



## **Análise da correlação de variáveis derivadas de luzes noturnas e o volume populacional no estado do Pará, Brasil**

Gabriel da Rocha Bragion<sup>1</sup>

Instituto Nacional de Pesquisas Espaciais – INPE Caixa Postal 515 – 12227-010-São

José dos Campos – SP, Brasil

[gabriel.bragion@inpe.br](mailto:gabriel.bragion@inpe.br)

**RESUMO:** Este trabalho apresenta uma análise de correlação linear simples entre o volume populacional de setores censitários e municípios do estado do Pará, Brasil, em relação a variáveis derivadas de dados de luzes noturnas dos sensores OLS e VIIRS. O sensor VIIRS obteve um desempenho superior, quando comparado ao sensor OLS, com  $R^2$  de 0,76 para municípios com população inferior a 200.000 habitantes e  $R^2$  de 0,9182 para a população urbana de municípios com menos de 200.000 habitantes. Apesar do coeficiente de determinação elevado, a regressão deve ser interpretada com ressalvas, visto que a população rural não apresenta grau de aderência significativo a correlação linear simples. A distribuição dos resíduos demonstrou que a faixa norte do estado do Pará tem comportamento mais estável em relação as diferentes variáveis testadas, salvo uma pequena porção do extremo nordeste. A porção sul do estado tem valores residuais maiores e variáveis entre as regressões lineares baseadas em diferentes variáveis.

**Palavras-chave:** Luzes noturnas, sensoriamento remoto, DMSP, SNPP.

### **1 Introdução**

As luzes noturnas são os vestígios mais evidentes da presença humana na superfície terrestre. Durante o dia, a Terra vista do espaço é um complexo arranjo de características geográficas e atmosféricas, de tal magnitude que a presença humana pode passar despercebida. Entretanto, em face a escura paisagem, as criações humanas se destacam junto aos outros fenômenos transientes, como raios, auroras e queimadas (CROFT, 1978).

Em 1970, o sensor *Operational Linescan System* (OLS) foi integrado ao Programa de Satélite de Defesa e Meteorologia (*Defense Meteorological Satellite Program* - DMSP). Em sua concepção inicial, o sensor era utilizado para o monitoramento de nuvens em condições de baixa irradiância, para fins meteorológicos (WELCH e ZUPKO, 1980), no entanto, tão logo seus dados foram disponibilizados ao mercado, outras aplicações puderam ser exploradas. No decorrer das décadas, diversos métodos de classificação de artefatos e fenômenos luminosos foram investigados e, dentre os mais óbvios, queima de resíduos, extensão urbana e fenômenos naturais (LEVIN e ZHANG, 2017).

A medida que estes dados de luzes noturnas eram explorados, mais evidente eram os problemas a eles relacionados. A sua resolução grosseira e limitações radiométricas, aliadas a deterioração do sensor e ausência de calibradores internos são traduzidas na superestimativa de áreas iluminadas e números digitais inconsistentes/saturados (TUTTLE, et al., 2014). Até a primeira década do século 21, diversos trabalhos propuseram técnicas de correção para estes dados, principalmente no tocante à remoção de fontes de luzes efêmeras, até que o Grupo de Observação da Terra (NOAA/EOG – *Earth Observation Group*) passou a disponibilizar os dados já filtrados, apenas com números digitais médios anuais de luzes persistentes, o que atendia a maioria das aplicações da época.

Ainda assim, o sensor OLS foi, por quase duas décadas, o único sensor orbital a capturar consistentemente imagens noturnas da superfície terrestre, até que, em 2011, o sensor *Visible Infrared Imaging Radiometer Suite* (VIIRS) é posto em órbita a bordo do satélite Suomi-NPP. Dentre suas 21 bandas, a *Day Night Band* foi empregada justamente com o intuito de registrar de modo contínuo a atmosfera e superfície durante a noite. Com essas melhorias, novas aplicações passaram a ser empregadas e a fronteira desses dados ainda não pôde ser identificada.

Neste sentido, este trabalho se propõe a analisar os aspectos mais relevantes envolvendo as aplicações dos sensores VIIRS e OLS ao longo da história, e identificar uma orientação que permita elucidar as próximas aplicações destas ferramentas.

## **1.2 Operational linescan system**

O OLS é um dos sensores que integram a série de satélites do Programa de Satélite de Defesa e Meteorologia dos Estados Unidos (DMSP). Lançado primeiramente em 1976, o equipamento é sensível a faixa do visível e infravermelho termal (0,5-0,9  $\mu\text{m}$ ) e, à noite, devido a presença de um tubo fotomultiplicador (PMT), é capaz de registrar níveis de radiância na ordem de  $10^{-10}$   $\text{W}/\text{cm}^2/\text{sr}/\mu\text{m}$  (ELVIDGE et al, 2013). Com uma faixa de imageamento de 3.000 km, o sensor pode atuar em duas resoluções espaciais: fina (0,56 km) e total (2,8 km). Os dados são disponibilizados em uma composição anual de luzes persistentes em uma resolução espacial de 5 km.

Croft (1978) traz o primeiro vislumbre sobre as potencialidades deste conjunto de dados inéditos sobre as condições de iluminação noturna da superfície terrestre, como

detecção de queimadas agrícolas, frotas marítimas, adensamentos populacionais e queima de gases. No mesmo ano, Brandli (1978) destaca que os detectores do sensor OLS também se mostraram sensíveis a presença de outros fenômenos naturais, como raios, vulcões ativos e a aurora boreal, além de evidenciar a presença de processos ópticos causadores de ruídos.

Welch e Zupko (1980), apresentaram o que foi talvez a primeira abordagem quantitativa sobre os dados de luzes noturnas. Os autores observaram que a intensidade da radiação registrada pelo OLS pode ser correlacionada com os padrões de consumo energético de áreas urbanas, mas apontaram que resoluções espaciais e radiométricas mais refinadas poderiam vir a contribuir para um melhor desempenho do modelo.

Mesmo após o seu evidente potencial, a década de 80 foi marcada por um hiato relativo a exploração destes dados, devido ao desinteresse pelo monitoramento de áreas urbanas, de forma que a utilização destes produtos como ferramenta científica fosse negligenciada (Limhof, 1997). Ainda disponível apenas em formato analógico, Sullivan (1989) produziu um mapa global de fontes emissoras de luzes noturnas com resolução espacial de 10 km, mas somente em 1992 os filmes passaram a ser compilados e digitalizados para a composição de um repositório global de fontes de luzes noturnas (ELVIDGE et al., 1997).

A disponibilização deste acervo digital motivou alguns pesquisadores a explorar mais a fundo os problemas discutidos por alguns trabalhos da área (e. g., Limhof et al., 1997; Elvidge et al., 1997). As principais limitações associadas a utilização dos dados do sensor OLS para o monitoramento de áreas urbanas ocorriam devido a persistência de luzes efêmeras que descaracterizavam assentamentos populacionais, a saturação dos valores de números digitais, devido à alta sensibilidade do sensor, e o chamado efeito *overgloom*, decorrente da reflexão de das luzes noturnas pelo *background*, principalmente em regiões costeiras ou próximas a espelhos d'água (Limhof et al., 1997).

Estes problemas também estavam associados a resolução espacial relativamente grosseira e a quantização de apenas seis bits. Elvidge et al. (1997) propuseram então um algoritmo para eliminação de fontes de luzes efêmeras em imagens noturnas, baseado na integração de dados temporais e análise automática do *background* local, como forma de identificar também manchas causadas pela reflexão de raios solares em latitudes mais elevadas. O efeito de *overgloom* ainda era uma problemática proeminente, até que Limhof

et al. (1997) propuseram um método de redução deste ruído a partir da análise empírica dos dados, que buscava identificar o limiar de números digitais que pudesse diferenciar áreas que fossem efetivamente fontes de luzes de regiões contaminadas pela reflexão ou espalhamento da radiação eletromagnética. Com o passar dos anos, a disponibilização do acervo dos dados OLS instigou estudos temporais que ainda estavam desprovidos de métodos de intercalibração entre os sensores OLS nas diferentes plataformas do programa DMSP, dando mais visibilidade a trabalhos como o de Elvidge et al. (2009) e Witmer e O'Loughlin (2011), motivados principalmente pela ausência de calibradores internos e a degradação dos equipamentos.

