

A GEOGRAFIA DO CENSO NO BRASIL: POTENCIALIDADES E LIMITAÇÕES NA EXECUÇÃO DE ANÁLISES ESPACIAIS

Maria do Carmo Dias Bueno*

Instituto Brasileiro de Geografia e Estatística (IBGE)**

Álvaro de Oliveira D'Antona***

Faculdade de Ciências Aplicadas (UNICAMP)****

Resumo: Neste artigo são abordados os limites e potencialidades do censo demográfico brasileiro para utilização em análises espaciais a partir de considerações sobre a representação gráfica e a resolução espacial dos seus resultados. Para esta finalidade, são consideradas as unidades geográficas utilizadas para coletar e disseminar dados pelo Instituto Brasileiro de Geografia e Estatística – IBGE, assim como a possibilidade de utilização de unidades mais desagregadas e com melhor resolução espacial. Apresenta-se uma visão geral dos censos demográficos e as limitações das unidades de disseminação, como as variações espacial e temporal e aquelas decorrentes de mudanças de suporte espacial. Destacam-se as características dos censos ao longo do tempo, indicando a incorporação das geotecnologias, as quais não apenas aumentam a qualidade dos resultados que são entregues à sociedade, como também viabilizam um novo paradigma para apresentação e análise de dados censitários: a grade estatística.

Palavras-chave: Censo. Disseminação de dados. Grade estatística. Análise espacial.

BRAZILIAN CENSUS GEOGRAPHY: POSSIBILITIES AND CONSTRAINTS FOR CONDUCTING A SPATIAL ANALYSIS

Abstract: The paper presents the constraints and possibilities of the Brazilian Population Census geography for conducting a spatial analysis considering their graphical representation as well as their spatial resolution. For this purpose, we considered the geographical units used to collect and disseminate data by the Brazilian Institute of Geography and Statistics – IBGE, and also the possibility of using units more disaggregate and with higher spatial resolution. An overview of the population censuses and the restrictions of the dissemination units are presented, such as spatial and temporal changes, and those problems arising from spatial support change. The main census features over time are highlighted, showing that the integration of geo-technologies into the census operation has increased the quality of the products delivered to society, and brought up a new paradigm for data presentation and analysis: the statistical grid.

Keywords: Census geography. Data dissemination. Population grid. Spatial analysis.

*Doutora em Demografia. Coordenadora de Projetos Especiais do Centro de Documentação e Disseminação de Informações do IBGE. Contato: maria.bueno@ibge.gov.br

**Endereço: R. Gal Canabarro, 706 - Maracanã, Rio de Janeiro - RJ, 20271-205, Brasil. Tel: (21) 2142-0123

***Doutor em Ciências Sociais. Professor da Faculdade de Ciências Aplicadas, do Programa de Pós-Graduação em Demografia e do Mestrado Interdisciplinar em Ciências Humanas e Sociais Aplicadas da Universidade Estadual de Campinas. Contato: alvaro.dantona@fca.unicamp.br

****Endereço: Rua Pedro Zaccaria, 1300 - Unicamp, Campus II - Caixa Postal 1068, Limeira - SP, 13484-350, Brasil. Tel: (19) 3701-6661

Introdução

O Censo de População e Habitação¹ pode ser definido como sendo o processo de coletar, compilar, avaliar, analisar e publicar os dados estatísticos, econômicos e sociais, relativos a todas as habitações e seus residentes, em um país ou região delimitada de um país (Nações Unidas, 2008).

Mesmo tendo características espaciais desde a sua origem, os dados dos censos são utilizados por diferentes áreas do conhecimento de distintas maneiras: ora como expressões do espaço, fazendo correspondência a uma determinada porção do território; ora sem considerar esta correspondência espacial, como nos casos das abordagens sociais sem significação espacial; ora como um índice semiótico, nos casos em que se utilizam as unidades político-administrativas (unidades de federação ou município) sem realmente pensá-las espacialmente.

Os avanços tecnológicos ocorridos nas duas últimas décadas levaram a um aumento da capacidade de processamento dos computadores, ao surgimento de novas tecnologias para armazenamento, à queda dos custos de hardwares e softwares e a um aumento significativo de dados espaciais, desde imagens de sensores remotos até dados de posicionamento coletados por uma grande diversidade de dispositivos portáteis. Todos esses fatores aumentaram o interesse pela análise espacial por profissionais de áreas que tradicionalmente não investiam nesta ferramenta, como, por exemplo, as Ciências Humanas e Sociais (Wachter, 2005). A análise espacial, tradicionalmente uma ferramenta utilizada principalmente por cartógrafos e geógrafos, atualmente está integrada a todas as áreas do conhecimento que utilizam a análise aplicada, a estatística e a modelagem para compreender os fenômenos e, então, propor atuações diretas sobre o espaço (Rosa, 2011).

Face ao exposto, este trabalho tem como objetivo apresentar uma breve discussão dos limites e potencialidades dos dados do censo demográfico brasileiro para análises demográficas geográficas e espaciais. A partir de uma visão geral dos censos e das suas unidades geográficas de coleta e de disseminação, são abordados dois aspectos distintos: um referente à representação e à resolução espacial dos dados demográficos e outro relativo às perspectivas para a utilização de dados mais desagregados espacialmente.

Com relação à representação e à resolução espacial, com base nas especificidades do censo brasileiro, são exploradas as limitações associadas à utilização das unidades de disseminação em análises espaciais demográficas, inclusive aquelas inerentes à própria

agregação de dados, que podem, inclusive, influenciar nos resultados obtidos.

Quanto às perspectivas futuras, graças ao uso crescente de geotecnologias por parte do Instituto Brasileiro de Geografia e Estatística – IBGE, mostra-se uma abordagem complementar que tem sido cada vez mais empregada pelos institutos de estatística de vários países, principalmente com o objetivo de integração entre geografia e estatística – a grade estatística.

Censos demográficos no Brasil

Uma visão geral dos censos

A partir da criação do IBGE em 1936, este passou a ser responsável pela realização dos censos demográficos no País (IBGE, 2010). O censo demográfico é uma operação estatística complexa, devido principalmente à sua abrangência, uma vez que investiga todos os domicílios do País. No Brasil, com uma área territorial de mais de 8,5 milhões de km², onde em 2010 se encontravam cerca de 60 milhões de domicílios, esta operação toma proporções gigantescas e complexas.

O questionário do censo procura refletir as preocupações e políticas correntes na época em que a pesquisa censitária é realizada, sempre buscando captar as mudanças que estão ocorrendo na sociedade. Os dois primeiros censos realizados no País, em 1872 e 1900, além de enumerarem a população se detinham em aspectos da divisão entre homens livres e escravos e na presença de estrangeiros no País; a partir do censo de 1920 foram incorporadas perguntas para possibilitar uma melhor caracterização da população, principalmente no que diz respeito à divisão social do trabalho; a partir do censo de 1940, então já sob a responsabilidade do IBGE, foram acrescentadas ao questionário questões referentes aos domicílios e relacionadas com o padrão habitacional (Oliveira, 2003).

As técnicas de investigação também são objeto de inovação, como no censo de 1960, momento em que a técnica de amostragem foi adotada pela primeira vez (Gonçalves, 1995). Com esta técnica, que permanece até hoje, são aplicados dois questionários: um curto, aplicado à totalidade de domicílios, e um longo, aplicado em uma amostra de domicílios, que permite investigar de forma mais detalhada algumas características da população. Uma das vantagens da amostragem é permitir um maior número de perguntas sem aumentar exageradamente o tempo de coleta e, conseqüentemente, o custo da operação.

