

IDENTIFICAÇÃO E CARACTERIZAÇÃO DE SUBCENTRALIDADES DE EMPREGO NA REGIÃO METROPOLITANA DE SÃO PAULO¹

Rodger Barros Antunes Campos²

Carlos Roberto Azzoni³

Este artigo identifica as regiões centrais de emprego na região metropolitana (RM) de São Paulo e aponta evolução temporal dessas regiões, bem como caracteriza essas áreas em função da escolaridade, do gênero e da complexidade das atividades econômicas ali localizadas. Todas as empresas do setor privado da base da Relação Anual de Informações Sociais (Rais) foram geocodificadas e distribuídas em células de 1,2 e 4 km², o que permitiu conhecer a distribuição geográfica de emprego em 2002, 2008 e 2014. Aplicou-se metodologia de identificação objetiva de subcentralidades, baseada em regressões geograficamente ponderadas, a qual supera outras abordagens empíricas que demandam conhecimento prévio da região estudada. Os resultados indicam a existência de três subcentralidades na RM de São Paulo em 2002 e apenas duas em 2014, quando utilizados critérios mais restritos de avaliação. Ao relaxar esses critérios, nota-se a ampliação do total de subcentros (*subcenter business districts* – SBDs) a ponto de criar uma grande e única mancha de emprego na região. Os resultados indicam movimento centrífugo do emprego, mas de forma concentrada, e isso sugere que a aglomeração de trabalhadores é um fator importante para a produtividade.

Palavras-chave: estrutura espacial; espraiamento do emprego; centros e subcentros de negócios.

IDENTIFICATION AND CHARACTERIZATION OF EMPLOYMENT SUB-CENTRALITY IN THE METROPOLITAN REGION OF SÃO PAULO

This article identifies the central regions of employment in the metropolitan region (MR) of São Paulo. It points to the temporal evolution of these regions, characterizing these areas according to schooling, gender, and the complexity of the economic activities. All private sector companies in the Relação Anual de Informações Sociais (Rais) database were geocoded and distributed into cells of 1, 2, and 4 km², unveiling the geographical distribution of employment in 2002, 2008, and 2014. An objective methodology for identifying subcentralities was applied, overcoming the limitations of other empirical approaches that require prior knowledge of the region. The results indicate three subcentralities in the MR of São Paulo, when more restricted evaluation criteria are used. As the evaluation criteria are relaxed, the number of subcenter business districts (SBD) increases, to the point of creating a large and unique stain of employment in the region. The results indicate a centrifugal movement of employment, but in a concentrated way, suggesting that the agglomeration is important for productivity.

Keywords: spatial structure; employment spreading; business centers and subcenters.

JEL: R00; R12.

1. DOI: <https://dx.doi.org/10.38116/ppe54n1art3>

2. Pesquisador associado no Núcleo de Economia Regional e Urbana da Universidade de São Paulo (Nereus/USP).
E-mail: rodgercampos@hotmail.com.

3. Professor titular no Departamento de Economia da USP. E-mail: cazzoni@usp.br.

1 INTRODUÇÃO

Uma extensa literatura discute o papel dos centros de negócios na economia (*central business districts* – CBDs) e dos SBDs, bem como a forma que essas áreas impactam salários, preço da terra e sistema de transporte em vários contextos. Li (2020) faz um apanhado de estudos referentes a cidades chinesas, que têm sido objeto de grande número de trabalhos. Nesse segmento, Kwon (2021) lida com Seoul, na Coreia; Hajrasouliha e Hamidi (2017), Liu *et al.* (2019), Manduca (2020) e Giuliano *et al.* (2021) analisam metrópoles americanas; Taubenböck *et al.* (2017) e Krehl (2018) tratam de cidades alemãs; Alidadi e Dadashpoor (2018) abordam o caso de Tehran, no Irã; Fernández-Maldonado *et al.* (2013) destacam cidades latino-americanas; e Aguirre-Nuñez e Marmolejo Duarte (2010) e Aguirre-Nuñez, Marmolejo Duarte e Vergara Perucich (2019) estudam cidades chilenas. No cenário brasileiro, Rodrigues, Silveira Neto e Miranda (2019) consideram as RMs em conjunto; Campos (2018) e Campos e Azzoni (2019; 2021) avaliam a RM de São Paulo; Cerqueira e Diniz (2022) lidam com a RM de Belo Horizonte; e Thomé e Porsse (2019) concentram-se no caso de Curitiba. No entanto, a definição do que seja SBDs não é consensual, sobretudo nos trabalhos empíricos que buscam identificá-los. A definição de CBDs é menos complexa relativamente, já que a maioria das cidades apresenta apenas um centro de negócios, o que garante a essa região tal denominação. A existência e a localização de CBDs basicamente são definidas historicamente pelo surgimento da primeira região central. Modelos teóricos monocêntricos indicam o CBD como a área com os mais elevados preços da terra e salários em comparação a quaisquer outras regiões intraurbanas.

A questão ganha maior relevância quando se trata de cidades com múltiplas concentrações de emprego. O processo de suburbanização das atividades econômicas dá lugar a uma cidade com múltiplos centros. Para conceituar essas áreas, os autores partem das características dos SBDs que consideram principais. Para um grupo de autores, é um lugar onde as atividades de produção, compras e acessibilidade do emprego contam com melhor sistema de transporte do que em qualquer outra área, o que proporciona facilidade de saída e entrada de insumos e produtos, bem como deslocamento pendular (Ogawa e Fugita, 1980; Fujita e Ogawa, 1982; Helsley e Sullivan, 1991; McMillen, 2001b). Nessa perspectiva, a centralização das atividades no SBD está relacionada a um bom sistema de transporte, que oferece acessibilidade eficiente. Alguns trabalhos empíricos utilizam informações sobre o fluxo de passageiros para identificar subcentralidades, considerando tal peculiaridade inerente às áreas centrais (Bourne, 1989; Burns, Moix Bergada e Cladera, 2001; Clark e Kuijpers-Linde, 1994; Gordon e Richardson, 1986; Aguilera e Mignot, 2004). Mais recentemente, outras formas de identificação das interações humanas são utilizadas, como ligações de telefones celulares (Yan *et al.*, 2021; Yan *et al.*, 2023) e o Google Places of Interest (Cerqueira e Diniz, 2022). Para um segundo

grupo de pesquisadores, SBD é a região cuja concentração de firmas apresenta fortes encadeamentos produtivos e de comunicação (Castells e Hall, 1994; Anas e Kim, 1996; Sasaki e Mun, 1996; Graham e Marvin, 1996). Nessa área tendem a se aglomerar atividades de alta ordem da administração pública e privada, bem como o sistema de transporte central. Essas atividades exigem contatos face a face e criam algumas ligações para trás e para frente, dada a concentração da atividade do setor de serviços. Yan, Wang e Zhang (2020) buscam conciliar essas linhas de pensamento, propondo um enfoque que trata simultaneamente da competição e da conexão espaciais.

Partindo dos modelos teóricos desenvolvidos por autores do primeiro grupo, os SBDs são áreas que impactam os preços relativos e concentram os mais elevados aluguéis e salários do que em qualquer outro lugar, exceto nos CBDs (Papageorgiou, 1971; White, 1976; 1988; 1990; 1999; Hartwick e Hartwick, 1974; Romanos, 1979; Sullivan, 1986; Wieand, 1987; Sivitanidou e Wheaton, 1992; Hotchkiss e White, 1993; Ross e Yinger, 1995; Wrede, 2015). Com base nessa característica, alguns autores usam o preço da terra como variável de resultado para identificar subcentralidades (Heikkila *et al.*, 1989; Wen e Tao, 2015; Campos e Chagas, 2021; Deng, 2020). Assim, devido à disponibilidade de infraestrutura, compras e proximidade a locais de trabalho, os preços da terra nessas localidades são relativamente maiores em comparação a outros espaços intraurbanos.

No fim dos anos 1980, vários pesquisadores lançaram esforços para identificar e avaliar a descentralização espacial do emprego na área intraurbana (Gordon, Richardson e Wong, 1986; McDonald, 1987; Cervero, 1989). Embora os trabalhos empíricos tenham buscado identificar SBDs, não existe uma estratégia empírica consolidada na literatura. Inicialmente, os estudos valiam-se do conhecimento prévio da região que era estudada (Bender e Hwang, 1985; Heikkila *et al.*, 1989; Richardson *et al.*, 1990), mas se demonstravam inconsistentes em relação aos dados observados (McMillen, 2001b). Campos e Azzoni (2020) sistematizaram as principais abordagens empíricas de identificação de SBDs embasados em métodos estatísticos objetivos. Para isso, os autores as dividiram em função das técnicas utilizadas, a saber: picos ou valor de corte, modelos econométricos, estatística espacial e fluxos.

Na abordagem de picos ou valor de corte, propõe-se um valor fixo para identificar áreas centrais, isto é, são escolhidas as zonas que têm densidade de emprego ou populacional acima de certo *cut-off* preestabelecido de forma *ad hoc*. No grupo de trabalhos que lidam com *modelos econométricos*, os autores estimam equações de densidade de emprego e/ou população, ou modelos de preço da terra, e consideram abordagens metodológicas de econometria paramétrica e/ou não paramétrica. Os trabalhos que aplicam *estatística espacial* geralmente utilizam técnicas de análise exploratória espacial aos dados, tal como análise de *clusters* que utiliza I de Moran.

A abordagem que considera os *fluxos* pode ser dividida em: i) *método de grafos*, que usa a técnica de grafos para avaliar as relações de deslocamento; ii) *método de saldo líquido*, que se baseia na avaliação do fluxo líquido de entrada e saída por motivo de trabalho em determinada região; iii) *cut-off de fluxo*, em que se identifica a zona central de negócios a partir da densidade de geração de viagens acima de determinado valor de corte; e iv) *atração de fluxo*, em que se utilizam dados sobre pendularidade casa-trabalho para identificar áreas centrais.

O método de valor de corte, proposto por McDonald (1987) e estendido por Giuliano e Small (1991), é a abordagem mais discutida e recorrente na literatura. McDonald (1987) identifica as subcentralidades pela concentração do emprego nas áreas adjacentes à analisada; e Giuliano e Small (1991) consideram como SBD um conjunto de áreas contíguas com densidade mínima de dez empregados por acre (4.040 m²) e, conjuntamente, dez mil empregados. Como destacado em Campos e Azzoni (2020), essa abordagem tem recebido muitas críticas devido à discricionariedade do valor do corte; à sensibilidade dos resultados ao corte escolhido; ao tamanho da área; e à dificuldade em generalizar o método. Uma vez que a abordagem está embasada em valor de corte *a priori*, o método demanda conhecimento prévio da área estudada.

Muitos trabalhos buscaram superar essas limitações utilizando técnicas econométricas. Inicialmente, estimava-se uma função de densidade de emprego ou da população e assumia-se uma cidade monocêntrica. Os resíduos positivos e significantes estatisticamente derivados dessa estimação identificam subcentralidades. A ausência de controles em um modelo de densidade de emprego ou população pode enviesar o erro estimado. Por consequência, os resíduos podem representar quaisquer outros fatores que não a concentração de emprego. Outro aspecto relevante é a sensibilidade dos termos de erro à unidade espacial tratada (McMillen, 2001b).

O espaço é considerado nos modelos de identificação de subcentralidades com a utilização de regressões ponderadas localmente – RPLs (McMillen, 2001a; 2001b; Redfearn, 2007). McMillen (2001a) utiliza os resíduos positivos e estatisticamente significativos para identificar SBDs; e McMillen (2003) adota abordagem mista e considera tanto valores de corte como RPLs. A estratégia consiste em estimar uma regressão do emprego espacialmente suavizada em dois estágios. No primeiro, consideram-se como candidatas a subcentralidades as áreas cujos resíduos sejam estatisticamente significantes. O segundo estágio identifica as áreas centrais, isto é, aquelas cujas vizinhanças somam, no mínimo, dez mil trabalhadores.⁴ Outros trabalhos que empregam econometria espacial restringem as avaliações à correlação espacial entre a taxa trabalhador-população e seus vizinhos. Esse segundo grupo

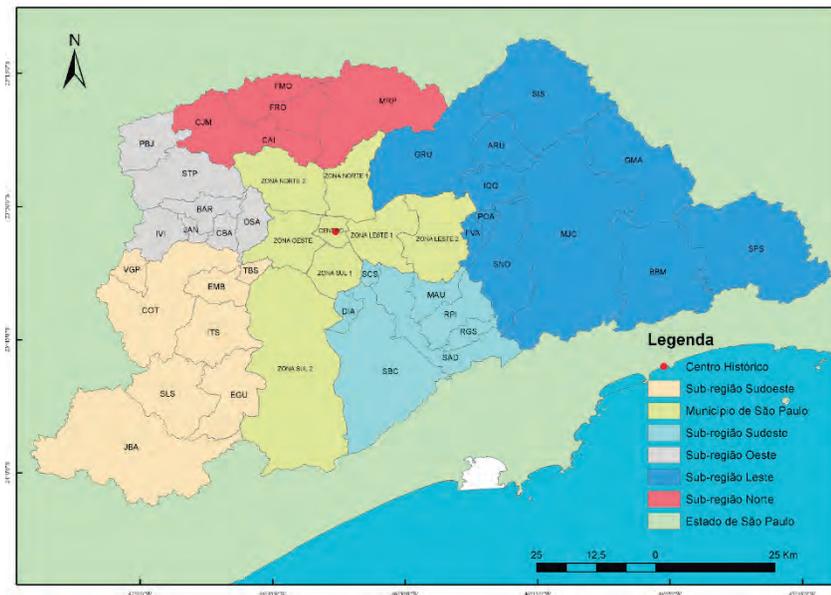
4. Para mais discussão sobre as metodologias e os artigos publicados na área, ver Campos e Azzoni (2020).

de trabalhos representa um esforço metodológico em aprimorar a identificação de SBDs. Todavia, a utilização de resíduos pode levar a resultados enviesados. Ademais, o uso de valores de corte apresenta um certo retrocesso em termos metodológicos devido à discricionariedade de sua escolha. Tais questões são pontos que precisam ser superados.

O objeto de análise neste estudo é a RM de São Paulo, que é composta por 39 municípios e ocupa área de 7.946 km². O mapa 1 expõe o mapa político dessa RM e de suas sub-regiões, definidas pela Lei Complementar Estadual nº 1.139/2011. O município de São Paulo é dividido em oito macrozonas, quais sejam, centro, zona leste 1, zona leste 2, zona norte 1, zona norte 2, zona oeste, zona sul 1 e zona sul 2, que são subdivididas em 96 distritos. A escolha dessa área deriva da sua importância na rede urbana brasileira, tanto em termos de nível quanto de crescimento de emprego. Ademais, a qualidade da informação sobre o arruamento possibilita melhor geocodificação dos dados.

MAPA 1

Mapa político-administrativo da RM de São Paulo e de suas sub-regiões



Elaboração dos autores.

Obs.: A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

O objetivo deste artigo é identificar regiões centrais de emprego na RM de São Paulo e apontar mudanças ocorridas entre 2002 e 2014, além de caracterizar as atividades presentes em tais áreas. Para alcançar esses propósitos, adota-se metodologia de identificação objetiva de subcentralidades e superam-se as abordagens

empíricas que demandam conhecimento prévio da região estudada. Além desta introdução, este artigo está dividido em quatro seções. Na seção 2, são discutidas as abordagens empíricas recorrentes na literatura. Ademais, sublinham-se as limitações e os avanços dos métodos, bem como as variáveis comumente utilizadas para identificar CBDs e SBDs. Na seção 3, verifica-se discussão sobre a base de dados e estatísticas descritivas. Adicionalmente, abordam-se a especificação econométrica, a identificação e o estimador utilizados neste trabalho. Na seção 4, debatem-se as principais localizações identificadas como SBDs de negócios, que são caracterizadas com base em características do mercado de trabalho formal da RM de São Paulo. Testes de sensibilidade são executados para avaliar o quanto as regiões identificadas são sensíveis às alterações do tamanho das células. Na seção 5, destacam-se as principais discussões desenvolvidas, bem como as limitações e as possíveis extensões da identificação de centralidades.

