

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

Agrupamento de municípios paulistas por densidade territorial urbana

Isabel Cristina Nunes de Sousa

Trabalho de Conclusão de Curso - MBA em Ciência de Dados (CEMEAI)

Isabel Cristina Nunes de Sousa

Agrupamento de municípios paulistas por densidade territorial urbana

Trabalho de conclusão de curso apresentado ao Centro de Ciências Matemáticas Aplicadas à Indústria do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para conclusão do MBA em Ciência de Dados

Área de concentração: Ciências de Dados

Orientador: Prof. Dr. Wallace Correa de Oliveira Casaca

Versão original

**São Carlos
2024**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

S725a Sousa, Isabel Cristina Nunes de
Agrupamento de municípios paulistas por densidade
territorial urbana / Isabel Cristina Nunes de
Sousa; orientador Wallace Correa de Oliveira
Casaca. -- São Carlos, 2024.
121 p.

Trabalho de conclusão de curso (MBA em Ciência
de Dados) -- Instituto de Ciências Matemáticas e de
Computação, Universidade de São Paulo, 2024.

1. Crescimento urbano. 2. Análise geoespacial.
3. Aprendizado não supervisionado. 4. IBGE. 5.
MapBiomass. I. Casaca, Wallace Correa de Oliveira,
orient. II. Título.

Isabel Cristina Nunes de Sousa

Clustering of São Paulo municipalities by urban land density

Term paper submitted to the Centro de Ciências Matemáticas Aplicadas à Indústria of the Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, in partial fulfillment of the requirements for the conclusion of the MBA in Data Science

Advisor: Prof. Dr. Wallace Correa de Oliveira Casaca

Original version

**São Carlos
2024**

“Computer science inverts the normal. In normal science, you’re given a world, and your job is to find out the rules. In computer science, you give the computer the rules, and it creates the world.”

Alan Kay

RESUMO

SOUZA, I. C. N. *Agrupamento de municípios paulistas por densidade territorial urbana.* 2024. 121p. Monografia (Trabalho de Conclusão de Curso) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2024.

Neste trabalho de conclusão de curso, exploramos a dinâmica do crescimento urbano em municípios paulistas de porte médio. Nosso objetivo era identificar padrões que pudessem ajudar a compreender os processos de expansão urbana, assim como a evolução da forma urbana das cidades. Para tanto, empregamos um método quantitativo em Python, utilizando uma função sigmoide em forma de S-invertido — uma ferramenta amplamente utilizada para caracterizar o declínio na densidade territorial urbana. Em seguida, examinamos as variações na densidade territorial urbana entre os municípios, empregando técnicas de aprendizado não supervisionado para agrupamento de dados. Identificamos padrões distintos de crescimento urbano nos municípios da amostra. Isso indica a viabilidade de reconhecer esses padrões por meio dos agrupamentos definidos, oferecendo informações relevantes sobre a dinâmica do crescimento urbano. Vale ressaltar que a adequação dos algoritmos variou em cada intervalo temporal, enfatizando a importância de testar diferentes métodos de agrupamento.

Palavras-chave: Crescimento urbano. Análise geoespacial. Aprendizado não supervisionado. IBGE. Mapbiomas.

ABSTRACT

SOUZA, I. C. N. **Clustering of São Paulo municipalities by urban land density.** 2024. 121p. Monograph (Conclusion Course Paper) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2024.

In this conclusion course paper, we explore the dynamics of urban growth in medium-sized municipalities in the State of São Paulo. Our aim was to identify patterns that could enhance our understanding of urban expansion processes and the evolving urban form of cities. To achieve this, we employed a quantitative method in Python, using an inverse S-shape sigmoid function — a widely used tool for characterizing the decline in urban land density. Following this, we examined variations in urban land density across municipalities, employing unsupervised learning techniques for data clustering. Our analysis uncovered distinct urban growth patterns within the sampled municipalities. This indicates the feasibility of identifying these patterns through the defined groupings, offering valuable insights into the dynamics of urban growth. It is noteworthy that the suitability of algorithms varied across time intervals, emphasizing the importance of testing different clustering methods.

Keywords: Urban growth. Geospatial analysis. Unsupervised learning. IBGE. Mapbiomas.

LISTA DE FIGURAS

Figura 1 – Gráfico do modelo da função sigmoide S-invertido.	43
Figura 2 – Interpretação do gráfico do modelo da função sigmoide S-invertido.	44
Figura 3 – Curvas da função sigmoide S-invertido com diferentes valores de alpha.	46
Figura 4 – Diagrama dos métodos empregados.	50
Figura 5 – Localização dos 15 municípios da amostra no Estado de São Paulo.	52
Figura 6 – <i>Box plots</i> da População residente, Área territorial e Densidade demográfica dos 15 municípios da amostra.	53
Figura 7 – Gráficos das curvas ajustadas para (a) Araçatuba, (b) Araraquara e (c) Assis.	55
Figura 8 – Gráficos das curvas ajustadas para (a) Barretos, (b) Bauru e (c) Birigui.	56
Figura 9 – Gráficos das curvas ajustadas para (a) Botucatu, (b) Catanduva e (c) Franca.	57
Figura 10 – Gráficos das curvas ajustadas para (a) Jaú, (b) Marília e (c) Ourinhos.	58
Figura 11 – Gráficos das curvas ajustadas para (a) Presidente Prudente, (b) São Carlos e (c) São José do Rio Preto.	59
Figura 12 – <i>Box plots</i> da variação do (a) α , (b) c e (c) D	61
Figura 13 – <i>Box plots</i> da variação do (a) r_1 , (b) r_2 e (c) f_1	62
Figura 14 – <i>Box plots</i> da variação do (a) f_2 , (b) k_s e (c) k_p	63
Figura 15 – Matrizes de correlação para (a) 1985 e (b) 1990.	64
Figura 16 – Matrizes de correlação para (a) 1995 e (b) 2000.	65
Figura 17 – Matrizes de correlação para (a) 2005 e (b) 2010.	66
Figura 18 – Matrizes de correlação para (a) 2015 e (b) 2020.	67
Figura 19 – Relação entre S_r e k_{ps} para os intervalos de (a) 1985-1990 e (b) 1990-1995.	69
Figura 20 – Relação entre S_r e k_{ps} para os intervalos de (a) 1995-2000 e (b) 2000-2005.	70
Figura 21 – Relação entre S_r e k_{ps} para os intervalos de (a) 2005-2010 e (b) 2010-2015.	71
Figura 22 – Relação entre S_r e k_{ps} para o intervalo de 2015-2020.	72
Figura 23 – Gráficos de densidade alisada do S_r para cada intervalo (a) com e (b) sem valores discrepantes.	74
Figura 24 – <i>Clusters</i> por algoritmo (1985-1990).	75
Figura 25 – <i>Clusters</i> por algoritmo (1990-1995).	76
Figura 26 – <i>Clusters</i> por algoritmo (1995-2000).	77
Figura 27 – <i>Clusters</i> por algoritmo (2000-2005).	78
Figura 28 – <i>Clusters</i> por algoritmo (2005-2010).	79
Figura 29 – <i>Clusters</i> por algoritmo (2010-2015).	80
Figura 30 – <i>Clusters</i> por algoritmo (2015-2020).	81

LISTA DE TABELAS

Tabela 1 – População residente, Área territorial e Densidade demográfica dos 15 municípios da amostra.	51
Tabela 2 – Estatísticas descritivas da População residente, Área territorial e Densidade demográfica dos 15 municípios da amostra.	53
Tabela 3 – Valores do S_r (<i>Sprawl ratio</i>) e do k_{ps} (<i>degree of sprawling</i>) por intervalos.	68
Tabela 4 – Estatísticas descritivas do S_r com e sem valores discrepantes para cada intervalo.	73
Tabela 5 – Integrantes por <i>clusters</i> para cada intervalo.	82
Tabela 6 – Municípios e respectivos <i>clusters</i> por algoritmo para os intervalos 1985-1990 e 1990-1995.	83
Tabela 7 – Municípios e respectivos <i>clusters</i> por algoritmo para os intervalos 1995-2000 e 2000-2005.	84
Tabela 8 – Municípios e respectivos <i>clusters</i> por algoritmo para os intervalos de 2005-2010 e 2010-2015.	85
Tabela 9 – Municípios e respectivos <i>clusters</i> por algoritmo para o intervalo 2015-2020.	86
Tabela 10 – Valores das métricas para cada algoritmo, por intervalos.	86
Tabela 11 – Densidade territorial urbana por raios para 1985 e 1990.	99
Tabela 12 – Densidade territorial urbana por raios para 1995 e 2000.	100
Tabela 13 – Densidade territorial urbana por raios para 2005 e 2010.	101
Tabela 14 – Densidade territorial urbana por raios para 2015 e 2020.	102
Tabela 15 – Parâmetros do ajuste da função sigmoide de densidade territorial urbana para 1985 e 1990.	103
Tabela 16 – Parâmetros do ajuste da função sigmoide de densidade territorial urbana para 1995 e 2000.	104
Tabela 17 – Parâmetros do ajuste da função sigmoide de densidade territorial urbana para 2005 e 2010.	105
Tabela 18 – Parâmetros do ajuste da função sigmoide de densidade territorial urbana para 2015 e 2020.	106
Tabela 19 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 1985.	107
Tabela 20 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 1990.	108
Tabela 21 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 1995.	108
Tabela 22 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 2000.	109

Tabela 23 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 2005.	109
Tabela 24 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 2010.	110
Tabela 25 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 2015.	110
Tabela 26 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 2020.	111
Tabela 27 – Estatísticas descritivas para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (α).	113
Tabela 28 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (α).	113
Tabela 29 – Estatísticas descritivas para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (c).	114
Tabela 30 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (c).	114
Tabela 31 – Estatísticas descritivas para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (D).	114
Tabela 32 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (D).	115
Tabela 33 – Estatísticas descritivas para as coordenadas dos pontos que caracterizam a forma das áreas urbanizadas por ano (f_1).	115
Tabela 34 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (f_1).	115
Tabela 35 – Estatísticas descritivas para as coordenadas dos pontos que caracterizam a forma das áreas urbanizadas por ano (f_2).	116
Tabela 36 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (f_2).	116
Tabela 37 – Estatísticas descritivas para as coordenadas dos pontos que caracterizam a forma das áreas urbanizadas por ano (r_1).	116
Tabela 38 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (r_1).	117
Tabela 39 – Estatísticas descritivas para as coordenadas dos pontos que caracterizam a forma das áreas urbanizadas por ano (r_2).	117
Tabela 40 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (r_2).	117
Tabela 41 – Estatísticas descritivas para as medidas de compacidade urbana por ano (k_s).	118

Tabela 42 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (k_s).	118
Tabela 43 – Estatísticas descritivas para as medidas de compacidade urbana por ano (k_p).	118
Tabela 44 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (k_p).	119
Tabela 45 – Taxas de crescimento de r_1 (δr_1) e r_2 (δr_2) por intervalos.	121

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Contextualização	21
1.2	Objetivos	22
1.3	Estrutura da Monografia	23
2	REFERENCIAL TEÓRICO	25
2.1	Função de densidade territorial urbana	25
2.2	Técnicas de agrupamento	27
2.2.1	<i>K-Means</i>	28
2.2.2	<i>K-Medoids</i>	30
2.2.3	<i>Mean Shift</i>	31
2.2.4	<i>Affinity Propagation</i>	32
2.3	Validação e interpretação dos <i>clusters</i>	33
2.3.1	Silhueta	34
2.3.2	Calinski–Harabasz	36
2.3.3	Davies–Bouldin	36
3	MÉTODOS	39
3.1	Base de Dados	39
3.2	Etapas da Pesquisa	40
3.2.1	Etapa 1: Coleta e tratamento dos dados	40
3.2.2	Etapa 2: Implementação da função sigmoide	41
3.2.3	Etapa 3: Análises de agrupamento	47
4	RESULTADOS	51
4.1	Caracterização da amostra	51
4.2	Ajuste da função sigmoide	52
4.3	Agrupamentos dos municípios	64
5	CONCLUSÃO	89
	REFERÊNCIAS	91

APÊNDICES	97
APÊNDICE A – DENSIDADES TERRITORIAIS URBANAS PARA CADA MUNICÍPIO, ANUAIS E POR RAIOS	99
APÊNDICE B – PARÂMETROS DO AJUSTE DA FUNÇÃO DE DENSIDADE TERRITORIAL URBANA	103
APÊNDICE C – COORDENADAS DOS PONTOS QUE CARACTERIZAM A FORMA DAS ÁREAS URBANIZADAS E MEDIDAS DE COMPACIDADE URBANA	107
APÊNDICE D – ESTATÍSTICAS DESCRIPTIVAS ANUAIS PARA OS PARÂMETROS DE AJUSTE DA FUNÇÃO DE DENSIDADE TERRITORIAL URBANA, DAS COORDENADAS DOS PONTOS QUE CARACTERIZAM A FORMA DAS ÁREAS URBANIZADAS E DAS MEDIDAS DE COMPACIDADE URBANA	113
APÊNDICE E – TAXAS DE CRESCIMENTO DE r_1 E r_2 PARA CADA INTERVALO	121

1 INTRODUÇÃO

Neste capítulo consta a introdução deste trabalho de conclusão de curso, cujo objetivo é identificar padrões de crescimento urbano em municípios paulistas de porte médio, via técnicas de aprendizado não supervisionado. A seção 1.1 abrange a contextualização do trabalho e a justificativa para seu desenvolvimento. Os objetivos gerais e específicos encontram-se na seção 1.2. Por fim, na seção 1.3 descrevemos a estrutura desta monografia.

1.1 Contextualização

O crescimento urbano pode ser definido como a expansão física das áreas urbanas (MAHTTA *et al.*, 2022). A análise da dinâmica do crescimento urbano ao longo do tempo, possibilita compreender os processos de expansão urbana, assim como as transformações na forma urbana das cidades. Isso é fundamental para identificar padrões e orientar o desenvolvimento de políticas públicas.

Com o advento do sensoriamento remoto, ampliaram-se as oportunidades de mapeamento e monitoramento da dinâmica espaço-temporal do crescimento urbano (REBA; SETO, 2020).

Nesse sentido, valendo-se de imagens obtidas via sensoriamento remoto, Jiao (2015) propôs uma função não linear de densidade territorial urbana (ou densidade construída) em formato de S-invertido (*inverse S-shaped*). Essa adaptação da função sigmoide descreve a diminuição da densidade consoante aumento da distância do centro da cidade, e possibilita caracterizar quantitativamente o crescimento urbano (LU *et al.*, 2023), viabilizando avaliações comparativas espaço-temporais. Dentre os aspectos que podem ser descritos pela função, a compacidade urbana se destaca.

Inicialmente desenvolvido na década de 1970, o conceito de cidade compacta se disseminou, especialmente em países desenvolvidos, à medida que se ampliava a compreensão das consequências sociais e ambientais adversas do chamado *urban sprawl* - espraiamento urbano, em tradução livre (ANABTAWI, 2023). O termo *urban sprawl* se originou na literatura científica anglo-saxônica, e relaciona-se com a expansão dos subúrbios norte-americanos após a 2ª Guerra Mundial, devido ao aumento no uso de automóveis e pela expansão das autoestradas interestaduais (SILVA; VERGARA-PERUCICH, 2021). Entretanto, desde a década de 1960 esse fenômeno não se restringe apenas à América do Norte e à Europa Ocidental, tendo se espalhado por cidades de países periféricos (AHANI; DADASHPOOR, 2021), assim como o conceito de cidade compacta, atualmente considerado uma abordagem universal (ANABTAWI, 2023).

Sob uma perspectiva socioambiental, *urban sprawl* refere-se a uma forma indesejável

de expansão territorial urbana (AMPONSAH *et al.*, 2022), pois envolve crescimento urbano disperso e de baixa densidade (POURTAHERIAN; JAEGER, 2022). Consequentemente, são pesquisadas alternativas mais sustentáveis de crescimento territorial urbano, com o modelo de cidade compacta sendo um dos mais explorados na literatura científica (ANABTAWI, 2023), dada sua associação com o uso mais eficiente do solo pela contenção da expansão territorial urbana e intensificação no uso e ocupação da terra (YAO *et al.*, 2022).

A definição do que constituiria uma mais cidade compacta, no entanto, ainda é debatida. Por meio de uma revisão bibliográfica, Anabtawi (2023) distinguiu cinco elementos que unificam a maioria das discussões ontológicas realizadas recentemente: Densidade, Diversidade (de usos do solo), Acesso (mobilidade/transporte), Forma (urbana) e Espaços Verdes (ANABTAWI, 2023). Com base em tal distinção, neste estudo nos concentraremos no elemento Densidade. Não obstante possa ser definida e calculada de diversas formas, utilizaremos a concepção de densidade territorial urbana, aqui entendida como áreas urbanizadas por área total, ou seja, densidade construída.

Em vista disso, considerando que São Paulo é o Estado com a maior extensão de áreas urbanizadas no Brasil (IBGE, 2022), e que as cidades médias experienciaram, nos últimos anos, um crescimento populacional relativo superior ao de cidades grandes e capitais (IBGE, 2023a), neste trabalho propomos uma análise das dinâmicas de crescimento urbano em municípios paulistas de porte médio.

O intuito é averiguar se existem padrões discerníveis de crescimento urbano que possibilitariam categorizar esses municípios. Tais padrões são identificados pela análise dos resultados da aplicação e ajuste da supramencionada função não linear de densidade territorial urbana.

O agrupamento dos municípios com base em seus padrões de crescimento urbano possibilitará investigar, em pesquisas futuras, quais características, para além das morfológicas, tornam esses municípios similares dentro de seus grupos. Em outras palavras, esses resultados ajudarão a compreender, de maneira mais aprofundada, os fatores sociais, políticos, econômicos e ambientais que exercem influência sobre o crescimento urbano.

As análises são guiadas por dados de acesso público, incluindo imagens oriundas de sensoriamento remoto, e pelo uso de técnicas de aprendizado não supervisionado via utilização da linguagem de programação Python, desenvolvida sob uma licença de código aberto e, portanto, livremente utilizável e distribuível, mesmo para uso comercial.

1.2 Objetivos

O objetivo central destes trabalho é identificar padrões de crescimento urbano em municípios paulistas de porte médio.

Para tanto, foram elaborados os seguinte objetivos específicos:

- Implementar o método quantitativo de caracterização não linear da variação da densidade territorial urbana via função sigmoide.
- Analisar as variações de densidade territorial urbana em municípios paulistas de porte médio por meio do ajuste da função sigmoide.
- Realizar análises de agrupamento para identificação de padrões de crescimento urbano nos municípios paulistas analisados.

1.3 Estrutura da Monografia

Além deste capítulo introdutório, esta monografia está organizada em mais quatro capítulos. No Capítulo 2 encontra-se a fundamentação teórica necessária para compreensão dos conceitos utilizados no decorrer do trabalho. No Capítulo 3 são descritos os dados e os métodos utilizados, bem como as técnicas para obtenção, tratamento e análise desses dados. No Capítulo 4 são apresentados os resultados da análise das variações de densidade territorial urbana e da aplicação de técnicas de agrupamento de dados para identificação de padrões de crescimento urbano nos municípios paulistas analisados. Finalmente, no Capítulo 5, são expostas as conclusões, com destaque para as limitações do trabalho e as possíveis contribuições para estudos futuros.

2 REFERENCIAL TEÓRICO

Neste capítulo abordamos os principais conceitos que nortearam a elaboração deste trabalho de conclusão de curso. Para tanto, apresentamos breves revisões bibliográficas sobre a função de densidade territorial urbana (seção 2.1), bem como algumas das técnicas de agrupamento de dados existentes na literatura (seção 2.2), com posterior apresentação de algumas formas de validação e interpretação dos *clusters* formados pelos algoritmos de agrupamento (seção 2.3).

2.1 Função de densidade territorial urbana

Uma das principais características que distingue zonas urbanas de rurais é a concentração espacial da população e das áreas construídas, de modo que, tanto a densidade populacional urbana quanto a densidade territorial urbana são significativamente superiores às densidades nas zonas rurais (ZHENG *et al.*, 2022).

Pesquisadores que utilizam a partição das cidades em anéis concêntricos para quantificar mudanças espaço-temporais frequentemente descobrem que elementos urbanos, como a densidade populacional urbana e a densidade territorial urbana (densidade construída), decaem gradualmente do centro para as regiões periféricas das cidades (ZHENG *et al.*, 2022).

A partição concêntrica de cidades foi comumente utilizada em diversos estudos, com muitos pesquisadores seguindo a abordagem de análise baseada em anéis concêntricos (JIAO, 2015). Essa análise de anéis concêntricos, também conhecida como análise de gradiente, se apoia no modelo monocêntrico de cidades (GUÉROIS; PUMAIN, 2008). Nesse tipo de análise, parte-se do pressuposto de um espaço uniforme e isotrópico, onde a medida radial é considerada mais apropriada e robusta para descrever padrões que exibem uma estrutura centro-periferia (GUÉROIS; PUMAIN, 2008). Embora tenha sido amplamente utilizada na modelagem urbana, sendo eficaz na caracterização do gradiente urbano-rural (WU *et al.*, 2021), as generalizações advindas dessa análise evidentemente assumem uma distribuição radial concêntrica do espaço construído (GUÉROIS; PUMAIN, 2008). Isso dificulta a captura de desenvolvimentos assimétricos, bem como das principais tendências direcionais desses desenvolvimentos (GUÉROIS; PUMAIN, 2008).

Foram formulados diversos modelos¹ para descrever o decaimento da densidade populacional urbana conforme o distanciamento dos centros urbanos: exponencial negativo

¹ Modelos são abstrações matemáticas do mundo real que possibilitam replicar um estado anterior (ou atual) e realizar simulações para prever cenários futuros (LI; GONG, 2016). Em geral, o procedimento de modelagem envolve determinar o comportamento do sistema e transformá-lo em equações (ABURAS; AHAMAD; OMAR, 2019).

(CLARK, 1951), densidade normal (TANNER, 1961), exponencial de potência negativa (SMEED, 1962), exponencial quadrático (NEWLING, 1969), *Splines* Cúbicos (ANDERSON, 1985), potência inversa (BATTY; KIM, 1992), Gaussiano (CHEN; FENG, 2012), dentre outros.

Em comparação com estudos sobre densidade populacional urbana, ainda é limitado o conhecimento acerca das características de concentração espacial e dos padrões de declínio da densidade territorial urbana (GUÉROIS; PUMAIN, 2008; ZHENG *et al.*, 2022). No entanto, adquirir informações precisas sobre mudanças no uso e cobertura da terra é mais simples do que obter dados sobre a distribuição populacional urbana, devido à ampla disponibilidade de dados multitemporais provenientes de sensoriamento remoto (JIAO, 2015).

Ao examinar a distribuição da densidade construída em 28 cidades da China, por meio de imagens de sensoriamento remoto, Jiao (2015) identificou um padrão de decaimento que geralmente segue uma forma de S-invertido, caracterizado por um declínio inicialmente lento, seguido por uma queda rápida e, finalmente, uma diminuição gradual. Baseando-se na função sigmoide, o autor propôs a função de densidade territorial urbana em forma de S-invertido para descrever esse comportamento de declínio na densidade construída à medida que a distância aumenta (JIAO, 2015).

O constatado decaimento da densidade territorial urbana em anéis concêntricos, conforme o distanciamento do centro, reflete o gradiente da intensidade do desenvolvimento urbano (JIAO, 2015).

Embora descreva o padrão de decaimento espacial da densidade construída, essa função proposta não incorpora a dimensão vertical, isto é, a altura dos edifícios. Tal informação poderia fornecer uma indicação mais precisa da intensidade de ocupação do solo urbano, dado que sua ausência acarreta na subestimação da ocupação das zonas mais densas, com prédios mais elevados (GUÉROIS; PUMAIN, 2008). Ainda assim, a função representa um modelo empírico relevante para instrumentalizar a análise da distribuição espacial das áreas urbanas. Sua aplicação viabiliza descrever quantitativamente, sob uma perspectiva espacial, as características do processo de crescimento urbano (ZHENG *et al.*, 2022), propiciando comparar uma cidade em momentos distintos, bem como diferentes cidades em um mesmo horizonte temporal.

A função foi amplamente utilizada para descrever o decaimento da densidade territorial urbana em diversos lugares, incluindo cidades de países como Tanzânia (SUMARI *et al.*, 2020), Tailândia (KEERATIKASIKORN, 2018), Laos (HUANG; ZHANG, 2022), Índia (SHELADIYA; PATEL, 2023), Irã (BAGHERI; SOLTANI, 2023), dentre outros (XU *et al.*, 2019). Embora também tenha sido aplicada em cidades de países latino-americanos (WU *et al.*, 2021), como o Brasil (MENZORI, 2021; MENZORI; SOUSA; GONÇALVES, 2023), desconhecemos abordagens que valeram-se dos resultados do ajuste dessa função

para categorização de cidades por padrões de crescimento urbano.

2.2 Técnicas de agrupamento

A análise de agrupamento (*clustering*) faz parte do campo mais amplo de Aprendizado de Máquina (*Machine Learning*). Segundo definido por Mitchell (1997), o Aprendizado de Máquina, se concentra no desenvolvimento de programas computacionais capazes de se aperfeiçoarem automaticamente por meio da experiência. Isso significa que um modelo de Aprendizado de Máquina, uma vez configurado, é capaz de descobrir, de maneira independente, estruturas nos dados ou prever dados novos (desconhecidos) (GÉRON, 2022).

Pichler e Hartig (2023) consideram que o Aprendizado de Máquina ramificou-se dos modelos estatísticos clássicos com o objetivo de aumentar a flexibilidade e complexidade da modelagem, abandonando a ideia de um modelo probabilístico. Em termos práticos, o Aprendizado de Máquina visa identificar padrões relevantes nos dados, conforme destacado por Rodrigues (2023), podendo ser dividido em: Supervisionado, Não Supervisionado e por Reforço (GÉRON, 2022).

- O Aprendizado **Supervisionado** usa dados rotulados no treinamento de algoritmos de Aprendizado de Máquina que farão previsões para novos dados. Uma maneira de avaliar a precisão de um modelo supervisionado é testá-lo em alguns dados pré-rotulados e, em seguida, comparar as previsões de classificação do modelo com o resultado original (GÉRON, 2022).
- O Aprendizado **Não Supervisionado** é aplicado a dados não rotulados, visando identificar estruturas potenciais (RODRIGUES, 2023), sendo comumente utilizado na análise exploratória dos dados (JAMES *et al.*, 2023).
- O Aprendizado por **Reforço** é utilizado em tarefas onde a aprendizagem depende das ações executadas e das consequências produzidas. Nesse caso, o treinamento do algoritmo de Aprendizado de Máquina ocorre pela interação com um ambiente virtual (PICHLER; HARTIG, 2023).

A ênfase deste trabalho está nos métodos de aprendizado não supervisionado, que incluem diversos algoritmos de agrupamento de dados (RODRIGUES, 2023; JAMES *et al.*, 2023). O objetivo da análise de agrupamento é encontrar uma estrutura nos dados que possibilite que itens alocados em um mesmo grupo sejam mais semelhantes entre si do que aqueles de grupos diferentes (RASCHKA; MIRJALILI, 2019; JAMES *et al.*, 2023). Em outras palavras, busca-se reunir o que é semelhante e separar o que é diferente (LOPES; GOSLING, 2020).

No algoritmo *K-Means*, por exemplo, uma das técnicas de agrupamento mais comuns existentes (CARVALHO *et al.*, 2011a; ALBON, 2018a), o intuito é partitionar as observações em um número pré-especificado de *clusters* (k), de maneira que o algoritmo atribua cada observação a exatamente um dos k *clusters*, ao passo que no algoritmo *MeanShift*, por outro lado, não é necessário definir previamente o número de *clusters* (JAMES *et al.*, 2023).

A análise de agrupamento não requer que os dados sigam distribuições exigidas em outros métodos, e viabiliza a obtenção de resultados consistentes mesmo com um conjuntos reduzido de dados (LOPES; GOSLING, 2020). No entanto, o agrupamento de dados apresenta diversos desafios (JAMES *et al.*, 2023; RODRIGUES, 2023). A ausência de rótulos, por exemplo, torna a avaliação dos resultados do agrupamento subjetiva, assim como a determinação do número ideal de *clusters* e a seleção dos algoritmos apropriados, dependentes do contexto (RODRIGUES, 2023). A avaliação também é problemática, devido à dificuldade das métricas existentes em capturar estruturas complexas, como *clusters* sobrepostos, com densidades variadas ou formas não esféricas (RODRIGUES, 2023).

2.2.1 *K-Means*

No *K-Means*, cada *cluster* é representado pelo seu centroide (Equação 2.2), sendo que a coordenada de cada centroide é a média (ou mediana) das coordenadas de pontos semelhantes (PARK; JUN, 2009; CARVALHO *et al.*, 2011a; RASCHKA; MIRJALILI, 2019; JAMES *et al.*, 2023). Nesse sentido, esse algoritmo atribui pontos iterativamente ao grupo mais próximo, usando a distância do centroide do grupo (ponto central) e formando k grupos (MOLIN; JEE, 2021a). A similaridade entre objetos é definida como sendo oposta à distância, sendo que k objetos representativos são selecionados para minimizar a dissimilaridade (Equação 2.1).

$$\frac{1}{n} \sum_{i=1}^n d(i, m(i)) \quad (2.1)$$

em que $d(i, m(i))$ é a dissimilaridade do objeto i ao seu objeto representativo mais próximo, expresso por $m(i)$ (ROUSSEEUW, 1987).

$$\bar{x}^{(k)} = \frac{1}{n_k} \sum_{x_i \in C_k} x_i \quad (2.2)$$

em que $C_k = \{x_1, x_2, \dots, x_{n_k}\}$ é um *cluster* com n_k objetos e $\bar{x}^{(k)}$ é o centroide do *cluster* C_k (CARVALHO *et al.*, 2011a).

A distância euclidiana (Equação 2.3) é comumente utilizada para agrupar exemplos com características contínuas (RASCHKA; MIRJALILI, 2019). Com base nessa métrica de

distância, o algoritmo *K-Means* pode ser descrito como um simples problema de otimização, em que se aplica uma abordagem iterativa para minimizar a inércia, isto é, a soma dos erros quadráticos (*sum of squared errors*) da distância dos pontos ao centroide do *cluster* mais próximo (RASCHKA; MIRJALILI, 2019). Para um agrupamento contendo k *clusters*, o erro quadrático equivale à soma da variação dentro dos *clusters* (CARVALHO *et al.*, 2011a), conforme Equação 2.4.

$$d(C_1, C_2) = \min_{\substack{x_i \in C_1 \\ x_j \in C_2}} d(x_i, x_j) \quad (2.3)$$

em que a distância entre *clusters* é determinada pela distância entre os objetos de dois *clusters* que estão mais próximos, isto é, a distância mínima entre quaisquer dois objetos, um de cada *cluster* (CARVALHO *et al.*, 2011a).

$$E = \sum_{j=1}^k \sum_{x_i \in C_j} d(x_i, \bar{x}^{(j)})^2 \quad (2.4)$$

em que $\bar{x}^{(j)}$ representa o centroide do *cluster* C_j (Equação 2.2), enquanto $d(x_i, \bar{x}^{(j)})$ denota a distância euclidiana entre um objeto x_i e o centroide $\bar{x}^{(j)}$ (CARVALHO *et al.*, 2011a).

O funcionamento do algortmo *K-Means* pode ser explicado por meio das seguintes etapas (CARVALHO *et al.*, 2011a; ALBON, 2018a):

1. *Inicialização*: K centroides são escolhidos (geralmente de maneira aleatória).
2. *Atribuição aos clusters*: Calcula-se a distância (usualmente euclidiana²) entre cada observação e os K centroides, atribuindo cada observação ao *cluster* do centroide mais próximo.
3. *Atualização dos centroides*: Os centroides são recalculados, movendo-se para as médias de seus respectivos *clusters*.
4. *Convergência*: Repete-se as etapas 2 e 3 até que não ocorram mais mudanças significativas nas atribuições das observações aos *clusters*.

Como a aplicação do *K-Means* envolve a utilização de uma métrica de distância euclidiana, os atributos medidos devem estar na mesma escala, sendo pertinente, portanto, aplicar alguma forma de padronização/normalização dos dados, quando necessário (RASCHKA; MIRJALILI, 2019). Adicionalmente, como o algoritmo *K-Means* escolhe aleatoriamente seus centroides iniciais, sendo possível obter resultados diferentes de *clusters*,

² A distância de Mahalanobis também pode ser utilizada para definição do erro quadrático (CARVALHO *et al.*, 2011a).

é necessário especificar uma semente, de modo a garantir a reproduzibilidade (MOLIN; JEE, 2021a).

Um aspecto do *K-Means* que merece atenção é sua suscetibilidade à presença de *outliers* no conjunto de dados. Isso porque cada centroide é uma média, ou seja, uma medida de tendência central afetada por valores extremos (LOPES; GOSLING, 2020). Dada a popularidade desse algoritmo, foram desenvolvidas alternativas mais apropriadas para lidar com o efeito de *outliers*, como o algoritmo *K-Medoids*, que é menos sensível à presença de valores extremos por não utilizar a média para definição dos grupos (LOPES; GOSLING, 2020).

2.2.2 *K-Medoids*

O *K-Medoids*, também chamado de PAM (do inglês *Partitioning Around Medoids*), é uma versão aprimorada do *K-Means* (XU; TIAN, 2015). Enquanto o *K-Means* busca minimizar a soma de distâncias euclidianas quadráticas entre cada ponto e o centroide de seu *cluster*, o *K-Medoids* busca minimizar a soma da distância euclidiana padrão entre cada ponto e o medoide de seu *cluster* (STEINLEY, 2015). O medoide é o objeto do *cluster* com a menor distância em relação aos demais membros de seu *cluster* (KAUFMAN; ROUSSEEUW, 2005).

O *K-Medoids* identifica um número k de medoides representativos nas observações do conjunto de dados, sendo que o primeiro objeto é aquele para o qual a soma das dissimilaridades com todos os demais objetos é a menor possível, isto é, o mais centralmente localizado no conjunto de objetos (KAUFMAN; ROUSSEEUW, 2005). A cada iteração, um novo objeto é escolhido, continuando esse processo até que sejam identificados k objetos que minimizem ao máximo a soma das dissimilaridades com os demais objetos, sendo formados *clusters*, portanto, pela atribuição de cada observação ao medoide mais próximo (KAUFMAN; ROUSSEEUW, 2005). O algoritmo realiza trocas sucessivas entre os medoides e seus objetos complementares, otimizando as funções objetivo até atingir o menor valor possível e, assim, interromper o processamento e apresentar o resultado final (KAUFMAN; ROUSSEEUW, 2005; LOPES; GOSLING, 2020). Segundo Reddy e Vinzamuri (2014), esse processo de trocas sucessivas funciona da seguinte forma:

1. Selecione aleatoriamente um objeto não representativo x_i .
2. Calcule o custo total S de troca do objeto representativo m por x_i .
3. Se $S < 0$, então troque m por x_i para formar o novo conjunto de K objetos representativos, até que o critério de convergência seja atendido.

Por utilizar medoides (pontos reais nos dados) em vez de centroides, o *K-Medoids*

é considerado mais robusto a *outliers*³ (LOPES; GOSLING, 2020), além de permitir uma maior interpretabilidade dos centros dos *clusters* do que o *K-Means* - no qual o centro de um *clusters* não é necessariamente um dos pontos dos dados de entrada, mas sim a média entre os pontos do *cluster* (KAUFMAN; ROUSSEEUW, 2005). No entanto, a complexidade computacional no *K-medoids* é maior do que no *K-Means*, sendo inadequado, portanto, para grandes conjuntos de dados (REDDY; VINZAMURI, 2014).

