

População, Espaço e Ambiente

Abordagens Espaciais em Estudos de População: Métodos Analíticos e Técnicas de Representação

Parte IV – Integração de Dados

De áreas a superfícies

Silvana Amaral

Antonio Miguel V. Monteiro

{silvana@dpi.inpe.br, miguel@dpi.inpe.br}



Modelos de População “Multivariados”

- Interpoladores Inteligentes:

- *Smarter SIM*

- uso de variáveis de posição, densidade e distância como indicadoras de população
 - uso de redes neurais para mapear as variáveis preditoras de entrada nas variáveis de interesse

- *Clever SIM – Smart SIM* melhorado:

- Pré-processamento mais sofisticado das variáveis espaciais preditoras
 - *Bootstrap* no treinamento- melhor resultado volta para o treinamento, recursivo
 - Treinamento: zero, alta e baixa densidade e combinando-as por inferência fuzzy.

TO READ....

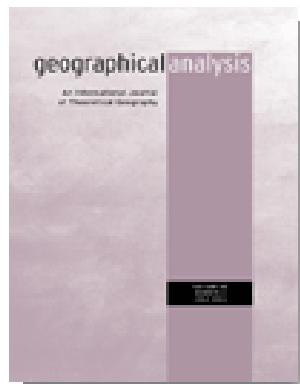


geographical analysis

Geographical Analysis



© The Ohio State University



Special Issue: Areal Interpolation and Dasymetric Modeling

July 2013

Volume 45, Issue 3

Pages 213–344

[Previous Issue](#) | [Next Issue](#)

An Evaluation of Small Area Population Estimation Techniques Using Open Access Ancillary Data

Mitchel Langford

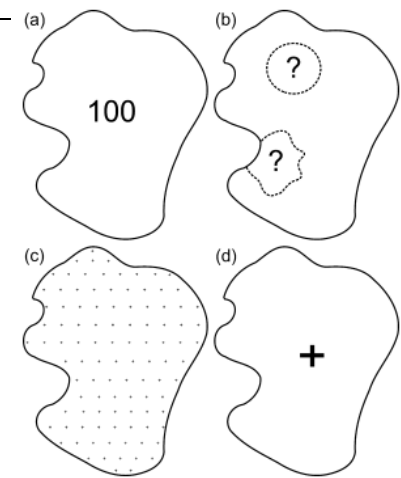
GIS Research Unit, Faculty of Advanced Technology, University of Glamorgan, Pontypridd, Wales, U.K.

Simple areal interpolation.

Aggregated data are available for source zone (a) but required for target zones (b).

Areal weighting distributes population uniformly across source zone (c), while a centroid places all population at a representative point (d).

Either distribution model may be used to provide estimates for target zones.



This study adds to our understanding of the relative merits of alternative methods by comparing dasymetric, street network, and surface-based models interpolating across two spatial resolutions.

It examines the importance of the ancillary data source used to drive the process, particularly the efficacy of open access products.

Results from an empirical study show that interpolation accuracy is influenced by the choice of ancillary data input as well as the methodological approach adopted.

The strongest overall performance is delivered by dasymetric mapping combined with openaccess data identifying the locations of buildings.

An Evaluation of Small Area Population Estimation Techniques Using Open Access Ancillary Data

Mitchel Langford

GIS Research Unit, Faculty of Advanced Technology, University of Glamorgan, Pontypridd, Wales, U.K.

Mitchel Langford

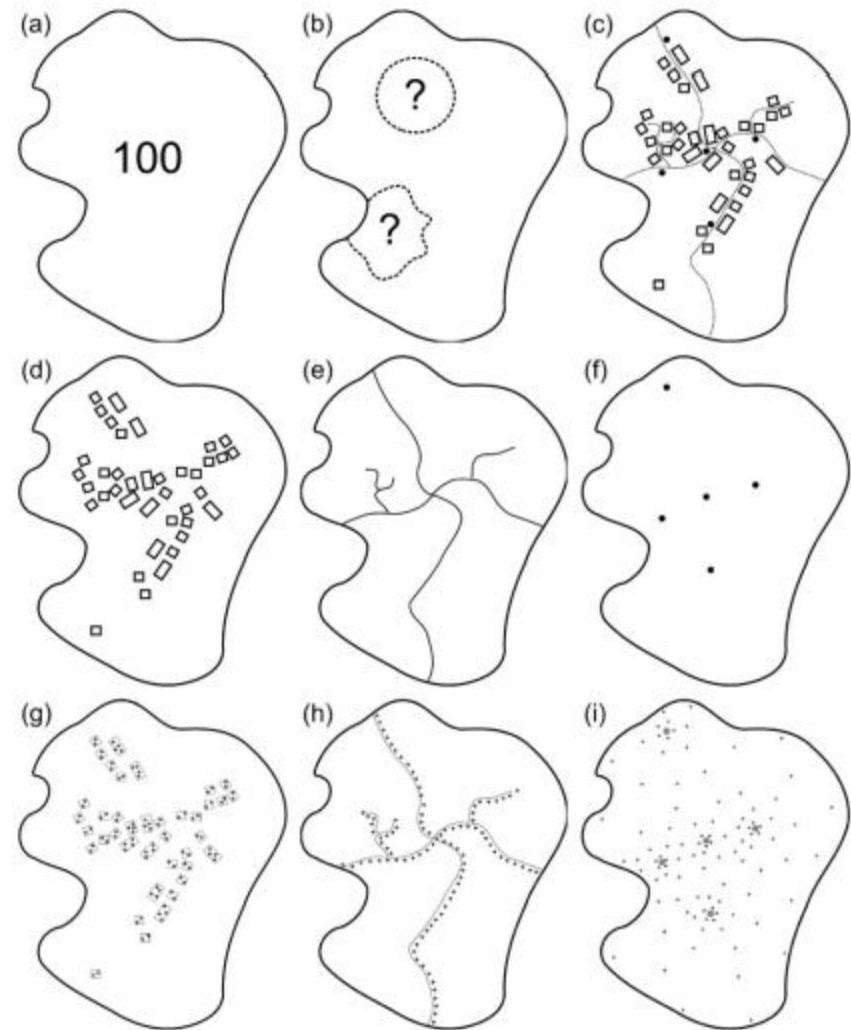
Small Area Population Estimation Techniques

Schematic example of intelligent areal interpolation.

Population aggregated for source zone (a) are required for target zones (b).

Ancillary data residing in a GIS (c) can be used to model the internal distribution of source zone population.

For example, the spatial distribution of buildings (d), of roads (e), or bus stops (f) can be used to yield intelligent estimates of source zone population distribution, shown in (g), (h), and (i), respectively.



(b) Área – Cardiff (UK)
Pop = 305.353 (2001)

(a) Escolas primárias

(c) Áreas residenciais: Landsat ETM+

(d) Polígonos das construções

(e) ruas

(f) pontos de ônibus

-OA – menor unidade 125 casas
(~300 pessoas)

-LSOA – agregados de AO
(~1500 pessoas)

-Constrói com LSOA e testa com
OA; Constrói com OA e testa
com UPC (unidade menor ainda)

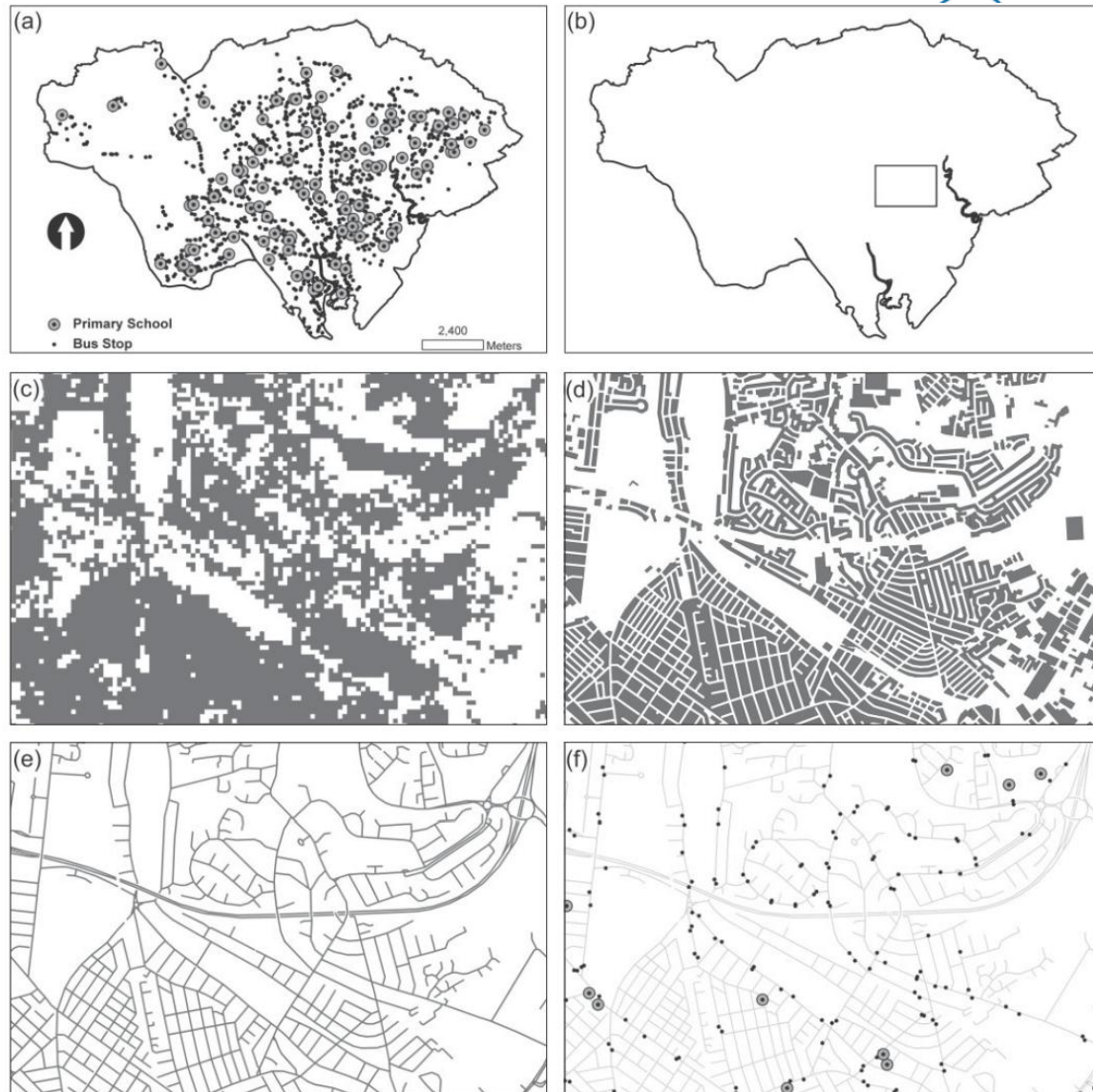


Figure 3. Ancillary data used in the intelligent areal interpolation models: (a) primary school and bus stop points within the study area; (b) location of the detail maps; (c) residential land cover from satellite imagery; (d) building polygons from OS VectorMap District; (e) street vectors from OS VectorMap District; (f) bus stops and primary schools.

Exemplo – superfícies

Pop Distr/model:

- a) LSAO com OA limites
- b) Ponderado pela área
- c) Densidades maiores próximo a escola e pto ônibus
- d) ETM – valor único, mas apenas para células populadas
- e) Polígonos dos distritos
- f) Baseado nos vetores - ruas

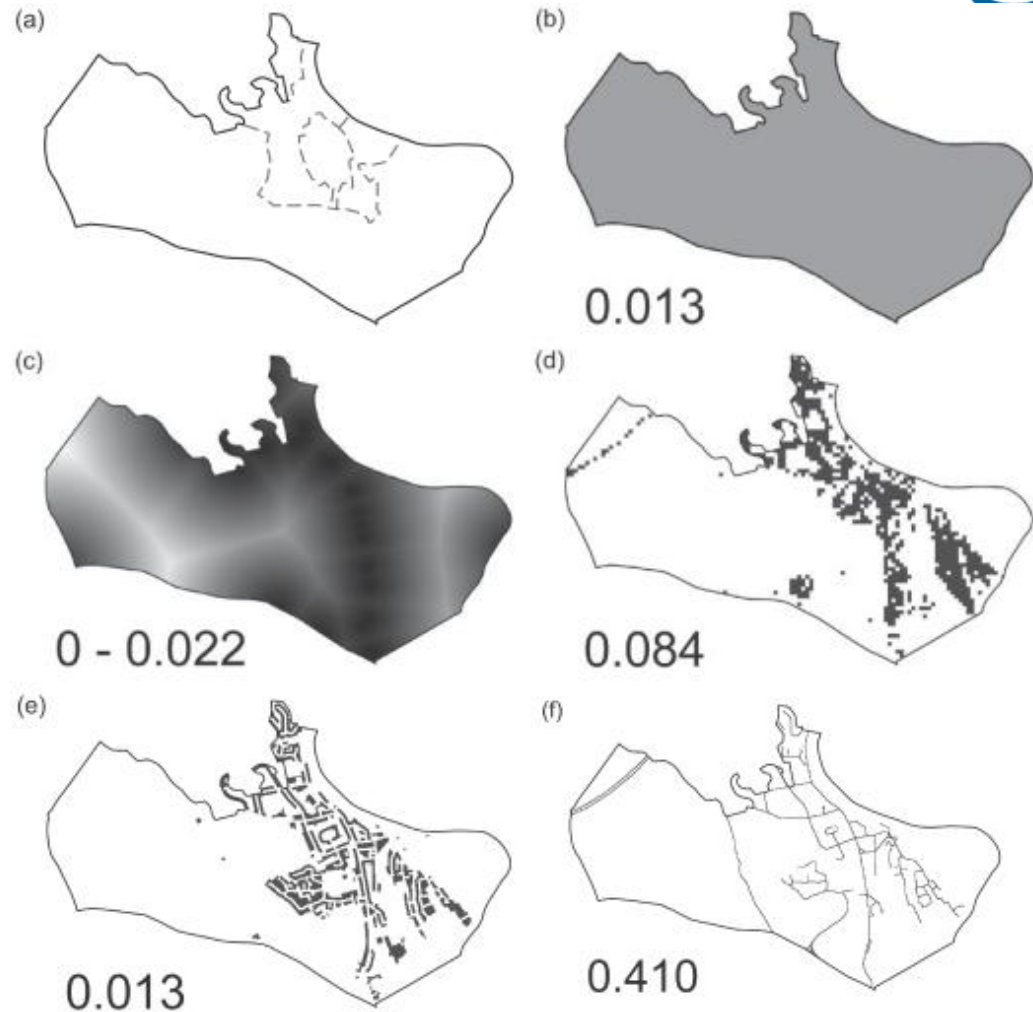


Figure 4. Population distribution modeling in action: (a) a selected LSAO zone and its internal OA boundaries; (b) areal weighting assigns a single low density of 0.013 to every cell; (c) the surface model assigns varying densities, 0–0.022, among all cells, and binary dasymetric models assign a single density but to populated cells only, (d) 0.084 for classified ETM+ data and (e) 0.130 for OS VectorMap District data; (f) street weighting assigns a high density of 0.410 to cells defining the road network.

