para a utilização destas matrizes conforme discutidas por Almeida (2012) e apresentadas a seguir:

- $0 \leq w_{ij} < \infty$: pesos não negativos e finitos, indicando que as relações entre duas regiões só fazem sentido dentro deste intervalo;
- $\sum_j w_{ij} \neq 0$: impede que na matriz de vizinhança existam observações sem vizinhos, além dessas observações isoladas tornarem a matriz singular impedindo a sua inversão;
- $w_{ii} = 0$: condição que a observação não pode ser vizinha de si mesma;
- $E(w_{ij}\varepsilon) = 0$: condição que os pesos são exógenos;

A segunda condição mencionada não é necessária nos modelos em que a dependência espacial está incorporada nas variáveis independentes, pois a inversão da matriz de vizinhança não é requerida. Em situações em que as observações representem indivíduos, assume-se uma observação isolada pela ausência de dados de seus vizinhos e esta é retirada das análises.

2.2.4 Estimadores

Pela complexidade relacionada à estimação dos modelos de escolha discreta espaciais, como problemas com endogeneidade e a necessidade de inversão da matriz espacial, um grande esforço computacional é necessário para determinar os valores dos parâmetros associados aos atributos (MCMILLEN, 1992; SMIRNOV, 2010). Segundo Santos (2020), três formas de estimação são possíveis pelos seguintes métodos: Métodos de Máxima Verossimilhança (*MVs*); Métodos Bayesianos; e Método Generalizado de Momentos (*Generalized Method of Moments* - *GMM* em inglês).

O primeiro estimador entre os *MVs* foi proposto por McMillen (1992), o *Expectation-Maximization (EM)* utilizado para os modelos de autorregressivos. Além desse, os estimadores *Recursive Importance Sampling (RIS)* (BERON; VIJVERBERG, 2004), *Gewere-Hajivassiliou-Keene (GHK)* (PACE; LESAGE, 2016) e o métodos *MV* parciais (BHAT, 2011; BILLÉ; LEORATO, 2020; WANG; IGLESIAS; WOOLDRIDGE, 2013) fazem uso de técnicas de simulação. Os métodos de aproximação se relacionam ao estimador *Composite Maximum Likelihood (CML)* para as questões da inversão da matriz e dos pesos espaciais (MARTINETTI; GENIAUX, 2017; MOOZHAROVSKY; VOGLER, 2016).

Pelos estimadores Bayesianos, LeSage (2000) propôs um estimador que utiliza a amostragem de *Gibbs*, ao passo que Zhou, Wang e Holguín-Veras (2016) propuseram uma forma para modelos *probit* espacial com matriz de vizinhança e pesos espaciais endógenos baseados no estimador de *Markov Chain Monte Carlo (MCMC)*.

Em relação aos estimadores do grupo *GMM*, a primeira estimação foi proposta para um modelo de escolha discreta do tipo *SEM* por Pinkse e Slade (1998), além de Klier e McMillen (2008) proporem uma linearização do estimador para modelos *logit*. No caso de modelos multinomiais, a proposição do estimador por *GMM* de Carrión-Flores, Flores-Lagunes e Guci (2018) é indicada como a mais adequada. Para modelos binários do tipo *SAR*, Calabrese e Elkink (2014) compararam os estimadores maximização da expectativa, amostragem de *Gibbs*, amostragem de importância recursiva, método generalizado de momentos, estimador de Pinkse e Slade, estimador de Klier e McMillen e simulação Monte Carlo, indicando suas vantagens e desvantagens.

2.2.5 Aplicações

Trabalhos empíricos desenvolvidos na temática de modelos de escolha discreta espacial estão concentrados em uso/parcelamento do solo, escolha de local de residência e problema de localização de firmas usando modelos do tipo *SAR* e *SEM* (CARRIÓN-FLORES; FLORES-LAGUNAS; GUCI, 2018; GUO; BHAT, 2004; MIYAMOTO *et al.*, 2004; WANG; KOCKELMAN; LEMP, 2012). Em relação aos modelos *SLX* foram identificados os trabalhos de Alamá-Sabater, Artal-Tur e Navarro-Azorín (2011) e Autant-Bernand (2006) na temática de localização de instalações industriais.

Em estudos da área de Engenharia de Transportes, o número de aplicações que consideram modelos com dependência espacial são consideráveis. As aplicações são principalmente em modelos lineares com dados do tipo transversal e em modelos com dados em painel, esses últimos em menor quantidade. As aplicações discutidas neste trabalho estão delimitadas aos estudos de modelos de escolha de discreta espaciais aplicados a problemas de escolha de modos de transportes.

Gotzke (2008) foi um dos pioneiros a explorar a questão de dependência espacial na análise de escolha de modo de transportes pela estimativa de um modelo *logit* binário espacial, para determinar os efeitos sociais de rede na escolha de transporte público em viagens a trabalho

na cidade de Nova York, EUA. A matriz de vizinhança utilizada foi definida pelos critérios de k -vizinhos mais próximos e o de distância. Os parâmetros para a definição da matriz foi considerar como vizinhos os 40 indivíduos mais próximos ou aqueles que estivessem a até 1,2 km de distância do indivíduo.

O autor estimou três modelos *logit* pelo método *MV*: sem considerar o efeito espacial; considerando o efeito espacial pelo termo ($W\gamma$); e acrescido de uma variável *dummy* para controlar o acesso de automóveis no centro da cidade. Os resultados indicaram que existem efeitos de dependências espacial nos dois últimos modelos espaciais, com parâmetros de sinal positivo e estatisticamente significativos, e que estes dois modelos apresentam melhor desempenho que o primeiro modelo não espacial em termos da log-verossimilhança (GOETZKE, 2008).

Sidharthan *et al.* (2011) estimaram um modelo *probit* multinomial espacial do tipo *SAR* para escolha de modos de transporte no contexto de viagens de crianças à escola na cidade de São Francisco, EUA. O modelo foi estimado considerando efeitos de interação espacial e social, sendo a matriz de vizinhança definida por critérios geográficos e sociais. Para a parte geográfica foram analisados diferentes formatos de matrizes espaciais, de distância inversa, de distância inversa exponencial e de distância e pesos binários. A determinação do critério social foi por meio de distância social, sendo determinada pela proximidade demográfica relacionada às variáveis renda e idade.

Os resultados indicaram dependência espacial positiva e estatisticamente significativa para os efeitos de interação espacial, mas não para os efeitos sociais. Entre as matrizes analisadas, a que apresentou melhor desempenho em termos da log-verossimilhança foi a de distância inversa. A estimação dos modelos foi pelo estimador *Maximum Approximate Composite Marginal Likelihood (MACML)* de Bhat (2011), o estimador para modelos de escolha discreta multinomiais autorregressivos espaciais mais disseminado na literatura. Os resultados da modelagem indicaram que existe uma dependência espacial na escolha de modos de transporte em viagens realizadas por crianças com destino à escola. Como consequência dessa dependência a propensão de uma criança escolher um determinado modo influencia na escolha de outras crianças para a escolha do mesmo modo. Um outro indício da importância de analisar a presença de dependência espacial é a indicação de que, em termos do valor de log-verossimilhança, o modelo espacial apresenta melhor desempenho comparado ao não espacial.

O estudo desenvolvido por Wang *et al.* (2015) propôs identificar a escolha de bicicleta em viagens realizadas por estudantes e funcionários da Universidade do Estado de Ohio, EUA, por meio de um modelo *probit* binário do tipo *SAR*. Para a definição da matriz de vizinhança foram utilizados dois critérios de distância com valor de corte igual a 1.000 metros, por pesos binários e por pesos espaciais iguais à distância inversa. O processo de estimação foi pelo estimador bayesiano *MCMC* utilizando simulação. O parâmetro espacial foi positivo, evidenciando a existência do efeito de dependência no uso da bicicleta. Essa condição estabelece que a propensão ao uso do modo bicicleta pelo indivíduo influencia seus vizinhos a tomarem a mesma decisão.

No contexto de viagens a trabalho e efeitos de dependência espacial, Sheng e Sharp (2019) estimaram um modelo *probit* binário do tipo *SAR* para analisar as viagens realizadas na região metropolitana de Auckland na Nova Zelândia, em que as alternativas consideradas foram transporte público e transporte privado. A matriz de vizinhança foi determinada pelo critério de distância com corte igual a 1.000 metros, sendo os pesos definidos pela distância inversa e a matriz padronizada por linha. O parâmetro espacial estimado apresentou sinal positivo e estatisticamente significativo, o que indica que a escolha do modo transporte coletivo por um indivíduo influencia os seus vizinhos a realizarem a mesma escolha.

No contexto do Brasil, foram encontrados três trabalhos que consideraram os efeitos de dependência espacial nos modelos de escolha de modo de transportes. O trabalho desenvolvido por Assirati (2018) utilizou o indicador *SivarG* para incluir a dependência espacial nas análises de probabilidade de pertencimento a grupos com um determinado comportamento por meio de clusterização *k-means*. Pela incorporação das variáveis espaciais, o modelo obteve melhores resultados de desempenho ao se comparar ao modelo não espacial. Lopes, Brondino e Silva (2014) estimaram modelos de geração de viagens considerando a dependência espacial ao estimar modelos do tipo *SAR* e *SEM*. Esses modelos apresentaram melhor desempenho comparados aos não espaciais.

O terceiro trabalho foi o de Caldeira (2022), que investigou os efeitos de dependência espacial na escolha de modos de transporte para viagens a trabalho no município de São Paulo por meio da estimação de um modelo *SLX*. O autor estimou as matrizes de vizinhança considerando dois critérios de vizinhança, a primeira relacionada à distância linear entre a observação e seus vizinhos com pesos binários e a segunda com base no *k*-ésimo vizinho mais próximo utilizando pesos binários.

O critério de corte analisado pelo autor para o limite de distância foi para variações entre 1.000 metros e 2.000 metros, variando em intervalos de 100 metros, sendo selecionado o limite igual a 1.000 metros. Com relação ao critério de k vizinhos mais próximos, o número de vizinhos variou entre 10 e 100 em intervalos de 10 vizinhos, sendo selecionado k igual a 40.

Foram estimados os modelos não espaciais e os espaciais do tipo *SLX*, comparando o desempenho pela abordagem da Teoria da Utilidade (*RUM*) e pela Teoria do Arrependimento (*Random Regret Minimization, RRM*). Os modelos não espaciais e espaciais estimados foram do tipo *logit*, apenas com os atributos dos modos e considerando esses atributos além daqueles que descrevem as características sociodemográficas dos indivíduos. A estimação foi pelo método da máxima verossimilhança.

Os modelos estimados pela abordagem *RRM* apresentaram melhores resultados em termos de log-verossimilhança, e os que consideraram os efeitos de dependência espacial tiveram ainda melhores desempenho. As variáveis dos atributos dos modos foram defasadas espacialmente e identificou-se a superestimação de parâmetros relacionados à alternativa automóvel, além dos atributos relacionados ao transporte público resultarem elásticos.

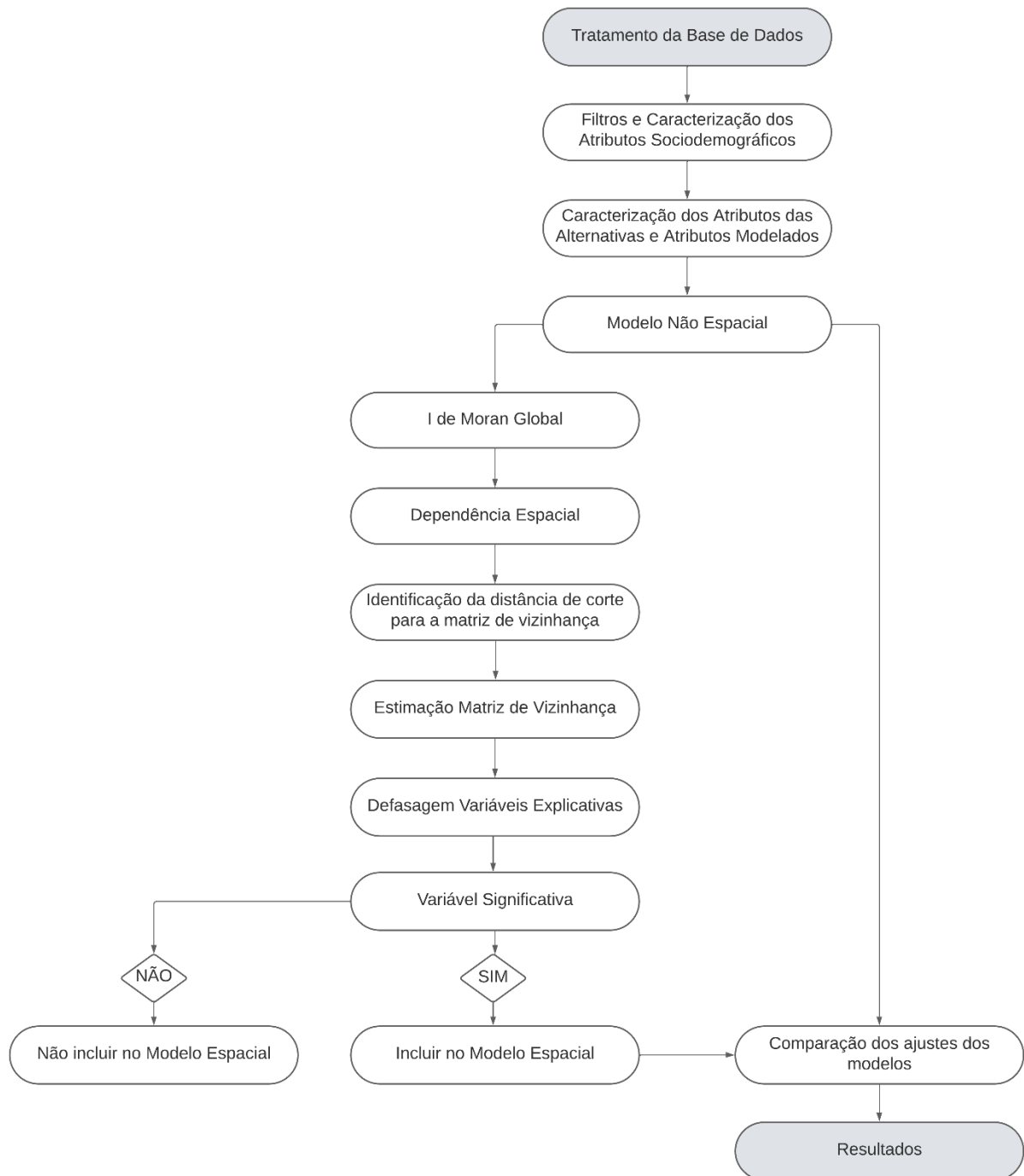
Além das pesquisas citadas, foi identificado o trabalho de Neves, Silva e Arruda (2021) já discutido anteriormente neste capítulo. Este estudo foi o único encontrado na literatura que investigou a existência de dependência espacial na análise de escolha do modo a pé, e os demais citados investigaram esses efeitos espaciais quanto aos demais modos. Portanto, ressalta-se a pertinência da presente pesquisa ao investigar especificamente os efeitos espaciais nas escolhas pelo modo a pé em viagens urbanas.

O estado da arte na literatura de Engenharia de Transportes quanto à escolha do modo a pé e à análise da dependência espacial evidenciam a escassez de pesquisas específicas sobre a dependência espacial em decisões sobre viagens pelo modo a pé, juntamente com limitações na definição de variáveis espaciais relevantes para a escolha desse modo. Adicionalmente, nota-se a necessidade de uma abordagem mais detalhada na análise dos aspectos sociodemográficos que exercem influência na dependência espacial. Diante dessas lacunas identificadas, esta pesquisa propõe-se a contribuir para a literatura ao identificar atributos espaciais, incluindo fatores sociodemográficos, que influenciam a escolha do modo a pé. Além disso, busca-se analisar a dependência espacial presente no contexto de viagens a pé com motivação relacionada a trabalho e estudo.

3. MÉTODO

Nesse capítulo é apresentado o processo de seleção e tratamento dos dados, além da caracterização dos atributos das alternativas e da modelagem pra atingir o objetivo da pesquisa. A esquematização do método é apresentada na Figura 1.

