

De Conceitos a Medidas Territoriais: A Construção de Índices Espaciais de Segregação Urbana

*Flávia F. Feitosa, Gilberto Câmara, Antônio Miguel V. Monteiro,
Thomas Koschitzki, Marcelino P. S. Silva*

1 Introdução

Com o progressivo processo de descentralização político-administrativo brasileiro, iniciado com a Constituição de 1988 e reforçado com o Estatuto da Cidade de 2001, o planejamento das cidades vem adquirindo um crescente reconhecimento de sua importância como instrumento indispensável na regulamentação das ações dos diversos agentes urbanos. Em concordância com esta tendência, observa-se uma preocupação em relação à compreensão da cidade e seus fenômenos, assim como ao embasamento teórico e técnico dos planejadores envolvidos no processo de intervenção sobre o espaço urbano.

Este esforço em busca de avanços na compreensão do espaço urbano requer a interação entre dois elementos básicos (Ramos, 2002): (a) um conjunto de conceitos e teorias que caracterizem os fenômenos que atuam sobre o espaço urbano, seus padrões, causas e implicações e (b) métodos e técnicas quantitativas de representação e análise do objeto de estudo em questão. A evolução no processo de apreensão de um fenômeno urbano em particular ocorre com a interação cíclica entre estes dois elementos: enquanto os conceitos e teorias fundamentam o processo de representação do objeto investigado, este permite a comprovação ou refutação dos conceitos e teorias existentes e a revelação de novas propriedades do fenômeno. Estas propriedades devem ser traduzidas em novos conceitos e teorias - ou no refinamento dos já existentes - que, por sua vez, realimentam o ciclo e viabilizam uma contínua mutação dos elementos básicos.

Este capítulo dedica-se a um dos elementos mencionados, mais especificamente, à construção de índices que visam representar um fenômeno de destaque nos atuais debates sobre políticas públicas: a *segregação urbana*.

O conceito de segregação urbana exprime a separação entre diferentes grupos sociais em um ambiente urbano. A segregação - seja ela de cunho racial, socioeconômico, religioso ou étnico - é comumente relacionada à imposição de inúmeros prejuízos a grupos populacionais menos privilegiados. Entre estes prejuízos, podem ser apontados: dificuldades de acesso a equipamentos, serviços e infra-estruturas, redução de oportunidades de emprego, acentuada discriminação e uma maior exposição à violência, desastres naturais e problemas de saúde (Luco e Rodríguez, 2003; Rodríguez, 2001; Sabatini et al., 2001; Torres, 2004).

Associados à segregação urbana, estão ainda impactos que afetam não apenas as famílias excluídas, mas todos os moradores da cidade, como o aumento da violência, do medo e a proliferação de soluções como condomínios exclusivos e *shopping centers*, que intensificam ainda mais a segregação e suas conseqüências (Caldeira, 2000). Além disto, estes empreendimentos contribuem para a deterioração da cidade ao fragmentá-la, renegar os espaços públicos e o convívio entre os diferentes.

Apesar do reconhecimento da importância da segregação urbana, ainda são poucos os trabalhos latino-americanos que utilizam índices de segregação como uma ferramenta de análise do fenômeno. Entre os raros trabalhos existentes, podemos citar os de Telles (1992; 1995), Sabatini (2001) e Torres (2004), que adotam o tradicional índice de dissimilaridade D (Duncan e Duncan, 1955), e os trabalhos de Rodríguez (2001) e Luco e Rodríguez (2003), que aplicam índices baseados em variância. No entanto, estes trabalhos baseiam-se em índices muito simples, incapazes de representar a complexidade dos atuais conceitos e teorias a respeito da segregação. Surge, portanto, a necessidade de utilização de medidas mais sofisticadas, que apresentem propriedades importantes para as análises de segregação. Entre estas propriedades, destaca-se a capacidade de incorporar informações espaciais, de revelar as variações do fenômeno ao longo da cidade e de elaborar análises em múltiplas escalas.

Este capítulo apresenta índices de segregação alternativos que superam algumas das principais limitações dos índices comumente utilizados. O capítulo trata ainda do

problema de validação dos índices de segregação e propõe a aplicação de um teste estatístico para a verificação da significância dos resultados destes índices. Posteriormente, características dos índices apresentados e aspectos importantes sobre como aplicá-los e interpretá-los serão exemplificados por intermédio de experimentos realizados sobre dados artificiais e reais.

2 Índices de Segregação Urbana: Uma Breve Revisão

Os primeiros estudos focados na mensuração da segregação tiveram início no final da década de 40 e início da década de 50, quando vários índices passaram a ser propostos e discutidos nos Estados Unidos. O índice mais popular desta primeira geração de medidas foi o índice de dissimilaridade D (Duncan e Duncan, 1955). O índice D , ainda utilizado por muitos pesquisadores, mede a relação existente entre a composição populacional das unidades de área (por exemplo, setores ou distritos) e a composição populacional de toda a área de estudo. O índice varia entre 0 e 1 e é definido como:

$$D = \frac{1}{2} \sum_{j=1}^J \frac{N_{jm}}{N_m} - \frac{N_{jn}}{N_n}, \quad (1)$$

onde N_{jm} e N_{jn} são, respectivamente, a população dos grupos m e n na unidade de área j , enquanto N_m e N_n representam a população total dos grupos m e n na área de estudo.

Além do índice de dissimilaridade, outros índices foram propostos neste período, como o índice de exposição/isolamento (Bell, 1954), o índice Gini (Cowell, 1977) e o índice Atkinson (Atkinson, 1970). Contudo, todos os índices propostos nesta fase limitaram-se a medir segregação entre dois grupos populacionais. Esta restrição reflete as preocupações sociais que motivaram o desenvolvimento de tais índices: a dicotômica realidade americana nos anos 50 aos 70 – era dos direitos civis – na qual a segregação entre negros e brancos era o foco de muitos estudos.

A partir da década de 70, os estudos raciais de segregação entre negros e brancos começaram a ceder espaço para estudos sobre a segregação entre múltiplos grupos, como a segregação entre distintas classes sociais e entre brancos, negros e hispânicos (Reardon e Firebaugh, 2002). Para atender a esta nova demanda, índices para medir a segregação entre vários grupos passaram a ser propostos. Entre estes, destacam-se o índice de entropia (Theil e Finizza, 1971), índice de dissimilaridade

generalizado $D(m)$ (Sakoda, 1981), índice Gini generalizado (Boisso et al., 1994) e o índice de segregação residencial ISR (Jargowsky, 1996; Rodríguez, 2001).

No entanto, todos os índices existentes até este momento apresentavam uma importante limitação: a incapacidade de considerar o arranjo espacial da população entre as unidades de área. Este fato leva ao que White (1983) descreveu como o “problema do tabuleiro de damas” (*checkerboard problem*): dados dois tabuleiros de damas – o primeiro dotado de quadrados pretos em uma de suas metades e quadrados brancos na outra metade, e o segundo com alternância entre quadrados brancos e pretos -, uma medida não espacial como o índice de dissimilaridade D mostraria o mesmo nível de segregação em ambas as situações (Figura 1).

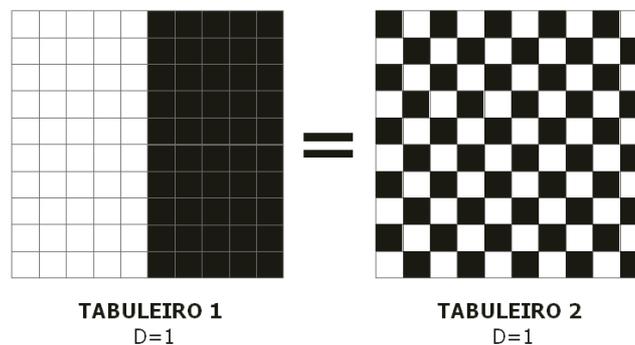


Figura 1 - Problema do tabuleiro de damas (*checkerboard problem*).

Para suprir esta inadequação, alguns pesquisadores passaram a dedicar-se ao desenvolvimento de medidas capazes de incorporar informações espaciais em suas formulações (Jakubs, 1981; Morgan, 1983; Morrill, 1991; Reardon e O'Sullivan, 2004; White, 1983; Wong, 1993, 1998).

Baseados nesta concepção espacial de mensuração da segregação, Reardon e O'Sullivan (2004) estabeleceram duas dimensões espaciais da segregação: *dispersão/agrupamento* e *exposição/isolamento* (Figura 2). A dimensão dispersão/agrupamento refere-se ao equilíbrio da distribuição dos distintos grupos populacionais no espaço residencial da cidade. A dimensão exposição/isolamento, por sua vez, diz respeito à capacidade de encontro entre membros de grupos distintos (ou do mesmo grupo, no caso do isolamento) (Reardon e O'Sullivan, 2004).

