

# ANÁLISE DO USO DOS DADOS DE LUZES NOTURNAS PARA ESTIMATIVA E DISTRIBUIÇÃO ESPACIAL DA POPULAÇÃO: ESTUDO PARA A REGIÃO METROPOLITANA DO VALE DO PARAÍBA E LITORAL NORTE, SÃO PAULO, BRASIL.

Ana Carolina de Faria Santos<sup>1</sup>

<sup>1</sup>Instituto Nacional de Pesquisas Espaciais - INPE  
Caixa Postal 515 - 12227-010 - São José dos Campos - SP, Brasil  
ana.faria@inpe.br

**Abstract.** The process of urbanization has intensified during this century, mainly with the increase of the population and increase of the urban use and coverage. Nighttime Light data have great potential to measure a spatial distribution of the urban area and can be correlated with various types of data and spatial modeling to generate products that support urban studies, such as population estimates. In this context, the objective of this work was to analyze the relationship between Nighttime Light data and population data for population estimation, using Linear (OLS) and Geographically Weighted Regression (GWR) Models, and to verify the influence of the application of land cover data in this estimation. The study was carried out at a regional level for a Metropolitan Region of the Vale do Paraíba and North Coast and at municipal level for São José dos Campos, being divided into three phases: Phase 0 that corresponded to the elaboration of a map with use of Land Use and Cover for the municipality and the aggregation of data in a cellular space of 5km<sup>2</sup> at a regional level and 1km<sup>2</sup> at a municipal level; Phase 1, at regional and municipal level, verified a relationship of Nighttime Light data and population data through the steps of Exploratory Analysis, OLS, Spatial Analysis and GWR; and Phase 2, which at a municipal level, verified an influence of the variable of land use and cover through the same steps of Phases 1. The study demonstrated that the three types of data are highly correlated, showing spatial autocorrelation in all phases and injuring as premises of OLS, so the GWR Models was more satisfactory in estimating the population. The insertion of Land Use and Cover data was significant and increased the robustness of the models, however, an overestimation and underestimation of the data were observed, so it is necessary to manage with other types of aggregation in order to minimize these impacts.

**Palavras-chave:** Estimativa Populacional. Luzes Noturnas. VIIRS. Uso e Cobertura do solo.

## 1. Introdução

Estima-se que o processo de urbanização tem crescido durante o século 21, sendo a expansão urbana e a densidade populacional seus principais representantes. Espera-se uma duplicação da população até meados deste século, elevando a concentração de pessoas e de atividades socioeconômicas nos centros urbanos. O aumento destes processos inclui também uma mudança na paisagem, podendo levar a um crescimento de até três vezes na cobertura urbana, gerando diversos impactos, não apenas socioeconômicos como também ambientais (OLIVEIRA et al., 2007; BAGAN E YAMAGATA, 2015; HUILEI et al., 2017).

Conforme a urbanização avança, pode causar alterações que modificam os sistemas naturais, motivando o desenvolvimento antrópico e as aceleradas mudanças da sociedade das últimas décadas, tornando os estudos sobre crescimento urbano extremamente importantes ante aos impactos gerados (TUCCI E CLARKE, 1997; BAGAN E YAMAGATA, 2015).

Técnicas de Geoprocessamento associadas a dados de sensoriamento remoto têm sido utilizadas para auxiliar nestes estudos de população, tornando possível, por exemplo, a estimativa da população, informação de extrema importância, principalmente em locais com dados escassos intercensos (DORIA et al., 2016). Diversas pesquisas têm usado imagens de luzes noturnas, em correlação com outros tipos de dados, para estimar a população, como Amaral et al. (2005) através da correlação entre dados de luzes noturnas e energia elétrica para a Amazônia, bem como Letu et al. (2010) o fizeram para a China e a Índia. Doria et al (2016) estimou a população para a Região Metropolitana de São Paulo através da relação dos dados de luzes noturnas com os de população do Censo Demográfico, bem como Bagan e

Yamagata (2015) que estimaram a população do Japão através de uma relação entre os dados de luzes noturnas e de uso e cobertura da terra com dados de população.

Alguns modelos estatísticos são amplamente utilizados para descrever estas relações, como a análise de Regressões Múltiplas utilizadas para modelagem espacial, na qual é possível quantificar a relação entre uma ou mais variáveis, de modo que uma variável resposta (dependentes) possa ser explanada pelas variáveis explicativas (independentes). A Regressão Linear é uma técnica de estimação com parâmetros globais que pressupõe a independência dos dados e sua homocedasticidade, ou seja, uma variância sempre constante. Todavia, quando este modelo é aplicado a dados espaciais, estes pressupostos são violados devido à incorporação do efeito espacial na distribuição dos dados. De forma a refletir as características espaciais da relação entre as variáveis, é utilizada a Regressão Geograficamente Ponderada, a qual permite estimar parâmetros locais ao computar em seu modelo coeficientes para cada localização, evitando a dependência espacial e distribuindo os resíduos de forma mais aleatória (DORIA et al., 2016; HUILEI et al., 2017).

Neste contexto, este trabalho visa analisar a relação dos dados de luzes noturnas e dados populacionais para estimativa populacional, utilizando Modelos de Regressão Linear e Geograficamente Ponderada, e verificar a relação e influencia da inserção de dados de uso e cobertura do solo nesta estimativa.

## 2. Material e Métodos

### 2.1. Área de Estudo

A Região Metropolitana do Vale do Paraíba e Litoral Norte (RMVPLN) do estado de São Paulo foi criada em 2012 e é composta por 39 municípios divididos em cinco (5) sub-regiões. Situada no eixo que liga duas das principais regiões metropolitanas do país, São Paulo e Rio de Janeiro, a RMVPLN possui uma área total de 16.178 km<sup>2</sup> e concentra 5,2% do Produto Interno Bruto (PIB) paulista (EMPLASA, 2017).

