

Instituto de Pesquisa Econômica Aplicada

**Secretaria Nacional de Habitação do
Ministério do Desenvolvimento Regional**

Termo de Execução Descentralizada n. 01/2019 SNH/MDR e Ipea

Pesquisa de Núcleos Urbanos Informais no Brasil

**Produto 12 - Relatório de Caracterização, Tipologias e Modelos para
Identificação de Núcleos Urbanos Informais**

Elaboração

**Flávia da Fonseca Feitosa
Gilmara Gonçalves
Luis Felipe Bortolatto da Cunha
Pedro Reis Simões
Cleandro Krause
Juliana Gomes Petrarolli**

Versão de 23 de agosto de 2021

Pesquisa de Núcleos Urbanos Informais no Brasil

Coordenação

Cleandro Henrique Krause (titular) e Marco Aurélio Costa (suplente) – Técnicos de Planejamento e Pesquisa da Diretoria de Estudos e Políticas Regionais, Urbanas e Ambientais (Dirur) do Ipea

Equipe da Pesquisa de Núcleos Urbanos Informais no Brasil – bolsistas e colaboradores

Alexandrina Saldanha Sobreira de Moura – FUNDAJ / Ipea (PNPD)
Ana Carolina Campos de Melo – Ipea (PNPD)
André Simionato Castro – Ipea (PNPD)
Bruno Gallina – UFRGS
Cátia Wanderley Lubambo – FUNDAJ / Ipea (PNPD)
David Melo Van Den Brule – Ipea (PNPD)
Elisa Escosteguy Utzig – Ipea (PNPD)
Fernanda Balestro – Ipea (PNPD)
Fernanda Carolina Vieira da Costa – Ipea (PNPD)
Flávia da Fonseca Feitosa – UFABC / Ipea (PNPD)
Gilmara Gonçalves – Ipea (PNPD)
Gabriel Moraes de Outeiro – UNIFESSPA / Ipea (PNPD)
Giuliana de Freitas – Ipea (PNPD)
Guilherme Frizzi Galdino da Silva – Ipea (PNPD)
Heleniza Ávila Campos – UFRGS / Ipea (PNPD)
Juliana Gomes Petrarolli – Ipea (PNPD)
Kaiena Thyelle Malaquias – FUNDAJ
Livia Salomão Piccinini – UFRGS
Luis Felipe Bortolatto da Cunha – Ipea (PNPD)
Manoela Guedes Ferreira Jordão de Vasconcelos – Ipea (PNPD)
Marcela Rodrigues Santos – Ipea (PNPD)
Mariana Roberti Bomtempo – Ipea (PNPD)
Miriam Francisca Rodrigues Couto – Ipea (PNPD)
Paulo Somlanyi Romeiro – Ipea (PNPD)
Pedro Reis Simões – Ipea (PNPD)
Rafael Gonçalves Gumiero – UNIFESSPA
Raquel de Mattos Viana – FJP / Ipea (PNPD)
Rosana Denaldi – UFABC / Ipea (PNPD)
Sergio Moreno Redón – UNIFESSPA / Ipea (PNPD)
Tatiana Mamede Salum Chaer – Ipea (PNPD)
Thaís Pires Rubioli – Ipea (PNPD)
Tiago Gonçalves da Silva – Ipea (PNPD)
Valéria Barroso da Silveira – Ipea (PNPD)

LISTA DE FIGURAS

- FIGURA 1. Imagem esquemática da integração de dados distintos em uma grade celular.
- FIGURA 2. Sobreposição da grade celular NUI com a grade estatística IBGE.
- FIGURA 3. Processo de construção das áreas ocupadas (presença humana), Polo Recife.
- FIGURA 4. Comparação entre os Modelos Digitais de Elevação JAXA e SRTM, Polo Brasília.
- FIGURA 5. Índice de forma.
- FIGURA 6. Processo metodológico para integração da grade celular
- FIGURA 7. Box-plot: (a) renda média dos responsáveis pelos domicílios (RenRespMedia) e (b) alfabetização dos responsáveis por domicílios (NRespAlf).
- FIGURA 8. Box-plot: (a) população dependente (RenPopDependente), (b) idade média dos responsáveis por domicílios (NRespldade) e (c) mulheres responsáveis por domicílios (NRespFem).
- FIGURA 9. Box-plot: (a) declividade (Declividade) e (b) proporção do NUI com APP de curso d'água (APP30m).
- FIGURA 10. Box-plot:(a) cobertura de vias carroçáveis (Vias50m) e (b) forma das quadras (IndiceForma).
- FIGURA 11. Box-plot: (a) abastecimento de água por rede geral (DomSAqua), (b) ausência de ligação à rede de esgoto ou fossa séptica (DomSRedeEsg), (c) esgoto a céu aberto (DomSEsg) e (d) ausência de banheiro de uso exclusivo dos moradores.
- FIGURA 12. Box-plot: (a) acesso à energia elétrica (DomSEne) e (b) ausência de medidor de uso exclusivo (DomSRedeEsg).
- FIGURA 13. Box-plot: (a) ausência de coleta de lixo por serviço de limpeza (DomSColLix) e (b) entorno com lixo acumulado nos logradouros (DomCLixAc).
- FIGURA 14. Box-plot - Porcentagem de domicílios em NUI: (a) sem pavimentação (DomSPav), (b) sem calçada (DomSCal) e (c) sem iluminação pública.
- FIGURA 15. Polo Belo Horizonte: Tipologias BH-1, BH-2 e BH-3.
- FIGURA 16. Polo Belo Horizonte: Tipologia BH-1.
- FIGURA 17. Belo Horizonte: Tipologia BH-2.
- FIGURA 18. Belo Horizonte: Tipologia BH-3.
- FIGURA 19. Brasília: Tipologias BSB-1, BSB-2, BSB-3 e BSB-4.
- FIGURA 20. Brasília: Tipologia BSB-1.
- FIGURA 21. Brasília: Tipologia BSB-2.
- FIGURA 22. Brasília: Tipologia BSB-3.
- FIGURA 23. Brasília: Tipologia BSB-4.
- FIGURA 24. Juazeiro do Norte: Tipologias JN-1, JN-2 e JN-3.
- FIGURA 25. Juazeiro do Norte: Tipologia JN-1.
- FIGURA 26. Juazeiro do Norte: Tipologia JN-2.
- FIGURA 27. Juazeiro do Norte: Tipologia JN-3.
- FIGURA 28. Marabá: Tipologias MB-1, MB-2 e MB-3.
- FIGURA 29. Marabá: Tipologia MB-1.
- FIGURA 30. Marabá: Tipologia MB-2.
- FIGURA 31. Marabá: Tipologia MB-3.
- FIGURA 32. Porto Alegre: Tipologias POA-1, POA-2, POA-3 e POA-4.
- FIGURA 33. Porto Alegre: Tipologia POA-1.
- FIGURA 34. Porto Alegre: Tipologia POA-2.

FIGURA 35. Porto Alegre: Tipologia POA-3.

FIGURA 36. Porto Alegre: Tipologia POA-4.

FIGURA 37. Recife: Tipologias RE-1, RE-2, RE-3 e RE-4.

FIGURA 38. Recife: Tipologia RE-1.

FIGURA 39. Recife: Tipologia RE-2.

FIGURA 40. Recife: Tipologia RE-3.

FIGURA 41. Recife: Tipologia RE-4.

FIGURA 42. Coeficiente Kappa dos modelos de acordo com o limiar de classificação.

FIGURA 43. Curva ROC dos modelos.

FIGURA 44. Superfícies de probabilidade (modelo "Y=AGSN") do polo Belo Horizonte (municípios Ribeirão das Neves, Contagem e Esmeraldas): (a) AGSN sobre a superfície de probabilidade; (b) NUI sobre a superfície de probabilidade.

FIGURA 45. Superfícies de probabilidade do modelo do polo Juazeiro do Norte (Bairro João Cabral): (a) AGSN sobre a superfície de probabilidade (modelo Y=AGSN); (b) NUI sobre a superfície de probabilidade (modelo Y=AGSN); (c) Imagem orbital da área; (d) NUI sobre a superfície de probabilidade (modelo Y=NUI).

FIGURA 46. Superfícies de probabilidade (modelo "Y=AGSN") do polo Brasília (Bairro Brasília Teimosa): (a) NUI sobre a superfície de probabilidade (modelo Y=AGSN); (b) Imagem orbital da área.

LISTA DE QUADROS

QUADRO 1. Dados levantados para a caracterização e identificação dos NUI.

QUADRO 2. Parâmetros da projeção adotados na grade celular NUI.

QUADRO 3. Integração do código da célula em que cada domicílio familiar cadastrado no cadúnico está localizado

QUADRO 4. Soma ou média das variáveis cadúnico condicionada pelo código da célula

QUADRO 5. Variáveis selecionadas para a análise de agrupamentos

QUADRO 6. Aspectos comparativos entre as metodologias CEM, MAPPA e NUI.

LISTA DE TABELAS

TABELA 1. Núcleos Urbanos Informais (NUI) nos polos da pesquisa: Número, total de domicílios, área total, área média e densidade média de domicílios.

TABELA 2. Tipologias dos polos e variáveis quantitativas. Todas as variáveis correspondem ao percentual de domicílios, com exceção de NUIDenDom (domicílios/ha), Declividade (%) e RenRespMédia (R\$).

TABELA 3. Tipologias do Polo Belo Horizonte.

TABELA 4. Tipologias do Polo Brasília.

TABELA 5. Tipologias do Polo Juazeiro do Norte.

TABELA 6. Tipologias do Polo Marabá.

TABELA 7. Tipologias do Polo Porto Alegre.

TABELA 8. Tipologias do Polo Recife.

TABELA 9. Aglomerados Subnormais dos polos da pesquisa classificados em tipologias.

TABELA 10. Núcleos Urbanos Informais dos polos da pesquisa classificados em tipologias.

TABELA 11. Parâmetros estimados dos modelos logísticos: Polos Belo Horizonte, Distrito Federal e Juazeiro do Norte, Marabá, Porto Alegre e Recife.

TABELA 12. Métricas obtidas pelos modelos.

LISTA DE APÊNDICES apresentados em separado

APÊNDICE 1 - Quadro das variáveis construídas e integradas na grade celular NUI

APÊNDICE 2 - Geocodificação do CadÚnico

APÊNDICE 3 - Análise Comparativa de Técnicas de Classificação

Sumário

1. Apresentação	8
2. Construção e Integração de Variáveis.....	9
2.1 Unidades Espaciais de Análise	10
2.2 Construção de Variáveis.....	13
2.3 Integração de Variáveis.....	19
3. Caracterização dos Núcleos Urbanos Informais a partir de Dados Secundários	22
3.1 Características Gerais dos NUI	22
3.2 Características da População	23
3.3 Características Físico-Ambientais e Unidades de Conservação.....	25
3.4 Cobertura de Vias Carroçáveis e Forma das Quadras.....	27
3.5 Infraestrutura e Entorno dos Domicílios.....	27
4. Tipologias de NUI	33
4.1 Polo Belo Horizonte	36
4.2 Polo Brasília.....	39
4.3 Polo Juazeiro do Norte.....	42
4.4 Polo Marabá.....	46
4.5 Polo Porto Alegre	50
4.6 Polo Recife	55
4.7 Comparação entre Aglomerados Subnormais e Núcleos Urbanos Informais	59
5. Modelos para a Identificação de NUI	64
5.1 Metodologias relacionadas e definição de diretrizes	64
5.2. Metodologia para Identificação de NUI.....	68
6. Considerações Finais	81
7. Referências	83

1. Apresentação

O presente relatório refere-se ao Produto 12 previsto no Plano de Trabalho da Pesquisa de Núcleos Urbanos Informais no Brasil, estabelecida por meio do Termo de Execução Descentralizada n. 01/2019 celebrado entre a Secretaria Nacional de Habitação (SNH/MDR) e o Instituto de Pesquisa Econômica Aplicada (IPEA). O relatório explora a utilização de dados provenientes de fontes diversas, complementares ao levantamento realizado em campo, para auxiliar na identificação e caracterização de Núcleos Urbanos Informais (NUI), além de propor tipologias de NUI, escopo básico do Produto 12.

A partir desses dados, devidamente sistematizados e integrados, foram conduzidas análises com os seguintes objetivos: (a) complementação da caracterização dos NUI dos seis polos da pesquisa (Belo Horizonte, Brasília, Juazeiro do Norte, Marabá, Porto Alegre e Recife) realizada no Produto 11, focando na comparação entre os polos (análise inter-polo); (b) comparação, para cada polo, das características dos NUI com as das demais áreas do polo (análise intra-polo, dentro e fora de NUI); (c) identificação e caracterização de tipologias de NUI para cada polo, explicitando as distintas facetas da precariedade (análise intra-polo, inter-NUI); e (d) desenvolvimento de uma metodologia para a identificação de NUI que possa ser aplicada em outras regiões do país para auxiliar futuros levantamentos de campo e/ou na análise da qualidade de mapeamentos existentes.

A caracterização dos NUI, identificação de tipologias, e desenvolvimento metodológico para a identificação de NUI baseou-se na utilização de um conjunto de variáveis obtidas por meio da integração de dados de natureza e fontes distintas, incluindo produtos de sensoriamento remoto, dados de Censo Demográfico e do Cadastro Único para Programas Sociais do Governo Federal (CadÚnico). O levantamento e processamento dos dados utilizados para a construção e integração dessas variáveis é descrito no **Capítulo 2**.

O **Capítulo 3** apresenta uma análise das características dos NUI dos seis polos da pesquisa, comparando as diferenças e similaridades entre os polos e destacando as dimensões da precariedade predominantes em cada um deles. Além da análise comparativa inter-polos, o capítulo também compara, para cada polo, as características dos NUI com as das demais áreas, analisando o quanto distintas são as condições das áreas localizadas dentro e fora dos NUI.

O **Capítulo 4** apresenta uma análise intra-polo, baseada na construção de tipologias de NUI para cada polo e na descrição e comparação dessas tipologias. Essa análise destaca subgrupos de NUI significativos em cada polo, evidenciando suas características e como se diferenciam dos demais subgrupos. O Capítulo também compara os dados dos NUI, classificados por tipologias, com os dos Aglomerados Subnormais (AGSN) do IBGE. Dado que os NUI identificados pela pesquisa são mais numerosos do que os AGSN, a comparação busca avaliar se os AGSN podem ser considerados "amostras" das distintas tipologias de NUI, passíveis de serem consideradas na construção de modelos para a identificação de NUI, objetivo do **Capítulo 5**.

O Capítulo 5 apresenta o desenvolvimento e aplicação de uma metodologia que visa auxiliar na identificação de NUI por intermédio da construção de modelos para a geração de superfícies de probabilidade relacionadas à presença de NUI. Essas superfícies de probabilidade podem ser utilizadas como um plano de informação para subsidiar trabalhos de campo e/ou para análise da qualidade de bases de dados existentes (tais como a dos AGSN ou informações fornecidas por prefeituras).

Por fim, o **Capítulo 6** apresenta as considerações finais deste relatório.

2. Construção e Integração de Variáveis

Este capítulo descreve o processo de levantamento e processamentos de dados considerados relevantes para a caracterização dos núcleos urbanos informais, bem como para o desenvolvimento de uma metodologia que auxilie na sua identificação. Esses dados são complementares às informações coletadas em campo e provêm de fontes diversas, conforme apresentado no Quadro 1. Os dados considerados são abertos (com exceção do CadÚnico, que demanda cuidados específicos para a manutenção de sigilo¹) e estão disponíveis para todo território nacional.

Os dados, por apresentarem fontes e naturezas distintas, são disponibilizados em unidades espaciais de análise diversas (tais como setores censitários, células ou pontos de localização), o que demandou processamentos específicos para a integração dos mesmos em uma base comum. A base utilizada como referência foi uma grade celular com resolução de 100 x 100m. A partir dessa base, foi possível detectar a diferença entre as áreas (células) classificadas como NUI e as não classificadas como NUI (**Capítulo 3**), bem como construir modelos para identificar áreas com maior probabilidade de serem NUI (**Capítulo 5**). Para a caracterização e comparação dos NUI existentes nos distintos polos (**Capítulo 3**) e posterior construção de tipologias (**Capítulo 4**), os dados da grade celular foram agregados por NUI.

QUADRO 1. Dados levantados para a caracterização e identificação dos NUI.

Dado	Fonte
Aglomerados Subnormais 2019.	IBGE, 2020
Núcleos Urbanos Informais, conforme pesquisa de campo.	IPEA, 2021.
Cadastro Único para Programas Sociais do Governo Federal (CadÚnico).	BRASIL, 2020b.
Censo Demográfico 2010 - Resultados do universo agregados por setores censitários.	IBGE, 2010.
Atlas do Desenvolvimento Humano	PNUD, IPEA e FJP, 2016.
Modelos Digitais de Terreno provenientes da Shuttle Radar Topography Mission (SRTM).	NASA, 2010.
Logradouros extraídos da plataforma OpenStreetMap.	OpenStreetMap, 2020.
Dados de hidrografia da Fundação Brasileira para o Desenvolvimento Sustentável.	FBDS, 2015.

¹ O CadÚnico é o instrumento de identificação e caracterização socioeconômica das famílias de baixa renda para seleção de beneficiários e a integração dessas pessoas a programas sociais do Governo Federal, podendo ser utilizado também nas esferas Estadual e Municipal. O Ipea recebeu as bases de dados do Cadastro Único do Ministério do Desenvolvimento Social e Combate à Fome (MDS), sucedido pelo Ministério da Cidadania. Esta instituição fornecedora das bases exigiu a assinatura de termo de compromisso de manutenção de sigilo específico dos pesquisadores que acessarem as bases. Tais termos dispõem que os pesquisadores manusearão as bases de dados identificados apenas por necessidade de serviço; manterão cautela quanto à exibição dos dados (em tela, impressora ou gravação); não se ausentarão do terminal sem encerrar a sessão de uso das bases; e manterão sigilo dos dados, abstendo-se de revelá-los ou divulgá-los. A divulgação de resultados com base nesses dados deve passar por um procedimento de desidentificação de qualquer indivíduo presente na base, contendo apenas estimativas e valores populacionais que não possam ser desagregados ao nível do elemento ou similares.

Unidades de conservação de proteção integral.	MMA, 2020.
Faixas de servidão de linhas de alta tensão.	ANEEL, 2020.
Faixas de servidão de dutos.	ANP, 2020.

Fonte: Elaboração própria, 2021.

A partir dos dados apresentados no Quadro 1 foram construídas variáveis que representam aspectos relevantes para a caracterização dos núcleos urbanos informais, a saber:

- a. **Forma urbana**, incluindo informações sobre a área do NUI, densidade de domicílios, irregularidade da forma de quadra ou bolsões de ocupação e cobertura de vias carroçáveis;
- b. **Características físico-territoriais**, tais como declividade e curvatura do terreno, bem como áreas com restrições à ocupação (unidades de conservação, faixa de servidão de linhas de alta tensão e dutovias);
- c. **Características das edificações, entorno e infraestrutura**.
- d. **Características sociodemográficas**, incluindo aspectos relacionados à renda, trabalho, educação e saúde da população.

O quadro contendo todas as variáveis construídas e integradas na grade celular encontra-se no Apêndice 1. Os dados e metadados do projeto, assim como os scripts desenvolvidos na linguagem R para a construção de variáveis, poderão ser disponibilizados posteriormente, em repositório cujas condições de acesso serão definidas conjuntamente pelo Ipea e pela SNH/MDR.

2.1 Unidades Espaciais de Análise

2.1.1 Geração da grade celular

A diversidade dos dados e variáveis considerados nas análises apresentadas neste relatório demanda esforços de integração dos mesmos em uma base de dados unificada, uma vez que são disponibilizados em formatos distintos. Tomemos, como exemplo, as variáveis censitárias que estão espacializadas de forma operacional a partir dos setores censitários, as variáveis de declividade que estão distribuídas em pixels (ou células) ou mesmo as unidades de conservação que são apresentadas na forma de polígonos condizentes com suas localizações e dimensões. Para que se possa realizar análises conjuntas a partir desses dados é necessário integrá-los em uma mesma base espaço-temporal, garantindo que todas as variáveis compartilhem da mesma unidade espacial de análise. Estabeleceu-se, para a presente pesquisa, que a integração desses dados consideraria a construção de uma grade celular composta por células regulares de 100 por 100 metros, cuja área é compatível com a de uma quadra média. Também optou-se pela construção de uma grade que fosse uma subdivisão da grade estatística do IBGE², com o intuito de facilitar a integração dos resultados

² A grade estatística do IBGE, publicada em 2016, foi gerada com o intuito de disseminar dados estatísticos a partir dos microdados do universo do Censo (IBGE, 2016). Disponibiliza dados populacionais em células de 200 por 200 metros nas áreas urbanas e de 1 por 1 km nas áreas rurais.

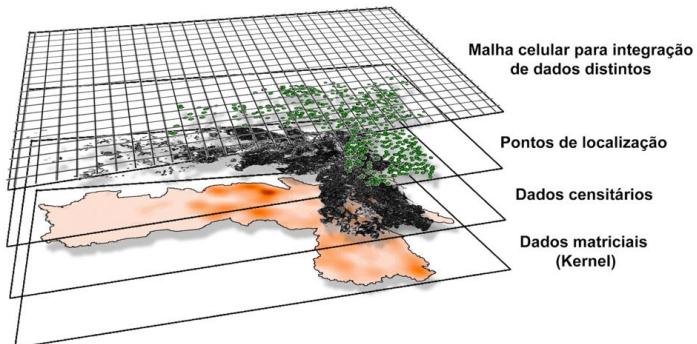
produzidos para ambas. Para a realização de análises restritas à caracterização dos núcleos urbanos informais, os dados foram posteriormente integrados aos polígonos dos NUI.

Uma grade celular é um dado vetorial formado por um conjunto de células, onde cada célula é representada por um polígono que tem a ele associados vários atributos. Cada polígono tem uma localização espacial e é representativo de uma porção territorial, mantendo assim suas propriedades e relações com as células vizinhas. Supõe-se também que a célula seja internamente homogênea, ou seja, quaisquer dois pontos distintos no interior de uma célula apresentarão as mesmas propriedades.

Além do potencial já descrito das grades celulares, de permitir a integração de dados provenientes de unidades geográficas distintas como as político-administrativas, físicas ou operacionais, a grade celular possui também outras duas vantagens. A primeira delas é a estabilidade espaço-temporal, uma vez que as suas unidades não estão sujeitas a alterações nos seus limites físicos. A segunda vantagem está relacionada a sua possibilidade de adaptação a recortes espaciais, dado que as células podem ser agregadas ou desagregadas, se adaptando assim a diferentes recortes espaciais.

A seguir, na Figura 1, é possível observar uma imagem esquemática que representa a forma como dados distintos podem ser integrados utilizando uma grade celular única.

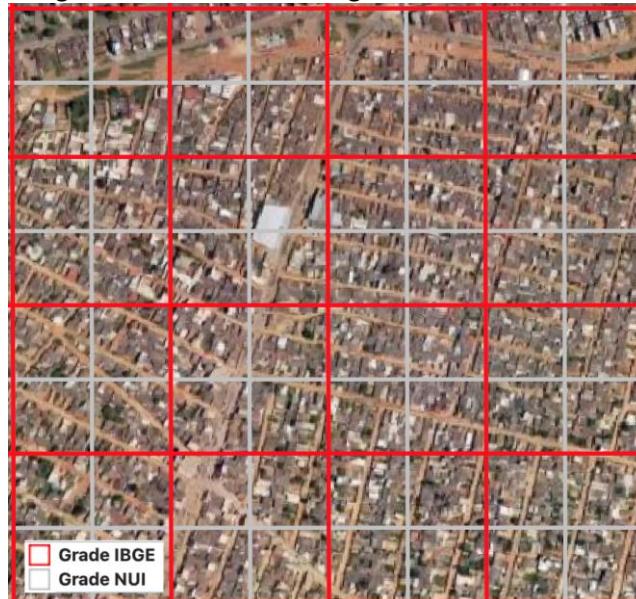
FIGURA 1. Imagem esquemática da integração de dados distintos em uma grade celular.



Fonte: Frizzi e Pinho (2017).

A escolha de tornar a grade celular dos modelos NUI uma subdivisão da grade estatística do IBGE se deu com o intuito de facilitar o intercâmbio de informações entre as duas bases. Dessa forma, é possível que variáveis obtidas a partir da grade estatística IBGE, tais como densidade populacional ou de domicílios, possam ser transferidas para a grade celular NUI de forma direta. Sendo as dimensões das células urbanas da grade estatística do IBGE de 200 por 200 metros, e as dimensões das células da grade celular NUI de 100 por 100 metros, uma célula urbana da grade estatística IBGE corresponde a quatro células da grade NUI (Figura 2).

FIGURA 2. Sobreposição da grade celular NUI com a grade estatística IBGE.



Fonte: elaboração própria, 2020.

A grade celular dos modelos NUI adotou a projeção Equivalente de Albers, a mesma utilizada na geração da grade estatística IBGE, uma vez que, apresentando como característica principal a equivalência em área, minimiza as distorções decorrentes da grande extensão territorial do país (IBGE, 2016). O *datum* horizontal adotado é o SIRGAS2000. Os parâmetros da projeção utilizada estão descritos no Quadro 2.

QUADRO 2. Parâmetros da projeção adotados na grade celular NUI.

Meridiano central	-54
Latitude de origem	-12
1º Paralelo padrão	-2
2º Paralelo padrão	-22
Origem E:	5.000.000
Origem N	10.000.000

Fonte: IBGE, 2016.

2.1.2 Áreas ocupadas

O processo de construção e integração das variáveis depende da utilização de uma variável indicadora de presença humana, seja para distribuição dos dados populacionais nas superfícies, seja para o recorte da grade celular e seleção das células que serão objeto de análise estatística. Sendo assim, distintas fontes de informação indicadoras de presença humana foram levantadas, analisadas e avaliadas quanto a sua adequabilidade aos objetivos e escala de análise da pesquisa.

No decorrer da análise, optou-se pela construção de base indicadora da presença humana a partir da combinação de distintas fontes de dados. Nesse processo, buscou-se a construção de uma base a mais inclusiva possível, dado que é comum a presença de NUI em áreas antes não ocupadas. As fontes de dados consideradas foram as seguintes:

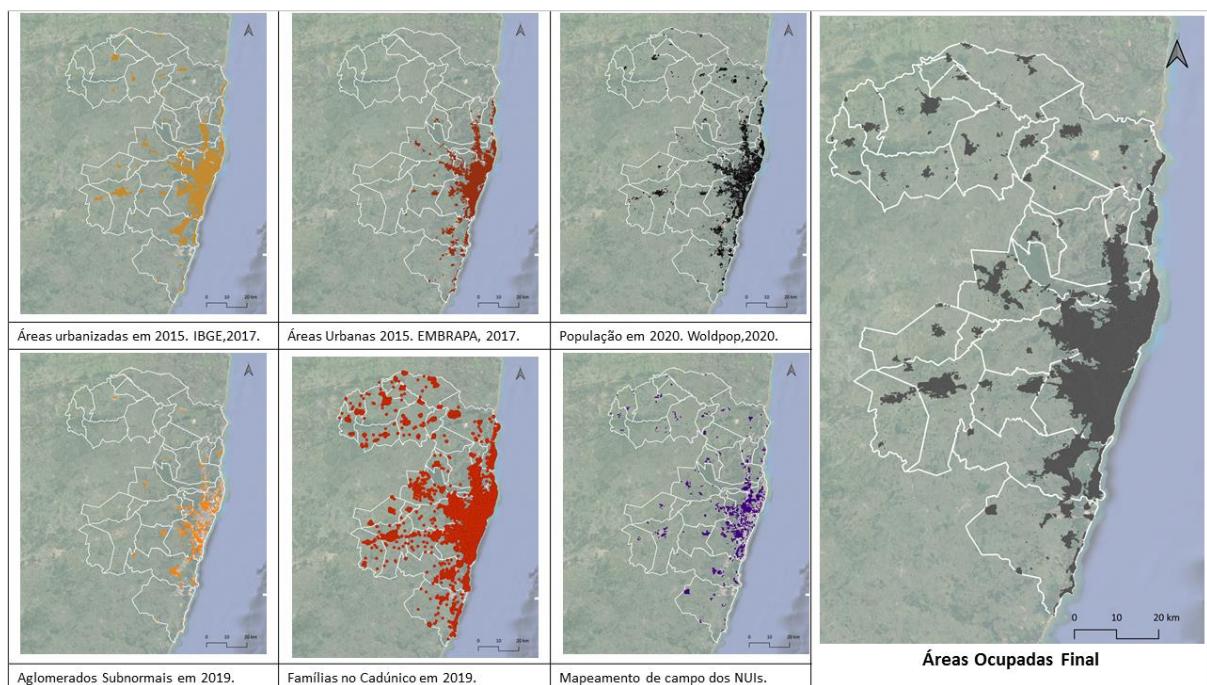
- Áreas urbanizadas do Brasil em 2015 (IBGE,2017);

- Áreas urbanas no Brasil em 2015 (EMBRAPA, 2017);
- Distribuição espacial da população em 2020 no Brasil (*The spatial distribution of population in 2020, Brazil*) (World Pop, 2020);
- Mapeamento de Aglomerados Subnormais em 2019 (IBGE, 2020);
- Células com presença de famílias cadastradas no CadÚnico em 2019;
- Mapeamento de campo dos NUIs realizado no âmbito da pesquisa.

Cabe salientar que muitos desses dados são passíveis de atualização frequente, tais como os oriundos de produtos de sensoriamento remoto e do CadÚnico.

Todos os dados, disponibilizados originalmente em formato vetorial, foram compatibilizados, projetados na projeção cônica de área igual de Albers (EPSG 102033) e, em seguida, transformados em dados matriciais de 30 por 30 metros para integração com a grade celular (Figura 3).

FIGURA 3. Processo de construção das áreas ocupadas (presença humana), Polo Recife.



Fonte: Elaboração própria, 2021.

2.2 Construção de Variáveis

2.2.1 Características físico-ambientais e unidades de conservação

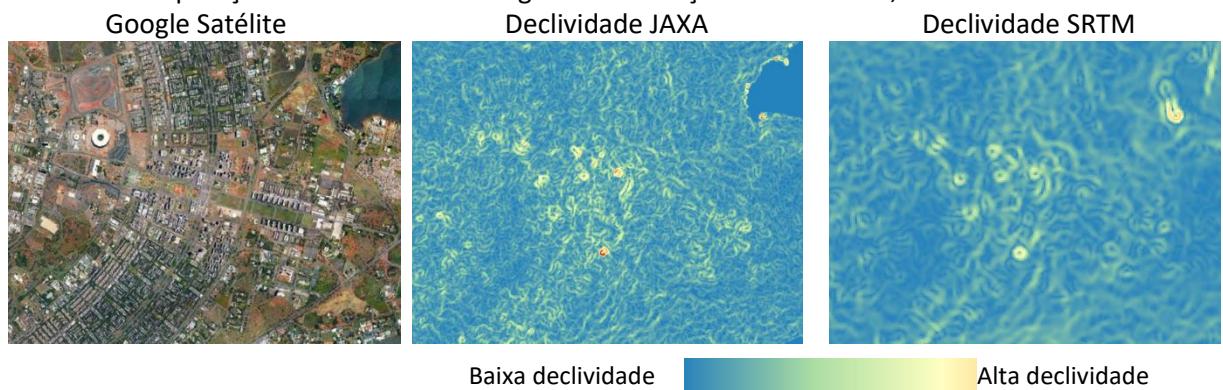
2.2.1.1 Declividade

Para a construção da variável de declividade do terreno é necessária uma base com informações de elevação. Inicialmente foram considerados dois Modelos Digitais de Elevação (MDE). Os MDE são dados do tipo matricial que representam de forma computacional o relevo da superfície terrestre (topografia). Os dois MDE analisados foram o MDE da Missão Topográfica Radar Shuttle (SRTM) (NASA, 2017) e o MDE Alos Palsar (JAXA, 2011).

