

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA DE TRANSPORTES

ALCEU DAL BOSCO JUNIOR

USABILIDADE DE PONTOS DE INTERESSE E
CENTRALIDADES DE REDE DE MAPAS COLABORATIVOS
PARA ANÁLISE DE ATRAÇÃO DE VIAGENS:
ESTUDO DE CASO DE CURITIBA

São Carlos

2020

ALCEU DAL BOSCO JUNIOR

USABILIDADE DE PONTOS DE INTERESSE E
CENTRALIDADES DE REDE DE MAPAS COLABORATIVOS
PARA ANÁLISE DE ATRAÇÃO DE VIAGENS:
ESTUDO DE CASO DE CURITIBA

**Programa de Pós-graduação em
Engenharia de Transportes da EESC-USP**

**Exemplar definitivo (corrigido). O exemplar
original está disponível na CPG da EESC-
USP**

São Carlos, 09/02/2021

RESOLUÇÃO CoPG N° 6018, DE 13 DE OUTUBRO DE 2011, artigo 5°

Dissertação apresentada ao Departamento de Engenharia de Transportes da Universidade de São Paulo (STT/EESC/USP) para a obtenção do título de Mestre em Ciências, Programa de Pós graduação em Engenharia de Transportes.

Área de concentração: Planejamento e Operação dos Sistemas de Transportes.

Orientador:

Prof. André Luiz Barbosa Nunes da Cunha

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da EESC/USP
com os dados inseridos pelo(a) autor(a).

D136u Dal-Bosco-Junior, Alceu
Usabilidade de pontos de interesse e centralidades
de rede de mapas colaborativos para análise de atração
de viagens: estudo de caso de Curitiba / Alceu Dal-
Bosco-Junior; orientador André Luiz Barbosa Nunes da
Cunha. São Carlos, 2020.

Dissertação (Mestrado) - Programa de
Pós-Graduação em Engenharia de Transportes e Área de
Concentração em Planejamento e Operação de Sistemas de
Transporte -- Escola de Engenharia de São Carlos da
Universidade de São Paulo, 2020.

1. Atração de viagens. 2. Mapas colaborativos.
3. Sintaxe Espacial. 4. Centralidades. 5. Pontos de
interesse. I. Título.

FOLHA DE JULGAMENTO

Candidato: Engenheiro **ALCEU DAL BOSCO JUNIOR**.

Título da dissertação: " Usabilidade de pontos de interesse e centralidades de rede de mapas colaborativos para análise de atração de viagens: estudo de caso de Curitiba".

Data da defesa: 10/12/2020.

Comissão Julgadora

Resultado

Prof. Dr. **André Luiz Barbosa Nunes da Cunha**
(Orientador)
(Escola de Engenharia de São Carlos – EESC/USP)

APROVADO

Prof. Dr. **Gustavo Garcia Manzato**
(Universidade Estadual Paulista "Júlio de Mesquita Filho" UNESP - Bauru)

APROVADO

Prof. Dr. **Bruno Vieira Bertoncini**
(Universidade Federal do Ceará/UFC)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia de Transportes:
Prof. Associado **Adalberto Leandro Faxina**

Presidente da Comissão de Pós-Graduação:
Prof. Titular **Murilo Araujo Romero**

Agradecimentos

Ao meu orientador André Luiz Cunha. Por apontar os destinos e permitir a exploração dos caminhos. Por todo o auxílio e suporte ao longo dos três anos de maior aprendizado de minha vida. Também aos professores Gustavo Garcia Manzato e Bruno Vieira Bertoncini e à professora Cira Souza Pitombo pelas contribuições essenciais à pesquisa.

À minha mãe Márcia e meu pai Alceu, que me apoiaram em todas as escolhas e fizeram de mim quem sou. Mesmo que as distâncias diminuam os almoços e jantares juntos, sempre me senti em casa por saber que tenho seu incentivo. A vocês, todo o amor e admiração.

Ao meu amor, Vanessa, que me acompanha nesta viagem. Também superamos distâncias e dificuldades, sempre com carinho, compreensão, emojis, músicas e planos. Espero ser para você o orgulho e o apoio que você é para mim. Te amo muito!

Aos parceiros de orientação André, Bruna, Elaine, Gabriel, Helena, Leandro, Natália e Paola; e aos colegas Francisco, Ana Luiza, Marcus, Marina, Renan, Tássio, Ricardo, Daniela, Tiago, Gustavo, Rosuel, Igor, Sérgio, Juliana, Aurenice, Jorge, Vanessa, Samuel, Thalita, as duas Luizas, os dois Matheus, os dois Felipes e o Piva. Grandes amizades que dão muito mais sentido ao tempo que dividimos na pesquisa, com suporte mútuo, parcerias, *happy hours* e ‘jogatinas’. O programa de pós-graduação não seria nada sem a família discente do primeiro andar.

Aos professores e técnicos do departamento, que levam cada um de nós, estudantes, ao patamar de pesquisadores internacionais, com profissionalismo, seriedade e atenção. E também com a dose certa de churrascos e informalidade. A vocês, todo o reconhecimento merecido.

A toda minha família, que sempre me faz entender a importância da educação. Por serem parte integrante de minhas conquistas e por encherem meu mundo de alegria.

Aos amigos de longa data Gabriel, Raul e Ping. Aos colegas do IPPUC Francisco e Paula. À(os) camaradas do Impulsão. E a todos os parceiros e professores do Grupo de Estudos em Transportes UFPR. Por serem todos responsáveis por minha escolha do mestrado, por estarem sempre comigo, por me darem incríveis oportunidades e lições e por rirem das minhas piadas.

Ao IPPUC e à SETRAN de Curitiba, pelo atendimento solícito e fornecimento dos dados.

Pelo fomento: O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

*If you're asking me
I got one thing to say
If I can't live here
Guess I'll go live on the internet
(All Them Witches – Internet)*

Resumo

DAL BOSCO JR., A. Usabilidade de Pontos de Interesse e Centralidades de Rede de Mapas Colaborativos para Análise de Atração de Viagens: Estudo de Caso de Curitiba. Dissertação (Mestrado). Escola de Engenharia de São Carlos, Universidade de São Paulo, 2020.

Entender a demanda por viagens é crucial para o planejamento sustentável de transportes. A atração de viagens é comumente analisada e modelada considerando-se a localização de atividades que induzem os deslocamentos – os Pontos de Interesse (*POIs*) – e aspectos de acessibilidade. Além disto, há mútua relação entre a rede e as atividades urbanas, portanto a rede, além de impactar a acessibilidade, é também reflexo da localização das atividades, como preconiza a Sintaxe Espacial através das medidas de centralidades. Contudo, estes dados geográficos de atividades e da rede viária, de suma importância para análises de atração de viagens, são muitas vezes escassos ou indisponíveis. Este trabalho, portanto, busca testar a usabilidade de dados obtidos do mapa colaborativo *OpenStreetMap* (*OSM*), uma plataforma de Informações Geográficas Voluntárias, para a explicação da atração de viagens em Curitiba. Para tanto, a relação entre as centralidades da rede com as atividades econômicas da cidade é verificada a nível desagregado com o auxílio da Estimativa de Densidade de Kernel e, posteriormente, a nível agregado das zonas de tráfego, mostrando que atividades de comércio e serviços se encontram em vias mais acessíveis e em zonas intermediárias dos caminhos preferenciais. Após, investiga-se a qualidade dos dados de *POIs* e da rede, adquiridos na plataforma colaborativa, em comparação com dados oficiais da autoridade local, a fim de atestar a aplicabilidade das informações. No geral, o mapa online traz uma rede viária com qualidade equiparável à oficial, mas os *POIs* apresentam incoerências e devem ser usados com cautela. Por fim, modelos explicativos são calibrados com dados oficiais e dados do *OSM*, confrontando-se o comportamento para ambos os bancos de dados. Os *POIs* do *OSM* geram modelos piores, mas as centralidades são correlatas para o *OSM* e os dados oficiais na explicação da atração, atuando como complemento de dados incompletos e adicionando a informação de acessibilidade prioritariamente por caminhos retilíneos, como indica a Sintaxe Espacial. Assim, o trabalho fomenta o uso de mapas colaborativos como fontes alternativas de dados, demonstrando a maior usabilidade da rede viária representada.

Palavras-chave: Atração de viagens, Mapas colaborativos, Sintaxe Espacial, Centralidades, Pontos de Interesse.

Abstract

DAL BOSCO JR., A. Usability of Points of Interest and network centralities of Collaborative Maps for Trip Attraction Analysis: Case study of Curitiba. Master Thesis. São Carlos School of Engineering, University of Sao Paulo, São Carlos –SP, 2020.

It is crucial to understand travel demand for a sustainable transport planning. Trip attraction, for instance, is commonly analyzed and modeled considering both attractive activities locations – the Points of Interest (POIs) – and accessibility aspects. Likewise, there is mutual relationship between the network and the urban activities, therefore the network not only impacts the accessibility, but also reflects the activities placement, as stated by the Space Syntax through network centrality measures. However, these geographic data of road network and urban activities, which are greatly important for trip attraction analyses, are often scarce or even unavailable. Therefore, this work tests the usability of data acquired from OpenStreetMap (OSM) collaborative map, a Voluntary Geographic Information platform, to explain trip attraction in Curitiba. Hence, it verifies the relationship between network centralities and the city's economic activities at disaggregated level with the help of Kernel Density Estimation and, after, at aggregated level in traffic analysis zones, showing that commerce and service activities are located in accessible roads and intermediate zones for preferential paths. Then, it investigates the OSM POIs and network quality compared to official databases acquired from the local authority, aiming to assess this information's applicability. In general, the collaborative map presents a network with a similar quality to the official data, but the POIs have incoherencies and should be used with caution. Lastly, it calibrates explanatory models with both OSM and official databases to confront their behaviors. The OSM POIs generate worst models, however the OSM and official network centralities are compatible when explaining trip attraction, acting as complement for insufficient data and adding information of rectilinear paths accessibility, as asserted by the Space Syntax. In conclusion, this work fosters collaborative map usage as alternative data sources, suggesting a better usability for the represented road network.

Key-words: Trip Attraction, Collaborative maps, Space Syntax, Centralities, Points of Interest.

Lista de Figuras

Figura 1 – Estrutura e etapas do trabalho.	7
Figura 2 - Exemplo de grafo não-direcionado (a) e direcionado (b).	13
Figura 3 - Diferenciação entre grafo primal (a) e grafos duais obtidos com visadas retilíneas (b) e por segmento de via (c).	17
Figura 4 – Esquema da relação entre as centralidades de rede a atração de viagens.	18
Figura 5 – Esquema do método geral adotado nas etapas do trabalho.	27
Figura 6 - Localização da Região Metropolitana de Curitiba e seus municípios.	28
Figura 7 - Esquema da concepção dos eixos estruturantes de Curitiba.	29
Figura 8 – Região conurbada e localização dos eixos estruturantes da cidade.	30
Figura 9 – Municípios e zonas de tráfego da região conurbada.	31
Figura 10 - Exemplo de visualização do <i>OSM</i> (a) e diferenciação dos dados obtidos da plataforma (b).	32
Figura 11 – Processo de correção da rede para exclusão de nós intermediários.	33
Figura 12 – Rede viária obtida do <i>OSM</i>	33
Figura 13 – Quantidade e posição dos <i>POIs</i> em Curitiba.	35
Figura 14 – Rede viária oficial obtida com a prefeitura de Curitiba.	36
Figura 15 – Distribuição dos alvarás comerciais, pontos de ônibus e paraciclos obtidos do banco de dados oficial.	38
Figura 16 – Processo de obtenção das distâncias métrica e angular para cada aresta do grafo dual e posterior cálculo das centralidades.	39
Figura 17 – Resultados para centralidades de Intermediação na rede oficial e do <i>OSM</i>	41
Figura 18 – Resultados para centralidades de Proximidade na rede oficial e do <i>OSM</i>	42
Figura 19 – Resultados para centralidade de Grau na rede oficial e do <i>OSM</i>	43
Figura 20 – Distribuições das centralidades para a rede oficial e do <i>OSM</i>	44
Figura 21 - Esquema de diferenciação da produção e atração de viagens de acordo com a presença ou não da residência.	45
Figura 22 – Atração de Viagens total, por motivo e por modo nas ZTs de Curitiba.	46
Figura 23 – Esquemas de KDE em pontos sem peso (a) e com peso (b).	49
Figura 24 – Processo de análise das correlações entre centralidades e atividades econômicas.	51
Figura 25 – Resultado das KDE de cada centralidade.	52
Figura 26 – Distribuição de probabilidades das KDE de cada centralidade.	52
Figura 27 – Resultados das KDE de cada classe de alvará (Parte 1 de 2).	53
Figura 28 – Resultados das KDE de cada classe de alvará (Parte 2 de 2).	54
Figura 29 – Distribuição de probabilidade das KDE das classes de alvará.	55
Figura 30 – Correlações entre <i>rasters</i> das KDE de centralidades e de classes de alvarás.	57

Figura 31 – Correlações de Spearman entre estatísticas de centralidades e soma das classes de alvarás em cada ZT (Parte 1 de 2).	59
Figura 32 – Correlações de Spearman entre estatísticas de centralidades e soma das classes de alvarás em cada ZT (Parte 2 de 2).	60
Figura 33 – Processo de verificação da qualidade dos dados do <i>OSM</i>	70
Figura 34 - Esquema de análise da acurácia posicional da rede do <i>OSM</i> em relação à oficial. ...	71
Figura 35 - Resultado da completude da rede do <i>OSM</i> em relação à oficial para cada zona.....	76
Figura 36 – Acurácia posicional: porcentagem de elementos da rede <i>OSM</i> abrangidos segundo distância de <i>buffer</i>	77
Figura 37 – Erros relativos entre KDEs de centralidades para ambas as redes.	79
Figura 38 – Sobreposição visual das redes com apontamento de algumas incoerências.....	80
Figura 39 – Relação entre o número de <i>POIs</i> e o número de pontos oficiais em cada ZT.....	81
Figura 40–Histogramas e testes de hipótese para proporções de <i>POIs</i> e pontos oficiais nas ZT.	82
Figura 41 – Relação entre a proporção de <i>POIs</i> e a proporção de pontos oficiais em cada ZT....	83
Figura 42 – Processo de obtenção e análise dos modelos de regressão MQO.....	92
Figura 43 – Exemplo de transformação para a variável Atração de Todas as Viagens.	95
Figura 44 – Gráficos Q-Q para as variáveis dependentes antes e após a transformação Box-Cox com ajuste à distribuição normal.	96
Figura 45 – Correlações de Pearson entre variáveis dependentes de atração e variáveis explicativas tradicionais.....	96
Figura 46 – Correlações entre variáveis dependentes de atração e variáveis explicativas de centralidades de rede.	98
Figura 47 – Comparação entre modelos com dados oficiais e do <i>OSM</i> calibrados somente com variáveis explicativas tradicionais.....	101
Figura 48 – Comparação entre modelos com dados oficiais e do <i>OSM</i> calibrados somente com variáveis explicativas de centralidades.	103
Figura 49 – Comparação de modelos com dados oficiais calibrados somente com variáveis explicativas tradicionais (“Trad.”) e considerando variáveis de centralidade (“c/ Centr.”).	105
Figura 50– Comparação de modelos com dados do <i>OSM</i> calibrados somente com variáveis explicativas tradicionais (“Trad.”) e considerando variáveis de centralidade (“c/ Centr.”).	107

Sumário

1	INTRODUÇÃO.....	1
1.1	PANORAMA GERAL.....	1
1.2	OBJETIVOS.....	4
1.3	JUSTIFICATIVA.....	5
1.4	ESTRUTURA DO TRABALHO.....	7
2	REVISÃO BIBLIOGRÁFICA.....	9
2.1	A ATRAÇÃO DE VIAGENS E A ACESSIBILIDADE.....	9
2.2	AS REDES COMPLEXAS E AS CENTRALIDADES.....	13
2.3	A SINTAXE ESPACIAL NO CONTEXTO DAS VIAGENS.....	15
2.4	INFORMAÇÕES GEOGRÁFICAS VOLUNTÁRIAS.....	19
2.4.1	Web 2.0 e as IGV.....	19
2.4.2	<i>OpenStreetMap</i>	21
3	MÉTODO GERAL E OBJETO DE ANÁLISE.....	25
3.1	MÉTODO GERAL.....	25
3.2	ESTUDO DE CASO: CURITIBA.....	27
3.3	BANCOS DE DADOS.....	31
3.3.1	Dados do <i>OpenStreetMap</i>	31
3.3.2	Dados oficiais.....	35
3.3.3	Centralidades.....	38
3.3.4	Atração de Viagens.....	44
4	RELAÇÃO ENTRE AS CENTRALIDADES DE REDE E AS ATIVIDADES ECONÔMICAS.....	47
4.1	CONTEXTUALIZAÇÃO.....	47
4.1.1	Relação entre centralidades e dinâmicas socioeconômicas.....	47
4.1.2	<i>Kernel Density Estimation</i>	48
4.2	MÉTODO E PREPARAÇÃO DOS DADOS.....	50
4.2.1	Obtenção das superfícies de análise.....	51
4.2.2	Correlações a nível de rua e de zonas.....	56
4.3	RESULTADOS.....	56
4.4	CONSIDERAÇÕES FINAIS.....	62

5	AVALIAÇÃO DA QUALIDADE DOS DADOS DO OPENSTREETMAP	65
5.1	CONTEXTUALIZAÇÃO.....	65
5.1.1	A qualidade dos dados do <i>OSM</i>	66
5.1.2	Quesitos para a usabilidade.....	68
5.2	MÉTODO PARA AVALIAÇÃO DA QUALIDADE DOS DADOS	69
5.2.1	Avaliação da rede.....	70
5.2.2	Avaliação dos Pontos de Interesse	72
5.3	RESULTADOS	75
5.3.1	Resultados da avaliação da rede.....	75
5.3.2	Resultados da avaliação dos Pontos de Interesse	81
5.4	CONSIDERAÇÕES FINAIS.....	83
6	COMPARAÇÃO DOS MODELOS DE ATRAÇÃO DE VIAGENS	87
6.1	CONTEXTUALIZAÇÃO.....	87
6.2	MÉTODO.....	90
6.2.1	Calibração dos modelos explicativos.....	90
6.2.2	Verificação de premissas de regressão.....	93
6.3	RESULTADOS	95
6.3.1	Modelos com Variáveis Tradicionais	100
6.3.2	Modelos Apenas com Centralidades	102
6.3.3	Modelos com Variáveis Tradicionais e Centralidades.....	105
6.4	CONSIDERAÇÕES FINAIS.....	109
7	CONCLUSÕES E RECOMENDAÇÕES.....	111
7.1	CONCLUSÕES.....	111
7.1.1	Centralidades de rede e atividades econômicas.....	111
7.1.2	A qualidade dos dados do <i>OpenStreetMap</i>	112
7.1.3	Modelos explicativos de viagens atraídas com dados do <i>OpenStreetMap</i>	113
7.1.4	A usabilidade do <i>OpenStreetMap</i> para análise de atração de viagens	115
7.2	SUGESTÕES PARA TRABALHOS FUTUROS	117
	REFERÊNCIAS BIBLIOGRÁFICAS	121
	APÊNDICES	133

APÊNDICE 1 – NÚMERO DE ALVARÁS POR ZONA.....	135
APÊNDICE 2 – ESTATÍSTICAS ZONAS DE CENTRALIDADES DA REDE OFICIAL.....	136
APÊNDICE 3 – ESTATÍSTICAS ZONAS DE CENTRALIDADES DA REDE DO <i>OSM</i>	137
APÊNDICE 4 – NÚMERO DE PONTOS POR ZONA DAS VARIÁVEIS TRADICIONAIS OFICIAIS E DO <i>OSM</i>	138
APÊNDICE 5 - RESULTADOS DAS CALIBRAÇÕES DOS MODELOS	139

1 INTRODUÇÃO

1.1 PANORAMA GERAL

A análise dos deslocamentos urbanos traz informações valiosas para o planejamento de transportes. Convencionalmente, modelos com base em pesquisas domiciliares que representam as dinâmicas das viagens são base para tomadas de decisões, porém, frequentemente dependem de informações geográficas sobre as atividades urbanas e a infraestrutura de transporte para que os deslocamentos sejam modelados. A localização de atividades cotidianas relativas ao destino dos viajantes, por motivos de trabalho, estudo ou até lazer, é importante para se entender a relação entre a indução da demanda por transportes e o ambiente urbano (MCNALLY; RINDT, 2008; ORTÚZAR; WILLUMSEN, 2011). Variáveis relativas a comércios e estabelecimento de ensino, por exemplo, são utilizadas como variáveis explicativas da atração de viagens (SASIDHAR; VINEETH; SUBBARAO, 2016; AL-TAEI; TAHER, 2006), assim como variáveis pertinentes ao tipo de uso do solo (ESCAMILLA; COS; CÁRDENAS, 2016). As informações geográficas que qualificam as regiões urbanas são parte integrante de análises da atração de viagens, sendo as variáveis mais tradicionalmente aplicadas para tanto.

Além da localidade das atividades, a oferta de infraestrutura de transporte também é importante para entender a demanda por viagens, especialmente por impactar a acessibilidade percebida pelos viajantes (VICKERMAN, 1974). Locais mais acessíveis, assim, são mais atrativos, a depender do modo de transporte utilizado e o motivo da viagem (LEAKE; HUZAYYIN, 1980; CORDERA *et al.*, 2016). Ou seja, para o planejamento urbano, principalmente para a modelagem das viagens urbanas, as variáveis tradicionais – informações geográficas de atividades econômicas e de uso do solo – podem explicar a atração das viagens, mas a rede de transporte também é crucial para as análises, pois acrescenta aspectos de acessibilidade que impactam na atração das viagens, impacto que é explorado desde modelos comportamentais desagregados (MCNALLY; RINDT, 2008), até para modelos de uso do solo (IACONO; LEVINSON; EL-GENEIDY, 2008). Portanto, a configuração da rede pode ser importante para a explicação da demanda por viagens e possivelmente das próprias atividades econômicas, para além da acessibilidade, como preconiza o marco teórico da Sintaxe Espacial.

Hillier e Hanson (1984) apresentam a Sintaxe Espacial como estrutura de análise para a arquitetura, que projeta “visadas retilíneas” do espaço, mais relacionadas com a percepção humana

do ambiente, para entender quantitativamente o ambiente construído. Mais tarde, Hillier (1996) demonstra que, em redes viárias, esta percepção impacta e é afetada pela configuração das ruas e, por sua vez, afeta os movimentos urbanos e gera a mútua relação do uso do espaço urbano e da rede. Com a premissa da “economia do movimento”, em que as dinâmicas das viagens e o desenho urbano explicam o posicionamento de locais de interesse, a Sintaxe Espacial vai ao encontro do argumento de reciprocidade entre o uso do solo e a infraestrutura de transporte, em que um fator afeta o outro, ou seja, a rede é produto do uso do solo tanto quanto as atividades deste uso são produtos da configuração da rede (KELLY, 1994; GUZMAN; GOMEZ; RIVERA, 2017). Assim, por si só, a rede viária pode trazer informações sobre atividades atrativas e acessibilidade, enriquecendo análises de atração de viagens, mesmo sendo um elemento estático e pouco mutável, por induzir a posição das atividades e do uso do solo (OMER; GOLDBLATT, 2016).

Para observar a relação entre a rede e as atividades urbanas, as medidas de centralidades, advindas da teoria dos grafos e redes complexas, permitem a caracterização de cada elemento da rede no seu contexto global, expressando a tendência de movimentos e a acessibilidade (KURAMOTO, 2019), ajudando, assim, a explicar o posicionamento de atividades econômicas (PORTA *et al.*, 2012; OMER; GOLDBLATT, 2016). Foi com esta relação que Jayasinghe, Sano e Rattanaporn (2017) calibraram modelos de atração de viagens utilizando somente medidas de centralidade da rede viária de Colombo, no Sri-Lanka, mostrando que apenas a rede viária permite a explicação das regiões mais atrativas aos viajantes. Dessa forma, as informações geográficas da rede urbana se tornam cada vez mais valiosas para a análise da atração das viagens e, conseqüentemente, para o planejamento urbano.

Reforça-se, assim, a importância de dados geográficos relativos às variáveis explicativas tradicionais para viagens atraídas – como a localização do comércio, instituições de ensino, dentre outros –, e tendo em vista o efeito da acessibilidade e a base teórica da Sintaxe Espacial, a rede viária urbana pode ter impacto relevante para a análise da atração de viagens, principalmente ao refletir o posicionamento das atividades. Contudo, muitos países carecem de dados georreferenciados, seja das atividades ou da rede, tendo-se em vista o declínio dos esforços de mapeamentos oficiais por parte de agências governamentais (ESTES; MOONEYHAN, 1994). É neste contexto que os mapas colaborativos, plataformas online que abrigam informações geográficas, se tornam mais comuns, massificando a disponibilidade de dados que estão associados aos locais indutores de viagens e à própria rede viária.

As plataformas colaborativas surgem no contexto da chamada Web 2.0 (O'REILLY, 2007), em que usuários, antigamente receptores de informações, passam a ter um papel de assistentes, como uma consciência coletiva. É o novo modelo de uso da internet que aproveita a capilaridade e descentralização dos usuários para convergir dados, provendo serviços de forma mais irrestrita, especializada e sujeita ao escrutínio. Este processo é apelidado de “wikização” de informações, referindo-se à famosa iniciativa da Wikipédia – enciclopédia online totalmente dependente de adições e correções dos usuários e voluntários.

Plataformas colaborativas de dados geográficos, chamadas de Informações Geográficas Voluntárias (IGV), complementam, e até mesmo substituem, práticas tradicionalmente burocráticas e centralizadas de mapeamento. Até 2020, a mais famosa e difundida IGV gratuita é o *OpenStreetMap* (*OSM*), fundado em 2004, que apresenta crescente engajamento (GOODCHILD, 2007; HAKLAY; WEBER, 2008), sendo, portanto, uma promissora fonte de informações para o planejamento urbano e de transportes, sobretudo por ser mais atualizável e possivelmente apresentar maior detalhamento das informações (MACHADO; CAMBOIM, 2019).

Dependente apenas de edições voluntárias através de dispositivos GPS ou imagens aéreas (HAKLAY; WEBER, 2008), o *OSM*, como muitos mapas colaborativos, é uma proponente fonte de dados relativos a locais de interesse cotidiano através dos Pontos de Interesse (*POIs*, do inglês *Points of Interest*), e de dados de redes viárias (THILL; KIM, 2005). Portanto, podem não somente fornecer variáveis que tradicionalmente explicam a atração de viagens, como também de variáveis que representam a configuração da rede de forma mais acessível e atualizada. Porém, a qualidade dos dados destes mapas colaborativos pode impactar sua usabilidade.

O *OSM* tem apresentado bons resultados para a qualidade das redes viárias quando comparadas a fontes oficiais de agências de mapeamento na Inglaterra (HAKLAY, 2010), na França (GIRRES; YOUYA, 2010), na Alemanha (NEIS; ZIELSTRA; ZIPF, 2012) e em diferentes grandes cidades do planeta, como relatado por Neis, Zielstra e Zipf (2013), mas nenhuma cidade brasileira foi observada pelo estudo. Já a comparação dos *POIs* é mais difícil devido à carência de informações detalhadas em fontes oficiais (HOCHMAIR; JUHÁSZ; CVETOJEVIC, 2018; MACHADO; CAMBOIM, 2019). A boa qualidade recorrente da rede viária nos mapas colaborativos enseja que as centralidades tenham um papel importante de complementação das informações. Os *POIs*, por sua vez, estão diretamente relacionados com as variáveis explicativas

tradicionais para atração de viagens, sendo imperativo que também se observem sua qualidade e usabilidade. Por conseguinte, ao se considerar o acesso facilitado aos dados através do *OSM*, somado à crescente importância da rede viária e suas centralidades, as análises da atração de viagens podem se tornar mais acessíveis e diretas.

Usualmente, modelos de regressão linear que explicam as viagens através do método dos mínimos quadrados ordinários (MQO) são muito adotados por planejadores e operadores locais para análises de demanda por viagens, em particular para se observar a relação entre as variáveis e a atração. Portanto, pode-se observar como as informações adquiridas do *OSM* se comportam nestes modelos em comparação com informações oficiais adquiridas com a autoridade local. Ainda que estes modelos não contemplem por completo a teoria do comportamento dos viajantes e necessitem de verificações de premissas estatísticas, eles são amplamente conhecidos e podem ser equivalentes a modelos mais complexos ao explicarem as viagens de forma agregada e observarem relações entre as variáveis explicativas (GEORGE; KATTOR, 2013; CHANG *et al.*, 2014) principalmente se contemplam aspectos de fenômenos espaciais (LOPES; BRONDINO; RODRIGUES DA SILVA, 2014). Desta forma, os modelos MQO são ferramentas satisfatórias para explorar os dados e variáveis advindas do *OSM* e indicar sua usabilidade, tanto para os *POIs* quanto para as centralidades da rede viária, entendendo como estas afetam a explicação das viagens.

1.2 OBJETIVOS

Considerando a disponibilidade de dados do *OSM* e, em especial, o potencial da própria rede viária no contexto da Sintaxe Espacial, o estudo se dispõe a responder à seguinte pergunta: como é possível utilizar os dados de *POIs* e da configuração da rede do *OSM*, através das centralidades, para análises de atração de viagens? Para tanto, o objetivo geral é verificar a usabilidade dos dados de *POIs* e da rede do *OSM* para a explicação da atração de viagens para a cidade de Curitiba, Brasil.

Por se tratar de aplicação das centralidades de rede, informações não tradicionalmente usadas em modelos de atração de viagem, é necessário explorar como as centralidades podem se relacionar com as atividades atrativas, discutindo-se como esta relação traduz a interdependência entre o uso do espaço urbano e a infraestrutura de transportes. Além disso, o *OSM* como fonte dos dados é a proposição apresentada e, dada a incipiente utilização dos mesmos, é preciso

verificar as possíveis incoerências em relação aos dados oficiais obtidos com a prefeitura da cidade. As etapas do trabalho, portanto, visam atingir aos seguintes objetivos específicos:

1. Observar as correlações entre diferentes atividades econômicas e as centralidades da rede viária a fim de entender o papel da configuração da rede no posicionamento das atividades-fim de viagens urbanas;
2. Avaliar a qualidade dos dados adquiridos com o *OSM* para o uso proposto, a fim de determinar diferenças e semelhanças em relação aos dados oficiais equivalentes; e
3. Comparar modelos MQO calibrados com dados oficiais e do *OSM* para explicação da atração de viagens a fim de atestar a usabilidade do mapa colaborativo e compreender como as centralidades de rede auxiliam no processo.

1.3 JUSTIFICATIVA

O *OSM* e as plataformas colaborativas de IGV como um todo podem auxiliar em aplicações para o planejamento, operação e, possivelmente, políticas urbanas. Este estudo se ampara nos quesitos pontuados a seguir.

1. Busca por fonte alternativa de dados geográficos

O *OSM* se mostra uma fonte alternativa de dados para a análise de viagens, com potencial de atualização e cobertura dificilmente reproduzíveis por bases de dados oficiais. As informações do *OSM* estão disponíveis em todos os continentes e contam com um número crescente de colaboradores e elementos adicionados (HAKLAY; WEBER, 2008), o que reduz a dependência de autoridades e agências de mapeamento. A busca por alternativas para analisar a atração de viagens não só facilita a tomada de decisões em metrópoles já estabelecidas, como permite mais informação e possibilidades de uso para cidades menores. Contudo, acessar as limitações dos dados, comparando-os a dados oficiais, é crucial para compreender a situação atual das informações disponíveis.

2. Valorização de dados e plataformas *opensource*

Verificar a usabilidade para explicação de viagens atraídas confere valor e ponderação aos dados do *OSM* e de mapas colaborativos como um todo, seja no contexto acadêmico ou profissional, pois, mesmo que ainda incompletos, estão disponíveis e representam locais e elementos

importantes para as dinâmicas em ambientes urbanos. É interessante não apenas aplicar os dados, mas demonstrar como a qualidade das informações pode melhorar para a finalidade desejada. Ou seja, observar a usabilidade significa fomentar não somente a aplicação dos dados colaborativos, mas também a conscientização em relação ao fenômeno da descentralização de informações geográficas através das IGV, manifestando a importância que a qualidade dos dados possui para os fins práticos, aumentando as possibilidades advindas das plataformas. Este trabalho, assim, representa tanto um referencial para quando e como utilizar os dados de *POIs* e da rede viária para a explicação da atração de viagens, quanto uma explanação de como melhorá-los.

3. Utilização da rede viária para explicação da atração de viagens

Ao considerar as premissas da Sintaxe Espacial para aplicar as centralidades da rede viária nos modelos explicativos, este estudo contribui para o conhecimento da configuração da rede na sua mútua influência sobre as dinâmicas urbanas, destacando o papel que a acessibilidade e os potenciais de movimentos desempenham nas atividades econômicas e nas próprias viagens. A rede, portanto, pode fornecer parte das informações necessárias para analisar a atração de viagens de forma global, complementando a análise. Assim, aprofunda-se a compreensão da rede como variável exógena da atração de viagem e apresentam-se indicativos de como a configuração da rede pode representar um reflexo, com o passar do tempo, das próprias viagens.

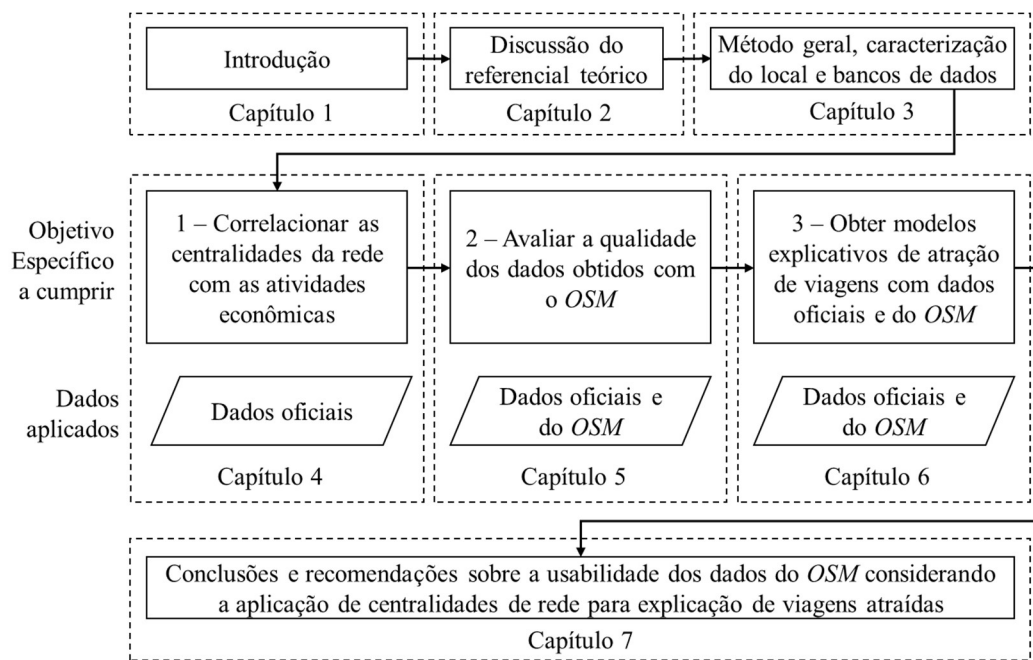
4. Busca por prática alinhada às tendências de massificação de informações

Dado o valor próspero dos dados em questão, o ponto principal do *OSM* é a disponibilidade massiva de *POIs* e de redes. Nesse sentido, a aplicação local dos dados colaborativos em Curitiba, além de estudo de caso para a atração das viagens, é parte da conjectura de *Big Data* para informações georreferenciadas, abundantes por consequência da capilaridade e frequência das IGV na Web 2.0. Esta abundância é benéfica para, e ao mesmo tempo é produto da, computação ubíqua, ou seja, da massificação de dispositivos programáveis, como GPS e celulares, e da Internet das Coisas, características importantes para a Indústria 4.0 (XU; XU; LI, 2018). Portanto, a aplicação dos dados do *OSM* está alinhada às novas tendências de serviços e processos, legitimando cada vez mais o *OSM* e os mapas colaborativos para análises práticas de transporte urbano.

1.4 ESTRUTURA DO TRABALHO

O trabalho inicia com esta introdução e posteriormente apresenta o referencial teórico que baseia as etapas subsequentes. Após, o método geral é constituído, caracterizando Curitiba como local das observações e apresentando os bancos de dados utilizados. Com o fim de atingir cada um dos objetivos específicos, os próximos capítulos segmentam as etapas necessárias, tal como demonstra a Figura 1.

Figura 1 – Estrutura e etapas do trabalho.



Fonte: O autor (2020).

A primeira etapa, Capítulo 4, busca replicar os estudos embasados na Sintaxe Espacial que visam a entender como a configuração da rede viária se relaciona com as dinâmicas urbanas, especialmente com as atividades econômicas. Assim, os resultados permitem constatar possíveis correlações entre as centralidades da rede e as atividades tradicionalmente explicativas da atração das viagens, corroborando os preceitos da Sintaxe Espacial, e o fato de a acessibilidade ser indutora de atividades e viagens. Com resultados positivos, a rede se torna aplicável para entender a atração de viagens. Esta etapa utiliza somente os dados oficiais, ou seja, fornecidos pela prefeitura de Curitiba através do Instituto de Pesquisa e Planejamento Urbano (IPPUC), órgão responsável pelo planejamento urbano do município. Desta forma, a relação entre centralidades e atividades econômicas é observada com dados já validados pela prática local.

Antes da etapa de calibração dos modelos explicativos, o Capítulo 5 avalia a qualidade dos dados obtidos com o *OSM*, comparando-os aos oficiais. Sabe-se que as diferenças entre os bancos de dados não necessariamente são falhas do *OSM*, pois alguns dados podem estar representados na plataforma colaborativa, mas não no banco de dados oficial. Porém, a comparação permite quantificar estas diferenças, tanto para a rede quanto para algumas classes de *POIs*, e entender possíveis incoerências nos modelos da última etapa.

Portanto, os modelos gerados na terceira etapa, Capítulo 6, complementam as duas primeiras etapas e permitem a comparação dos dados oficiais com os dados do *OSM* na explicação das viagens atraídas. Os resultados contribuem tanto para o contexto teórico da Sintaxe Espacial, ao adicionar as centralidades da rede como variáveis explicativas, quanto para a usabilidade das informações do *OSM* em análises de viagens, ressaltando-se que os dados aplicados não esgotam as variáveis explicativas utilizadas em modelos de atração, pois os modelos são utilizados como ferramentas de comparação dos dados em si, não tendo caráter preditivo. O objetivo, ressalte-se, é comparar os dados e seu comportamento na análise das viagens, ainda que os resultados auxiliem em aplicações mais extensivas da modelagem da demanda por viagens.

O trabalho finaliza com conclusões acerca da usabilidade dos dados do *OSM* bem como da aplicação de centralidades da rede viária para a análise da dinâmica das viagens, discutindo como as centralidades, o *OSM* e possivelmente outras IGV podem ser importantes ferramentas para análises de transporte. Também são sugeridas diferentes abordagens e aplicações possibilitadas pelas conclusões do presente trabalho.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo são apresentados os fundamentos teóricos sobre a atração de viagens e sua relação com a infraestrutura viária, mostrando como a Sintaxe Espacial pode auxiliar em análises globais da rede urbana e as atividades humanas, o que destaca as possibilidades dos dados de Informações Geográficas Voluntárias de explicarem a atração de viagens.

2.1 A ATRAÇÃO DE VIAGENS E A ACESSIBILIDADE

A análise de origens e destinos dos deslocamentos é parte importante do planejamento de transportes. Modelos que buscam replicar as dinâmicas das viagens são ferramentas para análises urbanas, como tradicionalmente faz o modelo de quatro etapas, o qual consiste na geração das viagens, distribuição, escolha modal e alocação. A primeira etapa, geração de viagens, se segmenta em produção e atração e é o ponto de partida para a análise da demanda por transportes (ORTÚZAR; WILLUMSEN, 2011). A produção advém de variáveis socioeconômicas relacionadas aos viajantes (ex.: renda, motorização e idade), ao passo que a atração é muito explicada pelas atividades-fim, ou seja, pelas variáveis socioeconômicas dos atrativos, como uso do solo (ESCAMILLA; COS; CÁRDENAS, 2016), número de empregados e de estabelecimentos comerciais (GEORGE; KATTOR, MALIK, 2013, SASIDHAR; VINEETH; SUBBARAO, 2016) e número de estabelecimentos de ensino (AL-TAEI; TAHER, 2006), e também por informações indiretas sobre locais atrativos, como estacionamentos (INNES; IRCHA; BADOE, 1990).

Modelos como o “quatro etapas” adotam uma abordagem baseada nas viagens, ou seja, observa as viagens de forma agregada e busca relacioná-las com variáveis explicativas. Por outro lado, esta abordagem não traz em sua estrutura a base teórica de que as viagens são produtos comportamentais das atividades desejadas pelos viajantes, as atividades-fim (KITAMURA, 1988). Reproduzir o comportamento dos viajantes, fundamentando-se nas atividades desejadas, possibilita maior desagregação da análise, porém, exige ajustes da relação entre indivíduos e o ambiente, em que as rotinas diárias estão em equilíbrio e são dependentes da acessibilidade que o ambiente fornece para que os indivíduos participem das atividades (LUNDBERG, 1988). McNally e Rindt (2008) discutem como a abordagem nas atividades fortalece a teoria basal do comportamento de viagens, trazendo, inclusive, melhorias para modelos como o “quatro etapas”, por expandir entendimentos quanto à relação entre locais atrativos e as impedâncias que

podem facilitar ou dificultar viagens, impactando até mesmo a distribuição, escolha modal e até a alocação dos deslocamentos.

A acessibilidade, de fato, era considerada importante mesmo para modelos agregados, como verifica Vickerman (1974), que considerou pressupostos econométricos de que a redução das impedâncias, ou a facilidade para atingir os objetivos (neste caso, as atividades-fim) é também fator para os deslocamentos, impactando o potencial de atração de viagens ou mesmo a geração da demanda por viagens. Leake e Huzayyin (1980) comentam que a acessibilidade pode se dar em relação à estrutura da rede de transporte, à resistência atribuída aos trechos da rede e ao nível de serviço ou provimento de transporte, sendo variável conforme a localização, modos utilizados e horários do dia. Ambos os estudos mostram que viagens opcionais – não relacionadas a trabalho ou estudo – são mais afetadas por medidas de acessibilidade, e que o uso de automóveis é inversamente proporcional a estas medidas. Os motivos das viagens, portanto, também impactam as escolhas dos destinos, assim como os modos utilizados.

A base teórica da influência da acessibilidade para a geração de viagens é de que as viagens são produtos secundários, e que a provisão de serviços e infraestrutura de transporte induz a demanda. Neste sentido, é comum que as medidas de acessibilidade se baseiem em distâncias relativas entre os extremos dos deslocamentos, ou que considerem taxas de atratividade dividida por impedâncias espaciais ou temporais, também conhecidas como acessibilidade gravitacional (THILL; KIM, 2005; CORDERA *et al.*, 2016). Isto mostra que a atratividade é intimamente relacionada à possibilidade dos deslocamentos. Não obstante, também foi demonstrado que variáveis que buscam traduzir a acessibilidade podem ter pouca significância (EWING; DEANNA; LI, 1996), ou que são necessárias diferentes medidas de acessibilidade para melhor observar as dinâmicas das viagens (THILL; KIM, 2005). Outro fenômeno recorrente em todos os estudos citados é a multicolinearidade entre as variáveis tradicionais e as variáveis de acessibilidade. Ou seja, há muita correlação entre estas variáveis, o que dificulta a aplicação em modelos, pois acabam representando um mesmo poder explicativo. Isto pode indicar que as medidas que já consideram a atratividade podem não ser razoáveis (LEAKE; HUZAYYIN, 1980), ou ainda que a acessibilidade, de modo geral, afeta as próprias variáveis tradicionais (ex.: que locais mais acessíveis induzem o posicionamento de lojas ou que uma maior acessibilidade, do ponto de vista de distâncias, reduz a necessidade de adquirir automóveis) (EWING; DEANNA; LI, 1996).

Cordera *et al.* (2016) buscaram responder se há, de fato, importância na acessibilidade para a geração das viagens, controlando possíveis colinearidades e dependências espaciais. Para tanto, utilizaram três medidas: tempo de viagem até o centro da cidade e acessibilidade gravitacional entre zonas com e sem o uso de automóvel. As duas últimas consistem em funções que agregam os atrativos da zona de destino e o tempo de viagem a partir da zona de origem. Os resultados corroboram novamente os de Vickerman (1974), de que as medidas de acessibilidade têm maior relação com viagens opcionais e relação inversa com viagens motorizadas privadas. O tempo de viagem ao centro coloca nas equações a propensão a este destino, existente em cidades com um núcleo comercial bem definido, e se mostra tão importante quanto as medidas gravitacionais. Os modelos que corrigiram as dependências espaciais deram menor importância às acessibilidades.

A dependência espacial é comum para dados com dimensão locacional, como é o caso da demanda por viagens, o que significa que existem correlações nas observações, e que as relações modeladas apresentam heterogeneidade espacial (LESAGE, 1999). A atração de viagens, por exemplo, tende a se concentrar em zonas vizinhas centrais, com as áreas periféricas tendo em comum a baixa atratividade, especialmente em cidades com centros bem definidos. A análise das relações entre as variáveis explicativas e a variável explicada pode não ficar clara quando este efeito não é considerado. Wang *et al.* (2012) apresentam uma revisão bibliográfica tratando do tema, pontuando que a existência de correlações espaciais pode advir da representação escolhida para agregar as informações (ex.: zonas), ou ainda ser indício da omissão de importantes variáveis para o evento modelado. Portanto, quando Cordera *et al.* (2016) controlam este efeito, as variáveis de acessibilidade respondem por enfraquecer seu poder explicativo, trazendo mais importância às variáveis tradicionais e fortalecendo a hipótese de que a acessibilidade tem o mesmo poder explicativo das variáveis socioeconômicas, possivelmente por explicarem as próprias variáveis tradicionais (VICKERMAN, 1974; EWING; DEANNA; LI, 1996).

Ademais, essa relação recíproca entre acessibilidade e as variáveis socioeconômicas é também base para modelos de interação de uso do solo e transporte (*Land Use Transport Interaction – LUTI*). Os primeiros modelos LUTI consideram medidas de impedância entre regiões que impactam o posicionamento de residências (LOWRY, 1964). Posteriormente, modelos econométricos de mercado imobiliário também consideram o posicionamento de atividades econômicas como variáveis endógenas – não fornecidas, mas resultantes dos próprios modelos –, como os

modelos MEPLAN, utilizados para a modelagem regional dos impactos do túnel entre a Inglaterra e a França (ECHENIQUE, 2004), e o MUSSA, para a cidade de Santiago no Chile (MARTINEZ, 1996). Portanto, a infraestrutura viária, que molda as características de acessibilidade, impacta no posicionamento de atividades econômicas e a população que, por sua vez, impactam a própria geração de viagens. Isto passa a ser visto não apenas como marco teórico (KELLY, 1994), mas em modelos que demonstram esta relação.

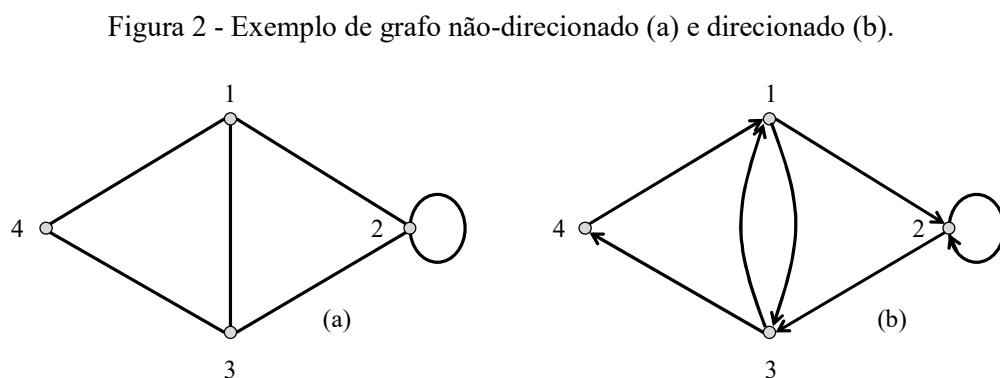
É interessante notar que estes modelos LUTI podem ter sua arquitetura formada por submodelos que alimentam as variáveis uns dos outros, por vezes modelos simples de regressão linear como o modelo aplicado na cidade de Bogotá, na Colômbia (GUZMAN; GOMEZ; RIVERA, 2017), que utilizou a abordagem de dinâmica de sistemas para estimativas de geração de viagens, com a produção observada por residência e a atração com agregação zonal, considerando-se diversos mecanismos de retroalimentação. O modelo teve bons resultados para previsão de viagens e era capaz de estimar tendências de crescimento urbano, uso do solo e de emissões. Sua confiabilidade reforçou a dependência de condições de acessibilidade para escolhas de posicionamento de atividades. Ademais, Iacono, Levinson e El-Geneidy (2008) trazem uma revisão sobre o surgimento e a aplicação de modelos LUTI, expondo inclusive algumas dificuldades na consideração de redes de transporte e aspectos de acessibilidade.