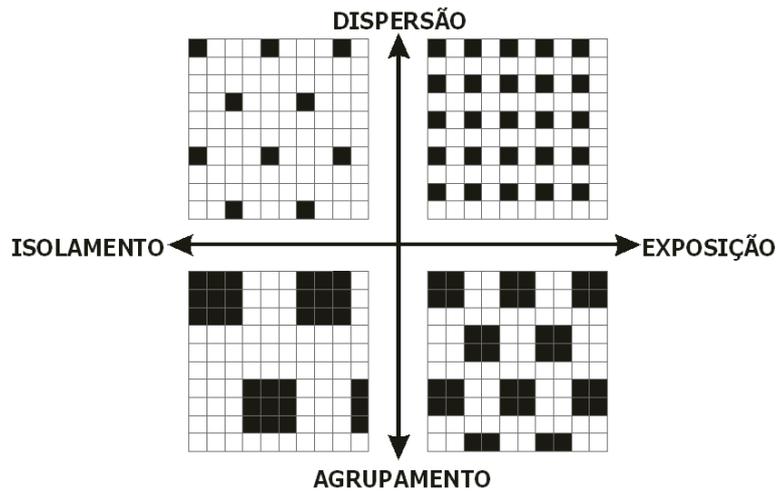


Figura 2 - Dimensões espaciais da segregação urbana.
Adaptada de Reardon e O'Sullivan (2004).

Apesar da existência de índices espaciais, a grande maioria dos estudos empíricos baseia-se em índices não-espaciais, que são facilmente computáveis em planilhas eletrônicas. Já os índices espaciais requerem a extração de informações geográficas, um procedimento que necessita de implementações computacionais nem sempre acessíveis aos pesquisadores da área.

3 Construção de Índices Espaciais de Segregação Urbana

A abordagem proposta para a construção dos índices de segregação parte da hipótese de que um ambiente urbano é composto por diferentes *localidades*, ou seja, locais onde as pessoas vivem e interagem com seus vizinhos. A intensidade destas interações, um aspecto chave para a mensuração da segregação, varia de acordo com o grau de proximidade entre as pessoas.

A abordagem apresentada foi desenvolvida para dados agregados em unidades de área, como setores censitários ou distritos. Cada unidade de área considerada na análise é associada a uma localidade, cujas características populacionais são expressas por sua *intensidade populacional local*. O cômputo da intensidade populacional local é realizado através do uso de um estimador kernel (Silverman, 1986). Estimadores kernel são funções que determinam a intensidade de um determinado evento nos diferentes pontos da área de estudo. Para calcular a intensidade populacional local de uma localidade j , o estimador kernel é posicionado no centróide da unidade j e computa a média ponderada dos dados populacionais da área de estudo (Figura 3). Os pesos

utilizados neste cômputo dependem dos parâmetros do estimador kernel escolhido pelo pesquisador. Existem inúmeros tipos de estimadores kernel, entre eles o kernel linear, gaussiano, polinomial e sigmóide (Schölkopf e Smola, 2002).

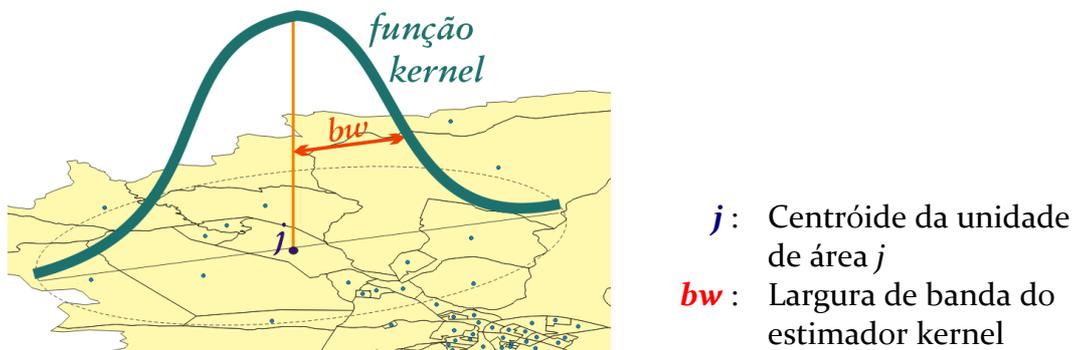


Figura 3 - Estimador kernel gaussiano.

3.1 Índices Globais Espaciais de Segregação

Índices globais de segregação são aqueles que expressam o grau de segregação da cidade como um todo. Esta seção apresenta índices espaciais globais de segregação construídos a partir do conceito de *intensidade populacional local*. São estes:

- a) Índice espacial de dissimilaridade generalizado $\check{D}(m)$, que é uma medida da diferença média entre a composição populacional das localidades e a composição populacional total da cidade;
- b) Índice espacial de exposição $\check{P}_{(m,n)}^*$, que mede o contato potencial entre os grupos *m* e *n*;
- c) Índice espacial de isolamento \check{Q}_m , que mede o contato potencial do grupo *m* com ele próprio;
- d) Índice espacial de segregação residencial \check{ISR} , que é uma medida das disparidades populacionais entre as distintas localidades da cidade.

O índice de dissimilaridade generalizado espacial $\check{D}(m)$, o índice espacial de exposição $\check{P}_{(m,n)}^*$ e o índice espacial de isolamento \check{Q}_m são mais adequados para estudos que utilizam variáveis categóricas, como os de segregação racial ou étnica. O índice espacial de segregação residencial \check{ISR} é apropriado para estudos que utilizam em variáveis contínuas, como os de segregação socioeconômica baseados em renda.

Todos os índices espaciais de segregação propostos requerem a estimativa da *intensidade populacional local de todas as localidades* da área de estudo. Em termos formais, a intensidade populacional local de uma localidade j (\tilde{L}_j) é definida como:

$$\tilde{L}_j = \sum_{j=1}^J k(N_j), \quad (2)$$

onde N_j é a população total na unidade de área j , J é o número total de unidades de área na cidade ou área metropolitana. O símbolo k representa a função do estimador kernel, que estima a influência das características populacionais de cada unidade de área sobre a localidade j . A intensidade populacional local do grupo populacional m na localidade j (\tilde{L}_{jm}) é calculada através da substituição da população total na unidade de área j (N_j) pela população do grupo m na unidade de área j (N_{jm}) na Equação (2).

$$\tilde{L}_{jm} = \sum_{j=1}^J k(N_{jm}). \quad (3)$$

3.1.1 Índice Espacial de Dissimilaridade Generalizado

O índice espacial de dissimilaridade generalizado $\tilde{D}(m)$ é uma versão espacial do índice de dissimilaridade generalizado $D(m)$ desenvolvido por Sakoda (1981). O índice espacial $\tilde{D}(m)$ mede a diferença média entre a composição populacional das várias *localidades* da cidade e a composição populacional da cidade como um todo. Este índice capta a dimensão espacial dispersão/agrupamento.

Dado um conjunto de grupos populacionais, o índice espacial de dissimilaridade generalizado $\tilde{D}(m)$ é definido como:

$$\tilde{D}(m) = \sum_{j=1}^J \sum_{m=1}^M \frac{N_j |\tilde{\tau}_{jm} - \tau_m|}{2NI} \quad (4)$$

$$\text{, onde } I = \sum_{m=1}^M \tau_m (1 - \tau_m) \quad \text{e} \quad \tilde{\tau}_{jm} = \frac{\tilde{L}_{jm}}{\tilde{L}_j} \quad (5)(6)$$

Nas Equações (4) e (5), N é a população total da cidade, N_j é a população total na unidade de área j , τ_m é a proporção do grupo m na cidade, $\tilde{\tau}_{jm}$ é a proporção do grupo m na localidade j e M é o número total de grupos populacionais. Na Equação (6), \tilde{L}_{jm} é a intensidade populacional local do grupo m na localidade j e \tilde{L}_j é a

intensidade populacional local na localidade j . O índice $\check{D}(m)$ varia entre 0 e 1, onde 0 equivale ao mínimo grau de segregação e 1 representa o máximo grau.

3.1.2 Índices Espaciais de Exposição e Isolamento

O índice espacial de exposição $\check{P}_{(m,n)}^*$ e o índice espacial de isolamento \check{Q}_m são versões espaciais dos índices de exposição e isolamento propostos por Bell (1954). Estes índices capturam a dimensão espacial exposição/isolamento. Dado dois grupos populacionais, o índice espacial de exposição do grupo m ao grupo n $\check{P}_{(m,n)}^*$ mede a proporção média do grupo n na *localidade* de cada membro do grupo m :

$$\check{P}_{(m,n)}^* = \sum_{j=1}^J \frac{N_{jm}}{N_m} \left(\frac{\check{L}_{jn}}{\check{L}_j} \right), \quad (7)$$

onde N_{jm} é a população do grupo m na unidade de área j , N_m é a população do grupo m na cidade, \check{L}_{jn} é intensidade populacional local do grupo n na localidade j e \check{L}_j é a intensidade populacional local da localidade j . O índice $\check{P}_{(m,n)}^*$ expressa o contato potencial entre dois grupos populacionais e varia entre 0 (mínima exposição) e 1 (máxima exposição).