Análises comparativas entre os dois modelos selecionados foram realizadas com o auxílio de imagens orbitais de 2017 disponibilizadas pelo Google Satélite. O MDE disponibilizado pela Agência Espacial

Japonesa (JAXA) possui resolução espacial de 12,5 metros, enquanto o MDE disponibilizado pela SRTM apresenta resolução espacial de 30 metros. A comparação entre a declividade gerada a partir das duas bases de MDE para uma região do Polo Brasília com o auxílio da imagem orbital de alta resolução pode ser observada na Figura 4.

FIGURA 4. Comparação entre os Modelos Digitais de Elevação JAXA e SRTM, Polo Brasília.



Fonte: Elaboração própria, 2021.

A despeito do MDE da JAXA possuir uma melhor resolução espacial em comparação com o MDE da SRTM, as imagens SRTM possuem uma maior cobertura espacial, de 15.800km², enquanto as imagens JAXA possuem cobertura de 5.000km². Dessa forma, a utilização de produtos SRTM diminui a quantidade de imagens necessárias para cobrir toda a extensão territorial de um polo e, por conseguinte, diminui os custos computacionais de processamento. Além disso, como a grade NUI é composta por células de 100 por 100 metros, a resolução SRTM de 30 metros apresenta-se adequada, compatível com a escala adotada na pesquisa. Assim, optou-se pela utilização do MDE SRTM.

A partir do MDE da SRTM os processamentos para a geração da base de declividade do terreno para os seis polos da pesquisa foram realizados no SAGA GIS 2.1.2³. (Sistema de Análises Geocientíficas Automatizado) segundo o método de 9 parâmetros de polígonos de segunda ordem (ZEVENBERGEN e THORNE, 1987).

2.2.1.2 Curvatura do terreno

A identificação das curvaturas do terreno (verticais e horizontais) apresenta-se como fator importante na distribuição e concentração de água no solo. Por exemplo, terrenos com formas côncavas e convergentes são mais propensos a ocorrências de deslizamentos de terra (GUERRA, 2001). Isso se deve, primeiramente, a que a base das encostas côncavas tende a acumular sedimentos da parte superior da vertente. No momento de uma chuva, as águas concentram-se na base dessa mesma encosta, infiltrando no sedimento, que se torna instável e tende a se desmobilizar, causando um movimento de massa (GALERÀ et al., 2017). Além da propensão a deslizamentos, as áreas de relevo côncavo, devido ao caráter convergente das águas, tendem a ser mais úmidas, criando bolsões de umidade usualmente denominados de “grotas” e “vales encaixados”, por onde fluem nascentes, fluxos de enxurrada e, frequentemente, ocorrem inundações rápidas (flash floods) (VARALLO et al., 2017).

³ O SAGA GIS é um software de sistema de informação geográfica (SIG), livre e de código aberto (SAGA-GIS, 2018).

Na grade NUI optou-se por incluir informações sobre as curvaturas totais, medida que considera conjuntamente a curvatura vertical e horizontal do terreno. Os valores acima de zero apresentam caráter convexo/divergente e as áreas com valores próximos a zero apresentam vertentes planares, ao passo que as áreas com curvatura negativa apresentam relevo côncavo/convergente. As curvaturas foram geradas a partir do MDE SRTM (NASA, 2017). O procedimento para a geração das curvaturas totais foi realizado na ferramenta Terrain Analysis - Morphometry do software SAGA GIS 2.1.2.

2.2.1.3 Hidrografia

Para a geração da variável relacionada à proximidade dos cursos d'água, foi considerada uma faixa de distância, ou *buffer*, dos cursos d'água de 30 metros. Os dados de hidrografia são provenientes da Fundação Brasileira para o Desenvolvimento Sustentável (FBDS). O mapeamento da hidrografia desenvolvido pela FBDS inicia-se com o levantamento das bases cartográficas já disponíveis para os Estados, essas são então complementadas e/ou adaptadas com base nas imagens *RapidEye*, na escala de visualização de 1:10.000. A variável relacionada à distância ao curso d'água foi gerada através do cálculo da porcentagem da célula com até 30 metros de distância do curso d'água.

2.2.1.4 Unidades de conservação de proteção integral

O Sistema Nacional de Unidades de Conservação da Natureza (SNUC), implantado em 18 de julho de 2000 por meio da Lei Federal n. 9.985, estabelece critérios e normas para criação, implantação e gestão das unidades de conservação em âmbitos municipal, estadual e federal, configurando-se assim como um dos principais mecanismos de gestão adotados para a proteção e conservação dos recursos naturais. A lei definiu princípios e práticas de conservação da natureza de acordo com o tipo de uso da área, criando assim dois tipos de grupos: a) Unidades de proteção integral: têm como objetivo a preservação da natureza, sendo admitido o uso indireto dos seus recursos naturais, com exceção dos casos previstos na Lei do SNUC; b) Unidades de uso sustentável: têm como objetivo compatibilizar a conservação da natureza com o uso direto de parcela dos seus recursos naturais.

Para as análises realizadas na pesquisa NUI foi considerada a variável dos limites das unidades de conservação do tipo integral, ou seja, unidades onde os recursos só podem ser utilizados indiretamente. Os limites das Unidades de Conservação de proteção integral foram extraídos do portal do Ministério do Meio Ambiente (MMA) na escala 1:250.000, tendo sido acessados para este projeto em janeiro de 2021 (MMA, 2020). Assim, a partir da base de UC foi possível a criação da variável de porcentagem (%) da célula ocupada por unidade de conservação.

2.2.1.5 Linhas de alta tensão

As faixas de servidão de linhas de alta tensão caracterizam-se como locais com limitações de uso e ocupação do solo e contribuem para garantir a plena operação e a execução dos serviços de manutenção da linha, bem como a preservação do meio ambiente e a segurança de pessoas e bens em suas proximidades. Segundo a norma técnica NBR 5422 – Projeto de Linhas Aéreas de Transmissão de Energia Elétrica – a faixa de servidão é a faixa de terra ao longo do eixo da linha aérea de domínio da concessionária e que está sujeita a desapropriação de terceiros por meio de medida judicial. A base das faixas de servidão de linhas de alta tensão foi disponibilizada pela Agência Nacional de Energia Elétrica (ANEEL, 2020). A variável de linhas de alta tensão corresponde a porcentagem (%) da célula ocupada pela faixa de servidão das linhas de alta tensão.

2.2.1.6 Dutovias

As faixas de servidão de dutos, assim como as de linhas de alta tensão, são áreas de terreno de largura definida, ao longo da diretriz dos dutos, legalmente destinada à construção, montagem, operação, inspeção e manutenção dos dutos. A ANP - Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – órgão regulador das atividades que integram as indústrias de petróleo e gás natural e de biocombustíveis no Brasil, estabelece no Regulamento Técnico de Dutos Terrestres RTDT Nº 2/2011 os requisitos essenciais e os mínimos padrões de segurança operacional para os Dutos Terrestres (Oleodutos e Gasodutos), por ele abrangidos, visando à proteção do público em geral e da força de trabalho da companhia operadora, bem como a proteção das instalações e do meio ambiente. Segundo a norma, é considerada como invasão qualquer ocupação irregular por terceiros em faixa de duto ou em área legalmente titulada pelo proprietário do duto sujeito a desapropriação. O limite das faixas de servidão foi disponibilizado pela ANP (ANP, 2020). A variável de dutovias corresponde a porcentagem (%) da célula ocupada pela faixa de dutovias.

2.2.2 Forma Urbana

2.2.1.1 Acesso a Vias Carroçáveis

As vias carroçáveis são aquelas adequadas ao tráfego de veículos motorizados. A construção de um indicador que mostre a presença e distribuição de vias carroçáveis pode auxiliar na identificação e caracterização de núcleos urbanos informais de distintas formas. Por um lado, a cobertura de vias carroçáveis é um forte indicador da presença humana e é comum a consolidação de NUI ao longo de vias. Por outro lado, alguns NUI apresentam quantidade expressiva de domicílios com acesso deficiente à vias carroçáveis (distâncias superiores a 50 metros), resultado da ausência de um traçado viário formal e hierarquizado.

A partir de uma análise comparativa das bases de logradouros disponíveis, *StreetBase* e *OpenStreetMap*, identificou-se a base do *Openstreetmap* como a mais adequada para prosseguir com a metodologia em virtude de sua ampla cobertura, escala apropriada e constante atualização.

O primeiro passo para a construção da variável foi a identificação do atributo de tipo da via, para a seleção das vias carroçáveis da base. Os tipos de via identificados como carroçáveis⁴ foram: *motorway*, *primary*, *secondary*, *tertiary*, *service*, *residential*, *trunk*, *unclassified* e *unknown*. Em seguida, foram gerados buffers de 50 metros ao longo das vias carroçáveis, que representam as áreas com facilidade de acesso a esse tipo de via. Esses dados foram convertidos em raster⁵ para que pudessem ser integrados na grade celular para a geração da variável "porcentagem da célula com cobertura de via carroçável".

2.2.2.2 Forma das Quadras

O índice de forma das quadras (*shape index*) é uma métrica representativa da organização do traçado viário na formação das quadras (ou bolsões de ocupação, no caso de ausência de quadras). A construção do indicador se inicia com a geração de polígonos fechados delimitados pelas vias

⁴ A definição de cada tipo de via pode ser acessada na documentação do OpenStreetMap. Disponíveis em: <https://wiki.openstreetmap.org/wiki/Key:highway>.

⁵ Dados do tipo raster são compostos por linhas (horizontais) e colunas (verticais) de pixels. Cada pixel representa uma região geográfica e o valor do pixel representa uma variável.

carroçáveis, as quadras. Quanto maior a cobertura das vias carroçáveis, menores são as quadras e quanto menor a cobertura de vias carroçáveis, maiores e mais amorfos são os polígonos de quadra (ou bolsões de ocupação).

A partir dos resultados de quadras, computou-se a métrica de índice de forma (*shape index*), para vetores cuja fórmula é apresentada a seguir (MCGARIGAL e MARKS, 1994):

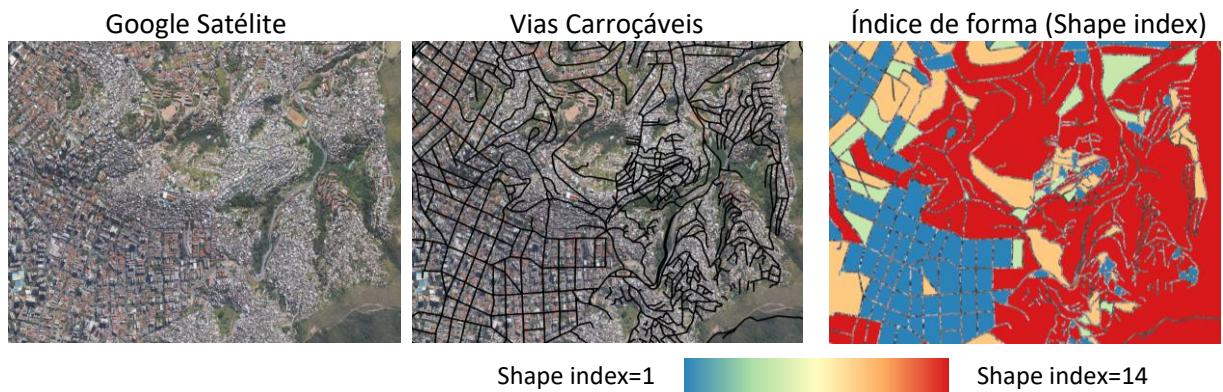
$$SHAPE = \frac{P_{ij}}{2\sqrt{\pi} \cdot A_{ij}}$$

Onde:

- Unidade: nenhum
- Variação: shape ≥ 1 , sem limite
- Descrição: shape é igual ao perímetro (P_{ij}), em metros, dividido pela raiz quadrada da área, em m^2 , ajustado por uma constante para ajustar ao padrão circular.

O índice de forma (*shape index*) resulta em 1 quando o polígono é circular e aumenta sem limite à medida que o formato do polígono se torna mais irregular. Ou seja, quanto mais regular (ortogonal) a quadra for, mais próximo de 1 é o resultado do índice. Os índices gerados para os 6 polos apresentaram grande heterogeneidade de valores, o que reflete a diversidade de formas e tamanhos das ocupações (Figura 5).

FIGURA 5. Índice de forma.



Fonte: Elaboração própria, 2021.

2.2.3 Censo Demográfico (Universo)

Os dados do Censo Demográfico de 2010 do IBGE foram considerados para a construção de variáveis populacionais, tais como renda, escolaridade e razão de dependência. Também foram considerados os dados censitários relativos às características dos domicílios e entorno que possam contribuir para a caracterização e identificação de NUI. Os dados do censo demográfico utilizados fazem parte do dados do universo agregados por setores censitários. As variáveis selecionadas do censo demográfico estão listadas no Apêndice 1.

2.2.4 Vulnerabilidade Social

Indicadores do Atlas de Desenvolvimento Humano (ADH), fruto de uma parceria entre IPEA, PNUD e FJP (PNUD; IPEA; FJP, 2016), também foram considerados na pesquisa. Esses indicadores foram

construídos a partir de dados da amostra do Censo Demográfico 2010 e da PNAD Contínua, ambos do IBGE, e disponibilizados por Unidades de Desenvolvimento Humano (UDH). As UDH são unidades territoriais análogas às áreas de ponderação do IBGE, para as quais dados da amostra do censo demográfico podem ser agregados. Entretanto, as UDH são áreas mais homogêneas, constituídas a partir da agregação de setores censitários com semelhanças socioeconômicas.

O ADH dispõe de diversas variáveis, disponibilizadas por UDH, que apoiam a construção de indicadores relacionados à vulnerabilidade social e que podem ser relevantes para a identificação e caracterização dos NUI. Os dados por UDH estão disponíveis apenas para os municípios integrantes de regiões metropolitanas, tendo sido considerados apenas para os polos de Belo Horizonte e Distrito Federal, uma vez que são os únicos polos em que a maioria dos municípios são integrantes de regiões metropolitanas⁶. As variáveis selecionadas nesta etapa estão listadas no Apêndice 1.

2.2.5 CadÚnico

O Cadastro Único para Programas Sociais (CadÚnico) tem como objetivo o cadastramento e manutenção de informações atualizadas das famílias de baixa renda (renda per capita inferior a ½ salário mínimo - SM ou renda familiar total de até 3 SMs) em todos os municípios brasileiros. O CadÚnico é o principal instrumento para a seleção e a inclusão de famílias de baixa renda em programas federais de concessão de benefícios como o Programa Bolsa Família, Tarifa Social de Energia Elétrica e o Programa Minha Casa Minha Vida (BRASIL, 2019).

O cadastro das famílias é responsabilidade dos municípios, que, além de planejar e definir as equipes de cadastradores e realizar as entrevistas, também são responsáveis por atualizar, monitorar e prezar pela veracidade e qualidade das informações preenchidas. Atualmente o CadÚnico tem mais de 26 milhões de famílias cadastradas e constitui uma importante base de dados para caracterização da parcela mais pobre da população brasileira, agregando informações de renda, características do domicílio, escolaridade, deficiência, entre outras (BRASIL, 2020b). As variáveis selecionadas desta base estão listadas no Apêndice 1.

Os dados do CadÚnico incluem o endereço das famílias cadastradas, o que potencializa o uso dessas informações para a identificação de núcleos urbanos informais. Entretanto, para que esses dados possam ser espacializados, é necessário conduzir procedimentos de geocodificação, conforme descritos no Apêndice 2. Uma vez realizada a geocodificação, dados do CadÚnico anonimizados puderam ser representados na forma de pontos (formato vetorial) e, posteriormente, agregados na grade celular.

Paralelamente à geocodificação dos endereços das famílias cadastradas no CadÚnico, foram realizados pré-processamentos das variáveis do cadastro. O pré-processamento consistiu na geração de variáveis intermediárias, necessárias para a etapa de integração das variáveis do CadÚnico à grade celular. Por exemplo, uma variável intermediária binária foi criada a partir das informações sobre domicílios improvisados. Nesse caso, a variável foi gerada para indicar se cada família vive em domicílio improvisado (1) ou não (0). Na etapa de integração das variáveis do CadÚnico à grade celular, essa

⁶ Dos 29 municípios do polo de Belo Horizonte, apenas Moeda não é integrante da Região Metropolitana de Belo Horizonte, enquanto para o polo do Distrito Federal, dos 13 municípios, apenas 3 não são integrantes da Área Metropolitana de Brasília: Cabeceiras, Cocalzinho de Goiás e Mimoso de Goiás. Os dados por UDH não foram utilizados para os polos de Recife e Porto Alegre porque eles abrangem muitos municípios não metropolitanos: 16 dos 30 municípios do polo Recife não integram a região metropolitana, enquanto 19 dos 45 municípios do polo Porto Alegre não são metropolitanos.

variável foi somada, resultando, assim, no total de domicílios improvisados por célula. O *script* da construção das variáveis intermediárias, a partir da base do CadÚnico (família como unidade de análise), poderá ser disponibilizado em repositório, cujas condições para acesso ainda serão definidas.

2.3 Integração de Variáveis

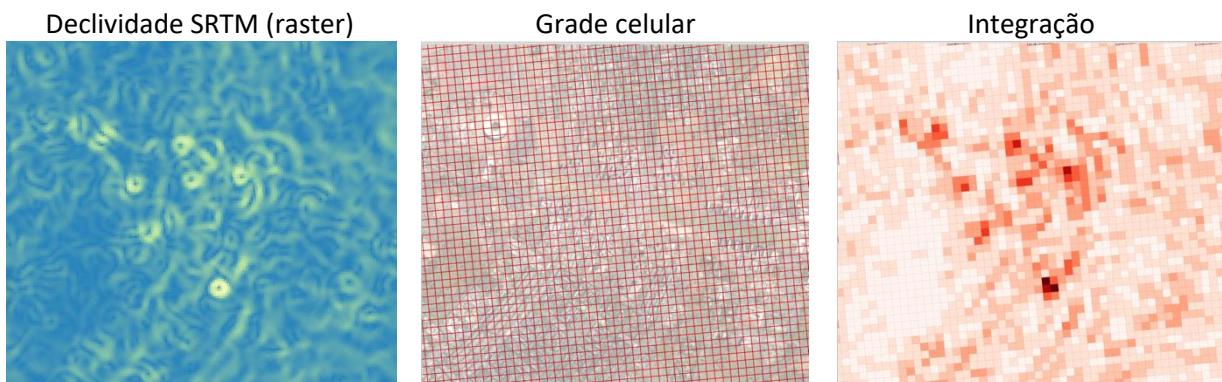
O processo de integração das variáveis com a grade celular NUI foi realizado a partir de duas estratégias distintas. A primeira delas, do tipo “raster-vetor”, baseou-se na transformação individual das variáveis em dados no formato raster para a subsequente integração desses rasters ao vetor da grade celular NUI. A segunda estratégia de integração utilizada, do tipo “vetor-vetor”, se apoia no cruzamento de informações de dados vetoriais distintos, sendo um deles o da variável de interesse e o outro a grade celular.

O processo de integração no primeiro caso, “raster-vetor”, se dá pelo uso da técnica conhecida como Estatísticas Zonais. Nesse caso, a cada célula da grade celular é atribuída a média da variável de interesse, ponderada pela área de intersecção entre as duas bases. Nessa primeira estratégia de integração é possível realizar a integração de forma que exige menor processamento computacional e, consequentemente, de maneira mais rápida quando comparada a outras técnicas de integração, como por exemplo a do tipo “vetor-vetor”. Entretanto, essa premissa é válida apenas para integrações na qual cada variável possui limites espaciais únicos, como é o caso das variáveis relativas a características físico-ambientais e unidades de conservação (Seção 2.2.1) e das variáveis de forma urbana (Seção 2.2.2), que envolvem unidades espaciais muito distintas.

No caso das variáveis provenientes do Censo Demográfico (Seção 2.2.3), Atlas do Desenvolvimento Humano (2.2.4) e CadÚnico (2.2.5), que são disponibilizadas em uma mesma unidade de análise (setor censitário, UDH ou pontos de localização das famílias), o processamento de integração a partir da técnica “vetor-vetor” tende a ser mais eficiente visto que, embora mais lento, pode ser realizado de uma única vez.

As variáveis referentes a características físico-ambientais, unidades de conservação e forma urbana compõem um total de 8 (oito) variáveis. Dado o baixo número de variáveis, foi possível realizar o processo de integração no software QGIS 3.4 (QGIS, 2020) a partir da ferramenta *Zonal Statistics*. A Figura 6 ilustra o resultado da integração da declividade SRTM à grade celular para uma área do Polo Brasília.

FIGURA 6. Processo metodológico para integração da grade celular



Fonte: Elaboração própria, 2021.

Dado o grande número de variáveis provenientes do Censo Demográfico e do Atlas do Desenvolvimento Humano (39 e 65, respectivamente), optou-se pela realização do procedimento de integração das variáveis no software R, que comporta processamentos mais robustos. O código R para a integração dessas variáveis na grade celular poderá ser disponibilizado em repositório, a definir.

Para a integração das variáveis do CadÚnico, optou-se por identificar a qual célula da grade celular NUI cada observação pertence. Esse processo foi realizado no software QGIS 3.4 a partir da ferramenta "Atributos pela Localização", no qual cada ponto de localização da família cadastrada recebeu o atributo de identificação (ID) da célula na qual estava localizado. A partir dessa informação, foi possível somar variáveis binárias previamente construídas (por exemplo, densidade excessiva) para gerar novas variáveis para as células, tais como "número de domicílios com densidade excessiva". Para gerar variáveis como "renda média per capita da família", os dados do CadÚnico foram agregados para células por intermédio de uma média simples.

Os Quadros 3 e 4 ilustram, de forma esquemática, o processo de integração do código da célula a cada família (Quadro 3) e o resultado da soma ou média simples condicionada pelo ID da célula (Quadro 4).

QUADRO 3. Integração do código da célula em que cada domicílio familiar cadastrado no CadÚnico está localizado.

Código familiar CadÚnico	Variáveis CadÚnico		Código da célula
	Binária	Não binária	
A	0	5	C001
B	1	10	C002
C	1	20	C002
D	1	10	C003

Fonte: elaboração própria, 2021.

QUADRO 4. Soma ou média das variáveis CadÚnico condicionada pelo código da célula

Código da célula	Variáveis CadÚnico	
	binária	Não binária
C001	0	5
C002	2	15
C003	1	10

Fonte: elaboração própria, 2021.

Os scripts desenvolvidos para integração das variáveis provenientes do Cadúnico na grade celular NUI poderão ser disponibilizados em repositório.

Para a condução de análises cuja unidade de análise são os próprios NUI, tais como a comparação entre os NUI dos distintos polos (Capítulo 3) e a construção de tipologias de NUI (Capítulo 4), as variáveis - originalmente integradas por células - foram agregadas aos NUI pela técnica “vetor-vetor”.

3. Caracterização dos Núcleos Urbanos Informais a partir de Dados Secundários

A partir das variáveis apresentadas no Capítulo 2, este capítulo complementa a caracterização dos núcleos urbanos informais apresentada no Produto 11. Essa análise, de caráter exploratório, tem como principal objetivo comparar as características dos NUI dos seis polos da pesquisa, analisando as dimensões da precariedade predominantes em cada região. Também são analisadas as diferenças entre áreas localizadas dentro e fora dos NUI. Para tanto, foram comparadas as características das células, avaliando o quanto divergentes são as condições das células classificadas como NUI em relação às demais.

3.1 Características Gerais dos NUI

O trabalho de campo conduzido no âmbito da pesquisa identificou um total de 4.968 NUI nos seis polos da pesquisa, que ocupam uma área total de 78.948 ha. Estima-se a presença de 1.486.725 domicílios nesses NUI, resultando em uma densidade média de 18,8 domicílios/ha. A Tabela 1 sumariza essas estatísticas para cada um dos polos da pesquisa.

TABELA 1. Núcleos Urbanos Informais (NUI) nos polos da pesquisa: Número, total de domicílios, área total, área média e densidade média de domicílios.

Polo	Núcleos Urbanos Informais				
	Número	Total de domicílios	Área total (ha)	Área média (ha)	Densidade média (domicílios/ha)
Belo Horizonte	1.413 (28,4%)	433.537 (29,2%)	27.460 (34,8%)	19,4	15,8
Brasília	420 (8,5%)	154.689 (10,4%)	14.393 (18,2%)	34,3	10,7
Juazeiro do Norte	219 (4,4%)	40.223 (2,7%)	5.602 (7,1%)	25,6	7,2
Marabá	212 (4,3%)	71.119 (4,8%)	7.252 (9,2%)	34,2	9,8
Porto Alegre	1.479 (29,8%)	173.413 (11,7%)	6.725 (8,5%)	4,5	25,8
Recife	1.225 (24,7%)	613.744 (41,3%)	17.514 (22,2%)	14,3	35,1
Total	4.968 (100%)	1.486.725 (100%)	78.948 (100%)	15,9	18,8

Fonte: Elaboração própria, 2021.

Belo Horizonte é o polo que apresenta a maior extensão de áreas classificadas como NUI (27.460 ha, 34,8% do total de área de NUI nos 6 polos). Com 1.413 NUI, é o segundo polo com maior número de domicílios em NUI (433.537), inferior apenas ao apresentado por Recife. A densidade média de

domicílios por NUI é de 15,8 domicílios/ha, inferior à apresentada por Recife e Porto Alegre e superior à de Brasília, Juazeiro do Norte e Marabá.

O Polo **Brasília** apresenta 420 NUI com um total estimado de 154.689 domicílios. Destaca-se por apresentar a maior média de área por NUI (média de 34,3 ha), o que se reflete em uma densidade média de domicílios nos NUI relativamente baixa (10,7 dom/ha), superior apenas às apresentadas pelos polos Juazeiro do Norte e Marabá.

Os polos **Juazeiro do Norte e Marabá** apresentam o menor número de NUI (219 e 212 NUI, respectivamente) e de domicílios em NUI (40.223 e 71.119 domicílios, respectivamente). Apresentam também as menores densidades médias de domicílios em NUI: 5,5 domicílios/ha no Polo Juazeiro do Norte e 9,8 domicílios/ha no Polo Marabá. Os NUI dos polos Juazeiro do Norte e Marabá apresentam, em média, dimensões de 25,6 ha e 34,2 ha, respectivamente, inferiores apenas às dos NUI do Polo Brasília.

Porto Alegre é o polo com maior número de NUI (1.479), embora apresentem, em média, menor dimensão (4,5 ha) e menor número de domicílios (117,4 domicílios por NUI). Em média, os NUI do Polo Porto Alegre apresentam alta densidade de domicílios (25,8 domicílios/ha), inferior apenas à apresentada pelo Polo Recife.

O Polo **Recife** apresenta 1.225 NUI e destaca-se por apresentar o maior número de domicílios em NUI (613.744, 41,3% do total de domicílios em NUI nos 6 polos). Apresenta a maior média de domicílios por NUI (501,2 domicílios por NUI), embora os NUI sejam, em média, relativamente pequenos (14,3 ha, média superior apenas à do Polo Porto Alegre). Essas características se refletem na densidade média dos NUI, que é a mais elevada entre os seis polos (35 domicílios/ha).

3.2 Características da População

Esta seção apresenta um panorama de um conjunto de características populacionais nos NUI de cada polo, comparando ainda a diferença dessas características nas células situadas dentro e fora dos NUI. Para a caracterização, foram consideradas as seguintes variáveis: renda média dos responsáveis pelos domicílios (*RenRespMedia*), alfabetização dos responsáveis por domicílios (*NRespAlf*), porcentagem da população dependente (*RenPopDependente*), idade média do responsável (*NRespldade*) e mulheres responsáveis pelos domicílios (*NRespFem*).

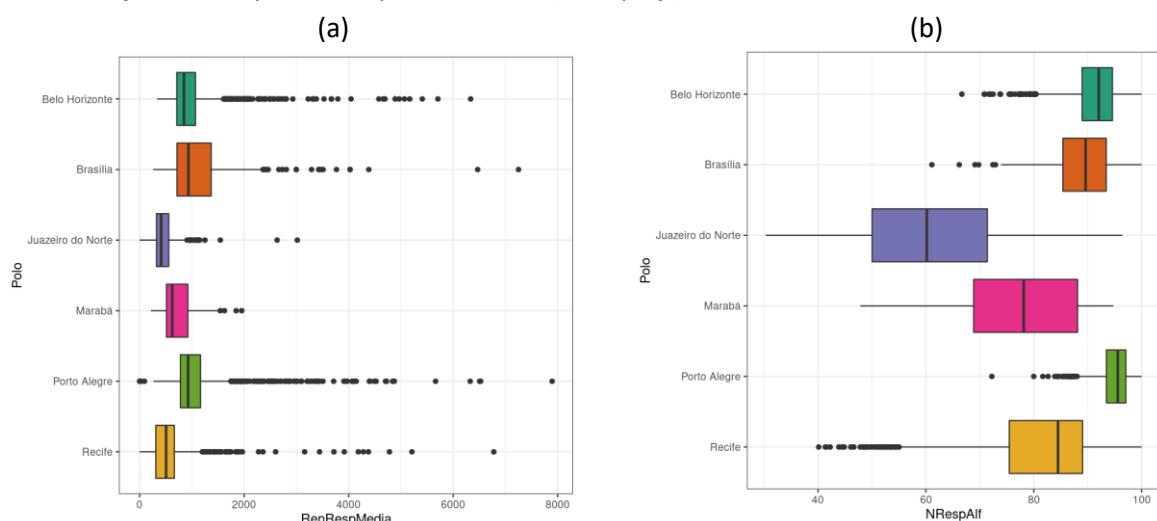
Considerando a **renda média dos responsáveis pelos domicílios** nos NUI, os polos de Juazeiro do Norte, Recife e Marabá apresentaram as menores médias, R\$464, R\$488 e R\$767, respectivamente. A maior média salarial é constatada no Polo Brasília (R\$1015), seguido pelos polos Porto Alegre (R\$919) e Belo Horizonte (R\$883). A variação da renda nos NUI dos seis polos pode ser observada por intermédio dos gráficos do tipo box-plot apresentados na Figura 7a. O box-plot permite identificar os valores entre o 1º e o 3º quartis (caixa que abrange 50% das observações), a mediana (linha que divide a caixa) e os valores extremos (pontos). Juazeiro é o polo que apresenta a menor dispersão dos dados de renda, enquanto o Polo Brasília apresenta a maior dispersão.

Comparando a renda média nas células dentro e fora dos NUI, Brasília destaca-se como o polo onde a diferença de renda é maior, sendo 2,4 vezes mais elevada fora dos NUI. O polo com menor diferença salarial é Marabá, no qual a renda média é apenas 1,18 vezes mais elevada fora dos NUI.