RESULTADOS

Table 2 Interpolation Results from LSOA to OA Zones

Interpolation method	RMSE	CoV
Areal weighting using zone boundaries only	225.2	0.731
Population density surface using primary schools	220.4	0.715
Population density surface using bus stops	217.5	0.706
Street weighting using OS VectorMap District roads (vector-mode processing)	180.9	0.587
Street weighting using OS VectorMap District roads (raster-mode processing)	178.6	0.580
Street weighting using OS VectorMap District roads with 15-m buffer (vector-mode processing)	170.5	0.553
Binary dasymetric using OS VectorMap District buildings (vector-mode processing)	152.9	0.496
Binary dasymetric using OS VectorMap District buildings (raster-mode processing)	152.9	0.496
Binary dasymetric using classified land cover from Landsat ETM+	144.9	0.471

Table 3 Interpolation Results from OA to UPC Zones*Note:* Mean population of target units is 308

Interpolation method	RMSE	CoV
Street weighting using OS VectorMap District roads (vector-mode processing)	39.3	0.745
Street weighting using OS VectorMap District roads (raster-mode processing)	35.5	0.674
Population density surface using primary schools	35.4	0.672
Population density surface using bus stops	33.8	0.642
Areal weighting using zonal boundaries only	30.3	0.575
Street weighting using OS VectorMap District roads with 15-m buffer (vector-mode processing)	30.3	0.574
Binary dasymetric using classified land cover from Landsat ETM+	23.8	0.451
Binary dasymetric using OS VectorMap District buildings (vector-mode processing)	21.6	0.410
Binary dasymetric using OS VectorMap District buildings (raster-mode processing)	21.6	0.410

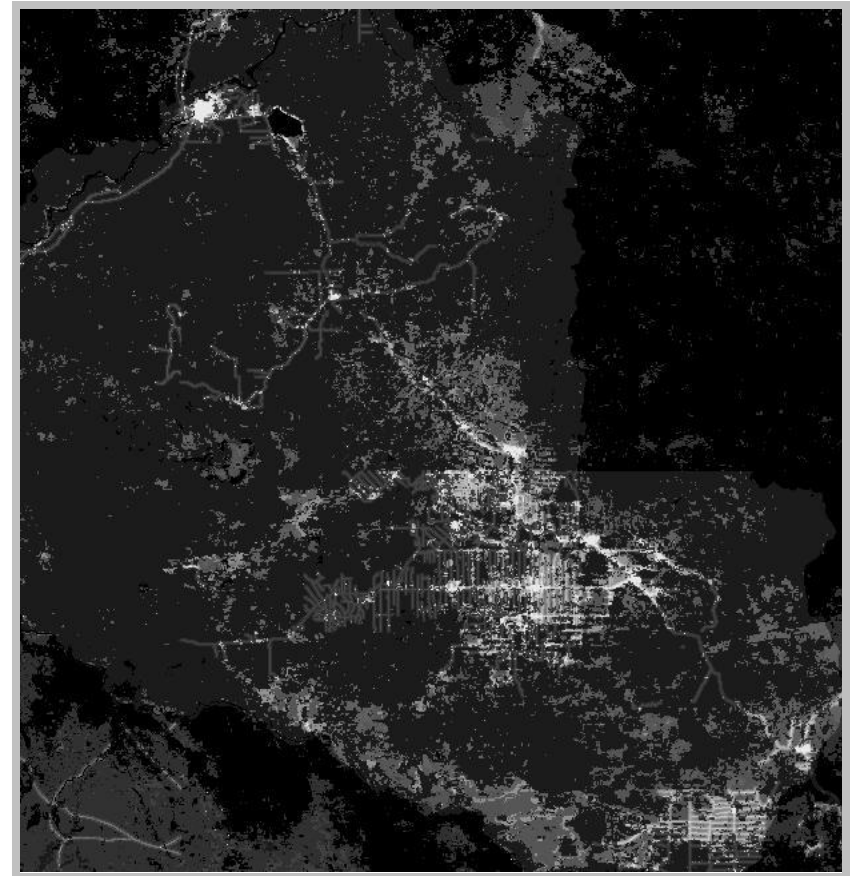
Note: Mean population of target units is 53.

Comentários:

- Dados auxiliares – importância da escolha e qualidade → impactos nos resultados;
- *The underlying rationale and theoretical assumptions on which any particular technique is based must be carefully considered because models that perform well in one environment may prove to be much less successful in another.*
 - Modelo baseado nos vetores de rua – foi ótimo no Texas e pobre em Cardiff
- Baseado em RS – grande potencial pela facilidade de acesso e processamento das imgs
- Todo método de interpolação tem erros e performance variando com condições específicas – preciso estar atento às restrições de cada método.
- Este artigo – apenas área urbana; áreas mais complexas, maior o desafio

Modelos de População “Multivariados”

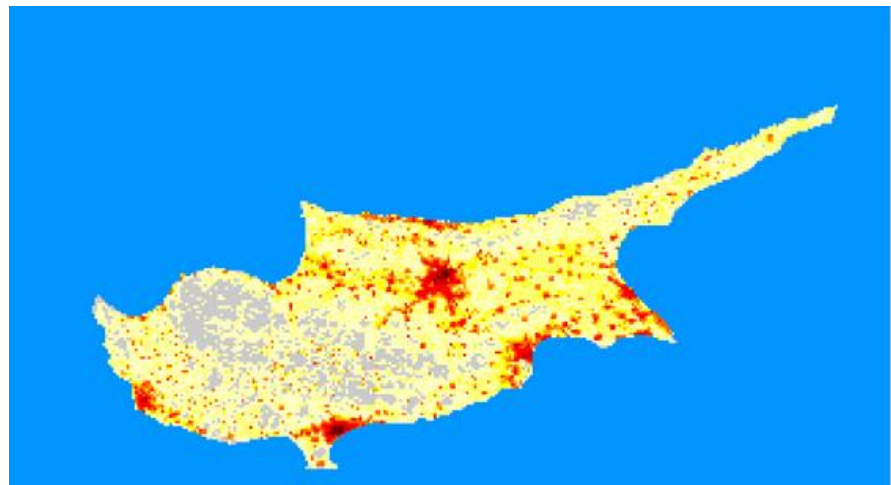
- Interpoladores Inteligentes:
 - Exemplo: *LandScan* – grade 1km, 1995
 - Modelo de População: uso do solo, proximidade de estradas, luzes noturnas => **coeficiente de probabilidade**
 - População em risco: Informação para medidas emergenciais – desastres naturais ou antropogênicos



Landscan - <http://web.ornl.gov/sci/landscan/index.shtml>

- Abordagem GIS & RS
- ORNL's LandScan™ - padrão para distribuição de população global.
- Resolução ~ 1 km resolution (30" X 30")
- Dados representando população ambiental (média 24 h).
- Algoritmo LandScan: dados espaciais e imagens em modelo dasimétrico multi-variável para desagregar contagens(censo) dos limites administrativos
- Ajusta as condições dos dados e natureza geográfica de cada país e região
- Ex: dado – Chipre (2011)
(N de pessoas/células)
30 arc sec; Geographic WGS84

- Fees – caso-a-caso !



LandScan™

Geographic Information Science & Technology

About

ORNL's LandScan™ is a community standard for global population distribution data. At approximately 1 km (30' X 30') spatial resolution, it represents an ambient population (average over 24 hours) distribution. The database is refreshed annually and released to the broader user community around October.

LandScan™ is now available at no cost to the research and education communities. The latest LandScan™ dataset available is LandScan Global 2016. Older LandScan Global data sets (LandScan 1998, 2000-2015) will soon be available from this site. These data set can be licensed for commercial and other applications through multiple third-party vendors.

LandScan is developed using best available demographic (Census) and geographic data, remote sensing imagery analysis techniques within a multivariate dasymetric modeling framework to disaggregate census counts within an administrative boundary. Since no single population distribution model can account for the differences in spatial data availability, quality, scale, and accuracy as well as the differences in cultural settlement practices, LandScan population distribution is essentially a combination of locally adoptive models that are tailored to match the data conditions and geographical nature of each individual country and region.

- Based upon the spatial data and the socioeconomic and cultural understanding of an area, cells are preferentially **weighted** for the possible occurrence of population during a day.
- Within each country, the population distribution model calculates a “**likelihood**” **coefficient** for each cell and applies the coefficients to the census counts, which are employed as control totals for appropriate areas.
- The total population for that area is then allocated to each cell proportionally to the calculated population coefficient. The resultant population count is an ambient or **average day/night population count**.
- The LandScan™ dataset was **first produced in 1998** as an improved resolution global population distribution database for estimating populations at risk

- The LandScan™ algorithm, an R&D 100 Award Winner, uses spatial data and imagery analysis technologies and a **multi-variable dasymetric** modeling approach to disaggregate census counts within an administrative boundary.
- is also referred to as a “**smart interpolation**” technique
- the typical dasymetric modeling is improved by integrating and employing **multiple ancillary or indicator data layers**.
- The modeling process uses sub-national level census counts for each country and primary geospatial input or ancillary datasets, including **land cover, roads, slope, urban areas, village locations, and high resolution imagery analysis**; all of which *are key indicators* of population distribution

- **Elevation and slope** are also important indicators of population distribution potential. Although on occasion humans will frequent the very highest elevations on the planet, there are elevation levels above which consistent habitation is impractical.
- Likewise, very steep slopes inhibit settlements, agriculture, and industrial development – the most likely areas of ambient population distribution.
- Various vector data layers used in the modeling algorithms include **roads, populated areas (urban boundaries), and populated points (towns and villages)**. Each data layer serves as an *indicator of likely population* locations

Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data

PLOS ONE | DOI:10.1371/journal.pone.0107042 February 17, 2015

Forrest R. Stevens^{1*}, Andrea E. Gaughan¹, Catherine Linard^{2,3}, Andrew J. Tatem^{4,5}

- Justificativa: população e em grade → ausência de dados
- Redistribuição dasimétrica – pop counts/dados de censo e ponderação baseado em modelagem preditiva não-paramétrica “Randon Forest”
- Dados geoespaciais e de SR de múltiplas escalas → 100m
- Estudo de Caso: Kenya (KEN), Vietnã (VNM) e Camboja (KHM)