Dado um grupo populacional, o índice espacial de isolamento do grupo m - \check{Q}_m - é um caso particular do índice espacial de exposição $\check{P}_{(m,n)}^*$. O índice \check{Q}_m mede a exposição do grupo m a ele próprio e é definido como:

$$\check{Q}_m = \sum_{j=1}^J \frac{N_{jm}}{N_m} \left(\frac{\check{L}_{jm}}{\check{L}_j} \right) \quad (8)$$

onde \check{L}_{jm} é intensidade populacional local do grupo m na localidade j e os demais parâmetros são definidos como na Equação (7). O índice de isolamento varia entre 0 (mínimo isolamento) e 1 (máximo isolamento).

Os índices de exposição e isolamento dependem da composição populacional da cidade como um todo, ou seja, as proporções de cada grupo são relevantes. Por exemplo, o índice de exposição do grupo m ao n tende a apresentar valores mais altos se a proporção do grupo n em toda a cidade for elevada, dado que a interação entre os indivíduos do grupo n e os demais grupos se torna mais provável. Em virtude desta

característica, o índice de exposição é considerada assimétrico, ou seja, $\tilde{P}_{(m,n)}^*$ não é igual a $\tilde{P}_{(n,m)}^*$, exceto se os grupos m e n apresentarem a mesma proporção populacional.

3.1.3 Índice Espacial de Segregação Residencial

O índice espacial de segregação residencial \tilde{ISR} é uma versão espacial do índice de segregação residencial ISR (Jargowsky, 1996; Rodríguez, 2001) e captura a dimensão dispersão/agrupamento. Diferentemente dos índices apresentados até o momento, o índice espacial de segregação residencial \tilde{ISR} é uma medida baseada em variâncias e permite a utilização da distribuição original de variáveis contínuas em seu cômputo. Desta forma, o \tilde{ISR} dispensa o estabelecimento de grupos populacionais, o que é particularmente problemático em estudos socioeconômicos (quais são os grupos socioeconômicos? quantos são?), e é considerado mais apropriado em investigações baseadas em variáveis contínuas, como renda dos chefes de família.

Considerando uma variável contínua X , o \tilde{ISR} baseia-se no fato de que a variância total de X ($\tilde{\sigma}_{total}^2$) na cidade equivale à soma da variância *interna* de cada *localidade* ($\tilde{\sigma}_{intra}^2$) e da variância *entre as localidades* ($\tilde{\sigma}_{entre}^2$), ou seja:

$$\tilde{\sigma}_{total}^2 = \tilde{\sigma}_{intra}^2 + \tilde{\sigma}_{entre}^2 \quad (9)$$

Quanto maior a participação da variância *entre* as distintas localidades da cidade na composição da variância total, menor será a probabilidade de interação entre os diferentes grupos populacionais. Seguindo esta lógica, a índice \tilde{ISR} equivale à proporção da variância *entre* as localidades ($\tilde{\sigma}_{entre}^2$) que contribui na composição da variância total de X nas localidades da cidade ($\tilde{\sigma}_{total}^2$):

$$\tilde{ISR} = \frac{\tilde{\sigma}_{entre}^2}{\tilde{\sigma}_{total}^2}. \quad (10)$$

A variância de X entre as diferentes localidades da cidade é definida como:

$$\tilde{\sigma}_{entre}^2 = \frac{\sum_{j=1}^J \tilde{L}_j (\tilde{X}_j^2 - \tilde{X}^2)}{\sum_{j=1}^J \tilde{L}_j}, \quad (11)$$

onde

$$\bar{\bar{X}}_j = \sum_{m=1}^M \bar{\tau}_{jm} X_m \quad \text{e} \quad \bar{\bar{X}} = \frac{\sum_{j=1}^J (\bar{L}_j \bar{\bar{X}}_j)}{\sum_{j=1}^J \bar{L}_j} . \quad (12)(13)$$

Nas Equações (11) e (13), J é o número total de unidades de área na cidade, \bar{L}_j é a intensidade populacional local da localidade j , $\bar{\bar{X}}_j$ é a média ponderada de X considerando a proporção de todos os grupos na localidade j e $\bar{\bar{X}}$ é a média de $\bar{\bar{X}}_j$ na cidade, ponderada pela intensidade populacional local das localidades. Na Equação (12), $\bar{\tau}_{jm}$ é proporção do grupo m na localidade j , X_m é o valor de X para o grupo m e M é o número de grupos populacionais na cidade.

A variância total da cidade, considerando as diferentes localidades, é definida como:

$$\bar{\sigma}_{total}^2 = \sum_{m=1}^M \bar{\tau}_m (X_m - \bar{\bar{X}})^2, \quad (14)$$

onde $\bar{\tau}_m$ é a proporção do grupo m na cidade, considerando a intensidade populacional local de todas as localidades. Assim como os demais índices, o \bar{ISR} varia entre 0 e 1: o valor 0 indica o grau mínimo de segregação e o valor 1 indica o máximo grau.

3.2 Índices Locais de Segregação

O cômputo dos índices espaciais apresentados na Seção anterior resulta em valores que sintetizam o grau de segregação da área de estudo como um todo. No entanto, a segregação urbana é um fenômeno cuja intensidade varia ao longo da cidade e as informações a respeito destas variações são de grande valia no processo de compreensão da segregação e de estabelecimento de políticas públicas que a minimizem.

Por conseguinte, os índices globais devem ser complementados por índices locais de segregação, que possibilitam a observação do grau de segregação nos distintos pontos da cidade por intermédio de “mapas de segregação”.

Em resposta a esta questão, esta Seção apresenta índices locais de segregação gerados a partir da decomposição dos índices espaciais globais $\bar{D}(m)$, $\bar{P}_{(n,m)}^*$ e \bar{Q}_m . Estes

índices locais indicam o quanto cada localidade contribui para a medida global da segregação da cidade. A versão local do índice espacial de dissimilaridade generalizado $\check{D}(m)$, denominada $\check{d}_j(m)$, é definida como:

$$\check{d}_j(m) = \sum_{m=1}^M \frac{N_j |\check{\tau}_{jm} - \tau_m|}{2N\tau_m(1-\tau_m)}, \quad (15)$$

onde os parâmetros são definidos como os da Equação (4).

Seguindo o mesmo raciocínio, o índice local de exposição - $\check{p}_{j(m,n)}^*$ - é obtido da seguinte maneira:

$$\check{p}_{j(m,n)}^* = \frac{N_{im}}{N_m} \left(\frac{\check{L}_{jn}}{\check{L}_j} \right), \quad (16)$$

onde os parâmetros são definidos como os da Equação (7).

Para a obtenção do índice local de isolamento $\check{q}_{j,m}$, basta substituir \check{L}_{jn} pela intensidade populacional local do próprio grupo m na área j (\check{L}_{jm}). O \check{ISR} , ao contrário dos demais índices espaciais, não permite a geração de índices locais a partir da abordagem de decomposição apresentada.

4 Validação dos Índices Espaciais de Segregação Urbana

Após o cômputo dos índices, é importante determinar o significado dos resultados obtidos. Uma alternativa muito utilizada neste processo é a adoção de limiares fixos para a interpretação dos resultados, assim como os estabelecidos por Massey e Denton (1987) para o índice de dissimilaridade D . De acordo com estes autores, um índice D com valores entre 0 e 0,3 indica uma segregação suave, de 0,3 a 0,6 indica uma segregação moderada e acima de 0,6 revela uma segregação severa.

No entanto, todos os índices são susceptíveis a uma interferência conhecida como “problema da unidade de área modificável” ou MAUP (Openshaw, 1984). Ou seja, os resultados dos índices variam com a alteração das fronteiras das unidades de análises consideradas em seu cômputo. Por exemplo, estes índices tendem a exibir valores mais elevados com o aumento do número de unidades de área computadas. Em virtude dos efeitos do MAUP, a comparação entre índices computados para unidades

de área com geometrias distintas não é aconselhável. Da mesma forma, o estabelecimento de limiares fixos não representa uma solução adequada para a interpretação dos índices.

Como contribuição a esta problemática, propomos a aplicação de um teste de hipótese, conhecido como *teste de pseudo-significância* (Anselin, 1995), para o estabelecimento da validade estatística dos índices espaciais de segregação. Para que se saiba o que esperar quando uma determinada hipótese é verdadeira, os testes de hipótese são geralmente conduzidos para uma hipótese contrária ao que de fato se quer provar, denominada *hipótese nula* (Freund e Perles, 1999). No caso dos índices de segregação, os testes deverão ser conduzidos para a hipótese nula de que a distribuição dos grupos populacionais na cidade *não* representa um arranjo segregado. Se a hipótese nula for rejeitada, aceita-se a hipótese alternativa de que o arranjo populacional *é* segregado.