2 METODOLOGIA

Esta seção apresenta a metodologia adotada para a identificação de subcentralidades neste trabalho. Primeiro, é apresentada a base de dados e o tratamento dado para o georreferenciamento. Segundo, são demonstrados os ajustes feitos para a estimação do modelo econométrico. As subseções seguintes apresentam a estratégia empírica, os modelos econométricos e os métodos utilizados.

2.1 Base de dados

Foram utilizados microdados da Rais, do Ministério do Trabalho, que registra todos os trabalhadores formais e detalha as características de empregados e empregadores. Considera-se a RM de São Paulo, que abrangia 7.744.879 empregados e 442.388 empresas em 2014, o que representava 10,7% do total de empregos formais do Brasil. O município de São Paulo representava 68,5% do emprego dessa RM no mencionado ano. Foram consideradas apenas empresas do setor privado, já que o setor público tem características singulares tanto em relação à remuneração quanto à escolha de localização. Não estão incluídos também setores cujos trabalhadores têm elevada mobilidade espacial durante a jornada de trabalho (setor de transportes, limpeza urbana, construção civil e captação, tratamento e distribuição de água). Essa característica da atividade impõe viés à identificação de subcentralidades, pois agrega todos os funcionários na sede do estabelecimento, enquanto o desenvolvimento do trabalho ocorre em outras localidades da área intraurbana.

A base de microdados da Rais, por seu formato longitudinal, permite estudar questões relevantes do mercado de trabalho sob uma abordagem de dados em painel, o que tem sido recorrentemente explorado nos diversos trabalhos sobre o mercado de trabalho formal no Brasil. Um potencial distinto não explorado nessa base de dados é sua característica espacial intraurbana. Com base em informações de

endereçamento dos estabelecimentos, foi possível identificar latitudes e longitudes, além de explorar abrangência geográfica nunca utilizada com a base de dados da Rais em trabalhos acadêmicos. Padronizaram-se os endereços das empresas pelos arquivos cartográficos georreferenciados (*shapefiles*) disponibilizados pelo Centro de Estudos da Metrópole⁵ (CEM), e o georreferenciamento foi conduzido mediante ArcMap 10.5.1. Utilizaram-se *address locators* com o estilo brasileiro, e consideraram-se o *shapefile* de arruamento disponibilizado pelo CEM de 2016,⁶ bem como, de forma complementar, o arruamento da base de dados do GeoNames de 2017.⁷

2.2 Modelo econométrico, método de estimação e estratégia de identificação

O procedimento-padrão para modelar cidades policêntricas tem sido embasado na abordagem de econometria clássica. Todavia, uma forma funcional incorretamente determinada tende a enviesar a inferência da regressão paramétrica (McMillen e McDonald, 1997). Ademais, estimações paramétricas têm baixo poder em áreas metropolitanas complexas, devido aos erros de especificação, quando os candidatos a SBDs são espacialmente correlacionados ou heterogêneos (Giuliano *et al.*, 2005). *Locally weighted regressions* (LWRs) têm sido recorrentemente utilizadas na identificação de subcentralidades. Tais estimações não paramétricas apresentam-se como alternativas ao procedimento clássico, oferecendo vantagens significativas sobre o método de regressão linear simples. Estimadores não paramétricos são flexíveis e permitem que o coeficiente estimado varie entre as áreas, bem como fornecem maior acurácia em relação às centralidades principais (SBDs na área central da cidade) e às centralidades periféricas quando comparados a outras abordagens, além de permitirem visualização em mapas das centralidades estimadas (McMillen e McDonald, 1997; McMillen, 2001b; Redfean, 2007).

Modelos não paramétricos fazem inferência sobre a superfície a partir da regressão dos dados, suavizam as observações e mantêm a complexidade necessária para fornecer estimativas imparciais. Ao lidar com modelos não paramétricos, algumas decisões arbitrárias são necessárias, como o tamanho da banda e o *kernel* ao estimar as regressões ponderadas localmente. O número de subcentralidades identificadas pelo procedimento não paramétrico e o nível de significância são influenciados pelo tamanho da janela. Um tamanho menor de banda leva a um menor alisamento da superfície da densidade do emprego, o que tende a reduzir a magnitude dos resíduos ao aumentar a variância estimada, e o efeito líquido dessas tendências opostas não é claro (McMillen, 2003).

5. Trata-se de instituição de pesquisa sediada na USP e no Centro Brasileiro de Análise e Planejamento (Cebap). O CEM investiga temáticas relacionadas às desigualdades e à formulação de políticas públicas nas metrópoles brasileiras contemporâneas. A instituição é financiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) e conta com grupo multidisciplinar de pesquisadores das áreas de demografia, ciências políticas, sociologia, geografia, economia e antropologia.

6. Disponível sob requerimento.

7. GeoNames é uma base de dados fornecida pelo ArcGIS, que compila o arruamento para todo o globo terrestre. A base está disponível no serviço *on-line* de georreferenciamento do ArcGIS, e sua utilização é taxada em função da quantidade de elementos georreferenciados.

A escolha por um estimador não paramétrico deriva de sua vantagem em controlar os padrões globais de concentração de emprego, pois sua distribuição no espaço não é isotrópica, o que demanda abordagem que considere essa heterogeneidade espacial. Regressões ponderadas localmente lidam com as referidas questões, usando critérios estatísticos-padrão para determinar o nível de emprego cuja concentração varia espacialmente dentro da cidade e em toda a área metropolitana. Em suma, a abordagem fornece desagregação local das estatísticas globais, lida com não estacionariedade espacial e permite avaliações espaciais dos dados (Fotheringham, Brunson e Charlton, 2002; Bivand, Pebesma e Gómez-Rubio, 2008). Assim, a identificação de subcentralidades está embasada na estimação da média espacial da concentração do emprego por uma abordagem não paramétrica, seguindo McMillen (2001a; 2003) e Kane, Hipp e Kim (2016), como discutido na subseção 2.3.

2.3 Especificação econométrica e métodos de estimação

Para tornar tratável a estimação de regressões ponderadas geograficamente (*geographically weighted regressions* – GWRs), optou-se por agregar os trabalhadores por células do *grid*. Foram criadas i células com 1 km² de área, em que e_i indica a soma dos trabalhadores de cada célula. Nesse sentido, associa-se o centroide da célula i a um total de trabalhadores, reduzindo significativamente o total de coeficientes locais a serem estimados. Evitam-se assim regiões com áreas distintas (áreas de ponderação do Censo Demográfico, áreas de mercado de trabalho etc.) e problemas decorrentes de densidade bruta e líquida do emprego (McDonald, 1987; McMillen, 2001a).⁸ O mapa 2 apresenta a formatação do *grid*. Devido à irregularidade da área oficial da RM de São Paulo, as células possuem área menor que 1 km² nas franjas da região. Essa diferença é pouco relevante, uma vez que pouco emprego é observado nessas partes da RM.

A estimativa do emprego médio espacial em um contexto de regressão local é dada por:

$$e_i = \beta_0(u_i, v_i) + \varepsilon_i \quad (1)$$

Na equação (1), e_i é o total de emprego na célula i ; o vetor (u_i, v_i) denota as coordenadas geográficas (latitude e longitude) dos i -ésimos centroides associados às células; $\beta_0(u_i, v_i)$ é a média espacial no ponto i , que usa métodos embasados em *kernels*; e ε_i é o termo de erro aleatório.

A estimativa GWRs para o parâmetro $\beta_0(u_i, v_i)$ da equação (1) é dada por:

$$\hat{\beta}_0(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) e \quad (2)$$

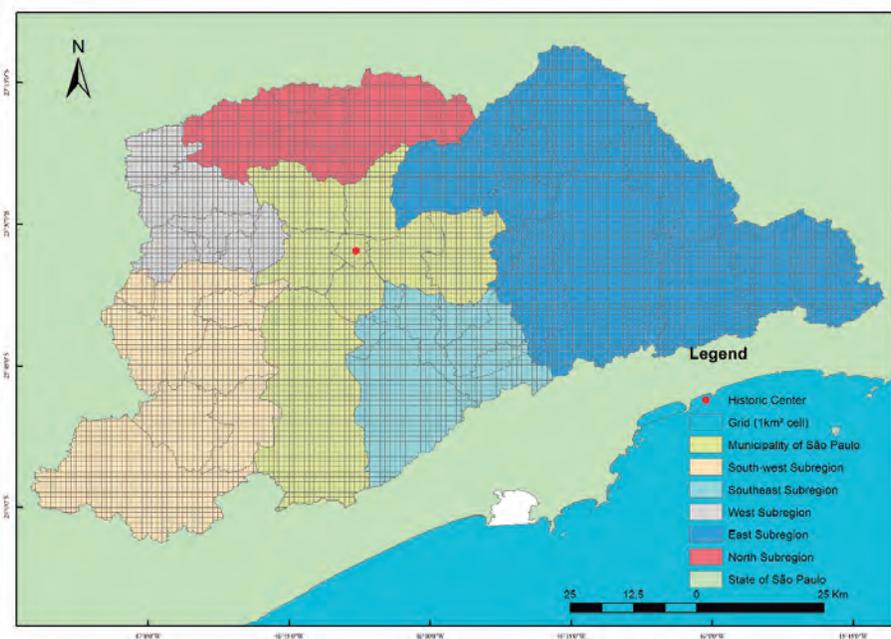
8. Mieszkowski e Smith (1991) indicam que o uso da densidade bruta tende a superestimar os coeficientes estimados, uma vez que áreas mais afastadas do centro histórico tendem a concentrar maiores extensões de terra não desenvolvidas. Redfearn (2007) destaca que dados agregados em áreas grandes mascaram as concentrações de empregos locais independentes.

Na equação (2), $\hat{\beta}_0$ representa a média estimada de β_0 ; $W(u_i, v_i)$ denota uma matriz de pesos $n \times n$, na qual os elementos da diagonal principal são pesos geográficos de cada uma das n observações contidas na regressão para o ponto i (ou célula i); e os elementos fora da diagonal principal são nulos. Já X é um vetor $n \times 1$ com valores unitários; e o sobrescrito T denota que o vetor está na transversal.

$$W(i) = \begin{bmatrix} w_{i1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & w_{in} \end{bmatrix} \tag{3}$$

Nesse sentido, $W(u_i, v_i)$ é a matriz de pesos que conecta as observações para cada subamostra local e tem função essencial na estimação de modelos do tipo GWRs. Tais conexões são especificadas por funções de pesos, em que observações mais próximas da célula i são ponderadas com maior peso em comparação a observações mais longínquas.

MAPA 2
Grid da RM de São Paulo



Elaboração dos autores.
Obs.: 1. Células com 1 km² de área.
2. A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

A GWR busca obter um estimador não paramétrico para funções contínuas de localização (u_i, v_i) mediante *kernels*. A *log-likelihood* para cada conjunto particular de estimativas não fornece solução única. A maneira de ajustar a otimização é considerar *log-likelihoods* locais e assumir observações próximas a i (Bowman e Azzalini,

1997; Fotheringham, Brunson e Charlton, 2002). A estimação das regressões envolve a seleção de bandas (ou janelas) para uma função de peso *kernel* isotrópica espacialmente. Entre as diferentes funções de ponderação que podem ser usadas para modelar o processo de dependência espacial, estão as funções gaussiana, tricúbica e quadrática. As funções de peso podem ser fixas ou adaptativas. O primeiro método considera o tamanho da banda constante no espaço, não importando o tamanho da amostra ao redor de i , enquanto o segundo adapta o tamanho da banda de acordo com o tamanho da amostra. Problemas potenciais podem derivar de funções de peso fixas, tais como calibração com poucos dados, o que dá origem a estimativas de parâmetros com elevados erros-padrão e, por conseguinte, gera superfícies pouco suaves. Em situações extremas, a estimação dos parâmetros pode ser comprometida devido à variação insuficiente em pequenas amostras (Fotheringham, Brunson e Charlton, 2002). A segunda abordagem busca superar tal limitação ao ajustar o tamanho das janelas e considerar a variação de densidade da amostra.

No caso deste trabalho, bandas calculadas pelo método adaptativo comprometem os resultados da identificação. Como regiões periféricas tendem a ter pouco emprego, ajustes que consideram o tamanho da amostra resultariam em uma banda excessivamente grande nessas regiões e menor no centro. Nesse sentido, apenas as bandas fixas serão exploradas. Por exemplo, regiões pouco representativas tenderão a não rejeitar a hipótese nula de média espacial diferente de zero devido ao elevado erro-padrão do coeficiente estimado, ou, no caso extremo, não será possível estimar a média espacial. De acordo com Fotheringham, Brunson e Charlton (2002), os resultados derivados de regressões locais são sensíveis à escolha do tamanho da banda. A escolha ótima do tamanho da banda envolve um *trade-off* entre viés e variância: bandas pequenas geram elevada variância, enquanto bandas grandes enviesam os resultados de regressões locais. Buscando superar escolhas arbitrárias,⁹ empregou-se o critério de informação *Akaike information criterion* (AIC) para a seleção das bandas.¹⁰ Os resultados da otimização AIC são sensíveis à função *kernel* escolhida, por isso serão utilizadas diferentes formas funcionais de *kernel*, tais como gaussiana, biquadrada e tricúbica, cuja escolha estará sujeita ao critério de informação AIC.

De forma prática, o modelo econométrico calcula para cada célula do *grid* médias espaciais do emprego total, bem como considera funções *kernel* e distâncias de bandas definidas por critérios de otimização. Serão consideradas como subcentralidades as

9. Em McMillen (2001a), Redfearn (2007) e Kane, Hipp e Kim (2016), o tamanho das bandas é arbitrariamente escolhido. O primeiro autor escolhe 50% da amostra, enquanto Redfearn (2007) escolhe 1% das observações mais próximas de cada uma das n regiões avaliadas. Kane, Hipp e Kim (2016) escolheram o valor que seleciona 120 *grids* como vizinhos.

10. Os métodos *cross validation* (CV) e *generalized cross validation* (GCV) não foram utilizados devido às limitações apresentadas, tais como elevada sensibilidade a pequenas amostras, *outliers* e pontos de influência. Ademais, o problema de otimização dos métodos CV e GCV apenas fornece mínimo global, e nenhuma informação relacionada a estimativas locais (Faber e Páez, 2007).

áreas cujas médias espaciais de emprego ($\hat{\beta}_0^{**}$) sejam estatisticamente significantes e maiores ou iguais ao valor crítico c . Buscando diminuir o papel da discricionariedade do valor crítico, este será igual ao percentil χ da distribuição de $\hat{\beta}_0$. Para o exercício, utilizam-se os percentis 99, 95 e 90. Embora tal corte continue ainda a ser discricionário, sua vantagem está em não exigir conhecimento prévio da região e permitir adequação temporal desse corte em função da própria distribuição das médias espaciais estimadas e estatisticamente significativas. Matematicamente, representa-se da seguinte forma:

$$SBD_i = \hat{\beta}_0^{**} \geq c \quad (4)$$

A regra de identificação da equação (4) dispensa a utilização do valor de corte predefinido e estático intertemporalmente. Tal flexibilidade decorre do fato de as médias espaciais estimadas serem representativas da concentração do emprego local. Logo, dispensa a tentativa de ajuste por um valor de corte sobre o total de emprego, como ocorre em estimações derivadas de densidade de emprego (Guiliano e Small, 1991; McMillen, 2003; Kane, Hipp e Kim, 2016). Outra vantagem advém da escolha endógena do valor de corte para a identificação, isto é, o *cut-off* deriva da própria distribuição das médias estimadas, e cabe ao pesquisador a escolha do nível de significância para a inferência estatística. Essa endogeneização da escolha do valor de corte permite comparabilidade temporal dos resultados, mesmo com a variação em nível do total de emprego por célula e generalização da abordagem para qualquer região.