Com relação ao mapeamento censitário, que pode ser definido como o conjunto de atividades necessárias

¹ A Organização das Nações Unidas denomina a operação sistemática de aquisição e disseminação de informações de pessoas e domicílios como Censo de População e Habitação (Nações Unidas, 2008). No Brasil esta operação é conhecida como Censo Demográfico (IBGE, 2010).

para realizar a divisão integral do País em unidades que permitam planejamento, coleta, tratamento e disseminação de informações estatísticas, também houve inovações. Inicialmente, os mapas utilizados nos recenseamentos não passavam de croquis desenhados à mão em papel. No censo de 1960, eram utilizados os Mapas Municipais oriundos da ampliação dos originais da Enciclopédia dos Municípios Brasileiros; no censo de 1970, pouca coisa mudou, havendo apenas uma atualização dos mapas utilizados no censo anterior; em 1980, passou-se a utilizar bases aerofotogramétricas reambuladas e desenhadas em papel vegetal (IBGE, 1981). Já no censo de 1991, o mapeamento censitário se apresentava mais estruturado, mas ainda em meio analógico; no censo de 2000, todo o mapeamento censitário foi digitalizado, mas encontrava-se dividido em vertente rural e vertente urbana, devido principalmente aos insumos utilizados para a sua elaboração (Barbuda, 2004); e no censo demográfico mais recente, em 2010, a base foi estruturada em uma arquitetura de banco de dados espacial, as duas vertentes foram integradas e foi feita a integração com o Cadastro de Endereços² (IBGE, 2013).

Na linha do tempo na Figura 1, podemos perceber como os censos brasileiros evoluíram ao longo do tempo, incorporando novas técnicas e tecnologias para melhorar a qualidade dos resultados disponibilizados. Da perspectiva dos usuários, entretanto, é necessário um conhecimento mais amplo das unidades geográficas que estão envolvidas na operação censitária para que se tenha consciência não apenas dos potenciais, mas também das limitações existentes.

Unidades geográficas de coleta

As unidades básicas de coleta dos censos demográficos são os indivíduos e os domicílios. A representação espacial dessas unidades é feita por objetos discretos vetoriais representados por pontos, através do uso de ferramentas de geoprocessamento. No caso dos domicílios, a representação é direta, com a utilização das coordenadas geográficas que melhor representam espacialmente o domicílio; no caso dos indivíduos, a representação é indireta, sendo utilizada a representação do domicílio como uma aproximação para representar as pessoas que residem no seu interior.

Para a realização dos censos faz-se necessária a delimitação de unidades geográficas para a coleta de dados, que são conhecidas como áreas de enumeração ou setores censitários. Essas áreas são atribuídas a um recenseador, que tem a tarefa de realizar a coleta

de dados em todas as edificações domiciliares ali incluídas, em um determinado período de tempo (IBGE, 2003). O critério utilizado para a delimitação dos setores censitários é a quantidade de domicílios existentes e que o recenseador é capaz de visitar no período de tempo destinado para a execução da coleta de dados, que é de aproximadamente um mês. Nas áreas urbanas, essa quantidade varia de 250 a 350 domicílios, enquanto nas áreas rurais esse valor é menor, variando entre 150 a 250 domicílios. Nas áreas rurais, além da quantidade de domicílios acrescenta-se o critério da extensão do setor censitário, sendo que um setor idealmente não deve ter mais do que 500 km² (IBGE, 2013).

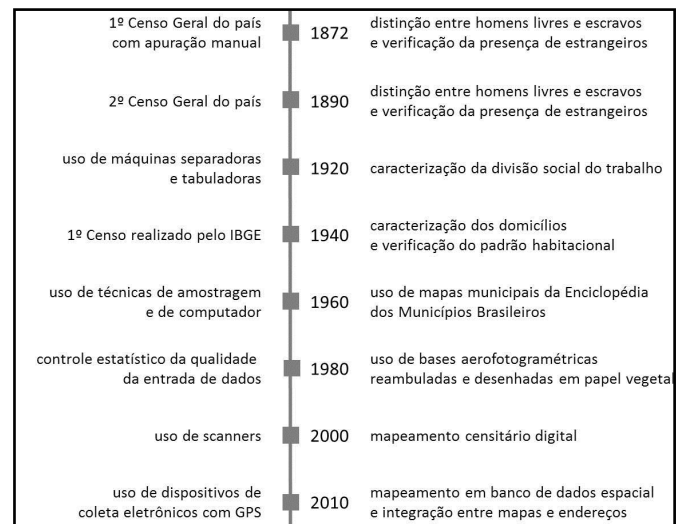


Figura 1 – Linha do tempo do censo demográfico no Brasil.
Fonte: elaborado pelos autores.

No caso das áreas urbanas, no interior de cada uma dessas áreas de enumeração existe uma malha de logradouros cujos trechos – as faces – formam quadras ou quarteirões. Nesses trechos de logradouros se localizam as edificações, que são as representações físicas dos domicílios, e, então, se pode afirmar que um agrupamento linear sequencial de domicílios forma uma face e que um agrupamento de faces forma uma quadra³.

A agregação das quadras gera o primeiro nível de unidades geográficas da operação censitária, que nada mais é do que a unidade operacional de coleta, denominada setor censitário ou área de enumeração.

No caso das áreas rurais, no interior de cada área de enumeração há uma malha de vias interligando fazendas e núcleos populacionais, mas esta malha não forma um desenho que permita o seu particionamento em quadras e faces ou algo similar. Dessa maneira, não há nenhuma divisão das unidades operacionais de coleta, sendo os domicílios agregados diretamente nos setores censitários.

Os setores censitários são classificados quanto

² O Cadastro de Endereços é uma listagem de todos os endereços das unidades visitadas durante a operação censitária e, que com o tempo, evoluiu para um cadastro digital integrado com o mapeamento censitário. Ele é utilizado por todas as pesquisas domiciliares realizadas pelo IBGE e sua atualização é permanente (IBGE, 2010).

³ Esta descrição é válida apenas para setores urbanos formais, sendo que nos demais casos pode ou não haver logradouros e, conseqüentemente, pode ou não haver faces e quadras.

a sua situação territorial de acordo com a legislação municipal vigente, e nos casos em que não exista nenhuma lei, de acordo com a tipologia da ocupação humana observada naquele local (IBGE, 2003). Assim, em uma primeira divisão, os setores são classificados como sendo urbanos ou rurais. No entanto, a adoção de uma classificação mais detalhada se faz necessária para melhor representar as situações encontradas *in situ*.

A Tabela 1 mostra a classificação dos setores censitários utilizada no Brasil até o momento, com as suas respectivas quantidades de setores povoados em 2010⁴. A classificação em tipologias de rural e urbano mostrada na Tabela 1 é utilizada desde o censo de 1991 (IBGE, 2003). Até o censo de 1970 era utilizada apenas a divisão entre rural e urbano, sem detalhamentos, sendo que no censo de 1980 foi implantada a primeira tentativa de tornar esta classificação mais detalhada (IBGE, 1981). Essas classificações são revistas periodicamente para verificar se houve alguma alteração no local e, portanto, se a classe utilizada para descrever a ocupação do território deve ser mantida ou alterada.

Situação			Setor censitário (N)	
Urbana	Cidade ou vila	Área urbanizada	227.250	
		Área não urbanizada	6.191	
	Área urbana isolada	3.273		
Total urbano			236.714	
Rural	Aglomeramento rural	De extensão urbana	1.514	
		Isolado	Povoado 9.200 Núcleo 236 Outros 1.281	
	Zona rural		61.175	
	Total rural			73.406
	Total			310.120

Tabela 1 – Setores censitários povoados por situação, Brasil, 2010. Fonte: IBGE, Censo Demográfico 2010. Elaborado pelos autores.

