

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.

População, Espaço e Ambiente

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

Aluna: Mariane Souza Reis

7 de agosto de 2017

Sobre a revista

- Fundada em 1969;
- Publicada pela Wiley em nome do departamento de Geografia da Universidade do Estado de Ohio;
- Publicações em Teoria geográfica, construção de modelos e métodos quantitativos aplicados a geografia e campos relacionados, com atenção a análise de dados espaciais.

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, Pontypridd, Wales, Reino Unido;
 - PhD. “*Applications of Digital Image Processing for Automation in Palynology*”, Departamento de Engenharia Eletrônica, Hull University (1989);
 - B.Sc Hons. Geografia física e geologia, Hull University (1984);
- Possui trabalhos na área de modelagem da distribuição da população, computação e modelagem analítica espacial, bancos de dados geográficos, engenharia de softwares, processamento de imagens, entre outros;



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 artigo

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

[PDF] [researchgate.net](#)

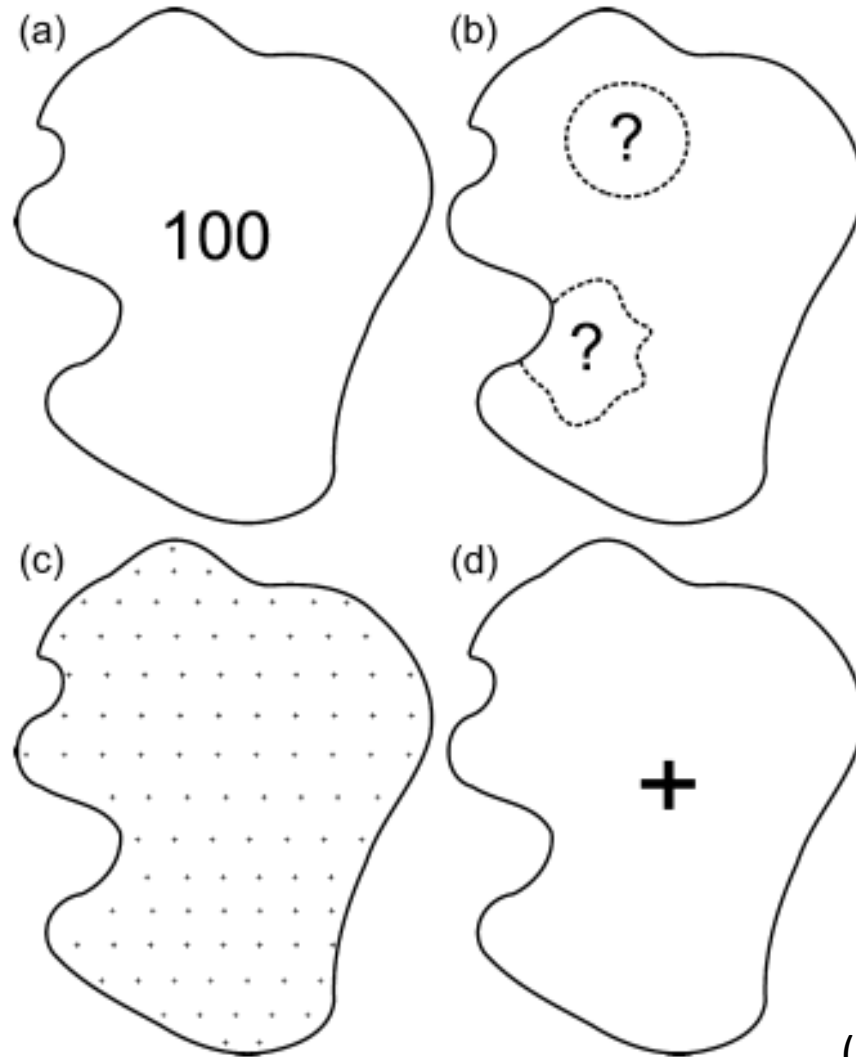
M Langford - [Geographical Analysis](#), 2013 - [Wiley Online Library](#)

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](#)