Diferentes resultados podem derivar de diferentes agregações espaciais dos dados, mesmo quando se utilizam as mesmas bases de dados e abordagem empírica (Giuliano e Small, 1991; McMillen, 2001a; Redfearn, 2007; Kane, Hipp e Kim, 2016). Tal questão é conhecida na literatura como problema de unidade de área modificável (*modifiable area unit problem* – Maup). O Maup surge devido às múltiplas variedades de agregação dos dados e à disponibilização destes por órgãos competentes, o que muitas vezes limita a pesquisa sobre agregação prévia dos dados. Segundo Wrigley *et al.* (1996), efeitos de escala e de zoneamento compõem o Maup. O efeito de escala está relacionado à tendência de se obterem resultados estatísticos distintos ao considerar o mesmo conjunto de dados quando a agregação geográfica é alterada. O efeito de zoneamento diz respeito à variabilidade dos resultados estatísticos em função de uma nova agregação derivada de zoneamentos para a coleta dos dados (por exemplo, alteração do número de setores censitários para a mesma região em diferentes períodos de coleta). Nesse sentido, neste trabalho as médias espaciais da concentração do emprego são também calculadas para *grids* com células de 2 e 4 km² de área, para fins de averiguar a sensibilidade dos resultados.

O acesso a dados sobre o total de área construída e utilizada pelas firmas – por unidade espacial – não é uma informação trivial de se obter, e estimar modelos econométricos em áreas muito diferentes, com base no nível de emprego, também

comprometeria os resultados. Dado o georeferenciamento realizado neste estudo, é possível criar áreas tão pequenas e homogêneas quanto o pesquisador requeira, com tamanhos idênticos ao longo da área metropolitana,¹¹ o que supera o viés potencial da normalização.¹² Em outras palavras, tal medida é uma maneira de superar as limitações decorrentes da normalização do total de emprego ou da ausência de normalização. Além de trabalhar com medidas mais flexíveis, a variável dependente não será sensível a diferentes valores no denominador a ponto de reverter os resultados.

3 RESULTADOS

3.1 Características da amostra

A tabela 1 traz informações sobre os totais de emprego; número de firmas; massa salarial e empregados por faixas de experiência; tipo de salário; grau de instrução; salário mínimo; e gênero para anos selecionados. Nota-se que houve espraiamento das firmas na RM de São Paulo. Em 2002, 43% das 9.071 células de 1 km² tinham pelo menos um estabelecimento formal com vínculo ativo, tendo aumentado para 48% em 2014. A área com emprego formal expandiu-se em 12% entre esses anos, o que representa 482 km² em termos absolutos. A média de emprego por célula em 2002 era de 340 vínculos, passando de 496 em 2008 para 589 em 2014. O máximo de empregos em uma célula era de 54.607 em 2002, elevando-se a 87.707 e 89.476 em 2008 e 2014, respectivamente. Em termos do número médio de firmas por célula, verifica-se também este crescimento: 26 em 2002; e 32 e 38 em 2008 e 2014, respectivamente. Em 2002, 4.272 firmas formais estavam concentradas em apenas uma célula do *grid*. Já em 2008 e 2014, os valores máximos aumentam para 4.450 e 4.919, respectivamente. Os dados apontam a presença de desigualdade salarial no espaço intraurbano, que pode derivar tanto da elevada concentração de trabalhadores quanto da não aleatoriedade espacial das firmas e dos trabalhadores mais produtivos, isto é, *sorting*. O desvio-padrão da massa salarial em 2002 era de R\$ 2.460.355, o que foi ampliado para R\$ 5.397.076 em 2008 e R\$ 10.092.028 em 2014.¹³

11. Kane, Hipp and Kim (2016) utilizaram a estratégia de *grids*, todavia a microlocalização se deu por zonas de códigos de endereçamento postal (CEPs). Isso pode gerar erros de medida, pois as zonas de CEPs divergem em extensão a depender da região em questão. Ademais, o georeferenciamento feito sob essa estratégia agrega todas as firmas no centroide da zona, gerando pesos artificiais em determinadas zonas de CEPs, em termos de concentração de emprego. Nesse sentido, o georeferenciamento que considera o endereço completo das firmas é mais adequado.

12. É importante destacar que o processo de criar *grids* idênticos obedeceu aos limites geográficos preestabelecidos pelo IBGE. Isto é, nas periferias municipais e distritais, os *grids* têm área menores do que 1 km². Todavia, tal questão não prejudica os resultados; pelo contrário, impõe restrição extra – em termos de área – para considerar uma subcentralidade.

13. Desconta-se a inflação e traz-se a valor de 2014 com o Índice de Preços ao Consumidor da Fundação Instituto de Pesquisas Econômicas (IPC/Fipe).

TABELA 1
Estatísticas descritivas

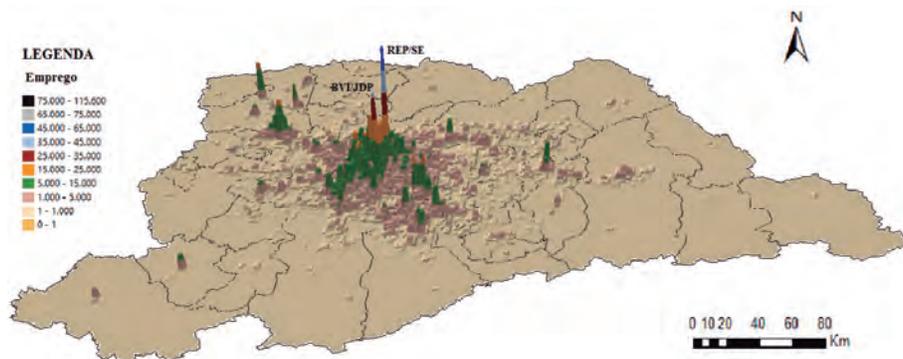
Variáveis	2002			2008			2014		
	Média	Desvio	Máximo	Média	Desvio	Máximo	Média	Desvio	Máximo
Emprego	340	1.532	54.607	496	2.291	87.707	589	2.672	89.476
Total de firmas	26	117	4.272	32	136	4.450	38	152	4.919
Massa salarial (R\$)	408.171	2.460.355	88.542.751	876.598	5.397.076	176.594.369	1.640.874	10.092.028	324.799.763
Número de células ocupadas		3.901			4.212			4.383	
Número total de células		9.071			9.071			9.071	

Elaboração dos autores.

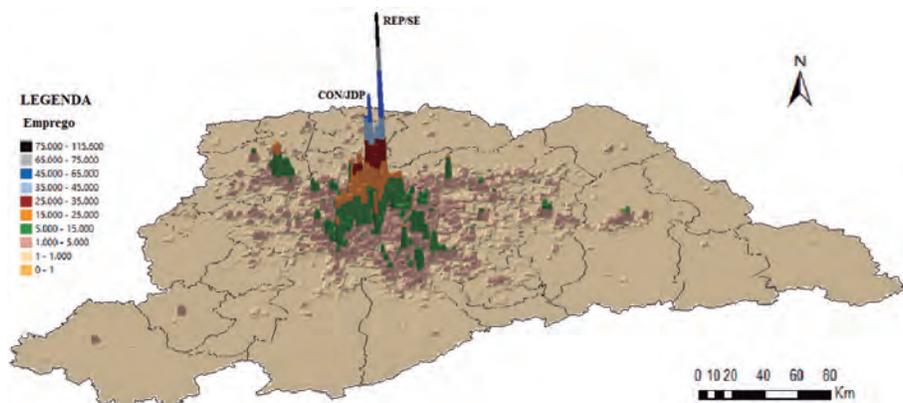
A localização desses empregos pode ser observada nos mapas 3 a 5, que revelam a elevada concentração do emprego no centro expandido do município de São Paulo. Ao longo do período avaliado, mesmo em face do aumento na área com emprego, percebe-se que houve mais aglomeração do que espraiamento. As novas localizações de emprego surgem na proximidade de firmas preexistentes, e a taxa de crescimento da área ocupada por emprego formal é inferior à taxa de crescimento do emprego na região como um todo. Destaca-se que os empresários apresentam preferência pela aglomeração em lugares específicos. A maior concentração de empregos formais está entre os limites dos distritos da República e da Sé, bem como dos distritos de Jardim Paulista e Bela Vista, mais precisamente na avenida Paulista. As zonas central e sul 1 do município de São Paulo apresentaram forte crescimento de empregos formais e se expandem em direção ao sul da cidade. Em 2014, a zona central do município continuava concentrando emprego, com destaque para o distrito do Itaim Bibi, que se tornou tão relevante quanto a concentração Jardim Paulista/Bela Vista. Pode-se notar a formação de uma cordilheira intraurbana de emprego que passa pelos distritos Bresser, Santa Cecília, República, Sé, Consolação, Jardim Paulista, Bela Vista, Liberdade, Pinheiros, Itaim Bibi, Santo Amaro.

Áreas relevantes em termos de emprego também podem ser observadas fora da capital. Na sub-região leste, destaca-se a aglomeração formada no município de Guarulhos, que apresenta concentrações de empregos formais tão elevadas quanto determinadas regiões do município de São Paulo, o que forma as cordilheiras intraurbanas de empregos. Na sub-região sudeste, destacam-se as concentrações formadas na região do ABCD (Santo André, São Bernardo, São Caetano do Sul e Diadema). Entre esses locais, é importante destacar a maior representatividade do município de São Caetano do Sul, que, desde 2002, apresenta aglomeração entre 15.001 e 35.000 empregos.

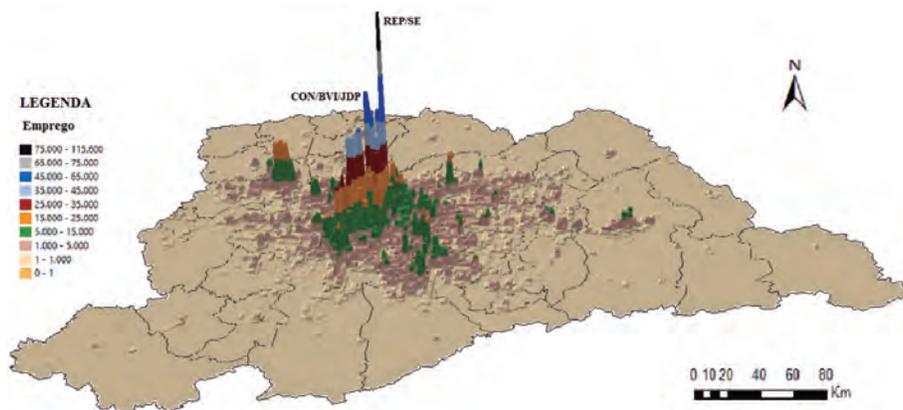
MAPA 3
Distribuição espacial de emprego estimado
3A – 2002



3B – 2008



3C – 2014



Elaboração dos autores.

Obs.: A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

3.2 Identificação dos subcentros

A forma objetiva de selecionar o tamanho ótimo da banda, derivado da otimização da função de máxima log-verossimilhança, é comparar critérios de informação. A tabela 2 apresenta os critérios de informação para as funções *kernel* gaussiana, biquadrada e tricúbica, e esta última apresenta o menor AIC para todos os anos, o que repete os resultados da literatura (McMillen e Smith, 2003). As bandas que minimizam o critério de informação nos três anos são de 2,3 km, 2,09 km e 2,06 km, respectivamente. Nota-se que os tamanhos das bandas se reduzem em função do espraiamento do emprego na RM de São Paulo.

TABELA 2
Critério de informação AIC e tamanho das bandas

Função <i>kernel</i>	2002		2008		2014	
	AIC	Banda (km)	AIC	Banda (km)	AIC	Banda (km)
Gaussiana	150.453,9	0,910	157.403,6	0,879	159.668,7	0,871
Biquadrada	150.289,6	2,271	157.169,2	2,157	159.429,8	2,129
Tricúbica	150.197,6	2,309	157.033,3	2,095	159.279,6	2,063

Elaboração dos autores.

A tabela 3 apresenta os resultados das estimações por quartil, tendo-se também estimado a média do emprego global por *ordinary least square* – OLS (global). No primeiro quartil, os coeficientes estimados são iguais a zero, e as medianas são bastante baixas, o que decorre da utilização de todo o espaço da RM de São Paulo, tais como áreas verdes, rios, áreas tipicamente residenciais etc.¹⁴ Esse resultado indica que, em 50% da área da região, a densidade de empregos é muito baixa, de modo que há grande homogeneidade em relação a essa variável. No terceiro quartil, a média estimada é de 97 empregos para 2002 e se eleva ao longo dos anos, alcançando 175 em 2014. A heterogeneidade do valor máximo de emprego estimado é patente, mesmo com a suavização espacial das funções *kernel*. A estimativa por OLS desconsidera a não isotropia espacial dos empregos, superestima as médias (340, 496 e 589 empregos) e, de certa forma, esconde a estrutura espacial heterogênea dos resultados.

14. Utilizar toda a área da RM incorpora maior realismo à distribuição espacial do emprego, a qual é limitada a fatores geográficos e políticas públicas. Tais fatores, por consequência, influenciam a própria aglomeração de emprego e seu espraiamento. Portanto, considerar apenas áreas de trabalho com ao menos um trabalhador acarretaria resultados enviesados, sobretudo em áreas de municípios pequenos, que tendem a apresentar apenas uma única concentração de emprego e arredores vazios.

TABELA 3
GWRs (tricúbica)

	2002	2008	2014
Mínimo	0	0	0
Primeiro quartil	0	0	0
Mediana	1,4	2,4	3,5
Terceiro quartil	97	147	175
Máximo	21.884	34.297	41.131
Global (OLS)	340	496	589
AIC	150.198	157.033	159.280
Quasi-global R ²	0,74	0,76	0,78
Número de observações	9.071	9.071	9.071

Elaboração dos autores.

Obs.: Os quartis se referem à distribuição dos coeficientes estimados por GWRs.

Diferentemente dos trabalhos que usam valores de corte *ad hoc* e fixos no tempo, a identificação dos SBDs deriva da distribuição da amostra de cada ano, considerando os limites da distribuição dos percentis 99, 95 e 90, e a significância estatística correspondente. Em 2002, 1% das células têm número de empregos superior ou igual a 5.764; para 5% e 10% das células, os números são de 1.876 e 819 empregos. Esses valores de corte aumentam substancialmente nos outros anos, passando a 10.070 (percentil 99), 3.132 (percentil 95) e 1.380 (percentil 90) em 2014. Identificam-se apenas 91 células como SBDs ao utilizar o primeiro corte, independentemente do ano; 363 novas células no segundo corte; e 454 novas áreas no terceiro corte (tabela 4). Alterações no grau de significância estatística não mudam os números de células identificadas, o que indica a robustez dos resultados. Como houve crescimento nas médias de emprego por célula e como o número de células não se altera, necessariamente os valores de corte aumentam.

TABELA 4
Valores de corte e número de SBDs identificados

Ano	Percentis	Valores de corte	Número de SBDs		
			10%*	5%*	1%*
2002	99	5.764	91	91	91
	95	1.876	454	454	454
	90	819	908	908	908
2008	99	8.484	91	91	91
	95	2.654	454	454	454
	90	1.161	908	908	908
2014	99	10.070	91	91	91
	95	3.132	454	454	454
	90	1.380	908	908	908

Elaboração dos autores.

