



Avaliação das Grades Popacionais Globais para Representar a Distribuição da População Residente na Amazônia Brasileira: o Caso do Baixo Tocantins (PA)

Evaluation of Global Gridded Population Data for Representing Resident Population Distribution in the Brazilian Amazon: The Case of the Baixo Tocantins Region

Gustavo Piva Lopes Salgado ¹, Ana Paula Dal'Asta ², Bruno Vargas Adorno ³ e Silvana Amaral ⁴

¹ Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brasil. gustavo.salgado@inpe.br.

ORCID: <https://orcid.org/0009-0004-8167-3780>

² Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brasil. ana.dalasta@inpe.br.

ORCID: <https://orcid.org/0000-0002-1286-9067>

³ Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brasil. bruno.adorno@inpe.br.

ORCID: <https://orcid.org/0000-0003-0302-7834>

⁴ Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brasil. silvana.amaral@inpe.br.

ORCID: <https://orcid.org/0000-0003-4314-7291>

Recebido: 10.2024 | Aceito: 05.2025

Resumo: Este trabalho avaliou quatro Grades Popacionais Globais (GPWv4, GHS-POP, HRSL e WorldPop) para representar a distribuição da população residente na Amazônia brasileira, com foco na concordância com dados populacionais e na aderência a uma *proxy* de densidade de endereços de domicílios. A área de estudo é a região do Baixo Tocantins, Pará, por sua diversidade de usos e coberturas da terra, frequência de remanescentes florestais e população rural dispersa por ambientes de ilhas e terra firme. As Grades Globais para 2020 foram avaliadas através de: i) análise de concordância com os dados de população residente, por município e setor censitário, provenientes de estimativa municipal para 2020 e Censo Demográfico de 2022, e ii) análise de aderência às estimativas de domicílios, com referência aos dados do Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE), a partir de uma *proxy* de densidade de domicílios. Erros absolutos percentuais médios avaliaram as estimativas das Grades e os dados oficiais. Os maiores erros de concordância nos dados populacionais foram registrados para as grades GPWv4 e WordPop. WorldPop e GHS-POP subestimam a ocorrência de domicílios nas áreas de ilhas e margens de rios. A grade HRSL apresentou a maior concordância em relação aos dados populacionais e maior aderência às estimativas de domicílios em comparação ao CNEFE. Assim, na ausência de dados populacionais oficiais desagregados recomenda-se a grade HRSL como opção preferencial para estudos que envolvam a distribuição da população na Amazônia brasileira.

Palavras-chave: Grade Populacional Global. Modelagem Demográfica. Região Amazônica. Análise de Aderência.

Abstract: This study evaluated four Global Gridded Population datasets (GPWv4, GHS-POP, HRSL, and WorldPop) to represent the distribution of the resident population in the Brazilian Amazon, focusing on their concordance with population data and adherence to a residence address density proxy. The study area was the Baixo Tocantins region, Pará, chosen for its diversity of land uses and coverages, frequency of forest remnants, and dispersed rural population across island and mainland environments. The Global Grids for 2020 were evaluated through: i) concordance analysis with resident population data, by municipality and census tract, derived from municipal estimates for 2020 and the 2022 Demographic Census, and ii) analysis of adherence to residences estimates, using data from the Brazilian National Address Register for Statistical Purposes (CNEFE), based on a residence density proxy. Mean absolute percentage errors evaluated the Grids' estimates and official data. The highest concordance errors in the population data were observed for the GPWv4 and WorldPop grids. WorldPop and GHS-POP underestimate the occurrence of residences in island and riverbank areas. The HRSL grid showed the highest concordance with Census population data and better adherence to residences estimates compared to CNEFE. Thus, in the absence of disaggregated official population data, the HRSL grid is recommended as the preferred option for studies involving population distribution in the Brazilian Amazon.

Keywords: Global Gridded Population. Demographic Modeling. Amazon Region. Adherence Analysis.

1 INTRODUÇÃO

Com cerca de 6,5 milhões de quilômetros quadrados de vastas áreas de florestas tropicais, a Amazônia abriga mais da metade da biodiversidade do planeta e cerca de 20% da água doce do mundo, além de desempenhar um papel na regulação dos padrões climáticos locais e regionais ao liberar na atmosfera cerca de 20 bilhões de toneladas de vapor d'água diariamente (Nobre et al., 2009; Fearnside, 2006). Ao mesmo tempo, abriga 26 milhões de habitantes, cerca de 13,1% da população do Brasil (IBGE, 2022a), distribuídos em uma diversidade de unidades socioespaciais (Cardoso et al., 2023) ao longo de 772 municípios (IBGE, 2022c). Com densidade demográfica de 5,3 pessoas por km² (IBGE, 2022a), a região tem parte de sua população concentrada principalmente em pequenos espaços urbanos, além de uma variedade de povos vivendo sob o dossel florestal, representando desafios para a representação da distribuição da população. A estes desafios, acrescenta-se a extensão das unidades administrativas e das unidades territoriais de publicação de dados populacionais que pressupõem a homogeneização de seus ambientes internos e estão sujeitas a alterações periódicas de seus limites.

Nas últimas décadas, diversas iniciativas têm desenvolvido métodos para a produção e disseminação de estimativas de (re)distribuição de dados espacializados da população em nível de grade, chamados de Grades Populacionais (Dahmm & Rabiee, 2020). Uma grade populacional é formada por células de tamanho padrão, podendo ser de abrangência global, nacional, regional ou local. Em cada célula da grade, é possível associar informações relativas à população, como: estimativas de população residente e de população ambiente (população presente durante determinado período, como em 24h); população por sexo, faixa etária e etnia, entre outras informações relacionadas à caracterização demográfica. Os censos demográficos nacionais constituem a principal fonte de dados demográficos para a produção das estimativas em grade, das quais também deriva a qualidade dessas representações. Leyk et al. (2019) sugerem que os usuários avaliem as características e qualidade dos dados populacionais dispostos em grades, quanto à resolução espacial e sua relação com o propósito do trabalho. Dados auxiliares, derivados especialmente de imagens satelitais, como indicativos espaciais de concentração populacional, podem produzir incertezas na medida em que invisibilizam particularidades regionais (Leyk et al., 2019; Gonçalves et al., 2021; Silva, 2023).

A crescente disponibilidade de grades que estimam densidades de população ampliou o uso destes produtos para diversas aplicações, a exemplo de Renner et al. (2018), Zhou et al. (2020) e Zhao & Wang (2023), aumentando assim a necessidade de avaliar suas acurácias e/ou consistências. Estudos têm buscado avaliar a qualidade das grades considerando as diferenças nos métodos e nos tipos de variáveis auxiliares utilizadas para sua construção, assim como nas particularidades dos contextos históricos e geográficos que podem dificultar a modelagem, como abordado em Bai et al. (2018), Archila Bustos et al. (2020), Calka & Bielecka (2020), Chen et al. (2020), Xu et al. (2021), Yin et al. (2021) e Kuffer et al. (2022). Silva (2023) construiu uma grade de população para a Região Metropolitana do Vale do Paraíba e Litoral Norte (SP), baseada em variáveis auxiliares adaptadas ao contexto regional, e comparou com as grades de grandes projetos globais mostrando as limitações desses dados ao representar contextos locais. Para duas áreas na Amazônia, Amaral et al. (2012) propuseram um método para redistribuir, para o nível de grade, a população agregada por setor censitário destacando os desafios da modelagem nesta região.

Assim, este trabalho discute a qualidade das principais Grades Populacionais Globais na representação da distribuição da população residente na Amazônia brasileira, a partir de dois questionamentos: 1) as estimativas das Grades Populacionais Globais são consistentes com os dados oficiais de população dos municípios da Amazônia brasileira? 2) estas Grades Populacionais Globais podem ser utilizadas como alternativas para representar a densidade e distribuição da população nos espaços intramunicipais da Amazônia brasileira? Para tanto, dois objetivos específicos foram propostos considerando o Baixo Tocantins (PA) como recorte espacial: (i) avaliar a concordância das estimativas de valores das Grades Populacionais Globais com os dados de referência da população residente por município; e (ii) analisar a aderência dos dados das grades globais a partir de um dado *proxy* de densidade de endereços de domicílios no nível de grade. Ao avaliar as principais Grades Populacionais Globais, quanto aos seus dados e métodos de construção, este estudo contribui para o entendimento das vantagens e limitações para a representação da distribuição da população na Amazônia.

2 GRADES POPULACIONAIS E A DISTRIBUIÇÃO DA POPULAÇÃO NO TERRITÓRIO

A distribuição da população em grades regulares permite a manipulação dos dados populacionais em um suporte espacial estável no tempo e no espaço, ou seja, independe dos limites administrativos e das unidades de coleta censitária, sujeitos a mudanças ao longo do tempo. O uso de grades também facilita a integração com outras fontes de dados multitemporais, uma vez que reduz problemas estatísticos associados a formas de agregação por áreas, como a unidade de área modificável (*Modifiable Areal Unit Problem – MAUP*)¹, e aumenta a disponibilidade de dados populacionais intercensitários para modelagens que requerem este tipo de entrada desagregada (Archila Bustos et al., 2020).

Lloyd et al. (2017) explicam que Grades Populacionais são especialmente úteis para o planejamento de políticas públicas, permitindo o monitoramento da distribuição e das dinâmicas populacionais, bem como o auxílio à formulação de intervenções adequadas em diferentes escalas territoriais. Estes autores elencam também aplicações específicas de grades, como para estimar a ocorrência de doenças, modelar epidemias, definir a alocação de recursos e auxiliar na gestão de desastres naturais, além da modelagem de índices de acessibilidade, planejamento urbano, mapeamento da pobreza e na avaliação de impactos ambientais.

As Grades Populacionais podem ser produzidas por métodos de agregação (*bottom-up*) ou desagregação (*top-down*). Os métodos *bottom-up* dependem de dados de contagens ou amostras populacionais georreferenciados para modelar a distribuição, sendo mais comuns na produção de grades locais, onde a escala reduzida permite uma disponibilidade exequível destes dados. Por outro lado, os métodos *top-down* utilizam como dados de entrada os agrupados em unidades administrativas ou censitárias, realizando assim a desagregação desses dados para unidades regulares (Dahmm & Rabiee, 2020). Dentre os métodos *top-down* existem aqueles que não usam dados auxiliares, como o método ponderado pela área (*areal weighting*), com redistribuição que considera apenas as áreas das células da grade que sobrepõe as unidades administrativas ou censitárias (Hallisey et al., 2017). A desagregação também pode ser realizada por técnicas de realocação, chamadas de “dasimétricas”, ou por estimativa/interpolação, como aquelas geradas por modelagem *Random Forest*. Segundo Nagle et al. (2013), ambas dependem de algum dado censitário de entrada para ser realocado ou interpolado, mas também de dados auxiliares para extrair os pesos que serão utilizados para essa ponderação. Derivadas de projetos globais que utilizam métodos *top-down*, as principais Grades Populacionais disponíveis são *Gridded Population of the World* (GPW), *Global Human Settlement Layer – Population* (GHS-Pop), *High Resolution Settlement Layer* (HRSL) e *WorldPop*.

A *Gridded Population of the World* (GPW) foi a primeira Grade Populacional Global produzida, sendo sua primeira versão lançada em 1995. Desde então, mapas de limites administrativos dos países e máscaras de corpos d'água e gelo são utilizados como principais dados auxiliares. Em sua versão mais recente (GPWv4), baseia-se nos censos nacionais obtidos entre 2005 e 2014 e produziu densidades de população residente para os anos de 2000, 2005, 2010, 2015 e 2020, ajustadas com as estimativas da população mundial da Organização das Nações Unidas (ONU), que por sua vez são obtidas dos censos e estimativas populacionais nacionais. Os conjuntos de estimativas populacionais são alocados proporcionalmente para as células de um arquivo raster usando uma abordagem uniforme de ponderação de áreas que intersectam fronteiras subnacionais ou nacionais. A população é desagregada com base na proporção da área de cada unidade localizada na célula de grade (CIESIN, 2018).