O número de trabalhos envolvendo métodos de classificação e detecção a partir dos dados do sensor OLS havia despontado (e. g., NAKAYAMA et al., 1999; DOLL et al., 2000; KOHIYAMA et al. 2004; SUTTON, 2003; LIU et al., 2012; e. g., ZHOU et al., 2014; MA et al., 2015; XIE e WENG, 2016; ZHENG et al., 2018), mas alguns trabalhos haviam retomado a análise quantitativa de dados espaciais derivados de luzes noturnas aplicados a centro urbanos para estimativa de consumo de energia e volume populacional (AMARAL et al., 2003; AMARAL et al., 2006). Embora estes trabalhos tenham identificado uma correlação linear entre variáveis demográficas, os métodos abordados diferem significativamente entre si, indicando que as componentes envolvidas na variação da distribuição espacial e intensidade de luzes noturnas são regionais e dependem de fatores como infraestrutura e padrões de uso e ocupação.

### **1.3 Visible infrared imaging radiometer suite**

Lançado em 2011 a bordo do *Suomi National Polar-orbiting Partnership* (SNPP), o sensor VIIRS compartilha algumas características básicas com o seu antecessor, como órbita, faixa de imageamento e o espectro sensível. No entanto, algumas implementações diferem significativamente os produtos gerados por estes sensores (Tabela 1). A resolução espacial do sensor VIIRS é superior que ao OLS (742 m) com quantização de 14 bits (contra 6 bits do OLS) e sensibilidade na ordem de  $10^{-11}$  W/cm<sup>2</sup>/sr/μm (ELVIDGE et al., 2013). Dessa forma, problemas como saturação e *overglom*, antes dominantes no sensor OLS, são substancialmente reduzidos (Figura 1).

Tabela 1 - Comparação entre DMSP-OLS e SNPP-VIIRS (CONTÍNUA)		
Variável	DMSP-OLS	SNPP-VIIRS
Órbita	Polar – 850 km altitude, 98,8 graus de inclinação, 102 minutos	Polar – 827 km altitude, 98,7 graus de inclinação, 102 minutos
Faixa de imageamento	3.000 km	3.000 km

Tabela 2 - Comparação entre DMSP-OLS e SNPP-VIIRS (CONTINUAÇÃO)		
Horário de passagem	~19h:30min	~01h:30min
Amplitude espectral de baixa intensidade	0,5 a 0,9 $\mu\text{m}$	0,5 a 0,9 $\mu\text{m}$
Quantização	6 bit	14 bit
Sensibilidade mínima	5E-10 W/cm <sup>2</sup> /sr/ $\mu\text{m}$	2E-11 W/cm <sup>2</sup> /sr/ $\mu\text{m}$

Fonte: Adaptado de Elvidge et al. (2013).

Além de dispor de um melhor detalhamento espacial, o sensor VIIRS conta com um calibrador a bordo, que permite gerar produtos baseados em grandezas físicas com maior confiabilidade. Estas características o tornam mais atrativo para a análise de fatores ambientais, físicos e econômicos associados as variações de distribuição espacial e intensidade de luzes noturnas, permitindo avanços em estudos antes tomados apenas como promissores a partir dos dados do OLS (ELVIDGE et al., 2013).

Apesar dessas qualidades, os produtos VIIRS ainda devem passar por uma série processamentos, a depender dos objetivos da análise. Luzes efêmeras e dados considerados como *background* ainda exercem influência em modelos de regressão para a estimativa do produto interno bruto e volume populacional, embora os modelos baseados no VIIRS apresentam R<sup>2</sup> consistentemente superiores aos baseados no OLS (SHI et al., 2014).

Outro problema, agora associado a alta sensibilidade do sensor VIIRS, é a maior influência de fontes de radiação difusa exógenas a superfície terrestre, como luminescência atmosférica e estelar. Segundo Miller et al. (2012), estes fatores são variáveis no tempo e espaço e, por isso, compõem desafios para abordagens qualitativas dos dados de luzes noturnas. Atualmente, o NOAA/NGDC disponibiliza estes dados já processados em composições mensais, com uma banda de luzes estáveis baseada em todos os produtos adquiridos em um intervalo de 30 dias, e outra com o número de imagens livres de nuvens utilizadas no processo da composição (disponível em:

[https://ngdc.noaa.gov/eog/viirs/download\\_dnb\\_composites.html](https://ngdc.noaa.gov/eog/viirs/download_dnb_composites.html)). Estas duas bandas integradas permitem o usuário associar um grau de incerteza as análises e desenvolver produtos melhor adaptados as condições atmosféricas regionais.

Procedimentos já adotados anteriormente para a detecção de postos de queima residual de gases da cadeia produtiva petrolífera podem agora ser associados a algoritmos, conforme o desenvolvido por Elvidge et al. (2015), como ferramenta de monitoramento e categorização do volume de queima desses gases. No âmbito socioeconômico, trabalhos como o de Chen et al. (2015) demonstram o potencial dos produtos VIIRS para estimar variáveis mais específicas, relativas a contextos locais, como a taxa de imóveis desocupados em áreas metropolitanas, além de processos mais complexos, como variações populacionais, crescimento e declínio econômico e urbanização, além de monitoramento de desastres e áreas de conflito (BENNET e SMITH, 2017).

Estas diversas aplicações abordadas por diversos autores evidenciam a complexidade de fatores que possam vir a influenciar a intensidade e distribuição de fontes de luzes noturnas na superfície terrestre. Mais além, considerando que as fontes de luzes noturnas estáveis estão de alguma forma associadas com a infraestrutura de adensamentos populacionais, como seria a resposta dessas variáveis em regiões onde os processos de evolução da paisagem são pouco convencionais? O volume populacional estaria associado a distribuição de luzes noturnas ou a intensidade dessas luzes?

Como objetivo de investigar estas questões, buscou-se analisar e comparar a resposta dos dados de luzes noturnas dos sensores VIIRS e OLS ao volume populacional dos municípios e setores censitários do estado do Pará, Brasil. A análise da correlação entre esses dados pode indicar a presença de outras variáveis, socioeconômicas, físicas ou antropogênicas, que atuem na variação dos dados de luzes noturnas.

## **2. Materiais e método**

### **2.1 Dados de luzes noturnas**

Todos os dados de luzes noturnas utilizados no desenvolvimento deste trabalho foram disponibilizados pela NOAA/NGDC (disponível em: <https://ngdc.noaa.gov/eog/download.html>). Para o sensor OLS, foram adquiridos os dados da composição anual de luzes estáveis médias livres de nuvens para o ano de 2010,

capturadas pelo sensor a bordo do bloco F182010. Os dados do sensor OLS têm quantização de seis bits.

Como já descrito anteriormente, o sensor VIIRS foi posto em órbita apenas em 2011, passando a obter dados apenas no ano de 2012. Até o momento, os dados do sensor VIIRS, para o ano de 2012, foram disponibilizados apenas em composições mensais médias de luzes estáveis livres de nuvens noturnas, de março até dezembro. Junto a estes dados, a NOAA/NGDC também disponibiliza uma matriz auxiliar mensal do número de imagens livres de nuvens processadas durante a composição dos dados mensais. Foram então adquiridas todas as composições mensais para o ano de 2012 e a composição anual para o ano de 2015 (a única composição anual disponível até o momento). A composição anual de 2015 e as imagens de total de observações livres de nuvens para os meses do ano 2012 foram então utilizadas para o processamento das imagens do sensor VIIRS e geração de um produto anual de luzes persistentes médias.