O município de São José dos Campos, considerado o 7º município mais populoso do Estado de São Paulo e o mais populoso da RMVPLN, com 629.921 habitantes segundo o Censo Demográfico 2010 (IBGE, 2017). Atualmente seu território abrange uma área de 1.099,409 km<sup>2</sup>, sendo que aproximadamente 30% de sua área é formada por um planalto onde se concentra toda a população urbana do Município (SOUZA, 2004)

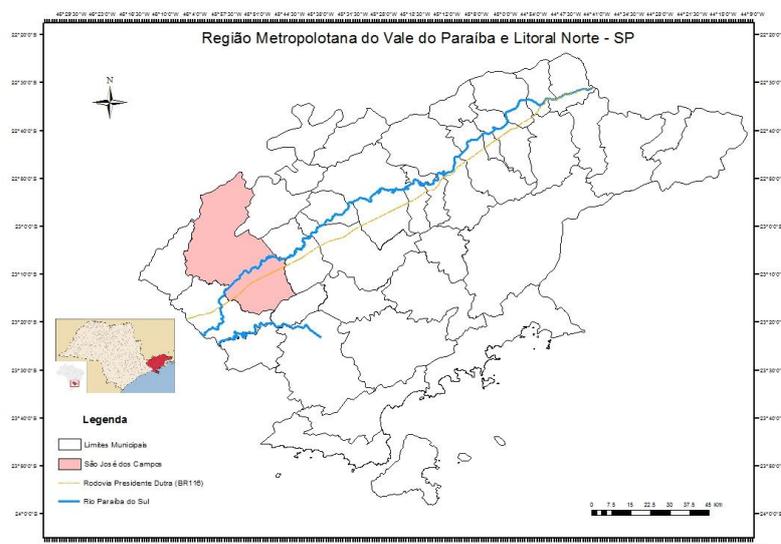


Figura 1. Região Metropolitana do Vale do Paraíba e Litoral Norte, no Estado de São Paulo, com destaque para o Município de São José dos Campos.

## 2.2. Dados

O sensor VIIRS opera desde 2011 junto ao satélite *Suomi National Polar-orbiting Partnership* (S-NPP), lançado pelo Programa *Joint Polar Satellite System* (JPSS), realizando o imageamento diurno e noturno da Terra. O VIIRS tem como objetivo principal a observação de nuvens e de variáveis da superfície da Terra, sendo que cada imagem recobre aproximadamente 3.000 km e provem imagens diárias do planeta. Este sensor possui 22 bandas espectrais, sendo que entre elas está a *Day/Night Band* (DNB) que abrange a faixa espectral de 0,5 a 0,9  $\mu\text{m}$ , sendo responsável pela aquisição dos dados de Luzes Noturnas.

Neste estudo foi utilizada a Versão 1 da Composição Mensal dos dados de Luzes Noturnas, produzida pelo *Earth Observations Group* (EOG), a qual é um produto de resolução espacial de aproximadamente 500 metros que contém a média dos valores de radiância do respectivo mês (nanoWatt/cm<sup>2</sup>/sr). Utilizou-se a composição referente ao mês de Julho de 2012, por ser a com menor cobertura de nuvens para a área de estudo.

Também foram utilizados o dado de população total advindo do Censo Demográfico 2010, produzidos pelo Instituto Brasileiro de Geografia e Estatística (IBGE), e dados temáticos de Uso e Cobertura do Solo produzidos especificamente para este estudo. Para a elaboração do Mapa Temático de Uso e Ocupação do Solo foram utilizadas seis cenas, do ano de 2012 (Tabela 1), do Sistema Alemão *Rapid Eye* para compor toda a área de estudo, as quais são disponibilizadas pelo Geocatálogo do Ministério do Meio Ambiente (MMA, 2017). Este sistema é composto por uma constelação de cinco satélites cujos sensores possuem cinco bandas espectrais, resolução radiométrica de 12 bits, resolução espacial de 5 metros (ortorretificadas, nível 3A), já devidamente registradas.

Tabela 1. Cenas do Sistema *Rapid Eye* utilizadas para compor a área de estudo.

Cena	Data de aquisição	Identificação do Satélite	Bandas
2328310	05/12/2012	RE4	
2328311	03/07/2012	RE2	Azul (440-510 nm)
2328410	17/08/2012	RE4	Verde (520-590 nm)
2328411	03/07/2012	RE2	Vermelho (630-690 nm) Red Edge (690-730 nm)
2328510	03/07/2012	RE2	Infravermelho Próximo (760-880 nm)
2328511	03/07/2012	RE2	

## 2.3. Metodologia

Este estudo foi dividido em três fases: (1) a Fase 0 corresponde à preparação dos dados e elaboração do mapa temático de Uso e Cobertura do Solo do Município de São José dos Campos; (2) a Fase 1, que corresponde à estimativa da população, a nível Regional e Municipal, utilizando a relação entre os dados de Luzes Noturnas e de População; e (3) a Fase 2, que corresponde à estimativa da população, apenas a nível Municipal, através da relação dos dados de luzes, de população e adicionando os de uso e cobertura do solo.

### 2.3.1. Fase 0

Para a elaboração do mapa temático, foi primeiramente criada uma chave de interpretação, que consiste na descrição de um conjunto de elementos que permitem distinguir determinados alvos na imagem de forma precisa e objetiva (SOUZA, 2004). A partir de um prévio conhecimento dos alvos nas imagens e visando caracterizar os adensamentos urbanos da área de estudo, foram determinadas as sete classes: Urbanização Densa (UD), Urbanização

Média (UM), Urbanização Baixa (UB), Área Industrial/Comercial (AI), Vegetação (VEG), Agrossilvipastoril (ASP) e Água (AG).

Assim, para a elaboração do mapa temático no software Spring 5.5.2, primeiramente foi realizado o processo de segmentação com o algoritmo de Crescimento de Regiões, utilizando limiar de similaridade de 25 e limiar de área de 50, os quais foram definidos empiricamente por particionar a área em regiões homogêneas. Posteriormente para a classificação, foram selecionadas para cada classe 20 amostras de aquisição e realizada uma classificação supervisionada por regiões utilizando o classificador Bhattacharyya com um limiar de aceitação de 99,9%. Apesar desta técnica de classificação ser bastante empregada e apresentar resultados confiáveis, foram constatadas dificuldades na classificação automática das classes UD, UM, UB e AI, devido a pouca diferença espectral entre elas, portanto estas classes foram digitalizadas visualmente, com o auxílio das próprias imagens do RapidEye, com o emprego de pistas visuais, como tonalidade, textura, forma, padrão e outras.

Uma etapa de pós-classificação foi necessária para a retirada de ruídos ou pequenos grupos de pixels que não pertenciam a classe que foram erroneamente classificados, essa edição foi realizada no software ArcGIS 10.2 através de interpretação visual. Por último, um mosaico foi criado com o objetivo de unir as seis cenas do sistema RapidEye para a formação do mapa temático de uso e ocupação do solo do município de São José dos Campos.

Com todos os dados adquiridos, os mesmos foram agregados a nível celular, sendo a nível regional com células de 5 km<sup>2</sup> e a nível municipal com células de 1 km<sup>2</sup>, valores estes determinados empiricamente e gerados pelo Software TerraView 5.2. Para os dados de luzes foi utilizada a variável Soma de Luzes Noturnas que corresponde à soma dos valores de radiância presentes em cada unidade celular. Os dados populacionais, fornecidos como população total por setor censitário, foram agregados de acordo com a Soma Ponderada pela Área dos setores em relação a unidade celular. Os dados temáticos foram utilizados de acordo com a porcentagem de cada classe na unidade celular. A variável dependente considerada nos modelos foi a população, e as variáveis independentes foram as Luzes Noturnas e as classe de uso e ocupação do solo.