A Grade Populacional *Global Human Settlement Layer – Population* (GHS-Pop) provê estimativas globais da população residente de 1975 a 2020, com intervalos de cinco anos e projeções populacionais para os anos de 2025 e 2030. A GPWv4, utilizada como dado de entrada, é desagregada pelo método dasimétrico binário para células refinadas, através da distribuição da densidade e classificação de áreas construídas. Para cada ano, a base *Global Human Settlement Layer Built* (GHSL-Built) correspondente é a referência de áreas construídas. Na desagregação dasimétrica da população, células definidas como "NoData" no dado auxiliar GHSL-Built são consideradas referência para a ausência de população. As células da grade GPWv4 definidas

¹ As diferentes formas de agregação de dados podem levar a resultados diferentes, influenciando as conclusões e interpretações de um estudo. Isso ocorre porque a escolha das unidades de agregação pode afetar a variabilidade dos dados e a distribuição espacial das observações (OPENSHAW, 1984).

como "desabitadas", mas que continham alguma superfície construída em 2020, foram revisadas quanto à presença de população residente, com auxílio de imagens de altíssima resolução dos serviços de mapeamento *web* Google Maps® e Bing®. Inconsistências entre os dados GPWv4 e evidências contraditórias de áreas residenciais foram detectadas e corrigidas (Schiavina et al., 2023).

A grade *High Resolution Settlement Layer* (HRSL) dispõe de estimativas de população para 2015 e 2020. É construída com o método de desagregação da dasimetria volumétrica, ou seja, estimativas de população são redistribuídas no espaço conforme dados auxiliares da densidade volumétrica de edificações. Esta construção ocorreu em quatro etapas: 1) o Centro da Rede Internacional de Informações sobre Ciências da Terra (CIESIN) usou a base de dados da GPW para modelar o crescimento populacional por país e em níveis subnacionais; 2) a Meta/Facebook treinou algoritmos de visão computacional, como detecção de bordas e duas redes neurais convolucionais (*Convolutional Neural Networks* – CNNs) independentes: SegNet e FeedbackNet, para processar imagens de satélite da DigitalGlobe®, com 50cm de resolução espacial, e assim identificar edificações. Obteve-se assim o mapeamento da presença ou ausência de áreas com infraestruturas construídas, inclusive com sensibilidade a edifícios individualizados, tornando possível a identificação de residências em áreas rurais; 3) calculou-se a densidade de edifícios em uma grade com células de 30 x 30 metros, verificando a quantidade e o tamanho dos edifícios para determinar a densidade total em cada área e período censitário; 4) a população total foi redistribuída nas células com edifícios, com base em um método dasimétrico do volume de edifícios, estimando a população de cada unidade e criando assim o mapa de densidade (Tiecke et al., 2017).

A grade *WorldPop* utiliza um conjunto de variáveis auxiliares, relacionadas à distribuição de assentamentos humanos na paisagem, para obter relações estatísticas entre essas variáveis e dados populacionais de censos e estimativas oficiais (WorldPop, n.d.). Locais de ocorrência de assentamentos, dados de uso e cobertura da terra, mapas de estradas e edifícios, locais de instalações de saúde, dados de luzes noturnas, mapeamentos de vegetação, topografia, entre outras variáveis são utilizadas para a modelagem semi-automatizada baseada em regressão *Random Forest*. O modelo *Random Forest* é usado para gerar uma superfície de ponderação que é utilizada para a redistribuição dos dados de contagens censitárias a nível nacional (Lloyd et al., 2017).

Uma síntese das características de construção e origem das principais Grades Populacionais Globais, disponíveis com estimativas de população residente para o ano de 2020, é apresentada no Quadro 1.

Quadro 1 – Principais Grades Populacionais Globais de dados de População Residente disponível para o ano de 2020.

Grade	Instituição	Método	Dados populacionais	Dados auxiliares
GPWv4	CIESIN (2018)	Ponderação pelas áreas	Censos populacionais nacionais ajustados com estimativas	Limites administrativos dos países Máscaras de corpos d'água/gelo permanente
GHS-POP	European Commission, Joint Research Centre [JRC]: Carioli et al. (2023)	Dasimétrico binário	Grade GPWv4	Mapas de volume construído (<i>GHSL-Build</i> , ajustado com estimativas de áreas urbanas da ONU)
HRSL	Facebook Connectivity Lab & CIESIN (2018)	Dasimétrico volumétrico	Grade GPWv4	Mapas de áreas construídas obtidos de imagens de satélite (50cm de resolução espacial)
WorldPop	University of Southampton: Bondarenko et al. (2020)	Modelagem <i>Random Forest</i>	Grade GPWv4	Uso do solo, edificações, luzes noturnas, vegetação, entre outras.

Elaboração: Os autores (2024).

No contexto brasileiro, o Instituto Brasileiro de Geografia e Estatística (IBGE) produziu a primeira Grade Estatística nacional utilizando os dados censitários de 2010, que incluem dados de população e domicílios ocupados. Fazendo uso de dispositivos com posicionamento GNSS (*Global Navigation Satellite System*) para capturar coordenadas geográficas dos domicílios nas áreas rurais e da integração das faces de

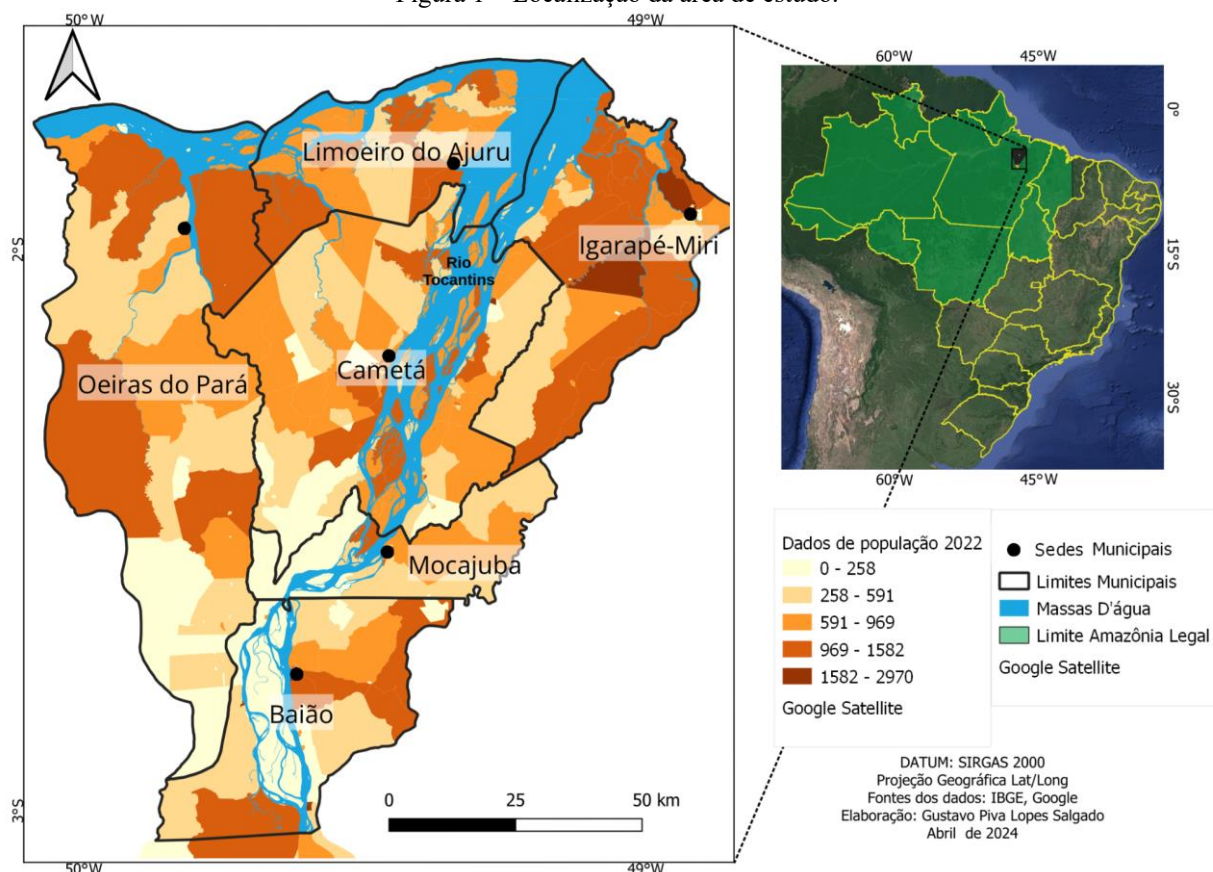
logradouros das áreas urbanas com o mapeamento censitário, o Censo de 2010 obteve dados inéditos de localização dos endereços. Apesar de muito útil para a operacionalização do processo censitário, os registros de endereços de 2010 apresentaram uma quantidade significativa sem dados de localização, o que fez com que o método de agregação não fosse suficientemente qualificado para tratar a totalidade dos dados na geração da Grade Estatística (IBGE, 2016). Deste modo, a Grade Estatística 2010 do IBGE utilizou uma abordagem híbrida, combinando agregação (*bottom-up*) e desagregação (*top-down*) de dados (IBGE, 2016). Nos setores censitários onde a ausência de localização é inferior a 50%, foi utilizada a abordagem de agregação; naqueles setores onde a ausência de localização é superior a 50%, a abordagem de desagregação foi utilizada (IBGE, 2016). Os métodos de desagregação foram baseados, em ordem de preferência: 1) dasimetria com proximidade de vias, se dados de vias estivessem disponíveis; 2) dasimetria com dados de uso e cobertura da terra, se estivessem disponíveis; 3) ponderação por área, se não existissem dados auxiliares disponíveis. O resultado foi uma grade com uma resolução de 200m para células em áreas urbanas e 1km para células em áreas rurais, que possui codificação do método utilizado em cada célula para fim de informação da incerteza associada (IBGE, 2016).

3 METODOLOGIA

3.1 Área de Estudo

A área de estudo localiza-se na região do Baixo Tocantins, no estado do Pará, distante cerca de 90 km da capital Belém e contém os municípios de Baião, Cametá, Igarapé-Miri, Limoeiro do Ajuru, Mocajuba e Oeiras do Pará (Figura 1). Apenas a porção centro-norte do município de Baião foi considerada neste estudo, em favor de se delimitar maior homogeneidade de características naturais no uso e ocupação da terra. As interações entre as características ambientais e socioeconômicas, em uma região de transição entre áreas fluviais e terra firme, definem a organização espacial das atividades humanas, influenciando a mobilidade e a distribuição da população local (Figura 1).

Figura 1 – Localização da área de estudo.



Elaboração: Os autores (2024).

Dentre os seis municípios, Cametá é o mais populoso, com mais de 130 mil habitantes, e apresenta a maior densidade populacional, com 43,55 hab/km² (Tabela 1). Contudo, registrou a segunda menor taxa de crescimento da região entre 2010 e 2022. Baião, por sua vez, obteve a maior taxa de crescimento populacional da área de estudo, que pode estar associado a atividades econômicas que atraíram migrantes (IBGE, 2022a). Oeiras do Pará, Limoeiro do Ajuru e Mocajuba apresentam os menores quantitativos populacionais nos Censos de 2010 e 2022, com população variando entre 27 e 34 mil habitantes, combinado com aumento da proporção de população residindo em áreas oficialmente urbanas, embora esse processo tenha ocorrido associado a um crescimento populacional em Limoeiro do Ajuru e Oeiras do Pará e a uma estagnação em Mocajuba. Mocajuba é o município com maior proporção da população vivendo em áreas urbanizadas, enquanto Igarapé-Miri, Limoeiro do Ajuru e Oeiras do Pará encontram-se com a maioria da população vivendo em áreas rurais. Oeiras do Pará, por sua vez, destaca-se por possuir a menor densidade populacional, com apenas 8,79 hab/km².