Figura 1 - Fluxograma esquemático do método

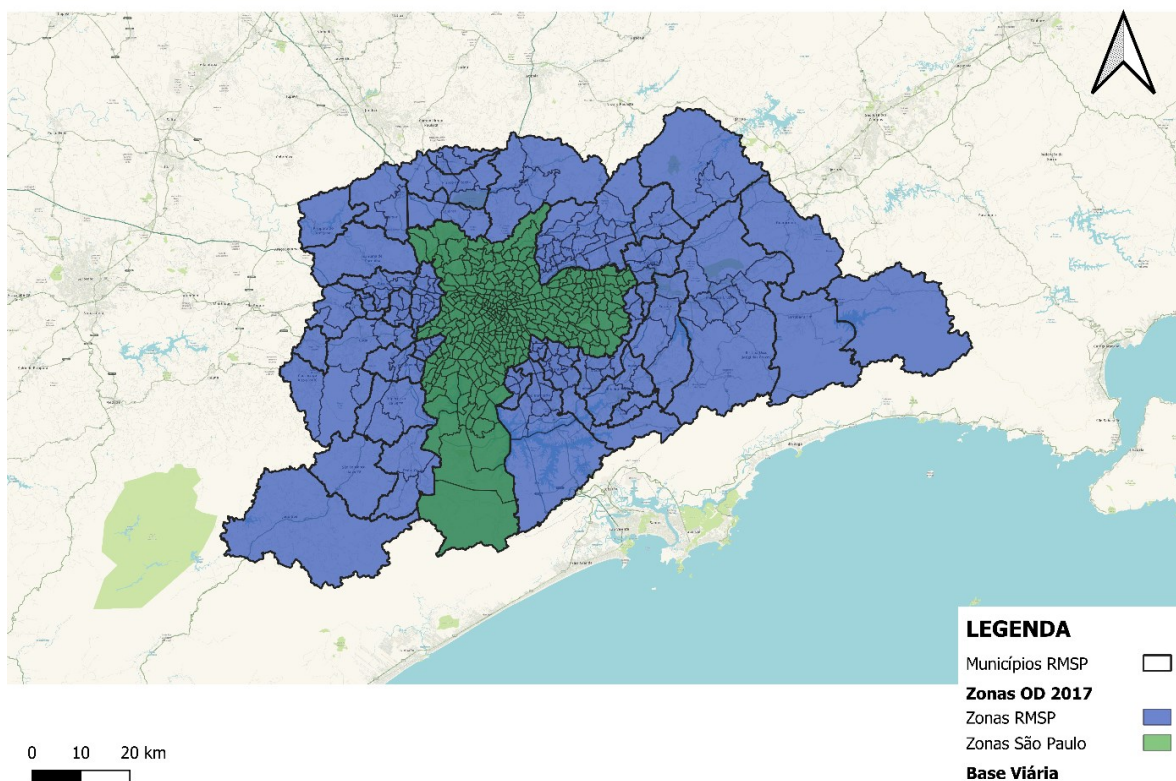


3.1 BASE DE DADOS

A base de dados utilizada nessa pesquisa é proveniente da Pesquisa Domiciliar Origem e Destino 2017 (Pesquisa OD 2017) realizada pela Companhia do Metropolitano de São Paulo (Metrô-SP), empresa que administra o sistema metroviário da cidade de São Paulo. A pesquisa buscou identificar as características dos deslocamentos no espaço urbano da Região Metropolitana de São Paulo (RMSP). As pesquisas acontecem a cada 10 anos, tendo sido a primeira realizada em 1967 e a mais recente em 2017/2018 (METRÔ, 2022).

A Pesquisa OD 2017 tem formato de coleta de dados em preferência revelada (PR), em que são coletadas as informações dos deslocamentos dos entrevistados em um dia típico da semana anterior ao de realização da entrevista. Além das informações que caracterizam as viagens, são coletadas informações socioeconômicas do domicílio e sociodemográficas do indivíduo para a caracterização da amostra. A RMSP é composta de 39 municípios, divididos em 517 zonas de tráfego conforme ilustrado na Figura 2. Apenas o município sede da região, São Paulo, possui 342 zonas.

Figura 2 - Zonas de Tráfego da pesquisa OD 2017 na RMSP



Fonte: Autor (2023)

Algumas limitações são encontradas para a estimação dos modelos propostos nesta pesquisa. Assim, para eliminar observações que inviabilizariam a estimação ou que não eram realistas, foram executados os procedimentos apresentados nas seções subsequentes deste documento. Um resumo das etapas desses procedimentos é apresentado na Figura 3.

Figura 3 - Procedimento para tratamento do banco de dados



Fonte: Autor (2023)

3.1.1 Caracterização da pesquisa OD 2017

A pesquisa foi realizada de forma domiciliar por amostragem aleatória estratificada da população da RMSP. Foram entrevistados 32 mil domicílios e 103 mil pessoas. Essa amostra não é um censo populacional, sendo que um fator de expansão foi calculado para as viagens.

Por esse procedimento, das 183.092 viagens caracterizadas pela realização da Pesquisa OD 2017, foram determinadas cerca de 42 milhões de viagens na RMSP em um dia útil da semana.

3.1.2 Seleção de viagens no município de São Paulo

Este filtro teve como propósito selecionar as viagens que ocorreram dentro do limite municipal, ou seja, aquelas que começaram e terminaram na cidade de São Paulo. A utilização desse procedimento se fez necessária pela impossibilidade de caracterização dos tempos de viagem fora do limite municipal em função das bases de dados disponíveis.

No caso da Pesquisa OD2017, o número de observações dentro do limite municipal de São Paulo é maior que nas demais localidades analisadas pelo formato de amostragem para coleta dos dados, o que colabora para uma maior quantidade de observações selecionadas por esse filtro. Sabe-se que, para o planejamento de sistemas de transportes urbanos, a consideração da escala metropolitana é importante, mas para o objetivo de análise deste estudo acredita-se que esta limitação espacial não gerou prejuízos para a discussão dos resultados.

3.1.3 Filtro por motivo de viagem

Foi realizado um filtro para obter as viagens por motivo estudo e trabalho, as quais correspondem às seguintes categorias na base de dados da Pesquisa OD 2017: “Escola/Educação”, “Trabalho Indústria”, “Trabalho Comércio” e “Trabalho Serviços”. Por este filtro as viagens com os demais propósitos foram eliminadas, a exemplo das viagens para compras, saúde e outros motivos.

3.1.4. Definição das alternativas

A etapa de determinação do conjunto de alternativas disponíveis para a escolha de modo é fundamental para a estimação dos modelos de escolha discreta. Nos modelos foram consideradas duas alternativas, o modo a pé (*WALK*) e os demais modos (*NO WALK*) que englobam os modos Ônibus (“Ônibus/micro-ônibus/perua do município de São Paulo”, “Ônibus/micro-ônibus/perua de outros municípios” e “Ônibus/micro-ônibus/perua metropolitano”), Trilhos (“Metrô”, “Trem” e “Monotrilho”), Fretado (“Transporte Fretado” e

“Transporte Escolar”), Automóvel (“Dirigindo Automóvel” e “Passageiro de Automóvel”), Táxi (“Táxi Convencional” e “Táxi não Convencional”), Moto (“Dirigindo Moto” e “Passageiro de Moto”), Bicicleta e Outros.

Outro fator importante é que a identificação do modo escolhido pelo indivíduo foi realizada pela variável “Modo Principal”, que entre os modos utilizados ao longo da viagem representa aquele de maior capacidade. Como exemplo, uma pessoa que utiliza os modos “a pé”, “ônibus” e “metrô” é classificada como tendo realizado a viagem pelo modo metrô.

Viagens com distâncias maiores do que 4.000 metros em que o indivíduo tinha escolhido o modo a pé na Pesquisa OD 2017 foram eliminadas, tendo em vista que a partir desta distância pode-se desconsiderar a caminhada como uma opção de modo segundo Vale *et al.* (2018). Uma limitação é introduzida ao restringir a classificação apenas aos modos *WALK* e *NOWALK*, pois a caracterização dos atributos associados ao *NOWALK* pode não ser suficientemente abrangente. Isso resulta na perda de informações relacionadas aos atributos dos níveis de serviço dos outros modos de transporte. Devido a essa condição, torna-se desafiador comparar a magnitude do impacto na escolha do modo *WALK* em relação, por exemplo, ao modo ônibus, diante dos resultados obtidos neste estudo.

3.1.5 Obtenção de dados sociodemográficos

O filtro referente aos atributos sociodemográficos consistiu em estabelecer um limite mínimo para as variáveis idade e renda familiar. O limite inferior para a variável idade foi de 16 anos, que se justifica pela idade mínima permitida para jovens trabalharem. Para a renda familiar, valores iguais a zero foram desconsiderados, já que indivíduos com emprego têm ao menos uma fonte de renda.

Para os atributos sociodemográficos idade, renda familiar, sexo, nível de escolaridade, número total de viagens do indivíduo e número de moradores do domicílio foram removidas observações com valores faltantes para cada atributo na base de dados.

3.1.6 Eliminação de observações duplicadas

Dadas as limitações das técnicas de Econometria Espacial consideradas neste estudo, se fez necessário o processo de eliminação das observações duplicadas. A Pesquisa OD 2017 coletou informações das viagens realizadas em um dia útil típico tal que uma pessoa pode ter realizado mais de uma viagem a trabalho, por exemplo, quando se considera que está envolvida no setor de serviços. Assim, esses indivíduos apresentam mais de uma observação no banco de dados.

No procedimento de modelagem de escolha discreta tradicional, esta propriedade não é um problema, mas quando se consideram as posições geográficas das observações, os indivíduos que têm mais de uma viagem por motivo ou se mais de um indivíduo de uma mesma residência respondeu a pesquisa, estes apresentarão mais de uma observação para a localização.

Um dos princípios definidos por Anselin (1988) é que um indivíduo não é vizinho de si mesmo e apenas uma viagem de um mesmo indivíduo pode ser considerada na análise. Assim, observações duplicadas foram eliminadas da base de dados para modelagem, ou seja, viagens pelo mesmo motivo e mesmo local de origem por indivíduo foram eliminadas por sorteio, mantendo apenas uma no banco de dados modelado por restrição dos modelos espaciais analisados.

3.1.7 Obtenção da distância mínima para matriz de vizinhança

Valores de corte devem ser definidos para a estimação da matriz de vizinhança pelo critério de distância entre as observações. Diferente do observado recorrentemente na literatura (WANG *et al.*, 2015; SHENG e SHARP; 2019; CALDEIRA, 2022), ao analisar o fenômeno de viagens a pé a distância de corte para a definição da matriz de vizinhança deve ser menor que 1.000 metros (NEVES, 2019). Assim, foram investigadas distâncias de corte pela aplicação do *I* de Moran na variável dependente considerando matrizes de vizinhança construídas a partir dos diferentes pontos de corte entre 50 e 4.000 metros. Com o objetivo de garantir que todas as observações tivessem ao menos um vizinho para uma distância mínima de 50 metros – distância mínima inicial para análise da dependência espacial deste fenômeno –, foram eliminadas as observações que não se enquadravam neste critério.

Assim, após a execução dos filtros descritos nas subseções anteriores, foi possível obter a base de dados final para modelagem de escolha discreta espacial proposta nesta pesquisa. A partir das 2.198 observações do banco de dados obtido até esta etapa, foram identificadas

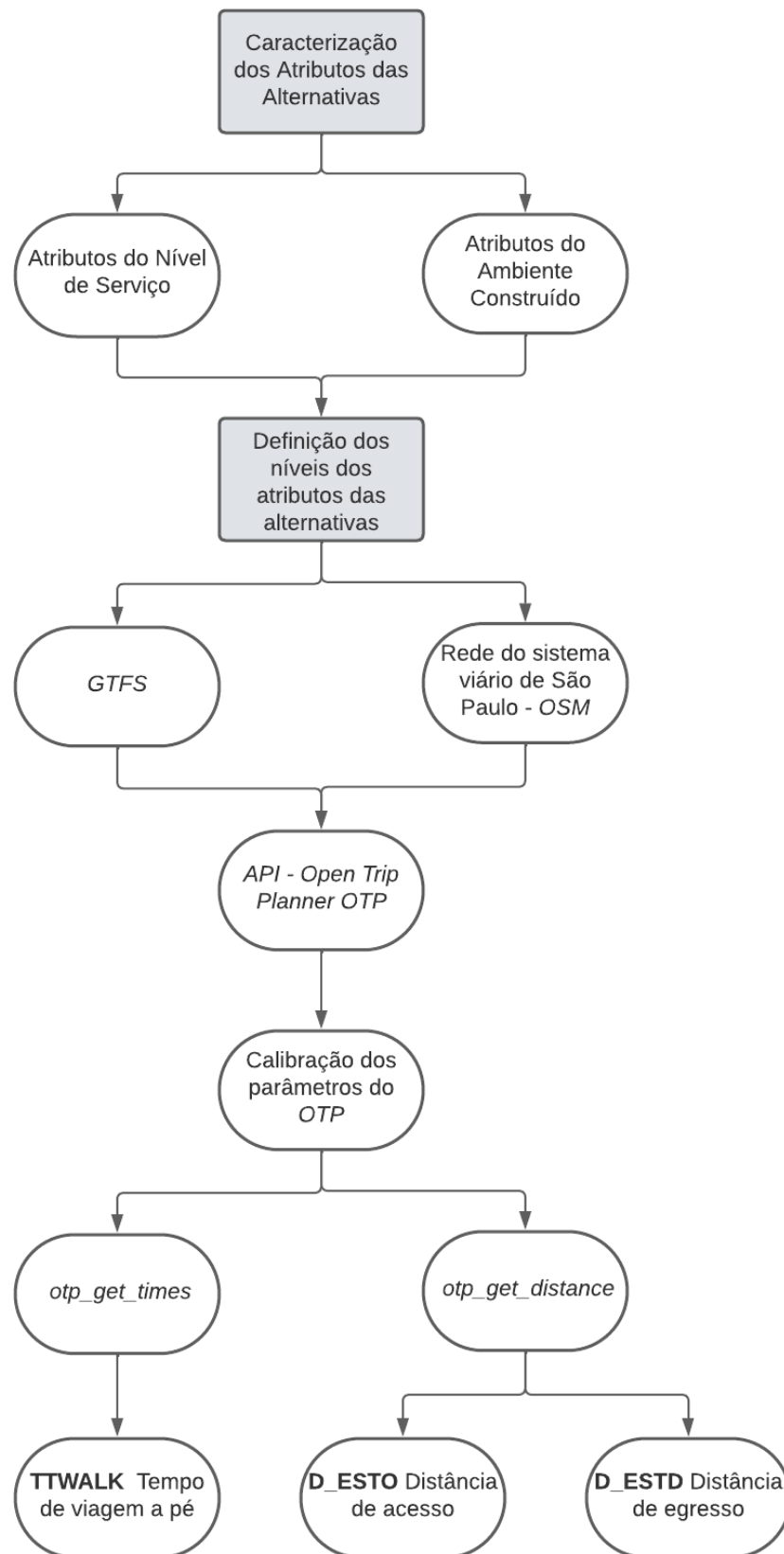
somente seis observações que resultaram em indivíduos sem nenhum vizinho com distância mínima de 50 metros entre si, resultando em um banco de dados final com 2.192 observações representando 232.646 viagens pela expansão das viagens caracterizadas (1,20% do banco de dados inicial).

3.2 CARACTERIZAÇÃO DOS ATRIBUTOS DAS ALTERNATIVAS

Como o formato da Pesquisa OD 2017 dispõe de informações apenas dos modos de transporte utilizados pelo entrevistado na viagem do dia anterior, os atributos das demais alternativas disponíveis, mas não escolhidas para a realização da viagem devem ser modeladas. Apesar da Pesquisa OD 2017 indicar valores para os atributos das alternativas escolhidas pelos indivíduos, tais informações não podem ser utilizadas nos modelos de escolha discreta pois referem-se à percepção dos indivíduos e geralmente não representam a realidade. Assim, Koppelman e Bhat (2006) recomendam modelar todos os atributos relacionados a todas as alternativas incluindo as escolhidas pelos indivíduos, sendo necessário estabelecer as alternativas disponíveis aos indivíduos conforme regras de acesso ao modo de transporte, como a distância máxima de viagens pelo modo a pé.

Nesta pesquisa, os atributos utilizados para caracterizar a alternativa que representa o modo a pé (*WALK*) foram: tempo de viagem, distância de acesso ao sistema de transporte público e distância de egresso ao sistema de transporte público, que se relacionam ao ambiente construído, sendo relacionados na literatura como explicativos da escolha do modo a pé (OLOJEDE, YOADE e OLUFEMI, 2017; NEVES, 2019; TON *et al.*, 2019; NEVES *et al.*, 2021). A caracterização dos atributos é apresentada nas próximas seções conforme procedimento apresentado na Figura 4.