Em relação à **alfabetização dos responsáveis por domicílios** nos NUI, Juazeiro do Norte destaca-se como o polo com os piores níveis, com 60% dos responsáveis alfabetizados, seguido pelos polos de Marabá e Recife, com 82% e 84%, respectivamente. O Polo Porto Alegre apresenta os melhores níveis de alfabetização, com 94% dos responsáveis alfabetizados, seguido dos polos de Belo Horizonte e Brasília, ambos com 90%. Os polos com piores níveis de alfabetização representam também os com maior dispersão dos dados, conforme é possível observar nos box-plots da Figura 7b.

A alfabetização dos responsáveis nas células dentro e fora de NUI varia entre 2 e 4 pontos percentuais, sendo superior fora dos NUI. O Polo Marabá, no entanto, representa uma exceção, dado que não foi constatada diferença entre o grau de alfabetização dentro e fora dos NUI.

FIGURA 7. Box-plot: (a) renda média dos responsáveis pelos domicílios (*RenRespMedia*) e (b) alfabetização dos responsáveis por domicílios (*NRespAlf*).



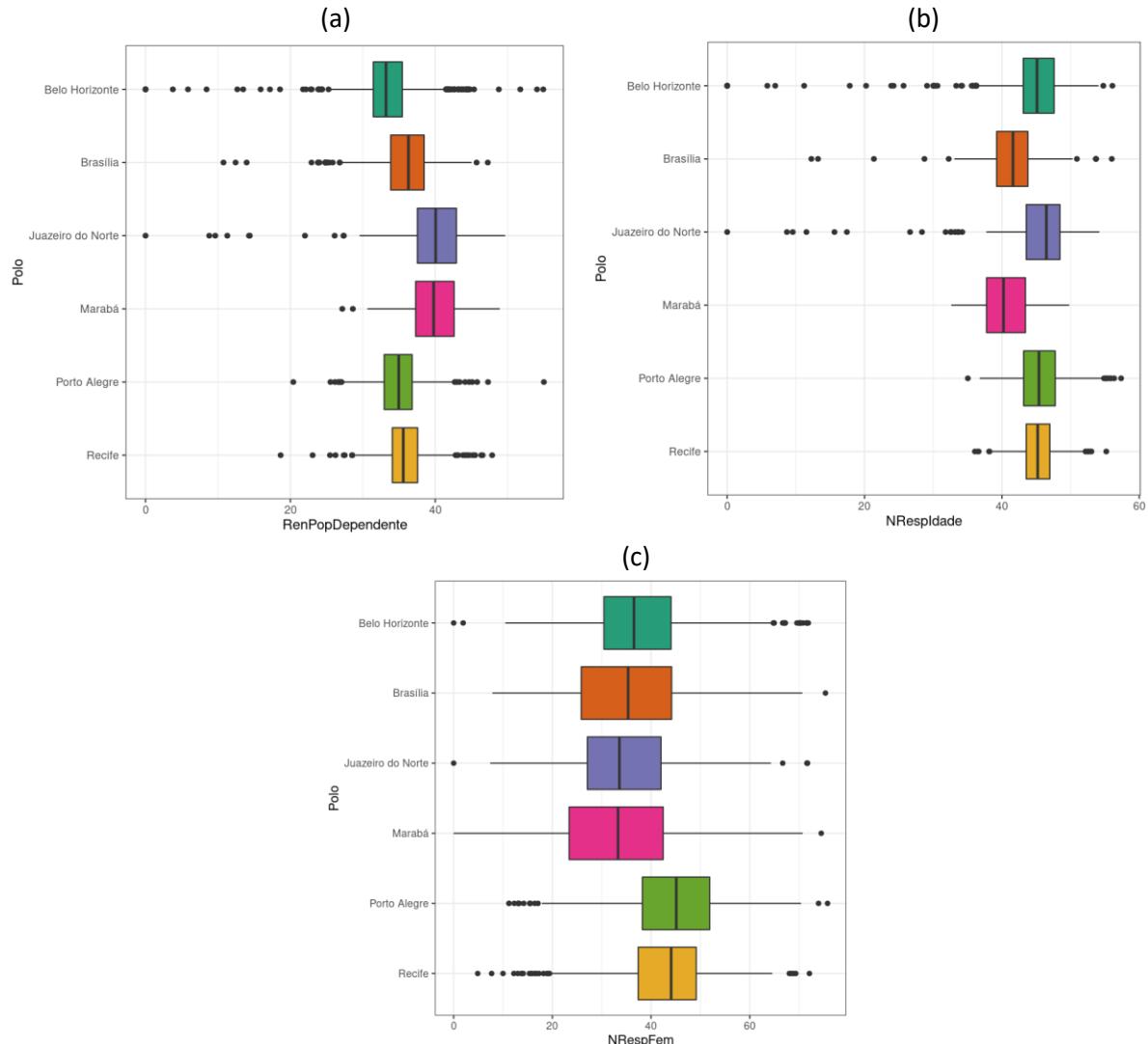
Fonte: Elaboração própria, 2021.

Quanto à **porcentagem da população dependente** (menores de 14 anos e maiores de 65 anos), os polos Juazeiro do Norte e Marabá são os que apresentam a maior presença de população dependente nos NUI (Figura 8a), com 39,5% e 38,5% de sua população, respectivamente, classificada como potencialmente inativa. Belo Horizonte é o polo que apresenta a menor porcentagem de população dependente (33,5%), seguido pelos polos Brasília (34,6%), Recife (35%) e Porto Alegre (35,4%). Comparando as células identificadas como NUI com as demais, observa-se que a proporção da população dependente é mais elevada nos NUI, com exceção do Polo Recife, onde a proporção é praticamente a mesma. Brasília foi o polo que apresentou a maior diferença entre as áreas, com cerca de 3 pontos percentuais a menos nas áreas não classificadas como NUI.

Em relação à **idade média do responsável** dos domicílios em NUI, Marabá é o polo que apresenta a menor média (39,4), seguido pelo Polo Brasília (40). Os demais polos apresentam distribuições semelhantes, com médias entre 44,3 e 45,5, sendo Juazeiro do Norte o polo que apresenta a maior média (Figura 8b). Considerando a chefia feminina, Porto Alegre e Recife se destacam como os polos com a maior proporção de **mujeres responsáveis pelos domicílios** em NUI, 46,5% e 45%, respectivamente. Os polos com menor proporção de mulheres responsáveis são Juazeiro do Norte e Marabá (35% e 36,7%), seguidos de Belo Horizonte e Brasília (40,3% e 41,2%). Com exceção de Marabá, em todos os polos a proporção de mulheres responsáveis é maior nas áreas identificadas

como NUI. Recife, em particular, destaca-se como o polo onde essa diferença é maior, chegando a 6 pontos percentuais.

FIGURA 8. Box-plot: (a) população dependente (*RenPopDependente*), (b) idade média dos responsáveis por domicílios (*NRespldade*) e (c) mulheres responsáveis por domicílios (*NRespFem*).



Fonte: Elaboração própria, 2021.

3.3 Características Físico-Ambientais e Unidades de Conservação

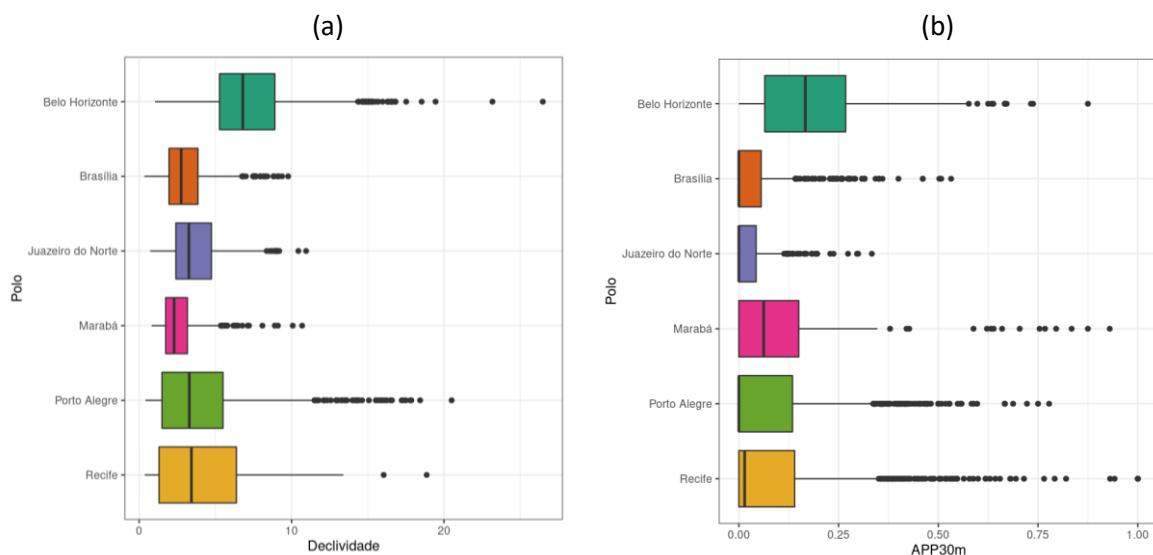
Para a caracterização de aspectos físico-ambientais dos NUI dos seis polos, considerou-se variáveis cuja construção foi apresentada na Seção 2.2.1 deste relatório, a saber: declividade (*Declividade*), presença de APPs de curso d'água (*APP30m*), presença de unidade de conservação ambiental de proteção integral (*UCIntegral*), faixa de servidão de linhas de alta tensão (*AltaTensao*) e dutovias (*Dutovias*).

O Polo Belo Horizonte é o que apresenta a maior média de **declividade** nos NUI (8%), com 25% de seus NUI apresentando declividade média superior a 9% (3º quartil, Figura 9). Os polos de Recife e Porto Alegre apresentam uma distribuição positivamente assimétrica dessa variável (Figura 9), com a maioria dos NUI caracterizados por declividades muito baixas, porém uma quantidade considerável

de NUI com declividades mais elevadas. Nos polos Marabá e Brasília predominam NUI de baixa declividade: mais de 75% dos NUI nesses polos apresentam declividades médias inferiores a 3,9%. Cabe ressaltar, entretanto, que essa variável retrata a declividade média dos NUI, o que representa uma limitação, principalmente em polos cujos NUI são maiores, como é o caso de Marabá, Brasília e Juazeiro do Norte.

Belo Horizonte também é o polo cujos NUI apresentam maior **presença de APPs de curso d'água**, no qual 15% das áreas ocupadas por NUI apresentam esse tipo de restrição. Belo Horizonte, Marabá e Recife são os polos que apresentam mediana superior a zero: 50% dos NUI desses polos apresentam, respectivamente, pelo menos 17%, 6% e 1% de seu território coberto por áreas com até 30 metros de distância de cursos d'água (Figura 9a). Apesar dos baixos valores de mediana, todos os polos apresentam NUI com APPs de curso d'água, sendo Juazeiro do Norte o polo que apresenta a menor incidência desse tipo de gravame ambiental.

FIGURA 9. Box-plot: (a) declividade (Declividade) e (b) proporção do NUI com APP de curso d'água (APP30m).



Fonte: Elaboração própria, 2021.

Outras restrições à ocupação observadas incluem a presença de NUI em Unidades de Conservação Ambiental de Proteção Integral (UC)⁷, faixa de servidão de linhas de alta tensão e dutovias. O Polo Juazeiro do Norte é o único polo que não apresenta NUI em Unidades de Conservação Ambiental de Proteção Integral e o Polo Marabá possui apenas 1 NUI com 6% de sua área em UC. O Polo Brasília apresenta 7 NUI com áreas que se sobrepõem a UCs, enquanto os polos Recife, Porto Alegre e Belo Horizonte apresentam, respectivamente, 38, 33 e 24 NUI com sobreposições dessa natureza.

Os polos Juazeiro do Norte e Marabá não apresentam NUI em faixas de serviço de dutovias, ao passo que os polos Belo Horizonte, Porto Alegre, Recife e Brasília apresentam, respectivamente, 19, 13, 10 e 6 NUI com alguma cobertura de dutovias. Todos os polos possuem NUI em faixas de servidão de linha de alta tensão, sendo Recife o que apresenta maior número de NUI com áreas nessa condição

⁷ Os dados aqui apresentados referem-se apenas às Unidades de Conservação de Proteção Integral disponibilizadas pelo Ministério do Meio Ambiente, disponível em: <https://antigo.mma.gov.br/areas- protegidas /cadastro-nacional-de-ucs.html>.

(270), seguido pelos polos Belo Horizonte (240), Porto Alegre (173), Brasília (30), Juazeiro do Norte (17) e Marabá (15).

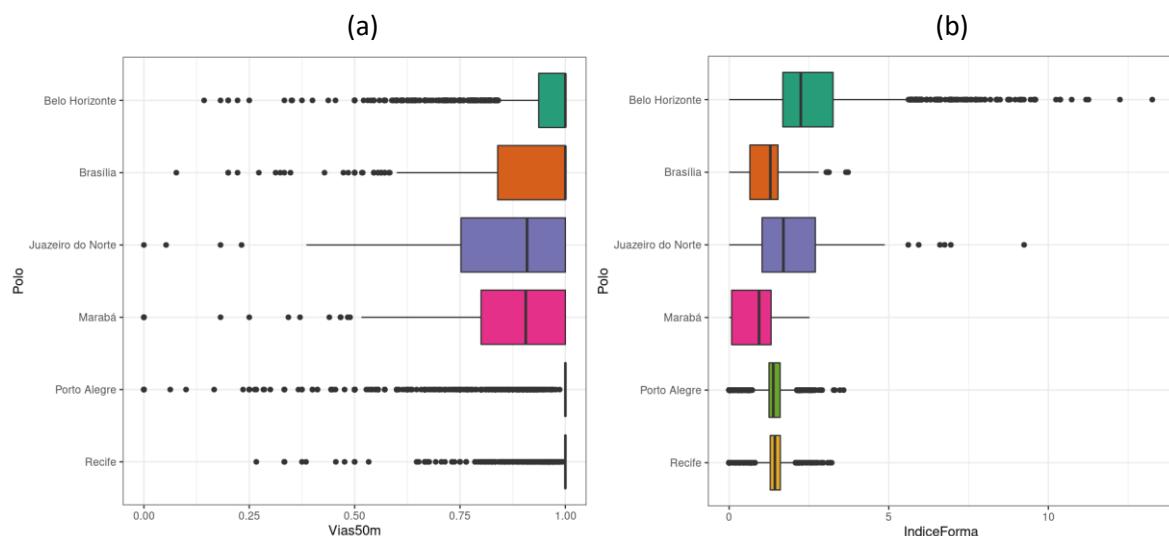
3.4 Cobertura de Vias Carroçáveis e Forma das Quadras

A análise de aspectos relacionados à forma urbana dos NUI considerou as variáveis cobertura de vias carroçáveis (*Vias50m*) e forma das quadras (*IndiceForma*), descritas na Seção 2.2.2 deste relatório.

O Polo Recife apresenta a melhor cobertura de vias carroçáveis em seus NUI (99%), o que é condizente com o fato da maioria dos núcleos estarem inseridos na malha urbana. Já Juazeiro do Norte e Marabá, com médias de 86% e 93%, respectivamente, são os únicos polos que apresentam a mediana da cobertura de vias carroçáveis inferior a 1 (Figura 10a). Esse resultado coincide com o fato de apresentarem elevada proporção de núcleos isolados.

Em relação à forma das quadras ou bolsões de ocupação, os polos de Marabá e Brasília apresentam maior regularidade, com as menores médias do índice "shape" (Seção 2.2.2). Nos polos de Porto Alegre e Recife também predominam NUI com relativa regularidade na forma das quadras. Já o Polo Belo Horizonte destaca-se por apresentar uma maior quantidade de NUI caracterizados por quadras ou bolsões de ocupação irregulares, segundo pelo Polo Juazeiro do Norte, conforme fica evidente nos diagramas de caixa apresentados na Figura 10b.

FIGURA 10. Box-plot:(a) cobertura de vias carroçáveis (*Vias50m*) e (b) forma das quadras (*IndiceForma*).



Fonte: Elaboração própria, 2021.

3.5 Infraestrutura e Entorno dos Domicílios

Para análise da infraestrutura e entorno dos domicílios foram consideradas variáveis do Censo Demográfico 2010 integradas por NUI (para a comparação entre polos) e por células (para a comparação entre áreas dentro e fora de NUI), conforme descrito na Seção 2.2.3.

Em relação ao **abastecimento de água**, Marabá destaca-se como o polo que apresenta as piores condições, no qual 62% dos domicílios em NUI não possuem abastecimento por rede geral (*DomSAqua*), predominando o acesso por meio de nascentes ou poço na propriedade (52,9%). Embora a carência de acesso à rede geral de água seja generalizada no Polo Marabá, esse percentual é menor

nas áreas fora de NUI, onde 48% dos domicílios não dispõem de acesso à rede geral de abastecimento de água.

O polo com melhor atendimento de rede geral de água é Belo Horizonte, no qual 4,8% dos domicílios em NUI não possuem acesso à rede geral, seguido pelo Polo Porto Alegre (8,9%), Polo Recife (13%), Polo Brasília (15,8%) e Polo Juazeiro do Norte (25,7%). A Figura 11a apresenta o box-plot da variável e revela a grande variabilidade de situações de abastecimento de água nos NUI dos polos Brasília, Juazeiro do Norte e Marabá.

Quanto ao esgotamento sanitário, os NUI dos polos Juazeiro do Norte e Marabá apresentam o maior percentual de domicílios **sem ligação à rede de esgoto ou fossa séptica** (78% e 75%, respectivamente), seguidos do Polo Recife (50%) e do Polo Brasília (48%). Nos três primeiros polos, a carência de esgotamento adequado nos domicílios é generalizada, e a diferença entre as células dentro e fora dos NUI é de, no máximo, 4 pontos percentuais. Já no Polo Brasília, a diferença entre as áreas dentro e fora de NUI é expressiva, de aproximadamente 23 pontos percentuais (25% dos domicílios fora de NUI apresentam condições inadequadas). O Polo Brasília também se destaca, juntamente com o Polo Recife, por uma grande diversidade de condições de esgotamento entre os NUI, que pode ser observada pelo box-plot da variável (Figura 11b).

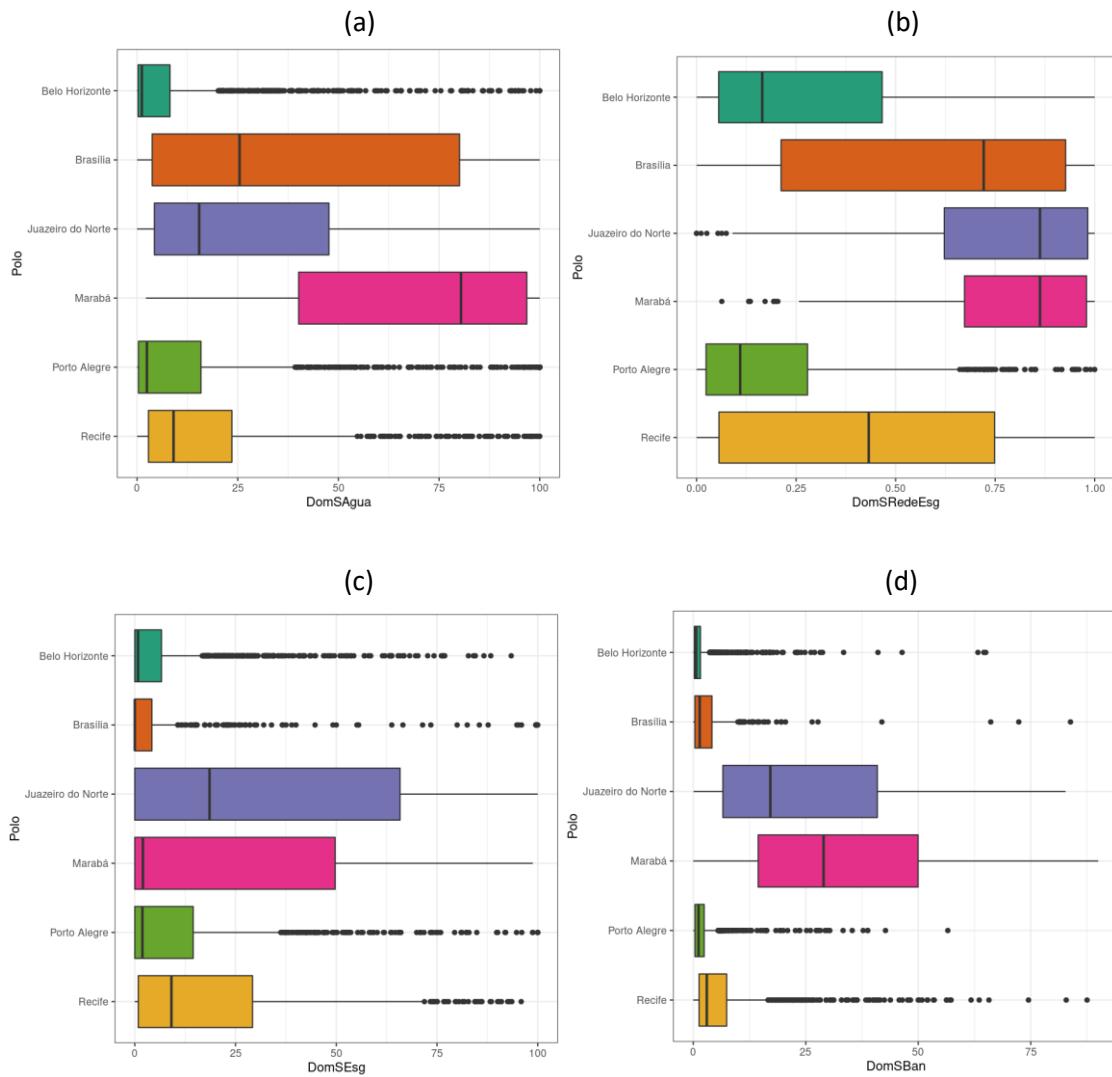
Os NUI dos polos Porto Alegre e Belo Horizonte apresentam menores proporções de domicílios com esgotamento inadequado, 18,3% e 24,5%, respectivamente. Já a diferença entre as condições de esgotamento dentro e fora dos NUI é mais díspar no Polo Belo Horizonte (cerca de 9 pontos percentuais) do que no Polo Porto Alegre (3 pontos percentuais).

Entre as alternativas inadequadas de esgotamento, a **fossa rudimentar** predomina em todos os polos, exceto Porto Alegre. É adotada em 60% dos domicílios nos NUI do Polo Marabá, 50% no Polo Juazeiro do Norte, 46,5% no Polo Brasília, 38% no Polo Recife, 15,5% no Polo Belo Horizonte e 7% no Polo Porto Alegre. Porto Alegre é o polo que apresenta o maior porcentual de NUI com esgotamento sanitário por vala (8,5%) e também o único no qual essa alternativa predomina entre as formas inadequadas de esgotamento.

As condições de esgotamento produzem reflexos no entorno dos NUI, em particular na presença de **esgoto a céu aberto**. Os NUI dos polos Marabá e Juazeiro do Norte apresentam a maior proporção de domicílios cujo entorno possui esgoto a céu aberto, 38,5% e 37%, respectivamente. Considerando essa variável, ambos os polos destacam-se por apresentar uma grande variedade de condições, conforme evidencia o box-plot da Figura 11c. O polo com menor proporção de domicílios com esgoto a céu aberto é Belo Horizonte (8,8%), seguido do Polo Porto Alegre (11,5%), Polo Brasília (11,8%) e Polo Recife (20%).

Outra variável que expressa o grau de precariedade dos NUI é a **ausência de banheiro de uso exclusivo dos moradores**. Novamente, os NUI dos polos Marabá e Juazeiro do Norte apresentam a maior proporção de domicílios nessa condição, 25% e 22%, respectivamente. Essa situação reflete a condição dos polos como um todo, dado que a diferença entre os domicílios dentro e fora dos NUI é de apenas 6% no Polo Marabá e 2% no Polo Juazeiro do Norte. Nos demais polos, a proporção de domicílios em NUI sem banheiro de uso exclusivo é baixa: 2,5% nos polos Belo Horizonte e Porto Alegre, 3,3% no Polo Brasília e 4,5% no Polo Recife (Figura 11d).

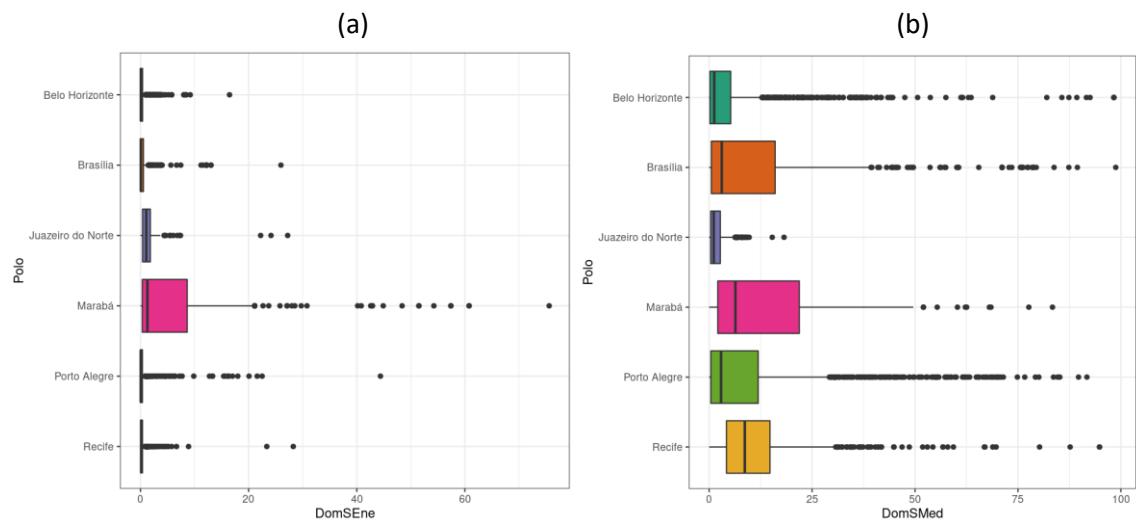
FIGURA 11. Box-plot: (a) abastecimento de água por rede geral (*DomSAqua*), (b) ausência de ligação à rede de esgoto ou fossa séptica (*DomSRedeEsg*), (c) esgoto a céu aberto (*DomSEsg*) e (d) ausência de banheiro de uso exclusivo dos moradores.



Fonte: Elaboração própria, 2021.

O acesso à energia elétrica é quase universal nos polos analisados. Nos NUI dos polos Recife, Belo Horizonte, Porto Alegre e Brasília, menos de 0,5% dos domicílios não têm acesso à energia elétrica. Já nos NUI dos polos Juazeiro do Norte e Marabá, esse percentual é de 1,3% e 2,7%, respectivamente, sendo Marabá o polo no qual se observa uma maior diversidade de condições dos NUI quanto ao acesso à energia elétrica (Figura 12a). O polo de Marabá também destaca-se quanto à ausência de medidor de energia elétrica de uso exclusivo, dado que quase 20% dos domicílios em NUI encontram-se nessa situação, e pela diversidade de condições dos NUI nesse quesito (Figura 12b). Em relação aos demais polos, esse percentual é de 14,8% no Polo Porto Alegre, 10,2% no Polo Brasília, 9,9% no Polo Recife, 6,9% no Polo Belo Horizonte e 1,7% no Polo Juazeiro do Norte.

FIGURA 12. Box-plot: (a) acesso à energia elétrica (*DomSEne*) e (b) ausência de medidor de uso exclusivo (*DomSRedeEsg*).



Fonte: Elaboração própria, 2021.

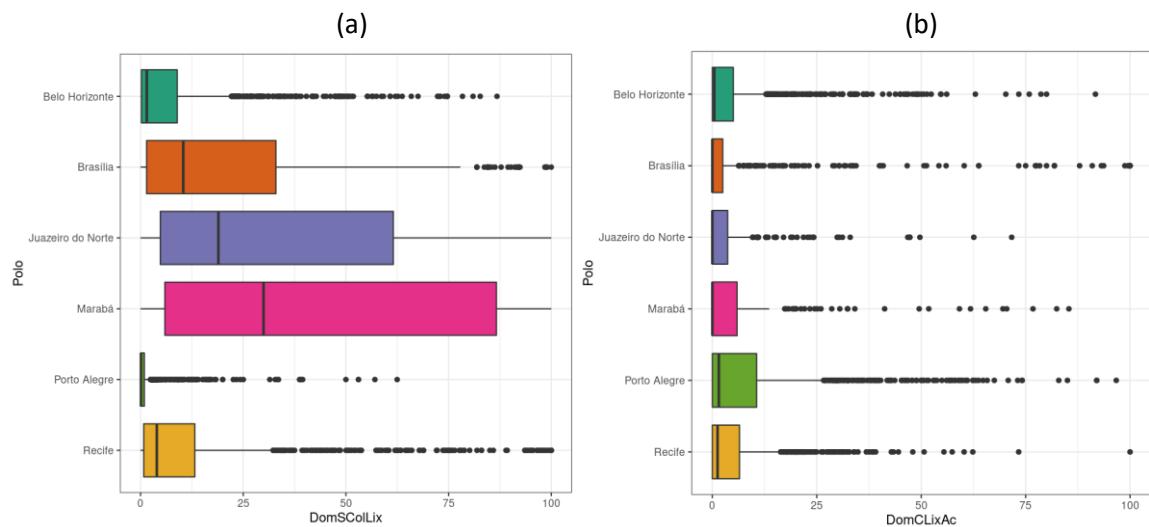
Os NUI dos polos Marabá e Juazeiro do Norte apresentam as piores condições de **coleta de lixo**: 30,7% e 25,9% de seus domicílios, respectivamente, não têm coleta de lixo por serviço de limpeza. São também os polos que apresentam maior diversidade de condições entre os NUI, o que pode ser constatado pela análise do box-plot (Figura 13a). A situação, embora bastante agravada no interior dos NUI, reflete a situação dos polos como um todo. No Polo Juazeiro do Norte, as medianas das células dentro e fora dos NUI são de 53% e 31%, respectivamente, ao passo que no Polo Marabá são de 37% dentro e 29% fora dos NUI.

Porto Alegre é o polo cujos NUI apresentam melhores condições, com apenas 1% dos domicílios sem coleta de lixo, seguido pelos polos Belo Horizonte (4,8%), Recife (7%) e Brasília (10%). Entre esses, o Polo Brasília é o que apresenta maior diversidade de condições dos NUI, assim como maior diferença entre as células dentro e fora de NUI (medianas de 5,6% e 2,3%, respectivamente).

Para boa parte dos domicílios sem coleta de lixo a solução adotada é queimá-lo na propriedade. Nos polos Marabá e Juazeiro do Norte cerca de 20% dos domicílios em NUI têm o lixo queimado na propriedade. No Polo Marabá, nota-se que a adoção dessa solução para o destino do lixo não diverge dentro e fora dos NUI (mediana das células é de 27% e 26%, respectivamente), já no Polo Juazeiro do Norte a mediana das células em NUI é de 24,8% e a das demais células é de 15,5%.