2.2.3 Mean Shift

O *Mean Shift* é um algoritmo de agrupamento baseado em densidade (XU; TIAN, 2015). A ideia básica por trás dos algoritmos baseados em densidade é que, no espaço de objetos, os *clusters* seriam regiões de alta densidade, separados por regiões de baixa densidade (CARVALHO *et al.*, 2011a).

De maneira simplificada, esse algoritmo funciona da seguinte forma (REDDY; VINZAMURI, 2014; XU; TIAN, 2015; ALBON, 2018a):

1. *Kernel de densidade*: Inicialmente, o algoritmo utiliza um *kernel* de densidade (Parzen, conforme Equação 2.5) para representar a distribuição espacial dos dados. O *kernel* é uma função de suavização que atribui pesos aos pontos de dados com base em sua proximidade espacial.
2. *Centroides iniciais*: Cada ponto de dado é considerado como um candidato a centroide inicial.
3. *Iteração*: Para cada ponto, um deslocamento (*shift*) é calculado com base na média ponderada dos pontos dentro de uma determinada vizinhança definida pelo *kernel* de densidade.
4. *Atualização do centroide*: O ponto é movido para a posição calculada pelo deslocamento na etapa anterior.
5. *Convergência*: Os passos 3 e 4 são repetidos até que os pontos alcancem convergência, ou seja, seus deslocamentos se tornem insignificantes.
6. *Agrupamento*: Os pontos que convergem para o mesmo centroide são agrupados no mesmo *cluster*.

$$f(x) = \frac{1}{Nh^d} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right) \quad (2.5)$$

³ Como mencionado por Steinley (2015), esta percepção de robustez do *K-Medoids* não é consensual na literatura, pois alguns autores o consideram insensível a valores discrepantes apenas em casos especiais.

em que, considerando N pontos de dados x_i , onde $i = 1, \dots, N$ em um espaço d -dimensional R^d , a estimativa multivariada de densidade do *kernel* da janela Parzen $f(x)$ é obtida com *kernel* $K(x)$ e raio de janela h (REDDY; VINZAMURI, 2014).

O algoritmo *Mean Shift*, por ser baseado em estimativas de densidade de kernel não paramétricas, faz poucas suposições sobre os dados, podendo modelar *clusters* não convexos (CARREIRA-PERPIÑÁN, 2015). Ademais, o *Mean Shift* não requer a definição prévia do número de *clusters*, sendo capaz de encontrar naturalmente os agrupamentos nos dados pelo ajuste dos centroides, conforme a densidade local (ALBON, 2018a). No entanto, a escolha do tamanho da janela do *kernel* (`bandwidth`) pode afetar a sensibilidade do algoritmo aos detalhes da estrutura dos dados. Em muitas das implementações desse algoritmo, esse hiperparâmetro pode ser definido manualmente, ou estimado automaticamente - com um aumento significativo no custo computacional (ALBON, 2018a). Adicionalmente, embora seja computacionalmente lento, o *Mean Shift* pode ser acelerado e dimensionado para grandes conjuntos de dados, conforme explicitado por Carreira-Perpiñán (2015).

2.2.4 *Affinity Propagation*

O algoritmo *Affinity Propagation* é um método de agrupamento relativamente simples, baseado na teoria dos grafos e fundamentado na troca de mensagens entre pontos de dados para identificação automática dos *clusters* em um conjunto de dados. A ideia central do *Affinity Propagation* é considerar todos os pontos de dados como potenciais centros de *cluster* (FREY; DUECK, 2007), sendo a afinidade o valor negativo da distância euclidiana entre dois pontos de dados (XU; TIAN, 2015). Quanto maior a soma da afinidade de um ponto de dados para outros pontos de dados, maior a probabilidade deste ponto ser o centro do *cluster* (XU; TIAN, 2015). Basicamente, ele funciona da seguinte forma (FREY; DUECK, 2007):

1. *Matriz de similaridade*: Inicialmente, o algoritmo começa com uma matriz de similaridade que calcula as semelhanças entre todos os pares de pontos de dados no conjunto.
2. *Responsabilidade (Responsibility) e Disponibilidade (Availability)*: Para cada ponto de dados são calculadas duas medidas: a responsabilidade (Equação 2.6), que mede o quanto adequado é que um ponto seja o exemplar de outro, e a disponibilidade (Equação 2.7), que indica o quanto bem um ponto está posicionado para ser escolhido por outro como seu exemplar.
3. *Atualização das mensagens*: As mensagens de responsabilidade e disponibilidade são atualizadas iterativamente com base nas mensagens recebidas dos outros pontos.

4. *Clusters exemplares*: Durante as iterações, os pontos começam a se agrupar. Pontos com alta responsabilidade para um ponto específico são propensos a serem agrupados no mesmo *cluster*.
5. *Convergência*: O algoritmo continua iterando até que os pontos alcancem um estado estável. Os *clusters* são identificados com base nos pontos exemplares, que são aqueles com a maior disponibilidade.

$$r(i, k) \leftarrow s(i, k) - \max_{k' \text{ s.t. } k' \neq k} \{a(i, k') + s(i, k')\} \quad (2.6)$$

$$a(i, k) \leftarrow \min \left(0, r(k, k) + \sum_{i' \text{ s.t. } i' \notin \{i, k\}} \max(0, r(i', k)) \right) \quad (2.7)$$

em que, $r(i, k)$ representa a responsabilidade do ponto de dado k ser o exemplar do ponto de dado i , $s(i, k)$ é a similaridade entre os pontos de dados i e k , e $a(i, k')$ representa a disponibilidade do ponto de dado k ser o exemplar do ponto de dado i (FREY; DUECK, 2007).

Em síntese, ao visualizar cada ponto de dados como um nó em uma rede, o *Affinity Propagation* utiliza a comunicação iterativa entre pontos de dados para identificar automaticamente os *clusters* (FREY; DUECK, 2007), sendo particularmente útil quando o número de *clusters* não é conhecido previamente.

Ao considerar simultaneamente todos os pontos de dados como candidatos a centros, esse algoritmo é capaz de evitar muitas das soluções inadequadas causadas por inicializações aleatórias, além de ser insensível a *outliers* (FREY; DUECK, 2007) e também permitir uma maior interpretabilidade dos centros dos *clusters*. Entretanto, uma desvantagem do *Affinity Propagation* é sua elevada complexidade computacional, que o torna inapropriado para grandes conjuntos de dados (XU; TIAN, 2015).

2.3 Validação e interpretação dos *clusters*

De acordo com Carvalho *et al.* (2011b), avaliar e comparar algoritmos de agrupamento difere do objetivo de validar as estruturas encontradas por tais algoritmos. Ambos propósitos são complexos, pois não existe um resultado correto a ser alcançado com qualquer algoritmo de agrupamento, visto o frequente desconhecimento da estrutura real dos *clusters* (CARVALHO *et al.*, 2011b). Além disso, como qualquer algoritmo de agrupamento encontrará uma possível estrutura para os dados, independente de sua existência, podem haver soluções múltiplas, isto é, o fato de um algoritmo encontrar um número de *clusters* diferente de outro algoritmo não implica, necessariamente, que apenas um deles esteja

correto (CARVALHO *et al.*, 2011b). Como existem diversas maneiras de aferir a similaridade/dissimilaridade entre objetos, variadas estruturas são possíveis (CARVALHO *et al.*, 2011b). Isso não impede, entretanto, que algum(ns) algoritmo(s) encontre(m) *clusters* mais apropriados do que outros (CARVALHO *et al.*, 2011b).

Dentre as estratégias para validar agrupamentos, medidas internas de validação podem ser úteis para escolher o algoritmo mais apropriado, e determinar o número ideal de *clusters*, especialmente em algoritmos que requerem ajuste desse hiperparâmetro, como *K-Means* e *K-Medoids* (XIONG; LI, 2014).

2.3.1 Silhueta

A Silhueta (*silhouette*) pode ser utilizada para definir o k mais adequado, com o número de *clusters* sendo determinado pela maximização do valor deste índice (ROUSSEEUW, 1987; CARVALHO *et al.*, 2011b; XIONG; LI, 2014). A Silhueta oferece uma medida quantitativa única que avalia duas características: o quanto agrupadas estão as observações dentro de um mesmo *cluster*, e quanto distantes estão as observações entre *clusters* diferentes (ALBON, 2018b). Nesse sentido, essa métrica é usada para avaliar a qualidade de um agrupamento, fundamentando-se na proximidade entre os objetos dentro de um *cluster* e na distância desses objetos em relação ao *cluster* mais próximo (CARVALHO *et al.*, 2011b).

A Silhueta também funciona como uma ferramenta gráfica para averiguar quanto agrupados estão os exemplos nos *clusters* (RASCHKA; MIRJALILI, 2019). O intervalo dessa medida varia entre -1 e 1, com a melhor partição dos dados se aproximando de 1 (CARVALHO *et al.*, 2011b; ALBON, 2018b), e o valor retornado por essa métrica sendo o coeficiente médio da Silhueta para todas as observações (ALBON, 2018b).

Para construir Silhuetas, dois elementos são necessários: a partição obtida por meio da aplicação de alguma técnica de agrupamento, e a coleção de todas as proximidades entre os objetos do conjunto de dados segmentado (ROUSSEEUW, 1987). Portanto, essa medida independe do algoritmo de agrupamento empregado, podendo ser aproveitada para comparação dos resultados de algoritmos diferentes aplicados a um mesmo conjunto de dados (ROUSSEEUW, 1987; CARVALHO *et al.*, 2011b). Isso porque a Silhueta avalia a adequação de um objeto ao seu *cluster*, ou seja, quanto bem ele foi classificado, assim como a qualidade de cada *cluster* individualmente (ROUSSEEUW, 1987; CARVALHO *et al.*, 2011b), sendo, portanto, uma métrica baseada em um compromisso entre homogeneidade dentro do *cluster* e separação entre *clusters* (HALKIDI; VAZIRGIANNIS; HENNIG, 2015).

Para medidas de distância entre objetos, as Silhuetas de um *cluster* e de uma partição são calculadas por meio das Equações 2.8 e 2.9, respectivamente (CARVALHO *et al.*, 2011b).

$$sil(C_k) = \frac{1}{|C_k|} \sum_{x_i \in C_k} sil(x_i) \quad (2.8)$$

$$sil(\pi) = \frac{1}{n} \sum_{i=1}^n sil(x_i) \quad (2.9)$$

em que n é o número de objetos agrupados e $sil(x_i)$ é a Silhueta do objeto x_i (CARVALHO *et al.*, 2011b).

A Silhuera do objeto x_i é expressa pela Equação 2.10, (ROUSSEEUW, 1987; CARVALHO *et al.*, 2011b).

$$sil(x_i) = \begin{cases} 1 - a(x_i, C_i)/b(x_i), & a(x_i, C_i) < b(x_i) \\ 0, & a(x_i, C_i) = b(x_i) \\ b(x_i)/a(x_i, C_i) - 1, & a(x_i, C_i) > b(x_i) \end{cases} \quad (2.10)$$

$$a(x_i, C_k) = \frac{1}{|C_k|} \sum_{\substack{x_i, x_j \in C_k \\ x_i \neq x_j}} d(x_i, x_j) \quad (2.11)$$

$$b(x_i) = \min_{\substack{x_i \in C_i \\ C_i \neq C_j}} a(x_i, C_j) \quad (2.12)$$

em que $a(x_i, C_i)$ (Equação 2.11) representa a distância média do objeto x_i em relação aos demais objetos do *cluster* C_i , ao qual o objeto x_i pertence, e $b(x_i)$ (Equação 2.12) representa a menor distância média de x_i em relação a todos os outros *clusters* (CARVALHO *et al.*, 2011b).

De maneira simplificada, portanto, a Silhueta é calculada pela subtração da média das distâncias entre os pontos em um determinado *cluster* e o *cluster* diferente mais próximo (b) pela média das distâncias entre cada dois pontos em um *cluster* (a), dividido pelo máximo dos dois, conforme Equação 2.13 (ROUSSEEUW, 1987; ALBON, 2018b; RASCHKA; MIRJALILI, 2019; MOLIN; JEE, 2021a).

$$\frac{b - a}{\max(a, b)} \quad (2.13)$$

Com base na Equação 2.13, observa-se que o coeficiente de Silhueta é 0 se a separação e a coesão do *cluster* forem iguais ($b = a$). Ademais, um coeficiente de Silhueta ideal (1) ocorre se $b > a$, já que b quantifica quão diferente um exemplo é de outros *clusters*, e a quantifica quão semelhante um exemplo é dos outros exemplos em seu próprio *cluster* (RASCHKA; MIRJALILI, 2019).

A Silhueta apresenta uma desvantagem ao considerar distâncias uniformes em todo o conjunto de dados, implicitamente assumindo um conceito de *cluster* esférico (HALKIDI; VAZIRGIANNIS; HENNIG, 2015). Por conseguinte, essa medida se destaca na identificação de *clusters* compactos e bem separados, especialmente os com formatos aparentemente esféricos (CARVALHO *et al.*, 2011b).

2.3.2 Calinski–Harabasz

O índice Calinski-Harabasz (CALIŃSKI; HARABASZ, 1974) baseia-se no equilíbrio entre a variação dentro do *cluster* (Equação 2.14) com a variação entre *clusters* (Equação 2.15).

$$W_{C_K} = \sum_{j=1}^K \sum_{c(i)=j} (\mathbf{x}_i - \bar{\mathbf{x}}_j)(\mathbf{x}_i - \bar{\mathbf{x}}_j)^t \quad (2.14)$$

$$B_{C_K} = \sum_{j=1}^K n_j (\bar{\mathbf{x}}_j - \bar{\mathbf{x}})(\bar{\mathbf{x}}_j - \bar{\mathbf{x}})^t \quad (2.15)$$

Nesse sentido, o critério de razão de variância é definido conforme a Equação 2.16 (HALKIDI; VAZIRGIANNIS; HENNIG, 2015).

$$CH(C_K) = \frac{\text{trace}(B_{C_K})}{\text{trace}(W_{C_K})} \times \frac{n - K}{K - 1} \quad (2.16)$$

em que C_K é a configuração do *cluster*, $\text{trace}(W_{C_K})$ é o traço da matriz de dispersão dentro do *cluster*, $\text{trace}(B_{C_K})$ é o traço da matriz de dispersão entre *clusters*, n é o número total de observações e K é o número de *clusters*.

No índice Calinski–Harabasz, valores mais elevados representam bons agrupamentos, pois indicam que os grupos são homogêneos e bem distintos entre si (HALKIDI; VAZIRGIANNIS; HENNIG, 2015).

2.3.3 Davies–Bouldin

O índice Davies-Bouldin (DAVIES; BOULDIN, 1979) é definido como a similaridade média entre cada *cluster* e seu *cluster* mais semelhante (XU; TIAN, 2015). Ele é calculado como a média das razões entre a soma das distâncias médias intra-*cluster* ($\sigma_i + \sigma_j$) e a distância entre os centros dos *clusters* ($d(c_i, c_j)$), para cada par de *clusters* ($i \neq j$), conforme Equações 2.17 a 2.21 (HALKIDI; VAZIRGIANNIS; HENNIG, 2015).

$$S_k = \left(\frac{1}{n_{1k}} \sum_{c(i)=k} \|x_i - \bar{x}_k\|_2^q \right)^{1/q} \quad (2.17)$$

em que a medida de dispersão no *cluster* k (S_k) é calculada pela distância euclidiana elevada à potência q entre cada ponto (x_i) do *cluster* k ao centroide (\bar{x}_k), elevado à potência $\frac{1}{q}$.

$$M_{ij} = \|\bar{x}_i - \bar{x}_j\|_p \quad (2.18)$$

em que a medida de distância entre os *clusters* i e j (M_{ij}) é calculada pela distância entre os centroides dos *clusters* i e j , sendo que p denota a forma de medir a distância entre dois pontos no espaço. O valor específico de p é determinado pelo contexto ou aplicação. Escolhas comuns incluem $p = 1$ (distância de *Manhattan*), $p = 2$ (distância Euclidiana), dentre outras.

$$R_{ij} = \frac{S_i + S_j}{M_{ij}} \quad (2.19)$$

em que a medida de separação (R_{ij}) entre os *clusters* i e j é calculada pela razão entre a soma das medidas de dispersão dos *clusters* i e j e sua medida de distância, fornecendo uma medida de quão bem separados estão dois *clusters*, considerando tanto suas dispersões quanto a distância entre seus centroides.

$$D_i = \max_{j \neq i} R_{ij} \quad (2.20)$$

em que a separação máxima do *cluster* i em relação a qualquer outro *cluster* é calculada pela medida de separação máxima de *clusters* (D_i).

$$DB(C_K) = \frac{1}{K} \sum_{i=1}^K D_i \quad (2.21)$$

em que o índice Davies-Bouldin para determinada solução de agrupamento C_K é calculado pela média das medidas máximas de separação entre *clusters*. Essa métrica proporciona uma avaliação global da qualidade da solução de agrupamento encontrada.

Como apontado por Davies e Bouldin (1979), a utilidade desse índice é restringida quando são formados *clusters* com apenas um membro, uma vez que nesses casos a dispersão é zero.

No índice Davies-Bouldin, valores mais baixos (até zero) indicam uma melhor separação e compactação dos *clusters* (HALKIDI; VAZIRGIANNIS; HENNIG, 2015).

3 MÉTODOS

Neste capítulo são apresentados os métodos empregados para realização deste trabalho de conclusão de curso. Na seção 3.1 consta a descrição dos conjuntos de dados utilizados para composição da base de dados estruturados da pesquisa. A seção 3.2 é composta pelo detalhamento do processo de coleta e tratamento dos conjuntos de dados (3.2.1) descritos na seção anterior (3.1), bem como os procedimentos para implementação (3.2.2) dos métodos de análise e interpretação dos resultados (3.2.3).

3.1 Base de Dados

A base de dados utilizada neste trabalho foi construída a partir de conjuntos de dados disponibilizados publicamente pelo Instituto Brasileiro de Geografia e Estatística - IBGE e pelo Mapeamento Anual da Cobertura e Uso do Solo do Brasil - MapBiomas.

Os dados provenientes do IBGE referem-se aos Resultados de População do Censo Demográfico de 2022 (IBGE, 2023a), ao Cadastro de Localidades Selecionadas de 2010 (IBGE, 2010), e à Malha Municipal de 2022 (IBGE, 2023b).

Os Resultados de População do Censo Demográfico de 2022 englobam as seguintes tabelas: a) População residente, Área territorial e Densidade demográfica; b) População residente e Taxa média geométrica de crescimento anual (IBGE, 2023a).

O Cadastro de Localidades Selecionadas de 2010 é um banco de dados históricos contendo informações sobre localidades selecionadas, classificadas em cidades, vilas, povoados, projetos de assentamento (agrovilas), aldeias indígenas e áreas urbanas isoladas (IBGE, 2010).

A Malha Municipal de 2022 retrata a situação vigente da Divisão Político Administrativa (DPA) dos municípios brasileiros, via representação vetorial das linhas definidoras das divisas estaduais e dos limites municipais (IBGE, 2023b).

Os dados da Coleção 8 do MapBiomas englobam mapas e dados anuais de cobertura e uso da terra para o período de 1985 a 2022, obtidos via sensoriamento remoto (imagens do satélite *Landsat*, com resolução de 30 metros, disponíveis gratuitamente na plataforma de análise geoespacial baseada na nuvem *Google Earth Engine* - GEE), com os respectivos Códigos da Legenda das classes de cobertura e uso da terra no Brasil (CUNHA, 2023).

O método utilizado pela equipe do MapBiomas para obtenção das classes de uso e cobertura da terra é baseado em aprendizado supervisionado, envolvendo a implementação de algoritmos de processamento de imagens no GEE, visando pré-processar todas as imagens *Landsat* e normalizá-las para treinar um classificador Floresta Aleatória (*Random*

Forest), de modo a mapear as classes de uso e cobertura da terra de todos os biomas brasileiros (SOUZA *et al.*, 2020).

3.2 Etapas da Pesquisa

Este trabalho divide-se em três etapas principais. A primeira etapa (3.2.1) envolveu a coleta e o tratamento dos conjuntos de dados descritos na seção anterior. A segunda etapa (3.2.2) abrangeu a implementação da mencionada função sigmoide em linguagem de programação Python. A terceira etapa (3.2.3) consistiu nas análises de agrupamento dos municípios da amostra, via técnicas de aprendizado não supervisionado.

3.2.1 Etapa 1: Coleta e tratamento dos dados

Inicialmente, utilizamos as tabelas com os Resultados de População do Censo Demográfico de 2022 (IBGE, 2023a) para obtenção da lista com os nomes dos municípios paulistas de porte médio da amostra. Obtivemos essa lista de nomes pela seleção dos municípios com população residente entre 100 e 500 mil habitantes, exceto cidades litorâneas e integrantes de regiões metropolitanas.

Esses critérios de exclusão foram adotados para diminuir as chances de incorporação de conurbações urbanas (expansão de duas ou mais cidades limítrofes que se unem), bem como de cidades policêntricas, uma vez que o método proposto por Jiao (2015) requer adaptações para aplicação em cidade policêntricas. Ademais, a supressão de cidades litorâneas se deu em função das especificidades do crescimento urbano, que não ocorre de maneira uniforme em todas as direções nessas cidades, tornando a análise por meio de círculos concêntricos equidistantes inadequada nesses casos.

Por outro lado, os critérios de inclusão garantem o enfoque em um estrato representativo, porém negligenciado, da realidade urbana no mundo. Isso porque, embora aproximadamente um terço da população mundial viva em aglomerados com menos de 500.000 habitantes, há uma desproporcionalidade no estudo das áreas urbanas de pequeno e médio porte, que são menos investigadas do que cidades grandes e metrópoles (REBA; SETO, 2020). Ademais, a maioria dos estudos foca em cidades do Norte Global, havendo uma manifesta demanda por compreender e caracterizar melhor o crescimento urbano no Sul Global¹ (REBA; SETO, 2020).

Nos pautamos nessa lista de cidades para extração dos centros históricos do Cadastro de Localidades Selecionadas de 2010 (IBGE, 2010), por meio das coordenadas do centroide (pontos) do setor censitário de referência das localidades, municípios da amostra nesse caso. Optamos por não utilizar o centro geométrico dos municípios, ou de suas áreas urbanizadas,

¹ Os termos Norte Global e Sul Global captam condições econômicas atuais, além de refletirem a economia política e a história de colonização das nações (REBA; SETO, 2020).

porque tais centroides nem sempre correspondem às origens de desenvolvimento urbano das cidades.

A mesma lista de nomes viabilizou a aquisição dos limites territoriais dos municípios analisados (polígonos), disponíveis no conjunto de dados da Malha Municipal de 2022 (IBGE, 2023b).

Por fim, a partir dos limites territoriais dos municípios, adquirimos as áreas urbanas da Coleção 8 do MapBiomass (CUNHA, 2023) para o período de 1985 a 2020, em intervalos de 5 anos, totalizando oito anos: 1985, 1990, 1995, 2000, 2005, 2010, 2015 e 2020. Para cada um desses oito anos, extraímos das imagens previamente classificadas somente dados referentes às áreas urbanizadas. Para tanto, foi utilizado o arquivo com os Códigos da Legenda da Coleção 8, para identificação do código referente especificamente às áreas urbanizadas (ID 24). Para facilitar as análises, convertemos os dados extraídos do formato matricial (*raster*) para o vetorial (pontos e polígonos).

Cabe destacar que as áreas urbanizadas também poderiam ter sido obtidas do IBGE. Isso porque, via interpretação de imagens de satélite, o IBGE publicou mapeamentos (vetorizações) das manchas urbanizadas brasileiras (IBGE, 2022). Esses dados foram disponibilizados no formato *shapefile*, em feições poligonais. Entretanto, optamos por não utilizar tais produtos por não estarem disponíveis em intervalos de tempo regulares, sendo encontrados apenas para os anos de 2005, 2015 e 2019.

Todos os conjuntos de dados utilizados foram importados em seu formato original para uma estrutura de dados tabular bidimensional (matriz).

3.2.2 Etapa 2: Implementação da função sigmoide

Para implementar a função não linear de densidade territorial urbana em linguagem de programação Python, foi necessário adaptar o método definido por Jiao (2015), que utilizou os softwares ENVI 4.5 e MATLAB R2010b para processamento de imagens e ajuste da função, respectivamente.

O método consiste em criar anéis concêntricos isométricos para cada cidade, a partir de seu centro urbano (JIAO, 2015). Com base nessa partição das cidades em anéis concêntricos, calcula-se a densidade territorial urbana (proporção de área urbana), ou seja, área construída dividida pela área total para cada anel (JIAO, 2015):

$$Dens = \frac{S_{urb}}{S} \quad (3.1)$$

em que $Dens$ é a densidade territorial urbana (área construída), S_{urb} é a área construída em um anel e S é a área total no mesmo anel.

Essa relação entre a área construída e a área total do anel (Equação 3.1) fornece

uma estimativa da densidade média construída a uma determinada distância do centro (GUÉROIS; PUMAIN, 2008).

Inspirado no formato em S da função sigmoide (Equação 3.2), Jiao (2015) propôs uma função modificada (função de densidade territorial urbana) com um formato de S-invertido para caracterizar o decaimento da densidade construída conforme o distanciamento do centro urbano da cidade.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3.2)$$

A função sigmoide proposta por Jiao (2015) é definida da seguinte forma:

$$f(r) = \frac{1 - c}{1 + e^{\alpha((2r/D)-1)}} + c \quad (3.3)$$

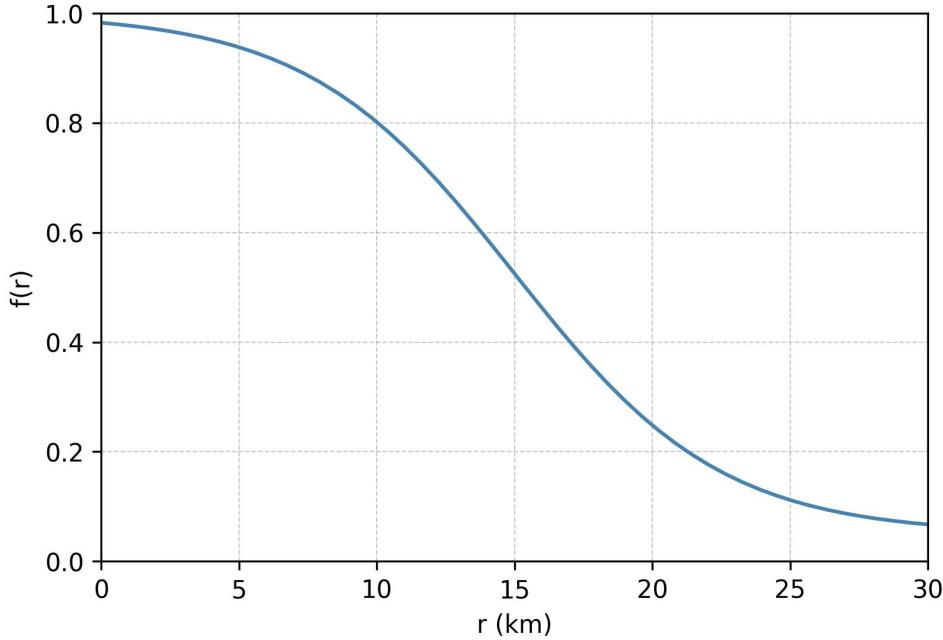
em que f representa a densidade territorial urbana, que é uma função da distância a partir do centro da cidade r , e representa o número de Euler, e α , c e D são constantes - parâmetros que representam as três características básicas da distribuição da densidade territorial urbana (CHAKRABORTY *et al.*, 2022), que podem ser ajustadas. Esses parâmetros possuem significado espacial explícito: o parâmetro α controla a forma da curva ajustada, sendo que um α maior indica uma forma urbana mais compacta; o parâmetro c representa a densidade mínima do local em que a equação está sendo aplicada; o parâmetro D denota o limite estimado da extensão urbana de uma cidade, que tende a aumentar com o tempo em razão da expansão urbana (WU *et al.*, 2021). O gráfico da função é mostrado na Figura 1, com as constantes $\alpha = 4$, $c = 0,05$ e $D = 30$.

Jiao (2015) empregou um método não linear de mínimos quadrados para ajuste da curva da função (Equação 3.3) aos dados das cidades analisadas em seu estudo. Esse método envolve o ajuste de uma função não linear aos dados observados, refinando os parâmetros por meio de iterações sucessivas. Existem diversos algoritmos para realização de tais ajustes, incluindo os métodos *Trust-region*, *Levenberg-Marquardt*, *Gauss-Newton* e *Powell's dog leg*, sendo que o mencionado autor utilizou o método *Trust-region*.

A função ajustada descreve quantitativamente a tendência de decaimento da densidade territorial urbana da área central para a periférica da cidade (JIAO, 2015). Sua aplicação possibilita caracterizar a forma urbana e a dinâmica de crescimento territorial urbano, viabilizando distinguir o padrão disperso da forma compacta (XU *et al.*, 2019).

Para compreender a diminuição da densidade territorial urbana do núcleo urbanizado para as áreas urbanas periféricas, isto é, a taxa decrescente e sua variação, Jiao (2015) examinou as derivadas das curvas ajustadas para a densidade territorial urbana. A primeira derivada da função descreve a taxa decrescente da densidade, e a segunda derivada descreve a taxa de variação da taxa decrescente da densidade.

Figura 1 – Gráfico do modelo da função sigmoide S-invertido.



Fonte: Adaptado de Jiao (2015, p. 29).

Ao examinar a primeira derivada, Jiao (2015) inferiu que as coordenadas do ponto decrescente mais rápido (denotado por p_0 na Figura 2) na curva da função de densidade territorial urbana são:

$$r_0 = \frac{D}{2}, f_0 = \frac{1+c}{2} \quad (3.4)$$

A segunda derivada tem dois extremos, que denotam os locais onde a taxa de diminuição muda mais fortemente. Os dois pontos correspondentes na curva de densidade territorial urbana são denotados por p_1 e p_2 na Figura 2. As coordenadas de p_1 e p_2 são representadas pelas Equações 3.5 e 3.6, respectivamente:

$$r_1 = \frac{D}{2} \left(\frac{-1.316957}{\alpha} + 1 \right), f_1 = 0.788675(1-c) + c \quad (3.5)$$

$$r_2 = \frac{D}{2} \left(\frac{1.316957}{\alpha} + 1 \right), f_2 = 0.211325(1-c) + c \quad (3.6)$$

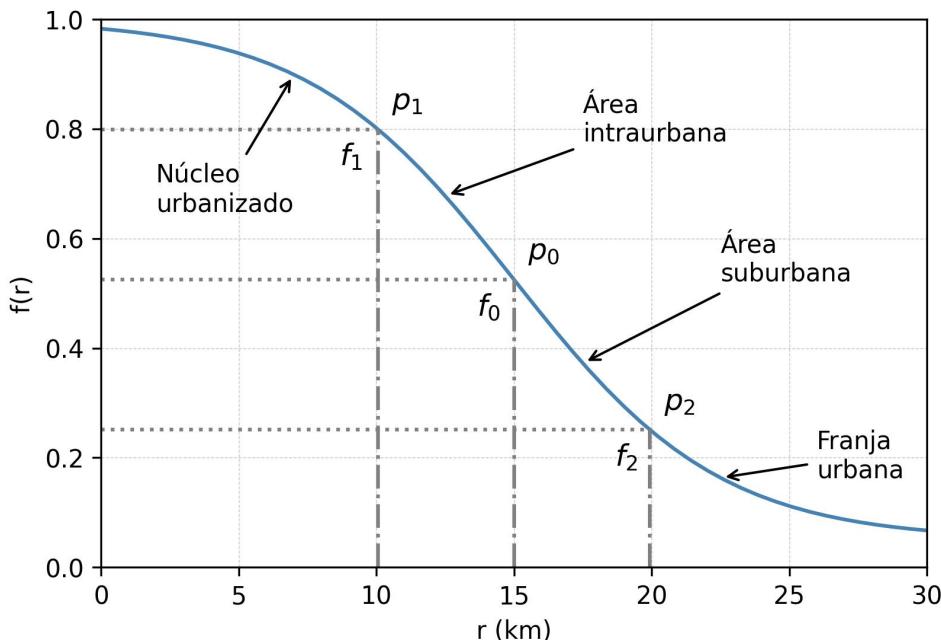
Sendo que:

$$r_1 + r_2 = D, f_1 + f_2 = 1 + c \quad (3.7)$$

em que r_1 e r_2 são raios que determinam, respectivamente, o limite do núcleo urbanizado (*urban core*) e o início da franja urbana (*urban fringe*, com densidade mínima

c). A extensão das áreas intraurbana (*inner urban area*, com densidade máxima f_1) e suburbana (*suburban area*, com densidade mínima igual a f_2) é caracterizada pela diferença entre esses dois raios. A extensão da área urbana principal da cidade analisada (incluindo o núcleo urbanizado, áreas intraurbanas, áreas suburbanas, e a franja urbana) é descrita pelo parâmetro D (JIAO, 2015). Evidentemente que r_0 , r_1 e r_2 são funções de D e α , e f_0 , f_1 e f_2 são funções de c . Portanto, os parâmetros da função denotam algumas características básicas da forma urbana da cidade (Figura 2).

Figura 2 – Interpretação do gráfico do modelo da função sigmoide S-invertido.



Fonte: Adaptado de Jiao (2015, p. 33).

É importante destacar que esse modelo partitiona a área urbana conforme a variação da densidade territorial urbana (JIAO, 2015). Nesse sentido, p_1 , p_0 e p_2 auxiliam a dividir nitidamente a área urbana, de maneira concêntrica, definindo quatro áreas circulares do centro urbano para a periferia: núcleo urbanizado, área intraurbana, área suburbana e franja urbana (Figura 2). Dessa forma, o núcleo urbanizado é a área central de uma cidade, com densidade elevada que diminui lentamente. A área intraurbana e a área suburbana possuem uma densidade em queda, sendo que a densidade da primeira é superior a f_0 . A franja urbana circunda a área urbana principal, possuindo densidade reduzida, que decresce de forma relativamente lenta. Em vista disso, essa função também é frequentemente adotada para delimitação do limite urbano, pois, por mais que as cidades tenham seus limites administrativos, as formas de delimitar tais limites variam significativamente entre países (CHAKRABORTY *et al.*, 2022).

A compacidade das cidades também pode ser caracterizada pela função sigmoide proposta. Como a densidade territorial urbana de uma cidade mais compacta é elevada

na área central e diminui acentuadamente para um valor reduzido na área circundante ao núcleo, as curvas de densidade construída nas cidades mais compactas são acentuadas, diminuindo lentamente nas cidades em expansão. Portanto, a inclinação da curva na função sigmoide reflete a densidade territorial urbana geral. Diante disso, Jiao (2015) propõe a seguinte forma de cálculo da inclinação:

$$k_s = \frac{f_1 - f_2}{r_2 - r_1} = \frac{0.57735(1 - c)\alpha}{1.316957D} \quad (3.8)$$

em que D , c e α são os parâmetros da função sigmoide ajustada.

Conforme salientado pelo autor, embora o k_s possa fornecer uma estimativa intuitiva sobre a densidade territorial urbana, esse índice pode levar a uma interpretação tendenciosa quando examinado em cidades com tamanhos diferentes (JIAO, 2015). Isso porque o k_s é determinado por três parâmetros, sendo o D uma variável mutável entre cidades. Como um D elevado ocasiona em um k_s reduzido, o valor desse índice para uma cidade tende a diminuir conforme seu crescimento urbano, com cidades maiores possuindo valores de k_s menores. Em outras palavras, o k_s é útil na comparação de densidades apenas quando usado para cidades com extensões territoriais iguais ou semelhantes.