### 2.3.2. Fases 1 e 2 - Etapas

As fases 1 e 2 foram compostas de quatro etapas cada uma, diferenciadas apenas pelas variáveis inseridas nos modelos. As etapas são: (a) Análise Exploratória dos dados; (b) Regressão Linear; (c) Análise Espacial; e (d) Regressão Geograficamente Ponderada. As três primeiras etapas foram realizadas no Software GeoDa e a última foi realizada no Software GWR4.

A etapa (a) consistiu em analisar a relação entre os dados para verificar se realmente possuíam uma correlação significativa, a qual foi feita através do diagnóstico de histogramas e gráficos de dispersão das variáveis que poderiam indicar uma relação linear.

A etapa (b) é a verificação da correlação, através da análise do Coeficiente de Determinação ( $R^2$ ), de modo a verificar o comportamento da estimativa através da estimação global dos parâmetros. O modelo da Regressão Linear utilizado neste estudo é representado pela Equação 1:

$$\text{Equação (1):} \quad Y_i = \beta_0 + \beta_1 \times X_i + \varepsilon$$

Onde  $Y_i$  é a variável dependente da  $i$ -ésima unidade celular,  $\beta_0$  é o coeficiente linear,  $\beta_1$  é o coeficiente angular,  $X_i$  são as variáveis independentes na  $i$ -ésima unidade celular e  $\varepsilon$  é o resíduo.

A etapa (c) representa a análise, principalmente, dos resíduos da etapa (b), de forma a avaliar se há uma estruturação espacial que contrarie os pressupostos da independência e da

homocedasticidade. Esta análise é feita através dos gráficos dos resíduos *versus* os valores estimados de população e pela espacialização dos resíduos para verificar a interferência do espaço e a presença de agrupamentos de erros positivos e negativos na área de estudo. Para mensurar e identificar estes agrupamentos foi realizada a análise do *Local Indicators of Spatial Association* (LISA) e calculado do Índice de Moran (*I*).

O Índice de Moran quantifica o grau de correlação espacial existente, variando seu valor -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. A interpretação deste índice está associada ao Diagrama de Espalhamento de Moran (ANSELIN, 1996), que relata espacialmente o relacionamento entre os valores do vetor de desvios e os valores das médias locais, indicando diferentes regimes espaciais presentes nos dados. O LISA apresenta a localização espacial dos agrupamentos e sua significância.

Por fim, se verificada a autocorrelação espacial, a etapa (d) consistirá na aplicação do Modelo de Regressão Geograficamente Ponderado (GWR), de maneira a ajustar o modelo de acordo com as observações mais próximas ao ponto avaliado. O modelo de regressão utilizado pelo GWR está representado pela Equação 2:

$$\text{Equação (2):} \quad Y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i) \times X_i + \varepsilon$$

Onde  $(u_i, v_i)$  representam as coordenadas geográficas de cada um dos centroides da unidade celular.

A realização do modelo de regressão local foi realizada no Software GWR4, no qual foi utilizado uma Função Gaussiana com largura de banda adaptável, isto é, largura menor em áreas com alta densidade de dados e maiores em áreas com menor densidade de dados, cuja largura de banda foi selecionada com auxílio do próprio software. Para avaliar qual o melhor modelo para os dados observados, avaliaram-se os valores de  $R^2$  Ajustado, que quanto mais elevado melhor a habilidade das variáveis explicativas interpretarem as variáveis respostas, e o *Akaike Information Criterion* (AIC), que quanto menor seu valor, melhor a capacidade do modelo em descrever os dados observados.

### 3. Resultados e Discussão

#### 3.1. Fase 1 – Estimativa Regional e Municipal da População

No nível Regional, observando o gráfico de dispersão (Figura 2A) percebe-se a relação entre os dados de luzes noturnas e população foi considerada bem significativa, com um  $R^2$  de 0,835. Analisando os histogramas das variáveis, notou-se que ambas possuem uma frequência similar, indicando uma relação espacial forte, fato que pode ser visualizado no Cartograma (Figura 2B), no qual os círculos maiores e vermelhos representam regiões com maiores somas de luz e maior quantidade de pessoas, os quais se concentram no eixo da Rodovia Presidente Dutra e nos centros urbanos.

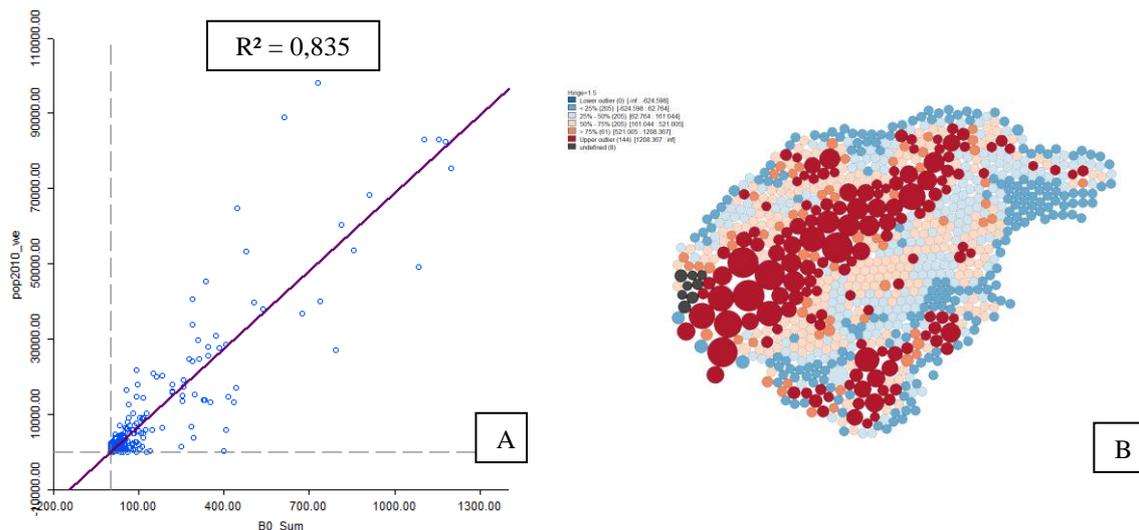


Figura 2. Relação entre os dados de luzes noturnas e os dados populacionais para a RMVPLN. A: Gráfico de dispersão das variáveis com seu respectivo Coeficiente de Determinação ( $R^2$ ); B: Cartograma das variáveis no qual o tamanho do círculo indica a variação de luzes e as cores indicam a variação da população.