A ausência de coleta de lixo por serviço de limpeza não apresenta relação direta com o acúmulo de lixo nos logradouros, conforme é possível observar na Figura 13. Os NUI do Polo Porto Alegre, polo que apresenta os melhores indicadores de coleta de lixo por serviço de limpeza, é o que apresenta maior percentual de domicílios cujo entorno possui lixo acumulado nos logradouros (10,7%). Os demais polos apresentam percentual de domicílios em NUI nessa condição entre 4,6% (Polo Recife) e 6,8% (Polo Marabá).

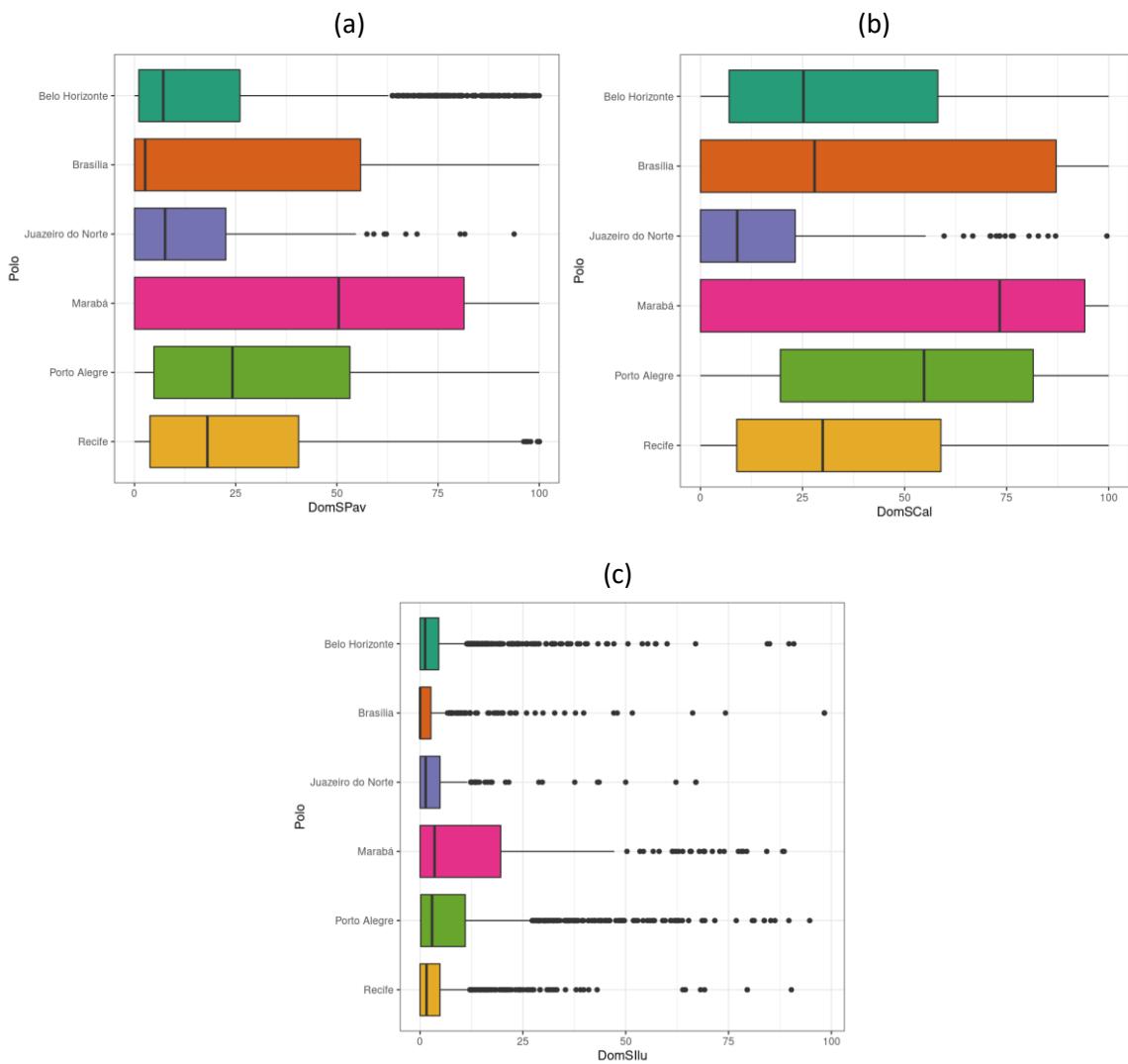
FIGURA 13. Box-plot: (a) ausência de coleta de lixo por serviço de limpeza (*DomSCollLix*) e (b) entorno com lixo acumulado nos logradouros (*DomCLixAc*).



Fonte: Elaboração própria, 2021.

Em relação à **ausência de pavimentação**, Marabá destaca-se como o polo com maior proporção de domicílios em NUI sem pavimentação (59%), sem calçada (70%) e sem iluminação pública (16,6%). É também o polo cujos NUI apresentam maior diversidade de condições desses três aspectos, conforme pode ser observado nos box-plot da Figura 14. O Polo Porto Alegre também apresenta elevada proporção de domicílios em NUI sem pavimentação (34%), sem calçada (55%) e sem iluminação pública (10%), seguido pelos polos Brasília (25%, 53% e 3,5%), Recife (22,8%, 34,7% e 3,5%) e Belo Horizonte (17,4%, 33,9% e 4%). Juazeiro do Norte é o polo que apresenta melhores níveis nos três indicadores, com 13,4% dos domicílios em NUI sem pavimentação, 16,5% sem calçada e 3% sem iluminação pública.

FIGURA 14. Box-plot - Porcentagem de domicílios em NUI: (a) sem pavimentação (*DomSPav*), (b) sem calçada (*DomSCal*) e (c) sem iluminação pública.



Fonte: Elaboração própria, 2021.

4. Tipologias de NUI

O presente capítulo busca analisar as características dos NUI e suas diferenças intra-polo a partir da construção de subgrupos significativos de NUI, os quais trataremos como tipologias. A definição de tipologias de NUI para cada polo foi realizada por meio de análise de agrupamentos, cujo objetivo principal é classificar uma amostra de entidades (no caso, NUI) em um número menor de grupos mutuamente excludentes, com base nas similaridades entre as entidades, considerando um conjunto de variáveis pré-estabelecidas (HAIR et al., 2009). Os grupos estabelecidos devem exibir elevada homogeneidade interna (dentro dos grupos) e elevada heterogeneidade externa (entre grupos).

Ao contrário das demais técnicas de análise multivariada, o conjunto de variáveis que representam as características usadas para comparar entidades em análise de agrupamentos não é estimada empiricamente na análise de agrupamentos, mas sim especificada pelo pesquisador, o que representa um passo crítico da análise (HAIR et al., 2009). O processo de seleção das variáveis a serem utilizadas na construção de grupos de NUI considerou, inicialmente, sua relevância na representação e caracterização de aspectos da precariedade habitacional. Para tanto, foram avaliadas as variáveis apresentadas e analisadas nos Capítulos 2 e 3. Observou-se, inicialmente, a relevância dessas variáveis na discriminação de distintos NUI, especialmente quando observadas de forma agregada (um valor para todo o NUI).

Outra questão considerada diz respeito à multicolinearidade entre as variáveis, verificada para cada polo por meio da análise de matrizes de correlação das variáveis. A multicolinearidade atua como uma forma de ponderação implícita na análise de agrupamentos, afetando os resultados da análise. Em caso de multicolinearidade, deve-se reduzir as variáveis a números iguais em cada conjunto de medidas correlacionadas (HAIR et al., 2009). As variáveis selecionadas para a condução da análise estão listadas no Quadro 5. As variáveis listadas foram utilizadas na construção de tipologias em todos os polos, com as exceções: (a) "Declividade", cuja versão agregada por NUI (média da declividade do NUI) manteve-se relevante apenas no polo de Recife; e (b) "AguaNascente" foi utilizada apenas na análise do polo de Juazeiro do Norte, substituindo a variável "DomSAqua", que apresentou problemas de multicolinearidade. Todas as variáveis foram padronizadas para a condução da análise de agrupamentos.

Para a determinação dos grupos, adotou-se o **algoritmo k-média**, uma abordagem não-hierárquica que utiliza uma técnica iterativa para partitionar um conjunto de dados (JOHNSON; WICHERN, 1995). Nesse processo, os elementos de um conjunto de dados são agrupados em k grupos, sendo k a quantidade de grupos estabelecida previamente. São atribuídas medidas de proximidade (similaridade) a todos os dados dos k grupos e entre os k grupos. Em um processo iterativo, o algoritmo busca minimizar a distância dos elementos de um conjunto de dados com k centros, de tal forma que as observações dentro de cada grupo estejam mais próximas entre si do que as observações de outros grupos.

Para a definição do número de grupos, foram conduzidos testes exploratórios nos quais os resultados obtidos para números distintos de grupos foram analisados de forma qualitativa e quantitativa. Em relação à última, observou-se, em particular, a relação entre o aumento do número de grupos e a razão entre a soma dos quadrados dos desvios *entre* os grupos e a soma dos quadrados total.

QUADRO 5. Variáveis selecionadas para a análise de agrupamentos

Variável	Descrição
NUIDenDom	Densidade domiciliar do NUI (IBGE, 2010)
Declividade	Declividade média do terreno
DomSBan	Domicílios particulares permanentes sem banheiro de uso exclusivo dos moradores;
DomSAgua	Domicílios sem abastecimento de água de rede geral;
AguaNascente	Domicílios particulares permanentes com abastecimento de água de poço ou nascente na propriedade
DomSRedeEsg	Domicílios sem ligação à rede de esgoto ou fossa séptica (%)
DomSEsg	Domicílios com esgoto a céu aberto (%)
DomSEne	Domicílios sem energia elétrica;
DomSMed	Domicílios sem medidor de uso exclusivo;
DomSCollLix	Domicílios sem coleta de lixo;
DomCLixAc	Domicílios com lixo acumulado nos logradouros;
DomSPav	Domicílios sem pavimentação;
DomSCal	Domicílios sem calçada;
DomSIden	Domicílios sem identificação do logradouro;
DomSIlu	Domicílios sem iluminação pública;
RenRespMed	Renda média do responsável pelo domicílio;

Fonte: Elaboração própria, 2021.

Como resultado das análises de agrupamentos, obtiveram-se, para cada um dos polos da pesquisa, diferentes tipologias. As características médias das distintas tipologias de cada polo, considerando as variáveis utilizadas na análise de agrupamentos, estão sumarizadas na Tabela 2.

A seguir, são descritas as características das tipologias identificadas para cada polo (Seções 4.1 a 4.6). Por fim, na Seção 4.7, as tipologias de NUI são comparadas aos dados dos aglomerados subnormais, com o objetivo de verificar se esse dado, que é disponibilizado pelo IBGE para todo o país, representa uma amostra adequada dos NUI.

TABELA 2. Tipologias dos polos e variáveis quantitativas. Todas as variáveis correspondem ao percentual de domicílios, com exceção de NUIDenDom (domicílios/ha), Declividade (%) e RenRespMedia (R\$).

POLOS	BELO HORIZONTE			BRASÍLIA				JUAZEIRO DO NORTE			MARABÁ			PORTO ALEGRE				RECIFE			
Variável	BH-1	BH-2	BH-3	BSB-1	BSB-2	BSB-3	BSB-4	JN-1	JN-2	JN-3	MB-1	MB-2	MB-3	POA-1	POA-2	POA-3	POA-4	RE-1	RE-2	RE-3	RE-4
NUIDenDom	25,49	20,91	3,87	24,92	8,86	15,10	4,98	26,71	9,85	4,09	14,18	8,89	6,00	34,22	24,11	24,87	16,94	48,86	32,10	31,43	17,70
Declividade																		4,10	4,90	4,88	5,53
DomSBan	1,82	4,40	3,71	3,32	3,37	1,52	4,35	6,80	13,93	42,13	19,43	26,61	42,70	2,51	3,09	2,96	2,55	4,18	3,61	5,04	9,48
DomSAqua	1,51	5,56	27,13	5,70	12,30	2,43	66,43				50,57	76,86	77,42	4,38	10,92	5,13	26,57	12,06	9,86	15,75	25,04
AguaNascent								1,09	3,67	11,87											
DomSRedeEsg	12,39	44,68	63,84	13,47	68,01	62,06	74,52	35,47	86,15	87,24	72,61	76,02	83,01	15,14	18,93	18,16	13,45	28,90	42,83	42,58	48,35
DomSEsg	3,49	32,19	7,52	0,69	12,80	68,72	1,88	45,45	58,15	6,06	38,70	49,55	11,54	11,12	11,59	13,40	9,06	19,98	16,02	24,06	29,78
DomSEne	0,15	0,36	1,11	0,12	0,54	0,09	1,22	0,80	1,02	2,02	1,08	1,13	12,81	0,29	0,44	0,27	0,37	0,17	0,20	0,24	0,70
DomSMed	3,65	19,35	8,47	6,23	4,60	37,44	20,08	0,92	1,36	2,54	17,81	26,12	11,34	17,53	11,26	20,24	7,61	10,73	9,51	9,78	8,76
DomSCollix	2,09	8,36	18,39	1,81	5,48	23,99	40,14	4,60	15,10	66,58	14,61	26,33	69,73	0,85	1,05	1,17	1,69	5,80	5,80	8,37	16,38
DomCLixAc	3,84	16,90	1,58	1,78	6,31	35,78	0,60	3,15	12,33	0,67	5,45	11,53	2,25	11,24	8,95	12,70	10,00	5,93	3,07	4,65	4,94
DomSPav	8,41	37,25	47,19	6,37	37,39	79,85	4,40	15,32	18,97	5,53	60,09	77,92	17,36	30,79	35,64	35,11	28,32	24,21	16,20	27,84	28,96
DomSCal	24,91	54,33	62,52	42,89	70,80	86,70	6,66	15,66	25,43	6,11	74,64	85,83	20,18	51,74	56,27	54,55	58,08	32,58	31,13	39,43	44,70
DomSIden	36,73	53,24	49,92	66,57	74,90	79,24	4,32	65,90	52,57	13,07	47,39	69,52	16,31	52,36	61,04	53,63	58,70	36,02	34,34	42,55	47,67
DomSIlu	2,54	8,44	7,31	2,42	5,60	3,78	0,21	3,59	3,71	1,98	10,59	31,93	6,97	13,42	7,03	12,00	7,55	3,60	3,03	3,62	4,64
RenRespMedia (R\$)	953	693	705	1.242	840	874	972	581	494	358	842	713	557	922	859	980	899	455	558	475	409

Fonte: Elaboração própria, 2021.

4.1 Polo Belo Horizonte

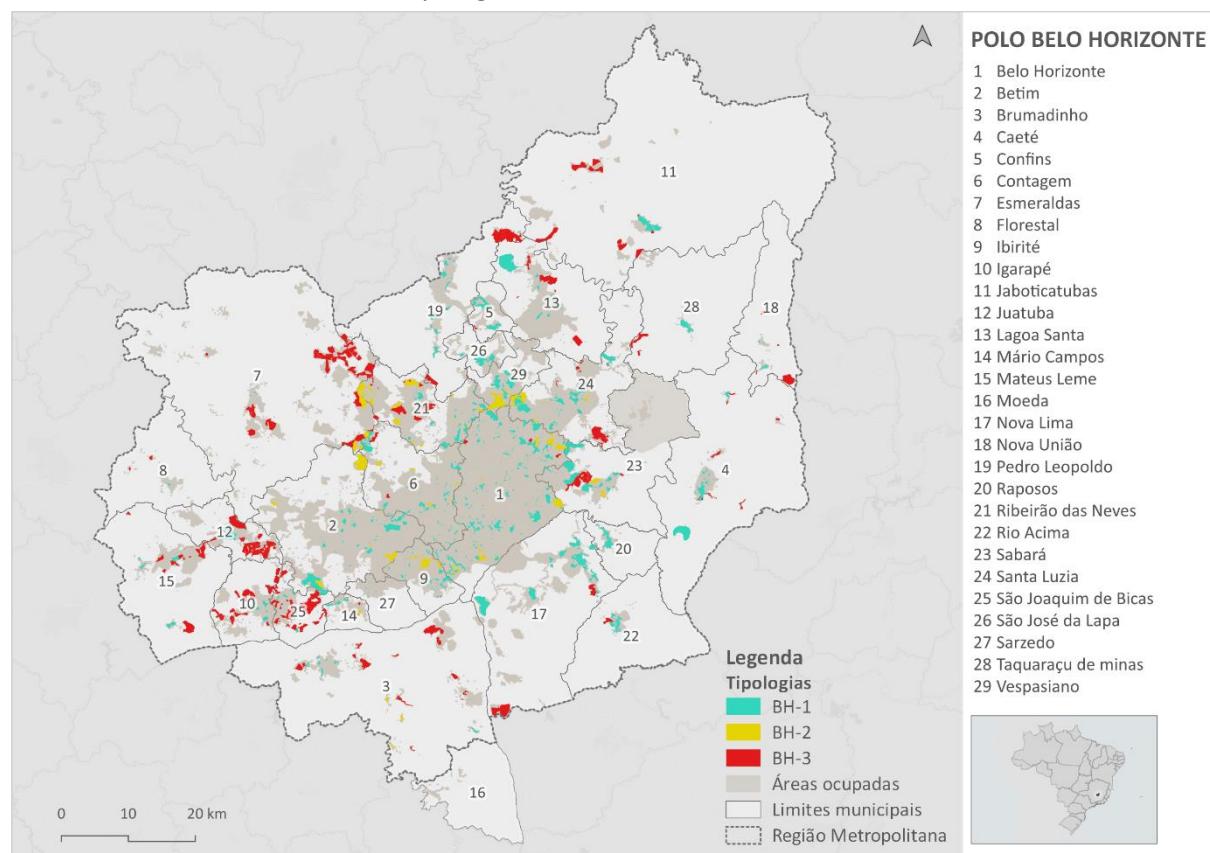
A análise de agrupamentos resultou em três tipologias para o polo Belo Horizonte: BH-1, BH-2 e BH-3 (Tabela 3 e Figura 15). As tipologias estão em ordem crescente de precariedade, considerando as variáveis incluídas na análise de agrupamento dos NUI, entre as quais destacam-se as que representam aspectos relacionados à infraestrutura.

TABELA 3. Tipologias do Polo Belo Horizonte

Polo Belo Horizonte	BH-1	BH-2	BH-3	Não Classificados	Total
Número de NUI	979 (69,3%)	176 (12,5%)	252 (17,8%)	6 (0,4%)	1.413 (100%)
Domicílios	313.566 (72,3%)	74.918 (17,3%)	44.762 (10,3%)	291 (0,1%)	433.357 (100%)

Fonte: Elaboração própria, 2021.

FIGURA 15. Polo Belo Horizonte: Tipologias BH-1, BH-2 e BH-3.



Fonte: Elaboração própria, 2021.

4.1.1 Tipologia BH-1

Os 979 NUI classificados na tipologia BH-1 são os que apresentam melhores índices de infraestrutura no polo, a maior renda média dos responsáveis pelos domicílios (R\$953) e também a maior densidade (25 domicílios/ha). São NUI inseridos completamente na malha urbana (Figura 16). Em relação aos índices de infraestrutura, a maioria dos domicílios possuem acesso à rede de abastecimento de água (98,5%), coleta de lixo por serviço de limpeza (98%) e ligação à rede de esgoto (88%). Essa tipologia destaca-se ainda por apresentar as menores proporções de domicílios sem pavimentação (8,4%), sem calçada (24%) e sem identificação de logradouros (37%).

FIGURA 16. Polo Belo Horizonte: Tipologia BH-1.



Fonte: Google Satellite, 2021.

4.1.2 Tipologia BH-2

A tipologia BH-2 abrange 176 NUI e apresenta a menor renda média dos responsáveis por domicílio (R\$ 693). Apresenta uma densidade média intermediária, de 20 domicílios/ha (Figura 17). Em relação a localização, são geralmente periféricos, mas ainda assim inseridos na malha urbana. Quanto ao abastecimento de água, 95% possuem acesso à rede de água. Por outro lado, 45% dos domicílios não possuem acesso à rede de esgoto e é a tipologia do polo de BH com a maior proporção de domicílios com esgoto a céu aberto (32%). Apenas 8% dos domicílios não possuem coleta de lixo, ao passo que 16% dos domicílios possuem lixo acumulado no entorno. Ainda a respeito da infraestrutura, os NUI da BH-2 possuem alta proporção de domicílios sem pavimentação (37%), sem calçada (54%) e sem identificação de logradouro (53%).

FIGURA 17. Belo Horizonte: Tipologia BH-2.



Fonte: Google Satellite, 2021.

4.1.3 Tipologia BH-3

A tipologia BH-3 abrange 252 NUI, cuja renda média dos responsáveis por domicílios é de R\$705. Destaca-se por apresentar a menor densidade da região (4 domicílios/ha) e localização predominantemente periférica (Figura 18). A maioria dos NUI desta tipologia estão localizadas nos municípios mais periféricos do polo, sendo a tipologia predominante de municípios como Brumadinho, Esmeraldas, Juatuba e Igarapé. Destaca-se a carência de infraestrutura dos NUI dessa tipologia, que apresentam a maior proporção de domicílios sem acesso à rede de água (27%), sem acesso a rede de esgoto (63%) e sem coleta de lixo (18%). A tipologia BH-3 também apresenta a maior proporção de domicílios sem calçada (62%), sem pavimentação (47%) e sem identificação de logradouros (49%).

FIGURA 18. Belo Horizonte: Tipologia BH-3.



Fonte: Google Satellite, 2021.

4.2 Polo Brasília

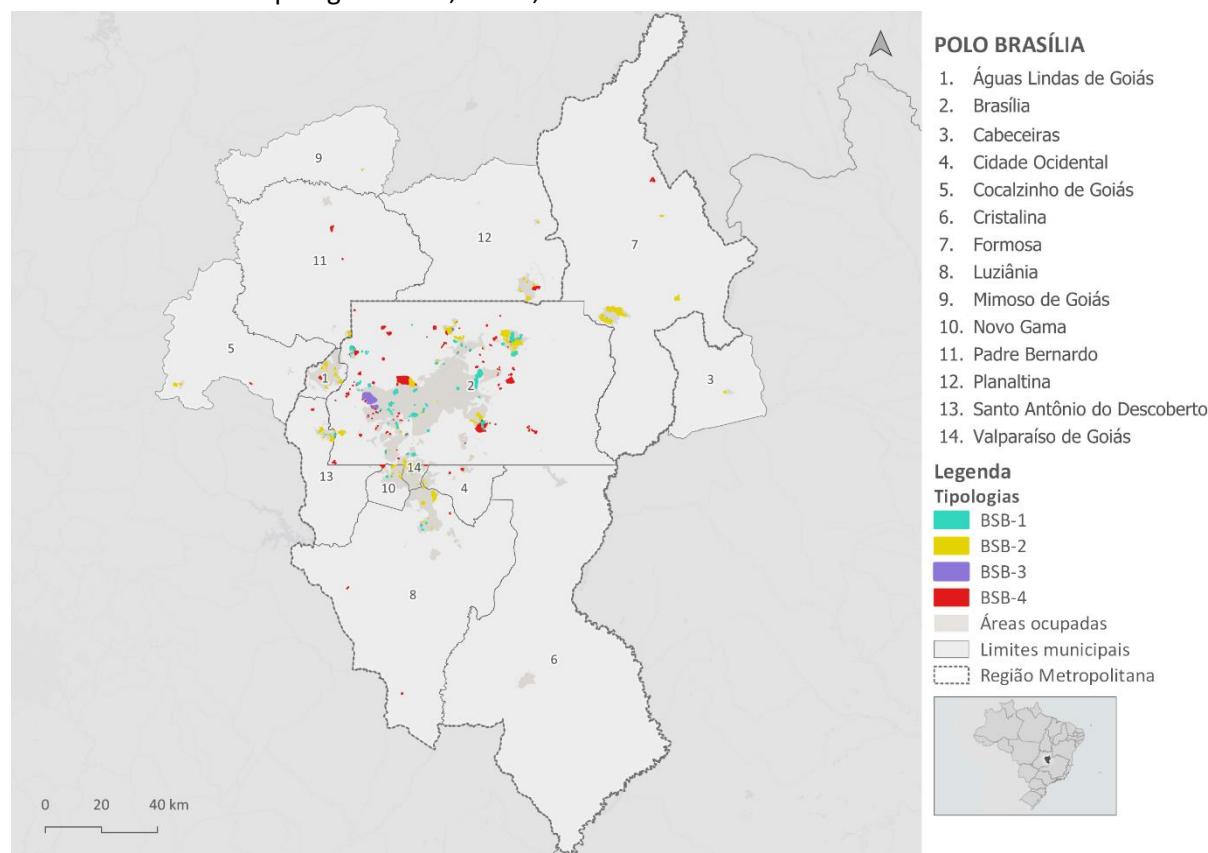
Foram estabelecidas quatro tipologias para o polo de Brasília: BSB-1, BSB-2, BSB-3 e BSB-4, sendo a primeira a menos precária e a última a que caracteriza núcleos mais isolados (Figura 19). Nas três primeiras predominam núcleos inseridos na malha urbana, sendo a tipologia BSB-1 a menos precária, enquanto a BSB-3 é a mais precária. Na tipologia BSB-4 predominam núcleos isolados, com os piores indicadores de infraestrutura e serviços urbanos. A tabela 4 apresenta a distribuição das tipologias do polo por número de NUI e de domicílios.

TABELA 4. Tipologias do Polo Brasília.

Polo Brasília	BSB-1	BSB-2	BSB-3	BSB-4	Não Classificados	Total
Número de NUI	107 (25,5%)	139 (33,1%)	24 (5,7%)	149 (35,5%)	1 (0,2%)	420 (100%)
Domicílios	59.517 (38,5%)	61.230 (39,6%)	14.060 (9,1%)	19.868 (12,8%)	14 (0,01%)	154.689 (100%)

Fonte: Elaboração própria, 2021.

FIGURA 19. Brasília: Tipologias BSB-1, BSB-2, BSB-3 e BSB-4.



Fonte: Elaboração própria, 2021.

4.2.1 Tipologia BSB-1

A tipologia BSB-1 abrange 107 NUI, entre os quais 94 (87%) localizam-se no Distrito Federal. São NUI predominantemente inseridos na malha urbana e com menor precariedade (Figura 20). Apresentam

a maior densidade de domicílios da região (24 domicílios/ha) e maior renda média dos responsáveis pelos domicílios (R\$ 1.242). A tipologia destaca-se também por apresentar as melhores condições de infraestrutura, com 96% dos domicílios com abastecimento de água por rede geral, 99% com coleta de lixo por serviço de limpeza e 87% com ligação à rede de esgoto ou fossa séptica. Também apresenta um percentual baixo de domicílios em ruas sem pavimentação (6%), embora o mesmo não possa ser afirmado em relação à ausência de calçada (43%) e identificação de logradouros (67%). Quanto aos dois últimos indicadores, cabe salientar que, embora ruins, apresentam-se superiores aos observados nas demais tipologias dos NUI de Brasília, com exceção da BSB-4.

FIGURA 20. Brasília: Tipologia BSB-1.



Fonte: Google Satellite, 2021.

4.2.2 Tipologia BSB-2

A tipologia BSB-2 abrange 139 NUI, a maioria inseridos na malha urbana (Figura 21). Predominam NUI com baixa densidade média (8 domicílios/ha) e consolidados (87% possuem tempo de estabelecimento acima de 10 anos). Essa tipologia apresenta a menor renda média do responsável pelo domicílio (R\$ 840) e elevada proporção de domicílios sem pavimentação (37%), sem calçada (70%) e sem identificação de logradouro (74%). Em relação à infraestrutura, 86% dos domicílios possuem acesso à rede de abastecimento de água, 95% possuem coleta de lixo por serviço de limpeza e 33% possuem esgotamento adequado.

FIGURA 21. Brasília: Tipologia BSB-2.



Fonte: Google Satellite, 2021.

4.2.3 Tipologia BSB-3

A tipologia BSB-3 apresenta a pior situação de precariedade do polo (Figura 22). Todos os 24 NUI dessa tipologia são loteamentos irregulares inseridos na malha urbana do município de Brasília, e em áreas de proteção dos mananciais. Há também o predomínio de NUI com situação de risco ambiental. A renda média do responsável pelo domicílio é de R\$874. Em relação à infraestrutura, apresenta a maior proporção de domicílios sem pavimentação (79%), sem identificação de logradouro (79%), sem calçada (86%), sem coleta de lixo por serviço de limpeza (23%). Destaca-se ainda a elevada proporção de domicílios com lixo acumulado nas ruas (35%), o predomínio de domicílios com esgoto a céu aberto (68%) e maior proporção de domicílios sem medidor de energia (37%).

FIGURA 22. Brasília: Tipologia BSB-3.



Fonte: Google Satellite, 2021.

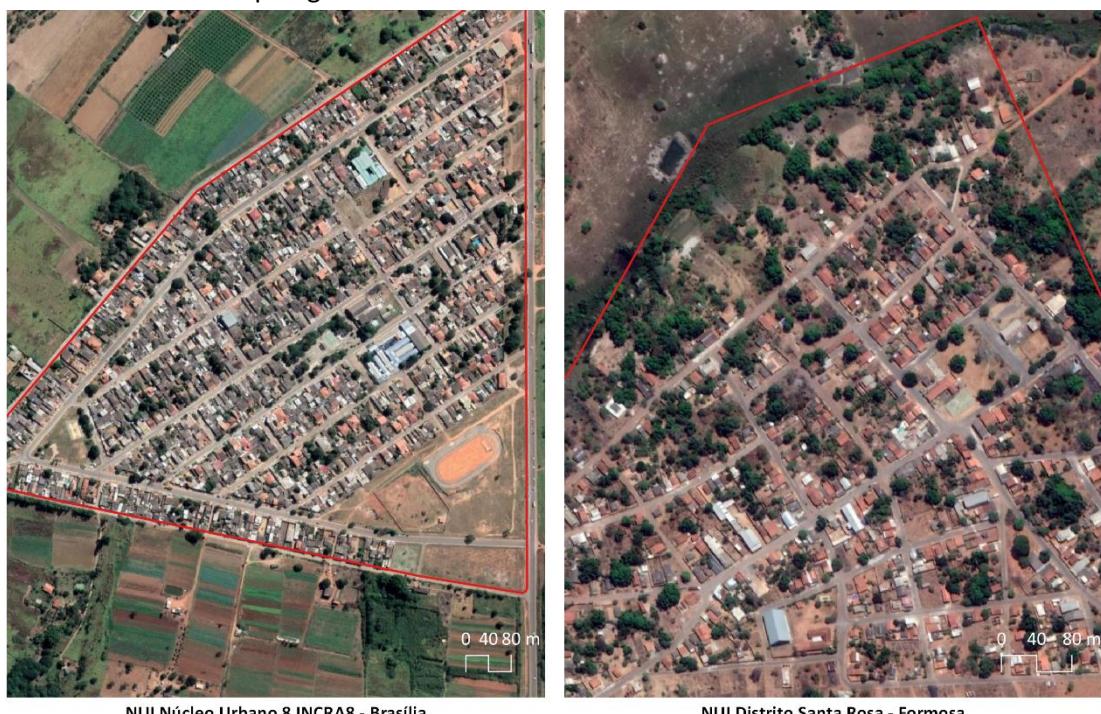
4.2.4 Tipologia BSB-4

A tipologia BSB-4 abrange 149 NUI, predominando assentamentos isolados da malha urbana, de baixa densidade (média de 5 domicílios/ha), majoritariamente em área de proteção dos mananciais e unidades de conservação de uso sustentável (85%) (Figura 23). Há, ainda, uma elevada proporção de NUI que apresentam situação de risco (49%) e menos de 10 anos de tempo de estabelecimento (49%). A tipologia BSB-4 possui a segunda maior renda média dos responsáveis pelos domicílios (R\$972).