Figura 4 - Etapas da caracterização dos atributos das alternativas



Fonte: Autor (2023)

3.2.1 Níveis dos atributos das alternativas

As Interfaces de Programação de Aplicações (*API – Application Programming Interface*) são um conjunto de rotinas de programação para acesso a um aplicativo de software ou plataforma baseada na internet (GRUS, 2016). Essas interfaces têm sido utilizadas para determinação dos níveis de atributos específicos de cada modo.

Neste estudo foi utilizada a *API* do *OTP (Open Trip Planner)*, um roteador multimodal desenvolvido em *Java* e com uso livre, capaz de determinar as rotas de menor caminho para o modo individual e as características da viagem pelos modos coletivos por meio de dados de *GTFIS (General Transit Feed Specification)*, tendo como base de dados uma rede do *Open Street Map (OSM, 2022)*. O *OTP* foi utilizado com o objetivo de determinar o tempo de viagem em rede pelo modo a pé, sendo que o mesmo permite a obtenção do tempo de viagem e a distância de viagem em rede. O *GTFIS* deve ser relacionado ao sistema de transporte da cidade ou região que se deseja analisar tendo sido utilizado o arquivo disponibilizado pela SPTrans (SPTrans, 2022).

O uso da ferramenta do *OTP* tem sido recorrente em análises de acessibilidade para o cálculo do tempo de viagem e em estudos de comportamento de viagens utilizando modelos de escolha discreta (MARCUS; BLAINEY, 2016; PEREIRA, 2018; PEREIRA *et al.*, 2019; PRITCHARD *et al.*, 2019; SLOVIC *et al.*, 2019). Neste trabalho, esta *API* foi utilizada para obter os tempos de viagem pelo modo a pé.

3.2.1.1 Distância de acesso/egresso

A partir dos dados da localização dos pontos de paradas e estações de metrô/trem obtidos pelas informações de *GTFIS* disponibilizados pela SPTrans, foi realizada uma operação espacial de “vizinho mais próximo” (*nearest object*) a fim de determinar as coordenadas dos pontos e estações de transporte público mais próximas das origens (local de residência) e destinos (local de trabalho ou estudo) das viagens observadas.

A distância determinada considera a rede urbana uma vez que este processo foi realizado pela *API* do *OTP*, permitindo determinar a menor distância entre estes locais e as origens ou destinos dos indivíduos.

3.2.1.2 Tempo de viagem *WALK*

O tempo de viagem para a alternativa *WALK* foi calculado utilizando o *OTP* considerando as informações das coordenadas das origens e dos destinos compatibilizadas com a rede viária disponível. Vale ressaltar que o *OTP* considera velocidade média de caminhada igual a 1,3 m/s o que corresponde à estabelecida por manuais como o *Highway Capacity Manual* (HCM, 2010).

3.3 RESUMO DOS ATRIBUTOS PARA MODELAGEM

Os atributos considerados para as análises e modelagem das alternativas foram estabelecidos com base em evidências da literatura abordada na seção anterior deste documento, referentes ao tempo de viagem (TV_W), distância de acesso (D_{acesso}) e egresso ($D_{egresso}$) ao transporte público para a alternativa *WALK*. Para a caracterização dos indivíduos foram considerados atributos socioeconômicos: idade (*IDADE*) dividido em três categorias (*IDADE0* – idade entre 16 e 30 anos como referência; *IDADE1* – idade entre 30 e 60 anos; *IDADE2* – idade maior que 60 anos), sexo (*SEXO*), com a categoria masculino como referência; número de moradores do domicílio (*MORADORES*); critério de renda (*CR*) dividido em três categorias como variável binária (*CR0* - critério Brasil de renda 1, mais alta, como referência; *CR1* – critérios Brasil de renda 2 e 3, médias; e *CR2* – critérios Brasil de renda 4, 5 e 6, baixas). Para o grau de instrução também foi utilizada uma variável binária com três classes (*GI0* - até Fundamental II incompleto como referência, *GI2* - Fundamental II completo até Superior incompleto e *GI3* - Superior completo). Também foram considerados os atributos QT_{AUTO} , QT_{MOTO} , e $QT_{BICICLE}$ relacionados à quantidade que o indivíduo possui de automóveis, motos e bicicletas, respectivamente. A última variável que descreve os indivíduos está relacionada à característica da viagem quanto ao período do dia, ou seja, viagens realizadas na hora pico da manhã (*PICO MANHÃ*) como variável binária (1 se realizada nesse período e 0 caso contrário). O resumo dos atributos considerados na modelagem é apresentado na Tabela 1.

Tabela 1 - Relação da sigla do atributo e sua descrição

Atributo	Descrição
TV_W	Tempo de viagem modo WALK
D_{acesso}	Distância de acesso ao transporte público
$D_{egresso}$	Distância de egresso do transporte público
$IDADE1$	Categoria de idade 1
$IDADE2$	Categoria de idade 2
$SEXO$	Sexo
$CR1$	Categoria do critério de renda 1
$CR2$	Categoria do critério de renda 2
$GI1$	Categoria do grau de instrução 1
$GI2$	Categoria do grau de instrução 2
$PICO\ MANH\tilde{A}$	Viagens que ocorrem na hora pico manhã
$MORADORES$	Número de moradores do domicílio
QT_{AUTO}	Quantidade de automóveis
QT_{MOTO}	Quantidade de motocicletas
QT_{BICILE}	Quantidade de bicicletas

Fonte: Autor (2023)

3.4 ESTATÍSTICAS ESPACIAIS

Para a identificação da dependência espacial é necessário realizar testes nos dados quanto às variáveis de interesse. Dessa forma, se utilizou o teste I de Moran univariado para a caracterização da dependência espacial. Como a variável de escolha é binária (*WALK/NOWALK*), foram calculadas as probabilidades de escolha do modelo não espacial como variável referência para o cálculo do I de Moran.

Tendo em vista o formato de análise, uma espacialização dos dados é necessária a fim de se adequar aos padrões da estatística espacial e ao processo de estimação dos modelos. Assim, ajustes nas coordenadas das observações foram necessários conforme o procedimento de filtragem aplicado na seção 3.1.6.

3.4.1 Matriz de vizinhança

Nas ciências regionais, o objeto de análise são regiões espaciais que dividem seus limites territoriais geralmente com outra região vizinha, e por essa razão o critério de contiguidade para definição de vizinhos para a matriz pode ser utilizado. Em modelos de escolha discreta os indivíduos são o objeto de estudo, os quais não possuem limites geográficos que compartilham com seus vizinhos. Assim, o critério para definição de matriz de vizinhança mais utilizado na literatura para esse fim é o de distância conforme a Equação (18), também utilizado nesta pesquisa.

Para verificar a robustez da matriz de vizinhança foram criadas dezenove matrizes pelo critério de distância máxima com *cutoffs* (d_{max}) entre 50 metros e 200 metros (em incrementos de 50 metros), entre 200 metros e 1.500 metros (em incrementos de 100 metros), 2.000 metros e 4.000 metros. Pelas análises dos estudos empíricos relacionados ao comportamento de viagens, o limite mais utilizado é de 1.000 metros, mas no que tange a viagens a pé o limite de 400 metros conforme indicado por Neves (2019). Foram considerados valores menores e maiores que essa referência a fim de comparar o desempenho dos modelos quanto ao fenômeno analisado conforme recomendado por Goetzke (2008).

Pelo critério de distância, a questão de algumas observações estarem isoladas não se torna um problema, já que com o filtro para a distância mínima todas as observações têm ao menos um vizinho com 50 metros. Este processo garante todas as propriedades necessárias para a matriz de pesos espaciais. Outra consideração importante é que a matriz é exógena, ou seja, as relações espaciais não se estabelecem a partir do fenômeno objeto de análise.

A matriz de vizinhança é o elemento que relaciona e explicita a dependência espacial presente nos dados em análise, sendo fundamental para capturar efeitos espaciais e a estimação de modelos espaciais consistentes. Sabe-se que diferentes matrizes podem ser estimadas pelos diferentes critérios de estimação disponíveis na literatura. Entretanto, dada a limitação de trabalhos sobre modelos de escolha discreta espaciais na literatura de comportamento de viagens que analisam o desempenho do modelo atrelado as diferentes formas de especificação da matriz, não se realizou variações entre matrizes estimadas por diferentes critérios. Assim, a mesma forma de especificação da matriz foi considerada para os diferentes atributos e modos.

3.5 MODELOS DE ESCOLHA DISCRETA

Um modelo de escolha discreta não espacial foi estimado considerando as alternativas e atributos relatados nas seções anteriores e diferentes modelos espaciais do tipo *SLX* com defasagem em alguns atributos foram estimados, buscam identificar a dependência espacial da escolha de modo a pé em relação às variáveis independentes.

3.5.1 Modelo *RUM* não espacial

A modelagem *RUM* não espacial por um modelo *logit* binomial considerou as funções utilidades indicadas nas equações (19) e (20).

$$\begin{aligned}
 U_{WALK} = & ASC_{WALK} + \beta_1 * TV_W + \beta_2 * D_{acesso} + \beta_3 * D_{egresso} + \beta_4 \\
 & * IDADE1 + \beta_5 * IDADE2 + \beta_6 * SEXO + \beta_7 * CR1 + \beta_8 \\
 & * CR2 + \beta_9 * GI1 + \beta_{10} * GI2 + \beta_{11} * PICO\ MANH\tilde{A} + \beta_{12} \quad (19) \\
 & * MORADORES + \beta_{13} * QT_{AUTO} + \beta_{14} * QT_{MOTO} + \beta_{15} \\
 & * QT_{BICICLE}
 \end{aligned}$$

$$U_{NOWALK} = 0 \quad (20)$$

Esse modelo resulta em informações importantes quanto ao comportamento relacionado a viagens. Porém, se houver dependência espacial nos dados analisados, os resultados dos parâmetros estimados podem estar comprometidos quanto à sua capacidade explicativa do fenômeno de forma completa.

3.5.2 Modelo *RUM* espacial *SLX*

O modelo *RUM* espacial do tipo *SLX* estabelecido pela formulação das equações (21) e (22) para a alternativa *WALK* foi utilizado para analisar a dependência espacial na escolha do modo a pé. Novamente, a alternativa *NOWALK* teve valor constante igual a zero.

$$U_{in} = \sum_{k=1}^K \beta_k x_{ikn} + \sum_{k=1}^K z_{ikn} \theta_k + \varepsilon_{in} \quad (21)$$

$$z_{ikn} = \left(\sum_{n=1}^N w_{pq} \right) x_{ikn} \quad (22)$$

em que U_{in} é a utilidade da alternativa i para o indivíduo n ;

x_{ikn} são os k atributos sob a premissa da abordagem *RUM* da alternativa i para o indivíduo n ;

z_{ikn} são os atributos defasados espacialmente sob a premissa da abordagem *RUM* da alternativa i para o atributo k para o indivíduo n ;

w_{pq} é o elemento $\langle p, q \rangle$ da matriz de pesos espaciais de dimensão $(n \times n)$;

β_k são parâmetros estimados sob a premissa da abordagem *RUM*;

θ_k são os parâmetros espaciais estimados sob a premissa da abordagem *RUM*;

ε_{in} o termo de erro aleatório IID que assume distribuição Valor Extremo do tipo I.

O modelo *SLX* considera os efeitos do tipo *spillovers* espaciais, em que os efeitos de dependência espacial estão relacionados as variáveis explicativas dos vizinhos, e com isso influenciam a escolha do modo de transporte. A facilidade de estimação pelo método de máxima verossimilhança desse modelo comparado ao *SAR* e *SEM* evita questões como heterocedasticidade no modelo e a necessidade de inversão da matriz de pesos espaciais, o que facilita o processo computacionalmente.

Ao se obter a estimação do parâmetro espacial e identificar que ele é significativo tem-se que a variável z influencia um indivíduo diretamente, e indiretamente os indivíduos na sua vizinhança. Outros atributos que em situação de dependência espacial podem influenciar a escolha são os relacionados as características sociodemográficas, em que condições de similaridades são relacionadas a dependência espacial positiva e dissimilaridades a dependência espacial negativa.

Os modelos foram estimados no intervalo de 200 a 600 metros, por ser neste intervalo os maiores valores de I de Moran por base do modelo de escolha discreta não espacial, sendo que a partir de 600 metros notou-se redução expressiva no valor para o índice.

3.6 SOFTWARE

Para o tratamento da base de dados, assim como a estimação dos modelos, o software estatístico utilizado nesta pesquisa foi o *R* (R CORE TEAM, 2022). Para a etapa de georreferenciamento e espacialização dos dados foi utilizado o pacote *sf* (PEBESMA, 2018). Para obtenção dos atributos das alternativas consideradas por meio do roteamento das viagens foi utilizada a *API* do *OTP* pelo pacote *otpr* (YOUNG, 2019). Para a realização dos procedimentos relacionados às estatísticas espaciais e a construção das matrizes de vizinhança foi utilizado o pacote *spdep* (BIVAND; PEBESMA; GÓMEZ-RUBIO, 2013; BIVAND; PIRAS, 2015). Para a estimação dos modelos de escolha discreta foi utilizado o pacote *apollo* (HESS; PALMA, 2019). O modelo espacial *SLX* não necessita de um pacote específico para a sua estimação, o procedimento de defasagem espacial das variáveis ocorre pela função *mutate()* relacionada as matrizes de vizinhança estimadas pelo *spdep* (MARTINETTI; GENIAUX, 2021). Um resumo das principais ações, bibliotecas e funções utilizadas nesta pesquisa é apresentado na Tabela 2.

Tabela 2 - Relação de bibliotecas e funções utilizadas

Ação	Biblioteca
Tratamento de dados	RStudio
Georreferenciamento e espacialização dos dados	<i>sf</i>
Roteamento das viagens	<i>otpr</i>
Estatísticas espaciais e construção das matrizes de vizinhança	<i>spdep</i>
Estimação modelo de escolha discreta não espacial	<i>apollo</i>
Estimação modelo de escolha discreta espacial	<i>apollo</i>
Defasagem espacial das variáveis explicativas	<i>spdep</i> – função <i>mutate</i>
Cálculo I de Moran	<i>spdep</i> – função <i>moran.test</i>
Filtro de distância mínima para matrizes de vizinhança	<i>spdep</i> – funções <i>dist</i> , <i>knearneigh</i> , <i>knn2nb</i> , <i>nbdist</i>

Fonte: Autor (2023)

4. RESULTADOS

Neste capítulo são apresentados e discutidos os resultados referentes às estatísticas descritivas, estatísticas espaciais, ao modelo *RUM* não espacial e aos modelos espaciais *SLX* estimados. Inicialmente são apresentados os resultados das estatísticas descritivas, os resultados obtidos para o teste I de Moran de propensão à escolha pelo modo a pé nas viagens em São Paulo, o critério para definição da distância da matriz de vizinhança, a topologia das matrizes de vizinhança, e o critério de seleção das variáveis a serem defasadas espacialmente. Os resultados gerais relacionados ao desempenho dos modelos pelas métricas estabelecidas são apresentados em seguida e, por fim, são apresentados e discutidos os parâmetros estimados em cada modelo.

4.1 RESULTADOS DAS ESTATÍSTICAS DESCRITIVAS

A Tabela 3 apresenta as estatísticas descritivas relativas aos atributos considerados nos modelos. Quanto às características das viagens a pé, têm duração de 28 minutos em média, com distâncias de acesso ao transporte coletivo inferior a 2.000 metros, e para egresso inferior a 1.500 metros, sendo que a maior parte dessas viagens ocorrem no pico da manhã. Quanto ao perfil sociodemográfico, observa-se que a maioria dos indivíduos são adultos, entre 30 e 60 anos, em sua maioria do sexo feminino, com rendas maiores e grau de instrução mais elevados, com escolaridade correspondente ao ensino superior completo. Quanto à posse de veículos, em média os indivíduos possuem um automóvel, há uma moto para cada 10 indivíduos da amostra e uma bicicleta para cada 2 indivíduos da amostra.