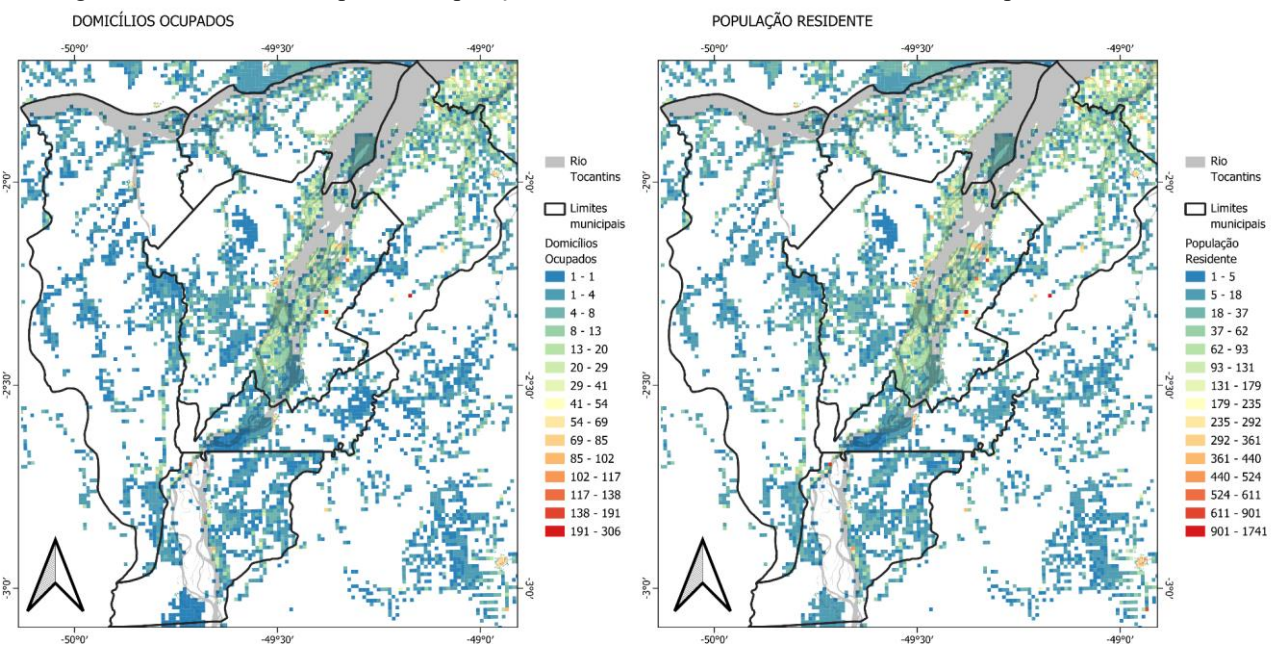
Tabela 1 – Dados demográficos dos municípios da área de estudo.

Município	População Total 2010	Proporção Rural 2010	População Total 2022	Proporção Rural 2022	Taxa de Crescimento Exponencial (2010-2022)	Densidade populacional 2022 (hab / km ²)
Baião	36.882	50 %	51.641	47 %	2,80%	13,73
Cametá	120.896	57 %	134.184	49 %	0,87%	43,55
Igarapé-Miri	58.077	55 %	64.831	55 %	0,92%	32,47
Limoeiro do Ajuru	25.021	76 %	29.569	70 %	1,39%	19,84
Mocajuba	26.731	32 %	27.198	29 %	0,14%	31,22
Oeiras do Pará	28.595	61 %	33.844	54 %	1,40%	8,79

Fonte: IBGE (2010; 2022a).

A Figura 2 apresenta a distribuição de domicílios ocupados e da população residente na área de estudo, conforme os dados da Grade Estatística IBGE 2010 (IBGE, 2016). Observa-se a concentração da população nas margens dos rios, na região de ilhas e nas áreas urbanas, principalmente nas sedes municipais, mas com extensas áreas sem a presença de população residente ou domicílios.

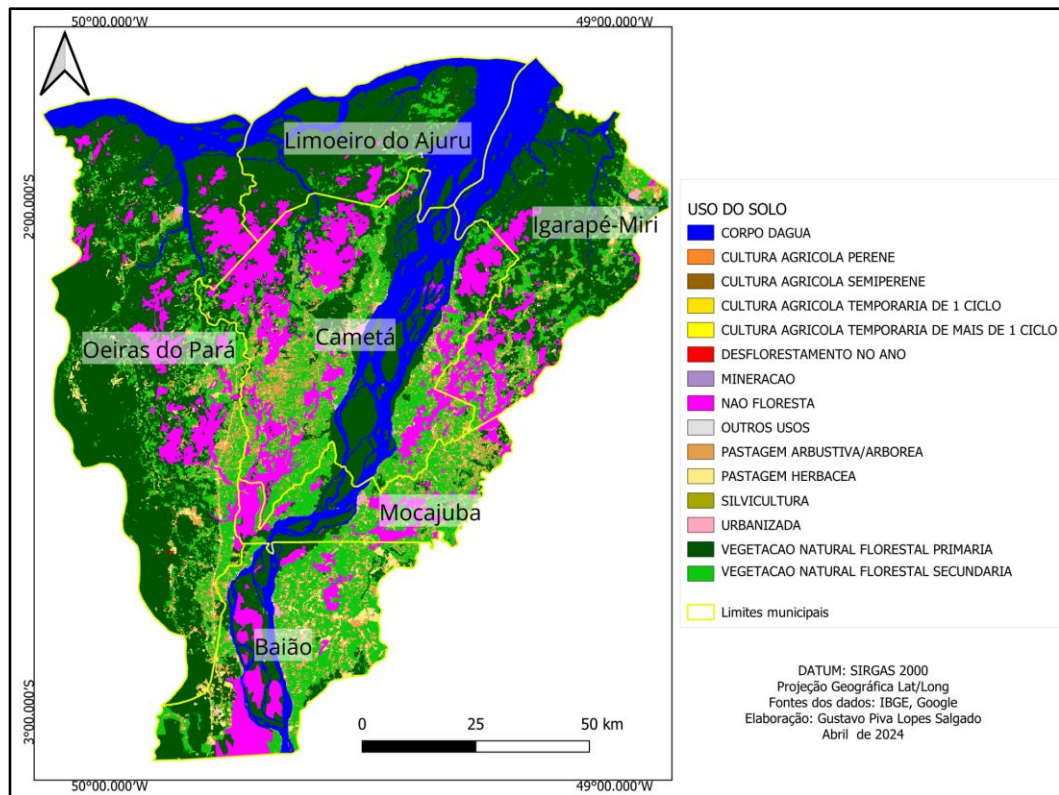
Figura 2 – Domicílios Ocupados e População Residente da Grade Estatística IBGE 2010 para a área de estudo.



Elaboração: Os autores (2024).

A floresta ombrófila densa primária cobre 43,5% da área de estudo (INPE, 2023), concentrando-se principalmente na região das ilhas e na várzea do rio Tocantins (Figura 3). Na região, cerca de 92% do desmatamento em floresta primária ocorreu até 2007 (INPE, 2023). Áreas naturais com fitofisionomia não florestal, conhecidas localmente como "campos da natureza", correspondem a 12,87% da região (INPE, 2023).

Figura 3 – Uso e cobertura da terra na área de estudo, segundo o projeto TerraClass (INPE & EMBRAPA, 2020).



Elaboração: Os autores (2024).

Segundo Souza et al. (2021), as atividades econômicas do Baixo Tocantins diferenciam-se em função dos ambientes de terra firme, várzea e ilhas. Em terra firme e várzea, além da produção do cacau, mandioca e pimenta do reino, há cultivo e coleta do açaí. Em terra firme, na última década, ocorreu a entrada dos sistemas de açaí plantado e irrigado (Souza, 2024). Pastagens ocorrem em menor escala, especialmente no município de Baião. Nas margens dos rios, principalmente em áreas de várzea próximas das comunidades ribeirinhas, a extração do açaí ocorre juntamente com outras palmeiras de menor valor comercial, como o buriti, babaçu e inajá. Nas ilhas, predomina a cobertura florestal com açaizais de manejo intensivo que, juntamente com a pesca de camarão, integram as principais atividades econômicas (Souza et al., 2021). Tais atividades representam sistemas agrários, principalmente camponeses, de grande importância para as populações e comunidades locais, resultantes de processos histórico-geográficos que deixam na paisagem sua expressão espacial (Souza et al., 2019). Entretanto, parte dessas atividades é de pequena escala, e muitas vezes associada a sistemas extrativistas ou agroflorestais sob o dossel florestal, nem sempre bem representadas nos mapeamentos sistemáticos por sensoriamento remoto de uso e cobertura da terra (Souza et al., 2019; 2023).

Deste modo, a região representa um desafio para o mapeamento das superfícies ocupadas, auxiliares na representação demográfica por grades, porque possui extensas áreas de cobertura florestal, sistemas agroflorestais de produção, diferentes proporções de população rural, heterogeneidade de ambientes e da ocupação em áreas não-florestais, além da dinâmica particular com os rios, que condicionam a mobilidade e distribuição da população residente. Adicionalmente, a região é representada por setores censitários extensos, com área média de 27,35 km², e densidades populacionais variadas de 0,14 a 15.400 hab/km² por setor (IBGE, 2022a; 2022b), que gera dificuldades de desagregação dos dados para áreas regulares de grades.

3.2 Materiais e Procedimentos Metodológicos

3.2.1 MATERIAIS

Quatro Grades Populacionais derivadas de projetos globais foram escolhidas para serem analisadas com base em dois critérios: possuir estimativas de população residente e disponibilidade de conjuntos de dados para o ano de 2020. As bases de dados utilizadas nas análises encontram-se descritas no Quadro 2.

Quadro 2 – Dados utilizados no trabalho.

Dado	Arquivo	Versão	Formato	Resolução
Estimativa populacional municípios	POP2020_20220905.xls	2020	Tabular	N/A
Censo Demográfico Nível Municipal	tabela_2_1_PA.xlsx	2022		
Censo Demográfico Nível Setores	Agregados_preliminares_por_setores_censitarios_PA.csv			
Grade Estatística IBGE	grade_id75.shp	2010	Vetorial Polígono	1km (áreas rurais) e 200m (áreas urbanas)
Malha de setores censitários do Pará	PA Malha _ Preliminar _ 2022.shp	2022		N/A
Malha territorial de municípios do Pará	PA Municípios _ 2022.shp			
CNEFE	15_PA.csv	2022	Vetorial Ponto	N/A
Gridded Population of World (GPW)	gpw_v4_population_count_rev11_2020_30_sec.tiff	2020	Raster	30 arcsec (~1 km no equador)
WorldPop	bra_ppp_2020_UNadj_constrained.tiff	2020	Raster	3 arcsec (~100 m no equador)
Global Human Settlement Pop Layer (GHS-POP)	GHS_POP_E2020_GLOBE_R2023A_4326_3ss_V1_0_R10_C13.tiff	2020	Raster	
	GHS_POP_E2020_GLOBE_R2023A_4326_3ss_V1_0_R10_C14.tiff	2020		
High-Resolution Settlement Layer (HRS�)	population_bra_northeast_2018-10-01.tiff	2020	Raster	1 arcsec (~ 30 m no equador)

Elaboração: Os autores (2024).

3.2.2 PROCEDIMENTOS METODOLÓGICOS

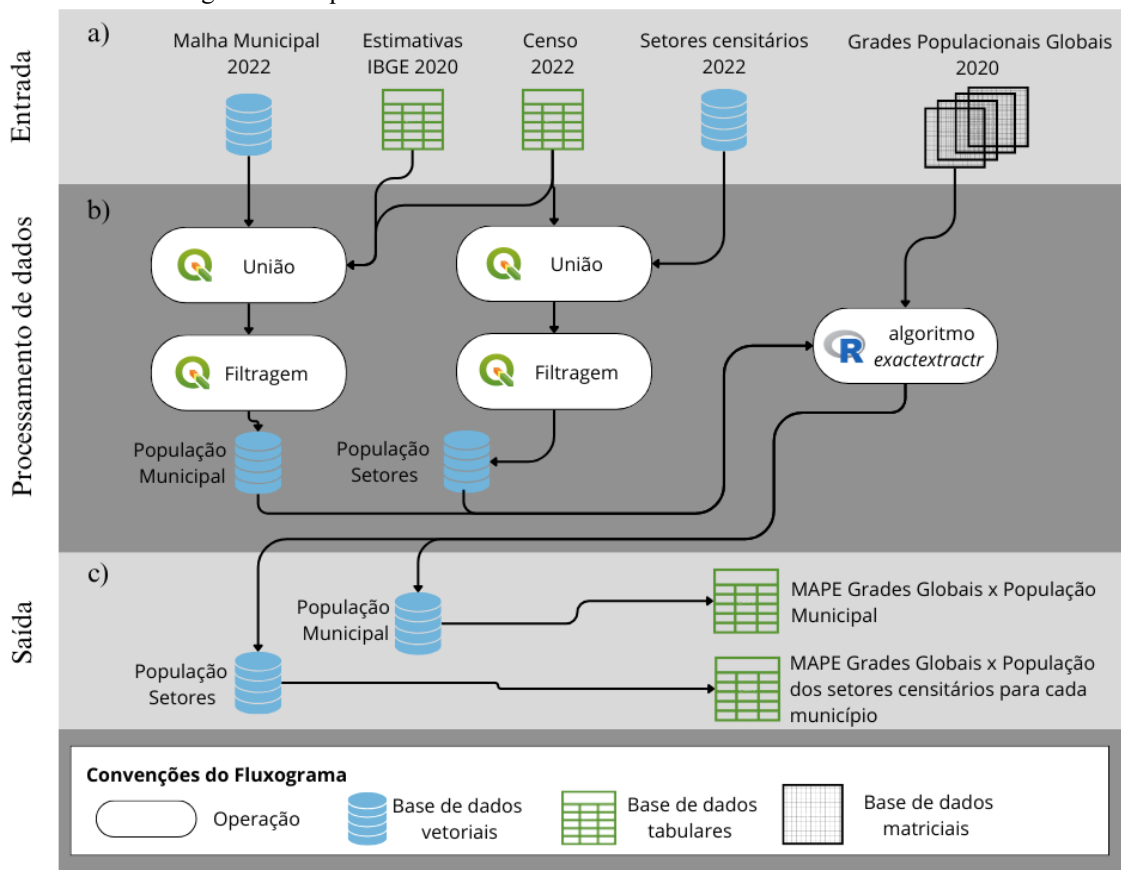
As estimativas de densidades populacionais das Grades Populacionais Globais foram analisadas quanto à concordância e aderência aos dados censitários mais recentes, ou seja, os dados do Censo Demográfico de 2022. A análise de 'concordância' compara os dados a partir da mesma unidade de medida, estimativa de contagem populacional neste caso, realizada por diferentes grades e em tempos diferentes. Por sua vez, a análise de 'aderência' descreve quão bem distribuído está o dado de população em relação ao dado altamente correlacionado à sua manifestação, neste caso, o número de domicílios. As operações de manipulação de dados espaciais foram realizadas no RStudio, um ambiente de programação da linguagem R, e no sistema de informações geográficas QGIS.