Em relação à infraestrutura e serviços, o acesso à água se dá predominantemente por poço na propriedade (ou nascente) (60%), 40% dos domicílios não possuem coleta de lixo por serviço de limpeza e 74% dos domicílios não possuem esgotamento adequado. Apesar disso, praticamente não se observa, no entorno, esgoto a céu aberto ou lixo acumulado nas ruas (1% e 0,6% dos domicílios, respectivamente). Quanto à energia elétrica, cujo acesso é praticamente universalizado no polo, destaca-se o número de domicílios sem medidor de energia elétrica (20%).

Apesar dos altos níveis de precariedade apresentados pelos indicadores acima mencionados, os NUI dessa tipologia apresentam a melhor situação quanto à pavimentação das ruas (96%), presença de calçada (93%) e identificação de logradouros (96%).

FIGURA 23. Brasília: Tipologia BSB-4.



Fonte: Google Satellite, 2021.

4.3 Polo Juazeiro do Norte

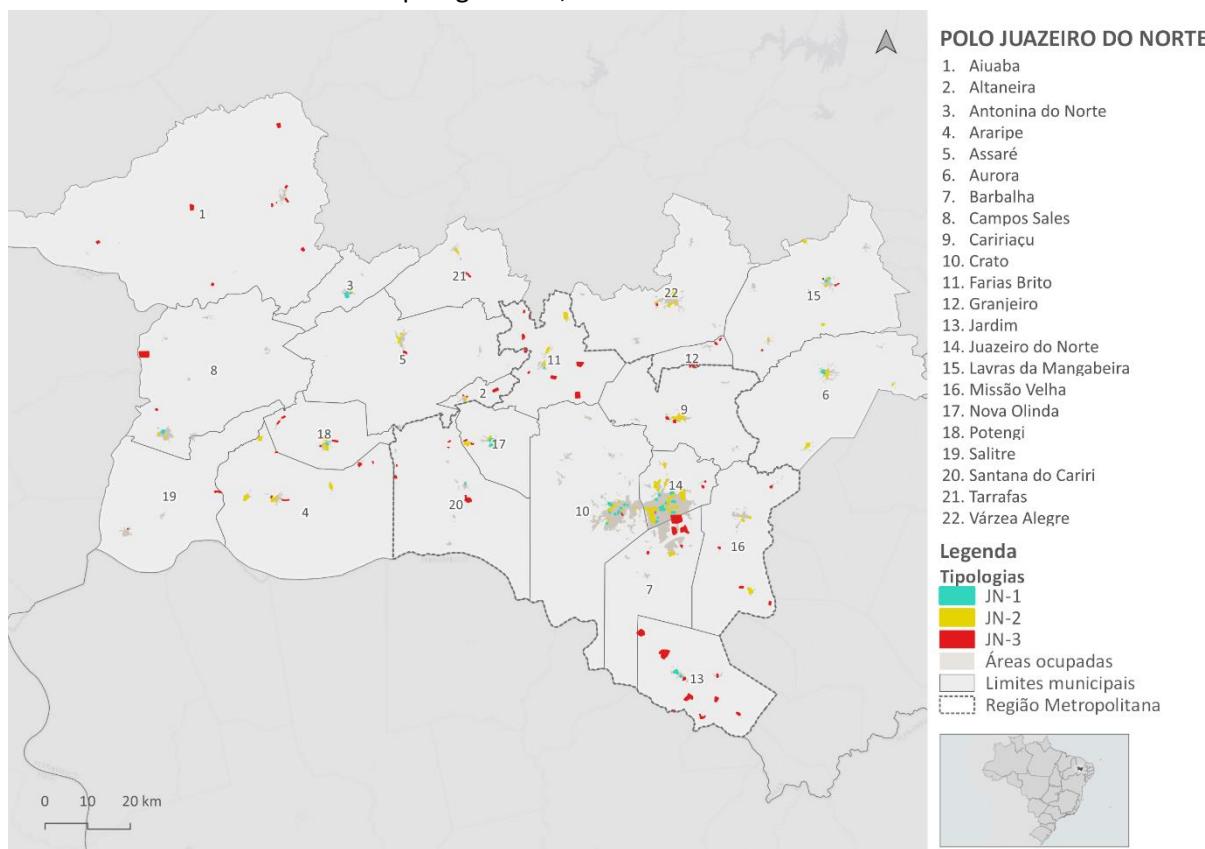
Foram estabelecidas três tipologias para o polo de Juazeiro do Norte: JN-1, JN-2 e JN-3, em ordem crescente de precariedade (Tabela 5 e Figura 24).

TABELA 5. Tipologias do Polo Juazeiro do Norte

Polo Juazeiro do Norte	JN-1	JN-2	JN-3	Não Classificados	Total
Número de NUI	39 (17,8%)	99 (45,2%)	79 (36,1%)	2 (0,9%)	219 (100%)
Domicílios	8.594 (21,4%)	17.372 (43,2%)	14.115 (35,1%)	142 (0,4%)	40.223 (100%)

Fonte: Elaboração própria, 2021.

FIGURA 24. Juazeiro do Norte: Tipologias JN-1, JN-2 e JN-3.



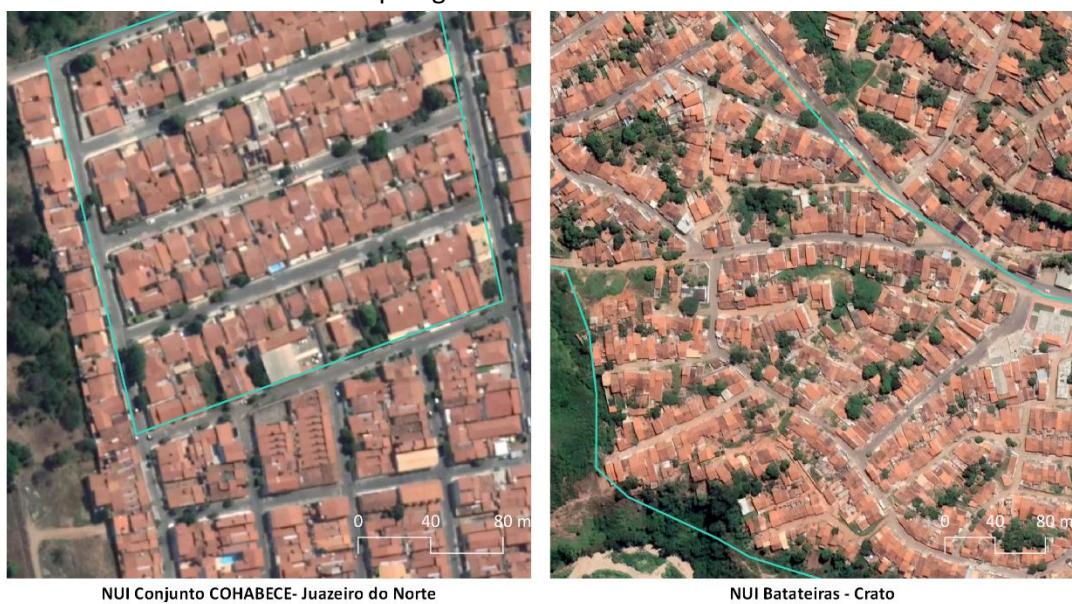
Fonte: Elaboração própria, 2021.

4.3.1 Tipologia JN-1

A tipologia JN-1 caracteriza-se pelo menor grau de precariedade e maior densidade de domicílios da região (27 domicílios/ha). Abrange 39 NUI, entre os quais predominam favelas ou loteamento irregulares, localizados na malha urbana, com traçado semelhante ao da cidade formal e lotes bem definidos (Figura 25). Apresenta a maior renda média dos responsáveis pelos domicílios (R\$581), bem como as melhores condições de infraestrutura: menor proporção de domicílios sem abastecimento de água (4%), sem esgotamento adequado (35%) e sem coleta de lixo (4,6%). Apesar de apresentar, em comparação com as demais tipologias, baixa proporção de domicílios sem esgotamento adequado, os dados de entorno revelam que 45% dos domicílios estão localizados em logradouros com esgoto a

céu aberto. Destaca-se ainda o elevado percentual de domicílios em logradouros sem pavimentação e calçada (15%) e sem identificação (66%).

FIGURA 25. Juazeiro do Norte: Tipologia JN-1.



Fonte: Google Satellite, 2021.

4.3.2 Tipologia JN-2

A tipologia JN-2 abrange 99 NUI, a maioria inseridos na malha urbana, porém em localização mais periférica do que os da tipologia JN-1. Em geral, apresentam lotes bem definidos e baixa densidade média de domicílios (10 domicílios/ha), revelando, por vezes, assentamentos ainda em processo de consolidação (Figura 26). Nesse sentido, apresenta qualidade de infraestrutura inferior à média da tipologia JN-1: 12% dos domicílios não possuem abastecimento por rede de água, 86% não possuem esgotamento adequado e 15% não possuem coleta de lixo realizada por serviço de limpeza. Em relação ao esgotamento sanitário, predominam as fossas rudimentares, utilizadas em mais de 77% dos domicílios nos NUI classificados como JN-2. É a tipologia que apresenta a maior proporção de domicílios em entorno caracterizado pela presença de esgoto a céu aberto (58%).

Em relação às condições das vias de circulação, apresenta alta proporção de domicílios em vias não pavimentadas (19%), sem calçada (25%) e sem identificação (53%). A média da renda dos responsáveis por domicílios é de aproximadamente R\$495, próxima à média da renda dos responsáveis residentes em NUI no polo de Juazeiro do Norte (R\$464).

FIGURA 26. Juazeiro do Norte: Tipologia JN-2.



Fonte: Google Satellite, 2021.

4.3.3 Tipologia JN-3

A tipologia JN-3 abrange 79 NUI, entre os quais predominam os do tipo distrito ou povoado (69%) (Figura 27). São NUI afastados da malha urbana e com densidade de domicílios muito baixa (média de 4 domicílios/ha). A média da renda dos responsáveis por domicílios é de aproximadamente R\$358, inferior às das demais tipologias do polo. Essa tipologia apresenta também as piores condições de infraestrutura: 55% dos domicílios não apresentam abastecimento de água por rede geral, 66,8% não possuem coleta de lixo por serviço de limpeza e 87% não possuem esgotamento adequado. Destaca-se ainda o fato de 42% dos domicílios não possuírem banheiro próprio.

Por outro lado, a tipologia JN-3 apresenta o menor percentual de domicílios em vias sem pavimentação (5,5%), sem calçadas (6%) ou sem identificação (13%). Também é baixa a presença de domicílios em vias com esgoto a céu aberto (6%) e praticamente inexistente a presença de lixo acumulado nas ruas.

FIGURA 27. Juazeiro do Norte: Tipologia JN-3.



Fonte: Google Satellite, 2021.

4.4 Polo Marabá

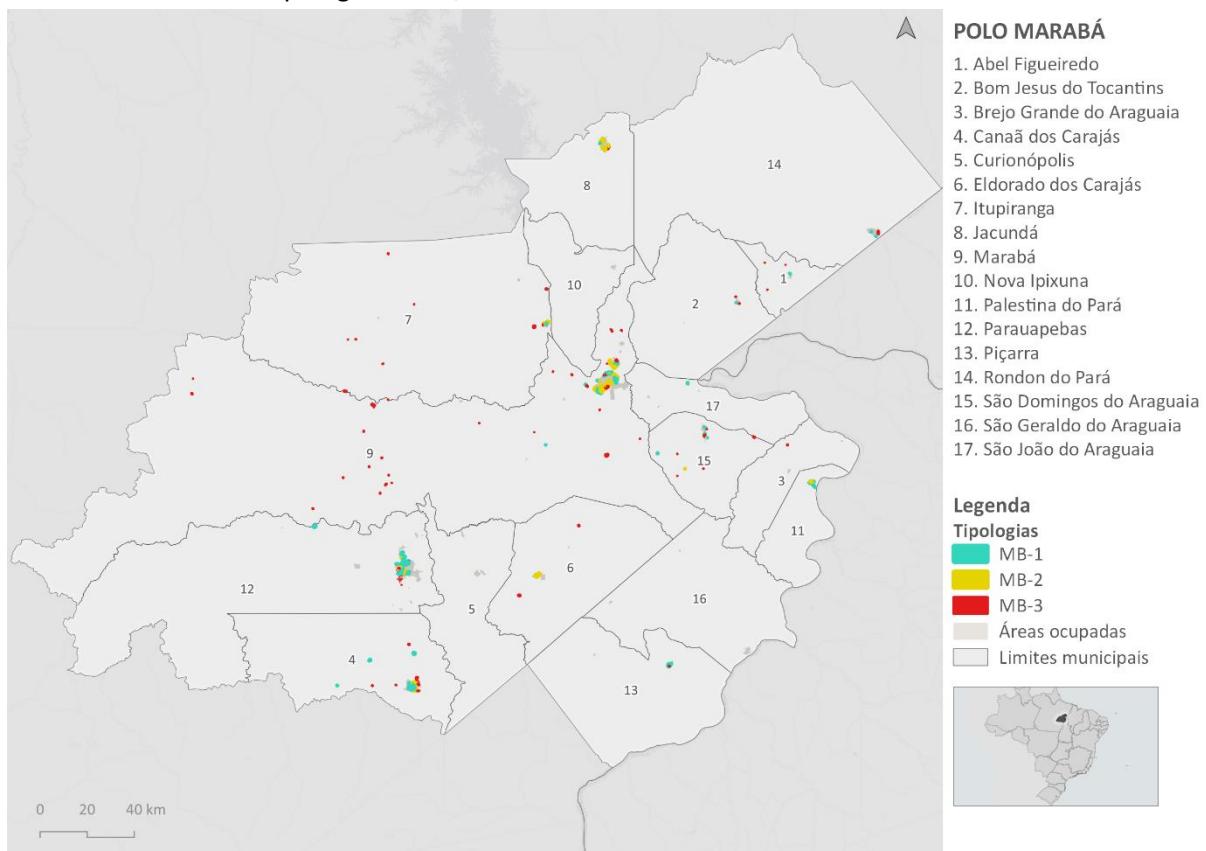
Foram estabelecidas três tipologias para o polo de Marabá: MB-1, MB-2 e MB-3 (Tabela 6 e Figura 28). Nas duas primeiras predominam NUI inseridos na malha urbana, sendo a primeira a menos precária. Na terceira tipologia predominam núcleos isolados, com os piores indicadores de infraestrutura e serviços urbanos.

TABELA 6. Tipologias do Polo Marabá.

Polo Marabá	MB-1	MB-2	MB-3	Não Classificados	Total
Número de NUI	83 (39,2%)	54 (25,5%)	72 (34,0%)	3 (1,4%)	212 (100%)
Domicílios	38.608 (54,3%)	20.426 (28,7%)	10.040 (14,1%)	2.045 (2,9%)	71.119 (100%)

Fonte: Elaboração própria, 2021.

FIGURA 28. Marabá: Tipologias MB-1, MB-2 e MB-3.



Fonte: Elaboração própria, 2021.

4.4.1 Tipologia MB-1

A tipologia MB-1 representa NUI inseridos na malha urbana e com os menores níveis de precariedade (Figura 29). Abrange 83 NUI que apresentam, em média, densidade de 14,2 domicílios/ha. Embora baixa, essa densidade é superior à observada nas demais tipologias. A renda média dos responsáveis dos domicílios localizados nos NUI dessa tipologia é de R\$842, superior à média nos NUI do Polo (R\$767).

Caracteriza-se por condições precárias de infraestrutura e serviços urbanos, embora, em comparação com as demais tipologias, apresente as melhores condições. Em relação ao abastecimento de água, 50,6% dos domicílios não dispõem de abastecimento por rede geral e aproximadamente 44% dos domicílios são abastecidos por intermédio de nascente ou poço na propriedade. A ausência de esgotamento adequado é generalizada nos NUI do Polo Marabá, e, embora a tipologia MB-1 apresente condições superiores à média, apenas 5% dos domicílios possuem atendimento por rede de esgoto e 22% possuem fossa séptica, ao passo que a maioria (61%) dispõe de fossa rudimentar. Quanto à coleta de lixo, 85% dos domicílios possuem coleta por serviço de limpeza (de forma direta ou indireta), enquanto em 11% dos domicílios o lixo é queimado. Cerca de 20% dos domicílios não possuem banheiro de uso exclusivo.

Em relação às vias de circulação, a maioria dos domicílios localizam-se em vias não pavimentadas (60%) e sem calçada (74%). Boa parte dos domicílios situam-se ainda em vias não identificadas (47%) e sem iluminação (11%).

FIGURA 29. Marabá: Tipologia MB-1.



Fonte: Google Satellite, 2021.

4.4.2 Tipologia MB-2

A tipologia MB-2 abrange NUI inseridos na malha urbana, com densidade média de 8,9 domicílios/ha e renda média dos responsáveis equivalente a R\$714 (Figura 30). Em comparação com a tipologia MB-1, os 54 NUI dessa tipologia apresentam, em média, níveis mais acentuados de precariedade em relação a todas as variáveis analisadas. Apenas 22% dos domicílios têm acesso à rede geral de abastecimento de água e 66% dos domicílios são abastecidos por intermédio de nascente ou poço na propriedade. Quanto ao esgotamento, 76% dos domicílios apresentam condições inadequadas, predominando a utilização de fossas rudimentares (58%). É a tipologia com a maior proporção de domicílios em cujo entorno observa-se esgoto a céu aberto (50%). Há ainda uma proporção expressiva de domicílios sem banheiro de uso exclusivo (26%). Em relação à coleta de lixo, a maioria dos domicílios apresenta coleta por meio de serviço de limpeza (74%), enquanto 20% dos domicílios têm seu lixo queimado. Constata-se lixo acumulado no entorno de 11% dos domicílios.

As condições das vias de circulação são ainda piores do que as observadas nos NUI classificados como MB-1: 78% dos domicílios estão em vias sem pavimentação, 86% em vias sem calçada, 70% em vias sem identificação e 32% sem iluminação.

FIGURA 30. Marabá: Tipologia MB-2.



Fonte: Google Satellite, 2021.

4.4.3 Tipologia MB-3

Na tipologia MB-3 predominam núcleos isolados da malha urbana, do tipo distrito ou povoado (Figura 31). Em comparação às demais tipologias, os 72 NUI pertencentes a essa tipologia apresentam a menor renda média dos responsáveis (R\$557) e os piores indicadores em relação ao abastecimento de energia elétrica, água, esgotamento sanitário e coleta de lixo.

Em relação à energia elétrica, 13% dos domicílios não têm acesso, um percentual elevado se comparado às demais tipologias, onde praticamente a totalidade dos domicílios é atendida. Quanto ao abastecimento de água, 77,4% dos domicílios não têm acesso à rede geral e predomina o acesso por intermédio de nascentes ou poços na propriedade (62%). Cerca de 83% dos domicílios não possui esgotamento adequado (rede de esgoto ou fossa séptica), predominando fossas rudimentares (58%). Destaca-se também a expressiva proporção de domicílios sem banheiro (43%). Em relação ao destino do lixo, 70% dos domicílios não são atendidos por serviço de limpeza e em 62% dos domicílios o lixo é queimado.

Apesar das condições inadequadas de esgotamento e atendimento de serviço de coleta de lixo, é a tipologia que apresenta menor percentual de domicílios cujo entorno possui esgoto à céu aberto (11%) e lixo acumulado nas ruas (2%). É também a que apresenta a menor proporção de vias sem pavimentação (17%), sem calçada (20%), sem identificação (16%) e sem iluminação pública (7%).

FIGURA 31. Marabá: Tipologia MB-3.



Fonte: Google Satellite, 2021.

4.5 Polo Porto Alegre

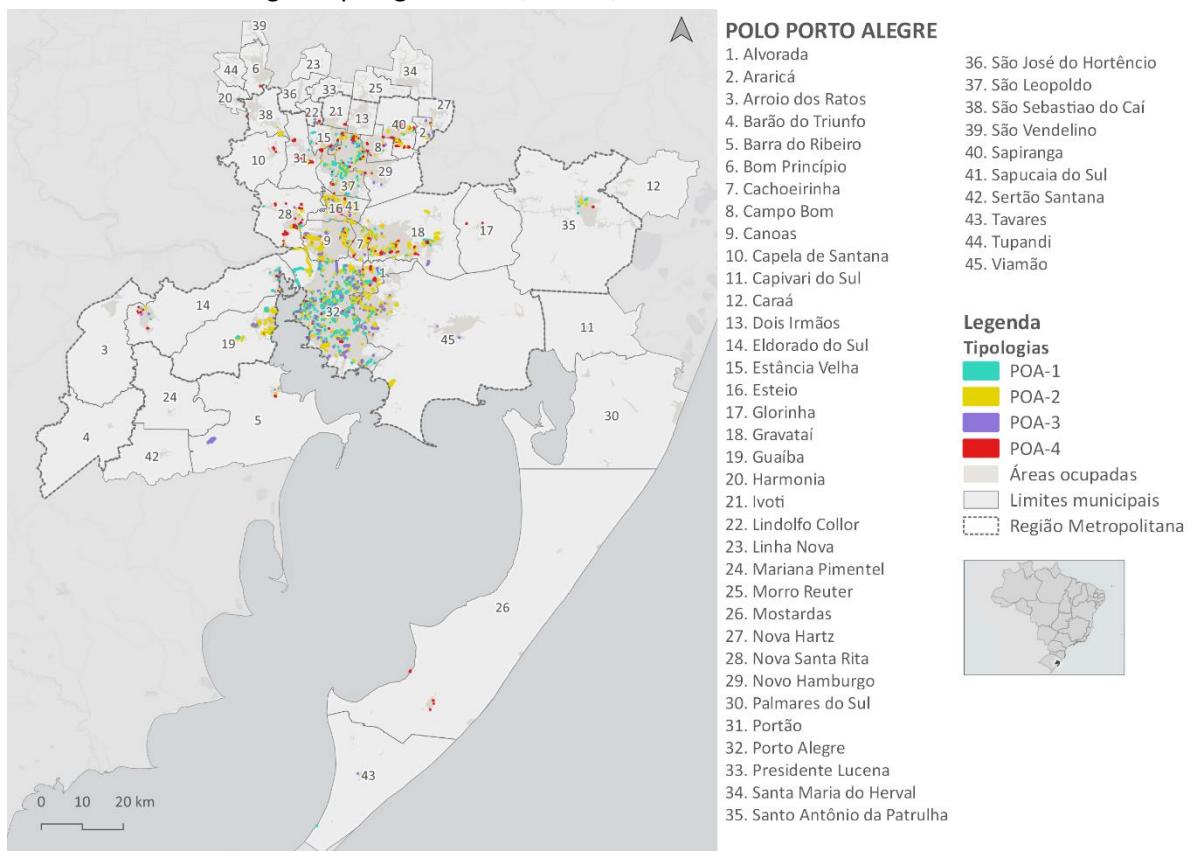
Foram definidas 4 tipologias: POA-1, POA-2, POA-3, POA-4, sendo a primeira menos precária e com maior densidade de domicílios, e a última a mais precária e com menor densidade (Figura 32). A tabela 7 apresenta a distribuição das tipologias do polo por número de NUI e de domicílios.

TABELA 7. Tipologias do Polo Porto Alegre.

Polo Porto Alegre	POA-1	POA-2	POA-3	POA-4	Não Classificados	Total
Número de NUI	406 (27,5%)	591 (40,0%)	263 (17,8%)	184 (12,4%)	35 (2,4%)	1.479 (100%)
Domicílios	53.393 (30,8%)	75.457 (43,5%)	28.031 (16,2%)	12.284 (7,1%)	4.248 (2,4%)	173.413 (100%)

Fonte: Elaboração própria, 2021.

FIGURA 32. Porto Alegre: Tipologias POA-1, POA-2, POA-3 e POA-4.



Fonte: Elaboração própria, 2021.

4.5.1 Tipologia POA-1

A tipologia POA-1 apresenta as melhores condições de infraestrutura, a maior densidade de domicílios (34 domicílios/ha) e a segunda maior renda média dos responsáveis (R\$ 922) (Figura 33). São 406 NUI do tipo POA-1, localizados predominantemente em áreas centrais do polo, principalmente no município de Porto Alegre (63% dos NUI).

A maioria dos domicílios é considerada adequada quanto à infraestrutura. O acesso à rede de abastecimento de água é quase universal (96%), assim como a coleta de lixo por serviço de limpeza (99%). Quanto ao esgotamento, 14% dos domicílios não possuem esgotamento adequado (desses 4% por fossa rudimentar e 9% vala). Em relação ao entorno, apresenta esgoto a céu aberto em 11% dos domicílios. Em relação às condições das vias de circulação, 30% dos domicílios não possuem pavimentação, 51% não possuem calçada e 52% não apresentam identificação de logradouros.

FIGURA 33. Porto Alegre: Tipologia POA-1.



Fonte: Google Satellite, 2021.

4.5.2 Tipologia POA-2

A tipologia POA-2 apresenta condições intermediárias de infraestrutura, renda média dos responsáveis de R\$859 e densidade de 24 domicílios/ha (Figura 34). São 591 NUI, frequentemente localizados nas periferias da malha urbana do polo (53% dos NUI).

Cerca de 11% dos domicílios não possuem acesso à rede de abastecimento de água, sendo que 9% tem acesso por meio de nascente ou poço na propriedade. A coleta de lixo por serviço de limpeza é quase universal. Quanto ao esgotamento, 11% não possuem esgotamento adequado, sendo 9% com fossa rudimentar e 8% vala. A tipologia apresenta elevada proporção de domicílios em ruas sem pavimentação (35%), sem calçada (56%) e sem identificação de rua (61%).

FIGURA 34. Porto Alegre: Tipologia POA-2.



Fonte: Google Satellite, 2021.

4.5.3 Tipologia POA-3

A tipologia POA-3, juntamente com a tipologia POA-4, apresenta as piores condições de precariedade (Figura 35). São 263 NUI com densidade domiciliar média de 25 domicílios/ha. Essa tipologia representa NUI predominantemente localizados na malha urbana, geralmente na periferia. A inserção dos NUI na malha urbana associa-se a uma baixa proporção de domicílios sem acesso à rede de abastecimento de água (5%). Quanto ao esgotamento, os NUI dessa tipologia apresentam cerca de 18% dos domicílios sem esgotamento adequado, com destaque para as soluções do tipo vala (9%) e fossa rudimentar (6%). Como consequência, é a tipologia que apresenta maior proporção de esgoto a céu aberto (13%). Embora o serviço de coleta de lixo predomine, apresenta a maior proporção de domicílios com lixo acumulado nas ruas (12%).

Embora o abastecimento de energia seja praticamente universal no polo, os NUI dessa tipologia apresentam 20% dos domicílios sem medidor de energia. Essa tipologia apresenta alta proporção de domicílios sem pavimentação (35%), sem calçada (54%), sem identificação de logradouro (54%) e sem iluminação pública (12%).

FIGURA 35. Porto Alegre: Tipologia POA-3.



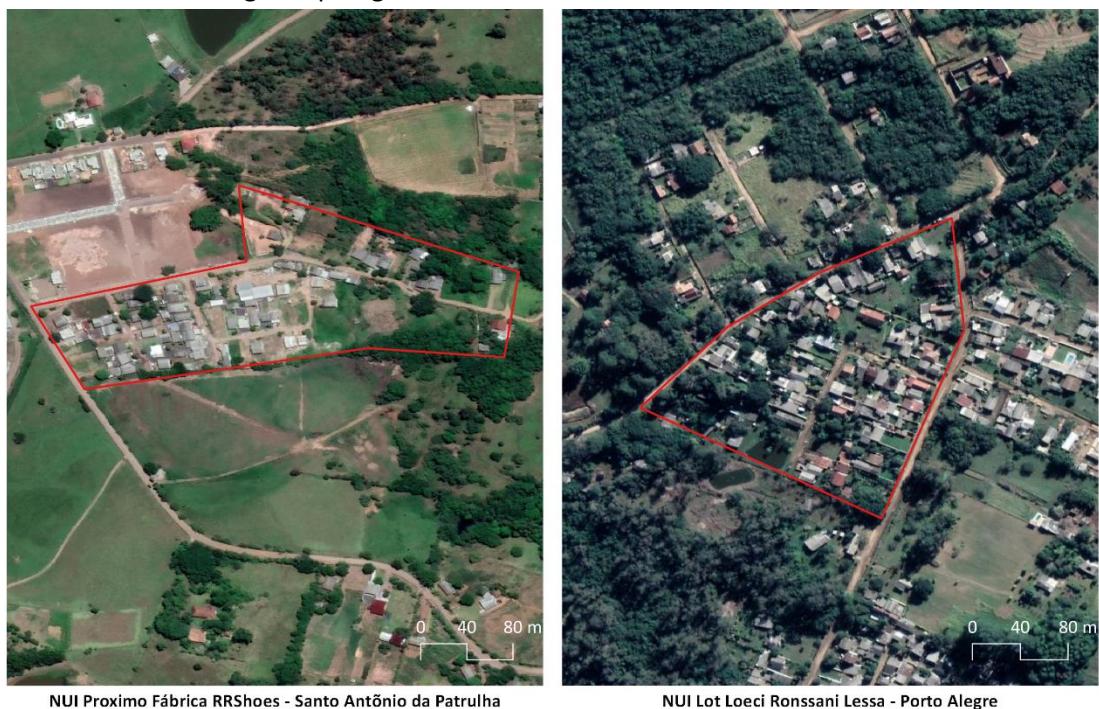
Fonte: Google Satellite, 2021.

4.5.4 Tipologia POA-4

Na tipologia POA-4 predominam NUI de baixa densidade (16 domicílios/ha) localizados frequentemente em locais afastados das áreas centrais (Figura 36). Abrange 184 NUI cuja renda média dos responsáveis pelos domicílios é de R\$899. Destaca-se pelo pior acesso a abastecimento de água por rede geral: 26% dos domicílios não possuem acesso à rede de água e, desses, 23% são abastecidos através de nascente ou poço na propriedade. Embora seja a tipologia com menor proporção de domicílios com acesso à rede de esgoto (58%, enquanto as tipologias POA-1, POA-2 e POA-3 apresentam 70%, 62% e 67%, respectivamente), apresenta melhores condições de esgotamento do que as demais tipologias graças à elevada porcentagem de domicílios com fossa séptica (26%). Consequentemente, apresenta a menor proporção de domicílios com esgoto à ceu aberto no entorno (9%).

Em relação à qualidade das vias de circulação, 28% dos domicílios estão em logradouros sem pavimentação, 58% sem calçada, 59% sem identificação do logradouro e 8% sem iluminação pública.

FIGURA 36. Porto Alegre: Tipologia POA-4.



Fonte: Google Satellite, 2021.