## **2.2 Derivação dos dados de luzes noturnas do sensor OLS**

Como os dados do sensor OLS já são disponibilizados em uma composição anual média de luzes estáveis, os únicos procedimentos adotados foram a geração de um subproduto com efeito de *overglom* reduzido a partir da aplicação de um limiar equivalente a 10 unidades de números digitais e o recorte do mapa baseado na extensão do estado do Pará. Assim, a todo pixel com valor de número digital inferior a 10, foi atribuído o valor de zero. Este procedimento foi baseado na observação de algumas extensões urbanas em imagens do catálogo BING e na extensão das manchas de luzes noturnas.

Seguiu-se a metodologia proposta por Limhof et al. (1997), de forma que o limiar fosse incrementado até o ponto no qual a integridade das manchas de luzes em centros urbanos fosse comprometida. Embora este procedimento busque reduzir o efeito de *overglom*, decorrente do processo de espalhamento atmosférico e efeito do *background* associado a fontes de luzes noturnas intensas, é inevitável que fontes de luzes fracas provenientes de adensamentos populacionais reduzidos venham a ser omitidos por este processo.

### **2.3 Derivação dos dados de luzes noturnas do sensor VIIRS**

Como descrito anteriormente, os dados de luzes noturnas para o ano de 2012 são disponibilizados apenas com composições mensais. Desta forma, utilizaram-se as imagens do número de observações livres de luzes noturnas como parâmetros para qualificar estas composições, de modo que apenas os produtos de melhor qualidade integrassem a composição anual.

Dessa forma, foram selecionados os meses de junho, agosto, setembro e outubro de 2012 para integrar a composição anual de luzes noturnas. As composições mensais não são submetidas ao mesmo processamento adotado na produção das composições anuais, o que resulta na presença de ruídos e luzes efêmeras nestas composições. Assim, adotou-se a metodologia proposta por Wu et al. (2018) para a adaptação desses dados. A metodologia se baseia no pressuposto de que, se um pixel tem valor igual a zero na composição anual de 2015, ele deve permanecer igual a zero para todos os anos. Além disso, todos os valores negativos foram também designados a zero.

A desvantagem é que, pixels da composição anual de 2015 que representam fontes de luzes noturnas que surgiram depois de 2012 permitirão que ruídos possam ser associados a composição anual desse ano, caso o ruído e a fonte de luz emergida nesse período correspondam ao mesmo local. Dos dados filtrados, foi então extraída a média aritmética das composições mensais, resultando em uma composição anual média livre de nuvens para os dados do sensor VIIRS do ano 2012.

Para a redução do efeito *overgloom*, utilizou-se o mesmo método aplicado aos dados OLS.

### **2.4 Dados demográficos**

Os dados demográficos correspondem ao volume populacional e subgrupos. Esta variável foi obtida a partir da consulta dos resultados do censo 2010 (IBGE, 2010). O estado do Pará é composto por 8.919 setores censitários, dispostos em 143 municípios, podendo ser classificados como rural ou urbano. O que define a sua classificação é a posição em relação ao perímetro urbano contido na legislação urbanística municipal. Os municípios que carecem de tal legislação, têm os seus setores classificados de acordo com um acordo firmado entre o instituto e as autoridades municipais, levando em consideração o estado de direito e o estado de fato (IBGE, 2015). Existem controvérsias

relativas a esta classificação, principalmente no tocante aos adensamentos populacionais que não estão incluídos nos limites urbanos legais e dispõem de infraestrutura urbana incompleta (PERA e BUENO, 2016). As análises envolvendo os diferentes tipos de setores obedecem a classificação utilizada pelo IBGE.

Os limites municipais considerados referem-se a divisão político administrativa do ano de 2010 (disponível em: [geoftp.ibge.gov.br/organizacao\\_do\\_territorio/malhas\\_territoriais/malhas\\_municipais/municipio\\_2010/pa/](http://geoftp.ibge.gov.br/organizacao_do_territorio/malhas_territoriais/malhas_municipais/municipio_2010/pa/)). Os polígonos dos setores censitários referem-se ao censo de 2010 (disponível em: [ftp.ibge.gov.br/Censos/Censo\\_Demografico\\_2010/Sinopse/Agregados\\_por\\_Setores\\_Censitarios/](http://ftp.ibge.gov.br/Censos/Censo_Demografico_2010/Sinopse/Agregados_por_Setores_Censitarios/)).

Ao todo, 134 setores censitários carecem de informações demográficas, sendo estes excluídos das análises. Apenas dois setores têm volume populacional igual a zero, sendo que estes foram incluídos nas análises. O volume populacional municipal foi calculado a partir da soma dos volumes de seus devidos setores censitários. A partir dos dados dos setores censitários, foram também discretizados o volume populacional urbano e rural de cada município.

## **2.5 Associação dos dados de luzes noturnas as unidades de análise**

Com o intuito de associar os dados de luzes noturnas a cada um dos setores censitários, foram estipuladas duas variáveis de análises derivadas dos dados de luzes noturnas: a) soma dos valores dos pixels (SVP) e b) contagem dos pixels válidos (CPV). As composições anuais médias dos sensores OLS e VIIRS foram reamostradas para uma resolução espacial de um arco-segundo (aproximadamente 30 metros), baseado no eixo menor do setor censitário com menor área. Essa reamostragem tem o intuito de garantir que cada setor censitário que seja sobreposto por fontes de luzes noturnas tenha pelo menos um pixel válido a ele associado. Este procedimento não garante que todos os setores censitários serão associados a valores de números digitais, radiância ou pixels válidos diferentes de zero. No entanto, este método permite a associação das derivadas de luzes noturnas a quaisquer setores, até mesmo àqueles cuja áreas são muito inferiores a resolução espacial das imagens de luzes noturnas.

As derivadas de luzes noturnas associadas aos setores foram então agrupadas por município e por tipo. Assim, como já citado, cada município foi associado a um volume populacional total, urbano e rural através da estatística zonal aplicada aos devidos polígonos.

## **2.6 Procedimentos estatísticos**

Com o intuito de analisar a relação entre as derivadas de luzes noturnas e o volume populacional, foi utilizada a regressão linear global (Equação 1). A análise dos resíduos das regressões foi feita através da normalidade dos quartis e da análise da homocedasticidade entre os valores previstos e os resíduos. Os valores residuais foram então associados aos municípios e investigados.