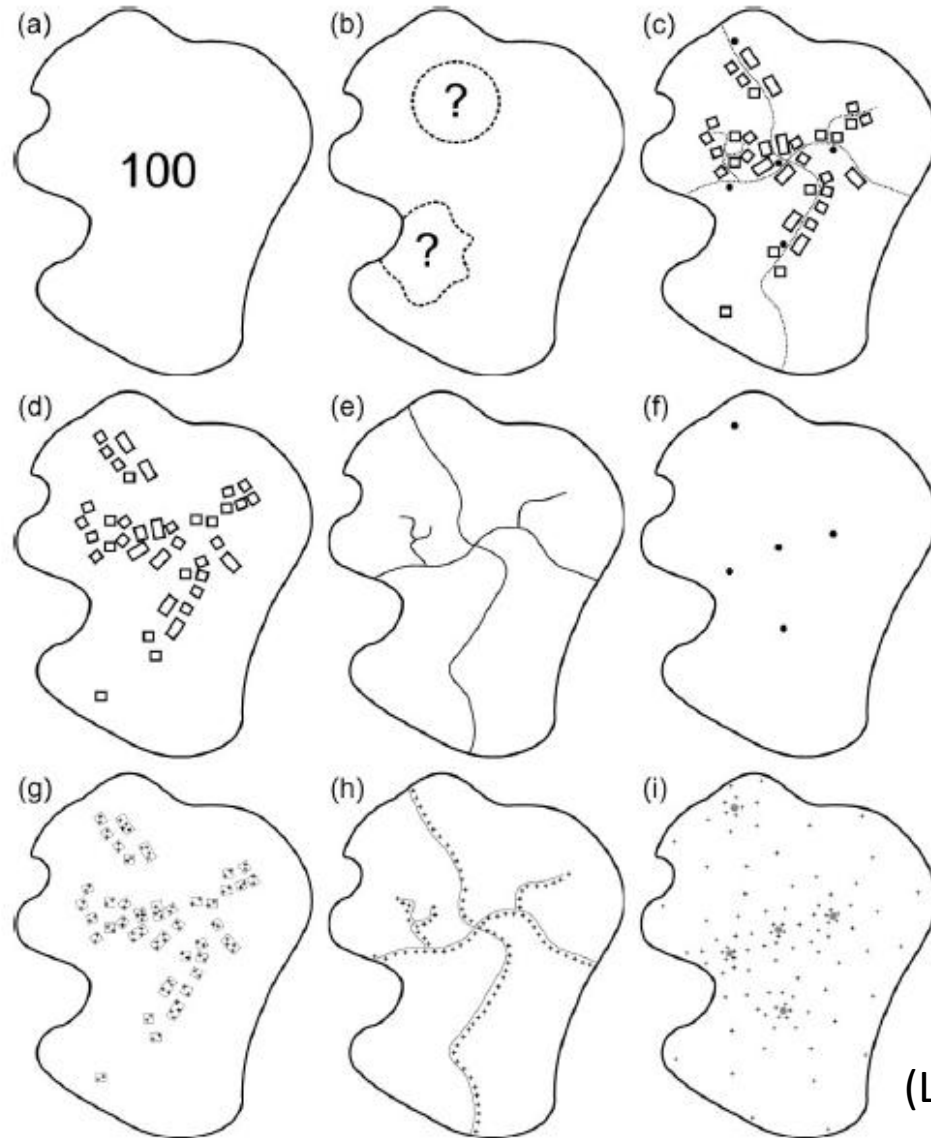
<https://scholar.google.com.br/>

Sobre o que se trata o artigo?