3.2.2.1 Análise de Concordância

Os dados do Censo Demográfico 2022, disponíveis para municípios e setores censitários, permitiram avaliar a concordância das estimativas das Grades Globais nessas unidades espaciais (Figura 4). Com base na

estatística zonal do pacote *exactextractr*², realizada em linguagem de programação R, as estimativas populacionais das Grades Globais foram contabilizadas para os limites municipais e dos setores censitários contidos na área de estudo. Para esse cálculo, foi considerada a área de intersecção de cada célula das grades aos limites, a partir das quais foram realizadas as somas ponderadas dos valores populacionais das grades, considerando a proporção da área de cada célula que se sobrepõe aos limites das unidades espaciais.

Figura 4 – Esquema da estrutura de trabalho da análise de concordância.



Elaboração: Os autores (2024).

Com os valores das estimativas de população das Grades Globais e dos dados de referência disponíveis por municípios e por setores censitários em arquivo tabular, calculou-se o *Mean Absolute Percentage Error* (MAPE) – Erros Absolutos Percentuais Médios – dos valores de cada Grade Populacional em relação às referências em nível municipal, tanto em malha municipal, quanto em setores censitários. O MAPE (Equação 1) avalia a precisão de acordo com a média do desvio absoluto dos valores de referência e dos valores estimados, padronizados através da divisão dos valores estimados pelos valores de referência.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Grade_i - Referência_i}{Referência_i} \right| \quad (1)$$

Especificamente para a análise em malha municipal foram utilizados dois dados de referência: 1) resultados do Censo Demográfico de 2022 (IBGE, 2022a), que é o dado de contagem mais recente disponível no Brasil, sendo assim o parâmetro mais confiável de comparação para dados populacionais, mas que possui dois anos de defasagem em relação ao ano de referência dos conjuntos de dados da Grades Globais, que é 2020; 2) estimativas de população residente nos municípios brasileiros, com data de referência em 1º de Julho de 2020 (IBGE, 2020), que são valores projetados, mas temporalmente correlatos às Grades Globais. Isso é importante para entender se os dados de saída das Grades Globais refletem as projeções oficiais. Ressalta-se

² É uma implementação de estatística zonal que consegue lidar com células de grade que são parcialmente cobertas por um polígono, ponderando a área de cobertura e extraíndo os valores através dessa ponderação (BASTON, 2023).

que este estudo não tem como propósito avaliar a precisão das estimativas populacionais do IBGE e das Grades Globais quando agregadas ao nível municipal, em comparação com uma população tida como “real” ou “exata”. As estimativas municipais para 2020 baseiam-se em projeções populacionais que apresentam limitações, especialmente em relação à capacidade dos modelos de capturar variações abruptas no ritmo de crescimento populacional em áreas menores (Campos, 2017). A qualidade dessas estimativas depende não apenas da fecundidade e da mortalidade, mas em grande parte da migração que, por sua vez, pode apresentar oscilações importantes, que incorrem em mudanças abruptas nos ritmos de crescimento (Smith & Morrison, 2005). Assim, considerando que as Grades Globais partem diretamente de estimativas da ONU, como com a GPWv4, ou indiretamente no caso das outras grades, a qualidade de seus dados já está sujeita à qualidade das estimativas oficiais. O propósito de se comparar as estimativas do IBGE para municípios em 2020 e a população agregada das grades nestes limites administrativos, portanto, é tão somente analisar a concordância dos valores das projeções populacionais realizadas pelos projetos globais para 2020, baseadas em estimativas das Nações Unidas (UN, 2015), e potenciais perdas em relação àquelas estimativas dadas pelo escritório nacional responsável pelas estatísticas populacionais brasileiras, agregada em recortes municipais (IBGE, 2020).

3.2.2.2 Análise de Aderência

A principal variável e mais comumente utilizada para explicar a distribuição da população residente no espaço é a variável que representa as infraestruturas construídas. Dados que mensuram esse tipo de variável têm sido utilizados como variável sintomática na criação das Grades Globais WorldPop, GHS-POP e HRSL. O conceito de ‘população residente’ está relacionado à ação humana de habitar determinado local, mensurado de forma indireta por imagens de satélite. As infraestruturas associadas aos locais de residência da população podem ser estimadas por sensoriamento remoto a partir de padrões de respostas espectrais de elementos comuns de áreas habitáveis, como superfícies impermeáveis (Zha et al., 2003; Lu & Weng, 2006), volumes de edificações (Tiecke et al., 2017) e luzes noturnas (Elvidge et al., 1997).

O instrumento mais detalhado que o IBGE dispõe para localizar a população no território brasileiro é o seu cadastro de endereços, para saber como a população está distribuída, onde ela está concentrada, e qual a utilização que é destinada a cada uma das edificações construídas no país (IBGE, 2024). O Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE) é o cadastro georreferenciado que contém os endereços do território nacional catalogados pelo IBGE durante suas pesquisas desde 2005 (IBGE, 2024a). Os dados do CNEFE são atualizados continuamente, de forma integral a cada censo demográfico, ou de forma pontual conforme as demandas das demais pesquisas do IBGE (IBGE, 2023). Os endereços na base do CNEFE são classificados em categorias, a fim de padronizar a identificação e análise estatística dos tipos de endereços no Brasil. Os códigos atribuídos a cada categoria de endereço são: 1 para Domicílio particular, 2 para Domicílio coletivo, 3 para Estabelecimento agropecuário, 4 para Estabelecimento de ensino, 5 para Estabelecimento de saúde, 6 para Estabelecimento de outras finalidades, 7 para Edificações em construção, e 8 para Estabelecimento religioso (IBGE, 2023). Estes dados são utilizados para a contagem de domicílios ocupados na Grade Estatística do IBGE 2010, onde domicílios de áreas rurais são os pontos de localização de domicílios do CNEFE, e domicílios de áreas urbanas são linhas vetoriais das faces de logradouros, com respectivos valores de contagem de domicílios associados (IBGE, 2016).

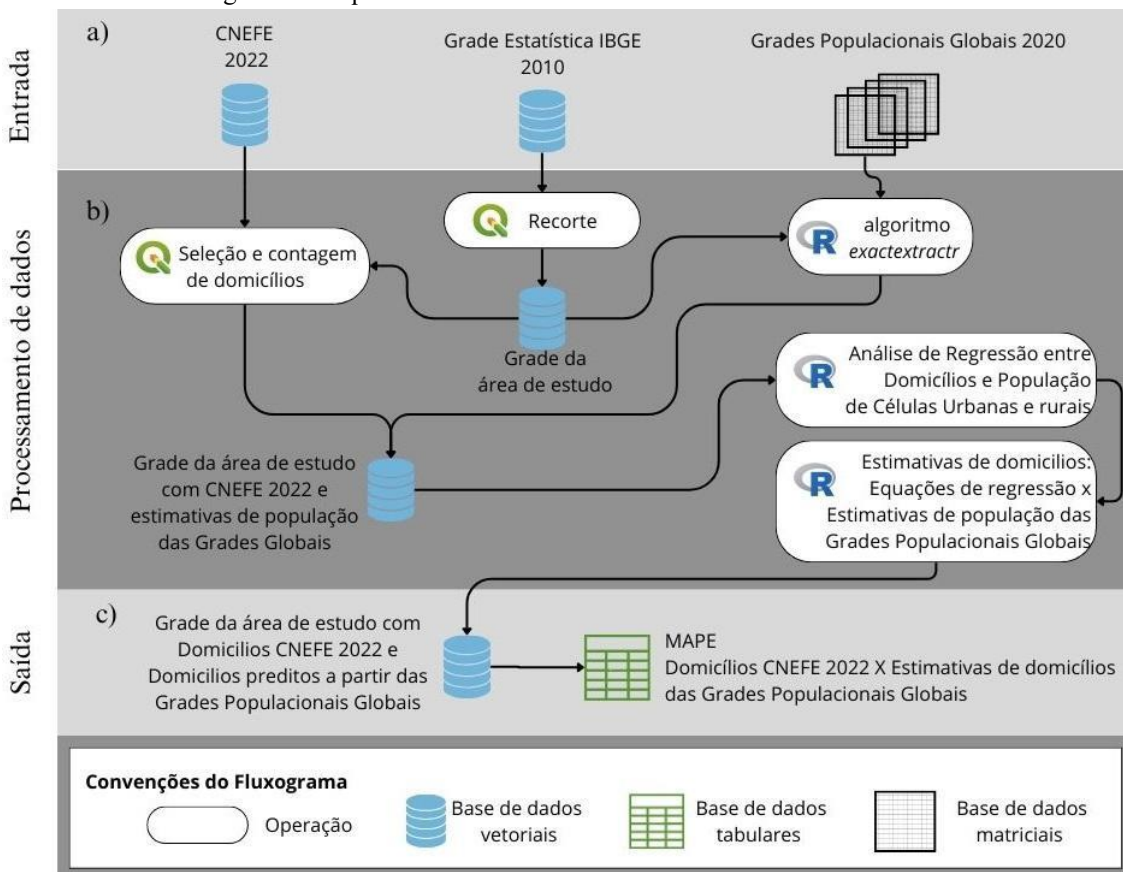
Quanto à qualidade dos dados de coordenadas geográficas do CNEFE no Censo Demográfico de 2022, ela foi garantida por um rigoroso processo de coleta, validação e correção. Os recenseadores coletaram coordenadas por meio de GNSS Global (Sistema de navegação por satélite) embarcados nos dispositivos móveis de coleta (DMC) dos recenseadores. Embora haja imprecisões associadas às características do instrumento de coleta e ao ambiente em que a coleta ocorreu, o erro quadrático médio (EQM) das coordenadas em condições ideais foram de 5,84 metros, podendo chegar até 11,71 metros em condições normais e até ultrapassar esse valor em áreas urbanas densas ou com dificuldades de acesso (IBGE, 2024a). Para garantir a qualidade, foram aplicadas técnicas de validação que ajustaram e corrigiram as coordenadas quando necessário, considerando a proximidade a pontos de referência, como faces de quadra e endereços próximos (IBGE, 2024b). O CNEFE também adotou um processo de geocodificação, atribuindo coordenadas válidas e,

em casos de falhas, utilizando informações alternativas como coordenadas anteriores ou pontos médios das faces de quadra (IBGE, 2024b). No entanto, 99,68% dos endereços do Brasil foram geocodificados ao nível de endereço, não sendo necessário nenhum outro processo de correção. No caso do estado do Pará, esse índice foi de 97,44% (IBGE, 2024b). A qualidade posicional dos dados do CNEFE reflete um alto grau de precisão e abrangência, o que lhe confere uma alta confiabilidade para ser utilizado neste estudo.

Deste modo, partindo do pressuposto que áreas construídas e residências são *proxy* da distribuição populacional, analisamos a aderência das estimativas das Grades Globais com os dados cadastrados dos domicílios CNEFE 2022 na área de estudo, considerando os recortes rural e urbano da Grade Estatística do IBGE. Em outras palavras, a análise de aderência das grades foi feita a partir da conversão das variáveis para um indicador da densidade de domicílios.