Por conseguinte, observar a atração de viagens e a acessibilidade proporcionada pela infraestrutura de transportes traz importantes ponderações. A acessibilidade chama a atenção por ser uma forte base teórica para a demanda de transporte, mas dificilmente observável com modelos empíricos. Contudo, as avaliações apontam para a reciprocidade entre as variáveis tradicionalmente explicativas das viagens e a infraestrutura (ex.: uso do solo e os sistemas de transporte, em especial a rede urbana), relação discutida há décadas e que começa a tomar forma para além dos marcos teóricos, como ensinava Kelly (1994). A rede de transporte é formada por dinâmicas urbanas, mas também molda estas dinâmicas, o que pode afetar o uso do solo e as atividades que influenciam as dinâmicas, levando a mudanças na própria rede, seja por meios informais ou impostos por autoridades administrativas. Hillier (1996) corrobora esta relação, trazendo o marco teórico da Sintaxe Espacial, que permite mensurar a configuração do espaço urbano, especialmente a rede viária, e ajuda a explicar grande parte dos eventos socioeconômicos das cidades.

Através de ferramentas das redes complexas, a Sintaxe Espacial busca representar a rede de transportes considerando princípios da percepção humana, traduzindo os elementos da rede conforme sua importância no contexto morfológico com as chamadas centralidades da rede. Esta abordagem permite obter informações posicionais da rede e confrontá-las com outras informações geográficas, como atividades econômicas e a própria atração de viagens (JAYASINGHE; SANO; RATTANAPORN, 2017). Portanto, a rede viária pode ser observada de forma global, e não somente como uma medida de oferta e impedância entre pontos no ambiente urbano, como comumente se aplica em modelos de demanda por viagens (CORDERA *et al.*, 2016). No contexto de massificação de informações geográficas em plataformas online colaborativas, este marco teórico propicia informações valiosas e facilmente adquiríveis. Porém, é preciso entender como as redes complexas podem ajudar nesta análise.

2.2 AS REDES COMPLEXAS E AS CENTRALIDADES

As redes complexas são comumente aplicadas nas ciências sociais e na antropologia com o intuito de caracterizar relações interpessoais de indivíduos e grupos (BAVELAS, 1948; WASSERMAN; FAUST, 1994). A ciência das redes complexas, no entanto, tem aplicação nas mais diversas áreas do conhecimento, como na biologia, na medicina, na internet, nas comunicações, na computação, na economia e nos transportes (WASSERMAN; FAUST, 1994; COSTA *et al.*, 2011). O objeto de análise das redes são os grafos, diagramas formados por nós e arestas como demonstrado na Figura 2.



Fonte: O autor (2020).

Um grafo $G = (N, E)$ é constituído por um conjunto N de nós e um conjunto E de pares de nós i e j do conjunto N , e esses pares definem arestas descritas como (i, j) (COSTA *et al.*, 2011; PORTA; CRUCITTI; LATORA, 2006a). A diferença entre os grafos da Figura 2(a) e da Figura

2(b) é que o primeiro representa um grafo não-direcionado, em que uma aresta representa uma conexão de duplo sentido entre os nós. Já a segunda demonstra um grafo direcionado, ou dígrafo, em que as arestas têm sentido definido e os caminhos somente podem percorrer tais sentidos, representação mais utilizada para redes urbanas, por exemplo.

Bavelas (1948) utilizou grafos para observar relações interpessoais de indivíduos, caracterizando a importância dos elementos, ou indivíduos, mais “centrais” em grupos com cadeias hierárquicas, observando a independência, influência e controle social de cada elemento (BIGGS; LLOYD; WILSON, 1998). Para tanto, o autor introduz as medidas de centralidades, que trazem informações de cada elemento no contexto da rede em que se encontra. Para redes de transporte terrestre, as centralidade são fundamentais para observar a vulnerabilidade e resiliência das redes viárias (TAYLOR; D’ESTE, 2003; HAWICK; JAMES, 2007; MORELLI, 2019), a conectividade e acessibilidade de redes de transporte coletivo (SCHEURER; CURTIS; PORTA, 2007; SOUZA; NETO; FERREIRA, 2018; PUMAR; BARBASTEFANO; CARVALHO, 2018), bem como para caracterizar as diferentes redes viárias quanto à própria topologia, dispersão de informação e à ordenação (JIANG; CLARAMUNT, 2004; ROSVALL *et al.*, 2005; BOEING, 2019). Também são os grafos e as centralidades de redes urbanas que permitem as análises de Sintaxe Espacial (HILLIER, 1996).

A ideia de centralidade de um nó k , por exemplo, pode ser intuitiva ao se observar a aglomeração em torno do mesmo, ou seja, o número de nós adjacentes a k . Também pode ser entendida como o número de arestas conectadas ao ponto (FREEMAN, 1979). Esta medida é chamada de grau (*degree* em inglês) (C_D), definido por

$$C_D(k) = \sum_{i=1}^N a(i, k) \quad [2.1]$$

em que:

k : nó em questão;

$a(i, k)$: igual a 1 quando o nó i é conectado a k por uma aresta e 0 caso contrário;

N : total de nós do grafo.

Outra forma de se observar a centralidade de um elemento é sua “proximidade” (*Closeness*, em inglês) (C_C) ao restante da rede. Ou seja, um nó k é mais central quando está a distâncias menores de todos os outros da rede (FREEMAN, 1979). Portanto,

$$C_C(k) = \frac{N - 1}{\sum_{i=1}^N d(i, k)} \quad [2.2]$$

em que:

$d(i, k)$: distância entre o nó k e qualquer outro nó i do grafo.

Há, também, a relação que observa a importância de um elemento da rede para os caminhos dentro dela, que é a intermediação (do inglês, *betweenness*) (C_B). Esta centralidade obtém quantitativamente o número de caminhos mínimos que passam por um nó k em relação ao total de caminhos mínimos na rede com

$$C_B(k) = \frac{\sum_{i, j \in N} g_{ij}(k)}{\sum_{i, j \in N} g_{ij}} \quad [2.3]$$

em que,

i e j : quaisquer outros nós do grafo satisfazendo: $i \neq j \neq k$;

g_{ij} : igual a 1 quando há um caminho entre i e j ; e 0, caso contrário.

$g_{ij}(k)$: igual a 1 quando há um caminho entre i e j que passa por k , e 0 caso contrário.

Ao se observarem redes urbanas, as medidas de centralidade permitem encontrar relações entre os elementos que podem não ser intuitivas quando analisadas independentemente da configuração espacial completa. De fato, as centralidades têm sido relacionadas a diversas dinâmicas das cidades, especialmente com base nos preceitos teóricos da Sintaxe Espacial (HILLIER, 1996). As centralidades, assim, não restringem as observações a medidas relativas entre zonas representadas por seus centroides, o que geralmente ocorre em modelos de demanda que buscam representar a acessibilidade (THILL; KIM, 2005), mas fornecem medidas desagregadas ao longo da rede, que podem ou não serem agregadas.

2.3 A SINTAXE ESPACIAL NO CONTEXTO DAS VIAGENS

O crescimento e a formação da malha viária urbana são produtos da relação recíproca entre a infraestrutura de transporte e o uso do solo, moldados também por escolhas de planejamento. Kelly (1994) discute como o uso do solo afeta e é afetado pela estruturação da rede, e diversos

modelos tentam replicar esta dinâmica (LOWRY, 1964; ECHENIQUE, 2004; GUZMAN; GOMEZ; RIVERA, 2017), especialmente buscando considerar a acessibilidade para auxílio na modelagem de posicionamento de atividades (IACONO; LEVINSON; EL-GENEIDY, 2008). A Sintaxe Espacial, neste sentido, relaciona a percepção do espaço e as dinâmicas sociais, demonstrando a dependência mútua entre o movimento das pessoas e a economia urbana, ao que se nomeia de premissa da “economia do movimento”. Isto é, as dinâmicas socioeconômicas moldam a rede urbana, que por sua vez molda os movimentos cotidianos; ao mesmo tempo, estes movimentos reforçam ou modificam o posicionamento de atividades econômicas que atraem as viagens, retroalimentando o ciclo (HILLIER, 1996).

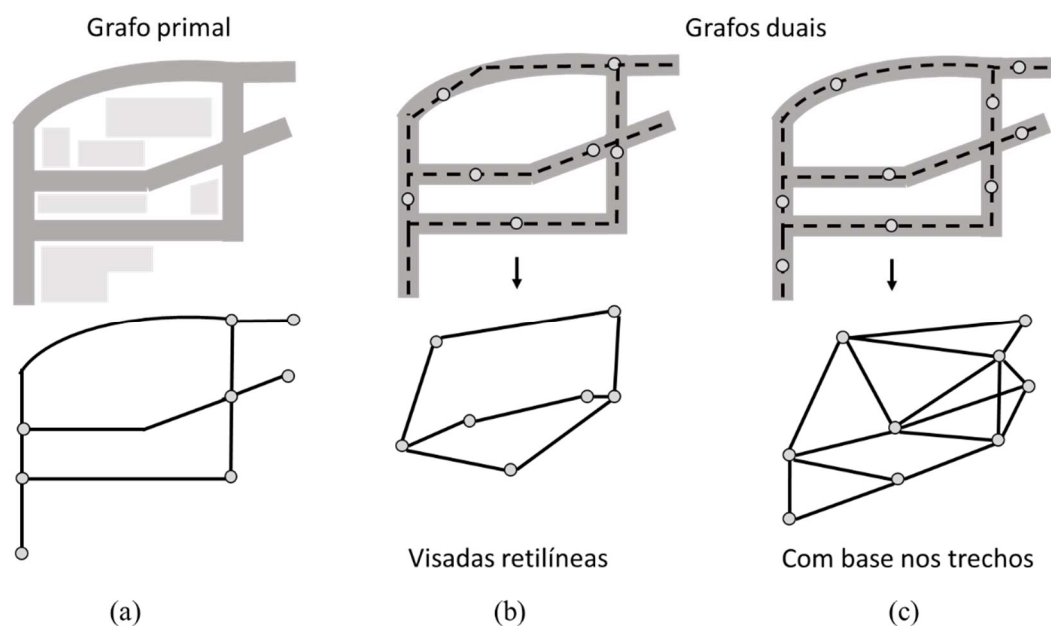
Jacobs (1961) já discutia que, para além da acessibilidade, muitas atividades econômicas buscam ao mesmo tempo um posicionamento que usufrua dos movimentos na cidade, ou seja, locais que apresentem circulação de pessoas a depender do modo e da natureza da atividade. Neste sentido, Jiang e Jia (2011) mostram como a configuração da rede viária pode explicar grande parte da distribuição do tráfego, ao compararem simulações baseadas em agentes com comportamento aleatório ao longo da rede de Londres com agentes que definem origens e destinos, buscando o menor caminho topológico. Jiang e Liu (2009) correlacionaram medidas de centralidade com o tráfego médio diário anual na cidade de Hong Kong, obtendo correlações maiores do que 0,8 para a centralidade de proximidade local, que considera até dois nós de distância do nó observado, muito utilizada como medida de acessibilidade em curtas distâncias na rede. Já Jiang e Jia (2011) demonstram que intermediação e conectividade podem melhor representar fluxos agregados do que a acessibilidade. As centralidades, assim sendo, por traduzirem a configuração da rede viária relativa à proximidade, intermediação e conectividade, podem auxiliar na explicação de movimentos na rede, acessibilidade e, como resultado, no posicionamento de atividades.

Os resultados de quaisquer estudos com centralidades, no entanto, dependem também da forma de representação dos grafos (TURNER, 2007; GAO *et al.*, 2013). Na economia do movimento, premissa importante para a Sintaxe Espacial, a percepção do ambiente pelos indivíduos, representada por linhas retas a partir do observador, é a base para a representação da rede urbana. As dinâmicas humanas e sua relação com o espaço físico estariam sujeitas a percepção retilínea. Em um ambiente urbano, o observador está restrito pela configuração das vias, pois as edificações comumente formam bloqueios visuais em ambos os lados dos caminhos. Os trabalhos de

Hillier e Hanson (1984) e Hillier (1996) argumentam que esta premissa se relaciona mais fielmente aos movimentos urbanos, ou “movimento natural” e seus derivados, como o posicionamento de atividades econômicas altamente dependente de passagem de clientes (ex.: varejo), observando principalmente a centralidade de proximidade, também chamada de integração, mas também a centralidade de intermediação. Porta, Crucitti e Latora. (2006b) expõem que, apesar de centralidades isoladas auxiliarem nos resultados, análises de centralidades múltiplas podem prover maiores explicações, mas se deve atentar para a representação utilizada da rede.

Dessa forma, diferente da representação tradicional das redes urbanas em SIG, em que as interseções e fins de rua são os nós, e os trechos de rua são as arestas dos grafos, chamada de representação *primal* (Figura 3(a)), os grafos originários para análises de Sintaxe Espacial se baseiam nas visadas retilíneas, transformando trechos retos em nós, e as conexões entre estes trechos em arestas (Figura 3(b)). Não obstante, há também a transformação dos trechos das vias em nós e suas conexões em arestas (Figura 3(c)). As duas últimas são chamadas de representações *duais* (PORTA; CRUCITTI; LATORA, 2006b). A última representação, por trechos das vias, é mais intuitiva e pode ser base para análises em Sintaxe Espacial, caso se adote a distância ou impedância angular para cálculo das centralidades, garantindo a percepção-base para a Sintaxe Espacial (TURNER, 2007).

Figura 3 - Diferenciação entre grafo primal (a) e grafos duais obtidos com visadas retilíneas (b) e por segmento de via (c).

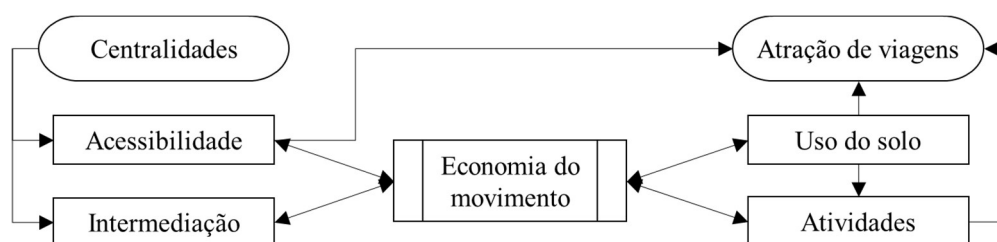


Fonte: O autor (2020).

Turner (2009) analisou as rotas de motociclistas em Londres, constatando que a minimização dos ângulos teve um papel cognitivo mais importante em comparação com a minimização de distâncias físicas. Serra e Hillier (2019), por sua vez, adotam a representação baseada em trechos da rede completa da Grã-Bretanha, comparando as centralidades com o tráfego de diferentes modos. Os autores obtêm as centralidades considerando a distância métrica, entendida como impedância energética (distância em metros dos caminhos) e a distância angular, ou impedância de informação (soma dos ângulos ao longo dos caminhos), mostrando que esta última tem correlações maiores do que aquela, aumentando em 21% a média dos coeficientes.

Estudos baseados na “economia do movimento” analisam outros fenômenos urbanos relacionados às centralidades de rede, como a dispersão de atividades econômicas (PORTA *et al.*, 2012; OMER; GOLDBLATT, 2016), uso do solo (RUI; BAN, 2014; LIU *et al.*, 2016) e, como este trabalho, a própria atração de viagens (JAYASINGHE; SANO; RATTANAPORN, 2017). Muitos deles utilizam Estimativa de Densidade de Kernel (KDE, do inglês *Kernel Density Estimation*), aplicada aos resultados, comparando as superfícies de densidade geradas, uma vez que a natureza geométrica dos resultados de atividades e centralidades é diferente (PORTA *et al.*, 2012; RUI; BAN, 2014; LIU *et al.*, 2016; JAYASINGHE; SANO; RATTANAPORN, 2017). Todos mostram que a configuração das redes urbanas explica, mesmo que parcialmente, as respectivas dinâmicas sociais observadas, corroborando que aspectos de acessibilidade e movimentos ajudam a explicar o posicionamento de locais atrativos. Também demonstram que a impedância angular pode trazer melhorias para estas análises, como sugerido pela Sintaxe Espacial. O papel das centralidades, portanto, está tanto relacionado à acessibilidade – especialmente quando considerada a centralidade de proximidade –, quanto ao potencial de movimentos – intermediação – (OMER; GOLDBLATT, 2016), refletindo, e possivelmente estimulando, os usos socioeconômicos do espaço. A Figura 4 esquematiza esta relação.

Figura 4 – Esquema da relação entre as centralidades de rede a atração de viagens.



Fonte: O autor (2020).

As centralidades, por expressarem acessibilidade e intermediação, têm relação com o uso do solo e as próprias atividades através da premissa da economia do movimento da Sintaxe Espacial. O posicionamento das atividades, por sua vez, explica a atração de viagens. Contudo, estas relações são mútuas, pois acessibilidade e uso da oferta de infraestrutura são impactados pelo próprio uso do solo e posicionamento das atividades. Ademais, a acessibilidade tem impacto direto na atração de viagens, pois locais mais acessíveis tendem a atrair mais viajantes. Assim, acessibilidade e intermediação podem ser indicadores de regiões atrativas, e a rede viária, observada a partir das centralidades, pode ser considerada fonte de dados para análise de viagens atraídas, não somente como a infraestrutura por onde as viagens acontecem, mas como o reflexo do uso do espaço.

Assim sendo, as redes viárias possivelmente são dados geográficos tão importantes quanto as variáveis mais tradicionalmente usadas para explicar a atratividade de um local. Isto traz ainda mais destaque aos mapas colaborativos, que disponibilizam informações de rede e de atividades cotidianas.

2.4 INFORMAÇÕES GEOGRÁFICAS VOLUNTÁRIAS

Tradicionalmente, o mapeamento de locais e suas características topográficas é papel atribuído a agências com estruturas hierárquicas, altamente dependente de especialistas e com atualizações lentas e custosas. Estes e Mooneyhan (1994) argumentam que há um declínio nos esforços governamentais para atualização e melhoramento de dados geográficos oficiais, com exceção de poucos países. Mesmo no Brasil, a cobertura de mapas e dados geográficos detalhados é escasso e desigual (CAMBOIM; BRAVO; SLUTER, 2015). Goodchild (2007) discute estas questões e demonstra o surgimento de alternativas que podem suplementar agências de mapeamento: as chamadas Informações Geográficas Voluntárias (IGV), que são parte das iniciativas que compõem a chamada Web 2.0. Informações geográficas gratuitas possibilitam a aplicação rápida e universalizada, sendo, portanto, valiosas para estudos em transporte, especialmente por trazerem dados relacionados à rede urbana e a pontos de interesse cotidiano.

2.4.1 Web 2.0 e as IGV

O termo Web 2.0 é utilizado por O'Reilly (2007) para definir o conjunto de padrões observados em produtos e plataformas online que deixam de utilizar a internet como uma mídia restrita

para disposição de páginas e sites, e passam a utilizar a estrutura e a permeabilidade da internet para atuação. Em outras palavras, enquanto a Web 1.0 era apenas um meio para adicionar dados e conteúdo de forma supervisionada e estática, a Web 2.0 considera os usuários como colaboradores e as relações entre eles como a ‘consciência coletiva’ da rede, onde os dados, e seu manejo, são partes fundamentais dos serviços prestados. Um exemplo seria o Google Search – que busca páginas mais acessadas, considerando a estrutura de links para melhores resultados. Outro seria a enciclopédia online Wikipedia – que confia integralmente na adição e controle desregulado das informações pelos usuários. O primeiro não apresenta dados abertos, diferentemente do segundo, mas ambos utilizam a abordagem *bottom-up* para atuarem, dependendo inteiramente dos usuários para a evolução dos serviços.

Nesse sentido, as IGV estão entre as tecnologias mais marcantes na Web 2.0, segundo Goodchild (2007), pois utilizam dados obtidos de forma desagregada por usuários ou voluntários e, com eles, moldam plataformas com dados georreferenciados. Alguns produtos, sob propriedade de companhias privadas, representam os conceitos de IGV: *MapMaker (Google)*, *Here Map Creator (Nokia)*, *Map Share (Tom-Tom)* e *Waze* (adquirido pela *Google*). Contudo, iniciativas como o *OpenStreetMap*¹ (*OSM*) e o *WikiMapia*², fundadas em 2004 e 2006, respectivamente, são IGV com dados abertos, editáveis e acessíveis publicamente (NEIS; ZIELSTRA, 2014).

As IGV, especialmente de dados abertos, são tecnologias advindas do “público” e não de “experts”, como menciona Elwood (2009). A autora discute como a facilidade para adição e aquisição dos dados, geralmente a partir de interfaces de aplicação de programação (*Application programming interfaces* – APIs), podem representar tanto avanços democráticos e de empoderamento cívico quanto erosão de privacidade e exclusão. A geolocalização massiva e facilitada de locais, fenômenos e até pessoas traz consigo diversas questões. Contudo, apesar da controvérsia, as IGV são fenômenos crescentes e trazem a possibilidade de uso extensivo de dados geográficos, concebidos de forma voluntária por usuários e colaboradores ao redor do mundo, ganhando muita atenção no meio acadêmico, sendo o exemplo mais bem sucedido do *OSM* (GOODCHILD, 2007; HAKLAY, 2010), hospedado por diversas organizações, incluindo as universidades, tais como a University College of London (UCL) e a Universidade de Zagreb (NEIS; ZIELSTRA, 2014).

¹ Openstreetmap.org

² Wikimapia.org

2.4.2 *OpenStreetMap*

O *OSM* é o mapa colaborativo gratuito mais conhecido e disseminado, especialmente por estar sob a *Open Data Commons Open Database License (ODbL)*³, regime de licença em que os dados podem ser copiados, distribuídos, transmitidos e adaptados com a única ressalva de que o *OSM*, e, portanto, os colaboradores, sejam creditados (NEIS; ZIELSTRA, 2014). Para adicionar ou editar dados na plataforma é necessário o cadastro do colaborador, para eventuais disputas licenciasais. Os dados podem ser adicionados através de arquivos GPX de aparelhos GPS ou de forma manual por plataformas online e off-line de edição de dados geográficos. Em 2006 a Yahoo garantiu acesso ao *OSM* das suas imagens de satélite, permitindo adições remotas com auxílio das fotografias aéreas (HAKLAY; WEBER, 2008). O *OSM* também se beneficia de dados oficiais de diversos lugares do mundo, seja por licenças de domínio público ou por doações (SWOCH, 2019). Arsanjani *et al.* (2015) mostram que o Brasil é um dos países em desenvolvimento onde mais elementos foram adicionados à plataforma em 2014.

A requisição de dados, por sua vez, é feita através de API, cuja exportação das informações se faz em formato dedicado *OSM XML*. Diversos prestadores de serviços e softwares de SIG têm incorporado ferramentas de importação do formato *OSM XML* em suas plataformas, a fim de facilitar o uso dos dados (HAKLAY; WEBER, 2008), favorecendo estudos como Hirata *et al.* (2013), que utilizou a plataforma para mapeamento colaborativo de alagamentos em São Paulo; ou Zhang e Pfoser (2019), que analisam a viabilidade do uso dos dados para modelar mudanças urbanas em Manhattan. Não obstante, outra ferramenta que facilitou extensiva utilização dos dados é a biblioteca *OSMnx*, em linguagem de programação *Python* (Boeing, 2017), que importa dos dados diretamente da API do *OSM* e os converte em objetos para manuseio em *Python*. A biblioteca permitiu, por exemplo, a aplicação das informações para estudos de espalhamento e acessibilidade urbanos (GERVASONI *et al.*, 2017), simulações de emissão de CO₂ (HOFER; JÄGER; FÜLLSACK, 2018) e análise morfológica de 27 mil regiões urbanas nos Estados Unidos (BOEING, 2018). No Brasil, os dados de rede do *OSM*, através da biblioteca *OSMnx*, foram utilizados por Kuramoto (2019) para análise da morfologia e acessibilidade de redes urbanas brasileiras, e por Morelli (2019), que estudou a resiliência de diversas malhas viárias. Os dados podem também ser obtidos com a biblioteca *osmdata*, para a linguagem R

³ opendatacommons.org

(PADGHAM *et al.*, 2017), por funções pré-definidas em *osmapi-overpass*, para Java⁴, ou diretamente de suas APIs⁵ utilizando outras linguagens.

Desta forma, o *OSM*, especialmente através da biblioteca *OSMnx*, representa uma promissora fonte de dados não somente por sua disponibilidade, mas pelo crescente uso e respaldo acadêmico. Contudo, a questão que se levanta é: a qualidade destes dados é suficiente para sua aplicação? Muitos estudos buscam responder, deixando claro que a qualidade dos dados pode variar de região para região do planeta, de locais urbanos para rurais e a depender do nível de envolvimento com a plataforma (HAKLAY, 2010; NEIS; ZIELSTRA; ZIPF, 2012; NEIS; ZIELSTRA; ZIPF, 2013; GIRRES; TOUYA, 2018). Segundo estudo de Camboim, Bravo e Sluter (2015), os dados brasileiros são mais completos e numerosos em áreas mais urbanizadas, como capitais e suas adjacências. O estudo teve como objeto a mesorregião metropolitana de Curitiba, que abrange a área observada no presente trabalho, mas não obteve comparações quantitativas aos bancos de dados oficiais, por não estarem disponíveis naquele momento.

A rede viária foi o primeiro tipo de informação que o *OSM* abrigava, mas logo passou a permitir que colaboradores demarquem regiões administrativas, lagos, rios, edificações e Pontos de Interesse (ZHANG; PFOSE, 2019), que indicam, através de polígonos ou pontos, a posição de lojas, instalações, parques e até mesmo estátuas. Os *POIs* são, portanto, diretamente relacionados às variáveis tradicionais usadas para explicação da atração de viagens, especialmente por representarem muitas das atividades econômicas e atividades-fim de viagens urbanas. Desta forma, a rede viária e os *POIs* representam o tipo de dado geográfico almejado para a aplicação deste trabalho. Deve-se destacar que a qualidade é dependente do uso ao que se propõem os dados (HAKLAY, 2010). Ou seja, a usabilidade será qualificada na medida em que os dados mostrarem a qualidade necessária para a aplicação na análise de atração de viagens.

Para a utilização das centralidades de rede para a explicação de atração de viagens deve-se, primeiramente, observar a relação das centralidades com as atividades-fim das viagens urbanas. Esta etapa deve ser verificada com dados controlados, ou seja, dados adquiridos com a autoridade local, com qualidade assegurada para aplicação em tomadas de decisão, evitando incertezas provenientes dos dados colaborativos. Após, então, estas incertezas podem ser avaliadas

⁴ github.com/westnordost/osmapi-overpass

⁵ wiki.openstreetmap.org/wiki/Databases_and_data_access_APIs

comparando-se os dados oficiais aos dados do *OSM*, o que permite não somente verificar as semelhanças e limitação da rede para a aquisição de centralidades, mas dos próprios *POIs*, que representam variáveis tradicionais de modelos de atração.

Com o entendimento da relação “atividades x centralidades”, e da verificação da qualidade entre dados “oficiais x *OSM*”, é possível a calibração de modelos explicativos com todas as informações disponíveis, uma vez que os resultados terão base para apuração de eventuais semelhanças e diferenças nos modelos obtidos. Assim, verifica-se tanto o uso dos dados do *OSM* para análise de atração de viagens quanto o papel que as centralidades de rede podem desempenhar, auxiliando os resultados. O próximo capítulo descreve o método geral para tanto e apresenta o estudo de caso caracterizando-se os bancos de dados utilizados.

3 MÉTODO GERAL E OBJETO DE ANÁLISE

Este capítulo sumariza o método geral da pesquisa em três etapas, que individualmente atingem os objetivos específicos do trabalho, e conjuntamente cumprem o objetivo geral de avaliar a usabilidade do *OSM* para análises de atração de viagens. Após, contextualiza-se Curitiba como local de análise e apresentam-se os bancos de dados considerados.

3.1 MÉTODO GERAL

O trabalho se segmenta em três etapas focadas a atingir cada um dos objetivos específicos. Para cada etapa foi estruturado um método tendo-se em vista o estudo da literatura efetuado, e os respectivos resultados foram apresentados, que em conjunto permitem atestar a usabilidade dos dados do *OSM* e as possibilidades que as centralidades representam.

Com base nos bancos de dados obtidos com a autoridade local – dados oficiais –, a primeira etapa correlaciona as centralidades da rede viária com as atividades econômicas da cidade. As atividades são segmentadas em classes que qualificam o uso dos locais representados e, posteriormente, aplica-se a Estimativa de Densidade de Kernel (KDE) nos pontos das atividades e nos resultados de centralidades de rede. Assim, as superfícies resultantes – que correspondem à densidade espacial dos dados – podem ser correlacionadas diretamente entre si para constatar a relação a nível desagregado (PORTA *et al.*, 2012; RUI; BAN, 2014; LIU *et al.*, 2016). Ademais, também se correlaciona a soma das atividades econômicas com estatísticas zonais das centralidades de rede por ZT, constatando assim a relação a nível zonal, agregação comum para análises urbanas (WANG *et al.*, 2012). Esta etapa, portanto, avalia a relação, estabelecida pela Sintaxe Espacial, de que a configuração da rede viária explica o posicionamento de parte das atividades econômicas pelo princípio da economia do movimento (HILLIER, 1996).

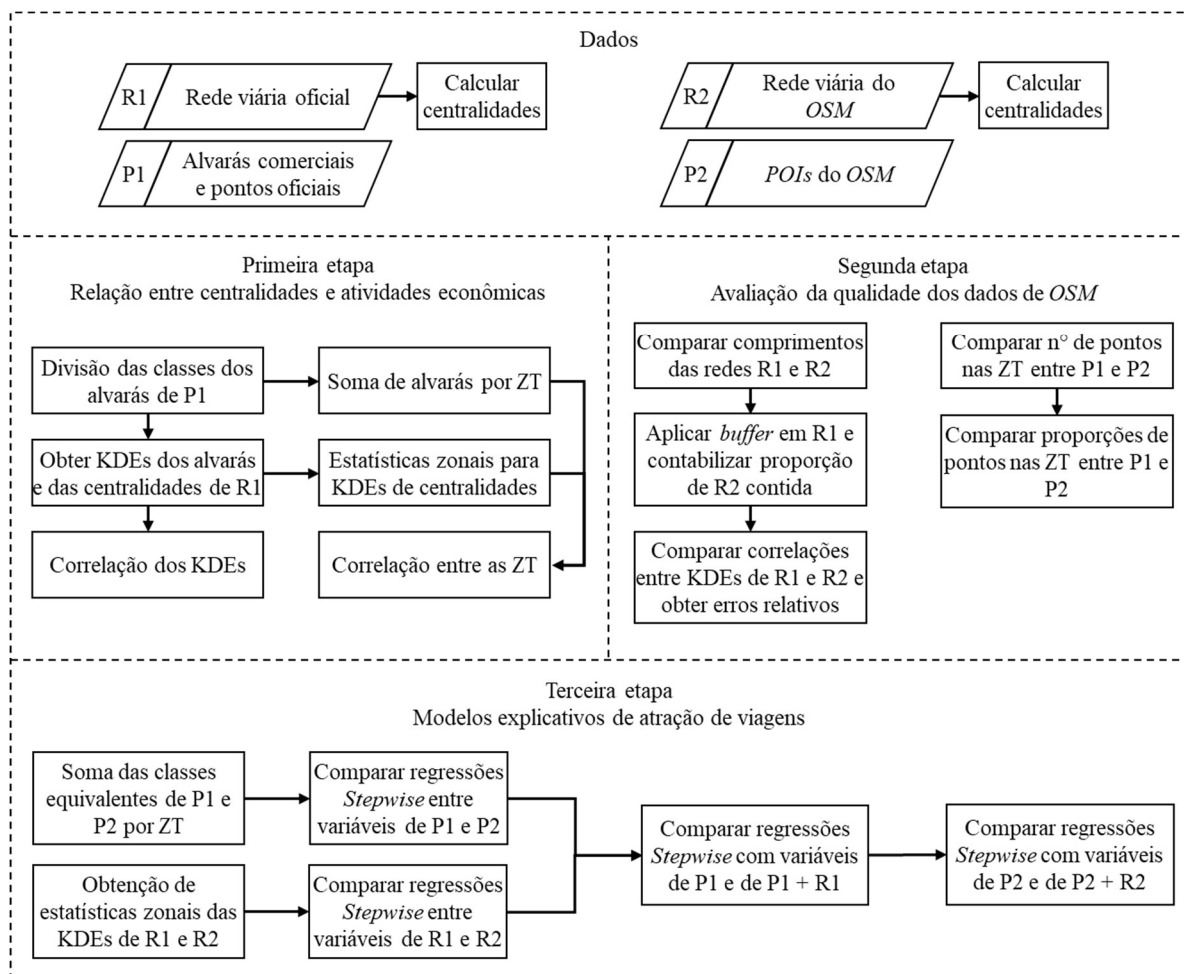
Posteriormente, a segunda etapa observa os bancos de dados oficiais e os bancos de dados obtidos com o mapa colaborativo, avaliando a qualidade das informações geográficas obtidas da plataforma online. Para tanto, alguns critérios são importantes para afirmar a usabilidade das informações, como apontam Van Oort (2006), Haklay (2010) e Girres e Touya (2018). A rede do *OSM* tem sua completude estabelecida por comparação de seu comprimento com o comprimento da rede oficial. Também é obtida a acurácia posicional ao se aplicarem *buffers* na rede oficial e ao se contabilizar a proporção de elementos da rede do *OSM* abrangidos. Da mesma

forma, é preciso verificar a consistência lógica entre as redes, e para isto as centralidades de ambas são calculadas e, após aplicação de KDE nos resultados, as superfícies são correlacionadas e os erros relativos são observados. Já os *POIs* são comparados aos alvarás e outros dados oficiais com classificação equivalente, observando-se a completude e a acurácia posicional de forma agregada nas ZT, comparando-se o número de pontos por zona e a proporção dos pontos por zona, respectivamente. Assim, pode-se averiguar se os dados obtidos do *OSM* têm qualidade suficiente para a aplicação proposta, estabelecendo-se as vantagens e limitações destas informações.

A terceira e última etapa obtém os modelos explicativos de atração de viagens através de regressão *Stepwise* com os seguintes conjuntos de variáveis explicativas: (1) apenas variáveis de locais de interesse, tradicionalmente adotadas em modelos de atração, comparando-se os modelos obtidos com dados oficiais e com dados do mapa colaborativo; (2) apenas com variáveis de centralidades de rede, similar ao efetuado por Jayasinghe, Sano e Rattanaporn (2017), comparando-se os modelos obtidos com centralidades da rede oficial e da rede do *OSM*; e (3) com variáveis de locais de interesse e de centralidades conjuntamente, tanto para dados oficiais quanto para dados do *OSM*, visando-se a entender como as variáveis de centralidade podem auxiliar nos modelos e qual é o papel que desempenham. Todos os modelos obtidos são verificados quanto a premissas da regressão linear (normalidade dos erros, homoscedasticidade e multicolinearidade) e de fenômenos espaciais (independência espacial global). Deste modo, os dados são comparados na aplicação final, de obtenção de modelos explicativos, consolidando as semelhanças e limitações da plataforma colaborativa para o uso proposto, além de esclarecer a possibilidade de aplicação das centralidades de rede como variáveis explicativas para locais atrativos.

A Figura 5 apresenta, de forma esquemática, o método geral segmentado por etapa. Note-se que em conjunto os resultados concretizam o objetivo do presente trabalho, mas individualmente são passíveis de replicação. Portanto, cada etapa tem seu método detalhado nos capítulos subsequentes, delineando-os como estudos específicos, que associados respondem à pergunta de pesquisa: é possível utilizar os dados de *POIs* e de centralidades da rede do *OSM* para análises de atração de viagens?

Figura 5 – Esquema do método geral adotado nas etapas do trabalho.



Fonte: O autor (2020).

O método pode ser aplicado para outras plataformas colaborativas e em outros contextos urbanos. Como apresentado anteriormente, este trabalho observa a cidade de Curitiba e avalia os dados obtidos do *OpenStreetMap*. Sendo assim, as próximas seções caracterizam o local de estudo de caso e apresentam os bancos de dados considerados.

3.2 ESTUDO DE CASO: CURITIBA

Capital do estado do Paraná, o município de Curitiba possui uma área de 435,27 km², e no censo de 2010 contava com 1,75 milhão de habitantes. A região metropolitana, por sua vez, possui 16.581,20 km² e uma população de 3,22 milhões de pessoas (IBGE, 2020). A Figura 6 mostra a localização da região e seus municípios, enquanto a Tabela 1 apresenta as populações e áreas dos respectivos municípios.

Figura 6 - Localização da Região Metropolitana de Curitiba e seus municípios.



Fonte: O autor (2020).

Tabela 1 - População e área dos municípios da Região Metropolitana de Curitiba (IBGE, 2019).

Município	População (2010)	Área (km ²)	Município	População (2010)	Área (km ²)	Município	População (2010)	Área (km ²)
Curitiba	1.751.907	435,27	Campo Magro	24.843	304,49	Pinhais	117.008	60,75
Adrianópolis	6.376	1.349,34	Cerro Azul	16.938	1.341,19	Piraquara	93.207	227,03
Agudos do Sul	8.270	192,23	Colombo	212.967	197,36	Quatro Barras	19.851	181,13
Almirante Tamandaré	103.204	194,75	Contenda	15.891	299,04	Quitandinha	17.089	447,03
Araucária	119.123	469,17	Doutor Ulysses	5.727	781,45	Rio Branco do Sul	30.650	812,33
Balsa Nova	11.300	348,97	Fazenda Rio Grande	81.675	116,68	Rio Negro	31.274	603,25
Bocaiuva do Sul	10.987	826,35	Itaperuçu	23.887	314,42	São José dos Pinhais	264.210	946,44
Campina Grande do Sul	38.769	538,97	Lapa	44.932	2.093,83	Tijucas do Sul	14.537	672,20
Campo do Tenente	7.125	1.249,42	Mandirituba	22.220	379,18	Tunas do Paraná	6.256	666,47
Campo Largo	112.377	275,57	Piên	11.236	254,90	Total	3.223.836	16.581,20

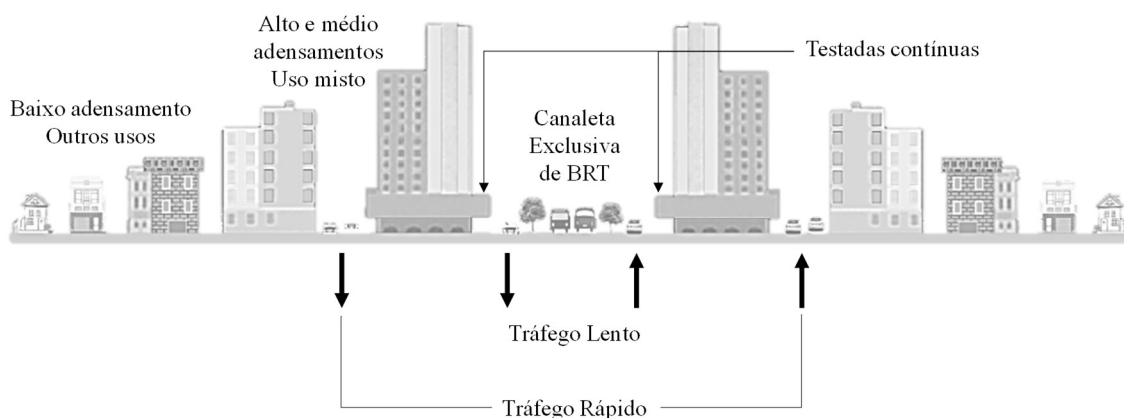
Curitiba é conhecida pelas soluções urbanas de Desenvolvimento Orientado ao Transporte Público (DOT), que busca o provimento de infraestrutura e operação de transporte a uma região

previamente à sua ocupação. Os esforços para a aplicação destas soluções se iniciaram na década de 1940, quando o arquiteto francês Alfred Agache conferiu à cidade um plano de crescimento circular baseado em centros funcionais, típico de cidades europeias (MIRANDA, 2010).

A principal caracterização de Curitiba, no entanto, começa na década de 1960, a partir do plano Wilhelm, em que eixos lineares de crescimento começam a surgir, unindo uso do solo, transporte coletivo e circulação, contrário a planos urbanistas de outras grandes cidades brasileiras, baseados em aumento contínuo de capacidade de vias e grandes quadras. Este tipo de planejamento também era contrário aos preceitos do urbanismo modernista, que buscava segregar a cidade de acordo com as atividades desempenhadas em cada região (MIRANDA, 2010). Assim, Curitiba buscou a coexistência de atividades econômicas e do transporte, fazendo com que os eixos garantissem a variabilidade de atividades-fim das viagens e, ao mesmo tempo, a acessibilidade.

Esses eixos estruturais se identificam por um conjunto de três vias, definindo o chamado sistema trinário, cuja via central é composta por canaleta do sistema BRT com duas vias locais nas laterais e um binário de tráfego mais rápido formados pelas vias paralelas. O zoneamento da cidade é baseado nestes eixos, com gabaritos maiores permitidos ao longo da via central e gradativa redução em vias mais afastadas. O uso misto do solo é incentivado ao longo dos eixos, de modo a garantir fachadas ativas e continuidade das testadas dos primeiros pavimentos (IPPUC, 2008; CURITIBA, 2015). A Figura 7 apresenta as características citadas.

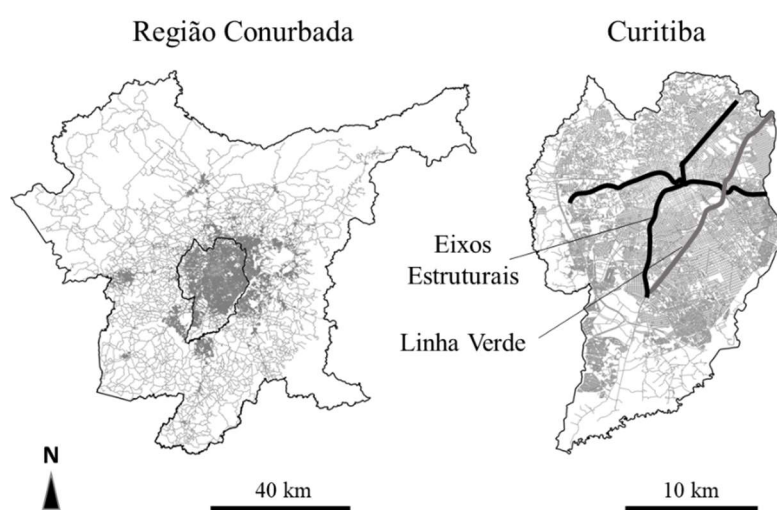
Figura 7 - Esquema da concepção dos eixos estruturantes de Curitiba.



Fonte: Adaptado de Ippuc (2020).

Deve-se considerar o caráter metropolitano da cidade, pois os municípios vizinhos têm grande impacto na dinâmica e no planejamento de Curitiba. A Figura 8 apresenta a região conurbada e a localização dos eixos estruturais para o município. A rede é disposta para demonstrar o aglomerado urbano limítrofe. A Linha Verde apontada é o mais recente eixo estruturante a ser implementado, com algumas variações em relação ao modelo apresentado anteriormente, mas ainda com a princípio de canaletas centrais e adensamento próximo ao sistema coletivo de transporte.

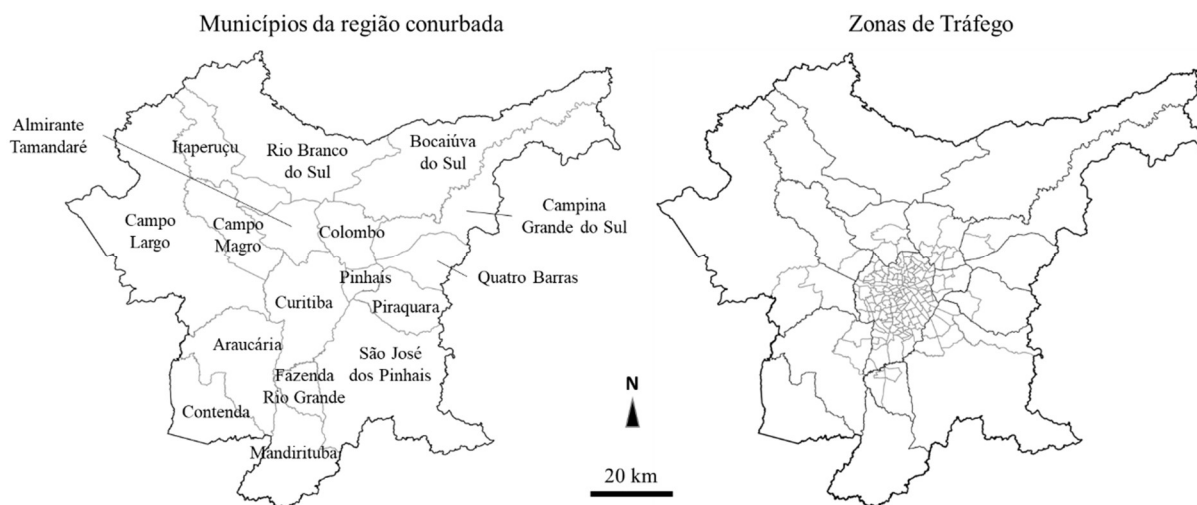
Figura 8 – Região conurbada e localização dos eixos estruturantes da cidade.



Fonte: Adaptado de Ippuc (2020).

No ano de 2018, a Prefeitura Municipal de Curitiba efetuou a primeira pesquisa Origem-Destino da cidade, por meio de pesquisa domiciliar (IPPUC, 2020). A região abrangida considera a região conurbada apontada anteriormente, contemplando uma população de aproximadamente 3,043 milhões de habitantes, numa área de 6.580,92 km² que foi dividida em zonas de tráfego (ZTs) pré-estabelecidas por setores censitários. Os municípios que compõem a região e as ZTs são apresentados na Figura 9. As zonas do município de Curitiba são menores e mais numerosas, num total de 135, enquanto externamente ao município são apenas 46 zonas de tamanhos maiores, totalizando 181 para toda a região.

Figura 9 – Municípios e zonas de tráfego da região conurbada.



Fonte: O autor (2020).

Este trabalho analisa o município de Curitiba, onde foi possível obter as informações de atividades econômicas, porém considera a rede viária para toda a região conurbada, uma vez que a configuração da rede deve garantir a representação global de sua estrutura, evitando efeitos de borda. As ZTs são utilizadas como unidades de agregação das informações, visto que são a base estabelecida para as viagens levantadas pela pesquisa OD. Assim, deve-se apresentar o banco de dados com as informações a serem utilizadas.

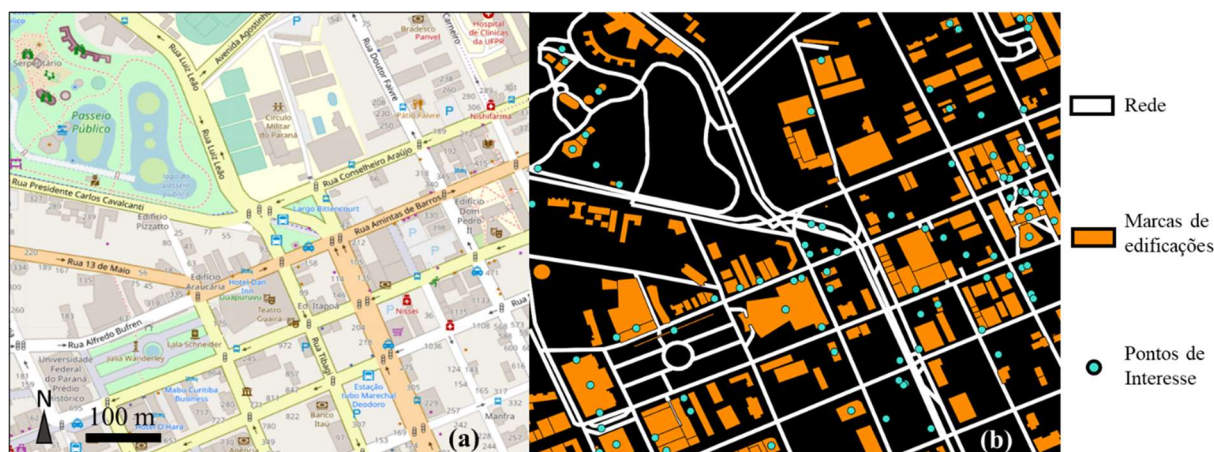
3.3 BANCOS DE DADOS

Para que as etapas demonstradas anteriormente sejam efetivadas, é importante descrever a aquisição e a preparação dos dados usados. Os dados obtidos do *OSM* e os bancos de dados oficiais – fornecidos pela prefeitura de Curitiba – são apresentados a seguir. Ressalta-se que o período de aquisição dos dados do *OSM* e da compilação dos dados oficiais foi próximo, entre o último trimestre de 2018 e o primeiro trimestre de 2019, mantendo-se a similaridade temporal das informações.

3.3.1 Dados do *OpenStreetMap*

Como já mencionado, os dados do *OSM* foram adquiridos com o auxílio da biblioteca *OSMnx*, que permite a importação da rede, *POIs* e marcas de edificações de forma simplificada a partir de polígonos definido por coordenadas ou a partir de áreas administrativas. A representação do mapa colaborativo no site e os respectivos elementos importáveis são mostrados na Figura 10.

