

Análise crítica do artigo

geographical analysis

Geographical Analysis (2013) 45, 324–344

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.

Mariane Souza Reis

População, Espaço e Ambiente

Professores: Dra. Silvana Amaral e Dr. Antônio Miguel Vieira Monteiro

9 de agosto de 2017

Sobre a revista

- Fundada em 1969.
- Publicada pela Wiley em nome do departamento de Geografia da Ohio State University.
- Publicações em teoria geográfica e construção de modelos.

Geographical Analysis

© The Ohio State University



Edited By: Rachel S. Franklin

Impact Factor: 1.905

ISI Journal Citation Reports © Ranking: 2016: 27/79 (Geography)

Online ISSN: 1538-4632

[http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1538-4632](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1538-4632)



Avaliação do triênio
2013-2016

Área de Avaliação	Classificação
ADMINISTRAÇÃO PÚBLICA E DE EMPRESAS, CIÊNCIAS CONTÁBEIS E TURISMO	A1
ECONOMIA	B4
ENGENHARIAS I	B2
ENGENHARIAS III	A2
INTERDISCIPLINAR	B1
MATEMÁTICA / PROBABILIDADE E ESTATÍSTICA	B2

<https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/veiculoPublicacaoQualis/listaConsultaGeralPeriodicos.jsf>

Sobre o autor



Mitchel Langford

<http://staff.southwales.ac.uk>

- *Reader* da Faculdade de Computação, Engenharia e Ciências da South Wales University;
- PhD. “*Applications of Digital Image Processing for Automation in Palynology*”, Hull University (1989);
- B.Sc Hons. Geografia física e geologia, Hull University (1984).