A estrutura de trabalho da análise de aderência está resumida na Figura 5. Inicialmente, a Grade Estatística 2010 do estado do Pará foi recortada para compreender a área de estudo, mantendo suas resoluções originais (1 km para áreas rurais e 250m para áreas urbanas). Da base de endereços (CNEFE 2022) foram filtrados apenas os endereços (pontos) definidos como domicílios particulares (Código 1) ou coletivos (Código 2). A grade recortada para a área de estudo foi preenchida com os pontos de domicílios por meio do operador “Contagem de pontos em polígonos” presente no QGIS. Então, a grade da área de estudo foi preenchida com os dados das Grades Populacionais por meio do algoritmo *exactextract*, considerando a contagem de valores de pixels dos rasters, ponderados pela área de sobreposição com as células da grade.

Figura 5 – Esquema da estrutura de trabalho da análise de aderência.



Elaboração: Os autores (2024).

Para instrumentalizar a análise de aderência, foi necessário mensurar a relação estatística entre os dados de domicílios ocupados ('DOM_OCU') e de população residente ('POP') da Grade Estatística do IBGE 2010. Uma análise de regressão entre a variável independente DOM_OCU e a variável dependente POP foi ajustada para a área de estudo, tanto para as células rurais quanto para as células urbanas. Para atender às premissas de adequação de um modelo de regressão, as equações de regressão entre 'domicílios' e 'população' da Grade Estatística do IBGE 2010 precisam possuir resíduos que apresentem normalidade e homocedasticidade (variância constante), para que suas distribuições estatísticas contemplem intervalos de

confiança válidos nos testes de hipóteses estatísticas (Breusch & Pagan, 1979). Assumindo-se que a relação estatística entre as variáveis não se alterou significativamente do ano base de contagem do Censo 2010 até o ano de 2022, a equação de regressão ajustada foi utilizada para prever valores de domicílios a partir das Grades Globais, que foram então comparados com o conjunto de dados de domicílios do CNEFE 2022 (IBGE, 2022a). Desta forma, avaliou-se indiretamente a aderência da distribuição das estimativas populacionais das Grades Globais, considerando os recortes intrarregionais ou intramunicipais.

Partindo da simplicidade de um modelo de regressão linear, e realizando as transformações da variável preditora e da variável resposta caso fosse necessário, as regressões entre as variáveis ‘domicílios’ e ‘população’, considerando as áreas urbanas e rurais separadamente, foram testadas de modo a se obter os modelos que atendessem às premissas de conformidade para testes estatísticos. O resultado foi o modelo de regressão para as áreas urbanas (Equação 2), que utilizou apenas uma transformação quadrática da variável resposta ‘população’. Para as áreas rurais (Equação 3), além da transformação quadrática da variável resposta, foi necessário ponderar as amostras do modelo com base no inverso da variância dos resíduos de um modelo inicial de regressão linear. Essa ponderação visou corrigir a heterocedasticidade presente nos resíduos do modelo inicial, garantindo que as amostras com maiores variâncias dos resíduos tivessem menor peso na estimativa dos parâmetros, resultando em um ajuste mais robusto, com normalidade e homoscedasticidade.

$$POP^2 = 4,866047 + (0,170471 * DOM_OCU) \quad (2)$$

$$POP^2 = 2,73807 + (0,33765 * DOM_OCU) \quad (3)$$

Os modelos de regressão obtidos de áreas urbanas e rurais apresentaram coeficientes de determinação (R^2) iguais à 0,93 e 0,90, respectivamente. Com as premissas de adequação de um modelo de regressão atendidas e com valores elevados de R^2 , pode-se afirmar que os modelos são bem ajustados aos dados. Invertendo-se as equações e isolando o termo de ‘domicílios’, foi possível obter equações capazes de calcular as densidades de domicílios na área de estudo a partir de densidades populacionais, como aquelas fornecidas pelas Grades Globais. As equações de predição de domicílios, partindo da variável população em áreas urbanas e rurais são, respectivamente:

$$\text{Predição de domicílios em áreas urbanas} = \frac{\text{Estimativa População Grade}^2 - 4,866047}{0,170471} \quad (4)$$

$$\text{Predição de domicílios em áreas rurais} = \frac{\text{Estimativa População Grade}^2 - 2,73807}{0,33765} \quad (5)$$

As equações resultantes do modelo de regressão foram utilizadas para se obter os dados de predição da densidade de domicílios em áreas urbanas e rurais para 2020 a partir dos dados de estimativas populacionais das Grades Globais. Assim, produziu-se um novo dado que constitui uma *proxy* passível de comparação direta com a densidade de domicílios por célula obtida a partir dos dados do CNEFE 2022. A partir destes dados, foram calculados Erros Absolutos Percentuais Médios (MAPEs) para analisar a distribuição dos erros dos conjuntos de dados na escala de grade celular.

4 RESULTADOS E DISCUSSÃO

4.1 Análise de Concordância

A Tabela 2 apresenta as divergências encontradas entre as estimativas populacionais derivadas das Grades Globais e as medidas de referência das Estimativas Municipais IBGE 2020 e do Censo demográfico 2022. Observa-se uma menor discrepância das Grades Populacionais Globais em relação às Estimativas Municipais IBGE para 2020. Isso se deve a todas as Grades Globais analisadas basearem-se direta, ou indiretamente, nas estimativas da ONU para população de 2020.

Tabela 2 – Erros Absolutos Percentuais Médios (MAPE) entre as estimativas de população das Grades Populacionais Globais e os dados de referência oficiais por município.

Grade Populacional 2020	Estimativa IBGE 2020	Censo 2022
GHS-Pop	4,02%	6,64%
GPWv4	11,42%	12,88%
HRSL	1,75%	4,77%
WorldPop	22,96%	24,34%

Elaboração: Os autores (2024).

A modelagem da GPWv4, com sua última revisão sendo lançada em 2018, foi realizada com base em contagens censitárias e previsões históricas e futuras de população disponíveis no relatório ‘*World Population Prospects* das Nações Unidas: revisão 11’ (CIESIN, 2018), que por sua vez é baseado nas estimativas dos escritórios de estatísticas nacionais, incluindo o IBGE (UN, 2015). Os conjuntos de contagens e estimativas da ONU para a população de 2020 são ajustadas inicialmente ao nível nacional, mas podem possuir níveis administrativos mais refinados a depender da disponibilidade dos dados de cada país (CIESIN, 2018). No caso do Brasil, os conjuntos de dados raster da GPWv4 foram construídos a partir das unidades administrativas ao nível do setor censitário de 2010 e o conjunto de dados analisados é uma projeção populacional destes recortes para o ano de 2020, desagregados ao nível de grade (CIESIN, 2018). Como GHS-Pop, HRSL e WorldPop utilizam a GPWv4 como entrada, estas primeiras estão sujeitas ao mesmo dado de referência projetado para 2020. Deste modo, entende-se que as projeções para o ano de 2020 que partem dos mesmos escritórios estatísticos, embora produzidas em momentos distintos – 2015 no caso das Grades Globais e 2020 para a estimativa municipal do IBGE, tendem a convergir para valores semelhantes, uma vez que se pressupõe similaridade dos tipos de dados, métodos utilizados e mesmo ano de projeção.

A GPWv4, por mais que seja utilizada como entrada para as outras grades, apresentou o segundo maior erro: 11,42% para as Estimativas IBGE 2020 e 12,88% para o Censo 2022, indicando a limitação da resolução de 1km em representar recortes administrativos municipais. O maior erro foi da WorldPop, que demonstrou diferenças de 22,96% e 24,34%, respectivamente. Estas diferenças podem estar associadas ao método que aplica alto grau de modelagem com diversas variáveis auxiliares, das quais algumas podem não ser sintomáticas de população na região, mas principalmente do uso do *Random Forest*, que não preserva o volume populacional de entrada (Lei et al., 2023). Observou-se as menores diferenças nas estimativas da HRSL: apenas 1,75% em relação às Estimativas IBGE 2020, o que demonstra uma maior conformidade dos dados de entrada para a escala municipal. Ela também obteve a menor diferença em relação às contagens do Censo 2022, o que mostra a robustez do projeto em captar padrões populacionais, com perdas mínimas em relação à contagem de dois anos posteriores ao seu dado de entrada projetado para 2020. Embora utilize os mesmos dados de entrada, a grade HRSL conseguiu resultados superiores aos da GHS-Pop, demonstrando as vantagens de uma maior resolução de grade – 30m em comparação a de 100m da GHS-Pop. Esse maior refinamento foi possível pela disponibilidade de imagens de satélite com resolução espacial centimétrica e identificação de áreas construídas por redes neurais convolucionais, gerando assim o principal dado auxiliar utilizado no modelo dasimétrico de criação da HRSL.

A Tabela 3 sumariza os MAPE entre as Grades Globais e os dados de referência de 2022 em nível de setores censitários. A comparação nesse nível territorial resultou em valores de erros ainda mais altos, o que demonstra a dificuldade dos conjuntos de dados em representar os volumes populacionais na escala intramunicipal. Apesar disso, a grade HRSL também foi a que obteve as menores discrepâncias em todos os municípios, seguida pela grade GHS-Pop. Este aspecto esteve evidente nos municípios de Cametá, Igarapé-Miri e Limoeiro do Ajuru e com erros praticamente idênticos em Oeiras do Pará. A grade GPWv4 obteve os maiores erros em quatro dos seis municípios, sendo superada pela WorldPop apenas em Limoeiro do Ajuru e Mocajuba, que possuem o maior e o menor percentual de população rural, respectivamente.

Tabela 3 – MAPE entre as Grades Globais 2020 e população por setor censitário do Censo de 2022, por município.

Município	GHS-Pop	GPWv4	HRSL	WorldPop
Baião	51,81 %	256,74 %	46,38 %	94,57 %
Cametá	48,14 %	135,19 %	47,28 %	103,20 %
Igarapé-Miri	46,28 %	65,03 %	44,10 %	62,47 %
Limoeiro do Ajuru	25,81 %	53,87 %	23,71 %	79,85 %
Mocajuba	62,37 %	156,47 %	49,73 %	191,77 %
Oeiras do Pará	42,67 %	99,52 %	42,51 %	84,90 %

Elaboração: Os autores (2024).

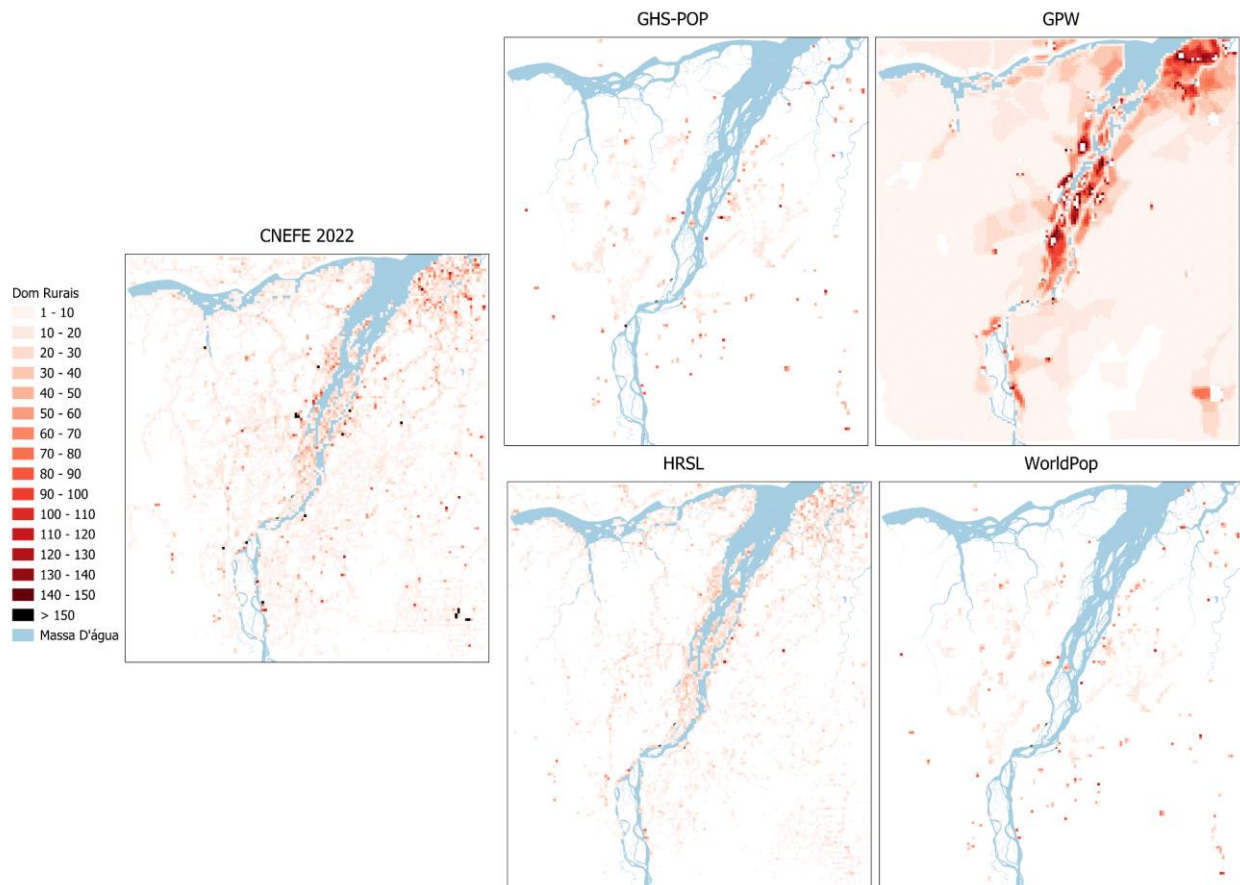
Segundo Doxsey-Whitfield et al. (2015) o principal benefício de usar ponderação de área como método de desagregação é a manutenção dos volumes dos dados de entrada. A principal desvantagem é a variabilidade da precisão das estimativas em nível de pixel, porque são uma função direta do tamanho da unidade de área de entrada, como os setores administrativos e censitários. Em regiões onde as unidades de entrada são muito grandes, a precisão dos pixels originados dessas unidades é degradada. Isso pôde ser observado nos resultados da grade GPWv4, com pixels de resolução de cerca de 1km e método de desagregação por ponderação de área. GPWv4 apresentou os segundos maiores Erros Absolutos Percentuais Médios (Tabela 2), enquanto que em referência às contagens populacionais dos setores censitários, as estimativas desta grade foram as que mais erraram em quatro dos seis municípios (Tabela 3).