Um outro índice foi proposto pelo autor para contornar esse viés. Tal índice (k_p) reflete o grau de dispersão (*degree of dispersion*) via proporção da parte da curva que decresce rapidamente, ou seja, a parte que denota as áreas urbanas e suburbanas internas (JIAO, 2015), sendo representado da seguinte forma:

$$k_p = \frac{r_2 - r_1}{D} = \frac{1.316957}{\alpha} \quad (3.9)$$

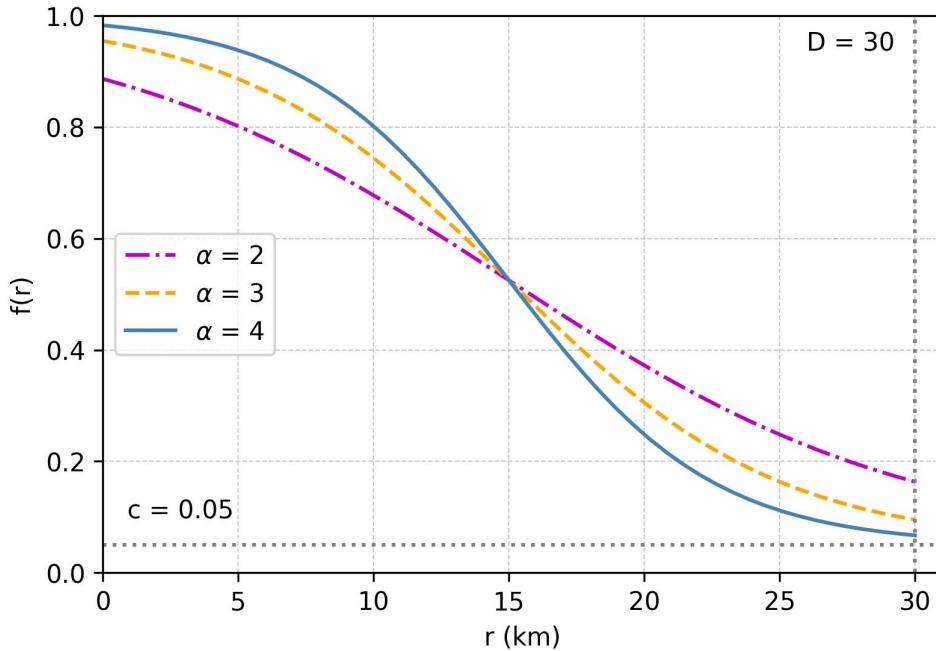
em que α é o parâmetro que controla a inclinação da curva da função de densidade territorial urbana (Figura 3).

Desta forma, a compacidade pode ser analisada com base no k_p . Valores menores desse índice indicam maior compacidade, enquanto valores mais elevados indicam maior dispersão. O k_p também pode ser utilizado para medir o grau de expansão (*degree of sprawling*), por meio da comparação de seus valores em dois momentos distintos (para diferenciação, esse índice comparativo será denominado k_{ps}). Um aumento no valor do k_{ps} de uma cidade no tempo implica uma diminuição em sua compacidade, indicando uma provável expansão de maneira mais dispersa (JIAO, 2015).

Por fim é calculado o índice *Sprawl Ratio* (S_r):

$$S_r = \frac{\delta r_2}{\delta r_1} = \frac{(r_2^i - r_2^{i-1})r_1^{i-1}}{r_2^{i-1}(r_1^i - r_1^{i-1})} \quad (3.10)$$

Figura 3 – Curvas da função sigmoide S-invertido com diferentes valores de alpha.



Fonte: Adaptado de Wu *et al.* (2021, p. 5).

em que S_r representa o grau de espraiamento urbano; δr_1 e δr_2 representam as taxas de crescimento de r_1 e r_2 , respectivamente; r_1^i e r_2^i são os raios do núcleo urbano e da zona suburbana de uma cidade no tempo i , respectivamente; e r_1^{i-1} e r_2^{i-1} são os raios em um ponto no tempo anterior. Importante destacar que o S_r pode produzir valores irrazoáveis em alguns poucos casos em que o δr_1 se aproximar de zero, ou se δr_1 e δr_2 tiverem valores negativos.

Por conseguinte, o S_r determina a magnitude do aumento ou diminuição no espraiamento urbano entre os cenários final (i) e inicial ($i - 1$) de um dado período (JIAO, 2015; MENZORI, 2021).

Neste estudo, criamos dez *buffers* (polígonos circulares) com espaçamento regular de 1 km de distância cada (raios), de forma crescente e gradual, a partir dos centros históricos das cidades da amostra de municípios. Após calcular a diferença entre os *buffers*, cada cidade passou a ter uma série de anéis concêntricos isométricos, estendendo-se a partir de seu centro urbano.

Para os anéis de cada cidade, calculamos a densidade territorial urbana como a proporção da área urbana, ou seja, área construída em um anel dividida pela área total do respectivo anel.

Escolhemos a distância de 10 km para o anel externo (D) por ser um tamanho suficiente para incorporar áreas urbanas funcionais contínuas nos municípios da amostra, um dos critérios definidos por Jiao (2015) para estipular a distância do limite final

de uma cidade. Para que os anéis externos de 10 km não ultrapassassem os limites territoriais de seus respectivos municípios, fizemos um recorte desses anéis pelos limites municipais correspondentes. Ademais, como em algumas cidades não havia área urbana em determinados anéis, foram ignoradas as distâncias dos anéis em que a densidade era igual 0 (zero), bem como adotado um limite de densidade para exclusão de anéis com densidade muito baixa, da seguinte forma: nos casos em que a densidade territorial urbana em determinado anel fosse igual ou inferior a 0,05, eram retidos apenas os valores de densidade até o anel adjacente, ignorando-se os posteriores.

Em seguida, ajustamos a função sigmoide, também por meio do algoritmo *Trust-region*, que tem sua implementação em Python baseada nos trabalhos de Byrd, Schnabel e Shultz (1988) e Branch, Coleman e Li (1999). E, na sequência, armazenamos os coeficientes e índices da função ajustada ($\alpha, c, D, r_1, r_2, f_1, f_2, k_s, k_p$) em uma matriz, com as cidades como observações, para cada um dos oito anos analisados. Os valores de S_r e k_{ps} foram calculados como a diferença entre determinado ano e o respectivo ano anterior, formando sete intervalos temporais: 1985-1990, 1990-1995, 1995-2000, 2000-2005, 2005-2010, 2010-2015, 2015-2020.

3.2.3 Etapa 3: Análises de agrupamento

Como mencionado anteriormente, o propósito deste trabalho é averiguar se existem padrões discerníveis de crescimento urbano nos municípios amostrados. Para tanto, fizemos análises de agrupamento dos municípios, com base nos valores do S_r (*Sprawl ratio*) e do k_{ps} (*degree of sprawling*). Tais índices são relativos, pois comparam valores entre anos distintos. O S_r retrata a taxa de espraiamento urbano e o k_{ps} indica a compacidade ou dispersão urbana em momentos distintos.

Essas análises de agrupamento se fundamentaram em quatro algoritmos de aprendizado não supervisionado: *K-Means*, selecionado por sua simplicidade e popularidade; *K-Medoids*, selecionado por sua robustez a valores extremos e interpretabilidade dos centros dos *clusters*; *Mean Shift*, selecionado por não requerer a definição prévia do número de *clusters*; e *Affinity Propagation*, selecionado por sua simplicidade e robustez a valores extremos, por não requerer a definição prévia do número de *clusters* e pela interpretabilidade dos centros dos *clusters*. A seguir são detalhados os hiperparâmetros ajustados para cada algoritmo (foi mantida a configuração padrão nos hiperparâmetros não mencionados):

- *K-Means*:
 - Método de inicialização do centroide (`init`): *k-means++*².

² No *k-means++*, o primeiro centroide é escolhido aleatoriamente, e os centroide seguintes são aqueles que estão mais distantes desse centroide inicial, sendo selecionados com uma probabilidade proporcional à sua distância quadrada do centroide existente mais próximo

- Número randômico para inicialização do centroide e para tornar a aleatoriedade determinística (`random_state`): 42.

- *K-Medoids*

- Algoritmo a ser utilizado (`method`): pam.
- Método de inicialização do medoide (`init`): *k-medoids++*³.
- Número máximo de iterações ao ajustar (`max_iter`): 0 (significa que o algoritmo deve ser executado até a convergência, sem um número máximo predefinido de iterações).
- Número randômico para inicialização do medoide e para tornar a aleatoriedade determinística (`random_state`): 42.

- *MeanShift*

- Define se as localizações iniciais do *kernel* são as de todos os pontos (configuração padrão, isto é, *False*), ou se as localizações iniciais do kernel são a da versão discretizada dos pontos (`bin_seeding`): *True* (esta opção acelera a execução do algoritmo, pois menos sementes são inicializadas).
- Número de processos simultâneos que devem ser usados para rotinas que são paralelizadas (`n_jobs`): -1 (significa usar todos os processadores).

- *Affinity Propagation*

- Gerador de números pseudo-aleatórios (randômicos) para controlar o estado inicial e tornar a aleatoriedade determinística (`random_state`): 42.

Inicialmente, fizemos análises prévias dos dados (exploratórias) para guiar as escolhas a serem feitas nas etapas do processo de agrupamento, via estatísticas descritivas e visualização de dados (MOLIN; JEE, 2021b).

Devido à discrepâncias de escala das variáveis (S_r e k_{ps}), antes de realizar as análises de agrupamento, redimensionamos as variáveis de entrada consoante à presença de valores negativos e/ou positivos, da seguinte forma: se apenas valores positivos estivessem presentes, o intervalo seria [0, 1]; se apenas valores negativos estivessem presentes, o intervalo seria [-1, 0]; e se valores negativos e positivos estivessem presentes, o intervalo seria [-1, 1]. Nessa forma de redução do impacto das diferentes dimensões dos dados, cada recurso é redimensionado pelo seu valor absoluto máximo (Equação 3.11), sendo suscetível

(REDDY; VINZAMURI, 2014). Em comparação com a inicialização aleatória, esse processo tende a resultar em desempenho superior, em termos de acurácia e velocidade (ARTHUR; VASSILVITSKII, 2007).

³ Segue uma abordagem baseada no *k-means++*.

à presença de valores discrepantes (*outliers*). No entanto, sua escolha se justifica pela natureza dos dados, em que interpretações distintas dos resultados são associadas a valores positivos ou negativos. Também definimos uma semente para garantia de reprodutibilidade do estudo. E utilizamos a Silhueta (*silhouette*) para determinar o número de *clusters* (k) no *K-Means* e no *K-Medoids*, visto que ambos algoritmos requerem a definição desse hiperparâmetro.

$$\frac{X_i}{|X_{max}|} \quad (3.11)$$

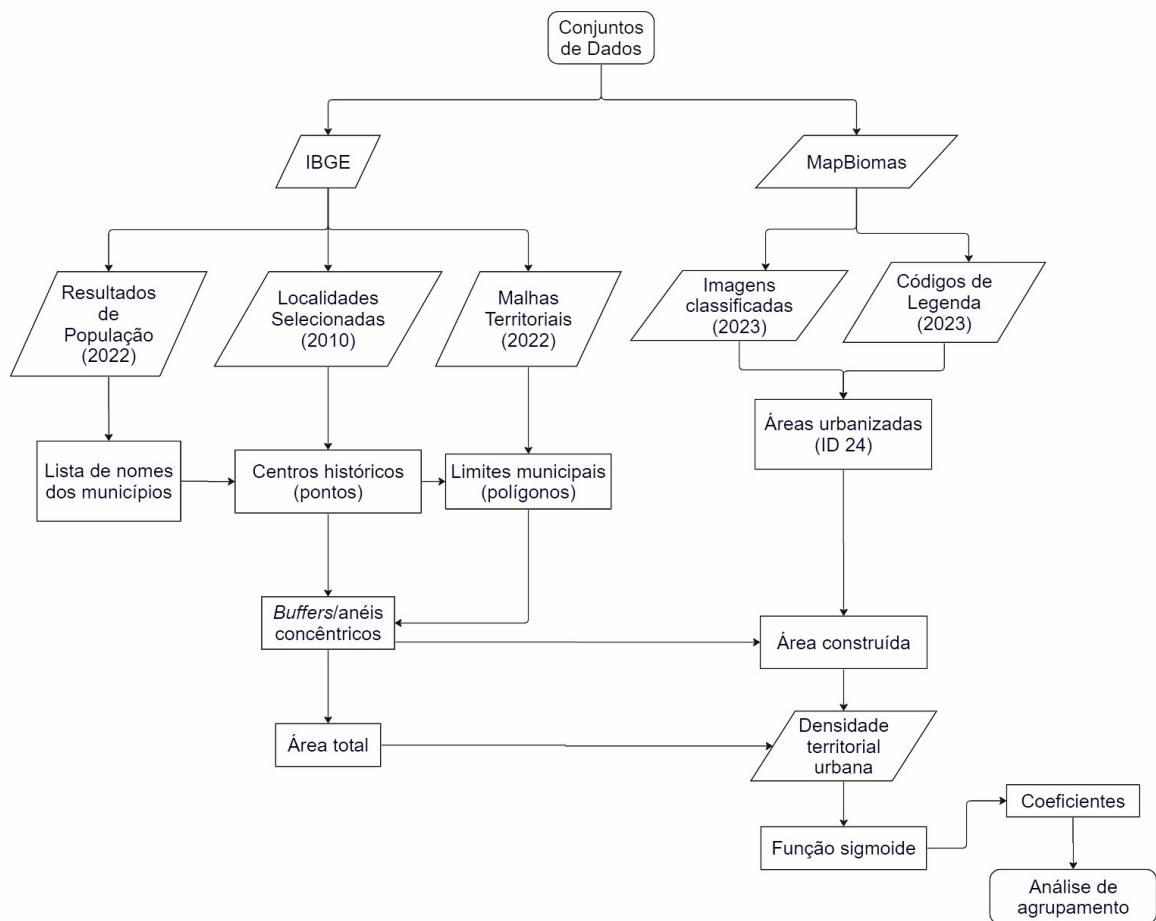
em que X_i é um dos pontos do conjunto de dados e $|X_{max}|$ é o valor absoluto máximo.

De modo a evidenciar os resultados mais consistentes dentre os obtidos pelos algoritmos de agrupamento selecionados, aplicamos índices de validação baseados em critérios internos como índices relativos (CARVALHO *et al.*, 2011b). Nesse sentido, adotamos as métricas Silhueta, Calinski–Harabasz e Davies–Bouldin como medidas de comparação entre os agrupamentos formados (HALKIDI; VAZIRGIANNIS; HENNIG, 2015).

Para o desenvolvimento deste trabalho, utilizamos a linguagem de programação *Python* (versão 3.10.12) no ambiente *JupyterLab* (versão 4.0.9). O material cartográfico⁴ foi produzido com auxílio do *software QGIS* (versão 3.28.10). A Figura 4 sintetiza as etapas envolvidas no método proposto.

⁴ Disponibilizado com os dados e códigos das análises no seguinte repositório do GitHub:
https://github.com/icn-sousa/TCC_MBA_CD.git

Figura 4 – Diagrama dos métodos empregados.



Fonte: Elaboração própria (2023).

4 RESULTADOS

Neste capítulo detalhamos os resultados deste trabalho de conclusão de curso. Na seção 4.1 descrevemos uma breve caracterização da amostra analisada. A seção 4.2 abrange a análise das variações de densidade territorial urbana nos municípios paulistas de porte médio, obtidos via implementação e ajuste da função sigmoide. A seção 4.3 traz os resultados da aplicação das técnicas de agrupamento de dados para identificação de padrões de crescimento urbano nos municípios paulistas analisados.

4.1 Caracterização da amostra

Conforme mencionado na seção 3.2.1, foram selecionados municípios paulistas com população entre 100 e 500 mil habitantes. Dessa listagem inicial de 69 municípios, foram excluídos 16 municípios litorâneos paulistas (CETESB, 2022), e 38 municípios integrantes das seis regiões metropolitanas¹ do Estado de São Paulo (IGC, 2023), totalizando 15 municípios na amostra (Tabela 1 e Figura 5).

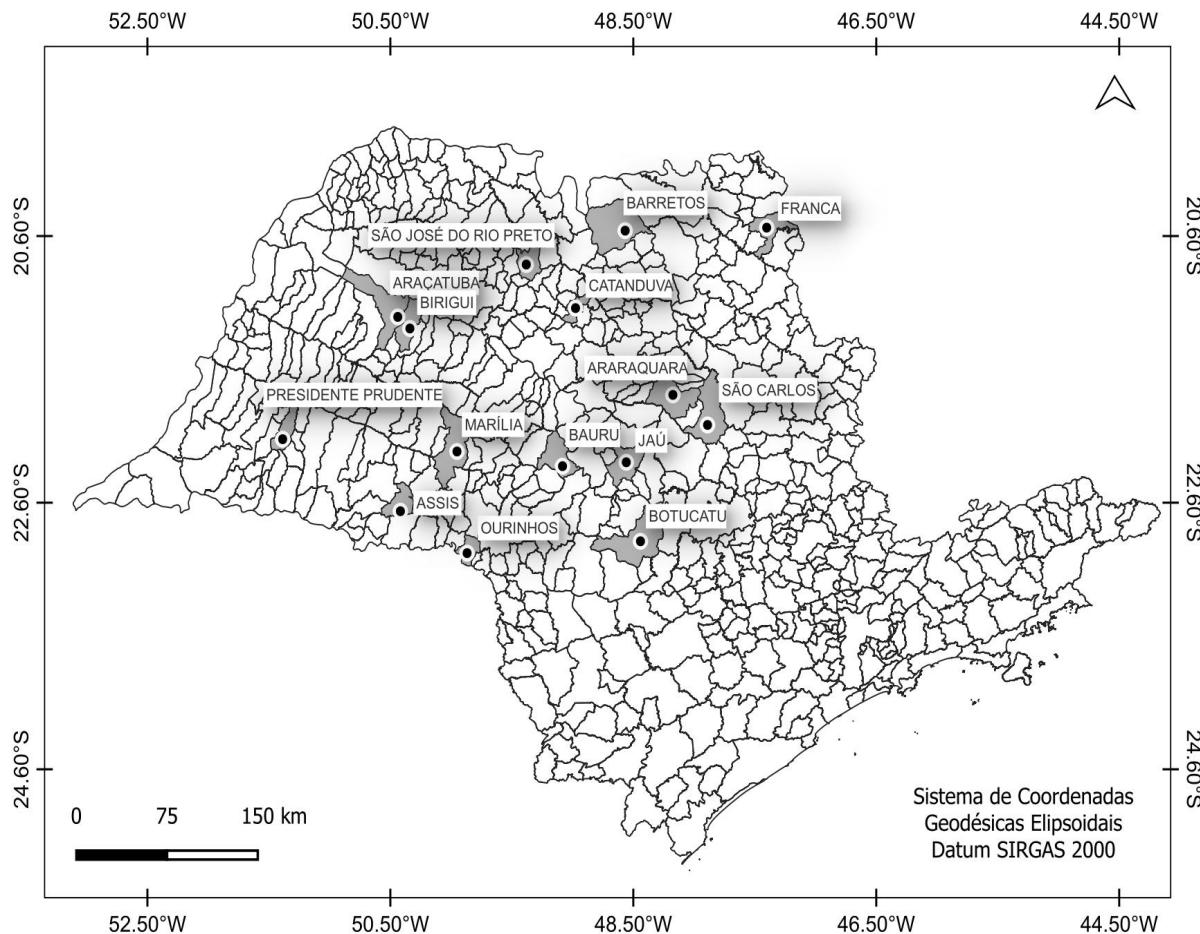
Tabela 1 – População residente, Área territorial e Densidade demográfica dos 15 municípios da amostra.

Município	População (hab.)	Área (km ²)	Densidade (hab./km ²)
ARAÇATUBA	200.124	1.167	171.47
ARARAQUARA	242.228	1.004	241.35
ASSIS	101.409	461	220.16
BARRETOS	122.485	1.566	78.21
BAURU	379.146	668	567.85
BIRIGUI	118.979	530	224.48
BOTUCATU	145.155	1.483	97.9
CATANDUVA	115.785	291	398.44
FRANCA	352.537	606	582.05
JAÚ	133.497	687	194.29
MARÍLIA	237.629	1.171	203.01
OURINHOS	103.970	296	351.47
PRESIDENTE PRUDENTE	225.668	561	402.52
SÃO CARLOS	254.822	1.137	224.14
SÃO JOSÉ DO RIO PRETO	480.439	432	1112.27

Fonte: Adaptado do IBGE (2023a).

¹ Região Metropolitana da Baixada Santista, Região Metropolitana do Vale do Paraíba e Litoral Norte, Região Metropolitana de Sorocaba, Região Administrativa e Metropolitana de Campinas, Região Metropolitana de Ribeirão Preto, Região Metropolitana de São Paulo (IGC, 2023).

Figura 5 – Localização dos 15 municípios da amostra no Estado de São Paulo.



Fonte: Elaboração própria, com base em dados de IBGE (2010), IBGE (2023b).

Na Tabela 2 são apresentadas as estatísticas descritivas da amostra e na Figura 6 os *box plots* da população residente (habitantes), área territorial (km^2), e densidade demográfica (hab./km^2). São José do Rio Preto destoa dos demais municípios da amostra em razão de seu elevado número de habitantes (maior da amostra, com quase 500 mil) e reduzida área territorial (terceira menor da amostra, com 432 km^2), resultando na maior densidade demográfica (hab./km^2) dentre todos os 15 municípios.

4.2 Ajuste da função sigmoide

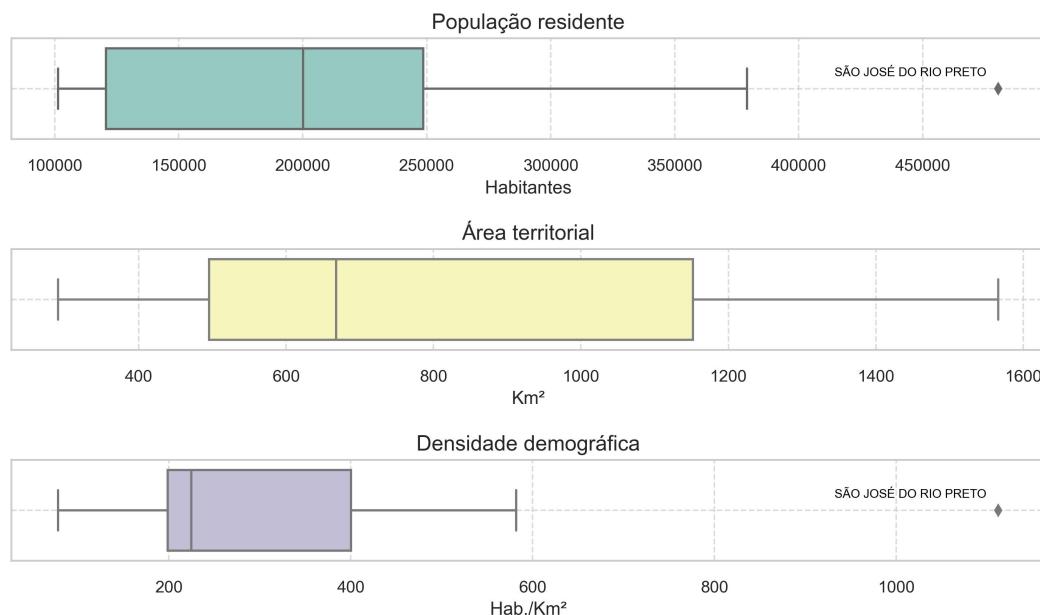
Conforme explicitado na seção 3.2.2, a função sigmoide de densidade territorial urbana foi ajustada para os 15 municípios da amostra em cada um dos oito anos analisados (1985, 1990, 1995, 2000, 2005, 2010, 2015 e 2020). No Apêndice A constam as tabelas com os valores de densidade territorial urbana dos municípios para cada raio, por ano. O Apêndice B engloba as tabelas com os parâmetros de ajuste da função (α , c e D), bem como o coeficiente de determinação (R^2) para cada município/ano, com destaque (sublinhado) para os maiores e menores valores dos coeficientes α e D . Nesse sentido, é

Tabela 2 – Estatísticas descritivas da População residente, Área territorial e Densidade demográfica dos 15 municípios da amostra.

Medida	População (hab.)	Área (km ²)	Densidade (hab./km ²)
Média	214258,2	804	337,974
Desvio padrão	114598,2381	417,6304	261,9579
Mín.	101409	291	78,21
Q1	120732	495,5	198,65
Mediana	200124	668	224,48
Q3	248525	1152	400,48
Máx.	480439	1566	1112,27
Amplitude	379030	1275	1034,06
Coeficiente de Variação (CV)	0,5349	0,5194	0,7751
Amplitude Interquartil (IQR)	127793	656,5	201,83
Coeficiente de dispersão de quartil (QCD)	120732,5142	496,0699	199,154

Fonte: Elaboração própria com base em dados do IBGE (2023a).

Figura 6 – *Box plots* da População residente, Área territorial e Densidade demográfica dos 15 municípios da amostra.



Fonte: Elaboração própria com base em dados do IBGE (2023a).

possível notar que, com base no método de ajuste não linear dos mínimos quadrados, a função se ajusta bem em todos os municípios amostrados (R^2 elevado para os oito anos).

Os gráficos das curvas ajustadas constam nas Figuras 7 a 11. Examinando essas curvas, é possível perceber que, assim como as 28 grandes cidades chinesas avaliadas por Jiao (2015), os 15 municípios da amostra deste estudo também seguem a chamada “regra do formato S-invertido” (“*Inverse S-shape Rule*”) na variação da densidade territorial

urbana conforme o distanciamento do centro urbano.

Uma análise comparativa entre os municípios, com base nas tabelas do Apêndice B, permite identificar aqueles com a forma urbana mais compacta para cada ano, isto é, com os maiores valores de α : Catanduva (1985, 1990 e 2020), Araraquara (1995), e Ourinhos (2000, 2005, 2010 e 2015). Alternativamente, os municípios com a forma urbana menos compacta para cada ano, isto é, com os menores valores de α são: Botucatu (1985, 1990 e 1995), e Marília (2000, 2005, 2010, 2015 e 2020). Como o α controla a inclinação da curva ajustada pela função sigmoide (conforme Figura 3), quanto mais próxima de um formato S-invertido for a curva, mais compacta é a forma urbana no município. Um exemplo disso pode ser visto ao comparar as curvas de 1985 para Botucatu (Figura 9a) e Catanduva (Figura 9b), em que o formato S-invertido é bem mais acentuado para Catanduva (maior α em 1985) do que para Botucatu (menor α em 1985).

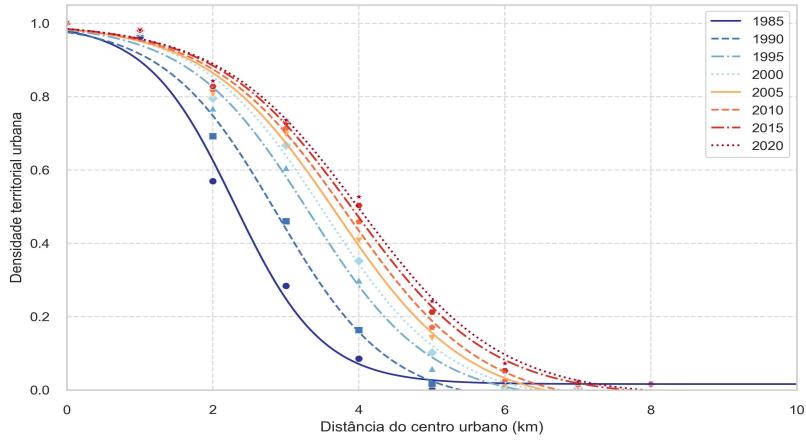
Cabe salientar a existência de alguns valores negativos para o parâmetro c (Apêndice B), que representa a densidade mínima do local em que a função sigmoide está sendo aplicada. Esses valores negativos podem ser explicados por uma limitação inerente à escala de 1 km adotada neste estudo. Isso porque esse distanciamento não confere precisão suficiente para ajustamento da função em todas as cidades da amostra. Em outras palavras, essa delimitação de raios de 1 km pode ser insuficiente para captar, com adequado refinamento, as transições entre as partições das áreas urbanas, pois o método (anéis concêntricos) assume uma homogeneidade do crescimento urbano, algo nem sempre compatível com a realidade dos municípios brasileiros, que têm seu crescimento urbano mais restringido por condicionantes do meio físico (relevo, corpos hídricos, etc.).

A análise comparativa entre os municípios, com base nas tabelas do Apêndice B, também possibilita identificar quais municípios possuem os maiores e menores limites estimados de extensão urbana. Nesse sentido, por meio do parâmetro D , constata-se que as áreas urbanas mais extensas, para todos os oito anos (1985, 1990, 1995, 2000, 2005, 2010, 2015 e 2020), encontram-se em São José do Rio Preto. Já com relação às áreas urbanas menos extensas, há uma maior variabilidade de municípios: Barretos (1985), Ourinhos (1990 e 1995), Birigui (2000) e Assis (2005, 2010, 2015 e 2020).

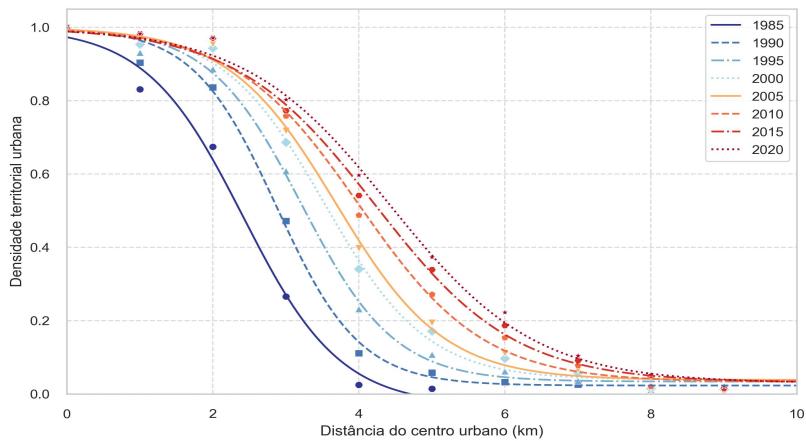
Nas tabelas do Apêndice C são elencados, para cada município/ano, os raios do limite do núcleo urbanizado (r_1) e do início da franja urbana (r_2), a extensão das áreas intraurbanas (f_1) e suburbanas (f_2), além da inclinação na função sigmoide, que reflete a densidade territorial urbana geral (k_s), e do grau de dispersão (k_p), que reflete a proporção da parte da curva que decresce rapidamente, denotando as áreas urbanas e suburbanas internas, consoante ilustrado na Figura 2 (seção 3.2.2).

Nas tabelas do Apêndice D são apresentadas, para cada um dos oito anos (1985, 1990, 1995, 2000, 2005, 2010, 2015 e 2020), estatísticas descritivas para os parâmetros de ajuste da função sigmoide de densidade territorial urbana (α , c e D), das coordenadas dos

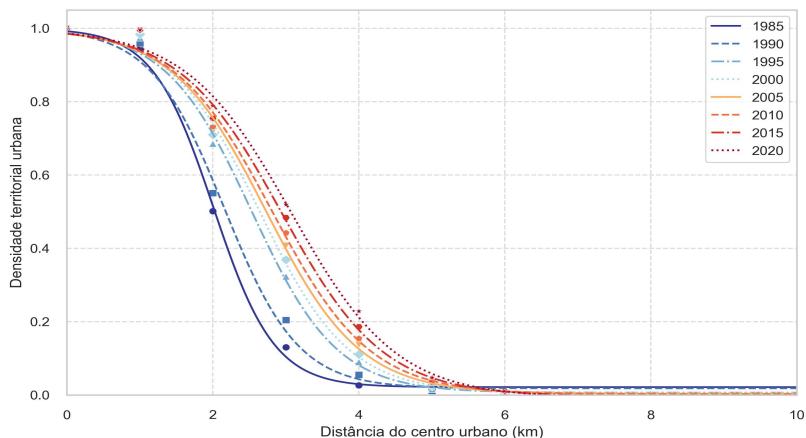
Figura 7 – Gráficos das curvas ajustadas para (a) Araçatuba, (b) Araraquara e (c) Assis.



(a) Araçatuba



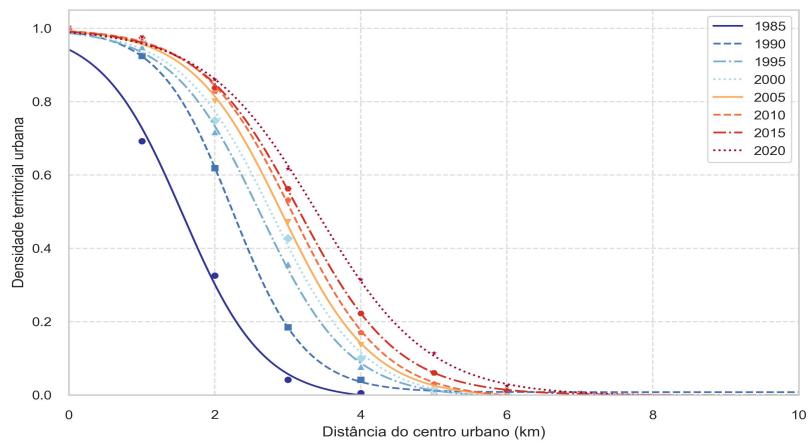
(b) Araraquara



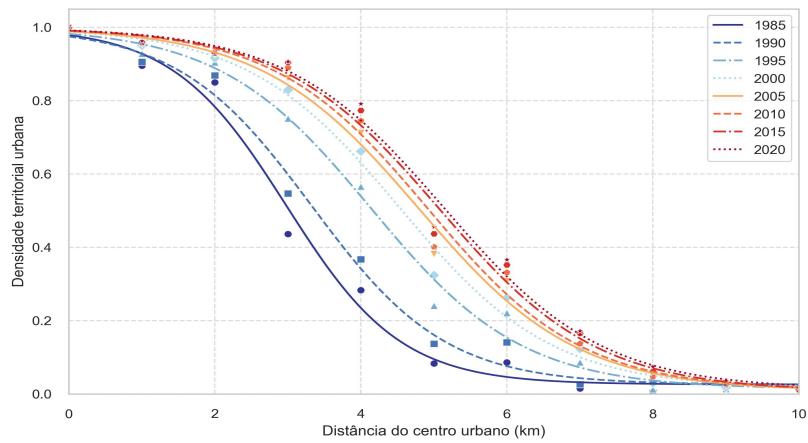
(c) Assis

Fonte: Elaboração própria

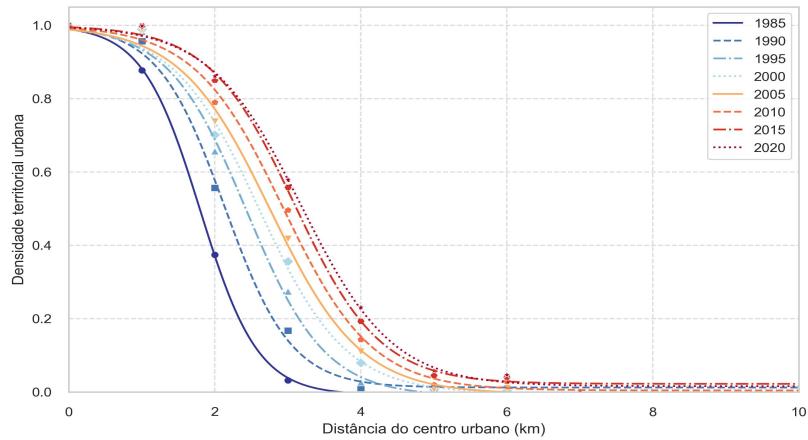
Figura 8 – Gráficos das curvas ajustadas para (a) Barretos, (b) Bauru e (c) Birigui.



