

ESTIMATIVA E DISTRIBUIÇÃO ESPACIAL DE POPULAÇÃO URBANA COM IMAGENS DE SATÉLITES DE LUZES NOTURNAS: UM ESTUDO PARA A REGIÃO METROPOLITANA DE SÃO PAULO, BRASIL, COM O SENSOR VISIBLE/ INFRARED IMAGING RADIOMETER SUITE (VIIRS)

Vinicius Etchebeur Medeiros DÓRIA^{1, 2}

Silvana AMARAL¹

Antônio Miguel Vieira MONTEIRO¹

Resumo

Este artigo avalia o potencial das imagens de luzes noturnas do sensor VIIRS (Visible/Infrared Imaging Radiometer Suite) embarcado no satélite S-NPP para estimar o volume e a distribuição da população urbana na Região Metropolitana de São Paulo, explorando seu uso como uma possibilidade complementar às informações demográficas produzidas pelos censos demográficos. Utilizou-se uma imagem mosaico de luzes noturnas de 2012, com resolução espacial de aproximadamente 500m, e as áreas de ponderação do censo 2010 do IBGE. Após a análise por regressão linear simples, procedeu-se o estudo da dependência espacial para os dados de população através da análise de autocorrelação espacial dos índices de Moran global e local. Uma Regressão Geograficamente Ponderada (GWR) foi então testada para considerar a dependência espacial nas relações entre os dados de luzes noturnas e os dados de população nas áreas de ponderação. Os resultados das estimativas de população pela GWR foram consistentes com os dados do censo de 2010. Dos coeficientes da regressão local, 32,8% mostraram-se estatisticamente significantes e, no geral, o modelo superestimou o valor de população em 5%. Os coeficientes do modelo GWR indicaram a existência de diferentes regimes espaciais na RMSP quando observamos a associação entre luzes e o volume de população. A análise dos diferentes padrões associados à distribuição espacial da população e os dados de luzes noturnas também se mostraram instrumentos auxiliares para observação e leitura das estratégias de ocupação humana das terras urbanas na região metropolitana.

Palavras-chave: Luzes noturnas. VIIRS. Regressão linear. GWR. População urbana.

¹ Instituto Nacional de Pesquisas Espaciais – INPE. Av. dos Astronautas, 1758, Jd. da Granja, CEP: 12227-010 - São José dos Campos, SP. E-mail: {silvana, miguel}@dpi.inpe.br

² Instituto Brasileiro de Geografia e Estatística – IBGE. Av. Brasil, 15671, Parada de Lucas, CEP: 21241-051-Rio de Janeiro, RJ. E-mail: vinicius.doria@ibge.gov.br

Abstract

Urban population estimate and spatial distribution from nighttime light data: a study for the São Paulo Metropolitan Region, Brazil, using Visible/Infrared Imaging Radiometer Suite (VIIRS) sensor

This paper assesses the potential of nighttime lights images from VIIRS sensor onboard S-NPP satellite to estimate the total urban population and its distribution in the São Paulo Metropolitan Region (RMSP). We explored the use of nighttime lights data as a complementary data source to the existing demographic census population data, particularly useful for the inter-census periods. A 2012 nighttime light image mosaic, with spatial resolution of approximately 500 m, was analyzed taking the weighted sampled areas from IBGE 2010 census tracts as unit of analysis. First, a model of simple linear regression was tested taking the sum of the digital number of nighttime light pixels as independent variable, and the total population as response variable. Secondly, we studied the spatial dependence based on the values of global and local Moran Indexes. Then, a Geographically Weighted Regression (GWR) was used to explore the spatial dependence in the relationships between nighttime lights and population on intra-urban spatial units. The population results estimated by GWR were consistent with data from the 2010 census. The model in general overestimated population values in 5%, and 32.8% of the GWR coefficients were statistically significant. GWR coefficients indicated the presence of different spatial regimes influencing the relationships between nighttime lights and total population at the RMSP. The analysis of the different patterns emerged from the spatial distribution of population and nighttime lights data showed also to be an auxiliary instrument to observe human settlement strategies of urban land in the metropolitan area.

Key words: Nighttime lights. VIIRS. Linear regression. GWR. Urban population.

INTRODUÇÃO

Embora as áreas urbanas ocupem uma pequena porção da superfície continental da Terra (SCHNEIDER et al., 2009), mais da metade da população mundial vive nestas áreas (UN, 2012). O crescimento urbano implica em transformações importantes na estrutura sócio-espacial das cidades, envolvendo a distribuição desta população e as mudanças de uso e cobertura da terra em área urbana.

Em meados da década de 1970, com o advento dos sensores da série Landsat e a possibilidade de coletar informações multiespectrais da superfície da Terra, incluindo o infravermelho termal, as pesquisas relacionadas à urbanização se firmaram em realizar classificações de uso e cobertura da terra para estimar algumas métricas urbanas, como, por exemplo, densidade de residências e índices de qualidade residencial (FORSTER, 1983).

A partir de dados de sensoriamento remoto também é possível se obter estimativas de população realizando contagem de unidades residenciais individuais, medidas de extensão de áreas urbanas, classificações de uso e cobertura da terra (JENSEN; COWEN, 1999). Estas estimativas são importantes em períodos intercensosos e em diversas partes do mundo onde os dados sobre a população são escassos. A partir de contagem de residências, utilizando imagens de 2,5 m de resolução espacial, Jensen e Cowen (1999), por exemplo, obtiveram correlação significativa ($R^2 = 0,81$) na estimativa da população da Carolina do Sul, E.U.A.. Gonçalves et al. (2004), utilizando

imagens ortorretificadas de 0,6 m de resolução espacial, e através de contagem de residências, observaram uma correlação linear significativa ($R^2 = 0,92$) entre a população inferida e a população observada para o município de São José dos Campos – SP, no ano 2000.

Dados de sensoriamento remoto e técnicas de geoprocessamento têm sido utilizados também para auxiliar os estudos relacionados à urbanização. O sensoriamento remoto óptico da Terra utiliza principalmente sensores que operam sob a luz do Sol refletida pela Terra. Os sensores ópticos de imageamento noturno trabalham basicamente nas regiões do espectro eletromagnético referentes ao infravermelho e microondas. Estes sensores surgiram na década de 1960 e possibilitaram observar alguns aspectos que durante o dia dificilmente seriam perceptíveis, como, por exemplo, auroras, focos de incêndios florestais, luzes de cidades e chamas provenientes da queima de gases (FOSTER, 1983). Desde então, uma parcela da comunidade científica tem direcionado esforços em compreender, estimar, monitorar e mapear dinâmicas relacionadas aos processos da urbanização utilizando dados de sensores de imageamento noturno (AMARAL et al., 2006; DOLL, 2010; SMALL; ELVIDGE, 2011; ZHANG; SETO, 2011).

Quanto aos dados populacionais, os censos demográficos são a fonte de informação oficial dos países e preferencialmente utilizadas para estudos em escala global. No Brasil, o Instituto Brasileiro de Geografia e Estatística (IBGE) realiza o censo a cada 10 anos e divulga as informações de contagem de população agregadas por setores censitários³. Considerando escalas regionais (níveis municipal e estadual) e globais (níveis federais) de análise, há alguns anos a academia científica tem explorado os dados de luzes noturnas, do sensor DMSP/OLS, de resolução espacial de 1 km, como alternativa para estimar a população e/ou o consumo de energia elétrica (AMARAL et al., 2005; 2006; DOLL; PACHAURI, 2010; LETU et al., 2010; MA et al., 2012).

Em 28 de outubro de 2011 entrou em operação o sensor *Visible Infrared/Imaging Radiometer Suite* (VIIRS), projetado para realizar o imageamento diurno e noturno do planeta Terra para obter informações relativas a 22 bandas espectrais com resolução espacial de 500 m, trazendo novas oportunidades para explorar seu uso no contexto dos estudos de população.

Este trabalho tem por objetivo apresentar uma avaliação de imagens de luzes do sensor VIIRS quanto ao seu potencial para estimar a população urbana da Região Metropolitana de São Paulo (RMSP) numa escala local, adotando como unidade espacial de análise as áreas de ponderação⁴ definidas pelo censo demográfico. As variações intraurbanas das relações espaciais entre as informações de luzes noturnas e os dados de população, observados nas áreas de ponderação, são apresentadas e discutidas.

Este trabalho pretende contribuir com o campo de estudos de população e ambiente, avaliando as possibilidades e as limitações da utilização das imagens de luzes noturnas deste novo sensor, como dado complementar e auxiliar para a estimativa de volumes populacionais em pequenas áreas, utilizando imagens para uma unidade espacial de análise até então inexplorada em dados de luzes noturnas.

³ Um setor censitário é a menor unidade de coleta e registro das informações obtidas durante o recenseamento e os setores são divididos de forma que facilite a coleta do agente de campo, respeitando os limites político-administrativos dos municípios (IBGE, 2010).