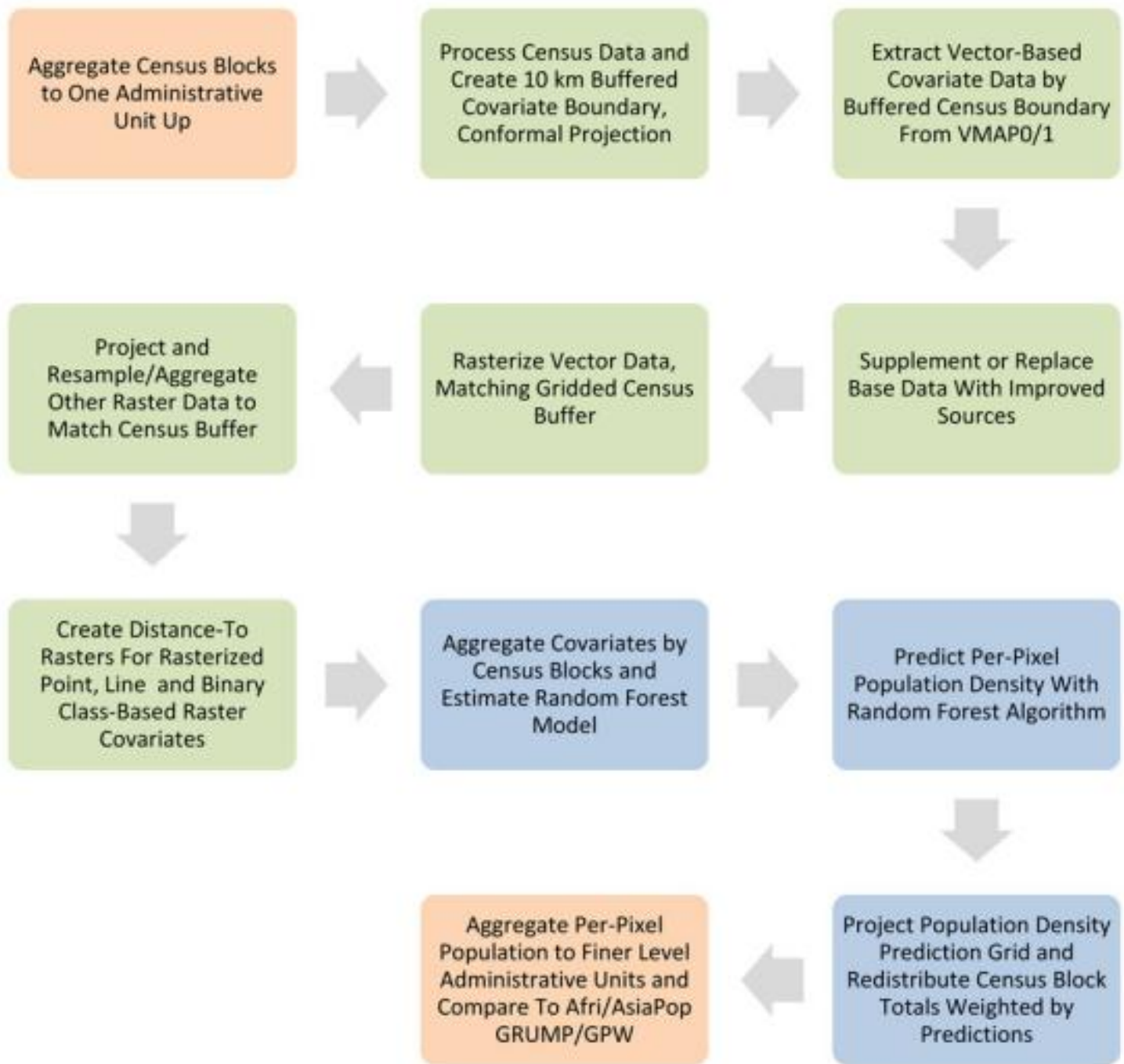
Metodologia

- Dados de censo e limites administrativos de níveis mais detalhados de cada país
 - Log da densidade de população como variável resposta
 - Distr Pop é sempre relacionado com tipos de cobertura da terra – Mapeamentos de LandCover local, do GeoCover(30m) e GlobCover (300m) → classes compatibilizadas e todos para pixels de ~100m
 - SRTM – DEM;
 - MOD17A3 – produção primária líquida (NPP)
 - Imagens de luzes noturnas – VIIRS
 - Clima: pp anual (BIO12) e temperatura média anual (BIO1)
 - Vetores>> redes de estradas, corpos d'água, locais de população ou assentamentos, áreas de proteção e facilidades (hospitais, escolas, clínicas de saúde)
- Pesos → Random Forest Trees → superfície de densidade

Laranja – este trabalho

Verde – pre-processamento

Azul – estimativa do modelo Random Forest, predição por-pixel e redistribuição dasimétrica dos counts do censo



Mapas para usuários finais → estimativas para outros anos 2010, 2015 e 2020

Projeções baseados na proporção de pop urbana e rural

P_{-} → ano de projeção

P_d → pop no ano do dado de entrada

t → n de anos entre o dado de entrada e o ano projetado

r → é a taxa de crescimento da pop urbana ou rural

$$P_{-}\{2010\} = P_{-}\{d\}e^{\{rt\}}$$

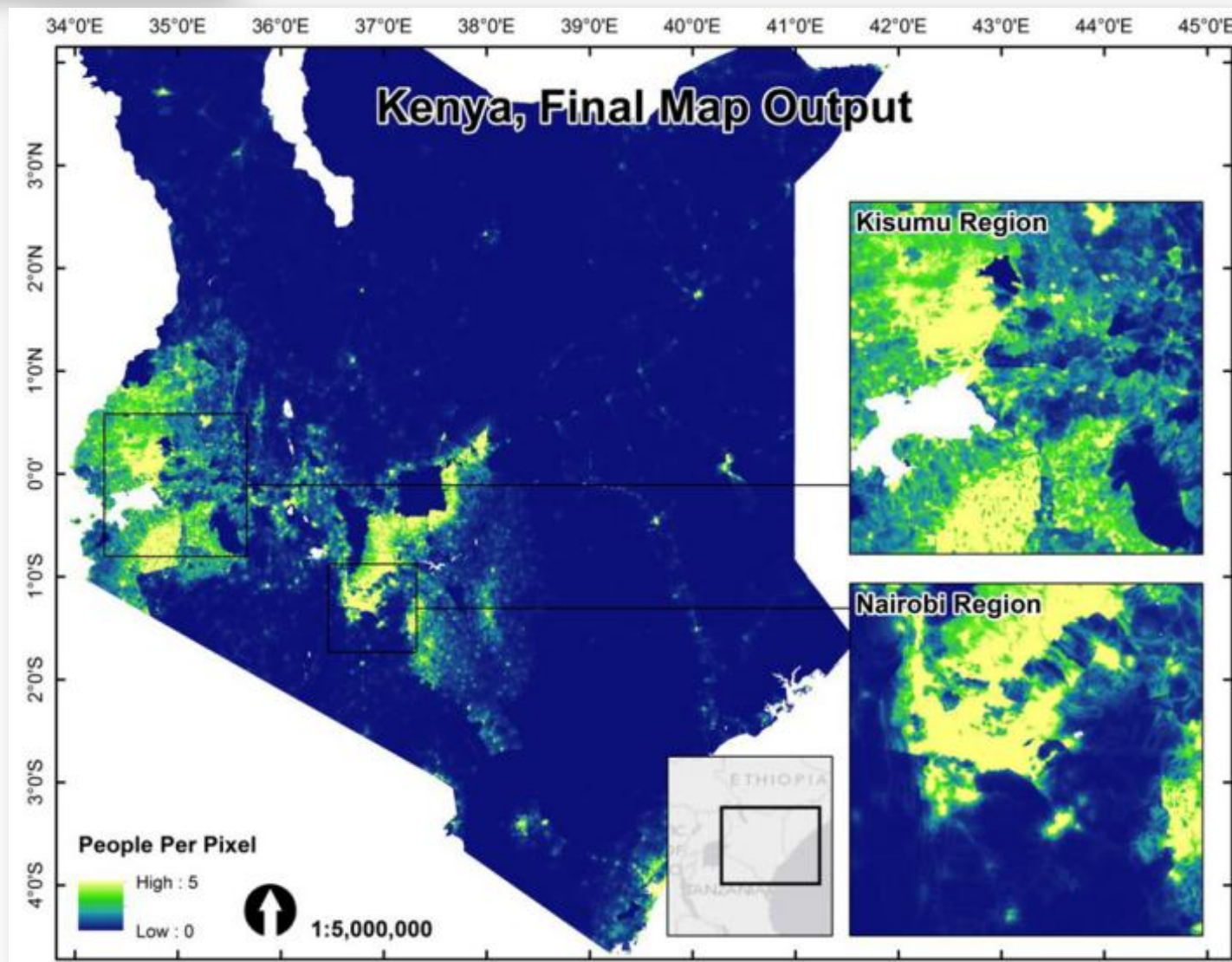
Comparações dos resultados do Random Forest (RF) com outras bases de pop em grades

- Global Rural Urban Mapping Project (GRUMP)
- Gridded Population of the World (GPW)
- AfriPop - freely-available gridded population data for Africa

Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data

Forrest R. Stevens^{1*}, Andrea E. Gaughan¹, Catherine Linard^{2,3}, Andrew J. Tatem^{4,5}

Kenya



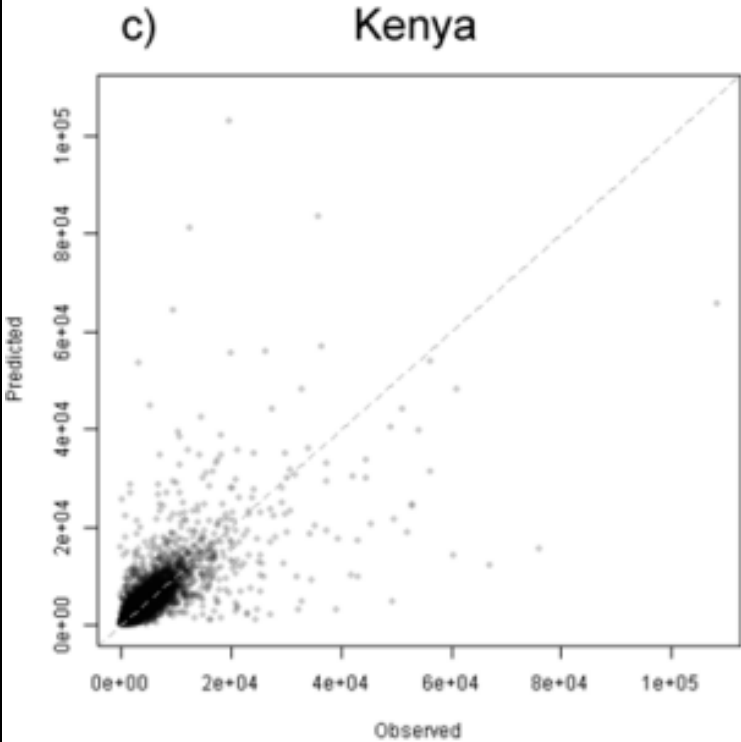


Fig 4. Observed census counts plotted from the finer census administrative units versus the summed grid cell values from the population map estimated using coarser administrative units for a) Cambodia, b) Vietnam, and c) Kenya.

Random Forest Variable Importance for Kenya Population Map Production

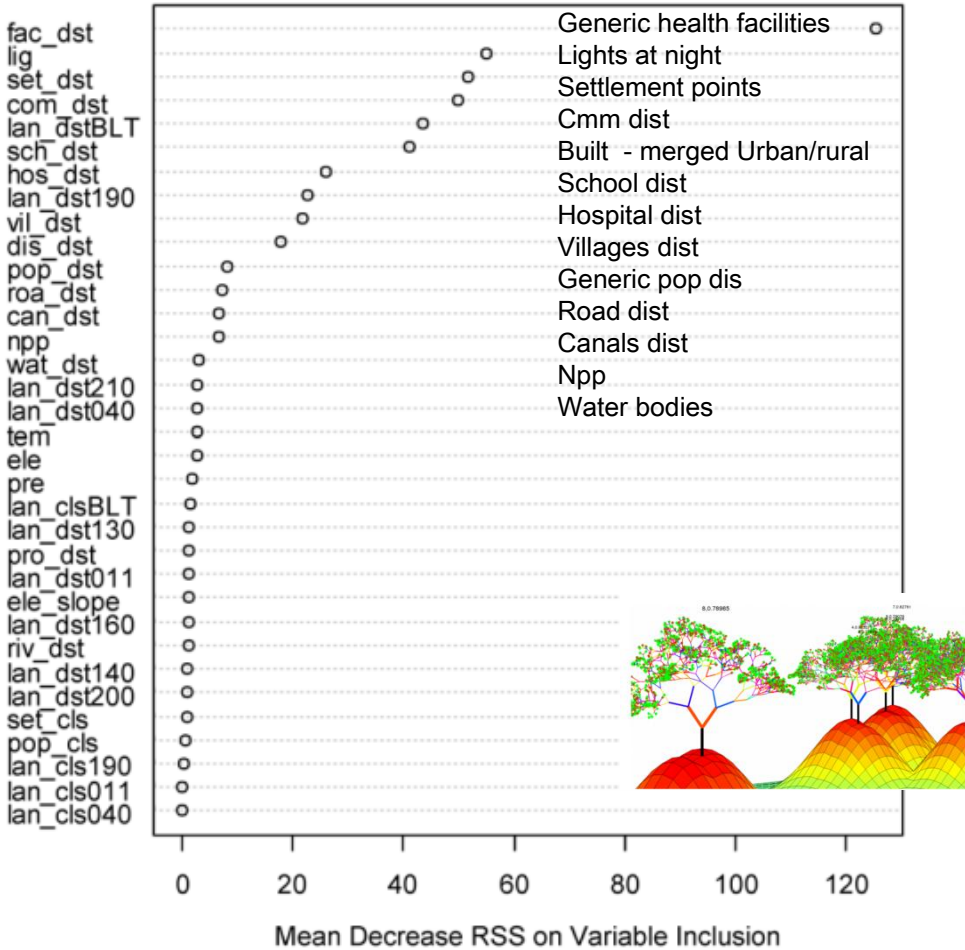


Fig 3. Variable importance for Random Forest regression, presented as the mean decrease in residual sum of squares when the variable is included in a tree split. The model including variables here was used to produce the density weighting layer for the dasymmetrically distributed population map in Fig 2. Variable names are defined and described in Table 1.

Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data

Forrest R. Stevens^{1*}, Andrea E. Gaughan¹, Catherine Linard^{2,3}, Andrew J. Tatem^{4,5}

Comparação

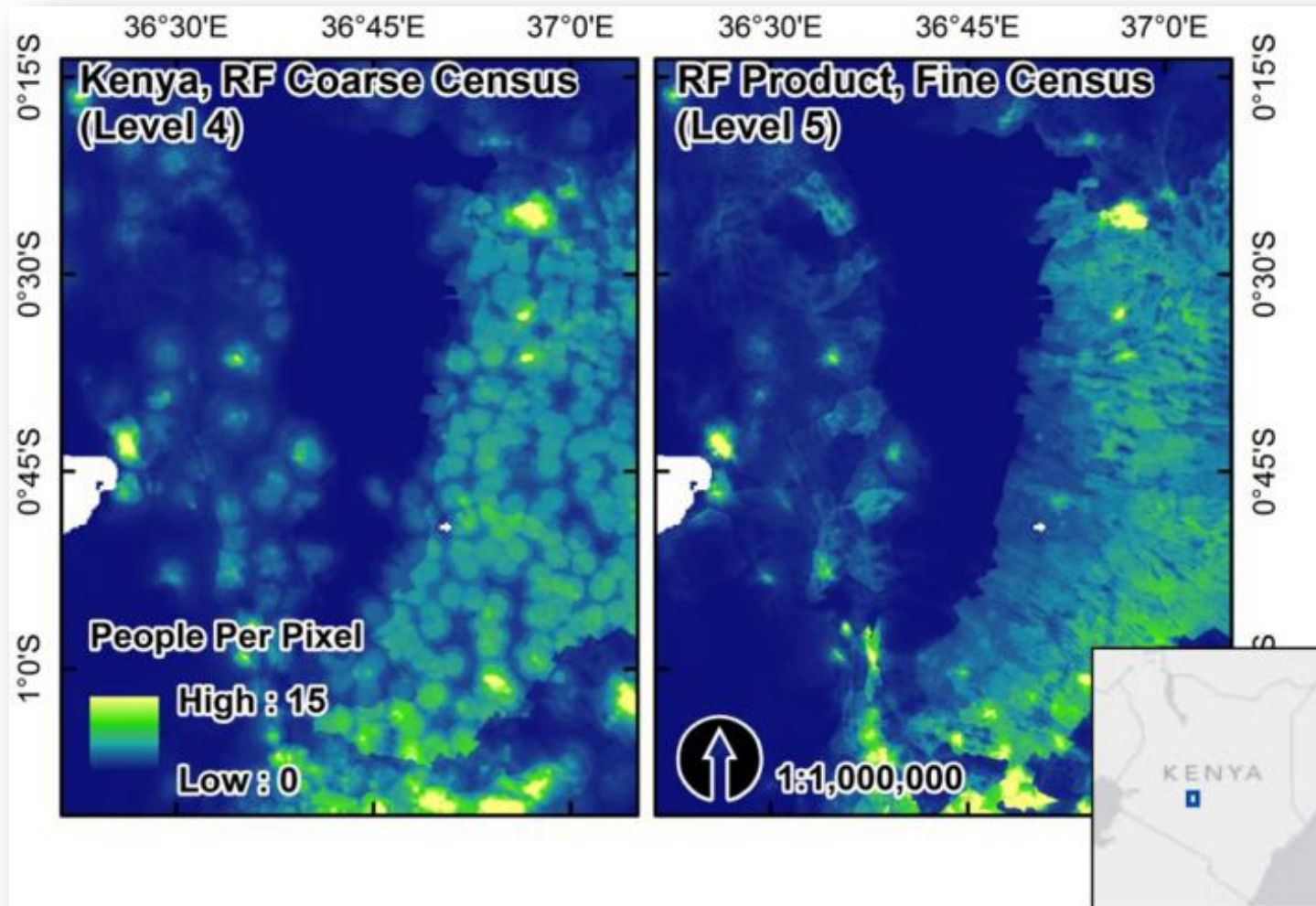


Fig 5. A visual comparison of Kenyan population maps for census data in 1999 produced at coarser administrative unit (Level 4) and finer-scale administrative unit (Level 5). The difference illustrates the finer gradations of RF model predictions for the density weighting layer when there are larger ranges of observed population densities present in training data (N = 505 for Level 4, N = 6622 for Level 5).

Comparaç o

Table 2. Accuracy assessment results for the RF, Afri/AsiaPop, GRUMP and GPW modeling methods for Cambodia, Vietnam and Kenya.

Country	Method	RMSE	% RMSE	MAE
Cambodia	RF	3209.40	38.84	2026.73
	AsiaPop	3834.51	46.40	2494.32
	GRUMP	6767.39	81.89	3889.39
	GPW	6794.88	82.22	4021.41
Vietnam	RF	4367.00	61.96	2778.80
	AsiaPop	4943.31	70.13	3007.04
	GRUMP	6523.77	92.56	3771.64
	GPW	7081.76	100.47	3844.47
Kenya	RF	3956.74	91.36	1685.82
	AsiaPop	5208.79	120.28	2184.64
	GRUMP	6294.61	145.35	2383.64
	GPW	6327.80	146.11	2304.70

Two different error assessment methods are presented: root mean square error (RMSE), also expressed as a percentage of the mean population size of the administrative level (% RMSE); and the mean absolute error (MAE).

- Global Rural Urban Mapping Project (GRUMP)
- Gridded Population of the World (GPW)
- AsiaPop - freely-available gridded population data for Asia
- AfriPop - freely-available gridded population data for Africa

Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data

Forrest R. Stevens^{1*}, Andrea E. Gaughan¹, Catherine Linard^{2,3}, Andrew J. Tatem^{4,5}

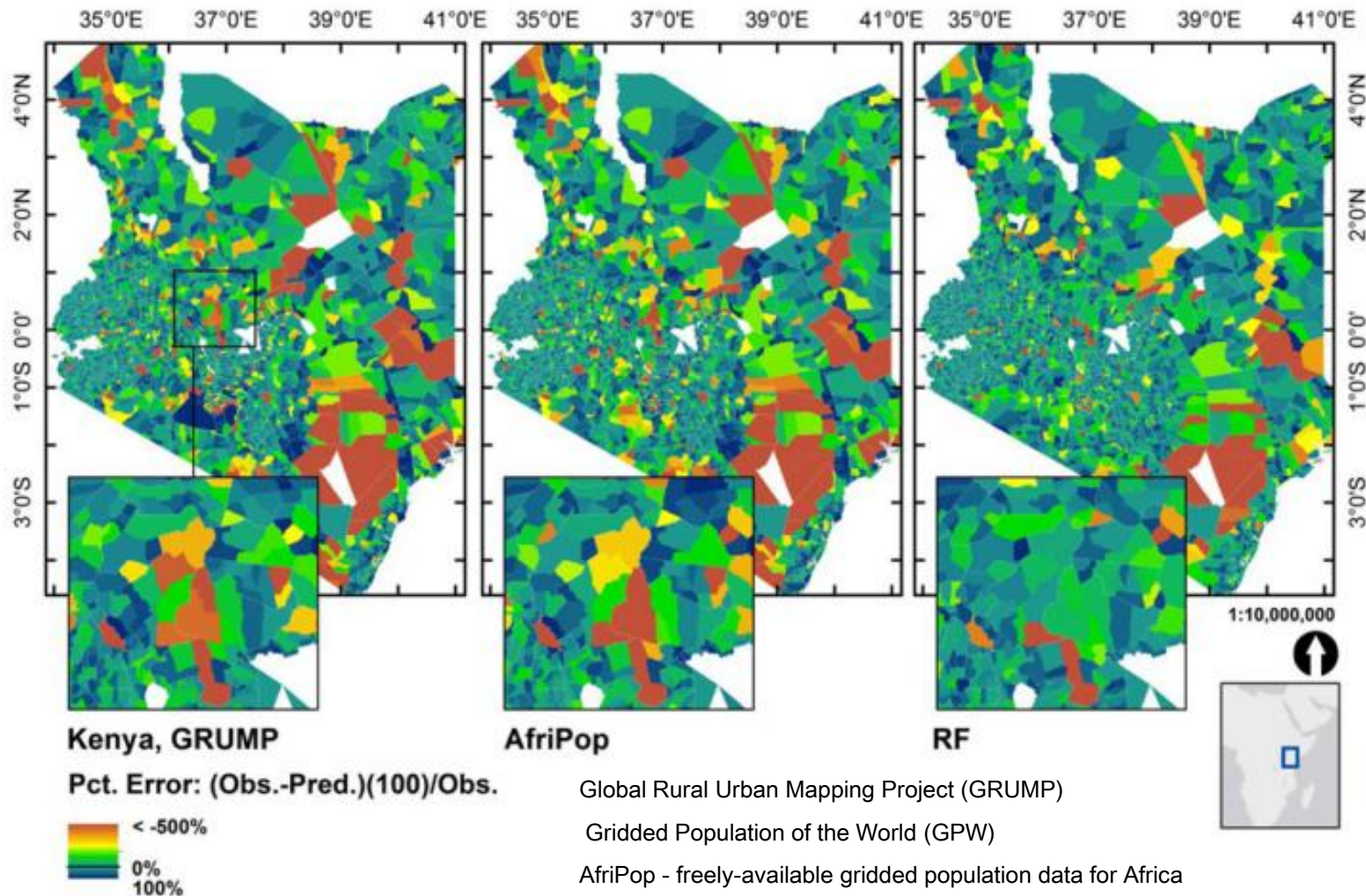


Fig 6. For Kenya we compared the summed predictions of population maps estimated using coarse census data (Level 4, "Division" level) to population counts at Level 5 (finer scale) units. We present prediction errors (observed minus predicted) as percentage values of the observed Level 5 census counts. The RF approach results in far fewer census units with extreme over predictions (negative percent residuals in yellows and oranges) and under predictions (positive percent residuals in dark blues).

Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data

Forrest R. Stevens^{1*}, Andrea E. Gaughan¹, Catherine Linard^{2,3}, Andrew J. Tatem^{4,5}

Kenya

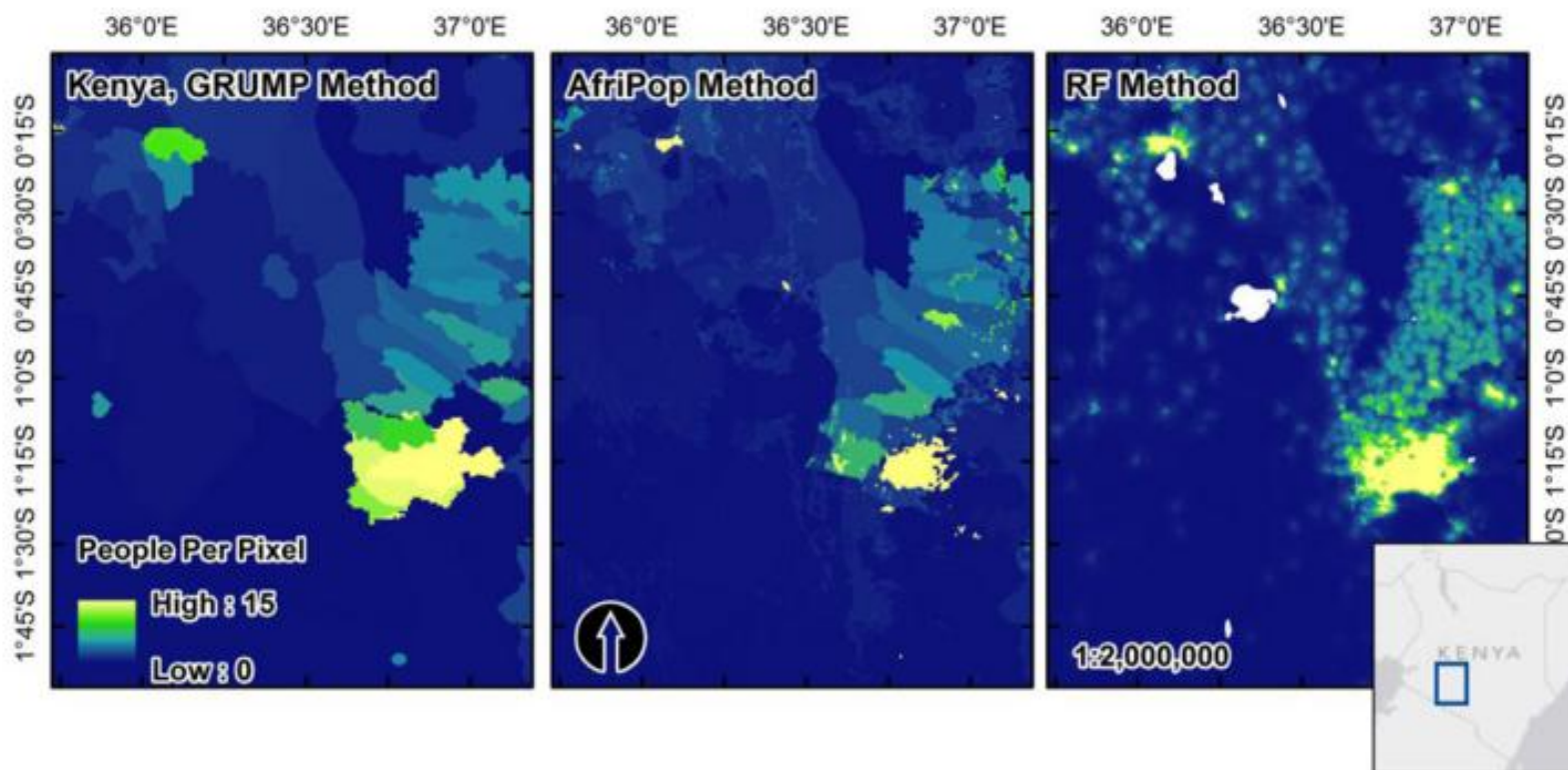


Fig 7. Visual comparisons of GRUMP, AfriPop and the RF-based population map from this study. Though this region northwest of Nairobi, Kenya is not a highly populated region this figure shows the results of the more detailed RF weighting layer versus the use of just urban areas (GRUMP) and land cover plus urban areas (AfriPop). The distinct edges in estimated people per pixel between census units are almost eliminated by the RF approach and it achieves greater consistency in predicted population density after census count redistribution.