Verificando-se essa alta correlação entre os dados, realizou-se a etapa da Regressão Linear, a qual teste teve uma significância estatística, com um  $R^2$  Ajustado de 0,834 e um AIC de 16140,97. Os resultados também demonstram a heterocedasticidade dos resíduos através do Teste Breusch-Pagan, o qual realiza uma regressão dos quadrados dos resíduos contra os quadrados dos valores estimados de Y, rejeitando a hipótese nula de que os resíduos possuem uma variância constante, o que pode ser melhor observado na nuvem de pontos do gráfico de dispersão dos resíduos em relação a variável predita (Figura 3B).

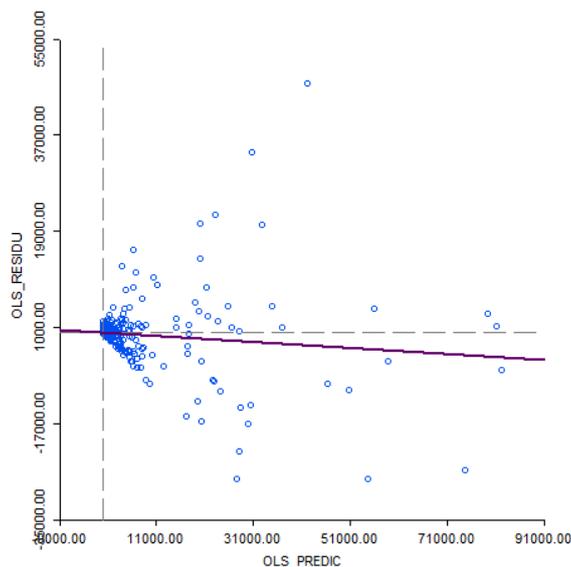


Figura 3. Gráfico de dispersão dos resíduos da Regressão Linear para a Região Metropolitana do Vale do Paraíba.

Considerando que os testes indicaram que a distribuição dos resíduos fere os pressupostos da Regressão Linear, foi realizada a etapa de Análise Espacial, na qual se obteve um Índice de Moran pouco significativo de 0,048 com um Diagrama de Espalhamento com uma nuvem de pontos aglomerada no centro (Figura 4A), indicando que os resíduos possuíam uma possível distribuição aleatória. Todavia, analisando espacialização do indicador LISA (Figura 4B)

nota-se agrupamentos significantes nos centros urbanos e próximo ao eixo da Rodovia Presidente Dutra, de modo que se cogitou realizar a etapa da GWR.

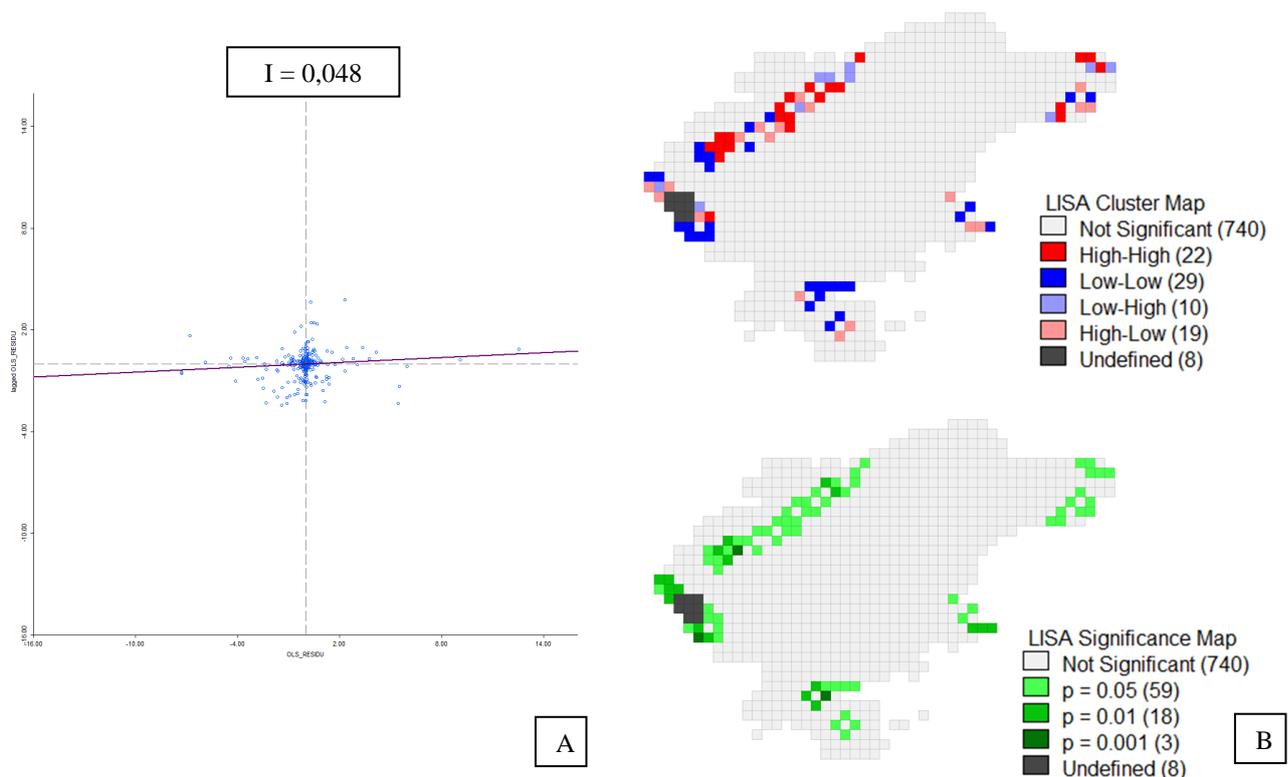


Figura 4. Análise espacial dos resíduos da Regressão Linear da RMVPLN. A: Diagrama de Espalhamento de Moran e seu respectivo índice; B: Mapa de Agrupamento e Mapa de Significância do indicador LISA.

Analisando os resultados estatísticos da regressão GWR notou-se que houve uma melhora pouco significativa do modelo em descrever as variáveis e dados observados, fato que pode ser verificado pelo baixo aumento de 0,015 do  $R^2$  ajustado (0,849) e uma leve diminuição de 55,95 do AIC (16085,02) em relação à Regressão Linear. Considerando que os pressupostos para realizar a Regressão Linear foram quebrados, apesar da pequena melhora, o modelo GWR pode descrever satisfatoriamente o fenômeno.