Tabela 3 - Descrição das variáveis analisadas estatisticamente

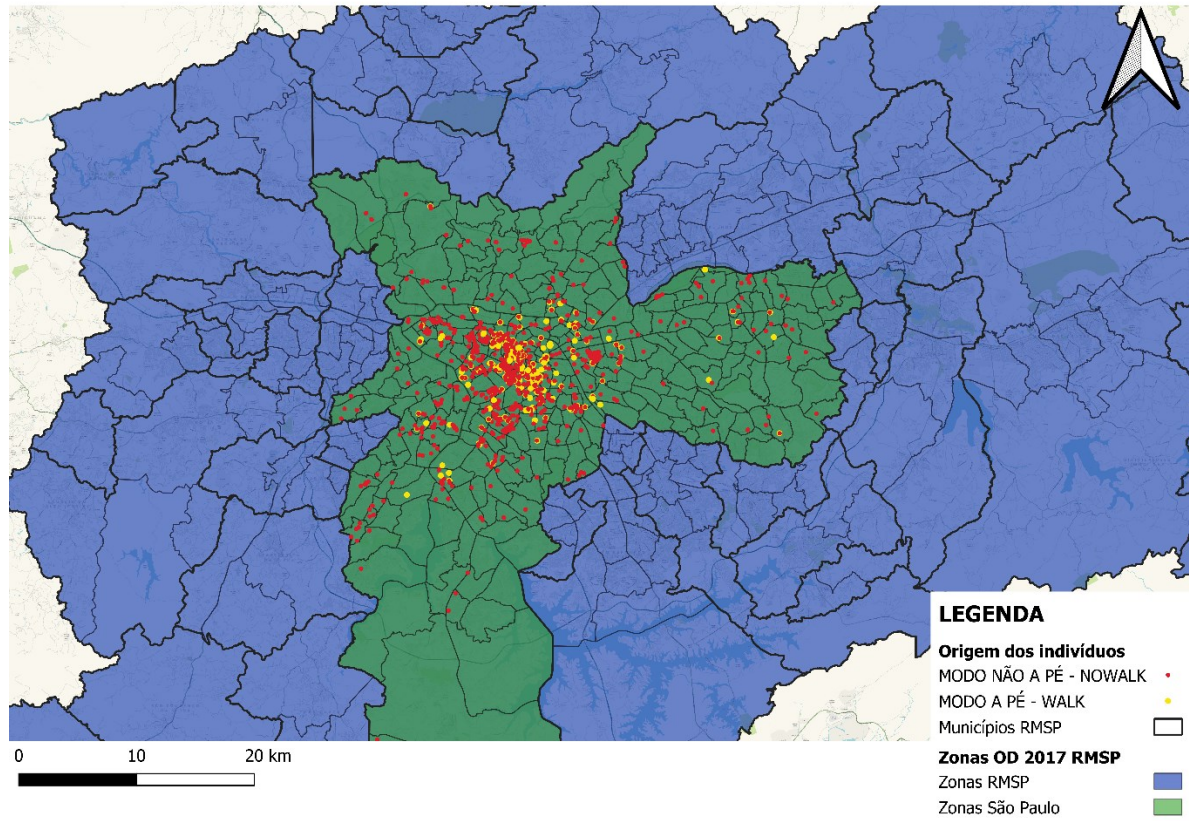
Atributo	Unidade	Média	Mínimo	Máximo	Desvio Padrão
TV_W	Minutos	28,27	0,75	59,95	16,69
D_{acesso}	Metros	1.909,10	0,00	23.179,00	1.873,29
$D_{egresso}$	Metros	1.403,00	0,00	11.386,10	1.436,37
$IDADE1$	%	67,06	-	-	-
$IDADE2$	%	11,54	-	-	-
$SEXO$	%	43,98	-	-	-
$CR1$	%	51,51	-	-	-
$CR2$	%	22,35	-	-	-
$GI1$	%	29,47	-	-	-
$GI2$	%	60,40	-	-	-
$PICO MANHÃ$	%	57,30	-	-	-
$MORADORES$	-	2,77	1	12	1,27
QT_{AUTO}	-	0,93	0	5	0,81
QT_{MOTO}	-	0,09	0	2	0,30
QT_{BICILE}	-	0,57	0	6	0,86

*A descrição das variáveis analisadas está na Tabela 1, página 50.

Fonte: Autor (2023)

Uma consideração importante se faz quanto à distribuição espacial das observações, já que a distância entre os indivíduos é importante para a análise e modelagem espacial. Na Figura 5 é apresentada a localização dos domicílios dos indivíduos da amostra e as indicações de suas escolhas perante as alternativas disponíveis para viagem (*WALK* e *NOWALK*).

Figura 5 - Escolha das alternativas da amostra analisada por origem



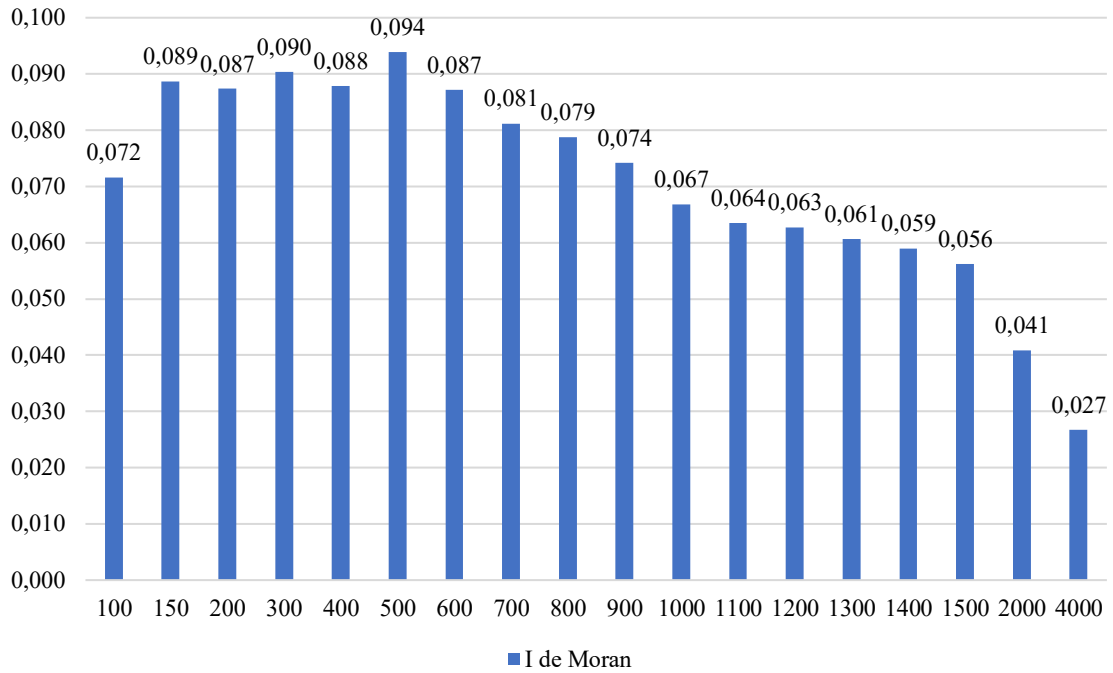
Fonte: Autor (2023)

4.2 RESULTADOS DAS ESTATÍSTICAS ESPACIAIS

Para a identificação de correlação espacial na variável dependente – *WALK/NOWALK*, utilizou-se a variável obtida pela estimação do modelo *RUM* não espacial referente à propensão de escolha. Essa condição foi necessária pois o teste I de Moran é realizado apenas com variáveis numéricas. Para o teste foram estimadas as matrizes de vizinhanças nos critérios especificados na seção 3.4.1, afim de capturar a relação espacial entre as observações.

Inicialmente foi realizado o teste I de Moran que captura a correlação espacial de forma global nas observações, a fim de identificar o grau dessa correlação espacial e a distância de defasagem que apresenta maior intensidade quanto a correlação. Os resultados do teste I de Moran considerando a propensão à escolha modal resultante do modelo *RUM* não espacial são apresentados na Figura 6. Todos os valores foram significativos a 95% de confiança, considerando o *p-value*.

Figura 6 - Resultado teste I de Moran variável propensão de escolha do modo



Fonte: Autor (2023)

Os resultados indicam que há dependência espacial na propensão à escolha do modo, porém todos os valores do I de Moran são pequenos, sendo o maior referente à distância de 500 metros (0,0939). Apesar dos resultados do teste indicarem uma baixa correlação espacial na escolha do modo a pé, entende-se que este valor máximo sugere presença de efeito espacial nos dados analisados. Assume-se, portanto, a premissa de que essa pequena dependência espacial sugere efeito espacial local, justificando a estimativa de modelos de escolha discreta do tipo *SLX* conforme conduzido nesta pesquisa. Assim, para compreender este fenômeno justifica-se a estimativa dos modelos de escolha discreta espacial defasados nas variáveis explicativas conforme propostos nesta pesquisa.

A fim de verificar a consistência das matrizes de vizinhança, foram analisadas as topologias de cada delas em função da distância de defasagem. Cabe destacar que, para garantir o pressuposto de que todas as observações tivessem ao menos um vizinho e a natureza do fenômeno em análise estar correlacionada a menores distâncias, se estabeleceu que todas as observações teriam pelo menos um vizinho para a distância mínima de corte de 50 metros. Os resultados da análise de topologia das matrizes são apresentados na Tabela 4.

Tabela 4 - Topologia das matrizes de vizinhança estimadas

Distância	Número de conexões não nulas	Porcentagem de pesos não nulos	Número médio de conexões	Regiões sem conexões	Uma conexão	Mais de uma conexão	Observações mais conectadas
50	3.888	0,08%	1,8	0	1420	778	3 (1339, 1340, 1356 com 12 conexões)
100	6.020	0,12%	2,7	0	960	1.238	1 (393 com 17 conexões)
150	8.850	0,18%	4,0	0	751	1.447	1 (374 com 25 conexões)
200	12.112	0,25%	5,5	0	603	1.595	2 (356, 370 com 31 conexões)
300	20.066	0,42%	9,1	0	434	1.764	2 (407, 411 com 37 conexões)
400	30.144	0,62%	13,7	0	304	1.894	1 (1114 com 54 conexões)
500	40.764	0,84%	18,5	0	245	1.953	3 (284, 305, 311 com 77 conexões)
600	53.398	1,10%	24,3	0	210	1.988	1 (300 com 100 conexões)
700	67.274	1,39%	30,6	0	174	2.024	1 (314 com 114 conexões)
800	83.156	1,72%	37,8	0	146	2.052	1 (333 com 135 conexões)
900	100.542	2,08%	45,7	0	127	2.071	1 (269 com 169 conexões)
1000	118.514	2,45%	53,9	0	111	2.087	1 (309 com 193 conexões)
1100	139.598	2,89%	63,5	0	88	2.110	1 (388 com 217 conexões)
1200	161.464	3,34%	73,5	0	84	2.114	1 (391 com 240 conexões)
1300	183.990	3,81%	83,7	0	69	2.129	5 (354, 357, 966, 967, 977 com 268 conexões)
1400	208.006	4,31%	94,6	0	61	2.137	1 (365 com 300 conexões)
1500	232.702	4,82%	105,9	0	56	2.142	1 (258 com 342 conexões)
2000	370.826	7,68%	168,7	0	28	2.170	1 (355 com 481 conexões)
4000	1.114.410	23,07%	507,0	0	4	2.194	2 (1021, 1042 com 1057 conexões)

Fonte: Autor (2023)

Perante os resultados da topologia das matrizes, identifica-se que a partir de 500 metros a porcentagem de pesos espaciais não nulos é superior a 0,84%, o número médio de conexões com vizinhos se torna superior a 18 (ou seja, em média cada observação tem 18 vizinhos), não existem observações sem vizinhos e menos de 11% das observações possuem só um vizinho. Ainda, as três observações mais conectadas possuem 77 vizinhos cada. Esses resultados garantem as propriedades pertinentes as matrizes mencionadas na seção 2.2.3.2.

Para identificar as variáveis explicativas que apresentam dependência espacial entre si, novamente o teste I de Moran foi realizado para cada uma das variáveis explicativas, considerando a matriz de vizinhança de 500 metros. As variáveis, resultados do teste e *p-values* são apresentados na Tabela 5.

Tabela 5 - Teste de dependência espacial I de Moran com matriz de vizinhança de 500m nas variáveis explicativas modeladas

Variável	I de Moran	<i>p-value</i>
<i>TV_W</i>	0,2676	0,0000
<i>D_{acesso}</i>	0,9798	0,0000
<i>D_{egresso}</i>	0,0340	0,0054
<i>SEXO*</i>	0,0079	0,2682
<i>CR1</i>	0,1196	0,0000
<i>CR2</i>	0,3431	0,0000
<i>IDADE1</i>	0,0238	0,0365
<i>IDADE2</i>	0,0345	0,0049
<i>GI1</i>	0,1215	0,0000
<i>GI2</i>	0,2818	0,0000
<i>MORADORES</i>	0,1346	0,0000
<i>PICO MANHÃ</i>	0,0333	0,0063
<i>QT_{AUTO}</i>	0,1867	0,0000
<i>QT_{MOTO}*</i>	-0,0016	0,5324
<i>QT_{BICICLE}</i>	0,0505	0,0001

*significativo a 95%

Fonte: Autor (2023)

A partir dos resultados pode-se observar que a maior parte das variáveis possuem valores de I de Moran muito baixos, inferiores a 0,2, sendo em alguns casos o parâmetro não significativo ao nível de confiança de 95%. As variáveis com maiores valores são *D_{acesso}* que se relaciona ao ambiente construído e *CR2* que se relaciona a categoria de menores rendas. Portanto, a definição das duas variáveis defasadas espacialmente (*D_{acesso}* e *CR2*) no modelo de escolha discreta foi pelo critério de corte (*cutoff*) mínimo de 0,3 para o I de Moran considerando a distância de defasagem para a matriz de vizinhança de 500 metros.

4.3 RESULTADOS GERAIS

Os resultados quanto ao desempenho dos seis primeiros modelos estimados, não espacial e espaciais com matrizes de vizinhança de 200, 300, 400, 500 e 600 metros ($RUM_{\text{não espacial}}$, RUM_{200} , RUM_{300} , RUM_{400} , RUM_{500} e RUM_{600}), para a amostra de 2.192 observações são apresentados na Tabela 6. Estas distâncias para as matrizes de vizinhança foram estabelecidas devido aos valores do I de Moran do modelo de escolha discreta não espacial, conforme resultados apresentados na Figura 6.

Tabela 6 - Resultados gerais dos modelos estimados não espacial e espacial de 200m, 300m, 400m, 500m e 600m.

Medida de desempenho	$RUM_{\text{não espacial}}$	RUM_{200}	RUM_{300}	RUM_{400}	RUM_{500}	RUM_{600}
LL_{inicial}	-516,39	-516,39	-516,39	-516,39	-516,39	-516,39
LL_{final}	-288,83	-293,13	-286,88	-286,91	-286,95	-286,87
ρ^2	0,4014	0,3925	0,4054	0,4054	0,4053	0,4054
ρ^2_{ajustado}	0,3682	0,3593	0,3723	0,3722	0,3721	0,3723
AIC	609,66	618,25	605,76	605,81	605,91	605,74
BIC	700,74	709,34	696,84	696,89	696,99	696,83
N de parâmetros	16	16	16	16	16	16
N de iterações	29	27	28	28	29	29
Tempo de estimação	2,76s	3,24s	2,4s	2,23s	2,32s	2,16s

Os ajustes dos modelos espaciais evidenciam que o modelo RUM_{600} apresenta o melhor desempenho, seguido do RUM_{300} , RUM_{400} , RUM_{500} e RUM_{200} . A melhora mais significativa ocorre em termos da log-verossimilhança entre os modelos RUM_{600} e RUM_{200} , sendo de 6,26 pontos de log-verossimilhança. Entre o modelo com melhor desempenho RUM_{600} e o modelo base RUM_{500} , a melhora em termos da log-verossimilhança é de apenas 0,08 pontos.

Por meio do teste de razão de verossimilhança, considerando os modelos RUM_{200} e RUM_{500} , obteve-se como resultado do $LR = 12,348 > \chi^2_9 = 3,325$ ($p\text{-value} < 5\%$), o que indica a rejeição da hipótese de que as diferenças entre os modelos não são significativas. Considerando

os modelos RUM_{300} , RUM_{400} e RUM_{600} , em comparação com o RUM_{500} obteve-se como resultado dos testes de razão de verossimilhança os valores de 0,1495, 0,0946 e 0,1622, respectivamente. Para todos esses resultados, a interpretação é a mesma, indica-se a rejeição da hipótese de diferença significativa entre os modelos espaciais. O único modelo que é considerado diferente em relação ao RUM_{500} (modelo RUM_{200}) é o que apresenta piores resultados quanto ao desempenho e ajuste se comparado aos demais modelos.