Para a execução do teste de pseudo-significância, os dados correspondentes a cada unidade de área devem ser permutados aleatoriamente inúmeras vezes entre as áreas, produzindo distintos arranjos espaciais dos grupos populacionais. As permutações geradas são utilizadas na construção de uma distribuição de referência que é confrontada ao valor do índice obtido a partir dos dados originais. Se o valor observado for superior aos valores dos índices computados para os arranjos populacionais aleatórios, a hipótese nula será rejeitada.

A Figura 4 apresenta um exemplo de aplicação do teste de pseudo-significância sobre o índice de dissimilaridade $\check{D}(m)$. Neste caso, a distribuição de referência, em laranja, foi construída a partir de 99 permutações aleatórias e o valor observado está representado pela cor azul. Observa-se na Figura 4 que o valor observado é superior aos calculados para os arranjos populacionais gerados aleatoriamente. Portanto, rejeita-se a hipótese nula de que o arranjo populacional *não é* segregado.

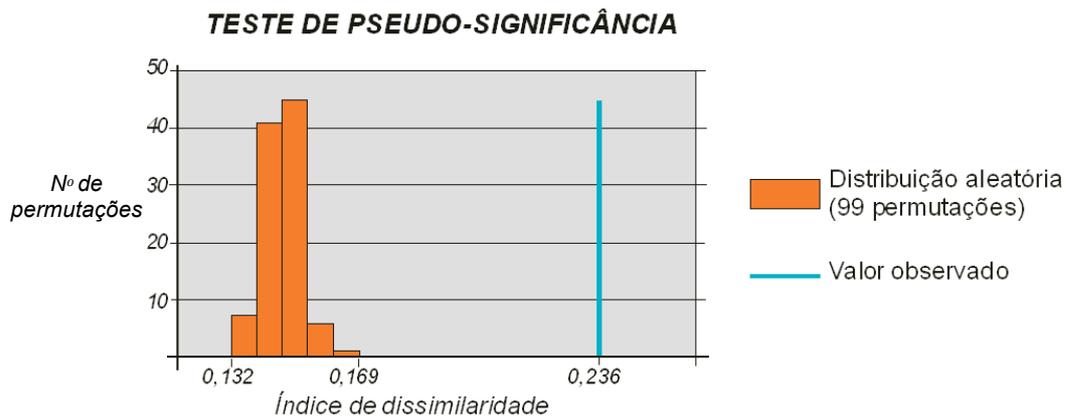


Figura 4 - Exemplo de distribuição simulada do indicador $\tilde{D}(m)$.

Na prática, a verificação da significância dos índices é efetuada através do cálculo do nível de pseudo-significância (*p-valor*) para cada índice computado. O *p-valor* representa a probabilidade de rejeitarmos a hipótese nula quando ela é verdadeira (Freund e Perles, 1999). Como o *p-valor* representa um erro, a condição para que um índice seja considerado significativo é a de que seu *p-valor* seja baixo. O valor 0,05 é comumente adotado como limiar neste processo (Freund e Perles, 1999), ou seja, um índice com *p-valor* inferior a 0,05 é considerado significativo. Este limiar, porém, pode ser alterado a critério do pesquisador.

Formalmente, o *p-valor* é calculado da seguinte maneira (Anselin, 2003):

$$p\text{-valor} = \frac{d + 1}{D + 1}, \quad (17)$$

onde *d* representa o número de permutações aleatórias cujo índice de segregação é superior ou igual ao índice calculado para o conjunto de dados observado, e *D* representa o número total de permutações geradas. Por representar uma probabilidade, o *p-valor* varia entre 0 e 1.

A aplicabilidade dos testes de pseudo-significância será apresentada na próxima Seção, na qual as propriedades dos índices apresentados serão ilustradas sobre dados artificiais e reais.

5 Aplicação e Interpretação dos Índices Espaciais de Segregação

Para que a utilização de índices subsidie análises qualificadas a respeito do fenômeno da segregação urbana, é necessário que se tenha consciência sobre as propriedades dos

índices selecionados, como devem ser aplicados e o que são capazes de revelar ou não. Por esta razão, destacamos nesta seção alguns aspectos básicos dos índices de segregação apresentados e os ilustramos com alguns exemplos.

5.1 Incorporação de Informações Espaciais

O primeiro aspecto, que representa a diferença essencial entre os índices apresentados e os comumente utilizados em estudos empíricos, diz respeito à capacidade dos índices de *incorporar informações espaciais* em suas formulações. A Figura 5 demonstra a diferença entre os índices espaciais globais apresentados e suas versões não-espaciais através da aplicação destes índices sobre dois conjuntos de dados artificiais: (a) o conjunto A, que configura um caso de extrema segregação, no qual cada área apresenta apenas indivíduos do mesmo grupo e as áreas caracterizadas pelo mesmo grupo estão agrupadas e (b) o conjunto B, onde cada área também apresenta apenas indivíduos do mesmo grupo, porém estas áreas estão dispersas de maneira equilibrada.

Para capturar a dimensão dispersão/agrupamento, foram calculados índices não-espaciais e espaciais de dissimilaridade ($D(m)$ e $\check{D}(m)$) e índices de segregação residencial (ISR e \check{ISR}). Estes índices variam entre 0 (mínimo agrupamento) e 1 (máximo agrupamento). Em relação à dimensão exposição/isolamento, foram calculados índices não-espaciais e espaciais de exposição do grupo Azul ao grupo Verde ($P_{(Azul,Verde)}^*$ e $\check{P}_{(Azul,Verde)}^*$) e de isolamento do grupo Azul (Q_{Azul} e \check{Q}_{Azul}). Os índices da dimensão exposição/isolamento dependem da composição populacional da cidade como um todo, ou seja, é importante que se tenha em mente a proporção de cada grupo durante o processo de interpretação destes índices. No caso dos conjuntos de dados apresentados, onde a proporção dos quatro grupos é de 0,25, um arranjo populacional integrado deverá apresentar índices de exposição e isolamento sempre próximos a 0,25. Isto significa dizer que a composição populacional das localidades da cidade permite um contato equitativo entre os membros de um grupo e os membros de todos os demais grupos. Já um arranjo segregado deverá apresentar índices de exposição próximos a 0 (não há exposição entre grupos distintos) e índices de isolamento próximos a 1 (máximo isolamento).

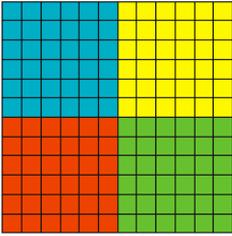
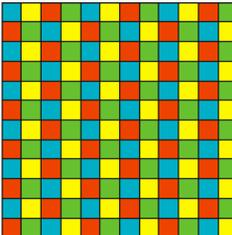
	Índices não-espaciais	Índices espaciais
<p>Conjunto A</p> 	$D(m)=1$ $ISR = 1$ $P_{(Azul,Verde)}^* = 0$ $Q_{Azul} = 1$	$\check{D}(m)=0,86$ (p-valor = 0,01) $\check{ISR}=0,82$ (p-valor=0,01) $\check{P}_{(Azul,Verde)}^* = 0,003$ (p-valor=1) $\check{Q}_{Azul} = 0,88$ (p-valor=0,01)
<p>Conjunto B</p> 	$D(m)=1$ $ISR = 1$ $P_{(Azul,Verde)}^* = 0$ $Q_{Azul} = 1$	$\check{D}(m)=0,05$ (p-valor = 1) $\check{ISR}=0,07$ (p-valor=1) $\check{P}_{(Azul,Verde)}^* = 0,24$ (p-valor=0,01) $\check{Q}_{Azul} = 0,24$ (p-valor=1)

Figura 5 - Comparação entre índices globais espaciais e não-espaciais.

Observa-se na Figura 5 que, embora o conjunto A apresente uma configuração muito mais segregada do que o conjunto B, os índices não espaciais foram incapazes de diferenciá-los. Todos os índices indicaram uma situação de máxima segregação para ambos os conjuntos. Estes resultados exemplificam o problema do tabuleiro de damas, ou seja, em qualquer caso onde as unidades de área estejam ocupadas apenas por indivíduos do mesmo grupo, os índices de segregação não-espaciais indicarão o grau máximo de segregação, independente do arranjo espacial destas unidades.

Os índices espaciais, por sua capacidade de considerar as unidades vizinhas, superam esta limitação e permitem a diferenciação entre os conjuntos A e B. Considerando a dimensão dispersão/agrupamento, os índices espaciais de dissimilaridade $\check{D}(m)$ e de segregação residencial \check{ISR} calculados sobre o conjunto A apresentaram valores elevados e significativos (p-valor=0,01). Neste caso, rejeita-se a hipótese nula de que a distribuição populacional *não* representa um arranjo segregado. Para o conjunto de dados B, os índices $\check{D}(m)$ e \check{ISR} apresentaram valores quase nulos e não significativos (p-valor=1), ou seja, aceita-se a hipótese nula de que o arranjo espacial não seja segregado.