Considerando o nível municipal, os dados de luzes noturnas e de população obtiveram uma correlação significativa, com um  $R^2$  de 0,502, embora com uma nuvem de pontos mais dispersa em torno da reta, como se pode observar na Figura 5A. Os histogramas das variáveis também possuíam uma frequência similar que indicavam uma relação espacial, fato que pode ser visualizado no Cartograma (Figura 5B), no qual os círculos maiores e vermelhos representam regiões com maiores somas de luz e maior quantidade de pessoas, os quais se concentram fortemente na região considerada urbana pelo Plano Diretor Integrado do Município de São José dos Campos, estabelecido pela Lei Complementar n° 306/2006.

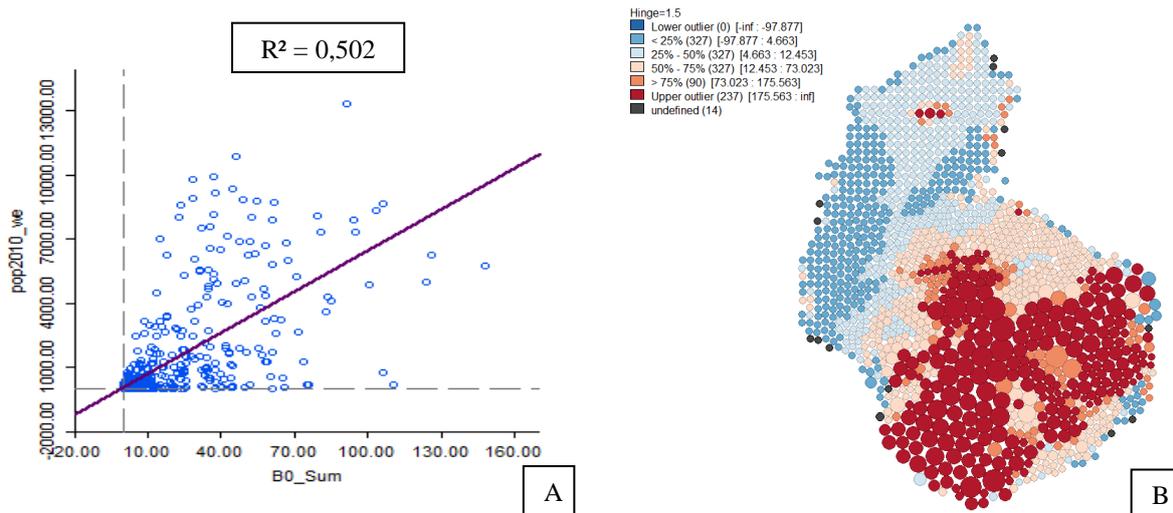


Figura 5. Relação entre os dados de luzes noturnas e os dados populacionais para o município de São José dos Campos. A: Gráfico de dispersão das variáveis com seu respectivo Coeficiente de Determinação ( $R^2$ ); B: Cartograma das variáveis no qual o tamanho do círculo indica a variação de luzes e as cores indicam a variação da população.

Obtendo-se esta correlação significativa, realizou a Regressão Linear que apresentou alta significância estatística, um  $R^2$  Ajustado de 0,501 e um AIC de 22213,8. A análise do Teste de Breusch-Pagan para heterocedasticidade demonstrou, assim como para a RMVPLN, uma variância não constante dos resíduos da regressão, o que pode ser mais bem explanado pelo gráfico de dispersão dos resíduos em relação à variável predita (Figura 6).

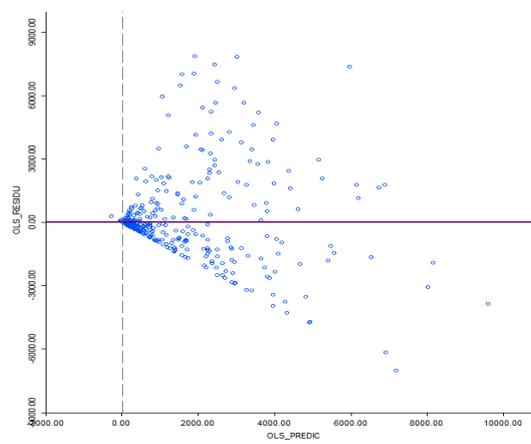


Figura 6. Gráfico de dispersão dos resíduos da Regressão Linear para o município de São José dos Campos.

Constatada a quebra dos pressupostos da Regressão Linear novamente, avaliou-se a possibilidade da dependência espacial, a qual foi evidenciada pelo Índice de Moran obtido de 0,373, que indica uma autocorrelação espacial positiva, fato corroborado pelo Diagrama de Espalhamento (Figura 7A) e pelos mapas do indicador LISA (Figura 7B) que demonstra uma alta significância para os agrupamentos localizados principalmente na área urbana do município.