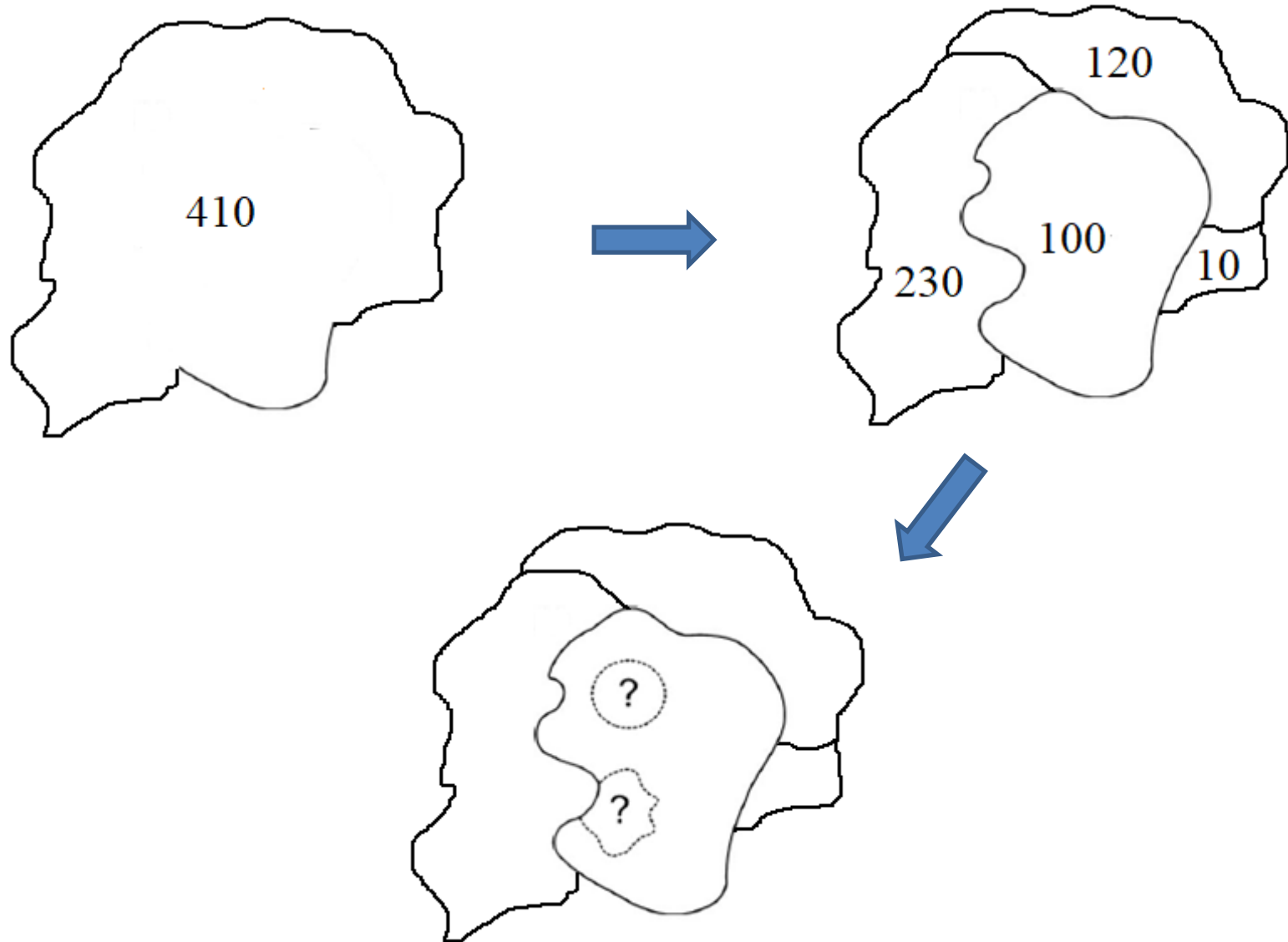
(Langford, 2013)

Sobre o que se trata o artigo?



(Langford, 2013)

Sobre o que se trata o artigo?



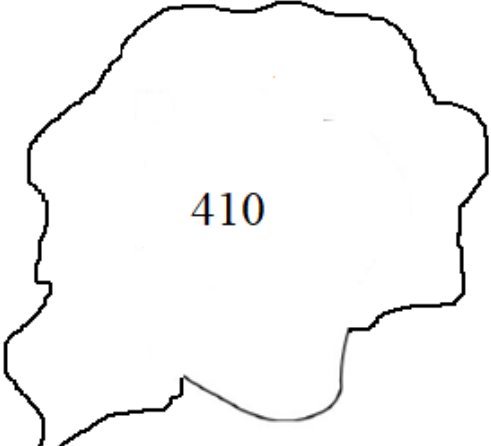
Objetivo

Contribuição para o conjunto de estudos que comparam o desempenho de métodos “inteligentes” de “interpolação por área”

Avaliar a exatidão dos métodos testados para estimar a população em “áreas pequenas”, pelo uso de dados de contagem de UPCs