(a) Barretos



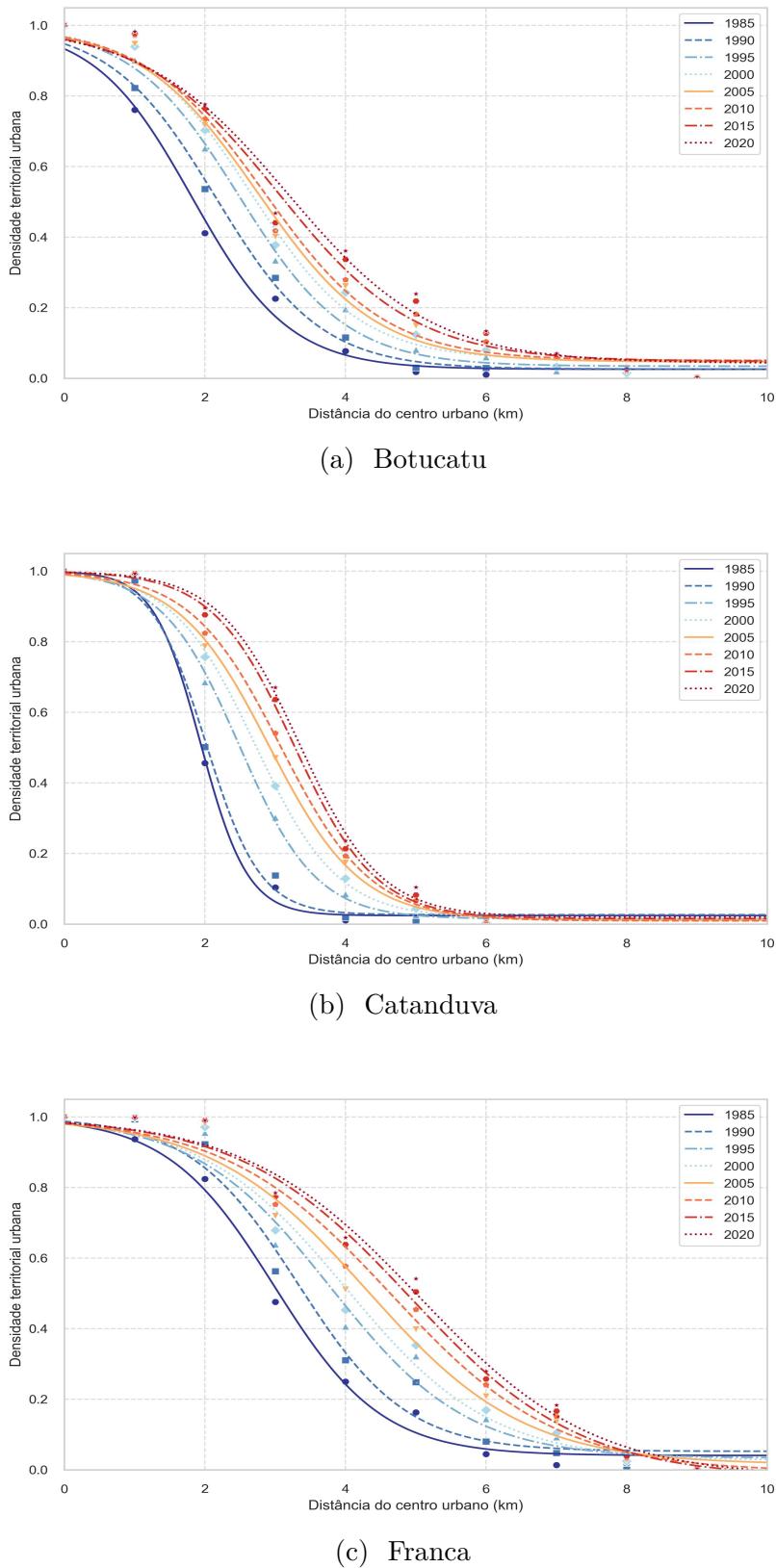
(b) Bauru



(c) Birigui

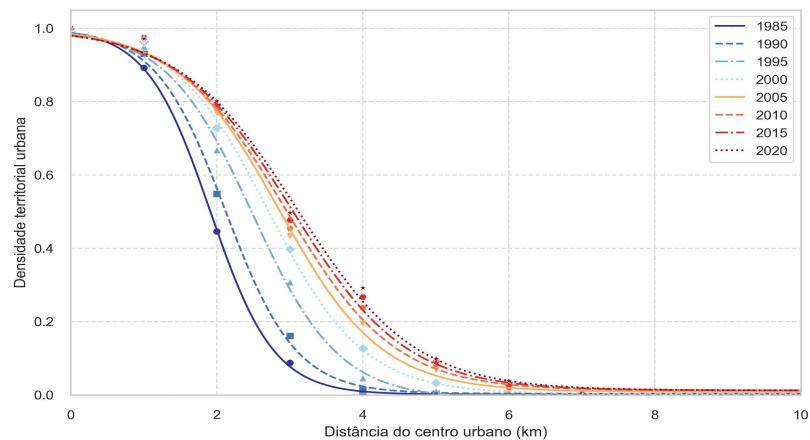
Fonte: Elaboração própria

Figura 9 – Gráficos das curvas ajustadas para (a) Botucatu, (b) Catanduva e (c) Franca.

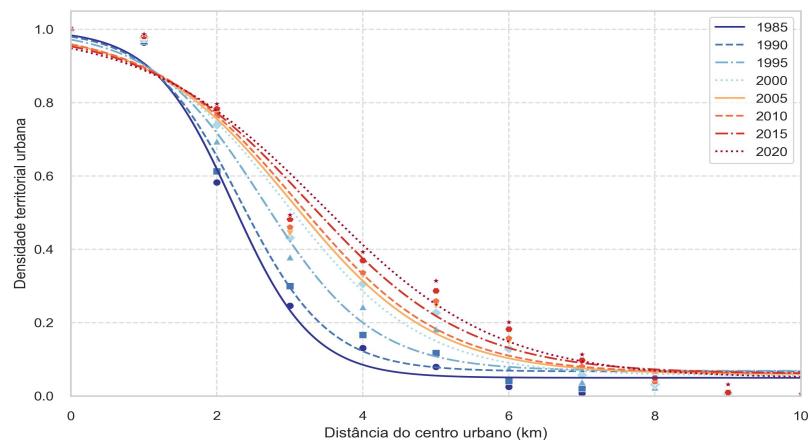


Fonte: Elaboração própria

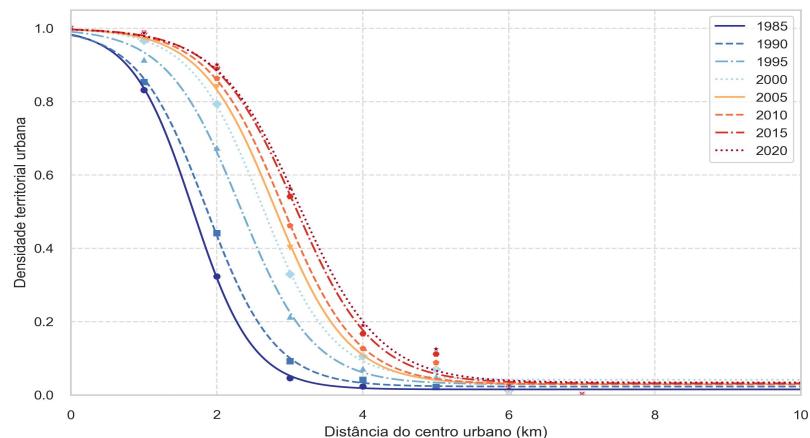
Figura 10 – Gráficos das curvas ajustadas para (a) Jaú, (b) Marília e (c) Ourinhos.



(a) Jaú



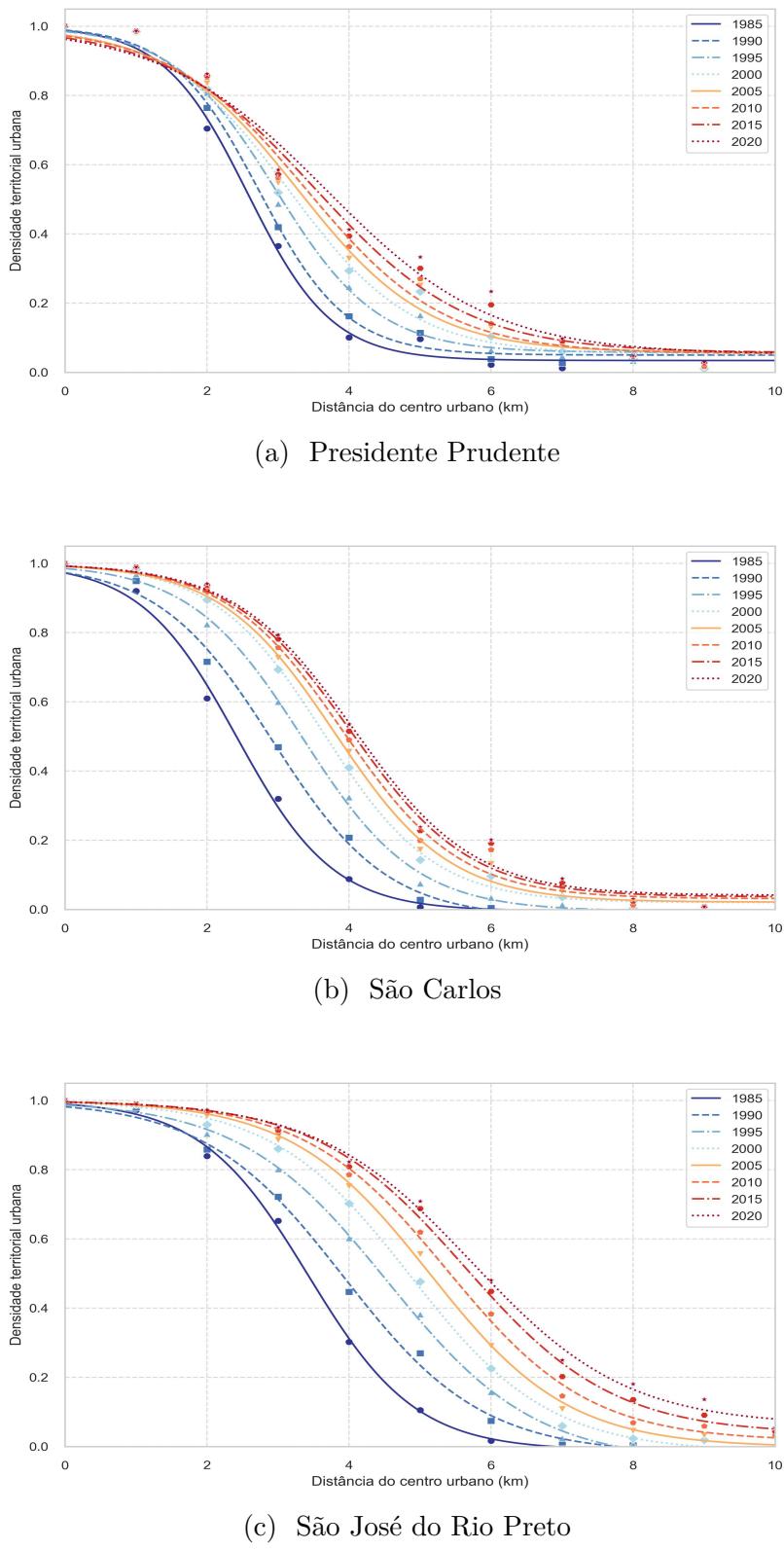
(b) Marília



(c) Ourinhos

Fonte: Elaboração própria

Figura 11 – Gráficos das curvas ajustadas para (a) Presidente Prudente, (b) São Carlos e (c) São José do Rio Preto.



Fonte: Elaboração própria

pontos que caracterizam a forma das áreas urbanizadas (r_1 , r_2 , f_1 e f_2), e das medidas de compacidade urbana (k_s e k_p).

As Figuras 12 a 14 representam os *box plots* dos parâmetros de ajuste da função de densidade territorial urbana (α , c e D), das coordenadas dos pontos que caracterizam a forma das áreas urbanizadas (r_1 , r_2 , f_1 e f_2), e das medidas de compacidade (k_s e k_p). Esses gráficos agrupam os oito anos analisados (1985, 1990, 2000, 2005, 2010, 2015 e 2020). Isso impossibilita discernir os 15 municípios da amostra, mas permite enfatizar as transformações ao longo dos anos. Para essas mesmas variáveis são apresentadas as matrizes de correlação (coeficiente de correlação de *Pearson*) para cada ano nas Figuras 15 a 18.

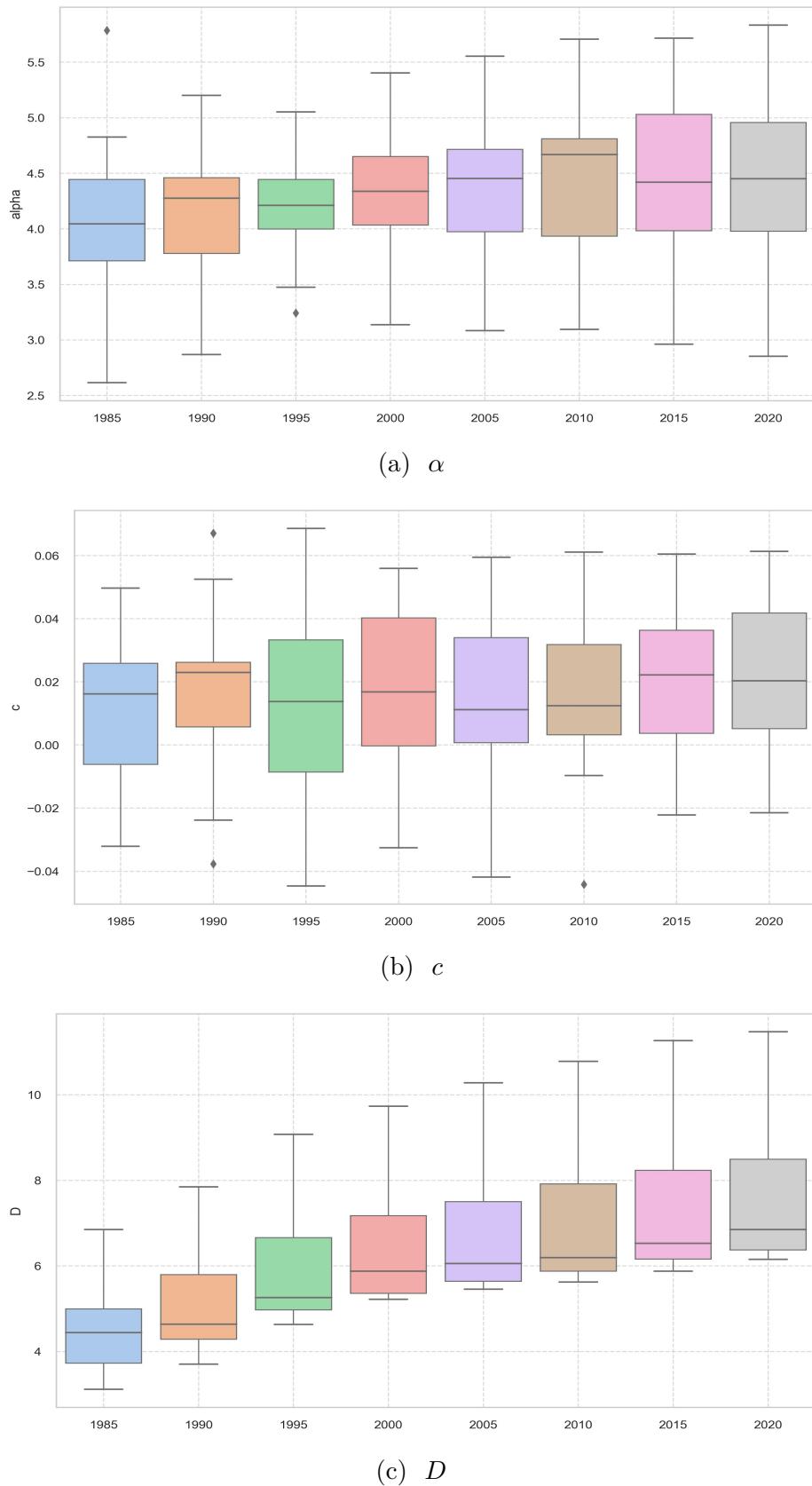
Constata-se que, de maneira geral, houve expansão física das áreas urbanas (crescimento urbano) nos municípios da amostra ao longo dos anos. Isso pode ser verificado pelo aumento gradual nos valores de D (Figura 12c). Esse incremento foi acompanhado pela esperada redução do k_s (Figura 14b), que tende a diminuir com o crescimento urbano (JIAO, 2015), conforme mostrado nas Figuras 15 a 18.

Dentre as constatações advindas da observação das matrizes de correlação para cada ano (Figuras 15 a 18), a perfeita correlação positiva entre c , f_1 e f_2 , bem como a forte correlação positiva entre D , r_1 e r_2 são explicadas pelo fato de r_1 e r_2 serem funções de D , assim como f_1 e f_2 serem funções de c (conforme Equação 3.7 e Figura 2).

Ademais, destaca-se uma relação inversamente proporcional muito forte entre o α e o k_p . Essa relação evidencia como um aumento na compacidade da forma urbana (maior α) resulta em uma consequente diminuição no grau de dispersão (menor k_p), e vice-versa.

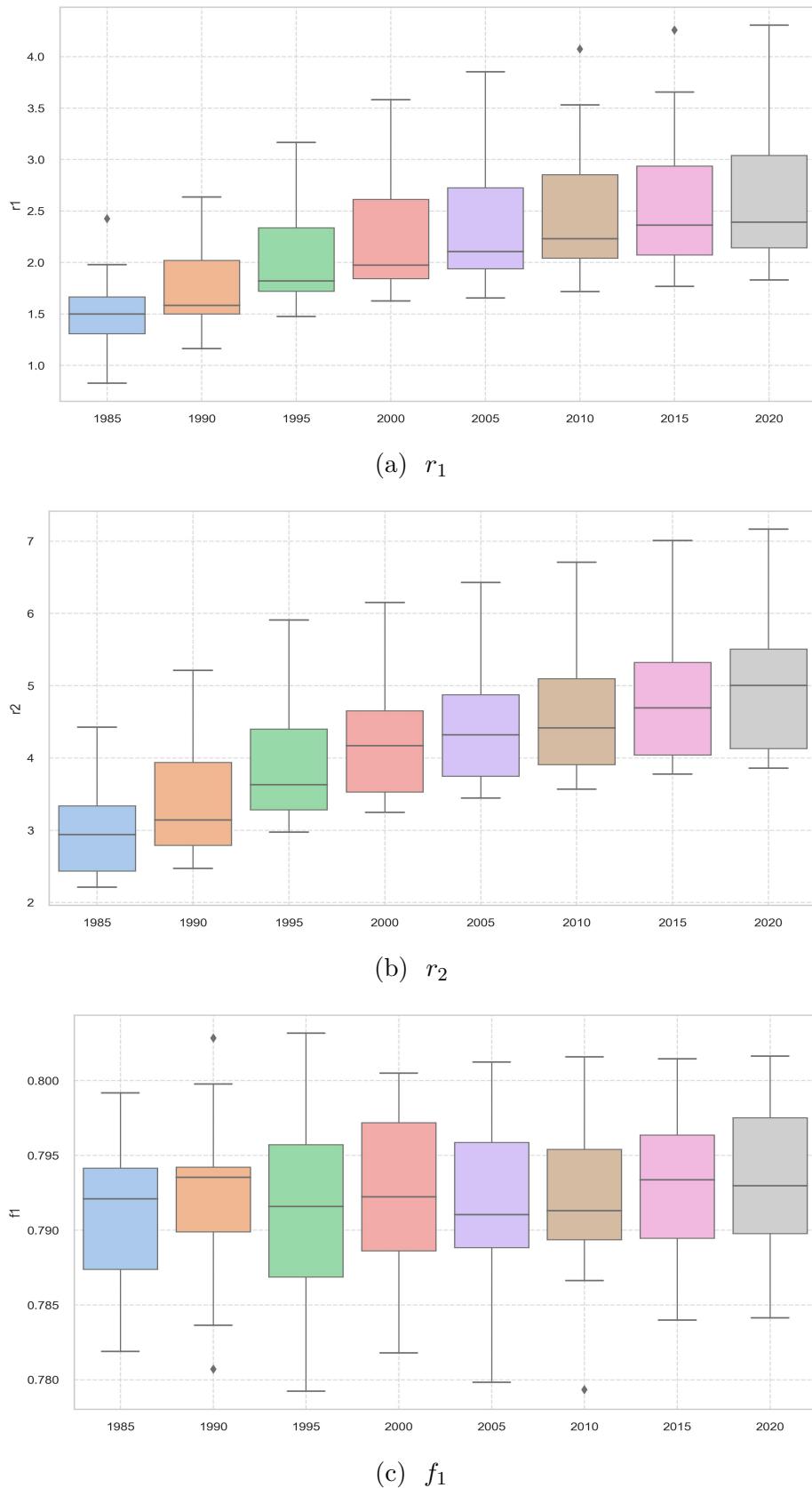
A relação inversamente proporcional (moderada a forte) entre r_2 e k_s é explicada pelo fato de a inclinação da curva ajustada pela função sigmoide, que reflete a densidade territorial urbana geral, poder ser calculada pelo k_s (Equação 3.8). Nesse sentido, quanto maior o r_2 e, consequentemente, menos acentuada a inclinação (k_s) da curva (conforme Figura 2), menor a densidade territorial urbana, indicando um crescimento possivelmente mais expansivo, visto que as curvas de densidade construída são mais acentuadas em cidades mais compactas. Complementarmente, como salientado por Jiao (2015), o k_s fornece uma estimativa intuitiva da densidade territorial urbana. No entanto, esse índice leva a uma interpretação tendenciosa quando é feita uma análise comparativa entre cidades com tamanhos diferentes, pois o k_s é determinado pelo D , uma variável que se modifica de acordo com a cidade (JIAO, 2015). Isso ajuda a explicar a relação negativa (moderada a forte) entre o D e o k_s .

Figura 12 – *Box plots* da variação do (a) α , (b) c e (c) D .



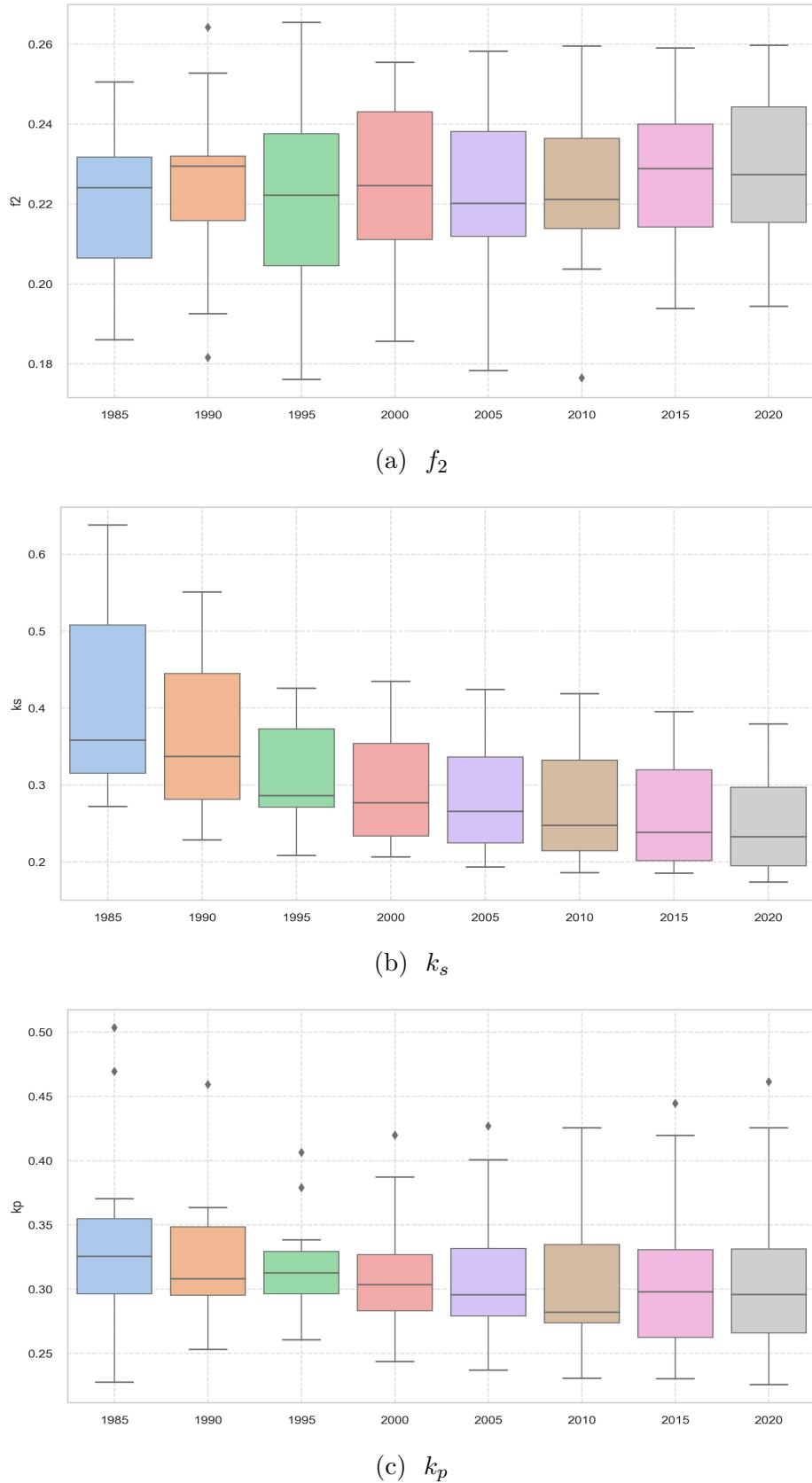
Fonte: Elaboração própria

Figura 13 – *Box plots* da variação do (a) r_1 , (b) r_2 e (c) f_1 .



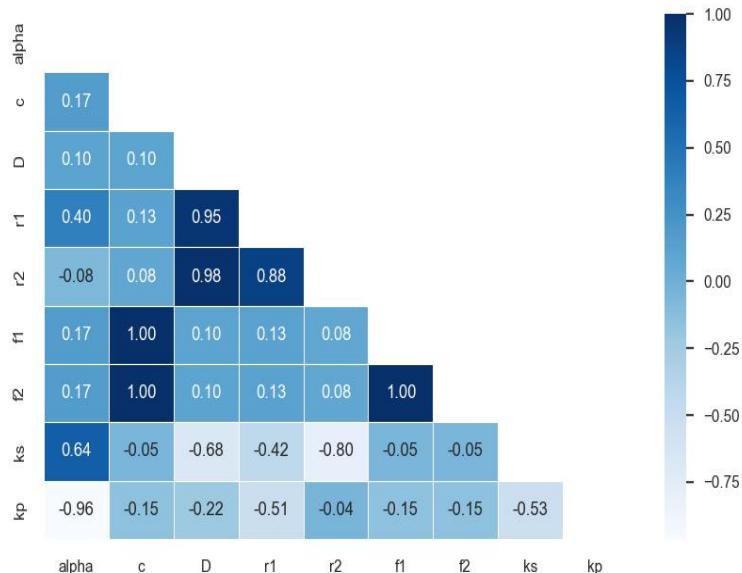
Fonte: Elaboração própria

Figura 14 – Box plots da variação do (a) f_2 , (b) k_s e (c) k_p .



Fonte: Elaboração própria

Figura 15 – Matrizes de correlação para (a) 1985 e (b) 1990.



(a) 1985



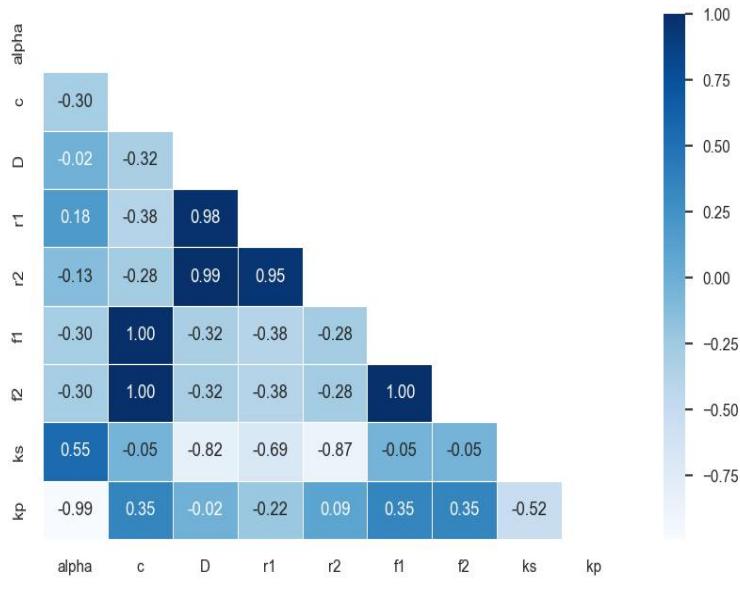
(b) 1990

Fonte: Elaboração própria

4.3 Agrupamentos dos municípios

Conforme mencionado na seção 3.2.3, as análises de agrupamento dos municípios foram feitas com base nos valores do S_r (*Sprawl ratio*) e do k_{ps} (*degree of sprawling*). Como tais índices são relativos, isto é, comparam valores entre anos distintos, a Tabela 3 traz os

Figura 16 – Matrizes de correlação para (a) 1995 e (b) 2000.



(a) 1995



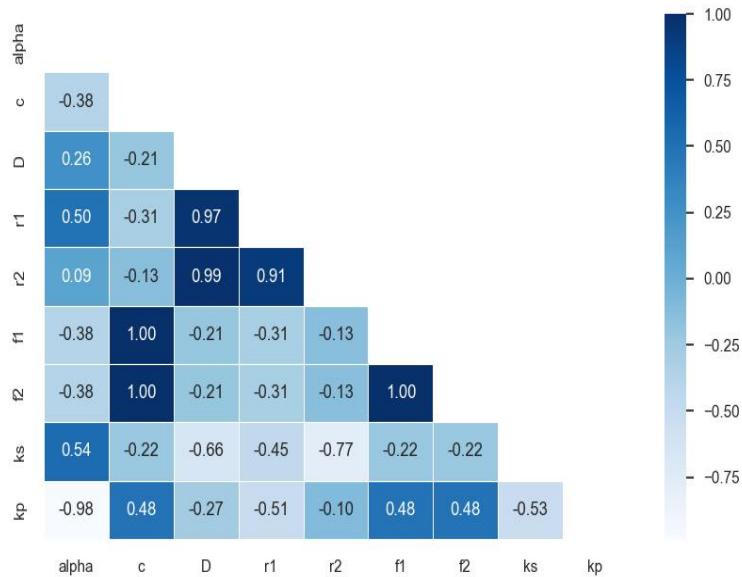
(b) 2000

Fonte: Elaboração própria

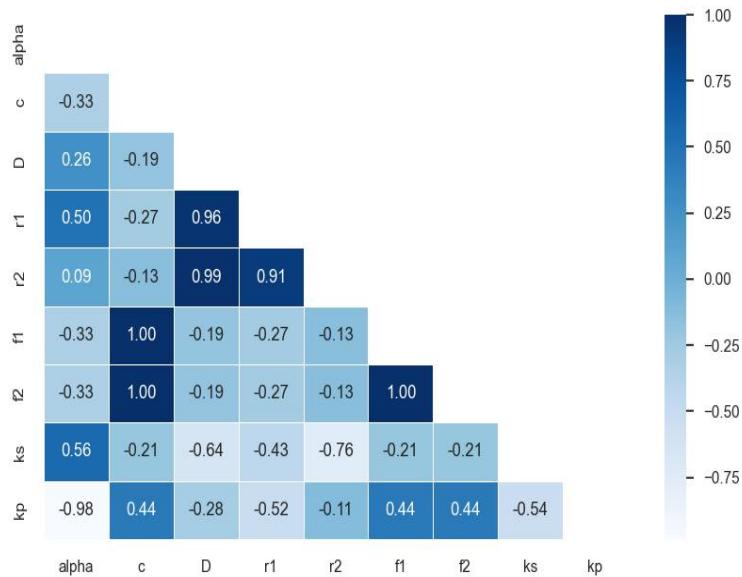
valores do S_r e do k_{ps} por intervalos.

O S_r retrata o quanto aumentou ou reduziu o espraiamento urbano entre dois períodos (MENZORI, 2021), de modo que valores mais elevados indicam aumento no espraiamento urbano (JIAO, 2015). O k_{ps} retrata a compacidade ou dispersão urbana

Figura 17 – Matrizes de correlação para (a) 2005 e (b) 2010.



(a) 2005



(b) 2010

Fonte: Elaboração própria

em momentos distintos, de maneira que seu aumento indica redução na compacidade e uma provável expansão de maneira mais dispersa (JIAO, 2015). Nas Figuras 19 a 22 são ilustradas as relações entre ambas variáveis em cada um dos sete intervalos temporais (1985-1990, 1990-1995, 1995-2000, 2000-2005, 2005-2010, 2010-2015 e 2015-2020). Nota-

Figura 18 – Matrizes de correlação para (a) 2015 e (b) 2020.



(a) 2015



(b) 2020

Fonte: Elaboração própria

se que, com exceção do intervalo de 1990-1995 (Figura 19b), todos os demais períodos apresentaram coeficiente de determinação (R^2) abaixo de 0,70.