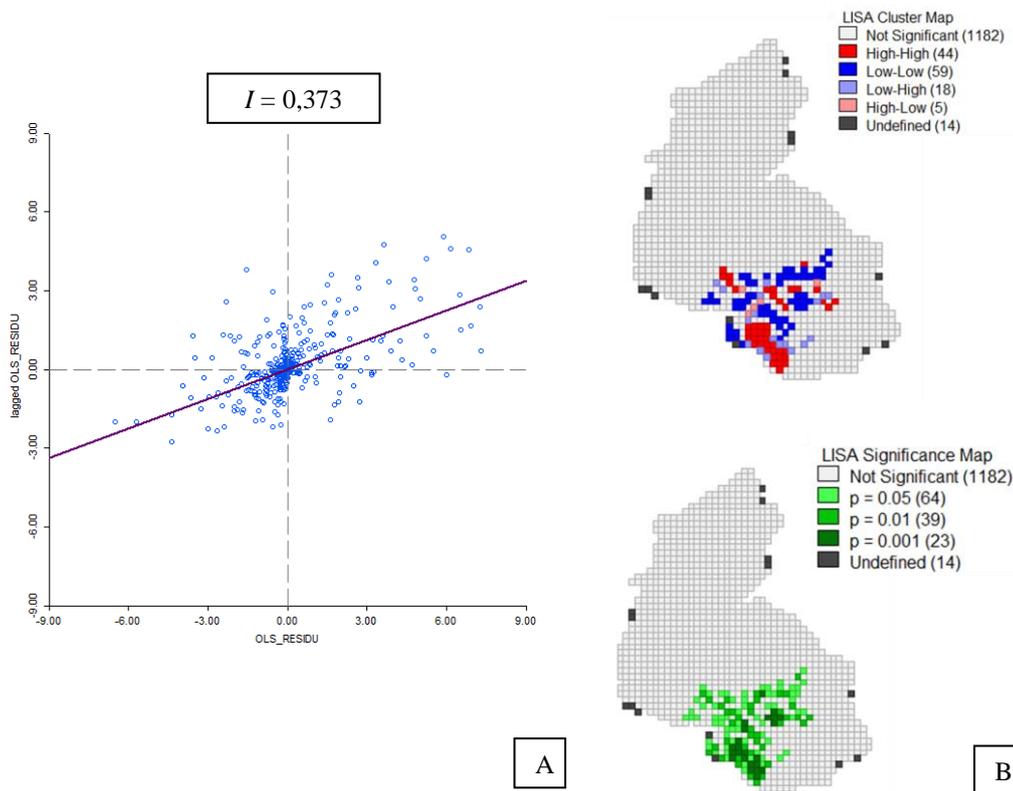


Figura 7. Análise espacial dos resíduos da Regressão Linear do Município de São José dos Campos. A: Diagrama de Espalhamento de Moran e seu respectivo índice; B: Mapa de Agrupamento e Mapa de Significância do indicador LISA.

Os resultados estatísticos da regressão GWR a nível Municipal demonstrou uma melhora significativa deste modelo em descrever as variáveis e dados observados, fato que pode ser verificado pelo aumento, embora pouco relevante, de 0,076 do  $R^2$  ajustado (0,578) e uma dedução significativa de 209,81 do AIC (22003,99) em relação à Regressão Linear.

### 3.2. Fase 2 – Inserindo a componente de Uso e Cobertura do Solo

A Figura 8 demonstra o Mapa de Uso e Cobertura do Solo para o ano de 2012 do município de São José dos Campos. A paisagem no município é predominantemente composta por pastagens, algumas culturas agrícolas e comerciais e vegetação. As áreas de vegetação nativa concentram-se, principalmente, na região norte do limite municipal, nas nascentes dos cursos d'água e ao longo do rio Paraíba do Sul, sendo também possível encontrar fragmentos junto a área urbana. As classes urbanas coincidiram, em sua maioria, com o zoneamento do município, conforme disposto no Plano Diretor Integrado do Município, incluindo as áreas vazias destinadas à expansão. Embora parte da área urbana exerça pressão sobre alguns córregos e ribeirões que cortam a cidade, boa parte das áreas de urbanização Densa e Média, estão fora das Áreas de Proteção Ambiental e de Controle à Impermeabilização, conforme indicado pelo Plano Diretor. Além disso, é possível notar um avanço das áreas de Urbanização Baixa ao redor dos grandes aglomerados já existentes, indicando novos loteamentos e bairros. Tais dados indicam a confiabilidade da classificação gerada e demonstram o bom planejamento e gestão direcionados ao crescimento urbano e industrial do município.

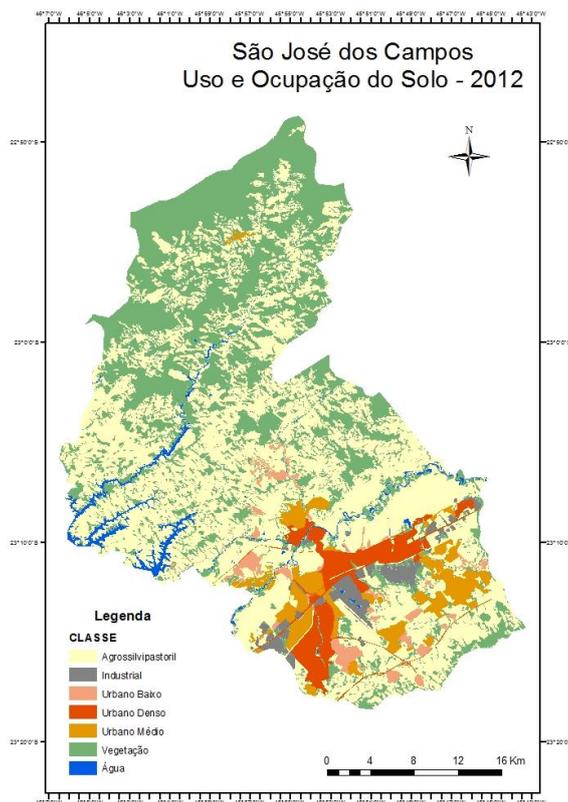


Figura 8. Mapa de Uso e Cobertura do Solo do município de São José dos Campos para o ano de 2012.

Considerando que as classes de interesse para este estudo eram as consideradas urbanas, as classes Agrossilvipastoril, Vegetação e Água foram agregadas em uma única classe denominada Outros, de modo a facilitar a modelagem destes dados. Analisando a relação da porcentagem das classes em relação aos dados populacionais (Figura 9A), nota-se uma forte relação, principalmente com a classe de Urbanização Densa que registrou um  $R^2$  de 0,671. Todavia, a classe Outros apresentou uma correlação negativa, fato que era esperado por englobar usos e coberturas não urbanos, o que levou a variável a ser retirada das análises. Concomitantemente, a relação das classes urbanas foi igualmente significativa considerando os dados de luzes noturnas (Figura 9B), de modo que os três tipos de dados, sendo correlacionados, poderiam ser utilizados nos modelos de regressão.

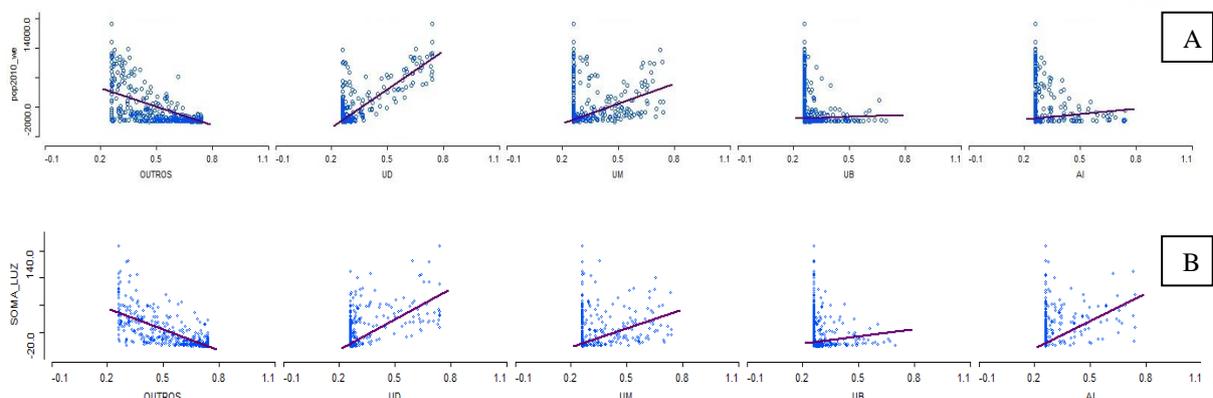


Figura 9. Relação dos dados de Uso e Cobertura do Solo do município de São José dos Campos em relação: A: aos dados populacionais; B: aos dados de luzes noturnas.