⁴ Uma área de ponderação abrange, no mínimo, 400 domicílios particulares ocupados que tenham respondido ao questionário da amostra, e é constituída por setores censitários circunvizinhos. Áreas de ponderação apenas são disponíveis para municípios que possuam mais do que 190 mil habitantes (IBGE, 2010).

DADOS DE LUZES NOTURNAS PARA ESTIMATIVA DE POPULAÇÃO

A possibilidade de aplicação do sensoriamento remoto noturno em estudos relacionados à urbanização iniciou-se na década de 1970 (p. ex. CROFT, 1978). O primeiro sensor a realizar o imageamento noturno foi o *Operational Linescan System* (OLS), sensor que orbita a bordo dos satélites do *Defense Meteorological Satellites Program* (DMSP). O DMSP é um programa do Departamento de Defesa Norte Americano, administrado pelo Centro de Sistemas Espaciais e Mísseis da Força Aérea dos Estados Unidos desde meados da década de 1960 (NASA, 2003). O OLS é um radiômetro de varredura oscilatória que adquire imagens diurnas e noturnas. Um tubo fotomultiplicador presente na banda do visível/infravermelho próximo aumenta a sensibilidade dos detectores em até quatro vezes durante o imageamento noturno, permitindo identificar nuvens iluminadas pela luz da lua e fontes que emitem pouca radiação do infravermelho próximo, como, por exemplo, as luzes de uma cidade (ELVIDGE et al., 2001).

Diversos estudos utilizaram essas imagens de luzes noturnas do OLS para realizar estimativas populacionais. Dentre estes estudos, Amaral et al. (2005) e Letu et al. (2010) encontraram significativa correlação entre as imagens de luzes, a população e o consumo de energia elétrica na Amazônia, e no Japão e China, respectivamente. Por sua vez, Doll e Pachauri (2010) verificaram a capacidade de se utilizar imagens de luzes noturnas para auxiliar o mapeamento e a estimativa de populações rurais que não são abastecidas por energia elétrica em países em desenvolvimento. Apesar dos resultados terem superestimado as populações que não possuem energia elétrica, a informação das luzes noturnas foi considerada uma referência inicial para estimativas de população em países de informação deficitária.

O VISIBLE INFRARED IMAGING RADIOMETER SUITE (VIIRS)

O *Joint Polar Satellite System* (JPSS) é um programa estabelecido em 2010 como resultado da reestruturação do *National Polar-Orbiting Environmental Satellite System* (NPOESS), para prover continuidade ao *Polar-orbiting Operational Environmental Satellite System* (POES) da *National Oceanic and Atmospheric Administration* (NOAA). Como resultado, foi lançado um projeto preparatório do programa JPSS, o satélite *Suomi National Polar-orbiting Partnership* (S-NPP) (CAO et al., 2014) em 28 de outubro de 2011 da base aérea estadunidense de Vandenberg, Califórnia (XIONG et al., 2012). Dentre os cinco sensores a bordo desta plataforma, encontra-se o *Visible Infrared Imaging Radiometer Suite* (VIIRS), cujo foco principal está na observação das nuvens e de variáveis da superfície terrestre, enquanto que os outros sensores foram desenvolvidos para medir o balanço radiativo e variáveis atmosféricas da Terra (CAO et al., 2014).

O sensor VIIRS é um radiômetro de varredura mecânica que adquire imagens diurnas e noturnas. Cada imagem recobre aproximadamente 3.000 km da superfície terrestre, provendo cobertura diária diurna e noturna da Terra. Possui 22 bandas espectrais distribuídas entre 0,40 e 12,43 μm . Entre elas encontra-se a banda *Day/Night Band* (DNB) que realiza medições durante a órbita diurna e noturna do satélite (LEE et al., 2006). Esta banda tem um intervalo espectral de 0,5 a 0,9 μm , possui um arranjo de quatro sensores CCD (*Charged-Coupled Device*) dedicados e as imagens têm resolução espacial de 750 metros. Ademais, a calibração da banda DNB realizada a bordo do satélite permite uma acurácia radiométrica superior ao do sensor OLS (LEE et al., 2006; CAO, 2013).

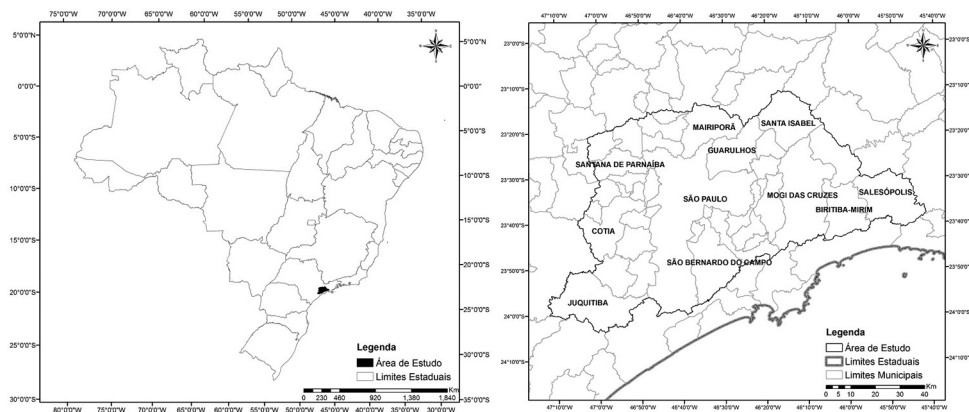
Utilizando os dados do novo sensor VIIRS, Li et al. (2013) utilizaram as imagens de luzes noturnas do OLS e do VIIRS como *proxies* do Produto Interno Bruto (PIB) de 31 províncias e 393 municípios da China. O objetivo deste trabalho era avaliar o desempenho das imagens noturnas do sensor VIIRS em comparação ao das imagens do sensor OLS, utilizando regressão linear simples. Os resultados indicaram um desempenho significativamente superior do VIIRS (R^2 de 0,86 e 0,85 para as províncias e municípios, respectivamente) em relação ao OLS (R^2 de 0,70 e 0,67 para as províncias e municípios, respectivamente).

ÁREA DE ESTUDO

A Região Metropolitana de São Paulo (RMSP) foi instituída pela Lei Complementar Federal n.14 de 1973 e disciplinada pela Lei Complementar Estadual n. 94 de 1974. Dos 645 municípios do estado de São Paulo, 39 compõem a RMSP: Arujá, Barueri, Biritiba-Mirim, Caieiras, Cajamar, Carapicuíba, Cotia, Diadema, Embu, Embu-Guaçu, Ferraz de Vasconcelos, Francisco Morato, Franco da Rocha, Guararema, Guarulhos, Itapeverica da Serra, Itapevi, Itaquaquecetuba, Jandira, Juquitiba, Mairiporã, Mauá, Mogi das Cruzes, Osasco, Pirapora do Bom Jesus, Poá, Ribeirão Pires, Rio Grande da Serra, Salesópolis, Santa Isabel, Santana de Parnaíba, Santo André, São Bernardo do Campo, São Caetano do Sul, São Lourenço da Serra, São Paulo, Suzano, Taboão da Serra e Vargem Grande Paulista.

Entre os anos de 2000 e 2010 a população total residente na RMSP cresceu 11,8%, cerca de 2 milhões habitantes, sendo que 98,8% da população é considerada urbana. O município que possui a maior densidade demográfica é Diadema, com 12.519 habitantes por km^2 , e Salesópolis, com 39 habitantes por km^2 , apresenta a menor densidade demográfica. São Paulo é o município mais populoso, com mais de 11 milhões de habitantes, e São Lourenço da Serra é o menos populoso, com 13.973 habitantes (IBGE, 2010).

A região de estudo deste trabalho compreende as 619 áreas de ponderação, consideradas urbanas pelo IBGE, dos 39 municípios da RMSP, totalizando uma área de aproximadamente 6.200 km^2 , no estado de São Paulo (Figura 1).



Fonte: IBGE, 2010. Elaborado pelos Autores.

MATERIAL

Considerando-se as áreas de ponderação urbanas da RMSP como unidades de análise, obteve-se o valor total de população residente a partir do censo demográfico 2010 (IBGE, 2010) e calculou-se o total de pixels de luzes noturnas a partir das imagens do VIIRS.

Neste trabalho, foram utilizadas somente as áreas de ponderação de situação urbana, ou seja, todos os setores censitários que compõem a área de ponderação são categorizados quanto a uma das seguintes situações: área urbanizada de vila ou cidade; área não urbanizada de vila ou cidade; ou área urbana isolada. Estas situações de setores são definidas pelo próprio IBGE e disponibilizadas juntamente com os resultados do Universo (IBGE, 2010).

O total de população para cada área de ponderação foi calculado a partir dos resultados do Universo do Censo Demográfico de 2010, realizado pelo IBGE, que disponibiliza os dados por setor censitário. Os dados de população total dos setores urbanos foram então somados e atribuídos para cada área de ponderação, adequando esta variável à unidade de análise.