Com exceção da grade WorldPop, as Grades Globais avaliadas apresentaram menos erros em Limoeiro do Ajuru do que nos outros municípios. Este município possui a maior proporção de população rural, que normalmente apresenta maior estabilidade ao longo do tempo e homogeneidade na distribuição territorial, em relação às áreas urbanas ou de transição rural-urbana, o que o torna mais previsível para as estimativas da grade. Por outro lado, Baião, Cametá e Mocajuba foram os que apresentaram as maiores discrepâncias em todos os conjuntos de dados. O município de Baião possui a maior parte da população urbana, com alta taxa de crescimento populacional para o período de 2010 a 2022. Mocajuba, por sua vez, por mais que tenha crescido pouco no período, é o mais urbanizado da área de estudo, o que aumenta os desafios associados à distribuição heterogênea da população. Nos três municípios ocorrem extensas áreas de fitofisionomias naturais não-florestais – os chamados “campos da natureza”. Em parte do ano, estas áreas apresentam resposta espectral com forte contribuição de solo exposto o que, em mapeamentos derivados de imagens de satélite, assemelham-se às áreas urbanizadas e superfícies construídas. Estas possíveis confusões entre classes de coberturas implicam em erros na redistribuição populacional dos conjuntos de dados das Grades Globais, que se baseiam na classe de superfícies construídas.

4.2 Análise de Aderência

A Figura 6 apresenta a distribuição dos domicílios preditos para as Grades Populacionais dos projetos globais pela regressão em células rurais.

Figura 6 – Domicílios CNEFE 2022 e domicílios Grades Globais 2020 preditos por regressão em células rurais.



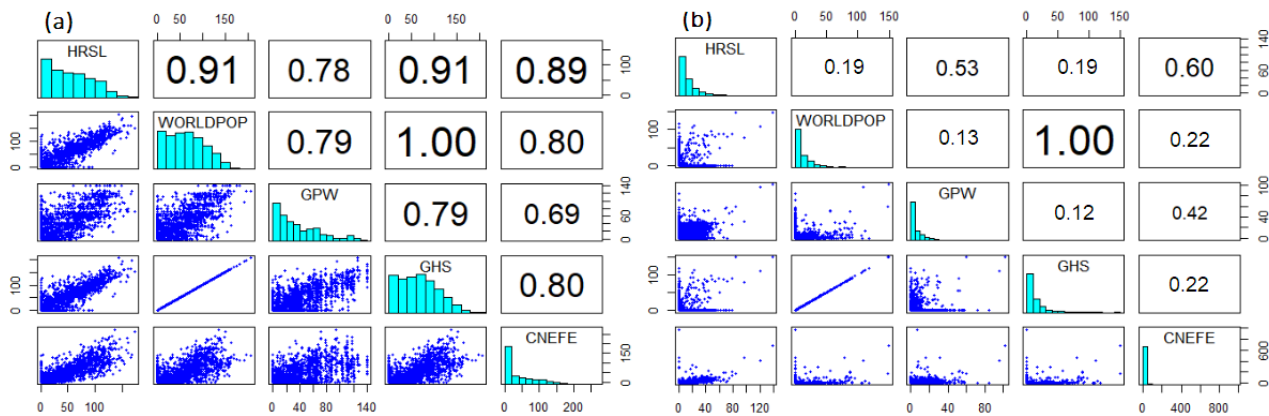
Elaboração: Os autores (2024).

Observa-se que o produto derivado da grade HRSL foi o que mais se aproximou de representar a distribuição de densidade de domicílios do CNEFE de 2022. Ressalta-se a semelhança da distribuição de domicílios nas áreas de ilhas e próximas às margens dos rios, que são os principais locais de concentração demográfica da região. Além da grade HRSL, a grade GPWv4 também capturou a concentração de domicílios associados à ocupação das margens e na região de ilhas do rio Tocantins.

Em contrapartida, as grades GHS-Pop e WorldPop subestimaram a densidade de domicílios próximos às margens e na região de ilhas, indicando que essas grades possuem limitações na modelagem para a área de estudo. A grade GHS-Pop apresentou bons valores na análise de concordância ao nível de escala do município e indica-se fortemente utilizá-la para esse fim. Seus resultados também foram muito próximos à grade HRSL na análise de concordância no nível de setor censitário. Contudo, devido aos seus problemas de subestimativas em áreas insulares e às margens de rios no teste de aderência, não é indicada a ser utilizada em análises de escalas mais refinadas, visto as incongruências na distribuição dos dados populacionais para o contexto analisado. A WorldPop, por outro lado, apresentou resultados inferiores tanto no teste de concordância quanto no de aderência, o que a torna uma Grade Populacional Global ainda menos indicada para ser utilizada em estudos de distribuição de população na Amazônia, principalmente em regiões com padrões similares de organização territorial e composição rural-urbana.

A Figura 7 apresenta as relações estatísticas entre o dado de referência de domicílios do CNEFE 2022 e os domicílios preditos através das Grades Populacionais Globais para áreas urbanas (a), com resolução de 200x200m, e para áreas rurais (b) com 1km de resolução. Observa-se que a grade HRSL apresentou os maiores coeficientes de correlação com o CNEFE: 0,89 em células urbanas e 0,60 em células rurais. Nas células urbanas, a GPWv4 obteve o menor coeficiente (0,69), enquanto nas rurais, as grades GHS-Pop e WorldPop obtiveram os menores valores (0,22 cada).

Figura 7 – Histogramas (diagonal), análise de dispersão (azul) e Coeficiente de Correlação (R) dos valores de domicílios preditos para as Grades Globais (2020) e o dado oficial CNEFE (2022) para as células urbanas (a) e rurais (b) na área de estudo.



Elaboração: Os autores (2024).

Segundo Bondarenko et al. (2020), o método de desagregação utilizado pela grade WorldPop possui a vantagem de que, quando os assentamentos e edifícios são mapeados com precisão, os resultados fornecem uma distribuição altamente precisa, evitando a previsão de valores em áreas provavelmente desabitadas. No entanto, essa vantagem também é uma limitação: se os elementos de edificações forem mapeados erroneamente (erro de inclusão), o modelo destina parte dos valores de população às células que contém esses erros, o que acarretará em uma subestimativa de população nas células com edificações identificadas corretamente, para preservar o volume dos dados. Por outro lado, se edificações reais forem omitidas em determinadas células da grade, pode ocorrer a superestimativa populacional em células que contém edificações identificadas corretamente. Isso pode explicar as grades WorldPop e GHS-Pop terem subestimado a predição de domicílios na região de ilhas e ocupação ribeirinha, enquanto apresentaram superestimação nos ambientes de terra firme. Segundo Gonçalves et al. (2021), os dados auxiliares associados às infraestruturas urbanas podem não mensurar adequadamente as infraestruturas em áreas de superfícies de vegetação florestal das ilhas e margens da massa d’água, por não possuírem um padrão típico de habitação encontrado em assentamentos humanos de maior densidade de residências. Falsos positivos também podem ocorrer, como nos casos em que áreas de vegetação não-florestais da região são confundidas com superfícies urbanas por algoritmos de classificação de imagens.

Os MAPE das predições de domicílios na área de estudo, por estrato urbano e rural encontram-se na Tabela 4. Observa-se que, para as células urbanas, os valores de MAPE variam consideravelmente entre as grades: HRSL apresentou o menor erro (11%), seguido por GHS-Pop e WorldPop (ambos com 32%) e, por fim, GPWv4 com erro significativamente maior (93%). Já para as células rurais, todas as Grades Populacionais apresentaram erros maiores, destacando-se o GPWv4 com 106%, seguido por WorldPop com 78%, GHS-Pop com 79%, e HRSL com 34%. Estes resultados indicam que a precisão das predições de domicílios é geralmente menor em áreas rurais em comparação com áreas urbanas, com a grade HRSL demonstrando uma performance superior em ambos os contextos.

Tabela 4 – Erros Absolutos Percentuais Médios (MAPE) das predições de domicílios das Grades Globais em referência aos domicílios CNEFE, por situação urbana e rural.

	GHS-Pop	GPWv4	HRSL	WorldPop
Células Urbanas	32%	93%	11%	32%
Células Rurais	79%	106%	34%	78%

Elaboração: Os autores (2024).

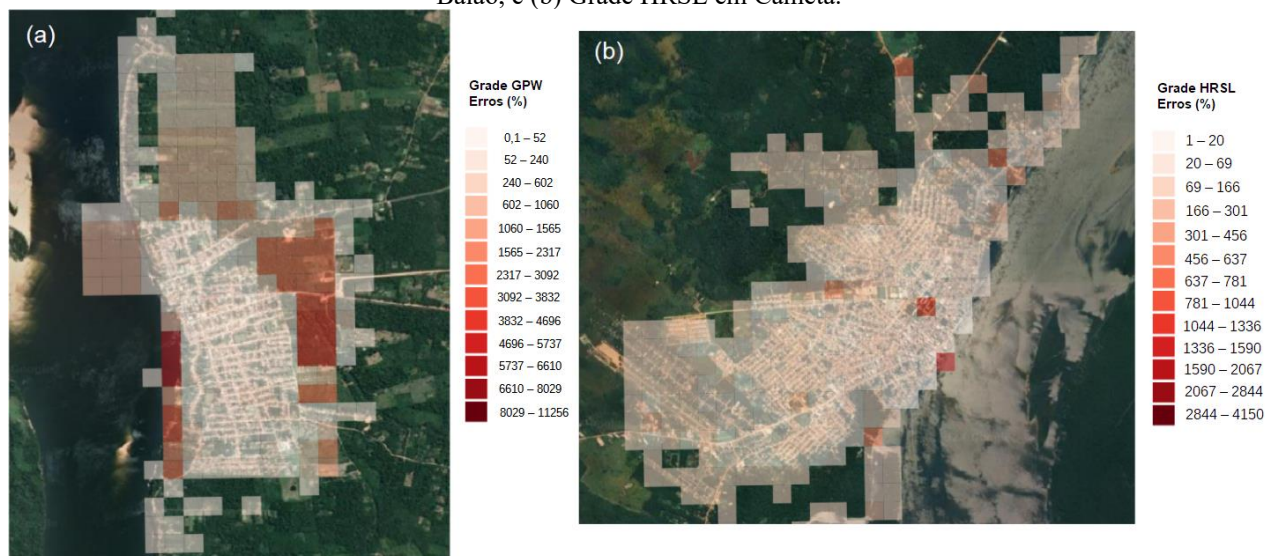
Isso está em concordância com a literatura internacional, que destaca diferentes medidas de acurácia dos modelos de representação populacional por grades conforme padrões distintos de concentração de

superfícies edificadas ou impermeáveis, mapeadas nos dados auxiliares de seus modelos. As indicações de agrupamentos populacionais variam a partir de áreas de maior adensamento urbano e diminuem ao longo de trajetórias urbano-rurais de maior dispersão de áreas construídas (Harvey, 2002; Leyk et al., 2018; Silva, 2023). A limitação da acurácia das grades em áreas rurais pode ser verificada por ambientes difíceis de serem diferenciados nos dados auxiliares de áreas impermeáveis por possuírem características de reflectância semelhantes a de outros elementos, como aqueles com categorias dominadas por grama (Wickham et al., 2013) ou solo exposto (Weng & Hu, 2008). Essas limitações ressaltam a necessidade de modelos que consigam lidar com misturas espectrais durante as estimativas de áreas construídas, bem como de sensores que tenham resolução espacial mais refinada para a delimitação desses tipos de ambientes nos dados auxiliares, para assim melhorar a representatividade das grades populacionais em contextos rurais.