Na etapa da Regressão Linear testou-se a possibilidade de usar as variáveis de uso e cobertura individualmente e com a sua soma por unidade celular, de modo a verificar se as variáveis urbanas juntas poderiam tornar o modelo mais robusto, todavia como pode-se observar na Tabela 2, as variáveis individualmente contribuíram de forma mais significativa para o modelo, possuindo um R<sup>2</sup> Ajustado maior e um AIC menor em relação ao uso da soma das variáveis. Desta forma, as etapas seguintes foram realizadas utilizando as variáveis urbanas individualmente.

Contudo, analisando os resíduos desta regressão, foi observado novamente uma tendência a heterocedasticidade, fato que novamente invalidaria o modelo de Regressão Linear por ferir um de seus pressupostos.

Tabela 2. Estatística geral da Regressão Linear realizada para as variáveis de uso e ocupação do solo individualmente e somadas.

	<i>Variáveis Individuais</i>	<i>Soma das Variáveis</i>																						
R <sup>2</sup> Ajustado	0,842	0,585																						
AIC	20692,8	21970,5																						
Probabilidade das Variáveis	<table border="1"> <thead> <tr> <th>variable</th> <th>Probability</th> </tr> </thead> <tbody> <tr> <td>CONSTANT</td> <td>0.67872</td> </tr> <tr> <td>SOMA_LUZ</td> <td>0.00000</td> </tr> <tr> <td>UD</td> <td>0.00000</td> </tr> <tr> <td>UM</td> <td>0.00000</td> </tr> <tr> <td>UB</td> <td>0.64463</td> </tr> <tr> <td>AI</td> <td>0.00149</td> </tr> </tbody> </table>	variable	Probability	CONSTANT	0.67872	SOMA_LUZ	0.00000	UD	0.00000	UM	0.00000	UB	0.64463	AI	0.00149	<table border="1"> <thead> <tr> <th>Variable</th> <th>Probability</th> </tr> </thead> <tbody> <tr> <td>CONSTANT</td> <td>0.06256</td> </tr> <tr> <td>SOMA_LUZ</td> <td>0.00000</td> </tr> <tr> <td>URBTOTAL</td> <td>0.00000</td> </tr> </tbody> </table>	Variable	Probability	CONSTANT	0.06256	SOMA_LUZ	0.00000	URBTOTAL	0.00000
	variable	Probability																						
CONSTANT	0.67872																							
SOMA_LUZ	0.00000																							
UD	0.00000																							
UM	0.00000																							
UB	0.64463																							
AI	0.00149																							
Variable	Probability																							
CONSTANT	0.06256																							
SOMA_LUZ	0.00000																							
URBTOTAL	0.00000																							

Avaliando-se a dependência espacial, obteve-se um Índice de Moran consideravelmente baixo de 0,027, porém ao se observar seu Diagrama de Espalhamento, notou-se uma nuvem dispersa de pontos principalmente nos quadrantes Q1 e Q3, indicando uma possível autocorrelação positiva, conforme pode ser visto na Figura 10A. Conjuntamente, analisando o indicador Bivariado LISA, que considerou tanto os resíduos quanto a variável predita, demonstra diversos agrupamentos de relativa significância em todo o território do município, principalmente concentrados na área urbana, demonstrando que realmente há uma componente de dependência espacial.

Deste modo, foi aplicado o modelo GWR que apresentou uma alta significância e um avanço considerável em representar os dados observados, com um aumento do R<sup>2</sup> Ajustado (0,880) de 0,038 e uma decaída relevante de 347,71 do AIC (20345,09) em relação ao modelo de Regressão Linear.

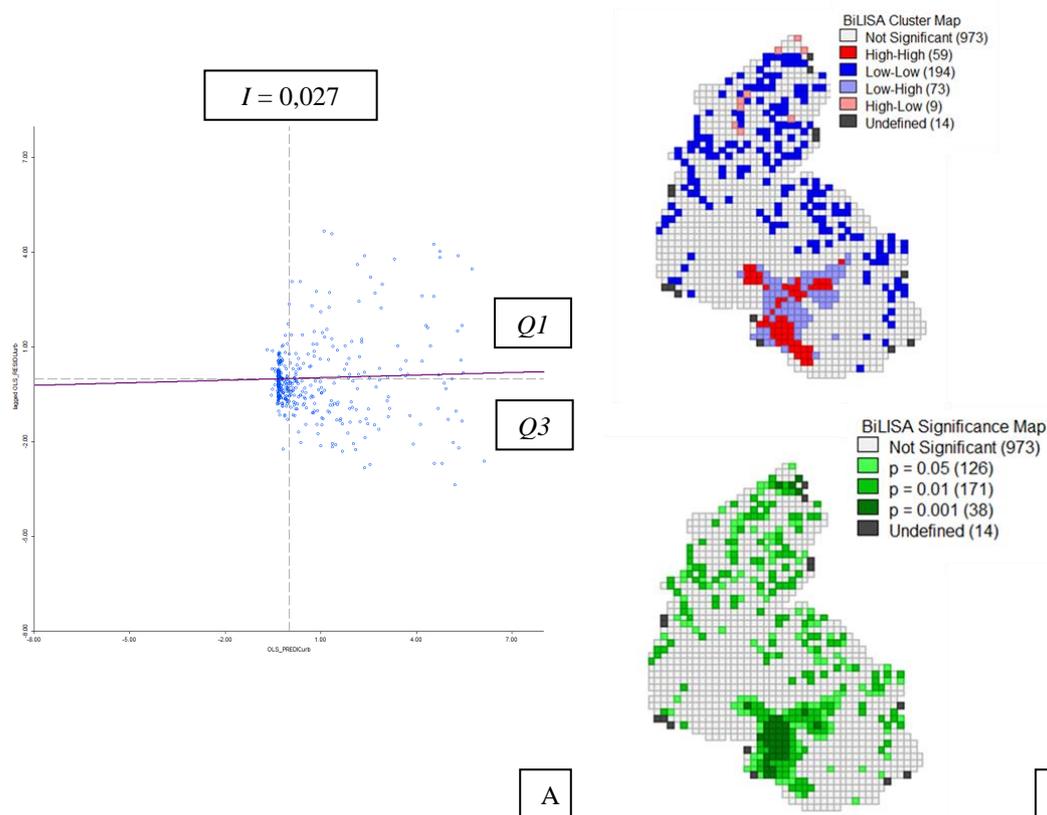


Figura 10. Análise espacial dos resíduos da Regressão Linear do Município de São José dos Campos. A: Diagrama de Espalhamento de Moran e seu respectivo índice; B: Mapa de Agrupamento e Mapa de Significância do indicador Bivariado LISA.

