
USO DE DADOS DE DOMÍNIO PÚBLICO PARA MAPEAMENTO DASIMÉTRICO

VITOR OLIVEIRA FRANÇA (1)
JULIA CELIA MERCEDES STRAUCH (1)

(1) Escola Nacional de Ciências Estatísticas
Programa de Pós-Graduação *Stricto Sensu* em Estudos Populacionais e Pesquisas Sociais
Rua André Cavalcanti, 106, sala 503 B
e-mail: {vitor.franca, julia.strauch}@ibge.gov.br

Informações detalhadas sobre a estrutura e a distribuição espacial das populações humanas são imprescindíveis para tomada de decisões em uma variedade de domínios. A principal e mais completa fonte dessas informações são os censos de população realizados por quase todos os países geralmente a cada decênio. As variáveis demográficas em conjunto com a base cartográfica do zoneamento censitário podem ser estruturadas em bancos de dados geográficos, tratadas e analisadas por meio dos Sistemas de Informação Geográfica (SIGs), não apenas para representação cartográfica, como também para obter estimativas quantitativas usadas em análises espaciais e no cálculo de taxas e indicadores. Entretanto, variáveis censitárias possuem limitações relacionadas às suas escalas espacial e temporal que comprometem sua utilização em algumas aplicações, como por exemplo, meio ambiente, defesa civil, saúde pública e *geomarketing*. Devido à inerente agregação dos dados censitários por questões operacionais e de sigilo da informação, a técnica cartográfica mais utilizada para representar e analisar dados socioeconômicos é a técnica coroplética. Todavia, dados agregados em zonas coropléticas inevitavelmente introduzem viés em análises espaciais, pois essa técnica assume o pressuposto da distribuição uniforme da variável no interior de cada unidade do zoneamento, o que raramente ocorre com a população. Quanto maior a unidade, maior o viés introduzido, maior a perda em detalhe e especificidade, maior a propagação de erros nos cálculos subsequentes. Dados agregados, e, por conseguinte, mapas coropléticos, estão fadados à Falácia Ecológica e ao Problema da Unidade de Área Modificável (MAUP). Além disso, os zoneamentos censitários não coincidem com outras regiões de análise, como unidades da paisagem, bacias hidrográficas, regiões administrativas de ensino e saúde etc. Essas limitações podem ser contornadas pela geração de dados demográficos com resolução espacial e/ou temporal mais detalhada, que representem a distribuição populacional de forma mais realista. Nesse sentido, muitos métodos de mapeamento dasimétrico foram propostos para desagregar e estimar variáveis populacionais em pequenas áreas. Na prática, esses métodos constituem uma interpolação zonal com uso de informações auxiliares. A interpolação dasimétrica possui diversas abordagens e pode ser realizada com diferentes fontes de dados, como classes de uso ou cobertura da terra, malha viária (eixos de logradouros) e cadastro de endereços. Este trabalho examina a eficácia da aplicação de Conjuntos de Dados Geoespaciais (CDGs) abertos ou de domínio público para desagregar a variável população total do Censo Demográfico 2010 em células de uma grade estatística regular. A área de estudo compreendeu o município de Marabá/PA. Foram empregados os métodos de interpolação ponderada por segmentos de logradouro (REIBEL & BUFFALINO, 2005) e por endereços (TAPP, 2010) para os quais foram utilizados os dados do *OpenStreetMap* e do Cadastro de Endereços para Fins Estatísticos (CNEFE). Foi possível calcular os pesos e estimar a população em cada segmento de logradouro e coordenada de endereço. As estimativas foram então agregadas em células regulares com resolução espacial de 250m. Foram elaborados algoritmos para automação dos procedimentos no aplicativo *ArcGIS*, empregando a ferramenta *ModelBuilder* que proporcionou a construção visual do modelo e a geração de um código em linguagem *Python* com a biblioteca *Arcpy*. Os resultados foram avaliados através do erro médio absoluto, coeficiente de variância e o Erro Médio Quadrático (RMSE). Devido à disponibilidade de dados, a interpolação dasimétrica por logradouro se mostrou mais eficiente nas áreas urbanas, enquanto que o cadastro de endereços foi mais eficiente nos setores rurais. Desta forma, este trabalho contribui para o estado da arte dos métodos dasimétricos, uma vez que a crescente oferta de dados abertos e os métodos descritos na literatura tornam viável a representação mais detalhada da distribuição espacial da população. Verificou-se que os CDGs abertos oferecem considerável potencial para o mapeamento dasimétrico. Algoritmos disponibilizados através de códigos e *plug-ins* para aplicativos SIG podem contribuir para a difusão das técnicas e facilitar a estimativa de variáveis sociodemográficas em pequenas áreas.