Pela análise da Tabela 3, constata-se que Catanduva e Presidente Prudente foram os municípios com aumentos mais expressivos no espraiamento urbano, especialmente entre

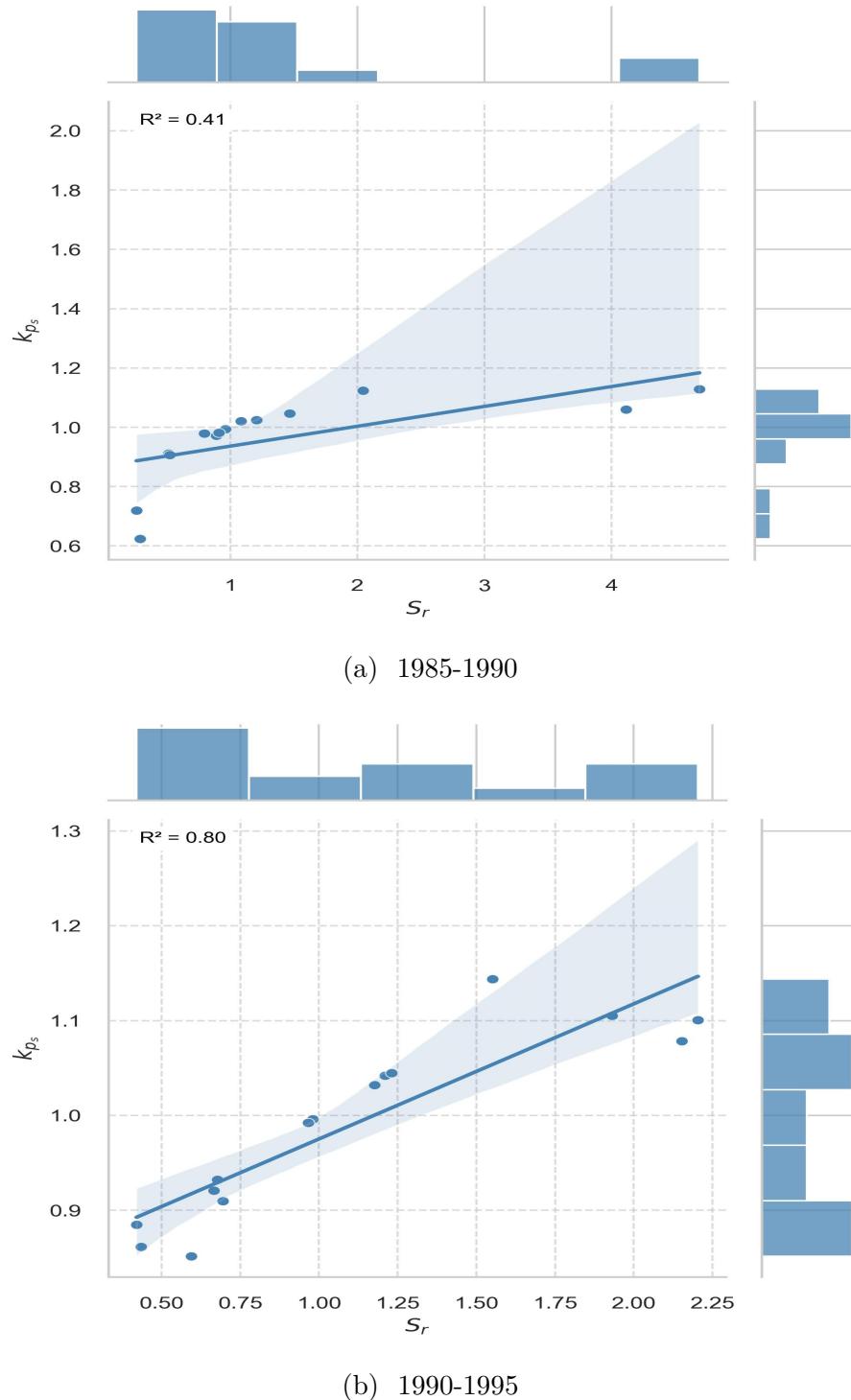
Tabela 3 – Valores do S_r (*Sprawl ratio*) e do k_{ps} (*degree of sprawling*) por intervalos.

S_r por intervalos.							
Município	1985-1990	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015	2015-2020
ARAÇATUBA	1.0850	0.6777	0.6428	1.1121	0.9272	0.7094	1.0247
ARARAQUARA	0.2649	0.9803	1.0180	1.1244	1.7268	2.1703	0.8575
ASSIS	4.6885	0.9656	1.0128	1.4707	1.2196	0.9847	0.7857
BARRETOS	0.2891	1.2104	0.7615	0.4251	0.6849	2.4667	1.7021
BAURU	1.4686	0.6950	0.5331	0.7301	0.3223	0.8972	0.4957
BIRIGUI	0.8927	1.1757	1.2799	0.9102	0.4598	0.3869	2.2530
BOTUCATU	0.5156	0.4206	0.5190	2.8351	0.7761	2.9290	1.4236
CATANDUVA	<u>13.9797</u>	1.5513	0.9066	1.2609	0.4990	0.2329	0.4901
FRANCA	0.5242	1.9318	0.9684	1.1205	0.8453	0.6327	0.8327
JAÚ	0.9615	1.2301	1.0771	1.3807	3.1006	1.3771	1.3813
MARÍLIA	4.1142	2.2038	3.4624	1.9836	0.8726	4.3976	3.0440
OURINHOS	1.2053	0.5939	0.4802	0.8042	0.7181	0.9882	1.8170
PRESIDENTE PRUDENTE	0.7970	2.1522	<u>362.0630</u>	1.6303	1.6000	<u>33.8249</u>	3.9460
SÃO CARLOS	0.9123	0.4339	0.4433	0.9751	0.7696	0.9637	0.6605
SÃO JOSÉ DO RIO PRETO	2.0443	0.6653	0.3133	0.5959	0.7530	0.9953	1.9326
k_{ps} por intervalos.							
Município	1985-1990	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015	2015-2020
ARAÇATUBA	1.0207	0.9321	0.9621	1.0070	0.9960	0.9896	1.0008
ARARAQUARA	0.7197	0.9962	1.0028	1.0088	1.0688	1.0574	0.9889
ASSIS	1.1291	0.9924	1.0008	1.0183	1.0082	0.9990	0.9838
BARRETOS	0.6241	1.0421	0.9755	0.9348	0.9687	1.0396	1.0554
BAURU	1.0471	0.9098	0.9202	0.9728	0.9552	0.9943	0.9844
BIRIGUI	0.9715	1.0320	1.0250	0.9921	0.9227	0.9129	1.0378
BOTUCATU	0.9118	0.8850	0.9527	1.0349	0.9913	1.0566	1.0138
CATANDUVA	1.1122	1.1440	0.9867	1.0241	0.9388	0.8524	0.9636
FRANCA	0.9069	1.1050	0.9977	1.0101	0.9836	0.9650	0.9931
JAÚ	0.9941	1.0449	1.0086	1.0227	1.0358	1.0111	1.0096
MARÍLIA	1.0603	1.1007	1.1075	1.0168	0.9969	1.0444	1.0382
OURINHOS	1.0247	0.8517	0.8576	0.9728	0.9731	0.9987	1.0204
PRESIDENTE PRUDENTE	0.9795	1.0785	1.1275	1.0198	1.0187	1.0577	1.0420
SÃO CARLOS	0.9814	0.8617	0.8959	0.9983	0.9868	0.9985	0.9905
SÃO JOSÉ DO RIO PRETO	1.1244	0.9209	0.8741	0.9492	0.9747	0.9996	1.0206

Fonte: Elaboração própria.

1985-1990 (Catanduva), 1995-2000 e 2010-2015 (Presidente Prudente). No entanto, os valores para esses intervalos (sublinhados na mencionada Tabela 3) são muito discrepantes de todos os demais. É mais provável, portanto, que se tratem de valores irrazoáveis (JIAO, 2015), que ocorrem devido ao denominador (δr_1) da fórmula utilizada para cálculo do S_r (Equação 3.10) se aproximar de 0 (zero) nesses três casos (valores sublinhados na Tabela 45 do Apêndice E).

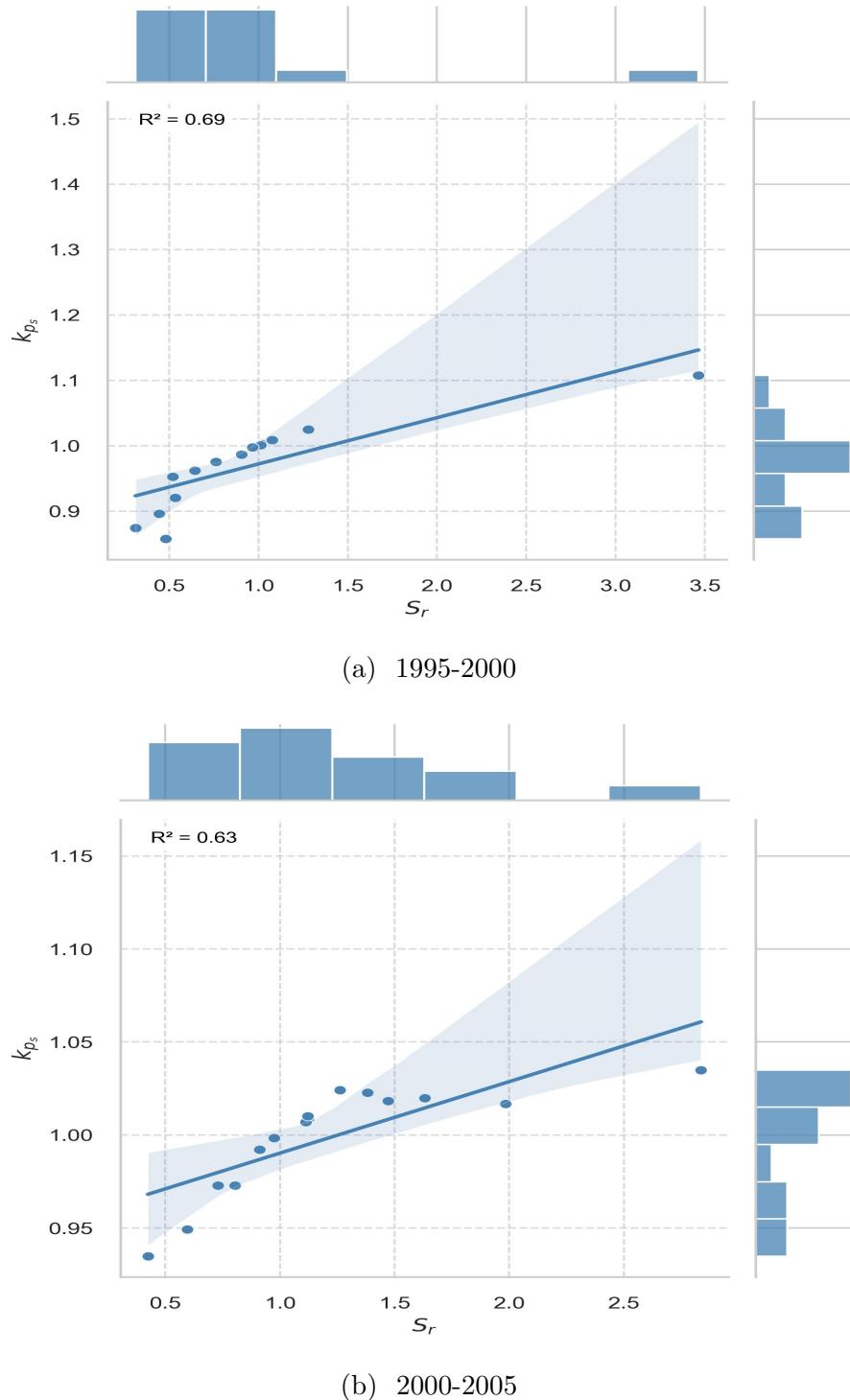
Figura 19 – Relação entre S_r e k_{ps} para os intervalos de (a) 1985-1990 e (b) 1990-1995.



Fonte: Elaboração própria.

Por tais motivos, para os supramencionados períodos, esses municípios foram desconsiderados nas análises de agrupamento, conforme ilustrado nos gráficos de densidade alisada (*kernel density estimate - KDE*) e nas estatísticas descritivas do S_r para cada intervalo, antes e após a remoção dos municípios de Catanduva (1985-1990) e Presidente

Figura 20 – Relação entre S_r e k_{ps} para os intervalos de (a) 1995-2000 e (b) 2000-2005.



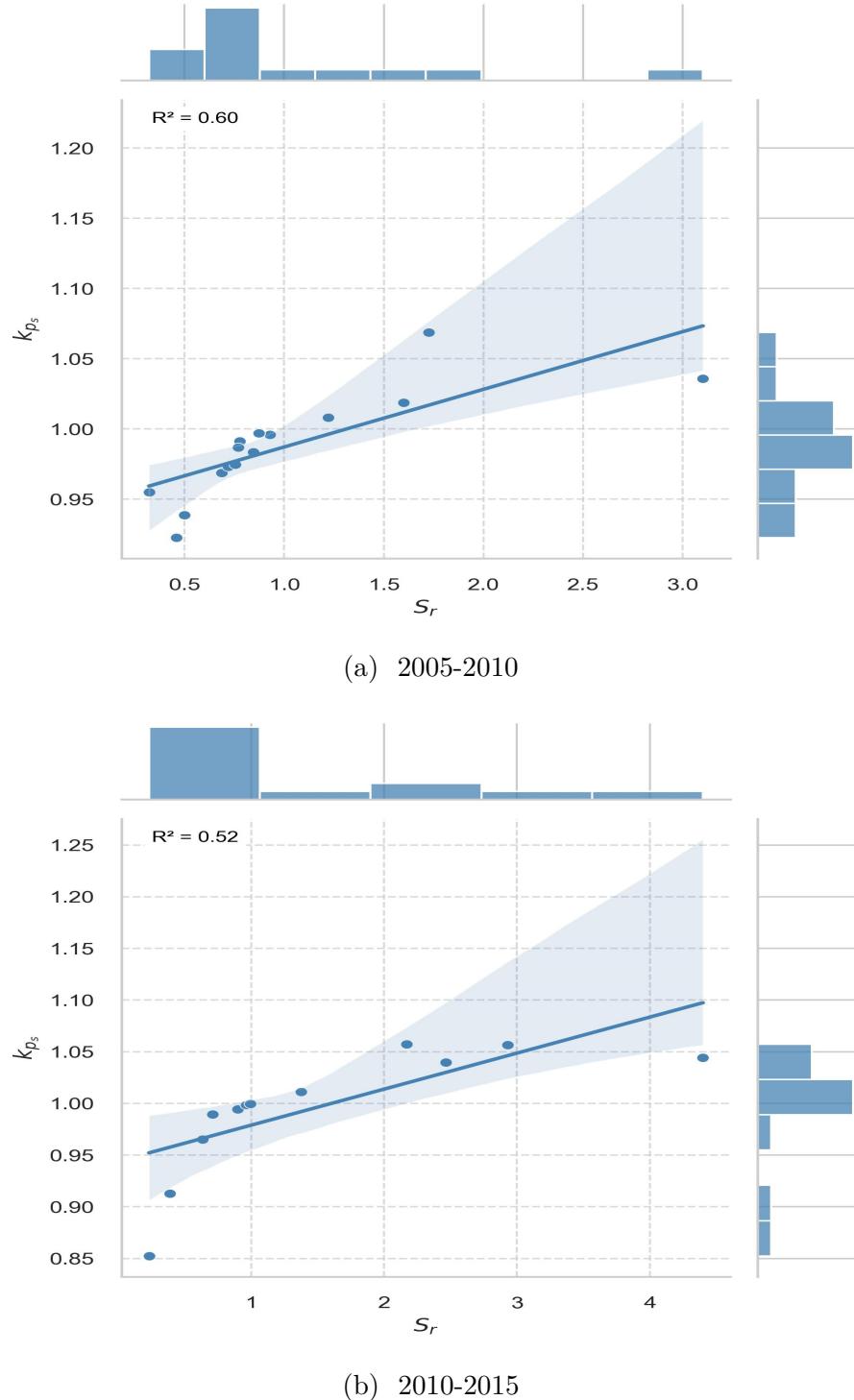
Fonte: Elaboração própria.

Prudente (1995-2000 e 2010-2015), conforme Tabela 4 e Figura 23.

Nas Figuras 24 a 30 constam os gráficos de dispersão dos *clusters* formados por cada algoritmo, por intervalos². Na Tabela 5 constam as quantidades de integrantes de

² Abreviaturas: S = Silhueta; CH = Calinski–Harabasz; DB = Davies–Bouldin.

Figura 21 – Relação entre S_r e k_{ps} para os intervalos de (a) 2005-2010 e (b) 2010-2015.

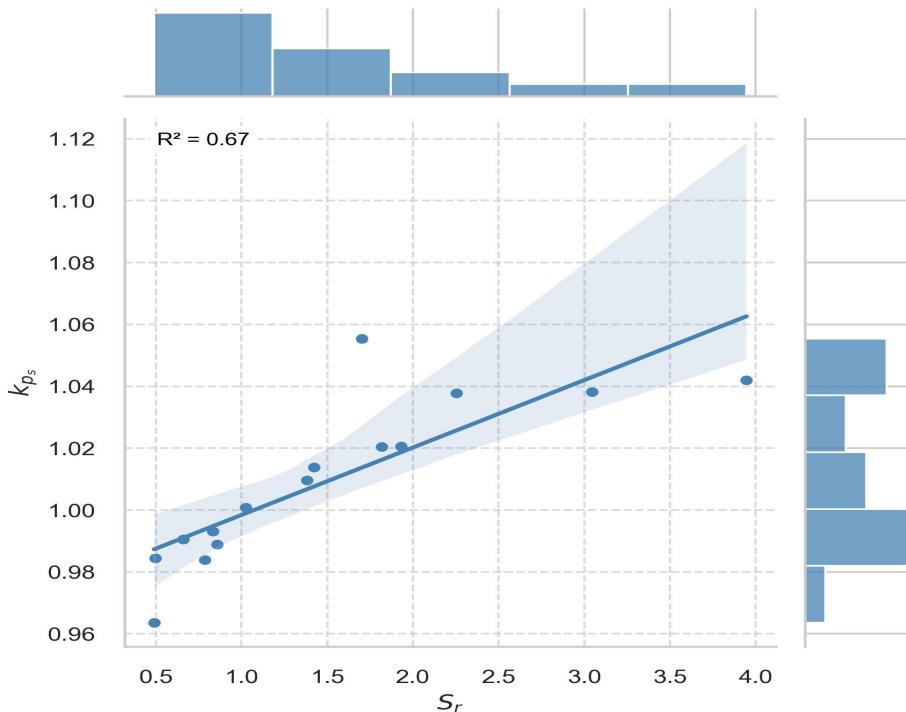


Fonte: Elaboração própria.

cada *cluster*, por algoritmo e para cada intervalo.

Como pode ser observado, o algoritmo *K-Means* produz *clusters* mais homogêneos, de maneira que grandes distâncias dentro dos *clusters* são menos recorrentes (HENNIG, 2015). Exceto quando valores extremos são mais sobressalentes, como no intervalo de

Figura 22 – Relação entre S_r e k_{ps} para o intervalo de 2015-2020.



Fonte: Elaboração própria.

1995-2000 (Figura 26 e Tabela 5), valores discrepantes são integrados a *clusters* distantes no *K-Means*, ocasionando maiores diferenças dentro dos *clusters* nesses casos (HENNIG, 2015), como pode ser observado no gráfico do intervalo de 2000-2005 (Figura 27).

O *K-Medoids* é semelhante ao *K-Means*, produzindo, em alguns casos, *clusters* idênticos aos formados pelo *K-Means* (exceto pela localização dos centros), como nos intervalos de 1985-1990, 1995-2000, 2005-2010 e 2010-2015 (Figuras 24, 26, 28 e 29, respectivamente). No entanto, o *K-Medoids* usa dissimilaridades não quadráticas, o que possibilita maiores dissimilaridades dentro dos *clusters*, e maior flexibilidade em relação a *outliers* e desvios da forma esférica do *cluster* (HENNIG, 2015). Isso pode ser observado nos intervalos de 2000-2005 e 2015-2020 (Figuras 27 e 30, respectivamente), em que os *clusters* formados pelo *K-Medoids* apresentam desempenho nas métricas inferior ao *K-Means*, possivelmente em razão dessa maior dissimilaridade dentro dos *clusters*.

Nos métodos de agrupamento baseados em densidade, como o algoritmo *Mean Shift*, busca-se identificar *clusters* com áreas de alta densidade (HENNIG, 2015). Essas áreas de alta densidade podem ter formatos bem flexíveis, variar de tamanho e incluir diferenças expressivas (HENNIG, 2015). Ademais, *outliers* não afetam excessivamente o *Mean Shift*, sendo transformados em *clusters* com apenas um integrante (CARREIRA-PERPIÑÁN, 2015), como evidenciado pela maior incidência de *clusters* unitários nesse algoritmo, em relação ao *K-Means* e ao *K-Medoids* (Tabela 5).

Tabela 4 – Estatísticas descritivas do S_r com e sem valores discrepantes para cada intervalo.

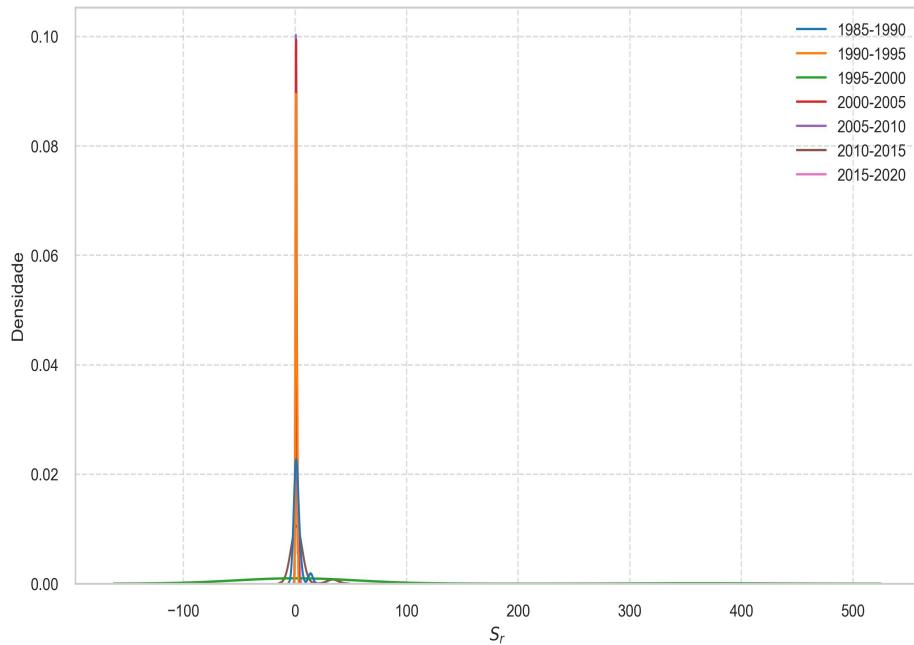
S_r com valores discrepantes								
Intervalo	N	Média	Desvio padrão	Mín.	Q1	Mediana	Q3	Máx.
1985-1990	15	2.2495	3.4978	0.2649	0.6606	0.9615	1.7564	13.97977
1990-1995	15	1.1258	0.5959	0.4206	0.6715	0.9803	1.3907	2.2038
1995-2000	15	25.0321	93.2398	0.3133	0.5260	0.9066	1.0475	362.0630
2000-2005	15	1.2239	0.6018	0.4251	0.8572	1.1205	1.4257	2.8351
2005-2010	15	1.0183	0.6936	0.3223	0.7015	0.7761	1.0734	3.1006
2010-2015	15	3.5971	8.4364	0.2329	0.8033	0.9882	2.3185	33.8249
2015-2020	15	1.5098	0.9859	0.4901	0.8092	1.3813	1.8748	3.9460
S_r sem valores discrepantes								
Intervalo	N	Média	Desvio padrão	Mín.	Q1	Mediana	Q3	Máx.
1985-1990	14	1.4117	1.3547	0.2649	0.5924	0.9369	1.4028	4.6885
1990-1995	15	1.1258	0.5959	0.4206	0.6715	0.9803	1.3907	2.2038
1995-2000	14	0.9585	0.7755	0.3133	0.5225	0.8340	1.0167	3.4624
2000-2005	15	1.2239	0.6018	0.4251	0.8572	1.1205	1.4257	2.8351
2005-2010	15	1.0183	0.6936	0.3223	0.7015	0.7761	1.0734	3.1006
2010-2015	14	1.4380	1.1582	0.2329	0.7564	0.9864	1.9720	4.3976
2015-2020	15	1.5098	0.9859	0.4901	0.8092	1.3813	1.8748	3.9460

Fonte: Elaboração própria.

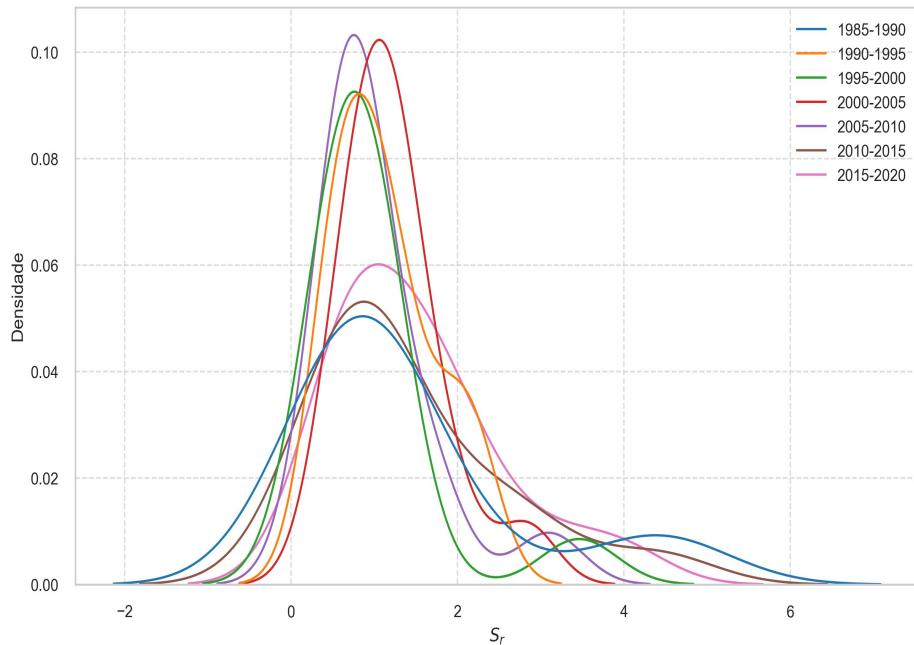
Métodos de agrupamento baseados na teoria dos grafos, como o *Affinity Propagation*, não são regidos por funções objetivas que tentam minimizar as dissimilaridades dentro dos *clusters* ou amplificar dissimilaridades entre os *clusters* (HENNIG, 2015). Nesse sentido, algoritmos como o *Affinity Propagation* se assemelham à *Single Linkage*, uma das formas como a distância entre dois *clusters* é calculada nos agrupamentos hierárquicos (*hierarchical clustering*), pois seus *clusters* “ideais” correspondem teoricamente a componentes conectados de um gráfico, enfatizando mais a separação entre *clusters*, isto é, em manter os pontos mais próximos de *clusters* diferentes separados uns dos outros (HENNIG, 2015). Essa maior ênfase na separação entre *clusters* se refletiu em uma maior quantidade de partições no *Affinity Propagation* em todos os intervalos, exceto para 1990-1995 (Figura 25 e Tabela 5).

Dentre os agrupamentos formados por cada algoritmo, apenas no intervalo de 1990-1995 (Figura 25) houve maior “concordância” entre os métodos, isto é, três algoritmos (*K-Means*, *Mean Shift* e *Affinity Propagation*) obtiveram partições idênticas, com uma separação bem delimitada entre os *clusters* formados.

Figura 23 – Gráficos de densidade alisada do S_r para cada intervalo (a) com e (b) sem valores discrepantes.



(a) S_r com valores discrepantes.



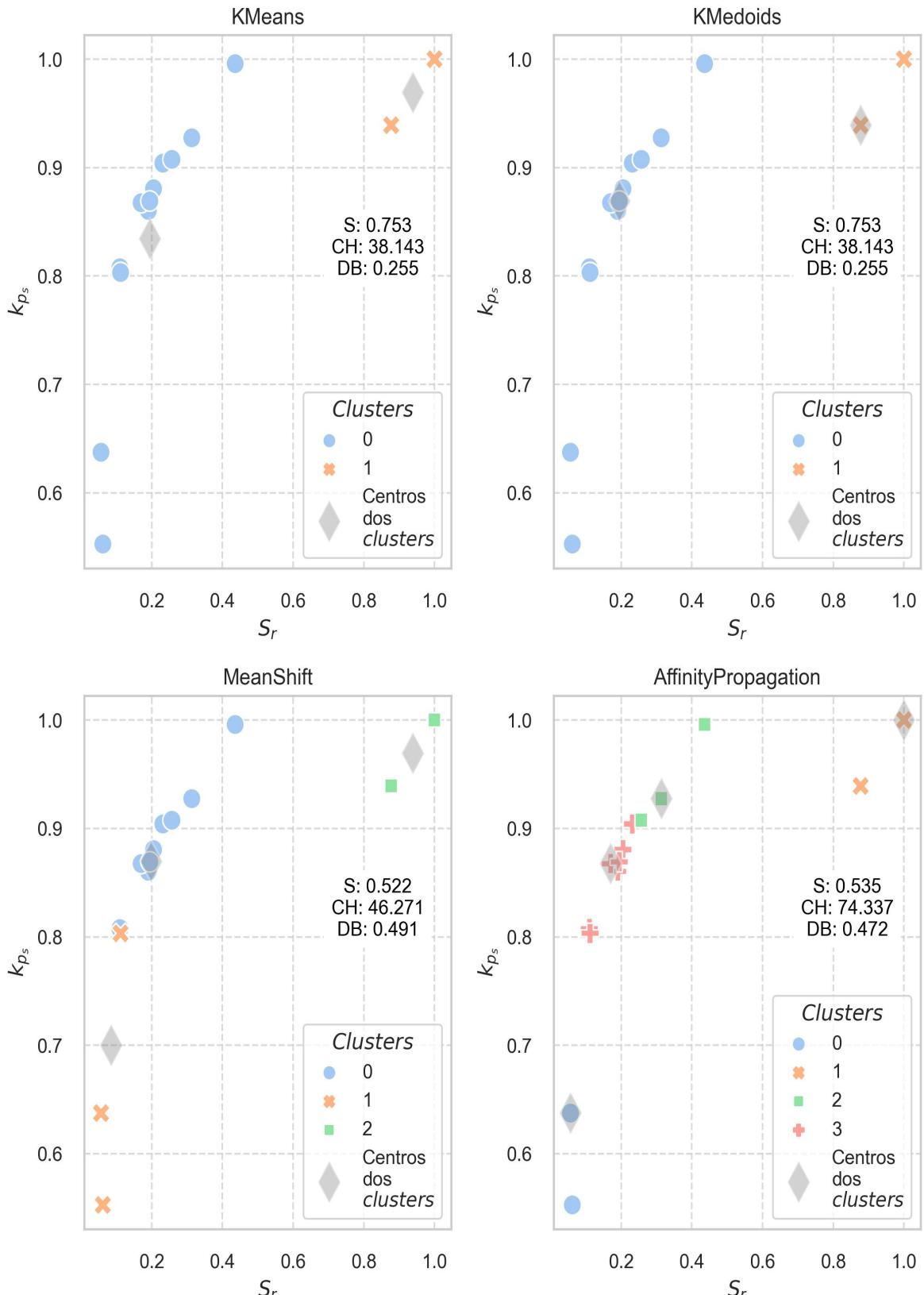
(b) S_r sem valores discrepantes.

Fonte: Elaboração própria.

Os municípios com os respectivos *clusters* formados pelos algoritmos, para cada intervalo, são apresentados nas Tabelas³ 6 a 9. Na Tabela 10 constam os valores obtidos

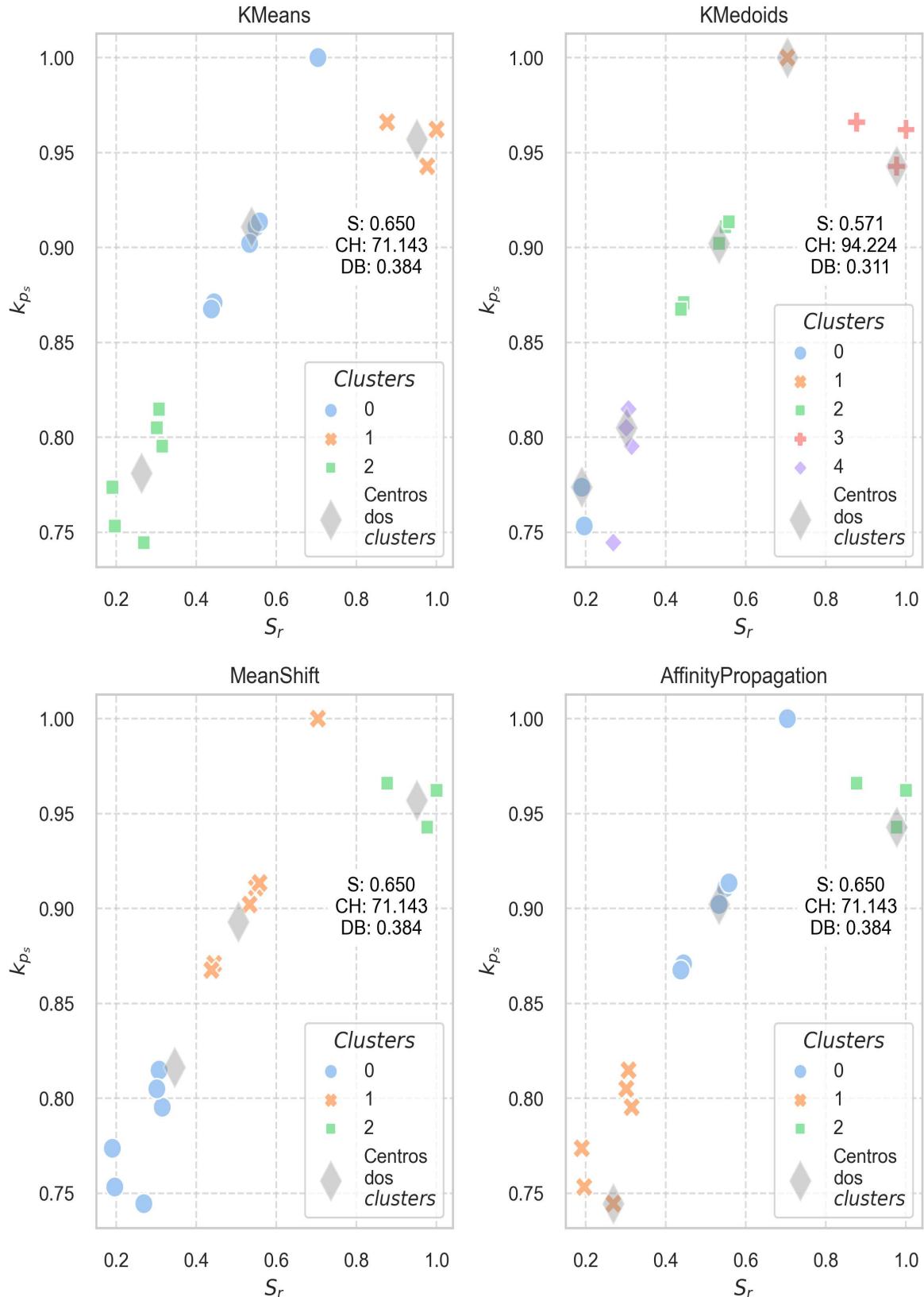
³ Abreviaturas: S_r red. e k_{ps} red. referem-se aos valores redimensionados de S_r e k_{ps} , respectivamente.

Figura 24 – Clusters por algoritmo (1985-1990).