Estes resultados são apresentados com o intuito de obter parâmetros de comparação e evidenciar a distância com melhor ajuste no modelo. No entanto, o modelo RUM_{500} permanece como o modelo principal para as análises, já que perante o critério de seleção da distância de defasagem o que resultou em maior valor de correlação espacial pelo teste I de Moran.

Considerando o modelo RUM_{500} para análise das medidas de desempenho em relação ao modelo $RUM_{não\ espacial}$, os resultados apresentados indicam que o modelo espacial possui melhor desempenho em todos os parâmetros de ajuste em detrimento do não espacial. No que tange à log-verossimilhança final, o modelo espacial melhora em 1,88 pontos. Em relação aos valores de ρ^2 e $\rho_{ajustado}^2$, o modelo RUM_{500} melhora em 0,0039 para ambos. Ao se analisar os valores de AIC, o modelo espacial melhora em 3,75 pontos. Por meio do teste de razão de verossimilhança, considerando os modelos $RUM_{não\ espacial}$ e RUM_{500} , obteve-se como resultado do $LR = 3,757 > \chi_9^2 = 3,325$ ($p\text{-value} < 5\%$), o que indica a rejeição da hipótese que a incorporação da componente espacial não é significativa. Assim, apesar da pequena diferença entre os modelos não-espacial e espaciais em termos de métricas de desempenho (verossimilhança, ρ^2 , $\rho_{ajustado}^2$, AIC e BIC), a inclusão de aspectos espaciais na análise melhora o ajuste do modelo e, por consequência, o poder explicativo do fenômeno de escolha de modos de transporte quando viagens a pé são uma alternativa plausível.

O teste de *Variance Inflation Factor* (*VIF*) foi executado para verificar a presença de multicolinearidade nos modelos. Os resultados apresentados na Tabela 7 indicam que não existe multicolinearidade nos dados pois para todos os parâmetros o valor do *VIF* é menor que 1,00.

Tabela 7 - Teste de multicolinearidade VIF

Parâmetro	VIF $RUM_{não\ espacial}$	VIF RUM_{500}
TV_W	0,6006	0,6771
D_{acesso}	0,6001	0,6770
$D_{egresso}$	0,6001	0,6770
$SEXO$	0,6243	0,7055
$CR1$	0,6511	0,7089
$CR2$	0,7028	0,6444
$IDADE1$	0,6081	0,6857
$IDADE2$	0,4153	0,5050
$GI1$	0,5983	0,7017
$GI2$	0,4610	0,6029
$MORADORES$	0,6042	0,6816
$PICO\ MANHÃ$	0,6254	0,7070
QT_{AUTO}	0,5899	0,6701
QT_{MOTO}	0,7123	0,8147
$QT_{BICICLE}$	0,6103	0,6883

Fonte: Autor (2023)

A fim de verificar o desempenho destes modelos perante a variação no *cutoff* do I de Moran para a defasagem das variáveis explicativas, foram estimados modelos com a distância de 500 metros para os *cutoffs* de 0,1; 0,2; 0,3 e 0,5. Para valores maiores do que 0,4 não foi necessária a estimação pois resultado foi igual o modelo de 0,5. Para o *cutoff* de 0,1 ($RUM_{500}^{0,1}$) foram defasadas as variáveis TV_W , D_{acesso} , $CR1$, $CR2$, $GI1$, $GI2$, $MORADORES$ e QT_{AUTO} . Para o *cutoff* de 0,2 ($RUM_{500}^{0,2}$) foram defasadas as variáveis TV_W , D_{acesso} , $CR2$, $GI2$. Para o *cutoff* de 0,3 ($RUM_{500}^{0,3}$) foram defasadas as variáveis D_{acesso} , $CR2$. Para o *cutoff* de 0,5 ($RUM_{500}^{0,5}$) foi defasada a variável D_{acesso} . Os resultados destes modelos estão apresentados na Tabela 8.

Tabela 8 - Resultados gerais dos modelos espaciais estimados de 500m, com *cutoffs* de 0,1; 0,2, 0,3 e 0,5

Amostra	Medida de desempenho	$RUM_{500}^{0,1}$	$RUM_{500}^{0,2}$	$RUM_{500}^{0,3}$	$RUM_{500}^{0,5}$
	$LL_{inicial}$	-516,39	-516,39	-516,39	-516,39
	LL_{final}	-438,6	-420,18	-286,95	-288,84
	ρ^2	0,091	0,1553	0,4053	0,4014
	$\rho^2_{ajustado}$	0,0578	0,1292	0,3721	0,3682
2.192 observações	AIC	909,21	872,35	605,91	609,68
	N de parâmetros	16	16	16	16
	N de iterações	26	27	29	29
	Tempo de estimação	4,24s	2,66s	2,32s	2,67s

Os parâmetros de ajuste dos modelos evidenciam que o modelo $RUM_{500}^{0,3}$ apresenta o melhor desempenho, seguido do $RUM_{500}^{0,5}$, $RUM_{500}^{0,2}$ e $RUM_{500}^{0,1}$. A melhora mais significativa ocorre em termos da log-verossimilhança entre os modelos $RUM_{500}^{0,3}$ e $RUM_{500}^{0,1}$, sendo de 151,65 pontos. Dentre o modelo com melhor desempenho, $RUM_{500}^{0,3}$, e o modelo base, $RUM_{500}^{0,5}$, a melhora em termos da log-verossimilhança é de 1,89 pontos.

Por meio do teste de razão de verossimilhança, considerando os modelos $RUM_{500}^{0,1}$ e $RUM_{500}^{0,3}$, obteve-se como resultado do $LR = 303,3 > \chi_9^2 = 3,325$ ($p\text{-value} < 5\%$), o que indica a rejeição da hipótese que as diferenças entre os modelos não são significativas. Da mesma forma, considerando os modelos RUM_{300} e $RUM_{500}^{0,3}$, RUM_{400} e $RUM_{500}^{0,3}$ obteve-se como resultado do LR os valores de 266,45 ($\chi_9^2 = 3,325$) e 3,7735 ($\chi_9^2 = 3,325$), respectivamente. Para todos esses resultados, a interpretação é a mesma indicando a rejeição da hipótese que as diferenças entre os modelos não são significativas. Os três modelos são considerados diferentes, o que é esperado já que existem diferentes variáveis defasadas espacialmente incorporadas em cada um; no entanto, por apresentarem piores resultados quanto ao desempenho e ajuste, estes foram desconsiderados para demais comparações e análises.

Estes resultados são apresentados com o intuito de evidenciar que o modelo $RUM_{500}^{0,3}$ possui melhor desempenho quanto aos parâmetros de ajustes, sendo importante para o estudo a defasagem espacial das duas variáveis D_{acesso} , $CR2$, através das quais se observam os maiores efeitos espaciais. Os efeitos desses parâmetros espaciais são apresentados no próximo tópico.

4.4 RESULTADOS DOS PARÂMETROS ESTIMADOS

Os resultados referentes aos parâmetros estimados pelos diferentes modelos são apresentados na Tabela 9 seguida de discussões. Como apresentado na seção anterior, os modelos RUM_{200} , RUM_{300} , RUM_{400} e RUM_{600} , relacionados às distâncias de defasagem, assim como os modelos $RUM_{500}^{0,1}$, $RUM_{500}^{0,2}$ e $RUM_{500}^{0,5}$, relacionados às variáveis explicativas defasadas, não serão analisados e discutidos quanto aos seus parâmetros. Entretanto, o Apêndice A contém as tabelas com os valores dos parâmetros para os referidos modelos.

Tabela 9 - Resultados dos parâmetros estimados para os modelos não espacial e espacial com defasagem da 500 metros e *cutoff* 0,3

Parâmetro	$RUM_{\text{não espacial}}$		$RUM_{500}^{0,3}$	
	Parâmetro	Teste-t	Parâmetro	Teste-t
<i>Atributos dos modos e ambiente construído</i>				
ASC_{WALK}	-4,7622	-6,52*	-4,6836	-7,73*
TV_W	-0,1461	-12,44*	-0,1470	-12,39*
D_{acesso}	-2,874e-05	-0,32	-	-
D_{egresso}	-2,316e-06	-0,02	1,502e-06	0,01
<i>Atributos sociodemográficos</i>				
$CR1$	0,1477	0,48	0,0537	0,24
$CR2$	0,2764	0,63	-	-
$SEXO$	-0,0893	-0,42	-0,1121	-0,53
$IDADE1$	0,6881	2,70*	0,7347	2,87*
$IDADE2$	1,3082	3,21*	1,2807	3,14*
$MORADORES$	-0,0179	-0,21	-0,0295	-0,34
$GI1$	0,7866	2,19*	0,6972	1,97*
$GI2$	1,2073	2,99*	1,0957	2,97*
$PICO MANHÃ$	0,0789	0,37	0,1315	0,60
QT_{AUTO}	0,9769	5,30*	0,9665	5,72*
QT_{MOTO}	-0,2840	-0,63	-0,2951	-0,65
QT_{BICICLE}	-0,0178	-0,13	-0,0224	-0,17
<i>Atributos espaciais</i>				
$LAG D_{\text{acesso}}$	-	-	-7,122-e-05	-0,76
$LAG CR2$	-	-	0,9078	2,04*

* significativo a 95%.

4.4.1 Parâmetros dos atributos dos modos e ambiente construído

Denota-se significância estatística ao nível de confiança de 95% para dois dos quatro parâmetros estimados, sendo eles a constante de cada modelo e o tempo de viagem $WALK$ para cada modelo.

As constantes em ambos os modelos apresentam sinal negativo, indicando menor propensão dos indivíduos à escolha do modo $WALK$. O valor obtido para o modelo $RUM_{\text{não espacial}}$ é 0,0786 maior que o obtido para o modelo RUM_{500} , evidenciando que os parâmetros estimados são muito próximos para ambos os modelos e causam o mesmo efeito direto na tomada de decisão.

Os parâmetros relacionados ao tempo de viagem $WALK$ apresentam sinal negativo para a alternativa, quanto maior o tempo de viagem menor a propensão de escolha do modo a pé. O parâmetro relativo ao tempo de viagem pelo modo $WALK$ possui efeito maior sobre a escolha no modelo RUM_{500} em 0,009 do que o do modelo $RUM_{\text{não espacial}}$, o que evidencia que os parâmetros estimados são muito próximos para ambos os modelos, causando o mesmo efeito direto na tomada de decisão.

Ao se analisar os parâmetros relativos à distância de acesso e egresso do transporte coletivo, atributos relacionados ao ambiente construído, no modelo $RUM_{não\ espacial}$ ambos apresentam sinal negativo, ao contrário do modelo RUM_{500} que apresenta sinal positivo. No entanto, nenhum destes parâmetros são significativos à 95%. A magnitude dos parâmetros é da mesma ordem de grandeza em ambos os modelos. O parâmetro da distância de acesso pelo modelo RUM_{500} será apresentada em tópicos subsequentes, já que é um atributo espacial.

4.4.2 Parâmetros dos atributos sociodemográficos

Perante os atributos sociodemográficos considerados, os modelos $RUM_{não\ espacial}$ e RUM_{500} apresentam cinco parâmetros significativos dos doze estimados, referentes à idade, grau de instrução e quantidade de automóveis.

Os resultados dos parâmetros relacionados à idade, $IDADE1$ e $IDADE2$, em ambos os modelos são significativos à 95% de confiança, e ambos apresentam sinal positivo, indicando que independentemente da idade as pessoas têm maior propensão a escolher o modo a pé. Cabe destacar, no entanto, que os indivíduos da categoria $IDADE2$ possuem uma maior propensão, indicado pelo maior valor do parâmetro em comparação à categoria $IDADE1$. A diferença no valor entre a categoria $IDADE2$ e $IDADE1$ é de 0,6201 no $RUM_{não\ espacial}$ e de 0,5460 no RUM_{500} . Esse resultado é interessante pois evidencia que indivíduos mais velhos possuem maior propensão de caminhar.

Quanto aos parâmetros relacionados ao grau de instrução, GII e $GI2$, ambos apresentam sinais positivos e são significativos estatisticamente a 95% de confiança. A magnitude desses dois parâmetros é similar. Em todos os modelos, independentemente do grau de instrução, existe uma maior propensão pela escolha do modo $WALK$, e à medida que se aumenta esse grau, a pessoa tem ainda maior propensão de fazer essa escolha. A diferença no valor entre a categoria $GI2$ e GII é de 0,4207 no $RUM_{não\ espacial}$ e de 0,3985 no RUM_{500} . É interessante evidenciar que pessoas com maior grau de instrução possuem também maior propensão a escolher o modo $WALK$ para viagens a trabalho e ou estudo, o que pode estar relacionado, no caso de São Paulo, a pessoas com maior grau de instrução morarem mais próximos de seus locais de trabalho ou estudo.

Os últimos parâmetros a serem analisados se relacionam a quantidade de veículos. No que tange a quantidade de automóveis (QT_{AUTO}), em ambos os modelos o sinal é positivo e significativo, indicando que quanto maior a quantidade de veículos, maior é a propensão de escolha do modo *WALK*. Esse fator pode estar relacionado ao fenômeno de mudança de estrutura familiar, a exemplo de casais que tem seus filhos, que produz uma mudança na disponibilidade de veículos frente ao aumento do número de atividades rotineiras da família, ou à associação com a renda do indivíduo, que em São Paulo concentra serviços que permitem acesso a pé em regiões de elevada renda e posse de automóvel. Já para a quantidade de motos (QT_{MOTO}) e bicicletas ($QT_{BICYCLE}$), em ambos os modelos o sinal é negativo, porém não significativos estatisticamente.

4.4.3 Parâmetros dos atributos espaciais

Perante o critério estabelecido para a definição das variáveis defasadas espacialmente, foram considerados os atributos D_{acesso} e $CR2$. O primeiro associa-se a um atributo de serviço e ambiente construído, enquanto o segundo a um atributo sociodemográfico.

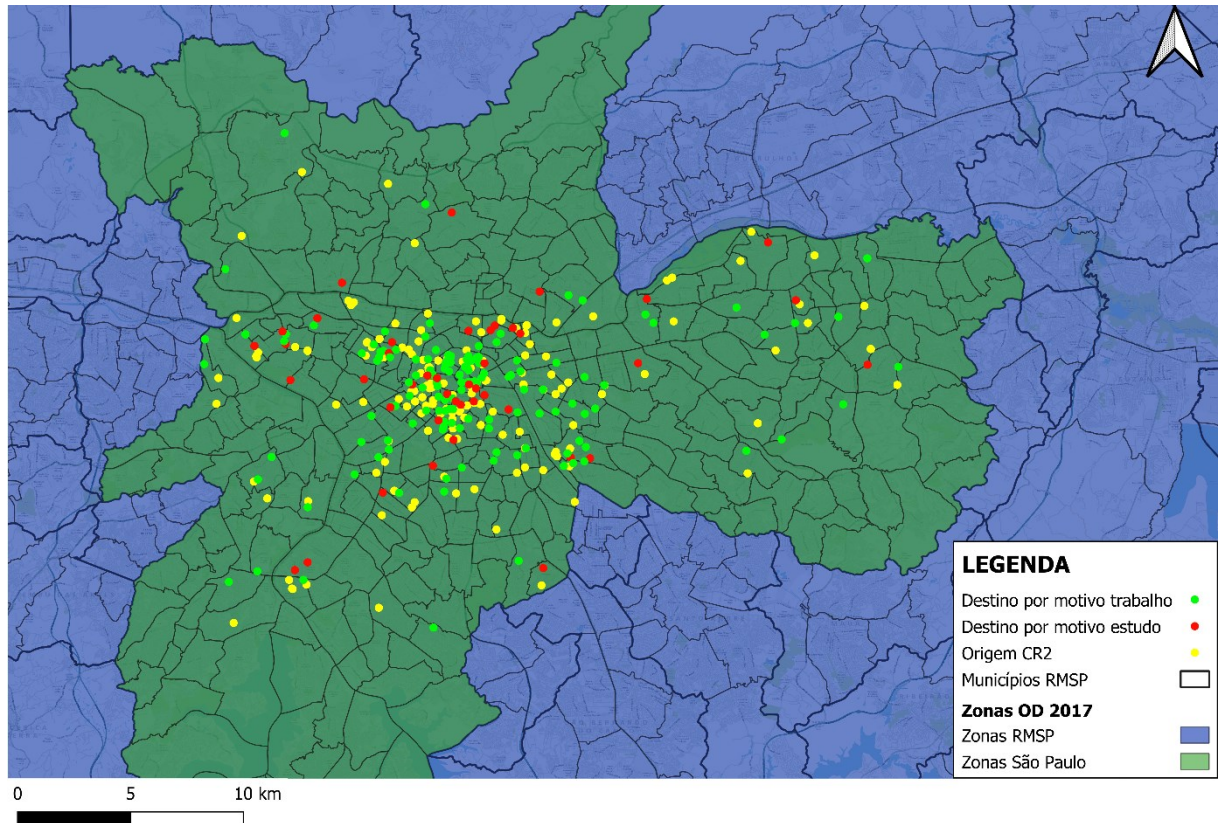
O parâmetro obtido pelo modelo espacial com defasagem de 500 metros e *cutoff* 0,3 associados a D_{acesso} , representa a defasagem espacial quando a distância de acesso ao transporte público. Esse parâmetro não é significativo estatisticamente ao nível de confiança de 95%.