Population datasets for 2010, 2015, and 2020 are freely available for download from the WorldPop Project (<http://www.worldpop.org.uk/>)

WorldPop: High resolution age-structured population distribution maps



What is WorldPop?

High spatial resolution, contemporary data on human population distributions are a prerequisite for the accurate measurement of the impacts of population growth, for monitoring changes and for planning interventions. The WorldPop project aims to meet these needs through the provision of detailed and open access population distribution datasets built using transparent approaches.

The WorldPop project was initiated in October 2013 to combine the AfriPop, AsiaPop and AmeriPop population mapping projects. It aims to provide an open access archive of spatial demographic datasets for Central and South America, Africa and Asia to support development, disaster response and health applications. The methods used are designed with full open access and operational application in mind, using transparent, fully documented and peer-reviewed methods to produce easily updatable maps with accompanying metadata and measures of uncertainty. [Click to download WorldPop flyer.](#)

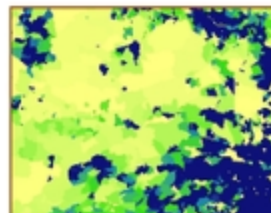
[Click to read WorldPop case studies](#) ➤

Brazil 100m Population

Brazil



Number of people
per grid square



WorldPop Americas dataset details

DATASET: Alpha version 2010, 2015 and 2020 estimates of numbers of people per pixel ('ppp') and people per hectare ('pph'), with national totals adjusted to match UN population division estimates (<http://esa.un.org/wpp/>) and remaining unadjusted.

REGION: Americas

SPATIAL RESOLUTION: 0.000833333 decimal degrees (approx 100m at the equator)

PROJECTION: Geographic, WGS84

UNITS: Estimated persons per grid square

MAPPING APPROACH: Random Forest

FORMAT: Geotiff (zipped using 7-zip (open access tool): www.7-zip.org)

FILENAMES: Example - MEX_ppp_v2b_2010_UNadj.tif = Mexico (MEX) population per pixel (ppp) map for 2010 (2010) adjusted to match UN national estimates (UNadj), version 2b (v2b).

DATE OF PRODUCTION: June 2015

Also included: (i) Metadata html file, (ii) 2 Google Earth files, and (iii) Population datasets produced using original census year data (2010).

[View Metadata](#)

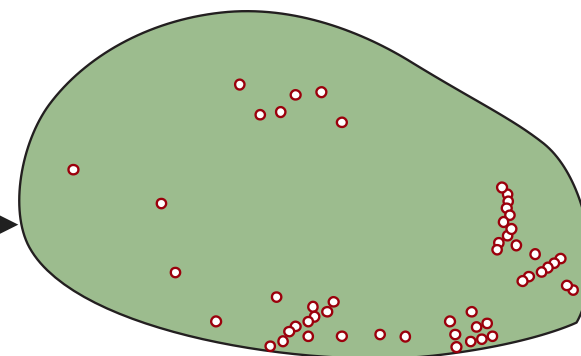
http://www.worldpop.org.uk/data/WorldPop_data/AllContinents/8_metadata.html

- Interpoladores Inteligentes e Variantes:
 - Importância da escolha de variáveis e conhecimento das inter-relações – “modelo”
 - Existência e qualidade dos dados auxiliares determinam a exatidão da superfície resultante.

GEOCODIFICAÇÃO DO CADASTRO NACIONAL DE ENDEREÇOS PARA FINS ESTATÍSTICOS (CNEFE)



row.names	Codigo.do.setor	nome.mun	end
1 908	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
2 909	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
3 910	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
4 912	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
5 913	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
6 914	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
7 921	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
8 923	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
9 924	350315805000005	ARAPEÍ	ESTRADA SP 68, 250, ARAPEÍ - SP, 12870-000, Brazil
10 943	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
11 945	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
12 947	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
13 953	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
14 954	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
15 955	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
16 956	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
17 958	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
18 960	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
19 961	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
20 962	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
21 972	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
22 976	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
23 987	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
24 991	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
25 993	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
26 994	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
27 995	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
28 996	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
29 1009	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
30 1013	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
31 109	350350505000001	AREIAS	RUA RIACHUELO, 250, AREIAS - SP, 12870-000, Brazil
32 2	350350505000001	AREIAS	RUA RIACHUELO, 296, AREIAS - SP, 12870-000, Brazil

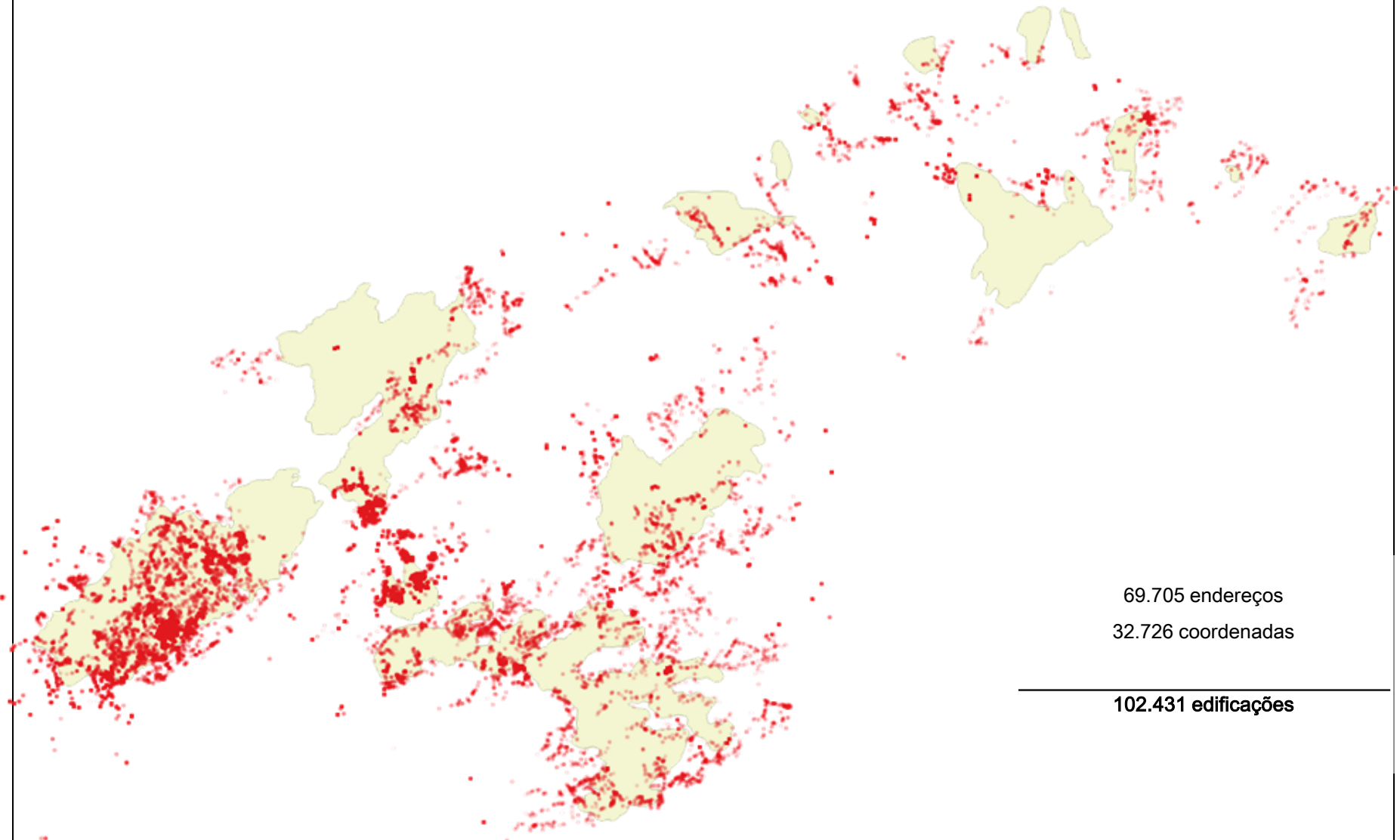


69.705 endereços

32.726 coordenadas

102.431 edificações

GEOCODIFICAÇÃO DO CADASTRO NACIONAL DE ENDEREÇOS PARA FINS ESTATÍSTICOS (CNEFE)