Na situação urbana, consideram-se as áreas urbanizadas e não urbanizadas internas ao perímetro urbano de cidades ou vilas e as áreas urbanas isoladas; a situação rural abrange toda a área situada fora desses limites, incluindo as extensões das áreas urbanas e os aglomerados rurais isolados, como povoados, núcleos e outros similares (IBGE, 2008). Associada a essas classes de situação há ainda uma classificação segundo a tipologia das construções, em que são consideradas algumas características que tornam necessária a adoção

⁴Em 2010, o número total de setores censitários do País foi de 317.380, sendo que 7.260 correspondem a áreas não habitadas cuja delimitação foi realizada exclusivamente para atendimento a requisitos operacionais.

de um tratamento diferenciado na coleta. Dessa maneira, os setores que englobam aglomerados subnormais, quartéis, alojamentos, embarcações, aldeias indígenas, presídios, asilos, orfanatos e outros são classificados como setores especiais (IBGE, 2008).

Unidades geográficas de disseminação

As unidades de coleta são a origem de um sistema hierárquico de unidades que são utilizadas para divulgar os dados censitários. Realizando operações de agregação sequenciais, cada nível dará origem a um nível superior, gerando a estrutura hierárquica geográfica completa do censo. Na prática, os setores censitários são criados de maneira que um conjunto dessas unidades operacionais ocupe completamente a superfície do menor nível administrativo municipal, permitindo, assim, a criação de uma malha hierárquica aninhada. A Figura 2 ilustra a hierarquia utilizada no censo demográfico no Brasil em 2010.

O primeiro nível da hierarquia, considerando-se o censo de população, é formado pelas pessoas que, quando agregadas, formam o nível imediatamente superior, o nível dos domicílios. No caso do censo de habitação, o primeiro nível é o de domicílios. Esses dois níveis não são passíveis de divulgação de dados, uma vez que seus dados fornecem características individualizadas e sua disseminação é vedada por legislação (Bianchini, 2012).

O próximo nível é o das faces de quadra, que são os trechos de logradouros onde as edificações dos domicílios estão alojadas. A seguir, temos o nível dos quarteirões ou quadras, que é o agrupamento de geralmente quatro faces, mas que pode ser no mínimo três faces e, teoricamente, sem quantidade máxima. Esses dois níveis são passíveis de disseminação, pois os dados já se encontram agregados, mas, embora essa possibilidade tenha sido cogitada, nunca chegou a ser implementada.

O nível seguinte é o dos setores censitários, que é a unidade utilizada para planejar e controlar a operação de coleta de dados. No Brasil, o setor censitário é a menor unidade geográfica utilizada para disseminar os dados do censo. Nem todos os países disseminam seus dados em unidades equivalentes aos nossos setores censitários. Geralmente a disseminação é feita apenas em unidades político-administrativas, como em Moçambique (INE Moçambique, s./d.) e Cabo Verde (INE Cabo Verde, s./d.), por exemplo, ou em unidades estatísticas especificamente criadas para este fim, como no caso da Inglaterra e do País de Gales (ONS, 2012), ou em unidades derivadas das áreas de enumeração, mas de maior extensão, como é o caso dos Estados Unidos, do Canadá e da Austrália (Nações Unidas, 2007).

As áreas de ponderação são unidades estatísticas

compostas por agrupamentos de setores censitários e têm o objetivo de disseminar os dados da amostra do censo demográfico. Como uma parte dos questionários foi aplicada a uma amostra da população, é necessário que se tenha uma quantidade mínima de questionários respondidos dentro de uma determinada área para que os dados possam ser expandidos para a população total daquela área com a devida confiança. Devido a esta condição, o tamanho das áreas de ponderação pode variar, sendo que em alguns municípios é possível existirem várias unidades, enquanto em outros o próprio município é a área de ponderação.

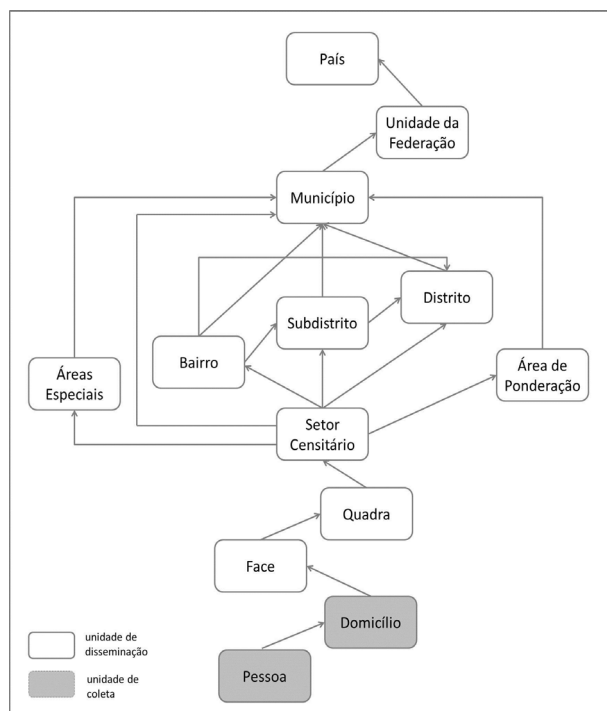


Figura 2 – Unidades de coleta e de disseminação do censo demográfico brasileiro.

Fonte: IBGE, Censo Demográfico 2010. Elaborado pelos autores.

A incorporação de geotecnologias no trabalho de diferentes órgãos do governo trouxe consigo a demanda de disseminação dos dados dos censos em unidades compatíveis com unidades específicas desses órgãos. O IBGE buscou se adaptar a essa realidade e alterar os limites de suas unidades de coleta para atender tais necessidades de disseminação para áreas especiais. Podemos citar como exemplo as áreas protegidas, como Unidades de Conservação Ambiental e Terras Indígenas, e os Aglomerados Subnormais (IBGE, 2008). No censo demográfico de 2010 essas unidades foram compostas por um ou mais setores censitários e, portanto, a totalização dos dados se tornou facilmente computável.

Com relação às divisões político-administrativas, elas são compostas por um número inteiro de setores censitários. Um setor ou um grupo de setores forma a menor divisão político-administrativa existente em determinado município, podendo ser um bairro, um distrito, um subdistrito ou até mesmo o próprio município,

nos casos em que não há divisão intramunicipal. Um conjunto destas unidades, por sua vez, forma a unidade hierarquicamente superior, e assim, até atingir o nível municipal. Grupos de municípios formam as Unidades da Federação, que, por sua vez, formam o nível político-administrativo mais alto, que é o País.

Diante da diversidade de unidades geográficas utilizadas na pesquisa censitária, as aplicações que as utilizam devem considerar as características e limitações de cada uma delas, uma vez que as mesmas podem causar dificuldades operacionais e analíticas, como apresentado no próximo item.

As limitações das unidades de disseminação

As unidades de disseminação dos censos são compostas por unidades operacionais (setores censitários), unidades estatísticas (áreas de ponderação), unidades especiais (aglomerados subnormais e áreas protegidas) e unidades político-administrativas (bairros, subdistritos, distritos, municípios, unidades da federação e país).

A utilização dessas unidades em análises espaciais é marcada por diversas limitações. Algumas dessas limitações surgem devido ao conflito entre o objetivo para o qual essas unidades geográficas foram criadas e a utilização dada pelo usuário; outras limitações surgem da necessidade de divulgar dados agregados; e outras surgem do emprego de técnicas inadequadas de representação gráfica de dados ou de técnicas de análise.