O parâmetro $CR2$ defasado espacialmente para o mesmo modelo representa o critério de renda da categoria 2, ou seja, indivíduos de menor renda. Este parâmetro é significativo estatisticamente a 95% de confiança e apresenta sinal positivo, indicando que as pessoas que pertencem a esta categoria possuem maior propensão a escolha do modo *WALK* para viagens a trabalho e estudo.

A análise deste resultado deve ser tal que o indivíduo pertencer ao critério de renda 2 influencia positivamente na escolha do modo a pé e que, se os seus vizinhos pertencem ao critério de renda 2, existe uma influência na escolha deles indiretamente. Essa condição indica que as escolhas das pessoas de renda mais baixa é semelhante aos seus vizinhos e existe um efeito espacial nas decisões de escolha de modo desses indivíduos pertencentes ao critério de renda 2. Essa condição é verificada pela concentração de indivíduos com baixa renda no município de São Paulo em bairros específicos. O mesmo não ocorre com o critério de renda

mais alta (*CR1*), já que a população de média renda se estabelece de forma mais heterogênea no espaço. A Figura 7 apresenta a distribuição das origens e destinos de viagens de acordo com o motivo e o modo escolhido para os indivíduos do critério de renda mais baixa (*CR2*).

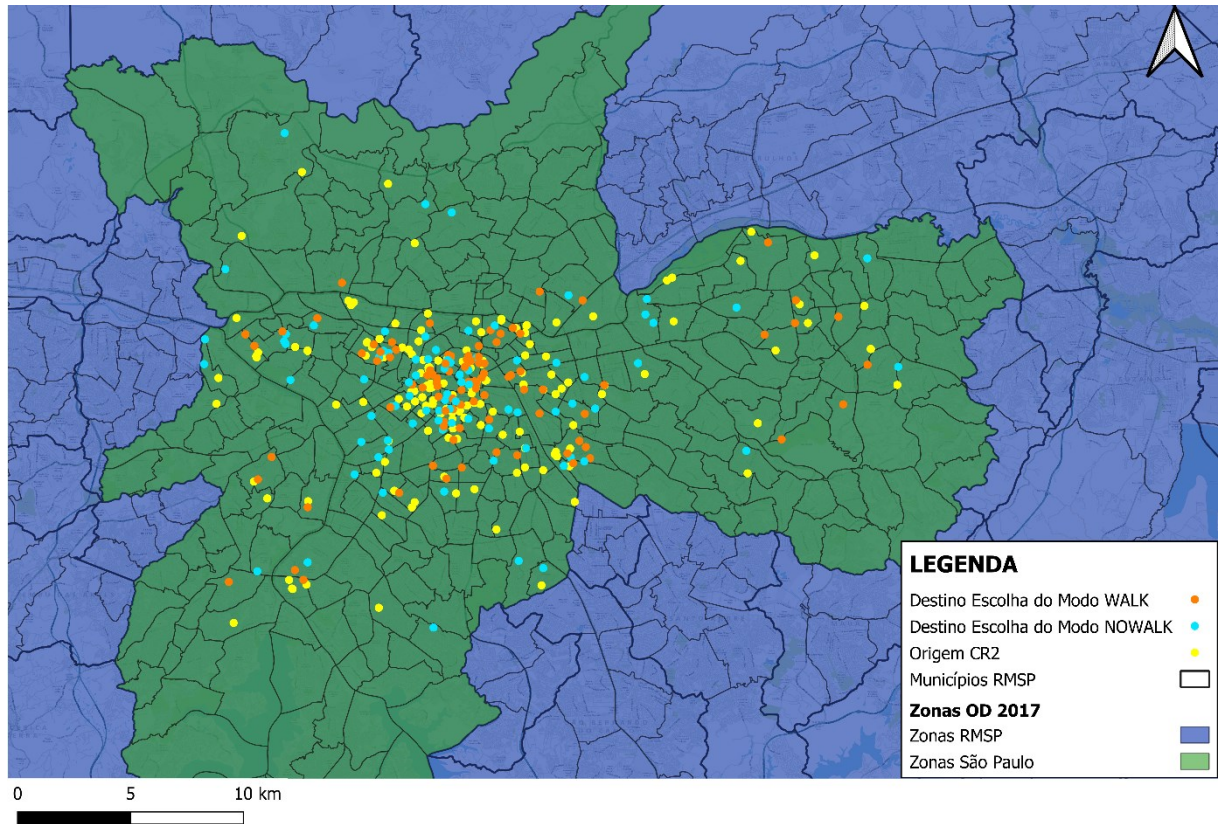
Figura 7 - Origens dos indivíduos classificados por critério de renda mais baixa (*CR2*) e destinos por motivo de viagem



Fonte: Autor (2023)

A distribuição dos destinos pelos respectivos motivos de viagem ocorre em todo o território da cidade quanto à origem dos indivíduos no estrato de menor renda, tal que grande parte dos destinos se concentram no centro expandido e na zona sudeste. As observações fora da concentração centro-sudeste têm origens ao longo dos principais corredores de transporte metropolitano. No que tange aos destinos, se observa uma concentração de viagens para o centro expandido com motivo trabalho. A Figura 8 evidencia a localização dos destinos em relação ao modo escolhido para as viagens.

Figura 8 - Origens dos indivíduos classificados por critério de renda mais baixa (CR2) e destinos por escolha de modo (WALK/NOWALK)



Fonte: Autor (2023)

É possível verificar que a maior parte das viagens pelo modo *WALK* dos indivíduos do CR2 ocorrem para a região central e do centro expandido, além desta região existem viagens pontuais que estão próximas de locais de origem dos indivíduos analisados. Contudo não é possível afirmar que a origem dessas viagens é próxima destes destinos pelo modo a pé, no entanto sabe-se que a origem está no máximo a um raio de 4.000 metros, como critério indicado para a disponibilidade da alternativa *WALK*.

4.5 CONSIDERAÇÕES FINAIS

A fim de comparar os resultados obtidos e aqueles encontrados na literatura, são discutidos os principais resultados em relação ao estudo de Neves (2019), que modelou o mesmo fenômeno com a base de dados da Pesquisa OD 2007 da RMSP com atributos associados apenas ao ambiente construído, além dos principais resultados das demais

referências apresentadas quanto a escolha do modo a pé que se relacionam aos resultados encontrados nesta pesquisa. Perante os resultados das estatísticas espaciais, ao se comparar os resultados obtidos para a distância de defasagem da matriz de vizinhança, observa-se que o *cutoff* de distância com maior valor do teste I de Moran foi de 400 metros e o valor de I de Moran foi de 0,116 no estudo de Neves (2019). O presente estudo apresentou o maior valor de I de Moran a 500 metros igual a 0,0939. Os valores são relativamente baixos e indicam correlação espacial positiva baixa. Quanto ao critério de defasagem das variáveis espaciais, Neves (2019) adotou o *cutoff* de 0,5 para o I de Moran das variáveis explicativas, enquanto para nesta pesquisa foi adotado o valor de 0,3, perante os melhores valores de ajuste do modelo *SLX*.

Ao se analisar as variáveis significativas, observa-se que Neves (2019) considera apenas variáveis que se relacionam aos atributos do ambiente construído, e no presente estudo considerou-se apenas duas variáveis, distância de acesso e de egresso ao transporte público, sendo que nenhuma delas foi significativa em nenhum dos modelos. Essas mesmas variáveis apresentaram significância no modelo *SLX* de Neves (2019), além de no modelo não espacial. Nesta pesquisa os parâmetros obtidos apresentam sinal positivo, ao contrário dos obtidos pelo autor. Olojede, Yoade e Olufemi (2017) e Ton *et al.* (2019) também identificaram de forma similar que a densidade de acesso ao transporte público é um fator que afeta a escolha pelo modo a pé para viagens.

Quanto aos atributos sociodemográficos, assim como neste estudo, os parâmetros modelados por Neves (2019) para idade, disponibilidade de veículos e grau de instrução foram significativos no modelo não espacial e no *SLX*. No entanto, enquanto nesta pesquisa apenas no modelo espacial o parâmetro de critério de renda é significativo, em ambos os modelos os parâmetros são significativos naquele modelo. Olojede, Yoade e Olufemi (2017) e Ton *et al.* (2019) também identificaram que os atributos posse de veículos e a renda apresentam coeficientes significativos e que afetam a escolha pelo modo a pé.

Ao se analisar os resultados obtidos com foco em políticas públicas, é possível indicar que indivíduos mais velhos, assim como as pessoas com maior grau de instrução, maiores rendas e maior quantidade de automóveis possuem uma maior propensão a caminhar. Esse resultado é específico para essa aplicação ao município de São Paulo, que possui estruturas sociodemográficas significativamente diferentes dos demais municípios brasileiros. Além disso, existe uma menor utilidade inicial pelo modo *WALK*. Esses fatores indicam que a caminhada é preferida por pessoas que possuem maiores rendas (relação entre maior renda, maior grau de instrução e maior quantidade de automóveis), sendo relacionado a maior

qualidade de infraestrutura do ambiente construído nas localidades de origem destes indivíduos, enquanto para os demais a qualidade é inferior, o que desincentiva a caminhada. Assim, o investimento na melhoria das condições do ambiente construído, assim como a maior disponibilidade de oportunidades de trabalho e estudo nas demais regiões do município, pode diminuir a distância de caminhada para acesso as atividades e aumentar a propensão da escolha do modo a pé nas viagens a trabalho e estudo, especialmente na zona Leste da capital onde a maior concentração do uso do solo é residencial.

5 CONCLUSÕES

O planejamento de transportes é uma tarefa de relevância no contexto das cidades. O processo de desenvolvimento de soluções que possibilitem melhorias na qualidade de vida das pessoas que vivem ou passam por um município permeia as responsabilidades de um planejador de transportes. Grandes são os desafios ao analisar cidades de países em desenvolvimento como o Brasil, em especial ao se considerar os desafios de promover transporte público de qualidade, frente ao incentivo e desenvolvimento de políticas e infraestruturas que privilegiem o modo individual. As consequências dessa escolha são visíveis nas externalidades como congestionamentos, poluição, alto número de sinistros no trânsito, além da baixa qualidade de vida atrelada à desigualdade de acesso às oportunidades e atividades da cidade.

Perante esta reflexão são propostas técnicas e métodos que auxiliam no planejamento dos sistemas de transporte, em especial no que tange a demanda por transportes. Sabe-se que essa demanda é dividida entre os diferentes modos de transporte disponíveis em uma cidade, desde os modos ativos, como caminhada e bicicleta, até os modos motorizados, como ônibus, metrô, trem, monotrilho, entre outros. Ao se considerar o problema das externalidades presentes pelo uso de veículos motorizados, os modos ativos surgem como uma solução estratégica, já que não geram poluição, congestionamentos no sistema viário, e por um planejamento estratégico do uso do solo podem melhorar o acesso as oportunidades e atividades da cidade. Esses modos ativos são incentivados e priorizados pela Política Nacional de Mobilidade Urbana – Lei nº 12.587/2012, e perante esta proposição legal o modo a pé foi considerado como objeto de análise neste estudo.

Entre as técnicas de previsão e modelagem do comportamento da demanda se destacam os modelos de escolha discreta derivados de teorias econométricas para avaliação da viabilidade de projetos e políticas públicas com foco na promoção de infraestrutura e bem-estar social. A abordagem *RUM*, estabelecida pela premissa que o indivíduo realiza suas escolhas baseado na maximização de sua utilidade perante um conjunto de alternativas disponíveis, é a mais utilizada na literatura de Engenharia de Transportes. Apesar da técnica ser amplamente utilizada, a modelagem de fenômenos sociais presentes no espaço social e geográfico é complexa, e uma solução em busca de melhorar a previsibilidade desses modelos econométricos tem sido pela inclusão de efeitos de dependência espacial presentes no fenômeno.

Assim, ficou estabelecida a seguinte pergunta de pesquisa: Qual a influência do efeito espacial local nas escolhas individuais na realização de viagens a pé? E assumiu-se como hipótese de que existem evidências de dependência espacial na escolha do modo a pé no contexto de viagens urbanas por motivo trabalho e estudo. A hipótese foi confirmada pelos resultados encontrados na aplicação conduzida nesta pesquisa no contexto da cidade de São Paulo.

O objetivo principal desta pesquisa foi analisar a dependência espacial em viagens a pé através de modelos de escolha discreta espaciais sob a perspectiva da Teoria da Utilidade com estudo de caso no município de São Paulo. Os seguintes objetivos específicos compreenderam: identificar os atributos espaciais que potencialmente afetam os usuários na escolha do modo em viagens urbanas pelo modo a pé; caracterizar o conjunto de atributos que potencialmente afetam a escolha do modo a pé com estudo de caso no município de São Paulo; e analisar a dependência espacial na escolha do modo em viagens urbanas a pé através de modelos de escolha discreta espaciais sob a perspectiva da Teoria da Utilidade.

Perante o primeiro objetivo específico foi identificado que o atributo *LAG CR2* quanto ao critério de renda mais baixa defasado espacialmente afeta potencialmente os usuários na escolha do modo a pé em viagens urbanas em São Paulo. Quanto ao segundo objetivo, pode-se estabelecer que indivíduos maiores de 15 anos, e especialmente os mais velhos, assim como as pessoas com maior grau de instrução, maiores rendas e maiores quantidade de automóveis, possuem uma maior propensão a escolher o modo a pé em viagens urbanas, assim como o inverso é verdadeiro. Quanto ao último objetivo, foi estimado o modelo espacial e identificou-se o atributo que apresenta dependência espacial (classe de renda mais baixa), assim como foi possível identificar a distância de corte para a matriz de vizinhança que captura o maior efeito espacial.

Foi utilizado o banco de dados da Pesquisa Origem-Destino de 2017 da RMSP realizada pelo METRÔ-SP. A pesquisa contém informações de preferência revelada dos indivíduos, quanto a sua viagem realizada em um dia útil da semana entre o período de 2017 e 2018. As informações sociodemográficas dos indivíduos foram utilizadas como variáveis explicativas do fenômeno, além da caracterização do nível de serviço do modo *WALK* por meio de uma *API* de roteamento multimodal, obtendo informações do tempo de viagem e da distância entre a origem e o destino em rede.

Baseado na abordagem *RUM*, foram estimados modelos *logit* não espacial e espaciais *SLX* incluindo variáveis relacionadas ao nível de serviço do modo a pé (tempo de viagem) e que impactam a escolha do modo transporte público (distância de acesso e egresso), ao ambiente construído e aos atributos sociodemográficos. Para o modelo espacial *SLX* foram considerados como atributos espaciais aqueles que indicaram valor superior a 0,3 para o teste de correlação espacial I de Moran, resultando na distância de acesso ao transporte público e categoria de renda mais alta.

Os resultados dos modelos indicaram que o modelo *logit* espacial *SLX* apresentou melhor desempenho em termos de log-verossimilhança do que o modelo *logit* não espacial, sendo estes modelos estatisticamente diferentes tal que a inclusão dos atributos espaciais é significativa na predição das escolhas. Quanto às matrizes de vizinhança, foram especificadas dezenove matrizes no intervalo de 50 metros a 4.000m. Perante a análise do teste I de Moran, a distância que apresentou maior correlação espacial quanto aos dados de predição da escolha do modelo não espacial foi a matriz de 500 metros, com valor do teste indicou uma correlação espacial positiva baixa.