Figura 10 - Exemplo de visualização do *OSM* (a) e diferenciação dos dados obtidos da plataforma (b).



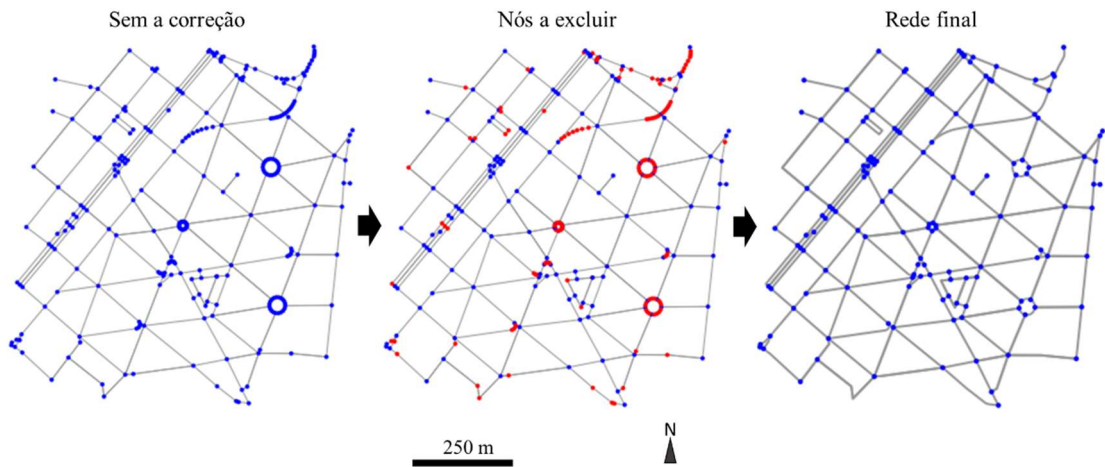
Fonte: O autor (2020).

As marcas de edificações são numerosas, mas não apresentam a classificação de uso de cada local, limitando a aplicação para o presente estudo, visto que o uso que se faz do local é importante para a avaliação. Isto ocorre porque tais dados foram possibilitados, em grande parte, através de adições remotas com o auxílio de imagens de satélite (HAKLAY; WEBER, 2008). Os *POIs*, por sua vez, comumente apresentam estas classificações, visto que sua adição muitas vezes tem a intenção de posicionar locais conhecidos dos colaboradores. Portanto, os *POIs*, assim como a rede viária, são passíveis de utilização.

3.3.1.1 Rede Viária

Com a ferramenta de importação da biblioteca *OSMnx*, a rede adquirida considera vias da rede “drive”, trafegáveis por todos os modos de transporte, e abrange todas os municípios da região conurbada de Curitiba, importante para o cálculo das centralidades de forma a evitar o efeito de borda (HILLIER; PENN, 1996) pois as redes das cidades vizinhas impactam o cálculo da rede de Curitiba. Como a representação se dá por nós e arestas retilíneas, as curvas são representadas por consecutivos nós e segmentos retos. A biblioteca permite a correção que exclui os nós intermediários e mantém somente os que representam interseções e fins de rua, reproduzindo a geometria original. Este processo é exemplificado na Figura 11.

Figura 11 – Processo de correção da rede para exclusão de nós intermediários.



Fonte: O autor (2020).

O grafo obtido (Figura 12) contém 176.757 arestas e 66.029 nós, sendo que 84.884 arestas e 31.612 nós se encontram no município de Curitiba. Cada aresta contém campos de atributos para nome da rua, largura, número de faixas, hierarquia e velocidade, enquanto os nós demonstram presença ou não de semáforo. Porém, muitos campos não apresentam informação, o que é comum para mapas do *OSM* (HAKLAY, 2010). A falta de informações dos atributos da rede não é restritiva, uma vez que apenas a morfologia e geometria da rede são importantes.

Figura 12 – Rede viária obtida do *OSM*.



Fonte: O autor (2020).

3.3.1.2 Pontos de Interesse

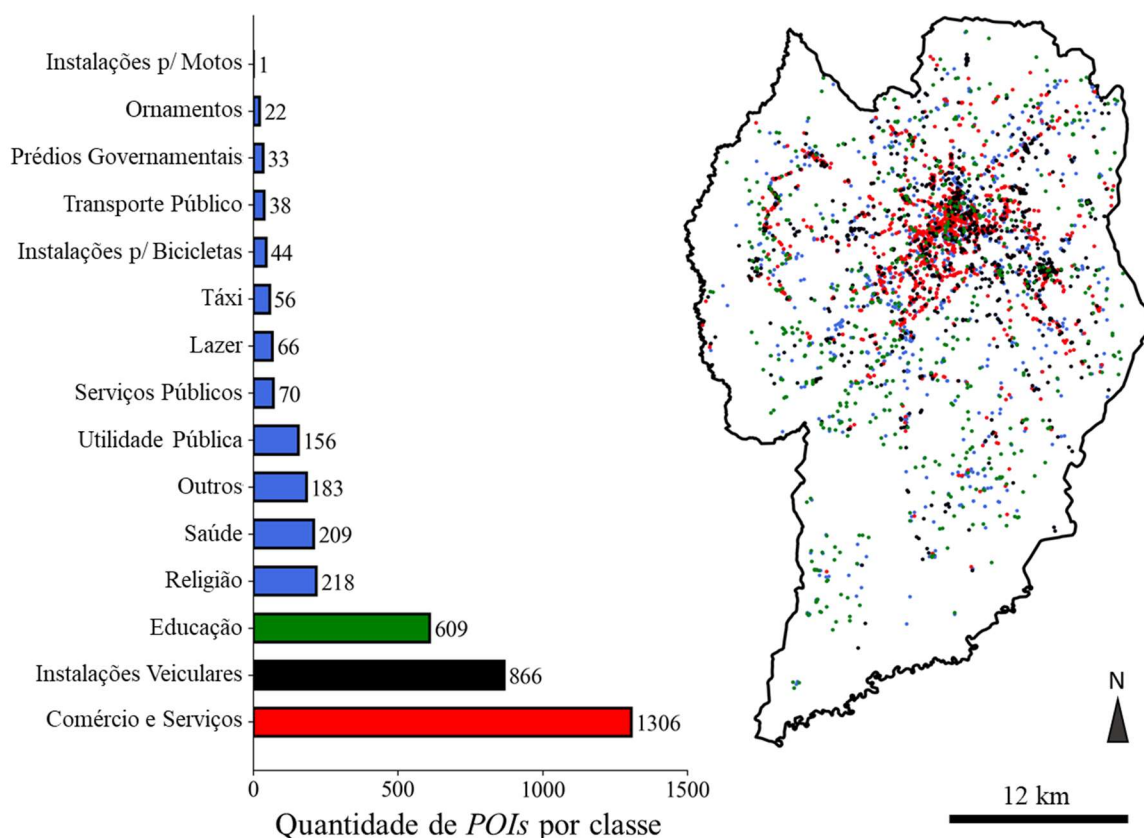
A área de importação dos *POIs* se restringe ao município de Curitiba, visto que os dados oficiais equivalentes não puderam ser obtidos para os outros municípios da região conurbada. Cada *POI* apresenta uma classificação pelo atributo “*amenity*”, que pode ser definida pelo colaborador ou por lista preconcebida⁶ para auxiliar no registro dos pontos. Como a classificação é muito numerosa, criam-se classes para agregar *POIs* similares quanto ao tipo de uso conforme a Tabela 2.

Tabela 2 - Divisão da classificação do banco de dados de *POIs* de acordo com classes.

Classe	Classificação (<i>amenity</i>)
Transporte Público	<i>Bus station</i>
Instalações Veiculares	<i>Fuel, parking_entrance, parking, car_rental, car_wash, charging_station, compressed air, parking space, driveway, car pooling</i>
Comércio e Serviços	<i>Fast food, restaurant, pharmacy, pub, bar, bank, post office, love hotel, food_court, cafe, marketplace, veterinary, ice_cream, nightclub, bbq, bureau_de_change, internet_cafe, pet, dojo, conference_centre, studio, scrapyard</i>
Utilidade pública	<i>Toilets, telephone, bench, atm, drinking_water, shelter, post_box, recycling, water_point, public_bath, vending_machine'</i>
Táxi	<i>Taxi</i>
Educação	<i>University, kindergarten, school, college, library, driving_school, childcare, music_school, language_school</i>
Religião	<i>Place_of_worship, monastery</i>
Lazer	<i>Cinema, theatre, social_centre, planetarium, community_centre, arts_centre</i>
Serviços Públicos	<i>Fire_station, police, grave_yard, crematorium, mortuary, animal_shelter</i>
Saúde	<i>Hospital, clinic, dentist, doctors, social_facility, nursing_home</i>
Prédios Governamentais	<i>Townhall, public_building, courthouse, embassy, register_office</i>
Instalações p/ bicicletas	<i>Bicycle_parking, bicycle_repair_station</i>
Ornamentos	<i>Fountain, clock</i>
Instalações p/ motos	<i>Motorcycle_parking</i>
Outros	<i>Auditorium, sem dados</i>

Com 3.877 registros, os *POIs* se distribuem no município conforme a Figura 13 a seguir, que também apresenta a quantidade de registros por classe. As três classes com mais pontos são destacadas para permitir sua visualização

⁶ <https://wiki.openstreetmap.org/wiki/OSMonto>

Figura 13 – Quantidade e posição dos *POIs* em Curitiba.

Fonte: O autor (2020).

A maioria dos pontos é da classe de comércio e serviços, sendo a distribuição concentrada no centro da cidade. É possível presumir que o número de *POIs* não reflete a quantidade real de locais de cada classe, cuja quantificação é apresentada no Capítulo 5.

3.3.2 Dados oficiais

Para a obtenção das centralidades, também se utiliza a rede viária urbana. Já em relação aos dados passíveis de comparação com os *POIs* do *OSM*, que representam variáveis tradicionalmente explicativas da atração de viagens, utilizam-se os bancos de dados geográficos de alvarás comerciais, pontos de ônibus e paraciclos. Todas as informações são fornecidas pela autoridade local.

3.3.2.1 Rede Viária

A rede viária se apresenta em dois arquivos geográficos *shapefile*, de forma vetorizada em pontos e linhas, abrangendo toda a região conurbada de Curitiba. Consideram-se somente vias em que automóveis particulares podem trafegar, o que abrange qualquer tipo de veículo e modos

não motorizados. Os arquivos são convertidos em grafo com a biblioteca *NetworkX*⁷ e passam pelo mesmo processo de correção da rede do *OSM*, mantendo-se somente pontos que representam fins de rua ou interseções.

O grafo resultante (Figura 14) contém 162.958 arestas e 64.571 nós, cuja parcela interna ao município de Curitiba representa 84.413 arestas e 33.761 nós. Cada aresta tem atributos relacionados à geometria, nome de rua, número de faixas, hierarquia e velocidade; já os nós informam sobre a presença ou não de semáforos. Contudo, os atributos não relacionados à geometria das vias não são importantes para a presente aplicação.

Figura 14 – Rede viária oficial obtida com a prefeitura de Curitiba.



Fonte: O autor (2020).

Note-se que em alguns pontos a rede excede os limites da região dos municípios vizinhos. Isto ocorre pois, diferentemente da rede do *OSM*, a rede oficial não foi limitada pelo polígono definido pela região. De qualquer forma, o número de elementos externos é pequeno e terá um impacto reduzido para o município de Curitiba, foco da análise das centralidades.

⁷ networkx.github.io

3.3.2.2 Alvarás comerciais, pontos de ônibus e paraciclos

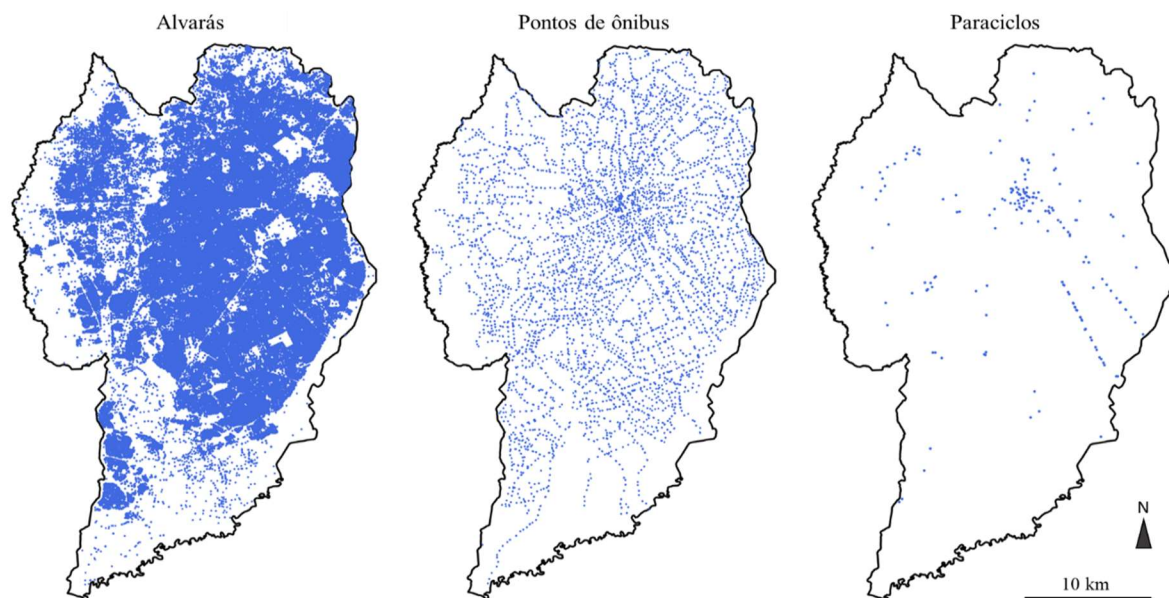
Fornecidos em três distintos *shapefiles* de pontos, os alvarás comerciais, os pontos de ônibus e os paraciclos abrangem o município de Curitiba. Os dois últimos estão disponíveis no site do IPPUC (IPPUC, 2020), enquanto que o primeiro é obtido diretamente com o órgão. Os alvarás compreendem 433.558 registros de pessoas jurídicas até o ano de 2019, sendo que 273.217 registros são alvarás com manifestação física, ou seja, em que há um local de atuação da atividade, não compreendendo pessoas jurídicas como motoristas profissionais, que não são atividades com local fixo e passíveis de atrair viagens. Portanto, os alvarás observados são desde microempresas até empresas de grande porte. A Tabela 3 mostra as classes dos alvarás com alguns exemplos de atividades abrangidas e a quantidade de registros em cada classe.

Tabela 3 – Classes de alvarás de pessoas jurídicas de Curitiba.

Classe de alvarás	Exemplos	Qtde.
A - Agropecuária e Florestal	Extração de madeira e cultivo de produtos agrícolas.	395
B - Indústria Extrativa	Extração de minérios e recursos naturais, beneficiadoras e atividades de apoio à extração.	75
C - Indústria de Transformação	Fabricantes de móveis e informáticas, usinagem e solda, impressoras e confecções.	16.504
D - Indústria Eletricidade	Geradoras e comercializadora de energia.	204
E - Água e Gestão de resíduos	Compostagem, coleta de resíduos e reciclagem.	376
F - Construção	Empreiteiras, incorporadoras e instaladoras.	20.024
G - Comércio	Varejo, atacados, padarias e representantes comerciais.	80.295
H - Transporte, armazenagem e correio	Entregadoras, transportadoras, fretagem e correios.	12.207
I - Alojamento e Alimentação	Restaurantes, bares, lanchonetes e hotéis.	14.293
J - Informação e Comunicação	Produtoras de imagem e som, suporte de TI, programadoras e editoras de livros.	12.083
K - Finanças e seguros	Corretoras, holdings financeiros e não-financeiros e sociedades de participação.	7.920
L - Imobiliárias	Imobiliárias, administradoras e loteadoras.	5.287
M - Científicas e técnicas	Serviços de engenharia, design de interiores, advocacias e publicitárias.	24.813
N - Administrativas e complementos	Limpeza de edifícios, serviços administrativos e Fotocópias	22.889
O - Adm Pública, defesa e seguridade	Companhia de abastecimento, previdência e empresas públicas.	11
P - Educação	Universidades, escolas e ensino de habilidades.	9.434
Q - Saúde e Serviços Sociais	Clinicas e hospitais	8.240
R - Arte, Cultura e Desportos	Produção musical e de eventos, recreação e gestão de instalações esportivas.	3.881
S - Outros serviços	Cabeleireiros, reparos eletrônicos e associações de moradores.	19.966
T - Serviços domésticos	Atividades de diarista.	381
U - Internacionais e extraterritoriais	Exemplo único: Instituto Cervantes.	1
X - Outros	Alvarás sem definição especificada	13.938

Há também 6.958 pontos, estações e terminais representados no arquivo de pontos de ônibus, demonstrados na Figura 15, juntamente com os 184 paraciclos e com a distribuição dos alvarás.

Figura 15 – Distribuição dos alvarás comerciais, pontos de ônibus e paraciclos obtidos do banco de dados oficial.



Fonte: O autor (2020).

Note-se que estes dados não equivalem a todas as classes de *POIs*, e o contrário também é verdade. Esta discussão é feita no Capítulo 5, com a comparação entre categorias que tradicionalmente são utilizadas como variáveis para explicar a atração de viagens.

3.3.3 Centralidades

O processo para a obtenção das centralidades da rede depende, primeiramente, da representação *dual*, obtida ao se transformarem os trechos de vias – arestas – em nós e as conexões entre estes trechos em arestas, como apontado anteriormente na Figura 3(c). Além disso, as centralidades de proximidade e intermediação, calculadas respectivamente pelas equações 2.2 e 2.3, dependem da medida de distância considerada, enquanto que a centralidade de grau (equação 2.1) depende somente do número de conexões de cada nó.

Ambas as redes, portanto, são reconstruídas como grafos *duais* e cada aresta destes adquire duas medidas de distâncias, a métrica e a angular, similar ao método adotado por Serra e Hillier (2019). Como cada aresta *dual* (uv) é constituída de duas arestas *primais* (u e v), a distância métrica ($d(m)$) é definida como a metade da soma das distâncias das duas arestas u e v (equação

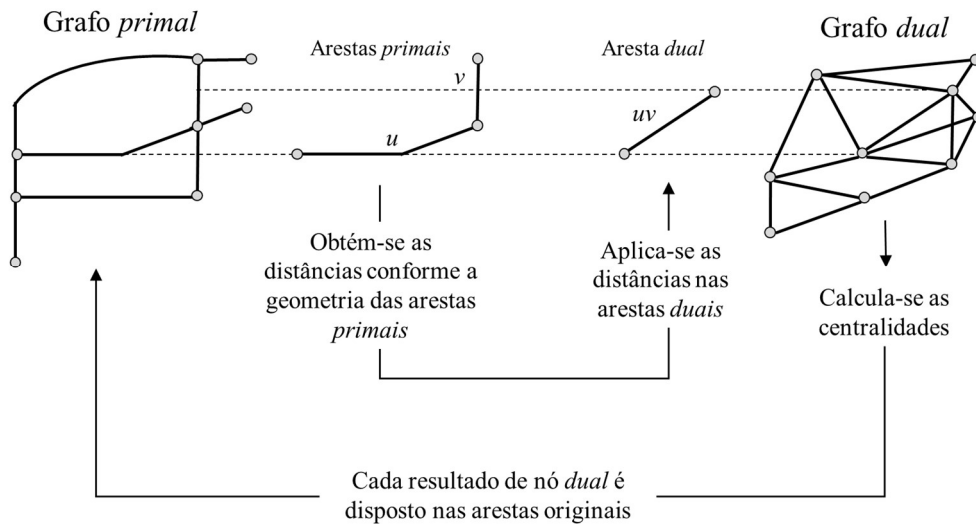
2.4), pois considera-se a distância média de deslocamento entre uma aresta e outra, ao passo que a distância angular ($d(a)$) é a metade da soma da diferença angular ($\delta(a)$) das duas arestas u e v somada à curvatura existente entre elas ($\Delta(a)$) (equação 2.5). As diferenças angulares e curvaturas são sempre em valores absolutos, evitando-se que direções opostas se anulem.

$$d(m)_{uv} = \frac{d(m)_u + d(m)_v}{2} \quad [2.4]$$

$$d(a)_{uv} = \frac{\delta(a)_u + \delta(a)_v}{2} + \Delta(a)_{uv} \quad [2.5]$$

Após a obtenção das distâncias para cada aresta *dual* (uv), faz-se o cálculo das centralidades. Os resultados são referentes a cada nó do grafo *dual*, o qual, reitera-se, representa uma aresta na rede *primal*. Esta, portanto, recebe os resultados, permitindo que as redes originais sejam analisadas conforme a representação física tradicional. A Figura 16 conceitua o processo descrito.

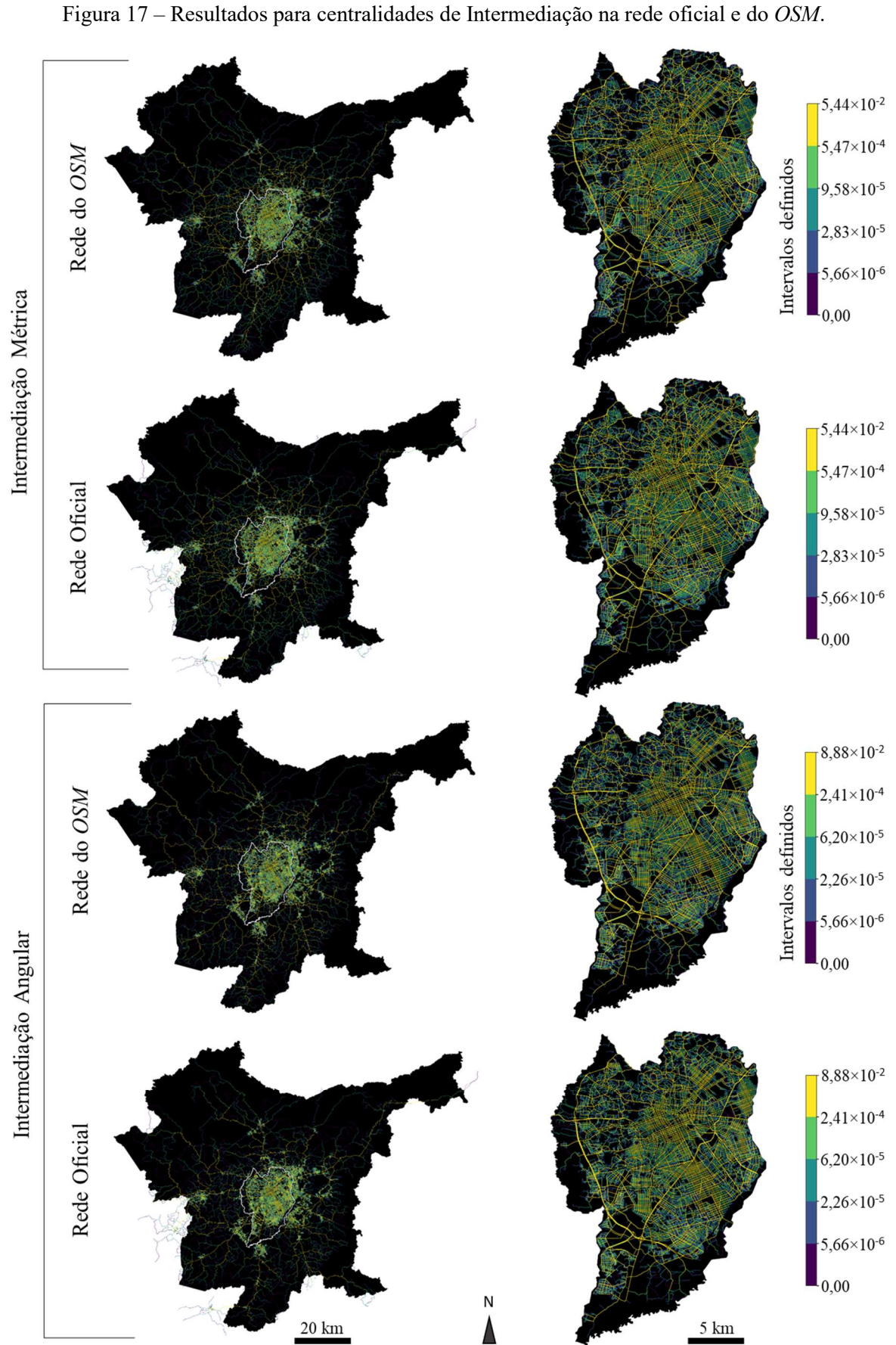
Figura 16 – Processo de obtenção das distâncias métrica e angular para cada aresta do grafo dual e posterior cálculo das centralidades.



Fonte: O autor (2020).

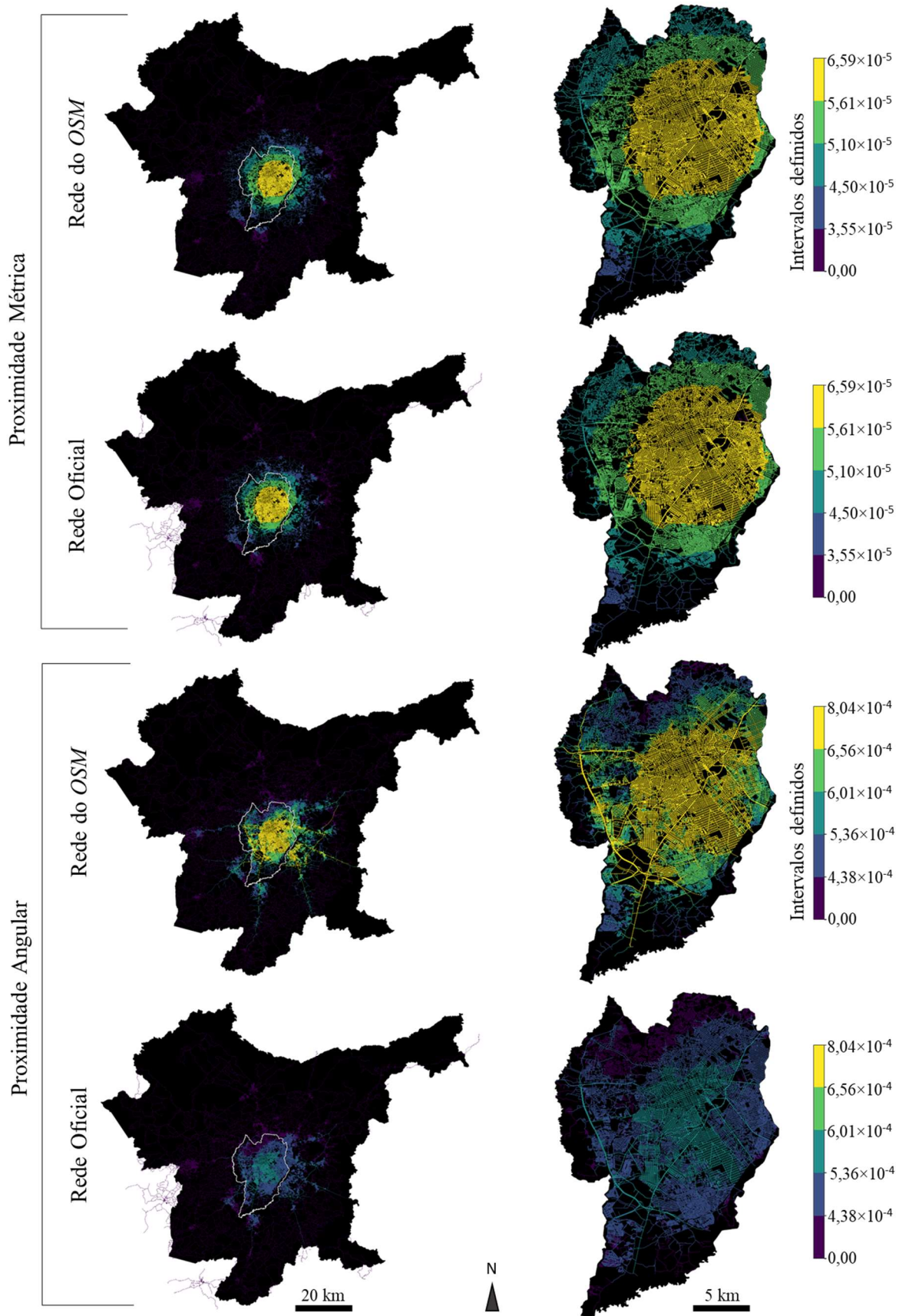
No total, são cinco resultados de centralidade para cada rede: proximidade métrica e angular, intermediação métrica e angular e grau. Os resultados são apresentados na Figura 17, na Figura 18 e na Figura 19. Os intervalos de cores são definidos pelos quintis (1/5) das arestas do OSM,

ou seja, cada intervalo contempla 20% das arestas e, para comparação, os mesmos intervalos são replicados para a rede oficial em cada centralidade.



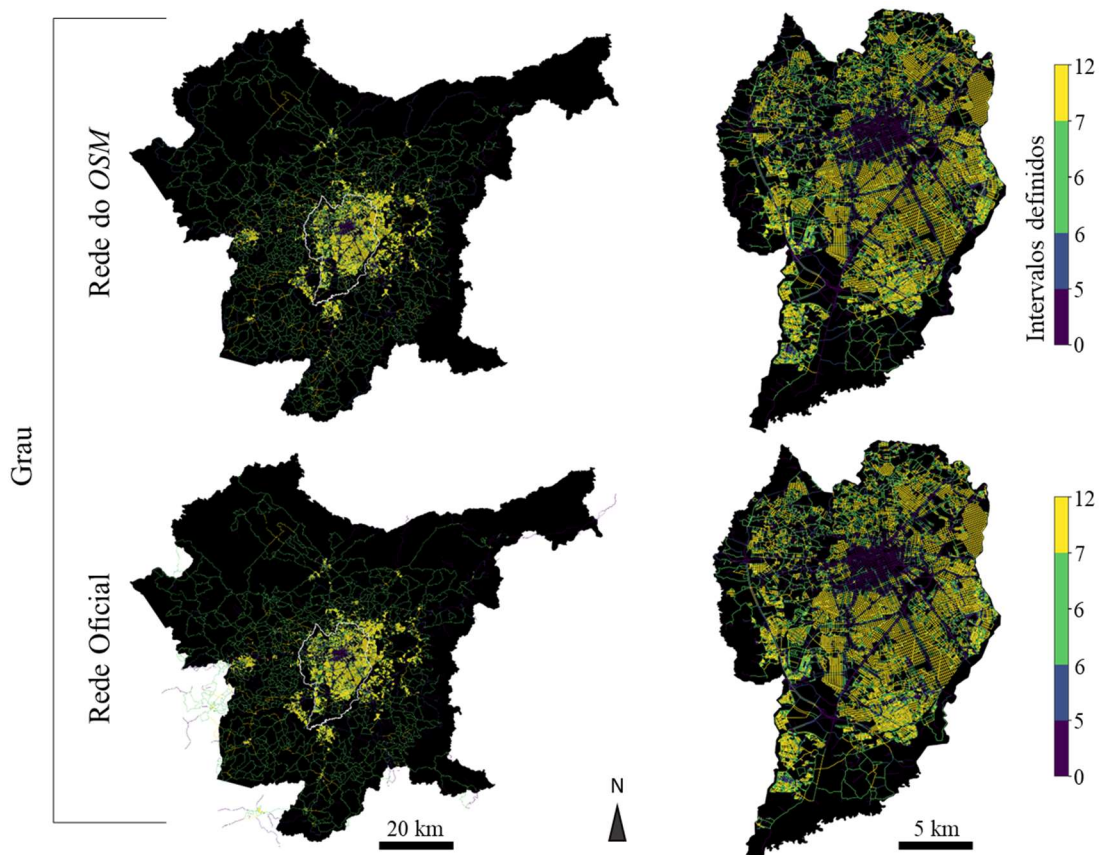
Fonte: O autor (2020).

Figura 18 – Resultados para centralidades de Proximidade na rede oficial e do *OSM*.



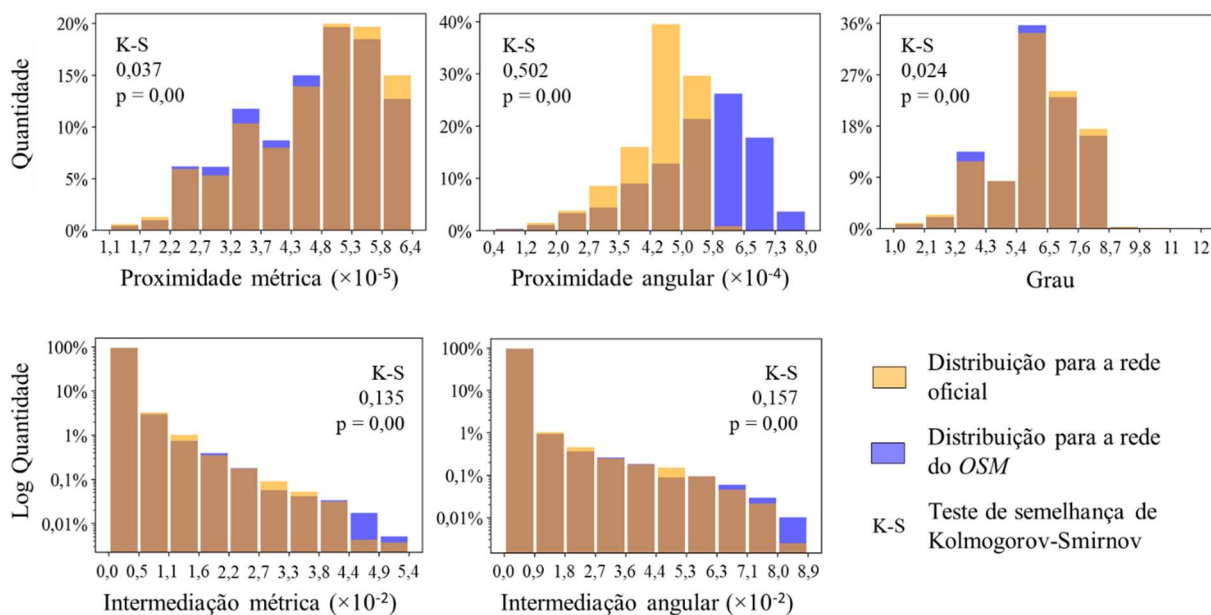
Fonte: O autor (2020).

Figura 19 – Resultados para centralidade de Grau na rede oficial e do *OSM*.



Fonte: O autor (2020).

Pode se observar que as centralidades de intermediação apresentam concentração dos maiores valores em poucas vias, demonstrando-se que a maioria dos caminhos dentro da rede prefere trechos semelhantes. Todas as centralidades apresentam comportamentos similares para a rede do *OSM* e a oficial, ainda que a proximidade angular da rede oficial tenha valores visivelmente menores do que a rede do *OSM*. A centralidade de grau tem por característica os valores maiores em locais com vias de duplo sentido e com muitos cruzamentos, mais comum em áreas residenciais, e menores valores em vias de sentido único, geralmente de maior hierarquia. A Figura 20 mostra as distribuições das centralidades para ambas as redes para comparação. As quantidades para as centralidades de intermediação estão em escala logarítmica por serem muito concentradas, melhorando a visualização.

Figura 20 – Distribuições das centralidades para a rede oficial e do *OSM*.

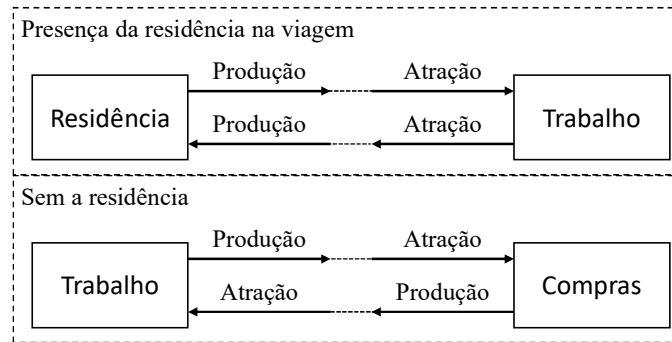
Com os gráficos, é possível notar a maior diferença para a proximidade angular percebida nos mapas. Esta diferença está vinculada ao número de elementos nas duas redes – maior na rede do *OSM* –, o que afeta o numerador da equação da centralidade e leva a maiores valores; também é vinculada a possíveis distinções na disposição e geometria dos elementos, que afetam a distância angular calculada mais do que a distância métrica. Através do teste de Kolmogorov-Smirnov, para se verificar a semelhança entre as distribuições, rejeita-se a hipótese nula de que as distribuições das duas redes sejam iguais para um nível de confiança de 99%, como verificado pelos valores-*p* demonstrados. Contudo, o comportamento das centralidades pode ser satisfatoriamente similar para a aplicação proposta, visto que elementos podem apresentar importância análoga mesmo que com valores distintos. A aplicabilidade e as incoerências são analisadas no Capítulo 5 com mais detalhes.

3.3.4 Atração de Viagens

Os valores de viagens atraídas foram embasados na Pesquisa de Origem e Destino de 2018 feita para a Cidade de Curitiba e municípios próximos, na região conurbada. A pesquisa considerou as ZTs da cidade para as entrevistas domiciliares e aponta fatores de extrapolação zonais para os deslocamentos. Assim, a atração de viagens é agregada nas ZTs e é definida conforme apontam Ortúzar e Willumsen (2011): deslocamentos que apresentam a residência em um dos ex-

tremos têm a atração no extremo oposto, ainda que a viagem termine na residência; já deslocamentos sem a residência tem a atração definida no extremo de destino, como aponta a Figura 21.

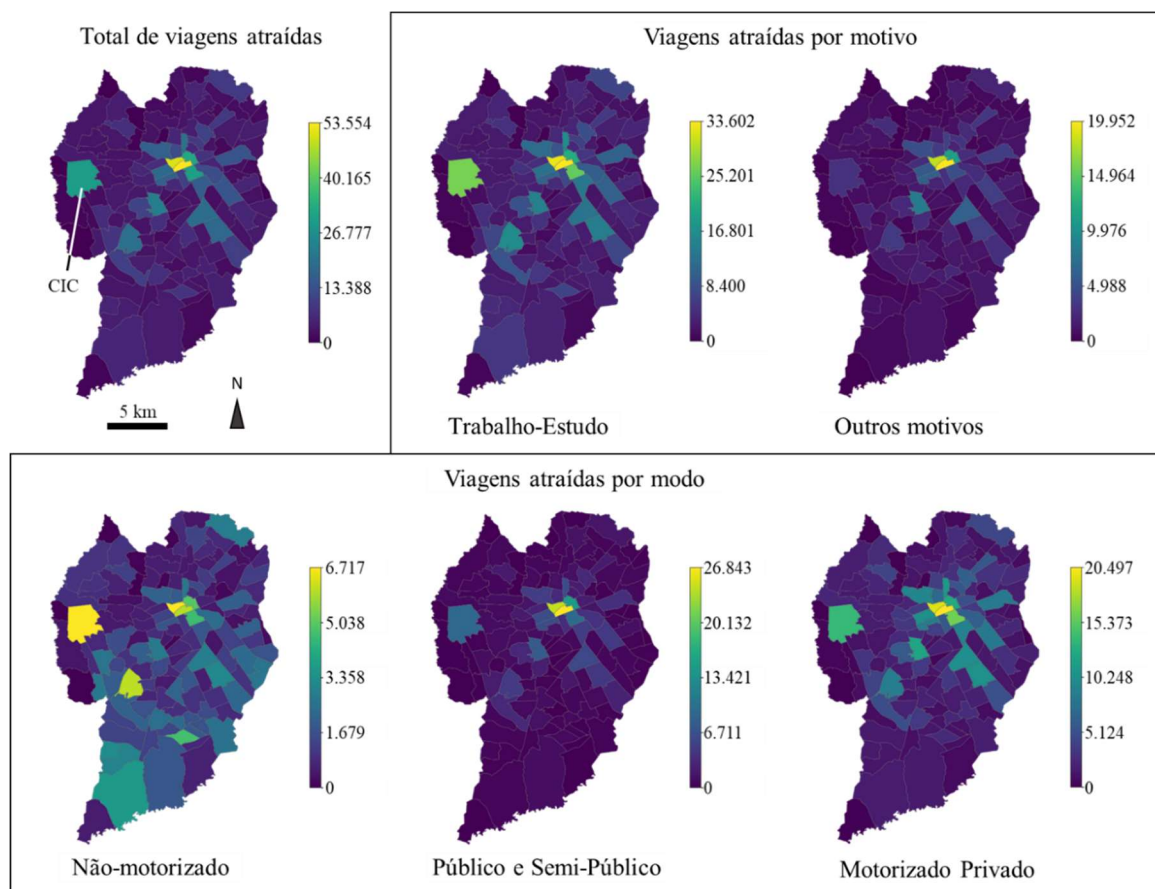
Figura 21 - Esquema de diferenciação da produção e atração de viagens de acordo com a presença ou não da residência.



Fonte: Adaptado de Ortúzar e Willumsen (2011).

A quantidade de viagens atraídas por ZT para o município de Curitiba é demonstrada na Figura 22. Para aprimorar as análises relativas à atratividade no Capítulo 6, as viagens também foram segmentadas segundo o motivo – trabalho e estudo ou outros motivos – e segundo o modo de transporte – não-motorizado, público e semipúblico ou motorizado privado, permitindo atestar diferenças na atração por motivos obrigatórios, como trabalho, ou opcionais, como compras ou lazer; também buscando-se entender se o modo utilizado é afetado pelas centralidades. Esta segmentação também viabiliza a ponderação quanto às variáveis originadas dos *POIs* – ligadas a variáveis tradicionais –, julgando a coerência dos resultados.

Figura 22 – Atração de Viagens total, por motivo e por modo nas ZTs de Curitiba.



Fonte: O autor (2020).

Todos os tipos de viagens se concentram próximas ao centro e ao longo dos eixos estruturais, exceto as não-motorizadas, que apresentam quantidades altas em ZTs afastadas do centro, o que reflete a tendência de se efetuar este tipo de viagem em locais mais próximos da residência, e também de que a escolha por modos não-motorizados é muito vinculada à renda – regiões mais centrais apresentam oportunidades mais próximas às residências, mas a renda, a motorização e a acessibilidade ao transporte público são mais elevadas. Também é possível perceber uma alta atratividade na Cidade Industrial de Curitiba (CIC) como já apontado, explicada em grande parte pelo motivo Trabalho-Estudo.

4 RELAÇÃO ENTRE AS CENTRALIDADES DE REDE E AS ATIVIDADES ECONÔMICAS

Neste capítulo, correlacionam-se as centralidades da rede viária de Curitiba e as atividades econômicas, representadas pelo posicionamento dos alvarás de pessoas jurídicas da cidade. A correlação visa a replicar, como estudo de caso brasileiro, a comparação entre a configuração da rede e as atividades econômicas no contexto da Sintaxe Espacial e da premissa da economia do movimento, em que a acessibilidade e os potenciais de movimento na rede explicam fenômenos urbanos como a economia. Assim, além de corroborarem a interdependência entre infraestrutura e uso do solo, os resultados ajudam a entender como as centralidades podem ser utilizadas para análises de atração de viagens.

4.1 CONTEXTUALIZAÇÃO

4.1.1 Relação entre centralidades e dinâmicas socioeconômicas

A presença de movimentos, a depender dos modos utilizados, tende a trazer vantagens a atividades econômicas, especialmente para as atividades cotidianas ou focadas em atendimento ao público (JACOBS, 1961). Melhores condições de acessibilidade também impactam nas escolhas de posicionamento das atividades (KELLY, 1994), como buscam replicar os modelos que agregam a infraestrutura de transporte e o uso do solo (IACONO; LEVINSON; EL-GENEIDY, 2008). O uso das medidas de centralidades pode auxiliar na observação destes fenômenos, como visto em Turner (2007), Jiang e Jia (2011) e Serra e Hillier (2019), que demonstram que grande parte dos movimentos nas cidades é produto da configuração da rede viária e se relaciona com ela; e Kim e Dong, 2002, Porta *et al.* (2012) e Scoppa e Peponis (2015), que mensuram a correlação entre as centralidades e a presença de atividades econômicas. Estas relações corroboram a premissa da “economia do movimento”, de que a economia da cidade é mutuamente moldada pela configuração da rede viária por meio dos movimentos que nela ocorrem e a acessibilidade propiciada por ela (HILLIER, 1996).

O estudo de Hillier e Penn (1996) foi o primeiro a testar a relação entre os princípios do “movimento natural” e da “economia do movimento” da Sintaxe Espacial. Os autores relacionam a centralidade de proximidade com os movimentos a pé no centro de Londres, considerando a rede de Londres completa para o cálculo. O estudo mostra que a configuração local, no contexto

urbano, pode explicar os movimentos e, possivelmente, a densidade de atividades e usos dos locais, apesar de não apresentarem resultados quantitativos para este último argumento. Neste sentido, outros estudos relacionam diretamente a intensidade e tipo do uso do solo com as centralidades das redes viárias (KIM; DONG, 2002; RUI; BAN, 2014; LIU *et al.*, 2016), mostrando que atividades econômicas, especialmente aquelas que se beneficiam de movimentos cotidianos, têm forte relação com a configuração da rede, contexto também observado por Porta *et al.* (2012) e Scoppa e Peponis (2015), mesmo para diferentes estruturas viárias, como constatado por Omer e Goldblatt (2016).

A comparação entre fluxos e centralidades é mais direta, visto que ambos são calculados na própria estrutura da rede. Já as comparações entre atividades econômicas e quaisquer informações advindas de redes viárias não são triviais, pois são comumente representadas por diferentes geometrias – atividades são representadas por pontos, e redes por linhas. Alguns estudos buscaram aliar as atividades às ruas mais próximas (SCOPPA; PEONIS, 2015; OMER; GOLDBLATT, 2016). Outros estudos, no entanto, aplicaram a Estimativa de Densidade de Kernel (KDE) para observar a relação das centralidades com as atividades (PORTA *et al.*, 2012; RUI; BAN, 2014; LIU *et al.*, 2016). Esta abordagem transforma ambos os bancos de dados em superfícies contínuas de mesma resolução, permitindo a correlação da superfície como um todo.

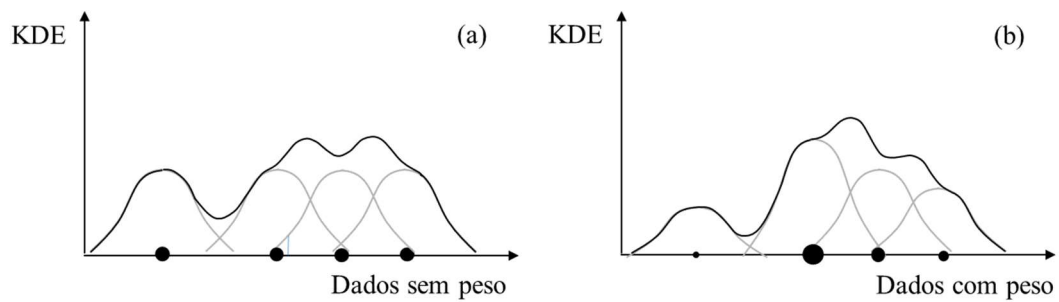
Nenhum destes estudos, no entanto, buscou agregar os dados em zonas de tráfego (ZTs), ou em zonas administrativas, elementos pelos quais autoridades locais e planejadores embasam as tomadas de decisão e políticas públicas urbanas (WANG *et al.*, 2012). A atração de viagens, em pesquisas de Origem e Destino, é um exemplo de informação que comumente é agregada por ZT. Jayasinghe, Sano e Rattanaporn (2017), com base na Sintaxe Espacial, buscaram relacionar a configuração da rede urbana com as viagens atraídas por zonas para a cidade de Colombo, no Sri-Lanka. Também utilizando KDE para as centralidades, obtiveram a média dos valores dos pixels em cada ZT pré-estabelecida pela autoridade local. Os resultados mostraram altas correlações com a densidade de viagens atraídas para cada zona. Porém, os autores só utilizaram a média das KDE e a densidade de viagens para a comparação, sem observarem diferentes parâmetros estatísticos em conjunto com a soma de viagens atraídas.

4.1.2 Kernel Density Estimation

A KDE visa a estimar a distribuição que um banco de dados pode seguir baseando-se em pontos disponíveis (SILVERMAN, 1986). Ela usa a premissa da geografia, em que a influência de

eventos ou objetos no espaço decaem continuamente na medida em que se distancia destes. Também é utilizada como método de suavização numérica ao longo do espaço (FOTHERINGHAM, 2000). Portanto, a partir de pontos conhecidos, é possível presumir como os pontos se distribuem, como mostra a Figura 23(a). Cada ponto pode conter um peso diferente, o que muda a distribuição prevista (Figura 23(b)). Os pesos podem ser, por exemplo, o valor das centralidades nas vias, para a obtenção das KDE das redes.

Figura 23 – Esquemas de KDE em pontos sem peso (a) e com peso (b).



Fonte: O autor (2020).