Avaliando-se a modelagem GWR aplicada nas duas fases propostas, a inserção da variável de uso e ocupação do solo se mostrou a com melhor habilidade das variáveis explicativas interpretarem as variáveis respostas e a que descreve melhor os dados observados, como se pode concluir através do resumo das análises das estatísticas do  $R^2$  Ajustado e do AIC descritas na Tabela 3.

Apesar dos resultados favoráveis da modelagem espacial, analisando a população estimada pelos modelos em relação ao dado populacional oficial fornecido pelo IBGE do Censo Demográfico de 2010, notou-se que os extremos foram subestimados e superestimados, fato que pode prejudicar estudos de dinâmica populacional. Sutton et al. (2001), em seu estudo de estimativa global de população utilizando dados de luzes noturnas também observou uma subestimação dos resultados que os autores atribuíram à relação entre a extensão da área e aos agrupamentos da população urbana. Doria et al. (2014), que realizou um estudo semelhante para áreas intraurbanas da Região Metropolitana de São Paulo, também obteve uma superestimação de 4,04% para mais de 50% de sua estimativa, tendo apresentado apenas 32,8% de coeficiente significativos, fatos estes que atribuiu a forma de agregação de seus dados.

Tabela 3. Resumo das estatísticas globais geradas para o Modelo GWR de todas as fases propostas no estudo.

	<b>FASE 1 (Regional)</b>	<b>FASE 1 (Municipal)</b>	<b>FASE 2</b>
<b>R<sup>2</sup> Ajustado</b>	0,578497	0,501949	0,880587
<b>AIC</b>	22003,99	22213,8	20345,09

#### 4. Conclusão

Este estudo demonstrou o potencial da correlação dos dados de luzes noturnas com dados populacionais para a estimativa populacional da Região Metropolitana do Vale do Paraíba e Litoral Norte e do Município de São José dos Campos. Também constatou que a inserção de dados de uso e cobertura do solo foi considerada significativa para a robustez da estimativa populacional a nível Municipal. A utilização da modelagem utilizando a Regressão Linear simples não foi possível, tendo em vista que os dados não atendiam aos pressupostos principais deste modelo, os quais são a homocedasticidade e a independência das variáveis, tendo em vista que foi constatada a dependência espacial dos resíduos.

No entanto, a aplicação do Modelo Geograficamente Ponderado foi satisfatória, embora se deva realizar mais pesquisas com diferentes modos de agregação dos dados de modo a minimizar os problemas de subestimação e superestimação do fenômeno avaliado. Além disso, este estudo indica a possibilidade do uso de dados de luzes registrados pelo Sensor VIIRS, cujas imagens de luzes possuem um grande potencial de utilização e com pesquisas relativamente escassas por ser um sistema considerado recente.

#### Referências Bibliográficas

AMARAL, S. et al. Estimating population and energy consumption in Brazilian Amazonia using DMSP night-time satellite data. **Computers, Environment and Urban Systems**, v. 29, n. 2, p. 179-195, 2005.

ATLAS BRASIL. **Atlas do Desenvolvimento Humano nos Municípios**. Disponível em [www.atlasbrasil.org.br/2013/pt/download](http://www.atlasbrasil.org.br/2013/pt/download). Acesso em 24 nov. 2017.

BAGAN, H.; YAMAGATA, Y. Analysis of urban growth and estimating population density using satellite images of nighttime lights and land-use and population data. **GIScience & Remote Sensing**, v. 57, n. 6, p. 765-780, 2015.

DORIA, V. E. M. et al. Avaliação do potencial das imagens de luzes noturnas para a estimativa e distribuição espacial de população urbana: um estudo para a Região Metropolitana de São Paulo, Brasil, com o sensor Visible/Infrared Imaging Radiometer Suite (VIIRS). In: XIX Encontro Nacional de Estudos Populacionais, São Pedro, SP, 2014. **Proceedings**. ABEP, v., p.20.

EMPLASA. Disponível em <https://www.emplasa.sp.gov.br/RMVPLN> Acesso em 24 nov. 2017.

HUILEI, L. et al. Urbanization impact on landscape patterns in Beijing City, China: A spatial heterogeneity perspective. **Ecological Indicators**, v. 82, p. 50-60, 2017.

**INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA.** Disponível em <https://cidades.ibge.gov.br/brasil/sp/sao-jose-dos-campos/panorama> Acesso em 24 nov 2017.

LETU, H. et al. Estimating energy consumption from night-time DMSP/OLS imagery after correcting for saturation effects. **International Journal of Remote Sensing**, v. 31, n. 16, p. 4443-4458, 2010.

MINISTÉRIO DO MEIO AMBIENTE (MMA). **GEOCATALOGO.** Disponível em <https://geocatalogo.mma.gov.br>. Acesso em 24 nov. 2017.

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION (NOAA). **Earth Observation Group (EOG).** Disponível em [https://ngdc.noaa.gov/eog/viirs/download\\_dnb\\_composites.html](https://ngdc.noaa.gov/eog/viirs/download_dnb_composites.html). Acesso em 24 nov. 2017.

OLIVEIRA, I. M. et al. **Segmentação para classificação de áreas urbanas a partir de imagem digital do Landsat7/ETM+: Estudo de caso – Cuiabá-MT.** In: Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, 2007, INPE, p. 6011-6018.

PREFEITURA DE SÃO JOSÉ DOS CAMPOS. **PLANO DIRETOR DE DESENVOLVIMENTO INTEGRADO DE 2006 (PDDI 2006).** LEI COMPLEMENTAR Nº 306, DE 17 DE NOVEMBRO DE 2006.

SOUZA, I. M. **Análise do espaço intra-urbano para Estimativa Populacional Intercensitária utilizando dados orbitais de Alta Resolução Espacial.** São José dos Campos: INPE, 2004, 108 p.

SUTTON, P.; ROBERTS, D.; ELVIDGE, C.; BAUGH, K. Census from Heaven: An estimate of the global human population using night-time satellite imagery. **International Journal of Remote Sensing**, v. 22, n. 16, p. 3061-3076, 2001.

TUCCI, C. E. M.; CLARKE, R. T. Impacto das mudanças da Cobertura Vegetal no escoamento: Revisão. **Revista Brasileira de Recursos Hídricos**, v. 2, n. 1, p. 135-152, 1997.