$$TPi = \beta_0 + \beta_1 Dnli + \epsilon_i \quad (1)$$

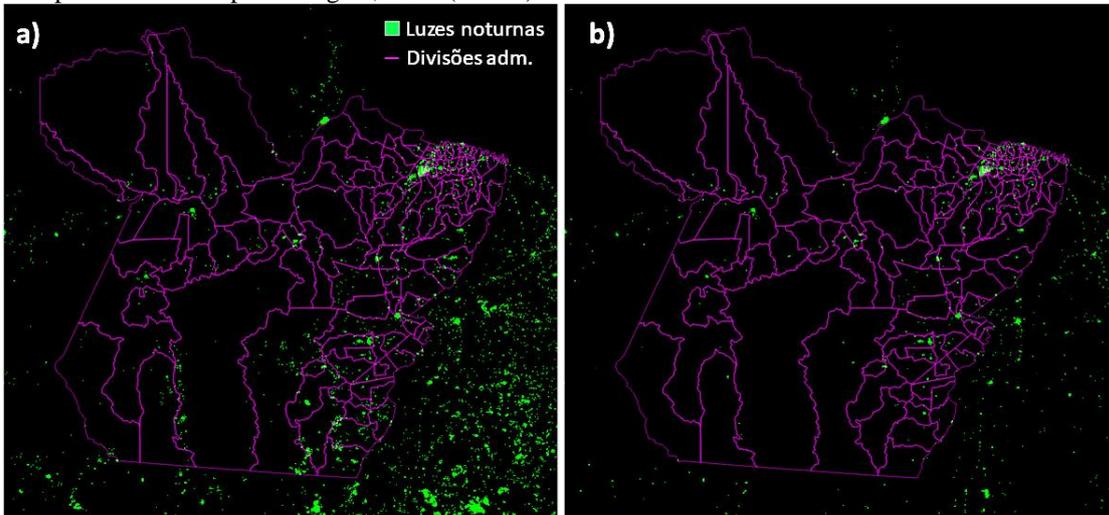
onde  $TPi$  é o total da população da  $i$ -ésima mancha de luz,  $\beta_0$  é o coeficiente linear,  $\beta_1$  é o coeficiente angular,  $Dnli$  é a soma das respectivas derivadas de luzes noturnas ( $i$ . e., SND e CPV (VIIRS/OLS) dos pixels do  $i$ -ésima setor/município e  $\epsilon_i$  é o erro.

## **3. Resultados e discussão**

### **3.1 Processamento das imagens de luzes noturnas**

Com o objetivo de analisar a relação entre variáveis derivadas de fontes de luzes noturnas e o volume populacional da população do estado do Pará, Brasil, foram adotados procedimentos para a geração de dados de luzes noturnas estáveis provenientes dos sensores OLS e VIIRS. A figura 1 demonstra o resultado da média aritmética obtida a partir das composições mensais selecionadas após a implementação da metodologia proposta por Wu et al. (2018) para a filtragem de dados espúrios e ruídos de composições mensais do sensor VIIRS.

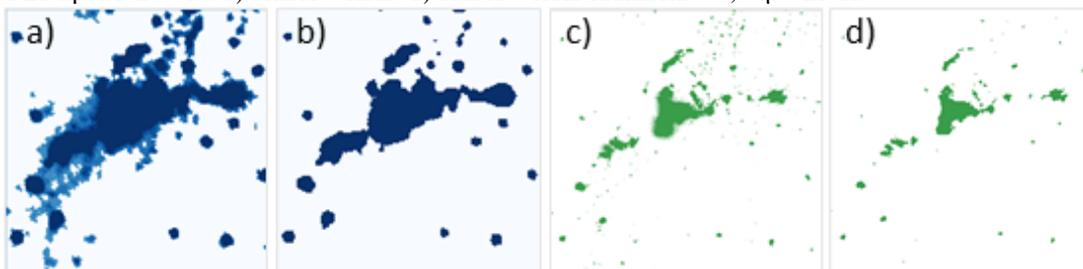
Figura 1 - Luzes noturnas da região do estado do Pará - Brasil. a) Comp. Setembro, 2012 (VIIRS). b) Comp. anual média após filtragem, 2012 (VIIRS).



Fonte: Autoria própria.

A partir da composição do sensor VIIRS (2012) e OLS (2010), foram produzidas outras duas composições, das quais os pixels com valores abaixo do limiar de *overgloom* foram associados ao zero (Figura 2). Para a composição do sensor OLS, foi utilizado um limiar de 10 unidades de número digital. Para o sensor VIIRS, foi utilizado um limiar de aproximadamente  $5,0 \eta\text{W}/\text{m}^2/\text{sr}$ . Na figura 2, é utilizada a mesma escala de representação para todos os itens. Nota-se a diferença entre a resolução espacial e radiométrica dos produtos gerados pelos dois sensores. A imagem do sensor OLS tem efeito *overgloom* mais acentuado que a do sensor VIIRS. Ao aplicar o limiar à imagem OLS, a extensão da feição de luz noturna é alterada drasticamente, enquanto que no VIIRS a mudança é menos proeminente.

Figura 2 – Amostra de luzes noturnas após redução do efeito de *overgloom*. a) Sensor OLS. b) sensor OLS após  $\text{ND} > 10$ . c) sensor VIIRS. d) sensor VIIRS radiância  $> 5,0 \eta\text{W}/\text{m}^2/\text{sr}$ .

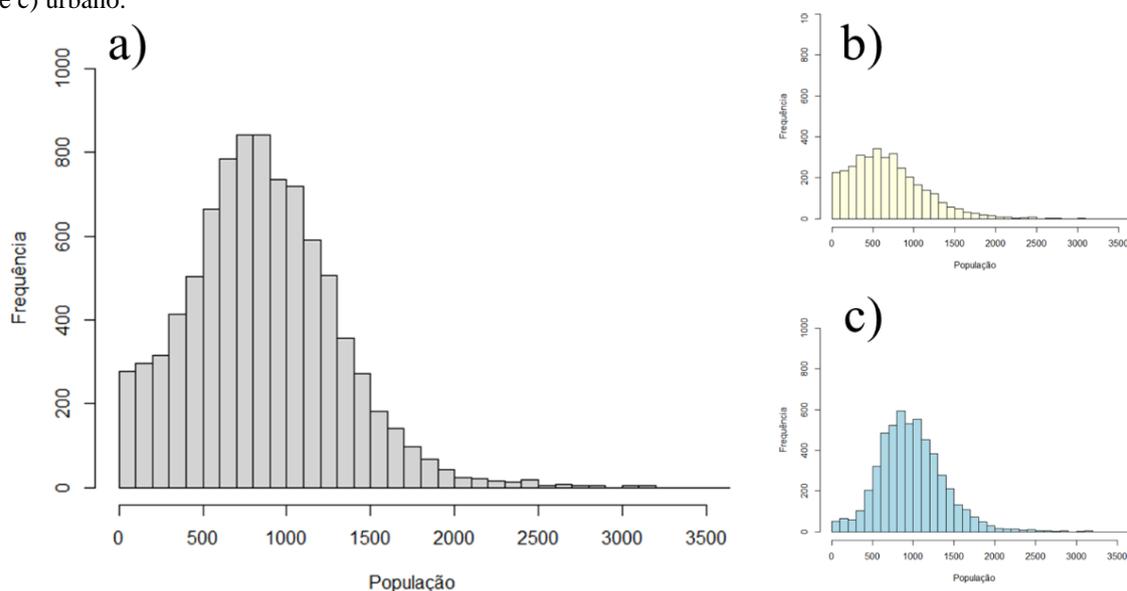


Fonte: Autoria própria.

### 3.2 Análise exploratória dos dados de luzes noturnas dos setores censitários

A população média por setor censitário para o estado do Pará é em torno de 850 habitantes. No entanto, a distribuição é desigual para setores do tipo urbano e rural (Figura 3). Os setores urbanos apresentam uma distribuição gaussiana com média centrada em 900 (Figura 3c), enquanto que os setores rurais são distribuídos de forma mais homogênea, o que caracteriza a curva menos acentuada do diagrama de frequência (Figura 3b).

Figura 3 – Distribuição de frequência do volume populacional por a) setor censitário, b) setor do tipo rural e c) urbano.



Fonte: Autoria própria.