65
Research items

5,708
Reads

1,337
Citations

Research · 65

Article · 34

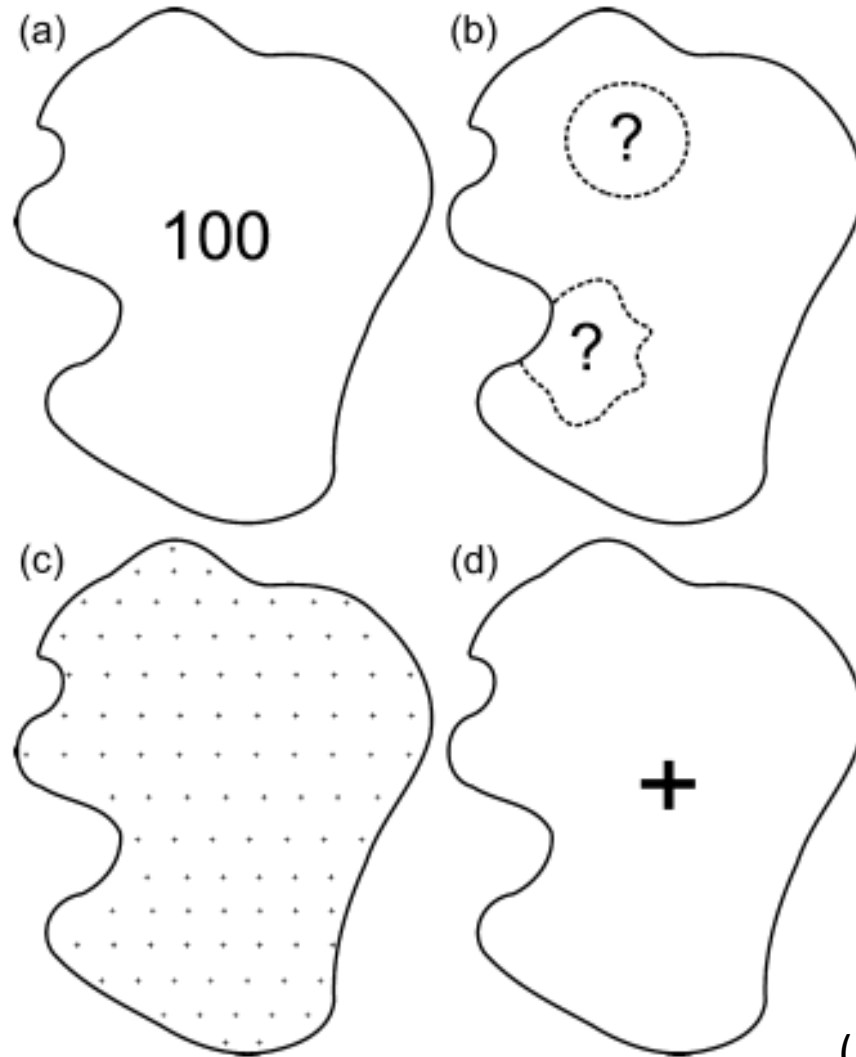
Chapter · 1

Conference Paper · 26

Data · 4

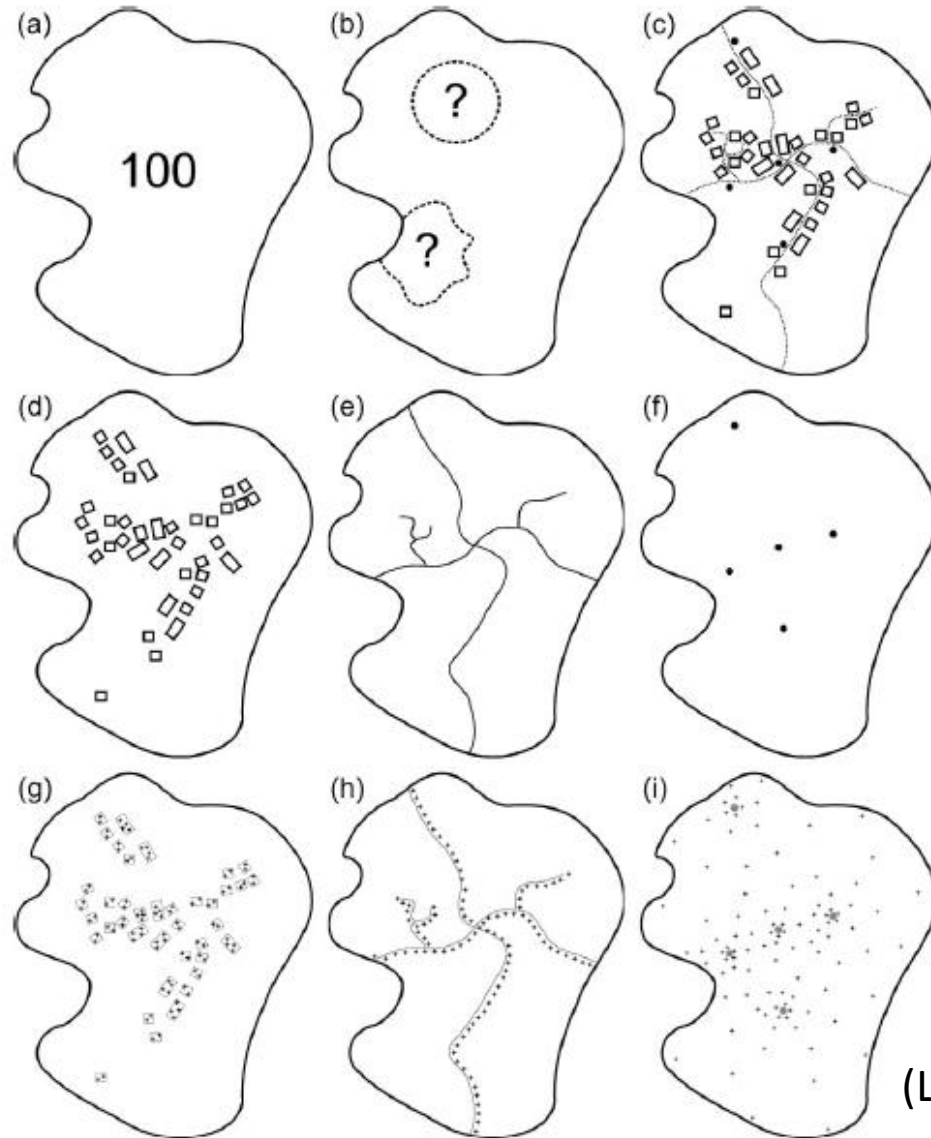
Full-texts · 29

Sobre o que se trata o artigo?