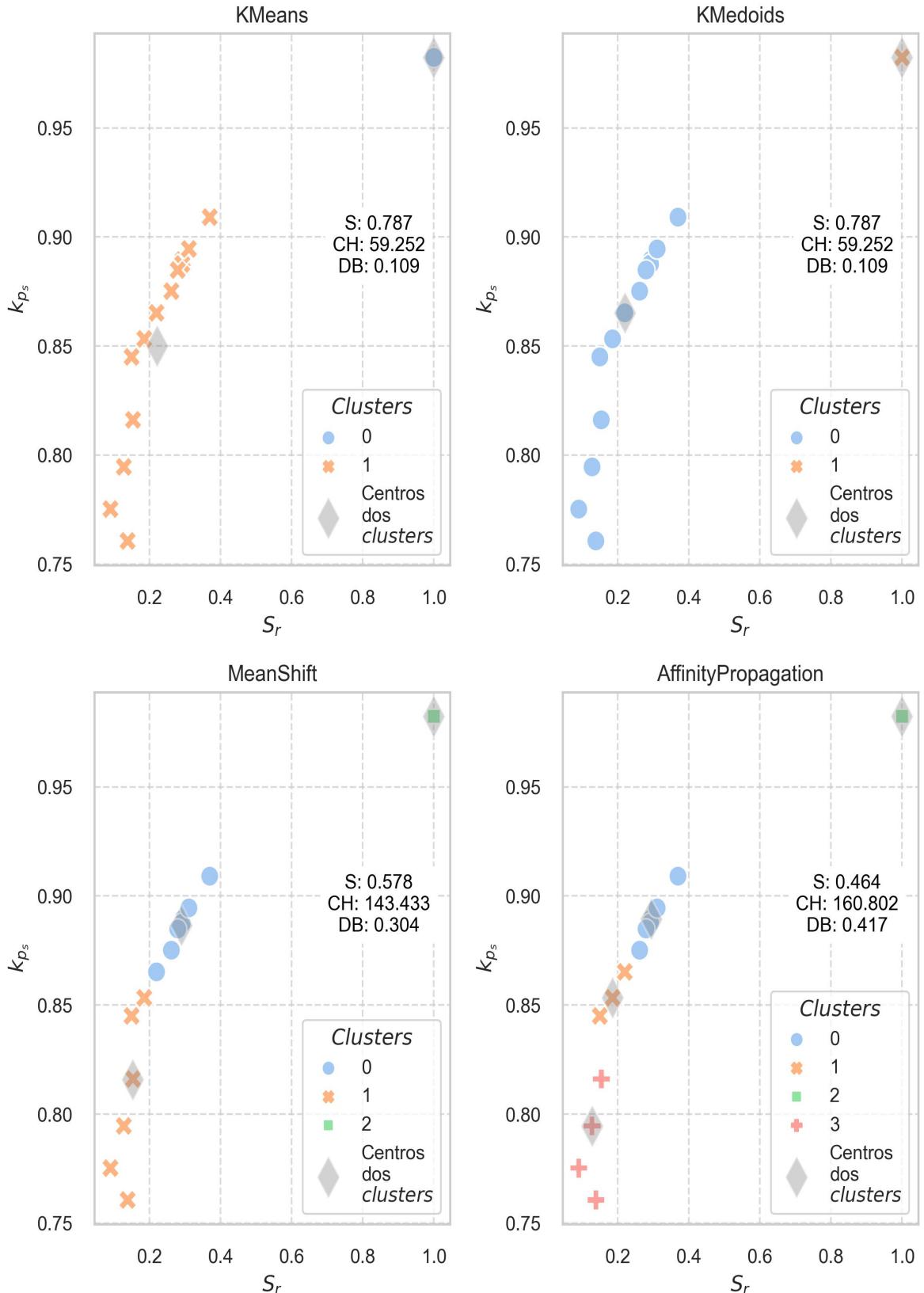
Fonte: Elaboração própria.

Figura 25 – Clusters por algoritmo (1990-1995).



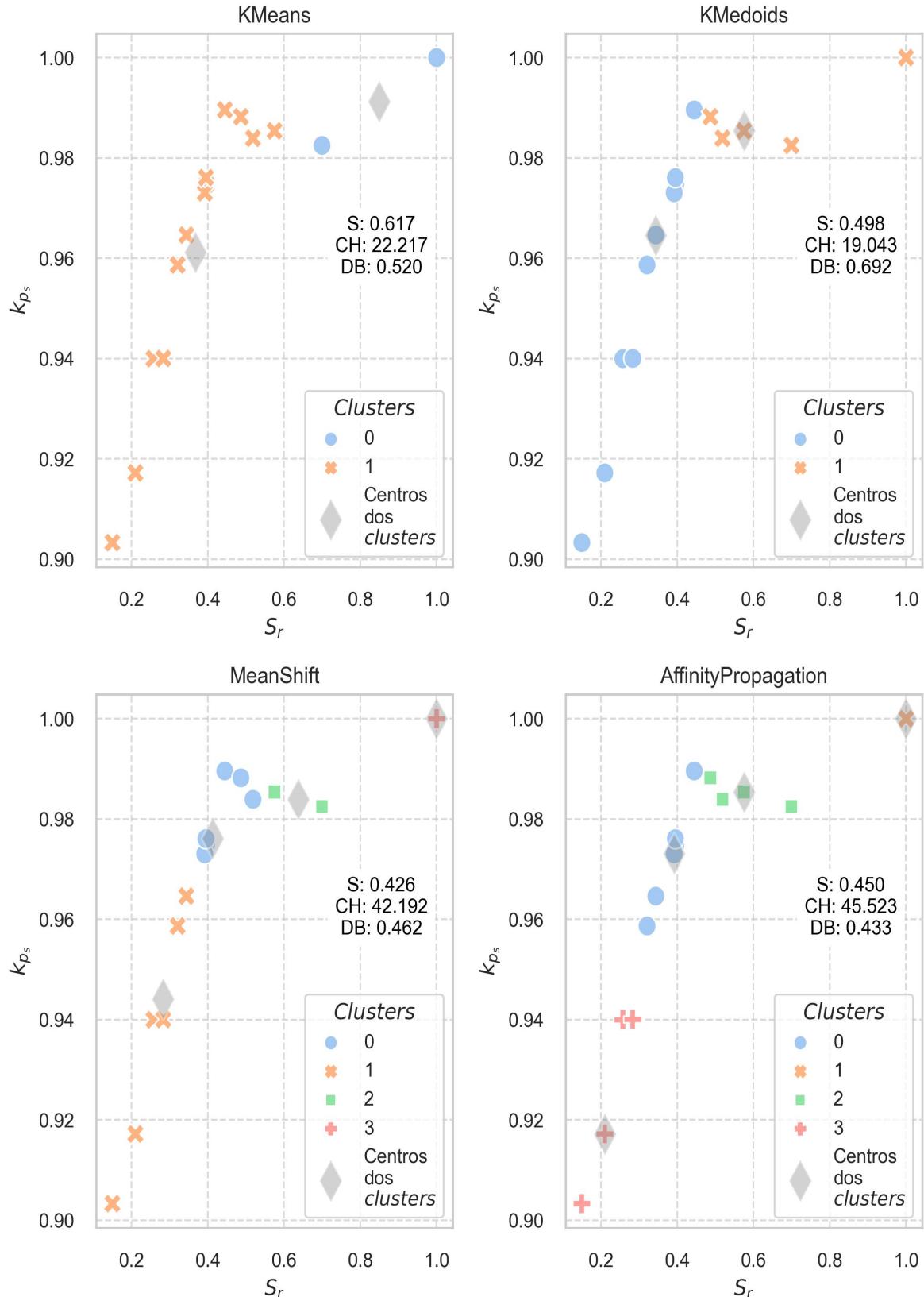
Fonte: Elaboração própria.

Figura 26 – *Clusters* por algoritmo (1995-2000).



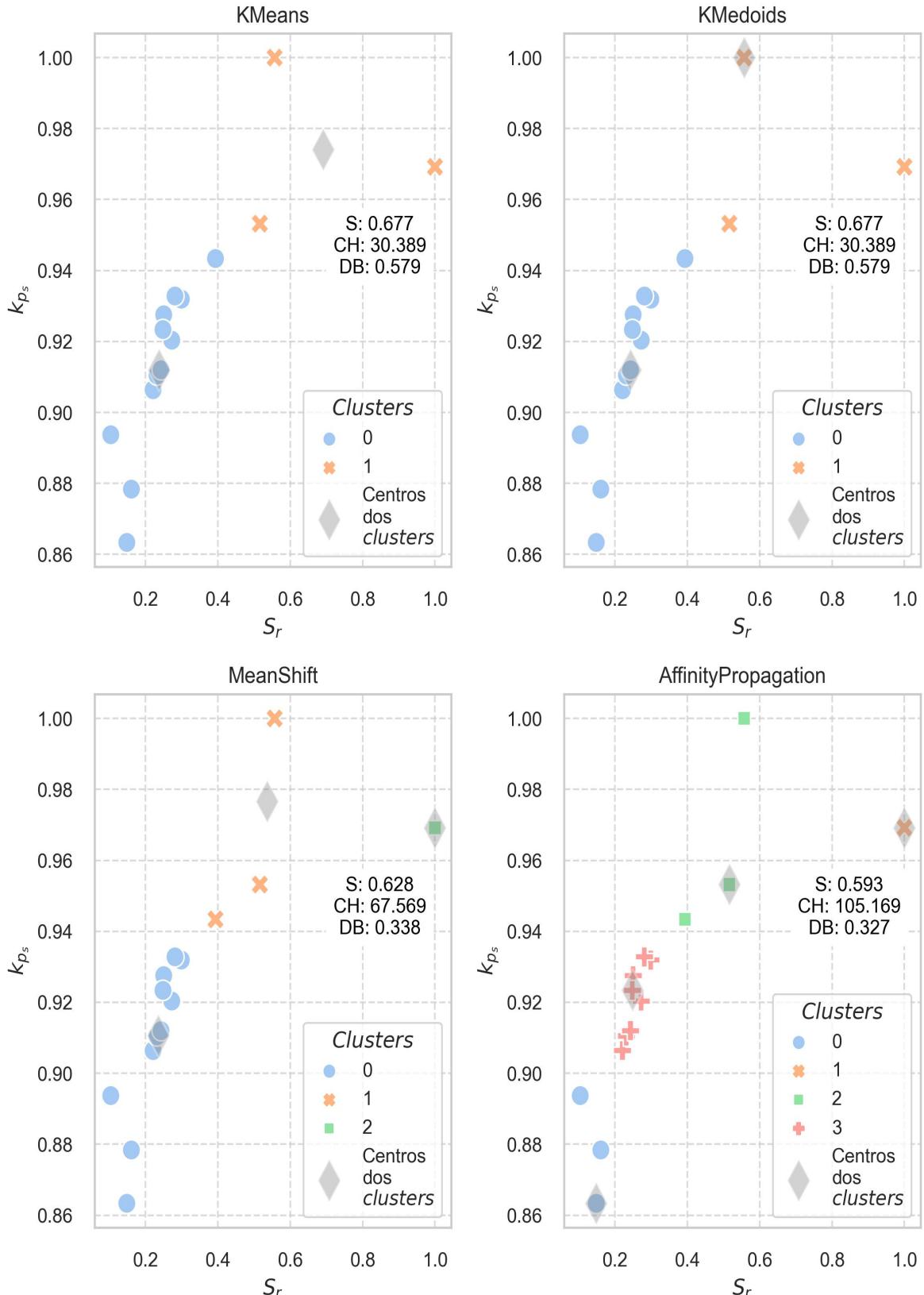
Fonte: Elaboração própria.

Figura 27 – Clusters por algoritmo (2000-2005).



Fonte: Elaboração própria.

Figura 28 – *Clusters* por algoritmo (2005-2010).

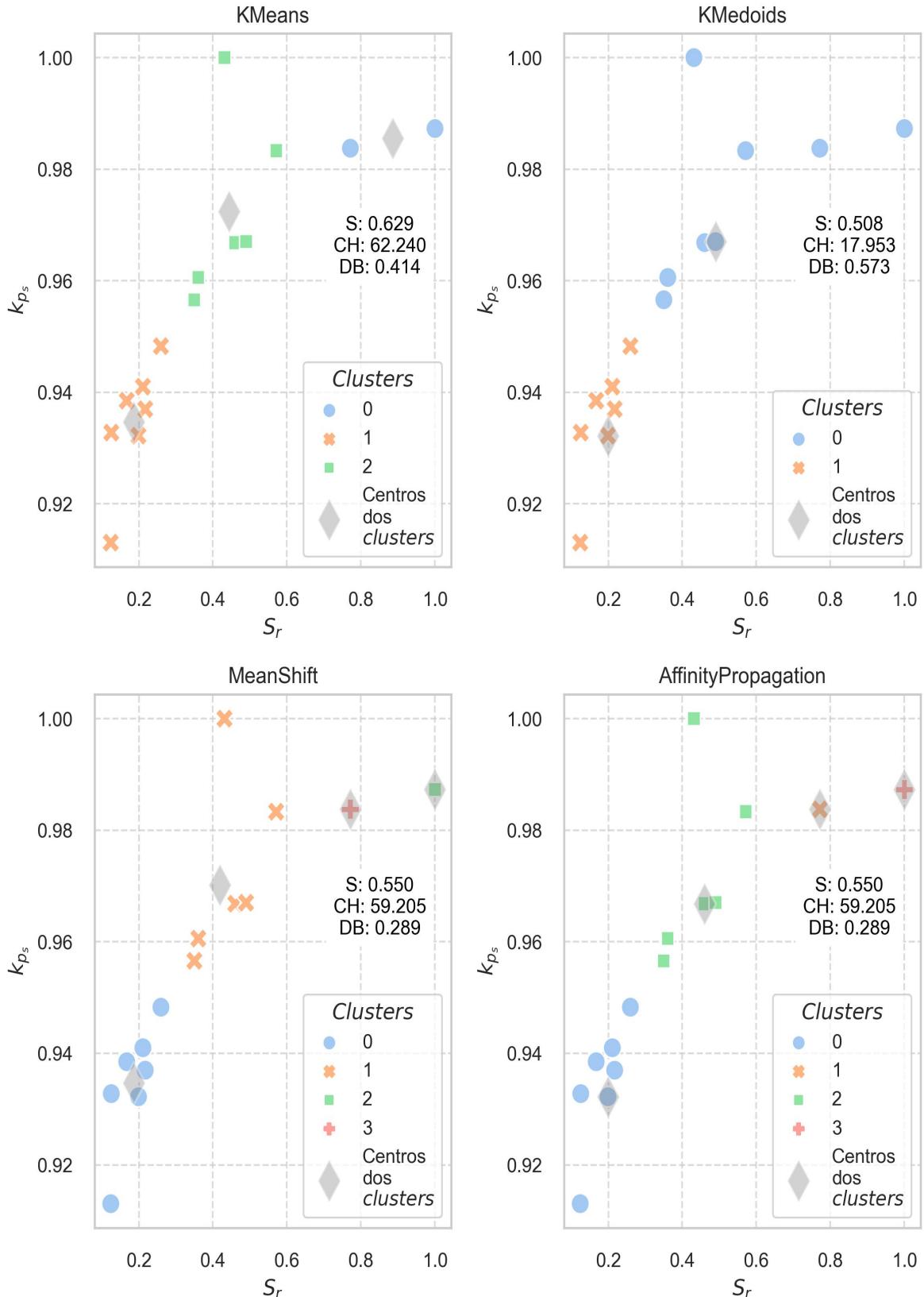


Fonte: Elaboração própria.

Figura 29 – *Clusters* por algoritmo (2010-2015).

Fonte: Elaboração própria.

Figura 30 – Clusters por algoritmo (2015-2020).



Fonte: Elaboração própria.

Tabela 5 – Integrantes por *clusters* para cada intervalo.

Algoritmo	<i>Clusters</i>	1985- 1990	1990- 1995	1995- 2000	2000- 2005	2005- 2010	2010- 2015	2015- 2020
<i>K-Means</i>	0	12	6	1	2	12	10	2
	1	2	3	13	13	3	4	7
	2		6				6	
Total		14	15	14	15	15	14	15
<hr/>								
Algoritmo	<i>Clusters</i>	1985- 1990	1990- 1995	1995- 2000	2000- 2005	2005- 2010	2010- 2015	2015- 2020
<i>K-Medoids</i>	0	12	2	13	10	12	4	8
	1	2	1	1	5	3	10	7
	2		5					
	3		3					
	4		4					
Total		14	15	14	15	15	14	15
<hr/>								
Algoritmo	<i>Clusters</i>	1985- 1990	1990- 1995	1995- 2000	2000- 2005	2005- 2010	2010- 2015	2015- 2020
<i>Mean Shift</i>	0	9	6	7	6	11	10	7
	1	3	6	6	6	3	3	6
	2	2	3	1	2	1	1	1
	3				1			1
Total		14	15	14	15	15	14	15
<hr/>								
Algoritmo	<i>Clusters</i>	1985- 1990	1990- 1995	1995- 2000	2000- 2005	2005- 2010	2010- 2015	2015- 2020
<i>Affinity Propagation</i>	0	2	6	6	6	3	3	7
	1	2	6	3	1	1	3	1
	2	3	3	1	4	3	1	6
	3	7		4	4	8	7	1
Total		14	15	14	15	15	14	15

Fonte: Elaboração própria.

nas métricas Silhueta (S), Calinski–Harabasz (CH) e Davies–Bouldin (DB) para cada algoritmo, por intervalos, com destaque para os melhores resultados de cada métrica por intervalo.

Por meio da Tabela 10, é possível distinguir os algoritmos com melhores desempenhos para cada intervalo, isto é, aqueles com melhores pontuações em pelo menos duas

Tabela 6 – Municípios e respectivos *clusters* por algoritmo para os intervalos 1985-1990 e 1990-1995.

1985-1990									
Município	S_r	k_{ps}	S_r red.	k_{ps} red.	KM	KMed	MS	AP	
ARAÇATUBA	1.0850	1.0207	0.2314	0.9040	0	0	0	3	
ARARAQUARA ²	0.2649	0.7197	0.0565	0.6374	0	0	1	0	
ASSIS ²	4.6885	1.1291	1.0000	1.0000	1	1	2	1	
BARRETOS	0.2891	0.6241	0.0617	0.5528	0	0	1	0	
BAURU ²	1.4686	1.0471	0.3132	0.9274	0	0	0	2	
BIRIGUI	0.8927	0.9715	0.1904	0.8605	0	0	0	3	
BOTUCATU	0.5156	0.9118	0.1100	0.8076	0	0	0	3	
FRANCA	0.5242	0.9069	0.1118	0.8032	0	0	1	3	
JAÚ	0.9615	0.9941	0.2051	0.8805	0	0	0	3	
MARÍLIA ¹	4.1142	1.0603	0.8775	0.9391	1	1	2	1	
OURINHOS	1.2053	1.0247	0.2571	0.9075	0	0	0	2	
PRESIDENTE P. ²	0.7970	0.9795	0.1700	0.8676	0	0	0	3	
SÃO CARLOS ¹	0.9123	0.9814	0.1946	0.8692	0	0	0	3	
SJRP	2.0443	1.1244	0.4360	0.9958	0	0	0	2	
1990-1995									
Município	S_r	k_{ps}	S_r red.	k_{ps} red.	KM	KMed	MS	AP	
ARAÇATUBA	0.6777	0.9321	0.3075	0.8148	2	4	0	1	
ARARAQUARA	0.9803	0.9962	0.4448	0.8708	0	2	1	0	
ASSIS	0.9656	0.9924	0.4381	0.8675	0	2	1	0	
BARRETOS ¹	1.2104	1.0421	0.5492	0.9109	0	2	1	0	
BAURU	0.6950	0.9098	0.3154	0.7953	2	4	0	1	
BIRIGUI ³	1.1757	1.0320	0.5335	0.9021	0	2	1	0	
BOTUCATU ¹	0.4206	0.8850	0.1909	0.7736	2	0	0	1	
CATANDUVA ¹	1.5513	1.1440	0.7039	1.0000	0	1	1	0	
FRANCA	1.9318	1.1050	0.8766	0.9660	1	3	2	2	
JAÚ	1.2301	1.0449	0.5582	0.9134	0	2	1	0	
MARÍLIA	2.2038	1.1007	1.0000	0.9621	1	3	2	2	
OURINHOS ²	0.5939	0.8517	0.2695	0.7445	2	4	0	1	
PRESIDENTE P. ³	2.1522	1.0785	0.9766	0.9427	1	3	2	2	
SÃO CARLOS	0.4339	0.8617	0.1969	0.7532	2	0	0	1	
SJRP ¹	0.6653	0.9209	0.3019	0.8050	2	4	0	1	

¹ Centros dos *clusters* formados pelo *K-Medoids*.

² Centros dos *clusters* formados pelo *Affinity Propagation*.

³ Centros dos *clusters* formados pelos algoritmos *K-Medoids* e *Affinity Propagation*.

Fonte: Elaboração própria.

métricas, exceto para o intervalo de 2010-2015, em que não houve um algoritmo com melhor desempenho em duas ou mais métricas. Dentre os quatro algoritmos utilizados, o *Mean Shift* foi o único a não atingir melhores pontuações em pelo menos duas métricas em nenhum dos intervalos, possivelmente em razão de sua maior aplicabilidade para tarefas

Tabela 7 – Municípios e respectivos *clusters* por algoritmo para os intervalos 1995-2000 e 2000-2005.

1995-2000								
Município	S_r	k_{ps}	S_r red.	k_{ps} red.	KM	KMed	MS	AP
ARAÇATUBA ²	0.6428	0.9621	0.1857	0.8533	1	0	1	1
ARARAQUARA ²	1.0180	1.0028	0.2940	0.8894	1	0	0	0
ASSIS	1.0128	1.0008	0.2925	0.8876	1	0	0	0
BARRETOS ¹	0.7615	0.9755	0.2199	0.8652	1	0	0	1
BAURU	0.5331	0.9202	0.1540	0.8161	1	0	1	3
BIRIGUI	1.2799	1.0250	0.3696	0.9091	1	0	0	0
BOTUCATU	0.5190	0.9527	0.1499	0.8450	1	0	1	1
CATANDUVA	0.9066	0.9867	0.2618	0.8751	1	0	0	0
FRANCA	0.9684	0.9977	0.2797	0.8848	1	0	0	0
JAÚ	1.0771	1.0086	0.3111	0.8945	1	0	0	0
MARÍLIA ³	3.4624	1.1075	1.0000	0.9823	0	1	2	2
OURINHOS	0.4802	0.8576	0.1387	0.7606	1	0	1	3
SÃO CARLOS ²	0.4433	0.8959	0.1280	0.7946	1	0	1	3
SJRP	0.3133	0.8741	0.0905	0.7753	1	0	1	3
2000-2005								
Município	S_r	k_{ps}	S_r red.	k_{ps} red.	KM	KMed	MS	AP
ARAÇATUBA ²	1.1121	1.0070	0.3922	0.9730	1	0	0	0
ARARAQUARA	1.1244	1.0088	0.3966	0.9748	1	0	0	0
ASSIS	1.4707	1.0183	0.5187	0.9839	1	1	0	2
BARRETOS	0.4251	0.9348	0.1499	0.9033	1	0	1	3
BAURU	0.7301	0.9728	0.2575	0.9400	1	0	1	3
BIRIGUI	0.9102	0.9921	0.3210	0.9587	1	0	1	0
BOTUCATU ²	2.8351	1.0349	1.0000	1.0000	0	1	3	1
CATANDUVA	1.2609	1.0241	0.4448	0.9896	1	0	0	0
FRANCA	1.1205	1.0101	0.3952	0.9761	1	0	0	0
JAÚ	1.3807	1.0227	0.4870	0.9882	1	1	0	2
MARÍLIA	1.9836	1.0168	0.6996	0.9825	0	1	2	2
OURINHOS	0.8042	0.9728	0.2837	0.9400	1	0	1	3
PRESIDENTE P. ³	1.6303	1.0198	0.5751	0.9854	1	1	2	2
SÃO CARLOS ¹	0.9751	0.9983	0.3440	0.9646	1	0	1	0
SJRP ²	0.5959	0.9492	0.2102	0.9172	1	0	1	3

¹ Centros dos *clusters* formados pelo *K-Medoids*.

² Centros dos *clusters* formados pelo *Affinity Propagation*.

³ Centros dos *clusters* formados pelos algoritmos *K-Medoids* e *Affinity Propagation*.

Fonte: Elaboração própria.

de visão computacional (COMANICIU; MEER, 2002).

Para os intervalos de 1990-1995 e 2010-2015 (Figuras 25 e 29, respectivamente), em vez de escolher o algoritmo com os agrupamentos mais apropriados a partir da comparação das métricas de validação interna de *clusters*, destacaremos apenas os motivos para

Tabela 8 – Municípios e respectivos *clusters* por algoritmo para os intervalos de 2005-2010 e 2010-2015.

2005-2010									
Município	S_r	k_{ps}	S_r red.	k_{ps} red.	KM	KMed	MS	AP	
ARAÇATUBA	0.9272	0.9960	0.2990	0.9319	0	0	0	3	
ARARAQUARA ¹	1.7268	1.0688	0.5569	1.0000	1	1	1	2	
ASSIS	1.2196	1.0082	0.3933	0.9434	0	0	1	2	
BARRETOS	0.6849	0.9687	0.2209	0.9064	0	0	0	3	
BAURU	0.3223	0.9552	0.1040	0.8937	0	0	0	0	
BIRIGUI ²	0.4598	0.9227	0.1483	0.8634	0	0	0	0	
BOTUCATU	0.7761	0.9913	0.2503	0.9275	0	0	0	3	
CATANDUVA	0.4990	0.9388	0.1609	0.8784	0	0	0	0	
FRANCA	0.8453	0.9836	0.2726	0.9204	0	0	0	3	
JAÚ ²	3.1006	1.0358	1.0000	0.9692	1	1	2	1	
MARÍLIA	0.8726	0.9969	0.2814	0.9328	0	0	0	3	
OURINHOS	0.7181	0.9731	0.2316	0.9105	0	0	0	3	
PRESIDENTE P. ²	1.6000	1.0187	0.5160	0.9532	1	1	1	2	
SÃO CARLOS ²	0.7696	0.9868	0.2482	0.9233	0	0	0	3	
SJRP ¹	0.7530	0.9747	0.2429	0.9120	0	0	0	3	
2010-2015									
Município	S_r	k_{ps}	S_r red.	k_{ps} red.	KM	KMed	MS	AP	
ARAÇATUBA	0.7094	0.9896	0.1613	0.9356	0	1	0	3	
ARARAQUARA	2.1703	1.0574	0.4935	0.9997	1	0	1	0	
ASSIS	0.9847	0.9990	0.2239	0.9445	0	1	0	3	
BARRETOS ²	2.4667	1.0396	0.5609	0.9829	1	0	1	0	
BAURU	0.8972	0.9943	0.2040	0.9401	0	1	0	3	
BIRIGUI ²	0.3869	0.9129	0.0880	0.8631	0	1	0	1	
BOTUCATU ¹	2.9290	1.0566	0.6660	0.9989	1	0	1	0	
CATANDUVA	0.2329	0.8524	0.0530	0.8059	0	1	0	1	
FRANCA	0.6327	0.9650	0.1439	0.9124	0	1	0	1	
JAÚ	1.3771	1.0111	0.3132	0.9560	0	1	0	3	
MARÍLIA ²	4.3976	1.0444	1.0000	0.9874	1	0	2	2	
OURINHOS ²	0.9882	0.9987	0.2247	0.9442	0	1	0	3	
SÃO CARLOS ¹	0.9637	0.9985	0.2191	0.9440	0	1	0	3	
SJRP	0.9953	0.9996	0.2263	0.9451	0	1	0	3	

¹ Centros dos *clusters* formados pelo *K-Medoids*.

² Centros dos *clusters* formados pelo *Affinity Propagation*.

Fonte: Elaboração própria.

escolha dos agrupamentos considerados mais adequados. Isso porque, além da existência de discrepâncias entre os resultados das métricas utilizadas, no caso do intervalo de 2010-2015, a definição do que seria um agrupamento “bom” é muitas vezes dependente do contexto (HALKIDI; VAZIRGIANNIS; HENNIG, 2015), no caso do intervalo de 1990-1995 isso se evidencia pelo fato de o algoritmo com melhor desempenho nas métricas ser o único com

Tabela 9 – Municípios e respectivos *clusters* por algoritmo para o intervalo 2015-2020.

Município	S_r	k_{ps}	S_r red.	k_{ps} red.	KM	KMed	MS	AP
ARAÇATUBA	1.0247	1.0008	0.2597	0.9483	1	1	0	0
ARARAQUARA	0.8575	0.9889	0.2173	0.9370	1	1	0	0
ASSIS ³	0.7857	0.9838	0.1991	0.9322	1	1	0	0
BARRETOS	1.7021	1.0554	0.4314	1.0000	2	0	1	2
BAURU	0.4957	0.9844	0.1256	0.9328	1	1	0	0
BIRIGUI	2.2530	1.0378	0.5710	0.9833	2	0	1	2
BOTUCATU	1.4236	1.0138	0.3608	0.9606	2	0	1	2
CATANDUVA	0.4901	0.9636	0.1242	0.9130	1	1	0	0
FRANCA	0.8327	0.9931	0.2110	0.9410	1	1	0	0
JAÚ	1.3813	1.0096	0.3501	0.9566	2	0	1	2
MARÍLIA ²	3.0440	1.0382	0.7714	0.9838	0	0	3	1
OURINHOS ²	1.8170	1.0204	0.4605	0.9668	2	0	1	2
PRESIDENTE P. ²	3.9460	1.0420	1.0000	0.9873	0	0	2	3
SÃO CARLOS	0.6605	0.9905	0.1674	0.9385	1	1	0	0
SJRP ¹	1.9326	1.0206	0.4898	0.9670	2	0	1	2

¹ Centros dos *clusters* formados pelo *K-Medoids*.² Centros dos *clusters* formados pelo *Affinity Propagation*.³ Centros dos *clusters* formados pelos algoritmos *K-Medoids* e *Affinity Propagation*.

Fonte: Elaboração própria.

Tabela 10 – Valores das métricas para cada algoritmo, por intervalos.

Métrica Algoritmo	1985-	1990-	1995-	2000-	2005-	2010-	2015-	
	1990	1995	2000	2005	2010	2015	2020	
S^1	<i>K-Means</i>	0.753	0.650	0.787	0.617	0.677	0.693	0.629
	<i>K-Medoids</i>	0.753	0.571	0.787	0.498	0.677	0.693	0.508
	<i>Mean Shift</i>	0.522	0.650	0.578	0.426	0.628	0.669	0.550
	<i>Affinity Propagation</i>	0.535	0.650	0.464	0.450	0.593	0.538	0.550
CH^2	<i>K-Means</i>	38.143	71.143	59.252	22.217	30.389	38.188	62.240
	<i>K-Medoids</i>	38.143	94.224	59.252	19.043	30.389	38.188	17.953
	<i>Mean Shift</i>	46.271	71.143	143.433	42.192	67.569	53.261	59.205
	<i>Affinity Propagation</i>	74.337	71.143	160.802	45.523	105.169	79.543	59.205
DB^2	<i>K-Means</i>	0.255	0.384	0.109	0.520	0.579	0.461	0.414
	<i>K-Medoids</i>	0.255	0.311	0.109	0.692	0.579	0.461	0.573
	<i>Mean Shift</i>	0.491	0.384	0.304	0.462	0.338	0.274	0.289
	<i>Affinity Propagation</i>	0.472	0.384	0.417	0.433	0.327	0.344	0.289

¹ Silhueta.² Calinski–Harabasz.³ Davies–Bouldin.

Fonte: Elaboração própria.

um agrupamento diferente do encontrado pelos demais algoritmos.

Cabe destacar que, idealmente, seria oportuno dividir os municípios da amostra

deste estudo em, pelo menos, dois grupos distintos: aqueles com espraiamento e crescimento urbano mais disperso (S_r e k_{ps} positivos); e aqueles sem espraiamento e com crescimento urbano mais compacto (S_r e k_{ps} negativos). No entanto, como em todos os intervalos houve espraiamento e crescimento urbano mais disperso (Tabela 3), torna-se interessante, portanto, agrupar os municípios por magnitude de espraiamento e dispersão, de maneira a evidenciar os valores mais destoantes.

Nesse sentido, a escolha dos *clusters* mais informativos, para os intervalos de 1990-1995 e 2010-2015, baseou-se na separação entre *clusters*, ou seja, não foi priorizada a homogeneidade dentro dos *clusters*, mas sim as diferenças entre *clusters*. Consideramos, portanto, que a estrutura do agrupamento formado pelos algoritmos *K-Means*, *Mean Shift* e *Affinity Propagation*, com $k = 3$ (Figura 25), possibilita uma explicação mais intuitiva dos grupos, no caso do intervalo de 1990-1995, assim como a estrutura do agrupamento formado pelo algoritmo *Mean Shift*, com $k = 3$ (Figura 29), possibilita uma explicação mais intuitiva dos grupos, no caso do intervalo de 2010-2015.

A seguir são elencados, para cada intervalo, os agrupamentos considerados mais adequados, com as respectivas categorizações dos grupos formados.

- **1985-1990:** *K-Means* e *K-Medoids*, ambos com $k = 2$ (Figura 24).
 - Grupo 0 - espraiamento: **reduzido a moderado**; crescimento disperso: **moderado a elevado**.
 - Grupo 1 - espraiamento e crescimento disperso: **elevados**.
- **1990-1995:** *K-Means*, *Mean Shift* e *Affinity Propagation*, com $k = 3$ (Figura 25). A seguinte caracterização é baseada nos grupos do algoritmo *K-Means*:
 - Grupo 0 - espraiamento: **moderado**; crescimento disperso: **moderado a elevado**.
 - Grupo 1 - espraiamento e crescimento disperso: **elevados**.
 - Grupo 2 - espraiamento: **reduzido**; crescimento disperso: **moderado**.
- **1995-2000:** *K-Means* e *K-Medoids*, ambos com $k = 2$ (Figura 26). A seguinte caracterização é baseada nos grupos do algoritmo *K-Means*.
 - Grupo 0 - espraiamento e crescimento disperso: **elevados**.
 - Grupo 1 - espraiamento: **reduzido a moderado**; crescimento disperso: **moderado a elevado**.
- **2000-2005:** *Affinity Propagation*, com $k = 4$ (Figura 27).
 - Grupo 0 - espraiamento: **reduzido a moderado** e crescimento disperso: **elevado**.

- Grupo 1 - espalhamento e crescimento disperso: **elevados**.
 - Grupo 2 - espalhamento: **moderado a elevado**; crescimento disperso: **elevado**.
 - Grupo 3 - espalhamento: **reduzido**; crescimento disperso: **elevado**.
- **2005-2010:** *Affinity Propagation*, com $k = 4$ (Figura 28).
 - Grupo 0 - espalhamento: **reduzido**; crescimento disperso: **moderado a elevado**.
 - Grupo 1 - espalhamento e crescimento disperso: **elevados**.
 - Grupo 2 - espalhamento: **moderado**; crescimento disperso: **elevado**.
 - Grupo 3 - espalhamento: **reduzido a moderado**; crescimento disperso: **elevado**.
 - **2010-2015:** *Mean Shift*, com $k = 3$ (Figura 29).
 - Grupo 0 - espalhamento: **reduzido a moderado**; crescimento disperso: **moderado a elevado**.
 - Grupo 1 - espalhamento: **moderado**; crescimento disperso: **elevado**.
 - Grupo 2 - espalhamento e crescimento disperso: **elevados**.
 - **2015-2020:** *K-Means*, com $k = 3$ (Figura 30).
 - Grupo 0 - espalhamento e crescimento disperso: **elevados**.
 - Grupo 1 - espalhamento: **reduzido**; crescimento disperso: **elevado**.
 - Grupo 2 - espalhamento: **reduzido a moderado**; crescimento disperso: **elevado**.

5 CONCLUSÃO

Neste trabalho de conclusão de curso, exploramos a dinâmica do crescimento urbano em municípios paulistas de porte médio. Nosso objetivo era identificar padrões que pudessem ajudar a compreender os processos de expansão urbana, assim como a evolução da forma urbana das cidades. Para tanto, empregamos um método quantitativo em Python, utilizando uma função sigmoide em forma de S-invertido — uma ferramenta amplamente utilizada para caracterizar o declínio na densidade territorial urbana. Em seguida, examinamos as variações na densidade territorial urbana entre os municípios, empregando técnicas de aprendizado não supervisionado para agrupamento de dados.

Como todo modelo matemático, a função sigmoide S-invertido é uma simplificação da realidade. A divisão de uma cidade em quatro partições, por exemplo, pode ser considerada reducionista e incapaz de englobar todo espectro da complexidade urbana. No entanto, por se fundamentar em uma regra intrínseca aos dados de densidade construída, representa uma abordagem apropriada para estudos comparativos, sendo precisa e menos suscetível à subjetividades, algo reforçado pelo adequado ajuste da função aos dados dos municípios paulistas.