A KDE pode ser descrita de forma geral por

$$\hat{f}(p) = \frac{1}{nh^2} \sum_{i=1}^n K\left(\frac{d_i}{h}\right) \quad [3.1]$$

Na qual,

- p : é o local onde a densidade será estimada;
- n : é o número de pontos existentes;
- h : raio ou largura de banda (parâmetro de suavização);
- d_i : distância euclidiana entre o ponto existente e o local p ;
- K : a função de densidade (função Kernel).

A função K , que define o formato da curva de densidade, pode variar, mas os resultados não são tão afetados por esta escolha; diferentemente da largura de banda (h), que tem impacto significativo na suavização da KDE (EPANECHNIKOV, 1969). A função gaussiana é mais comumente utilizada. Para cada $x = d_i / h$, a função é

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad [3.2]$$

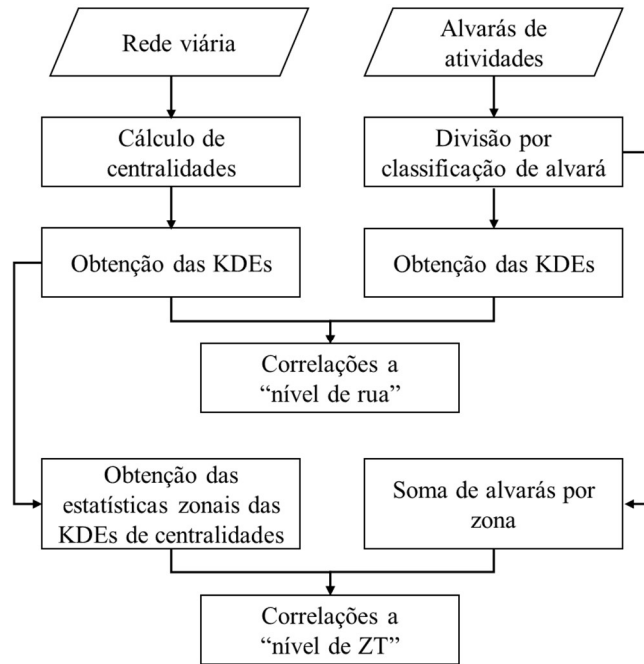
Porta *et al.* (2012) verificaram boas correlações entre as KDEs de centralidades – e compósitos, ou somas, de centralidades – e as atividades econômicas da cidade de Barcelona, na Espanha. As KDEs da rede foram obtidas atribuindo-se peso aos elementos, conforme as centralidades, enquanto que as KDEs das atividades não tiveram pesos atribuídos. Os autores adotaram a largura de banda de 300 metros, por ser o tamanho tradicional de vizinhança (PERRY, 1998). Rui e Ban (2014) utilizaram a largura adaptativa, que varia proporcionalmente à distância entre os pontos observados. Liu *et al.* (2016) demonstram que quanto maior a largura de banda, melhores são os resultados das correlações. Contudo, a variação da largura restringe a influência dos pontos onde eles são numerosos, ao passo que grandes larguras tendem a diluir o efeito dessa influência, dificultando a análise (CARLOS *et al.*, 2010). Portanto, a escolha com base em preceitos urbanísticos de vizinhança, como feita por Porta *et al.* (2012), permite considerar a intensidade dos elementos, bem como sua influência no espaço.

4.2 MÉTODO E PREPARAÇÃO DOS DADOS

Para observar a relação entre a configuração da rede viária, as atividades econômicas da cidade e a atração de viagens, são utilizados os dados oficiais obtidos com a prefeitura de Curitiba e não os dados do *OSM*. Assim, os resultados refletem apenas a relação entre as informações já em uso por autoridades locais, evitando possíveis efeitos de inconsistências do *OSM*. Além disso, como já mencionado, a análise se restringe ao município de Curitiba, uma vez que só existem dados de alvarás comerciais para esta localidade. Mesmo assim, a rede considerada no cálculo das centralidades é para toda a região conurbada de Curitiba, evitando-se efeitos de borda, já que a rede da cidade afeta e é afetada por redes vizinhas.

A Figura 24 esquematiza o método de análise da relação entre centralidades e atividades econômicas, iniciado com a obtenção das superfícies KDE para a correlação mais desagregada e, posteriormente, a nível agregado por ZT. As próximas subseções detalham todo o processo.

Figura 24 – Processo de análise das correlações entre centralidades e atividades econômicas.



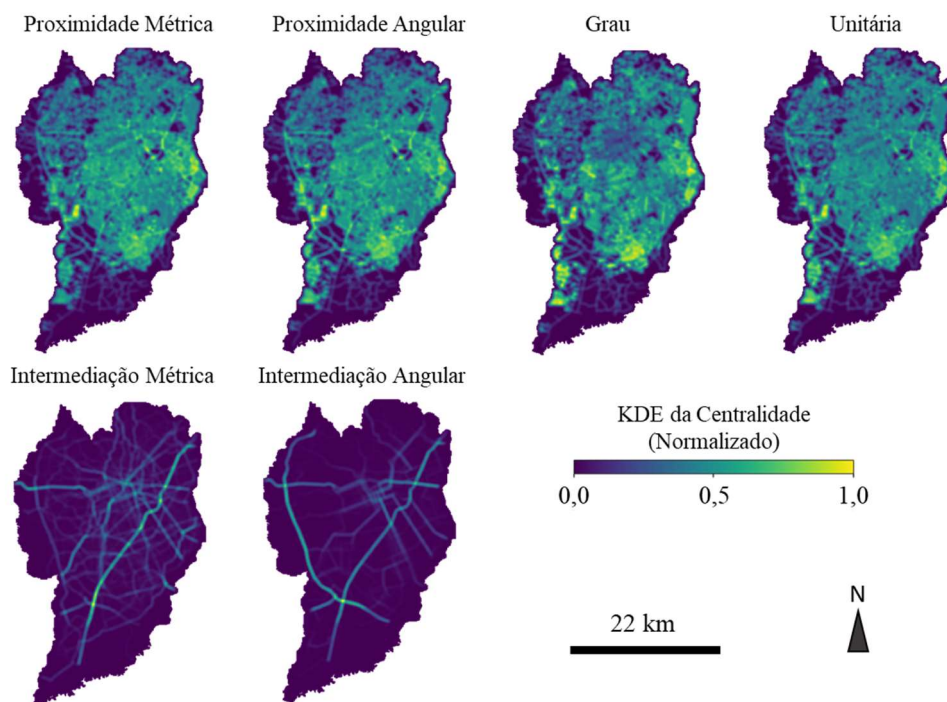
Fonte: O autor (2020).

4.2.1 Obtenção das superfícies de análise

O processo de preparação das KDE segue o padrão adotado por Porta *et al.* (2012). Após serem obtidas as centralidades da rede, aplica-se a KDE tanto para estes resultados quanto para cada classe de atividade econômica, criando-se superfícies contínuas. A resolução para as densidades é de 10 x 10 metros, o que possibilita a análise a “nível de rua”, ou seja, a observação dos dados de forma desagregada, verificando-se o comportamento localizado. A largura de banda é de 300 metros, o tamanho de vizinhança tradicionalmente usado no planejamento urbano (PERRY, 1998).

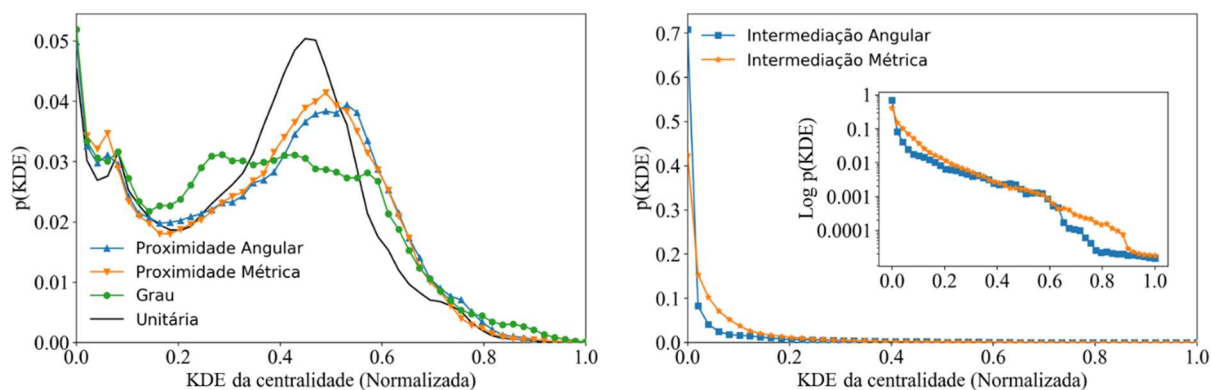
Admite-se, também, a centralidade fictícia chamada de “unitária” como controle, introduzindo-se o valor 1 para todos os nós do grafo *dual* da rede (arestas da rede *primal*). Assim, é possível julgar o efeito que as centralidades causam nas correlações, comparando-as às correlações da centralidade unitária, que não considera estes efeitos. O resultado das KDEs para as centralidades, normalizado entre 0 e 1, é mostrado na Figura 25. As distribuições de probabilidade destes resultados são demonstradas na Figura 26.

Figura 25 – Resultado das KDE de cada centralidade.



Fonte: O autor (2020).

Figura 26 – Distribuição de probabilidades das KDE de cada centralidade.



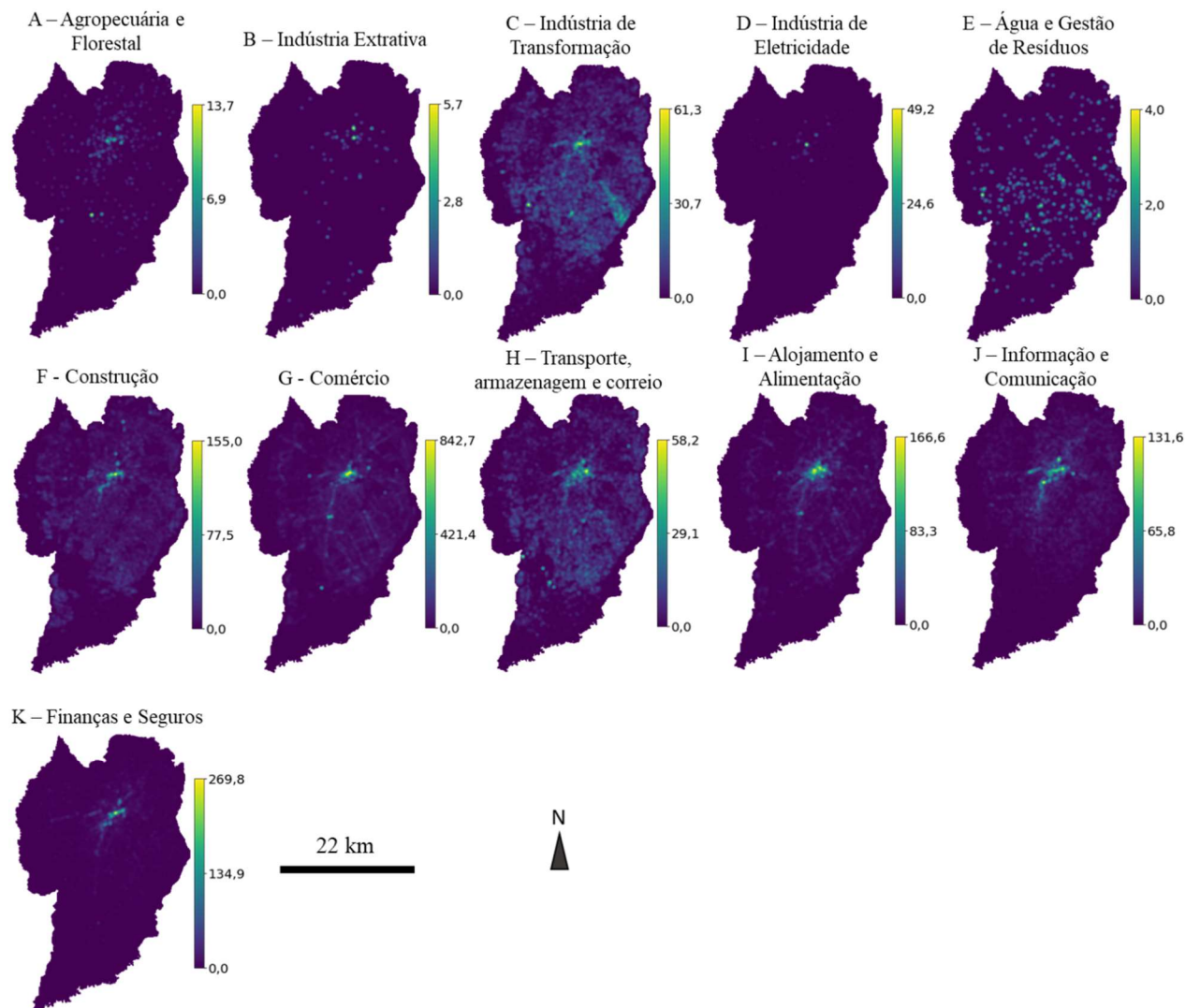
Fonte: O autor (2020).

As distribuições das centralidades de intermediação têm comportamento exponencial, concentrando-se em poucas vias, especialmente nos eixos estruturais da cidade quando a distância métrica é considerada. As centralidades de proximidade métrica e angular originam mapas de KDE similares, com alguns locais de valor elevado longe do centro urbano quando o peso angular é considerado. Já a centralidade de grau apresenta um resultado mais descontínuo, com reduções bruscas próximas de vias principais. Isto ocorre porque esta centralidade é maior onde as vias têm sentido duplo e com mais conversões permitidas, pois o grafo *dual* apresenta mais

arestas nestes locais. A centralidade unitária tem resultado parecido com as de proximidade, porém, reflete apenas a densidade das vias, ignorando-se a configuração espacial. Todas as três – proximidade, grau e unitária – têm maior frequência de menores valores, mas com comportamento similar ao normal para valores maiores.

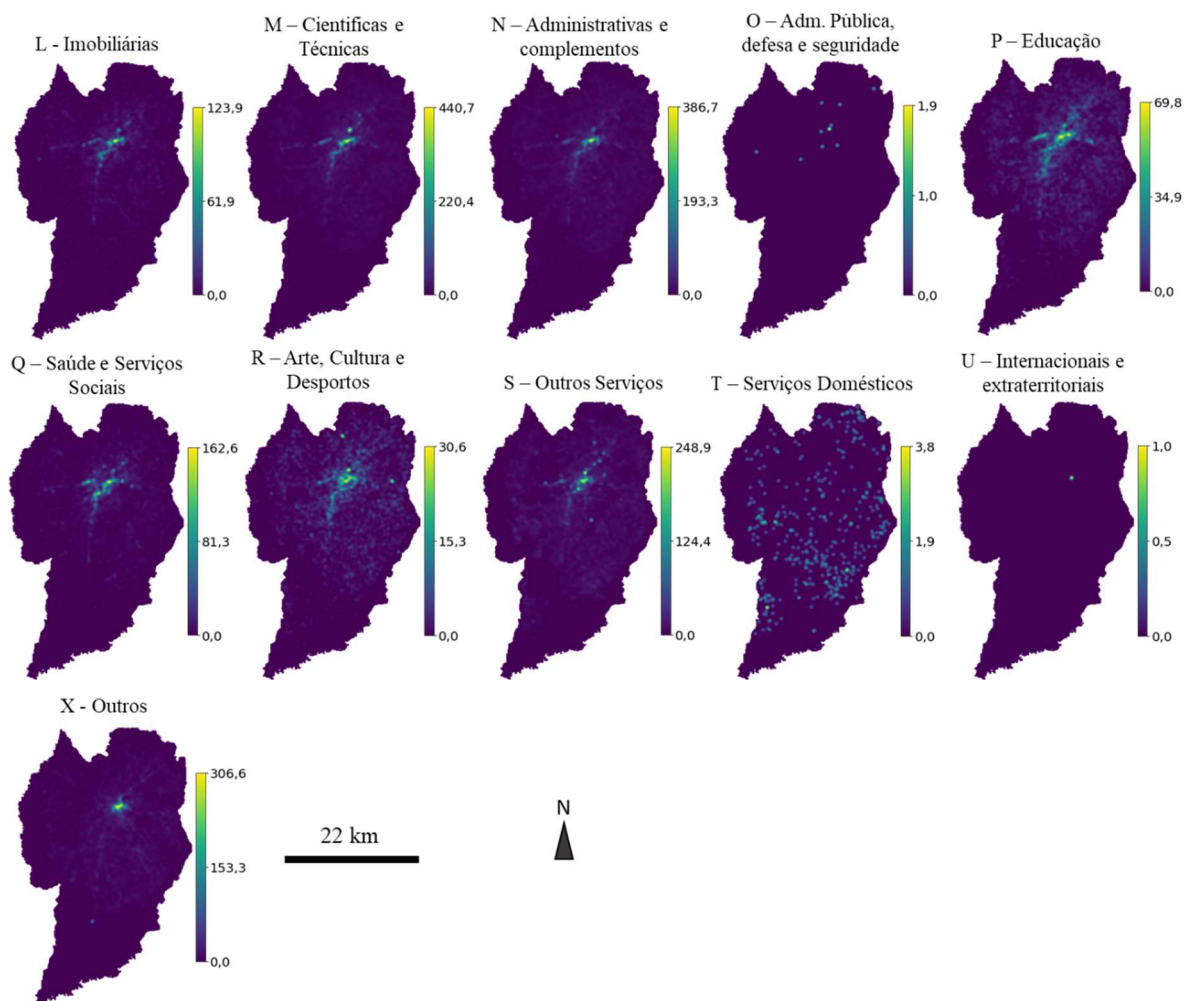
A Figura 27 e a Figura 28 apresentam os resultados de KDE para as classes de alvarás.

Figura 27 – Resultados das KDE de cada classe de alvará (Parte 1 de 2).



Fonte: O autor (2020).

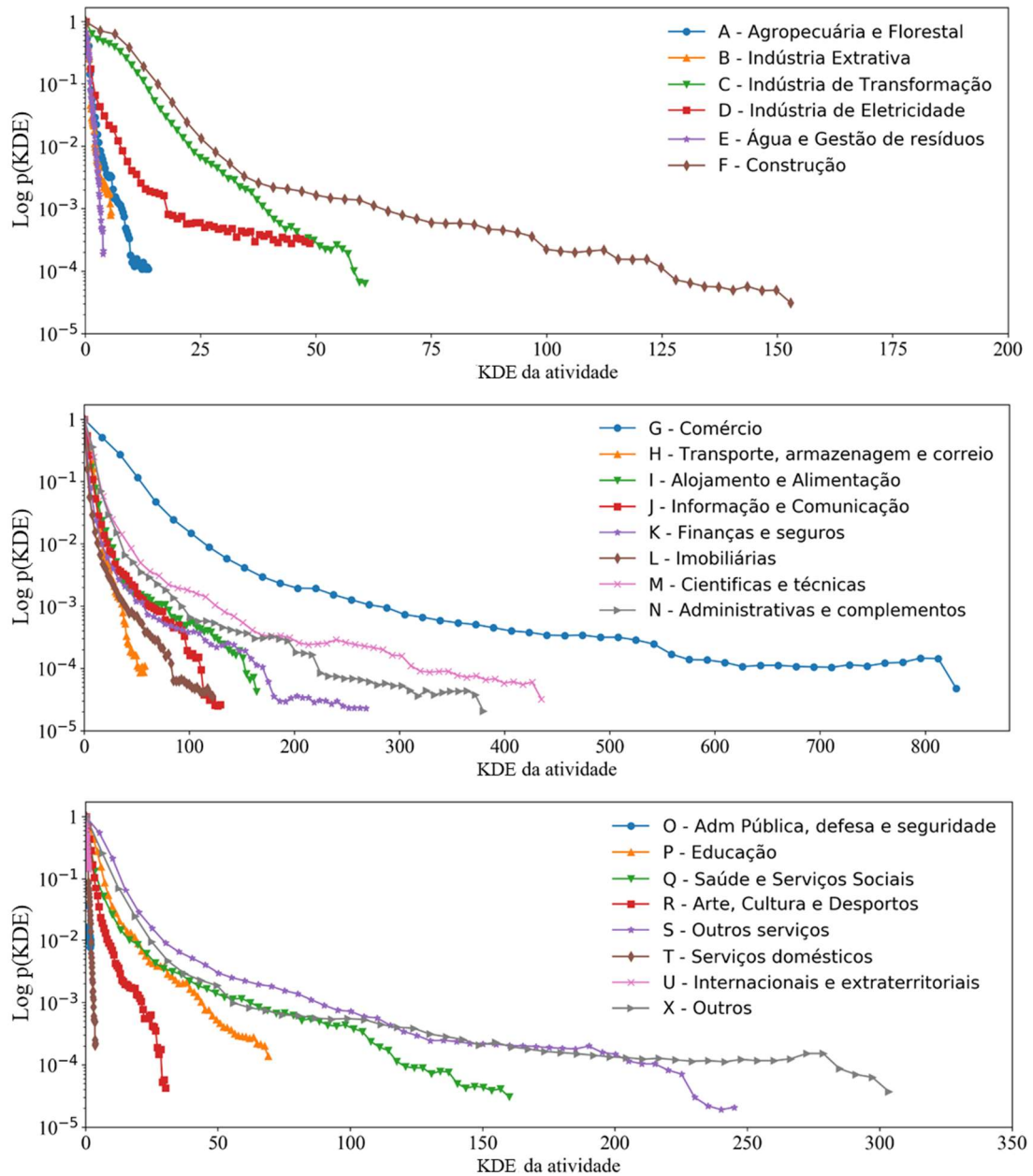
Figura 28 – Resultados das KDE de cada classe de alvará (Parte 2 de 2).



Fonte: O autor (2020).

As superfícies obtidas apresentam comportamento similar, com concentração dos valores preferencialmente no centro da cidade e, frequentemente, ao longo dos eixos estruturais próximos ao centro. Cada mapa tem escala de cores que se adequa ao máximo valor de KDE para cada classe observada. Assim, é possível perceber a dispersão de cada classe, demonstrando-se similaridade entre elas mesmo que com valores diferentes (ex.: Outros Serviços e Comércio), pois algumas classes têm valores muito menores do que outras, o que dificultaria a observação. As distribuições dos valores das KDE podem ser observadas na Figura 29.

Figura 29 – Distribuição de probabilidade das KDE das classes de alvará.



Fonte: O autor (2020).

Todas as atividades econômicas apresentam muitos valores baixos, com concentração de valores altos principalmente no centro e nos eixos estruturais conforme os mapas apontam, exceto para atividades industriais, extrativas e de serviços domésticos, que são mais dispersas, mesmo que apresentem distribuições estatísticas similares às demais classes de alvarás. Através das distribuições não normalizadas pode-se notar as classes com maior número de alvarás, como exemplo, o Comércio.

4.2.2 Correlações a nível de rua e de zonas

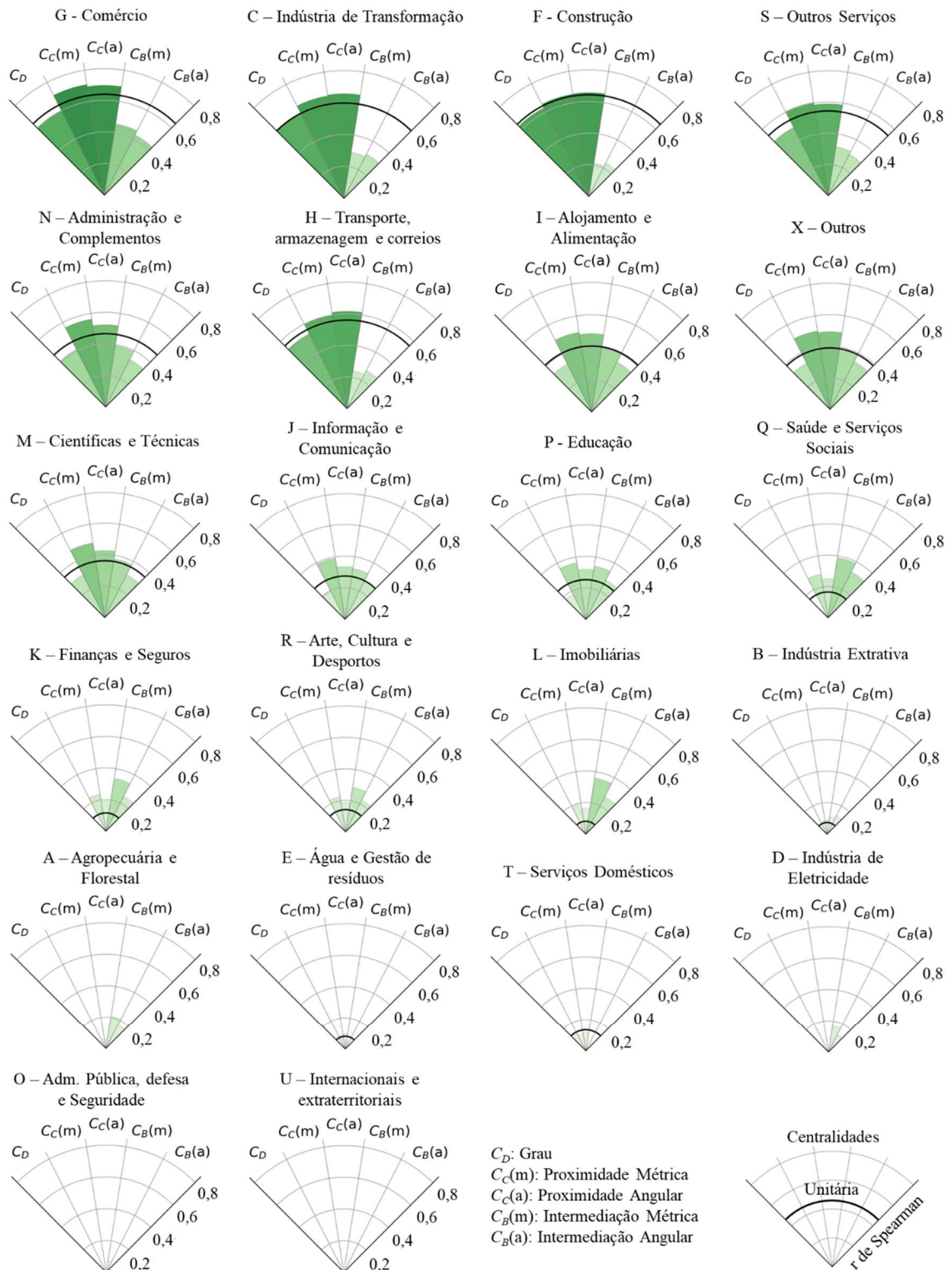
A correlação é observada pelo coeficiente não-paramétrico Spearman (r_s), visto que não se pode supor a normalidade das distribuições demonstradas anteriormente. Cada classe de alvará comercial será correlacionada a todas as centralidades, inclusive à centralidade unitária, para controle, considerando as KDEs resultantes para aferir os resultados “a nível de rua”, ou seja, observando a concentração contínua no espaço de ambas as distribuições.

Para se avaliarem as relações nas ZTs, primeiramente obtém-se o número de alvarás de cada classe nas zonas (Apêndice 1). Para as centralidades, as KDEs ainda são consideradas, porém, é necessário representá-las de forma agregada. Como cada zona tem diferentes tamanhos, as centralidades não são medidas cumulativas e não se pode pressupor o comportamento normal dos *rasters*, portanto, opta-se por agregar as KDEs em quatro parâmetros estatísticos: a média (\bar{X}), a mediana (M_d), o terceiro quartil (Q3) e 85º percentil (P85) das KDEs em cada zona. O resultado das estatísticas por zona é encontrado no Apêndice 2. Cada parâmetro é correlacionado com o total de alvarás de cada classe.

4.3 RESULTADOS

A Figura 30 apresenta os resultados por meio de gráficos de radar. Para cada classe, observam-se as correlações com as diferentes centralidades e com a centralidade unitária. A ordem dos gráficos é decrescente conforme a soma das correlações.

Figura 30 – Correlações entre *rasters* das KDE de centralidades e de classes de alvarás.



Fonte: O autor (2020).

De modo geral, as centralidades de proximidade têm maiores correlações – superiores a 0,60 para as quatro primeiras e a sexta classe de atividade –, seguidas da centralidade de grau – entre

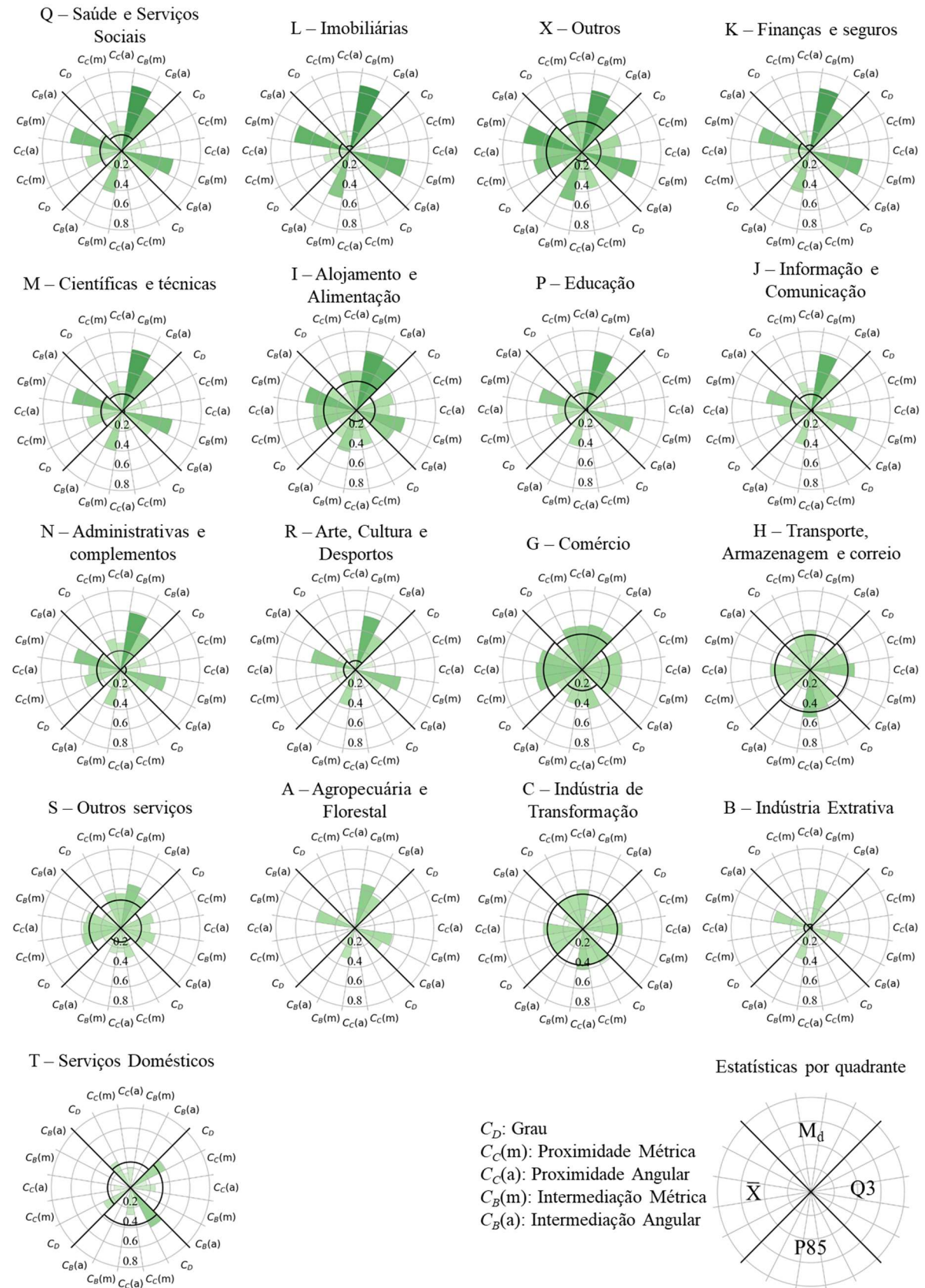
0,40 e 0,60 para as seis primeiras classes. A sequência dos gráficos acaba por refletir as quantidades de alvarás em cada classe, o que afeta as correlações. Contudo, a importância da centralidade para as atividades se dá na comparação com a centralidade unitária. Neste caso, as centralidades de proximidade têm diferenças maiores, e mais importância, na medida em que as correlações diminuem, demonstrando-se que quando se observam classes com menos pontos, estes pontos tendem a se encontrar em ruas mais integradas à rede. A proximidade métrica se mostra um pouco superior à angular para a maioria das classes, com algumas exceções, como as classes C – Indústria de Transformação e H – Transporte, armazenagem e correios.

As centralidades de intermediação apresentam correlações inferiores à unitária para as primeiras classes demonstradas, porém, este comportamento se inverte para classes com menos pontos, como para Q – Saúde e serviços sociais, K – Finanças e seguros e L – Imobiliárias. Mesmo as correlações sendo baixas, os valores são similares para estas três classes, entre 0,35 e 0,40, enquanto que a centralidade unitária não supera 0,20. A importância relativa que a intermediação, especialmente a métrica, apresenta para classes com menos pontos sugere que o posicionamento em vias importantes para os caminhos dentro da rede também é considerado, mas a correlação se dissipa por ser uma centralidade muito concentrada.

Muitas atividades com maiores correlações estão relacionadas ao comércio ou aos serviços ao público, com exceção da C – Indústria de Transformação e da N – Administrativos e Complementos. Contudo, estas duas classes também apresentam atividades relacionadas ao atendimento comercial, como copiadoras e fábrica-loja de móveis na classe C e lotéricas, locadoras de veículos e agências de viagem na classe N. Portanto, ainda apresentam atividades dependentes de “visibilidade”, fator esperado para o princípio da economia do movimento (HILLIER, 1996). A classe F – Construção, relacionada a serviços para edificações, apresenta a correlação da centralidade unitária muito similar às outras centralidades, próximas a 0,64, além de correlação maior com o grau, diferentemente das outras classes, demonstrando-se que esta classe de atividade é numerosa e distribuída de forma que sua relação com as vias não aparenta estar associada à configuração da rede.

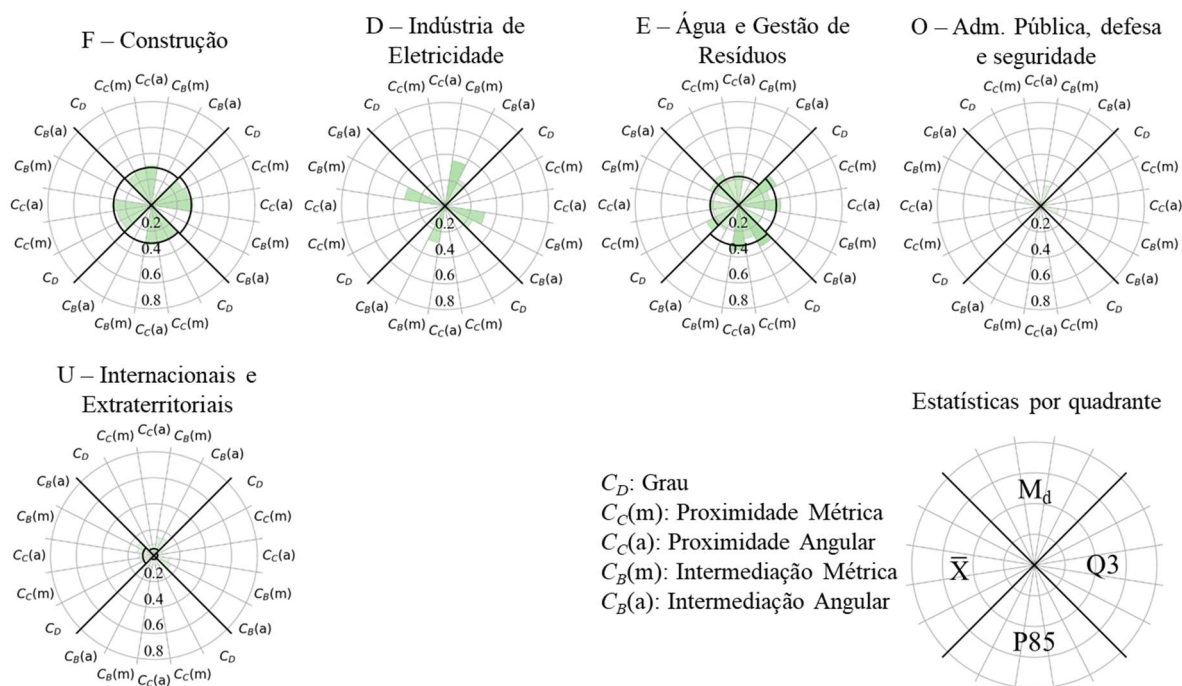
As correlações entre número de atividades e as estatísticas representativas nas ZT são apresentadas na Figura 31 e Figura 32. Os resultados também são dispostos em gráficos de radar, mas agora cada quadrante dos gráficos demonstra os resultados para um dos parâmetros estatísticos considerados.

Figura 31 – Correlações de Spearman entre estatísticas de centralidades e soma das classes de alvarás em cada ZT (Parte 1 de 2).



Fonte: O autor (2020).

Figura 32 – Correlações de Spearman entre estatísticas de centralidades e soma das classes de alvarás em cada ZT (Parte 2 de 2).



Fonte: O autor (2020).

Os maiores valores de correlação ocorrem principalmente para as centralidades de intermediação, especialmente com impedância métrica, com valores iguais ou superiores a 0,60 para as sete primeiras classes. Estes valores se encontram no quadrante da mediana, parâmetro mais representativo para dados não-paramétricos. Observando-se cada classe individualmente, as centralidades de proximidade, grau e unitária têm maiores valores para a média das KDEs nas zonas, o que pode ser um reflexo do comportamento próximo ao normal para valores elevados de KDE. No geral, a centralidade unitária apresenta correlações menores do que as centralidades de intermediação e proximidade, independentemente da estatística, demonstrando que o efeito das centralidades na rede tem importância para as zonas com maior número de atividades. A centralidade de grau não apresenta distinção notável em comparação com a unitária, tendo quase sempre uma correlação menor.

Desta vez, a ordem das classes não segue o tamanho dos bancos de dados, deixando evidente que classes mais vinculadas a comércio e serviços ao público apresentam maiores correlações com zonas intermediárias. A classe G – Comércio em si tem suas maiores correlações com as

médias das centralidades de proximidade, superiores a 0,45, e a mediana das centralidades intermediação, superiores a 0,40, com pouca diferença em relação a centralidade unitária. Talvez por terem o maior banco de dados, os locais de comércio se distribuem mais no espaço, não se concentrando somente nas zonas mais intermediárias como muitas das outras classes. Mesmo assim, é visível a importância dos dois tipos de centralidades – intermediação e proximidade. Já as classes subsequentes, a partir da H – Transporte, armazenagem e correios, têm correlações menores e inconstantes, com algumas sendo mais correlatas ao terceiro quartil e ao 85º percentil das KDEs em cada zona, ou sem maiores efeitos se comparados à centralidade unitária, o que ocorre também para a classe C – Indústria de Transformação e novamente para F – Construção, como ocorreu para a análise a ‘nível de rua’. Estas classes, portanto, demonstram ser pouco impactadas pela configuração da rede.

Outrossim, é interessante observar que muitas classes apresentam grande importância das centralidades de intermediação, mas algumas também demonstram correlações mais expressivas para centralidades de proximidade, principalmente para as classes X – Outros e I – Alojamento e Alimentação. Isto demonstra que há importância tanto da intermediação na rede, ligada a movimentos preferenciais, quanto da proximidade, ligada a acessibilidade, e em ambas, concomitantemente, a centralidade unitária também apresenta maiores correlações. Considerando-se que este comportamento não segue a quantidade de alvarás em cada classe, é um indicativo de que estas atividades estão mais dispersas na rede, mas optam por locais acessíveis e intermediários, diferentemente de Q – Saúde e Serviços Sociais e L – Imobiliárias, por exemplo, que buscam posicionamentos mais concentrados em zonas apenas intermediárias.

Ademais, a impedância angular tanto para a intermediação quanto para a proximidade se mostrou importante, ainda que um pouco inferior à impedância métrica. Portanto, a premissa de economia do movimento é visível a nível agregado de ZT, mas as políticas de uso do solo podem ter influência nos resultados, visto que o planejamento de Curitiba busca fomentar as atividades econômicas ao longo dos eixos estruturais, mais relacionados com as distâncias métricas que são tradicionalmente adotadas para o planejamento urbano.

4.4 CONSIDERAÇÕES FINAIS

Os resultados das correlações diretas entre as KDEs de centralidades e de atividades, a ‘nível de rua’, corroboram estudos como Porta *et al.* (2012) e Omer e Goldblatt (2016), em que a proximidade – acessibilidade ou integração em rede – tem maior correlação, mas demonstram que este resultado é muito dependente do tamanho do banco de dados e da presença de vias, fatores observados com a centralidade fictícia unitária de controle. Já a nível agregado de ZT, a mediana das centralidades de intermediação é expressivamente relacionada ao número de atividades se comparada à centralidade de controle. Isto demonstra que as atividades tendem a se concentrar em zonas intermediárias e em ruas mais acessíveis, mas não necessariamente em ruas intermediárias.

A importância das centralidades com impedância métrica é ligeiramente superior à angular, especialmente para os resultados zonais. Vale se atentar para a influência do planejamento da cidade neste nível de análise. Os conceitos de Desenvolvimento Orientado ao Transporte Público (DOT) de Curitiba encorajam atividades ao longo dos eixos estruturantes, que são também preferenciais para caminhos de menor distância métrica. Os resultados zonais não refletem apenas as escolhas dos comerciantes e donos das atividades, mas o processo de planejamento da cidade. Ainda assim, as distâncias angulares se mostraram importantes se comparadas à centralidade de controle. Os resultados são semelhantes para outras cidades com contextos de crescimento “natural”, em que as atividades deliberadamente escolhem regiões intermediárias e com boa acessibilidade (OMER; GOLDBLATT, 2016). Ou seja, a estruturação de Curitiba, seguindo preceitos de DOT, encoraja o posicionamento das atividades conforme seu desejo “natural”, que busca acessibilidade e movimento da região.

Observar as relações entre centralidades e atividades econômicas não só permite verificar a importância da configuração da rede para o posicionamento das atividades, mas também auxilia no entendimento de efeitos secundários deste posicionamento. Por exemplo, nota-se maior concentração de atividades relacionadas a P – Educação em zonas intermediárias, em detrimento de zonais mais integradas, diferentemente de I – Alojamento e Alimentação, com maior dispersão, o que indica que as viagens relativas à educação se concentram nestas zonas, e as relativas a hotéis e restaurantes se distribuem mais. Outro exemplo é da classe G – Comércio, mais distribuída nas zonas, como atestado pela importância relativa com a centralidade unitária, o que indica que os movimentos podem ser menores para que os habitantes atinjam estes locais. Ou

mesmo para a intensidade de transporte de cargas, que é mais distribuído para atender o comércio do que restaurantes.

Assim sendo, pode-se afirmar que as centralidades têm relação com as atividades atrativas de viagens. A acessibilidade, juntamente com o princípio da economia do movimento da Sintaxe Espacial, afeta a localização das atividades através de uma relação mútua entre uso do solo e oferta de infraestrutura de transporte. Ao se corroborar a Sintaxe Espacial, fica mais clara a relação colinear entre a atratividade de viagens e a acessibilidade. A dificuldade encontrada para modelar a demanda por viagens considerando-se variáveis de acessibilidade pode ser justificada pelo posicionamento de atividades optado, ou incentivado, em locais mais acessíveis (EWING; DEANNA; LI, 1996), especialmente do ponto de vista global. Deve-se saber, no entanto, como as medidas de centralidades se comportam ao explicarem a atração de viagens juntamente com as atividades e entender como este dado – a rede urbana – pode ser benéfico para esta explicação.

O planejamento de transporte e as políticas de uso e ocupação do solo podem se beneficiar com os resultados apresentados, entendendo-se que ruas mais integradas incentivam usos ligados ao público, e que regiões não intermediárias podem inibir certas atividades econômicas. Mais do que isso, a rede viária e sua configuração passam a ser importantes fontes de informação usadas para promover o uso sustentável do espaço urbano. No contexto de massificação de dados geográficos através das IGV, os presentes resultados valorizam as redes urbanas disponibilizadas. Portanto, é importante avaliar a qualidade dos dados disponíveis para possíveis aplicações.

5 AVALIAÇÃO DA QUALIDADE DOS DADOS DO OPENSTREETMAP

Após verificar que as centralidades e, portanto, a rede, podem auxiliar na explicação das atividades e possivelmente das viagens urbanas, este capítulo busca avaliar a qualidade dos dados obtidos do *OSM*. Ao se compararem os *POIs* com dados georreferenciados oficiais, e a rede do *OSM* com a rede oficial, é possível aferir semelhanças e diferenças que impactam nos modelos explicativos, e que são a aplicação desejada.

5.1 CONTEXTUALIZAÇÃO

Com a rápida expansão de sistemas de informação geográfica (SIG), a partir anos 1980, a atenção dada à qualidade destas informações também cresceu (VAN OORT, 2006), recebendo normativas da *International Organization for Standardization* (ISO) de controle de qualidade após 2002, com a ISO 19113 e a ISO 19114, que, a partir de 2013, foram unificadas na norma ISO 19157 (NEIS; ZIELSTRA, 2014). O *OpenStreetMap* entra neste contexto como um banco de dados aberto e massivo, cuja qualidade se propõe a igualar ao nível de confiança de bancos de dados oficiais (GOODCHILD, 2007), tendo como princípio a Lei de Linus, a qual determina que grandes números de usuários e colaboradores permitem autocorreções e redução dos erros dos dados (NEIS; ZIELSTRA, 2014).

Van Oort (2006) traz uma revisão que aponta diferentes quesitos de avaliações de qualidade de dados geográficos, com base nas normativas citadas e em diversos outros estudos. Com base nestes padrões, os dados do *OSM* são avaliados. Os quesitos são sumarizados na Tabela 4. Deve-se, no entanto, considerar o uso para o qual os dados estão sendo avaliados (HAKLAY, 2010; GIRRES; TOUYA, 2018), ou seja, a qualidade esperada para o uso neste estudo pode ser diferente para outros tipos de análises e aplicações. Assim, se ao observar cada quesito, os pressupostos para a aplicabilidade devem ser discutidos.

Tabela 4 – Quesitos para avaliação da qualidade de informações geográficas.

Tópico	Descrição
<i>Lineage</i> ou Linhagem	Histórico do banco de dados, com descrição da origem e das modificações nos dados produzidos.
Qualidade Temporal	Validade das mudanças nos dados ao longo do tempo em relação ao que se considera a realidade.
Acurácia Semântica	Fidelidade da representação dos objetos ao significado e interpretação que eles devem ter.
Acurácia de Atributos	Precisão dos outros atributos que não posicionais e temporais, como atributos nominais.
Consistência Lógica	É a fidelidade das relações internas do banco de dados (ex.: ruas sem saídas não podem ter sentido único).
Compleitude	Medida de ausência ou excesso de dados em relação ao que se considera a realidade.
Acurácia Posicional	Precisão da posição dos elementos nas coordenadas geográficas.
Usabilidade, Propósito e Restrições	Possibilidade de aplicação dos dados a certo uso ou propósito, estando dependente do usuário dos dados.

Fonte: van Oort (2006), Haklay (2010) e Girres e Touya (2018).

A seguir, faz-se uma breve descrição dos principais quesitos observados em estudos para os dados do *OSM*, com exceção do quesito usabilidade em si, que é posteriormente discutido para auxiliar na decisão dos quesitos necessários para a aplicação deste trabalho.

5.1.1 A qualidade dos dados do *OSM*

5.1.1.1 *Lineage*

O quesito *lineage*, ou linhagem, dos dados visa contextualizar o banco de dados, trazendo informações sobre sua construção. Para o *OSM*, a linhagem é parcialmente conhecida e aberta ao público. Há certa transparência quanto ao histórico da edição dos dados (HAKLAY; WEBER, 2008), porém pode haver usuários anônimos, portanto não identificáveis, e os métodos para aquisição dos dados também não são conhecidos, o que reduz a qualidade da linhagem (GIRRES; TOUYA, 2018). Em contrapartida, a qualidade temporal dos dados pode assim ter vantagens, uma vez que o número de contribuições na adição e edição torna o *OSM* dinâmico e mais facilmente atualizável (NEIS; ZIELSTRA; ZIPF, 2013).

5.1.1.2 *Qualidade temporal*

A qualidade temporal é observada com a modificação do banco de dados em relação ao que ocorre na realidade. Como exemplo, Zhang e Pfoser (2019) constataram boa qualidade temporal dos dados de *POIs* para cafeterias em Manhattan, Nova York, ainda que a acurácia posicional e a completude não fossem aceitáveis. Ou seja, a taxa de modificação e adição dos pontos seguiu o esperado para o período. O presente trabalho não observa a mudança temporal dos dados, mas busca garantir a equivalência temporal dos dados oficiais com o *OSM*. Portanto, os dados da rede oficial foram concebidos ao final de 2018 e os bancos de dados de pontos e atividades são do início de 2019, enquanto que todos os dados do *OSM* foram obtidos no início de 2019.