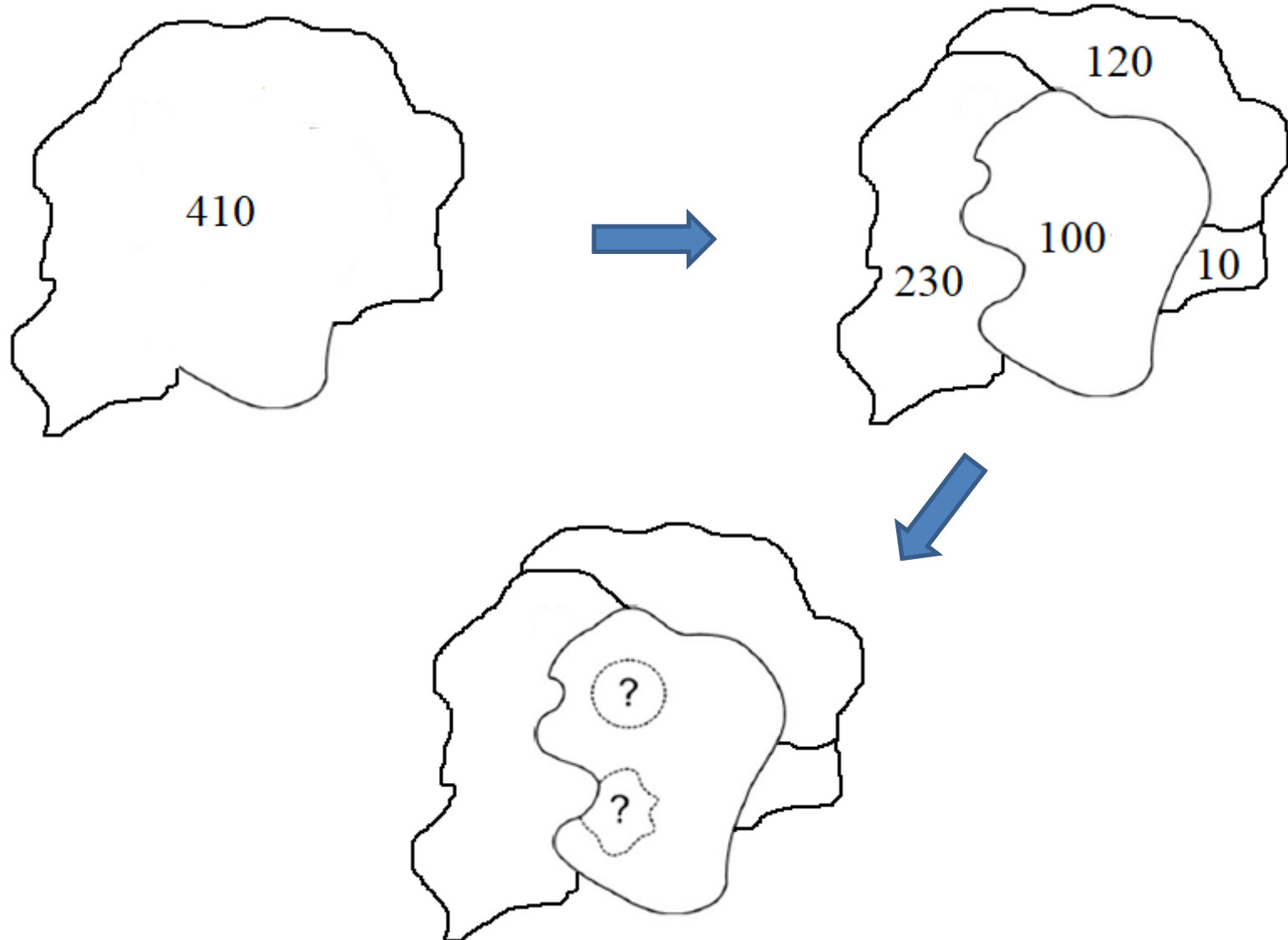
(Langford, 2013)

Sobre o que se trata o artigo?



(Langford, 2013)

Sobre o que se trata o artigo?



Objetivo

Avaliar a exatidão de métodos para estimar a população em “áreas pequenas”.



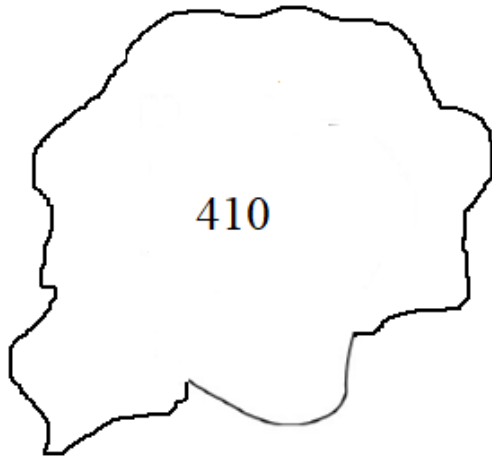
Avaliar dados auxiliares utilizados nos processos de interpolação: dados *open access*.