Os dados de luzes noturnas correspondem a um produto primário do sensor VIIRS disponibilizado no início de 2013, pelo *Earth Observing Group* da NOAA-NGDC – *National Geophysical Data Center*. Este produto é o resultado de um mosaico de diversas imagens livres de nuvens coletadas nos meses de abril e outubro de 2012, com resolução espacial de 15 segundos de arco, o equivalente a aproximadamente 500 m. A composição foi gerada usando dados coletados pela banda DNB (BAUGH et al., 2013). No entanto, este produto VIIRS não foi processado para remover as luzes associadas aos eventos efêmeros, tais como auroras, vulcões, focos de incêndios florestais, e os ruídos das imagens. Estes ruídos são aparentemente pequenos conjuntos de pixels com valores anômalos (muito altos) espalhados aleatoriamente por toda a imagem (Figura 2).

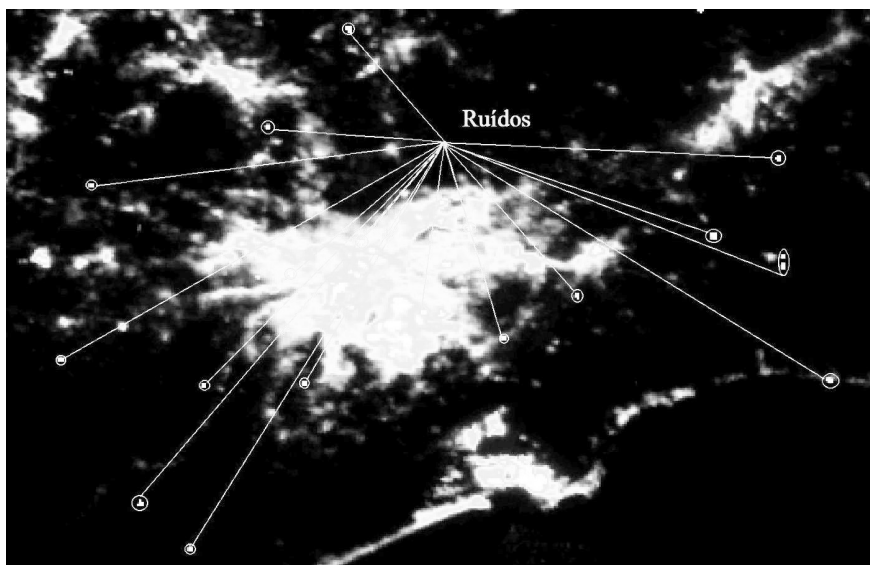


Figura 2 - Indicação de ruídos presentes na Imagem de luzes do VIIRS

Fonte: NOAA/NGDC/EOG, 2013. Elaborado pelos Autores.

O IBGE realiza e divulga anualmente estimativas de população para todos os municípios brasileiros para o cálculo das cotas do Fundo de Participação dos Estados e Municípios, para propostas de constituição de novos municípios e também para alterações de limites municipais ocorridos entre um ano e outro. De acordo com a estimativa populacional para o conjunto de municípios da RMSP, a população cresceu apenas 1,38% entre 2010 e 2012, cerca de 272 mil habitantes (IBGE, 2012). Por ser uma divulgação municipal, não é possível observar a distribuição deste crescimento em uma escala intramunicipal. Ademais, as imagens de luzes noturnas são consideradas indicadores indiretas da distribuição do crescimento populacional em áreas urbanizadas (CROFT, 1978; ELVIDGE et al., 1997; AMARAL et al., 2006; MA et al., 2012). Como não houve um crescimento populacional muito expressivo e as estimativas são divulgadas somente por município, justifica-se o uso dos dados do censo de 2010, pois somente os censos decenais possuem o nível de desagregação das informações na escala definida neste trabalho.

MÉTODOS

Antes de se calcular o valor de número de pixel de luzes noturnas para cada área de ponderação, a imagem mosaico de 2012 foi processada de modo a remover os ruídos originais. Neste procedimento, os ruídos foram eliminados por identificação visual e verificação de seus valores anômalos, comparativamente aos valores médios das luzes, de modo a evitar erros de superestimação da variável de luzes "soma dos pixels de luzes" noturnas (SPL) na comparação com dados de população. Como a área de estudo é uma região metropolitana de ocupação predominantemente urbana, não houve necessidade de se eliminar os pixels espúrios resultantes de eventos efêmeros, uma vez que as luzes presentes na imagem estão todas praticamente associadas às atividades humanas.

Os limites vetoriais das áreas de ponderação foram sobrepostos à imagem de luzes noturnas de modo a possibilitar a soma dos valores de Número Digital (ND) dos pixels de luzes compreendidos em cada dessas áreas. Desta forma, para cada área de ponderação foi calculado um valor para a variável SPL e o valor correspondente à variável total de população (TP).

Duas técnicas de regressão foram utilizadas para avaliar a relação entre as imagens de luzes e o total de população residente na RMSP: a regressão linear simples (global) e a Regressão Geograficamente Ponderada (*Geographically Weighted Regression* – GWR). Estas técnicas são brevemente descritas abaixo, juntamente com as justificativas para a escolha das mesmas.

A análise de regressão tem por objetivo quantificar a relação existente entre duas ou mais variáveis, de forma que uma variável (variável resposta) possa ser explicada por outra ou mais variáveis (variáveis explicativas). No caso deste estudo, TP é a variável resposta e a variável explicativa é a SPL.

É importante ressaltar que a regressão linear pressupõe que as variáveis resposta e explicativa, modeladas como variáveis aleatórias, são independentes. O modelo de regressão linear simples utilizado para descrever a relação entre TP e SPL é apresentado na Equação 1.

$$TP_i = \beta_0 + \beta_1.SPL_i + \varepsilon_i \quad (\text{Eq. 1})$$

onde TP_i é o total da população da i -ésima área de ponderação, β_0 é o coeficiente linear, β_1 é o coeficiente angular, SPL_i é a soma dos valores dos pixels de luzes da i -ésima área de ponderação e ε_i é o erro.

A escolha do modelo de regressão linear simples se justifica primeiramente pela observação do diagrama de dispersão das variáveis que indicou uma suposta relação linear, ainda que discreta, entre as variáveis. Segundo, a bibliografia pertinente a este estudo sugere que os dados de população e as informações de luzes apresentam uma associação que, como nos trabalhos de Amaral et al. (2005, 2006), Sutton et al. (2001) e Letu et al. (2010), podem ser caracterizadas por uma relação linear.

Porém, deve-se observar se os resíduos da regressão linear não apresentam estruturação espacial, o que indicaria que a localização destas variáveis não traz nenhum problema à hipótese da independência e portanto à utilização da regressão linear. Para isso, os índices de Moran global e local (ANSELIN, 1995; 1996) sobre os resíduos foram calculados para verificar a possível existência de autocorrelação espacial. Estes índices foram calculados e analisados, uma vez que o gráfico dos resíduos *versus* os valores estimados de população apresentou uma tendência à heterocedasticidade, isto é, a variância dos resíduos aparentemente não se apresentou constante. Os resíduos foram, então, espacializados para verificar a interferência do espaço nestes resultados, onde foi possível observar a presença de *clusters*, ou agrupamentos, de erros positivos e negativos na área de estudo. Os índices de Moran calculados neste trabalho têm como objetivo mensurar e identificar a presença destes agrupamentos, corroborando e consolidando a necessidade do uso de técnicas que considerem a localização geográfica das variáveis na análise deste estudo.

O índice global de Moran é análogo ao coeficiente de correlação convencional e quantifica o grau de correlação espacial existente. O valor deste índice varia de -1 a +1, onde -1 indica uma autocorrelação espacial negativa ou inversa, 0 indica aleatoriedade, isto é, quando não há autocorrelação espacial, e +1 indica autocorrelação espacial positiva ou direta.

O índice local de Moran pode ser considerado uma decomposição do índice global e produz um valor para cada área (polígono) da região de estudo. A interpretação deste índice está associada ao Diagrama de Espalhamento de Moran (Figura 3) (ANSELIN, 1996), que relata espacialmente o relacionamento entre os valores do vetor de desvios (Z) e os valores das médias locais (WZ), indicando possíveis diferentes regimes espaciais presentes nos dados. Os quadrantes Q1 ($Z+$, $WZ+$) e Q2 ($Z-$, $WZ-$) indicam pontos de associação espacial positiva, no sentido de que uma localização possui vizinhos com valores semelhantes. Os quadrantes Q3 ($Z+$, $WZ-$) e Q4 ($Z-$, $WZ+$) indicam pontos de associação espacial negativa, no sentido de que uma localização possui vizinhos com valores distintos.