Identificamos padrões distintos de crescimento urbano nos municípios da amostra. Isso indica a viabilidade de reconhecer esses padrões por meio dos agrupamentos definidos, oferecendo informações relevantes sobre a dinâmica do crescimento urbano. Vale ressaltar que a adequação dos algoritmos variou em cada intervalo temporal, enfatizando a importância de testar diferentes métodos de agrupamento.

Trabalhos futuros poderiam superar uma das limitações deste estudo via definição de anéis concêntricos com raios menores, cerca de 250 a 500 metros, por exemplo, para mitigar a ocorrência de valores negativos para o parâmetro c . Outras aplicações também se beneficiariam do uso de imagens de melhor resolução, como as da Coleção BETA do MapBiomas 10 metros (MAPBIOMAS, 2023), que inclui mapas anuais de cobertura e uso da terra para o período de 2016 a 2022, com base em imagens do satélite *Sentinel-2* (resolução espacial de 10 metros). Embora se trate de uma série temporal menor (2016-2022) do que a utilizada neste estudo, essa Coleção BETA permitiria incluir informações mais detalhadas, especialmente caso se pretenda reduzir o espaço temporal analisado (anual, por exemplo), o que poderia aperfeiçoar as análises de densidade territorial urbana.

Por fim, visando aprimorar as análises, a ampliação da amostra pela inclusão de municípios de outros Estados brasileiros possibilitaria realizar estudos comparativos para identificar potenciais *clusters* regionais. Além disso, a aplicação de abordagens estatísticas tradicionais poderia ser considerada para avaliar a presença de autocorrelações espaciais.

REFERÊNCIAS

- ABURAS, M. M.; AHAMAD, M. S. S.; OMAR, N. Q. Spatio-temporal simulation and prediction of land-use change using conventional and machine learning models: a review. **Environmental Monitoring and Assessment**, v. 191, n. 205, 2019.
- AHANI, S.; DADASHPOOR, H. Urban growth containment policies for the guidance and control of peri-urbanization: A review and proposed framework. **Environment, Development and Sustainability**, v. 23, p. 14215–14244, 2021.
- ALBON, C. Clustering. In: **Machine learning with Python cookbook: Practical solutions from preprocessing to Deep Learning**. Sebastopol, CA: O'Reilly Media, Inc., 2018. p. 287–296.
- ALBON, C. Model evaluation. In: **Machine learning with Python cookbook: Practical solutions from preprocessing to Deep Learning**. Sebastopol, CA: O'Reilly Media, Inc., 2018. p. 179–208.
- AMPONSAH, O. *et al.* Global urban sprawl containment strategies and their implications for rapidly urbanising cities in ghana. **Land Use Policy**, v. 114, p. 105979, 2022.
- ANABTAWI, S. Towards a better understanding of compact cities. **Journal of Planning Literature**, v. 38, n. 4, p. 531–547, 2023.
- ANDERSON, J. E. The changing structure of a city temporal changes in cubic spline urban density patterns. **Journal of Regional Science**, v. 25, n. 3, p. 413–425, 1985.
- ARTHUR, D.; VASSILVITSKII, S. K-means++ the advantages of careful seeding. In: **Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms**. [S.l.: s.n.], 2007. p. 1027–1035.
- BAGHERI, B.; SOLTANI, A. The spatio-temporal dynamics of urban growth and population in metropolitan regions of iran. **Habitat International**, v. 136, p. 102797, 2023.
- BATTY, M.; KIM, K. S. Form follows function: reformulating urban population density functions. **Urban studies**, v. 29, n. 7, p. 1043–1069, 1992.
- BRANCH, M. A.; COLEMAN, T. F.; LI, Y. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. **SIAM Journal on Scientific Computing**, v. 21, n. 1, p. 1–23, 1999.
- BYRD, R. H.; SCHNABEL, R. B.; SHULTZ, G. A. Approximate solution of the trust region problem by minimization over two-dimensional subspaces. **Mathematical programming**, v. 40, n. 1-3, p. 247–263, 1988.
- CALIŃSKI, T.; HARABASZ, J. A dendrite method for cluster analysis. **Communications in Statistics-Theory and Methods**, v. 3, n. 1, p. 1–27, 1974.

CARREIRA-PERPIÑÁN, M. Á. Clustering methods based on kernel density estimators: mean-shift algorithms. In: CHRISTIAN, H. et al. (ed.). **Handbook of cluster analysis**. Boca Raton, FL: Chapman & Hall/CRC Handbooks of Modern Statistical Methods CRC Press, 2015. p. 383–417.

CARVALHO, A. et al. Algoritmos de agrupamentos. In: **Inteligência Artificial—uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC, 2011. p. 208–218.

CARVALHO, A. et al. Avaliação de modelos descritivos. In: **Inteligência Artificial—uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC, 2011. p. 236–256.

CETESB. **Relatórios de Qualidade das Praias e Águas Litorâneas do Estado de São Paulo. Apêndice A - Características do litoral de São Paulo**. 2022. Disponível em: <https://cetesb.sp.gov.br/praias/publicacoes-relatorios/>.

CHAKRABORTY, S. et al. Building in or out? examining urban expansion patterns and land use efficiency across the global sample of 466 cities with million+ inhabitants. **Habitat International**, v. 120, p. 102503, 2022.

CHEN, Y.; FENG, J. Fractal-based exponential distribution of urban density and self-affine fractal forms of cities. **Chaos, Solitons & Fractals**, v. 45, n. 11, p. 1404–1416, 2012.

CLARK, C. Urban population densities. **Journal of the Royal Statistical Society. Series A (General)**, Oxford University Press, v. 114, n. 4, p. 490–496, 1951.

COMANICIU, D.; MEER, P. Mean shift: A robust approach toward feature space analysis. **IEEE Transactions on pattern analysis and machine intelligence**, v. 24, n. 5, p. 603–619, 2002.

CUNHA, A. L. **Collection 8 of the Annual Land Cover and Land Use Maps of Brazil (1985-2022)**. MapBiomas Data, 2023. Disponível em: <https://doi.org/10.58053/MapBiomas/VJIJCL>.

DAVIES, D. L.; BOULDIN, D. W. A cluster separation measure. **IEEE transactions on pattern analysis and machine intelligence**, n. 2, p. 224–227, 1979.

FREY, B. J.; DUECK, D. Clustering by passing messages between data points. **Science**, v. 315, n. 5814, p. 972–976, 2007.

GÉRON, A. **Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow**. United States of America: O'Reilly Media, Inc., 2022.

GUÉROIS, M.; PUMAIN, D. Built-up encroachment and the urban field: a comparison of forty european cities. **Environment and Planning A**, v. 40, n. 9, p. 2186–2203, 2008.

HALKIDI, M.; VAZIRGIANNIS, M.; HENNIG, C. Method-independent indices for cluster validation and estimating the number of clusters. In: CHRISTIAN, H. et al. (ed.). **Handbook of cluster analysis**. Boca Raton, FL: Chapman & Hall/CRC Handbooks of Modern Statistical Methods CRC Press, 2015. p. 595–618.

HENNIG, C. Clustering strategy and method selection. In: CHRISTIAN, H. et al. (ed.). **Handbook of cluster analysis**. Boca Raton, FL: Chapman & Hall/CRC Handbooks of Modern Statistical Methods CRC Press, 2015. p. 703–730.

- HUANG, C.; ZHANG, C. Characterizing urban growth in vientiane from 2000 to 2019 using time-series optical and sar-based estimates of urban land. **International Journal of Applied Earth Observation and Geoinformation**, v. 109, p. 102798, 2022.
- IBGE. **Cadastro de Localidades Selecionadas - 2010**. Rio de Janeiro: IBGE, 2010. Disponível em: <https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial/27385-localidades.html?=&t=acesso-ao-produto>.
- IBGE. **Áreas urbanizadas do Brasil : 2019**. Rio de Janeiro: IBGE, 2022. Disponível em: <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2101973>.
- IBGE. **Censo Demográfico 2022: População e Domicílios - Primeiros Resultados**. Rio de Janeiro: IBGE, 2023. Disponível em: <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2102011>.
- IBGE. **Malha Municipal Digital e Áreas Territoriais 2022: Informações Técnicas e Legais para a Utilização dos Dados Publicados**. Rio de Janeiro: IBGE, 2023. Disponível em: <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2101998>.
- IGC. **Mapas Individuais das Regiões Administrativas e Metropolitanas**. 2023. Disponível em: http://www.igc.sp.gov.br/produtos/mapas_ra.html.
- JAMES, G. *et al.* Unsupervised learning. In: **An introduction to statistical learning: With applications in Python**. [S.l.: s.n.]: Springer, 2023. p. 503–556.
- JIAO, L. Urban land density function: A new method to characterize urban expansion. **Landscape and Urban Planning**, v. 139, p. 26–39, 2015.
- KAUFMAN, L.; ROUSSEEUW, P. J. Partitioning around medoids (program pam). In: **Finding groups in data: an introduction to cluster analysis**. New Jersey: John Wiley & Sons, 2005. p. 68–125.
- KEERATIKASIKORN, C. A comparative study on four major cities in northeastern thailand using urban land density function. **Geo-spatial information science**, v. 21, n. 2, p. 93–101, 2018.
- LI, X.; GONG, P. Urban growth models: progress and perspective. **Science bulletin**, v. 61, n. 21, p. 1637–1650, 2016.
- LOPES, H. E. G.; GOSLING, M. d. S. Cluster analysis in practice: Dealing with outliers in managerial research. **Revista de Administração Contemporânea**, v. 25, p. e200081, 2020.
- LU, H. *et al.* Study on urban expansion and population density changes based on the inverse s-shaped function. **Sustainability**, v. 15, n. 13, 2023.
- MAHTTA, R. *et al.* Urban land expansion: the role of population and economic growth for 300+ cities. **npj Urban Sustainability**, v. 2, n. 1, p. 5, 2022.
- MAPBIOMAS. **Projeto MapBiomas – Coleção Beta de Mapas Anuais de Cobertura e Uso da Terra do Brasil**. 2023. Acessado em: 29 dez. 2023. Disponível em: <https://brasil.mapbiomas.org/mapbiomas-cobertura-10m/>.

MENZORI, I. D. **Dinâmicas territoriais e métricas espaciais: avaliação de expansão urbana em cidade de porte médio.** 2021. 259 p. Tese (Doutorado em Engenharia Urbana) — Universidade Federal de São Carlos, São Carlos, 2021. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/15020>.

MENZORI, I. D.; SOUSA, I. C. N. de; GONÇALVES, L. M. Local government shift and national housing program: Spatial repercussions on urban growth. **Land Use Policy**, v. 126, p. 106548, 2023.

MITCHELL, T. M. **Machine Learning.** [S.l.: s.n.]: McGraw-Hill, 1997.

MOLIN, S.; JEE, K. Getting started with machine learning in python. In: **Hands-On Data Analysis with Pandas: A Python data science handbook for data collection, wrangling, analysis, and visualization.** Birmingham—Mumbai: Packt Publishing Ltd, 2021. p. 537–620.

MOLIN, S.; JEE, K. Introduction to data analysis. In: **Hands-On Data Analysis with Pandas: A Python data science handbook for data collection, wrangling, analysis, and visualization.** Birmingham—Mumbai: Packt Publishing Ltd, 2021. p. 3–46.

NEWLING, B. E. The spatial variation of urban population densities. **Geographical Review**, v. 59, n. 2, p. 242–252, 1969.

PARK, H.-S.; JUN, C.-H. A simple and fast algorithm for k-medoids clustering. **Expert Systems with Applications**, v. 36, n. 2, p. 3336–3341, 2009.

PICHLER, M.; HARTIG, F. Machine learning and deep learning—a review for ecologists. **Methods in Ecology and Evolution**, v. 14, n. 4, p. 994–1016, 2023.

POURTAHERIAN, P.; JAEGER, J. A. How effective are greenbelts at mitigating urban sprawl? a comparative study of 60 european cities. **Landscape and Urban Planning**, v. 227, p. 104532, 2022.

RASCHKA, S.; MIRJALILI, V. Working with unlabeled data – clustering analysis. In: **Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2.** Birmingham: Packt Publishing Ltd, 2019. p. 353–382.

REBA, M.; SETO, K. C. A systematic review and assessment of algorithms to detect, characterize, and monitor urban land change. **Remote sensing of environment**, v. 242, p. 111739, 2020.

REDDY, C. K.; VINZAMURI, B. A survey of partitional and hierarchical clustering algorithms. In: AGGARWAL, C. C.; REDDY, C. K. (ed.). **Data clustering: algorithms and applications.** Boca Raton, FL: Chapman & Hall/CRC data mining and knowledge discovery series, 2014. p. 87–110.

RODRIGUES, F. A. Machine learning in physics: a short guide. **Europhysics Letters**, 2023.

ROUSSEEUW, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. **Journal of Computational and Applied Mathematics**, v. 20, p. 53–65, 1987.

- SHELADIYA, K. P.; PATEL, C. R. The impacts of urban growth drivers on the spatial and temporal pattern of city expansion. **Journal of the Indian Society of Remote Sensing**, v. 51, n. 8, p. 1631–1655, 2023.
- SILVA, C.; VERGARA-PERUCICH, F. Determinants of urban sprawl in latin america: evidence from santiago de chile. **SN social sciences**, v. 1, n. 202, p. 1–35, 2021.
- SMEED, R. J. The effect of some kinds of routeing systems on the amount of traffic in the central areas of towns:(with particular reference to journeys into and out of the central areas). 1962.
- SOUZA, C. M. *et al.* Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. **Remote Sensing**, v. 12, n. 17, 2020.
- STEINLEY, D. K-medoids and other criteria for crisp clustering. In: CHRISTIAN, H. *et al.* (ed.). **Handbook of Cluster Analysis**. Boca Ration, FL: Chapman & Hall/CRC Handbooks of Modern Statistical Methods CRC Press, 2015. p. 55–66.
- SUMARI, N. S. *et al.* On the absurdity of rapid urbanization: Spatio-temporal analysis of land-use changes in morogoro, tanzania. **Cities**, v. 107, p. 102876, 2020.
- TANNER, J. C. **Factors affecting the amount of travel**. [S.l.: s.n.]: HM Stationery Office, 1961.
- WU, S. *et al.* Characterizing urban expansion combining concentric-ring and grid-based analysis for latin american cities. **Land**, v. 10, n. 5, p. 444, 2021.
- XIONG, H.; LI, Z. Clustering validation measures. In: AGGARWAL, C. C.; REDDY, C. K. (ed.). **Data clustering: algorithms and applications**. Boca Raton, FL: Chapman & Hall/CRC data mining and knowledge discovery series, 2014. p. 571–606.
- XU, D.; TIAN, Y. A comprehensive survey of clustering algorithms. **Annals of Data Science**, v. 2, p. 165–193, 2015.
- XU, G. *et al.* Urban expansion and form changes across african cities with a global outlook: Spatiotemporal analysis of urban land densities. **Journal of Cleaner Production**, v. 224, p. 802–810, 2019.
- YAO, Y. *et al.* Do compact cities have higher efficiencies of agglomeration economies? a dynamic panel model with compactness indicators. **Land Use Policy**, v. 115, p. 106005, 2022.
- ZHENG, J. *et al.* A comparison of two models to measure the distance decay of urban land density to the city center. In: IEEE. **2022 29th International Conference on Geoinformatics**. [S.l.: s.n.], 2022. p. 1–7.

APÊNDICES

APÊNDICE A – DENSIDADES TERRITORIAIS URBANAS PARA CADA MUNICÍPIO, ANUAIS E POR RAIOS

Tabela 11 – Densidade territorial urbana por raios para 1985 e 1990.

1985								
Município	1km	2km	3km	4km	5km	6km	7km	8km
ARAÇATUBA	0.9626	0.5700	0.2842	0.0858	0.0005			
ARARAQUARA	0.8307	0.6741	0.2661	0.0249	0.0145			
ASSIS	0.9441	0.5017	0.1300	0.0263	0.0078			
BARRETOS	0.6924	0.3255	0.0415	0.0051				
BAURU	0.8954	0.8502	0.4366	0.2835	0.0834	0.0866	0.0153	0.0014
BIRIGUI	0.8770	0.3745	0.0315	0.0017				
BOTUCATU	0.7609	0.4112	0.2257	0.0776	0.0184	0.0105		
CATANDUVA	0.9716	0.4558	0.1040	0.0104	0.0062			
FRANCA	0.9371	0.8245	0.4755	0.2500	0.1627	0.0448	0.0137	
JAÚ	0.8926	0.4459	0.0877	0.0031	0.0017			
MARÍLIA	0.9636	0.5826	0.2459	0.1312	0.0792	0.0247	0.0073	
OURINHOS	0.8318	0.3232	0.0464	0.0233				
PRESIDENTE PRUDENTE	0.9825	0.7043	0.3652	0.1010	0.0959	0.0220	0.0121	
SÃO CARLOS	0.9207	0.6098	0.3193	0.0878	0.0069	0.0018		
SÃO JOSÉ DO RIO PRETO	0.9717	0.8395	0.6521	0.3023	0.1056	0.0170	0.0034	
1990								
Município	1km	2km	3km	4km	5km	6km	7km	8km
ARAÇATUBA	0.9749	0.6921	0.4608	0.1633	0.0173			
ARARAQUARA	0.9036	0.8363	0.4717	0.1116	0.0584	0.0323	0.0258	
ASSIS	0.9590	0.5511	0.2042	0.0550	0.0114	0.0024		
BARRETOS	0.9252	0.6190	0.1852	0.0418	0.0059			
BAURU	0.9062	0.8689	0.5469	0.3675	0.1368	0.1410	0.0263	0.0029
BIRIGUI	0.9565	0.5569	0.1678	0.0114				
BOTUCATU	0.8230	0.5365	0.2852	0.1161	0.0306	0.0299		
CATANDUVA	0.9734	0.5014	0.1378	0.0189	0.0098			
FRANCA	0.9930	0.9225	0.5631	0.3106	0.2482	0.0804	0.0475	0.0086
JAÚ	0.9314	0.5483	0.1613	0.0142	0.0038			
MARÍLIA	0.9677	0.6127	0.2998	0.1668	0.1177	0.0407	0.0209	
OURINHOS	0.8536	0.4413	0.0922	0.0414	0.0217			
PRESIDENTE PRUDENTE	0.9833	0.7647	0.4195	0.1621	0.1143	0.0384	0.0272	
SÃO CARLOS	0.9491	0.7159	0.4686	0.2078	0.0281	0.0046		
SÃO JOSÉ DO RIO PRETO	0.9797	0.8583	0.7217	0.4468	0.2700	0.0743	0.0134	0.0035

Fonte: Elaboração própria.

Tabela 12 – Densidade territorial urbana por raios para 1995 e 2000.

1995									
Município	1km	2km	3km	4km	5km	6km	7km	8km	9km
ARAÇATUBA	0.9790	0.7685	0.6063	0.2992	0.0586	0.0016	0.0039		
ARARAQUARA	0.9312	0.8861	0.6081	0.2322	0.1078	0.0624	0.0362	0.0123	
ASSIS	0.9753	0.6853	0.3238	0.0903	0.0170	0.0026			
BARRETOS	0.9494	0.7174	0.3571	0.0770	0.0127	0.0008			
BAURU	0.9283	0.9049	0.7517	0.5661	0.2408	0.2210	0.0866	0.0127	0.0137
BIRIGUI	0.9823	0.6569	0.2741	0.0267	0.0002				
BOTUCATU	0.9004	0.6512	0.3344	0.1960	0.0809	0.0599	0.0203	0.0036	
CATANDUVA	0.9908	0.6855	0.3016	0.0850	0.0251	0.0021			
FRANCA	0.9984	0.9556	0.6387	0.4064	0.3223	0.1442	0.0923	0.0190	0.0014
JAÚ	0.9506	0.6684	0.3086	0.0462	0.0106				
MARÍLIA	0.9733	0.6944	0.3791	0.2431	0.1828	0.0756	0.0387	0.0234	
OURINHOS	0.9154	0.6749	0.2144	0.0725	0.0503	0.0099			
PRESIDENTE PRUDENTE	0.9847	0.8077	0.4874	0.2467	0.1655	0.0647	0.0464	0.0332	
SÃO CARLOS	0.9680	0.8246	0.5997	0.3240	0.0751	0.0342	0.0135		
SÃO JOSÉ DO RIO PRETO	0.9874	0.9042	0.8014	0.6021	0.3822	0.1573	0.0244	0.0074	
2000									
Município	1km	2km	3km	4km	5km	6km	7km	8km	9km
ARAÇATUBA	0.9805	0.7955	0.6675	0.3523	0.1022	0.0071	0.0055		
ARARAQUARA	0.9535	0.9430	0.6864	0.3409	0.1713	0.0967	0.0584	0.0155	0.0120
ASSIS	0.9839	0.7101	0.3690	0.1107	0.0249	0.0038			
BARRETOS	0.9651	0.7488	0.4274	0.1001	0.0159	0.0016			
BAURU	0.9507	0.9166	0.8293	0.6621	0.3239	0.2630	0.1232	0.0302	0.0178
BIRIGUI	0.9897	0.7013	0.3564	0.0798	0.0061	0.0033			
BOTUCATU	0.9404	0.7033	0.3783	0.2429	0.1242	0.0800	0.0333	0.0154	
CATANDUVA	0.9915	0.7571	0.3916	0.1294	0.0440	0.0047			
FRANCA	0.9984	0.9711	0.6785	0.4532	0.3532	0.1693	0.1042	0.0247	0.0018
JAÚ	0.9638	0.7278	0.3974	0.1270	0.0336	0.0034			
MARÍLIA	0.9780	0.7392	0.4317	0.3062	0.2271	0.1285	0.0575	0.0322	0.0070
OURINHOS	0.9664	0.7939	0.3297	0.1058	0.0707	0.0122			
PRESIDENTE PRUDENTE	0.9852	0.8187	0.5197	0.2942	0.2330	0.1036	0.0604	0.0393	0.0108
SÃO CARLOS	0.9827	0.8948	0.6931	0.4097	0.1434	0.0960	0.0360	0.0034	
SÃO JOSÉ DO RIO PRETO	0.9896	0.9302	0.8605	0.7023	0.4764	0.2261	0.0601	0.0236	0.0190

Fonte: Elaboração própria.

Tabela 13 – Densidade territorial urbana por raios para 2005 e 2010.

2005										
Município	1km	2km	3km	4km	5km	6km	7km	8km	9km	10km
ARAÇATUBA	0.9807	0.8058	0.6967	0.4073	0.1413	0.0114	0.0071			
ARARAQUARA	0.9672	0.9553	0.7185	0.3980	0.1946	0.1117	0.0666	0.0164	0.0120	
ASSIS	0.9946	0.7196	0.4078	0.1369	0.0270	0.0050				
BARRETOS	0.9713	0.8008	0.4716	0.1364	0.0243	0.0029				
BAURU	0.9547	0.9268	0.8617	0.7124	0.3817	0.3084	0.1327	0.0412	0.0225	
BIRIGUI	0.9893	0.7379	0.4181	0.1102	0.0077	0.0107				
BOTUCATU	0.9479	0.7174	0.4016	0.2615	0.1495	0.0960	0.0544	0.0207	0.0011	
CATANDUVA	0.9933	0.7869	0.4705	0.1746	0.0565	0.0080	0.0125			
FRANCA	0.9984	0.9801	0.7199	0.5110	0.3988	0.2087	0.1363	0.0281	0.0020	
JAÚ	0.9699	0.7682	0.4336	0.1943	0.0678	0.0048	0.0017			
MARÍLIA	0.9795	0.7582	0.4439	0.3289	0.2463	0.1468	0.0731	0.0356	0.0076	
OURINHOS	0.9855	0.8399	0.4017	0.1207	0.0819	0.0138	0.0013			
PRESIDENTE	0.9844	0.8355	0.5476	0.3287	0.2510	0.1297	0.0820	0.0423	0.0141	
PRUDENTE										
SÃO CARLOS	0.9827	0.9132	0.7274	0.4561	0.1727	0.1317	0.0514	0.0056	0.0051	
SÃO JOSÉ DO	0.9904	0.9553	0.8866	0.7524	0.5564	0.2908	0.1080	0.0453	0.0328	
RIO PRETO										
2010										
Município	1km	2km	3km	4km	5km	6km	7km	8km	9km	10km
ARAÇATUBA	0.9810	0.8188	0.7144	0.4592	0.1713	0.0249	0.0141			
ARARAQUARA	0.9823	0.9693	0.7584	0.4875	0.2718	0.1536	0.0781	0.0208	0.0123	
ASSIS	0.9950	0.7319	0.4426	0.1548	0.0348	0.0079				
BARRETOS	0.9731	0.8291	0.5318	0.1701	0.0306	0.0056				
BAURU	0.9563	0.9281	0.8899	0.7451	0.4012	0.3317	0.1377	0.0477	0.0250	
BIRIGUI	0.9982	0.7904	0.4958	0.1430	0.0199	0.0285				
BOTUCATU	0.9719	0.7352	0.4184	0.2792	0.1821	0.1041	0.0577	0.0218	0.0015	
CATANDUVA	0.9933	0.8241	0.5409	0.1934	0.0666	0.0096	0.0140			
FRANCA	0.9984	0.9903	0.7524	0.5770	0.4547	0.2408	0.1515	0.0330	0.0021	
JAÚ	0.9714	0.7804	0.4544	0.2388	0.0822	0.0199	0.0027			
MARÍLIA	0.9798	0.7757	0.4603	0.3365	0.2595	0.1573	0.0802	0.0396	0.0081	
OURINHOS	0.9876	0.8638	0.4623	0.1275	0.0885	0.0197	0.0015			
PRESIDENTE	0.9856	0.8519	0.5624	0.3637	0.2705	0.1409	0.0849	0.0450	0.0176	
PRUDENTE										
SÃO CARLOS	0.9893	0.9300	0.7566	0.4904	0.1990	0.1726	0.0661	0.0128	0.0070	
SÃO JOSÉ DO	0.9896	0.9639	0.9059	0.7853	0.6194	0.3836	0.1460	0.0693	0.0588	0.0334
RIO PRETO										

Fonte: Elaboração própria.

Tabela 14 – Densidade territorial urbana por raios para 2015 e 2020.

2015										
Município	1km	2km	3km	4km	5km	6km	7km	8km	9km	10km
ARAÇATUBA	0.9810	0.8279	0.7305	0.5038	0.2128	0.0531	0.0210	0.0165		
ARARAQUARA	0.9823	0.9705	0.7724	0.5419	0.3391	0.1868	0.0920	0.0468	0.0158	
ASSIS	0.9956	0.7547	0.4844	0.1867	0.0394	0.0097				
BARRETOS	0.9750	0.8380	0.5630	0.2228	0.0605	0.0125	0.0045			
BAURU	0.9594	0.9301	0.9040	0.7735	0.4370	0.3522	0.1658	0.0620	0.0290	0.0124
BIRIGUI	0.9982	0.8505	0.5585	0.1930	0.0452	0.0406				
BOTUCATU	0.9767	0.7654	0.4406	0.3374	0.2187	0.1283	0.0669	0.0261	0.0023	
CATANDUVA	0.9933	0.8765	0.6362	0.2137	0.0825	0.0169	0.0153			
FRANCA	0.9984	0.9905	0.7747	0.6390	0.5048	0.2574	0.1663	0.0388	0.0028	
JAÚ	0.9746	0.7916	0.4774	0.2676	0.0916	0.0276	0.0056			
MARÍLIA	0.9813	0.7838	0.4827	0.3700	0.2875	0.1826	0.0974	0.0491	0.0097	
OURINHOS	0.9876	0.8932	0.5419	0.1676	0.1119	0.0235	0.0015			
PRESIDENTE	0.9864	0.8565	0.5724	0.3938	0.3007	0.1953	0.0913	0.0470	0.0272	
PRUDENTE										
SÃO CARLOS	0.9894	0.9369	0.7819	0.5151	0.2281	0.1918	0.0778	0.0241	0.0077	
SÃO JOSÉ DO RIO PRETO	0.9899	0.9685	0.9142	0.8091	0.6878	0.4484	0.2025	0.1353	0.0911	0.0439
2020										
Município	1km	2km	3km	4km	5km	6km	7km	8km	9km	10km
ARAÇATUBA	0.9818	0.8435	0.7364	0.5271	0.2422	0.0730	0.0238	0.0183		
ARARAQUARA	0.9826	0.9716	0.8027	0.5969	0.3733	0.2225	0.1048	0.0528	0.0217	
ASSIS	0.9953	0.7892	0.5197	0.2287	0.0476	0.0105				
BARRETOS	0.9764	0.8609	0.6173	0.3149	0.1142	0.0251	0.0048			
BAURU	0.9634	0.9318	0.9070	0.7919	0.4561	0.3663	0.1718	0.0749	0.0359	0.0174
BIRIGUI	0.9982	0.8611	0.5795	0.2308	0.0582	0.0470	0.0032			
BOTUCATU	0.9828	0.7761	0.4683	0.3614	0.2394	0.1342	0.0702	0.0287	0.0036	
CATANDUVA	0.9934	0.8959	0.6706	0.2362	0.1045	0.0174	0.0157			
FRANCA	0.9984	0.9905	0.7845	0.6584	0.5414	0.2795	0.1836	0.0526	0.0034	0.0050
JAÚ	0.9722	0.8024	0.4969	0.2922	0.0989	0.0385	0.0066			
MARÍLIA	0.9873	0.7967	0.4938	0.3930	0.3142	0.2013	0.1134	0.0659	0.0319	0.0067
OURINHOS	0.9881	0.9011	0.5635	0.1900	0.1258	0.0267	0.0015			
PRESIDENTE	0.9875	0.8630	0.5859	0.4133	0.3337	0.2343	0.0980	0.0488	0.0305	
PRUDENTE										
SÃO CARLOS	0.9894	0.9418	0.7937	0.5350	0.2378	0.2015	0.0895	0.0303	0.0082	
SÃO JOSÉ DO RIO PRETO	0.9899	0.9733	0.9226	0.8225	0.7088	0.4809	0.2498	0.1812	0.1360	0.0548

Fonte: Elaboração própria.

APÊNDICE B – PARÂMETROS DO AJUSTE DA FUNÇÃO DE DENSIDADE TERRITORIAL URBANA

Tabela 15 – Parâmetros do ajuste da função sigmoide de densidade territorial urbana para 1985 e 1990.

1985				
Município	α	c	D	R^2
ARAÇATUBA	3.812178	0.016193	4.591658	0.989400
ARARAQUARA	3.622040	-0.032035	4.829736	0.992273
ASSIS	4.827234	0.021492	4.020387	0.998342
BARRETOS	2.806045	-0.012688	<u>3.116652</u>	0.992604
BAURU	3.802923	0.026231	5.964926	0.988053
BIRIGUI	4.541682	-0.007759	3.599867	0.999735
BOTUCATU	<u>2.615209</u>	0.025604	3.633995	0.990038
CATANDUVA	<u>5.784738</u>	0.024033	3.878590	0.997133
FRANCA	3.900918	0.040574	5.977223	0.993451
JAÚ	4.375291	0.001482	3.822674	0.999681
MARÍLIA	4.054692	0.049665	4.441052	0.991717
OURINHOS	4.044486	0.015277	3.337119	0.999516
PRESIDENTE PRUDENTE	4.335721	0.034189	5.146679	0.994986
SÃO CARLOS	3.555486	-0.004551	4.841151	0.996417
SÃO JOSÉ DO RIO PRETO	4.512765	-0.010600	<u>6.851743</u>	0.998845

1990				
Município	α	c	D	R^2
ARAÇATUBA	3.734796	-0.037716	5.755702	0.991013
ARARAQUARA	5.032964	0.023100	5.745527	0.995946
ASSIS	4.275457	0.018191	4.318561	0.995566
BARRETOS	4.496217	0.007761	4.478411	0.999805
BAURU	3.631712	0.025800	6.658368	0.989590
BIRIGUI	4.674804	0.012237	4.267611	0.996779
BOTUCATU	<u>2.868209</u>	0.026006	4.306930	0.995320
CATANDUVA	<u>5.201176</u>	0.026431	4.028546	0.996098
FRANCA	4.301587	0.052545	6.654106	0.984196
JAÚ	4.401231	0.003706	4.242017	0.998768
MARÍLIA	3.824041	0.067047	4.640984	0.989121
OURINHOS	3.947091	0.023023	<u>3.705653</u>	0.999329
PRESIDENTE PRUDENTE	4.426266	0.049923	5.466514	0.996578
SÃO CARLOS	3.623020	-0.023810	5.842755	0.996068
SÃO JOSÉ DO RIO PRETO	4.013610	-0.019533	<u>7.850585</u>	0.997372

Fonte: Elaboração própria.

Tabela 16 – Parâmetros do ajuste da função sigmoide de densidade territorial urbana para 1995 e 2000.

1995				
Município	α	c	D	R^2
ARAÇATUBA	4.006830	-0.029919	6.635944	0.994307
ARARAQUARA	<u>5.052158</u>	0.033813	6.434897	0.997645
ASSIS	4.308026	0.006413	5.057062	0.997363
BARRETOS	4.314612	-0.007035	5.241675	0.999194
BAURU	3.991751	0.013806	8.277492	0.991441
BIRIGUI	4.529933	-0.011304	4.871403	0.995987
BOTUCATU	<u>3.240791</u>	0.034233	4.969307	0.994883
CATANDUVA	4.546561	0.014625	4.974375	0.996640
FRANCA	3.892666	0.032843	7.583232	0.982845
JAÚ	4.211931	-0.010108	4.976544	0.997761
MARÍLIA	3.474264	0.068635	5.263327	0.984795
OURINHOS	4.634600	0.029428	<u>4.632002</u>	0.998621
PRESIDENTE PRUDENTE	4.104174	0.057653	5.927107	0.994770
SÃO CARLOS	4.204511	-0.005344	6.682535	0.997988
SÃO JOSÉ DO RIO PRETO	4.358511	-0.044667	<u>9.078149</u>	0.998595
2000				
Município	α	c	D	R^2
ARAÇATUBA	4.164787	-0.032559	7.057167	0.994918
ARARAQUARA	5.037927	0.038920	7.060898	0.996323
ASSIS	4.304708	0.006250	5.272473	0.996897
BARRETOS	4.422898	-0.012014	5.566445	0.998366
BAURU	4.337931	0.015600	9.068822	0.992265
BIRIGUI	4.419537	-0.004011	<u>5.220809</u>	0.996324
BOTUCATU	3.401533	0.051617	5.305079	0.991986
CATANDUVA	4.607730	0.016845	5.433080	0.997704
FRANCA	3.901764	0.026096	8.026445	0.984267
JAÚ	4.176160	0.003448	5.414808	0.998460
MARÍLIA	<u>3.136893</u>	0.055967	5.876369	0.980533
OURINHOS	<u>5.403994</u>	0.041609	5.221131	0.998542
PRESIDENTE PRUDENTE	3.640060	0.052254	6.308769	0.989036
SÃO CARLOS	4.693110	0.019792	7.282824	0.998372
SÃO JOSÉ DO RIO PRETO	4.986292	-0.014919	<u>9.734482</u>	0.998701

Fonte: Elaboração própria.

Tabela 17 – Parâmetros do ajuste da função sigmoide de densidade territorial urbana para 2005 e 2010.