Contribuição para literatura.



Reamostragem: dados Censo 2001, Cardiff



Lower Super Output Areas (LSOAs):
•≈1500 pessoas.



Output Areas (OAs):
•≈300 pessoas ou 125 casas.



Output Areas (OAs):
•≈300 pessoas ou 125 casas.



Unit Postcodes (UPCs):
•≈15 casas.

Técnicas e dados utilizados

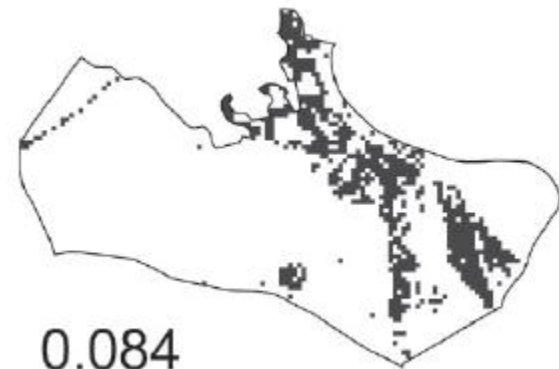
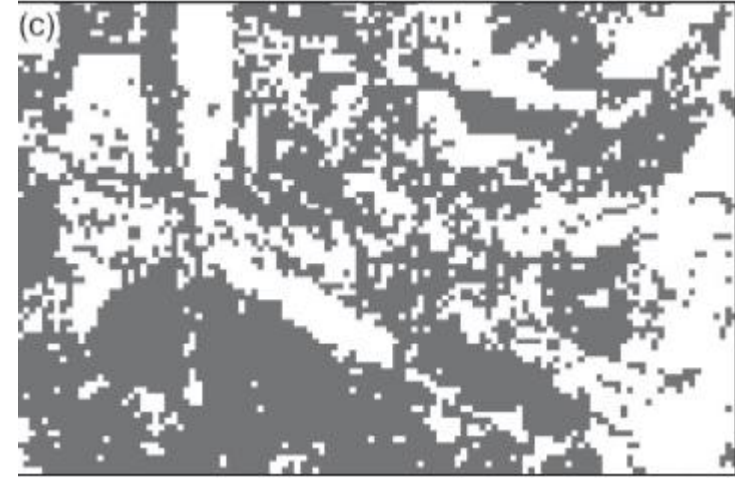
- Interpolação por área



Técnicas e dados utilizados

- Dasimétrico binário:

- Áreas habitadas (áreas residenciais) ou não habitadas segundo classificação de uma imagem ETM+ de 1999.



Modificado de Langford (2013)

Técnicas e dados utilizados

- Dasimétrico binário:
- Áreas habitadas (construções) ou não habitadas segundo polígonos de construções de dados *OS VectorMap District*.



Modificado de Langford (2013)

Técnicas e dados utilizados

- Ponderação por ruas:
- População realocada ao longo de ruas e rodovias.
- Experimentos com dados vetoriais, matriciais (5m) e *buffer* de 15m ao redor das vias.



0.410

Modificado de Langford (2013)

Técnicas e dados utilizados

- Superfícies de densidade:
 - Baseados no inverso da distância entre a “célula” e um ponto que sirva como indicador da concentração da população.



Modificado de Langford (2013)

Resultados obtidos

Método	Dado	LSOA - OA		AO - UPC	
		RMSE	Cov	RMSE	CoV
Interpolação por área	-	225,2	0,731	30,3	0,575
Superfície de densidade	Escolas primárias	220,4	0,715	35,5	0,674
	Pontos de ônibus	217,5	0,706	33,8	0,642
Ponderação por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (matriz)</i>	178,6	0,580	35,5	0,674
	<i>OS Vector Map District roads (buffer)</i>	170,5	0,553	30,3	0,574
Dasimétrico binário	<i>OS Vector Map District buildings (vetor)</i>	152,9	0,496	21,6	0,410
	<i>OS Vector Map District buildings (matriz)</i>	152,9	0,496	21,6	0,410
	Classificação ETM+	144,9	0,471	23,8	0,451