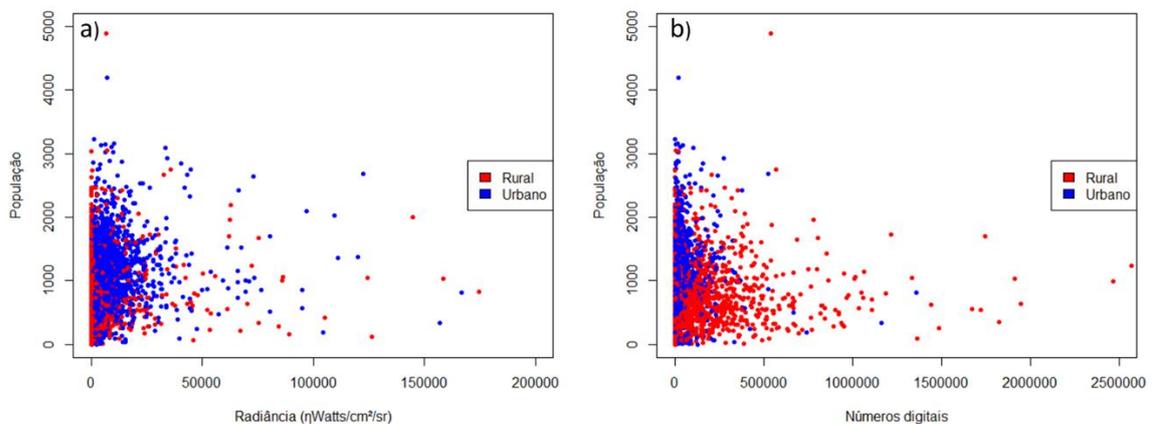
A comparação dos três diagramas mostra também que setores com menos de 500 habitantes são tipicamente rurais. Apesar de setores com mais de 900 habitantes serem mais frequentes na classe urbana, os setores rurais têm uma contribuição significativa para a frequência total de setores com volume populacional dessa ordem, sendo que a frequência de setores urbanos e rurais com mais de 2.000 habitantes é praticamente idêntica.

A distribuição de frequências pode ser explicada pelo método de determinação da extensão geográfica dos setores censitários. O tamanho de um setor censitário é determinado a partir de uma necessidade operacional. Quanto maior a quantidade de

domicílios, menor será o setor censitário. Pelo simples fato de um setor ser urbano, é menos provável que o volume populacional desse setor será pequeno. No entanto, o contrário parece não ser válido. Setores rurais podem apresentar um volume populacional baixo, pelo simples fato de serem rurais, mas também podem apresentar um volume consideravelmente maior, provavelmente devido a extensão desse setor.

A distribuição das derivadas de luzes noturnas não possibilitou uma análise precisa dos dados (Figura 4 e 5).

Figura 4 – Distribuição das derivadas de luzes noturnas e o volume populacional por setor censitário. a) sensor VIIRS (SVP em radiância). b) sensor OLS (SVP em número digitais).

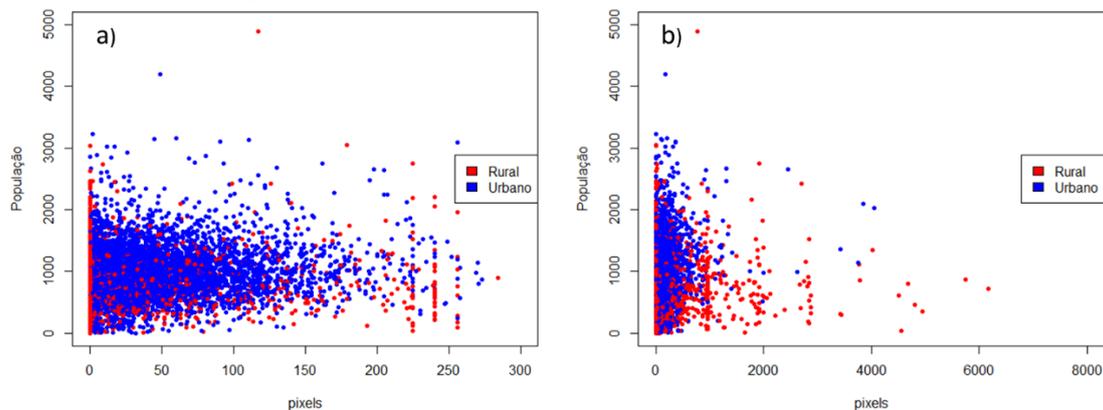


Fonte: Autoria própria.

A variável SVP (Soma dos valores de pixels) não demonstrou uma tendência linear de distribuição em relação ao volume populacional. Para o sensor VIIRS, os setores do tipo rural e urbano estão agrupados de forma similar, próximos ao eixo das variáveis independentes (Figura 4a). Já a relação entre SVP e população, os setores rurais estão mais distribuídos que os urbanos ao longo do quadrante 1, porém de forma pouco significativa. A frequência de setores com SVP inferiores a 500.000 é muito grande, porém com um volume populacional muito variável.

Os valores de CPV também não apresentaram resultados satisfatórios (Figura 5), ao menos no que tange a correlação linear entre o volume populacional e a contagem de pixels válidos.

Figura 5 - Distribuição das derivadas de luzes noturnas e o volume populacional por setor censitário. a) CPV sensor VIIRS e b) CPV sensor OLS.



Fonte: Autoria própria.

A contagem de pixels válidos do sensor VIIRS é bem distribuída, porém varia em uma taxa superior ao volume populacional, de modo que não haja uma proporção nas variações, descaracterizando uma correlação linear. Dada a característica da relação entre os dados, priorizou-se a análise do comportamento das derivadas de luzes noturnas em escala municipal.

### 3.3 Correlação das derivadas de luzes noturnas e volume populacional municipal

A estatística zonal a partir das informações dos limites municipais e dos dados de luzes noturnas permitiu a análise da correlação linear entre as variáveis a partir de um modelo de regressão linear simples. No caso da população urbana e total, foram considerados também apenas os municípios com volume populacional inferior a 200.000 habitantes, dada a presença de municípios que, apesar de constituírem parte do conjunto universo analisado, podem ser considerados outliers devido a disparidade do volume populacional em relação à amostragem como um todo.

Nesta seção serão destacados apenas as regressões com um melhor desempenho, avaliadas a partir do valor-p, do coeficiente de determinação ( $R^2$ , Tabela 2) e das análises da homocedasticidade baseadas nos respectivos QQ-Plots e distribuição dos resíduos vs valores preditos.

Tabela 2 – Índices da correlação linear simples entre o volume populacional e as variáveis derivadas de luzes noturnas. SVP: Soma dos valores de pixels; CPV: Contagem dos pixels válidos.