4.6 Polo Recife

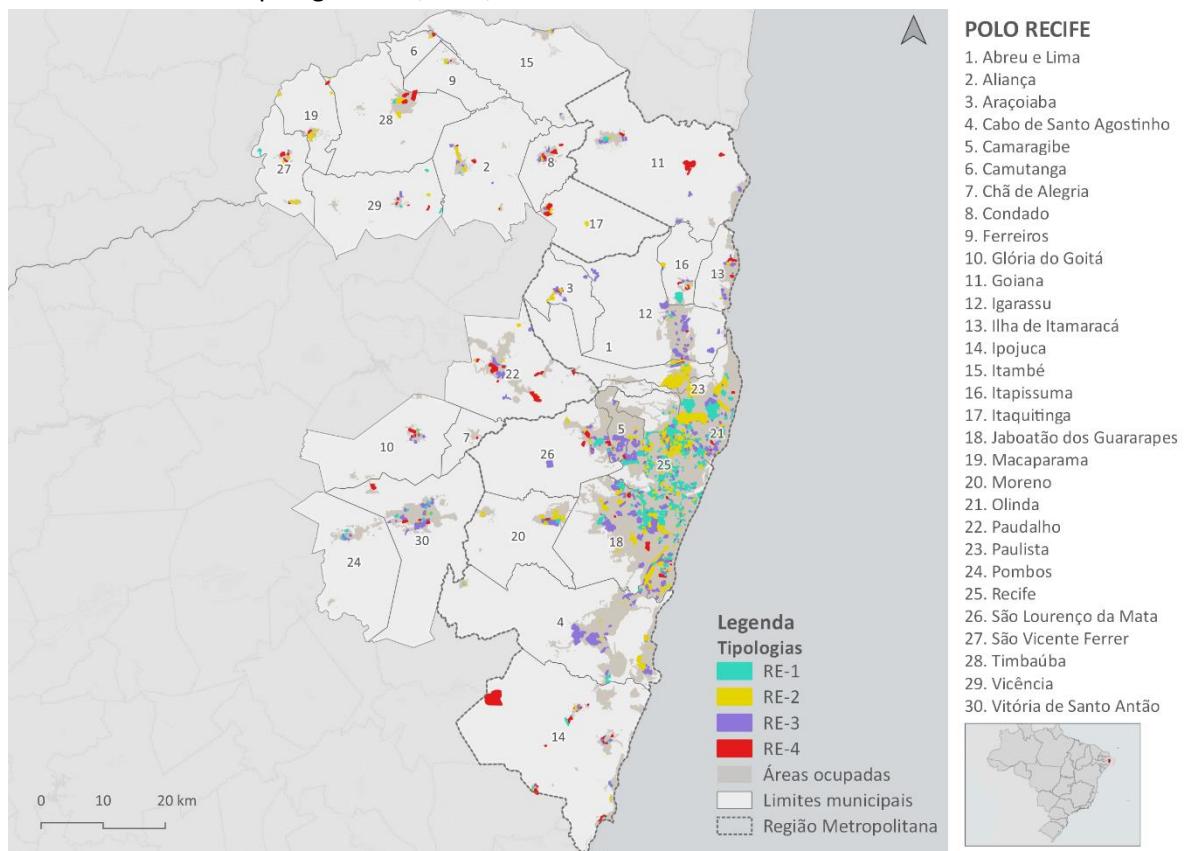
Foram definidas 4 tipologias: RE-1, RE-2, RE-3, RE-4, sendo a primeira menos precária e com maior densidade de domicílios, e a última a mais precária e com menor densidade (Figura 37). A tabela 8 apresenta a distribuição das tipologias do polo por número de NUI e de domicílios.

TABELA 8. Tipologias do Polo Recife

Polo Recife	RE-1	RE-2	RE-3	RE-4	Não Classificados	Total
Número de NUI	385 (31,4%)	336 (27,4%)	390 (31,8%)	104 (8,5%)	10 (0,8%)	1.225 (100%)
Domicílios	204.434 (33,3%)	194.803 (31,7%)	173.178 (28,2%)	37.308 (6,1%)	4.021 (0,7%)	613.744 (100%)

Fonte: Elaboração própria, 2021.

FIGURA 37. Recife: Tipologias RE-1, RE-2, RE-3 e RE-4.



Fonte: Elaboração própria, 2021.

4.6.1 Tipologia RE- 1

A tipologia RE-1 abrange 385 NUI, localizados predominantemente na malha urbana (89%). Possuem a maior densidade média do Polo (48 domicílios/ha) e a renda média dos responsáveis pelos domicílios é R\$ 455 (Figura 38). É a tipologia com maior proporção de NUI com menos de 5 anos de tempo de estabelecimento (25%).

Em relação à infraestrutura, possui a menor proporção de domicílios inadequados quanto ao esgotamento (29%) e 20% dos domicílios possuem esgoto à céu aberto em seu entorno. Quanto aos resíduos sólidos, apresenta a menor proporção de domicílios sem coleta de lixo por serviço de limpeza (6%), mas a maior proporção de domicílios cujo entorno possui lixo acumulado nas ruas (6%). Em relação à água, 12% dos domicílios não possuem abastecimento por rede geral.

Quanto às vias de circulação, 24% dos domicílios não têm pavimentação, 33% sem calçada, 36% sem identificação de logradouro e 4% sem iluminação pública.

FIGURA 38. Recife: Tipologia RE-1.



Fonte: Google Satellite, 2021.

4.6.2 Tipologia RE-2

A tipologia RE-2 abrange 336 NUI, com a maior renda média dos responsáveis pelos domicílios (R\$558) e densidade domiciliar média inferior à RE-1 (32 domicílios/ha) (Figura 39). Em comparação com as demais tipologias, apresentam melhores condições das vias de circulação: somente 16% dos domicílios estão em vias não pavimentadas, enquanto 31% não têm calçada, 34% não têm identificação de logradouro e 3% não têm iluminação pública.

A tipologia RE-2 apresenta também menor proporção de domicílios sem abastecimento por rede geral de água (10%) e sem coleta de lixo por serviço de limpeza (6%). Por outro lado, possui elevada proporção de domicílios com esgotamento inadequado (43%), com destaque para o uso de fossas rudimentares (41%). A proporção de domicílios com esgoto a céu aberto, entretanto, é mais baixa do que a observada nas demais tipologias do polo (16%).

FIGURA 39. Recife: Tipologia RE-2.



Fonte: Google Satellite, 2021.

4.6.3 Tipologia RE-3

A tipologia RE-3 abrange 390 NUI, localizados um pouco mais afastados (periféricos), ainda que a maioria esteja completamente inserido na malha urbana (63%) (Figura 40). Possui renda média de R\$475 e densidade média de 31 domicílios/ha. Em comparação com as tipologias RE-1 e RE-2, apresenta piores indicadores de infraestrutura e serviços urbanos: 16% dos domicílios não têm acesso à rede de abastecimento de água, 43% possuem esgotamento inadequado e 8% não têm coleta de lixo por serviço de limpeza. Possui, ainda, 24% dos domicílios cujo entorno apresenta esgoto a céu aberto. As condições das vias de circulação também são piores do que as das tipologias RE-1 e RE-2: 28% dos domicílios não estão em vias pavimentadas, 39% sem calçada, 43% sem identificação de logradouro e 4% sem iluminação pública.

FIGURA 40. Recife: Tipologia RE-3.



Fonte: Google Satellite, 2021.

4.6.4 Tipologia RE-4

A tipologia RE-4 é composta por 104 NUI, localizados predominantemente isolados ou na periferia da malha urbana (54%) (Figura 41). É a tipologia com piores condições de infraestrutura do polo, menor densidade domiciliar (17 domicílios/ha) e menor renda média dos responsáveis pelos domicílios (R\$409). Em relação à infraestrutura de saneamento, possuem a maior proporção de domicílios sem atendimento da rede de abastecimento de água (25%), sem esgotamento adequado (48%), sem banheiro (9%) e sem coleta de lixo (16%). É a tipologia com maior proporção de domicílios com esgoto a céu aberto (30%).

As condições das vias de circulação também são piores do que nas demais tipologias: 29% sem pavimentação, 45% sem calçada, 48% sem identificação e 5% sem iluminação pública.

FIGURA 41. Recife: Tipologia RE-4.



Fonte: Google Satellite, 2021.

4.7 Comparação entre Aglomerados Subnormais e Núcleos Urbanos Informais

Análises que envolvem a identificação e delimitação de assentamentos precários frequentemente baseiam-se nos dados dos aglomerados subnormais (AGSN) disponibilizados pelo IBGE. Segundo o IBGE, os aglomerados subnormais são “formas de ocupação irregular de terrenos de propriedade alheia (públicos ou privados) para fins de habitação em áreas urbanas e, em geral, caracterizados por um padrão urbanístico irregular, carência de serviços públicos essenciais e localização em áreas que apresentam restrição à ocupação” (IBGE, 2020, p. 5).

A metodologia de delimitação dos AGSN vem sendo aprimorada no decorrer dos anos e leva em consideração informações das prefeituras municipais (quando disponibilizadas), análises de campo, registros administrativos e bases de dados geoespaciais (IBGE, 2020). Entretanto, diversos autores vêm apontando para a subestimação desses assentamentos nos levantamentos realizados até 2010 (DENALDI, 2013; TASHNER, 2008; PEDRO; QUEIROZ, 2019). Análises preliminares dos AGSN divulgados

recentemente pelo IBGE (2020) foram realizadas para o município de São Paulo e apontaram uma melhora na qualidade dos dados. No entanto, ainda é possível identificar assentamentos precários ausentes na base, entre esses não apenas favelas pequenas (que, por definição metodológica⁸, não são delimitadas pelo IBGE) mas também assentamentos consolidados e de grandes dimensões (FERREIRA; FEITOSA, 2020).

Em que pesem as limitações mencionadas, os dados sobre AGSN estão disponíveis para todo o território nacional, o que os posicionam como uma referência relevante para a identificação de NUI. Considerando a importância desse dado, esta seção compara a delimitação dos AGSN e dos NUI, classificados por tipologia, nos seis polos da pesquisa. Como os NUI identificados pela pesquisa são mais numerosos do que os AGSN, busca-se avaliar se os AGSN podem ser considerados como "amostras" de NUI e, em caso positivo, se representam de forma não-enviesada as distintas tipologias identificadas. Essa análise é relevante para subsidiar o desenvolvimento de uma metodologia para a identificação de NUI, foco do Capítulo 5.

Cabe mencionar que, embora não existam AGSN em todos os municípios abarcados pela pesquisa, os mesmos foram considerados no processo da pesquisa de campo, com maior ou menor rigor, a depender do polo. No Polo Juazeiro do Norte 7 dos 22 municípios contam com algum AGSN e, no Polo Belo Horizonte, 18 dos 29 municípios possuem polígonos de AGSN. Nesses dois polos todos os polígonos de AGSN foram verificados junto aos técnicos municipais. Parte dos AGSN correspondiam em alguma medida a outros polígonos de áreas irregulares e/ou precárias definidas por planos ou políticas locais. Nesses casos, quase sempre foram considerados como NUI os polígonos dos planos e políticas locais, pois na grande maioria das vezes representavam o NUI de forma mais precisa. Alguns polígonos de AGSN não se sobreponham com áreas delimitadas localmente e foram incorporados ao mapeamento de NUI. Outros, entretanto, não correspondiam a Núcleos Urbanos Informais de Interesse Social de acordo com a percepção dos técnicos municipais e pesquisadores de campo e, por conseguinte, não foram incorporados ao mapeamento de NUI.

Nos polos Recife (20 de 30 municípios com AGSN) e Porto Alegre (21 de 45 municípios com AGSN) os AGSN foram analisados em alguns municípios e em outros não. Em municípios grandes, com histórico de planejamento e intervenção em áreas informais e precárias, como Recife/PE, Porto Alegre/RS, Novo Hamburgo/RS e São Leopoldo/RS, considerou-se como NUI os mapeamentos fornecidos pelas prefeituras e os AGSN não foram analisados. Nos demais municípios dos polos Porto Alegre e Recife os AGSN foram analisados e, quando conveniente, incorporados como NUI nos mapeamentos, de acordo com a mesma lógica descrita para os polos Juazeiro do Norte e Belo Horizonte. Vale ressaltar que em alguns municípios do Polo Recife, fora da área metropolitana, todos os AGSN foram considerados como NUI, por ausência de outras sistematizações locais.

No Polo Marabá, 4 dos 17 municípios contam com AGSN. Nos municípios de Marabá e Jacundá os AGSN foram analisados e incorporados aos mapeamentos quando julgou-se que os mesmos correspondiam a NUI. Nos trabalhos de campo conduzidos para os municípios de Parauapebas e Itupiranga os AGSN não foram avaliados. No Polo Brasília, 12 dos 14 municípios possuem delimitação

⁸ Em 2010, foram classificados como aglomerados subnormais exclusivamente setores censitários constituídos de, no mínimo, 51 unidades habitacionais. Em 2019, foi considerada a informação obtida pelas prefeituras municipais, incluindo 2.100 aglomerados que não constituem setores censitários exclusivamente subnormais (possuem menos de 51 domicílios, tendo o IBGE atribuído o número de 30 domicílios a esses casos).

de AGSN, entretanto em nenhum deles avaliou-se, durante o trabalho de campo, se os mesmos correspondiam a NUI ou não.

Considerando os seis polos, a pesquisa identificou um total de 4.968 NUI, ao passo que o IBGE aponta a existência de apenas 1.569 AGSN (IBGE, 2020), o que corrobora estudos anteriores que apontam a subenumeração do dado disponibilizado pelo IBGE (DENALDI, 2013; PASTERNAK, 2008; PEDRO; QUEIROZ, 2019; FERREIRA; FEITOSA, 2020). As Tabelas 9 e 10 apresentam quantos AGSN e NUI, respectivamente, são classificados em cada uma das tipologias apresentadas neste capítulo.

TABELA 9. Aglomerados Subnormais dos polos da pesquisa classificados em tipologias.

Polo	Tipologias				Não classificado	Total
	1	2	3	4		
Belo Horizonte	308	78	27	-	13	426
Brasília	17	62	3	12	45	139
Juazeiro do Norte	8	14	11	-	16	49
Marabá	14	23	2	-	3	42
Porto Alegre	119	150	53	25	46	393
Recife	199	114	153	18	36	520

Fonte: Elaboração própria, 2021.

TABELA 10. Núcleos Urbanos Informais dos polos da pesquisa classificados em tipologias.

Polo	Tipologias				Não classificado	Total
	1	2	3	4		
Belo Horizonte	979	176	252	-	6	1413
Brasília	107	139	24	149	1	420
Juazeiro do Norte	39	99	79	-	2	219
Marabá	83	58	72	-	1	212
Porto Alegre	406	591	263	184	35	1479
Recife	385	336	390	104	10	1225

Fonte: Elaboração própria, 2021.

O Polo Belo Horizonte apresenta um total de 426 AGSN e 1.413 NUI. Dos 426 AGSN, 413 (97%) apresentam sobreposição com NUI, indicando que, em que pese a subenumeração dos AGSN, há coincidência entre as duas bases. Em relação às tipologias, dos 413 AGSN coincidentes com NUI, 74% são da tipologia BH-1, 19% da tipologia BH-2 e 7% da tipologia BH-3. Comparando com a distribuição dos NUI em relação às mesmas tipologias (70% - 12% - 18%), conclui-se que os AGSN podem ser considerados como boa representação amostral dos NUI, embora seja constatada certa sub-

representação da tipologia BH-3, na qual predominam NUI de baixa densidade, localização periférica e com grande carência de infraestrutura.

O Polo Brasília apresenta 139 AGSN e 420 NUI. Entre os AGSN, 95 (68%) coincidem com NUI, uma porcentagem inferior à verificada em Belo Horizonte, embora ainda expressiva. Dos 95 AGSN coincidentes com NUI, 17% são da tipologia BSB-1, 65% da BSB-2, 3% da BSB-3 e 13% da BSB-4. Essa distribuição é bastante distinta da observada nos NUI mapeados pela pesquisa, que possuem 25%, 33%, 6%, e 35% das áreas classificadas como BSB-1, BSB-2, BSB-3 e BSB-4, respectivamente. Nota-se, assim, que os AGSN apresentam, em sua maioria, características compatíveis com a tipologia BSB-2, que abrange áreas precárias inseridas na malha urbana, e que se destacam, para os padrões do polo, por apresentar a menor renda média dos responsáveis pelo domicílios e pela elevada proporção de domicílios sem pavimentação, calçada e sem identificação de logradouros.

O polo de Juazeiro do Norte possui 49 AGSN e 219 NUI. Dos 49 AGSN, 33 (67%) coincidem com NUI, uma porcentagem semelhante à observada em Brasília. Destes 33 AGSN, 24% são da tipologia JN-1, 42% da JN-2 e 33% da JN-3, uma distribuição próxima da observada nos NUI, que é de 18%, 45% e 36% de JN-1, JN-2 e JN-3, respectivamente.

O polo Marabá apresenta 42 AGSN e 214 NUI. Entre os AGSN, 39 (93%) apresentam sobreposições com os NUI e, destes, 36% são da tipologia MA-1, 59% da MA-2 e 5% da MA-3. Essa distribuição, em comparação com a observada nos NUI (39%, 27% e 34% de MA-1, MA-2 e MA-3, respectivamente), revela que os AGSN apresentam características predominantemente compatíveis com os núcleos inseridos na malha urbana, principalmente os mais precários (JN-2), e sub-representam os NUI mais afastados, do tipo distrito ou povoado, que são típicos da tipologia JN-3.

O polo Porto Alegre apresenta 393 AGSN e 1.479 NUI. Entre os AGSN, 347 (88%) coincidem com os NUI e, destes, 34% são da tipologia POA-1, 43% da POA-2, 15% da POA-3 e 7% da POA-4. Essa distribuição é próxima à observada nos NUI, que possuem 28%, 40%, 18% e 13% das áreas classificadas como POA-1, POA-2, POA-3 e POA-4, respectivamente.

O polo de Recife possui 520 AGSN e 1.225 NUI. Entre os AGSN, 484 (93%) apresentam sobreposições com NUI e, destes, 41% são da tipologia RE-1, 24% da RE-2, 32% da RE-3 e 4% da RE-4. Essa distribuição assemelha-se à observada nos NUI, que é de 32%, 28%, 32% e 9% das tipologias RE-1, RE-2, RE-3 e RE-4, respectivamente.

A comparação entre os dados dos AGSN e NUI nos seis polos revela que os AGSN podem representar os NUI na condição de amostra, embora a qualidade dessa representação difira de forma substancial entre os polos, sendo inferior para Juazeiro do Norte e superior para Recife.

Em todos os polos analisados, os AGSN representam amostras dos NUI levantados, sendo maior nos polos de Recife, Belo Horizonte, Porto Alegre e Brasília (40%, 29%, 23% e 23%, respectivamente). Já em Juazeiro do Norte e Marabá esse percentual é menor do que nos demais polos (15% e 18%, respectivamente), o que se torna objeto de atenção em virtude do baixo número de AGSN nessas áreas (apenas 33 e 39 AGSN coincidentes com NUI, respectivamente).

Nos polos de Belo Horizonte, Recife, Marabá e Porto Alegre a totalidade dos AGSN representa uma amostra com pouco ruído (apenas 3%, 7%, 7% e 12% dos AGSN, respectivamente, não coincidem com

NUI), enquanto o mesmo não pode ser afirmado para os polos de Brasília e Juazeiro do Norte (33% e 32% dos AGSN, respectivamente, não coincidem com NUI). Quanto às tipologias, observa-se que os AGSN representam todas as tipologias de forma satisfatória nos polos de Recife, Porto Alegre e Juazeiro do Norte. Nos polos de Belo Horizonte, Brasília, Marabá observou-se sub-representação, com maior ou menor intensidade, de alguma tipologia de NUI, particularmente aquelas relacionadas a NUI afastados da malha urbana e/ou com menor grau de precariedade.

Observa-se, em suma, que os dados dos AGSN disponíveis para centros urbanos maiores, mais densos e consolidados, e com predominância de NUI do tipo favela e loteamentos irregulares tendem a apresentar qualidade superior à de centros menores, mais precários e com maior proporção de NUI isolados (distritos e povoados). Considerando o objetivo dessa análise, que é o de verificar se os AGSN podem ser utilizados como amostras de NUI na construção de modelos que auxiliem na identificação de NUI, tais características dos dados tendem a influenciar a qualidade dos modelos, conforme será confirmado nos experimentos apresentados no próximo capítulo. Nesses experimentos, o modelo estimado para o Polo Porto Alegre apresentou melhores resultados, ao passo que os modelos de Juazeiro do Norte e Marabá apresentaram os piores ajustes.

5. Modelos para a Identificação de NUI

As dificuldades e desafios inerentes à obtenção de informações confiáveis, consistentes e atualizadas sobre núcleos urbanos informais precários vêm sendo retratadas por inúmeros autores (DENALDI, 2013; MARQUES, 2007; MATION; NADALIN; KRAUSE, 2014; SMOLKA; BIDERMAN, 2009), assim como pela presente pesquisa. Buscando contribuir nessa direção, a pesquisa propõe uma metodologia, aqui denominada "Metodologia NUI", que objetiva auxiliar na identificação de núcleos urbanos informais por meio da construção de modelos que integrem dados secundários de naturezas distintas (Capítulo 2) para a geração de superfícies de probabilidade relacionadas à presença de NUI. Os modelos podem ser construídos para outras regiões do país e as superfícies de probabilidade deles resultantes podem ser utilizadas para subsidiar trabalhos de campo e/ou para análise da qualidade de bases de dados existentes (tais como a dos AGSN ou informações fornecidas por prefeituras).

O desenvolvimento da Metodologia NUI partiu de uma revisão de metodologias relacionadas, voltadas à identificação de assentamentos precários no contexto brasileiro. A partir da análise das características dessas metodologias frente aos objetivos da presente pesquisa foram estabelecidas diretrizes para a construção da Metodologia NUI. A Seção 5.1 apresenta os resultados dessa análise. Em seguida, a Seção 5.2 apresenta a Metodologia NUI e os resultados de sua aplicação para os seis polos da pesquisa.

5.1 Metodologias relacionadas e definição de diretrizes

Embora a informalidade e precariedade atinjam uma notável parcela da população brasileira e estejam posicionadas como tema de reconhecida relevância no âmbito acadêmico e das políticas públicas, a disponibilidade e qualidade das informações sobre núcleos urbanos informais ainda é muito limitada. Considerando os resultados da Seção 4.7, que apontam a existência de uma relação entre a presença de AGSN e NUI, dada a sua condição de precariedade, cabe destacar duas metodologias baseadas na construção de modelos estatísticos para a identificação de áreas precárias: a do Centro de Estudos da Metrópole (MARQUES, 2007) e a Metodologia para Identificação e Caracterização de Assentamentos Precários em Regiões Metropolitanas Paulistas - MAPPA (CDHU; UFABC; FEITOSA et al., 2019). Essa seção apresenta uma breve descrição e comparação dessas metodologias, ressaltando suas limitações e potencialidades em relação aos objetivos da presente pesquisa. A partir dessa análise comparativa, foi estabelecido um conjunto de critérios e diretrizes que guiou a construção da Metodologia NUI.

5.1.1. Metodologia CEM e MAPPA

Em 2007, o Centro de Estudos da Metrópole desenvolveu e divulgou, a pedido da Secretaria Nacional de Habitação, uma metodologia pioneira para a identificação de assentamentos precários. A metodologia baseia-se na utilização da técnica de análise discriminante para a identificação de setores censitários com características semelhantes às dos aglomerados subnormais, mas que não haviam sido classificados pelo IBGE como tais (MARQUES, 2007).

A análise discriminante tem como objetivo discriminar categorias através do uso de diferentes variáveis que as descrevem. No contexto da metodologia proposta pelo CEM, significa utilizar variáveis que descrevem as características de setores do tipo AGSN para tentar identificar setores censitários não classificados que possuem as mesmas características, classificando-os assim como assentamentos

precários. A metodologia, aplicada inicialmente para os dados censitários de 2000 para os municípios brasileiros, foi posteriormente aplicada para os municípios da Macrometrópole Paulista a partir dos dados de 2010 (CEM/FUNDAP, 2013).

Outra metodologia relacionada à identificação de áreas precárias é a Metodologia para Identificação e Caracterização de Assentamentos Precários em Regiões Metropolitanas Paulistas (MAPPA), desenvolvida em pesquisa contratada pela Companhia de Desenvolvimento Habitacional e Urbana do Estado de São Paulo (CDHU) e executada pela Universidade Federal do ABC (UFABC). A MAPPA propõe procedimentos para a identificação, classificação e caracterização de distintas tipologias de assentamentos precários a partir da utilização intensiva de técnicas de geoprocessamento, estatística e sensoriamento remoto (CDHU; UFABC; FEITOSA et al., 2019).

O desenvolvimento da MAPPA teve como ponto de partida a definição e caracterização de tipologias de tecidos de assentamentos precários (TECs) a partir da análise de características urbanas e morfológicas de assentamentos precários da Região Metropolitana da Baixada Santista, área caracterizada por grande diversidade de ocupações. As TECs definidas foram: morros, palafitas, áreas úmidas (APPs de rios e córregos ou aterros de mangues e restingas), ocupação desordenada (sem traçado regulador), ocupação ordenada por traçado regulador, e ocupação esparsa ou pouco consolidada.

Para a classificação das TECs, a MAPPA adota modelos de regressão logística para geração de superfícies de probabilidade da presença de cada tipologia de tecido de assentamentos precários. A partir das superfícies de probabilidade, constrói-se uma árvore de classificação das tipologias de assentamentos. A regressão logística é uma técnica estatística multivariada utilizada para a classificação de unidades de análise e que se assemelha à análise discriminante, adotada na metodologia proposta pelo CEM.

Apesar dos avanços metodológicos apresentados pela MAPPA, é incerta sua aplicabilidade para outras regiões do país, dadas as especificidades da classificação tipológica sobre a qual se baseia, bem como pelo fato de alguns dos dados considerados na metodologia estarem disponíveis apenas para o Estado de São Paulo. Ressalta-se, portanto, a necessidade de contribuições metodológicas que, assim como a MAPPA, baseiem-se na **integração de dados de fontes diversas**, mas que apresentem **abrangência nacional, maior possibilidade de generalização e explorem a utilização de dados passíveis de atualização mais frequente**.

5.1.2 Desenvolvimento da Metodologia NUI: Diretrizes

A partir da análise das metodologias desenvolvidas pelo CEM e pela UFABC/CDHU para a classificação de assentamentos precários, é possível realizar uma comparação de diferentes aspectos inerentes a esses estudos, um balanço sobre os avanços metodológicos alcançados, assim como as dificuldades e lacunas metodológicas remanescentes. A partir dessa comparação e buscando conciliar as contribuições das duas metodologias, foram estabelecidos critérios e diretrizes que foram considerados na construção da Metodologia NUI. O Quadro 6 sintetiza a comparação entre as metodologias CEM, MAPPA e NUI.

QUADRO 6. Aspectos comparativos entre as metodologias CEM, MAPPA e NUI.

Aspecto Analisado	Metodologia CEM (MARQUES, 2007)	MAPPA (UFABC/CDHU, 2019)	Metodologia NUI
Objetivo	Identificar assentamentos precários	Identificar assentamentos precários e classificá-los em distintas tipologias de tecido urbano (TECs)	Identificar núcleos urbanos informais
Abrangência geográfica	Território nacional	Território do Estado de São Paulo	Território nacional
Disponibilidade dos dados	Dados abertos para o território nacional	Dados abertos para o Estado de São Paulo	Dados abertos para o território nacional
Tipo de dados	Uso exclusivo de dados censitários	Dados provenientes de fontes e naturezas diversas	Dados provenientes de fontes e naturezas diversas
Atualização dos resultados	Dependente da realização de levantamentos censitários (geralmente atualizados a cada 10 anos)	Dados com diferentes resoluções temporais, incluindo alguns com atualização frequente (por exemplo, sensoriamento remoto e dados abertos de logradouros)	Dados com diferentes resoluções temporais, incluindo alguns com atualização frequente. Inova ao considerar dados georreferenciados do CadÚnico
Unidade espacial de análise	Setor censitário: caráter operacional, que não necessariamente dialoga com a forma urbana; não é adequado para integração de dados	UHCT e célula: tem aderência à forma urbana (UHCT) ou flexibilidade para tal (célula); facilita integração de dados (célula)	Célula: adequada para integração de dados e tem flexibilidade para a representação de NUI
Apresentação dos resultados	Classificação binária (precário ou não)	Resultados apresentados de forma contínua (superfícies de probabilidade) e de forma categórica (6 classes, uma para cada tipologia)	Resultados apresentados de forma contínua (superfícies de probabilidade da presença de NUI) e de forma categórica (após a escolha de um limiar de probabilidade)
Técnica de classificação	Análise discriminante	Análise de regressão logística	Análise de regressão logística

Fonte: Elaboração própria, 2021.

Os primeiros aspectos analisados dizem respeito aos **objetivos** e **abrangência geográfica** das metodologias. A Metodologia CEM busca identificar assentamentos precários, sendo possível de aplicação em todo o território nacional. A MAPPA foi desenvolvida para o Estado de São Paulo e busca não apenas identificar assentamentos precários como também classificá-los em tipologias de tecido urbano. Essas tipologias, embora possam servir de referência para distintas regiões, foram especificadas a partir da realidade de São Paulo. Já a Metodologia NUI, apesar de construída a partir

dos seis polos da pesquisa, é passível de generalização para todo o país. Assim como a metodologia CEM, adota uma abordagem mais abrangente para a identificação de núcleos urbanos informais.

A abrangência geográfica das metodologias tem relação com a **disponibilidade e tipo de dados** considerados no processo de classificação. A Metodologia CEM utiliza apenas dados censitários, que estão disponíveis para todo o Brasil. A MAPPA integra dados de fontes diversas para construção de variáveis representativas de distintos aspectos territoriais: aspectos físico-ambientais e unidades de conservação; malha urbana; população, domicílios e entorno dos domicílios (que incluem variáveis obtidas a partir do Censo Demográfico de 2010) (CDHU e UFABC, 2019a). Quanto à disponibilidade dos dados, a MAPPA utiliza tanto dados disponíveis para todo o país como dados disponíveis apenas para o Estado de São Paulo, visto que o Estado possui uma infraestrutura de dados espaciais mais completa do que a maioria dos estados brasileiros. A Metodologia NUI, assim como a MAPPA, também considera dados de fontes e naturezas diversas. Entretanto, dada sua abrangência, considera somente variáveis que possam ser obtidas a partir de dados disponíveis para todo território nacional.

Os dados utilizados também estão relacionados à **possibilidade de atualização** dos resultados. Todas as metodologias utilizam dados censitários, cujo levantamento deve ocorrer a cada dez anos (atualmente atrasado por conta da pandemia e restrições orçamentárias). Entretanto, ao contrário da metodologia CEM que se baseia exclusivamente em dados censitários, as metodologias MAPPA e NUI avançam ao incluir a utilização de outros dados que podem ser atualizados com mais frequência, como os provenientes de produtos de sensoriamento remoto. A metodologia NUI, em particular, **inova ao utilizar dados georreferenciados do Cadastro Único para Programas Sociais do Governo Federal (CadÚnico)**, que são passíveis de atualização constante.