Resultados obtidos

Método	Dado	LSOA - OA		AO - UPC	
		RMSE	Cov	RMSE	CoV
Interpolação por área	-	225,2	0,731	30,3	0,575
Superfície de densidade	Escolas primárias	220,4	0,715	35,5	0,674
	Pontos de ônibus	217,5	0,706	33,8	0,642
Ponderação por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (matriz)</i>	178,6	0,580	35,5	0,674
	<i>OS Vector Map District roads (buffer)</i>	170,5	0,553	30,3	0,574
Dasimétrico binário	<i>OS Vector Map District buildings (vetor)</i>	152,9	0,496	21,6	0,410
	<i>OS Vector Map District buildings (matriz)</i>	152,9	0,496	21,6	0,410
	Classificação ETM+	144,9	0,471	23,8	0,451

Interpolação por área

Não há
necessidade de
dados auxiliares.



Supõe
homogeneidade
interna das zonas
alvo.



Piores resultados
para LSOAs-OAs.






Resultados obtidos

Método	Dado	LSOA - OA		AO - UPC	
		RMSE	Cov	RMSE	CoV
Interpolação por área	-	225,2	0,731	30,3	0,575
Superfície de densidade	Escolas primárias	220,4	0,715	35,5	0,674
	Pontos de ônibus	217,5	0,706	33,8	0,642
Ponderação por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (matriz)</i>	178,6	0,580	35,5	0,674
	<i>OS Vector Map District roads (buffer)</i>	170,5	0,553	30,3	0,574
Dasimétrico binário	<i>OS Vector Map District buildings (vetor)</i>	152,9	0,496	21,6	0,410
	<i>OS Vector Map District buildings (matriz)</i>	152,9	0,496	21,6	0,410
	Classificação ETM+	144,9	0,471	23,8	0,451

Superfície de densidade: escolas primárias

Supõe a existência de escolas primárias próximas à áreas de maior densidade populacional:

<p>Verdade para estudo realizado por Zhang e Qiu (2007) em Collin County.</p> 	<p>Modelo não explica distribuição populacional na área de estudo analisada (Cardiff).</p> 	<p>Resultados suportados por Bakilla et al. (2014) (Hamburgo).</p> 
---	---	--

Resultados obtidos

Método	Dado	LSOA - OA		AO - UPC	
		RMSE	Cov	RMSE	CoV
Interpolação por área	-	225,2	0,731	30,3	0,575
Superfície de densidade	Escolas primárias	220,4	0,715	35,5	0,674
	Pontos de ônibus	217,5	0,706	33,8	0,642
Ponderação por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (matriz)</i>	178,6	0,580	35,5	0,674
	<i>OS Vector Map District roads (buffer)</i>	170,5	0,553	30,3	0,574
Dasimétrico binário	<i>OS Vector Map District buildings (vetor)</i>	152,9	0,496	21,6	0,410
	<i>OS Vector Map District buildings (matriz)</i>	152,9	0,496	21,6	0,410
	Classificação ETM+	144,9	0,471	23,8	0,451

Superfície de densidade: pontos de ônibus

Supõe a existência de pontos de ônibus a uma certa distância dos centros de concentração populacional

Podem estar associados a outros destinos importantes.



Grande distância entre data dos dados.






Resultados obtidos

Método	Dado	LSOA - OA		AO - UPC	
		RMSE	Cov	RMSE	CoV
Interpolação por área	-	225,2	0,731	30,3	0,575
Superfície de densidade	Escolas primárias	220,4	0,715	35,5	0,674
	Pontos de ônibus	217,5	0,706	33,8	0,642
Ponderação por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (matriz)</i>	178,6	0,580	35,5	0,674
	<i>OS Vector Map District roads (buffer)</i>	170,5	0,553	30,3	0,574
Dasimétrico binário	<i>OS Vector Map District buildings (vetor)</i>	152,9	0,496	21,6	0,410
	<i>OS Vector Map District buildings (matriz)</i>	152,9	0,496	21,6	0,410
	Classificação ETM+	144,9	0,471	23,8	0,451

Ponderação por ruas

- Supõe densidade populacional homogênea ao longo das ruas;