Os índices da dimensão exposição/isolamento também apontaram o conjunto A como segregado e o conjunto B como integrado. O índice espacial de exposição do

grupo Azul ao Verde ($\tilde{P}_{(Azul,Verde)}^*$), que neste caso representou mais um índice de “integração” do que de “segregação”, apresenta um valor não-significativo para o conjunto A (p-valor=1) e um valor significativo para o conjunto B (p-valor=0,01). Este resultado significa afirmar que a hipótese nula de que o grupo Azul *não está exposto* ao grupo Verde foi aceita para o conjunto A e rejeitada para o conjunto B. O índice espacial de isolamento do grupo Azul \tilde{Q}_{Azul} revelou-se significativo para o conjunto A (p-valor=0,01) e não-significativo para o conjunto B (p-valor=1). Ou seja, a hipótese nula de que *não há isolamento* do grupo Azul foi rejeitada para o conjunto A e aceita para o conjunto B.

Os resultados dos índices espaciais globais apresentados são complementados e melhor compreendidos por intermédio de um outro aspecto que deve ser considerado no momento de aplicação de índices de segregação: a *realização de análises locais*.

5.2 Análises Locais

Embora a grande maioria dos trabalhos que utilizam índices de segregação baseie-se apenas em índices globais, é fundamental que índices capazes de capturar as variações do fenómeno ao longo da cidade também sejam adotados. Estes índices locais permitem a elaboração de mapas nos quais as áreas mais segregadas da cidade podem ser identificadas.

A Figura 6 apresenta a espacialização das versões espaciais e não-espaciais do índice local de dissimilaridade generalizado ($d_j(m)$ e $\check{d}_j(m)$) para os conjuntos de dados apresentados na Seção anterior. A espacialização destes índices locais mostra o quanto cada unidade de área contribui na composição de seus respectivos índices globais. O índice local não-espacial $d_j(m)$ não aponta qualquer variação do grau de segregação ao longo da área analisada, o que reafirma o observado por intermédio do índice global $D(m)$. Já o índice local espacial $\check{d}_j(m)$ explicita espacialmente as distinções apontadas pelo índice global $\check{D}(m)$:

- (a) no conjunto A, as unidades de áreas mais segregadas encontram-se nas bordas, enquanto as unidades mais integradas localizam-se no centro, onde os diferentes grupos estão próximos uns dos outros;

(b) no conjunto B, todas as unidades de áreas apresentam um índice $\check{d}_j(m)$ muito baixo, visto que apresentam uma vizinhança diversificada.

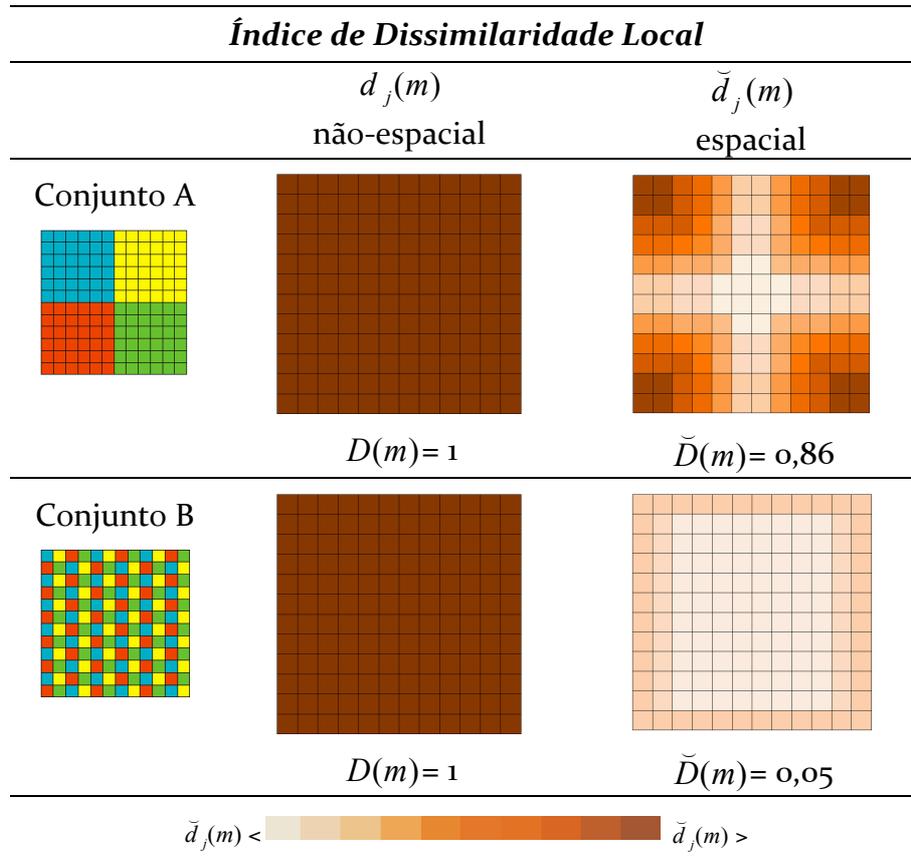
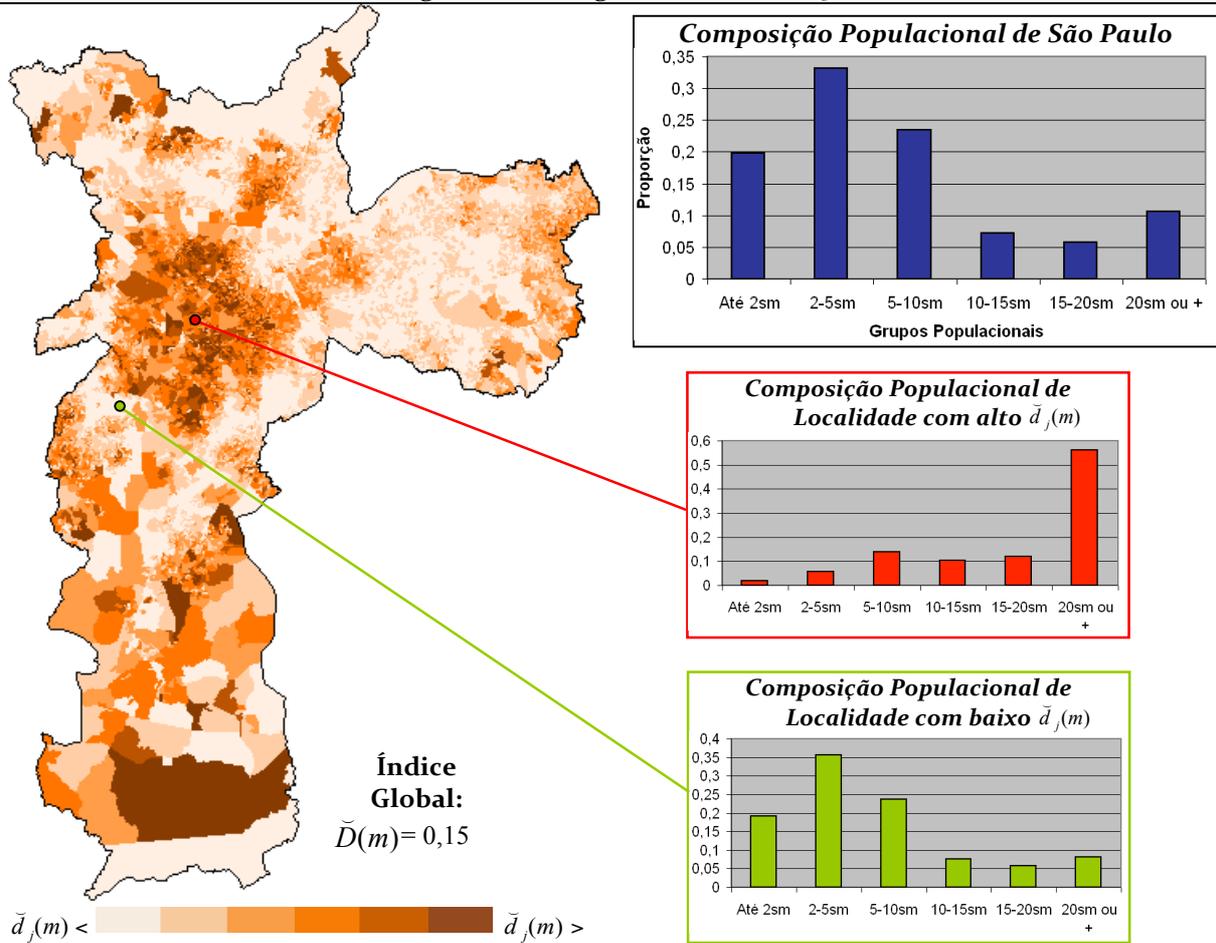


Figura 6 - Cálculo dos índices locais $d_j(m)$ (não-espacial) e $\check{d}_j(m)$ (espacial) sobre os conjuntos A e B.