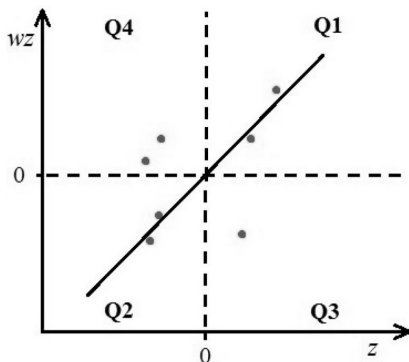


Figura 3 - Modelo do Diagrama de Espalhamento de Moran

Fonte: Anselin, 1996. Elaborado pelos Autores.

O processo de ocupação humana do solo urbano tem muitos fatores de formação e varia ao longo do tempo e do espaço, tendo características diferentes ao longo da região de estudo, apresentando uma heterogeneidade espacial. Para tratar este aspecto e incluir a localização das variáveis na análise, aplicou-se a técnica de Regressão Geograficamente Ponderada, GWR (BRUNSDON et al., 1996; FOTHERINGHAM et al., 1997). Esta técnica ajusta um modelo de regressão a cada ponto observado ponderando as demais observações em função da distância a este ponto. As observações mais próximas do ponto onde o modelo de regressão local foi ajustado terão peso maior do que as observações mais distantes. O modelo de regressão GWR utilizado é apresentado na Equação 2.

$$TP_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i).SPL_i + \varepsilon_i \quad (\text{Eq. 2})$$

onde u_i e v_i representam as coordenadas geográficas de cada um dos centroides das áreas de ponderação.

Na ponderação das observações em cada modelo de regressão local, utilizouse a função gaussiana com largura de banda adaptável, ou seja, largura menor em áreas com alta densidade de dados e largura maior em áreas com menor densidade de dados, por ter apresentado resultados melhores do que a largura de banda fixa. O software GWR4 (NAKAYA et al., 2014) foi utilizado para desenvolver o modelo, e neste, a largura de banda foi definida tomando-se como critério o menor valor do *Akaike Information Criterion* (AIC) (AKAIKE, 1973). A largura de banda adaptável neste caso deve buscar a menor distância para abranger os 25 vizinhos mais próximos para a estimativa do modelo.

Todos os mapas foram produzidos utilizando o sistema de coordenadas geográficas e o sistema de referência SIRGAS 2000.

RESULTADOS E DISCUSSÃO

A sobreposição das áreas de ponderação da RMSP sobre a imagem de luzes VIIRS após a eliminação dos ruídos (Figura 4) evidencia a redução da intensidade das luzes do centro para a periferia e as diferenças das áreas iluminadas nas áreas de ponderação.

Do diagrama de dispersão entre as variáveis SPL e TP (Figura 5) observa-se associação linear e positiva ($R^2 = 0,10$), e estatisticamente significativa (estatística $t = 8,49$), indicando que o modelo estimado é capaz de descrever 10% da variabilidade observada na variável TP.

A partir da visualização do histograma dos resíduos normalizados (Figura 6) observa-se uma distribuição próxima a uma distribuição normal. O diagrama de dispersão dos resíduos *versus* os valores estimados de população (Figura 7) apresentou a leve formação de um cone, uma estrutura indicativa da heterocedasticidade dos resíduos, indicando que quanto maior os valores de população, maior o erro adicionado ao modelo de regressão.

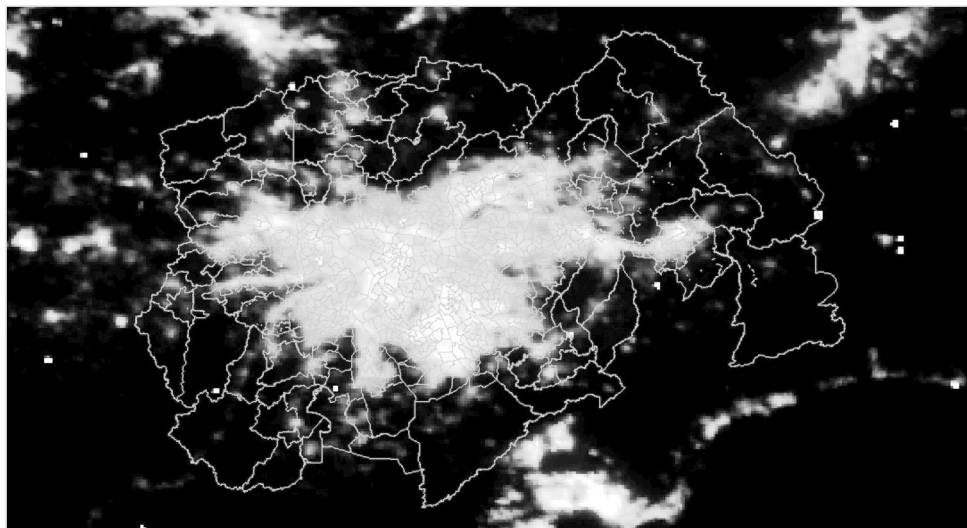


Figura 4 - Imagem mosaico de luzes do VIIRS 2012 e áreas de ponderação (vermelho) do censo IBGE 2010 da Região Metropolitana de São Paulo

Fonte: NOAA/NGDC/EOG, 2013 / IBGE, 2010. Elaborado pelos Autores.

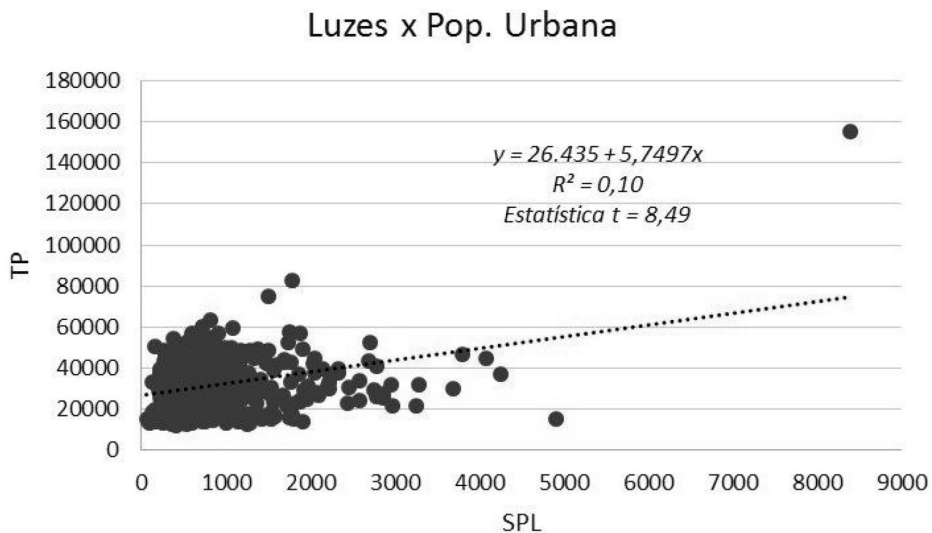


Figura 5 - Diagrama de dispersão entre as variáveis soma dos pixels de luzes noturnas (SPL) e total de população residente (TP) para as áreas de ponderação da RMSP

Fonte: NOAA/NGDC/EOG, 2013 / IBGE, 2010. Elaborado pelos Autores

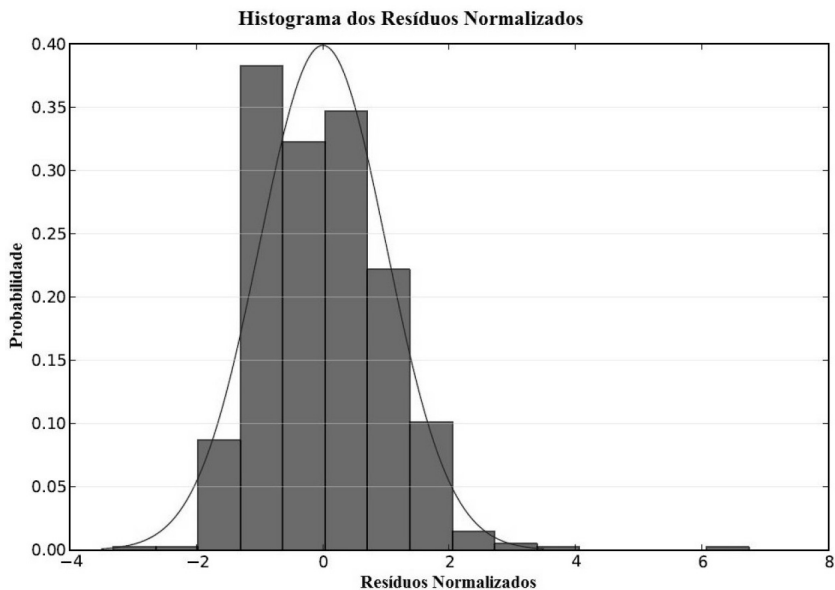


Figura 6 - Histograma dos resíduos da regressão linear simples; a linha contínua representa o histograma de uma distribuição normal
Elaborado pelos Autores.