Obs.: * nível de significância estatística considerado na identificação das células.

Embora o percentual da área total com empregos não se altere, a localização das células não é necessariamente a mesma. O comportamento dinâmico do mercado de trabalho envolve fechamento e abertura de firmas, expansão e retração da economia etc., e isso faz com que SBDs possam desaparecer por perda de participação relativa, enquanto outros podem emergir. O mapa 4 traz as disposições espaciais das subcentralidades por ordem do percentil, mostrando apenas as áreas estatisticamente significantes. Esses mapas ilustram os centros que emergiram, morreram, tiveram suas áreas aumentadas e/ou reduzidas, o que contrasta com as avaliações mais agregadas e recorrentes na literatura.

Passa-se agora da identificação de células relevantes individualmente para a caracterização dos SBDs de emprego, definidos com um conjunto de células que atendem aos critérios de identificação utilizados. Para facilitar a análise, identificam-se os SBDs pelo município ou pelo conjunto de municípios em que se localiza o *cluster*.

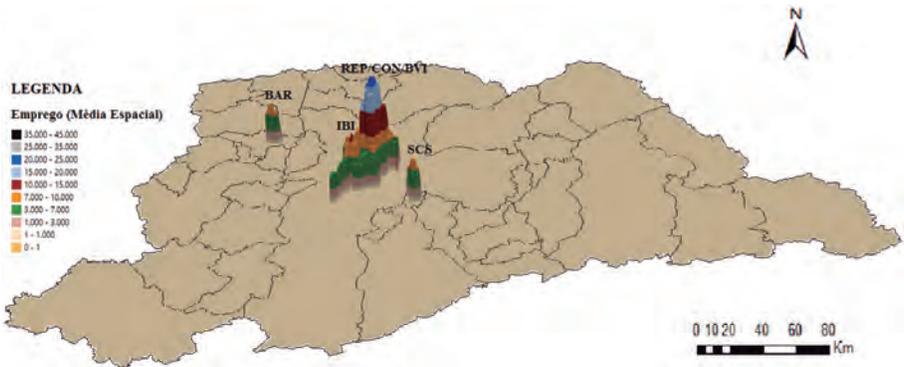
O mapa 4 demonstra as posições geográficas das médias espaciais de emprego para o ano-base 2002. Usando-se o limite mais restritivo para aceitação (percentil 99), em 2002 identificam-se como SBDs as seguintes áreas: Barueri, São Caetano do Sul e São Paulo. O SBD de Barueri ocupa 5 km², e a sua célula, com máxima média estimada pelo modelo, apresenta entre 7.000 e 10.000 empregos. O SBD de São Caetano do Sul abrange 7 km², com ao menos uma célula que apresenta média espacial entre 7.000 e 10.000 empregos. Já o SBD de São Paulo ocupa 79 km², desde a zona central, passando por parte das zonas oeste, sul 1 e sul 2. Nessa área estão localizados dois dos picos mais relevantes em termos de quantidade de emprego. O primeiro e mais importante apresenta entre 20.000 e 35.000 empregos, localizando-se entre os distritos da República, da Consolação e de Bela Vista. O segundo pico dessa distribuição concentra entre 10.000 e 15.000 empregos no distrito do Itaim Bibi. Em termos de concentração de emprego, a região do Itaim Bibi é mais aglomerada em comparação com os SBDs de Barueri e São Caetano do Sul.

Relaxando-se o valor de corte (percentil 95), a mancha de empregos espalha-se substancialmente ao redor dos três SBDs já identificados. Em alguns casos, forma-se uma única grande mancha uniforme, por exemplo, São Caetano do Sul, que se estende em direção aos municípios de Santo André e de São Paulo. Subcentralidades não conectadas às regiões pré-identificadas surgem na sub-região leste da RM de São Paulo, nos municípios de Guarulhos, Poá e Moji das Cruzes, além de São Bernardo do Campo e Diadema, na sub-região sudeste. Na sub-região oeste, dois SBDs passam a ser observados nos municípios de Santana do Parnaíba e Osasco e um na sub-região norte (Cajamar). Os municípios de Taboão da Serra e Embu das Artes, na sub-região sudoeste, também apresentam concentrações relevantes de médias de empregos formais. Todavia, essas regiões não são homogêneas em termos de concentração de emprego. Os SBDs de Santo André,

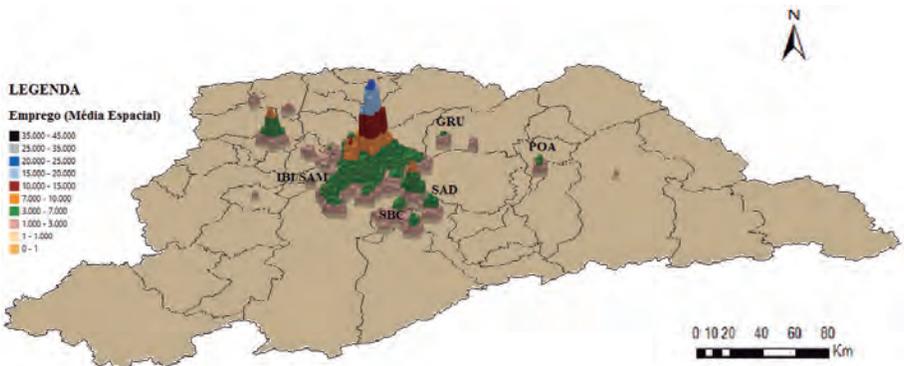
São Bernardo do Campo, Guarulhos e Poá apresentam maior relevância em termos de aglomeração de emprego, com pelo menos uma das células entre 3.000 e 7.000 empregos formais, enquanto Moji das Cruzes, Santana do Parnaíba, Cajamar, Osasco, Taboão da Serra e Embu das Artes situam-se no intervalo entre 1.000 e 3.000 empregos. O SBD de São Paulo cresce em direção às áreas periféricas do município de São Paulo e ocupa 240 km² adicionais.

Afrouxando ainda mais o critério para o percentil 90, os SBDs se espriam ainda mais em direção às franjas da RM de São Paulo. Percebe-se a formação de uma única mancha de SBD que une municípios da região de São Paulo, ABCD, Guarulhos, Osasco, Taboão da Serra e Embu das Artes. A área de Barueri é ampliada, e a nova configuração tende a diminuir a distância entre os SBDs de Barueri e a grande mancha formada. Cabe destacar, na capital, o aparecimento de um SBD separado da grande mancha, no distrito de Cidade Dutra, na zona sul 1, e outro nos limites políticos dos distritos de Itaquera e Parque do Carmo, na zona leste 2.

MAPA 4
SBDs identificados (2002)
4A – Percentil 99



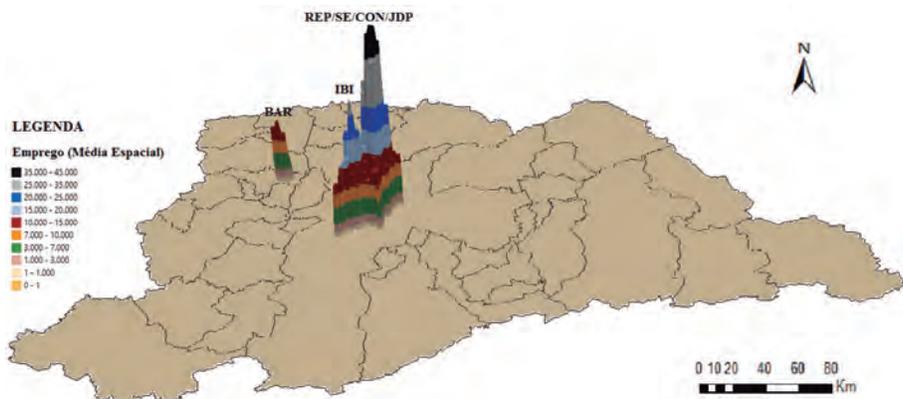
4B – Percentil 95



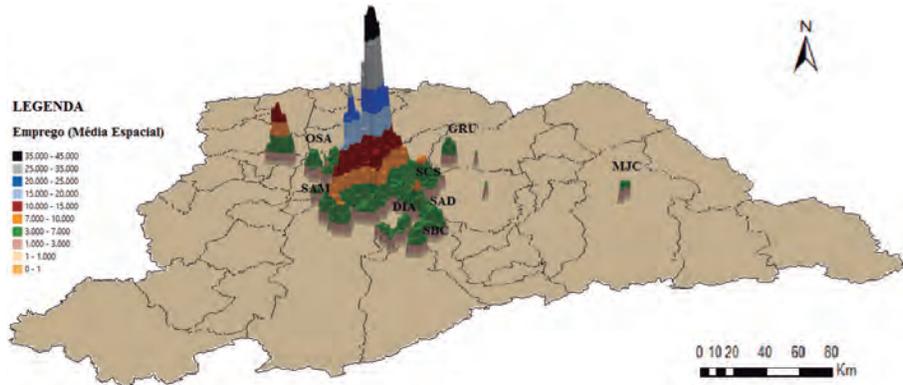
35.000 empregos). Ao relaxar o valor de corte para o percentil 95, destacam-se os SBDs de Taboão da Serra, Osasco, Guarulhos I e Guarulhos II, Moji das Cruzes, São Bernardo do Campo e Diadema. São Caetano do Sul passa a ter valor de corte menor. Outra mudança relevante é o aparecimento de subcentralidade em Itaquera. Seguindo-se o padrão do ano anterior, com valor de corte menos restritivo, obtém-se um SBD muito mais extenso.

O padrão encontrado nesses resultados é o relaxamento do valor de corte que resulta no aparecimento de novas centralidades derivadas ao redor das áreas com maior força de aglomeração de trabalhadores. Poucos são os SBDs que surgem em áreas distantes ou isoladas, e poucos são identificados por cortes mais restritivos. Esse resultado deriva claramente da abordagem empírica utilizada, em que a concentração de emprego dos vizinhos eleva ou reduz a aglomeração estimada de empregos, mesmo em face de funções *kernel* que ajustam o distanciamento da célula às suas vizinhas. Considerando-se a elevação dos valores de corte ao longo do tempo e o limitado espraiamento para novas áreas, é possível caracterizar o caminho espacial do emprego na RM de São Paulo por dois movimentos simultâneos: consolidação do adensamento e descentralização concentrada do emprego. Claramente, o efeito do adensamento tem domínio sobre o efeito da descentralização, tanto na capital como nos demais municípios. Por esse motivo, considera-se como uma desconcentração concentrada. Como discutido, a descentralização (novas áreas para aninhar o emprego) não altera a característica concentradora de empregos das centralidades, tendo pouco impacto na RM de São Paulo como um todo. Entretanto, em avaliação mais focalizada e restrita ao município de São Paulo, podem-se encontrar indícios de descentralização. Isto é, empregos formais se direcionam às áreas menos aglomeradas, todavia não distantes das áreas com concentração expressiva, salvo a concentração em Itaquera, e isso apenas com valor de corte mais flexível.

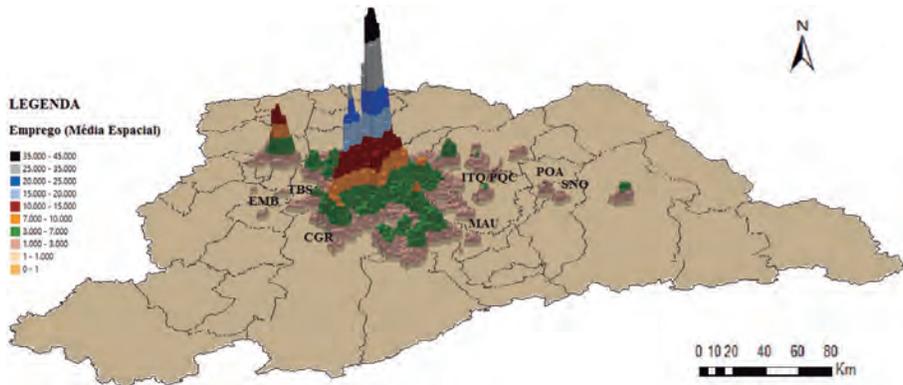
MAPA 5
SBDs identificados (2014)
5A – Percentil 99



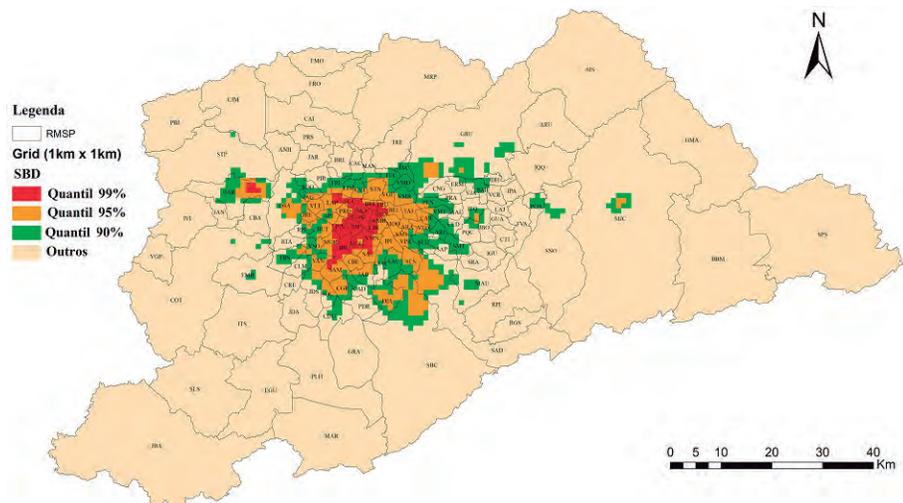
5B – Percentil 95



5C – Percentil 90



5D – Todos os percentis



Elaboração dos autores.

Obs.: A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

3.3 Caracterização dos SBDs

Nesta subseção exploram-se algumas características dos SBDs identificados, o que se restringe ao percentil de ordem 99. A tabela 5 apresenta as estatísticas descritivas para total de empregos, firmas e massa salarial concentrados em cada SBD. Há um total de 934.834 empregos – distribuídos entre 71.544 firmas –, em área de 91 km², no ano de 2002. Em 2008, o número total de empregos se eleva em 54,7% (1.446.631 postos de trabalho), alocados em 83.519 firmas (aumento percentual de 16,7% em relação a 2002). No período seguinte, a taxa de crescimento do emprego acumulado foi de 16,4% (1.684.185 postos de trabalho).

TABELA 5
Empregos, firmas, massa salarial e área

Estatísticas	Emprego			Número de firmas			Massa salarial (R\$ 1 mil) ¹		
	2002	2008	2014	2002	2008	2014	2002	2008	2014
Total	934.843	1.446.631	1.684.185	71.544	83.519	89.962	1.372.410	3.169.201	6.179.594
Média	35.956	53.579	67.367	2.752	3.093	3.598	52.785	117.378	247.184
Desvio-padrão	32.526	51.969	64.781	2.532	3.015	3.462	61.736	143.389	294.543
Mínimo	3.467	4.551	6.127	181	256	268	3.036	5.406	12.418
Máximo	139.220	228.516	252.890	9.991	11.690	12.985	242.061	534.020	1.017.398

Elaboração dos autores.