5.1.1.3 *Acurácia Semântica e Acurácia de Atributos*

A semântica dos dados do *OSM* é associada à estrutura de ‘tags’, ou atributos, com as informações necessárias para caracterizar cada elemento. As informações buscam certas predefinições observadas na [wiki.openstreetmap](https://wiki.openstreetmap.org/)⁸, porém, os usuários podem adicionar atributos, e as nomenclaturas não são controladas. Esta liberdade, segundo Girres e Touya (2018), leva locais com usos e atribuições semelhantes a terem diferentes representações nos atributos e, portanto, baixa qualidade semântica, como mensurado para redes viárias na França. Alguns métodos foram propostos para avaliar e reduzir estas falhas através de modelos para agregação de nomenclaturas e disposição espacial dos *POIs* (MÜLLIGANN *et al.*, 2011; BALLATORE; BERTOLOTTI, 2013). Contudo, estes métodos observam o banco de dados do *OSM* em si, sem comparação com outras bases.

A comparação semântica automática pode ser impraticável, uma vez que não há relação direta entre os elementos comparados, especialmente para *POIs*, que podem representar locais não existentes em bancos de dados oficiais, ou utilizar atributos que não permitem a relação direta (GIRRES; TOUYA, 2018). Machado e Camboim (2019), nesse sentido, categorizaram manualmente dados oficiais e do *OSM* para uma região de Curitiba, constatando maior variabilidade de informações e detalhamento do mapa colaborativo do que os dados oficiais, recomendando a integração dos dados para vantagens no mapeamento da cidade e, portanto, no planejamento urbano.

⁸ <https://wiki.openstreetmap.org/wiki/OSMonto>

5.1.1.4 *Consistência Lógica*

A consistência lógica é relacionada com a geometria e inter-relação dos elementos. Neis, Zielstra e Zipf (2012) relacionam a topologia da rede a esta verificação, considerando que as rotas e caminhos possíveis da rede estão relacionadas à qualidade lógica destes elementos. Os autores encontraram falhas nestes quesitos, mas uma melhora ao longo do tempo, entre 2007 e 2011 para a rede da Alemanha. Mais tarde, Girres e Touya (2018) também encontraram erros topológicos na rede da França, apesar de não demonstrarem numericamente. Os erros são atribuídos à falta de restrições para o manejo dos elementos. Os autores também observaram erros lógicos para elementos relacionados a lagos, limites administrativos e costas.

5.1.1.5 *Compleitude e Acurácia Posicional*

Grande parte dos estudos que avaliam a qualidade dos dados do *OSM* foca na completude e acurácia posicional da rede viária. Estes elementos comumente apresentam boa completude, mas heterogeneidade na acurácia posicional. Em áreas urbanas ou mais ricas a qualidade do banco de dados se assemelha a bases de dados comerciais, porém, para regiões rurais e mais pobres a qualidade pode cair drasticamente (HAKLAY, 2010; GIRRES; TOUYA, 2018; NEIS; ZIELSTRA; ZIPF, 2012). Estas diferenças também são observadas em diferentes regiões do planeta, com a qualidade sendo relacionada à renda per capita dos países e também ao número de contribuições para os mapas (NEIS; ZIELSTRA; ZIPF, 2013). Para *POIs*, estas avaliações são mais difíceis devido às dificuldades semânticas mencionadas e à falta de dados oficiais validados para a comparação, por vezes restando comparações entre outros bancos de dados online, como fizeram Hochmair, Juhász e Cvetojevic (2018), que compararam os pontos do *OSM* com a base de dados do *Google Maps* e de redes sociais para cidades americanas e europeias. A avaliação mostrou menor densidade para o *OSM*, com concentração em locais de atividade acadêmica, porém, uma maior coerência para ambos os mapas em comparação com as redes sociais.

5.1.2 **Quesitos para a usabilidade**

Para a usabilidade dos dados na análise de atração de viagens, deve-se considerar a agregação das viagens em *ZTs*, ou seja, a análise deve considerar que as informações são agregadas em zonas. Tanto para a rede quanto para os *POIs*, a acurácia posicional e a completude são pré-requisitos claros. A acurácia semântica, ou de atributos, deve guiar a verificação da acurácia posicional e a completude dos *POIs*, porque se trata do reconhecimento do uso que se faz de cada local. Contudo, dado o nível de agregação necessário, não é preciso realizar a comparação

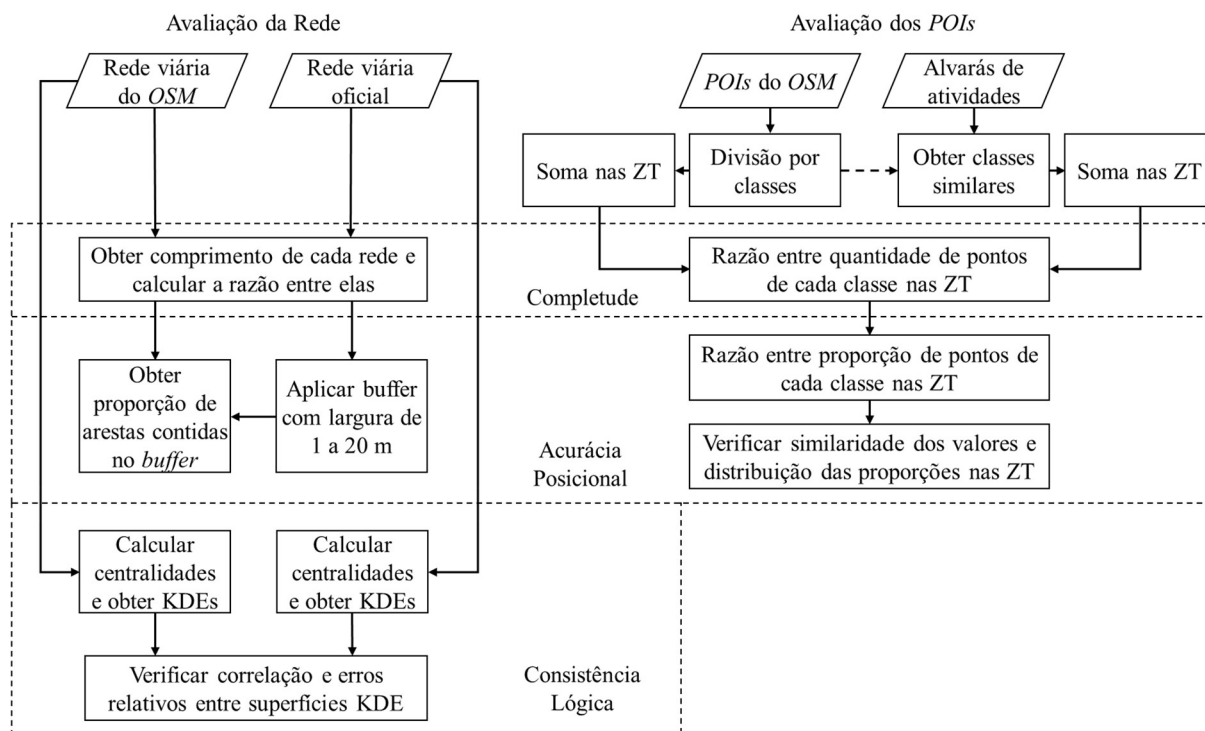
semântica ponto a ponto com os dados oficiais, mas sim a comparação das distribuições de ambos os bancos de dados nas ZT como aplicado em Zhang e Pfoser (2019), fixando-se uma semântica de classificação dos pontos.

Para a rede, a completude e a acurácia posicional devem ser observadas de forma desagregada, uma vez que a configuração é importante no cálculo das centralidades. Porém, diferentemente dos *POIs*, os atributos e, portanto, a acurácia semântica, não são relevantes para a aplicação proposta, visto que o cálculo das centralidades não é impactado pelo nome da via, hierarquia ou mesmo número de faixas. Não obstante, um fator importante para a rede é a consistência lógica, pois está intimamente relacionada com a navegação e topologia (GIRRES; TOUYA, 2018; NEIS; ZIELSTRA; ZIPF, 2012), influenciando, portanto, os resultados das centralidades.

Note-se que atentar para certos quesitos em detrimento de outros possibilita focar na qualidade desejada para o uso em questão, ou seja, permite avaliar a usabilidade para o presente propósito. Outras aplicações podem necessitar que outros quesitos sejam verificados, ainda que não expressos neste trabalho.

5.2 MÉTODO PARA AVALIAÇÃO DA QUALIDADE DOS DADOS

Os bancos de dados da rede e dos *POIs* do *OSM* são avaliados conforme a completude e a acurácia posicional, além da consistência lógica da rede, tendo em vista a usabilidade proposta. Para tanto, adotam-se os dados oficiais adquiridos com as autoridades locais como base da realidade a ser avaliada, ou seja, quanto mais próxima dos dados oficiais, maior a qualidade dos dados do *OSM*, como efetuado por diversos estudos (GIRRES; TOUYA, 2018; HAKLAY, 2010; GOODCHILD *et al.*, 1992; HUNTER, 1999). Mesmo que haja possíveis divergências dos dados oficiais com a realidade física, eles são utilizados como informações válidas para tomadas de decisões das agências locais, portanto, adotá-los como base para comparação é pressupor sua qualidade prática. O método de análise de cada um dos bancos de dados é esquematizado na Figura 33, sendo detalhado nas subseções a seguir.

Figura 33 – Processo de verificação da qualidade dos dados do *OSM*.

5.2.1 Avaliação da rede

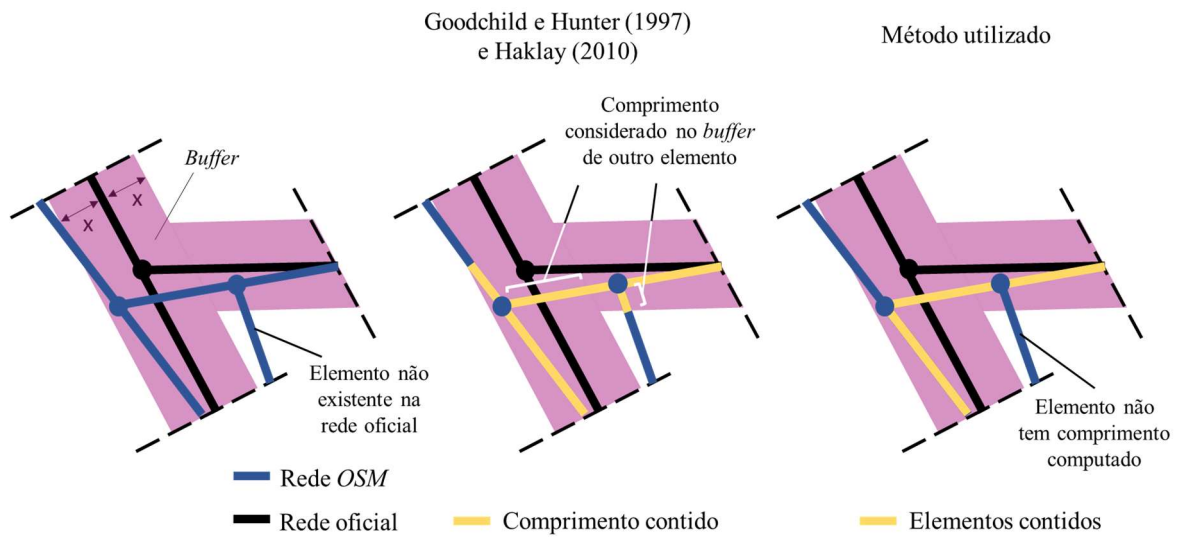
5.2.1.1 Compleitude

Para observar a diferença numérica de elementos entre a rede oficial e do *OSM*, adota-se o critério das distâncias de rede, como utilizado por Haklay (2010) e Goodchild *et al.* (1992). Portanto, dentro de cada ZT faz-se a razão entre o comprimento da rede do *OSM* e o comprimento da rede oficial, ou seja, se a rede do *OSM* tiver menor comprimento, portanto, menos informações, do que a rede oficial, a relação é menor do que 1, e maior caso contrário. O desejado é que todas as zonas de tráfego apresentem essa relação próxima ou igual a 1.

5.2.1.2 Acurácia Posicional

Goodchild e Hunter (1997) e Haklay (2010) utilizaram zonas de *buffer* ao redor da rede viária oficial e calcularam a proporção do comprimento da rede a ser testada contida neste *buffer* – neste caso, a rede do *OSM*. A largura do *buffer* aumenta gradativamente a fim de estimar a diferença entre as malhas. Neste trabalho, porém, opta-se por observar o número de elementos, mais especificamente das arestas da rede do *OSM* que estão completamente contidas pelos *buffers*. A proporção do comprimento contido pode trazer erros relacionados a elementos não análogos que estejam próximos, como apontado na Figura 34. Erros advindos da diferença do tamanho das redes podem ser explicados pela análise anterior referente à completude.

Figura 34 - Esquema de análise da acurácia posicional da rede do *OSM* em relação à oficial.



Fonte: O autor (2020).

A largura da zona de *buffer* é variada, de $x = 1$ m até $x = 10$ metros em cada lado dos elementos, com intervalos de 1 m, e mais dois buffers de $x = 15$ m e $x = 20$ m. Ao se calcular, para cada valor de *buffer*, o número dos elementos do *OSM* que estão contidos no *buffer* e dividindo este número pelo total de elementos da rede *OSM*, encontra-se a proporção da rede contida.

5.2.1.3 Consistência Lógica

A consistência lógica foi relacionada com a topologia dos elementos, pois através desta é possível verificar se as inter-relações são similares para as redes comparadas, como embasado por Neis, Zielstra e Zipf (2012) e Girres e Touya (2018). No entanto, nenhum estudo aponta métodos para se observarem numericamente estas relações, pois possíveis diferenças na completude e na acurácia posicionais dificultam este processo. Neste trabalho, contudo, a aplicação é muito dependente desta avaliação, pois a topologia afeta diretamente o resultado das centralidades.

O ponto positivo da aplicação final do presente estudo é a utilização da Estimativa de Densidade de Kernel (KDE) aplicada aos resultados das centralidades. A resolução por células de 10×10 m absorve erros posicionais diretos de até 10 metros; já a largura de banda de 300 m suaviza possíveis erros maiores. Idealmente, ambas as redes seriam idênticas e, portanto, as KDE seriam idênticas, e a correlação entre elas seria de 1. Erros ao longo da rede podem afetar os valores

de centralidades, mas ainda assim apresentar comportamentos similares, portanto, os resultados normalizados seriam similares.

Neste sentido, a observação se restringe ao município de Curitiba, onde a análise final da atração de viagens é feita. Calculam-se as centralidades de proximidade e de intermediação tanto métrica quanto angular e a centralidade de grau para toda a região metropolitana. Após, obtêm-se as KDE de cada resultado para as duas redes, e as superfícies são normalizadas. Faz-se, então, a correlação de Pearson e de Spearman para averiguar a similaridade das KDE, e plota-se a diferença relativa, o que permite a observação dos locais com maiores erros relativos. O erro relativo é calculado com a seguinte equação

$$ER_i = \frac{KDE(osm)_i - KDE(oficial)_i}{KDE(oficial)_i} \quad [4.1]$$

Em que,

ER_i : Erro relativo na célula i ;

$KDE(osm)_i$: KDE normalizada na célula i na rede do *OpenStreetMap*;

$KDE(oficial)_i$: KDE normalizada na célula i na rede oficial.

O resultado ideal seria um raster com valor zero para todas as células; quaisquer diferenças relativas, positivas ou negativas, mostram erros que podem advir da completude e da acurácia posicional, mas principalmente das diferenças topológicas, pois os valores das centralidades são muito impactados por estas.

5.2.2 Avaliação dos Pontos de Interesse

A qualidade dos *POIs* é averiguada em comparação com os alvarás comerciais e outros dados georreferenciados obtidos com a prefeitura de Curitiba, limitando-se a análise apenas para este município. Contudo, a área observada conta com a maioria das ZT (135). Ressalta-se que a avaliação é feita de forma agregada por zonas. Uma verificação ponto a ponto necessitaria que o banco de dados oficial e os *POIs* contassem com semelhança de atributos muitas vezes inexistentes, o que também prejudicaria o uso de informações que, mesmo não idênticas, são análogas nas classes e que, de forma agregada, permitem a utilização proposta. Esta abordagem é semelhante à de Zhang e Pfoer (2019), pois mesmo que os *POIs* do *OSM* não tenham representado todos os locais existentes, a análise agregada detectou as mudanças urbanas observadas.

Como os dados de alvarás apresentam classificações mais segmentadas do que os dados de *POIs*, agrupam-se algumas classes de alvarás de forma a fixar uma equivalência semântica suficiente para a comparação agregada em ZT. Algumas classes podem ser comparadas diretamente entre o banco de dados oficial e do *OSM*. Outras classes de *POIs*, porém, não têm representação análoga nos dados oficiais obtidos, ficando de fora da avaliação. A Tabela 5 a seguir apresenta o agrupamento feito para comparação. As classes de *POIs* não citadas não são avaliadas devido à impossibilidade da comparação (ex.: não existem bancos de dados oficiais de ornamentos, instalações para motocicletas e de locais religiosos para serem comparados a estas mesmas classes de *POIs*).

Tabela 5 – Classes de *POIs* e classes de alvarás e de outros bancos de dados a serem comparadas com as respectivas siglas utilizadas.

Comparação com banco de dados dos alvarás		
Classe de <i>POIs</i>	Classes respectivas dos alvarás	Sigla
Comércio e Serviços	F – Construção	C & S
	G – Comércio	
	H – Transporte, armazenagem e correio	
	I – Alojamento e Alimentação	
	J – Informação e Comunicação	
	K – Finanças e seguros	
	L – Imobiliárias	
	M – Científicas e técnicas	
	S – Outros serviços	
Educação	P – Educação	E
Saúde	Q – Saúde e Serviços Sociais	S
Lazer	R – Arte, Cultura e Desportos	L
Comparação com outros bancos de dados georreferenciados		
Classe de <i>POIs</i>	Banco de dados respectivo	
Transporte Público	Localização das estações de ônibus	TP
Instalações p/ bicicletas	Localização de paracielos	B

Note-se que o agrupamento feito para a classe de Comércio e Serviços tem o intuito de agregar atividades que atraem viagens pelo motivo de trabalho e que, dentro das classificações dos alvarás, apresentem locais de interesse comercial ou de serviço. Classes como F – Construção e H – Transporte, Armazenagem e Correio abrangem atividades de atendimento ao público, como

empregadas e transportadoras, que têm relação distinta com o público em relação a locais como restaurantes e escritórios de advocacia, das classes I – Alojamento e Alimentação, e M – Científicas e Técnicas, respectivamente. A escolha das classes é impactada pelo método de classificação oficial utilizado. Um maior desmembramento das atividades reduziria o banco de dados referente aos *POIs* do *OSM* para comparação, mesmo que tenham um efeito similar na atração de viagens por trabalho. Contudo, é importante destacar que esta agregação pode influenciar nos resultados, e diferentes agrupamentos podem trazer outras perspectivas. Os resultados zonais das classes definidas acima são encontrados no Apêndice 4, tanto para os dados oficiais quanto para os *POIs* do *OSM*.

5.2.2.1 *Compleitude*

Ao se observarem as quantidades dos bancos de dados, é notável que os *POIs* não representam todos os pontos existentes, sendo possível que alguns *POIs* representem pontos inexistentes nos bancos de dados oficiais. Apesar da verificação de cada ponto isoladamente ser impraticável para o presente estudo, a comparação agrupada por classes e por zona de tráfego pode demonstrar a magnitude dessa diferença. Ou seja, a completude, nesta etapa, é obtida ao se calcular o total de *POIs* do *OSM* e dividi-lo pelo total de pontos dos bancos de dados oficiais, para toda a cidade e por zona de tráfego, considerando-se cada classe de ponto.

5.2.2.2 *Acurácia Posicional*

Sabendo da deficiência nos dados de *POIs*, a precisão na localização dos pontos é feita de forma agregada nas ZT, através das classes estipuladas. Análises de atração de viagens podem considerar a dispersão de atividades no território, mesmo não se considerando o valor absoluto das atividades. Portanto, as proporções nas ZT de cada classe podem ser úteis, mesmo que o banco de dados esteja incompleto. A avaliação, assim, consiste em comparar a distribuição espacial dos dados do *OSM* com a distribuição dos dados oficiais. Para cada classe, ambos os bancos de dados são somados por ZT e divididos pelo total da classe, obtendo-se a proporção de pontos que cada zona detém. Após, dois testes de hipóteses são feitos: (i) Kolmogorov-Smirnov (K-S) – para observar a semelhança das distribuições destas proporções – e (ii) Wilcoxon – para observar a semelhança dos valores das proporções em cada ZT.

O teste de Kolmogorov-Smirnov compara duas funções de probabilidade acumuladas, obtendo a maior diferença de probabilidade entre elas. O teste busca verificar se dois conjuntos de dados seguem uma mesma distribuição. O teste de Wilcoxon, por sua vez, é um teste não-paramétrico

para verificação de similaridade de valores de diferentes grupos. Ou seja, este teste observa os valores dos grupos e constata se há uma diferença significativa entre eles, se um grupo tem valores maiores ou menores que o outro, por exemplo (HEUMANN; SHALABH, 2016).

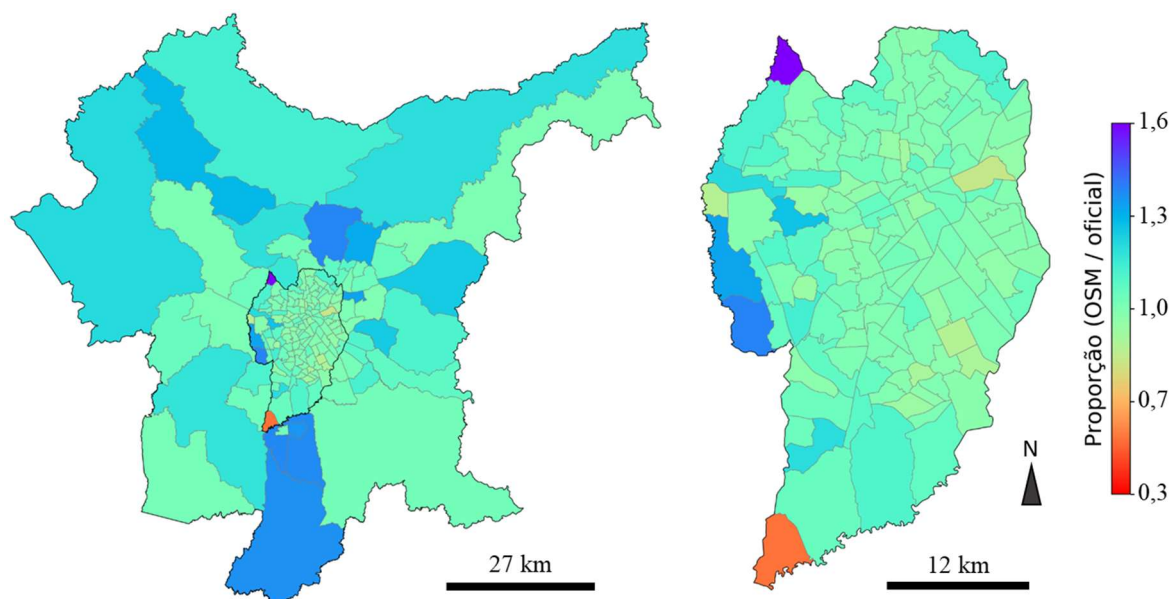
Após os dois testes serem efetuados para cada classe de *POIs*, observa-se a relação entre as proporções dos dois bancos de dados no espaço, isto é, em cada *ZT*, a proporção de *POIs* é dividida pela proporção de pontos dos dados oficiais na mesma zona. Faz-se isso para todas as zonas a fim de se verificar como ocorrem as diferenças entre os dados ao longo do espaço. Portanto, quando esta divisão é igual a 1, significa que a proporção de *POIs* da zona na classe em questão é igual à proporção dos pontos oficiais de mesma classe. Quando os valores são menores do que 1, indicam uma proporção de *POIs* inferior à esperada, e quando maiores do que 1, indicam uma proporção de *POIs* excedente ao esperado de acordo com os dados oficiais. Para além dos testes estatísticos mencionados anteriormente, esta análise visual permite verificar a dispersão espacial dos pontos em ambos os bancos de dados.

5.3 RESULTADOS

5.3.1 Resultados da avaliação da rede

A Figura 35 mostra a relação do comprimento das vias entre a rede *OSM* e a rede oficial nas *ZT*. A menor relação encontrada foi de 0,55 e a maior foi de 1,80, ambas no município de Curitiba. A relação para o município inteiro é de 1,02, enquanto que para o restante da região metropolitana é de 1,12, o que mostra uma representação com mais informações do que a rede oficial, principalmente da rede do contorno da cidade principal. A relação para a rede completa é de 1,09.

Figura 35 - Resultado da completude da rede do *OSM* em relação à oficial para cada zona.

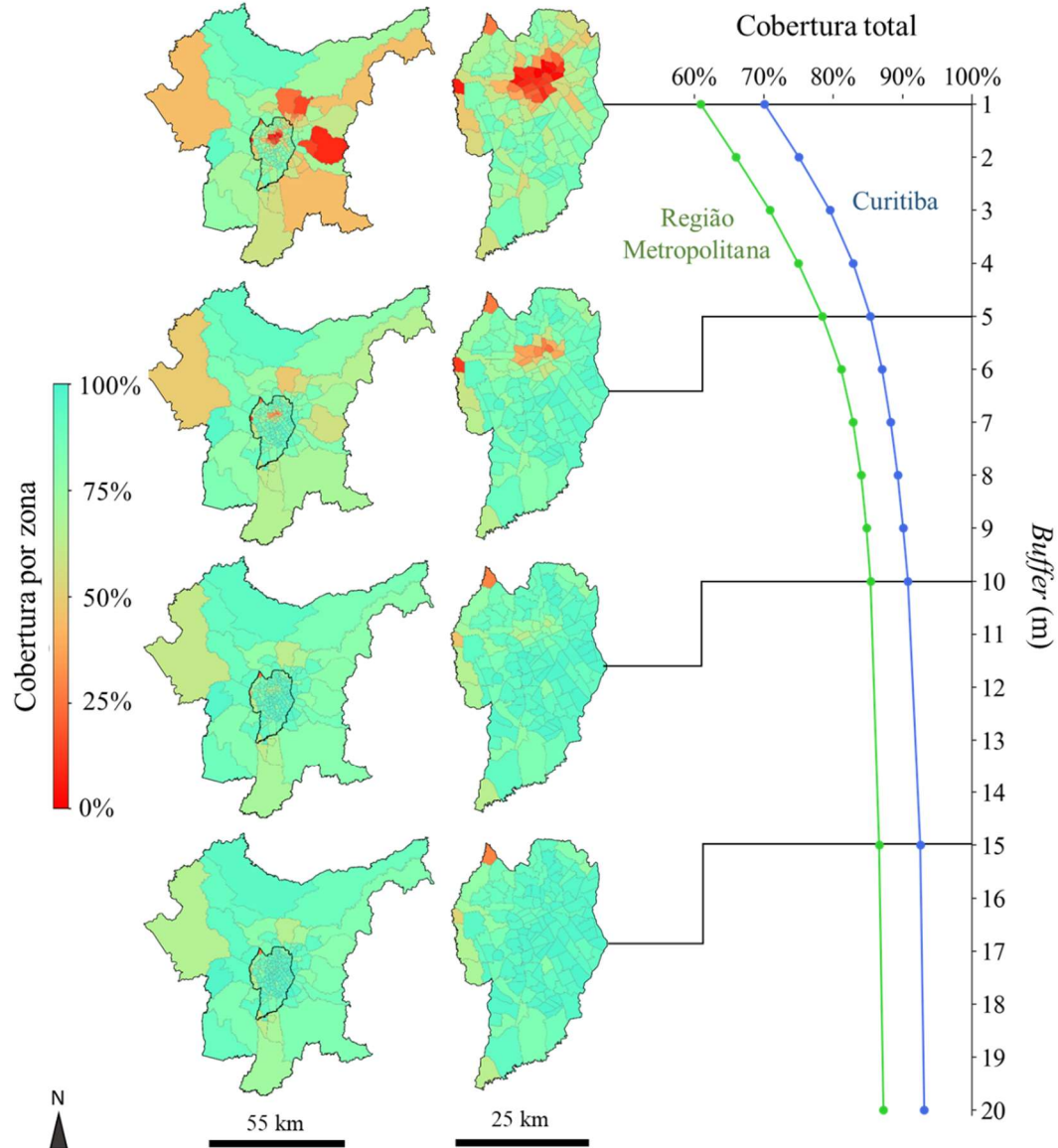


Fonte: O autor (2020).

É possível notar que, no geral, a proporção está próxima de 1, em especial nas ZT de Curitiba. Duas zonas chamam a atenção por serem os extremos calculados, ambas dentro do município de Curitiba – ao sul em laranja e ao noroeste em roxo. As duas são impactadas pelo número reduzido de vias que apresentam, ou seja, as proporções calculadas nelas são mais sensíveis às diferenças entre as redes. Algumas zonas com maiores áreas fora da cidade têm proporções maiores, mostrando que algumas vias rurais estão mais bem representadas pela rede do *OSM*. A completude, assim, é satisfatória.

No quesito acurácia posicional, a Figura 36 apresenta a porcentagem das arestas da rede que é abrangida pelo *buffer*, que varia de 1 a 20 m com mapas das proporções por zona para 1 m, 5 m, 10 m e 15 m de largura de *buffer*. Nota-se que a rede *OSM* do município de Curitiba tem uma maior porcentagem abrangida para todas as distâncias de *buffer*, com 70% da rede abrangida a 1 metro da rede oficial, enquanto que na rede da região metropolitana a proporção é de 60%. Na medida em que se aumenta o *buffer*, o acréscimo de elementos abrangidos decai, estabilizando-se próximo dos 90% para Curitiba e dos 85% para a região metropolitana.

Figura 36 – Acurácia posicional: porcentagem de elementos da rede *OSM* abrangidos segundo distância de *buffer*.



Fonte: O autor (2020).

Lembre-se que a existência de um número diferente de elementos, como visto anteriormente, afeta as proporções, o que pode explicar tanto o motivo de não se atingirem os 100% de cobertura após a estabilização, quanto a menor proporção de cobertura da região metropolitana, visto que a rede *OSM* tem mais elementos do que a rede obtida com a prefeitura. Haklay (2010) considera 1 metro de buffer como a sobreposição ideal para rodovias e entre 3,75 m e 5,6 m para vias urbanas da maior para a menor hierarquia. Isto posto, já em 5 metros de *buffer* aproximadamente 85% da rede de Curitiba e 75% da região metropolitana são abrangidas.

Os mapas constataam que o centro da cidade apresenta baixa abrangência dos elementos para *buffers* menores, mas com cobertura próxima de 75% para 10 m de largura. As zonas dos municípios ao norte de Curitiba e a sudoeste apresentam boas coberturas já para 1 m de largura, o que também ocorre para as zonas de Curitiba, exceto o centro. Note-se também que a zona a noroeste, que se mantém em laranja mesmo para *buffers* maiores, é a mesma que apresenta a máxima proporção entre a rede do *OSM* e a oficial, mostrando que parte da rede do *OSM* não contida nos *buffers* é de elementos inexistentes na rede oficial.

A consistência lógica, através da comparação entre as KDE das centralidades das duas redes, é observada apenas para o município de Curitiba, local de aplicação deste estudo. Os resultados das correlações entre as KDEs normalizadas são apresentados na Tabela 6 a seguir.

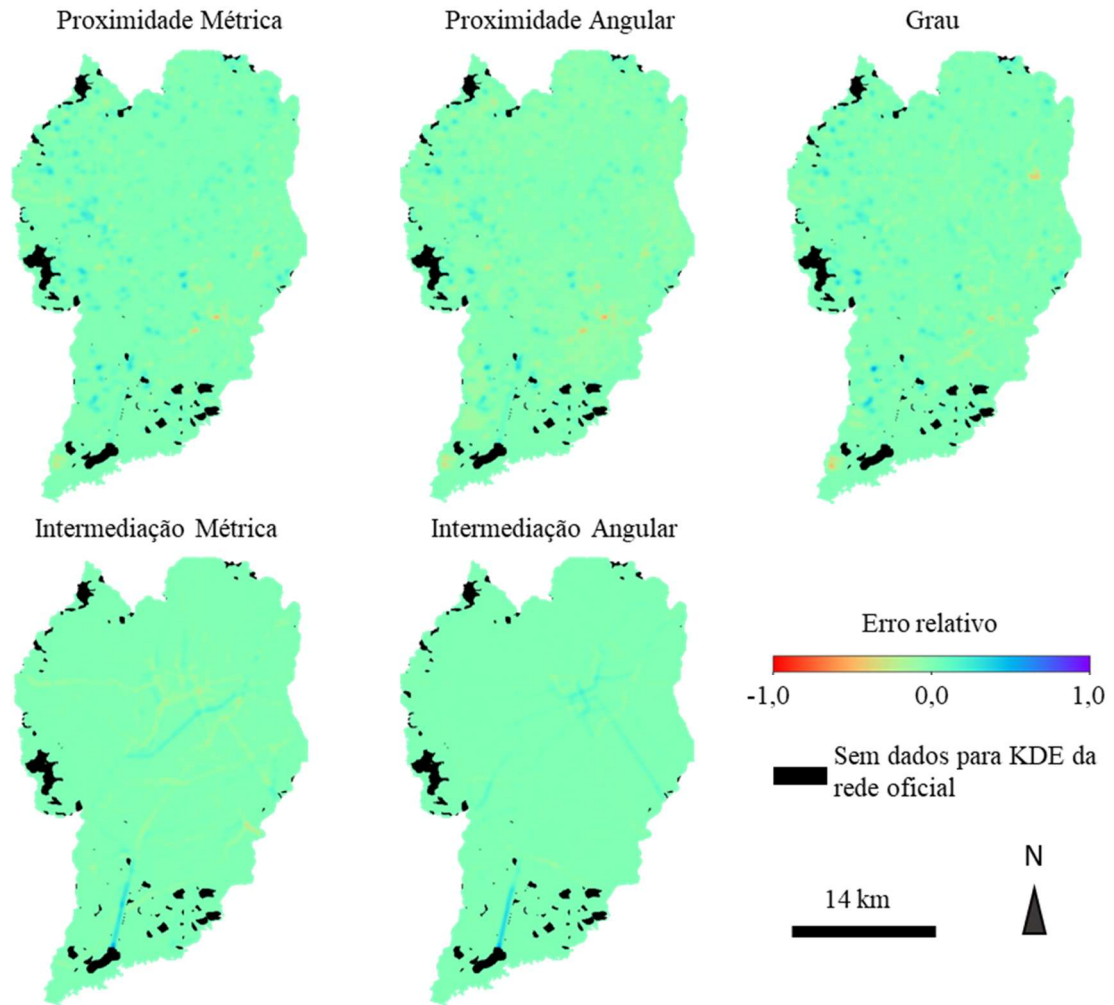
Tabela 6 – Correlações entre KDEs normalizadas das centralidades para ambas as redes.

KDE de Centralidade	Correlações	
	Pearson (r_p)	Spearman (r_s)
Proximidade métrica	0,989	0,985
Proximidade angular	0,989	0,985
Intermediação métrica	0,948	0,977
Intermediação angular	0,967	0,981
Grau	0,988	0,985

Todos os resultados são significativos para um nível de 99%.

Os mapas da Figura 37 demonstram os resultados dos erros relativos entre as superfícies das KDEs obtidas em cada centralidade. As regiões em preto representam locais onde há resultados para a rede do *OSM*, mas não para a rede oficial, ou seja, são locais com elementos não existentes na rede oficial.

Figura 37 – Erros relativos entre KDEs de centralidades para ambas as redes.



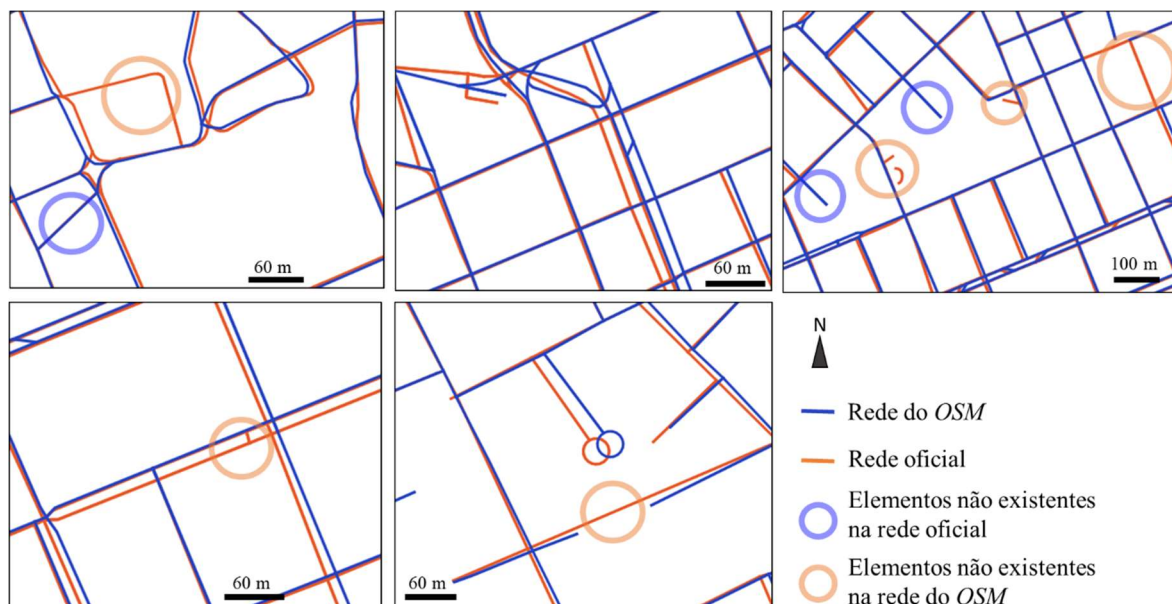
Fonte: O autor (2020).

Todos os erros estão próximos de zero, com valores mais altos e baixos em pontos isolados para as centralidades de proximidade e grau. Para as centralidades de intermediação, há segmentos com valores baixos quando a impedância métrica é considerada nos eixos estruturais da cidade, e altos tanto para a angular quanto para a métrica, mas em segmentos esporádicos. As diferenças observadas nos KDEs das centralidades de proximidade e de grau são muito semelhantes e pontuais – não definem segmentos de vias –, podendo ser explicadas por elementos existentes em apenas uma das redes. No geral, os resultados mostram boa consistência lógica entre as redes, observados pelos resultados das correlações na Tabela 6.

Considerando-se a completude, acurácia posicional e consistência lógica, a diferença mais visível é na região central, com baixas coberturas dos *buffers*, inferiores a 10 metros, indicando pior acurácia posicional em comparação com o restante da cidade. Além disso, os erros pontuais

citados indicam elementos não análogos entre as redes. Nesse sentido, a Figura 38 apresenta algumas destas inconsistências observadas.

Figura 38 – Sobreposição visual das redes com apontamento de algumas incoerências.



Fonte: O autor (2020).

Todas as imagens foram retiradas da região central da cidade. Em muitos locais há elementos na rede oficial que não existem para a rede do *OSM*. Também ocorre o contrário, mas com menos frequência. É perceptível também a distância entre os elementos, atestada pela análise dos *buffers* na acurácia posicional. As verificações utilizadas foram suficientes para caracterizar as incoerências apontadas.

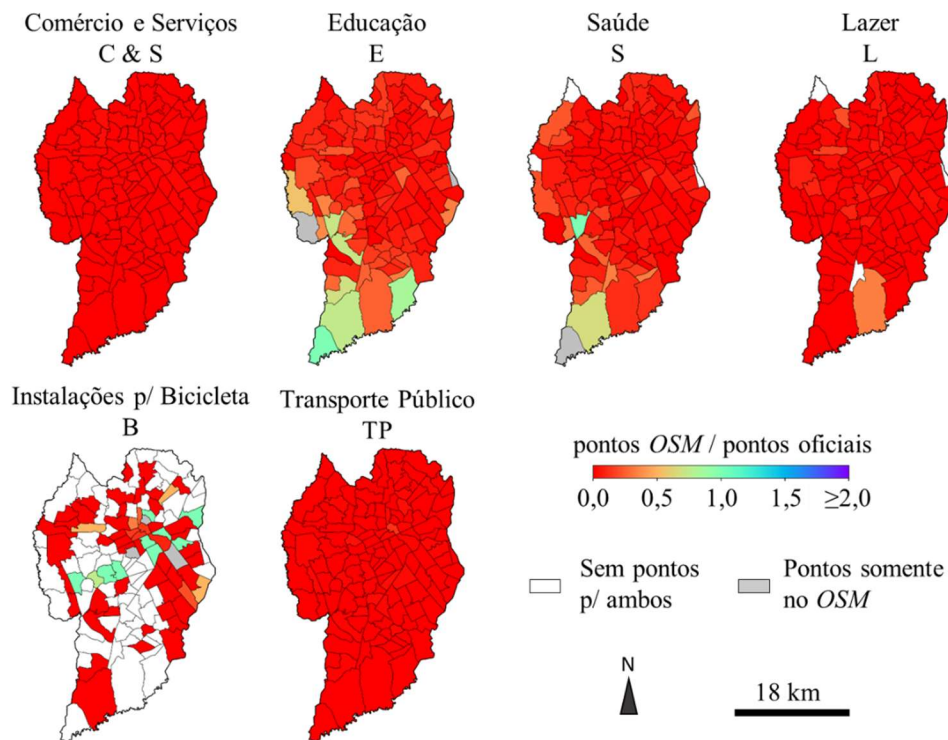
A falta ou excesso de elementos e inconformidades posicionais não significam falhas somente da rede do *OSM*, uma vez que a rede oficial também pode apresentar incoerências. Vale lembrar que a comparação busca quantificar as diferenças dos elementos do *OSM* para com os dados utilizados pela autoridade local, não sendo possível atestar a verdade fundamental para os dados. Estas inconsistências apenas devem fomentar cuidados em aplicações para condições mais locais de análise. Para a escala de análise global, como a deste trabalho, as diferenças não causam impacto significativo e, portanto, a qualidade da rede do *OSM* é satisfatória.

5.3.2 Resultados da avaliação dos Pontos de Interesse

5.3.2.1 Completude

A seguir, a Figura 39 apresenta os resultados por zona da completude dos *POIs* relativos aos dados dos alvarás e aos bancos de dados de paraciclos e pontos de ônibus. As classes são comparadas segundo a Tabela 5 definida anteriormente.

Figura 39 – Relação entre o número de *POIs* e o número de pontos oficiais em cada ZT.



Fonte: O autor (2020).

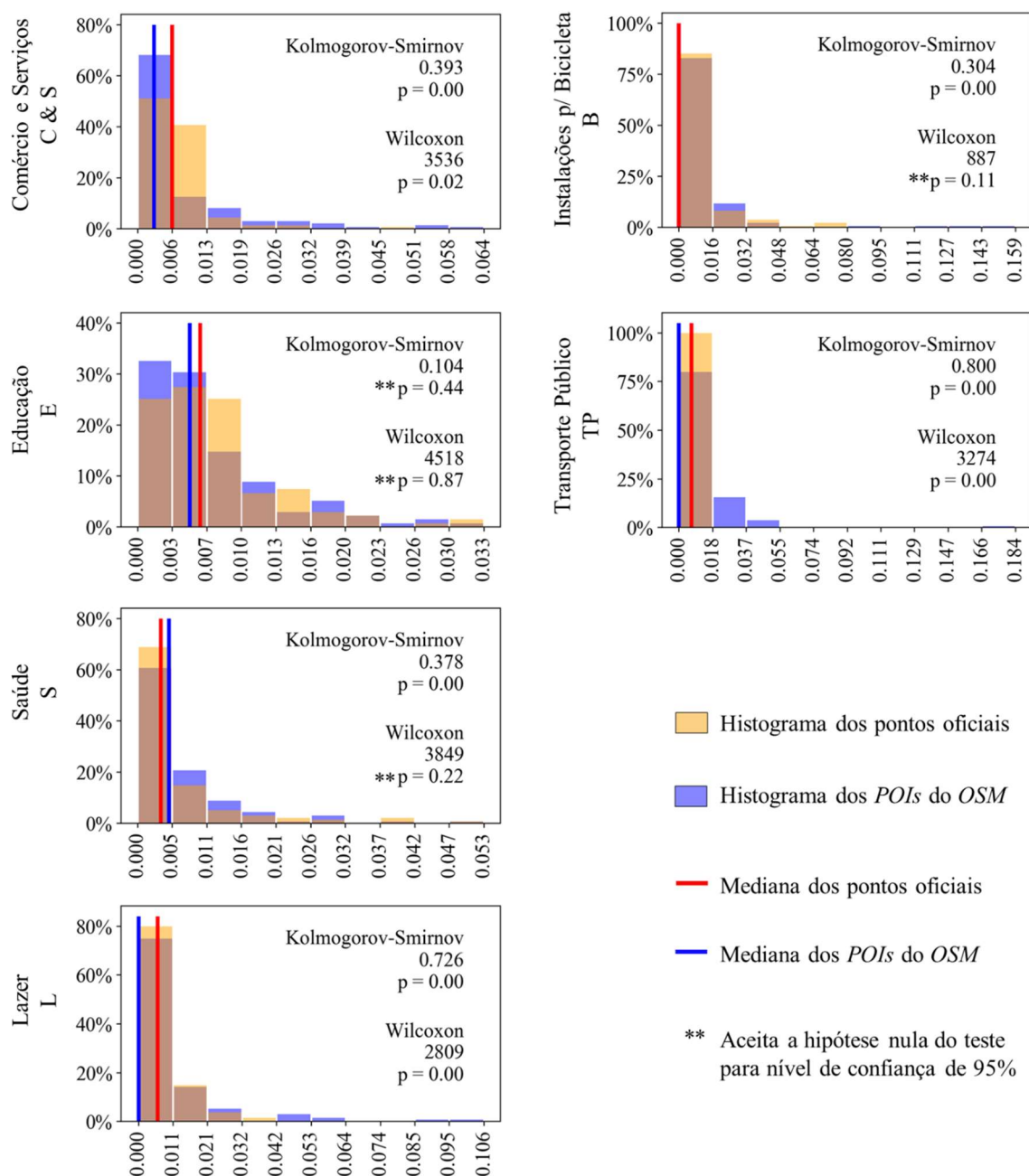
Nota-se que o número total de *POIs* de cada classe é muito inferior ao número dos pontos oficiais, levando a uma completude ruim, com poucas zonas próximas da proporção igual a 1 para Educação (E), Saúde (S) e Instalações para bicicleta (B), mas com a maioria das ZT com proporções próximas de 0. Este resultado é esperado devido à diferença constatada no número de pontos em cada banco de dados.

5.3.2.2 Acurácia Posicional

Como mencionado anteriormente, a acurácia posicional para a presente aplicação é observada de acordo com a proporção de pontos em cada ZT, ou seja, deve-se observar se a proporção dos *POIs* de certa classe nas zonas é próxima da proporção de pontos oficiais de mesma classe nesta

zona. Assim, a Figura 40 apresenta os histogramas e resultados dos testes de hipótese de similaridade das distribuições – Kolmogorov-Smirnov – e de similaridade dos valores – Wilcoxon. Novamente, as classes são comparadas segundo a Tabela 5.

Figura 40–Histogramas e testes de hipótese para proporções de *POIs* e pontos oficiais nas ZT.

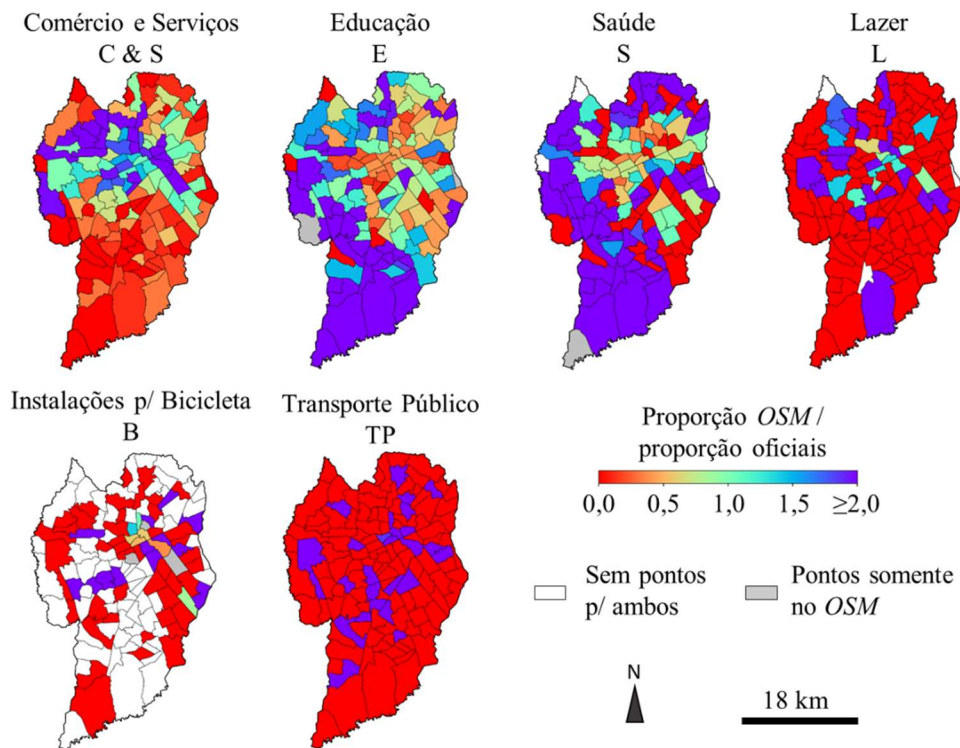


Fonte: O autor (2020).