VIIRS	valor-p	$R^2$	OLS	valor-p	$R^2$
SVP	< 2,2e-16	0,9349	SVP	< 2,2e-16	0,5921

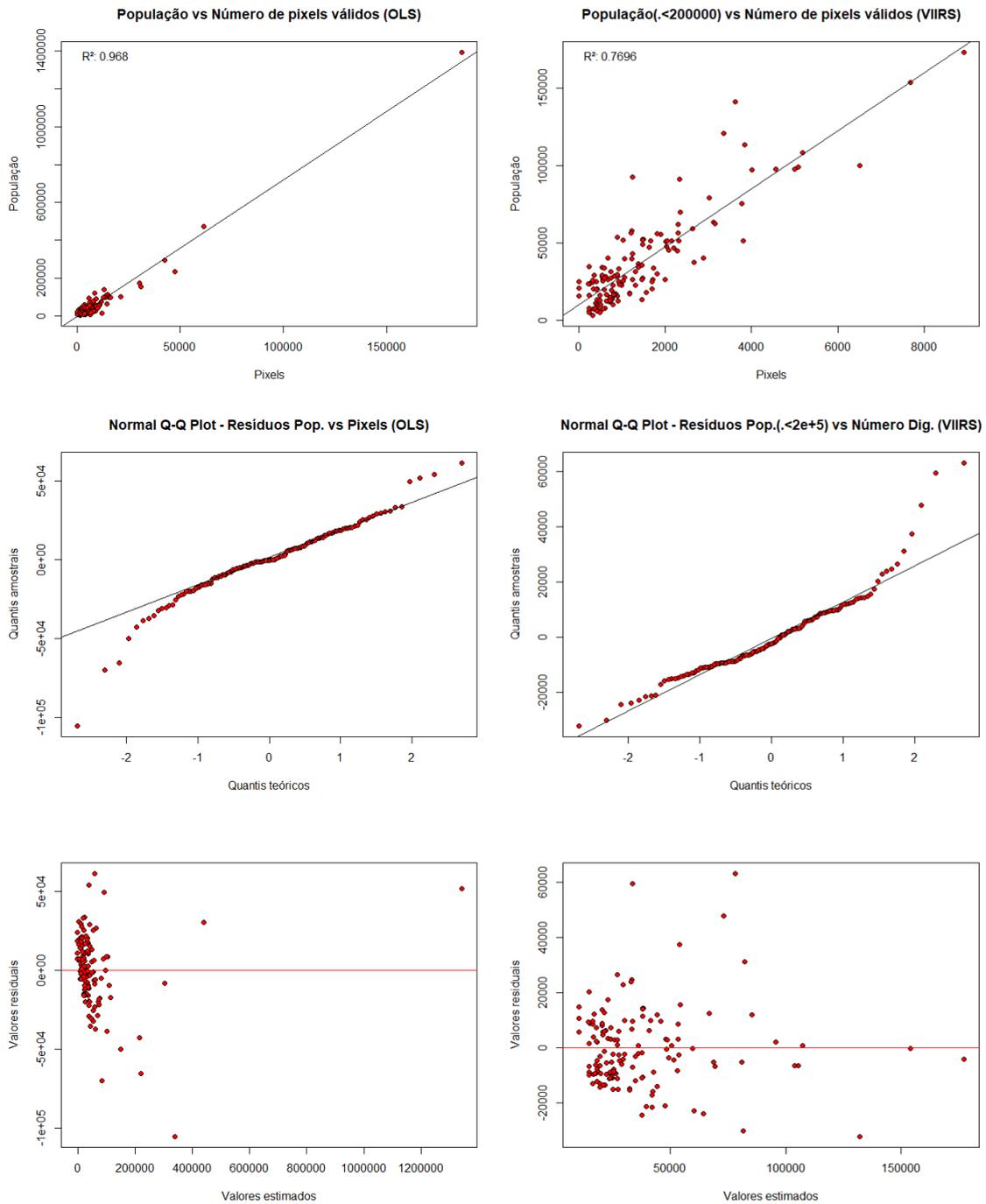
SVP (pop, até 200k)	< 2,2e-16	0,495	SVP (pop, até 200k)	< 2,2e-16	0,5631
<b>SVP Urbana</b>	<b>&lt; 2,2e-16</b>	<b>0,9849</b>	SVP Urbana	< 2,2e-16	0,8881
SVP Urbana (pop, até 200k)	< 2,2e-16	0,8718	SVP Urbana (pop, até 200k)	< 2,2e-16	0,8187
SVP Rural	0,0006138	0,0736	SVP Rural	5,60E-10	0,234
CPV	< 2,2e-16	0,9645	<b>CPV</b>	<b>&lt; 2,2e-16</b>	<b>0,968</b>
<b>CPV (pop, Até 200k)</b>	<b>&lt; 2,2e-16</b>	<b>0,7696</b>	CPV (pop, Até 200k)	< 2,2e-16	0,6642
CPV Urbana	< 2,2e-16	0,9768	CPV Urbana	< 2,2e-16	0,9825
<b>CPV Urbana (pop, até 200k)</b>	<b>&lt; 2,2e-16</b>	<b>0,9182</b>	CPV Urbana (pop, até 200k)	< 2,2e-16	0,8796
<b>CPV Rural</b>	<b>3,17E-08</b>	<b>0,19</b>	CPV Rural	9,41E-05	0,0965

Fonte: Autoria própria.

A variável CPV calculada a partir do sensor VIIRS se sobressaiu as outras. Sozinha, a variável obteve um melhor desempenho na estimativa da população municipal total, população municipal total de municípios com volume inferior a 200.000 habitantes, população urbana de municípios com população urbana inferior a 200.000 habitantes e população rural. A população municipal urbana foi melhor associada a variável SVP (Soma dos valores de pixels) do sensor VIIRS. A população total foi a única que foi melhor associada a uma variável do sensor OLS (CPV).

Na regressão baseada no CPV-OLS (Figura 8), observa-se que há uma boa concordância entre a regressão e os pontos plotados. A distribuição dos quartis demonstra que valores mais afastados dos centros estão mais sujeitos a ter a população subestimada. Como a maioria da população encontra-se no quartil inferior da regressão, os dados devem ser analisados com maior cautela. A regressão baseada apenas nos municípios com população, apesar de  $R^2$  inferior, tem os resíduos do primeiro quartil mais próximos à normal, o que indica um melhor ajuste desse conjunto à regressão. O custo disso fica evidenciado na distribuição dos resíduos vs valores preditos, o formato de cone indica que, à medida que o volume populacional aumenta, o erro também tende a aumentar.

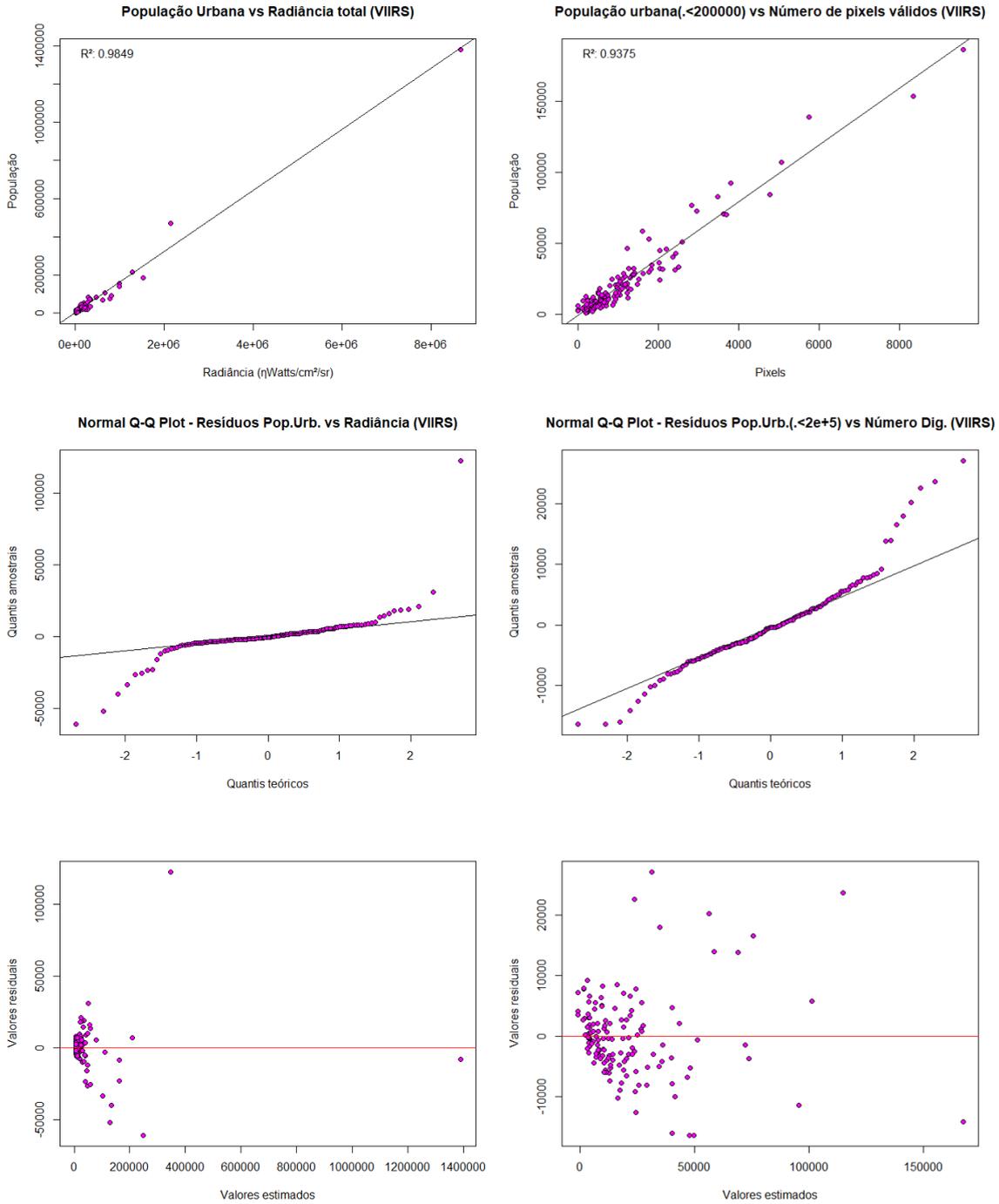
Figura 6 – Regressão linear simples da população municipal vs derivadas de luzes noturnas.