2005				
Município	α	c	D	R^2
ARAÇATUBA	4.135794	-0.041857	7.406790	0.994724
ARARAQUARA	4.993987	0.038388	7.376583	0.996502
ASSIS	4.227524	0.003550	<u>5.459353</u>	0.995420
BARRETOS	4.731226	-0.007300	5.837538	0.999500
BAURU	4.459314	0.008593	9.602700	0.992293
BIRIGUI	4.454603	-0.005560	5.518808	0.996703
BOTUCATU	3.286810	0.046442	5.522535	0.989867
CATANDUVA	4.499282	0.011176	5.824175	0.998102
FRANCA	3.862564	0.014954	8.624633	0.985438
JAÚ	4.083505	0.008674	5.712958	0.997628
MARÍLIA	<u>3.085129</u>	0.059446	6.055176	0.978231
OURINHOS	<u>5.554943</u>	0.029566	5.573172	0.997633
PRESIDENTE PRUDENTE	3.569458	0.057221	6.553469	0.988867
SÃO CARLOS	4.701168	0.021128	7.596935	0.997117
SÃO JOSÉ DO RIO PRETO	5.253063	-0.002103	<u>10.280799</u>	0.999025
2010				
Município	α	c	D	R^2
ARAÇATUBA	4.152442	-0.044139	7.692780	0.994668
ARARAQUARA	4.672658	0.032654	8.002919	0.995725
ASSIS	4.192967	0.002240	<u>5.626012</u>	0.995335
BARRETOS	4.884036	-0.009643	6.130865	0.999522
BAURU	4.668662	0.010770	9.839091	0.990389
BIRIGUI	4.827656	0.004257	5.872654	0.996836
BOTUCATU	3.315655	0.050298	5.700794	0.986020
CATANDUVA	4.792667	0.009276	6.149472	0.998734
FRANCA	3.926796	-0.006235	9.305098	0.987881
JAÚ	3.942179	0.012469	5.885711	0.996389
MARÍLIA	<u>3.094572</u>	0.061090	6.196057	0.978147
OURINHOS	<u>5.708274</u>	0.029188	5.802549	0.997821
PRESIDENTE PRUDENTE	3.503824	0.054092	6.809631	0.988425
SÃO CARLOS	4.763829	0.031041	7.828559	0.994772
SÃO JOSÉ DO RIO PRETO	5.389343	0.015575	<u>10.780099</u>	0.998261

Fonte: Elaboração própria.

Tabela 18 – Parâmetros do ajuste da função sigmoide de densidade territorial urbana para 2015 e 2020.

2015				
Município	α	c	D	R^2
ARAÇATUBA	4.196270	-0.018121	7.853774	0.995211
ARARAQUARA	4.419180	0.031916	8.446547	0.995999
ASSIS	4.196968	-0.003120	5.878638	0.995688
BARRETOS	4.698113	-0.000361	6.322450	0.999696
BAURU	4.695372	0.007707	10.160896	0.992235
BIRIGUI	5.288254	0.022227	6.188670	0.998793
BOTUCATU	3.138093	0.048184	6.096773	0.981281
CATANDUVA	5.622684	0.015459	6.531324	0.998665
FRANCA	4.069079	-0.022185	9.839433	0.990521
JAÚ	3.898825	0.012650	6.053726	0.995885
MARÍLIA	<u>2.963135</u>	0.060498	6.497763	0.976013
OURINHOS	<u>5.715901</u>	0.031910	6.138067	0.996850
PRESIDENTE PRUDENTE	3.312684	0.056679	7.066030	0.984681
SÃO CARLOS	4.771217	0.036396	8.028696	0.994412
SÃO JOSÉ DO RIO PRETO	5.391454	0.036229	<u>11.267756</u>	0.997610
2020				
Município	α	c	D	R^2
ARAÇATUBA	4.192953	-0.017166	8.035944	0.996402
ARARAQUARA	4.468936	0.029868	8.857383	0.997043
ASSIS	4.265862	-0.007464	<u>6.156178</u>	0.996665
BARRETOS	4.451562	-0.002193	6.795182	0.999503
BAURU	4.769610	0.012939	10.291868	0.992450
BIRIGUI	5.095754	0.015799	6.370457	0.998803
BOTUCATU	3.095446	0.042939	6.371389	0.981592
CATANDUVA	<u>5.834992</u>	0.020298	6.693564	0.998234
FRANCA	4.097291	-0.021454	10.106861	0.991712
JAÚ	3.861883	0.012544	6.202644	0.995791
MARÍLIA	<u>2.854007</u>	0.048834	6.849202	0.974867
OURINHOS	5.601764	0.034143	6.252083	0.996121
PRESIDENTE PRUDENTE	3.179285	0.052034	7.369185	0.980622
SÃO CARLOS	4.817107	0.040710	8.135550	0.993849
SÃO JOSÉ DO RIO PRETO	5.282821	0.061367	<u>11.475492</u>	0.997126

Fonte: Elaboração própria.

**APÊNDICE C – COORDENADAS DOS PONTOS QUE CARACTERIZAM A
FORMA DAS ÁREAS URBANIZADAS E MEDIDAS DE COMPACIDADE
URBANA**

Tabela 19 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 1985.

Município	r_1	r_2	f_1	f_2	k_s	k_p
ARAÇATUBA	1.502711	3.088947	0.792097	0.224096	0.358081	0.345460
ARARAQUARA	1.536833	3.292903	0.781905	0.186060	0.339306	0.363595
ASSIS	1.461776	2.558610	0.793217	0.228276	0.515065	0.272818
BARRETOS	0.826959	2.289693	0.785994	0.201318	0.399714	0.469329
BAURU	1.949633	4.015294	0.794218	0.232013	0.272167	0.346301
BIRIGUI	1.278005	2.321862	0.787035	0.205206	0.557384	0.289971
BOTUCATU	0.902001	2.731995	0.794086	0.231518	0.307415	0.503576
CATANDUVA	1.497794	2.380796	0.793754	0.230279	0.638135	0.227661
FRANCA	1.979651	3.997572	0.797249	0.243324	0.274503	0.337602
JAÚ	1.336027	2.486647	0.788988	0.212494	0.501029	0.300999
MARÍLIA	1.499303	2.941749	0.799170	0.250494	0.380379	0.324798
OURINHOS	1.125247	2.211872	0.791903	0.223374	0.523207	0.325618
PRESIDENTE PRUDENTE	1.791698	3.354980	0.795900	0.238289	0.356693	0.303746
SÃO CARLOS	1.523991	3.317160	0.787713	0.207736	0.323437	0.370401
SÃO JOSÉ DO RIO PRETO	2.426102	4.425641	0.786435	0.202965	0.291802	0.291829

Fonte: Elaboração própria.

Tabela 20 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 1990.

Município	r_1	r_2	f_1	f_2	k_s	k_p
ARAÇATUBA	1.863069	3.892634	0.780705	0.181579	0.295199	0.352618
ARARAQUARA	2.121058	3.624469	0.793557	0.229544	0.375156	0.261666
ASSIS	1.494163	2.824397	0.792519	0.225672	0.426126	0.308027
BARRETOS	1.583335	2.895076	0.790315	0.217446	0.436724	0.292903
BAURU	2.121932	4.536436	0.794127	0.231672	0.232948	0.362627
BIRIGUI	1.532683	2.734928	0.791261	0.220976	0.474350	0.281714
BOTUCATU	1.164687	3.142243	0.794171	0.231835	0.284359	0.459157
CATANDUVA	1.504252	2.524294	0.794261	0.232171	0.551045	0.253204
FRANCA	2.308456	4.345650	0.799779	0.252766	0.268513	0.306156
JAÚ	1.486350	2.755667	0.789458	0.214248	0.453166	0.299225
MARÍLIA	1.521341	3.119643	0.802844	0.264203	0.337008	0.344389
OURINHOS	1.234626	2.471027	0.793540	0.229483	0.456209	0.333653
PRESIDENTE PRUDENTE	1.920025	3.546490	0.799225	0.250698	0.337251	0.297532
SÃO CARLOS	1.859466	3.983290	0.783643	0.192546	0.278317	0.363497
SÃO JOSÉ DO RIO PRETO	2.637314	5.213270	0.784547	0.195920	0.228508	0.328123

Fonte: Elaboração própria.

Tabela 21 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 1995.

Município	r_1	r_2	f_1	f_2	k_s	k_p
ARAÇATUBA	2.227427	4.408516	0.782352	0.187729	0.272627	0.328678
ARARAQUARA	2.378749	4.056148	0.795821	0.237993	0.332555	0.260672
ASSIS	1.755563	3.301499	0.790030	0.216383	0.371068	0.305698
BARRETOS	1.820874	3.420801	0.787188	0.205777	0.363399	0.305232
BAURU	2.773292	5.504199	0.791593	0.222213	0.208495	0.329920
BIRIGUI	1.727586	3.143816	0.786286	0.202410	0.412275	0.290723
BOTUCATU	1.474967	3.494340	0.795909	0.238323	0.276118	0.406369
CATANDUVA	1.766749	3.207627	0.791766	0.222859	0.394833	0.289660
FRANCA	2.508846	5.074386	0.795616	0.237228	0.217649	0.338318
JAÚ	1.710257	3.266288	0.786539	0.203353	0.374791	0.312673
MARÍLIA	1.634103	3.629224	0.803179	0.265455	0.269519	0.379061
OURINHOS	1.657892	2.974111	0.794894	0.234534	0.425735	0.284158
PRESIDENTE PRUDENTE	2.012602	3.914506	0.800858	0.256794	0.286063	0.320882
SÃO CARLOS	2.294700	4.387835	0.787546	0.207110	0.277304	0.313225
SÃO JOSÉ DO RIO PRETO	3.167559	5.910590	0.779236	0.176098	0.219880	0.302158

Fonte: Elaboração própria.

Tabela 22 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 2000.

Município	r_1	r_2	f_1	f_2	k_s	k_p
ARAÇATUBA	2.412802	4.644365	0.781794	0.185646	0.267144	0.316212
ARARAQUARA	2.607560	4.453339	0.796900	0.242020	0.300621	0.261408
ASSIS	1.829722	3.442751	0.789996	0.216254	0.355692	0.305934
BARRETOS	1.954493	3.611951	0.786136	0.201850	0.352519	0.297759
BAURU	3.157805	5.911018	0.791972	0.223628	0.206429	0.303591
BIRIGUI	1.832542	3.388267	0.787827	0.208162	0.372602	0.297985
BOTUCATU	1.625567	3.679512	0.799583	0.252034	0.266584	0.387166
CATANDUVA	1.940113	3.492967	0.792235	0.224610	0.365536	0.285815
FRANCA	2.658645	5.367800	0.794190	0.231906	0.207549	0.337529
JAÚ	1.853621	3.561187	0.789404	0.214044	0.336947	0.315351
MARÍLIA	1.704651	4.171718	0.800502	0.255465	0.220925	0.419828
OURINHOS	1.974369	3.246762	0.797468	0.244141	0.434871	0.243701
PRESIDENTE	2.013143	4.295627	0.799718	0.252536	0.239731	0.361795
PRUDENTE						
SÃO CARLOS	2.619577	4.663247	0.792858	0.226935	0.276915	0.280615
SÃO JOSÉ DO	3.581727	6.152755	0.785522	0.199559	0.227910	0.264116
RIO PRETO						

Fonte: Elaboração própria.

Tabela 23 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 2005.

Município	r_1	r_2	f_1	f_2	k_s	k_p
ARAÇATUBA	2.524126	4.882664	0.779830	0.178313	0.255038	0.318429
ARARAQUARA	2.715658	4.660925	0.796787	0.241600	0.285404	0.263709
ASSIS	1.879329	3.580025	0.789425	0.214125	0.338274	0.311520
BARRETOS	2.106317	3.731221	0.787132	0.205568	0.357907	0.278354
BAURU	3.383380	6.219320	0.790491	0.218102	0.201834	0.295327
BIRIGUI	1.943615	3.575193	0.787500	0.206940	0.355827	0.295640
BOTUCATU	1.654885	3.867650	0.798489	0.247953	0.248800	0.400679
CATANDUVA	2.059708	3.764467	0.791037	0.220140	0.334885	0.292704
FRANCA	2.842015	5.782618	0.791835	0.223119	0.193401	0.340954
JAÚ	1.935246	3.777712	0.790508	0.218166	0.310639	0.322507
MARÍLIA	1.735194	4.319982	0.801237	0.258208	0.210086	0.426873
OURINHOS	2.125947	3.447226	0.794923	0.234643	0.424044	0.237078
PRESIDENTE	2.067778	4.485691	0.800767	0.256454	0.225117	0.368952
PRUDENTE						
SÃO CARLOS	2.734387	4.862547	0.793140	0.227988	0.265559	0.280134
SÃO JOSÉ DO	3.851688	6.429112	0.788231	0.209667	0.224474	0.250703
RIO PRETO						

Fonte: Elaboração própria.

Tabela 24 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 2010.

Município	r_1	r_2	f_1	f_2	k_s	k_p
ARAÇATUBA	2.626498	5.066282	0.779347	0.176513	0.247085	0.317152
ARARAQUARA	2.873675	5.129244	0.795576	0.237078	0.247608	0.281843
ASSIS	1.929477	3.696535	0.789148	0.213091	0.325998	0.314087
BARRETOS	2.238853	3.892012	0.786637	0.203720	0.352608	0.269645
BAURU	3.531818	6.307273	0.790951	0.219819	0.205780	0.282084
BIRIGUI	2.135314	3.737340	0.789575	0.214683	0.358853	0.272794
BOTUCATU	1.718237	3.982556	0.799304	0.250994	0.242152	0.397194
CATANDUVA	2.229842	3.919630	0.790635	0.218641	0.338501	0.274786
FRANCA	3.092191	6.212907	0.787357	0.206408	0.186159	0.335377
JAÚ	1.959741	3.925970	0.791310	0.221159	0.289972	0.334068
MARÍLIA	1.779601	4.416456	0.801585	0.259505	0.205578	0.425570
OURINHOS	2.231921	3.570628	0.794843	0.234345	0.418686	0.230710
PRESIDENTE PRUDENTE	2.125072	4.684559	0.800106	0.253986	0.213371	0.375863
SÃO CARLOS	2.832180	4.996379	0.795235	0.235806	0.258492	0.276449
SÃO JOSÉ DO RIO PRETO	4.072920	6.707179	0.791966	0.223609	0.215756	0.244363

Fonte: Elaboração própria.

Tabela 25 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 2015.

Município	r_1	r_2	f_1	f_2	k_s	k_p
ARAÇATUBA	2.694473	5.159301	0.784846	0.197034	0.238480	0.313840
ARARAQUARA	2.964698	5.481848	0.795420	0.236496	0.222046	0.298009
ASSIS	2.016997	3.861641	0.788016	0.208864	0.313964	0.313788
BARRETOS	2.275082	4.047367	0.788599	0.211041	0.325883	0.280316
BAURU	3.655485	6.505411	0.790304	0.217403	0.201023	0.280480
BIRIGUI	2.323739	3.864931	0.793372	0.228855	0.366286	0.249034
BOTUCATU	1.769076	4.327697	0.798857	0.249326	0.214776	0.419668
CATANDUVA	2.500772	4.030552	0.791942	0.223517	0.371573	0.234222
FRANCA	3.327451	6.511982	0.783987	0.193828	0.185320	0.323650
JAÚ	2.004440	4.049286	0.791348	0.221302	0.278772	0.337783
MARÍLIA	1.804925	4.692838	0.801460	0.259038	0.187825	0.444447
OURINHOS	2.361921	3.776146	0.795418	0.236491	0.395218	0.230402
PRESIDENTE PRUDENTE	2.128465	4.937565	0.800653	0.256026	0.193879	0.397550
SÃO CARLOS	2.906303	5.122393	0.796366	0.240029	0.251044	0.276021
SÃO JOSÉ DO RIO PRETO	4.257705	7.010051	0.796331	0.239898	0.202167	0.244267

Fonte: Elaboração própria.

Tabela 26 – Coordenadas dos pontos que caracterizam a forma das áreas urbanizadas e medidas de compacidade urbana para 2020.

Município	r_1	r_2	f_1	f_2	k_s	k_p
ARAÇATUBA	2.755974	5.279970	0.785047	0.197786	0.232671	0.314088
ARARAQUARA	3.123594	5.733788	0.794987	0.234882	0.214584	0.294691
ASSIS	2.127821	4.028357	0.787098	0.205438	0.306050	0.308720
BARRETOS	2.392442	4.402739	0.788212	0.209595	0.287826	0.295842
BAURU	3.725069	6.566799	0.791409	0.221530	0.200540	0.276114
BIRIGUI	2.362031	4.008425	0.792014	0.223785	0.345135	0.258442
BOTUCATU	1.830341	4.541047	0.797749	0.245190	0.203843	0.425450
CATANDUVA	2.591414	4.102151	0.792964	0.227333	0.374407	0.225700
FRANCA	3.429150	6.677712	0.784141	0.194404	0.181538	0.321421
JAÚ	2.043727	4.158917	0.791326	0.221218	0.269530	0.341014
MARÍLIA	1.844348	5.004854	0.798995	0.249839	0.173756	0.461441
OURINHOS	2.391119	3.860964	0.795890	0.238252	0.379385	0.235097
PRESIDENTE	2.158322	5.210863	0.799671	0.252363	0.179296	0.414231
PRUDENTE						
SÃO CARLOS	2.955679	5.179871	0.797278	0.243432	0.249010	0.273392
SÃO JOSÉ DO	4.307380	7.168111	0.801643	0.259724	0.189434	0.249291
RIO PRETO						

Fonte: Elaboração própria.

APÊNDICE D – ESTATÍSTICAS DESCRIPTIVAS ANUAIS PARA OS PARÂMETROS DE AJUSTE DA FUNÇÃO DE DENSIDADE TERRITORIAL URBANA, DAS COORDENADAS DOS PONTOS QUE CARACTERIZAM A FORMA DAS ÁREAS URBANIZADAS E DAS MEDIDAS DE COMPACIDADE URBANA

Tabela 27 – Estatísticas descritivas para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (α).

Ano	Média	Desvio padrão	Mín.	Q1	Mediana	Q3	Máx.
1985	4.0394	0.7759	2.6152	3.7125	4.0445	4.4440	5.7847
1990	4.1635	0.5966	2.8682	3.7794	4.2755	4.4612	5.2012
1995	4.1914	0.4491	3.2408	3.9993	4.2119	4.4442	5.0522
2000	4.3090	0.6153	3.1369	4.0333	4.3379	4.6504	5.4040
2005	4.3266	0.6916	3.0851	3.9730	4.4546	4.7162	5.5549
2010	4.3890	0.7445	3.0946	3.9345	4.6687	4.8102	5.7083
2015	4.4251	0.8696	2.9631	3.9840	4.4192	5.0297	5.7159
2020	4.3913	0.8907	2.8540	3.9796	4.4516	4.9564	5.8350

Fonte: Elaboração própria.

Tabela 28 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (α).

Ano	Amplitude Variação (CV)	Amplitude Interquartil (IQR)	Coeficiente de dispersão de quartil (QCD)
1985	3.1695	0.1921	0.7315
1990	2.3330	0.1433	0.6818
1995	1.8114	0.1072	0.4450
2000	2.2671	0.1428	0.6171
2005	2.4698	0.1599	0.7432
2010	2.6137	0.1696	0.8757
2015	2.7528	0.1965	1.0458
2020	2.9810	0.2028	0.9768

Fonte: Elaboração própria.

Tabela 29 – Estatísticas descritivas para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (c).

Ano	Média	Desvio padrão	Mín.	Q1	Mediana	Q3	Máx.
1985	0.0125	0.0227	-0.0320	-0.0062	0.0162	0.0259	0.0497
1990	0.0170	0.0286	-0.0377	0.0057	0.0230	0.0262	0.0670
1995	0.0122	0.0313	-0.0447	-0.0086	0.0138	0.0333	0.0686
2000	0.0177	0.0270	-0.0326	-0.0003	0.0168	0.0403	0.0560
2005	0.0162	0.0270	-0.0419	0.0007	0.0112	0.0340	0.0594
2010	0.0169	0.0274	-0.0441	0.0032	0.0125	0.0318	0.0611
2015	0.0211	0.0253	-0.0222	0.0037	0.0222	0.0363	0.0605
2020	0.0215	0.0257	-0.0215	0.0052	0.0203	0.0418	0.0614

Fonte: Elaboração própria.

Tabela 30 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (c).

Ano	Amplitude Variação (CV)	Amplitude Interquartil (IQR)	Coeficiente de dispersão de quartil (QCD)
1985	0.0817	1.8192	1.2313
1990	0.1048	1.6844	0.7871
1995	0.1133	2.5635	1.2486
2000	0.0885	1.5276	1.0067
2005	0.1013	1.6705	0.9794
2010	0.1052	1.6247	0.9012
2015	0.0827	1.2027	0.9025
2020	0.0828	1.1933	0.8814

Fonte: Elaboração própria.

Tabela 31 – Estatísticas descritivas para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (D).

Ano	Média	Desvio padrão	Mín.	Q1	Mediana	Q3	Máx.
1985	4.5369	1.0822	3.1167	3.7283	4.4411	4.9939	6.8517
1990	5.1975	1.1999	3.7057	4.2873	4.6410	5.7992	7.8506
1995	6.0403	1.3764	4.6320	4.9755	5.2633	6.6592	9.0781
2000	6.5233	1.4745	5.2208	5.3599	5.8764	7.1719	9.7345
2005	6.8630	1.5773	5.4594	5.6431	6.0552	7.5019	10.2808
2010	7.1748	1.6703	5.6260	5.8792	6.1961	7.9157	10.7801
2015	7.4914	1.7286	5.8786	6.1634	6.5313	8.2376	11.2678
2020	7.7309	1.7178	6.1562	6.3709	6.8492	8.4965	11.4755

Fonte: Elaboração própria.

Tabela 32 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (D).

Ano	Amplitude	Coeficiente de Variação (CV)	Amplitude Interquartil (IQR)	Coeficiente de dispersão de quartil (QCD)
1985	3.7351	0.2385	1.2656	3.9818
1990	4.1449	0.2309	1.5120	4.5480
1995	4.4461	0.2279	1.6838	5.2283
2000	4.5137	0.2260	1.8119	5.6126
2005	4.8214	0.2298	1.8588	5.8908
2010	5.1541	0.2328	2.0366	6.1365
2015	5.3891	0.2307	2.0743	6.4152
2020	5.3193	0.2222	2.1255	6.6211

Fonte: Elaboração própria.

Tabela 33 – Estatísticas descritivas para as coordenadas dos pontos que caracterizam a forma das áreas urbanizadas por ano (f_1).

Ano	Média	Desvio padrão	Mín.	Q1	Mediana	Q3	Máx.
1985	0.7913	0.0048	0.7819	0.7874	0.7921	0.7942	0.7992
1990	0.7923	0.0060	0.7807	0.7899	0.7935	0.7942	0.8028
1995	0.7913	0.0066	0.7792	0.7869	0.7916	0.7957	0.8032
2000	0.7924	0.0057	0.7818	0.7886	0.7922	0.7972	0.8005
2005	0.7921	0.0057	0.7798	0.7888	0.7910	0.7959	0.8012
2010	0.7922	0.0058	0.7793	0.7894	0.7913	0.7954	0.8016
2015	0.7931	0.0054	0.7840	0.7895	0.7934	0.7963	0.8015
2020	0.7932	0.0054	0.7841	0.7898	0.7930	0.7975	0.8016

Fonte: Elaboração própria.

Tabela 34 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (f_1).

Ano	Amplitude	Coeficiente de Variação (CV)	Amplitude Interquartil (IQR)	Coeficiente de dispersão de quartil (QCD)
1985	0.0173	0.0061	0.0068	0.7959
1990	0.0221	0.0076	0.0043	0.7953
1995	0.0239	0.0084	0.0089	0.7980
2000	0.0187	0.0072	0.0086	0.7994
2005	0.0214	0.0072	0.0070	0.7977
2010	0.0222	0.0073	0.0060	0.7970
2015	0.0175	0.0068	0.0069	0.7981
2020	0.0175	0.0069	0.0077	0.7995

Fonte: Elaboração própria.

Tabela 35 – Estatísticas descritivas para as coordenadas dos pontos que caracterizam a forma das áreas urbanizadas por ano (f_2).

Ano	Média	Desvio padrão	Mín.	Q1	Mediana	Q3	Máx.
1985	0.2212	0.0179	0.1861	0.2065	0.2241	0.2318	0.2505
1990	0.2247	0.0226	0.1816	0.2158	0.2295	0.2320	0.2642
1995	0.2210	0.0247	0.1761	0.2046	0.2222	0.2376	0.2655
2000	0.2253	0.0213	0.1856	0.2111	0.2246	0.2431	0.2555
2005	0.2241	0.0213	0.1783	0.2119	0.2201	0.2381	0.2582
2010	0.2246	0.0216	0.1765	0.2139	0.2212	0.2364	0.2595
2015	0.2279	0.0200	0.1938	0.2142	0.2289	0.2400	0.2590
2020	0.2283	0.0203	0.1944	0.2154	0.2273	0.2443	0.2597

Fonte: Elaboração própria.

Tabela 36 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (f_2).

Ano	Amplitude	Coeficiente de Variação (CV)	Amplitude Interquartil (IQR)	Coeficiente de dispersão de quartil (QCD)
1985	0.0644	0.0809	0.0253	0.3156
1990	0.0826	0.1004	0.0162	0.2855
1995	0.0894	0.1117	0.0330	0.3436
2000	0.0698	0.0945	0.0320	0.3427
2005	0.0799	0.0950	0.0262	0.3220
2010	0.0830	0.0962	0.0226	0.3093
2015	0.0652	0.0877	0.0257	0.3215
2020	0.0653	0.0888	0.0289	0.3337

Fonte: Elaboração própria.

Tabela 37 – Estatísticas descritivas para as coordenadas dos pontos que caracterizam a forma das áreas urbanizadas por ano (r_1).

Ano	Média	Desvio padrão	Mín.	Q1	Mediana	Q3	Máx.
1985	1.5092	0.4144	0.8270	1.3070	1.4993	1.6643	2.4261
1990	1.7569	0.4105	1.1647	1.4992	1.5833	2.0205	2.6373
1995	2.0607	0.4839	1.4750	1.7189	1.8209	2.3367	3.1676
2000	2.2511	0.5711	1.6256	1.8431	1.9744	2.6136	3.5817
2005	2.3706	0.6304	1.6549	1.9394	2.1063	2.7250	3.8517
2010	2.4918	0.6755	1.7182	2.0424	2.2319	2.8529	4.0729
2015	2.5994	0.7129	1.7691	2.0727	2.3619	2.9355	4.2577
2020	2.6692	0.7199	1.8303	2.1431	2.3924	3.0396	4.3074

Fonte: Elaboração própria.

Tabela 38 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (r_1).

Ano	Amplitude	Coeficiente de Variação (CV)	Amplitude Interquartil (IQR)	Coeficiente de dispersão de quartil (QCD)
1985	1.5991	0.2746	0.3573	1.5217
1990	1.4726	0.2337	0.5213	1.7572
1995	1.6926	0.2348	0.6178	1.9833
2000	1.9562	0.2537	0.7705	2.1379
2005	2.1968	0.2659	0.7856	2.2277
2010	2.3547	0.2711	0.8105	2.3265
2015	2.4886	0.2742	0.8628	2.3666
2020	2.4770	0.2697	0.8966	2.4380

Fonte: Elaboração própria.

Tabela 39 – Estatísticas descritivas para as coordenadas dos pontos que caracterizam a forma das áreas urbanizadas por ano (r_2).

Ano	Média	Desvio padrão	Mín.	Q1	Mediana	Q3	Máx.
1985	3.0277	0.7005	2.2119	2.4337	2.9417	3.3361	4.4256
1990	3.4406	0.8160	2.4710	2.7900	3.1422	3.9380	5.2133
1995	3.9796	0.9086	2.9741	3.2839	3.6292	4.3982	5.9106
2000	4.2722	0.9327	3.2468	3.5271	4.1717	4.6538	6.1528
2005	4.4924	0.9828	3.4472	3.7478	4.3200	4.8726	6.4291
2010	4.6830	1.0340	3.5706	3.9058	4.4165	5.0978	6.7072
2015	4.8919	1.0715	3.7761	4.0390	4.6928	5.3206	7.0101
2020	5.0616	1.0661	3.8610	4.1305	5.0049	5.5069	7.1681

Fonte: Elaboração própria.

Tabela 40 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (r_2).

Ano	Amplitude	Coeficiente de Variação (CV)	Amplitude Interquartil (IQR)	Coeficiente de dispersão de quartil (QCD)
1985	2.2138	0.2314	0.9023	2.7042
1990	2.7422	0.2372	1.1479	3.0815
1995	2.9365	0.2283	1.1143	3.5372
2000	2.9060	0.2183	1.1267	3.7692
2005	2.9819	0.2188	1.1248	3.9787
2010	3.1366	0.2208	1.1919	4.1396
2015	3.2339	0.2190	1.2816	4.2798
2020	3.3071	0.2106	1.3763	4.3805

Fonte: Elaboração própria.

Tabela 41 – Estatísticas descritivas para as medidas de compacidade urbana por ano (k_s).

Ano	Média	Desvio padrão	Mín.	Q1	Mediana	Q3	Máx.
1985	0.4026	0.1152	0.2722	0.3154	0.3581	0.5080	0.6381
1990	0.3623	0.0991	0.2285	0.2813	0.3373	0.4450	0.5510
1995	0.3135	0.0732	0.2085	0.2711	0.2861	0.3729	0.4257
2000	0.2955	0.0706	0.2064	0.2338	0.2769	0.3541	0.4349
2005	0.2821	0.0688	0.1934	0.2248	0.2656	0.3366	0.4240
2010	0.2738	0.0698	0.1862	0.2146	0.2476	0.3322	0.4187
2015	0.2632	0.0734	0.1853	0.2016	0.2385	0.3199	0.3952
2020	0.2525	0.0714	0.1738	0.1950	0.2327	0.2969	0.3794

Fonte: Elaboração própria.

Tabela 42 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (k_s).

Ano	Amplitude Variação (CV)	Amplitude Interquartil (IQR)	Coeficiente de dispersão de quartil (QCD)
1985	0.3660	0.2863	0.6946
1990	0.3225	0.2734	0.6490
1995	0.2172	0.2334	0.5442
2000	0.2284	0.2390	0.5735
2005	0.2306	0.2439	0.5569
2010	0.2325	0.2550	0.5688
2015	0.2099	0.2790	0.5715
2020	0.2056	0.2829	0.5383

Fonte: Elaboração própria.

Tabela 43 – Estatísticas descritivas para as medidas de compacidade urbana por ano (k_p).

Ano	Média	Desvio padrão	Mín.	Q1	Mediana	Q3	Máx.
1985	0.3382	0.0708	0.2277	0.2964	0.3256	0.3549	0.5036
1990	0.3230	0.0508	0.2532	0.2952	0.3080	0.3485	0.4592
1995	0.3178	0.0367	0.2607	0.2964	0.3127	0.3293	0.4064
2000	0.3119	0.0479	0.2437	0.2832	0.3036	0.3269	0.4198
2005	0.3122	0.0535	0.2371	0.2792	0.2956	0.3317	0.4269
2010	0.3088	0.0561	0.2307	0.2738	0.2821	0.3347	0.4256
2015	0.3096	0.0665	0.2304	0.2625	0.2980	0.3307	0.4444
2020	0.3130	0.0708	0.2257	0.2659	0.2958	0.3312	0.4614

Fonte: Elaboração própria.

Tabela 44 – Estatísticas descritivas (medidas de dispersão) para os parâmetros de ajuste da função sigmoide de densidade territorial urbana por ano (k_p).

Ano	Amplitude	Coeficiente de Variação (CV)	Amplitude Interquartil (IQR)	Coeficiente de dispersão de quartil (QCD)
1985	0.2759	0.2094	0.0585	0.4613
1990	0.2060	0.1574	0.0533	0.4481
1995	0.1457	0.1153	0.0329	0.3962
2000	0.1761	0.1536	0.0437	0.4168
2005	0.1898	0.1713	0.0525	0.4375
2010	0.1949	0.1818	0.0609	0.4558
2015	0.2140	0.2148	0.0682	0.4687
2020	0.2357	0.2261	0.0653	0.4631

Fonte: Elaboração própria.

APÊNDICE E – TAXAS DE CRESCIMENTO DE r_1 E r_2 PARA CADA INTERVALO

Tabela 45 – Taxas de crescimento de r_1 (δr_1) e r_2 (δr_2) por intervalos.

r_1 (δr_1)								
Município	1985-1990	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015	2015-2020	
ARAÇATUBA	1.1131	1.4183	0.8172	0.5170	0.4998	0.3444	0.3173	
ARARAQUARA	1.9238	0.9340	0.9281	0.4814	0.7365	0.4669	0.8710	
ASSIS	0.0829	0.7383	0.2448	0.1708	0.1795	0.3235	0.4280	
BARRETOS	1.7319	0.6877	0.4571	0.5484	0.4945	0.1410	0.4750	
BAURU	0.6918	2.9549	2.1164	1.3334	0.9232	0.7800	0.4527	
BIRIGUI	0.5913	0.5330	0.3300	0.3763	0.6854	0.7042	0.1480	
BOTUCATU	0.7177	0.9750	0.5262	0.1079	0.2450	0.2025	0.2651	
CATANDUVA	<u>0.0154</u>	0.6626	0.5561	0.4177	0.6405	1.0619	0.3653	
FRANCA	1.3144	0.8708	0.7601	0.9843	1.4467	1.4616	0.6623	
JAÚ	0.3738	0.6170	0.4683	0.2907	0.0925	0.1755	0.1591	
MARÍLIA	0.0648	0.3518	0.2560	0.1274	0.1918	0.1118	0.1850	
OURINHOS	0.2419	1.0459	0.9412	0.4921	0.3653	0.4642	0.1103	
PRESIDENTE PRUDENTE	0.4305	0.3283	<u>0.0021</u>	0.2347	0.2570	<u>0.0159</u>	0.1474	
SÃO CARLOS	1.1128	1.7337	1.4255	0.5354	0.4755	0.3703	0.2529	
SÃO JOSÉ DO RIO PRETO	0.9348	2.7643	2.4480	1.6610	1.4223	1.2394	0.3482	
r_2 (δr_2)								
Município	1985-1990	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015	2015-2020	
ARAÇATUBA	1.2077	0.9611	0.5253	0.5750	0.4635	0.2443	0.3251	
ARARAQUARA	0.5096	0.9156	0.9448	0.5413	1.2718	1.0133	0.7469	
ASSIS	0.3885	0.7129	0.2480	0.2512	0.2190	0.3186	0.3363	
BARRETOS	0.5006	0.8324	0.3481	0.2331	0.3387	0.3478	0.8085	
BAURU	1.0160	2.0535	1.1282	0.9736	0.2976	0.6998	0.2244	
BIRIGUI	0.5279	0.6267	0.4223	0.3426	0.3152	0.2724	0.3334	
BOTUCATU	0.3700	0.4101	0.2731	0.3058	0.1902	0.5930	0.3774	
CATANDUVA	0.2149	1.0279	0.5041	0.5267	0.3196	0.2473	0.1791	
FRANCA	0.6891	1.6823	0.7361	1.1029	1.2229	0.9248	0.5515	
JAÚ	0.3594	0.7590	0.5044	0.4014	0.2869	0.2417	0.2197	
MARÍLIA	0.2667	0.7752	0.8865	0.2527	0.1674	0.4918	0.5632	
OURINHOS	0.2916	0.6211	0.4520	0.3958	0.2623	0.4587	0.2003	
PRESIDENTE PRUDENTE	0.3431	0.7066	0.7670	0.3826	0.4112	0.5377	0.5817	
SÃO CARLOS	1.0152	0.7522	0.6320	0.5221	0.3659	0.3569	0.1670	
SÃO JOSÉ DO RIO PRETO	1.9109	1.8391	0.7671	0.9898	1.0710	1.2336	0.6730	

Fonte: Elaboração própria.