Nota-se que as proporções seguem valores similares para E, S e B, verificadas pelo teste de Wilcoxon, com valores-*p* iguais a 0,87, 0,22 e 0,11 respectivamente. Porém, apenas a classe E

aceita a hipótese nula de semelhança entre as distribuições estatísticas, verificada pelo teste de Kolmogorov-Smirnov com um valor- p igual a 0,44. Não obstante, os mapas da Figura 41 demonstram que nenhuma classe tem a distribuição no espaço fiel aos dados oficiais. Enquanto Educação (E) e Saúde (S) demonstram proporções inferiores em zonas centrais e ao longo dos eixos estruturais, com aumento gradativo para zonas mais afastadas, a classe Comércio e Serviços (C & S) tem comportamento oposto, com *POIs* mais concentrados na região central, quando comparados com os dados dos alvarás. As classes Lazer (L), Instalações para bicicleta (B) e Transporte Público (TP) têm concentração mais dispersa, mas também diferente dos dados oficiais.

Figura 41 – Relação entre a proporção de *POIs* e a proporção de pontos oficiais em cada ZT.



5.4 CONSIDERAÇÕES FINAIS

Os resultados são próximos aos encontrados na literatura. Enquanto Hakley (2010) encontrou uma média de 6 metros entre a rede obtida com a agência local e a obtida com o *OSM*, neste trabalho, mais de 80% da rede *OSM* de Curitiba e 70% da rede da região conurbada estão a 5 metros da rede oficial, corroborando que a cidade mais central comumente apresenta melhor acurácia posicional (GIRRES; TOUYA, 2010; NEIS; ZIELSTRA; ZIPF, 2012). No entanto, no presente trabalho a rede do *OSM* para a região externa a Curitiba tem um comprimento 12%

maior do que a rede oficial, representando uma melhor completude, enquanto que internamente a Curitiba, a rede do *OSM* tem comprimento 2% maior do que a oficial. Além disso, a região mais central de Curitiba também apresentou erros posicionais maiores do que outras zonas no próprio município. Contudo, mesmo indo de encontro com o esperado, estes erros na acurácia posicional não são preocupantes quando observados os erros relativos das KDEs originadas com as centralidades.

Os *POIs*, apesar de não serem verificáveis quanto à sua posição exata, podem ser comparados de forma agregada pela semântica das classes. De forma geral, a completude é deficitária, não permitindo que os dados sejam comparáveis numericamente aos dados oficiais. Ao se observar a proporção de pontos de cada classe nas zonas, nota-se que pontos relativos à classe Educação (E) têm distribuição estatística e valores similares aos alvarás comerciais, mas a acurácia posicional não é coerente para todas as ZT, com maiores concentrações de *POIs* longe do centro em relação às concentrações de alvarás. Ademais, nenhuma classe demonstra acurácia posicional satisfatória nas ZTs, apenas valores similares das proporções para classes de Saúde (S) e Instalações para bicicletas (B), além de Educação (E) como já mencionado, atestado pelo teste de Wilcoxon com valores-*p* maiores do que 0,10.

No geral, talvez os *POIs* sejam interessantes indicadores de tendências urbanas no tempo e até de uso do solo, como observado por Zhang e Pfoser (2019), mas, para estudos mais dependentes de acurácia posicional e completude, as deficiências podem reduzir a confiabilidade de modelos preditivos, limitando-os a modelos explicativos com aplicação operacional. Para o presente estudo, os *POIs* talvez demonstrem coerência na explicação das viagens, mas não podem indicar o real papel de cada tipo de atividade ou ponto nas explicações. Já a rede demonstra boa completude, acurácia posicional e consistências lógicas, permitindo que as informações adquiridas do *OSM* sejam equiparáveis ao banco de dados oficial, o que permite que seja utilizada para análises globais mais detalhadas e completas sobre sua configuração, substituindo redes oficiais, como feito por muitos estudos (BOEING, 2018; KURAMOTO; CUNHA, 2019; MORELLI; CUNHA, 2019), mas com cautela para análises locais mais sensíveis a inconsistências no número e na disposição dos elementos.

As centralidades, portanto, se mostraram as informações mais confiáveis a serem adquiridas do *OSM*, com a ressalva para possíveis diferenças locais. Assim, espera-se que as centralidades apresentem comportamentos similares na aplicação para modelos explicativos de atração de

viagens. Resta observar se, através das centralidades, a rede pode complementar as possíveis falhas observadas nos *POIs* para modelos calibrados somente com dados do *OSM*, visto que as centralidades apresentam relação com a posição das atividades econômicas, muito vinculadas à atração de viagens – como demonstrado no Capítulo 3.

6 COMPARAÇÃO DOS MODELOS DE ATRAÇÃO DE VIAGENS

Neste capítulo, é realizada a terceira e última etapa do trabalho, em que os modelos explicativos de atração de viagens são calibrados com dados do *OSM* e com dados oficiais. Desta forma, comparam-se os modelos obtidos com variáveis de ambas as fontes de dados, tanto ignorando-se e quanto considerando-se as variáveis de centralidade de rede, o que possibilita compreender como as centralidades de rede podem complementar a explicação das viagens atraídas e avaliar a usabilidade do *OSM* para o planejamento de transportes em relação à atração de viagens.

6.1 CONTEXTUALIZAÇÃO

A localização das atividades econômicas, ou atividades-fim, define os destinos mais comuns das viagens urbanas. Não obstante, a condição de acessibilidade dos destinos pode também impactar na preferência de certos locais ante outros (THILL; KIM, 2005; ORTÚZAR; WILLUMSEN, 2011), o que pode ser difícil de averiguar, visto que o posicionamento das atividades busca também levar em conta sua acessibilidade. Ou seja, duas informações que explicam a atração de viagens – local das atividades e acessibilidade – são relacionadas entre si (VICKERMAN, 1974; EWING; DEANNA; LI, 1996). Todavia, além da acessibilidade, muitas atividades buscam se beneficiar de locais com maior presença de pessoas para seu posicionamento, isto é, sua localização também é produto dos fluxos ao longo da rede viária, especialmente para atividades cotidianas (JACOBS, 1961).

As relações de movimento, acessibilidade e posição das atividades econômicas são a base para a premissa da “economia do movimento” da Sintaxe Espacial (HILLIER, 1996), cuja ferramenta para mensuração das dinâmicas nas redes viárias são as centralidades. De fato, há correlações observadas entre centralidades e movimentos ao longo das redes (HILLIER; PENN, 1996; TURNER, 2007; JIANG; LIU, 2009; JIANG; JIA, 2011; GAO *et al.*, 2013), e com o uso do solo e atividades econômicas (PORTA *et al.*, 2012; RUI; BAN, 2014; LIU *et al.*, 2016; OMER; GOLDBLATT, 2016), como também foi demonstrado no Capítulo 4 deste trabalho. Assim, as centralidades podem auxiliar na explicação da atração de viagens por propiciarem quantitativamente a relação estabelecida pela Sintaxe Espacial entre proximidade, intermediação e até mesmo conectividade com as atividades-fim das viagens. Jayasinghe, Sano e Rattanaporn (2017) buscaram substituir variáveis tradicionais por centralidades para modelos de atração de viagens para a cidade de Colombo, no Sri-Lanka, obtendo resultados satisfatórios,

mas argumentando que as centralidades podem ser usadas como variáveis explicativas principalmente em locais com pouca disponibilidade de dados.

As centralidades, portanto, têm relação com as atividades econômicas, corroborando as premissas da Sintaxe Espacial de mútua influência entre a infraestrutura de transporte e o uso do solo. Assim, a rede de transporte se torna um insumo para análises de atração de viagens não apenas como a oferta de transporte, mas como o reflexo dos próprios atrativos, tanto pela acessibilidade que representam, especialmente a centralidade de proximidade, quanto pela intermediação e “economia do movimento”, que Hillier (1996) aponta como indutora de dinâmicas urbanas. As centralidades podem desempenhar um papel importante especialmente quando há falta de dados, ou dados insuficientes. Os mapas colaborativos, portanto, ao trazerem informações de *POIs* e redes viárias, contemplam dados que tradicionalmente são usados para explicar a atração de viagens, e fornecem os dados para a obtenção das centralidades de rede. A crescente qualidade dos mapas do *OSM*, como averiguada anteriormente para o caso da rede viária de Curitiba, possibilita a aplicação destes dados de maneira mais ampla e irrestrita.

Para tanto, é comum a utilização de modelos de regressão linear na explicação da atração de viagens. Para além de previsores, os modelos são ferramentas para explorar as relações entre variáveis explicativas e o fenômeno explicado. O método de regressão mais utilizado é dos mínimos quadrados ordinários (MQO), que busca minimizar a soma dos quadrados dos resíduos entre as estimativas e os valores reais. Contudo, estes modelos devem ser utilizados com cautela para fenômenos espaciais, uma vez que esta dimensão não é abrangida em sua estrutura (WANG *et al.*, 2012). Apesar de modelos como o Espacial Auto Regressivo (EAR) e o Condicional Auto Regressivo (CAR) considerarem a espacialidade das variáveis, Lopes, Brondino e Rodrigues da Silva (2014) demonstram que modelos MQO podem ter resultados equiparados, ou até melhores, caso considerem variáveis espaciais que traduzam dependências locais ou globais.

Os modelos que observam as atrações agregadas por regiões têm sua abordagem baseada nas viagens, ou seja, observam o número total de viagens atraído nas zonas e traçam correlações com as variáveis explicativas. Já modelos que buscam replicar a base teórica das escolhas dos viajantes, com maior fidelidade aos preceitos de indução da demanda por viagens, têm sua abordagem baseada nas atividades, ou em jornadas (KITAMURA, 1988; INNES; IRCHA;

BADDOE, 1990; GOULIAS; PENDYALA; KITAMURA, 1991). Estes modelos visam a quantificar o potencial de escolha dos viajantes, adicionando a percepção da minimização de custos para os deslocamentos, possibilitando maior desagregação. Contudo, apresentam maior complexidade, sendo mais propícios em aplicações de simulação e com abundância de dados (MCNALLY; RINDT, 2008). Chang *et al.* (2014) demonstraram que modelos baseados nas viagens trazem resultados agregados semelhantes aos com abordagem mais baseada nas atividades, e de forma menos onerosa. Além disso, a abordagem tradicional tem historicamente auxiliado planejadores e autoridades locais em tomadas de decisões (MCNALLY; RINDT, 2008; ORTÚZAR; WILLUMSEN, 2011), sendo útil para aplicações de novas variáveis, como é o caso das centralidades.

Neste sentido, Jayasinghe, Sano e Rattanaorn (2017) consideraram o vínculo entre as atividades econômicas e a rede, à luz da Sintaxe Espacial, para criar modelos MQO de atração de viagens para a cidade de Colombo, no Sri-Lanka, somente com variáveis explicativas de centralidades da rede. Os modelos calibrados relacionaram a densidade de viagens atraídas – tendo, portanto, abordagem baseada nas viagens de forma agregada – e a média das KDE resultantes em cada ZT, obtendo ajustes com R^2 maiores do que 0,70. O estudo, no entanto, apresenta alguns modelos com alta multicolinearidade, não controlando a dependência espacial ou outras premissas da regressão, como normalidade dos erros e homoscedasticidade, apesar de garantirem a linearidade das variáveis ao aplicarem a transformação exponencial.

Observar somente as centralidades como variáveis explicativas pode significar substituir variáveis tradicionais, uma vez que há relação entre elas, mas também pode significar impor novas informações. Modelos que considerem ambos os tipos de variáveis podem trazer esclarecimentos neste sentido, portanto, o método proposto busca calibrar: (i) modelos somente com as variáveis tradicionais; (ii) somente com variáveis de centralidades – semelhante ao estudo de Jayasinghe, Sano e Rattanaorn (2017) –; e (iii) agregando os dois tipos de variáveis. Isto possibilita entender se a falta de certas informações, como a atestada por falhas dos dados de *POIs*, pode ou não ser complementadas pelas centralidades.

6.2 MÉTODO

6.2.1 Calibração dos modelos explicativos

Com os dados de *POIs* do *OSM* e dos bancos de dados oficiais, definidos anteriormente na Tabela 5, é possível obter os modelos multilineares com variáveis tradicionais de atividades de Comércio e Serviços (C & S), Educação (E), Saúde (S) e Lazer (L), e de acessibilidade ao Transporte Público (TP) e à Bicicleta (B) – o número de pontos por zona destas variáveis é mostrado no Apêndice 4. Outras classificações de locais não são utilizadas, pois algumas classes não apresentam dados equivalentes em um dos dois bancos de dados – oficial ou *OSM*. Os modelos, portanto, se limitam à explicação de viagens atraídas a partir destas variáveis. A limitação não impacta o escopo do estudo, que busca explorar as semelhanças e diferenças dos dados em relação à atração de viagens, e não encontrar modelos preditivos.

Outros dois modelos são calibrados somente com variáveis explicativas das centralidades da rede, um para os resultados da rede oficial e outro para a rede obtida do *OSM*. No entanto, diferentemente de Jayasinghe, Sano e Rattanaporn (2017), não se considera a densidade de atração por zona, e sim o total de viagens atraídas. E para se agregarem os resultados de centralidade por ZT, além da média (\bar{X}) das KDE, são considerados outros três parâmetros estatísticos zonais: mediana (M_d), terceiro quartil (Q3) e 85º percentil (P85). Os mesmos parâmetros são utilizados nas análises zonais do Capítulo 3, visto que a média pode não ser a melhor representação dos valores em cada ZT. Os valores estatísticos de cada centralidade são apresentados no Apêndice 2 para a rede oficial e no Apêndice 3 para a rede do *OSM*. Assim, a Tabela 7 demonstra todas as variáveis aplicadas para os diferentes modelos, com as respectivas abreviações utilizadas.

Tabela 7 – Lista de variáveis tradicionais e de centralidades a serem aplicadas nas regressões.

Variáveis tradicionais	Variáveis de centralidade					
C & S (Comércio e Serviços)	Intermediação		Proximidade		Grau	Unitária
E (Educação)	Métrica	Angular	Métrica	Angular		
S (Saúde)	$C_B(m) \bar{X}$	$C_B(a) \bar{X}$	$C_C(m) \bar{X}$	$C_C(a) \bar{X}$	$C_D \bar{X}$	Unit. \bar{X}
L (Lazer)	$C_B(m) M_d$	$C_B(a) M_d$	$C_C(m) M_d$	$C_C(a) M_d$	$C_D M_d$	Unit. M_d
TP (Transporte Público)	$C_B(m) Q3$	$C_B(a) Q3$	$C_C(m) Q3$	$C_C(a) Q3$	$C_D Q3$	Unit. Q3
B (Instalações p/ Bicicletas)	$C_B(m) P85$	$C_B(a) P85$	$C_C(m) P85$	$C_C(a) P85$	$C_D P85$	Unit. P85

No total, são seis variáveis tradicionais e 24 variáveis de centralidades, inclusive a variável controle unitária, o que permite verificar se os modelos escolhem os resultados que consideram

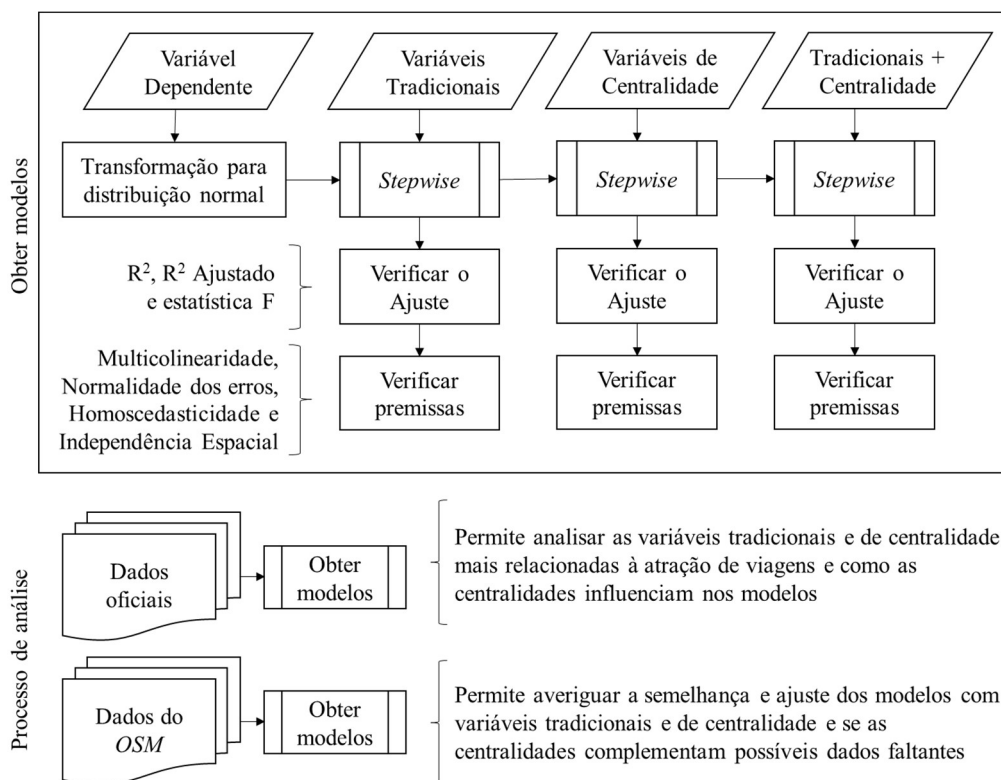
os efeitos das centralidades ou apenas a densidade da rede. Todas as variáveis, dependentes e independentes, são divididas pelo valor máximo de cada uma, normalizando os valores entre 0 e 1, porém o zero representa a ausência de dados, não sendo definido pelo valor mínimo. Isto facilita a análise dos parâmetros de calibração. Após a obtenção dos modelos apenas com variáveis tradicionais, e outros apenas com variáveis de centralidades, calibra-se um terceiro conjunto de modelos, desta vez considerando-se os dois grupos de variáveis (tradicionais e centralidades), tanto com os dados oficiais quanto com os dados do *OSM*. Assim, é possível observar como as centralidades podem complementar o poder explicativo dos modelos, e se elas podem substituir variáveis tradicionais.

O método de escolha das variáveis de cada grupo é feito com regressão *Stepwise*, um processo iterativo de adição e subtração das variáveis independentes que melhor explicam a variável dependente. O processo adiciona as variáveis que apresentam coeficientes significativamente diferentes de zero, através do valor da estatística *t* mais significativa, ou seja, com menor valor-*p*, e que representam o “poder explicativo” da variável. Após cada adição, todas as variáveis são novamente observadas e aquelas com os maiores valores-*p* – menor poder explicativo – são retiradas (KUTNER *et al.*, 2004). As variáveis escolhidas compõem os modelos finais para análise.

Um modelo MQO é generalizável na medida em que garante os pressupostos de resíduos com variância constante – homoscedasticidade –, independência e distribuição normal com média igual a zero. Estes fenômenos refletem a linearidade do fenômeno e a distribuição normal da variável dependente, quesitos importantes para a regressão linear múltipla (KUTNER *et al.*, 2004). Além disso, deve-se garantir a não multicolinearidade entre as variáveis explicativas, quesito sensível em análises de acessibilidade em modelos de demanda (EWING; DEANNA; LI, 1996), e a independência espacial, pois trata-se de um fenômeno com forte relação no espaço (FOTHERINGHAN; BRUNSDON; CHARLTON, 2000).

A Figura 42 esquematiza a obtenção dos modelos. A transformação da variável dependente e os testes de hipótese das premissas indicadas são descritos posteriormente.

Figura 42 – Processo de obtenção e análise dos modelos de regressão MQO.



Fonte: O autor (2020).

Ao se obterem os modelos com dados oficiais, ou seja, dados com validação prática da prefeitura, é possível entender quais variáveis mais se relacionam com a atração de viagens. Considerando-se os resultados de usabilidade dos *POIs*, espera-se que os modelos tradicionais com estes dados não apresentem ajustes semelhantes aos obtidos com dados oficiais. Contudo, a rede obtida com o *OSM* apresenta bons resultados de usabilidade, portanto, é esperado que modelos calibrados apenas com centralidades tenham resultados similares para ambos os bancos de dados. Os modelos com dados oficiais, e do *OSM* com os dois tipos de variáveis (tradicional e centralidades), devem ser analisados de formas distintas. Com dados oficiais é possível entender como as centralidades influenciam os resultados, melhorando-os ou trazendo pouca mudança. Já com dados do *OSM*, pode-se observar se as centralidades atuam como complemento aos dados, considerando a possível falta de informações dos *POIs*.

O primeiro passo para possibilitar a calibração dos modelos é a transformação da variável dependente a fim de tornar sua distribuição próxima da normal. Para tanto, utiliza-se a transformação proposta por Box e Cox (1964), que busca um expoente λ a ser aplicado aos valores da

variável para encontrar $y' = y^\lambda$ cuja distribuição é normal. Para tanto, o procedimento considera a função

$$y(\lambda) = \begin{cases} \frac{y^\lambda - 1}{\lambda} & \text{se } \lambda \neq 0 \\ \ln y & \text{se } \lambda = 0 \end{cases} \quad [5.1]$$

Em que,

$y(\lambda)$ é a variável transformada;

y é a variável original;

λ é o expoente aplicado para a transformação

O expoente λ da transformação Box-Cox é encontrado ao se considerar que $y(\lambda)$ segue uma distribuição gaussiana. Assim, maximiza-se a verossimilhança logarítmica de $y(\lambda)$ variando λ , isto é, encontra-se um λ tal que $y(\lambda)$ apresente a melhor aderência à curva gaussiana com média e desvio padrão resultantes de $y(\lambda)$. Esta etapa é feita com a ajuda da biblioteca SciPy⁹ para a linguagem Python de programação.

Com a transformação e posterior obtenção dos modelos através do já mencionado método *Stepwise*, é preciso prosseguir com as verificações de multicolinearidade, normalidade dos erros, homoscedasticidade e independência espacial.

6.2.2 Verificação de premissas de regressão

A multicolinearidade é averiguada com o Fator de Inflação da Variância (FIV), que mede o quanto uma variável explicativa pode ser explicada pelas outras (KUTNER *et al.*, 2004). O FIV pode constatar formas mais sutis de colinearidade, já que não observa apenas a correlação entre duas variáveis, mas entre compostos. Então, para cada variável X_i das n variáveis explicativas, o FIV será

$$FIV(X_i) = \frac{1}{1 - R^2} \quad [5.2]$$

Em que, dado $k, k+1, \dots, n-1 \neq i$, o R^2 é o coeficiente de determinação resultante do modelo

$$X_i = b_0 + b_k X_k + b_{k+1} X_{k+1} + \dots + b_{n-1} X_{n-1} \quad [5.3]$$

⁹ docs.scipy.org

Assim, quanto mais uma variável independente é explicada pelas outras, maior é o valor de FIV. Quanto menor o valor, menos colinear é a variável em questão em relação ao conjunto das demais. Um valor limite para o FIV comumente adotado é igual a 10 (KUTNER *et al.*, 2004; BELSLEY; KUH; WELSCH, 2004).

A normalidade dos erros, por sua vez, é verificada pelo teste Omnibus de D'Agostino e Pearson (1973). Este teste combina os resultados dos testes para curtose e assimetria. Quanto maior o valor da estatística, mais distantes os dados se encontram da distribuição normal. Para o teste de hipótese, o valor- p fornecido vem da probabilidade qui-quadrado (χ^2) entre os valores observados e os esperados em uma distribuição normal, aceitando-se a hipótese nula se $p > 0,05$. Já a heteroscedasticidade é verificada pelo teste de White (1980), que segue a distribuição χ^2 para obter seu valor- p . O teste obtém um novo modelo que busca estimar os resíduos (ε) resultantes ao quadrado a partir da fórmula

$$\varepsilon^2 = \delta_0 + \delta_1 \hat{Y} + \delta_2 \hat{Y}^2 \quad [5.4]$$

Em que

\hat{Y} é o conjunto de variáveis explicativas do modelo original;

δ_i são os coeficientes de calibração

O valor encontrado ao se multiplicar o número de observações pelo R^2 resultante é comparado ao valor χ^2 de mesmo grau de liberdade, visando a reter ou a rejeitar a hipótese nula de que a variância dos erros é constante, ou seja, homoscedástica, para um nível de confiança de 95%.

Os testes FIV, Omnibus e White são obtidos também com a biblioteca SciPy. O teste para se verificar a independência espacial, por sua vez, é feito com o auxílio da biblioteca PySAL¹⁰, também em linguagem Python, utilizada para análises espaciais. O teste aplicado é o índice de Moran, ou I de Moran (MORAN, 1950), que verifica a independência espacial global dos resíduos gerados pelos modelos calibrados ao calcular

$$I = \frac{N \sum_i \sum_j \omega_{ij} (x_i - \bar{x})(x_j - \bar{x})}{W \sum_i (x_i - \bar{x})^2} \quad [5.5]$$

Em que

N é o número de elementos espaciais; listados por i e j desde que $i \neq j$;

¹⁰ pysal.org

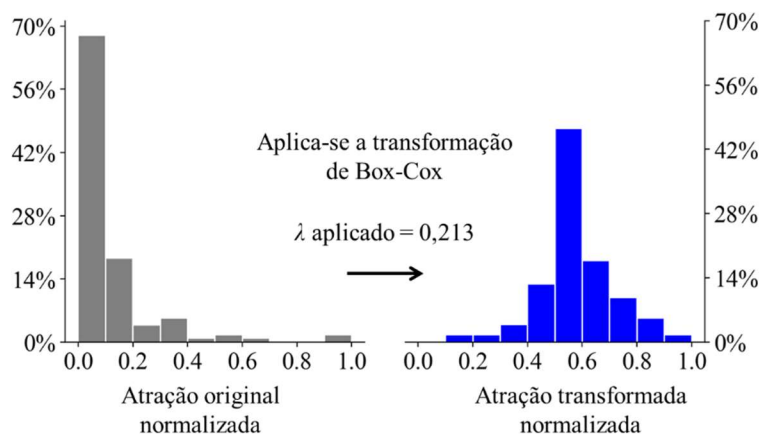
ω_{ij} é o peso entre os elementos i e j : igual a 1 se forem vizinhos e 0 caso contrário;
 x é o valor da variável, neste caso, os resíduos, em cada elemento, sendo \bar{x} a média;
 W é a soma de todos ω_{ij} .

São consideradas vizinhas as ZTs que apresentem ao menos um ponto em comum. Assim, os valores dos resíduos de cada modelo nas ZTs foram comparados a uma distribuição de valores I obtidos com 100 simulações, em que os mesmos resíduos são randomicamente distribuídos nas ZT, ou seja, obtém-se uma distribuição com 100 valores I de Moran com disposições randômicas dos resíduos, e verifica-se onde o valor real de I se encontra nesta distribuição, obtendo-se um pseudo-valor p para o teste de hipótese. A hipótese nula de independência espacial é rejeitada caso $p < 0,05$.

6.3 RESULTADOS

Antes de se obterem os modelos, faz-se a transformação das variáveis dependentes de atração de viagem. A Figura 43 exemplifica a transformação de Box-Cox para a variável dependente de atração de todas as viagens, demonstrando-se como a distribuição é modificada.

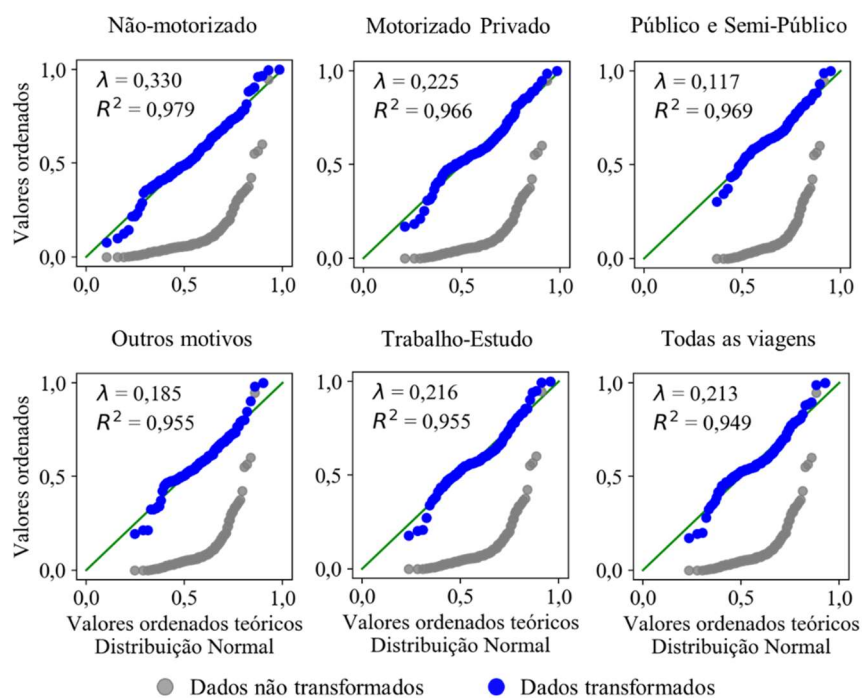
Figura 43 – Exemplo de transformação para a variável Atração de Todas as Viagens.



Fonte: O autor (2020).

A Figura 44 apresenta os gráficos Q-Q com os valores originais e após a transformação Box-Cox ao se aplicarem os valores λ apresentados. Os valores ordenados são comparados aos valores teóricos para uma distribuição normal de mesma média e desvio padrão, obtendo-se o coeficiente de determinação (R^2) para a comparação entre o resultado da transformação e a distribuição ideal.

Figura 44 – Gráficos Q-Q para as variáveis dependentes antes e após a transformação Box-Cox com ajuste à distribuição normal.



Fonte: O autor (2020).

Nota-se que a distribuição original de todas as variáveis dependentes é similar, o que também explica os valores próximos dos coeficientes de λ transformação encontrados. A Figura 45, por conseguinte, apresenta a correlação (r) de Pearson entre cada variável tradicional e as atrações de viagens, tanto para os dados oficiais quanto para os *POIs* do *OSM*.

Figura 45 – Correlações de Pearson entre variáveis dependentes de atração e variáveis explicativas tradicionais.

Atração de Viagens	Bancos de dados oficiais						Pontos de Interesse do <i>OSM</i>					
	C & S	L	E	TP	S	B	C & S	L	E	TP	S	B
	Não-motorizado	0,55 (0,00)	0,37 (0,00)	0,31 (0,00)	0,58 (0,00)	0,28 (0,00)	0,41 (0,00)	0,25 (0,00)	0,26 (0,00)	0,42 (0,00)	0,32 (0,00)	0,34 (0,00)
Motorizado Privado	0,74 (0,00)	0,72 (0,00)	0,69 (0,00)	0,51 (0,00)	0,58 (0,00)	0,48 (0,00)	0,59 (0,00)	0,41 (0,00)	0,37 (0,00)	0,36 (0,00)	0,51 (0,00)	0,33 (0,00)
Público e Semi-Púb.	0,76 (0,00)	0,74 (0,00)	0,72 (0,00)	0,53 (0,00)	0,61 (0,00)	0,49 (0,00)	0,60 (0,00)	0,43 (0,00)	0,36 (0,00)	0,41 (0,00)	0,52 (0,00)	0,32 (0,00)
Outros motivos	0,80 (0,00)	0,74 (0,00)	0,72 (0,00)	0,48 (0,00)	0,63 (0,00)	0,50 (0,00)	0,60 (0,00)	0,39 (0,00)	0,30 (0,00)	0,41 (0,00)	0,51 (0,00)	0,31 (0,00)
Trabalho-Estudo	0,72 (0,00)	0,66 (0,00)	0,61 (0,00)	0,56 (0,00)	0,53 (0,00)	0,48 (0,00)	0,53 (0,00)	0,40 (0,00)	0,42 (0,00)	0,37 (0,00)	0,50 (0,00)	0,30 (0,00)
Todas as viagens	0,76 (0,00)	0,69 (0,00)	0,65 (0,00)	0,55 (0,00)	0,57 (0,00)	0,50 (0,00)	0,56 (0,00)	0,41 (0,00)	0,40 (0,00)	0,39 (0,00)	0,52 (0,00)	0,31 (0,00)

Valores- p de significância entre parênteses

Fonte: O autor (2020).

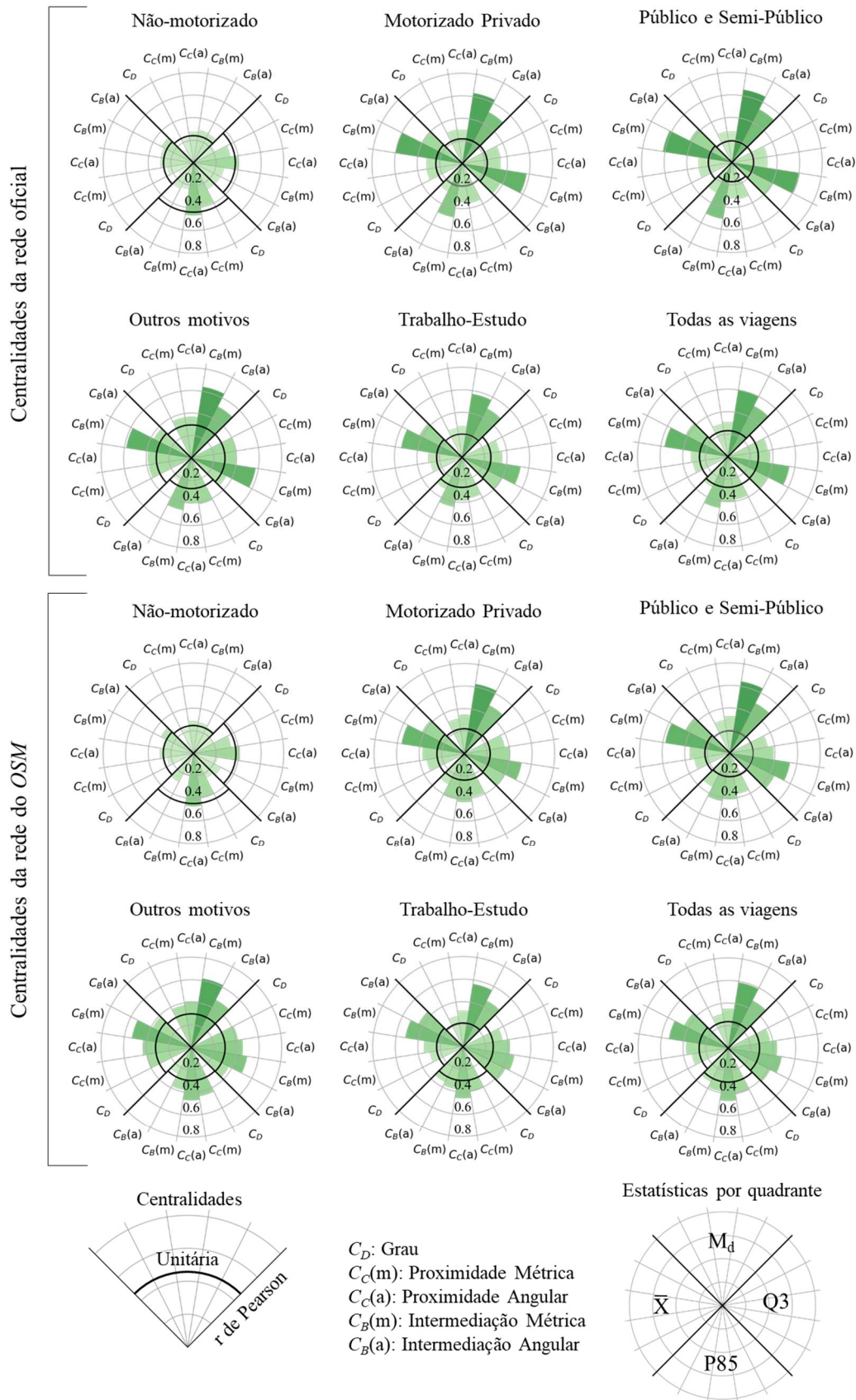
Para os dados oficiais, a classe Comércio e Serviços (C & S) apresenta as maiores correlações, seguida de Lazer (L), Educação (E) e Saúde (S). Os paraciclos, instalações para bicicletas (B) tem as piores correlações, mas é a terceira principal correlação para as viagens não-motorizadas, atrás de C & S e dos pontos de Transporte Público (TP). Mesmo não representando atrativos de viagens não-motorizadas, os pontos de ônibus geralmente se encontram em locais de atratividade, sendo uma variável indiretamente relacionada.

Para os *POIs*, as maiores correlações são para C & S, como nos dados oficiais, seguidas das correlações para Saúde (S) para todos os tipos de viagens, exceto as não-motorizadas. A classe Educação (E) é a mais correlacionada às não-motorizadas. Inclusive, os dados de *POIs* apresentam comportamento similar, mesmo que com valores de correlação menores, para C & S e S, mas, para as demais classes, o comportamento é distinto, levando à possibilidade de os modelos escolherem classes diferentes a depender do banco de dados utilizado.

As classes Educação (E) e Transporte Público (TP) chamam a atenção pelo comportamento nos dados oficiais. A primeira tem menor correlação com as viagens a trabalho ou estudo e a segunda se relaciona mais com viagens não-motorizadas do que com modos públicos ou semipúblicos, mesmo que com diferenças pequenas. Isto ocorre porque as classes de pontos podem trazer correlações indiretas. Como o conjunto de dados explicativo não é exaustivo a ponto de reduzir relações indiretas e apontar as diretas, é possível que haja incoerências nos modelos ao se tratarem de parâmetros e variáveis escolhidas. Ainda assim, todas as variáveis representam atividades e locais atrativos, mantendo-se a coerência teórica dos dados.

A Figura 46 apresenta as correlações entre as variáveis de centralidade das redes e a atração de viagens. As correlações são apresentadas em gráfico polar para facilitar a visualização e comparação.

Figura 46 – Correlações entre variáveis dependentes de atração e variáveis explicativas de centralidades de rede.



Fonte: O autor (2020).

Os comportamentos das correlações são similares para ambos os bancos de dados, demonstrando maiores correlações para as centralidades de intermediação métrica, seguida da intermediação angular, que são as centralidades mais relacionadas com as atividades econômicas, como constatado no Capítulo 4. A maior distinção é para correlações com o 85º Percentil (P85) das intermediações métricas – cujo maior valor de correlação é para a rede oficial –, o que pode refletir as diferenças lógicas para trechos específicos da rede como observado no Capítulo 5.

É interessante notar a importância relativa que as centralidades de proximidade, especialmente a angular, ganham na medida em que se observam os maiores registros de valores de centralidades nas zonas (Q3 e P85) – com correlações mais distintas da correlação com a centralidade de controle –, ou seja, as zonas que apresentam picos maiores de proximidades, muito relacionadas com a acessibilidade, são mais correlacionadas à atração de viagem observada. Além disso, todas as centralidades de intermediação e de proximidade são mais elevadas em comparação com a centralidade controle unitária, apontando para uma concentração de viagens atraídas ao longo dos eixos estruturais e do centro da cidade, seguindo-se a concentração de atividades econômicas. Este comportamento se repete para modos público e semipúblico e motorizados privados, e ocorre independente do motivo.

Já para viagens não-motorizadas, as correlações maiores se dão com zonas de maiores densidades de via (valores altos nas zonas para a centralidade unitária). A concentração de viagens atraídas não-motorizadas, portanto, não segue o padrão das demais, não aparentando preferência por zonas intermediárias. A centralidade de proximidade angular, no entanto, tem correlação ligeiramente superior às demais, o que também demonstra a importância da acessibilidade com a impedância angular para viagens não-motorizadas. Contudo, as diferenças com a centralidade unitária para estas viagens mostram que sua concentração não é influenciada pela configuração da rede, ao menos a nível global. Isto é coerente com o entendimento de que a acessibilidade por modos não-motorizados ocorre a nível mais localizado. Ou seja, estes modos se distribuem mais na rede e ocorrem em locais mais próximos de suas origens, portanto, possuem comportamentos distintos daqueles modos de transporte que permitem maior deslocamento e, assim, refletem diferentemente a relação com a rede de forma global.

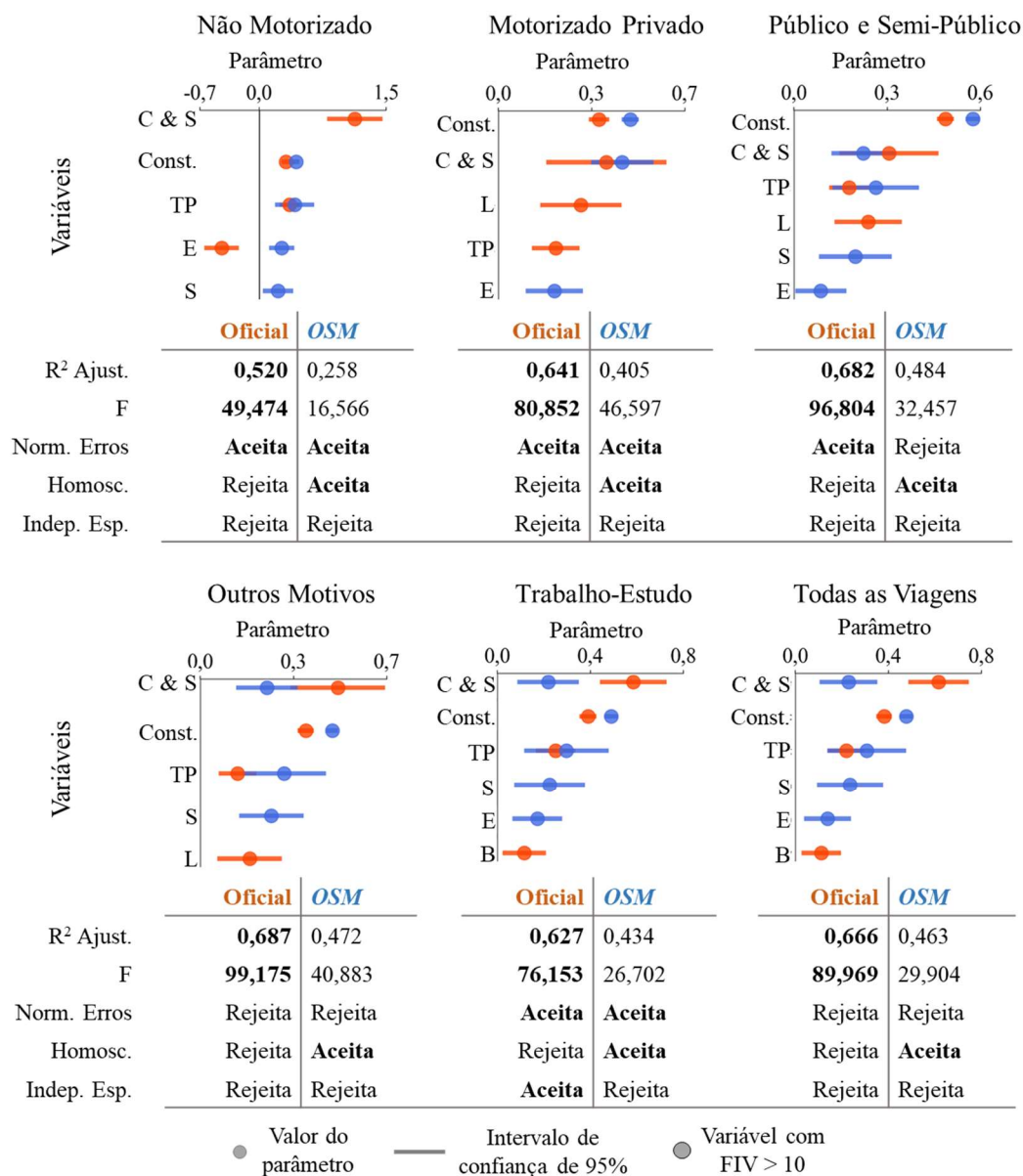
Resta, portanto, obter os modelos para ser possível analisar os comportamentos resultantes. O Apêndice 5 apresenta as tabelas com os resultados das calibrações dos trinta e seis modelos

explicativos. Cada tabela mostra os seis modelos calibrados para cada tipo de viagem atraída, demonstrando-se o expoente λ da transformação Box-Cox, as variáveis escolhidas para cada modelo com o coeficiente, o teste FIV para cada variável e os demais testes de hipótese descritos anteriormente, apontando-se a aceitação ou rejeição da premissa. Os resultados são aqui apresentados de forma resumida, com as características mais importantes para a análise.

6.3.1 Modelos com Variáveis Tradicionais

Para cada tipo de viagens atraídas, a Figura 47 apresenta as comparações entre os modelos com dados oficiais e dados do *OSM*, com o R^2 ajustado, a estatística F e os testes para normalidade dos erros, homoscedasticidade e independência espacial. Os gráficos em floresta indicam os coeficientes escolhidos para cada variável, mostrando o intervalo de confiança de 95% para cada coeficiente de calibração. Assim, é possível comparar as variáveis escolhidas para cada regressão e visualizar a conformidade entre os coeficientes, permitindo-se observar a diferença resultante entre os dados de *POIs* e os dados oficiais de alvarás comerciais, pontos de ônibus e paraciclos. Pontos circulados apontam as variáveis com FIV maiores do que 10, indicando multicolinearidade com as demais variáveis escolhidas.

Figura 47 – Comparação entre modelos com dados oficiais e do *OSM* calibrados somente com variáveis explicativas tradicionais.



Fonte: O autor (2020).

Os modelos com dados oficiais tiveram melhores ajustes, explicando melhor as viagens atraídas em comparação com os dados de *POIs* do *OSM*. Todos os valores de F são significativos para 99,9% (valores- $p \leq 0,001$). Como já demonstrado no Capítulo 5, as classes de *POIs* não têm boa acurácia posicional nas ZT, portanto, o baixo ajuste já era esperado. Algumas variáveis são escolhidas para ambos os modelos, independentemente dos dados utilizados para calibração, como Comércio e Serviços (C & S) e Transporte Público (TP). Com a observação das correlações pela Figura 45, a classe C & S de fato tem os maiores valores para ambos os bancos de

dados, mas TP não necessariamente representa conformidade entre os dados oficiais e os dados do *OSM*.

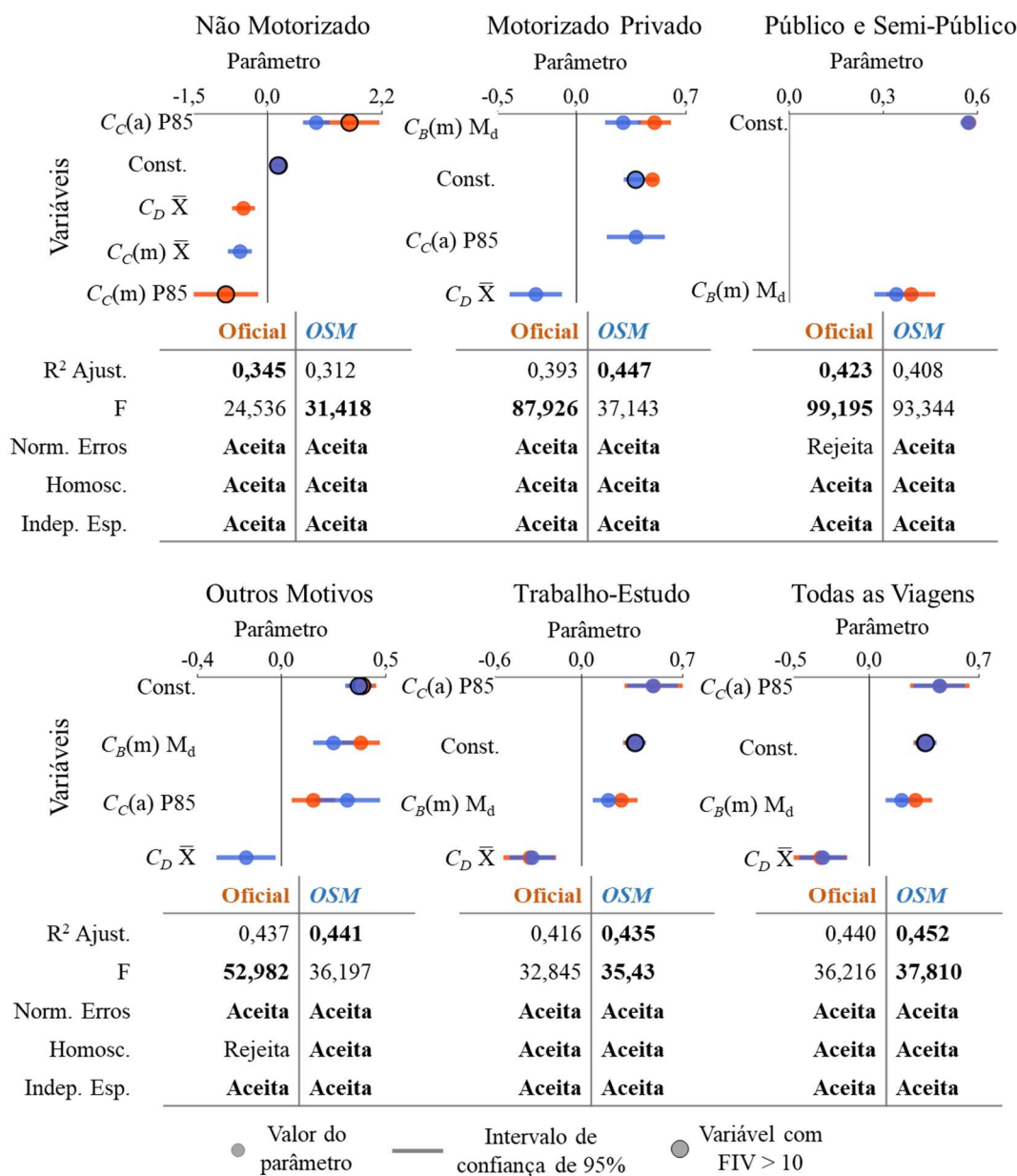
Note-se que o modelo com dados oficiais para viagens não-motorizadas escolhe um parâmetro negativo para a classe Educação (E). É esperado que instituições de ensino atraiam viagens por modos não-motorizados, mas o modelo acaba por refletir a incoerência de haver muitos pontos desta classe em regiões centrais, onde a maior parte das viagens atraídas é por modos privados ou públicos motorizados; já regiões periféricas representam grande parte das viagens atraídas não-motorizadas. Não obstante, este modelo tem o menor ajuste dentre os seis, mostrando a dificuldade de explicar estas viagens apenas com as variáveis em questão.