Fonte: Autoria própria.

A regressão linear entre a população urbana municipal e as variáveis derivadas de dados noturnos obtiveram um ótimo desempenho. O comportamento das duas variáveis em relação a população urbana é virtualmente idêntico ao anterior, no entanto, em proporções menores, evidenciando uma melhor aderência ao modelo linear.

Figura 7 - Regressão linear simples da população municipal urbana vs derivadas de luzes noturnas.

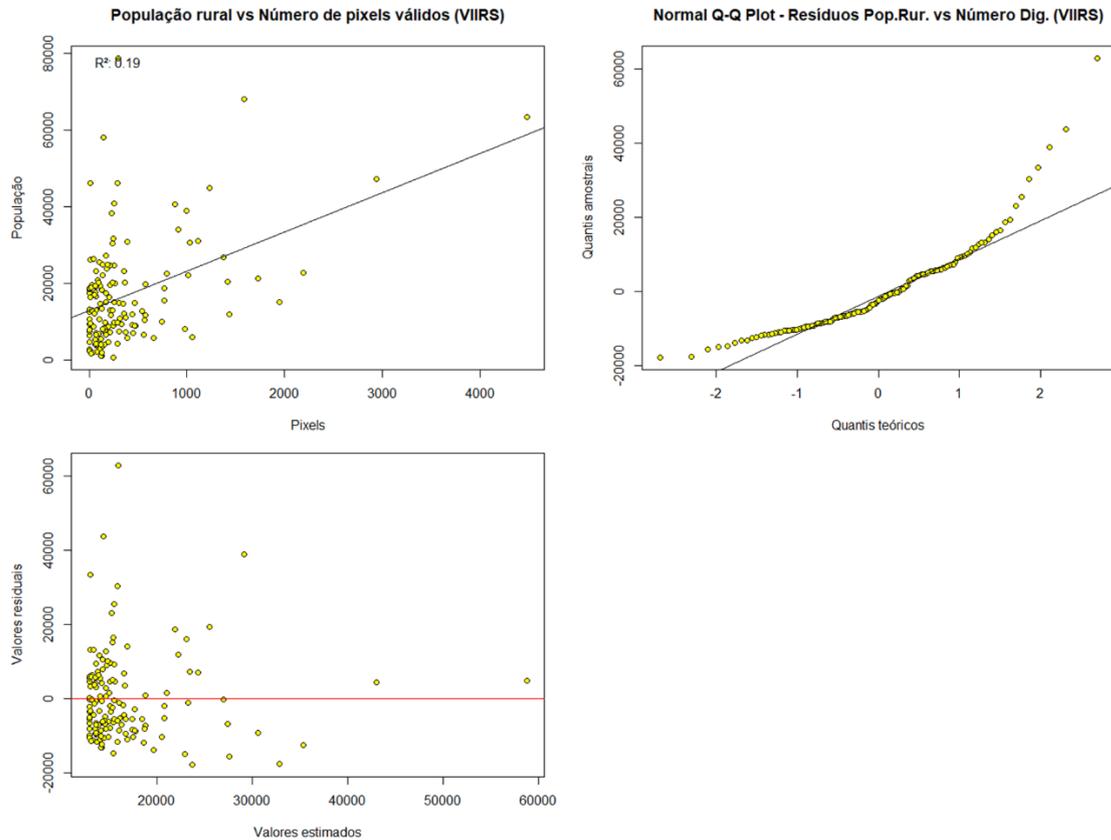


Fonte: Autoria própria.

Por fim, a população rural municipal não obteve um bom desempenho em nenhuma das quatro regressões, obtendo um  $R^2$  máximo de 0,19. Embora a variável tenha os quartis dos resíduos não muito afastados da normal e variância sem correlação aparente

com o valor predito, os valores das derivadas de luzes noturnas são agrupados próximo a origem, sem que haja uma relação clara entre as variáveis.

Figura 8 - Regressão linear simples da população municipal urbana vs derivadas de luzes noturnas.



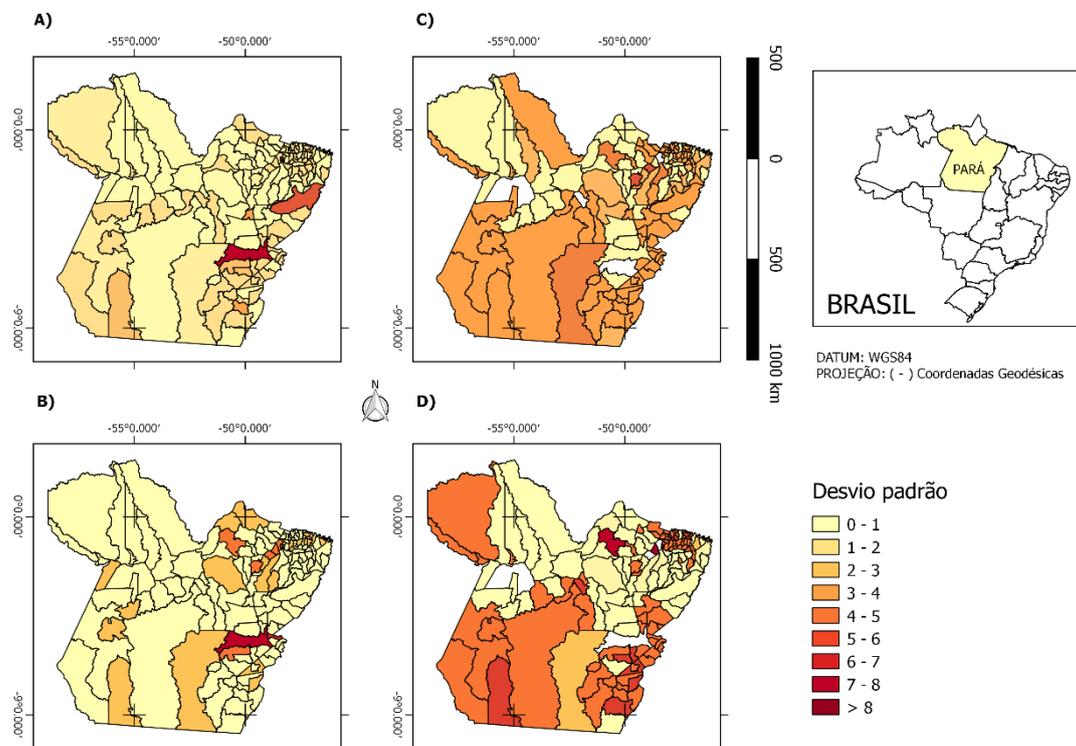
Fonte: Autoria própria.

Amaral et al. (2005) obtiveram também obtiveram valores de  $R^2$  em correlações lineares para a região amazônica próximos aos constatados. A alta correlação desses dados a nível municipal derivado de dados censitários e a ausência de correlação a nível de setores censitários indica que há ainda uma escala intermediária de análise que pode ser explorada.