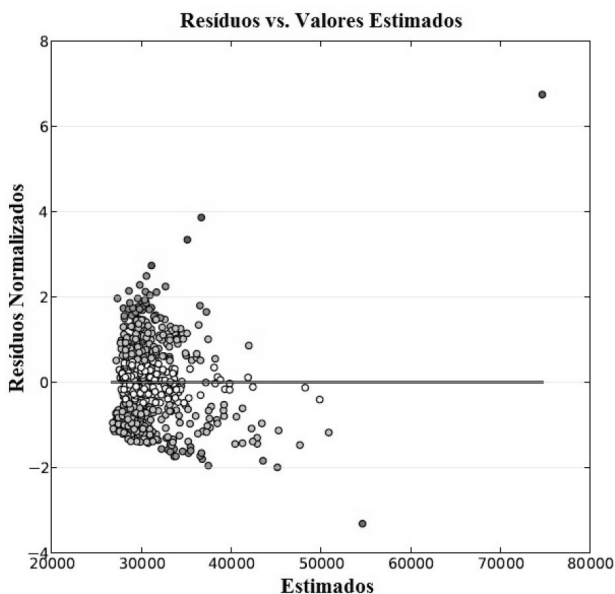


Figura 7 - Diagrama de dispersão dos resíduos para os valores de população estimados pelo modelo de regressão linear
Elaborado pelos Autores.

A distribuição espacial dos resíduos, apresentada no mapa dos resíduos da regressão (Figura 8), indicou a existência de *clusters* de regiões onde o modelo de regressão apresentou erro, tanto para mais quanto para menos, bastante considerável. A presença desses agrupamentos indica a influência do espaço nos resultados, ou seja, que os resultados são espacialmente dependentes. Neste caso, observa-se que existe uma concentração na região periférica da área de estudo, onde o erro associado é menor do que -1,5 desvios padrões. Percebem-se ainda regiões concentradas nas áreas mais urbanizadas da RMSP, onde o erro dos resíduos está entre 0,5 e 2,5 desvios padrões. Os agrupamentos espaciais observados no mapa dos resíduos indicam que o espaço está influenciando no resultado da regressão.

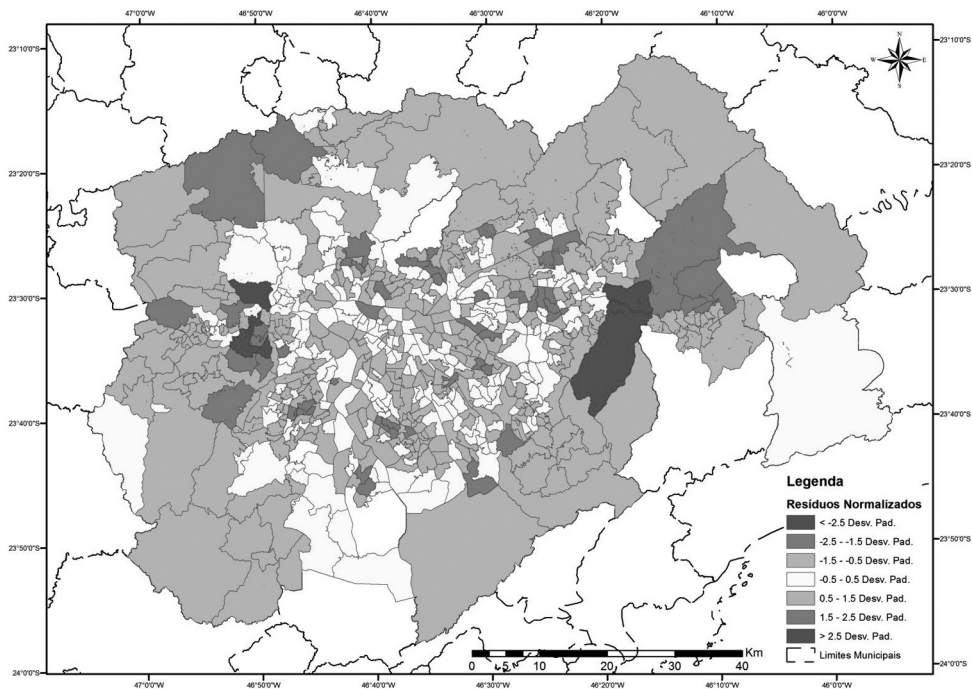


Figura 8 - Mapa da distribuição espacial dos resíduos

Fonte: IBGE, 2010. Elaborado pelos Autores.

A dependência espacial nas estimativas é corroborada com o valor de 0,45 de índice global de Moran para o estudo da dependência espacial dos resíduos. Ou seja, há indícios de que existe uma autocorrelação espacial positiva e estatisticamente significativa (p -valor = 0,01) entre as variáveis do modelo.

Explorando-se a autocorrelação espacial indicada pelo Índice Global de Moran e o Diagrama de Espalhamento de Moran dos resíduos (Figura 9a), percebe-se que os indícios de aglomerações espaciais observados no mapa dos resíduos (Figura 8) são ratificados no Box Mapa (Figura 9b), com associações predominantemente positivas nas áreas mais centrais e na região periférica da área de estudo.

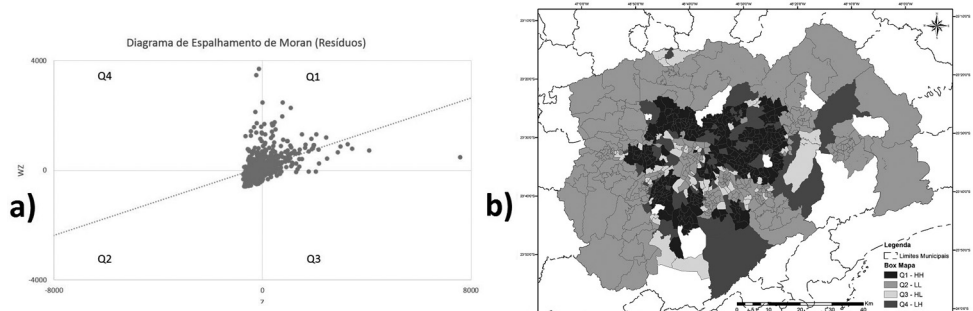


Figura 9 - a) Diagrama de Espalhamento de Moran dos Resíduos e b) Box Mapa dos Resíduos

Fonte: IBGE, 2010. Elaborado pelos Autores.

O resultado do Mapa de Moran dos resíduos (Figura 10) apresenta somente as áreas do Box Mapa cujos indicadores locais foram estatisticamente significantes, a 5% de significância.

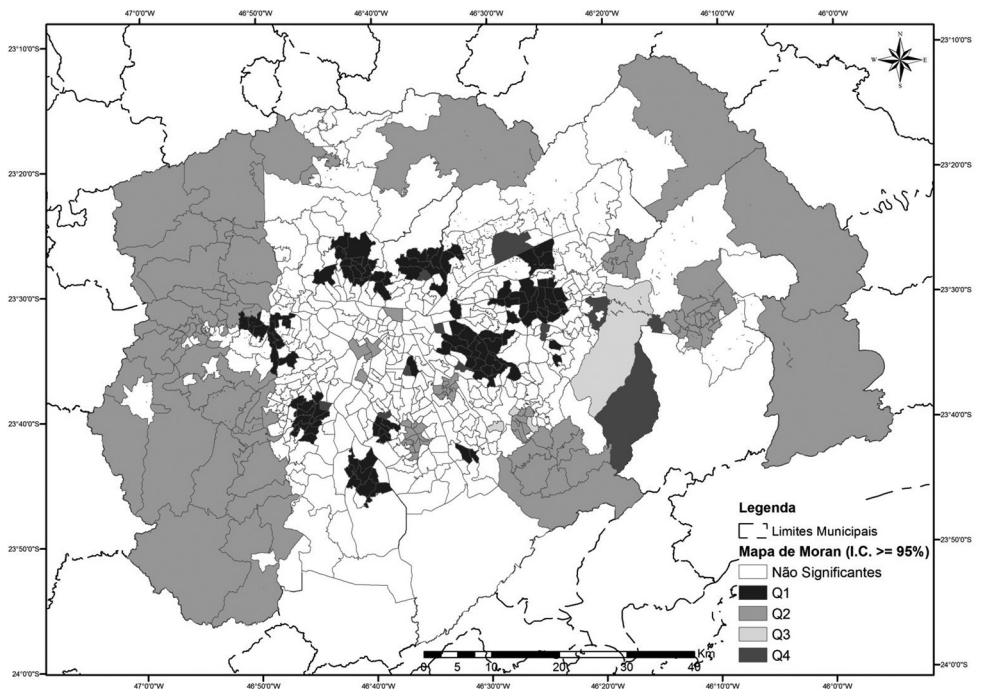


Figura 10 - Mapa de Moran dos resíduos da regressão linear simples

Fonte: IBGE, 2010. Elaborado pelos Autores.