Os dados considerados estão relacionados ainda a outro aspecto metodológico relevante: a **unidade espacial de análise**. A metodologia CEM adota o setor censitário como unidade espacial de análise. A escolha pelo setor censitário é coerente, visto que a metodologia utiliza apenas variáveis provenientes do questionário básico do censo demográfico, já agregadas por setor censitário. Entretanto, cabe salientar que a delimitação do setor censitário segue princípios operacionais, que não dialogam necessariamente com a forma urbana e, por conseguinte, pode apresentar geometrias muito divergentes das poligonais dos assentamentos precários. Outra limitação do uso do setor censitário como unidade espacial de análise é que ele dificulta a integração de dados de fontes distintas, um aspecto que não era relevante no caso da metodologia CEM.

A MAPPA, por utilizar dados diversos, envolve a problemática da integração de dados e demanda maior atenção na escolha da unidade espacial de análise. A metodologia explora duas unidades espaciais de análise distintas: as Unidades Homogêneas de Uso e Cobertura da Terra - UHCT (SÃO PAULO, 2014) e células de 100x100m. As UHCT são unidades delimitadas por meio de interpretação visual de ortofotos, utilizando critérios morfológicos. Por conseguinte, apresenta a vantagem de dialogar diretamente com a forma urbana e frequentemente coincide com os limites de assentamentos e loteamentos, incluindo os precários. No entanto, assim como o setor censitário, a UHCT não representa uma unidade espacial adequada para a integração de dados diversos. Outra limitação da UHCT diz respeito ao fato de ela estar disponível apenas para o Estado de São Paulo. Já as células podem ser geradas facilmente para qualquer área de estudo e são unidades mais adequadas para a integração de dados. Caso não tenham dimensões muito grandes, as células constituem

unidades espaciais flexíveis, "pixels" que podem indicar a presença de NUI e que, aglomerados, representam também sua abrangência em termos de área.

Dado que a metodologia NUI utiliza dados diversos, que demandam integração, a célula apresenta-se como alternativa mais adequada e foi adotada no desenvolvimento da metodologia, bem como em estudos para a caracterização e construção de tipologias de NUI (Capítulos 3 e 4). Em particular, optou-se por adotar uma grade celular compatível com a grade estatística do IBGE, embora com resolução mais refinada (100x100m), conforme apresentado no Capítulo 2.

Outro aspecto metodológico importante diz respeito à **forma de apresentação dos resultados da classificação**. Na metodologia CEM, os resultados da classificação são apresentados de forma binária, ou seja, os setores censitários são classificados como assentamentos precários ou não. Na MAPPA, os resultados são apresentados de duas formas distintas: contínua, como superfícies que indicam a probabilidade da presença de determinada tipologia de assentamento precário; e categórica, na qual as UHCTs ou células são classificadas como não precárias ou como uma das tipologias de assentamento precário estabelecidas. A metodologia NUI prioriza a representação contínua, por meio de superfícies de probabilidade da presença de NUI. Caso considere importante, o analista pode escolher um limiar de probabilidade para obter uma classificação binária (presença/ausência de NUI). A metodologia NUI propõe procedimentos para a escolha desse limiar.

Relacionada à forma de apresentação dos resultados, está a escolha da **técnica de classificação** adotada na metodologia. Na metodologia CEM optou-se pela análise discriminante, ao passo que a MAPPA adotou a análise de regressão logística. Embora envolvam procedimentos estatísticos distintos, as duas técnicas são análogas quanto ao tipo de resultado produzido, ambas permitem a estimativa de probabilidades e os resultados são de fácil interpretação. Quando as suposições básicas de ambas as técnicas são atendidas, seus resultados classificatórios e preditivos tendem a ser semelhantes. Entretanto, a regressão logística tem a vantagem de ser menos afetada quando as suposições básicas não são satisfeitas, em específico, a suposição referente à distribuição normal das variáveis (HAIR et al., 2009 apud CDHU e UFABC, 2019a, p. 60).

Para a seleção da técnica de classificação a ser detalhada na metodologia NUI, uma série de testes comparando distintas técnicas foram realizados a partir dos dados dos seis polos da pesquisa (Apêndice 3). A partir dos resultados optou-se pela utilização da regressão logística.

5.2. Metodologia para Identificação de NUI

A Metodologia NUI consiste das seguintes etapas: (1) construção e integração de variáveis potencialmente relevantes para a identificação de NUI; (2) construção de modelos de classificação; e (3) geração de superfícies de probabilidade de presença de NUI.

A Etapa 1, descrita no Capítulo 2, envolve a construção de variáveis a partir de dados provenientes de fontes diversas, disponíveis para o território nacional e preferencialmente abertos. Entre as fontes de dados adotadas, estão algumas passíveis de atualização frequente, como as provenientes de sensoriamento remoto e os dados do CadÚnico. Para essa pesquisa foram construídas variáveis para os seis polos analisados, que foram integradas em uma base celular com resolução de 100 metros, compatível com a grade estatística do IBGE.

Para a definição da Etapa 2 foi necessária a seleção de uma técnica de classificação para a identificação de Núcleos Urbanos Informais. Para tanto, três diferentes técnicas de classificação foram analisadas e comparadas sob as mesmas condições, ou seja, considerando o mesmo conjunto de variáveis e as mesmas áreas. Duas dessas técnicas, regressão logística e análise discriminante, respondem diretamente às diretrizes estabelecidas para o desenvolvimento da Metodologia NUI por serem adequadas para a construção de superfícies de probabilidade da presença de NUI e permitirem a fácil interpretação dos parâmetros gerados pelo modelo. Já a terceira técnica, a árvore de decisão (algoritmo C5.0), foi analisada em virtude do destaque que as técnicas de *machine learning* vem recebendo em estudos internacionais de identificação de assentamentos precários (MAHABIR et al., 2018; FRIESEN et al., 2018; RIBEIRO, 2015). Os resultados apontam que a regressão logística é o modelo com maior potencial para identificar núcleos urbanos informais sob as condições analisadas, por apresentar resultados mais inteligíveis (interpretáveis), robustos (maior potencial de concordância) e adequados para apresentação na forma de superfícies de probabilidade. Dessa forma, essa foi a técnica escolhida para a modelagem da probabilidade de ocorrência de NUI.

A construção dos modelos logísticos demanda informações sobre a presença de NUI, ainda que incompletas. Conforme apresentado na seção 4.7, a principal fonte de dados a nível nacional (aglomerados subnormais, ou AGSN, do IBGE) tende a subestimar a presença de assentamentos precários. Isso foi observado em todos os polos de pesquisa, embora existam diferenças substanciais entre eles – enquanto em Recife os AGSN representam uma amostra de 40% dos NUI, em Juazeiro do Norte esse percentual é de 15%. Apesar disso, todas as tipologias de NUI puderam ser identificadas na base de AGSN, o que confirma a possibilidade de utilização dessa base como “amostra” dos Núcleos Urbanos Informais e, portanto, como variável dependente dos modelos logísticos.

Esta seção apresenta modelos de regressão logística construídos a partir de dados dos seis polos da pesquisa, utilizando como variável dependente tanto os aglomerados subnormais do IBGE ($Y=AGSN$) quanto os NUI levantados pela pesquisa ($Y=NUI$). Na prática, deve-se usar como variável dependente a informação mais completa possível. Nesta seção, as duas variáveis foram consideradas com o objetivo de comparar os resultados dos modelos e avaliar o uso da informação dos AGSN em regiões onde não há disponibilidade de dados mais completos sobre NUI.

A Etapa 3 consiste em apresentar os resultados dos modelos logísticos estimados na forma de superfícies de probabilidade de presença de NUI. Caso seja relevante apresentar os resultados de forma binária (presença ou ausência de NUI), pode-se selecionar um limiar de probabilidade para a classificação. Cabe salientar, entretanto, que uma análise prévia para a seleção desse limiar deve ser adotada, conforme será demonstrado nesta Seção.

A Tabela 11 apresenta os parâmetros estimados para os modelos dos seis polos da pesquisa, considerando como variável dependente tanto a presença de AGSN quanto os NUI levantados pela pesquisa de campo. Para cada polo, foram consideradas variáveis significativas em ambos os modelos ($Y=AGSN$ e $Y=NUI$), ou seja, que apresentaram consistência na caracterização dos núcleos urbanos informais precários.

TABELA 11. Parâmetros estimados dos modelos logísticos: Polos Belo Horizonte, Distrito Federal e Juazeiro do Norte, Marabá, Porto Alegre e Recife.

Variáveis	Parâmetros estimados e erro padrão - b(EP)											
	Belo Horizonte		Distrito Federal		Juazeiro do Norte		Marabá		Porto Alegre		Recife	
	(Y=AGSN)	(Y=NUI)	(Y=AGSN)	(Y=NUI)	(Y=AGSN)	(Y=NUI)	(Y=AGSN)	(Y=NUI)	(Y=AGSN)	(Y=NUI)	(Y=AGSN)	(Y=NUI)
Constante	-11,703 (0,057)	-5,962 (0,057)	-6,516 (0,063)	-5,514 (0,063)	-6,212 (0,127)	-3,445 (0,127)	-5,503 (0,078)	-3,156 (0,078)	-9,758 (0,136)	-8,201 (0,136)	-4,401 (0,03)	-2,23 (0,03)
Declividade	0,083 (0,001)	0,047 (0,001)			0,128 (0,004)	0,025 (0,004)	0,105 (0,005)	0,044 (0,005)			0,088 (0,002)	0,067 (0,002)
APP30m	0,972 (0,023)	0,303 (0,023)					1,139 (0,05)	0,498 (0,05)				
AltaTensao					1,491 (0,245)	1,494 (0,245)						
Vias50m	1,513 (0,016)	0,568 (0,016)	1,183 (0,024)	1,281 (0,024)	0,978 (0,03)	0,672 (0,03)	1,453 (0,035)	0,908 (0,035)	0,422 (0,029)	0,928 (0,029)	0,961 (0,024)	0,726 (0,024)
IndiceForma									0,468 (0,013)	0,455 (0,013)	0,634 (0,011)	0,382 (0,011)
DomSIden							0,009 (4e-04)	0,002 (4e-04)				
DomSIIu							0,026 (6e-04)	0,018 (6e-04)				
DomSPav									0,004 (3e-04)	0,004 (3e-04)		
DomSArb	0,01 (2e-04)	0,009 (2e-04)										
DomSMed	0,016 (5e-04)	0,008 (5e-04)							0,022 (6e-04)	0,01 (6e-04)	0,042 (6e-04)	0,009 (6e-04)
DomSEsg	0,008 (4e-04)	0,005 (4e-04)	0,029 (5e-04)	0,007 (5e-04)	0,009 (4e-04)	0,004 (4e-04)			0,011 (6e-04)	0,007 (6e-04)		
DomRedeEsg									0,252 (0,036)	0,248 (0,036)	1,119 (0,02)	0,152 (0,02)
DomCLixAc					0,014 (1e-03)	0,005 (1e-03)			0,013 (6e-04)	0,003 (6e-04)		
DomApto				-0,09 (0,001)	-0,028 (0,001)							

DomImpr			0,085 (0,002)	0,048 (0,002)								
DomAdeq	-0,012 (2e-04)	-0,005 (2e-04)	-0,007 (3e-04)	-0,001 (3e-04)								
DomNBanDom			-0,957 (0,023)	-0,487 (0,023)								
DomNBanHab									-3,291 (0,095)	-1,888 (0,095)		
AguaRede							-0,021 (4e-04)	-0,009 (4e-04)				
LixoQueimado											-0,049 (4e-04)	-0,013 (4e-04)
LixoAterrado					0,274 (0,008)	0,136 (0,008)						
LixoCacamba			0,019 (3e-04)	0,005 (3e-04)								
RenMeioSM					0,015 (0,002)	0,034 (0,002)						
Ren3SM							0,008 (9e-04)	0,016 (9e-04)	0,022 (8e-04)	0,011 (8e-04)		
RenRespMedia					-3e-04 (8e-05)	-0,002 (8e-05)					-0,001 (1e-05)	-4e-04 (1e-05)
NDenDom											6e-06 (4e-06)	1e-04 (4e-06)
NDenPop	2e-04 (2e-06)	6e-05 (2e-06)					4e-05 (4e-06)	8e-05 (4e-06)	5e-06 (2e-06)	4e-05 (2e-06)		
NMoradores	0,209 (0,013)	0,312 (0,013)							1,296 (0,033)	1,155 (0,033)		
t_mort2			0,204 (0,002)	0,201 (0,002)								
t_fectot	2,372 (0,018)	1,462 (0,018)										
P2	0,092 (0,005)	0,131 (0,005)							0,283 (0,007)	0,327 (0,007)		
P10	0,243 (0,014)	0,197 (0,014)							0,163 (0,008)	0,163 (0,008)		

P35											0,184 (0,008)	0,091 (0,008)
P40					0,304 (0,038)	0,357 (0,038)					0,478 (0,019)	0,683 (0,019)
AIC	38146,828	183816,698	40870,028	98554,718	9673,072	37715,379	13050,824	38481,574	41626,124	84717,611	58217,311	110212,826
Cox & Snell R2	0,087	0,13	0,095	0,126	0,027	0,101	0,102	0,096	0,072	0,106	0,148	0,161
Nagelkerke R2	0,332	0,189	0,339	0,252	0,111	0,153	0,269	0,139	0,291	0,255	0,288	0,227

Todos os coeficientes são significativos ao nível de 1%.

Fonte: Elaboração própria, 2021.

Nos modelos estimados para o polo Belo Horizonte, as variáveis físico-territoriais relacionadas à alta declividade (Declividade), presença de Área de Proteção Permanente (APP30) e proximidade a vias carroçáveis (Vias50m) mostraram significância na identificação de NUI. No que se refere às variáveis do CadÚnico, os modelos indicaram que a presença de NUI está relacionada ao total de pessoas em famílias cadastradas (P2) e à presença de edificações com alvenaria sem revestimento (P10). Também apresentaram efeito positivo as variáveis maior proporção de domicílios sem arborização (SArb), sem medidor de uso exclusivo (SMed), com esgoto a céu aberto (SEsg), densidade populacional (NDenPop), média do número de moradores (NMoradores) e taxa de fecundidade (t_fectot). Uma variável apresentou efeito com sinal negativo, ou seja, é uma variável que diminui a probabilidade de presença de NUI: a porcentagem de domicílios com moradia adequada (DomAdeq).

No polo do Distrito Federal, áreas próximas a vias carroçáveis (Vias50m), com esgoto a céu aberto (SEsg), com alta proporção de domicílios improvisados (DomImpr), com lixo coletado em caçamba de serviço de limpeza (LixoCacamba) e com maiores taxas de mortalidade infantil (t_mort2) indicam maior probabilidade de presença de NUI. Os modelos indicam ainda que uma maior proporção de domicílios do tipo apartamento (DomApto), com moradia adequada (DomAdeq) e com maior média do número de banheiros por domicílio (DomNBanDom) diminuem a chance de presença de NUI.

Nos modelos construídos para o polo Juazeiro do Norte, destacaram-se, com maior tamanho de efeito, as variáveis que indicam a precariedade no destino do lixo (LixoAterrado), no caso do modelo Y=AGSN, e a baixa renda dos moradores (RenMeioSM), no modelo Y=NUI. Os modelos indicaram ainda que áreas com alta declividade (Declividade), localizadas em faixas de servidão de linhas de alta tensão (AltaTensao) e próximas a vias carroçáveis (Vias50m) têm maior probabilidade de serem NUI. A presença de famílias cadastradas no CadÚnico também está relacionada à presença de NUI (P40). No que diz respeito à infraestrutura, a presença de NUI está associada à presença de esgoto a céu aberto (DomSEsg) e lixo acumulado nas ruas (DomCLixAc). Uma maior concentração de responsáveis de baixa renda (RenRespMedia) também aumenta a probabilidade da presença de NUI.

No caso do polo Marabá, a variável "domicílios sem iluminação" (DomSIlu) apresentou o maior tamanho de efeito em ambos os modelos (Y=AGSN e Y=NUI), estando associada à presença de NUI. Os modelos também indicaram que apresentam maiores chances de serem NUI as áreas com ausência de rede de água (AguaRede) e identificação de logradouros (DomSlden), elevada densidade populacional (NDenPop), próximas a vias carroçáveis (Vias50m) e com presença de APP (APP30m), altas declividades (Declividade). Assim como em Juazeiro, a predominância de responsáveis pelo domicílio de baixa renda (Ren3SM) aumenta a probabilidade de presença de NUI em Marabá.

Nos modelos estimados para o polo Porto Alegre, a variável que representa a proporção de domicílios sem medidor de eletricidade (DomSMed) destacou-se para ambas variáveis dependentes (AGSN e NUI). Os modelos indicaram ainda que a presença de famílias cadastradas no CadÚnico (P40) e, em particular, aquelas que residem em edificações de alvenaria sem revestimento (P10), está relacionada à presença de NUI. Áreas com alta densidade populacional (NDenPop), forma irregular de quadras ou bolsões de ocupação (IndiceForma), com esgoto a céu aberto (DomSEsg), sem ligação à rede de esgoto ou fossa séptica (DomSRedeEsg), com lixo acumulado nos logradouros (DomCLixAc), sem pavimentação (DomSPav) e próximas de vias carroçáveis (Vias50m) também relacionam-se à maior probabilidade de presença de NUI. Situações relacionadas à alta densidade no domicílio, representadas nos modelos pelas variáveis "média de moradores por domicílio" (NMoradores) e

"número de banheiros por habitante" (DomNBanHab), também aumentam a probabilidade de presença de NUI. Por fim, os modelos indicam ainda a presença de responsáveis com renda de até 3 salários mínimos (Ren3SM) como característica dos NUI.

No caso do polo Recife, as variáveis que apresentaram maior poder explicativo foram as indicadoras da irregularidade da forma das quadras ou bolsões de ocupação (IndiceForma) e da presença de famílias cadastradas no CadÚnico (P40), especialmente aquelas com precariedade no acesso à eletricidade (P35, energia elétrica sem medidor). As áreas com moradores mais pobres (RenRespMedia), de alta densidade de domicílios (NDenDom), alta declividade (Declividade) e próximas a vias carroçáveis (Vias50m) apresentam maior probabilidade de serem NUI. Em relação às variáveis de infraestrutura, destacou-se a relativa à ausência de medidor de eletricidade (DomSMed) e ligação à rede de esgoto ou fossa séptica (DomSRedeEsg). Chamou ainda atenção o coeficiente negativo associado à variável "porcentagem de domicílios com lixo queimado na propriedade" (LixoQueimado), associado a áreas mais rurais, que comumente apresentam piores condições de renda e infraestrutura, mas que não estão necessariamente identificadas como NUI.

A avaliação dos modelos foi realizada com base em duas métricas: o coeficiente de concordância Kappa e a Área Sob a Curva (AUC). A Curva característica de Operação do Receptor (ROC) é uma figura que ilustra a performance de um classificador binário de acordo com diferentes limiares de classificação. Ela é influenciada por duas métricas: a razão de verdadeiros positivos e a razão de falsos positivos. Um resultado de classificação perfeito possui uma taxa de verdadeiro positivo igual a 1 e uma taxa de falso positivo igual a 0, produzindo um ponto no canto superior esquerdo da figura, enquanto um resultado de classificação "aleatório" é coincidente com a linha diagonal e indica a incapacidade de discriminação do modelo. Portanto, quanto maior a área sob a curva (AUC), que varia de 0 a 1, maior a capacidade preditiva do modelo. Para este estudo, o uso da AUC torna-se particularmente interessante pelo fato de tratar-se de uma métrica invariante em escala (não trabalha com valores absolutos e sim com a precisão das classificações) e por medir a qualidade das previsões do modelo independente do limiar de classificação.

Já o coeficiente Kappa é uma métrica desenhada para avaliar a concordância entre dois avaliadores, levando em consideração a probabilidade de a concordância ocorrer ao acaso. Ela é uma métrica muito utilizada para avaliar modelos quando a distribuição de classes é desigual – sendo este o caso da presença de Núcleos Urbanos Informais. Sua fórmula é $Kappa = (O-E)/(1-E)$, onde O é a acurácia observada e E é a acurácia esperada com base nos totais marginais da matriz de confusão⁹. O coeficiente Kappa pode apresentar valores entre -1 e 1, sendo que 0 representa a ausência de concordância entre as classes observadas e previstas, enquanto o valor 1 indica a concordância perfeita entre a previsão do modelo e as classes observadas. Valores negativos indicam que a predição está na direção oposta da verdade, mas raramente ocorrem em modelos preditivos (KUHN, 2013).

Ao contrário da ROC e AUC, o coeficiente de concordância Kappa avalia uma classificação categórica (Presença/Ausência de NUI), sendo o limiar de classificação geralmente utilizado igual a 0,5. Entretanto, em situações em que se tem conhecimento da presença do evento, mas há incertezas sobre a ausência, denominadas na análise de classificação como "pseudo-ausência" (HIJMANS; ELITH, 2017), os limiares relativos à presença/ausência do evento tornam-se muito incertos se definidos

⁹ Matriz de confusão é um recurso utilizado em análise preditiva que consiste em uma tabela que relata o número de falsos positivos, falsos negativos, verdadeiros positivos e verdadeiros negativos. No caso da presente pesquisa, os totais marginais da matriz de confusão indicam o número de células mapeadas como NUI e Ausência-NUI (totais marginais de linha) e o número de células cujos modelos classificaram como NUI e Ausência-NUI (totais marginais de coluna).

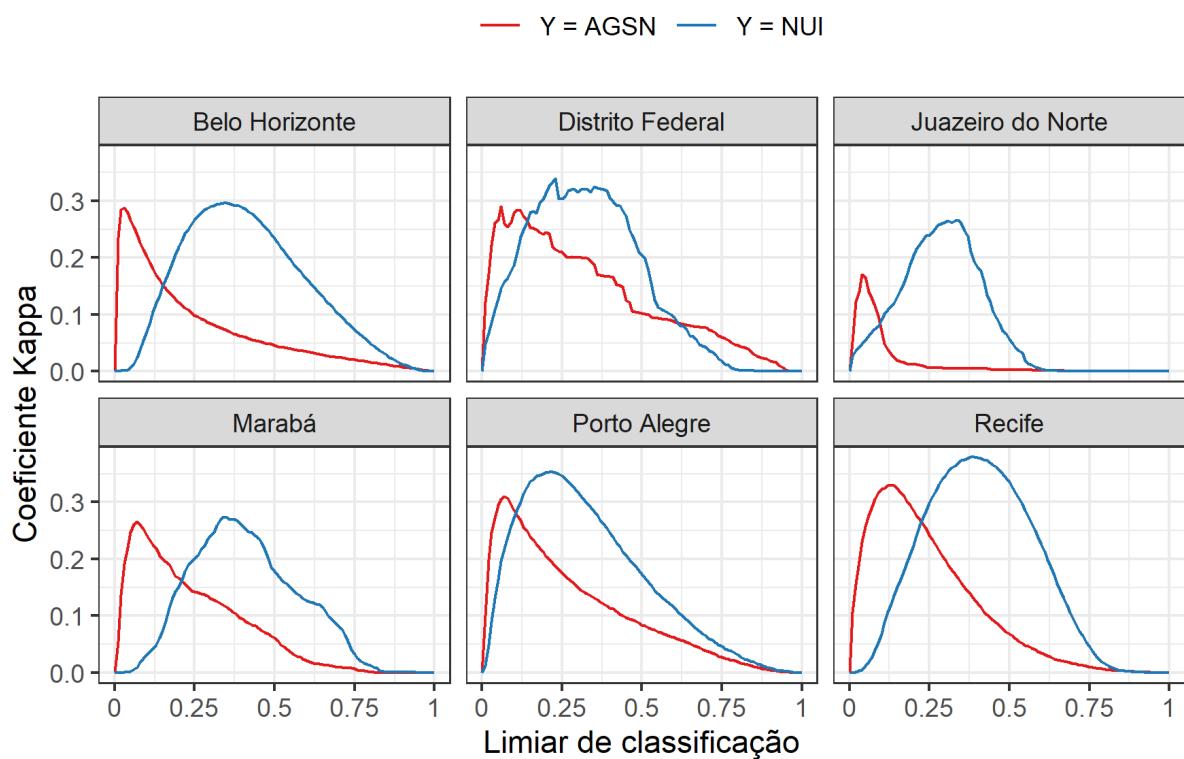
apenas pela base de dados (PHILIPS et al., 2009). Nesses casos, Philips e Elith (2011) recomendam que os limiares de classificação das superfícies de probabilidade sejam complementados por informações externas à base de dados original. Considerando que uma recomendação da Metodologia NUI é apresentar os resultados na forma de superfícies de probabilidade, o coeficiente Kappa foi computado para cada limiar de classificação (Figura 42), sendo reportada a maior concordância encontrada quando consideradas as superfícies de probabilidade. Sabe-se que, em uma situação real, os dados do levantamento em campo não estariam disponíveis para cálculo do Kappa potencial, tendo em vista que a Metodologia NUI almeja fornecer subsídios para a identificação de núcleos urbanos informais a partir de dados incompletos (AGSN), mas a análise das superfícies de probabilidade somada ao conhecimento local e outras fontes de dados permitiria alcançar uma concordância similar à reportada pelo Kappa potencial.

O Kappa foi utilizado neste trabalho por conta de sua popularidade, embora apresente uma série de limitações. Feinstein and Cicchetti (1990) destacam, por exemplo, a alta sensibilidade do coeficiente Kappa aos totais marginais da matriz de confusão. Mesmo com altos valores de acurácia observada (O), baixos valores de Kappa são observados quando os totais marginais são drasticamente desbalanceados, o que é o caso dos experimentos apresentados, dado que a maior parte das células não são classificadas como NUI. Por conseguinte, o principal objetivo do uso do coeficiente Kappa neste estudo é o de identificar limiares de classificação mais adequados, dado que, nesses casos, são comparados resultados cujas condições dos experimentos são as mesmas. Na Figura 42 identifica-se, por exemplo, que os limiares ótimos para a classificação de NUI são sempre inferiores a 0,5, especialmente nos modelos "Y = AGSN".

Considerando a indisponibilidade de dados detalhados sobre NUI, os limiares podem ser estabelecidos a partir da própria informação de AGSN. De todo modo, constata-se que a análise da superfície de probabilidade permite alcançar uma concordância muito maior do que a classificação binária (ausência/presença de NUI, com limiar de classificação = 0,5), o que justifica o uso da superfície de probabilidade.

Na maioria dos polos, a diferença do coeficiente Kappa entre os modelos "Y=AGSN" e "Y=NUI" é muito pequena, indicando que os aglomerados subnormais podem ser uma boa amostra para a identificação de NUI. Essa diferença varia, sendo muito menos nos polos de Belo Horizonte e Marabá e maior no polo Juazeiro do Norte. A diferença constatada no polo de Juazeiro era esperada, dado que a análise apresentada na Seção 4.7 apontava que os dados dos AGSN de Juazeiro do Norte apresentavam limitações maiores para seu uso como amostra de NUI, destacando-se o ruído da amostra (32% dos AGSN não são NUI), combinado com a baixa representatividade (os AGSN representam 15% dos NUI) e o baixo número de AGSN no polo.

FIGURA 42. Coeficiente Kappa dos modelos de acordo com o limiar de classificação.



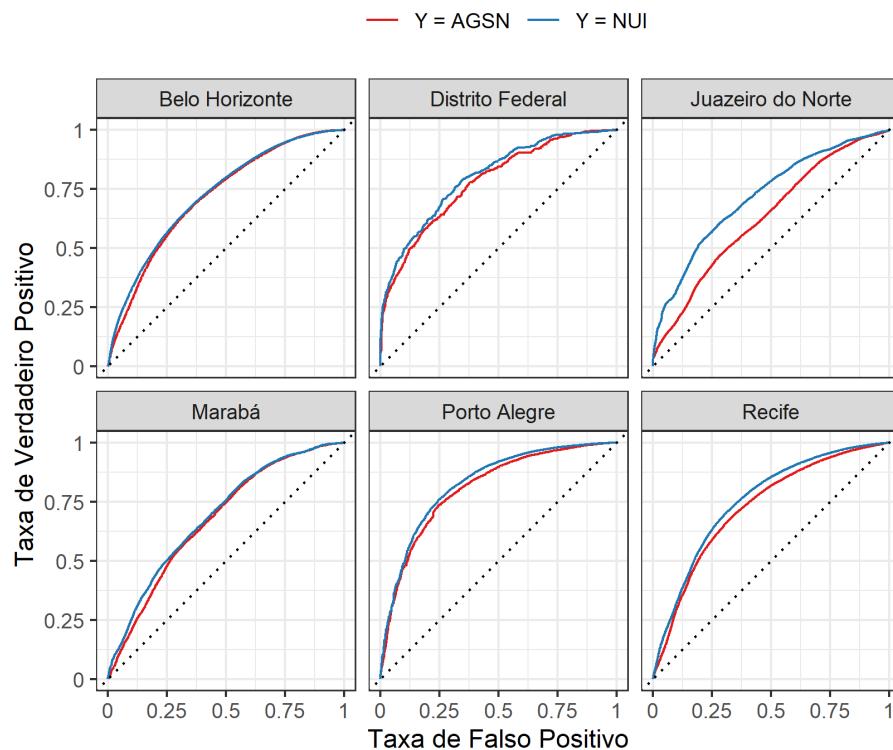
Fonte: Elaboração própria, 2021.

A Figura 43 apresenta as curvas ROC e a Tabela 12 sumariza as métricas AUC e Kappa obtidas para cada um dos modelos estimados. As métricas AUC destacam, novamente, que é pequena a diferença entre os modelos "Y=AGSN" e "Y=NUI" nos polos, com exceção de Juazeiro do Norte. De maneira geral, a AUC dos modelos variou entre 0,63 e 0,83. Os modelos estimados para Porto Alegre apresentaram os melhores resultados, o que pode ser justificado pelo fato de tratar-se de um centro urbano denso, consolidado, com muitas amostras de AGSN/NUI. Também é o polo com melhores condições de infraestrutura e onde as diferenças entre áreas precárias e não-precárias tendem a ser bem definidas.

O modelo construído para Brasília, mesmo o baseado nos AGSN, cuja amostra apresenta muito ruído (33% dos AGSN não coincidem com NUI), apresentou bons resultados em comparação com os demais polos. É possível que isso resulte do fato dos NUI da região apresentarem características bastante distintas das áreas não delimitadas como NUI.

Os modelos estimados para Juazeiro do Norte e Marabá apresentaram os piores resultados. Ambos os polos representam regiões menores, menos densas e com muitos NUI isolados. Apresentam também piores indicadores socioeconômicos e de infraestrutura, que, como caracterizam os polos como um todo, incluindo áreas não delimitadas como NUI, dificultam a distinção dos NUI. O modelo "Y = AGSN" de Juazeiro do Norte destaca-se como o modelo com pior ajuste. Trata-se de um resultado esperado, pois a amostra de AGSN do polo, além de pequena, também apresenta muito ruído (32% dos AGSN não coincidem com NUI).

FIGURA 43. Curva ROC dos modelos



Fonte: Elaboração própria, 2021.