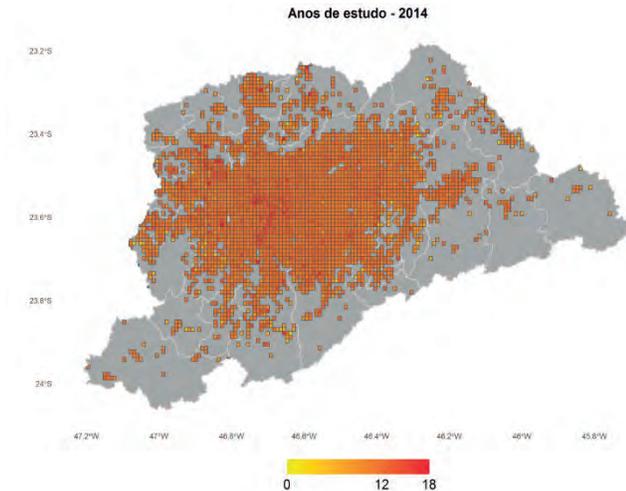
Nota: ¹ Valores em preços correntes.

As figuras a seguir permitem comparar aspectos relevantes da composição dos SBDs. Apresenta-se a distribuição espacial dos valores e os *clusters* de células com valores similares de cada variável, que podem ser combinações de valores alto-alto e baixo-baixo. Essa classificação resulta de análise exploratória de dados espaciais, em que se procura avaliar se os valores observados em uma célula são similares (ou dissimilares) aos valores observados em células vizinhas.¹⁵ Como consequência, são identificadas eventuais aglomerações de células com valores similares.

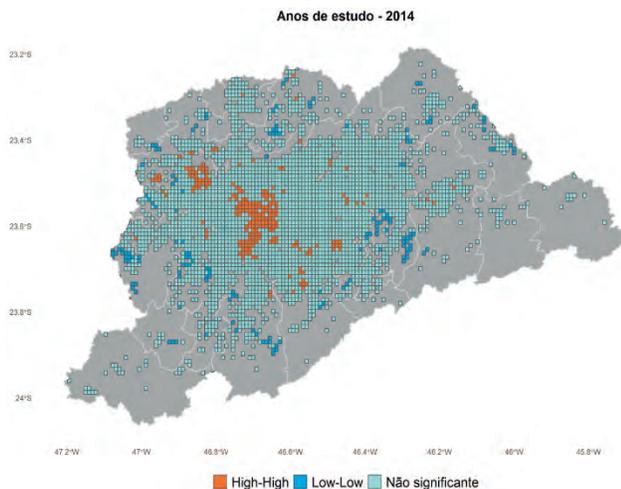
A figura 1 apresenta os níveis médios de anos de estudos. Pode-se observar concentração de valores alto-alto nos SBDs identificados e alguns aglomerados de valores baixo-baixo nas regiões mais distantes do centro. Embora não apresentado no mapa, verifica-se que o SBD de Barueri, no extremo oeste, aumenta o número de células entre os anos, enquanto os *clusters* observados na região mais ao leste perdem participantes.

15. Para mais detalhes, ver Almeida (2012).

FIGURA 1
Escolaridade (2014)
 (Em número médio de anos de estudos)
 1A – Distribuição espacial



1B – Análise de *clusters* espaciais

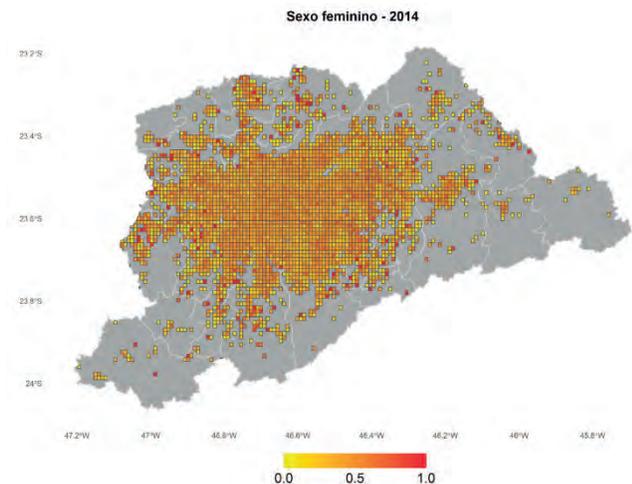


Elaboração dos autores.

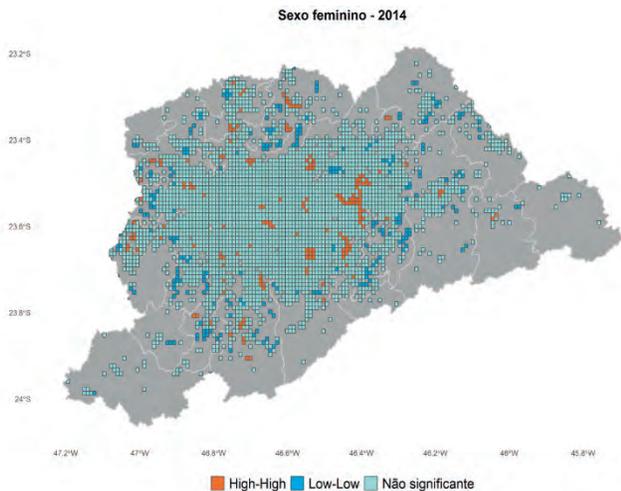
Obs.: A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

Na figura 2, dispõe-se a concentração de participação de mulheres nos empregos oferecidos em cada célula. O padrão é bem mais disperso do que no caso da escolaridade, em que se registra maior incidência e frequência de *clusters* nas áreas mais periféricas, ainda que no centro também se observem alguns. Todavia, eles praticamente desaparecem ao longo do período enquanto ganham importância nas regiões mais distantes, particularmente na zona leste da região.

FIGURA 2
Participação feminina nos empregos (2014)
2A – Distribuição espacial



2B – Análise de *clusters* espaciais



Elaboração dos autores.

Obs.: A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

As próximas figuras apresentam dimensões da complexidade das ocupações presentes nas células. Cada uma das 2.646 ocupações em que os trabalhadores estão classificados exige certo grau de habilidades, expresso em escala quantitativa no intervalo entre zero e um, conforme apresentado em Neves Júnior, Azzoni e Chagas (2021). As habilidades podem ser motoras (força, destreza, capacidade de desenvolver atividades manuais), sociais (negociação, argumentação) e cognitivas

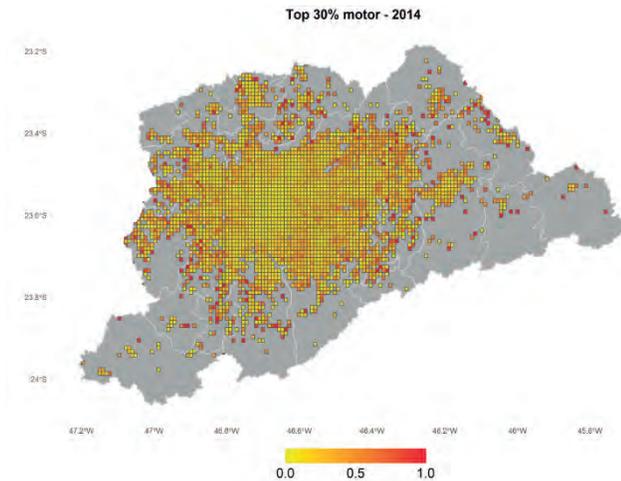
(raciocínio lógico, pensamento analítico). Cada trabalhador recebe um valor para a habilidade que sua ocupação requer. Os mapas apresentam a parcela de trabalhadores da célula cujas ocupações estão entre aquelas que mais requerem habilidades. Dessa forma, indicam o grau de sofisticação das ocupações da área e, assim, um indicador da sofisticação das atividades presentes.

A figura 3 demonstra os resultados para as habilidades motoras, que se dispersam por áreas periféricas, sendo muito pouco relevantes na área, a qual concentra *clusters* de células com valores baixos desse tipo de habilidade, circundadas por células nas mesmas condições. A evolução indica que essa área baixo-baixo se expandiu no período. Quanto às habilidades sociais, dispostas na figura 4, o padrão é inverso, com concentração no SBD identificado. Verifica-se crescimento ao longo do período, com densificação na área do SBD principal. Por fim, a figura 5 apresenta os resultados para as habilidades cognitivas, registrando-se grande concentração na área do SBD, concentração essa que aumentou ao longo do período. As habilidades sociais apresentam algumas pequenas aglomerações fora do SBD principal, mas isso não acontece com as habilidades cognitivas.

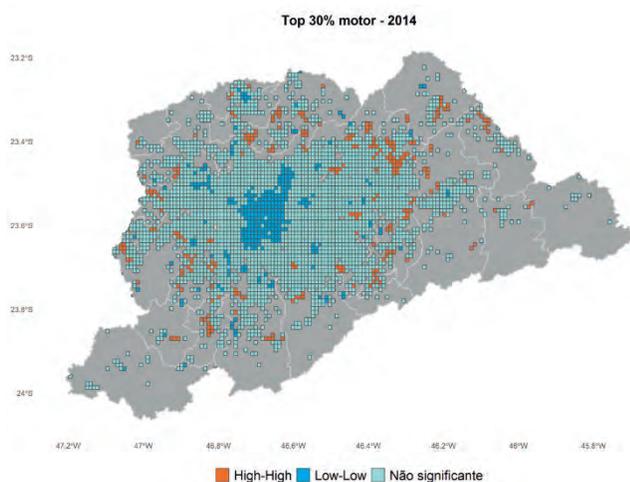
FIGURA 3

Intensidade de habilidades motoras das ocupações oferecidas (2014)

3A – Distribuição espacial



3B – Análise de *clusters* espaciais



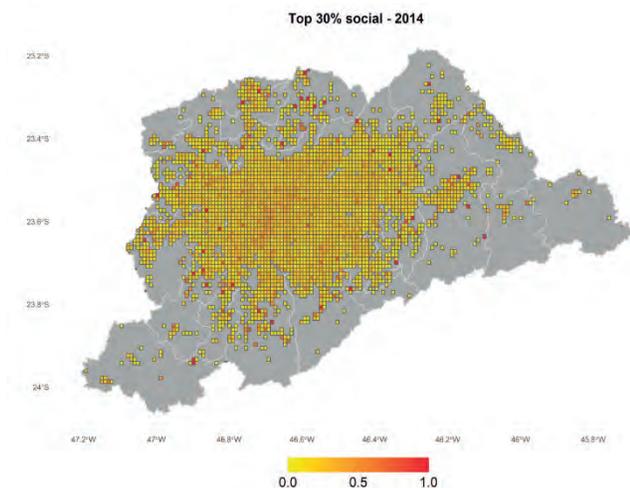
Elaboração dos autores.

Obs.: A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

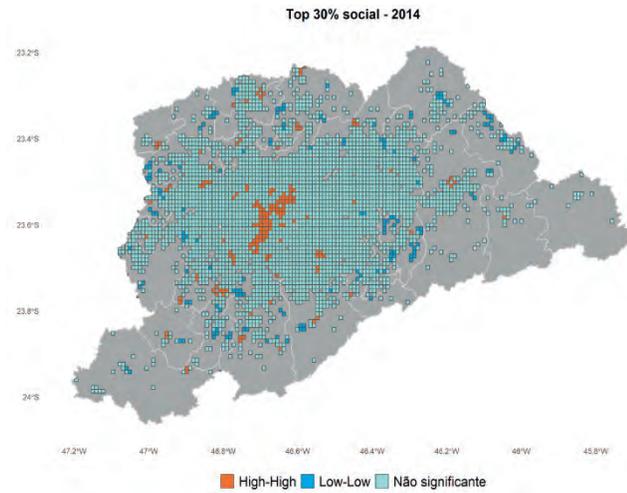
FIGURA 4

Intensidade de habilidades sociais das ocupações oferecidas (2014)

4A – Distribuição espacial



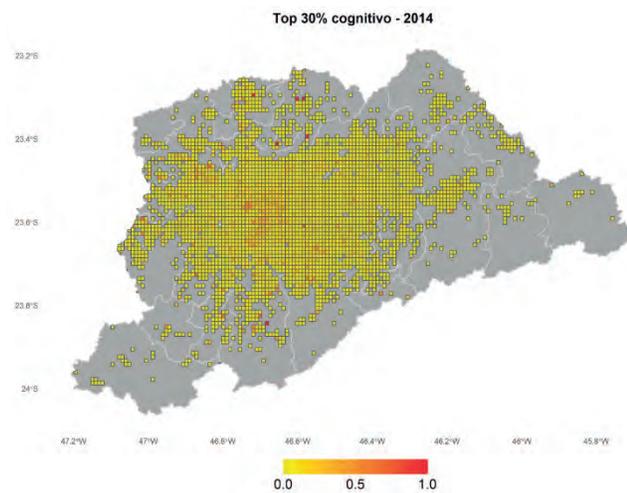
4B – Análise de *clusters* espaciais

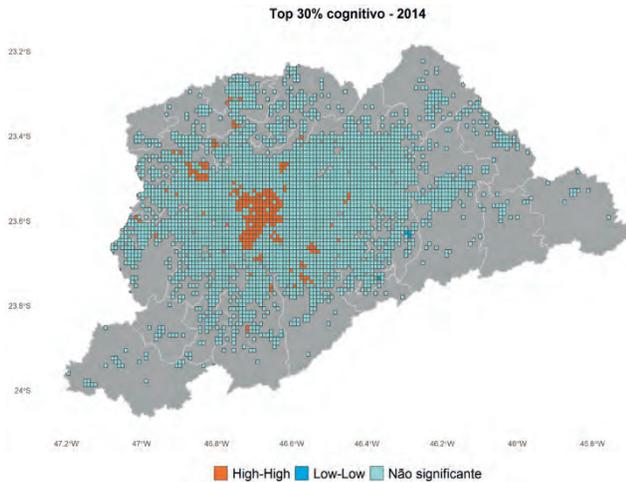


Elaboração dos autores.

Obs.: A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

FIGURA 5
Intensidade de habilidades cognitivas das ocupações oferecidas (2014)
 5A – Distribuição espacial



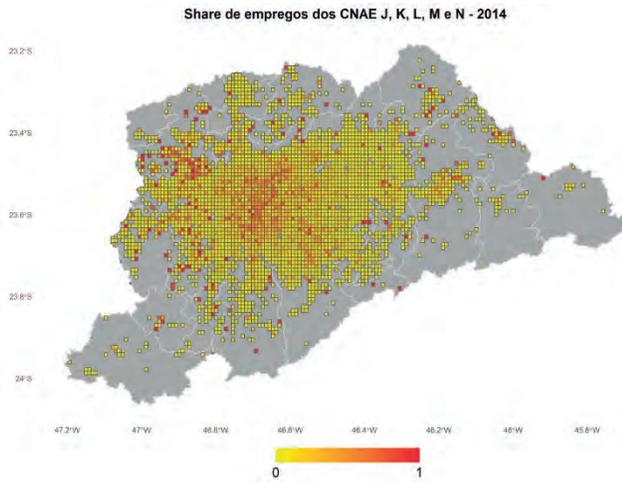
5B – Análise de *clusters* espaciais

Elaboração dos autores.

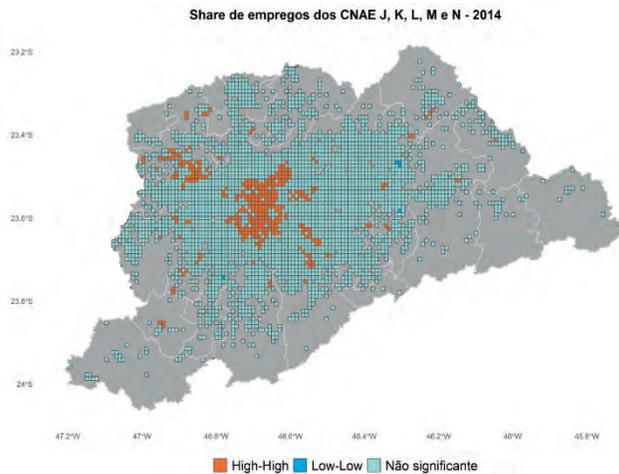
Obs.: A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

Finalmente, a figura 6 apresenta a dispersão espacial dos chamados serviços de apoio a empresas intensivos em conhecimento (*knowledge intensive business services* – KIBS), atividades sofisticadas de apoio às demais atividades econômicas. Foram assim identificados os setores CNAE J, K, L, M e N, que são: informação e comunicação; atividades financeiras, de seguros e serviços relacionados; atividades imobiliárias; atividades profissionais, científicas e técnicas; e atividades administrativas e serviços complementares. Nota-se semelhante padrão de concentração encontrado com escolaridade e habilidades sociais e cognitivas, com dispersão espacial um pouco maior no caso dos KIBS.