Índice espacial de dissimilaridade local $\check{d}_j(m)$ - São Paulo (2000)

Kernel gaussiano, largura de banda de 500 m



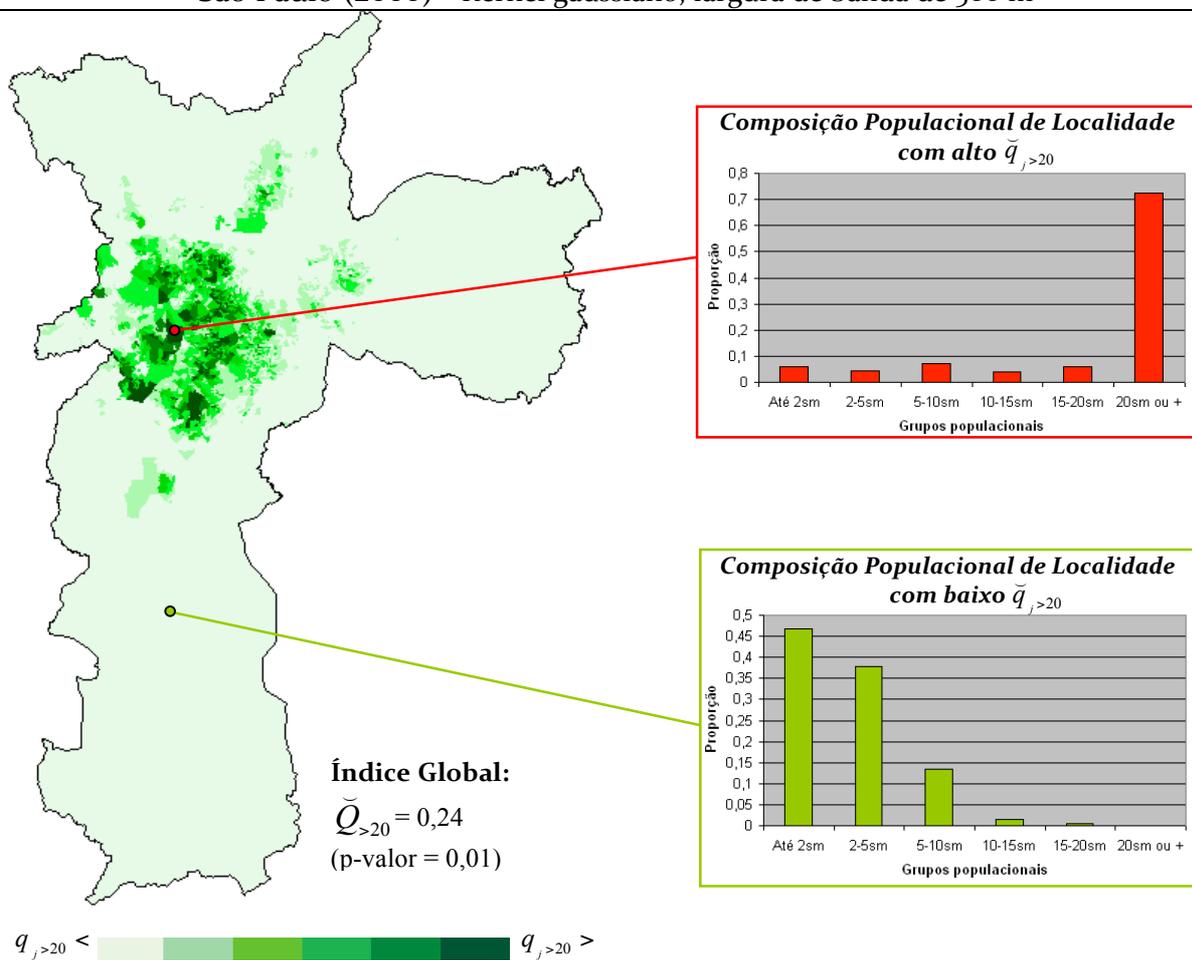
sm : salário(s) mínimo(s)

Figura 7 - Espacialização dos índices locais de dissimilaridade computados para grupos populacionais classificados quanto à renda dos chefes de família em São Paulo (2000).

A Figura 7 apresenta a espacialização do índice $\check{d}_j(m)$ para os dados censitários de renda dos chefes de família da cidade de São Paulo no ano de 2000. O índice de dissimilaridade $\check{D}(m)$ captura a dimensão dispersão/agrupamento, ou seja, o quão equilibrada é a distribuição populacional da área analisada. Por intermédio da espacialização de $\check{d}_j(m)$, é possível visualizar as localidades cuja composição populacional mais se distingue da composição populacional da cidade de São Paulo como um todo. Para ilustrar o que os mapas de dissimilaridade local ressaltam, a Figura 7 apresenta também um gráfico da composição populacional de São Paulo e gráficos da composição populacional de duas localidades - uma realçada como muito

segregada (alto $\check{d}_j(m)$) e outra como muito integrada (baixo $\check{d}_j(m)$). A comparação entre estes gráficos permite observar que a localidade de baixo $\check{d}_j(m)$ apresenta uma composição populacional muito semelhante à da cidade, enquanto a localidade de alto $\check{d}_j(m)$ dispõe de uma composição populacional muito distinta, com uma grande concentração de chefes de família pertencentes ao grupo populacional de maior renda (20 salários mínimos ou mais).

Índice espacial de isolamento dos chefes de família com renda superior a 20 sm $\check{q}_{j>20}$
 São Paulo (2000) - Kernel gaussiano, largura de banda de 500 m



sm : salário(s) mínimo(s)

Figura 8 - Espacialização dos índices locais de isolamento dos chefes de família com renda superior a 20 sm em São Paulo (2000).

A Figura 8 apresenta a espacialização do índice local de isolamento $\check{q}_{j>20}$ dos chefes de família com renda superior a 20 salários mínimos (sm) no município de São

Paulo. O índice global de isolamento $\check{Q}_{>20}$ mede a dispersão exposição/isolamento, mais especificamente, a proporção média do grupo dos chefes com renda superior a 20 sm na localidade de cada chefe de família na mesma faixa de renda. A espacialização do índice local $\check{q}_{j>20}$ indica o quanto cada localidade contribuiu para a composição do índice $\check{Q}_{>20}$. A Figura 8 apresenta o gráfico de localidades com alto e baixo índice local $\check{q}_{j>20}$. Na localidade com alto $\check{q}_{j>20}$, os chefes de família com renda superior a 20 sm representam a grande maioria, enquanto a localidade com baixo $\check{q}_{j>20}$ não apresenta nenhum indivíduo pertencente a este grupo.

5.3 Análises em Múltiplas Escalas

Ao considerarmos o fato de que a representação da segregação urbana através de índices consiste na análise da composição populacional dos locais que compõe a área de estudo, a questão da escala torna-se fundamental para qualquer estudo sobre o fenômeno. Sob esta perspectiva, os tradicionais índices de segregação tornam-se limitados, visto que sua capacidade de capturar distintas escalas restringe-se à adoção de unidades de áreas diferentes, como setores censitários e distritos. No entanto, estas unidades pré-estabelecidas não correspondem necessariamente às escalas mais apropriadas para a captura dos padrões de segregação das cidades.

Os índices apresentados neste capítulo respondem a esta limitação por permitirem a escolha dos parâmetros dos estimadores kernel que definem as localidades a serem consideradas no cômputo. Em virtude desta flexibilidade, os índices são capazes de partir da informação mais detalhada possível (geralmente fornecida por setores censitários) e generalizá-la para a análise da segregação em escalas mais abrangentes.

Para ilustrar esta questão, kernels gaussianos com larguras de banda de 500m e 1500m foram adotadas no cálculo dos índices locais $\check{d}_j(m)$ para os dados de renda dos chefes de família da cidade de São Paulo (Figura 9).

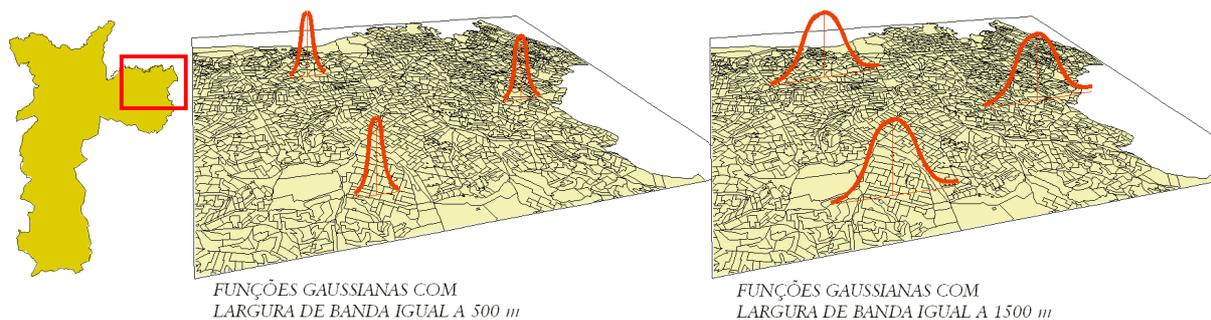


Figura 9 - Abrangência aproximada das funções gaussianas utilizadas no cômputo dos índices locais de dissimilaridade para os dados de São Paulo.