TABELA 12. Métricas obtidas pelos modelos

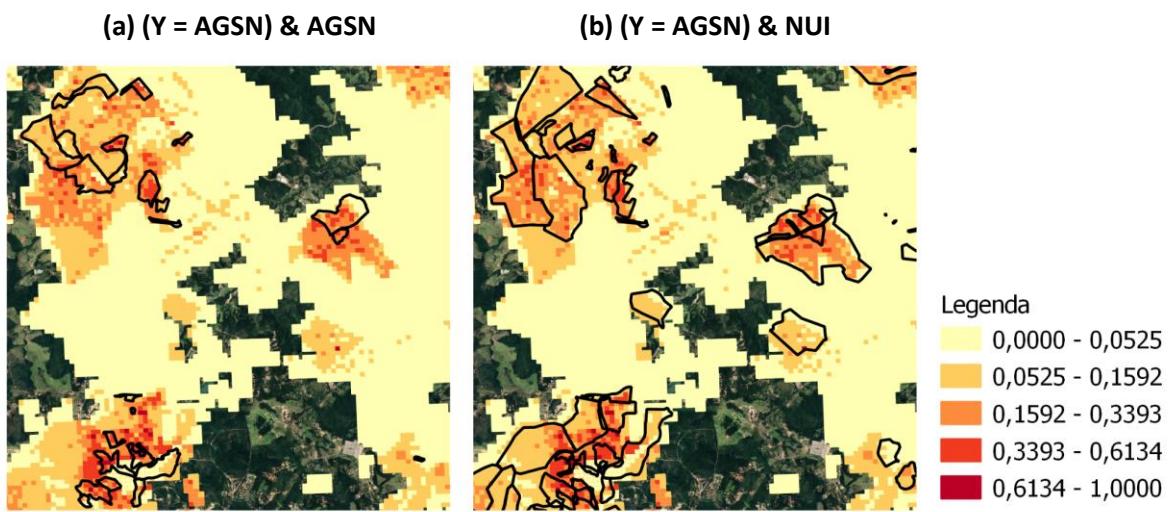
Métrica	Belo Horizonte		Distrito Federal		Juazeiro do Norte	
	(Y=AGSN)	(Y=NUI)	(Y=AGSN)	(Y=NUI)	(Y=AGSN)	(Y=NUI)
AUC	0,717	0,727	0,78	0,803	0,632	0,72
Kappa	0,288	0,296	0,29	0,339	0,171	0,267
Métrica	Marabá		Recife		Porto Alegre	
	(Y=AGSN)	(Y=NUI)	(Y=AGSN)	(Y=NUI)	(Y=AGSN)	(Y=NUI)
AUC	0,676	0,692	0,726	0,755	0,808	0,828
Kappa	0,266	0,274	0,33	0,379	0,31	0,353

Fonte: Elaboração própria, 2021.

Os resultados dos modelos estimados podem ser mapeados e apresentados como superfícies de probabilidade, conforme apresentado nas Figuras 44, 45 e 46. As Figuras revelam áreas indicadas com alta probabilidade de presença de NUI e que não foram originalmente mapeadas pelo IBGE como AGSN, mas que são, de acordo com a pesquisa de campo, núcleos urbanos informais.

A Figura 44 mostra os resultados do modelo "Y = AGSN" de Belo Horizonte apresentado na forma de superfície de probabilidade. A imagem ilustra como a superfície pode auxiliar na identificação de inúmeros NUI não presentes na base de AGSN.

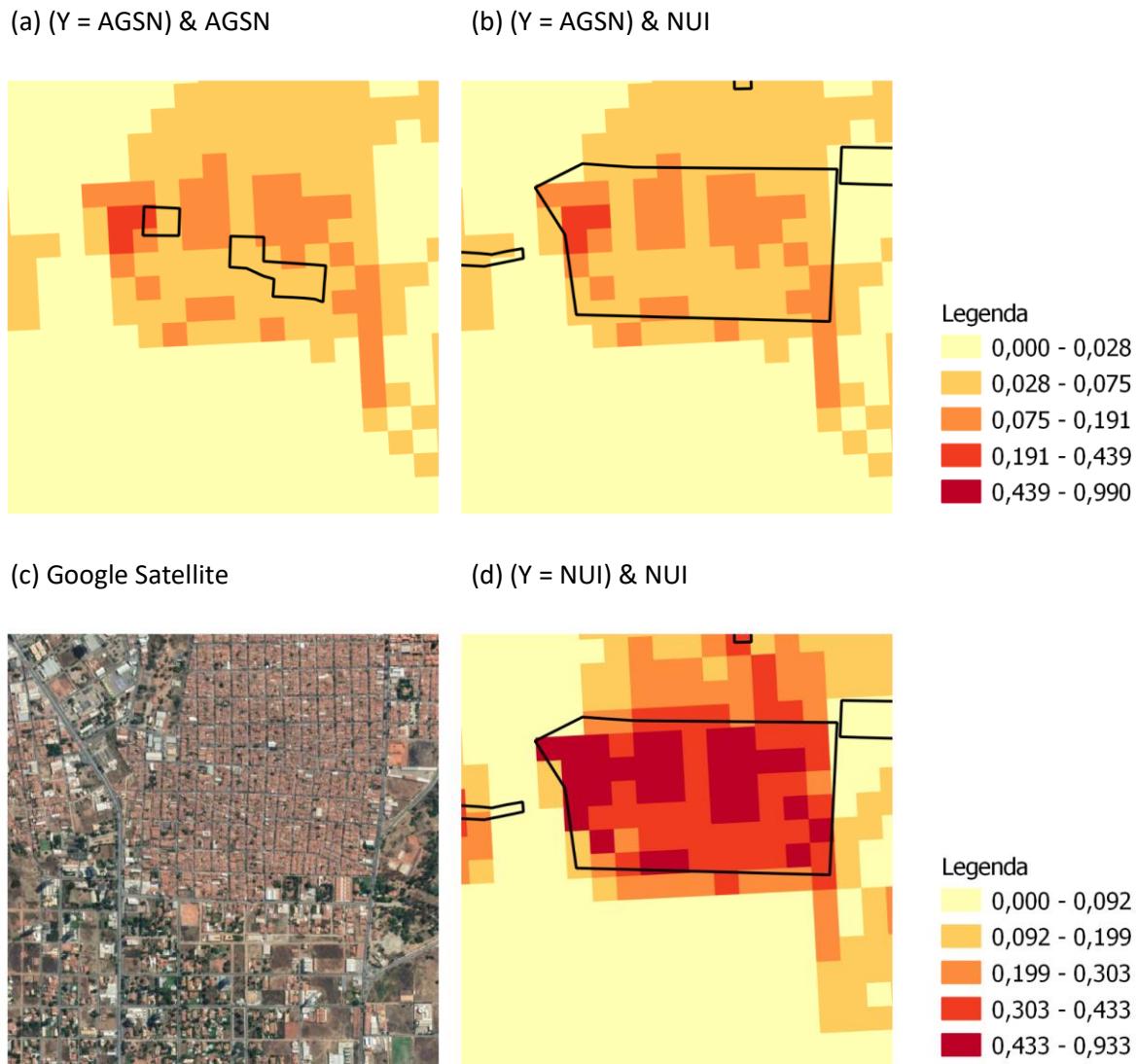
FIGURA 44. Superfícies de probabilidade (modelo "Y=AGSN") do polo Belo Horizonte (municípios Ribeirão das Neves, Contagem e Esmeraldas): (a) AGSN sobre a superfície de probabilidade; (b) NUI sobre a superfície de probabilidade.



Fonte: Elaboração própria, 2021.

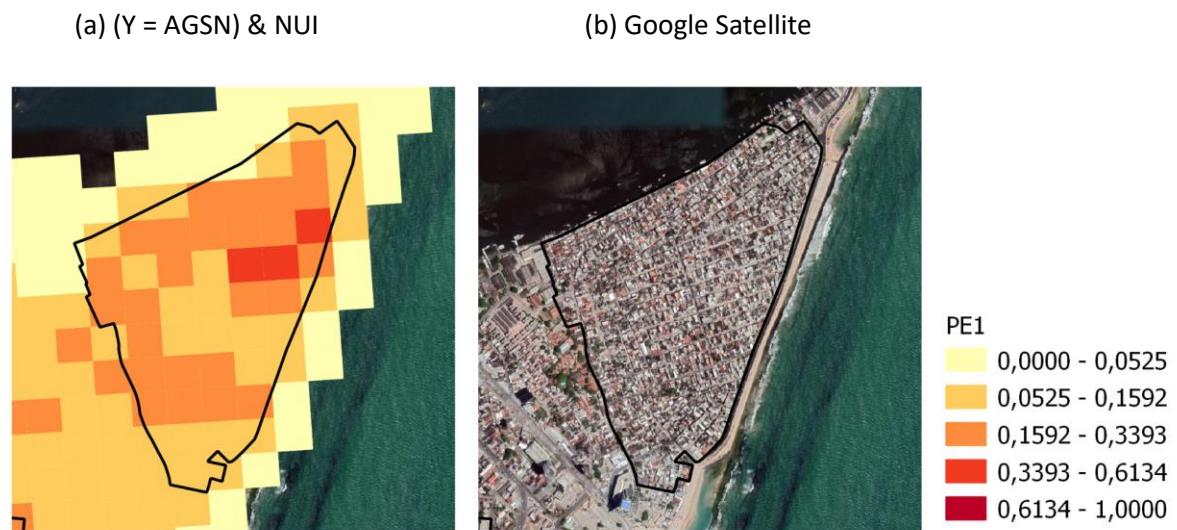
A Figura 45 destaca o bairro João Cabral, em Juazeiro do Norte, onde reside boa parte da população do município. O loteamento, caracterizado pela informalidade, possui calçamento, abastecimento de água e energia elétrica, mas o esgotamento sanitário é parcial. De acordo com o IBGE, apenas algumas porções dessa área são identificadas como AGSN (Figura 45a). Apesar de sua morfologia aproximar-se bastante à da cidade formal (Figura 45c), as superfícies de probabilidade indicam alta probabilidade de presença de NUI (Figura 45b e d).

FIGURA 45. Superfícies de probabilidade do modelo do polo Juazeiro do Norte (Bairro João Cabral): (a) AGSN sobre a superfície de probabilidade (modelo Y=AGSN); (b) NUI sobre a superfície de probabilidade (modelo Y=NUI); (c) Imagem orbital da área; (d) NUI sobre a superfície de probabilidade (modelo Y=NUI)



A Figura 46 apresenta o bairro Brasília Teimosa, situado na zona sul de Recife, que surgiu com a ocupação de uma área antes denominada Areal Novo, iniciada em 1947. Esse NUI, embora ausente na base de AGSN, foi destacado pela superfície de probabilidade estimada a partir desse dado.

FIGURA 46. Superfícies de probabilidade (modelo "Y=AGSN") do polo Brasília (Bairro Brasília Teimosa): (a) NUI sobre a superfície de probabilidade (modelo Y=AGSN); (b) Imagem orbital da área.



6. Considerações Finais

O presente relatório explora a utilização de dados provenientes de fontes diversas para auxiliar na caracterização e identificação de NUI de maneira complementar ao levantamento realizado em campo pela pesquisa. Esses dados foram integrados em uma base comum, que serviu, inicialmente, como referência para aprofundar a caracterização dos NUI apresentada no Produto 11. Por meio dessa análise, foi possível iluminar facetas diversas da precariedade vigente nos NUI de cada polo, com especial ênfase dada a características populacionais, físico-ambientais, forma urbana, infraestrutura de saneamento e sistema viário, bem como condições do entorno. A análise incluiu uma comparação das condições das áreas dentro e fora dos NUI, que permitiu constatar como as diferenças das características dos NUI - com maior destaque para algumas dimensões da precariedade em certos polos do que outros - refletem as condições da região em que estão inseridos.

Foi conduzida também, para cada polo, uma análise voltada à identificação de distintas tipologias de NUI. Ficou evidente, assim, como variam as características dos NUI em um mesmo polo, incluindo, por exemplo, desde aqueles melhor inseridos na malha urbana e relativamente semelhantes ao seu entorno, até aqueles distantes e isolados, menos densos e carentes de infraestrutura e serviços urbanos, mas que por vezes surpreendem quanto às condições relativamente adequadas do entorno dos domicílios e/ou de suas vias de circulação.

Os dados levantados para este relatório também foram utilizados na construção da Metodologia NUI, que objetiva subsidiar levantamentos para a identificação de NUI, bem como análises da qualidade de levantamentos existentes. Entre os aspectos considerados importantes para a construção da metodologia NUI estão a capacidade de ser generalizável para todo o território nacional, o que demanda a utilização de dados disponíveis para todo o Brasil e de procedimentos aplicáveis a diferentes realidades, tais como a estimativa de modelos que possam combinar diferentes variáveis dependendo da área de estudo.

Outras decisões importantes foram tomadas em relação ao tipo de resultado esperado da Metodologia NUI. Para além de classificações binárias, que apontam se determinada área é ou não um NUI, optou-se pelo desenvolvimento de uma metodologia cujos resultados possam ser apresentados de forma contínua, na forma de superfícies de probabilidade da presença de NUI, o que explicita tanto os diferentes níveis de precariedade quanto as incertezas inerentes aos resultados da classificação. Caso o analista prefira uma categorização binária, poderá fazê-lo a partir do estabelecimento de um limiar de probabilidade que considere adequado. Para além dos resultados da modelagem, considerou-se também fundamental a obtenção de modelos de fácil interpretação, que destacam características que predominam nos NUI de cada área analisada.

Os produtos resultantes da aplicação da metodologia NUI deverão servir como um plano de informação complementar às informações existentes sobre NUI (por exemplo, levantamentos municipais e perímetros dos aglomerados subnormais do IBGE). A combinação das superfícies de probabilidade a esses dados deverá dar origem a planos de informação adicionais, que indicam as áreas onde é possível ter maior/menor incerteza sobre a presença/ausência de NUI. Esses produtos deverão configurar uma importante referência para o planejamento de verificações em campo ou remota (por meio de análise de imagens orbitais).

Complementar ao trabalho apresentado no presente relatório, um aspecto desafiador de pesquisas dessa natureza diz respeito à comunicação e divulgação de seus resultados. Por ora, dados, *scripts* e resultados da pesquisa estão organizados em um repositório GitHub. Complementarmente, a equipe do projeto vem trabalhando no desenvolvimento de estratégias que permitam a divulgação do trabalho de forma interativa, por meio de *dashboards* com mapas e gráficos que facilitem tanto a divulgação e acesso aos resultados, como também sua análise e interpretação.

Espera-se, assim, por meio da combinação de diferentes estratégias, que os resultados da pesquisa promovam a produção e disseminação de informações sobre os NUI, que evidenciem a multidimensionalidade e diversidade de condições de irregularidade e precariedade de maneira a subsidiar, de forma efetiva, a elaboração e aprimoramento de políticas e programas habitacionais.

7. Referências

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA (ANEEL). **Shapefile das linhas de alta tensão em território nacional**, 2020.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NACIONAL E COMBUSTÍVEL (ANP). **Shapefile das dutovias em território nacional**, 2020.

AMORE, C. S.; MORETTI, R. “Gelo não é pedra!” – informalidade urbana e alguns aspectos da regularização fundiária de interesse social na lei 13.465/2017. **Cadernos da Defensoria Pública do Estado de São Paulo**, v. 3, n. 17, p. 73–83, 2018.

ANSELIN, L. (1995) Local indicators of spatial association – LISA. **Geographical analysis**, 27(2), p. 93-115. Disponível em: <http://dces.wisc.edu/wp-content/uploads/sites/30/2013/08/W4_An_selin1995.pdf>. Acesso em: 4 jul. 2020.

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**. WorldPop, University of Southampton, UK. doi:10.5258/SOTON/WP00685

BRASIL. **Lei 13.465/2017**, 11 jul. 2017. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2017/Lei/L13465.htm>. Acesso em: 22 jul. 2020.

BRASIL. Ministério do Meio Ambiente. **Cadastro Nacional de Unidades de Conservação**, 2020. BRASIL, 2020a.

BRASIL. Ministério do Desenvolvimento Social. **Cadastro Único**. Brasília: BRASIL, 2019. Disponível em : <<https://www.gov.br/cidadania/pt-br/acoes-e-programas/cadastro-unico>>. Acesso em: 7 abr. 2021.

BRASIL. Ministério do Desenvolvimento Social. **Famílias cadastradas no Cadúnico em 2019**. Brasília: BRASIL, 2020b.

CÂMARA, G.; CARVALHO, M. S.; CRUZ, O. G.; CORREA, V. (2002) Análise espacial de áreas. In: FUKS, S.D.; CARVALHO, M.S.; CÂMARA, G. (Orgs.) **Análise espacial de dados geográficos**. Disponível em: <<http://www.dpi.inpe.br/gilberto/livro/analise/cap5-areas.pdf>>. Acesso em: 4 jul. 2020.

CDHU – COMPANHIA DE DESENVOLVIMENTO HABITACIONAL E URBANO DO ESTADO DE SÃO PAULO; UFABC – UNIVERSIDADE FEDERAL DO ABC. FEITOSA, F. F. (org.) **Desenvolvimento e aplicação de metodologia para identificação, caracterização e dimensionamento de assentamentos precários**. São Bernardo do Campo: UFABC, 2019a. Relatório de Pesquisa. Produto V.

CDHU – COMPANHIA DE DESENVOLVIMENTO HABITACIONAL E URBANO DO ESTADO DE SÃO PAULO; UFABC – UNIVERSIDADE FEDERAL DO ABC. FEITOSA, F. F. (org.) **Sumário Executivo: MAPPA “Metodologia para Identificação e Caracterização de Assentamentos Precários em Regiões**

Metropolitanas Paulistas". São Bernardo do Campo: UFABC, 2019b. E-book - ISBN: 978-65-5040-006-4

CDHU- COMPANHIA DE DESENVOLVIMENTO HABITACIONAL E URBANO; UNIVERSIDADE FEDERAL DO ABC (UFABC); FEITOSA, F. et al. **Desenvolvimento e Aplicação de Metodologia para Identificação, Caracterização e Dimensionamento de Assentamentos Precários.** São Bernardo do Campo: UFABC, 2019.

CENTRO DE ESTUDOS DA METRÓPOLE (CEM); FUNDAÇÃO DE DESENVOLVIMENTO ADMINISTRATIVO (FUNDAP). **Diagnóstico dos Assentamentos Precários nos Municípios da Macrometrópole Paulista.** São Paulo: CEM, 2013.

CONSÓRCIO INTERMUNICIPAL DA REGIÃO DO GRANDE ABC (CIGABC); UNIVERSIDADE FEDERAL DO ABC (UFABC); DENALDI, R. et al. **Diagnóstico Habitacional Regional do ABC.** São Bernardo do Campo: UFABC, 2016. Disponível em: <<http://lepur.com.br/wp-content/uploads/2017/12/03-Diagnostico-Habitaciona I-Regional-do-Grande-ABC-Relat%C3%B3rio-Final.pdf>>. Acesso em: 23 jun. 2020.

DAVIS, M. **Planeta favela.** São Paulo: Boitempo, 2007.

DENALDI, R. (ed.). **Planejamento habitacional: notas sobre a precariedade e terra nos planos locais de habitação.** 1a ed. São Paulo, SP, Brasil: Annablume, 2013.

EMBRAPA, 2017. **Áreas urbanas do Brasil em 2015.** Disponível em:<https://geoinfo.cnpm.embrapa.br/layers/geonode%3Aareas_urbanas_br_15#category-more>.

FEINSTEIN, A. R.; CICCETTI, D.V. High agreement but low kappa: I. the problems of two paradoxes. **Journal of clinical epidemiology**, 43, 6, p. 543–549, 1990.

FEITOSA, F. F.; JACOVINE, T.; ROSEMBACK, R. G. Small area housing deficit estimation: A spatial microsimulation approach. **Revista Brasileira de Cartografia**, v. 68, p. 1157–1169, 2016.

FERRACINI, K. R. Localizando a vulnerabilidade social na dimensão infraestrutura (IVS-I) versus os investimentos do Programa de Aceleração do Crescimento (PAC) nos municípios pernambucanos. **Ciência&Trópico / Fundação Joaquim Nabuco.** -Vol.1, no.1. Recife: Editora Massangana, 2018.

FERREIRA, N. J.; Feitosa, F. F. . Cartografia das favelas: uma análise comparativa. In: **I Seminário Nacional - Urbanismo, Espaço e Tempo. Temática: Cidade, Pandemia e Cotidiano**, 2020, Belo Horizonte. Assentamentos populares informais - antigas demandas, novos riscos. Belo Horizonte: Revista Políticas Públicas & Cidades, 2020.

FIELD, A.; MILES, J.; FIELD, Z. **Discovering statistics using R.** 2012.

FJP – FUNDAÇÃO JOÃO PINHEIRO. **Déficit Habitacional no Brasil 2015.** Belo Horizonte: Fundação João Pinheiro – Centro de Estatística e Informações, 2018. Disponível em: <<http://fjp.mg.gov.br/index.php / docman/direi-2018/871-6-serie-estatistica-e-informacoes- deficit-habitacional-no-brasil- 2015291118 /file>>. Acesso em: 20 jul. 2020.

FOODY, GILES M. (2020). Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. **Remote Sensing of Environment**, 239(), 111630–. doi:10.1016/j.rse.2019.111630

FRIESEN, J.; RAUSCH, L.; PELZ, P.F.; FÜRNKRANZ, J. Determining Factors for Slum Growth with Predictive Data Mining Methods. **Urban Sci.** 2018, 2, 81. <https://doi.org/10.3390/urbansci2030081>

FRIZZI, G. **Índice de acesso à moradia adequada como instrumento para o planejamento territorial: elaboração metodológica e análise.** 2019. 228 p. Dissertação (Mestrado) - Programa de Planejamento e Gestão do Território, Universidade Federal do ABC, São Bernardo do Campo, 2019.

FRIZZI, G.; PINHO, C. M. D. Índice de vulnerabilidade socioecológica para avaliação das remoções na Cidade de São Paulo. In: **XVII Encontro Nacional da ANPUR.** Desenvolvimento, crise e resistência: Quais os caminhos do Planejamento Urbano e Regional? 2017, São Paulo. Anais do XVII ENANPUR, 2017.

FUNDAÇÃO BRASILEIRA PARA O DESENVOLVIMENTO SUSTENTÁVEL (FBDS). **Dados de hidrografia,** 2015.

FUNDAÇÃO JOÃO PINHEIRO - FJP. **Déficit habitacional municipal no Brasil 2010.** Belo Horizonte: FJP, 2013.

FUNDAÇÃO JOÃO PINHEIRO - FJP. **Déficit habitacional no Brasil 2015.** Belo Horizonte: FJP, 2018.

GARETH, J.; WITTEN, D.; HASTIE, T. e TIBSHIRANI, R. 2014. **An Introduction to Statistical Learning: With Applications in R.** Springer Publishing Company, Incorporated. 2014.

HIJMANS, R. J.; ELITH, J. (2017). **Species distribution modeling with R.** CRAN. <https://cran.r-project.org/web/packages/dismo/vignettes/sdm.pdf>

HILBE, J. M. **Logistic regression models.** CRC press, 2009.

HIRSCH, R. P. (2016). **Introduction to Biostatistical Applications in Health Research with Microsoft Office Excel.** [S.I.]: Wiley. p. 150. 392 páginas.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Áreas urbanizadas do Brasil em 2015.** Rio de Janeiro: IBGE, 2017. Disponível em <<https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2100639>>.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Aglomerados Subnormais 2019:** Classificação preliminar e informações de saúde para o enfrentamento à COVID-19. Rio de Janeiro: IBGE, 2020. Disponível em: <<https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2101717>>. Acesso em: 23 jun. 2020.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Censo Demográfico: Brasil, 2010.** Rio de Janeiro: IBGE, 2010

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Grade Estatística**. Rio de Janeiro: IBGE, 2016.

IPEA. **Atlas da vulnerabilidade social nos municípios brasileiros**. Ed. COSTA, M. A. MARGUTTI, B. O. Brasília: IPEA, 2015

IPEA. **Dados IVS**. Disponível: http://ivs.ipea.gov.br/images/publicacoes/lvs/publicacao_atlas_ivs.pdf. Acesso em: 10 ago. 2020.

IPEA. **Métodos e conceitos para o cálculo do índice de vulnerabilidade social com base nas PNADS e desagregações**. Brasília, 2018.

IPEA. **Territórios em números: insumos para políticas públicas a partir da análise do IDHM e do IVS de municípios e Unidades da Federação brasileira**, livro 1 / organizadores: Bárbara Oliveira Marguti, Marco Aurélio Costa, Carlos Vinícius da Silva Pinto. – Brasília: IPEA: INCT, 2017.

JAXA - JAPAN AEROSPACE EXPLORATION AGENCY. **ALOS PALSAR digital elevation model data. 2011**. Disponível em: <<https://bit.ly/2Eny5Vv>>.

JOHNSON, R.; WICHERN, D. **Applied Multivariate Statistical Analysis**. New Jersey: Prentice-Hall; 1995.

KOWARICK, L. **Viver em risco: sobre a vulnerabilidade socioeconômica e civil**. 1a ed. São Paulo, SP: Editora 34, 2009.

KUHN, M., et al. **Applied predictive modeling**. New York: Springer, 2013.

LUZARDO, A. J. R.; CASTAÑEDA FILHO, R. M.; RUBIM, I. B. Análise espacial exploratória com o emprego do índice de Moran. **GEOgraphia**, Niterói – RJ, v. 19, n. 40, 2017.

MAHABIR, R.; AGOURIS, P.; STEFANIDIS, A.; CROITORU, A.; CROOKS, A. T. (2018): Detecting and mapping slums using open data: a case study in Kenya, **International Journal of Digital Earth**, DOI: 10.1080/17538947.2018.1554010

MARICATO, E. Informalidade urbana no Brasil: a lógica da cidade fraturada. In: WANDERLEY, L. E. W.; RAICHELIS, R. (eds.). **A cidade de São Paulo: relações internacionais e gestão pública**. São Paulo: EDUC: Neplam, 2009. p. 269–294.

MARICATO, E. **O impasse da política urbana no Brasil**. Petrópolis: Editora Vozes, 2011.

MARQUES, E. (ed.). **Assentamentos precários no Brasil urbano**. São Paulo/Brasília: Centro de Estudos da Metrópole / CEBRAP ; Secretaria Nacional de Habitação / Ministério das Cidades, 2007.

MATION, L. F.; NADALIN, V. G.; KRAUSE, C. **Favelização no Brasil entre 2000 e 2010: Resultados de uma classificação comparável**. Brasília: IPEA, 2014.

MORAIS, M. P.; KRAUSE, C.; LIMA NETO, V. C. (eds.). **Caracterização e tipologia de assentamentos precários: estudos de caso brasileiros**. Brasília: IPEA, 2016.

NASA, National Aeronautics and Space Administration. **U.S. Releases Enhanced Shuttle Land Elevation Data**. Disponível em <<https://www2.jpl.nasa.gov/srtm/>> Acesso em set. de 2020.

OpenStreetMap©, contribuidores do OpenStreetMap, 2020.

PASTERNAK, S. O desafio da mensuração. In: BRASIL; MINISTÉRIO DAS CIDADES. **Política Habitacional e Integração Urbana de Assentamentos Precários**: Parâmetros conceituais, técnicos e metodológicos. Brasília: Ministério das Cidades, 2008. p. 93–110.

PHILIPS, S. J.; DUDÍK, M.; ELITH, J.; GRAHAM, C. H.; LEHMANN, A.; LEATHWICK, J.; FERRIER, S. (2009). Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. **Ecological Applications**, 19(1), 181–197, 2009. Disponível em: <https://doi.org/10.1890/07-2153.1>

PHILIPS, S. J.; ELITH, J. **Logistic methods for resource selection functions and presence-only species distribution models**. AAAI (Association for the Advancement of Artificial Intelligence), 2011.

PNUD, IPEA, FJP. **Índice de Desenvolvimento Humano Municipal Brasileiro**. – Brasília: PNUD, Ipea, FJP, 2013. Disponível em: <https://www.ipea.gov.br/portal/images/stories/PDFs/130729_Atlas_PNUD_2013.pdf>. Acesso em: 4 set. 2020.

PNUD, IPEA, FJP. **Desenvolvimento humano nas macrorregiões brasileiras: 2016**. – Brasília: PNUD, IPEA, FJP, 2016. Disponível em:<http://www.atlasbrasil.org.br/2013/data/rawData/19FEV_IDHM_WEB.pdf>. Acesso em: 4 set. 2020.

QGIS - Quantum GIS. **Software QGIS** / versão 3.4. Disponível em: <https://www.qgis.org/pt_BR/site/forusers/index.html>.

QUINLAN, J. R. Induction of decision trees. **Machine learning**, 1(1), p. 81-106, 1986.

QUINLAN, J. Ross. **C4. 5: programs for machine learning**. Elsevier, 1993.

RIBEIRO, B. M. G. (2015). [IEEE 2015 Joint Urban Remote Sensing Event (JURSE) - Lausanne, Switzerland (2015.3.30-2015.4.1)] **2015 Joint Urban Remote Sensing Event (JURSE)** - Mapping informal settlements using WorldView-2 imagery and C4.5 decision tree classifier. , (), 1–4. doi:10.1109/jurse.2015.7120470

SAGA-GIS. **Module Library Documentation v2.2.1**. Disponível em: <http://www.sagagis.org/saga_tool_doc/2.2.1/shapes_polygons.html>.

SANTOS, M. **O espaço do cidadão**. São Paulo, SP, Brasil: EDUSP, 2007.

SÃO PAULO (Estado). Empresa Paulista de Planejamento Metropolitano SA (EMPLASA). **Sistema de Informações Metropolitanas do Estado de São Paulo (SIM)**, s.d. Disponível em: <http://www.sim.emplasa.sp.gov.br/mapa>. Acesso em: 15 nov. 2018.

SÃO PAULO (Estado). Instituto Geológico / Coordenadoria de Planejamento Ambiental. **Unidades Homogêneas de Uso e Ocupação do Solo Urbano (UHCT) do Estado de São Paulo**, 2014.

SMOLKA, M. O.; BIDERMAN, C. Measuring informality in housing settlements: Why Bother? **Land Lines**, v. 21, n. 2, p. 14–19, 2009.

SOUZA, M.C.P.; ROSSI, M.T.B.; RUDGE, M.S. Mapeamento colaborativo de assentamentos precários em regiões metropolitanas paulistas. In: **III Seminário sobre Urbanização de Favelas**, Salvador, BA, Brasil, 21 a 23 de novembro de 2018. Anais..., Salvador: Universidade Católica de Salvador, 2018.

STEHMAN, S. V. (1997). **Selecting and interpreting measures of thematic classification accuracy**, 62(1), 77–89. doi:10.1016/s0034-4257(97)00083-7.

UNITED NATIONS – UN Committee on Economic, Social and Cultural Rights (CESCR). **General Comment No. 4: The Right to Adequate Housing** (Art. 11 (1) of the Covenant, 1991).

WHO – WORLD HEALTH ORGANIZATION. **Health principles of housing. Report of WHO Consultation**. Geneva: 1989. Disponível em: <https://apps.who.int/iris/bitstream/handle/10665/39847/9241561270_eng.pdf?sequence=1&isAllowed=y> Acesso em: 08 jul. 2020.

WU, X. ET AL. Top 10 algorithms in data mining. **Knowl Inf Syst**, 14(1), 1–37. 2008. doi:10.1007/s10115-007-0114-2

ZEVENBERGEN, L.W., THORNE, C.R. (1987). Quantitative analysis of land surface topography. **Earth Surface Processes and Landforms**, 12: 47-56.