FIGURA 6
Participação de KIBS (2014)
 6A – Distribuição espacial



6B – Análise de *clusters* espaciais



Elaboração dos autores.

Obs.: A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

Ainda que a distribuição espacial das atividades econômicas abranja muitas partes do território metropolitano, o coração econômico da região situa-se em área limitada, correspondendo a apenas 1% da área (91 células de um conjunto possível de mais de 9.071 células de 1 km²). O panorama constituído pelos mapas apresentados indica alto grau de concentração de escolaridade, habilidades sociais e cognitivas, bem como setores KIBS em torno do SBD principal identificado neste trabalho. Isso indica que não apenas há concentração de volume de atividades, mas

também de sofisticação, o que confere grande importância econômica às duas áreas de SBDs identificadas. Ainda que claramente não se possa dizer que a RM de São Paulo constitua um caso de distribuição monocêntrica de atividades, tampouco se pode ignorar essa incrível concentração no território metropolitano.

3.4 Discussão dos resultados

Esses resultados intrametropolitanos precisam ser situados no cenário regional mais amplo da realidade brasileira. Ocorrem em contexto de perda da importância quantitativa em relação à região Sudeste e à sua principal economia – São Paulo. Internamente a esse estado, a produção é bastante concentrada na RM, que também perde participação na produção industrial nacional (Azzoni e Haddad, 2018; Diniz e Mendes, 2021). Também se verifica a situação em que as atividades terciárias, portanto urbanas, ganham importância gradativa, tanto em termos da parcela da população empregada como das atividades econômicas. Em simultâneo, atividades industriais e ligadas a recursos naturais (agropecuária e mineração) buscam outras porções no território nacional para atuar. As primeiras, por melhoria do sistema de transporte, aumento no custo da terra, redução da qualidade da mão de obra requerida, padronização dos processos produtivos, entre outros fatores; as segundas, pela proximidade dos recursos naturais, ainda que os custos de transporte possam ser significativos.

Nesse contexto, a experiência da principal metrópole brasileira apresenta pontos relevantes para reflexão. Autor (2019) mostra que, nas grandes cidades americanas, há uma tendência de concentração de atividades de baixo conteúdo de habilidades, o que requer poucos conhecimentos e níveis educacionais (serviços às famílias), em paralelo com a formação de uma elite minoritária que desempenha atividades altamente sofisticadas e bem remuneradas (KIBS). Entre essas duas situações, praticamente não há muitas atividades com demanda de conhecimentos e habilidades intermediários. Esse seria o resultado do chamado *skill-biased technical change*, o qual representa uma mudança nas tecnologias de produção que favorece o trabalho qualificado, por elevar sua produtividade e, em consequência, sua demanda relativa. Rubinton (2020) indica que o retorno do crescimento das disparidades regionais naquele país, após 1980, pode ser explicado pelas respostas das firmas aos choques tecnológicos que favoreceram atividades mais sofisticadas e padronizaram atividades básicas.

Parece ser também o caso na RM de São Paulo, como mostram os resultados. Sim, há dispersão de atividades em geral, principalmente das que pouco requerem níveis educacionais e habilidades específicas, as quais evidentemente proveem remuneração baixa. Há também concentração espacial de atividades nobres, com grandes requisitos de conhecimentos/talentos e consequentes altas remunerações. Silva e Azzoni (2021) mostram que o prêmio salarial (indicador de produtividade)

crece com o tamanho das cidades, sendo explicado pelas qualidades das firmas e dos trabalhadores nesses locais. Neves Júnior, Azzoni e Chagas (2021) indicam que o prêmio salarial por habilidades sociais e cognitivas adicionais também cresce com o tamanho urbano, e isso indica que o alto retorno por unidade adicional de habilidade é uma característica dos grandes centros urbanos. Essas duas evidências apontam para maior competitividade das grandes cidades como um todo, sem considerar como esse retorno varia internamente em suas regiões.

Os resultados desta pesquisa mostram que a heterogeneidade interna à região metropolitana paulista é muito grande. As oportunidades são muito diferentes para as pessoas com maior nível educacional do que para a maioria da população pouco instruída, tanto em termos setoriais como espaciais. Contudo, a escolaridade traduz apenas o lado da oferta de trabalho, que tende a ser mais homogênea no espaço metropolitano do que a demanda por trabalho, esta resultante das escolhas locais das firmas. Para empresas, as oportunidades de sucesso também não são homogêneas no espaço metropolitano. Os dados de concentração espacial dos postos de trabalho intensivos em habilidades sociais e cognitivas dão clara medida dessa realidade. Igualmente, os retornos às características dos trabalhadores são desigualmente distribuídos no território metropolitano. Portanto, aquela distribuição em forma de “U”, indicada por Autor (2019), tem um rebatimento espacial na concentração revelada neste estudo, e o *skill-biased technical change*, que tem implicações sobre a distribuição salarial e sobre a desigualdade, terá também um desdobramento na configuração da estrutura urbana.

4 SENSIBILIDADE DOS RESULTADOS AO TAMANHO DAS CÉLULAS

Como exposto em seção 2, os resultados podem estar sendo influenciados pela escala e pelo formato das áreas, o chamado Maup (Openshaw e Taylor, 1979). A escala diz respeito à resolução espacial (ou ao tamanho) das unidades, enquanto o formato se refere ao desenho dos limites das áreas (Andersson, Klaesson e Larsson, 2016). Essa questão tem sido extensivamente discutida na literatura de geografia mas tem sido pouco abordada nas análises em economia espacial (Burger, Van Oort e Van Der Knaap, 2010; Briant, Combes e Lafourcade, 2010).

Para averiguar se os resultados podem ser influenciados pelo tamanho escolhido para as áreas, foram refeitos os cálculos para células com 2 km² (4.611 células) e 4 km² (2.341 células). Como anteriormente, utilizou-se o critério de informação AIC, que indicou a função *kernel* tricúbica nos dois casos, conforme disposto nas tabelas 6 e 7.

TABELA 6
Critério de informação AIC e tamanho das bandas

Função <i>kernel</i>	2002		2008		2014	
	AIC	Banda (km)	AIC	Banda (km)	AIC	Banda (km)
Células com 2 km ²						
Gaussiana	81.793	1,2335	85.192	1,2685	86.573	1,2592
Biquadrada	81.685	2,9479	85.114	3,0281	86.483	3,0543
Tricúbica	81.605	2,8863	85.036	2,9182	86.403	2,9183
Células com 4 km ²						
Gaussiana	44.093	1,6664	45.901	1,6647	46.664	1,6221
Biquadrada	43.978	4,2328	45.781	4,2479	46.539	4,0656
Tricúbica	43.940	4,2479	45.742	4,2754	46.492	4,0206

Elaboração dos autores.

TABELA 7
Teste de sensibilidade para GWRs

Função <i>kernel</i>	Células com 2 km ²			Células com 4 km ²		
	2002	2008	2014	2002	2008	2014
	I	II	III	IV	V	VI
	Tricúbica	Tricúbica	Tricúbica	Tricúbica	Tricúbica	Tricúbica
Mínimo	0,00	0,00	0,00	0,00	0,00	0,00
Primeiro quartil	0,00	0,00	0,02	0,35	0,51	0,83
Mediana	5,18	3,47	12,78	21,62	36,07	42,66
Terceiro quartil	239,97	353,88	414,92	572,07	740,72	972,64
Máximo	35.652,49	55.238,39	64.651,00	56.490,29	86.777,00	104.840,00
Global	668,85	975,69	1.159,00	1.316,90	1.921,00	2.281,80
AIC	81.604,9	85.036,2	86.403,2	43.939,8	45.742,3	46.491,6
<i>Quasi-global</i> R ²	0,7757	0,7833	0,7892	0,7990	0,8072	0,8208
Número de observações	4.611	4.611	4.611	2.342	2.342	2.342

Elaboração dos autores.

No *grid* formado por células com áreas de 1 km², cada regressão estimada incluiu, em média, uma área de 25 km², com janela de 2 km em média. Portanto, ao dobrar a área, cada regressão GWRs considera o dobro da área, isto é, 50 km² e 100 km², embora mantenha fixo o número de células. Observa-se que as médias espaciais do primeiro quartil e da mediana não se alteram significativamente. Os resultados do terceiro quartil apresentam maior concentração em relação às células menores em termos de área. Como esperado, os resultados apresentam maior concentração de trabalhadores, mas a densidade da média espacial estimada

é muito próxima à das estimações anteriores. Nesse sentido, o Maup aparenta ser pouco expressivo no desvio dos resultados estimados.

Os valores de corte utilizados no processo de identificação naturalmente aumentam, como observado na tabela 8. No *grid* com células de 2 km², são identificadas 47 células, pouco mais da metade das células do *grid* de 1 km², independentemente do valor de corte. No *grid* de 4 km², identificam-se 24 células, quase um terço do total das identificadas com 1 km². Resta saber se as mesmas células identificadas em um caso o são também nos demais.

TABELA 8
Valores de corte e total de células consideradas como SBDs

Quartil	2002		2008		2014	
	Valor de corte	SBDs*	Valor de corte	SBDs*	Valor de corte	SBDs*
Células com 2 km ²						
99	10.544	47	15.697	47	18.914	47
95	3.668	231	5.397	231	6.351	231
90	1.628	462	2.292	462	2.806	462
Células com 4 km ²						
99	18.861	24	28.890	24	33.843	24
95	7.482	118	11.410	118	12.572	118
90	3.310	235	4.677	235	5.622	235

Elaboração dos autores.

Obs.: SBDs* – refere-se ao total de células do *grid* identificadas a 1% de significância estatística. As mesmas células são identificadas ao relaxar a significância estatística.

A tabela 9 apresenta as informações relevantes para esse julgamento. Em 2002, para o percentil 99, ao ampliar a área para 2 km², 21 das 91 células anteriormente identificadas perdem essa condição, e outras 25 são aceitas no SBD. Ao ampliar para 4 km², 31 células saem e 36 áreas são aceitas. Essas discrepâncias diminuem ao longo do tempo, resultando em 29 células (2 km²) e 35 células (4 km²) em 2014. Para 2 km², as novas áreas centrais representam 75% das áreas de negócios identificadas com células de 1 km²; para 4 km², a sobreposição espacial das áreas é de 63%. As taxas de aderência se alteram ao longo do tempo, alcançando valores maiores, em relação a 2002.

TABELA 9

Matriz de aderência dos SBDs por diferentes tamanhos de área das células
(Em %)

	2002			2008			2014		
	$SBD_{1Km^2}^{Q99}$	$SBD_{2Km^2}^{Q99}$	$SBD_{4Km^2}^{Q99}$	$SBD_{1Km^2}^{Q95}$	$SBD_{2Km^2}^{Q95}$	$SBD_{4Km^2}^{Q95}$	$SBD_{1Km^2}^{Q90}$	$SBD_{2Km^2}^{Q90}$	$SBD_{4Km^2}^{Q90}$
$SBD_{1Km^2}^{Q99}$	100,0	-	-	100,0	-	-	100,0	-	-
$SBD_{2Km^2}^{Q99}$	75,0	100,0	-	77,1	100,0	-	76,9	100,0	-
$SBD_{4Km^2}^{Q99}$	63,8	78,3	100,0	72,5	86,5	100,0	64,9	82,9	100,0
$SBD_{1Km^2}^{Q95}$	100,0	-	-	100,0	-	-	100,0	-	-
$SBD_{2Km^2}^{Q95}$	86,1	100,0	-	86,1	100,0	-	83,9	100,0	-
$SBD_{4Km^2}^{Q95}$	79,1	85,7	100,0	80,4	86,8	100,0	80,2	85,4	100,0
$SBD_{1Km^2}^{Q90}$	100,0	-	-	100,0	-	-	100,0	-	-
$SBD_{2Km^2}^{Q90}$	88,7	100,0	-	89,3	100,0	-	87,9	100,0	-
$SBD_{4Km^2}^{Q90}$	81,5	87,1	100,0	82,2	87,8	100,0	82,7	87,5	100,0

Elaboração dos autores.

Os mapas no apêndice B exibem as novas localizações para as áreas na categoria SBD^{Q99} , considerando as novas dimensões das células. Nota-se que os SBDs fora da capital são mais sensíveis ao tamanho das células. Esses resultados derivam do padrão espacial dos empregos na capital, que são concentrados e cerceiam os grandes SBDs. Assim, ampliar a área das células implica acréscimos maiores para essas localidades relativamente aos vizinhos dos SBDs fora do município de São Paulo. Portanto, as médias espaciais são relativamente menores quando se consideram células maiores nas regiões fora do município de São Paulo, devido à baixa extensão do deslocamento do emprego.

5 CONSIDERAÇÕES FINAIS

Este artigo apresenta alguns passos para a discussão espacial do mercado de trabalho. O primeiro desafio vencido foi o georreferenciamento da base de dados da Rais, que envolveu grande esforço e proporcionou o insumo necessário à realização da pesquisa. Objetivou-se contribuir para a discussão do espraiamento das atividades econômicas na RM de São Paulo; identificar subcentralidades; e descrever suas respectivas características. O resultado amplia o conhecimento que se tem da distribuição das atividades econômicas no espaço metropolitano e, por seu aspecto temporal, torna possível avaliar as mudanças ocorridas no período considerado.

Outro importante avanço está na proposta de metodologia capaz de identificar centralidades de emprego utilizando GWRs. Como se baseia em coeficientes estimados estatisticamente significantes, é superior às demais metodologias utilizadas em trabalhos com objetivos semelhantes, o que supera a subjetividade, a necessidade de conhecimento prévio da área por parte do pesquisador e os valores de corte mantidos fixos no tempo. A abordagem proposta pode ser vista como uma adequação dos valores de corte à dinâmica do mercado de trabalho, o que permite endogeneizar a métrica utilizada para identificar os SBDs com base na média espacial estimada do emprego, e não nos resíduos estimados. Embora o pesquisador arbitre o percentil escolhido, essa métrica é objetiva, passível de atualização temporal, em função dos dados, além de ser aplicável a qualquer região estudada. Nesse sentido, a metodologia proposta constitui um avanço na abordagem proposta por Greene (1980), McDonald (1987) e Giuliano e Small (1991).

Considerando-se o corte mais restrito, que se restringe ao percentil 99, observa-se uma grande cordilheira de alta densidade de emprego, formada nas zonas centro, sul e oeste do município. Nos municípios de Barueri e São Caetano do Sul, é possível observar SBDs tão importantes quanto os elencados em São Paulo. À medida que o valor de corte é relaxado, a concentração de emprego se eleva e se expande nas margens dessas regiões, mantendo a conexão com o contexto urbano previamente estabelecido. Mais do que a concentração de emprego, observa-se grande concentração de atividades mais sofisticadas, que requisitam maiores níveis de escolaridade e demandam mais habilidades dos trabalhadores. Igualmente, registra-se grande concentração dos chamados KIBS, que são serviços de apoio a empresas com intensidade de conhecimento.