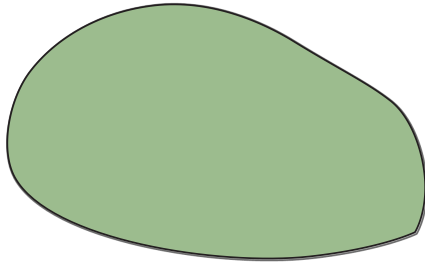
69.705 endereços
32.726 coordenadas

102.431 edificações

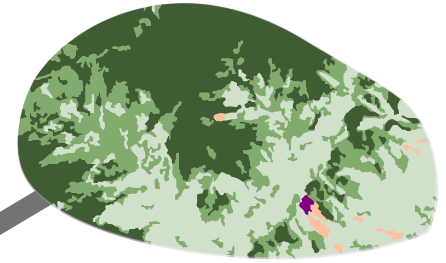
INTEGRAÇÃO DOS DADOS NO ESPAÇO CELULAR



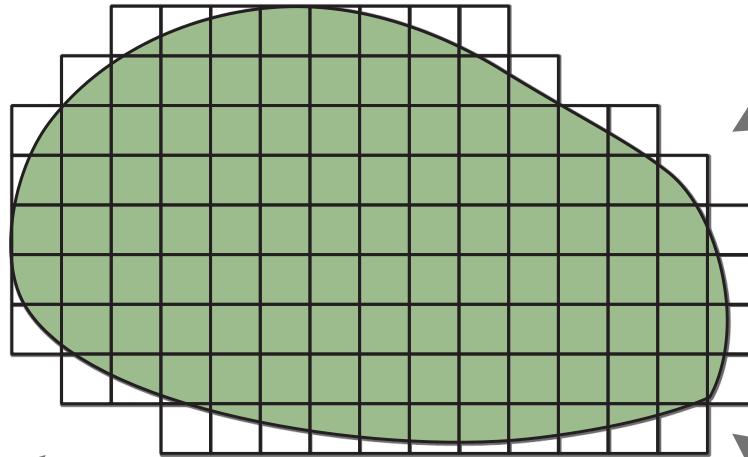
APA-MRPS



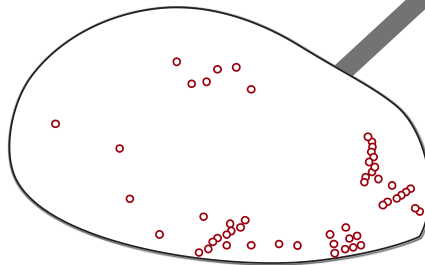
Uso e cobertura da terra



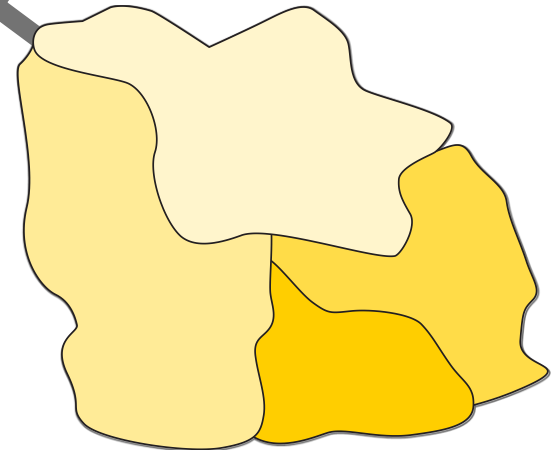
Espaço celular



CNEFE



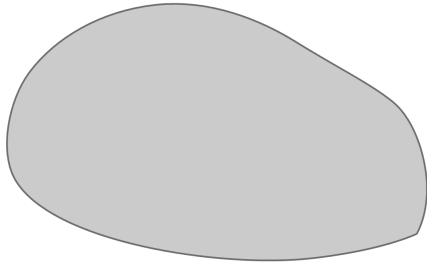
Setores censitários



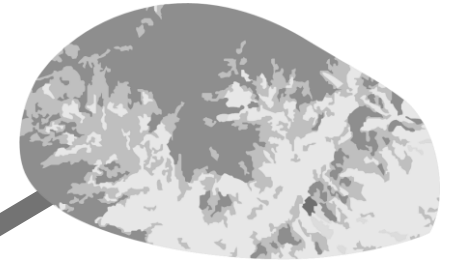
INTEGRAÇÃO DOS DADOS NO ESPAÇO CELULAR



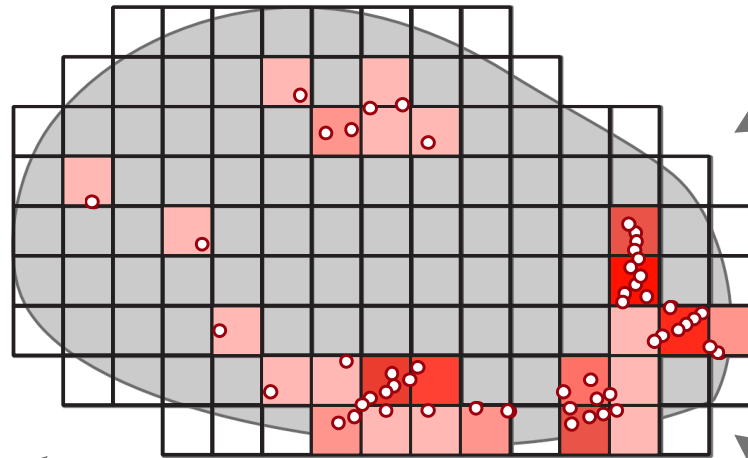
APA-MRPS



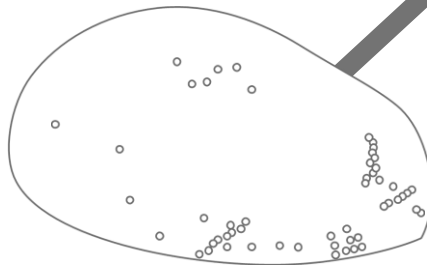
Uso e cobertura da terra



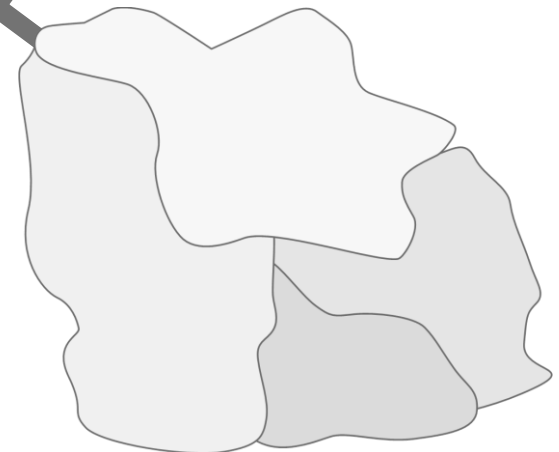
Espaço celular



CNEFE



Setores censitários



- Testaram método para aumentar a resolução de dados censitários e apresentá-los agregados em uma grade regular
- Metodologia:
 - agrega domicílios (coordenadas geográficas) -Contagem da População do Instituto Brasileiro de Geografia e Estatística (IBGE) 2007
 - desagrega por setor censitário - proporcionalidade de área.
- Estima pop residente em 114 UCs na AmzLegal Bras

STATISTICS FINLAND, 2010- dissemina dados em grades

- Validade no uso de grades regulares

(+) prescinde dos limites político-administrativos para apresentar as informações

(-) dificuldade de determinar e corrigir erros espaciais na coleta, como os decorrentes do registro incorreto de coordenadas geográficas ou até mesmo da falha na captura das coordenadas dos domicílios → novo campo(?)

Métodos de construção:

- técnicas de desagregação (*top-down*)
 - Dados agregados em áreas e métodos estatísticos para desagregar na grade (krigagem)
 - Variáveis auxiliares para a desagregação
 - Problemas conhecidos: não preserva volume e não representa espaços sem população
 - Outras técnicas – referências conhecidas

Métodos de construção:

- agregação (*bottom-up*)
 - quantificação dos dados discretos referentes à localização das residências, em células de uma grade regular que abrange a região de estudo
 - tem resultados superiores a qualquer método estatístico utilizado para desagregação dos volumes populacionais (óbvio, não?)

Brasil, agregação viável a partir de 2007 - uso sistemático e integrado de geotecnologias - espacialização do Cadastro de Endereços para Fins Estatísticos e seu relacionamento com os Mapas Censitários nas operações de 2010.

Contagem e Censo Agropecuário (2007) e Censo 2010 →
coordenadas geográficas das unidades situadas nas áreas rurais, sendo as mesmas complementares aos endereços, já que nessas áreas não existe um endereçamento formal como nas áreas urbanas (IBGE, 2009).

As coordenadas dos domicílios visitados é a base para a técnica de agregação em grades regulares, possibilitando a conversão de objetos discretos – os **pontos** representando os domicílios – em pequenas unidades de **áreas, as células**, que agregam as informações individuais.

METODOLOGIA

1) Agregação:

- da localização dos domicílios visitados na Contagem 2007 + Cadastro de Endereços para Fins Estatísticos (áreas exclusivamente rurais), exclusive as áreas de aglomerado rural– e relativos a domicílios particulares.

2) Desagregação:

- Malha vetorial de setores censitários - aglomerados rurais (povoados e vilas) e urbanos, associando-se os dados demográficos dos setores com as células da grade por meio de equivalência de área.
- Células da grade = 1min e 15 seg ou $\sim 5,3 \text{ km}^2$.
- A Amazônia Legal = 953.383 células. (testes)

- Operação espacial – atribuiu ao ponto a localização associada a grade (Junção espacial da grade vetorial e pontos)
- A estimativa do volume da população por domicílio = média de moradores no setor censitário onde o mesmo se localiza.

$$\begin{array}{lcl} \text{Média de} & & \text{população no setor censitário, residente em} \\ \text{moradores por} & & \text{domicílios permanentes ou improvisados} \\ \text{domicílio em} & = & \\ \text{cada setor} & & \hline \text{censitário} & & \text{número de domicílios, permanentes ou} \\ & & \text{improvisados, existentes no setor.} \end{array}$$

- Posteriormente, esse valor foi agregado para cada célula.

- Subenumeração relativa ao registro das coordenadas geográficas da Contagem 2007.
 - Não é erro de cobertura censitária, mas perda de dados quando da realização da espacialização.
 - O total de domicílios na grade regular (877.771) foi inferior ao Total do universo da pesquisa (1.113.258).
 - A “subenumeração espacial” - o erro não se distribui de modo uniforme no espaço - varia entre os setores censitário
 - Precisaria de campo para corrigir....

Então....

- nas células que **apresentam população**, os dados foram ajustados aplicando-se o percentual de erro calculado para o setor onde a mesma se encontra;
- nos casos em que a célula localizava-se **em mais de um setor censitário**, foi considerado o erro do setor com maior área dentro da célula;
- nas células onde a metodologia apontou **ausência de população**, o erro relativo à falta de domicílios com coordenadas não foi corrigido, assumindo-se que a população não localizada por meio de coordenadas da Contagem 2007 esteja totalmente distribuída por áreas onde a presença de população foi detectada.

Nas células de aglomerados rurais e urbano (polígonos)

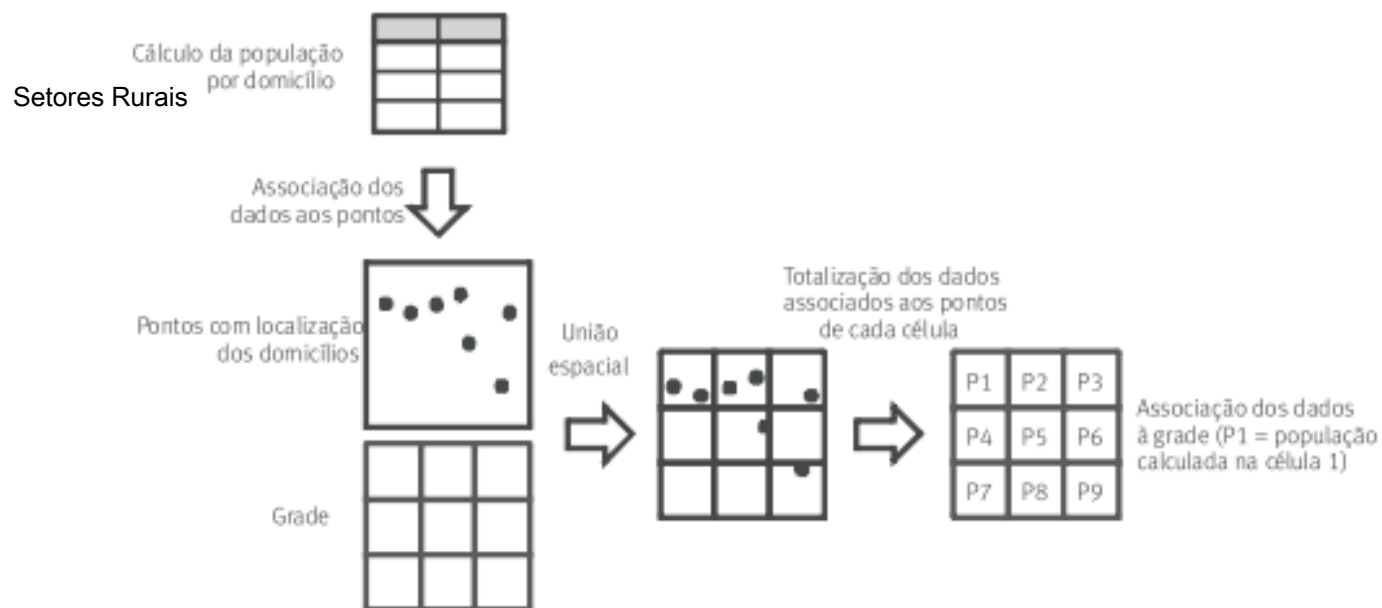
- metodologia de proporcionalidade de área,
 - supõe que a população esteja distribuída homogeneamente dentro do setor censitário.

Assim, a parcela de população equivalente à área do setor censitário que se encontra dentro de uma célula da grade é agregada a esta célula.

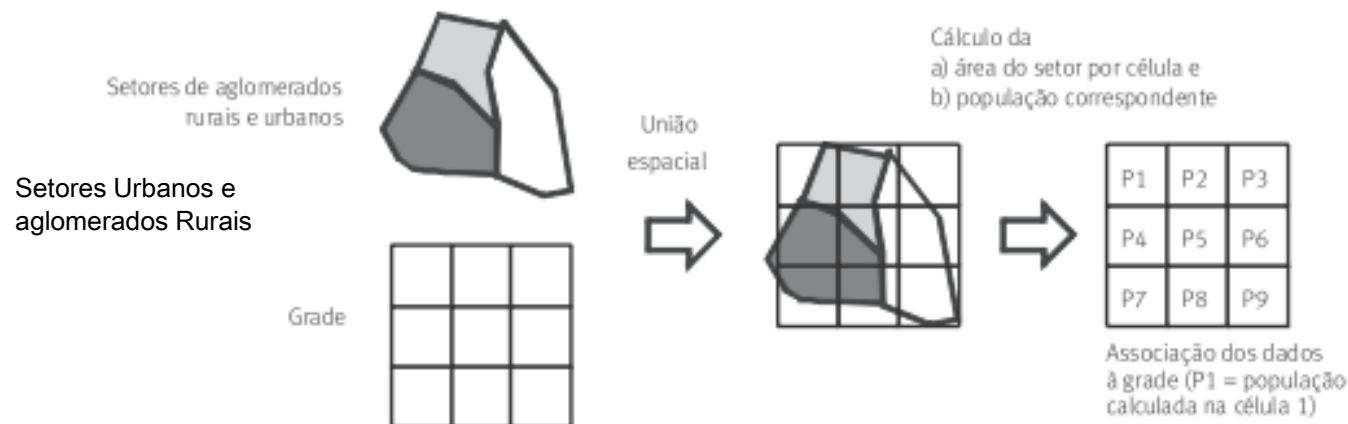
* pressuposto de homogeneidade

D'Antona, A.O.; Bueno, M.C. Dagnino, R.S. (2013) Estimativa da população em unidades de conservação na Amazônia Legal brasileira – uma aplicação de grades regulares a partir da Contagem 2007. REBEP (30) 2, p. 401-428.

A – Geração de dados por agregação (*bottom-up*)



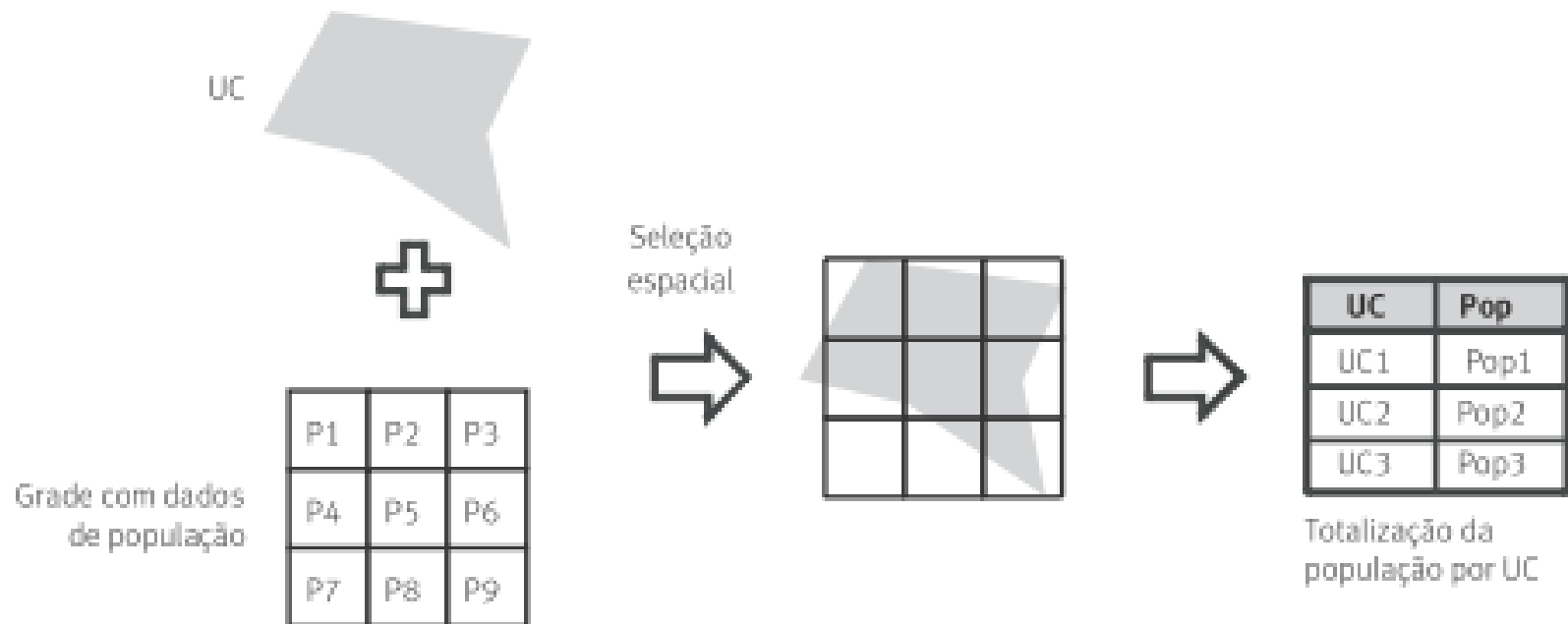
B – Geração de dados por desagregação (*top-down*)



Sobre a grade consolidada – limites dos polígonos das Ucs

- 121 Ucs federais (43 UCPI e 70 UCUS)
- Buffer 10km para computar pop no entorno tb;

C – Estimativa da população por Unidade de Conservação (UC)



D'Antona, A.O.; Bueno, M.C. Dagnino, R.S. (2013) Estimativa da população em unidades de conservação na Amazônia Legal brasileira – uma aplicação de grades regulares a partir da Contagem 2007. REBEP (30) 2, p. 401-428.