<p>Resultados abaixo do esperado pelo autor: limites coincidentes entre zonas alvo e vias.</p> 	<p>Transformação para matriz atua como um pequeno <i>buffer</i>.</p> <p><i>Buffer</i> de 15m apresenta melhores resultados para o método.</p> 	<p>Dúvidas acerca da validade das delimitações utilizadas.</p> <p>Possíveis problemas com data dos dados.</p> <p>Uso de rodovias na análise.</p> 
--	--	--

Resultados obtidos

Método	Dado	LSOA - OA		AO - UPC	
		RMSE	Cov	RMSE	CoV
Interpolação por área	-	225,2	0,731	30,3	0,575
Superfície de densidade	Escolas primárias	220,4	0,715	35,5	0,674
	Pontos de ônibus	217,5	0,706	33,8	0,642
Ponderação por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (matriz)</i>	178,6	0,580	35,5	0,674
	<i>OS Vector Map District roads (buffer)</i>	170,5	0,553	30,3	0,574
Dasimétrico binário	<i>OS Vector Map District buildings (vetor)</i>	152,9	0,496	21,6	0,410
	<i>OS Vector Map District buildings (matriz)</i>	152,9	0,496	21,6	0,410
	Classificação ETM+	144,9	0,471	23,8	0,451

Dasimétrico binário: *OS Vector Map District buildings*

Resultados
idênticos para
vetor e matriz?



Segundo melhor
resultado para
LSOAs-Oas.

Melhor resultado
para OAs-UPCs.



Não há distinção
entre tipos de
construções.

Possíveis
problemas com
data dos dados.



Resultados obtidos

Método	Dado	LSOA - OA		AO - UPC	
		RMSE	Cov	RMSE	CoV
Interpolação por área	-	225,2	0,731	30,3	0,575
Superfície de densidade	Escolas primárias	220,4	0,715	35,5	0,674
	Pontos de ônibus	217,5	0,706	33,8	0,642
Ponderação por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (matriz)</i>	178,6	0,580	35,5	0,674
	<i>OS Vector Map District roads (buffer)</i>	170,5	0,553	30,3	0,574
Dasimétrico binário	<i>OS Vector Map District buildings (vetor)</i>	152,9	0,496	21,6	0,410
	<i>OS Vector Map District buildings (matriz)</i>	152,9	0,496	21,6	0,410
	Classificação ETM+	144,9	0,471	23,8	0,451

Dasimétrico binário: classificação de imagem ETM+

- Supõe população uniformemente distribuída em áreas “habitadas”.

Baseado na
classificação de
uma imagem
ETM+ de 1999.



Melhor resultado
para LSOAs-Oas.
Segundo melhor
resultado para
OAs-UPCs.



Não há
informações
sobre
metodologia de
classificação ou
exatidão dos
resultados.



Dasimétrico binário: classificação de imagem ETM+

- Autor sugere uso de imagens de maior resolução espacial.

Jega et al. (2017):
imagens com
30m (EG=87,89%),
10m (EG=83,20%) e
3m (EG=82,03%).



Dependente de imagens
e técnicas utilizadas.



Conclusões

Diferentes dados auxiliares levam a diferentes resultados.

Lógica por trás da escolha dos métodos e dados auxiliares.

Resultados dependem da área de estudo.



Resultados suportados por Zandbergen e Ignizio (2010).



Conclusões

Informação de cobertura derivada da classificação da imagem LANDSAT ETM+ trata-se de um dos melhores candidatos a dados auxiliares para métodos inteligentes de interpolação.

Informações de uso necessárias.

Imagens oferecem informações de cobertura.



Reprodutibilidade da metodologia.

Influência da qualidade da classificação.



Resultados comparáveis encontrados em Jega et al. (2017) para Leicester.



Conclusões

Pelo menos no contexto do Reino Unido, uso de dados *open access* contendo informações sobre construções é preferível.

Resultados variam segundo área de estudo.

É possível afirmar para outras áreas?



Não há considerações sobre mudanças na região ao longo dos anos.

Para quais datas a afirmação é válida?



Conclusões

Resultados apresentados relacionam-se apenas à uma área de estudo predominantemente urbanizada. Estudos em áreas rurais esparsamente povoadas oferecem um ambiente mais desafiador.

Quais as lógicas válidas para essa áreas?



Quais dados auxiliares poderiam ser usados?



Conclusões

Resultados apresentados apenas para interpolação entre diferentes resoluções e não entre diferentes geografias em uma mesma resolução.

Qual a implicação?



Conclusões

Estudo contribui para o aumento da literatura sobre o assunto. Atividade necessária pela quantidade de resultados conflitantes.