Os resultados encontrados indicam dois movimentos espaciais do emprego na RM de São Paulo: consolidação do adensamento e descentralização concentrada do emprego, o que significa dizer que as forças de aglomeração são relevantes para as firmas. Ao longo do período estudado, os empregos se espalharam pela RM, contudo a expansão se deu próxima das áreas com elevada concentração de emprego prévia. Todavia, o crescimento do emprego manteve-se concentrado em pontos focais, uma vez que o efeito de deslocamento concentrado domina sobre o efeito de descentralização pulverizado, seja na capital, seja nos demais municípios.

Obviamente, não se trata de um caso de cidade monocêntrica, até mesmo pela existência de SBDs secundários. Porém, os resultados indicam um poder magnético muito grande do SBD principal, quantitativa e qualitativamente. Essa configuração está correlacionada aos ganhos de aglomeração e aos custos de transporte. Assim, a aglomeração fornece o ambiente para a difusão do conhecimento, o acesso ao mercado de trabalho e a redução de custo fixo na oferta de bens públicos, como meios de transporte de massa, com reflexos sobre a remuneração dos trabalhadores.

REFERÊNCIAS

- AGUILERA, A.; MIGNOT, D. Urban sprawl, polycentrism and commuting: a comparison seven french urban areas. **Urban Public Economics Review**, n. 1, p. 93-113, 2004.
- AGUIRRE-NÚÑEZ, C. A.; MARMOLEJO DUARTE, C. R. Hacia un método integrado de identificación de subcentros a escala municipal: un análisis para la región metropolitana de Barcelona. **ACE: Architecture, City and Environment**, ano 5, n. 14, p. 99-122, out. 2010.
- AGUIRRE-NUÑEZ, C. A.; MARMOLEJO DUARTE, C. R.; VERGARA-PERUCICH, J. F. Centralidad y subcentralidad en ciudades con baja regulación, el caso Antofagasta y la Serena en Chile. *In*: INTERNATIONAL CONFERENCE ON VIRTUAL CITY AND TERRITORY, 8., 2019, Barcelona, Espanha. **Anais...** Barcelona: CPSV, 2019.
- ALIDADI, M.; DADASHPOOR, H. Beyond monocentricity: examining the spatial distribution of employment in Tehran metropolitan region, Iran. **International Journal of Urban Sciences**, v. 22, n. 1, p. 38-58, jan. 2018.
- ALMEIDA, E. **Econometria espacial aplicada**. Campinas: Alínea, 2012.
- ANAS, A.; KIM, I. General equilibrium models of polycentric land use with endogenous congestion and job agglomeration. **Journal of Urban Economics**, v. 40, n. 2, p. 232-256, set. 1996.
- ANDERSSON, M.; KLAESSON, J.; LARSSON, J. P. How local are spatial density externalities? Neighbourhood effects in agglomeration economies. **Regional Studies**, v. 50, n. 6, p. 1082-1095, 2016.
- AUTOR, D. **Work of the past, work of the future**. Cambridge, Estados Unidos: NBER, fev. 2019. (Working Paper, n. 25588). Disponível em: <http://www.nber.org/papers/w25588>.
- AZZONI, C. R.; HADDAD, E. A. Regional disparities. *In*: AMANN, E.; AZZONI, C.; BAER, W. (Ed.). **Oxford handbook of the Brazilian economy**. 1. ed. Oxford: Oxford University Press, 2018. p. 422-445.
- BENDER, B.; HWANG, H.-S. Hedonic housing price indices and secondary employment centers. **Journal of Urban Economics**, v. 17, n. 1, p. 90-107, jan. 1985.
- BIVAND, R. S.; PEBESMA, E. J.; GÓMEZ-RUBIO, V. **Applied spatial data analysis with R**. Nova York: Springer, 2008.
- BOURNE, L. S. Are new urban forms emerging? Empirical tests for canadian urban areas. **Canadian Geographer**, v. 33, n. 4, p. 312-328, dez. 1989.

BOWMAN, A. W.; AZZALINI, A. **Applied smoothing techniques for data analysis: the kernel approach with S-plus illustrations**. Oxford: Oxford University Press, 1997.

BRIANT, A.; COMBES, P.; LAFOURCADE, M. Dots to boxes: do the size and shape of spatial units jeopardize economic geography estimations? **Journal of Urban Economic**, v. 67, p. 287-302, 2010.

BURGER, M. J.; VAN OORT, F. G.; VAN DER KNAAP, G. A. A treatise on the geographical scale of agglomeration externalities and the modifiable areal unit problem. **Scienze Regionali**, v. 9, p. 19-40, 2010.

BURNS, M. C.; MOIX BERGADÀ, M.; CLADERA, J. R. Contrasting indications of polycentrism within Spain's metropolitan urban regions. *In: EUROPEAN ESTATE SOCIETY CONFERENCE*, 8., 2001, Alicante, Espanha. **Anais...** 2001.

CAMPOS, R. B. A. **Subcentralidades e prêmio salarial intraurbano na região metropolitana de São Paulo**. 2018. Tese (Doutorado) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2018. Disponível em: <chrome-extension://efaidnbmninnibpcajpcgclefindmkaj/https://teses.usp.br/teses/disponiveis/12/12140/tde-26112018-122557/publico/CorrigidoRodger.pdf>.

CAMPOS, R. B. A.; AZZONI, C. R. Dispersão concentrada do emprego intraurbano: modelos teóricos e abordagens empíricas. **Revista Brasileira de Estudos Regionais e Urbanos**, v. 13, n. 4, p. 606-627, 2019.

CAMPOS, R. B. A.; AZZONI, C. R. Empirical identification of intra-urban subcentralities: a new methodological approach with an application for a developing country. **Journal Economía**, v. 44, n. 87, p. 56-73, 2021.

CAMPOS, R. B. A.; CHAGAS, A. L. S. Employment sub-centers of a megacity in a developing country: the case of the municipality of São Paulo, Brazil. **Nova Economia**, v. 31, n. 3, p. 839-867, 2021.

CASTELLS, M.; HALL, P. **Technopoles of the world: the making of twenty-first-century industrial complexes**. Nova York: Routledge, 1994.

CERQUEIRA, E. D. V.; DINIZ, A. M. A. Identifying centers and subcenters in the metropolitan region of Belo Horizonte through Google places of interest. **Mercator**, Fortaleza, v. 21, p. 1-14, 2022. Disponível em: <https://www.scielo.br/j/mercator/a/J69BmCc9Yz55XVh3kjpR4vw/?format=pdf&lang=en>.

CERVERO, R. **America's suburban centers: the land use-transportation link**. Winchester: Unwin Hyman, 1989.

CLARK, W. A. V.; KUIJPERS-LINDE, M. Commuting in restructuring urban regions. **Urban Studies**, v. 31, n. 3, p. 465-483, 1994.

DENG, F. Intrametropolitan variation of urban structure and housing price: the case of Chongqing, China. **Asia-Pacific Journal of Regional Science**, v. 4, n. 3, p. 639-655, out. 2020.

DINIZ, C. C.; MENDES, P. S. **Tendências regionais da indústria brasileira no século XXI**. Rio de Janeiro: Ipea, abr. 2021. (Texto para Discussão, n. 2640). Disponível em: https://portalantigo.ipea.gov.br/portal/index.php?option=com_content&view=article&id=37759.

FABER, S.; PÁEZ, A. A systematic investigation of cross-validation in GWR model estimation: empirical analysis and Monte Carlo simulations. **Journal of Geographical Systems**, v. 9, n. 4, p. 371-396, 2007.

FERNÁNDEZ-MALDONADO, A. M. *et al.* Polycentric structures in Latin American metropolitan areas: identifying employment sub-centres. **Regional Studies**, v. 48, n. 12, p. 1954-1971, maio 2013.

FOTHERINGHAM, A. S.; BRUNSDON, C.; CHARLTON, M. **Geographically weighted regression: the analysis of spatially varying relationship**. Nova York: Wiley, 2002.

FUJITA, M.; OGAWA, H. Multiple equilibria and structural transition of non-monocentric urban configurations. **Regional Science and Urban Economics**, v. 12, n. 2, p. 161-191, 1982.

GIULIANO, G.; SMALL, K. A. Subcenters in the Los Angeles region. **Regional Science and Urban Economics**, v. 21, n. 2, p. 163-182, 1991.

GIULIANO, G. *et al.* **Not all sprawl: evolution of employment concentrations in Los Angeles, 1980-2000**. Los Angeles: Lusk Center for Real Estate/USC, fev. 2005. (Working Paper, n. 2005-1002).

GIULIANO, G. *et al.* Polycentricity and the evolution of metropolitan spatial structure. **Growth and Change**, v. 53, n. 3, p. 593-627, dez. 2021.

GORDON, P.; RICHARDSON, H. W.; WONG, H. L. The distribution of population and employment in a polycentric city: the case of Los Angeles. **Environment and Planning A**, v. 18, n. 2, p. 161-173, 1986.

GRAHAM, S.; MARVIN, S. **Telecommunications and the city: electronic spaces, urban places**. Londres: Routledge, 1996.

GREENE, D. Recent trends in urban spatial structure. **Growth and Change**, v. 10, p. 29-40, 1980.

- HAJRASOULIHA, A. H.; HAMIDI, S. The typology of the american metropolis: monocentricity, polycentricity, or generalized dispersion? **Urban Geography**, v. 38, n. 3, p. 420-444, mar. 2017.
- HARTWICK, P. G.; HARTWICK, J. M. Efficient resource allocation in a multinucleated city with intermediate goods. **Quarterly Journal of Economics**, v. 88, n. 2, p. 340-352, 1974.
- HEIKKILA, E. *et al.* What happened to the CBD-distance gradient? Land rents in a polycentric city. **Environment and Planning A**, v. 21, n. 2, p. 221-232, 1989.
- HELSEY, R. W.; SULLIVAN, A. M. Urban subcenter formation. **Regional Science and Urban Economics**, v. 21, n. 2, p. 255-275, 1991.
- HOTCHKISS, D.; WHITE, M. J. A simulation model of a decentralized metropolitan area with two-worker, traditional and female-headed household. **Journal of Urban Economics**, v. 34, n. 2, p. 159-185, 1993.
- KANE, K.; HIPPE, J. R.; KIM, J. H. Los Angeles employment concentration in the 21st century. **Urban Studies**, v. 55, n. 4, p. 1-26, 2016.
- KREHL, A. Urban subcentres in German city regions: identification, understanding, comparison. **Papers in Regional Science**, v. 97, p. 79-104, mar. 2018.
- KWON, K. Polycentricity and the role of government-led development: employment decentralization and concentration in the Seoul metropolitan area, 2000-2015. **Cities**, v. 111, p. 1-10, 2021.
- LI, Y. Towards concentration and decentralization: the evolution of urban spatial structure of Chinese cities, 2001-2016. **Computers Environment and Urban Systems**, v. 80, p. 1-10, 2020.
- LIU, X. *et al.* Analyzing the changes of employment subcentres: a comparison study of Houston and Dallas. **Urban Studies**, v. 56, n. 12, p. 2532-2548, 2019.
- MANDUCA, R. A. The spatial structure of US metropolitan employment: new insights from administrative data. **Environment and Planning B: Urban Analytics and City Science**, v. 48, p. 1357-1372, 2020.
- MCDONALD, J. F. The identification of urban employment subcenters. **Journal of Urban Economics**, v. 21, n. 2, p. 242-258, 1987.
- MCMILLEN, D. P.; MCDONALD, J. F. A nonparametric analysis of employment density in a polycentric city. **Journal of Regional Science**, v. 37, n. 4, p. 591-612, nov. 1997.
- MCMILLEN, D. P.; SMITH, S. C. The number of subcenters in large urban areas. **Journal of Urban Economics**, v. 53, n. 3, p. 321-338, 2003.

MIESZKOWSKI, P.; SMITH, B. Analyzing urban decentralization: the case of Houston. **Regional Science and Urban Economics**, v. 21, n. 2, p. 183-199, 1991.

OGAWA, H.; FUJITA, M. Equilibrium land use patterns in a non-monocentric city. **Journal of Regional Science**, v. 20, n. 4, p. 455-475, 1980.

OPENSHAW, S.; TAYLOR, P. J. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In: WRIGLEY, N. (Ed.). **Statistical application in the spatial science**. London: Pion, 1979. p. 127-144.

PAPAGEORGIU, G. J. Rent distribution models within a multicentral framework. **Environment and Planning A**, v. 3, p. 267-282, 1971.

REDFEARN, C. The topography of metropolitan employment: identifying centers of employment in a polycentric urban area. **Journal of Urban Economics**, v. 61, n. 3, p. 519-541, maio 2007.

RICHARDSON, H. W. *et al.* Residential property values, the CBD, and multiple nodes: further analysis. **Environment and Planning A**, v. 22, p. 829-833, 1990.

RODRIGUES, F. A. da C.; SILVEIRA NETO, R. da M.; MIRANDA, F. Identification of employment subcenters in Brazilian metropolitan regions. In: ENCONTRO NACIONAL DE ECONOMIA, 47., 2019, São Paulo. **Anais...** São Paulo: Anpec, 2019. Disponível em: https://www.anpec.org.br/encontro/2019/submissao/files_I/i10-5bbc4f56d19a715e4bcb780479f6a76a.pdf.

ROMANOS, M. C. Household location in a linear multi-center metropolitan area. **Regional Science and Urban Economics**, v. 7, n. 3, p. 233-250, 1979.

ROSS, S.; YINGER, J. Comparative static analysis of open urban models with a full labor market and suburban employment. **Regional Science and Urban Economics**, v. 25, p. 575-605, 1995.

RUBINTON, H. **The geography of business dynamism and skill biased technical change**. Saint Louis: Federal Reserve Bank of Saint Louis, abr. 2020. (Working Paper 2020-020E).

SASAKI, K.; MUN, S. A dynamic analysis of a multiple-center formation in a city. **Journal of Urban Economics**, v. 40, n. 3, p. 257-278, 1996.

SILVA, D. L. G.; AZZONI, C. R. Worker and firm heterogeneity, agglomeration, and wages in Brazil. **Papers in Regional Science**, p. 1-27, 2021.

SIVITANIDOU, R.; WHEATON, W. C. Wage and rent capitalization in the commercial real estate market. **Journal of Urban Economics**, v. 31, n. 2, p. 206-229, 1992.

SULLIVAN, A. A general equilibrium model with agglomerative economies and decentralized employment. **Journal of Urban Economics**, v. 20, n. 1, p. 55-74, 1986.

TAUBENBÖCK, H. *et al.* Measuring morphological polycentricity: a comparative analysis of urban mass concentrations using remote sensing data. **Computers, Environment and Urban Systems**, v. 64, p. 42-56, jul. 2017.

THOMÉ, E. M.; PORSE, A. A. Urban employment subcenters: the case of Curitiba, Brazil. *In*: ENCONTRO NACIONAL DE ECONOMIA, 47., 2019, São Paulo. **Anais...** São Paulo: Anpec, 2019. Disponível em: https://www.anpec.org.br/encontro/2019/submissao/files_I/i10-1ef07e40df2f8bca3379063e1d77fa9d.pdf.

WEN, H.; TAO, Y. Polycentric urban structure and housing price in the transitional China: evidence from Hangzhou. **Habitat International**, v. 46, n. 2, p. 138-146, 2015.

WHITE, M. J. Firm suburbanization and urban subcenters. **Journal of Urban Economics**, v. 3, n. 4, p. 323-343, 1976.

WHITE, M. J. Location choice behavior and commuting behavior in cities with decentralized employment. **Journal of Urban Economics**, v. 24, n. 2, p. 129-152, 1988.