Para as áreas rurais, a grade GPWv4 apresentou o segundo maior coeficiente de correlação (0,42), e em contrapartida apresentou o maior erro percentual médio para as células rurais (106%). O coeficiente de correlação indicou que a força e direção das estimativas da grade GPWv4 seguem a mesma tendência dos dados do CNEFE, e com ainda mais força do que as grades GHS-Pop e WorldPop. Mas o alto MAPE indica que, em média, essas estimativas estão numericamente mais distantes dos valores reais, possivelmente devido ao viés sistemático da baixa resolução da grade de estimativas e método de redistribuição de população ponderado pela área (*areal weighting*), com apenas a máscara d'água como auxiliar e ausência de um dado auxiliar de infraestrutura para orientar a modelagem de densidades.

O problema da resolução da grade também ocorre em células urbanas (Figura 8a), onde os altos valores de densidade disponíveis em baixa resolução, quando desagregados em grade de resolução mais fina, geram erros maiores no entorno das áreas construídas que estavam contidas no recorte de células únicas da grade GPWv4, com resolução de 1 km. A Figura 8b ilustra a variabilidade dos erros da Grade HRSL, que foi menor do que a observada nas demais Grades Populacionais Globais analisadas, com a maioria dos valores concentrados entre 1% e 69%. Observa-se que os erros mais elevados estão distribuídos de forma heterogênea na área urbana de Cametá, mas também com uma pequena tendência de concentrar-se em áreas periféricas, o que sugere a limitação da HRSL, ainda que menor, em capturar a densidade correta em áreas de menor densidade populacional ou em transição urbana-rural. Algumas poucas células no centro e na área portuária também apresentaram erros maiores, o que indica uma certa dificuldade das imagens de satélite auxiliares em capturar com precisão as características complexas e densas de certos padrões de edificações.

Figura 8 – Erros Percentuais Absolutos (APEs) das predições de domicílios em áreas urbanas: (a) Grade GPWv4 em Baião, e (b) Grade HRSL em Cametá.



Elaboração: Os autores (2024).

5 CONSIDERAÇÕES FINAIS

A qualidade das principais Grades Populacionais Globais para a representação da distribuição da população na região do Baixo Tocantins foi avaliada em termos de concordância aos dados populacionais de

referência e aderência à distribuição dos dados de domicílios. Apesar da disponibilidade dos dados censitários de 2022 em nível municipal e de setores, um melhor entendimento da qualidade dos dados das Grades Globais requer a análise direta a grades com dados populacionais oficiais (des)agregados em unidades de tamanhos padronizados, tal como a Grade Estatística resultante do Censo de 2010.

A análise de aderência deste trabalho realizou uma avaliação indireta, com os dados do CNEFE já divulgados pelo Censo Demográfico de 2022, apresentando uma alternativa inédita de análise baseada na localização dos domicílios como uma *proxy* de avaliação da qualidade da distribuição das estimativas de população das Grades Globais. Essa análise de aderência detalhada destacou padrões específicos de subestimativa ou superestimativa relacionados às características de distribuição populacional local, avançando no entendimento dos desafios e limitações das Grades Globais na representação da distribuição da população residente em escalas mais refinadas e em contextos particulares.

Observou-se neste estudo que a qualidade das Grades Populacionais Globais varia entre os municípios, demandando atenção especial às implicações que a escolha da grade pode ter sobre o estudo. A generalização dos resultados deste estudo é limitada, de modo que estão intimamente ligados às características específicas e às particularidades da região estudada. Vale ressaltar que as características dos dados e métodos baseados em sensoriamento remoto para a modelagem populacional, em regiões como a do Baixo Tocantins, podem ser limitados na representação da infraestrutura domiciliar e das particularidades da ocupação territorial que definem o mapeamento das variáveis auxiliares utilizadas nos modelos globais, calibrados para representar padrões de uma urbanização industrial-capitalista. Assim, generalizações para outras regiões, mesmo na Amazônia brasileira, devem ser realizadas com cautela, especialmente em áreas que diferem substancialmente em composição rural-urbana, densidade populacional, infraestrutura e dinâmica no uso e cobertura da terra. Contudo, o estudo é capaz de oferecer um referencial para identificar potencialidades e limitações das principais Grades Globais em contextos que se distanciam dos padrões de concentração populacional em cidades, como o de esquemas típicos da ocupação tradicional amazônica.

A grade HRSL demonstrou consistência e superioridade em todas as áreas avaliadas e experimentos realizados. Além disso, o método e dados utilizados para gerá-la adequam-se melhor para captar as particularidades e especificidades da distribuição populacional nos territórios. Assim, quando os dados de população do Censo 2022 em Grade Estatística estiverem disponíveis, será relevante replicar este estudo, a fim de validar os resultados com base nos dados oficiais refinados como referência.

Espera-se que o trabalho contribua com o debate sobre a representatividade das Grades Populacionais dos grandes projetos globais em contextos particulares, como na Amazônia brasileira, amplamente sub-representada. Dessa forma, estudos como o nosso buscam contribuir no entendimento das limitações e potencialidades dessas Grades, oferecendo resultados e subsídios metodológicos para a avaliação crítica na escolha do produto mais confiável para a representação da distribuição da população em contextos de especificidades regionais.

Agradecimentos

Ao Programa de Pós-Graduação em Sensoriamento Remoto do Instituto Nacional de Pesquisas Espaciais (INPE). À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES – Código financeiro 001). Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, número de processo: 130659/2023-1). Ao Projeto Harmonize, financiado pela Fundação Wellcome Trust (número de processo: 224694/Z/21/Z).

Contribuição dos Autores

Conceptualização: G. P. L. S., A. P. D. e S. A.; Curadoria dos dados: G. P. L. S.; Análise formal: G. P. L. S.; Aquisição de financiamento: CAPES, CNPq e Wellcome Trust; Investigação: G. P. L. S.; Metodologia: G. P. L. S.; Administração do projeto: G. P. L. S., A. P. D. e S. A.; Recursos: G. P. L. S.; Software: G. P. L. S.; Supervisão: A. P. D. e S. A.; Validação: G. P. L. S.; Visualização: G. P. L. S.; Redação: G. P. L. S., A. P. D. ; Redação–revisão e edição: G. P. L. S., A. P. D., B. V. A. e S. A.

Conflitos de Interesse

Os autores informam não haver conflitos de interesses.

Referências

- Amaral, S., Gavlak, A. A., Escada, M. I. S. & Monteiro, A. M. V. (2012). Using remote sensing and census tract data to improve representation of population spatial distribution: case studies in the Brazilian Amazon. *Population and Environment*, 34, 142–170. <https://doi.org/10.1007/s11111-012-0168-2>.
- Archila Bustos, M. F., Hall, O., Niedomysl, T. & Ernstson, U. (2020). A pixel level evaluation of five multitemporal global gridded population datasets: a case study in Sweden, 1990–2015. *Population and Environment*, 42, 255–277. <https://doi.org/10.1007/s11111-020-00360-8>.
- Bai, Z., Wang, J., Wang, M., Gao, M., & Sun, J. (2018). Accuracy assessment of multi-source gridded population distribution datasets in China. *Sustainability*, 10(5), 1363. <https://doi.org/10.3390/su10051363>
- Baston, D. (2023). *exactextractr version (v0.10.0)* [R package]. GitHub. <https://github.com/isciences/exactextract>.
- Bondarenko, M., Kerr, D., Sorichetta, A. & Tatem, A. J. (2020). *Census/projection-disaggregated gridded population datasets, adjusted to match the corresponding UNPD 2020 estimates, for 183 countries in 2020 using Built-Settlement Growth Model (BSGM) outputs* [Data set]. WorldPop: University of Southampton, UK. <https://hub.worldpop.org/geodata/summary?id=49921>
- Breusch, T. S. & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, 47(5), 1287–1294. <https://doi.org/10.2307/1911963>
- Calka, B., & Bielecka, E. (2020). GHS-POP accuracy assessment: Poland and Portugal case study. *Remote Sensing*, 12(7), 1105. <https://doi.org/10.3390/rs12071105>
- Campos, J. (2017). *Estimativas populacionais a partir de dados orbitais de média resolução espacial: aplicações em municípios da Região Metropolitana de Belo Horizonte*. [Tese de doutorado, Universidade Federal de Minas Gerais] Repositório digital de teses e dissertações da UFMG. https://repositorio.ufmg.br/bitstream/1843/FACE-B27P7Y/1/tese___j_rvis_campos.pdf.
- Cardoso, A. C. D., Dal'Asta, A. P. & Monteiro, A. M. V. (2023). O que é o urbano na Amazônia contemporânea? Implicações para a vigilância em saúde no bioma. *Cadernos de Saúde Pública*, 39(9), e00129723. <https://doi.org/10.1590/0102-311XPT129723>
- Carioli, A., Schiavina, M., Freire, S. & MacManus, K. (2023). *GHS-POP R2023A - GHS population grid multitemporal (1975-2030)*. European Commission, Joint Research Centre (JRC) [Dataset]. <https://doi.org/10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE>
- Center for International Earth Science Information Network [CIESIN], Columbia University. (2018). *Documentation for the Gridded Population of the World, Version 4 (GPWv4), Revision 11 Data Sets*. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <https://doi.org/10.7927/H45Q4T5F>.
- Chen, R., Yan, H., Liu, F., Du, W., & Yang, Y. (2020). Multiple Global Population Datasets: Differences and Spatial Distribution Characteristics. *ISPRS International Journal of Geo-Information*, 9(11), 637. <https://doi.org/10.3390/ijgi9110637>.
- Dahmm, H. & Rabiee, M. (2020). *Leaving no one off the map: A guide for gridded population data for sustainable development*. Thematic Research Network on Data and Statistics (TREnds) of the UN Sustainable Development Solutions Network (SDSN), in support of the POPGRID Data Collaborative. https://www.popgrid.org/sites/default/files/documents/Leaving_no_one_off_the_map.pdf
- Doxsey-Whitfield, E., MacManus, K., Adamo, S. B., Pistolesi, L., Squires, J., Borkovska, O. & Baptista, S. R. (2015). Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4. *Papers in Applied Geography*, 1(3), 226–234.