Os resíduos relativos as regressões julgadas como de boa qualidade foram associados ao seus respectivos municípios. Como forma de analisar a distribuição espacial dos resíduos e comparar as diferentes interpolações, optou-se por expor o desvio padrão dos valores residuais, já que todos obtiveram média zero (Figura 9).

Nos mapas, é possível observar que as regressões que consideraram apenas os municípios com população inferior a 200.000 habitantes obtiveram um comportamento do erro similar. Embora a figura 9c trate da população total e a 9d apenas da população urbana, existe uma coerência nas distribuições dos desvios padrões residuais, principalmente na porção sul. O mesmo ocorre quando comparando a população total e população urbana (incluindo municípios com volume populacional acima de 200.000 hab., Figura 9a e 9b). O desvio padrão das regressões baseadas na população como um todo é mais estável.

Figura 9 – Desvio padrão dos erros relativos a cada uma das regressões abordadas, exceto para a população rural. A) população municipal total, B) população urbana total. C) População total acima de 200.000 habitantes. D) População urbana acima de 200.000 habitantes.



Fonte: Autoria própria.

Deve-se atentar ao fato de que o desvio padrão varia de acordo com a amplitude do dado. Dito isto, o mapa apresentado não tem caráter qualitativo, apenas busca identificar a concordância da distribuição dos valores residuais de cada município a partir de diferentes modelos de regressão.

#### 4. Conclusões

Este trabalho tem como objetivo analisar e comparar a resposta dos dados de luzes noturnas dos sensores VIIRS e OLS ao volume populacional dos municípios e setores censitários do estado do Pará, Brasil. A análise da distribuição dos setores censitários vs dados de luzes noturnas indicou que não há correlação entre essas variáveis. No entanto, a utilização dos dados dos setores censitários e a sua classificação quanto ao tipo, urbano ou rural, permitiram juntas identificar um alto grau de correlação entre o volume populacional da população urbana dos municípios do estado do Pará. No entanto, municípios com grande volume populacional (considerado 200.000 habitantes), não apresentam uma boa correlação com os dados de luzes noturnas, ao menos não tão boa quando considerada a outra parcela dos municípios. A população rural também não apresenta indícios de correlação com os dados de luzes noturnas, ao menos não linear. A distribuição dos resíduos das regressões demonstrou que a parcela sul do estado é menos estável quanto aos valores populacionais estimados, indicando que não apenas as variáveis derivadas de dados de luzes noturnas não respondem diretamente ao volume populacional, mas pode estar associada a outros fatores regionais, a serem estudados.

#### Referências Bibliográficas

- 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. **Computers, Environment and Urban Systems**. V. 29, n. 2, p. 179-195, 2005.
- AMARAL, S.; MONTEIRO, A. M.; CÂMARA, G.; QUINTANILHA, J. A. DMSP/OLS night-time light imagery for urban population estimates in the Brazilian Amazon. **International Journal of Remote Sensing**. V. 27, n. 05, p. 855-870, 2006.
- BENNETT, M. M.; SMITH, L. C. Advances in using multitemporal night-time lights satellite imagery to detect, estimate and monitor socioeconomic dynamics. **Remote Sensing of the Environment**. V. 192, p. 176-187, 2017.
- BRANDLI, H. W. **The night eye in the sky**. Photogrammetric Engineering and Remote Sensing, v. 44, n. 4, p. 503 – 505, 1978.
- CHEN, Z., et al. Estimating House Vacancy Rate in Metropolitan Areas Using NPP-VIIRS Nighttime Light Composite Data. **IEEE Journal of Selected Topics (...)**. V. 8, n. 5, p. 2188 – 2197, 2015.

CROFT, T. A. Nighttime Images of the Earth from Space. Scientific America, INC. p. 86 – 101, 1978.

ELVIDGE, C. D., et al. Mapping City Lights with Nighttime Data from the DMSP Operational Linescan System. **Photogrammetric Engineering and Remote Sensing**. V. 63. N. 6, 1997. p. 727 – 734.

ELVIDGE, C. D., et al. Methods for Global Survey on Natural Gas Flaring from Visible Infrared Imaging Radiometer Suite Data. **Energies**. V. 9, n. 14, p. 1-15, 2016.

ELVIDGE, C. D., et al. Why VIIRS data are superior to DMSP for mapping nighttime lights. **Proceedings of the Asia-Pacific Advanced Network**. V. 35, p. 62-69, 2013.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Metodologia do Censo Demográfico 2010. Vol. 41 da Série Relatórios metodológicos. Rio de Janeiro, 2013.

IMHOFF, W. T., et al. A Technique for Using Composite DMSP/OLS "City Lights" Satellite Data to Map Urban Area. **Remote Sensing of the Environment**. V. 61, p. 361 – 370.1997.

KOHIYAMA, M.; HAYASHI, H.; MAKI, N.; HIGASHIDA, M.; KROEHL, H. W.; ELVIDGE, C. D.; HOBSON, V. R. Early damaged area estimation system using DMSP-OLS night-time imagery. **International Journal of Remote Sensing**. V.25, n.11 p. 2015-2036, 2004.

LIU, Z.; HE, C.; ZHANG, Q.; HUANG, Q.; YANG, Y. Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008. **Landscape and Urban Planning**. V. 106, n.1, p. 62-72, 2012.

MA, T.; ZHOU, Y.; ZHOU, C.; HAYNIE, S.; PEI, T.; XU, T. Night-time light derived estimation of spatio-temporal characteristics of urbanization dynamics using DMSP/OLS satellite data. **Remote Sensing of Environment**. V. 158, p. 453-464, 2015.

MILLER, S. D., et al. Suomi satellite brings to light a unique frontier of nighttime sensing capabilities. **PNAS**. v. 109, n. 39, p. 157 – 15711, 2012.

NAKAYAMA, M.; MAKI, M.; ELVIDGE, C. D.; LIEW, S. C. Contextual algorithm adapted for NOAA-AVHRR fire detection in Indonesia. **International Journal of Remote Sensing**. V. 20, n. 17, p. 3415-3421, 1999.

PERA, C. K. L.; BUENO, L. M. M. Revendo o uso de dados do IBGE para pesquisa e planejamento territorial: reflexões quanto à classificação da situação urbana e rural. **Cad. Metro**. V. 18, n. 37, p. 712-742. São Paulo, 2016.

SHI, K., et al. Evaluating the Ability of NPP-VIIRS Nighttime Light Data to Estimate the Gross Domestic Product and the Electric Power Consumption of China at Multiple Scales: A comparison with DMSP-OLS Data. **Remote Sensing**. V. 6, p. 1705-1724, 2014.

SUTTON, P. C. A scale-adjusted measure of “Urban sprawl” using nighttime satellite imagery. **Remote Sensing of Environment**. V. 86, n.3, p. 353-369, 2003.

TUTTLE, B. T., et al. Aladdin's magic lamp: Active target calibration of the DMSP OLS. **Remote Sensing**. V.6, p. 12708 – 12722, 2014.

WELCH, R.; ZUPKO, S. Urbanized Area Energy Utilization Patterns from DMSP Data. **Photogrammetric Engineering and Remote Sensing**. V. 44, n. 2, p. 201 – 207, 1980.

XIE, Y.; WENG, Q. Updating urban extents with nighttime light imagery by using an object-based thresholding method. **Remote Sensing of Environment**. V. 187, p. 1-13, 2016.

ZHENG, Q.; JIANG, R.; WANG, K.; HUANG, L.; YE, Z.; GAN, M.; JI, B. Monitoring the trajectory of urban nighttime light hotspots using a Gaussian volume model. **International Journal of Applied Earth Observation and Geoinformation**. V. 65, p. 24-34, 2018.