Esta análise exploratória de regressão linear evidenciou a autocorrelação espacial dos resíduos, o que justificou realizar-se a regressão geograficamente ponderada. Ao visualizar-se a distribuição dos coeficientes de determinação locais (R^2) para as áreas de ponderação da região de estudo (Figura 11), observa-se que com os dados de luzes foi possível explicar 50% ou mais da variação da população, indicadas nas áreas mais escuras do mapa. O valor de R^2 global da regressão de 0,75 indica uma melhora significativa em relação à regressão linear simples, de R^2 igual a 0,10.

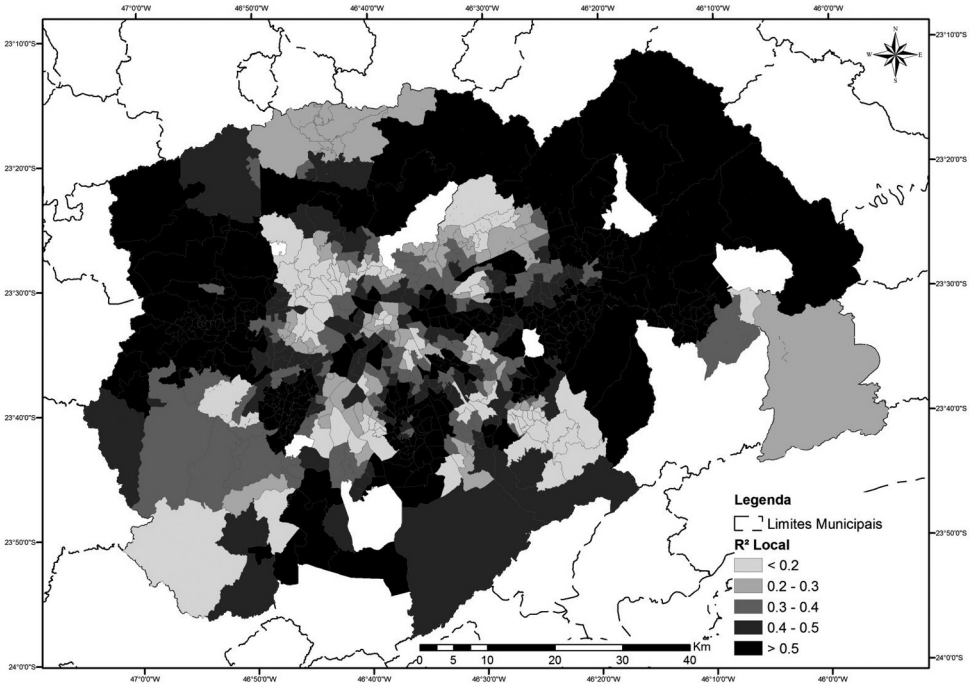


Figura 11 - Mapa dos coeficientes de determinação locais do modelo GWR

Fonte: IBGE, 2010. Elaborado pelos Autores.

A figura 12 apresenta o mapa dos coeficientes angulares locais (β_1) onde evidencia-se a variação espacial da relação entre os dados de luzes do VIIRS e o total da população. Apenas 32,8% dos coeficientes da variável SPL foram estatisticamente significantes, a 5% de significância, conforme apresentados na figura 13.

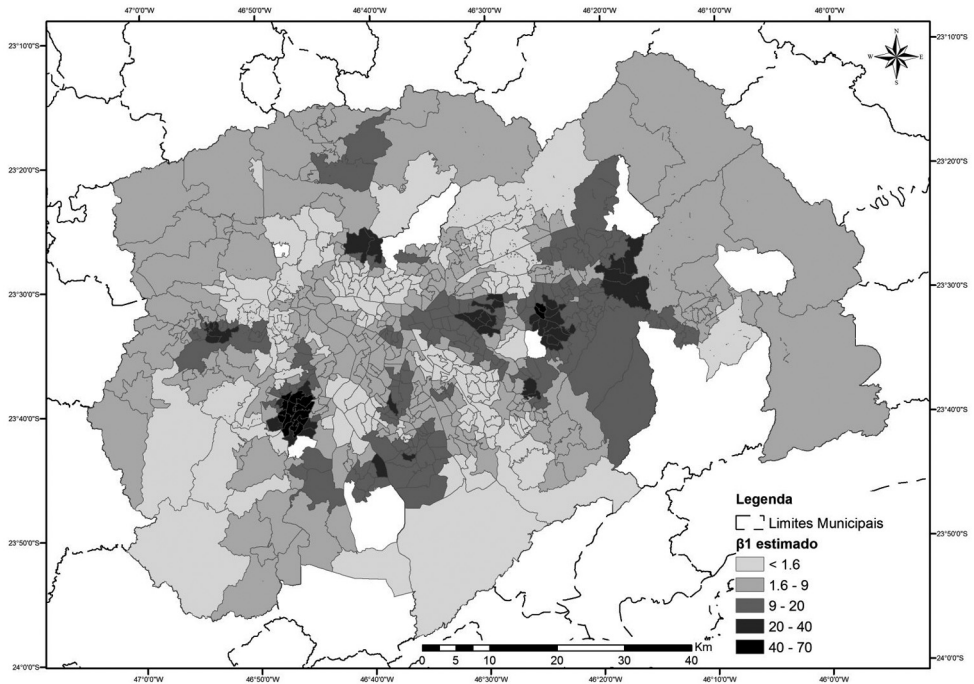


Figura 12 - Mapa de β_1 locais estimados

Fonte: IBGE, 2010. Elaborado pelos Autores.

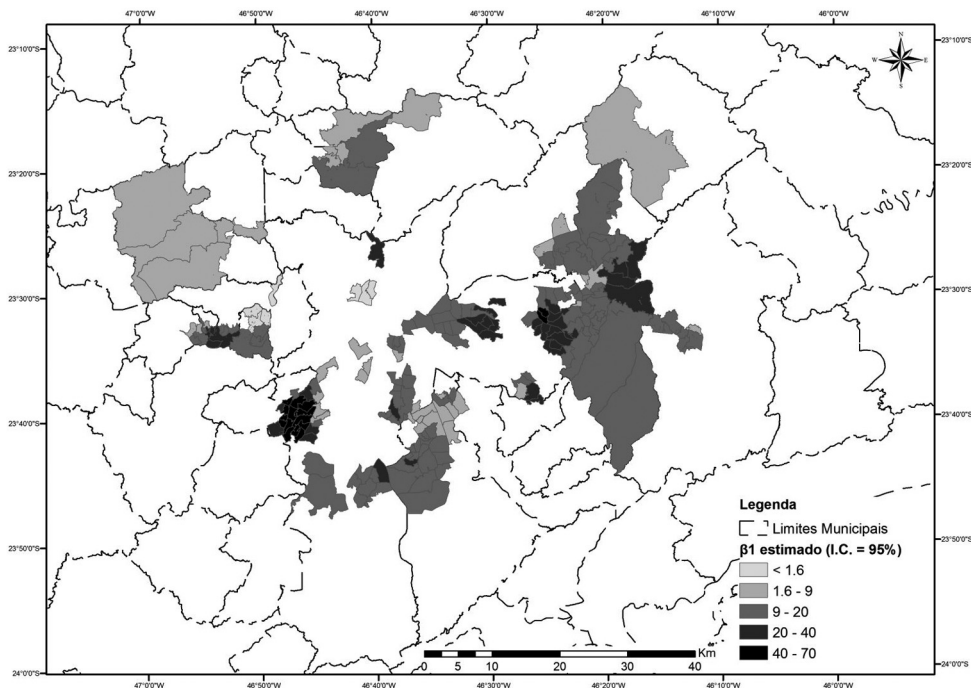


Figura 13 - Mapa de β_1 locais estimados considerando 5% de significância

Fonte: IBGE, 2010. Elaborado pelos Autores.

A variação espacial do β_1 local estimado indica diferentes padrões de ocupação urbana, isto é, onde o β_1 estimado é maior, existe um maior adensamento de ocupação urbana (poucas áreas abertas), e vice-versa, como mostra o exemplo da figura 14.

De modo geral, considerando-se que a data de obtenção das variáveis diferiu em dois anos (população referente ao censo 2010 e imagens de luzes do VIIRS relativas a 2012), a estimação da população a partir das imagens de luzes do VIIRS apresentaram resultados consistentes. Esta análise baseia-se no cálculo do erro relativo do valor de população estimado pelo modelo em comparação ao valor de população obtido do censo 2010 para cada unidade de ponderação. As duas áreas de ponderação com os maiores erros negativos (-40,38% e -39,50%) de estimativa, ou seja, onde o modelo subestimou a população, encontram-se no município de Francisco Morato. Estas áreas caracterizam-se por uma região construída predominantemente urbana e outra região maior desocupada. Estas áreas urbanas são, em sua maioria, caracterizadas como de baixa renda e carentes de investimento público em infraestrutura, equipamentos urbanos e projetos sociais (MORAES, 2007). Com a baixa qualidade da iluminação pública, a intensidade das luzes será menor do que o esperado para a quantidade de moradores da região, o que pode justificar a subestimação da população pelo modelo.