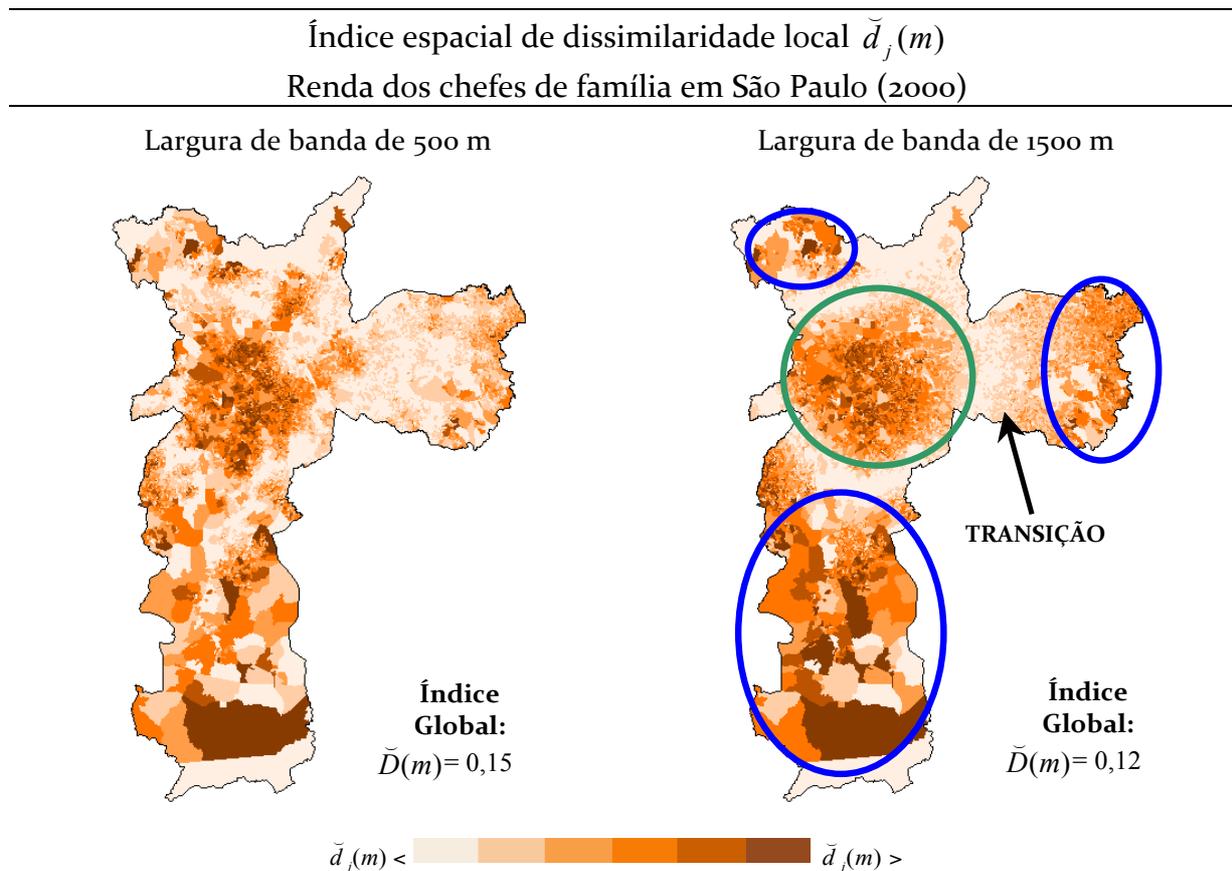


Figura 10 - Espacialização dos índices locais de dissimilaridade computados com estimadores gaussianos com larguras de banda de 500m e 1500m.

A adoção de uma largura de banda mais estreita implica no estabelecimento de localidades que serão significativamente influenciadas pelas unidades de área mais próximas, enquanto a adoção de larguras de banda maiores implica no estabelecimento de localidades mais abrangentes, que também são influenciadas pela composição populacional de unidades de área mais distantes. Por esta razão, é possível

verificar na Figura 10 que o mapa dos índices $\check{d}_j(m)$ baseado em larguras de banda de 500 m permite a detecção de maiores detalhes sobre o grau de segregação nas distintas localidades da cidade, ao passo que o mapa baseado em larguras de banda de 1500m indica grandes regiões segregadas da cidade. Neste último mapa é possível verificar, por exemplo, que São Paulo apresenta um grande núcleo segregado na região Central (circulado em verde), cercado por uma área de baixa segregação que representa uma transição para outros núcleos de segregação que se localizam nas áreas periféricas da cidade (em azul).

Com o aumento da largura de banda, os índices de segregação tendem a apresentar valores menores. No exemplo de São Paulo, o índice espacial global de dissimilaridade $\check{D}(m)$ calculado para uma largura de banda de 500m é igual a 0,15, enquanto o calculado para uma largura de 1500m é de 0,12. Este resultado é esperado, visto que com o aumento da largura de banda, as localidades incorporam as características populacionais de um número maior de unidades de área, o que geralmente implica em uma composição populacional mais diversificada.

5.4 Análises de Distintas Dimensões da Segregação

É consenso entre os pesquisadores o fato de que a segregação urbana não possui uma única dimensão e sua representação requer a aplicação de diferentes índices para cada dimensão considerada (Massey e Denton, 1988; Reardon e O'Sullivan, 2004). Conforme já mencionado, este trabalho apresenta índices de segregação para as dimensões espaciais dispersão/agrupamento e exposição/isolamento, estipuladas por Reardon e O'Sullivan (2004). Estas dimensões apontam aspectos distintos da segregação e os índices referentes a elas devem ser aplicados complementarmente. Enquanto os índices que capturam a dimensão dispersão/agrupamento medem o equilíbrio da composição populacional da cidade, os índices da dimensão exposição/isolamento nos fornecem indícios a respeito dos grupos populacionais que promovem possíveis desequilíbrios constatados na primeira dimensão.

A Figura 11 ilustra a complementaridade dos índices calculados para as duas dimensões. Os índices locais de dissimilaridade (dimensão dispersão/agrupamento) mostram as áreas onde a composição populacional mais se diferencia da composição média da cidade, enquanto os índices de isolamento (dimensão exposição/isolamento)

ressaltam quais os grupos predominam nas áreas indicadas como segregadas pelos índices de dissimilaridade. É possível observar que no núcleo segregado localizado na área central predomina o isolamento dos chefes de família com rendimentos mais elevados, enquanto os núcleos segregados periféricos são caracterizados pelo isolamento dos chefes com rendimentos inferiores.

Espacialização de índices espaciais locais de segregação da cidade de São Paulo (2000)