TABELA 1
População estimada, segundo tipos e categorias de unidades de conservação federais
Amazônia Legal brasileira – 2007

Tipos e categorias de UCs	Unidades de conservação (UC)						Entorno (1)			
	Qtde.	Área (km ²)	População (hab.) (2)				População (hab.) (2)			
			Rural (3)	Urbana (4)	Total (5)	Total (ajustada) (6)	Rural (3)	Urbana (4)	Total	Total (ajustada) (6)
Proteção integral	40	277.173	23.593	1.283	24.877	27.705	42.353	50.867	93.220	96.738
Estações Ecológicas	11	55.908	679	0	679	905	6.866	14.425	21.291	21.771
Parques Nacionais	21	185.819	15.937	1.085	17.022	19.228	26.770	35.927	62.697	65.082
Reservas Biológicas	8	35.447	6.978	198	7.176	7.572	8.717	515	9.232	9.886
Uso sustentável	74	261.239	119.347	164.293	283.640	297.693	237.791	657.148	894.938	923.499
Áreas de Proteção Ambiental	2	20.800	2.766	615	3.381	4.013	8.257	33.059	41.316	41.701
Áreas de Relevante Interesse Ecológico	3	200	213	0	213	296	4.861	9.568	14.429	14.629
Florestas Nacionais	29	146.429	28.905	34.163	63.068	65.196	48.673	120.035	168.708	174.699
Reservas de Desenvolvimento Sustentável	1	673	1.375	0	1.375	1.499	1.966	47	2.013	2.506
Reservas Extrativistas	39	93.137	86.088	129.515	215.603	226.689	174.033	494.439	668.472	689.964
Total	114	538.412	142.940	165.577	308.517	325.398	280.144	708.015	988.158	1.020.237

Fonte: IBGE. Microdados da Contagem Populacional de 2007.

(1) Definido pela delimitação de uma faixa de 10 km de largura ao redor de cada uma das UCs consideradas.

(2) Estimativa a partir da Contagem Populacional 2007.

(3) População em domicílios com localização registrada em setores censitários rurais, incluindo os aglomerados rurais.

(4) População em domicílios com localização registrada em setores censitários urbanos.

(5) População total (rural + urbana). Sujeito à subenumeração, conforme o setor censitário.

(6) População total ajustada para minimizar a subenumeração.

D'Antona, A.O.; Bueno, M.C. Dagnino, R.S. (2013) Estimativa da população em unidades de conservação na Amazônia Legal brasileira – uma aplicação de grades regulares a partir da Contagem 2007. REBEP (30) 2, p. 401-428.

(continuação)

Unidades de conservação	Unidade de conservação							Entorno (1)			
	Área (km²)	Células		População (hab.) (2)				População (hab.) (2)			
		Total	Ocup. (%) (3)	Rural (4)	Urbana (5)	Total (6)	Total (ajustada) (7)	Rural (4)	Urbana (5)	Total (6)	Total (ajustada) (7)
Mata Grande	125	40	80,00	1.315	15.317	16.632	16.682	8.565	215.639	224.204	224.481
Médio Juruá	2.536	572	5,59	1.699	0	1.699	1.713	1.043	0	1.043	1.072
Quilombo Flexal	87	24	45,83	679	1.483	2.162	2.174	6.526	7.212	13.738	14.267
Rio Cajari	5.201	1.062	10,36	2.430	0	2.430	2.824	1.344	0	1.344	1.490
Rio do Cautário	743	151	8,61	111	0	111	114	134	0	134	138
Rio Iriri	3.992	897	0,67	29	0	29	45	27	0	27	30
Rio Jutai	2.741	602	5,81	1.537	0	1.537	1.548	241	0	241	245
Rio Ouro Preto	2.037	479	12,53	486	0	486	490	436	0	436	461
Rio Unini	8.416	1.863	2,79	1.126	0	1.126	1.172	1.102	0	1.102	1.113
Riozinho da Liberdade	3.492	767	3,26	1.112	0	1.112	1.396	1.161	0	1.161	1.405
Riozinho do Anfrísio	7.351	1.511	1,26	187	0	187	250	505	0	505	565
São João da Ponta	32	30	53,33	1.294	2.293	3.587	3.592	12.874	18.543	31.417	32.820
Terra Grande-Pracuúba	1.949	434	5,99	1.045	0	1.045	1.400	3.600	9.874	13.475	14.315
Verde Para Sempre	12.860	2.548	11,07	6.086	0	6.086	7.391	4.407	26.447	30.854	31.368
Total	538.412	113.447	3,96	142.940	165.577	308.517	325.398	280.144	708.015	988.158	1.020.237

Fonte: IBGE. Microdados da Contagem Populacional de 2007.

(1) Definido pela delimitação de uma faixa de 10km de largura ao redor de cada uma das UCs consideradas.

(2) Estimativa a partir da Contagem Populacional 2007.

(3) Percentual de células com população em relação ao total de células no interior da UC.

(4) População em domicílios com localização registrada em setores censitários rurais, incluindo os aglomerados rurais.

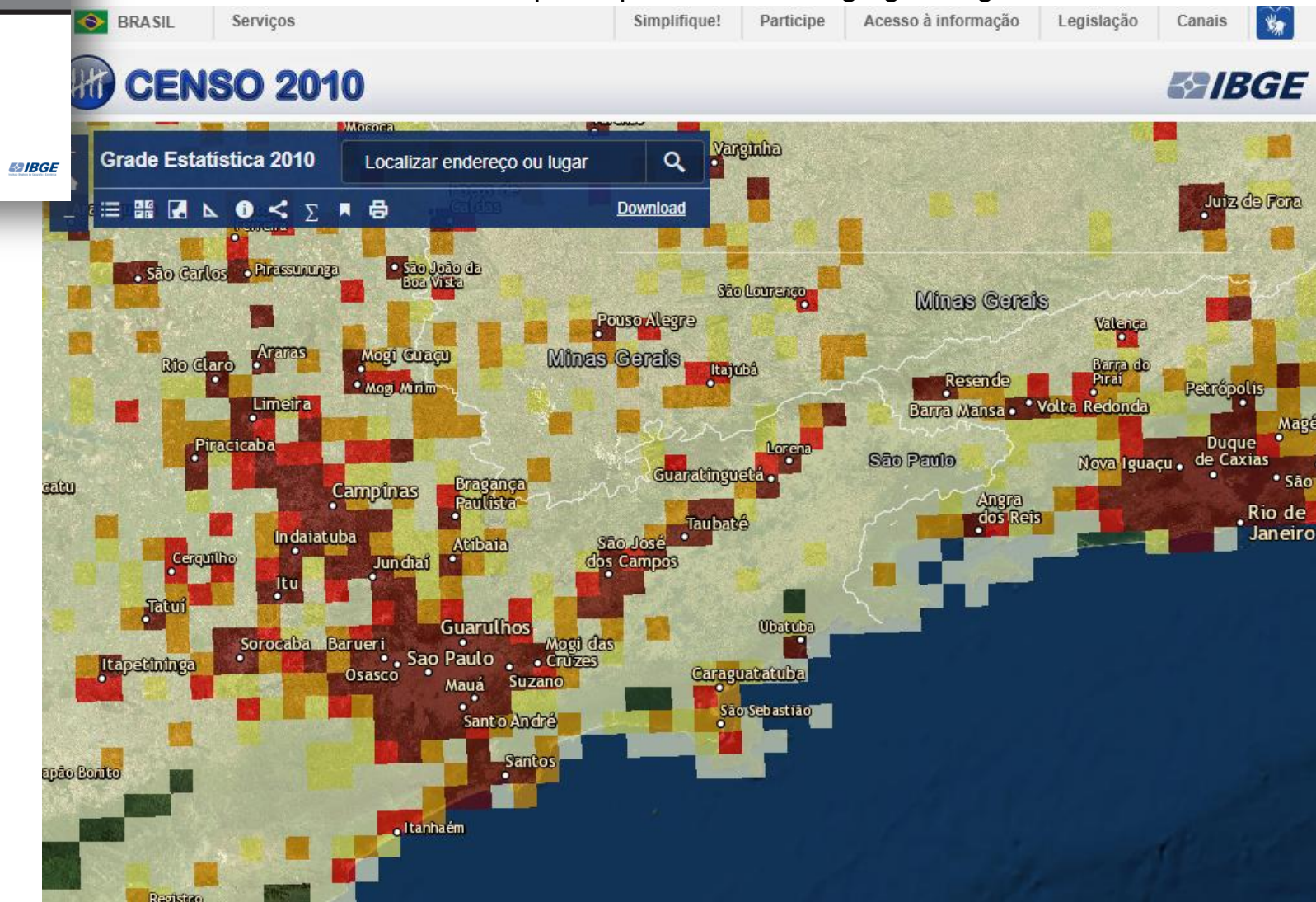
(5) População em domicílios com localização registrada em setores censitários urbanos.

(6) População total (rural + urbana). Sujeito à subenumeração, conforme o setor censitário.

(7) População total ajustada para minimizar a subenumeração.

GRADE
ESTATÍSTICA

<http://mapasinterativos.ibge.gov.br/grade/default.html>



Dados

Dados estatísticos - microdados do universo do Censo Demográfico 2010

Dados vetoriais - aos pontos de localização dos domicílios nas áreas rurais - CNEFE
linhas vetoriais das faces de logradouros e às divisões de setores censitários – Base Territorial.

Informação auxiliar - classificações de uso e cobertura das terras:

- Projeto de Monitoramento do Desmatamento dos Biomas Brasileiros por Satélite – PMDBBS, (imagens Landsat-5/TM - MMA 2007);
- Projeto TERRACCLASS 2010 – (Landsat-5/TM e MODIS – EMBRAPA e INPE

Classes agrupadas, sendo classificadas como “povoada” - com características antrópicas e como “não povoada” - de características naturais.

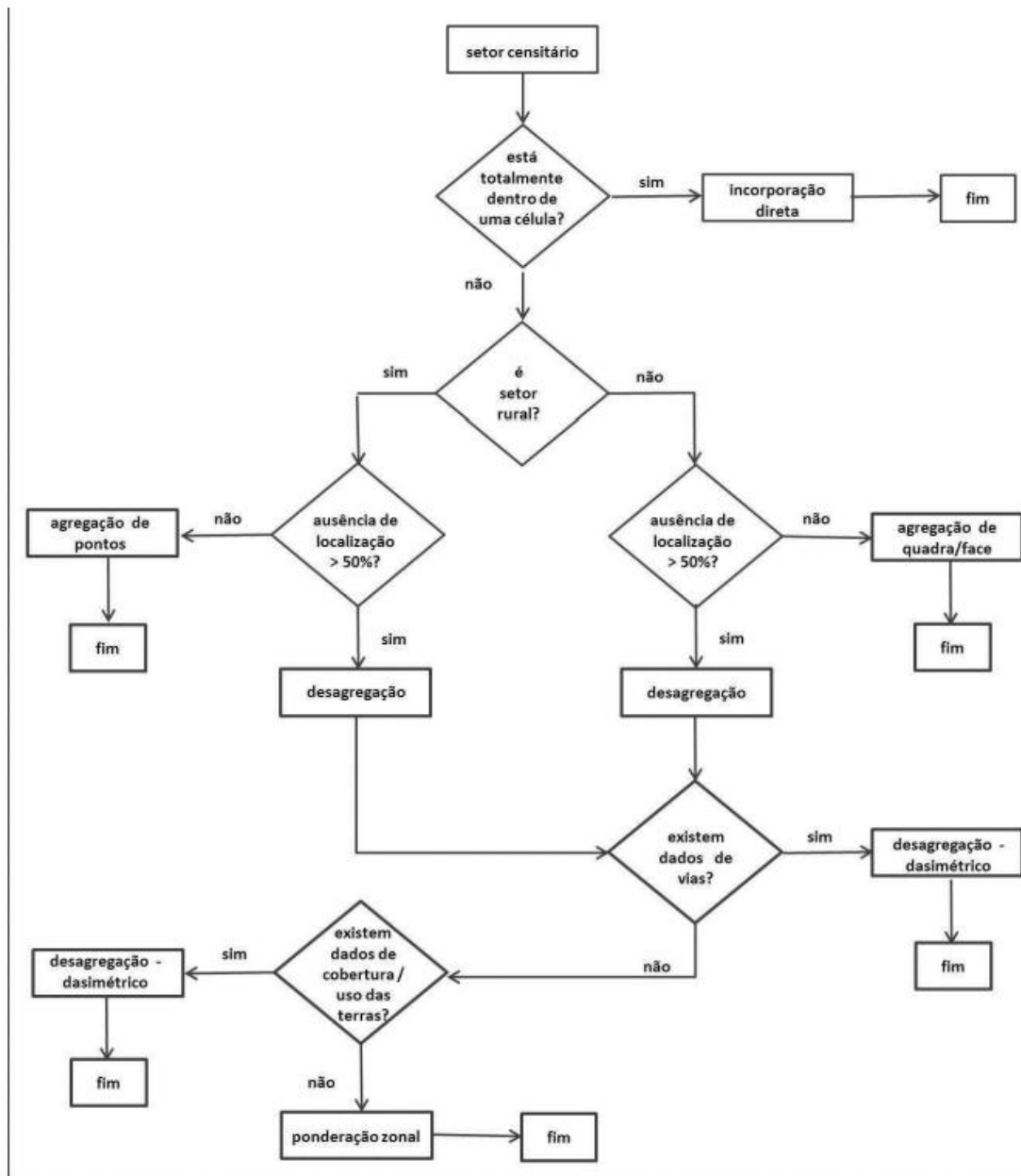
- Projeção Equivalente de Albers (produto final foi convertido para Projeção Geográfica)
- Datum horizontal SIRGAS2000.
- Células de 1 km x 1 km nas áreas rurais e 200 m x 200 m nas áreas urbanas.