WHITE, M. J. Commuting and congestion: a simulation model of a decentralized metropolitan area. **Journal of the American Real Estate and Urban Economics Association**, v. 18, n. 3, p. 335-368, 1990.

WHITE, M. J. Urban areas with decentralized employment: theory and empirical work. *In*: CHESIRE, P.; MILLS, E. S. (Ed.). **Handbook of regional and urban economics**. 1. ed. Amsterdã: North-Holland, 1999. v. 3, p. 1375-1412.

WIEAND, K. F. An extension of the monocentric urban spatial equilibrium model to a multicenter setting: the case of the two-center city. **Journal of Urban Economics**, v. 21, n. 3, p. 259-271, 1987.

WREDE, M. A continuous spatial choice logit model of polycentric city. **Regional Science and Urban Economics**, v. 53, p. 68-73, 2015.

WRIGLEY, N. *et al.* Analyzing, modelling, and resolving the ecological fallacy. *In*: LONGLEY, P. A.; BATTY, M. (Ed.). **Spatial analysis: modelling in a GIS environment**. Nova York: John Wiley Sons, 1996. p. 25-41.

YAN, L.; WANG, D.; ZHANG, S. Theoretical foundation and framework for understanding urban centers. **Progress in Geography**, v. 39, n. 9, p. 1576-1586, 2020.

YAN, L. *et al.* Understanding urban centers in Shanghai with big data: local and non-local function perspectives. **Cities**, v. 113, p. 1-12, jun. 2021. Disponível em: <https://doi.org/10.1016/j.cities.2021.103156>.

YAN, L. *et al.* A new approach for identifying urban employment centers using mobile phone data: a case study of Shanghai. **International Journal of Geographical Information Science**, v. 37, n. 5, p. 1-28, 2023.

BIBLIOGRAFIA COMPLEMENTAR

BESSA, V. **Estrutura dos Knowledge-Intensive Business Services (KIBS), no município de São Paulo**. São Paulo: PUC-SP, [s.d.]. Disponível em: https://www.pucsp.br/sites/default/files/download/eitt/vi_ciclo_vagner_bessa_kibs.pdf. Acesso em: 17 nov. 2020.

BRASIL. Ministério do Trabalho e Emprego. **Relação Anual de Informações Sociais (Rais) e Cadastro Geral de Empregados e Desempregados (Caged)**. Brasília: MTE, 2010.

MCDONALD, J. F.; MCMILLEN, D. P. Employment subcenters and land values in a polycentric urban area: the case of Chicago. **Environment and Planning A**, v. 22, n. 12, p. 1561-1574, 1990.

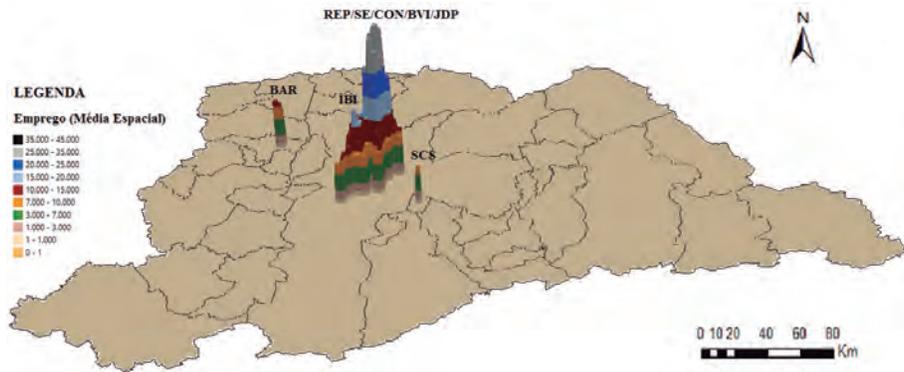
MCDONALD, J. F.; PRATHER, P. J. Suburban employment centres: the case of Chicago. **Urban Studies**, v. 31, n. 2, p. 201-218, 1994.

NEVES JÚNIOR, E. C. das; AZZONI, C. R.; CHAGAS, A. L. S. Habilidades e prêmio salarial urbano. **Pesquisa e Planejamento Econômico**, Brasília, v. 51, n. 2, p. 9-45, ago. 2021.

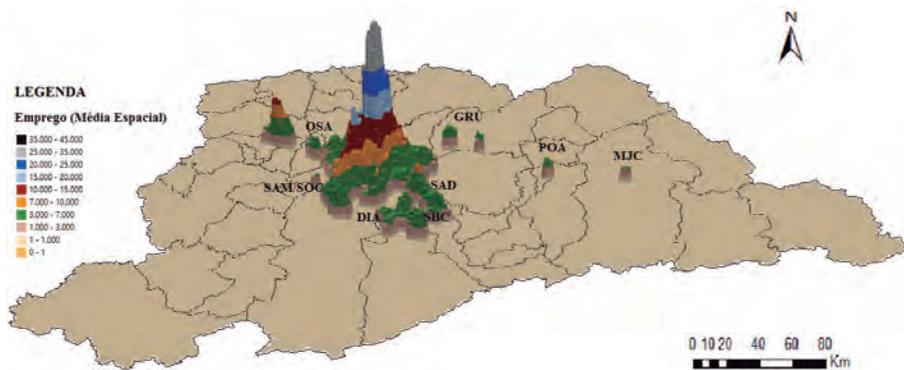
YINGER, J. Urban models with more than one employment center. **Journal of Urban Economics**, v. 31, n. 2, p. 181-205, 1992.

APÊNDICE A

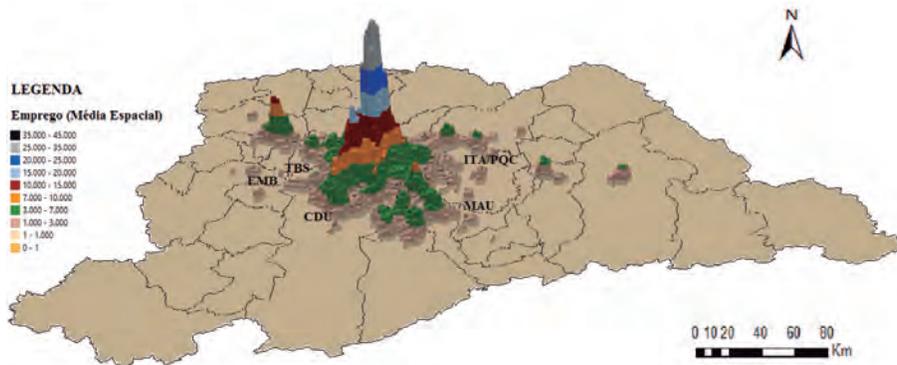
MAPA A.1
SBDs identificados em 2008 com *grid* de 1 km²
 A.1A – Percentil 99



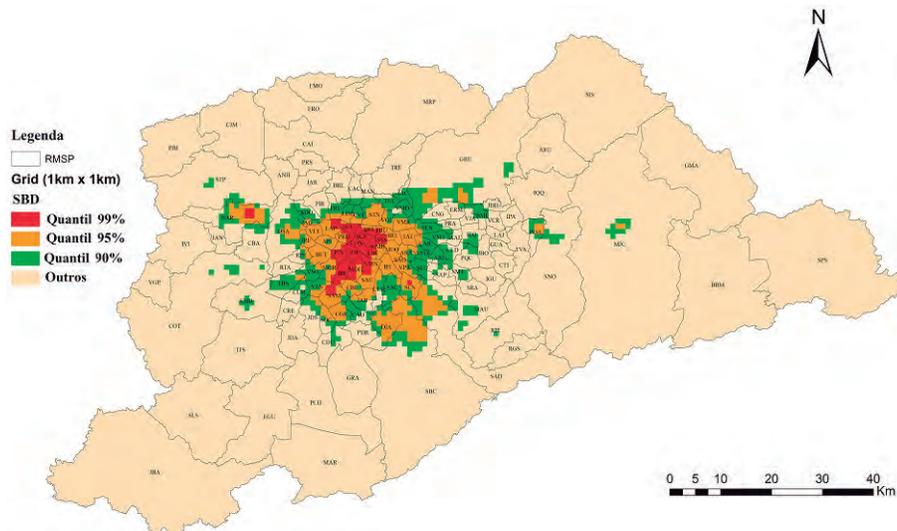
A.1B – Percentil 95



A.1C – Percentil 90



A.1D – Todos os percentis



Fonte: Resultados derivados dos modelos de regressões ponderadas geograficamente (*geographically weighted regressions – GWRs*).
 Elaboração dos autores.

Obs.: 1. SBDs – subcentros (*subcenter business districts*).

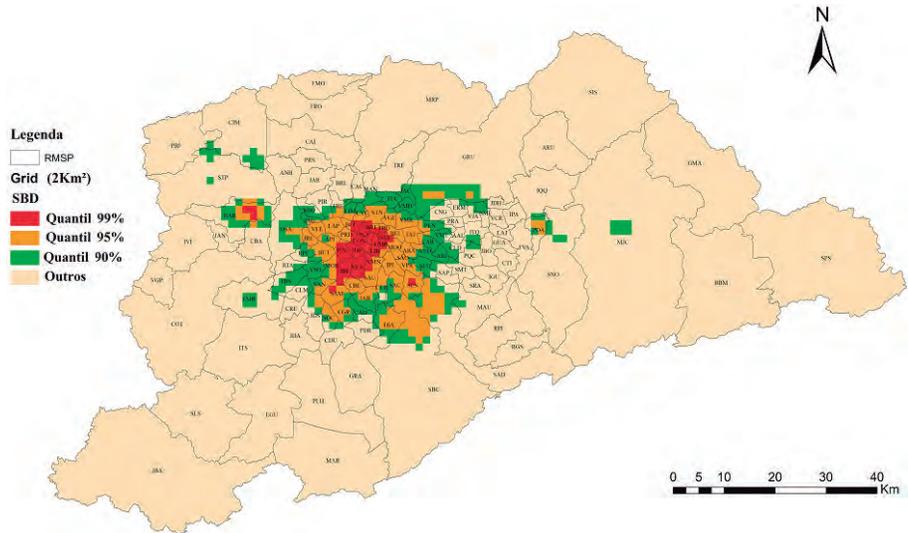
2. A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

APÊNDICE B

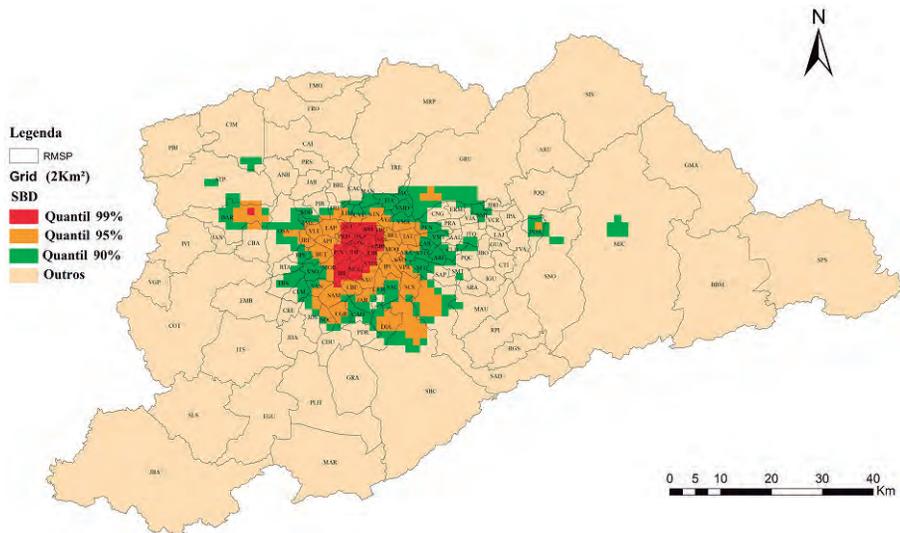
MAPA B.1

SBDs identificados com *grid* de 2 km² (2002 e 2014)

B.1A – SBDs (2002)



B.1B – SBDs (2014)



Fonte: Resultados derivados dos modelos GWRs.

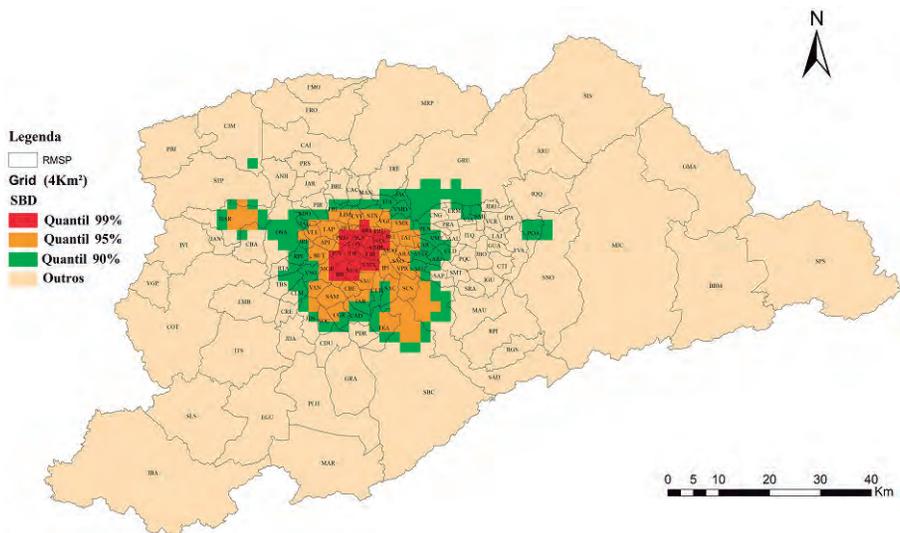
Elaboração dos autores.

Obs.: A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

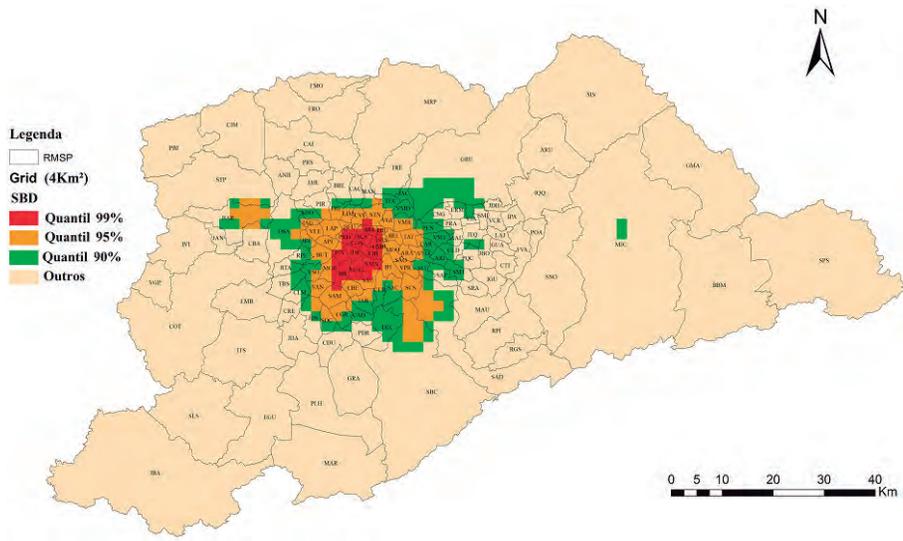
MAPA B.2

SBDs identificados com *grid* de 4 km² (2002 e 2014)

B.2A – SBDs (2002)



B.2B – SBDs (2014)



Fonte: Resultados derivados dos modelos GWRs.

Elaboração dos autores.

Obs.: A figura não pôde ser padronizada e revisada em virtude das condições técnicas dos originais (nota do Editorial).

Originais submetidos em: set. 2019.

Última versão recebida em: maio 2023.

Aprovada em: mar. 2023.