Função gaussiana, largura de banda 500 m

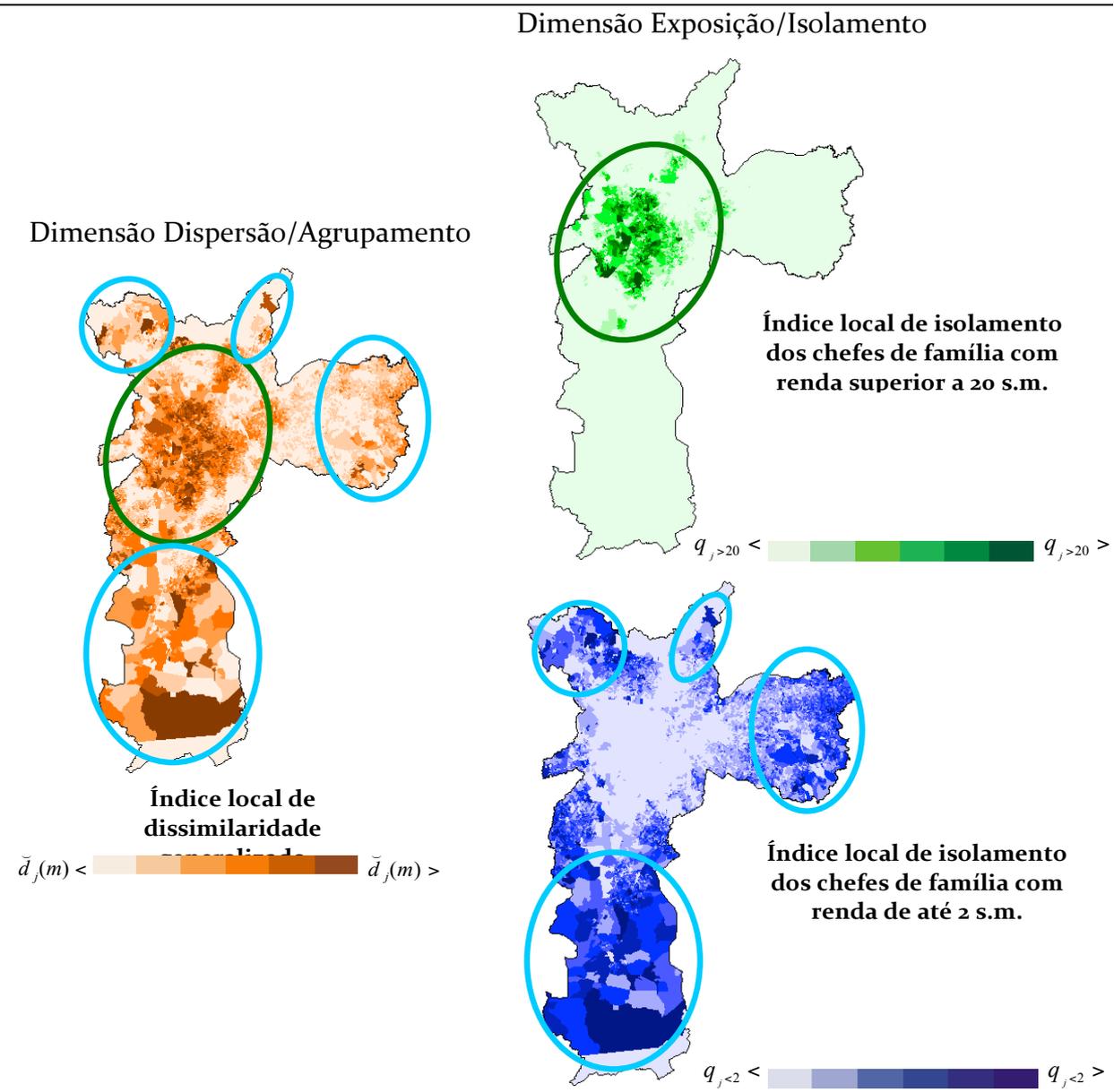


Figura 11 – Complementaridade dos índices das dimensões dispersão/agrupamento e exposição/isolamento.

6 Conclusões

Índices de segregação urbana são ferramentas importantes para a compreensão de padrões e tendências da segregação. No entanto, a maioria dos trabalhos sobre segregação permanece alheia a este recurso ou baseia-se na aplicação de índices muito simples, que são incapazes de capturar o fenômeno de maneira apropriada. Este capítulo apresenta índices de segregação alternativos que buscam minimizar as limitações dos índices mais populares. Para tanto, índices globais de segregação que incorporam informações geográficas em suas formulações foram construídos a partir da extensão de índices existentes.

A abordagem proposta para a construção destas medidas espaciais baseia-se nos conceitos de *localidade* e *intensidade populacional local*, o que possibilita a utilização de qualquer estimador kernel no computo dos índices. Esta flexibilidade permite a adoção de diferentes conceitos de *interação* entre grupos populacionais, assim como análises em distintas escalas. Considerando o fato de que as duas dimensões de segregação postuladas por Reardon e O'Sullivan (2004) contribuem de maneira diferenciada para a compreensão deste fenômeno, foram construídos índices espaciais para cada uma delas. Cabe salientar ainda que abordagem utilizada é genérica e pode ser adotada na extensão de outros índices não-espaciais.

O Capítulo também apresenta índices locais de segregação, obtidos a partir da decomposição dos índices espaciais globais. Estes índices podem ser espacializados na forma de mapas e permitem a visualização da intensidade da segregação em distintos locais da cidade. Este Capítulo recomenda ainda o uso de testes de pseudo-significância para a validação dos índices espaciais de segregação. Embora este teste não suporte afirmações sobre a intensidade da segregação em determinada área de estudo, possibilita uma maneira confiável de testar se um arranjo populacional é segregado ou não. Os testes de pseudo-significância também podem ser aplicados sobre índices locais, o que permitiria identificar as áreas da cidade que apresentam níveis significativos de segregação.

Referências Bibliográficas

ANSELIN, L. Local Indicators of Spatial Association - LISA. **Geographical Analysts**, v. 27, n.2, p. 93-115, 1995.

- ANSELIN, L. **GeoDa 0.9 User's Guide**. Urbana-Champaign: University of Illinois, 2003. 82 p.
- ATKINSON, A. B. On the measures of inequality. **Journal of Economic Theory**, v. 2, n.3, p. 244-263, 1970.
- BELL, W. A Probability Model for the Measurement of Ecological Segregation. **Social Forces**, v. 32, p. 337-364, 1954.
- BOISSO, D.; HAYES, K.; HIRSCHBERG, J.; SILBER, J. Occupational segregation in the multidimensional case. **Journal of Econometrics**, v. 61, n.1, p. 161-171, 1994.
- CALDEIRA, T. **Cidade de Muros. Crime, segregação e cidadania em São Paulo**. São Paulo: Ed. 34/ Edusp, 2000. 399 p.
- COWELL, F. A. **Measuring Inequality**. Oxford: Philip Allan, 1977. 193 p.
- DUNCAN, O. D.; DUNCAN, B. A methodological analysis of segregation indexes. **American Sociological Review**, v. 20, n.3, p. 210-217, 1955.
- FREUND, J. E.; PERLES, B. M. **Statistics: A First Course**. New Jersey: Prentice Hall, 1999. 532 p.
- JAKUBS, J. F. A distance-based segregation index. **Journal of Socio-Economic Planning Sciences**, v. 61, p. 129-136, 1981.
- JARGOWSKY, P. A. Take the money and run: Economic segregation in U.S. metropolitan areas. **American Journal of Sociology**, v. 61, n.6, p. 984-999, 1996.
- LUCO, C. A.; RODRÍGUEZ, J., 2003, **Segregación residencial en áreas metropolitanas de América Latina: magnitud, características, evolución e implicaciones de política**, in UNIDAS, N., ed., Población y desarrollo, Santiago de Chile, Proyecto Regional de Población CELADE/Fondo de Población de las Naciones Unidas (UNFPA).
- MASSEY, D. S.; DENTON, N. A. Trends in the Residential Segregation of Blacks, Hispanics, and Asians: 1970-1980. **American Sociological Review**, v. 52, n.6, p. 802-824, 1987.
- MASSEY, D. S.; DENTON, N. A. The dimensions of residential segregation. **Social Forces**, v. 67, n.2, p. 281-315, 1988.
- MORGAN, B. S. A distance-decay interaction index to measure residential segregation. **Area**, v. 15, p. 211-216, 1983.
- MORRIL, R. L. On the measure of spatial segregation. **Geography Research Forum**, v. 11, p. 25-36, 1991.
- OPENSHAW, S. **The Modifiable Areal Unit Problem**. Norwich, United Kingdom: Geo Books, 1984.
- RAMOS, F. R. **Análise Espacial de Estruturas Intra-Urbanas: O Caso de São Paulo**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 2002. Dissertação de Mestrado, Mestrado em Sensoriamento Remoto, 2002.
- REARDON, S.; O'SULLIVAN, D. Measures of Spatial Segregation. **Sociological Methodology**, v. 34, n.1, p. 121-162, 2004.
- REARDON, S. F.; FIREBAUGH, G. Measures of multigroup segregation. **Sociological Methodology**, v. 32, n.1, p. 33-67, 2002.
- RODRÍGUEZ, J., 2001, **Segregación residencial socioeconómica: que és?, cómo de mide?, que está pasando?, importa?**, Santiago de Chile, Centro Latinoamericano y Caribeño de Demografía (CELADE), Division de Población.

- Proyecto Regional de Población CELADE - FNUAP (Fondo de Población de las Naciones Unidas). p. 80.
- SABATINI, F.; CÁCERES, G.; CERDÁ, J. Segregación residencial en las principales ciudades chilenas: tendencias de las tres últimas décadas y posibles cursos de acción. **EURE (Santiago)**, v. 27, n.82, p. 21-42, 2001.
- SAKODA, J. A generalized index of dissimilarity. **Demography**, v. 18, n.2, p. 245-250, 1981.
- SCHÖLKOPF, B.; SMOLA, A. J. **Learning with Kernels : Support Vector Machines, Regularization, Optimization, and Beyond**. Cambridge, MA ; London: MIT Press, 2002. 626 p.
- SILVERMAN, B. W. **Density estimation for statistics and data analysis**. London ; New York: Chapman and Hall, 1986. 175 p.
- TELLES, E. E. Residential segregation by skin color in Brazil. **American Sociological Review**, v. 57, n.2, p. 186-198, 1992.
- TELLES, E. E. Structural sources of socioeconomic segregation in Brazil. **American Journal of Sociology**, v. 100, n.5, p. 199-223, 1995.
- THEIL, H.; FINIZZA, A. J. A note on the measurement of racial integration of schools by means of informational concepts. **Journal of Mathematical Sociology**, v. 1, n.3, p. 187-194, 1971.
- TORRES, H. D. G. Segregação Residencial e Políticas Públicas: São Paulo na década de 1990. **Revista Brasileira de Ciências Sociais**, v. 19, n.54, p. 41-56, 2004.
- WHITE, M. J. The measurement of spatial segregation. **American Journal of Sociology**, v. 88, n.4, p. 1008-1018, 1983.
- WONG, D. W. S. Spatial indices of segregation. **Urban Studies**, v. 30, n.3, p. 559-572, 1993.
- WONG, D. W. S. Measuring multiethnic spatial segregation. **Urban Geography**, v. 19, n.1, p. 77-87, 1998.