Ao se analisar os parâmetros os que relacionam aos atributos do modo, tempo de viagem pelo modo a pé apresentou efeitos negativos sobre a propensão de escolha do modo e foram significativos em ambos os modelos. Em relação aos atributos do ambiente construído, a distância de acesso ao transporte público não foi significativa no modelo não espacial, assim como a distância de egresso não foi significativa em nenhum dos modelos.

Em relação aos atributos sociodemográficos, indivíduos maiores que 15 anos, e especialmente os mais velhos, assim como as pessoas com maior grau de instrução, maiores rendas e maiores quantidade de automóveis possuem uma maior propensão a caminhar. Os demais atributos sociodemográficos não resultaram em parâmetros significativos. Perante os atributos espaciais, apenas aquele referente ao nível de renda mais alto (*CR2*) apresentou significância estatística a 95% de confiança, indicando dependência espacial positiva.

Os resultados apresentados evidenciam a importância da incorporação de efeitos espaciais nas análises no contexto do município de São Paulo, ao melhorar a capacidade de previsão dos modelos em termos de log-verossimilhança, além de evidenciar aspectos relacionados ao comportamento das viagens ou a caracterização sociodemográfica dos indivíduos que estão correlacionadas espacialmente no ambiente da cidade. O método empregado evidenciou valores de corte e a topologia das matrizes de vizinhança como meios de caracterizar a relação espacial

entre os dados, além de um formato de caracterização dos atributos do nível de serviço dos modos e a aplicabilidade do modelo *SLX* no contexto de escolha de modo. De forma geral os resultados indicam que a caminhada é preferida por pessoas que possuem maiores rendas (relação entre maior renda, maior grau de instrução e maior quantidade de automóveis), sendo relacionado à maior qualidade de infraestrutura do ambiente construído nas localidades de origem destes indivíduos, enquanto para os demais a qualidade é inferior, o que desincentiva a caminhada. Assim, o investimento na melhoria das condições do ambiente construído, assim como a maior disponibilidade de oportunidades de trabalho e estudo nas demais regiões do município, pode diminuir a distância de caminhada para acesso as atividades e aumentar a propensão da escolha do modo a pé nas viagens a trabalho e estudo.

5.1 LIMITAÇÕES E RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Os resultados fornecidos foram restritos exclusivamente à cidade de São Paulo, devido às limitações relacionadas à caracterização dos atributos do ambiente construído, à distância acesso e egresso ao transporte público e às informações disponíveis no momento da criação do banco de dados da pesquisa. Essa delimitação foi uma medida necessária devido à falta de dados *GTFIS* das outras cidades da Região Metropolitana de São Paulo na época em que a pesquisa foi conduzida.

Além disso, a exclusão de viagens que ocorrem com distâncias maiores de 4.000 metros diminuiu a amplitude de análise do estudo, já que pelo reportado no banco de dados original existem pessoas que realizam viagens a pé com distâncias entre a origem e o destino superior a esse limiar indicado por Vale *et al.* (2018). Essa condição de viagens realizadas com distâncias maiores que a predisposição comum pode estar relacionada a fatores sociodemográficos de renda, como a pobreza extrema com pessoas que não têm condição de pagar pela tarifa do transporte público.

Ainda são necessários aprofundamentos e estudos que se correlacionam ao tema desta dissertação, pela melhora no entendimento do comportamento relacionado a viagens a pé de forma a gerar subsídios para formulação de políticas públicas que melhorem a qualidade de vida das pessoas.

No que tange a consideração do fenômeno, demais estudos podem ser analisados considerando apenas modos ativos, identificando as principais diferenças nos atributos que

influenciam na escolha dos modos para viagem, podendo ser estimado um modelo binário perante as alternativas caminhada e escolha da bicicleta, além da modelagem multinomial considerando os modos ativos como referência, e se relacionando a demais motivos como lazer, compras e saúde. Aconselha-se que mesmas modelagens incorporem a análise de dependência espacial uma vez que traz melhorias no desempenho do modelo.

No que diz respeito à análise da dependência espacial, é importante observar que há a possibilidade de utilizar diferentes testes para identificar essa dependência, além de analisar homogeneidade espacial e estacionaridade. Além disso, pode-se explorar outras variáveis numéricas que estejam correlacionadas com a escolha como entrada para o teste, em vez de se restringir apenas à propensão de escolha por um modelo não espacial. Outro fator a se analisar é a construção de matrizes de vizinhança baseada em critérios sociais, como renda e grau de instrução. A partir disso, é possível considerar diferentes modelos espaciais propostos na literatura como o *SAR* e *SEM* em função dos efeitos espaciais identificados.

Além disso, o fenômeno caminhar a pé pode estar relacionado a efeitos sociais e atitudinais, evidenciado pelas variáveis explicativas sociodemográficas de renda e grau de instrução, podendo ser incorporadas aos modelos para melhor representar as influências da escolha no espaço social nos diferentes contextos empíricos.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABREU, V. H. S.; ALMEIDA, I. M.; TURINI, L. R. Utilização do transporte ativo para aumento da qualidade de vida da população urbana durante a pandemia do covid-19 à luz de abordagem bibliométrica. 2021. **Brazilian Journal of Production Engineering**, (2021), 139-152, 7(3)
- ALAMÁ-SABATER, L.; ARTAL-TUR, A.; NAVARRO-AZORÍN, J. M. Industrial location, spatial discrete choice models and the need to account for neighbourhood effects. **Annals of Regional Science**, v. 47, n. 2, p. 393–418, out. 2011.
- ALMEIDA, E. Econometria Espacial Aplicada. **Editora Alínea**, v. 1, 2012.
- ANSELIN, L. **Regional economic policy: measurement of its effect**. [s.l.] Martinus Nijhoff, 1988.
- ANSELIN, L. Spatial externalities, spatial multipliers, and spatial econometrics. **International Regional Science Review**, v. 26, n. 2, p. 153–166, abr. 2003.
- ANSELIN, L. Under the hood Issues in the specification and interpretation of spatial regression models. **Agricultural Economics**, v. 27, 2002.
- ANTONINI, G. **A discrete choice modeling framework for pedestrian walking behavior with application to human tracking in video sequences**. Tese—Lausanne: École Polytechnique Fédérale de Lausanne, 2005.
- ANTONINI, G.; BIERLAIRE, M.; WEBER, M. Discrete choice models of pedestrian walking behavior. **Transportation Research Part B: Methodological**, v. 40, n. 8, p. 667–687, 2006.
- ASSIRATI, L. **Análise da influência da vizinhança no comportamento individual relativo a viagens através de dados em painel**. São Carlos: Universidade de São Paulo, 2018.
- AUTANT-BERNARD, C. Where Do Firms Choose to Locate Their R&D? A Spatial Conditional Logit Analysis on French Data. **European Planning Studies**. v. 14, n.9, p. 1187-1208, 2006.
- BEN-AKIVA, M. E.; LERMAN, S. R. **Discrete Choice Analysis: Theory and Application to Travel Demand**. 1. ed. London: Press, The MIT, 1985.

BERON, K. J.; VIJVERBERG, W. P. M. Probit in a Spatial Context: A Monte Carlo Analysis. Em: [s.l: s.n.]. p. 169–195.

BHAT, C. R. The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. **Transportation Research Part B: Methodological**, v. 45, n. 7, p. 923–939, 2011.

BHAT, C.; GOVINDARAJAN, A.; PULUGURTA, V. Disaggregate Attraction-End Choice Modeling Formulation and Empirical Analysis. **Transportation Research Record**, p. 60–68, 1998.

BIERLAIRE, M.; ANTONINI, G.; WEBER, M. **Behavioral dynamics for pedestrians**. 10th International Conference on Travel Behavior Research. **Anais.2003**.

BILLÉ, A. G.; ARBIA, G. Spatial Limited Dependent Variable Models: A Review Focused on Specification, Estimation, and Health Economics Applications. **Journal of Economic Surveys**, v. 33, n. 5, p. 1531–1554, 1 dez. 2019.

BILLÉ, A. G.; LEORATO, S. Partial ML estimation for spatial autoregressive nonlinear probit models with autoregressive disturbances. **Econometric Reviews**, v. 39, n. 5, p. 437–475, 27 maio 2020.

BIVAND, R. S.; PEBESMA, E. J.; GÓMEZ-RUBIO, V. **Applied Spatial Data Analysis with R**. 2013.

BIVAND, R.; PIRAS, G. Comparing Implementations of Estimation Methods for Spatial Econometrics. **Journal of Statistical Software**, v.63, n.18, p. 1-36, 2015.

BORST, H. C.; VRIES, A. I.; GRAHAM, J. M. A.; DONGEN, J. E. F.; BAKKER, I.; MIEDEMA, H. M. E. Influence of environmental street characteristics on walking route choice of elderly people. **Journal of Environmental Psychology**, v. 29, n. 4, p. 477–484, dez. 2009.

BOVY, P. H. L.; STERN, E. **Route Choice: Wayfinding in Transport Networks**. Kluwer Academic ed. Dordrecht: Springer Netherlands, 1990. v. 9.

BRASIL. Casa Civil. LEI Nº 12.587, DE 3 DE JANEIRO DE 2012. Brasília, DF. 2012. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2012/Lei/L12587.htm>. Acesso em: 23 de maio de 2022.

CALABRESE, R.; ELKINK, J. A. Estimators of binary spatial autoregressive models: A Monte Carlo study. **Journal of Regional Science**, v. 54, n. 4, p. 664–687, 2014.

CALDEIRA, G. P. **Dependência espacial sob a abordagem de minimização do arrependimento em modelos de escolha de modo de transporte**. Dissertação – Mestrado. Departamento de Engenharia de Transportes – USP. São Paulo. (2022), 1-143.

CALDEIRA, G. P.; IMAI, R. A. M.; MELO, L. E. A.; CUNHA, C. B.; ISLER, C. A. Fatores que afetam a caminhabilidade de adultos mais velhos em áreas urbanas brasileiras: uma abordagem multinível. **Transportes**. V. 29, n. 3, 2657, 2021.

CARRIÓN-FLORES, C. E.; FLORES-LAGUNES, A.; GUCI, L. An estimator for discrete-choice models with spatial lag dependence using large samples, with an application to land-use conversions. **Regional Science and Urban Economics**, v. 69, p. 77–93, 1 mar. 2018.

CHORUS, C. G. What about behaviour in travel demand modelling? An overview of recent progress. **Transportation Letters**, v. 4, n. 2, p. 93–104, 2012.

CHORUS, C. G.; ROSE, J. M.; HENSHER, D. A. Regret minimization or utility maximization: It depends on the attribute. **Environment and Planning B: Planning and Design**, v. 40, n. 1, p. 154–169, 2013.

CLIFTON, K. J.; SINGLETON, P. A.; MUHS, C. D.; SCHNEIDER, R. J. Development of destination choice models for pedestrian travel. **Transportation Research Part A: Policy and Practice**, v. 94, p. 255–265, 1 dez. 2016.

CORRADO, L.; FINGLETON, B. Where is the economics in spatial econometrics? **Journal of Regional Science**, v. 52, n. 2, p. 210–239, maio 2012.

DAVIS, M. **Planeta de favelas - A involução urbana e proletariado informal**. [s.l: s.n.]. Boitempo Editorial, 2006.

DING, C.; CHEN, Y.; DUAN, J.; LU, Y.; CUI, J. Exploring the Influence of Attitudes to Walking and Cycling on Commute Mode Choice Using a Hybrid Choice Model. **Journal of Advanced Transportation**, v. 2017, p. 1–8, 2017.

DUGUNDJI, E. R.; PÁEZ, A.; ARENTZE, T. A.; WALKER, J. L.; CARRASCO, J. A.; MARCHAL, F.; NAKANISHI, H. Transportation and social interactions. **Transportation Research Part A: Policy and Practice**, v. 45, p. 239-247, 2011.

ELHORST, J. P.; VEGA, S. H. El modelo SLX: ampliación de la forma general, y sensibilidad de los desbordamientos espaciales a la especificación de la W. **Papeles de Economía Española**, v. 152, p. 34–50, 2017.

EWING, R.; CERVERO, R. Travel and the built environment. **Journal of the American Planning Association**, v. 76, n. 3, p. 265–294, jun. 2010.

GIBBONS, S.; OVERMAN, H. G. Mostly pointless spatial econometrics? **Journal of Regional Science**, v. 52, n. 2, p. 172–191, maio 2012.

GOETZKE, F. **Are Travel Demand Forecasting Models Biased because of Uncorrected Spatial Autocorrelation?** Philadelphia, PA: [s.n.]. Disponível em: <<https://www.researchgate.net/publication/253696403>>. 2003

GOETZKE, F. 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, fev. 2008.

GREENE, W. H.; HENSHER, D. A. A latent class model for discrete choice analysis: Contrasts with mixed logit. **Transportation Research Part B: Methodological**, v. 37, n. 8, p. 681–698, 2003.

GRUS, J. **Data Science do Zero**. [s.l.]. Alta Books Editora, 2016.

GUO, J. Y.; BHAT, C. R. Modifiable Areal Units: A Problem or a Matter of Perception in the Context of Residential Location Choice Modeling? **Transportation Research Record**. 2004.

HCM. Highway Capacity Manual - 2010. **Transportation Research Board**, 2010.

HENSHER, D. A.; GREENE, W. H. The Mixed Logit model: The state of practice. **Transportation**, (2003), 133-176, 30.

HESS, S.; PALMA, D. Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. **Journal of Choice Modelling**, v. 32, p. 100170, 2019.

HESS, S.; STATHOPOULOS, A.; DALY, A. Allowing for heterogeneous decision rules in discrete choice models: An approach and four case studies. **Transportation**, v. 39, n. 3, p. 565–591, 2012.

HILL, M. R. **Spatial Structure and Decision-Making Aspects of Pedestrian Route Selection through an Urban Environment**. [s.l.: s.n.]. Disponível em:

<<http://digitalcommons.unl.edu/geographythesis>>. (1982), 1-280

HOOGENDOORN, S. P.; BOVY, P. H. L. Pedestrian route-choice and activity scheduling theory and models. **Transportation Research Part B: Methodological**, v. 38, n. 2, p. 169–190, 2004.

KAMARGIANNI, M.; POLYDOROPOULOU, A. Hybrid choice model to investigate effects of teenagers' attitudes toward walking and cycling on mode choice behavior. **Transportation Research Record**, n. 2382, p. 151–161, 1 dez. 2013.

KLIER, T.; MCMILLEN, D. P. **Clustering of Auto Supplier Plants in the U.S.: GMM Spatial Logit for Large Samples**. (2008), 1-38

KOPPELMAN, F. S.; BHAT, C. **A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models**. (2006), 1-249.

LESAGE, J. P. Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models. **Geographical Analysis**, v. 32, p. 1–17, 2000. Disponível em: <<http://www.econ.utoledo.edu>>.

LOPES, S. B.; BRONDINO, N. C. M.; SILVA, A. N. R. GIS-based analytical tools for transport planning: Spatial regression models for transportation demand forecast. **ISPRS International Journal of Geo-Information**, v. 3, n. 2, p. 565–583, 1 jun. 2014.

MANESS, M.; CIRILLO, C.; DUGUNDJI, E. R. Generalized behavioral framework for choice models of social influence: Behavioral and data concerns in travel behavior. **Journal of Transport Geography**, v. 46, p. 137–150, 1 jun. 2015.

MARCUS, Y.; BLAINEY, S. Defining probability-based rail station catchments for demand modelling. In: **48th Annual UTSG Conference**. Bristol: 8 jan. 2016. Disponível em: <https://eprints.soton.ac.uk/384539/1/UTSG_paper_young_blaineyv2.pdf>. Acesso em: 8 junho. 2022.

MARTINETTI, D.; GENIAUX, G. Approximate likelihood estimation of spatial probit models. **Regional Science and Urban Economics**, v. 64, p. 30–45, 1 maio 2017.

McFADDEN, D. Conditional logit analysis of qualitative choice behaviour. In: Zarembka, P. (Ed.), **Frontiers in Econometrics**. Academic Press, New York, pp. 105–142. 1974.

MCMILLEN, D. P. Probit With Spatial Autocorrelation Journal of Regional Science. **Journal Of Regional Science**, (1992), 335-348, 32(3).