Variação espacial e temporal

Para atender as demandas dos governos por dados que possam direcionar e monitorar as políticas públicas, as informações censitárias são historicamente distribuídas em unidades coincidentes com as divisões político-administrativas (Guzman; Schensul; Zhang, 2013). No caso do Brasil, para melhorar a granularidade das informações, são utilizadas algumas unidades menores, como os setores censitários, mas a sua geografia está alinhada com os limites dessas divisões político-administrativas, conforme já indicado.

Contudo, se o objetivo da utilização dos dados não estiver diretamente relacionado com a administração territorial, essas unidades podem não ser as mais adequadas, sendo preferível utilizar critérios ambientais, físicos, gerenciais ou de acessibilidade para a delimitação das unidades geográficas. Um exemplo desta inadequação são os estudos ambientais, em que a unidade de agregação dos dados deveria possuir um

recorte mais próximo de unidades geográficas relevantes para a compreensão das relações demográficas e ambientais, como, por exemplo, as bacias hidrográficas (Hogan, 1991). Já no caso de a análise estar relacionada com a vulnerabilidade a algum evento físico, como deslizamento de encosta ou inundação, outra unidade geográfica baseada em relevo, por exemplo, seria mais indicada. Infelizmente, raramente os dados demográficos são reportados em unidades relacionadas com o meio físico ou ambiental (Deichmann, 1996). As unidades de disseminação convencionais também podem ser insuficientes naqueles casos em que outras divisões, inclusive de caráter administrativo, não coincidem com as unidades disponibilizadas, como, por exemplo, áreas de gerenciamento de saúde (Muller; Cubas; Bastos, 2010).

Segundo Judson (2007), o grande problema relacionado com dados no século XXI é que, do ponto de vista dos usuários, as agências de estatística deveriam disponibilizar os dados com maior rapidez, agregados em níveis geográficos cada vez menores e devidamente integrados ou com possibilidades de integração com outros tipos de dados. Na verdade, as agências de estatística não têm capacidade para atender essas demandas, considerando-se os recursos humanos e financeiros que hoje elas têm a sua disposição.

Um problema que decorre da utilização de unidades operacionais ou estatísticas como unidades de disseminação de dados é a dificuldade de realizar comparações temporais. Como mencionado anteriormente, os setores censitários têm o objetivo de facilitar o planejamento e o gerenciamento da operação censitária, não sendo criados objetivando a disseminação de dados ao público. Assim, de um censo para outro, essas unidades usualmente são alteradas e, na maioria dos casos, esta alteração se faz pela divisão de um setor em dois ou mais setores. Isso se dá principalmente em áreas de expansão urbana, locais em que a densidade populacional aumenta dia após dia e onde a transformação de áreas rurais em áreas urbanas ocorre frequentemente.

A consequência direta para o usuário dos dados censitários é a dificuldade de comparação temporal, já que as unidades se alteram com o passar do tempo e há a necessidade de se realizar uma compatibilização entre as unidades antigas e as novas. Isso, no entanto, não impede que a análise seja efetuada, desde que as informações de comparabilidade, que permitem que esta compatibilização seja realizada, sejam divulgadas juntamente com os dados. Na prática, mesmo havendo a divulgação das informações que permitam a comparabilidade temporal das unidades geográficas, a realização desta compatibilização é uma tarefa árdua e demorada.

As unidades político-administrativas também são passíveis de alteração ao longo do tempo, embora com

uma frequência menor, e essas alterações ocasionam dificuldades em análises temporais da mesma maneira que as demais unidades já comentadas.

Como as áreas de ponderação são formadas pela junção de setores censitários ou por unidades político-administrativas (municípios), sua alteração ao longo do tempo é consequência direta da alteração das unidades originais. Além disso, como sua delimitação depende da quantidade de questionários da amostra respondidos, à medida que uma região se torna mais densamente povoada, a tendência é que as áreas de ponderação sejam menores, assim como os setores censitários.

Mudança de suporte espacial

Toda informação geográfica está associada a um conjunto de pontos de localização que são modelados utilizando-se objetos. Esses objetos e as métricas utilizadas para medir a interação entre eles são conhecidos como “suporte espacial da informação” (OLEA, 1991). Como exemplo de suporte espacial de pontos, pode ser mencionada a localização geográfica (latitude e longitude) dos domicílios de uma pesquisa domiciliar; para suporte espacial de área, os setores censitários e as divisões municipais são exemplos adequados. A mudança de suporte de uma variável cria uma nova variável que, apesar de estar relacionada com a original, tem propriedades estatísticas e espaciais diferentes (Gotway; Young, 2002).

Todos os problemas relacionados com mudança de escala, agregação, desagregação e inferência entre unidades de suporte diferente fazem parte de um grupo de problemas que é denominado Problemas de Mudança de Suporte ou Change of Support Problem (COSP) (Gotway; Young, 2002, 2007; Plumejeaud et al., 2010). No contexto dos censos, comumente ocorrem dois tipos de mudança de suporte: de ponto (dados domiciliares e individuais) para área (dados agregados em unidades geográficas) e de área (setores censitários, por exemplo) para área (outro recorte espacial qualquer).

Os dois problemas de mudança de suporte mais conhecidos e discutidos estão relacionados com a agregação de dados individualizados em áreas: o problema da unidade de área modificável e a falácia ecológica – ambos considerados como inerentes aos dados agregados em áreas. Apesar de serem de difícil detecção e mensuração e, conseqüentemente, de difícil correção ou eliminação, eles não devem ser ignorados (Openshaw, 1984; Grasland; Madelin, 2006, Flowerdew, 2011).

Problema da unidade de área modificável

Este problema foi amplamente estudado por Openshaw (1984) e por Openshaw e Taylor (1979), sendo que este último criou o nome pelo qual ele é conhecido até hoje – Modifiable Area Unit Problem ou MAUP.

O MAUP envolve dois efeitos que ocorrem simultaneamente: o efeito de escala e o efeito de zoneamento (Wong, 1996). O efeito de escala, como o próprio nome sugere, é devido à mudança de escala, o que na prática significa que a quantidade de unidades é alterada. À medida que a escala diminui, a quantidade de unidades também diminui, já que a extensão de cada unidade aumenta. Dessa maneira, os dados originais são agregados em uma quantidade menor de unidades, levando à diminuição da variância dos dados à medida que a agregação aumenta. Isso causa um efeito de suavização, ocasionada pela perda de heterogeneidade interna e pode levar a uma distorção na interpretação dos dados.

A Figura 3 mostra dois mapas cujo tema é a densidade populacional. O primeiro mapa foi construído para as 63 microrregiões do estado de São Paulo, enquanto o segundo foi feito para as 15 mesorregiões do estado. Percebe-se que houve uma mudança de escala devido à alteração da quantidade de unidades mapeadas (de 63 para 15), o que causou uma suavização nos dados e, como consequência, surge um mapa temático diferente para a mesma região.

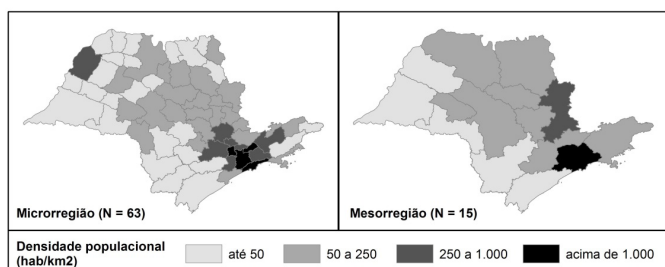


Figura 3 – Influência do efeito de escala na percepção de um fenômeno. Fonte: IBGE, Censo Demográfico 2010. Elaborado pelos autores.