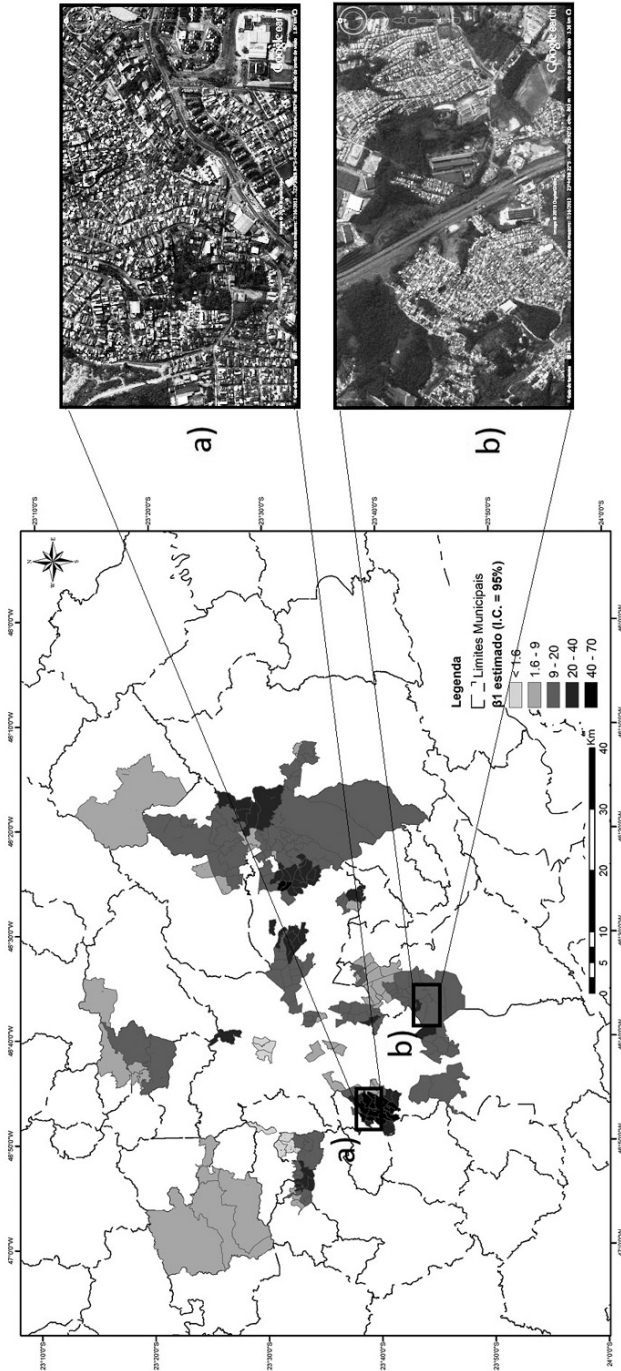


Figura 14 - Exemplo da associação entre β_1 local estimado e os diferentes padrões de ocupação urbana: a) Parte da região oeste do município de São Paulo; b) Parte da região norte do município de São Bernardo do Campo

Fonte: IBGE, 2010 / Google Earth, 2015. Elaborado pelos Autores.

Por sua vez, as duas áreas de ponderação que apresentaram os maiores erros positivos (116,43% e 105,12%), ou seja, locais em que o modelo superestimou a população, encontram-se no município de Cotia. Essa superestimativa pode ser atribuída ao fato de na região destas duas áreas de ponderação existirem diversos empreendimentos industriais que emitem luz durante a noite, contribuindo para um valor alto na estimativa, sem contudo possuírem população residente.

O erro geral das estimativas do modelo em relação ao valor observado no censo de 2010 foi de aproximadamente 5%. A porcentagem dos erros relativos das estimativas de população pela regressão espacial são apresentados estratificando-se em três classes (LI et al., 2013): de 0-25% considerando-se exatidão alta; de 25-50% exatidão moderada; e > 50% inexato (Tabela 1).

Tabela 1 - Níveis de exatidão da população estimada usando o modelo GWR

| Exatidão de erro relativo de população estimada por GWR | Áreas de ponderação (%) |
|---|-------------------------|
| Exatidão Alta (0 - 25%) | 77,87 |
| Exatidão Moderada (25 - 50%) | 18,09 |
| Inexato (> 50%) | 4,04 |

Do total das áreas de ponderação consideradas na regressão GWR, 77,87% das áreas apresentou um desvio de até 25% na estimativa de população em relação ao valor observado pelo censo, 18,09% apresentou um desvio entre 25% e 50% em relação ao valor observado. Por fim, apenas 4,04% das áreas de ponderação apresentou desvios maiores do que 50%. Este resultado indica que os dados das imagens de luzes noturnas VIIRS apresentaram bom desempenho em estimar a população, considerando os dados utilizados e a variabilidade espacial deste trabalho.

Utilizando regressão linear simples Sutton et al. (2001) estimaram a população de diversos países a partir de imagens de luzes noturnas do sensor OLS. O coeficiente de determinação (R^2) para o Brasil foi de 0,87, superestimando a população brasileira em 19%. Apesar dos valores de R^2 obtidos neste trabalho para a RMSD terem sido menores (0,10 e 0,75 utilizando regressão linear simples e GWR, respectivamente), o erro das estimativas, em geral, foi de aproximadamente 5%.

Também por regressão linear simples, Amaral et al. (2006) obtiveram R^2 maior que 0,82 entre as luzes noturnas do sensor OLS e os valores de população urbana para os anos de 1996 e 2000, considerando todos os municípios da Amazônia Legal Brasileira. Ainda que os contextos das áreas de estudo sejam muito diferentes, e as escalas e unidades de análise incomparáveis, o resultado de coeficientes de determinação foram semelhantes aos deste trabalho.

Ainda que outros estudos tenham utilizado imagens de luzes noturnas para estimar a população de uma dada região (p. ex. ELVIDGE et al., 1997; SUTTON et al., 2001; AMARAL et al., 2005; 2006) até o momento não se verificou resultados para abordagens na escala da unidade de análise deste trabalho, e tampouco estudos utilizando imagens de luzes noturnas VIIRS, por se tratar de um sensor relativamente recente.

CONSIDERAÇÕES FINAIS

Este trabalho realiza uma abordagem baseada em estatística espacial para avaliar o potencial das imagens de luzes do VIIRS em estimar a população de pequenas áreas, consideradas intraurbanas neste caso. A utilização de uma técnica de regressão linear simples não foi suficiente para observar alguma relação entre as imagens de luzes noturnas do VIIRS e os dados de população oficiais. No entanto, ao se considerar a heterogeneidade espacial incluída na técnica de regressão através do emprego da Regressão Geograficamente Ponderada (GWR), apresentam-se algumas contribuições relevantes:

- a) Ao se modelar fenômenos distribuídos no espaço deve-se verificar a existência de autocorrelação espacial das variáveis para que os resultados não sejam viesados. Neste trabalho, a análise dos resíduos da regressão linear simples indicou a presença da autocorrelação espacial, o que permitiu modelar a relação dos dados de luzes noturnas como inferência de população urbana pelo modelo de regressão espacial GWR.
- b) Observou-se relações entre as variáveis luzes noturnas e população residente, indicado por um R^2 global de 0,75 e por valores de população estimada consistentes com os dados censitários oficiais observados, ainda que somente um terço dos coeficientes tenham se apresentado estatisticamente significantes.
- c) Das análises dos mapas dos coeficientes da GWR, verificou-se que as intensidades desses coeficientes estão diretamente relacionadas aos padrões de adensamento populacional no espaço.

Da observação de que apenas 32,8% dos coeficientes da regressão terem sido estatisticamente significativos, sugere-se a alteração da unidade de análise para outra mais generalizada (p. ex. distritos da RMSP) para uma melhor adequação do modelo, mantendo ainda a escala local e intraurbana.

Finalmente, este trabalho indicou uma possibilidade do uso de dados de luzes do VIIRS para estimativas populacionais em escalas locais. Contribui-se assim para registrar o potencial das imagens de luzes noturnas VIIRS tanto em escala global, quanto em escala local, podendo auxiliar os modelos de estimativa anual da população, desde que sejam consideradas suas particularidades e heterogeneidade espacial, de acordo com a escala de análise e peculiaridades regionais.