METRÔ. **Pesquisa Origem e Destino 2017**. Disponível em: <<http://www.metro.sp.gov.br/pesquisa-od/resultado-das-pesquisas.aspx>>. Acesso em: 17 mai. 2022.

METRÔ. **Tarifa de integração de ônibus, CPTM e Metrô terá novo valor a partir do dia 15 de abril**. Disponível em: <<https://www.metro.sp.gov.br/noticias/11-04-2017-tarifa-de-integracao-de-onibus-cptm-e-metro-tera-novo-valor-a-partir-do-dia-15-de-abr.fss>>. Acesso em: 23 de abril de 2022.

MIYAMOTO, K.; VICHIANAN, V.; SHIMOMURA, N.; PÁEZ, A. Discrete Choice Model with Structuralized Spatial Effects for Location Analysis. **Transportation Research Board**, (2004), 183-190, 1898, 2004.

MOZHAROVSKIY, P.; VOGLER, J. Composite marginal likelihood estimation of spatial autoregressive probit models feasible in very large samples. **Economics Letters**, v. 148, p. 87–90, 1 nov. 2016.

NEVES, C. E. T. Análise da relação entre o ambiente construído e a escolha pelo modo a pé por meio de modelo espacial de escolha discreta. Dissertação – Mestrado. **Faculdade de Tecnologia**, Universidade de Brasília – UNB. Brasília. 2019.

NEVES, C. E. T.; SILVA, A. R.; ARRUDA, F. S. DE. Exploring the link between built environment and walking choice in São Paulo city, Brazil. **Journal of Transport Geography**, v. 93, 1 maio 2021.

OLOJEDE, O.; YOADE, A.; OLUFEMI, B. Determinants of walking as an active travel mode in a Nigerian city. **Journal of Transport and Health**, v. 6, p. 327–334, 1 set. 2017.

ORTÚZAR, J.D.; WILLUMSEN L.G. **Modelling Transport**. John Wiley & Sons, 4^a ed., New Delhi, 2011.

OSM. **OpenStreetMap**. Disponível em: <https://www.openstreetmap.org/#map=4/-15.11/-53.17>. Acesso em: 22 de jun de 2022.

PACE, R. K.; LESAGE, J. P. Fast simulated maximum likelihood estimation of the spatial probit model capable of handling large samples. **Advances in Econometrics**, v. 37, p. 3–34, 2016.

PÁEZ, A.; SCOTT, D. M. Social influence on travel behavior: A simulation example of the decision to telecommute. **Environment and Planning A**, v. 39, n. 3, p. 647–665, 2007.

PÁEZ, A.; SCOTT, D. M.; VOLZ, E. A discrete-choice approach to modeling social influence on individual decision making. **Environment and Planning B: Planning and Design**, v. 35, n. 6, p. 1055–1069, 2008.

PAPADIMITRIOU, E.; YANNIS, G.; GOLIAS, J. A critical assessment of pedestrian behaviour models. **Transportation Research Part F: Traffic Psychology and Behaviour**, v. 12, n. 3, p. 242–255, 2009.

PEBESMA, E. Simple Features for R: Standardized Support for Spatial Vector Data. **The R Journal**, v. 10, n. 1, p. 439–446, jul. 2018.

PENN, A.; TURNER, A. **Space Syntax Based Agent Simulation**. London: 2002.

PEREIRA, R. H. M.; BANISTER, D.; SCHWANEN, T.; WESSEL, N. Distributional effects of transport policies on inequalities in access to opportunities in Rio de Janeiro. **Journal of Transport and Land Use**, v. 12, n. 1, p. 741–764, 2019.

PEREIRA, R. H. M. Transport legacy of mega-events and the redistribution of accessibility to urban destinations. **Cities**, v. 81, p. 45–60, 2018.

PINJARI, A. R.; BHAT, C. R. Activity-based travel demand analysis. In: ANDRÉ DE PALMA et al. (Eds.). **A Handbook of Transport Economics**. [s.l] Edward Elgar Publishing Ltd., 2011. P. 213-248.

PINKSE, J.; SLADE, M. E. Contracting in space: An application of spatial statistics to discrete-choice models. **Journal of Econometrics**, v. 85, n. 1, p. 125-154, 1998.

PRITCHARD, J. P.; TOMASIELLO, D.; GIANNOTI, M.; GEURS, K. An international comparison of equity in accessibility to jobs: London, São Paulo and the Randstad Transport Findings. **Transport Findings**, 2019. Disponível em: <<https://research.utwente.nl/en/publications/an-international-comparison-of-equity-in-accessibility-to-jobs-lo>>. Acesso em: 13 jun. 2022.

R CORE TEAM. **R: The R Project for Statistical Computing** ViennaR Foundation for Statistical Computing, 2022. Disponível em: <<https://www.r-project.org/>>. Acesso em: 22 jun. 2022.

- ROBIN, T.; ANTONINI, G.; BIERLAIRE, M.; CRUZ, J. Specification, estimation and validation of a pedestrian walking behavior model. **Transportation Research Part B: Methodological**, v. 43, n. 1, p. 36–56, 2009.
- RODRIGUEZ, D.A., JOO, J. The relationship between non-motorized mode choice and the local physical environment. **Transportation Research Part D: Transport Environ.** 9 (2), 151–173. <http://dx.doi.org/10.1016/j.trd.2003.11.001>. 2004.
- RÜTTENAUER, T. Spatial Regression Models: A Systematic Comparison of Different Model Specifications Using Monte Carlo Experiments. **Sociological Methods and Research**, v. 51, n. 2, p. 728–759, 2019.
- SAELENS, B. E.; HANDY, S. L. Built environment correlates of walking: A review. **Medicine and Science in Sports and Exercise**, 2008.
- SAELENS, B. E.; SALLIS, J. F.; FRANK, L. D. Environmental Correlates of Walking and Cycling: Findings from the Transportation, Urban Design, and Planning Literatures. **Environment and Physical Activity**, v. 25, n. 2, p. 80–91, 2003.
- SANTOS, L. F. Á. da S. **Essays in Spatial Econometrics**. Ph. D.—Lisboa: Universidade de Lisboa, jul. 2020.
- SCHNEIDER, R. J.; ARNOLD, L. S.; RAGLAND, D. R. Pilot model for estimating pedestrian intersection crossing volumes. **Transportation Research Record**, n. 2140, p. 13–26, 2009.
- SCHRECKENBERG, M. AND S. SHARMA *Pedestrian and Evacuation Dynamics*, **Springer Verlag**. (Eds.), 2002.
- SENEVARANTE, P.N., MORALL, J.F., 1986. Analysis of factors affecting the choice of route of pedestrians. **Transportation Planning and Technology** 10, 147–159.
- SHENG, M.; SHARP, B. Commuter’s transport mode preferences and social network effects in New Zealand. **Journal of Transport Economics and Policy**, v. 53, n. January, p. 19–46, 2019.
- SIDHARTHAN, R.; BHAT, C. R.; PENDYALA, R. M.; GOULIAS, K. G. Model for children’s school travel mode choice. **Transportation Research Record**, n. 2213, p. 78–86, 2011.

SLOVIC, A. D.; TOMASIELLO, D. B.; GIANNOTTI, M.; ANDRADE, M. de F. The long road to achieving equity: Job accessibility restrictions and overlapping inequalities in the city of São Paulo. **Journal of Transport Geography**, v. 78, p. 181–193, 2019.

SMIRNOV, O. A. Modeling spatial discrete choice. **Regional Science and Urban Economics**, v. 40, n. 5, p. 292–298, set. 2010.

SPTRANS. **GTFS** | **SPTrans**. Disponível em: <https://www.sptrans.com.br/desenvolvedores/login-desenvolvedores/>. Acesso em: 25 de maio de 2022.

TOBLER, W. R. A Computer Movie Simulating Urban Growth in the Detroit Region. **Economic Geography**, v. 46, p. 234, 1970.

TON, D.; DUIVES, D. C.; CATS, O.; HOOGENDOORN-LANSER, S.; HOOGENDOORN, S. P. Cycling or walking? Determinants of mode choice in the Netherlands. **Transportation Research Part A: Policy and Practice**, v. 123, p. 7–23, 1 maio 2019.

TRAIN, K. E. **Discrete Choice Methods with Simulation**. 2nd. ed. New York: CAMBRIDGE UNIVERSITY PRESS, 2009.

TRIBBY, C. P.; MILLER, H. J.; BROWN, B. B.; WERNER, C. M.; SMITH, K. R. Analyzing walking route choice through built environments using random forests and discrete choice techniques. **Environment and Planning B: Urban Analytics and City Science**, v. 44, n. 6, p. 1145–1167, 1 nov. 2017.

VALE, D. S., PEREIRA, M.; VIANA, C. M. 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, 2018.

VASCONCELLOS, E. A. Urban transport policies in Brazil: The creation of a discriminatory mobility system. **Journal of Transport Geography**, v. 67, p. 85–91, 1 fev. 2018.

VEGA, S. H.; ELHORST, J. P. The slx model. **Journal of Regional Science**, v. 55, n. 3, p. 339–363, 2015.

WANG, C. H.; AKAR, G.; GULDMANN, J. M. 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, 2015.

WANG, H.; IGLESIAS, E. M.; WOOLDRIDGE, J. M. Partial maximum likelihood estimation of spatial probit models. **Journal of Econometrics**, v. 172, n. 1, p. 77–89, 2013.

WANG, X. C.; KOCKELMAN, K. M.; LEMP, J. D. The dynamic spatial multinomial probit model: Analysis of land use change using parcel-level data. **Journal of Transport Geography**, v. 24, p. 77–88, 2012.

YOUNG, M. **An R Wrapper for the “OpenTripPlanner” REST API [R package otpr version 0.4.1]**. Comprehensive R Archive Network (CRAN), 2019. Disponível em: <<https://cran.r-project.org/package=otpr>>. Acesso em: 17 jul. 2021.

ZHOU, Y.; WANG, X.; HOLGUÍN-VERAS, J. Discrete choice with spatial correlation: A spatial autoregressive binary probit model with endogenous weight matrix (SARBP-EWM). **Transportation Research Part B: Methodological**, v. 94, p. 440–455, 1 dez. 2016.

APÊNDICE A– RESULTADOS DOS MODELOS

MODELOS ESPACIAIS DISTÂNCIA DE DEFASAGEM

Parâmetro	$RUM_{espacial}^{200m}$		$RUM_{espacial}^{300m}$	
	Parâmetro	Teste-t	Parâmetro	Teste-t
<i>Atributos dos modos e ambiente construído</i>				
ASC	-3,5427	-5,98*	-4,7050	-7,83*
TT_W	-0,1454	-12,54*	-0,1471	-12,39*
D_{acesso}	-	-	-	-
$D_{egresso}$	-8,788e-06	-0,08	-6,327e-06	-0,06
<i>Atributos sociodemográficos</i>				
CR1	-0,1008	-0,46	0,0528	0,24
CR2	-	-	-	-
SEXO	-0,0928	-0,44	-0,1103	-0,52
AGE1	0,7656	3,02*	0,7258	2,83*
AGE2	1,2107	3,07*	1,2973	3,18*
HOUSE	-0,0404	-0,48	-0,0239	-0,28
G11	-0,1105	-0,47	0,7235	2,06*
G12	-	-	1,1221	3,09*
PICO MANHÃ	0,1113	0,52	0,1328	0,61
QT_{AUTO}	1,1071	6,66*	0,9785	5,81*
QT_{MOTO}	-0,2796	-0,61	-0,3012	-0,66
$QT_{BICICLE}$	-0,0036	-0,03	-0,0166	-0,12
<i>Atributos espaciais</i>				
LAG D_{acesso}	-1,612e-06	-0,02	-5,300-e-05	-0,58
LAG CR2	-1,0028	-2,42*	0,8077	2,09*
LAG G12	-0,2793	-0,70	-	-

* significativo a 95%.

Parâmetro	$RUM_{espacial}^{400m}$		$RUM_{espacial}^{600m}$	
	Parâmetro	Teste-t	Parâmetro	Teste-t
<i>Atributos dos modos e ambiente construído</i>				
ASC	-4,7164	-7,83*	-4,6684	-7,70*
TT_W	-0,1471	-12,39*	-0,1466	-12,38*
D_{acesso}	-	-	-	-
$D_{egresso}$	2,549e-07	-0,01	4,886e-06	-0,05
<i>Atributos sociodemográficos</i>				
CR1	0,0488	0,22	0,0495	0,22
CR2	-	-	-	-
SEXO	-0,1154	-0,54	-0,1071	-0,50
AGE1	0,7352	2,87*	0,7408	2,89*
AGE2	1,2913	3,16*	1,2892	3,17*
HOUSE	-0,0316	-0,37	-0,0326	-0,38
G11	-0,7137	2,02*	0,7024	1,99*
G12	1,1167	3,05*	1,1017	3,00*
PICO MANHÃ	0,1267	0,58	0,1459	0,66
QT_{AUTO}	0,9691	5,74*	0,9603	5,68*
QT_{MOTO}	-0,3001	-0,66	-0,2943	-0,64
$QT_{BICICLE}$	-0,0203	-0,15	-0,0223	-0,17
<i>Atributos espaciais</i>				
LAG D_{acesso}	-6,293e-05	-0,68	-7,058-e-05	-0,75
LAG CR2	0,8446	-2,06*	0,9593	2,08*
LAG G12	-	-	-	-

* significativo a 95%.

MODELOS ESPACIAIS DEFASAGEM DE VARIÁVEIS EXPLICATIVAS

Parâmetro	$RUM_{cutoff\ 0,1}^{500m}$		$RUM_{cutoff\ 0,2}^{500m}$	
	Parâmetro	Teste-t	Parâmetro	Teste-t
<i>Atributos dos modos e ambiente construído</i>				
ASC	1,0012	-1,01	-0,3142	-0,58
TT_W	-	-	-	-
D_{acesso}	-	-	-	-
$D_{egresso}$	5,405e-05	0,64	5,830e-05	0,67
<i>Atributos sociodemográficos</i>				
CR1	-	-	-0,1501	-0,86
CR2	-	-	-	-
SEXO	-0,2843	-1,70	-0,2179	-1,27
AGE1	0,7365	3,84*	0,5393	2,70*
AGE2	1,0514	3,38*	0,7932	2,44*
HOUSE	-	-	0,0760	1,14
GI1	-	-	-0,0875	-0,46
GI2	-	-	-	-
PICO MANHÃ	-4,352e-04	-0,01	0,0790	0,45
QT_{AUTO}	-	-	0,8732	6,48*
QT_{MOTO}	0,5182	1,37	0,1596	0,42
$QT_{BICICLE}$	0,1243	1,16	0,0401	0,37
<i>Atributos espaciais</i>				
LAG TT_W	-0,0064	-2,39*	-0,0062	-2,42*
LAG D_{acesso}	-8,926e-05	-0,64	-9,521e-05	-1,25
LAG CR1	0,5321	0,91	-	-
LAG CR2	1,8067	2,30*	1,3323	2,76*
LAG GI1	0,0026	0,01	-	-
LAG GI2	-0,5719	0,71	-0,1834	-0,39
LAG HOUSE	0,3931	2,45*	-	-
LAG QT_{AUTO}	0,7136	2,15*	-	-

* significativo a 95%.

Parâmetro	$RUM_{cutoff\ 0,5}^{500m}$	
	Parâmetro	Teste-t
<i>Atributos dos modos e ambiente construído</i>		
ASC	-4,7573	-6,52*
TT _W	-0,1461	-12,43*
D _{acesso}	-	-
D _{egresso}	-3,487e-06	-0,03
<i>Atributos sociodemográficos</i>		
CR1	0,1491	0,49
CR2	0,2769	0,64
SEXO	-0,0896	-0,42
AGE1	0,6874	2,69*
AGE2	1,3079	3,21*
HOUSE	-0,0179	-0,21
GI1	0,7841	2,19*
GI2	1,20	2,99*
PICO MANHÃ	0,0796	0,37
QT _{AUTO}	0,9779	5,31*
QT _{MOTO}	-0,2831	-0,62
QT _{BICICLE}	-0,0182	-0,13
<i>Atributos espaciais</i>		
LAG TT _W	-	-
LAG D _{acesso}	-2,689e-05	-0,30
LAG CR1	-	-
LAG CR2	-	-
LAG GI1	-	-
LAG GI2	-	-
LAG HOUSE	-	-
LAG QT _{AUTO}	-	-

* significativo a 95%.