Aumento da disponibilidade de dados *open access* oferece maior potencial de uso de ferramenta de interpolação inteligente da população.

Adoção de métodos inteligentes seria incentivada com a disponibilidade de ferramentas implementadas em Sistemas de Informação Geográficas comumente utilizados.

Considerações sobre o trabalho

Objetivos propostos foram cumpridos.

Autor se preocupa em utilizar dados com pouca ou nenhuma modificação.



Falta de uma segunda área de estudo.

Considerações sobre outras áreas são vagas.

Não menciona aplicabilidade para outras datas.



Considerações sobre o trabalho

Avaliação individual dos dados permite avaliar o potencial de cada um.



Não avalia o potencial dos dados de forma conjunta (Su et al., 2010; (Amaral et al., 2012).

Não utiliza métodos dasimétricos multi-classes (Su et al., 2010 e Bakillah et al., 2014)



Considerações sobre o trabalho

Índices utilizados apenas para comparação de métodos.

Não avalia utilidade dos resultados gerados.



Não há comparação com outros valores encontrados na literatura.



Valores similares obtidos por Jega et al. (2017)



Considerações sobre o trabalho

Não menciona melhorias no resultado com o refinamento da escala de análise.



Exclui da análise as “ruas verticais”.

Não apresenta esses resultados.



Problema revisitado por autores que citam esse trabalho.



Repercussões do trabalho

[An evaluation of small area population estimation techniques using open access ancillary data](#)

[PDF] [researchgate.net](#)

M Langford - [Geographical Analysis](#), 2013 - [Wiley Online Library](#)

Abstract National census data represent the "gold standard" for authoritatively portraying a country's residential population distribution, but their aggregated counts for fixed administrative areas present problems for many geographic information system (GIS) analyses. Intelligent areal interpolation algorithms assist by transferring data from one zonal system to another using ancillary data to improve accuracy. All areal interpolation ...

Citado por 26 [Artigos relacionados](#) [Todas as 3 versões](#) [Citar](#) [Salvar](#)