Procedimento

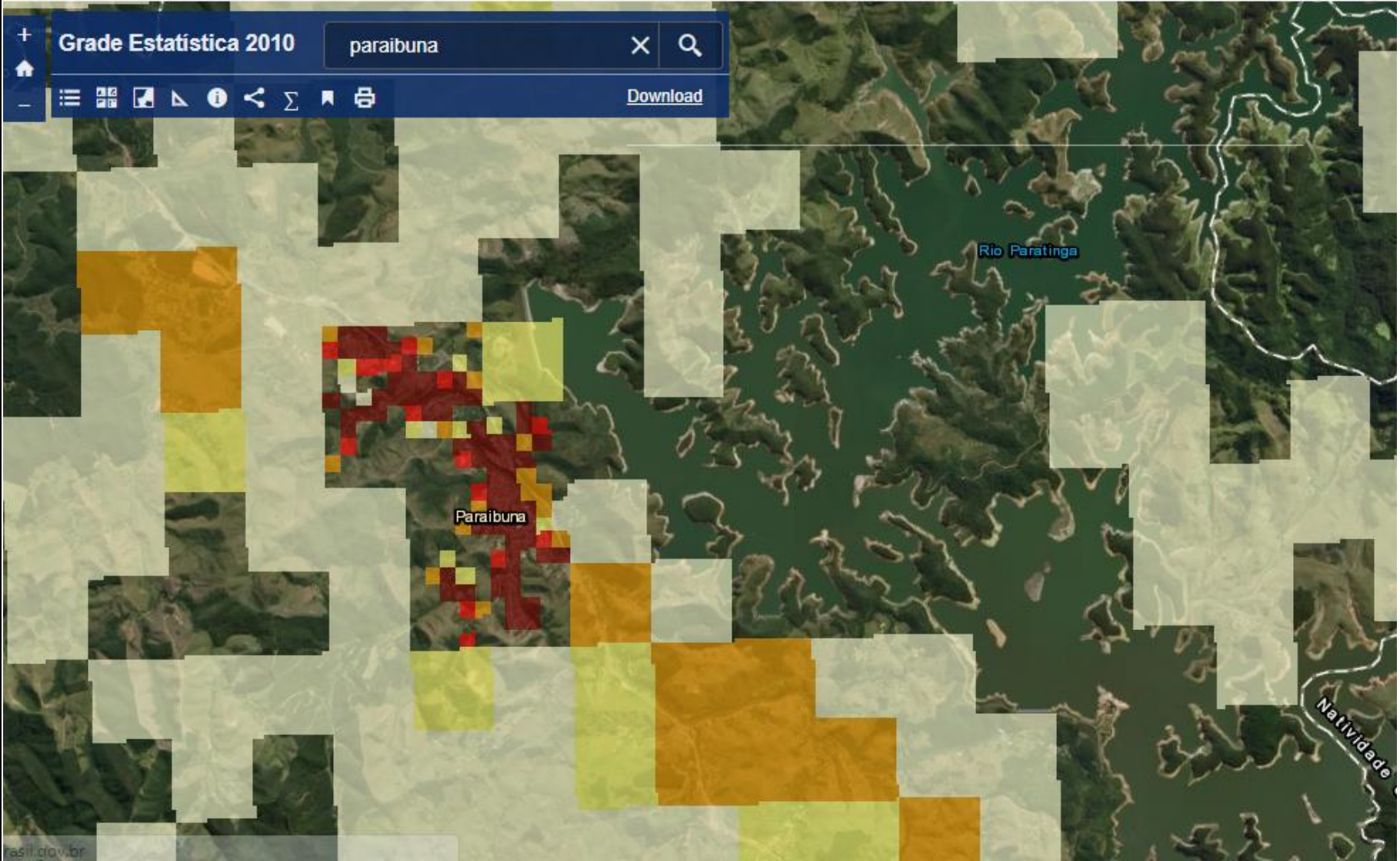
Árvore de decisão

para

Desagregação

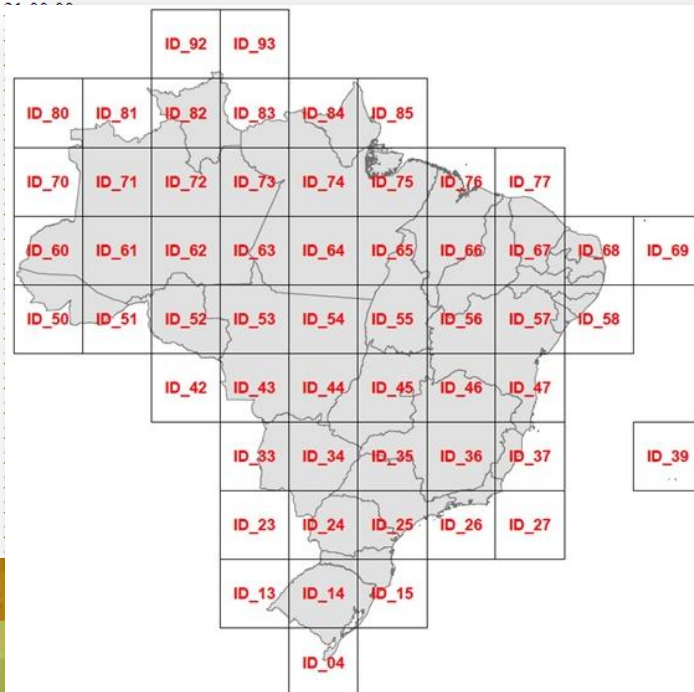


Fonte: Adaptado de BUENO, 2014.



 [\[diretório pai\]](#)

	Nome	Tamanho	Data da modificação
☐	articulacao.jpg	81.6 kB	31/05/2016 21:00:00
☐	grade_estadistica.pdf	771 kB	31/05/2016 21:00:00
☐	grade_id04.zip	3.7 MB	05/10/2016 21:00:00
☐	grade_id13.zip	1.8 MB	05/10/2016 21:00:00
☐	grade_id14.zip	31.1 MB	05/10/2016 21:00:00
☐	grade_id15.zip	17.1 MB	05/10/2016 21:00:00
☐	grade_id23.zip	155 kB	05/10/2016 21:00:00
☐	grade_id24.zip	22.3 MB	05/10/2016 21:00:00
☐	grade_id25.zip	46.9 MB	05/10/2016 21:00:00
☐	grade_id26.zip	23.8 MB	05/10/2016 21:00:00
☐	grade_id27.zip	156 kB	05/10/2016 21:00:00
☐	grade_id33.zip	8.0 MB	05/10/2016 21:00:00
☐	grade_id34.zip	20.6 MB	05/10/2016 21:00:00
☐	grade_id35.zip	31.6 MB	05/10/2016 21:00:00
☐	grade_id36.zip	36.8 MB	05/10/2016 21:00:00
☐	grade_id37.zip	11.8 MB	05/10/2016 21:00:00
☐	grade_id39.zip	161 kB	05/10/2016 21:00:00
☐	grade_id42.zip	608 kB	05/10/2016 21:00:00
☐	grade_id43.zip	16.5 MB	05/10/2016 21:00:00
☐	grade_id44.zip	20.0 MB	05/10/2016 21:00:00
☐	grade_id45.zip	25.9 MB	05/10/2016 21:00:00
☐	grade_id46.zip	20.7 MB	05/10/2016 21:00:00
☐	grade_id47.zip	15.4 MB	05/10/2016 21:00:00
☐	grade_id50.zip	5.3 MB	05/10/2016 21:00:00
☐	grade_id51.zip	8.9 MB	05/10/2016 21:00:00
☐	grade_id52.zip	17.5 MB	05/10/2016 21:00:00
☐	grade_id53.zip	20.4 MB	05/10/2016 21:00:00



Para Visitar...

- <http://mapasinterativos.ibge.gov.br/grade/default.html>
- <https://landscan.ornl.gov/>
- <http://sedac.ciesin.columbia.edu/data/collection/gpw-v4>
- <http://www.geog.leeds.ac.uk/people/a.turner/projects/medalus3/home.htm>
- <http://www.geog.ucsb.edu/~tobler/publications/reprints.html>
- http://www.worldpop.org.uk/data/data_sources/

Exemplos de aplicação:

- JONES E.E.; DeWITTE, S.N. (2012) Using spatial analysis to estimate depopulation for Native American populations in northeastern North America, AD 1616–1645. *Journal of Anthropological Archaeology*, 31 83–92.
- LIU, X. H.; KYRIAKIDIS, P. C.; GOODCHILD, M. F. Population-density estimation using regression and area-to-point residual kriging (2008) *International Journal of Geographical Information Science*. Vol. 22, No. 4, April 2008, 431–447
- YOO, E.; KYRIAKIDIS, P.C.; TOBLER, W. (2010) Reconstructing Population Density Surfaces from Areal Data: A Comparison of Tobler's Pycnophylactic Interpolation Method and Area-to-Point Kriging. *Geographical Analysis*, 42: 78–98.

Referências



- Tobler, W.R. (1979). Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association*, 74, 519-530.
- Tobler, W.R., Deichmann, U., Gottsegen, J. & Maloy, K. (1995). *The Global Demography Project*. Santa Barbara, CA: National Center for Geographic Information and Analysis.
- Martin, D (1989). Mapping population data from zone centroid locations. *Transactions of the Institute of British Geographers NS*, 14, 90-97.
- Martin, D. (1996). *Geographic Information Systems and their Socioeconomic Applications*, London: Routledge.
- Martin, D. (2002). Census Population Surfaces. In Rees, P., Martin, D., Williamson, P. (Eds.), *The Census Data System* (pp. 139-148). Chichester - England: John Wiley & Sons.
- Martin, D., Langford, M. & Tate, N.J. (2000). Refining Population Surface Models: Experiments with Northern Ireland Census Data. *Transactions in GIS*, 4, 343-360.
- Deichmann, U., Balk, D. & Yetman, G. (2001). *Transforming Population Data for Interdisciplinary Usages: From census to grid*. Palisades, N.Y.: Center for International Earth Science Information Network (CIESIN) - Columbia University.
- Dobson, J.E., Bright, E.A., Coleman, P.R., Duree, R.C. & Worley, B.A. (2000). LandScan: A Global Population Database for Estimating Populations at Risk. *Photogrammetric Engineering and Remote Sensing*, 66, 849-857.
- Faure, J.F., Tran, A., Gardel, A. & Polidori, L. (2003). *Sensoriamento remoto das formas de urbanização em aglomerações do litoral Amazônico: elaboração de um índice de densidade populacional*. Paper presented at XI Simpósio Brasileiro de Sensoriamento Remoto, Belo Horizonte, 5-10 Abril 2003.

Referências



- Wright, John K. 1936. A method of mapping densities of population with Cape Cod as an example. *Geographical Review* 26:103–10.
- Teresa Sá Marques, R.S.; Silva, F.B.; Delgado C. A ocupação edificada: delimitação de áreas de densidade homogênea. Departamento de Geografia, FLUP / CEGOT
<<http://repositorio-aberto.up.pt/handle/10216/19849>
- Filipe Batista e Silva. 2009. Modelação cartográfica e ordenamento do território: um ensaio metodológico de cartografia dasimétrica aplicado à região oeste e Vale do Tejo. Dissertação de Mestrado. Faculdade de Letras da Universidade do Porto
<http://repositorio-aberto.up.pt/handle/10216/18045>
- Langford, Mitchel and Higgs, Gary(2006) 'Measuring Potential Access to Primary Healthcare Services: The Influence of Alternative Spatial Representations of Population', *The Professional Geographer*, 58: 3, 294 — 306. DOI: 10.1111/j.1467-9272.2006.00569.x. < <http://dx.doi.org/10.1111/j.1467-9272.2006.00569.x>>