O segundo efeito do MAUP, conhecido como efeito de zoneamento ou de agrupamento, ocorre numa mesma escala e, portanto, a quantidade de unidades não se modifica; a alteração acontece na forma com que as unidades são agrupadas, gerando arranjos espaciais diferentes e variâncias dos dados também diferentes.

A Figura 4 ilustra o efeito de zoneamento, mostrando as mesorregiões originais do estado de São Paulo e as mesorregiões alteradas, onde o agrupamento das microrregiões foi feito de forma diferente da original. A quantidade de unidades não foi alterada, e, portanto, não há influência da escala no fenômeno. A diferença entre os mapas temáticos se deve exclusivamente às alterações de agrupamento das microrregiões.

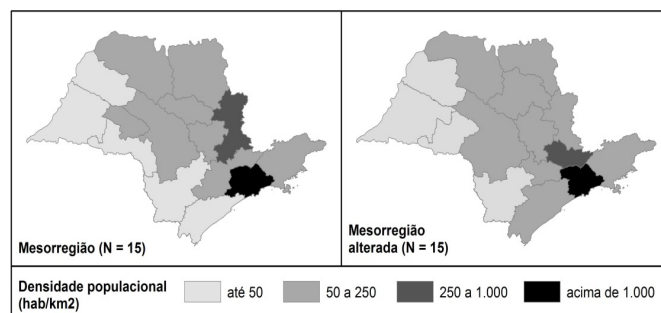


Figura 4 – Influência do efeito de zoneamento na percepção de um fenômeno. Fonte: IBGE, Censo Demográfico 2010. Elaborado pelos autores.

Na prática, os dois aspectos do MAUP interagem e ficam mais severos à medida que a quantidade de unidades diminui, ou seja, com o aumento da agregação.

De acordo com o que foi apresentado, é recomendável que as conclusões de análises que envolvam dados agregados em unidades zonais sejam cuidadosamente analisadas para verificar se os resultados são consistentes em outras escalas e outros arranjos espaciais, de modo a descartar a possibilidade de que possam ser resultantes do problema da unidade de área modificável (Fotheringham, 1998).

Falácia ecológica

A correlação entre duas variáveis é estabelecida de acordo com o nível de agregação dos dados, e por isso uma análise baseada em dados agregados pode levar a conclusões diferentes daquelas obtidas com a análise dos mesmos dados no nível individual ou em um nível diferente de agregação (Martin, 1996). Esse erro inferencial é conhecido como “falácia ecológica” e ocorre porque as análises que utilizam dados agregados em áreas não fazem distinção entre as associações espaciais criadas pelo agrupamento dos dados e as associações reais dos dados individuais existentes antes do agrupamento (Openshaw, 1984).

Por exemplo, sabe-se que o índice de suicídio em países de religião protestante é alto, mas não se pode dizer que é alta a probabilidade de pessoas protestantes se suicidarem (Freedman, 1999). Utilizando um exemplo relacionado com censo demográfico, uma análise utilizando setores censitários identificou a existência de uma correlação positiva entre as variáveis “rendimento” e “cor/raça”, de modo que a quantidade de pessoas com um baixo rendimento é grande para indivíduos que se declararam de cor preta. Essa correlação foi estabelecida no nível de setor censitário, mas não pode, no entanto, ser estendida com um alto grau de certeza para o nível individual, de modo que se possa afirmar que todos os indivíduos de cor preta têm rendimento baixo.

Variação na forma e na extensão

Devido aos critérios utilizados para a delimitação e a classificação adotada, os setores censitários são muito heterogêneos, tanto em forma quanto em extensão. Até mesmo setores que compartilham a mesma classificação variam de acordo com o local em que estão localizados, pois são influenciados diretamente pelos padrões de ocupação do território. Uma diferença marcante está no tamanho dos setores censitários de zona rural da Região Norte do País em comparação com o mesmo tipo de setor nas demais regiões. Enquanto os setores de zona rural da Região Norte têm em média 470km², os da Região Sudeste, por exemplo, têm 66km², e os da Região Sul, 44km². Esses valores podem ser explicados pela diferença na forma de ocupação do território, havendo muito mais espaços não ocupados por população na Região Norte do que nas Regiões Sudeste ou Sul. A maior quantidade de espaços sem população leva à necessidade de se ter um setor mais extenso, de modo que se possa atingir a quantidade de população média que idealmente um setor censitário deve ter.

A heterogeneidade na forma dos setores censitários também é bastante acentuada e tem origem na necessidade de se ter, sempre que possível, elementos físicos para demarcar os seus limites, bem como nos próprios critérios de delimitação. Assim, nas áreas urbanas, o traçado viário é geralmente o principal elemento delimitador dos setores, enquanto nas áreas rurais com frequência encontramos cursos d'água ou limites de propriedades na descrição do perímetro dos setores censitários.

O mesmo raciocínio é válido para as demais unidades de disseminação, como as áreas de ponderação e as unidades político-administrativas. A delimitação dessas unidades considera objetivos diversos, períodos temporais diferentes, a incorporação de elementos naturais e culturais, além de diversos outros fatores que as tornam heterogêneas quanto à forma e à extensão.

Essas diferenças na forma e na extensão das unidades de disseminação podem causar problemas na interpretação de mapas que mostram variáveis demográficas em esquemas coropléticos. Os dados demográficos podem ser expressos como números absolutos ou como taxas ou proporções, e esses valores são considerados homogêneos dentro das unidades geográficas, embora na realidade não o sejam. Os valores absolutos são dependentes da unidade geográfica para a qual são reportados, ou seja, se a unidade tiver sua área alterada, esses números também irão se alterar. Já as taxas e proporções não se alteram com mudanças na área ou perímetro da unidade geográfica, pois não dependem do tamanho da unidade. O primeiro tipo de dado é conhecido como espacialmente extensivo, enquanto o segundo é chamado de espacialmente

intensivo (Buckley, 2013). Dessa maneira, um mapa coroplético executado com os valores absolutos de uma variável censitária qualquer e com base em unidades geográficas com forma e extensão muito heterogêneas pode levar a interpretações equivocadas.

Por exemplo, comparando-se os mapas para as mesorregiões do estado de São Paulo na parte superior e inferior da Figura 5, percebe-se que o primeiro nos leva a crer que a maior parte da população se concentra em uma faixa que se inicia ao norte e vai até a porção sudeste do estado. No entanto, ao se observar o mapa tematizado com a variável densidade populacional percebe-se que a população está realmente concentrada nas mesorregiões de Campinas (9) e Metropolitana (14). Com relação aos mapas confeccionados para as microrregiões, ao observar o mapa feito com a variável população percebe-se, por exemplo, que a microrregião localizada no extremo oeste da faixa central do estado, microrregião de Presidente Prudente, tem um volume populacional maior que as suas áreas vizinhas, no entanto esta superioridade não é comprovada ao se observar o mapa de densidade populacional. Pode-se concluir, portanto, que o mais apropriado é utilizar taxas, proporções ou densidades para a criação de mapas coropléticos, já que assim os valores independem da forma e da extensão da unidade geográfica para a qual eles são reportados, minimizando a chance de mal-entendidos.

Uma dificuldade que também surge da heterogeneidade de forma e tamanho das unidades está na realização de análises espaciais que consideram a vizinhança como relacionamento espacial. A quantidade de vizinhos de uma unidade depende do tamanho e da forma dessa unidade, bem como dessas mesmas características dos seus vizinhos. Por exemplo, uma unidade pode ter muitos vizinhos, se estes forem de porte pequeno, mas pode ter poucos vizinhos, se estes forem maiores do que a unidade considerada. Consequentemente, a quantidade de informação potencialmente considerada na análise pode variar muito de acordo com a extensão e a forma das unidades que estão sendo analisadas e, nesses casos, não é aconselhável utilizar a vizinhança como forma de relacionamento espacial na análise, sendo recomendada a seleção de outra forma de interação espacial (ESRI, 2013).