Nenhuma variável demonstrou alta multicolinearidade. No entanto, os testes de normalidade dos erros, homoscedasticidade e independência espacial não são aceitos conjuntamente por qualquer dos modelos, o que demonstra que as variáveis tradicionais aplicadas mantêm certos vieses no modelo de regressão MQO, indicando que modelos generalizados seriam mais representativos para os dados aplicados, porém não é escopo do estudo buscar o modelo mais representativo para previsão das viagens, e sim a comparação do comportamento dos dados nos modelos mais tradicionais.

6.3.2 Modelos Apenas com Centralidades

As calibrações feitas apenas com as variáveis de centralidades por ZT são apresentadas na Figura 48 da mesma forma que os modelos com variáveis tradicionais, o que possibilita a comparação entre os modelos com as centralidades adquiridas da rede oficial e da rede do *OSM*.

Figura 48 – Comparação entre modelos com dados oficiais e do *OSM* calibrados somente com variáveis explicativas de centralidades.



Fonte: O autor (2020).

Dada a semelhança entre as redes e os resultados de centralidade constatados no Capítulo 5, os modelos calibrados com as informações do *OSM* têm ajustes similares, por vezes maiores, em comparação aos obtidos com informações oficiais. Há, também, variáveis iguais escolhidas para ambos, com valores de parâmetros muito próximos. As variáveis mais escolhidas foram a centralidade de intermediação métrica, através da mediana ($C_B(m) M_d$) e o 85º percentil da proximidade angular ($C_C(a) P85$), a primeira muito relacionada às atividades econômicas, como mostrado no Capítulo 4, e a segunda com importantes correlações observadas na Figura 46,

sendo mais vinculada com a acessibilidade. Ressalta-se que a explicação da atração por modos públicos e semipúblicos escolheu apenas a variável relativa à mediana da intermediação métrica ($C_B(m) M_d$), que, além de melhor se correlacionar às atividades econômicas, acompanha os eixos estruturais e o centro, regiões por onde passam as principais linhas do sistema de BRT da cidade.

Alguns modelos escolheram variáveis com parâmetros negativos, em sua maioria a média da centralidade de grau ($C_D \bar{X}$). A escolha do grau com parâmetros negativos pode refletir a súbita queda dos valores na região central e ao longo dos eixos estruturais, como visto no Capítulo 5. Note-se que na representação usada, por trechos de vias, esta centralidade indica regiões com muitos cruzamentos e ruas com dois sentidos, comumente encontradas em áreas residenciais e não em vias com maior hierarquia e com grandes volumes de tráfego.

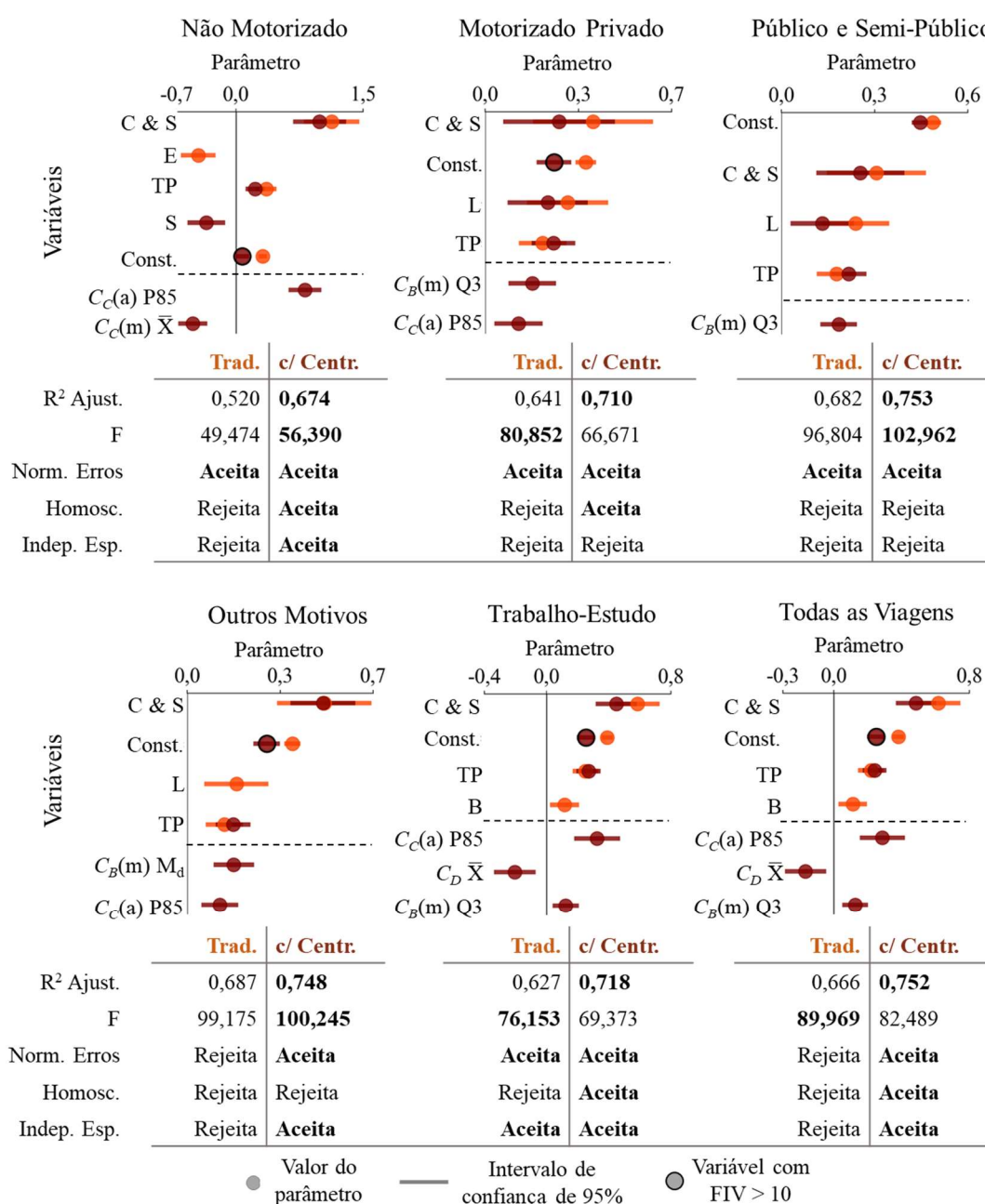
Para modos não-motorizados, o modelo com dados do *OSM* adicionou a média da proximidade métrica ($C_C(m) \bar{X}$), e o modelo com dados oficiais adicionou, além da média do grau, o 85º percentil da proximidade métrica ($C_C(m) P85$), esta última com alto valor de FIV, indicando colinearidade com as demais variáveis, em especial o 85º percentil da proximidade angular ($C_C(a) P85$). A escolha por variáveis relacionadas à proximidade métrica com parâmetros negativos reflete um problema comum em modelos que buscam aplicar variáveis que representem acessibilidade relativa à distância ou ao tempo, por vezes elas podem ser ambíguas e portanto desconsideradas (CORDERA *et al.*, 2016). A multicolinearidade apontada na $C_C(m) P85$ e na $C_C(a) P85$ reforça essa ambiguidade e demonstra como a impedância angular pode melhor se relacionar com a atração de viagens.

Todos os modelos apresentam estatística F significativa para um nível de confiança de 99,9%, com alguns modelos mostrando aceitação dos testes para normalidade dos erros, homoscedasticidade e independência espacial, ou seja, as premissas são aceitas para a explicação do fenômeno. Porém os ajustes são inferiores aos modelos com variáveis tradicionais, sendo o pior ajuste para as viagens não-motorizadas. A maioria dos modelos também apresenta altos valores de FIV para as constantes, ou interceptos. Isto não é preocupante, pois indica apenas que a combinação das outras variáveis pode prever uma variável constante. Inclusive, é comum não se adicionar o intercepto em análises de FIV para multicolinearidade (BELSLEY; KUH; WELSCH, 2004).

6.3.3 Modelos com Variáveis Tradicionais e Centralidades

Após a calibração de modelos apenas com variáveis tradicionais, e apenas com variáveis de centralidade, calibram-se modelos considerando-se os dois conjuntos de variáveis, tanto para os dados oficiais quanto para os dados do *OSM*. A Figura 49 apresenta os resultados para as calibrações com informações oficiais, comparando-se os modelos com variáveis tradicionais, já demonstrados, com os modelos em que as centralidades são também aplicadas ao *Stepwise*.

Figura 49 – Comparação de modelos com dados oficiais calibrados somente com variáveis explicativas tradicionais (“Trad.”) e considerando variáveis de centralidade (“c/ Centr.”).



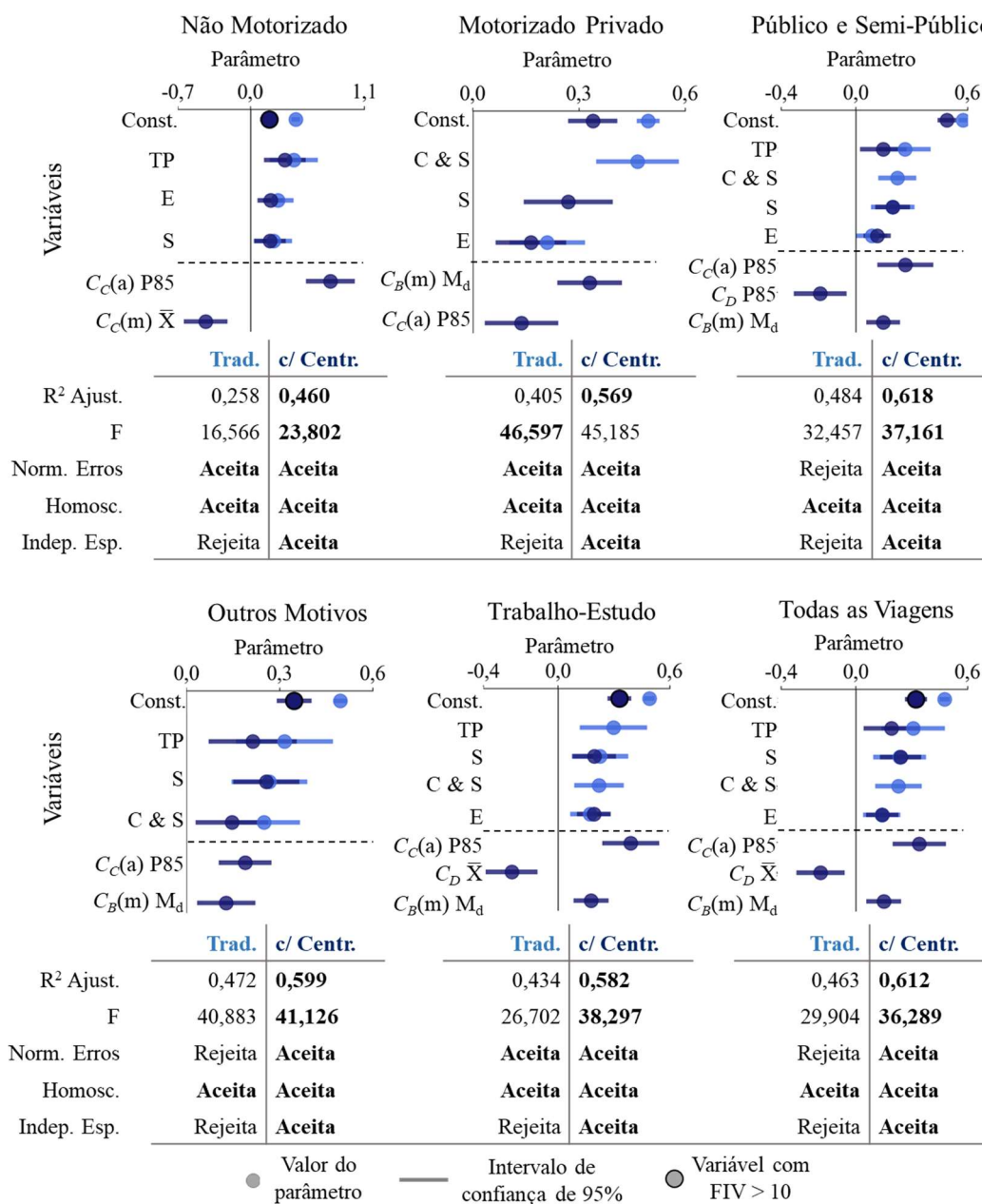
Fonte: O autor (2020).

A consideração das centralidades gera aumento do ajuste para todos os modelos, mesmo que algumas variáveis tradicionais sejam excluídas para viagens não-motorizadas, a trabalho-estudo, outros motivos e totais. Dentre as variáveis de centralidade mais escolhidas estão o 85º percentil da proximidade angular ($C_C(a)$ P85), o terceiro quartil da intermediação métrica ($C_B(m)$ Q3) e, para viagens com outros motivos, a mediana da intermediação métrica ($C_B(m)$ Md). A proximidade angular, portanto, se mostra importante para explicar as viagens mesmo em modelos com variáveis tradicionais, especialmente as viagens não-motorizadas, pois o parâmetro escolhido se equipara ao da variável de comércio e serviços (C & S). Já com parâmetros negativos, novamente a proximidade métrica e o grau são escolhidos, através de suas médias ($C_C(m) \bar{X}$ e $C_D \bar{X}$), o que demonstra ambiguidade na explicação das viagens por estas variáveis.

Muitos dos modelos passam a apresentar alta colinearidade no intercepto, o que não é problemático. O modelo para viagens não-motorizadas é o único que seleciona uma variável tradicional com parâmetro negativo, que representa os locais relacionados à Saúde (S), e exclui a variável tradicional de Educação (E), que também apresentava parâmetro negativo. Assim, apesar das centralidades levarem ao melhor aumento de ajuste dentre os seis modelos – um aumento de 0,154 no R^2 ajustado –, as viagens não-motorizadas necessitariam de outras variáveis para melhor explicá-las, visto que algumas incoerências persistem. Mesmo assim, a adição das centralidades tornou o ajuste similar aos ajustes da explicação das outras categorias de viagens.

O único modelo que não adiciona variáveis relativas à proximidade angular é para atração de viagens por modos públicos e semipúblicos. Ou seja, a acessibilidade angular parece não ter efeito para estas viagens. Todos os modelos têm estatística F com nível de significância de 99,9%, e muitas das premissas passam a ser aceitas com a inclusão das centralidades como variáveis, como visto para o modelo de atração de todas as viagens, que passa a aceitar as premissas de normalidade dos erros, homoscedasticidade e independência espacial. Assim, resta observar como a adição das centralidades afeta os modelos calibrados com dados do OSM. A Figura 50 apresenta a comparação com os modelos apenas com variáveis tradicionais.

Figura 50— Comparação de modelos com dados do *OSM* calibrados somente com variáveis explicativas tradicionais (“Trad.”) e considerando variáveis de centralidade (“c/ Centr.”).



Fonte: O autor (2020).

Os modelos obtidos com dados do *OSM*, que adicionam as variáveis relativas às centralidades, também levam a melhores ajustes se comparados com os que somente consideram os *POIs*, resultando em R² ajustados próximos de modelos com variáveis tradicionais calibrados com dados oficiais. Ou seja, modelos obtidos com informações de *POIs* e de centralidades de rede do *OSM* podem ter um poder explicativo semelhante a modelos tradicionais obtidos com dados oficiais. Todos os modelos também apresentam estatística F significativa para um nível de 99,9%. Porém, nota-se mais uma vez a dificuldade em explicar as viagens não-motorizadas,

pois o modelo tem o pior ajuste mesmo com a adição das centralidades; mas vale destacar o ganho em relação ao modelo somente com variáveis tradicionais, que é o maior – aumento de 0,202 – dentre os seis tipos de viagens atraídas, como ocorreu com os dados oficiais.

É interessante observar como todos os parâmetros das variáveis tradicionais diminuem ao se adicionarem as centralidades, o que ocorre independentemente do banco de dados utilizados – oficial ou do *OSM* –, com exceção da classe TP para alguns modelos com dados oficiais. Contudo, para a maioria dos modelos obtidos com dados do *OSM*, nota-se que a classe Comércio e Serviço (*C & S*) é excluída, e a mediana da intermediação métrica ($C_B(m)$ M_d) é adicionada, o que não ocorreu para os modelos com informações oficiais. Esta variável de centralidade é a mais correlacionada com as atividades econômicas e pode estar suprimindo informações faltantes nos modelos devido às deficiências dos *POIs*. Não obstante, a variável relativa ao 85º percentil da proximidade angular ($C_C(a)$ P85) foi adicionada a todos os modelos do *OSM*, e também foi adicionada à maioria dos modelos com dados oficiais. Esta variável traduz a acessibilidade no contexto da Sintaxe Espacial, indicando que pontos mais acessíveis de forma retilínea são também mais atrativos.

Em alguns modelos com dados do *OSM*, variáveis de centralidades relativas ao grau (C_D) e à proximidade métrica ($C_C(m)$) são escolhidas com parâmetros negativos, similar ao ocorrido com modelos de dados oficiais. Estas centralidades, portanto, podem não ser representativas para a explicação das viagens atraídas através dos modelos utilizados. A proximidade métrica corrobora a ambiguidade muito observada em modelos que tentam aplicar variáveis que traduzem acessibilidade por distância (CORDERA *et al.*, 2016); já a centralidade de grau, que indica acessibilidade local, pode ter este comportamento devido à representação *dual* escolhida para a rede, baseada nos trechos de vias.

Não há variáveis com valores de FIV maiores do que 10, apenas algumas das constantes passam a demonstrar multicolinearidade ao se adicionarem as variáveis de centralidades, o que não é um problema para os modelos (BELSLEY; KUH; WELSCH, 2004). Os testes de hipótese para normalidade dos erros, homoscedasticidade e independência espacial demonstram que, aplicando-se as variáveis de centralidades, os modelos passam a aceitar algumas das premissas. Nenhum modelo passa a recusar as premissas. Isto ocorre tanto para os modelos obtidos com dados oficiais quanto para os obtidos com dados do *OSM*, mas de forma mais evidente para estes últimos, pois todos aceitam as premissas dentro do nível de confiança de 95%.

6.4 CONSIDERAÇÕES FINAIS

Os resultados para os modelos calibrados apenas com variáveis tradicionais expõem novamente que os *POIs* do *OSM* não representam de maneira satisfatória os pontos de cada classe quando comparados com bancos de dados oficiais da prefeitura. Para estes modelos, os dados do *OSM* levam a coeficientes de determinação (R^2 ajustados) entre 0,258 e 0,484, com grandes diferenças nos parâmetros de calibração e nas variáveis escolhidas em relação aos modelos com dados oficiais, que tiveram R^2 ajustados entre 0,520 e 0,687. Contudo, os modelos apenas com variáveis de centralidade tanto da rede oficial quanto da rede do *OSM* são muito similares, com a maior diferença existente entre os R^2 ajustados sendo de 0,054 para os modelos de viagens motorizadas privadas. Todos apresentam valores de R^2 ajustados entre 0,312 e 0,452. Esta similaridade reforça que a completude, a acurácia posicional e a consistência lógica são satisfatórias para a rede do *OSM*.

A adição das centralidades para os modelos com dados oficiais leva aos melhores ajustes, com R^2 ajustados entre 0,672 e 0,757, mostrando que as centralidades podem auxiliar na explicação das viagens atraídas, especialmente através dos maiores valores de centralidade de proximidade angular em cada zona – o 85º percentil desta centralidade foi um dos mais escolhidos. Isto também ocorre para os modelos com dados do *OSM* com variáveis tradicionais e de centralidades, cujos R^2 ajustados ficam entre 0,460 e 0,618, aproximando-se dos ajustes dos modelos com dados oficiais que consideram apenas variáveis tradicionais. Além disso, a centralidade de intermediação, representada pela mediana nas *ZT*, foi muito escolhida nos modelos do *OSM*, demonstrando-se que, diferente dos dados oficiais, os dados do *OSM* foram complementados por esta que é a variável mais relacionada com as atividades econômicas. Ou seja, as centralidades auxiliam em contexto de falta de dados ou dados com menos informações. Ademais, nenhum modelo que utilizou informações de centralidades optou por variáveis relativas à centralidade unitária de controle, mostrando que o efeito da configuração da rede é mais relacionado com a atração das viagens do que simplesmente com a presença ou não de vias.

As viagens com maiores ganhos nos ajustes após a adição das centralidades foram as não-motorizadas, independentemente da origem dos dados observados, cuja variável relativa à proximidade angular teve maior importância, expressando que a acessibilidade pode estar mais vin-

culada com viagens que não dependam de automóveis, como relatam Vickerman (1974) e Cordera *et al.* (2016). Outrossim, a proximidade angular, ao ser escolhida em diversos modelos, independentemente do motivo da viagem, corrobora os preceitos da Sintaxe Espacial de Hillier (1996) de que a percepção retilínea pode melhor explicar a escolha dos caminhos e as dinâmicas urbanas. Apenas para atração de viagens por modos públicos e semipúblicos, quando considerados os dados oficiais, a proximidade angular não é importante, indicando que estes modos podem não ser tão impactados pela impedância angular como os outros são. Ademais, a centralidade de proximidade angular também teve grande importância para os modelos obtidos por Jayasinghe, Sano e Rattanaorn (2017), contudo, os modelos apenas com variáveis de centralidade para Curitiba trazem a centralidade de intermediação métrica como uma das mais importantes. Talvez isto reflita a diferença de se explicar o total de viagens atraídas e não a densidade de viagens atraídas, como fizeram os autores, mas também pode refletir o planejamento da cidade por eixos estruturantes segundo preceitos DOT, que, como visto no Capítulo 4, reproduz a busca por regiões intermediárias por parte das atividades econômicas (OMER; GOLDBLATT, 2016).

7 CONCLUSÕES E RECOMENDAÇÕES

Cessadas as etapas de obtenção dos resultados relativos aos objetivos específicos do trabalho, este capítulo sintetiza as principais conclusões de cada etapa e, por consequência, cumpre o objetivo principal de avaliar a usabilidade do *OSM* para análises de atração de viagens. As discussões abrangem o uso proposto e extrapolam as possibilidades que os achados permitem, evidenciando as vantagens e limitações dos dados do mapa colaborativo.

7.1 CONCLUSÕES

7.1.1 Centralidades de rede e atividades econômicas

Os resultados do Capítulo 4 corroboram estudos embasados na Sintaxe Espacial que relacionaram as centralidades da rede e características urbanas de uso e ocupação do solo (RUI; BAN, 2014; LIU *et al.*, 2016), especialmente para atividades econômicas (PORTA *et al.*, 2012; OMER; GOLDBLATT, 2016). Contudo, este trabalho adicionou uma centralidade fictícia de controle, definindo valor 1 para todas as arestas da rede. Isto demonstrou que a “nível de rua” a localização específica das atividades mais voltadas ao público, como comércio e serviços, busca vias mais integradas – com maior centralidade de proximidade –, porém esta preferência é mais perceptível para classes de atividades com menor número de registros. Ou seja, a centralidade de controle mostra que o efeito das centralidades é notável para atividades voltadas ao público quando não muito dispersas – com menor número de pontos. Também houve pouca diferença entre a impedância métrica e angular das centralidades observadas, mostrando que, para o caso de Curitiba, a impedância energética – métrica – e de informação – angular – têm importância similar para a localização das atividades.

Não obstante, ao se observar a nível de ZT, a relação fica mais evidente. A mediana dos valores das centralidades em cada zona expressa maior importância da intermediação para a localização das atividades, ou seja, as atividades estão dispostas nas zonas por onde a maior parte dos caminhos mínimos da rede ocorre. Mesmo comparadas à centralidade de controle, as centralidades de intermediação têm correlações maiores, principalmente com atividades voltadas ao público. A impedância métrica, ou energética, para a centralidade de intermediação mostrou maior correlação para a maioria das classes de alvarás, mas vale destacar a maior influência do planejamento urbano neste nível de análise, isto é, os resultados refletem não apenas a intenção das atividades, mas de dispositivos de uso e ocupação do solo de Curitiba. Isto é coerente com o

histórico de urbanização da cidade por meio de eixos estruturantes e princípios de DOT (MIRANDA, 2010), em que tradicionalmente as distâncias métricas são adotadas para as tomadas de decisão.

Os resultados também corroboram a economia do movimento da Sintaxe Espacial (HILLIER, 1996), que destaca como a configuração da rede é moldada e também molda as dinâmicas urbanas, levando as atividades econômicas a preferirem locais que usufruem desta configuração. Curitiba, por apresentar altas correlações das atividades com ruas mais integradas e, ao mesmo tempo, regiões mais intermediárias, se compara a cidades com crescimento “natural”, como visto por Omer e Goldblatt (2016). O primeiro objetivo específico, portanto, é atingido e garante maior valor à rede viária em análises de atração de viagens, especialmente com os resultados zonais diretamente relacionados à agregação das atividades destino das viagens.

7.1.2 A qualidade dos dados do *OpenStreetMap*

O Capítulo 5, que comparou os dados fornecidos oficialmente pela prefeitura de Curitiba e os dados adquiridos pelo *OSM*, trouxe resultados adversos para os *POIs*, mas satisfatórios para a rede viária obtida na plataforma. A completude dos *POIs* é evidentemente baixa, ou seja, o número de pontos demonstrados no *OSM* não se aproxima do número real observado nos bancos de dados oficiais segundo cada classe definida, com proporções próximas de 0 entre os dados do *OSM* e os oficiais para a maioria das ZT; a acurácia posicional agregada por zona, que observa a proporção por zona das classes, mostra que a dispersão dos *POIs* também não é equiparável à dispersão dos pontos oficiais nas ZT. Apenas as classes Saúde (S), Instalações para bicicleta (B) e Educação (E) apresentaram valores semelhantes – pelo teste de Wilcoxon –, mas somente a classe Educação (E) teve também uma distribuição semelhante – pelo teste de Kolmogorov-Smirnov –, o que pode refletir a tendência de maior mapeamento acadêmico ou próximo de ambientes acadêmicos, também notado por Hochmair, Juhász e Cvetojevic (2018).

A análise por classes dos *POIs* permite uma comparação semântica quase impraticável se feita ponto a ponto, devido às inconsistências dos atributos e de nomenclatura entre o banco de dados oficial e do *OSM*. Ou seja, apesar de não se verificar a qualidade semântica de cada *POI* – similaridade da interpretação dos locais –, ao se segmentá-los por classes este problema é minorado, pois a aplicação proposta não necessita de detalhes de cada ponto, mas de sua classe. A rede, por sua vez, apresenta completude e acurácia posicional satisfatórias para as análises

macroscópicas efetuadas neste trabalho, apesar de ter o centro da cidade com resultados ligeiramente piores para acurácia posicional. As falhas notadas, no entanto, não impactam significativamente a consistência lógica resultante das KDE de cada centralidade. Ainda que as centralidades não sejam exatamente iguais entre as duas redes, o comportamento ao longo dos elementos é similar, e resulta em superfícies de KDE muito correlacionadas e com erros próximos de 0. Portanto, a rede do *OSM* tem boa qualidade para a aplicação proposta.

O segundo objetivo, portanto, de avaliar a qualidade dos dados do *OSM* em relação aos dados oficiais é atingido. A rede do *OSM* se apresenta como substituta satisfatória para a aplicação desejada; já os *POIs* apresentam deficiências tanto na completude quanto na acurácia posicional, o que pode limitar análises de atração de viagens que visem à previsão, mas podem permitir usos explicativos das viagens limitados ao intervalo temporal de extração dos dados.

7.1.3 Modelos explicativos de viagens atraídas com dados do *OpenStreetMap*

Através do método *Stepwise*, os modelos calibrados no Capítulo 6 com variáveis tradicionais, ou seja, apenas com dados relativos a pontos atrativos, mostram que os *POIs* não permitem ajustes similares aos modelos com dados oficiais, o que inibe aplicações que visem a estabelecer a real relação entre a atração de viagem e cada classe de atividade. Porém, os modelos calibrados com *POIs* seguem uma coerência se comparados com os dados oficiais, especialmente para pontos de Comércio e Serviços e de Transporte Público.

Os modelos MQO calibrados apenas com variáveis das centralidades de rede refletem a similaridade entre a rede do *OSM* e a rede oficial. Ainda que com ajustes menores do que modelos com variáveis tradicionais, ambas as redes levaram a modelos com calibrações semelhantes. As centralidades escolhidas frequentemente são as mais relacionadas às atividades econômicas – intermediação – e, como também constatado por Jayasinghe, Sano e Rattanaporn (2017), à acessibilidade no contexto da Sintaxe Espacial – proximidade angular –, em que locais mais acessíveis por caminhos menos tortuosos melhor ajudam a explicar a atração de viagens. Há certa ambiguidade, no entanto, para proximidade métrica e grau, por vezes escolhidos com parâmetros de calibração negativos. A primeira pode estar associada a eventuais inconsistências de medidas de acessibilidade por distância ou tempo (CORDERA *et al.*, 2016); a segunda, no entanto, tem reflexo na representação *dual* utilizada, em que vias com muitos cruzamentos e duplo sentido apresentam maior grau, já que comumente estão em áreas mais residenciais.

Tanto para dados oficiais quanto para dados do *OSM*, adicionar centralidades aos modelos aumentou seus ajustes. Nota-se que, para os dados do *OSM*, as centralidades de intermediação – mais relacionadas às atividades econômicas – manifestam um efeito de substituição da variável de *POIs* de Comércio e Serviços, ou seja, a centralidade pode complementar dados faltantes ou deficitários relacionados a atividades desta classificação; já a centralidade de proximidade angular foi adicionada tanto para dados oficiais quanto para dados do *OSM*, indicando que sua escolha está vinculada ao papel de variável de acessibilidade, como comentado anteriormente, mais uma vez corroborando a Sintaxe Espacial na premissa de que caminhos retilíneos são mais importantes do que somente a distância percorrida. Neste sentido, a junção dos *POIs* e das centralidades da rede originou modelos com dados do *OSM* de ajustes semelhantes a modelos somente com variáveis tradicionais de dados oficiais. Mais do que isso, mostrou que a rede pode auxiliar os modelos, seja complementando dados incompletos, seja adicionando informações de acessibilidade ou de eventuais atividades não contempladas.

O ganho no poder explicativo com a adição das centralidades é parecido para o motivo de trabalho ou estudo e outros, mas foi mais acentuado para viagens não-motorizadas, o que é esperado em modelos que adicionam informações sobre acessibilidade (VICKERMAN, 1974; CORDERA *et al.*, 2016). Para os dados oficiais, as variáveis que trazem a informação de acessibilidade levam a um maior ganho de poder explicativo para atração de viagens não-motorizadas, cujo modelo escolhe a variável de proximidade angular com maior parâmetro, e apresenta maior aumento no ajuste. Viagens por modos públicos e semipúblicos, no entanto, possivelmente não se beneficiam tanto pela acessibilidade angular, como visto no modelo com dados oficiais, mas o modelo com dados do *OSM* a escolheu mesmo assim, o que indica que ela também pode ter caráter de substituição de informações faltantes.

Em suma, todos os modos e motivos de viagens se beneficiaram da adição das centralidades de proximidade e de intermediação, relacionadas à acessibilidade em rede – principalmente por caminhos retilíneos – e à concentração de atividades relacionadas ao público, respectivamente. A premissa de “economia do movimento” (HILLIER, 1996) é corroborada, portanto, não apenas pela relação da configuração da rede com as atividades econômicas, mas também com os potenciais de deslocamentos na cidade, verificados pela atração de viagens. Locais com maior integração angular são mais atrativos, bem como locais com intermediação métrica, por se correlacionarem com as atividades-fim de muitas das viagens. Ainda, as variáveis de centralidades em muitos dos modelos MQO garantiram aceitação de testes de hipótese para as premissas de

normalidade dos erros, homoscedasticidade e independência espacial. Nota-se, assim, que as informações de centralidades atuaram também minorando problemas recorrentes da modelagem de eventos espaciais. O terceiro objetivo, portanto, é atingido, demonstrando o papel que as centralidades de rede podem ter, e trazendo as ressalvas para a utilização dos *POIs*.

7.1.4 A usabilidade do *OpenStreetMap* para análise de atração de viagens

Voltando-se ao objetivo geral do trabalho: verificar a usabilidade dos dados de *POIs* e da rede do *OSM* para a explicação da atração de viagens em Curitiba. A verificação foi efetuada e traz alguns apontamentos importantes. É possível utilizar os dados de *POIs* e de centralidades da rede do *OSM* para análises de atração de viagens, desde que as devidas precauções sejam consideradas, principalmente em relação aos *POIs*.

- As centralidades de rede podem ser utilizadas como variáveis explicativas para a atração de viagens, atentando-se para o papel de complemento ou de acessibilidade que cada variável pode ter. No caso de Curitiba, a intermediação é muito relacionada a atração por também se relacionar com o posicionamento das atividades econômicas, mas este resultado pode estar relacionado ao histórico de planejamento da cidade, que pode variar a depender do contexto urbano (OMER; GOLDBLATT, 2016). Já a proximidade angular parece independe do contexto urbano, relacionando-se às atividades e à atração em maior e menor grau. Neste caso, está se adicionando aos modelos a informação de acessibilidade, ainda que também possa se relacionar com o posicionamento de atividades;
- A rede do *OSM* tem uma qualidade satisfatória para esta aplicação. Já os *POIs* devem ser usados de forma cautelosa, pois não têm boa completude, e a dispersão na cidade não leva a uma boa acurácia posicional se comparada a dados oficiais, mesmo que observados de forma agregada em ZT. Em decorrência disto, modelos calibrados apenas com dados dos *POIs* não tiveram ajustes semelhantes aos calibrados com dados oficiais, não podendo ser utilizados para afirmar a real relação entre cada classe e a atração de viagem, mas apenas para inferir as classes que auxiliam na atração. Modelos apenas com centralidades, por sua vez, foram muito similares entre a rede oficial e do *OSM*, com certa diferença maior para a centralidade de grau, mais suscetível a variações locais;
- Por conseguinte, os modelos explicativos calibrados com dados do *OSM* não têm caráter preditivo, mas permitem o uso para aplicações operacionais, como o planejamento de

rota e frota de transporte público para curto prazo, avaliação de impacto imediato de polo produtor de viagens, definição de regiões preferenciais para intervenções de acessibilidade – calçadas e travessias –, dentre outras utilidades momentâneas que exijam informações sobre locais atrativos muitas vezes não disponíveis.

A ubiquidade que os dados geográficos vêm adquirindo com mapas colaborativos pode ter papel fundamental no planejamento urbano e de transportes, sobretudo com o aumento de sua qualidade. Este trabalho demonstrou semelhanças e limitações dos dados do *OSM* para a cidade de Curitiba, mas promove um método que permite esta avaliação para outros mapas colaborativos e outras cidades. Lembrando que há informações no *OSM* que não são encontradas em dados oficiais, como outras classes de *POIs* e as marcas de edificações. As observações se limitaram a comparar dados semelhantes de ambos os bancos de dados. Ou seja, há informações que podem se somar às garantidas pela autoridade local.

Outrossim, o trabalho corrobora a Sintaxe Espacial, mostrando que a rede reflete e induz as dinâmicas econômicas urbanas, suscitando maior importância para a infraestrutura viária das cidades no impacto do uso e ocupação do solo e na geração das viagens, indicando que o banco de dados da rede deve ser valorizado para análises de atração de viagens. Outras cidades, neste sentido, podem obter o grafo da rede viária e calcular as centralidades, especialmente de proximidade e de intermediação, para complementar análises operacionais de atração de viagens, ou mesmo distribuição, com a ressalva de que a relação entre centralidades, atividades e viagens não é hermética, isto é, contextos diferentes de planejamento levam a relações distintas, com diversas variáveis que podem impactá-las, apesar de haver fortes indícios de que a acessibilidade angular persiste para diferentes contextos.

Os resultados também permitem algumas observações colaterais, como o potencial das centralidades de rede e de sua junção com alvarás comerciais para a definição de regiões atrativas de forma geral, ou até mesmo como indicadoras de uso do solo, como apontado por Liu *et al.* (2016), não necessariamente fixadas por zonas de tráfego oficialmente estabelecidas. Além disso, nota-se que, para a atração de viagens, a acessibilidade calculada com a impedância angular teve maior importância do que a calculada em metros, o que indica que há preferência por caminhos retilíneos, como aponta a Sintaxe Espacial (HILLIER, 1996; SERRA; HILLIER, 2019). Assim, além de distância em rede, distância euclidiana e tempo de viagens, como alguns

modelos aplicam, a diferença angular dos caminhos percorridos pode ter papel importante em medidas de acessibilidade para as viagens, principalmente para modos não-motorizados.

Ademais, a rede do *OSM* tem qualidade suficiente para aplicações globais, fortalecendo usos em larga escala dos dados geográficos, como em estudos morfológicos das redes (BOEING, 2018; KURAMOTO, 2019; MORELLI, 2019), podendo até mesmo ser utilizadas em análises microscópicas, desde que considerado o tratamento posterior à aquisição. A possibilidade de importação gratuita em formato de grafos é mais uma vantagem das IGV, pois isto facilita aplicações em transportes; muitas vezes o mapeamento oficial se restringe à representação dos eixos viários segmentados por logradouros, sem considerar a lógica de rede como no *OSM*. Até mesmo os *POIs* podem auxiliar em análises localizadas, informando pontos importantes próximos, mesmo que não em sua totalidade.

Nesse sentido, a integração entre dados oficiais e de mapas colaborativos pode ser muito benéfica, tanto para cidades com políticas de mapeamento vigentes quanto para cidades sem iniciativas similares. O compartilhamento de dados pelo *OSM* – a exemplo, a disponibilização dos dados oficiais de Curitiba na plataforma colaborativa – permite participação popular no mapeamento e em tomadas de decisões, reduzindo custos e possibilitando atualizações mais constantes, até mesmo levando a um maior escrutínio das informações e ao aumento de sua qualidade, como apontam Machado e Camboim (2019), o que facilita e expande realizações no âmbito do planejamento urbano e de transportes, bem como da análise das viagens, muito dependente do devido mapeamento dos fenômenos da cidade.

7.2 SUGESTÕES PARA TRABALHOS FUTUROS

Como já apontado, o método proposto incentiva aplicações em diferentes cidades, facilitando sua replicação, mas também para diferentes plataformas colaborativas, visando à comparação de bancos de dados e contextos urbanos. Especial atenção pode ser dada para IGVs comerciais que disponham de dados de viagens por dispositivos de GPS (ex.: *GoogleMap* e *TomTom*), não somente para verificar a usabilidade dos *POIs* e da rede, mas para utilizar os dados de viagens para inferir a própria atração de viagens, ou o comportamento do viajante, de forma mais desagregada, desobrigando-se a necessidade de pesquisas domiciliares para tanto. Dados dos deslocamentos não somente permitem obter a atração, mas a geração e também a alocação das via-

gens. A limitação das IGVs comerciais comumente são os dados não gratuitos, mas as potencialidades do ponto de vista de planejadores e acadêmicos, especialmente em largas escalas, são amplas.

Diferentes abordagens podem ser utilizadas na agregação para a verificação das centralidades e sua relação com atividades urbanas, bem como para a verificação da qualidade dos dados do *OSM*. Este trabalho utilizou o zoneamento oficial por ser o mesmo nível de agregação da demanda por viagens, mas diferentes zoneamentos podem atestar se as relações se mantêm independentemente das unidades zonais. Além disto, iniciativas de observação de regiões homogêneas, considerando-se atividades e centralidades, podem trazer mais esclarecimento de construção de zonas urbanas de análise, base para este e outros objetivos.

A verificação da qualidade dos dados do *OSM* pode ser complementada com estudos de longo prazo para se acessar a qualidade temporal das informações. Diferentes métodos podem ser explorados para este e outros quesitos de verificação. Os *POIs*, neste sentido, são um fértil conjunto de dados a ser explorado, pois abrangem localidades mais diversas e por vezes com semântica mais detalhada do que de dados oficiais (MACHADO; CAMBOIM, 2019), mas com avaliação desafiadora, por consequência. O maior entendimento dos dados do *OSM* permite diferentes bases para sua utilização e maiores aplicações no planejamento urbano e no contexto acadêmico.

Dentre as conclusões concomitantes, as relações encontradas das centralidades com as atividades econômicas e a atração de viagens podem ser base para estudos que visem a inferir o uso e a ocupação do solo. A rede pode ser um indicativo de acessibilidade e de intensidade de movimentos, que ao caracterizarem a economia do movimento (HILLIER, 1996), estimulam ou inibem certos usos da região. Estas relações têm grande potencial para exploração de outras dinâmicas urbanas, como o transporte de cargas urbano e o planejamento de depósitos, uma vez que as centralidades podem apontar a concentração ou dispersão de atividades econômicas, sendo indicativos de pontos de abastecimento.

Outrossim, dado a histórica estruturação de Curitiba por preceitos de desenvolvimento orientado ao transporte público (DOT), e ao se considerar a similaridade de Curitiba em relação a cidades que tiveram seu crescimento “natural” e, portanto, geralmente mais integradas e inteli-

gíveis (OMER; GOLDBLATT, 2016), a relação entre centralidades e uso do solo pode ser utilizada como indicador de DOT, principalmente para cidades de grande porte. Ou seja, cidades que apresentem maiores correlações entre atividades, ou mesmo atração de viagens, com regiões intermediárias e ao mesmo tempo acessíveis, indicam maior coerência dos princípios de DOT. Este é, portanto, um raciocínio a ser testado, inclusive considerando-se apenas a rede de transporte público, caso disponível.

Mais análises também são necessárias para compreender como as centralidades e a impedância angular atuam na escolha dos destinos. Este trabalho constatou que locais acessíveis por caminhos retilíneos são mais relacionados à atração de viagem através de modelos MQO baseados em viagens. Diferentes modelos podem ser utilizados e, dentre eles, modelos mais desagregados e com abordagens baseadas em atividades, adicionando variáveis relativas à distância angular e comparando-as a variáveis de acessibilidade mais comuns, como distância métrica e tempo de viagem. Simulações podem ser ferramentas interessantes para estes testes, desde que respaldadas em comportamentos observados. Considerar informações sobre o viajante e sobre a atividade, inclusive cadeias de atividades, pode auxiliar no entendimento do impacto das centralidades e de caminhos retilíneos, que baseiam a Sintaxe Espacial, na geração e distribuição das viagens, em especial para diferentes modos utilizados, e até mesmo para a alocação dos movimentos.

Outra possível iniciativa que vale ser registrada é a integração entre dados oficiais e de mapas colaborativos, mais especificamente a disponibilização de dados em mapas colaborativos por parte de agências reguladoras. Para além de experimento de gerência comunitária dos dados, isto pode se tornar prática benéfica para a cidade, garantindo-se acesso aos dados e constantes atualizações. Sugere-se, portanto, um estudo de caso em parceria com o poder público local para disponibilização de informações geográficas sobre a cidade, com permissão de modificações. Os resultados e experiências acumuladas podem liderar aplicações semelhantes, visando maior integração de dados e modernização de processos decisórios.

REFERÊNCIAS BIBLIOGRÁFICAS

AL-TAEI, A.K., TAHER, A.M. Trip attraction development statistical model in Dohuk city residential area. **Al-Rafidain Engineering**, v. 14 n. 2, p. 12-23. 2006.

ARSANJANI, J. J.; ZIPF, A.; MOONEY, P.; HELBICH, M. An Introduction to OpenStreetMap in Geographic Information Science: Experiences, Research, and Applications. In: ARSANJANI, J. J.; ZIPF, A.; MOONEY, P.; HELBICH, M. **OpenStreetMap in GIScience. Lecture Notes in Geoinformation and Cartography**. Springer Cham Heidelberg New York Dordrecht London, p. 1–15, 2015.

BALLATORE, A.; BERTOLOTTI, M. Geographic Knowledge Extraction and Semantic Similarity in OpenStreetMap. **Knowledge and Information Systems**, v. 37, n. 1, p. 1–21, 2013.

BAVELAS, A. A Mathematical Model for Group Structures. **Applied Anthropology**, v. 7, p. 16–30, 1948.

BELSLEY, D. A.; KUH, E.; WELSCH, R. E. **Regression Diagnostics – Identifying Influential Data and Sources of Collinearity**. New Jersey: John Wiley & Sons, Inc., 2004.

BIGGS, N. L.; LLOYD, E. K.; WILSON, R. J. **Graph Theory 1736-1936**. New York: Oxford University Press, 1998.

BOEING, G. *OSMnx*: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. **Computers, Environment and Urban Systems**, v. 65, p. 126-139, 2017.

_____. A multi-scale analysis of 27,000 urban street networks: Every US city, town, urbanized area, and Zillow neighborhood. **Environment and Planning B: Urban Analytics and City Science**, 8 ago. 2018.

_____. Urban spatial order: street network orientation, configuration, and entropy. **Applied Network Science**, v. 4 n. 67, p. 1-19, 2019.

BOX, G.E.P.; COX, D.R. An Analysis of Transformations. **Journal of the Royal Statistical Society: Series B**, v. 26, n. 2, p. 211-243, 1964.

CAMBOIM, S. P.; BRAVO, J. V. M.; SLUTER, C. R. An Investigation into the Completeness of, and the Updates to, OpenStreetMap Data in a Heterogeneous Area in Brazil. **International Journal of Geo-Information**, v. 4, n. 3, p. 1366–1388, 2015.

CARLOS, H. A.; SHI, X.; SARGENT, J.; TANSKI, S.; BERKE, E. M. Density estimation and adaptive bandwidths: A primer for public health practitioners. **International Journal of Health Geographics**, v. 39, n. 9, p. 1–8, 2010.

CHANG, J. S.; JUNG, D.; KIM, J.; KANG, T. Comparative analysis of trip generation models: results using home-based work trips in the Seoul metropolitan area. **Transportation Letters**, v. 6, n. 2, p. 78–88, 2014.

CORDERA, R.; COPPOLA, P.; DELL'OLIO, L.; IBEAS, A. Is accessibility relevant in trip generation? Modeling the interaction between trip generation and accessibility taking into account spatial effects. **Transportation**, v. 44, n. 6, p. 1577–1603, 2016.

COSTA, L. D. F.; OLIVEIRA, O. N.; TRAVIESO, G.; RODRIGUES, F. A.; VILLAS BOAS, P. R.; ANTIQUEIRA, L.; VIANA, M. P.; CORREA ROCHA, L. E. Analyzing and modeling real-world phenomena with complex networks: A survey of applications. **Advances in Physics**, v. 60, n. 3, p. 329–412, 2011.

CURITIBA. **Lei nº 14.771**, Revisão do Plano Diretor de Curitiba. 2015.

D'AGOSTINO, R.; PEARSON, E. S. Tests for departure from normality. *Biometrika*, v. 60, n. 3, p. 613-622, 1973.

ECHENIQUE, M., H. Econometric models of land use and transportation. In: HENSHER, D., A.; BUTTON, K., J.; HAYNES, K., E.; STOPHER, P., R. **Handbook of transport geography and spatial systems**, v. 5, 185-202. Amsterdam: Pergamon, 2004.

ELWOOD, S. Geographic Information Science: new geovisualization technologies – emerging questions and linkages with GIScience research. **Progress in Human Geography**, v. 33, n. 2, p. 256–263, 2009.

EPANECHNIKOV, V. Nonparametric estimation of a multivariate probability density. **Theory of Probability and Its Applications**, v. 14, n. 1, p. 153–158, 1969.

ESCAMILLA, J.; COS, C.; CÁRDENAS, J. Contesting Mexico City's alleged polycentric condition through a centrality- mixed land-use composite index. **Urban Studies**, v. 53 n. 11, p. 2380-2396. 2016.

ESTES, J. E.; MOONEYHAN, W. Of maps and myths. **Photogrammetric Engineering and Remote Sensing**, v. 60 n. 5, p. 517-524, 1994.

EWING, R.; DEANNA, M.; LI, S. Land Use Impacts on Trip Generation Rates. **Transport Research Record**, v. 1518, n. 1, p. 1–6, 1996.