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

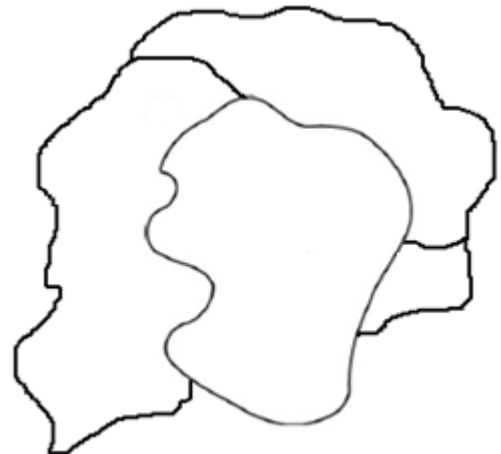
Estratégias de reamostragem: dados do Censo 2001



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.

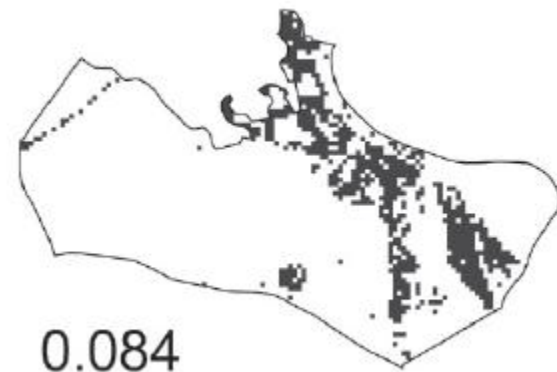
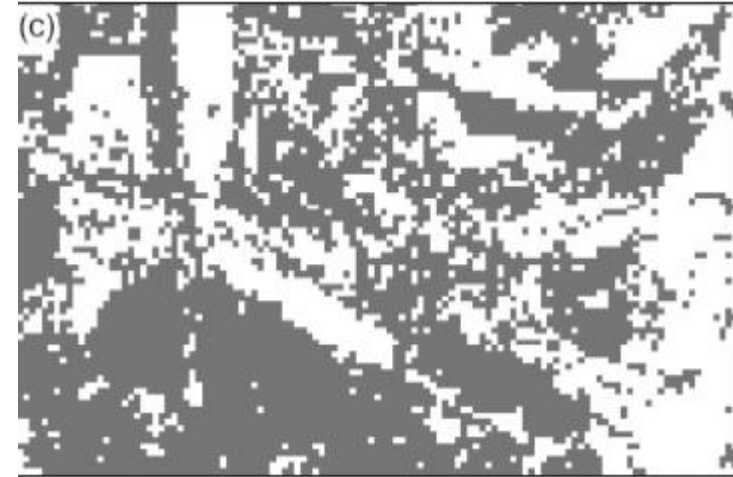
Técnicas e dados utilizados

- Interpolação por área



Técnicas e dados utilizados

- Dasimétrico binário:
- Áreas povoadas (áreas residenciais) ou não povoadas segundo classificação de uma imagem ETM+ de 1999.



Modificado de Langford (2013)

Técnicas e dados utilizados

- Dasimétrico binário:

- Áreas povoadas (construções) ou não povoadas segundo polígonos de construções de dados *OS VectorMap District*.



Modificado de Langford (2013)

Técnicas e dados utilizados

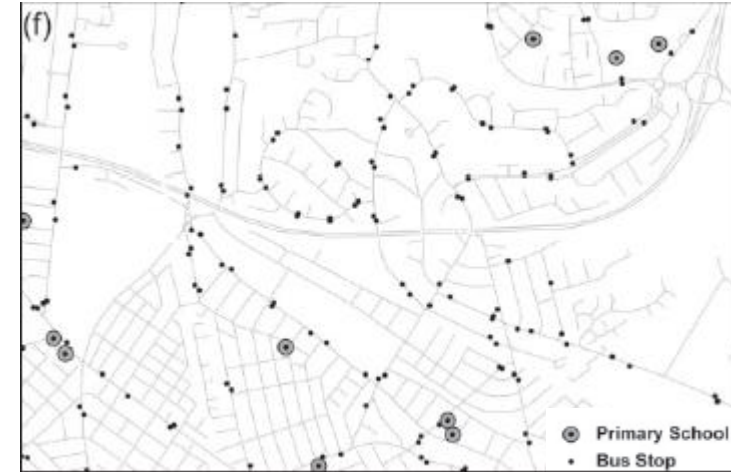
- Ponderamento por ruas:
- População realocada ao longo de ruas e rodovias;
- Experimentos com dados vetoriais, matriciais (5m) e buffer de 15m ao redor das vias.