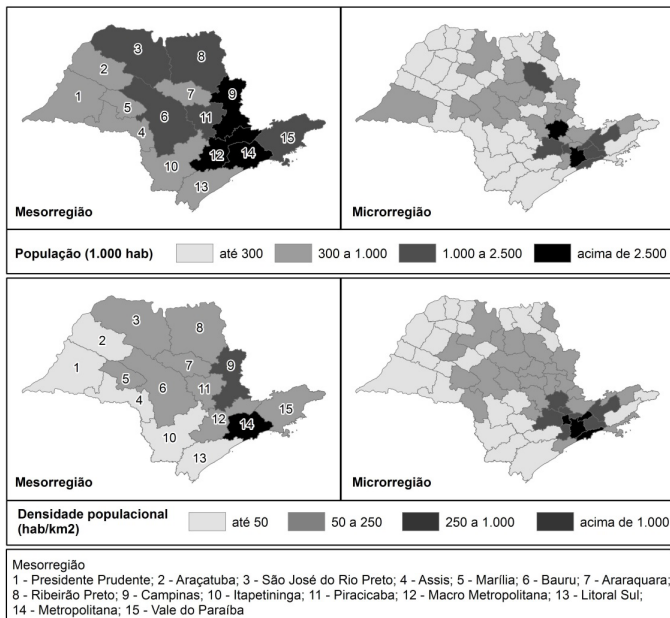


Figura 5 – Mapas coropléticos e variáveis demográficas.
Fonte: IBGE, Censo Demográfico 2010. Elaborado pelos autores.

Enfrentando as limitações

Contornando os obstáculos

Uma primeira possibilidade para lidar com as limitações apontadas é tentar contorná-las, encontrando formas de minimizar as adversidades, embora sem necessariamente suprimi-las. As considerações a seguir têm por base a utilização das unidades geográficas de disseminação de dados tradicionalmente oferecidas pelo IBGE.

Com relação ao não ajuste entre recortes geográficos e a necessidade de utilizar dados secundários, por exemplo, no caso de a área de estudo abranger porções de diversos municípios, uma opção é trabalhar com os dados totais dos municípios envolvidos, mesmo sabendo que esses dados abrangem uma região que extrapola a área de estudo delimitada. No caso de haver necessidade de trabalhar com dados mais detalhados e específicos para a área de estudo, a opção é a realização de pesquisas amostrais, embora isso implique um maior consumo de recursos técnicos e financeiros, bem como de tempo.

Quando se trata de análises temporais, a solução é ajustar os dados utilizando informações que permitam estabelecer a comparabilidade entre as informações para cada período, caso elas existam. Embora teoricamente esta tarefa pareça simples, ela implica o emprego de grande esforço e de tempo, podendo levar a resultados não muito precisos, dependendo do grau de comparabilidade conseguido.

No que diz respeito à heterogeneidade relacionada com a forma e a extensão das unidades, uma solução é evitar as técnicas de análise espacial e de mapeamento temático que possam ser comprometidas devido a essa característica, optando por soluções que não sejam

dependentes e/ou influenciadas pelas características físicas das unidades de agregação de dados.

Quando se trata da falácia ecológica e do MAUP, o fato de reconhecer que existe um problema nos leva a ter consciência de estar trabalhando com números imprecisos. Entre as soluções propostas para a solução desses dois problemas podemos citar a utilização de dados individuais ou não agregados, a definição de um agrupamento ótimo ou a modelagem com variáveis agrupadas (Grasland; Madelin, 2006).

A proposta de utilização de dados individuais é com certeza a melhor solução, mas está longe de poder ser adotada porque, na prática, a disseminação de dados estatísticos oficiais individualizados não é possível devido às questões de confidencialidade.

Uma possível opção é a definição automatizada de agrupamentos, que é feita com o estabelecimento de regras estatísticas e computacionais para gerar soluções ótimas para uma análise específica, conforme proposta de Openshaw (1977). A dificuldade para uma criação efetiva desses agrupamentos está na necessidade de um grande conhecimento dos fenômenos envolvidos, sendo que o resultado final é válido apenas para o caso particularmente analisado ou outro que possa ser considerado similar. De certa forma, esses são fatores limitantes para uma aplicação generalizada desta solução, apesar de não impedir a sua utilização como um meio para minimizar os efeitos do MAUP e da falácia ecológica.

Outra proposta para tratar os efeitos do MAUP e da falácia ecológica é a criação de uma estrutura de modelagem hierárquica com a utilização de grupos de variáveis que são medidas no nível individual e, então, são utilizadas para ajustar as variâncias e covariâncias no nível agrupado (Steel; Holt, 1996). O principal problema desta proposta está na disponibilidade das variáveis individuais, tendo em vista as questões relacionadas com a restrição de disseminação de dados individualizados.

A técnica proposta por Gotway e Young (2007) para solucionar os problemas relacionados com a mudança de suporte espacial de dados utilizando técnicas de geoestatística tem por objetivo tratar os problemas derivados da agregação e da desagregação espacial de dados, bem como realizar a predição de valores para unidades espaciais sobrepostas. Essa técnica é relativamente recente e existem poucos casos publicados para que seja possível avaliar as suas limitações e aplicabilidade.

Outro caminho também adotado em relação ao MAUP vai em direção contrária ao proposto nos parágrafos anteriores: não encarar o MAUP como sendo um problema, mas como sendo um fato que colabora para aumentar o conhecimento sobre o relacionamento entre as variáveis e as unidades que compõem a análise. Jelinski e Wu (1996) argumentam que o MAUP não deve

ser visto como um problema, já que oferece informações importantes para o entendimento dos sistemas reais. Esse ponto de vista também é defendido por Grasland e Madelin (2006), que propõem uma maior exploração dos dados, criando, por exemplo, representações multiescalares do mesmo fenômeno, de modo a expor a sua complementaridade, como proposto por Sui (1999) e Mennis (2002). Sui (1999) ressalta que estudos realizados em uma única escala e utilizando uma unidade zonal específica não podem levar a resultados confiáveis, já que não é possível saber exatamente o grau com que os resultados foram afetados pela agregação de dados. Um exemplo prático desta solução pode ser visto em Bueno e D'Antona (2012).

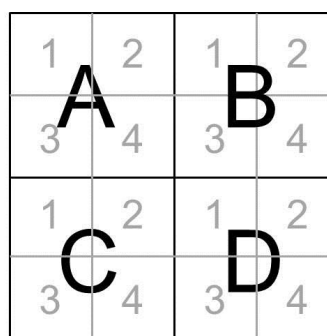
E, finalmente, ainda há outro caminho possível: ignorar o problema. Claramente, esta é a opção mais simples e fácil, e a consequência é a existência de incertezas nos resultados das análises. Infelizmente, este caminho é bastante utilizado, já que a mensuração dos efeitos do MAUP e da falácia ecológica não é uma tarefa simples (Dark; Bram, 2007).