REFERÊNCIAS

- AKAIKE, H. Information theory and an extension of maximum likelihood principle. In: PETROV B. N.; CSAKI, F. (Eds.). **Second International Symposium on Information Theory**. Budapest: Akademiai Kiado, 1973, p. 267-281.
- AMARAL, S.; CÂMARA, G.; MONTEIRO, A. M. V.; QUINTANILHA, J. A.; ELVIDGE, C. D. Estimating population and energy consumption in Brazilian Amazonia using DMSP night-time satellite data. **Comput. Environ. Urban Syst.**, v. 29, n. 2, p. 179-195, 2005.
- AMARAL, S.; MONTEIRO, A. M. V.; CAMARA, G.; QUINTANILHA, J. A. DMSP/OLS night-time light imagery for urban population estimates in the Brazilian Amazon. **Int. J. Remote Sens.**, v. 27, n. 5, p. 855-870, 2006.
- ANSELIN, L. Local Indicators of Spatial Association - LISA. **Geogr. Anal.**, v. 27, n. 2, p. 93-115, 1995.

ANSELIN, L. The Moran scatterplot as ESDA tool to assess local instability in spatial association. In: FISHER, M.; SCHOLTEN, H. J.; UNWIN, D. **Spatial analytical perspectives on GIS**. London: Taylor & Francis, 1996, p. 111-125.

BAUGH, K.; HSU, F. C.; ELVIDGE, C.; ZHINZHIN, M. Nighttime lights composing using the VIIRS Day-Night Band: preliminary results. **Proceedings of the Asia-Pacific Advanced Network**, v. 35, p. 70-86, 2013.

BRUNSDON, C.; FOTHERINGHAM, A. S.; CHARLTON, M. E. Geographically Weighted Regression: a method for exploring spatial nonstationarity. **Geogr. Anal.**, v. 28, n. 4, p. 281-298, 1996.

CAO, C.; XIONG, X.; WOLFE, R.; DeLUCCIA, F.; LIU, Q.; BLONSKI, S.; LIN, G.; NISHIHAMA, M.; POGORZALA, D.; OUDRARI, H. **Visible/Infrared Imager Radiometer Suite (VIIRS) Sensor Data Record (SDR) user's guide**. Washington: NOAA, 2013. 40 p. (Technical Report NESDIS 142). Disponível em: <https://cs.star.nesdis.noaa.gov/pub/NCC/UsersGuideVIIRS/VIIRS_SDR_USERS_GUIDE_NOAA_TechReport142.pdf>. Acesso em: 02.jun.2013.

CAO, C.; De LUCCIA, F.; XIONG, X.; WOLFE, R.; WENG, F. Early on-orbit performance of the Visible Infrared Imaging Radiometer Suite (VIIRS) onboard the Suomi National Polar-orbiting Partnership (S-NPP) satellite. **IEEE Trans. Geosci. Remote Sens.**, v. 52, n. 2, p. 1142-1156, 2014.

CROFT, T. A. Nighttime images of the earth from space. **Sci. Am.**, v. 238, p. 68-79, 1978.

DOLL, C. N. H. Population detection profiles of DMSP-OLS night-time imagery by regions of the world. In: Asia-Pacific Advanced Network Meeting, 30., 2010, Hanoi, Vietnam. **Proceedings...** Hanoi: APAN, 2010, p. 191-207.

DOLL, C. N. H.; PACHAURI, S. Estimating rural populations without access to electricity in developing countries through night-time light satellite imagery. **Energy Policy**, v. 38, n. 10, p. 5661-5670, 2010.

ELVIDGE, C. D.; BAUGH, K. E.; KIHN, E. A.; KROEHL, H. W.; DAVIS, E. R. Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. **Int. J. Remote Sens.**, v. 18, n. 6, p. 1373-1379, 1997.

ELVIDGE, C. D.; IMHOFF, M. L.; BAUGH, K. E.; HOBSON, V. R.; NELSON, I.; SAFRAN, J.; DIETZ, J. B.; TUTTLE, B. T. Night-time lights of the world: 1994-1995. **ISPRS J. Photogramm. Remote Sens.**, v. 56, n. 2, p. 81-99, 2001.

FOSTER, J. L. Observations of the Earth using nighttime visible imagery. In: Optical Engineering for Cold Environments, 0414., 1983, Arlington, USA. **Proceedings...** Arlington: SPIE, 1983. p. 187-193.

FORSTER, B. Some urban measurements from Landsat data. **Photogramm. Eng. Remote Sens.**, v. 49, n. 12, p. 1693-1707, 1983.

FOTHERINGHAM, A. S.; CHARLTON, M.; BRUNSDON, C. Two techniques for exploring non-stationarity in geographical data. **Geogr. Sys.**, v. 4, p. 59-82, 1997.

GONÇALVES, C. D. A. B.; SOUZA, I. M.; PEREIRA, M. N.; FREITAS, C. C. **Análise do ambiente urbano visando a inferência populacional a partir do uso de dados de sensoriamento remoto orbital de alta resolução**. São José dos Campos: INPE, 2004, 65 p. (INIPE-10459-RPQ/754). Disponível em: <<http://mtc-m12.sid.inpe.br/col/sid.inpe.br/sergio/2004/03.15.09.33/doc/publicacao.pdf>>. Acesso em: 05 jun. 2014.

IBGE. **Censo Demográfico**. 2010. Disponível em: <<http://censo2010.ibge.gov.br/>>. Acesso em: 17 fev.2014.

- IBGE. **Estimativas de População**. 2012. Disponível em: <<http://www.ibge.gov.br/home/estatistica/populacao/estimativa2012/default.shtm>>. Acesso em: 12.jun.2015.
- JENSEN, J. R.; COWEN, D. C. Remote sensing of urban/suburban infrastructure and socio-economic attributes. **Photogramm. Eng. Remote Sens.**, v. 65, n. 5, p. 611-622, 1999.
- LEE, T. E.; MILLER, S. D.; TURK, F. J.; SCHUELER, C.; JULIAN, R.; DEYO, S.; DILLS, P.; WANG, S. The NPOESS VIIRS day/night visible sensor. **Bull. Am. Meteorol. Soc.**, v. 87, n. 2, p. 191-199, 2006.
- LETU, H.; HARA, M.; YAGI, H.; NAOKI, K.; TANA, G.; NISHIO, F.; SHUHEI, O. Estimating energy consumption from night-time DMSP/OLS imagery after correcting for saturation effects. **Int. J. Remote Sens.**, v. 31, n. 16, p. 4443-4458, 2010.
- LI, X.; XU, H.; CHEN, X.; LI, C. Potential of NPP-VIIRS nighttime light imagery for modeling the regional economy of China. **Remote Sens.**, v. 5, p. 3057-3081, 2013.
- MA, T.; ZHOU, C.; PEI, T.; HAYNIE, S.; FAN, J. Quantitative estimation of urbanization dynamics using time series of DMSP/OLS nighttime light data: a comparative case study from China's cities. **Remote Sens. Environ.**, v. 124, p. 99-107, 2012.
- MORAES, C. A. **A periferização da pobreza e da degradação sócio-ambiental na Região Metropolitana de São Paulo, o caso de Francisco Morato**. 2007. 142 p. Dissertação (Mestrado em Planejamento Urbano e Regional) - Faculdade de Arquitetura e Urbanismo da Universidade de São Paulo, São Paulo, 2007.
- NAKAYA, T.; CARLTON, M.; LEWIS, P.; BRUNSDON, C.; YAO, J.; FOTHERINGHAM, S. **GWR4 User Manual**. Arizona, US: ASU, 2014. 39 p. Disponível em: <https://geodacenter.asu.edu/drupal_files/gwr/GWR4manual.pdf>. Acesso em: 10.jun.2015.
- NASA. **Defense Meteorological Satellites Program (DMSP) series**. 2003. Disponível em: <<http://heasarc.nasa.gov/docs/heasarc/missions/dmsp.html>>. Acesso em: 21 fev. 2014.
- SCHNEIDER, A.; FRIEDL, M. A.; POTERE, D. A new map of global urban extent from MODIS satellite data. **Environ. Res. Lett.**, v. 4, n. 4, p. 1-11, 2009.
- SMALL, C.; ELVIDGE, C. D. Mapping decadal change in anthropogenic night light. **Procedia Environ. Sci.**, v. 7, p. 353-358, 2011.
- SUTTON, P.; ROBERTS, D.; ELVIDGE, C.; BAUGH, K. Census from Heaven: An estimate of the global human population using night-time satellite imagery. **Int. J. Remote Sens.**, v. 22, n. 16, p. 3061-3076, 2001.
- UNITED NATIONS. **World urbanization prospects: the 2011 revision**. New York: United Nations, 2012.
- XIONG, X.; CHIANG, K.; McINTIRE, J.; OUDRARI, H.; WU, A.; SCHWALLER, M.; BUTLER, J. Early assessment of VIIRS on-orbit calibration and support activities. In: 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2012, Munich, Germany. **Proceedings...** Munich: IEEE, 2012, p. 7189-7192.
- ZHANG, Q.; SETO, K.C. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. **Remote Sens. Environ.**, v. 115, n. 9, p. 2320-2329, 2011.

Recebido em novembro de 2015

Revisado em março de 2016

Aceito em abril de 2016