FOTHERINGHAM, A. S.; BRUNSDON, C.; CHARLTON, M. **Quantitative Geography: Perspectives on Spatial Data Analysis**. London: Sage, 2000.

FREEMAN, L. C. Centrality in Social Networks. **Social Networks**, v. 1, n. 3, p. 215–239, 1979.

GAO, S.; WANG, Y.; GAO, Y.; LIU, Y. Understanding urban traffic-flow characteristics: A rethinking of betweenness centrality. **Environment and Planning B: Planning and Design**, v. 40, n. 1, p. 135–153, 2013.

GEORGE, P.; KATTOR, G. J.; MALIK, A. K. V. Prediction of trip attraction based on commercial land use characteristics. **International Journal of Innovative Research in Science, Engineering and Technology** v. 2 n. 1, p. 352-359. 2013.

GERVASONI, L.; BOSCH, M.; FENET, S.; STURM, P. Calculating spatial urban sprawl indices using open data. **15th International Conference on Computers in Urban Planning and Urban Management**, Adelaide, Austrália:, 2017.

GIRRES, J.; TOUYA, G. Elements of quality assessment of French OpenStreetMap data. **Transactions in GIS**, v. 14, n. 4, p. 435–459, 2018.

GOODCHILD, M. F. Editorial: Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0. **International Journal of Spatial Data Infrastructures Research**, v. 2, p. 24–32, 2007.

GOODCHILD, M.F.; HUNTER, G.J. A simple positional accuracy measure for linear features. **International Journal of Geographical Information Science**, v. 11, n. 3, p. 299–306, 1997.

GOULIAS, K. G.; PENDYALA, R. M.; KITAMURA, R. Practical Method for the Estimation of Trip Generation and Trip Chaining. **Transportation Research Record**, v. 1285, Reprint No. 62, p. 47–56, 1991.

GUZMAN, L. A.; GOMEZ, A. M.; RIVERA, C. A Strategic Tour Generation Modeling within a Dynamic Land-Use and Transport Framework: A Study of Bogota, Colombia. **Transportation Research Procedia**, v. 25, p. 2536–2551, 2017.

HAKLAY, M. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. **Environment and Planning B: Planning and Design**, v. 37, n. 4, p. 682–703, 2010.

HAKLAY, M.; WEBER, P. OpenStreetMap: User-Generated Street Maps. **IEEE Pervasive Computing**, v. 7, n. 4, p. 12–18, 2008.

HAWICK, K. A.; JAMES, H. A. Node Importance Ranking and Scaling Properties of some Complex Road Networks. **Res. Lett. Inf. Math. Sci.**, v. 11, p. 23–32, 2007.

HEUMANN, C.; SHALABH, M. S. **Introduction to Statistics and Data Analysis – with exercises, solutions and applications in R**. Switzerland: Springer International, 2016.

HILLIER, B. **Space is the Machine**. London: Press Syndicate of the University of Cambridge, 1996.

HILLIER, B.; HANSON, J. **The social logic of space**. Cambridge university press, 1984.

HILLIER, B.; PENN, A. Cities as movement economies. **Urban Design International**, v. 1 n. 1, p. 41–60, 1996.

HIRATA, E.; GIANNOTTI, M. A.; LAROCCA, A.P.C.; QUINTANILHA, J. A. Mapeamento dinâmico e colaborativo de alagamentos na cidade de São Paulo. **Boletim de Ciências Geodésicas**, v. 19, n. 4, p. 602–623, dez. 2013.

HOCHMAIR, H. H.; JUHÁSZ, L.; CVETOJEVIC, S. Data Quality of Points of Interest in Selected Mapping and Social Media Platforms. In: **Progress in Location Based Services 2018**, Springer, Cham: 2018.

HOFER, C.; JÄGER, G.; FÜLLSACK, M. Large scale simulation of CO2 emissions caused by urban car traffic: An agent-based network approach. **Journal of Cleaner Production**, v. 183, p. 1–10, 2018.

HUNTER, G. J. New tools for handling spatial data quality: Moving from academic concepts to practical reality. **URISA Journal**, v. 11, n. 2, p. 25–34, 1999.

IACONO, M.; LEVINSON, D.; EL-GENEIDY, A. Models of Transportation and Land Use Change: A Guide to the Territory. **Journal of Planning Literature**, v. 22, n. 4, p. 1–18, 2008.

INNES, J. D.; IRCHA, M. C.; BADOE, D. A. Factors affecting automobile shopping trip destinations. **Journal of Urban Planning and Development**, v. 116 n. 3, p. 126-136. 1990.

IPPUC. **Plano de Mobilidade Urbana e Transporte Integrado**, PlanMob Curitiba. Curitiba. 60 p. 2008.

_____. **Site oficial**. Disponível em: <<https://ippuc.org.br/>>. Acessado em: 12 ago. 2020.

JACOBS, J. **The Death and Life of Great American Cities**. Random House, Nova York, 1961.

JAYASINGHE, A.; SANO, K.; RATTANAPORN, K. Application for developing countries: Estimating trip attraction in urban zones based on centrality. **Journal of Traffic and Transportation Engineering (English Edition)**, v. 4, n. 5, p. 464–476, 2017.

JIANG, B; CLARAMUNT, C. Topological analysis of urban street networks. **Environment and Planning B: Planning and design**, v. 31, n. 1, p. 151-162, 2004.

JIANG, B.; JIA, T. Agent-based simulation of human movement shaped by the underlying street structure. **International Journal of Geographical Information Science**, v. 25, n. 1, p. 51–64, 2011.

JIANG, B.; LIU, C. Street-based Topological Representations and Analyses for Predicting Traffic Flow in GIS. **International Journal of Geographical Information Science**, v. 23, n. 9, p. 1119–1137, 2009.

KELLY, E. D. The Transportation Land-Use Link. **Journal of Planning Literature**, v. 9, p. 128–145, 1994.

KIM, H.; DONG, W. S. An analysis of the relationship between land use density of office buildings and urban street configuration: Case studies of two areas in Seoul by space syntax analysis. **Cities**, v. 19 n. 6, p. 409–418, 2002.

KITAMURA, R. An Evaluation of Activity-based travel analysis. **Transportation**, v. 15, p. 9–34, 1988.

KURAMOTO, B. **Exploração de dados de mapas colaborativos em avaliações de morfologias urbanas brasileiras**. Dissertação (Mestrado em Ciências: Planejamento e Operação de Sistemas de Transportes). São Carlos: Universidade de São Paulo, Escola de Engenharia de São Carlos; 2019.

KURAMOTO, B; CUNHA, A. L. Usabilidade e limitações de dados de mapas colaborativos em análises de acessibilidade. **33º Congresso de Pesquisa e Ensino Em Transporte Da ANPET**. Camboriú: ANPET, 2019

KUTNER, M. H.; NACHTSHEIM, C. J.; NETER, J.; LI, W. **Applied Linear Statistical Models**. 5ª ed. Nova York: McGraw-Hill/Irwin, 2004.

LEAKE, G. R.; HUZAYYIN, A. S. Importance of accessibility measures in trip production models. **Transport Planning and Technology**, v. 6, p. 9–20, 1980.

LESAGE, J. P. **Spatial Econometrics**. , Department of Economics, University of Toledo, 1999.

LIU, Y.; WEI, X.; JIAO, L.; WANG, H. Relationships between Street Centrality and Land Use Intensity in Wuhan, China. **Journal of Urban Planning and Development**, v. 142, n. 1, p. 1–14, 2016.

LOPES, S. B.; BRONDINO, N. C. M.; RODRIGUES DA SILVA, A. N. GIS-Based Analytical Tools for Transport Planning: Spatial Regression Models for Transportation Demand Forecast. **International Journal of Geo-Information**, v. 3, p. 565–583, 2014.

LOWRY, I., S. **A Model of Metropolis**. Santa Monica, CA. RAND Corporation, 1964.

LUNDBERG, C. G. On the Structuration of Multiactivity Task-environments. **Environmental and Planning A: Economy and Space**, v. 20, n. 12, p. 1603-1621, 1988.

MACHADO, A. A.; CAMBOIM, S. P. Mapeamento colaborativo como fonte de dados para o planejamento urbano: desafios e potencialidades. **Urbe, Rev. Bras. Gest. Urbana**, v. 11, n. e20180142, p. 1–21, 2019.

MARTINEZ, F. MUSSA: Land Use Model for Santiago City. **Transportation Research Record**, v. 1552, n. 1, p. 126–34, 1996.

MCNALLY, M. G.; RINDT, C. R. The Activity-Based Approach. **UC Irvine: Center for Activity Systems Analysis**, 2008.

MIRANDA, H. F. **Mobilidade Urbana Sustentável e o Caso de Curitiba**. 178 p. Dissertação (Mestrado em Ciências: Planejamento e Operação de Sistemas de Transportes). Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos. 2010.

MORAN, P. A. P. Notes on Continuous Stochastic Phenomena. **Biometrika**, v. 37, n. 1/2, p. 17-23, 1950.

MORELLI, A. B. **Análise exploratória de resiliência em redes viárias urbanas**. Dissertação (Mestrado em Ciências: Planejamento e Operação de Sistemas de Transportes). São Carlos: Universidade de São Paulo, Escola de Engenharia de São Carlos, 2019

MORELLI, A. B.; CUNHA, A. L. Verificação de Vulnerabilidades Em Redes de Transporte: Uma Abordagem Pela Teoria Dos Grafos. **33º Congresso de Pesquisa e Ensino Em Transporte Da ANPET**. Camboriú: ANPET, 2019.

MÜLLIGANN, C.; JANOWICZ, K.; YE, M.; LEE, W. C. Analyzing the Spatial-Semantic Interaction of Points of Interest in Volunteered Geographic Information. In: EGENHOFER M., GIUDICE N., MORATZ R., WORBOYS M. (Eds.) **Spatial Information Theory. COSIT 2011**. Berlin: Springer, p. 350-370, 2011.

NEIS, P.; ZIELSTRA, D. Recent Developments and Future Trends in Volunteered Geographic Information Research: The Case of OpenStreetMap. **Future Internet**, v. 6, p. 76–106, 2014.

NEIS, P.; ZIELSTRA, D.; ZIPF, A. The Street Network Evolution of Crowdsourced Maps: OpenStreetMap in Germany 2007–2011. **Future Internet**, v. 4, p. 1–21, 2012.

_____. Comparison of Volunteered Geographic Information Data Contributions and Community Development for Selected World Regions. **Future Internet**, v. 5, p. 282–300, 2013.

OMER, I.; GOLDBLATT, R. Spatial patterns of retail activity and street network structure in new and traditional Israeli cities. **Urban Geography**, v. 37, n. 4, p. 629–649, 2016.

O'REILLY, T. What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. **Communications & Strategies**, n. 1, p. 17-37, 2007.

ORTÚZAR, J. de D.; WILLUMSEN, L. G. **Modelling Transport**. 4. ed. Chichester, UK: John Wiley & Sons, Ltd, 2011.

PADGHAM, M.; RUDIS, B.; LOVELACE, R.; SALMON, M. (2017). osmdata. **Journal of Open Source Software**, v. 2, n. 14, 2017.

PERRY, C. A. **The neighborhood unit: from the Regional survey of New York and its environs**, volume VII, Neighborhood and community planning. London, Routledge/Thoemmes Press, 1998.

PORTA, S.; CRUCITTI, P.; LATORA, V. The network analysis of urban streets: A primal approach. **Environment and Planning B: Planning and Design**, v. 33, n. 5, p. 705–725, 2006a.

_____. The network analysis of urban streets: A dual approach. **Physica A: Statistical Mechanics and its Application**, v. 369, n. 2, p. 853–866, 2006b.

PORTA, S.; LATORA, V.; WANG, F.; RUEDA, S.; STRANO, E.; SCELLATO, S.; CARDILLO, A.; BELLI, E.; CÀRDENAS, F.; CORMENZANA, B.; WANG, F. Street Centrality and the Location of Economic Activities in Barcelona. **Urban Studies**, v. 49, n. 7, p. 1471–1488, 2012.

PUMAR, L.; BARBASTEFANO, R.; CARVALHO, D. Bus transport network analysis in Rio de Janeiro based on topological models using Social Networks. **arXiv**, p. 1–22, 2018. Disponível em: <<https://arxiv.org/abs/1808.05692>>. Acessado em: 12 ago. 2020.

ROSVALL, M.; GRÖNLUND, A.; MINNHAGEN, P.; SNEPPEN, K. Searchability of networks. **Physical Review E - Statistical, Nonlinear, and Soft Matter Physics**, v. 72, n. 4, p. 1–8, 2005.

RUI, Y.; BAN, Y. Exploring the relationship between street centrality and land use in Stockholm. **International Journal of Geographical Information Science**, v. 28, n. 7, p. 1425–1438, 2014.

SASIDHAR, K.; VINEETH, Y.; SUBBARAO, S. Trip attraction rates of commercial land use: a case study. **Indian Journal of Science and Technology**, v. 9 n. 30, p. 1-5. 2016.

SCHEURER, J.; CURTIS, C.; PORTA, S. Spatial Network Analysis of Multimodal Transport Systems: Developing a Strategic Planning Tool to Assess the Congruence of Movement and Urban Structure in Australian Cities. **30th Australasian Transport Research Forum**, January, 2007.

SCOPPA, M. D.; PEPONIS, J. Distributed attraction: The effects of street network connectivity upon the distribution of retail frontage in the City of Buenos Aires. **Environment and Planning B: Planning and Design**, v. 42, n. 2, p. 354–378, 2015.

SERRA, M.; HILLIER, B. Angular and Metric Distance in Road Network Analysis: A nationwide correlation study. **Computers, Environment and Urban Systems**, v. 74, p. 194–207, 2019.

SILVERMAN, B. W. **Density Estimation for Statistics and Data Analysis**. Londres: Chapman and Hall. 1986.

SOUSA, S. F.; NETO, C. R.; FERREIRA, F. F. Structure and robustness of São Paulo public transport network. **arXiv**, p. 1–21, 2018. Disponível em: <<http://arxiv.org/abs/1808.08117>>. Acessado em: 12 ago. 2020.

SZWOCH, G. Combining Road Network Data from OpenStreetMap with an Authoritative Database. **Journal of Transportation Engineering, Part A: Systems**, v. 145, p. 1–27, 2019.

TAYLOR, M. a. P.; D'ESTE, G. M. Concepts of network vulnerability and applications to the identification of critical elements of transport infrastructure. **26th Australasian Transport Research Forum Wellington, New Zealand**, p. 1–15, 2003.

THILL, J.-C.; KIM, M. Trip making, induced travel demand, and accessibility. **Journal of Geographic Systems**, v. 7, p. 229–248, 2005.

TURNER, A. From axial to road-centre lines: a new representation for space syntax and a new model of route choice for transport network analysis. **Environment and Planning B: Planning and Design**, v. 34, n. 3, p. 539–555, 2007.

TURNER, A. The Role of Angularity in Route Choice. In: HORNSBY, K. S. CLARAMUNT, C.; DENIS, M.; LIGOZAT, G. (Eds.). . **Spatial Information Theory. COSIT 2009. Lecture Notes in Computer Science**. Berlin: Springer, Berlin, Heidelberg. p. 489–504, 2009.

VAN OORT, P. A. J. **Spatial data quality: from description to application**. Wageningen University, Holanda, 2006.

VICKERMAN, R. W. Accessibility, Attraction, and Potential: A Review of Some Concepts and Their Use in Determining Mobility. **Environment and Planning A: Economy and Space**, v. 6, n. 6, p. 675–691, 1974.

WANG, C.; QUDDUS, M.; RYLEY, T.; ENOCH, M.; DAVIDSON, L. Spatial models in transport: a review and assessment of methodological issues. **91st TRB Annual Meeting**, n. January, p. 1–17, 2012.

WASSERMAN, S.; FAUST, K. **Social Network Analysis**. Cambridge: Cambridge University Press, 1994.

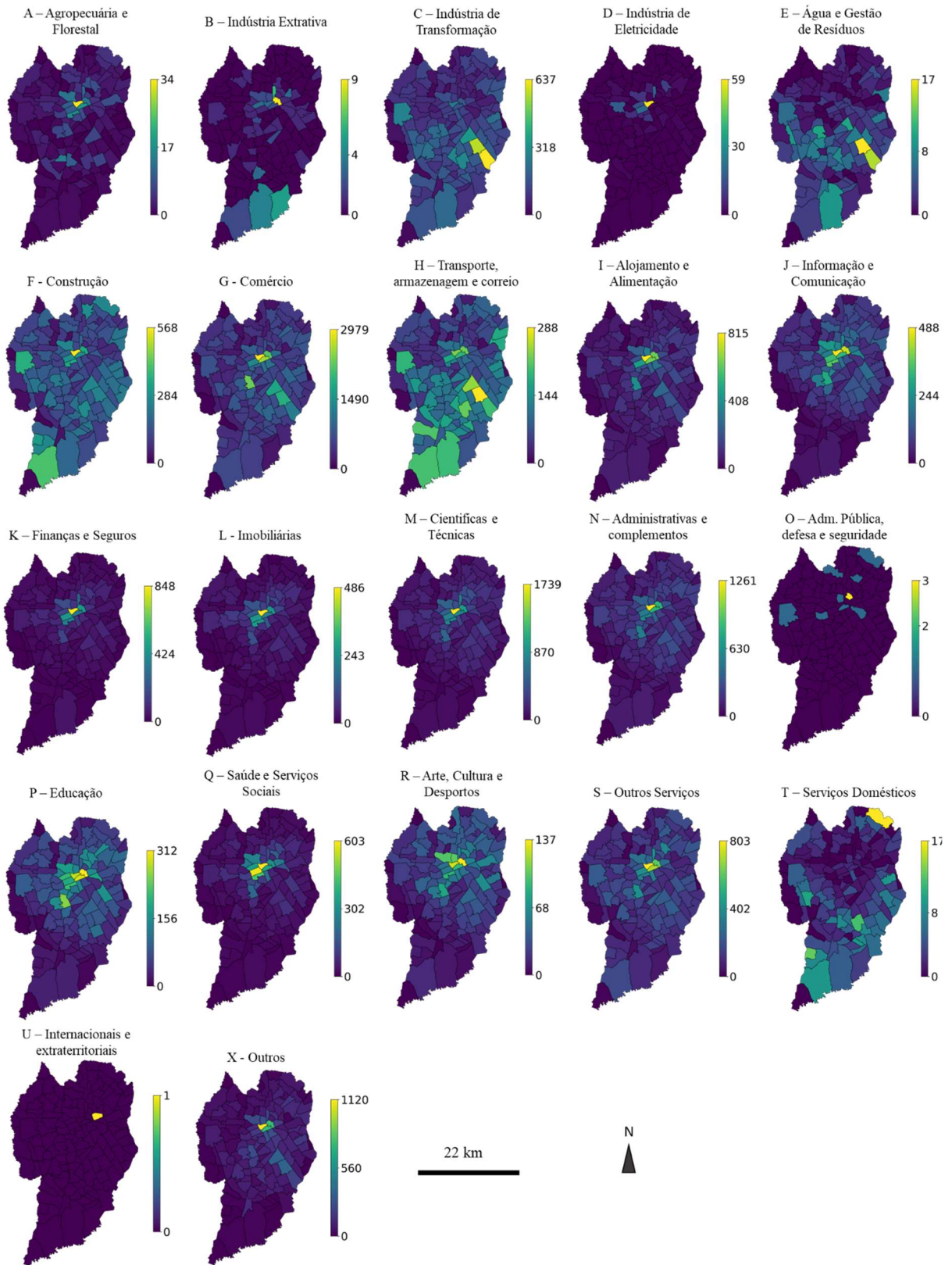
WHITE, H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. **Econometrica**, v. 48, n. 4, p. 817–838, 1980.

XU, L. D.; XU, E. L.; LI, L. Industry 4.0: state of the art and future trends. **International Journal of Production Research**, v. 56, n. 8, p. 2941–2962, 2018.

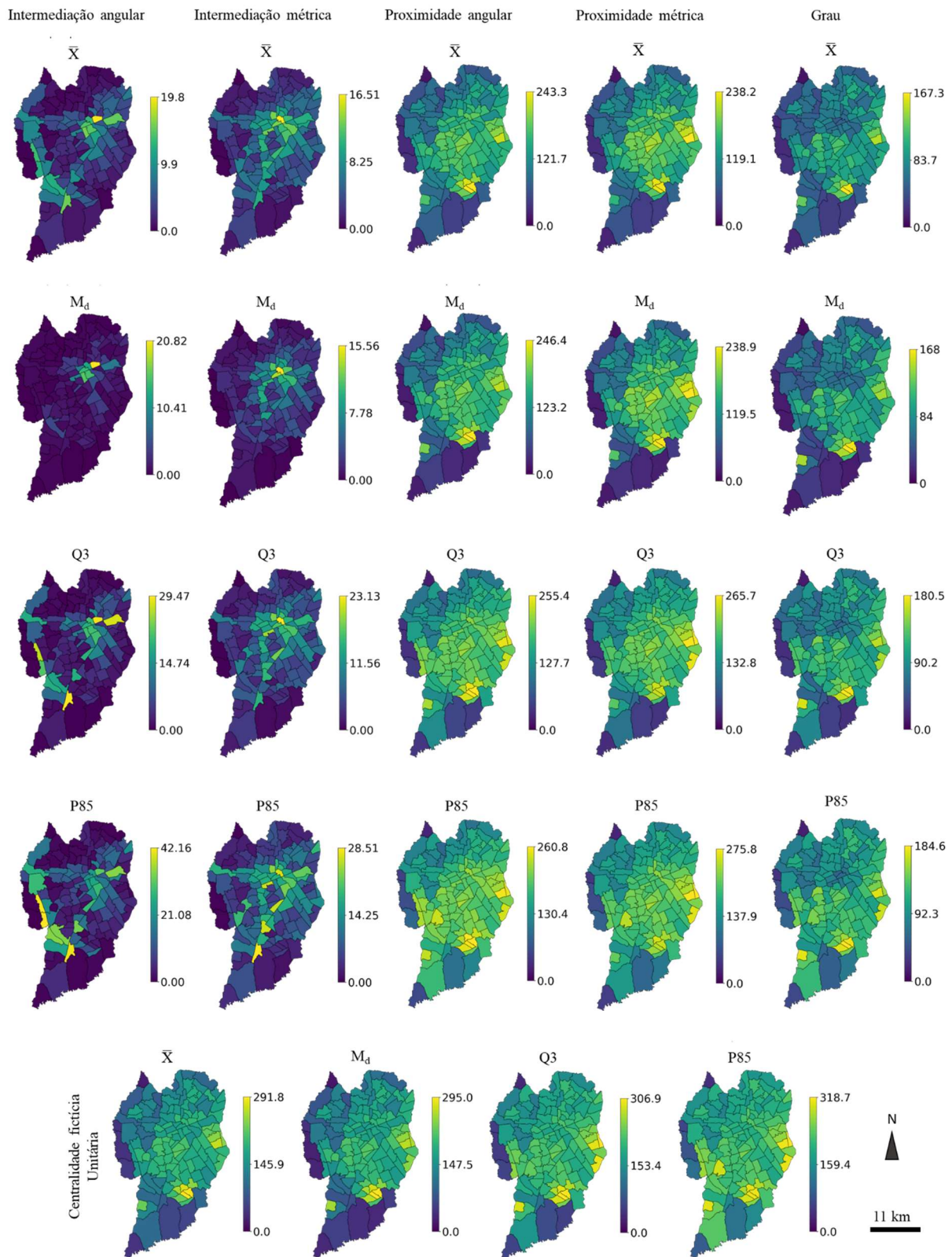
ZHANG, L.; PFOSE, D. Using OpenStreetMap point-of-interest data to model urban change - A feasibility study. **PLoS One**, v. 14, n. 2, p. 1–34, 2019.

APÊNDICES

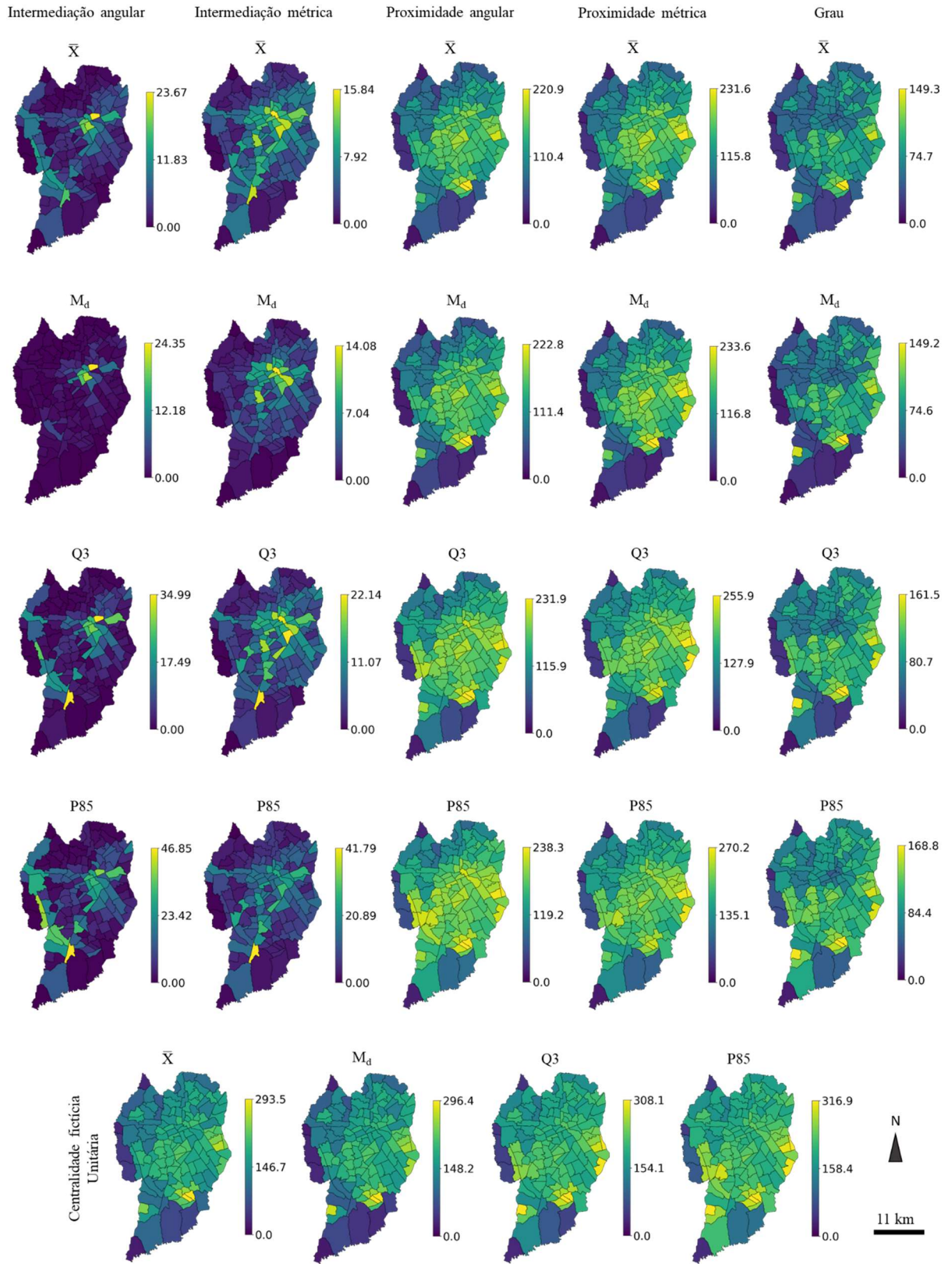
APÊNDICE 1 – NÚMERO DE ALVARÁS POR ZONA



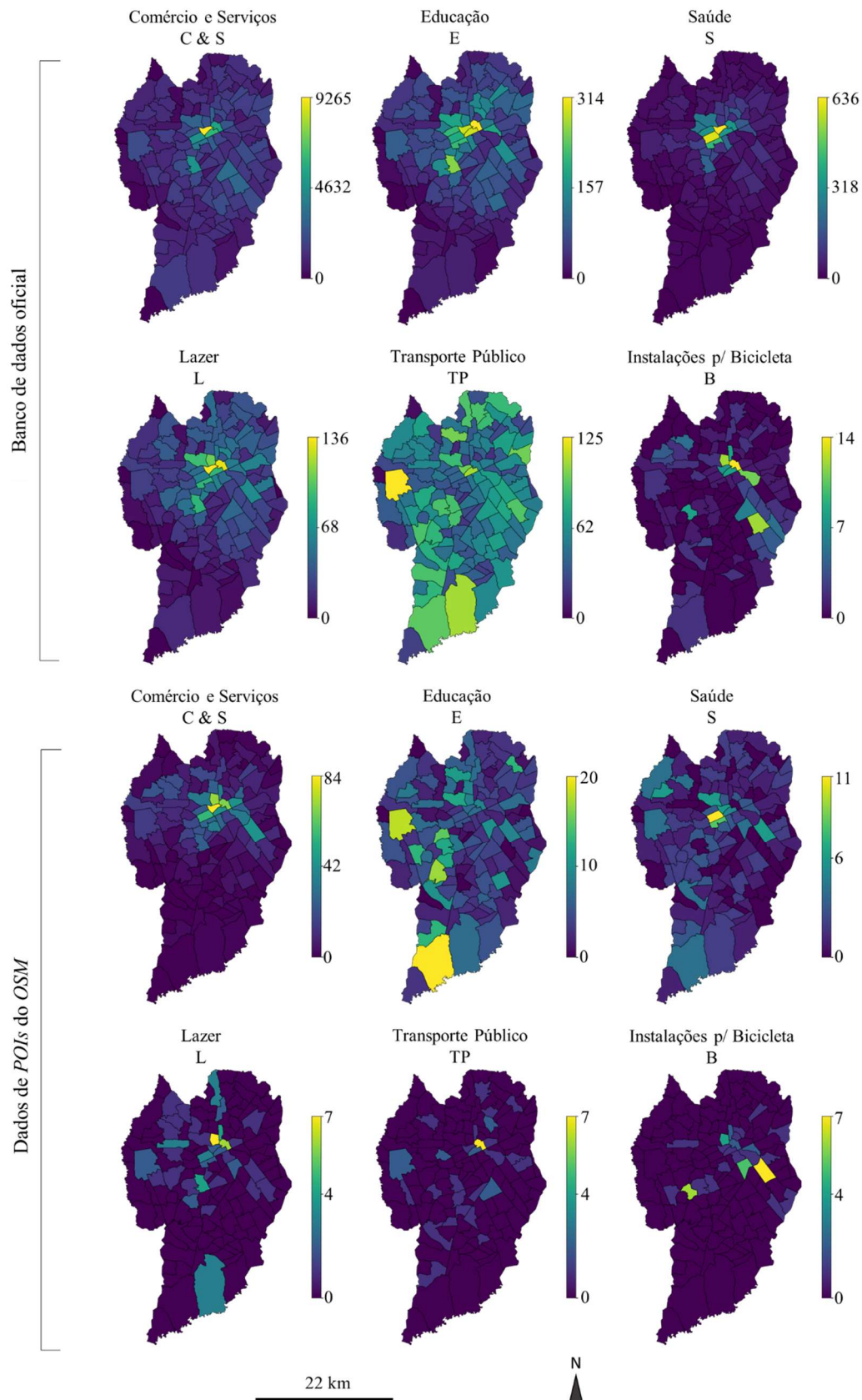
APÊNDICE 2 – ESTATÍSTICAS ZONAIS DE CENTRALIDADES DA REDE OFICIAL



APÊNDICE 3 – ESTATÍSTICAS ZONAIS DE CENTRALIDADES DA REDE DO OSM



APÊNDICE 4 – NÚMERO DE PONTOS POR ZONA DAS VARIÁVEIS TRADICIONAIS OFICIAIS E DO *OSM*



APÊNDICE 5 - RESULTADOS DAS CALIBRAÇÕES DOS MODELOS

		Variável Independente: Atração Outros Motivos						Transformação de Box-Cox: $\lambda = 0,185$				
		Tradicionalis			Centralidades			Tradicionalis + Centralidades				
		Var.	Param.	FIV	Var.	Param.	FIV	Var.	Param.	FIV		
Com dados Oficiais	Variáveis Explicativas	Const	0,396 (0,000)	6,161	Const	0,385 (0,000)	17,685	Const	0,299 (0,000)	21,346		
		C & S.	0,516 (0,000)	3,333	C _{B(m)} M _d	0,380 (0,000)	1,136	C & S.	0,511 (0,000)	1,955		
		TP	0,140 (0,000)	1,171	C _{C(a)} P85.	0,153 (0,004)	1,136	TP	0,173 (0,000)	1,239		
		L	0,185 (0,003)	3,123				C _{C(a)} P85.	0,123 (0,001)	1,142		
								C _{B(m)} M _d	0,175 (0,000)	1,794		
	Ajuste	R ²	0,694			R ²	0,445			R ²	0,755	
		R ² Ajust.	0,687			R ² Ajust.	0,437			R ² Ajust.	0,748	
		F	99,175 (0,000)			F	52,982 (0,000)			F	100,245 (0,000)	
	Testes	Omnibus (Norm.)	10,852 (0,004)	Rejeita	Omnibus (Norm.)	2,472 (0,291)	Aceita	Omnibus (Norm.)	0,379 (0,827)	Aceita		
		White (Homosc.)	56,138 (0,000)	Rejeita	White (Homosc.)	17,574 (0,004)	Rejeita	White (Homosc.)	31,502 (0,005)	Rejeita		
I de Moran (Ind. Esp.)		0,132 (0,007)	Rejeita	I de Moran (Ind. Esp.)	0,003 (0,414)	Aceita	I de Moran (Ind. Esp.)	0,002 (0,408)	Aceita			
Com dados do OSM	Variáveis Explicativas	Const	0,495 (0,000)	1,944	Const	0,371 (0,000)	16,206	Const	0,346 (0,000)	16,523		
		C & S.	0,250 (0,000)	1,604	C _{B(m)} M _d	0,250 (0,000)	1,603	S	0,257 (0,000)	1,419		
		S	0,267 (0,000)	1,410	C _{C(a)} P85.	0,316 (0,004)	3,124	TP	0,213 (0,004)	1,278		
		TP	0,315 (0,000)	1,173	C _D \bar{X}	-0,169 (0,020)	2,528	C & S.	0,146 (0,015)	2,168		
								C _{B(m)} M _d	0,127 (0,008)	2,122		
								C _{C(a)} P85	0,189 (0,000)	1,264		
	Ajuste	R ²	0,483			R ²	0,453			R ²	0,614	
		R ² Ajust.	0,472			R ² Ajust.	0,441			R ² Ajust.	0,599	
		F	40,883 (0,000)			F	36,197 (0,000)			F	41,126 (0,000)	
	Testes	Omnibus (Norm.)	9,798 (0,007)	Rejeita	Omnibus (Norm.)	1,665 (0,435)	Aceita	Omnibus (Norm.)	0,640 (0,726)	Aceita		
White (Homosc.)		10,729 (0,295)	Aceita	White (Homosc.)	14,486 (0,106)	Aceita	White (Homosc.)	20,256 (0,442)	Aceita			
I de Moran (Ind. Esp.)		0,146 (0,005)	Rejeita	I de Moran (Ind. Esp.)	0,027 (0,257)	Aceita	I de Moran (Ind. Esp.)	-0,010 (0,497)	Aceita			

Variável Independente: Atração Trabalho-Estudo			Transformação de Box-Cox: $\lambda = 0,216$							
Tradicionais			Centralidades			Tradicionais + Centralidades				
Var.	Param.	FIV	Var.	Param.	FIV	Var.	Param.	FIV		
Com dados Oficiais	Variáveis Explicativas	Const	0,390 (0,000)	6,083	Const	0,367 (0,000)	18,335	Const	0,254 (0,000)	21,549
		C & S.	0,584 (0,000)	1,485	C _{B(m)} M _d	0,274 (0,000)	1,371	C & S.	0,448 (0,002)	1,670
		TP	0,251 (0,000)	1,171	C _{c(a)} P85.	0,498 (0,000)	3,421	TP	0,270 (0,000)	1,297
		B	0,116 (0,015)	1,330	C _D \bar{X}	-0,356 (0,000)	3,012	C _{B(m)} Q3	0,123 (0,004)	1,862
								C _{c(a)} P85	0,325 (0,000)	3,821
							C _D \bar{X}	-0,205 (0,003)	3,383	
	Ajuste	R ²	0,635		R ²	0,429		R ²	0,729	
		R ² Ajust.	0,627		R ² Ajust.	0,416		R ² Ajust.	0,718	
		F	76,153 (0,000)		F	32,845 (0,000)		F	69,373 (0,000)	
	Testes	Omnibus (Norm.)	4,545 (0,103)	Aceita	Omnibus (Norm.)	3,467 (0,177)	Aceita	Omnibus (Norm.)	0,771 (0,680)	Aceita
White (Homosc.)		27,32 (0,001)	Rejeita	White (Homosc.)	10,619 (0,303)	Aceita	White (Homosc.)	23,752 (0,253)	Aceita	
I de Moran (Ind. Esp.)		0,068 (0,082)	Aceita	I de Moran (Ind. Esp.)	-0,004 (0,459)	Aceita	I de Moran (Ind. Esp.)	0,027 (0,250)	Aceita	
Com dados do OSM	Variáveis Explicativas	Const	0,489 (0,000)	2,691	Const	0,371 (0,000)	16,206	Const	0,330 (0,000)	16,803
		C & S.	0,219 (0,001)	1,617	C _{B(m)} M _d	0,186 (0,001)	1,603	E	0,192 (0,000)	1,249
		E	0,172 (0,002)	1,255	C _{c(a)} P85.	0,492 (0,000)	3,124	S	0,195 (0,001)	1,367
		TP	0,297 (0,001)	1,202	C _D \bar{X}	-0,340 (0,000)	2,528	C _{B(m)} M _d	0,177 (0,000)	1,649
		S	0,225 (0,004)	1,692				C _{c(a)} P85	0,390 (0,000)	3,243
							C _D \bar{X}	-0,249 (0,001)	2,647	
	Ajuste	R ²	0,451		R ²	0,448		R ²	0,597	
		R ² Ajust.	0,434		R ² Ajust.	0,435		R ² Ajust.	0,582	
		F	26,702 (0,000)		F	35,43 (0,000)		F	38,297 (0,000)	
	Testes	Omnibus (Norm.)	5,358 (0,069)	Aceita	Omnibus (Norm.)	3,504 (0,173)	Aceita	Omnibus (Norm.)	4,475 (0,107)	Aceita
White (Homosc.)		7,136 (0,929)	Aceita	White (Homosc.)	9,252 (0,414)	Aceita	White (Homosc.)	16,562 (0,681)	Aceita	
I de Moran (Ind. Esp.)		0,116 (0,019)	Rejeita	I de Moran (Ind. Esp.)	0,001 (0,417)	Aceita	I de Moran (Ind. Esp.)	0,015 (0,326)	Aceita	

Variável Independente: Atração Não-Motorizadas			Transformação de Box-Cox: $\lambda = 0,330$					
Tradicionais			Centralidades			Tradicionais + Centralidades		
Var.	Param.	FIV	Var.	Param.	FIV	Var.	Param.	FIV
Com dados Oficiais								
Variáveis Explicativas								
Const	0,317 (0,000)	6,259	Const	0,213 (0,000)	18,941	Const	0,073 (0,067)	22,071
TP	0,359 (0,000)	1,193	Cc(a) P85.	1,574 (0,000)	16,753	TP	0,225 (0,000)	1,577
C & S.	1,130 (0,000)	4,025	C _D \bar{X}	-0,460 (0,000)	2,608	C & S.	0,987 (0,000)	5,422
E	-0,444 (0,000)	3,684	Cc(m) P85.	-0,795 (0,012)	17,319	S	-0,349 (0,002)	4,294
						Cc(a) P85	0,815 (0,000)	3,878
						Cc(m) \bar{X}	-0,508 (0,000)	4,295
Ajuste								
R ²	0,531		R ²	0,360		R ²	0,686	
R ² Ajust.	0,520		R ² Ajust.	0,345		R ² Ajust.	0,674	
F	49,474 (0,000)		F	24,536 (0,000)		F	56,390 (0,000)	
Testes								
Omnibus (Norm.)	0,871 (0,647)	Aceita	Omnibus (Norm.)	2,370 (0,306)	Aceita	Omnibus (Norm.)	0,334 (0,846)	Aceita
White (Homosc.)	29,754 (0,000)	Rejeita	White (Homosc.)	7,335 (0,602)	Aceita	White (Homosc.)	10,631 (0,955)	Aceita
I de Moran (Ind. Esp.)	0,109 (0,014)	Rejeita	I de Moran (Ind. Esp.)	0,022 (0,269)	Aceita	I de Moran (Ind. Esp.)	0,050 (0,138)	Aceita
Com dados do OSM								
Variáveis Explicativas								
Const	0,436 (0,000)	2,658	Const	0,205 (0,000)	16,206	Const	0,181 (0,000)	16,740
E	0,265 (0,001)	1,246	Cc(a) P85.	0,939 (0,000)	3,735	E	0,195 (0,000)	1,294
TP	0,419 (0,001)	1,019	Cc(m) \bar{X}	-0,528 (0,000)	3,735	TP	0,333 (0,001)	1,045
S	0,223 (0,016)	1,225				S	0,187 (0,018)	1,233
						Cc(a) P85	0,771 (0,000)	3,972
						Cc(m) \bar{X}	-0,431 (0,000)	3,895
Ajuste								
R ²	0,275		R ²	0,322		R ²	0,480	
R ² Ajust.	0,258		R ² Ajust.	0,312		R ² Ajust.	0,460	
F	16,566 (0,000)		F	31,418 (0,000)		F	23,802 (0,000)	
Testes								
Omnibus (Norm.)	0,326 (0,849)	Aceita	Omnibus (Norm.)	4,434 (0,109)	Aceita	Omnibus (Norm.)	1,170 (0,557)	Aceita
White (Homosc.)	9,334 (0,407)	Aceita	White (Homosc.)	5,719 (0,335)	Aceita	White (Homosc.)	20,074 (0,453)	Aceita
I de Moran (Ind. Esp.)	0,191 (0,001)	Rejeita	I de Moran (Ind. Esp.)	0,080 (0,059)	Aceita	I de Moran (Ind. Esp.)	0,055 (0,128)	Aceita

Variável Independente: Atração Públicas e Semi-Públicas			Transformação de Box-Cox: $\lambda = 0,117$							
Tradicionais			Centralidades			Tradicionais + Centralidades				
Var.	Param.	FIV	Var.	Param.	FIV	Var.	Param.	FIV		
Com dados Oficiais	Variáveis Explicativas	Const	0,487 (0,000)	6,161	Const	0,573 (0,000)	2,346	Const	0,447 (0,000)	7,960
		C & S.	0,306 (0,000)	3,333	C _{B(m)} M _d	0,388 (0,000)	1,000	C & S.	0,254 (0,001)	3,379
		TP	0,177 (0,000)	1,171		TP	0,217 (0,000)	1,234		
		L	0,239 (0,000)	3,123		L	0,131 (0,012)	3,530		
						C _{B(m)} Q3	0,185 (0,000)	1,605		
	Ajuste	R ²	0,689		R ²	0,427		R ²	0,760	
		R ² Ajust.	0,682		R ² Ajust.	0,423		R ² Ajust.	0,753	
		F	96,804 (0,000)		F	99,195 (0,000)		F	102,962 (0,000)	
	Testes	Omnibus (Norm.)	2,063 (0,356)	Aceita	Omnibus (Norm.)	6,951 (0,031)	Rejeita	Omnibus (Norm.)	0,471 (0,790)	Aceita
White (Homosc.)		36,431 (0,000)	Rejeita	White (Homosc.)	0,621 (0,733)	Aceita	White (Homosc.)	29,488 (0,009)	Rejeita	
I de Moran (Ind. Esp.)		0,211 (0,001)	Rejeita	I de Moran (Ind. Esp.)	0,003 (0,400)	Aceita	I de Moran (Ind. Esp.)	0,156 (0,003)	Rejeita	
Com dados do OSM	Variáveis Explicativas	Const	0,575 (0,000)	2,691	Const	0,570 (0,000)	2,479	Const	0,488 (0,000)	17,330
		C & S.	0,222 (0,000)	1,617	C _{B(m)} M _d	0,341 (0,000)	1,000	S	0,198 (0,000)	1,448
		E	0,086 (0,040)	1,255		E	0,114 (0,002)	1,324		
		TP	0,263 (0,000)	1,202		TP	0,147 (0,021)	1,308		
		S	0,197 (0,001)	1,692		C _{B(m)} M _d	0,147 (0,002)	2,581		
				C _{c(a)} P85		0,265 (0,001)	5,241			
	Ajuste	R ²	0,500		R ²	0,412		R ²	0,635	
		R ² Ajust.	0,484		R ² Ajust.	0,408		R ² Ajust.	0,618	
		F	32,457 (0,000)		F	93,344 (0,000)		F	37,161 (0,000)	
	Testes	Omnibus (Norm.)	6,789 (0,034)	Rejeita	Omnibus (Norm.)	5,555 (0,062)	Aceita	Omnibus (Norm.)	1,922 (0,382)	Aceita
White (Homosc.)		9,903 (0,769)	Aceita	White (Homosc.)	0,816 (0,665)	Aceita	White (Homosc.)	33,212 (0,190)	Aceita	
I de Moran (Ind. Esp.)		0,173 (0,003)	Rejeita	I de Moran (Ind. Esp.)	0,013 (0,337)	Aceita	I de Moran (Ind. Esp.)	0,064 (0,092)	Aceita	

Variável Independente: Atração todas as viagens			Transformação de Box-Cox: $\lambda = 0,213$							
Tradicionais			Centralidades			Tradicionais + Centralidades				
Var.	Param.	FIV	Var.	Param.	FIV	Var.	Param.	FIV		
Com dados Oficiais	Variáveis Explicativas	Const	0,382 (0,000)	6,083	Const	0,355 (0,000)	18,335	Const	0,250 (0,000)	21,549
		C & S.	0,616 (0,000)	1,485	C _{B(m)} M _d	0,295 (0,000)	1,371	C & S.	0,484 (0,000)	1,670
		TP	0,219 (0,000)	1,171	C _{c(a)} P85.	0,450 (0,000)	3,421	TP	0,242 (0,000)	1,297
		B	0,112 (0,010)	1,330	C _D \bar{X}	-0,309 (0,000)	3,012	C _{B(m)} Q3	0,126 (0,001)	1,862
								C _{c(a)} P85	0,287 (0,000)	3,821
							C _D \bar{X}	-0,167 (0,007)	3,383	
	Ajuste	R ²	0,673		R ²	0,453		R ²	0,761	
		R ² Ajust.	0,666		R ² Ajust.	0,440		R ² Ajust.	0,752	
		F	89,969 (0,000)		F	36,216 (0,000)		F	82,489 (0,000)	
	Testes	Omnibus (Norm.)	6,094 (0,047)	Rejeita	Omnibus (Norm.)	3,266 (0,195)	Aceita	Omnibus (Norm.)	0,342 (0,842)	Aceita
White (Homosc.)		34,812 (0,000)	Rejeita	White (Homosc.)	12,664 (0,178)	Aceita	White (Homosc.)	27,664 (0,118)	Aceita	
I de Moran (Ind. Esp.)		0,080 (0,049)	Rejeita	I de Moran (Ind. Esp.)	-0,004 (0,476)	Aceita	I de Moran (Ind. Esp.)	0,030 (0,230)	Aceita	
Com dados do OSM	Variáveis Explicativas	Const	0,477 (0,000)	2,691	Const	0,359 (0,000)	16,206	Const	0,322 (0,000)	16,825
		C & S.	0,229 (0,000)	1,617	C _{B(m)} M _d	0,206 (0,000)	1,603	S	0,240 (0,000)	1,406
		E	0,139 (0,007)	1,255	C _{c(a)} P85.	0,447 (0,000)	3,124	E	0,141 (0,002)	1,290
		TP	0,307 (0,000)	1,202	C _D \bar{X}	-0,296 (0,000)	2,528	TP	0,191 (0,012)	1,293
		S	0,235 (0,001)	1,692				C _{B(m)} M _d	0,150 (0,002)	1,896
							C _{c(a)} P85	0,340 (0,000)	3,250	
							C _{c(m)} \bar{X}	-0,190 (0,004)	2,671	
	Ajuste	R ²	0,479		R ²	0,464		R ²	0,630	
		R ² Ajust.	0,463		R ² Ajust.	0,452		R ² Ajust.	0,612	
		F	29,904 (0,000)		F	37,810 (0,000)		F	36,289 (0,000)	
Testes	Omnibus (Norm.)	6,559 (0,038)	Rejeita	Omnibus (Norm.)	3,039 (0,219)	Aceita	Omnibus (Norm.)	3,188 (0,203)	Aceita	
	White (Homosc.)	8,429 (0,866)	Aceita	White (Homosc.)	10,794 (0,290)	Aceita	White (Homosc.)	20,312 (0,817)	Aceita	
	I de Moran (Ind. Esp.)	0,132 (0,012)	Rejeita	I de Moran (Ind. Esp.)	0,004 (0,391)	Aceita	I de Moran (Ind. Esp.)	0,008 (0,370)	Aceita	