Um novo paradigma

No que diz respeito às limitações relacionadas com análises temporais, o ideal seria trabalhar com unidades de disseminação que não se alterassem ao longo do tempo, e para isso elas teriam de ser independentes das unidades de coleta e também das unidades político-administrativas. No caso dos problemas ocasionados pela heterogeneidade de forma e extensão das unidades, a solução seria utilizar unidades com forma regular e de mesmo tamanho. Já no caso das dificuldades relacionadas com a adequação a recortes espaciais diversos, uma possível solução seria a utilização de unidades de pequenas dimensões, de modo que a agregação destas pudesse se aproximar do recorte desejado. Essa solução também seria a mais adequada para minimizar as questões relacionadas com a agregação de dados (MAUP e falácia ecológica), uma vez que unidades pequenas tendem a ser mais homogêneas internamente do que unidades de maior extensão.

Essas sugestões levam a concluir que para resolver ou minimizar os problemas apresentados anteriormente seria conveniente o uso de unidades geográficas pequenas e regulares, em vez de unidades político-administrativas e de coleta. Para tornar a solução ainda mais vantajosa, poderiam ser adotados vários conjuntos de unidades com dimensões diferentes, dispostas de forma hierárquica, formando uma "família" de unidades espaciais, como pode ser visto na Figura 6. Essa solução já é conhecida e utilizada há algum tempo, sendo denominada Grade Estatística – grade, devido a sua forma regular, e estatística, devido aos dados relacionados a cada célula.

A utilização da Grade Estatística é feita há vários anos por vários países do norte europeu e pelo Japão, mas tem se intensificado cada vez mais devido à disseminação do uso de geotecnologias em todas as etapas de elaboração dos censos de população (Nações Unidas, 2000; 2009), fato que, além de permitir efetivamente a sua construção, possibilita a utilização de técnicas que resultam em dados cada vez mais acurados. No Brasil, a partir do censo demográfico de 2010, a sua utilização passa a ser uma possibilidade real, e o IBGE lançou oficialmente produtos de disseminação com essa nova abordagem no início de 2016⁵. Considerando a trajetória de incorporação de novas técnicas nos censos demográficos brasileiros, o uso das geotecnologias desde a coleta dos dados viabiliza o que se poderia definir como uma nova forma, ou paradigma, de disseminação e análise espacial dos resultados dos censos.



Família 1: células A, B, C e D.
Família 2: células A1, A2, A3, A4, B1, B2, B3, B4, C1, C2, C3, C4, D1, D2, D3 e D4.

Dentre as principais vantagens da grade estatística destacam-se a adequação a recortes espaciais diversos, devido às suas pequenas dimensões, e a sua estabilidade temporal, uma vez que as células da grade não precisam ser alteradas ao longo do tempo por não terem significado administrativo ou territorial, sendo simplesmente um repositório para dados estatísticos. Além disso, o uso da grade como dado de entrada é vantajoso em modelos de simulação que utilizam dados espacializados, já que a maioria desses modelos trabalha com dados no formato matricial, que é semelhante à estrutura de uma grade regular (Bueno, 2014). Tais atributos, somados à facilidade com que a grade permite a integração de dados físicos e ambientais oriundos de classificação de imagens orbitais, propiciam a integração de duas áreas do conhecimento anteriormente distanciadas: a geografia e a estatística. Essa integração favorece, por exemplo, os estudos e análises no campo de População & Ambiente, permitindo compreender melhor as dimensões humanas das mudanças ambientais.

Considerações finais

Este artigo apresentou as unidades geográficas de coleta e de disseminação dos censos demográficos e

⁵Ver <<http://mapas.ibge.gov.br/interativos/grade>>.

explorou as suas limitações com relação à produção de mapas temáticos e análises espaciais. Essas limitações são em grande parte decorrentes das diferenças entre os objetivos para os quais essas unidades foram criadas e os objetivos das análises que os usuários desejam realizar.

Para cada limitação foram apresentadas sugestões para o seu enfrentamento, sendo algumas paliativas, já que alguns desses obstáculos são inerentes ao tipo de dado que está sendo trabalhado, uma vez que os dados coletados individualmente são agrupados para fins de disseminação.

Como parte das sugestões apresentadas para enfrentar a limitação das unidades comumente utilizadas para disseminação de dados censitários está a apresentação de um novo paradigma: a utilização de grades regulares como unidade de agregação. A regularidade da forma das células associada às suas pequenas dimensões, quando comparadas às das unidades político-administrativas, propicia uma melhor resolução espacial, maior adequação a recortes espaciais diversos, facilidade de integração com outras fontes de dados, além de facilitar as análises estatísticas espaciais.

No quadro mais geral dos avanços dos censos demográficos do IBGE, a incorporação das geotecnologias vem aumentando a qualidade dos resultados disseminados e viabilizando a existência de grades estatísticas definidas para a agregação de dados. Espera-se que as unidades convencionais (estatísticas e político-administrativas) e a grade estatística possam conviver harmonicamente, sem que uma seja preterida em função da outra, uma vez que elas atendem a objetivos diferentes e cabe ao usuário a escolha daquela que melhor se adéque ao seu trabalho.

Referências

- BARBUDA, M. M. S. (2004). A atualização cartográfica na Base Territorial Rural visando a Contagem da População 2005 e o Censo Agropecuário 2006. In: *CONGRESSO BRASILEIRO DE CADASTRO TÉCNICO MULTIFINALITÁRIO E GESTÃO TERRITORIAL*, 6, 2004, Florianópolis. Anais. Florianópolis: Universidade Federal de Santa Catarina.
- BIANCHINI, Z. M. (2012). Práticas de proteção da confidencialidade de informações adotadas no IBGE. In: *SEMINÁRIO DE METODOLOGIA DO IBGE*, 1, 2012, Rio de Janeiro. Resumos. Disponível em: <<http://www.smi2012.ibge.gov.br/SiteAssets/SitePages/listaConferencias/03Praticas%20de%20prote%C3%A7%C3%A3o%20Confidencialidade%20%20IBGE%20-%20ZeliaBianchini.pdf>>. Acesso em: 25 jul. 2015.
- BUCKLEY, A. (2013). Understanding Statistical Data for Mapping Purposes. *Arc User*, Winter 2013. Disponível em: <<http://www.esri.com/esri-news/arcuser/winter-2013/understanding-statistical-data-for-mapping-purposes>>. Acesso em: 15 jul. 2015.
- BUENO, M. C. D. (2014). Grade estatística: uma abordagem para ampliar o potencial analítico de dados censitários. Tese (Doutorado). Campinas: Instituto de Filosofia e Ciências Humanas, UNICAMP.
- BUENO, M. C. D.; D'ANTONA, A. O. (2012). Utilização de grades regulares para análises espaciais intramunicipais a partir de dados censitários – Teste para Limeira/SP, 2010. In: *ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS*, XVIII, 2012, Águas de Lindóia. Anais... Disponível em: <www.abep.nepo.unicamp.br/xviii/anais>. Acesso em: 14 mar. 2015.
- DARK, S. J.; BRAM, D. (2007). The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical*

Geography, vol. 31, n. 5, p. 471-479.

DEICHMANN, Uwe (1996). *A Review of Spatial Population Database Design and Modelling*. Technical Report TR-96-3. National Center for Geographic Information and Analysis (NCGIA), University of California, Santa Barbara, USA. Disponível em: <http://www.ncgia.ucsb.edu/Publications/Tech_Reports/96/96-3.PDF>. Acesso em: 16 ago. 2013.

ESRI (2013). *ArcGIS Desktop Help*.

FLOWERDEW, R. (2011). *How serious is the Modifiable Areal Unit Problem for analysis of English census data?* *Population Trends*, vol. 145, p. 106.

FOTHERINGHAM, S. (1998). *Trends in quantitative methods II: stressing the computational*. *Progress in Human Geography*, vol. 22, n. 2, p. 283-292.