<https://scholar.google.com.br/>

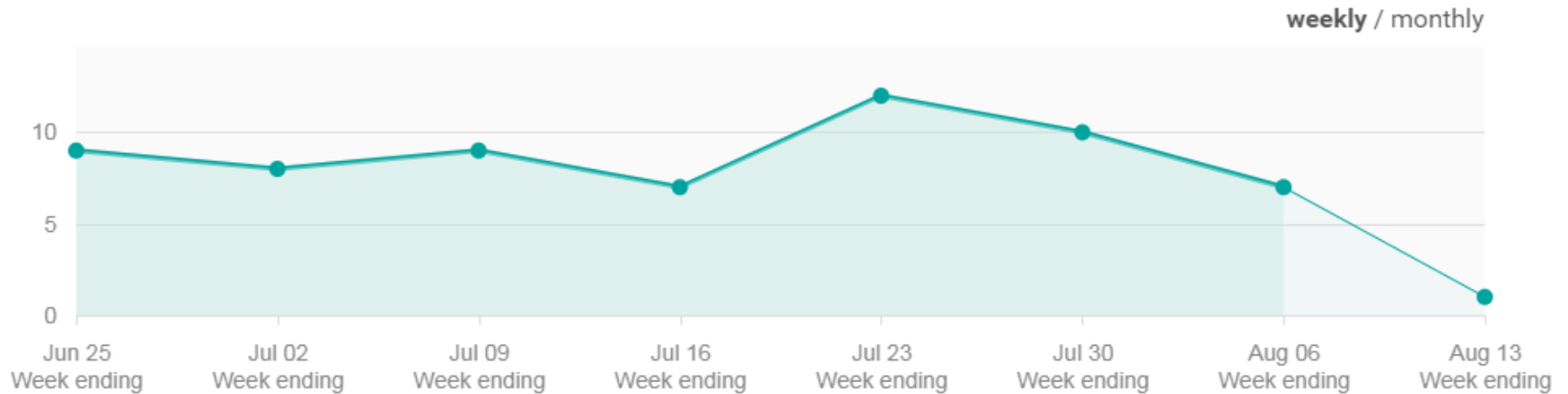


21 Citations

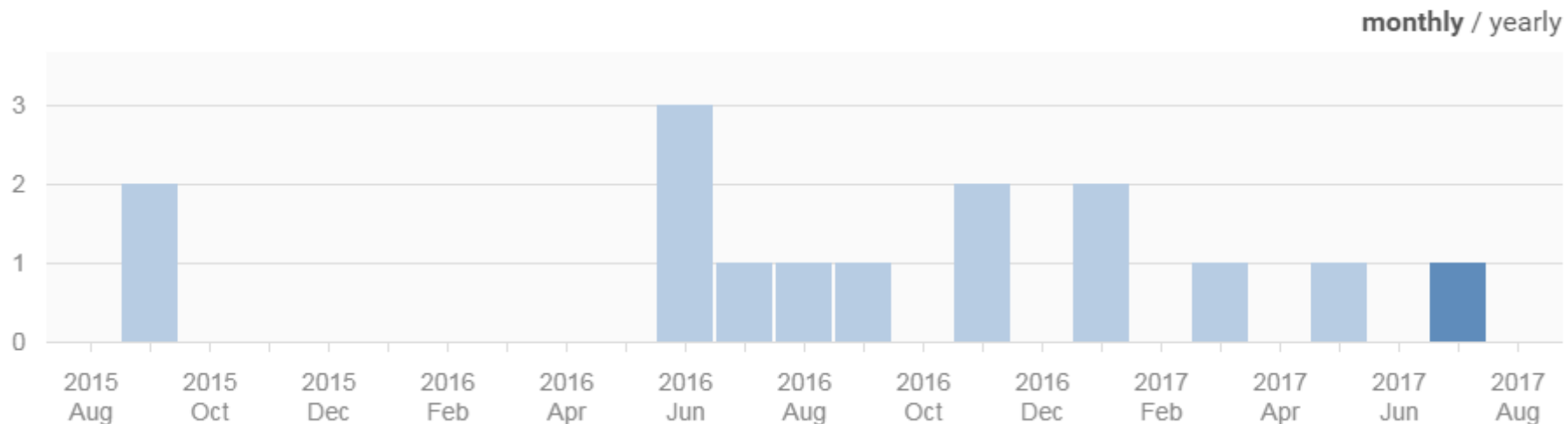
- ☹️ Editorial da revista em que foi publicado
- 😊 2 teses de doutorado (1 em japonês);
- ☹️ Pelo menos 4 trabalhos listados incorretamente.
- 😊 Trabalhos citam questionamentos levantados no artigo, baseiam-se na metodologia e/ou comparam os resultados obtidos.
- 😊 Artigo mais lido da última semana do perfil do autor no ResearchGate.

Repercussões do trabalho

258 Reads



21 Citations 21 Excluding self-citations



Grata pela atenção

Mariane Souza Reis

Trabalhos citados na apresentação:

- AMARAL, S.; GAVLAK, A. A., ESCADA, M. I. S.; MONTEIRO, A. M. V. Using remote sensing and census tract data to improve representation of population spatial distribution: case studies in the Brazilian Amazon. **Population and Environment**, v. 34, n.1, pp. 142–170, 2012.
- BAKILLA, M.; LIANG, S.; MOBASHERI, A.; ARSANJANI, J.J.; ZIPF, A. Fine-resolution population mapping using OpenStreetMap points-of-interest. **International Journal of Geographical Information Science**, v. 28, n. 9, pp.1940-1963, 2014.
- JEGA, I.M.; COMBER, A.J.; TATE, N.J. A comparison of methods for spatial interpolation across different spatial scales. **SSRG International Journal of Geo Informatics and Geological Science**, v.4, n.3, pp. 12-22, 2017.
- LANGFORD, M. An evaluation of small area population estimation techniques using open access ancillary data. **Geographical Analysis**, v.45, pp.324-344, 2013.
- SU, M.; LIN, M.; HSIEH, H.; TSAI, B.; LIN, C. Multi-layer multi-class dasymetric mapping to estimate population distribution. **Science of the Total Environment**, v. 208, pp. 4807-4816, 2010.
- ZANDBERGEN, P.A.; IGNIZIO, D. Comparison of Dasymetric Mapping Techniques for Small-Area Population Estimates. **Cartography and Geographic Information Science**, v. 37, n. 3, pp. 199-241, 2010.
- ZHANG, C.; QIU,F. A point-based intelligent approach to areal interpolation. **The Professional Geographer**, v.63, pp 262-276, 2011.