<https://doi.org/10.1080/23754931.2015.1014272>

- Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., Davis, E. R. & Davis, C. W. (1997). Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. *International Journal of Remote Sensing*, 18(6), 1373-1379. <https://doi.org/10.1080/014311697218485>
- Facebook Connectivity Lab & Center for International Earth Science Information Network [CIESIN] - Columbia University. (2018). *High Resolution Settlement Layer (HRSL) - Source imagery for HRSL © 2016 DigitalGlobe - Brazil Population Density Maps and Demographic Estimates (2015-2020) [Data set]*. <https://data.humdata.org/dataset/brazil-high-resolution-population-density-maps-demographic-estimates>.
- Fearnside, P. M (2006). Desmatamento na Amazônia: dinâmica, impactos e controle. *Acta Amazonica*, Manaus, 36(3), 395-400. <https://doi.org/10.1590/S0044-59672006000300018>
- Gonçalves, G. C., Oliveira, L. M. de, Dal'Asta, A. P. & Amaral, S. (2021). Geoinformação Para a Visibilidade das Áreas Urbanas de Cidades Amazônicas. *Revista Geoaraguaia*, 11(Especial), 149-165. <https://periodicoscientificos.ufmt.br/ojs/index.php/geo/article/view/12742>.
- Hallisey, E., Tai, E., Berens, A., Wilt, G., Peipins, L., Lewis, B., Graham, S., Flanagan, B. & Lunsford, N. (2017). Transforming Geographic Scale: A Comparison of Combined Population and Areal Weighting to Other Interpolation Methods. *International Journal of Health Geographics*, 16(29). <https://doi.org/10.1186/s12942-017-0102-z>
- Harvey, J. T. (2002). Estimating census district populations from satellite imagery: Some approaches and limitations. *International Journal of Remote Sensing*, 23(10), 2071-2095. <https://doi.org/10.1080/01431160110075901>
- Instituto Brasileiro de Geografia e Estatística [IBGE]. (2010). *Censo Demográfico 2010 [Dataset]*. Rio de Janeiro: IBGE. <https://censo2010.ibge.gov.br>.
- Instituto Brasileiro de Geografia e Estatística [IBGE]. (2016). *Grade Estatística IBGE 2010*. Rio de Janeiro: IBGE. <https://biblioteca.ibge.gov.br/visualizacao/livros/liv102043.pdf>
- Instituto Brasileiro de Geografia e Estatística [IBGE]. (2020). Estimativas de população residente por município, data de referência em 1º de Julho de 2020. *Diário Oficial da União em 27 de agosto de 2020*. <https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html?edicao=28674>.
- Instituto Brasileiro de Geografia e Estatística [IBGE]. (2022a). *Censo Demográfico 2022 [Dataset]*. Rio de Janeiro: IBGE. <https://censo2022.ibge.gov.br>.
- Instituto Brasileiro de Geografia e Estatística [IBGE]. (2022b). *Malha de setores censitários do Pará - Edição 2022 [Dataset]*. Rio de Janeiro: IBGE. https://www.ibge.gov.br/geociencias/downloads-geociencias.html?caminho=organizacao_do_territorio/malhas_territoriais/malhas_de_setores_censitarios__divisoes_intramunicipais/censo_2022_preliminar/setores/shp/UF.
- Instituto Brasileiro de Geografia e Estatística [IBGE]. (2022c). *Malha territorial municipal brasileira - Edição 2022 [Dataset]*. Rio de Janeiro: IBGE. <https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/15774-malhas.html>.
- Instituto Brasileiro de Geografia e Estatística [IBGE]. (2023). *Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE) do Censo 2022*. Rio de Janeiro: IBGE. <https://www.ibge.gov.br/estatisticas/sociais/populacao/38734-cadastro-nacional-de-enderecos-para-fins-estatisticos.html>.
- Instituto Brasileiro de Geografia e Estatística [IBGE]. (2024a). *Censo Demográfico 2022 - Coordenadas geográficas dos endereços no Censo Demográfico 2022: nota metodológica n. 01*. Rio de Janeiro: IBGE. <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2102063>
- Instituto Brasileiro de Geografia e Estatística [IBGE]. (2024b). *Censo Demográfico 2022 - Cadastro Nacional de Endereços para Fins Estatísticos: Nota metodológica n. 04*. Rio de Janeiro: IBGE. <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2102091>

- Instituto Nacional de Pesquisas Espaciais [INPE]. (2023). *Programa de Monitoramento da Amazônia e Demais Biomas [PRODES]: Desmatamento anual [Dataset]*. <https://terrabrasilis.dpi.inpe.br/downloads/>
- Instituto Nacional de Pesquisas Espaciais [INPE] & Empresa Brasileira de Pesquisa Agropecuária [EMBRAPA]. (2020). *Uso e cobertura da terra do Bioma Amazônia - TerraClass [Data set]*. <http://terrabrasilis.dpi.inpe.br/downloads/>
- Kuffer, M., Owusu, M., Oliveira, L., Sliuzas, R. & van Rijn, F. (2022). The missing millions in maps: Exploring causes of uncertainties in global gridded population datasets. *ISPRS International Journal of Geo-Information*, 11(7), 403. <https://doi.org/10.3390/ijgi11070403>
- Lei, Z., Xie, Y., Cheng, P. & Yang, H. (2023). From auxiliary data to research prospects, a review of gridded population mapping. *Transactions in GIS*, 27(1), 3–39. <https://doi.org/10.1111/tgis.13020>
- Leyk, S., Uhl, J. H., Balk, D. & Jones, B. (2018). Assessing the accuracy of multi-temporal built-up land layers across rural-urban trajectories in the United States. *Remote Sensing of Environment*, 204, 898–917. <https://doi.org/10.1016/j.rse.2017.08.035>
- Leyk, S., Gaughan, A. E., Adamo, S. B., de Sherbinin, A., Balk, D., Freire, S., Rose, A., Stevens, F. R., Blankespoor, B., Frye, C., Comenetz, J., Sorichetta, A., MacManus, K., Pistolesi, L., Levy, M., Tatem, A. J. & Pesaresi, M. (2019). The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use. *Earth System Science Data*, 11(3), 1385–1409. <https://doi.org/10.5194/essd-11-1385-2019>
- Lloyd, C. T., Sorichetta, A. & Tatem, A. J. (2017). High resolution global gridded data for use in population studies. *Scientific Data*, 4(1), 170001. <https://doi.org/10.1038/sdata.2017.1>
- Lu, D. & Weng, Q. (2006). Use of impervious surfaces in urban land-use classification. *Remote Sensing of Environment*, 102(1-2), 146–156. <https://doi.org/10.1016/j.rse.2006.02.010>
- Nagle, N. N., Battenfield, B. P., Leyk, S. & Spielman, S. (2013). Dasymetric Modeling and Uncertainty. *Annals of the American Association of Geographers*, 104(1), 80–95. <https://doi.org/10.1080/00045608.2013.843439>
- Nobre, C. A., Obregón, G. O., Marengo, J. A., Fu, R. & Poveda, G. (2009). Characteristics of Amazonian Climate: Main Features. In: Keller, M., Bustamante, M., Gash, J. & Silva Dias, P. (Ed.) *Amazonia and Global Change*. American Geophysical Union. <https://doi.org/10.1029/2009GM000903>
- Openshaw, S. (1984). *The modifiable areal unit problem*. Norwich: Geo Books.
- Renner, K., Schneiderbauer, S., Pruß, F., Kofler, C., Martin, D. & Cockings, S. (2018). Spatio-temporal population modelling as improved exposure information for risk assessments tested in the Autonomous Province of Bolzano. *International Journal of Disaster Risk Reduction*, 27, 470–479. <https://doi.org/10.1016/j.ijdrr.2017.11.011>
- Schiavina, M., Melchiorri, M., Pesaresi, M., Politis, P., Carneiro Freire, S. M., Maffenini, L., Florio, P., Ehrlich, D., Goch, K., Carioli, A., Uhl, J., Tommasi, P. & Kemper, T. (2023). *GHS-L data package 2023 (JRC133256)*. Publications Office of the European Union. <https://doi.org/10.2760/098587>
- Silva, D. (2023). *VaLin-Pop: uma nova grade populacional para a região metropolitana do Vale do Paraíba e Litoral Norte, SP*. [Dissertação de Mestrado, Instituto Nacional de Pesquisas Espaciais] Repositório digital de teses e dissertações do INPE. <http://mtc-m21d.sid.inpe.br/col/sid.inpe.br/mtc-m21d/2023/10.19.20.37/doc/thisInformationItemHomePage.html>
- Smith, S. K. & Morrison, P. A. (2005). Small-Area and Business Demography. In: Poston, D. L. & Micklin, M. (Orgs.) *Handbook of Population*. Handbooks of Sociology and Social Research. Springer, Boston, MA. https://doi.org/10.1007/0-387-23106-4_26
- Souza, A. R., Escada, M. I. S., Marujo, R. de F. B. & Monteiro, A. M. V. (2019). Cartografia do Invisível: Revelando a Agricultura de Pequena Escala com Imagens Rapideye na Região do Baixo Tocantins, PA. *Revista Do Departamento De Geografia*, 38, 137–153. <https://doi.org/10.11606/rdg.v38i1.151603>
- Souza, A. R., Adorno, B. V., Gonçalves, G. C., Bragion, G. da R., Oliveira, K. D., Escada, M. I. S., Reis, M. S., Sant'Anna, S. J. S. & Amaral, S. (2021). *Paisagens e uso da terra em núcleos populacionais e*

- estabelecimentos rurais da região do Baixo Tocantins - Pará*. São José dos Campos-SP: INPE. 81 p. Relatório Técnico de Atividade de Campo de 2018 e 2019 dos Cursos de Pós-Graduação em Sensoriamento Remoto e em Ciências do Sistema Terrestre do INPE, em parceria com a Universidade Federal do Pará. <http://mtc-m21d.sid.inpe.br/col/sid.inpe.br/mtc-m21d/2021/06.18.16.50/doc/thisInformationItemHomePage.html>.
- Souza, A. R., Amaral, S., Dal'Asta, A. P., Escada, M. I. S. & Monteiro, A. M. V. (2023) Modelos dasimétricos para desagregar dados do Censo Agropecuário. In: *Anais do XX Simpósio Brasileiro de Sensoriamento Remoto*. São José dos Campos: INPE. <http://marte2.sid.inpe.br/col/sid.inpe.br/marte2/2023/05.16.12.12/doc/156100.pdf>
- Souza, A. R. (2024). *Os sistemas do açaí na paisagem amazônica: elementos para análise da economia do açaí da região do Baixo Tocantins-PA*. [Tese de Doutorado, Instituto Nacional de Pesquisas Espaciais]. Repositório digital de teses e dissertações do INPE. <https://mtc-m21d.sid.inpe.br/col/sid.inpe.br/mtc-m21d/2024/09.03.22.32/doc/publicacao.pdf>.
- Tiecke, T. G., Liu, X., Zhang, A., Gros, A., Li, N., Yetman, G., Kilic, T., Murray, S., Blankespoor, B., Prydz, E. B. & Dang, H. H. (2017). *Mapping the world population one building at a time*. arXiv. <https://doi.org/10.48550/arXiv.1712.05839>.
- United Nations [UN] (2015). *UN World Population Prospects: Revision 11 [Data Booklet]*. <https://www.un.org/en/development/desa/publications/world-population-prospects-2015-revision.html>.
- Weng, Q. & Hu, X. (2008). Medium Spatial Resolution Satellite Imagery for Estimating and Mapping Urban Impervious Surfaces Using LSMA and ANN. *IEEE Transactions on Geoscience and Remote Sensing*, 46(8), 2397-2406. <https://doi.org/10.1109/TGRS.2008.917601>
- Wickham, J. D., Stehman, S. V., Gass, L., Dewitz, J., Fry, J. A., & Wade, T. G. (2013). Accuracy assessment of NLCD 2006 land cover and impervious surface. *Remote Sensing of Environment*, 130, 294–304. <https://doi.org/10.1016/j.rse.2012.12.001>
- WorldPop. (n.d.). *WorldPop methods*. University of Southampton. Retrieved [June 2024], from <https://www.worldpop.org/methods/>
- Xu, Y., Ho, H. C., Knudby, A. & He, M. (2021). Comparative assessment of gridded population data sets for complex topography: A study of Southwest China. *Population and Environment*, 42, 360–378. <https://doi.org/10.1007/s11111-020-00366-2>
- Yin, X., Li, P., Feng, Z., Yang, Y., You, Z. & Xiao, C. (2021). Which Gridded Population Data Product Is Better? Evidences from Mainland Southeast Asia (MSEA). *ISPRS International Journal of Geo-Information*, 10(10), 681. <https://doi.org/10.3390/ijgi10100681>
- Zha, Y., Gao, J. & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, 24(3), 583-594. <https://doi.org/10.1080/01431160304987>
- Zhao, T. & Wang, W. (2023). Coordination Dynamics between Population Change and Built-Up Land Expansion in Mainland China during 2000–2020. *Sustainability*, 15(22), 16059. <https://doi.org/10.3390/su152216059>
- Zhou, Y., Ma, M., Shi, K. & Peng, Z. (2020). Estimating and Interpreting Fine-Scale Gridded Population Using Random Forest Regression and Multisource Data. *ISPRS International Journal of Geo-Information*, 9(6), 369. <https://doi.org/10.3390/ijgi9060369>

Biografia do autor principal



Gustavo Piva Lopes Salgado nasceu em 1990 no município de Campinas, São Paulo. É bacharel em Geografia pela Universidade de Campinas (UNICAMP - 2020), especialista em Geoprocessamento pelo Centro Universitário SENAC de São Paulo (2023) e mestrando em Sensoriamento Remoto pelo Instituto Nacional de Pesquisas Espaciais (INPE) em São José dos Campos/SP. Atualmente atua como intérprete de imagens de satélite no Projeto de Monitoramento Anual da Supressão da Vegetação Nativa Brasileira (PRODES) do INPE. Desenvolve automação de processos, seleção e análise visual de imagens, validação de dados, assim como realiza análises de dados espaciais.



Esta obra está licenciada com uma Licença [Creative Commons Atribuição 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/) – CC BY. Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original.