FREEDMAN, D. A. (1999). *Ecological Inference and the Ecological Fallacy*. Department of Statistics, University of California Berkeley, Technical Report n. 549. Disponível em: <<http://www.stat.berkeley.edu/~census/549.pdf>>. Acesso em: 6 set. 2015.

GONÇALVES, J. M. M. (1995). *IBGE: um retrato histórico*. Documentos para Disseminação. Memória Institucional, n. 5. Rio de Janeiro: IBGE.

GOTWAY, C. A.; YOUNG, L. J. (2002). *Combining Incompatible Spatial Data*. *Journal of the American Statistical Association*, vol. 97, n. 458, p. 632-648.

_____. (2007). *A Geostatistical Approach to Linking Geographically Aggregated Data From Different Sources*. *Journal of Computational and Graphical Statistics*, vol. 16, n. 1, p. 115-135.

GRASLAND, C.; MADELIN, M. (2006). *The Modifiable Areas Unit Problem – Final Report*. ESPON – European Spatial Planning Observation Network. Disponível em: <http://www.espon.eu/export/sites/default/Documents/Projects/ESPON2006Projects/StudiesScientificSupportProjects/MAUP/espon343_maup_final_version2_nov_2006.pdf>. Acesso em: 25 mai. 2012.

GUZMÁN, J. M.; SCHENSUL, D.; ZHANG, S. (2013). *Understanding Vulnerability and Adaptation Using Census Data*. In: MARTINE, G.; SCHENSUL, D. (eds.). *The Demography of Adaptation to Climate Change*. New York, London and Mexico City: UNFPA, IIED e El Colegio de México.

HOGAN, D. J. (1991). *Crescimento demográfico e meio ambiente*. *Revista Brasileira de Estudos de População*, vol. 8, n. 1- 2, p. 61-77.

Instituto Nacional de Estatística de Cabo Verde – INE Cabo Verde (s./d.). *IV Recenseamento Geral da População e de Habitação – Censo 2010: Resumo dos Principais Resultados por Meio de Residência e Concelho*. Disponível em: <<http://www.ine.cv/censo/files/Resumo%20indicadores%20RGPH%202010.pdf>>. Acesso em: 16 ago. 2015.

Instituto Nacional de Estatística De Moçambique – INE Moçambique (s./d.). *II Recenseamento Geral da População e de Habitação 1997*. Disponível em: <<http://mozdata.microdatahub.com/index.php/catalog/7/study-description>>. Acesso em: 23 jul. 2015.

Instituto Brasileiro de Geografia e Estatística – IBGE (1981). *Departamento de Coordenação de Métodos. A base setorial dos censos de 80*. Coleção Ibgeana. Rio de Janeiro: IBGE.

_____. (2003). *Metodologia do Censo Demográfico 2000. Série Relatórios Metodológicos*, vol. 25. Rio de Janeiro: IBGE.

_____. (2010). *Censo 2010 – Síntese das Etapas da Pesquisa*. Rio de Janeiro: IBGE. Disponível em: <http://censo2010.ibge.gov.br/images/pdf/censo2010/sintese/sintese_censo2010_portugues.pdf>. Acesso em: 23 mar. 2011.

_____. (2013). *Metodologia do Censo Demográfico 2010. Série Relatórios Metodológicos*, vol. 41. Rio de Janeiro: IBGE.

JELINSKI, D. E.; WU, J. (1996). *The modifiable areal unit problem and implications for landscape ecology*. *Landscape Ecology*, vol. 11, n. 3, p. 129-140.

JUDSON, D. H. (2007). *Information integration for constructing social statistics: history, theory and ideas towards a research programme*. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 170, n. 2, p. 483-501.

MARTIN, D. (1996). *An Assessment of Surface and Zonal Models of Population*. *International Journal of Geographical Information Systems*, v. 10, n. 8, p. 973-989.

MENNIS, J. (2002). *Using Geographic Information Systems to Create and Analyze Statistical Surfaces of Population*

and Risk for Environmental Justice Analysis. *Social Science Quarterly*, vol. 38, n. 1, p. 281-297.

MULLER, E. P. L.; CUBAS, M. R.; BASTOS, L. C. (2010). Georreferenciamento como instrumento de gestão em unidade de saúde da família. *Revista Brasileira de Enfermagem*, vol. 63, n. 6.

NAÇÕES UNIDAS (2000). Handbook on geographic information systems and digital mapping. *Studies in Methods, Series F*, n. 79. New York: Nações Unidas.

_____. Divisão de Estatística (2007). Report of the Expert Group Meeting on Contemporary Practices in Census Mapping and Use of Geographical Information Systems. Disponível em: <http://unstats.un.org/unsd/demographic/meetings/egm/CensusMappingEGM07/docs/Final_Report.pdf>. Acesso em: 14 jun. 2010.

_____. Divisão de Estatística (2008). Principles and Recommendations for Population and Housing Census. *Statistical Papers*, Série M, n. 67/Rev. 2. New York: Nações Unidas.

_____. Divisão de Estatística (2009). Handbook on geospatial infrastructure in support of census activities. *Studies in Methods*, Série F, n. 103. New York: Nações Unidas.

OFFICE FOR NATIONAL STATISTICS – ONS (2012). Changes to Output Areas and Super Output Areas in England and Wales, 2001 to 2011. Disponível em: <<http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/output-area-oas-/index.html>>. Acesso em: 23 ago. 2013.

OLEA, R. A. (ed.) (1991). *Geostatistical Glossary and Multilingual Dictionary*. New York: Oxford University Press.

OLIVEIRA, J. S. (2003). Brasil mostra a tua cara: imagens da população brasileira nos censos demográficos de 1872 a 2000. *Textos para discussão*, n. 6. Rio de Janeiro: Escola Nacional de Ciências Estatísticas. 75 p.

OPENSHAW, S. (1977). A geographical solution to scale and aggregation problems in region-building, partitioning and spatial modelling. *Transactions of the Institute of British Geographers*, New Series, vol. 2, n. 4, p. 459-472.

_____. (1984). The modifiable areal unit problem. *Concepts and Techniques in Modern Geography*, n. 38. Norwich: Geo Books.

_____; TAYLOR, P. (1979). A Million or so Correlation Coefficients. In: WRIGLEY, N. (ed.). *Statistical Methods in the Spatial Sciences*. London: Pion.

PLUMEJEAUD, C.; PRUD'HOMME, J.; DAVOINE, P. A.; GENSEL, J. (2010). Transferring Indicators into Different Partitions of Geographic Space. In: TANIAR, D.; GERVASI, O.; MURGANTE, B.; PARDEDE, E.; APDUHAN, B. O. (eds.). *ICCSA'10 Proceedings of the 2010 International Conference on Computational Science and Its Applications – Volume Part I*. Berlin, Heidelberg: Springer-Verlag.

ROSA, R. (2011). Análise Espacial em Geografia. *Revista da ANPEGE*, vol. 7, n. 1, p. 275-289.

STEEL, D. G.; HOLT, D. (1996). Analyzing and adjusting aggregation effects: The ecological fallacy revisited. *International Statistical Review*, vol. 64, n. 1, p. 39-60.

SUI, D. (1999). GIS, Environmental Equity Analysis, and the Modifiable Areal Unit Problem. In: CRAGLIA, M.; ONSRUD, H. (eds.). *Geographic Information Research: Trans-Atlantic Perspectives*. London, UK: Taylor and Francis.

WACHTER, K. W. (2005). Spatial Demography Special Feature: Spatial Demography. *PNAS*, vol. 102, n. 43.

WONG, D. W. S. (1996). Aggregation effects in geo-referenced data. In: GRIFFITHS, D. A. (ed.). *Advanced Spatial Statistics*. Boca Raton, Florida: CRC Press.