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 *proxy* da concentração da população.



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
Ponderamento por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (raster)</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 (raster)</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
Ponderamento por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (raster)</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 (raster)</i>	152,9	0,496	21,6	0,410
	Classificação ETM+	144,9	0,471	23,8	0,451

Interpolação por área

- Não há a necessidade de dados auxiliares;
- Supõe homogeneidade interna das zonas alvos;
- LSOAs – OAs: piores resultados;
- OAs – UPCs: melhor que superfícies de densidade e ponderamento por ruas (vetor e raster).

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
Ponderamento por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (raster)</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 (raster)</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:
 - pessoas preferem morar perto das escolas;
 - escolas construídas para minimizar a distância percorridas pelos estudantes;
- Verdade para estudo realizado em Collin County (Zhang e Qiu, 2011);
- Modelo não explica distribuição populacional na área de estudo analisada (Cardiff):
 - Qualidade do dado auxiliar;
 - Aderência do modelo;
 - Validade da suposição.

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
Ponderamento por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (raster)</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 (raster)</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 as datas 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
Ponderamento por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (raster)</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 (raster)</i>	152,9	0,496	21,6	0,410
	Classificação ETM+	144,9	0,471	23,8	0,451

Ponderamento por ruas

- Supõe densidade populacional homogênea ao longo das ruas;
- Segundo o autor: resultados abaixo do esperado;
 - possivelmente relacionados com o fato dos limites entre zonas e ruas coincidirem;
 - transformação para *raster* atua como operação de *buffer*;
 - *buffer* de 15m apresentam os melhores resultados;
- **Dúvidas acerca da validade das delimitações das zonas utilizadas;**
- **Possíveis problemas com data dos dados;**
- **Não faz considerações a respeito de eliminar rodovias da análise.**

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
Ponderamento por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (raster)</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 (raster)</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 *raster*?
- LSOAs – OAs: segundo melhor resultado;
- OAs – UPCs: melhor resultado;

- **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
Ponderamento por ruas	<i>OS Vector Map District roads (vetor)</i>	180,9	0,587	39,3	0,745
	<i>OS Vector Map District roads (raster)</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 (raster)</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 “povoadas”;
- Baseado na classificação de uma imagem ETM+ de 1999 (diferença de dois anos do dado censitário);
- LSOAs – OAs: melhor resultado;
- OAs – UPCs: segundo melhor resultado.
- **Não fornece informações acerca da metodologia de classificação, exatidão dos resultados ou da escolha da imagem (data);**
- **Não justifica o não uso de uma imagem de data mais próxima ou avalia possíveis mudanças ocorridas no período.**

Dasimétrico binário: classificação de imagem ETM+

- Autor sugere capacidade de separar tipos de uso para melhor estimar regiões povoadas e que dados de resolução espacial mais fina apresentariam melhores resultados;
- **Jega et al. (2017): testes com classificações de imagens com 30m (EG=87,89%), 10m (EG=83,20%) e 3m (EG=82,03%). Melhores resultados com 30m;**
- **Afirmação dependente das imagens e técnicas utilizadas.**

Conclusões

Diferentes dados auxiliares levam a diferentes resultados.

- Importante considerar a lógica por trás da escolha de métodos e dados auxiliares;
- Resultados podem variar dependendo da área de estudo.

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.

- Classificação não se trata de cobertura, mas sim de uso da terra: áreas residenciais e não residenciais;
- Dúvidas sobre reprodutibilidade da metodologia, não propriamente descrita;
- Dúvidas acerca do efeito da exatidão da classificação nos resultados obtidos.

Conclusões

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

- Conforme observado por Zandbergen e Ignizio (2010) e relatado pelo próprio autor, o desempenho de diferentes dados e metodologias podem variar dependendo da área de estudo:
 - É possível afirmar a utilidade desse dado para todo o Reino Unido? Para quais anos de análise?

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 lógicas se aplicariam a essas regiões? Quais os dados auxiliares potencialmente utilizados?

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.