População, Espaço e Ambiente

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

Parte V – Estimativa de População Inferência por Sensoriamento Remoto

> Silvana Amaral Antonio Miguel V. Monteiro

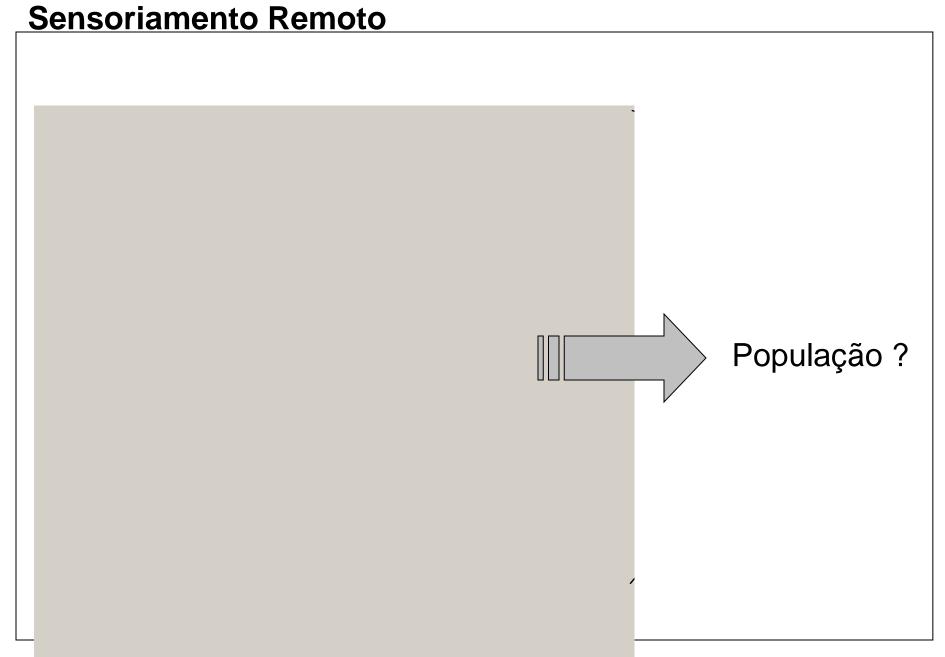
{silvana@dpi.inpe.br, miguel@dpi.inpe.br}



Estimação de população por Sensoriamento Remoto



Estimação de população de pequenas áreas por Sensoriamento Remoto

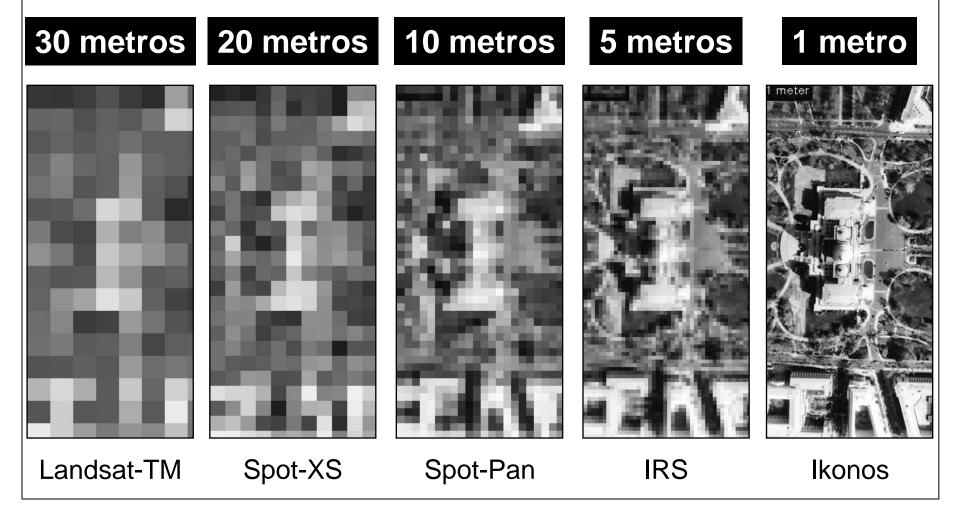


Estimação de população de pequenas áreas por Sensoriamento Remoto



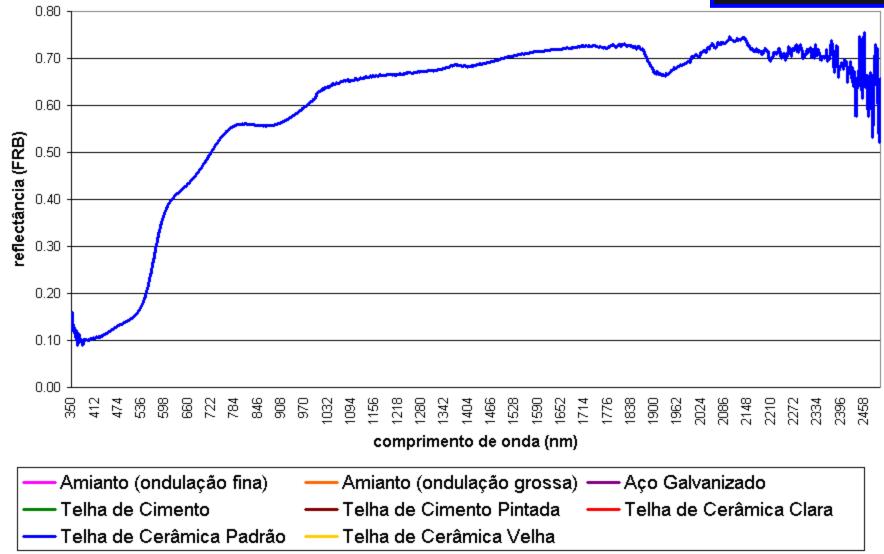
Avanço tecnológico → novos sensores → resolução espacial, temporal, radiométrica e espectral dos dados

Novas possibilidades de uso de imagens orbitais para estudos urbanos.



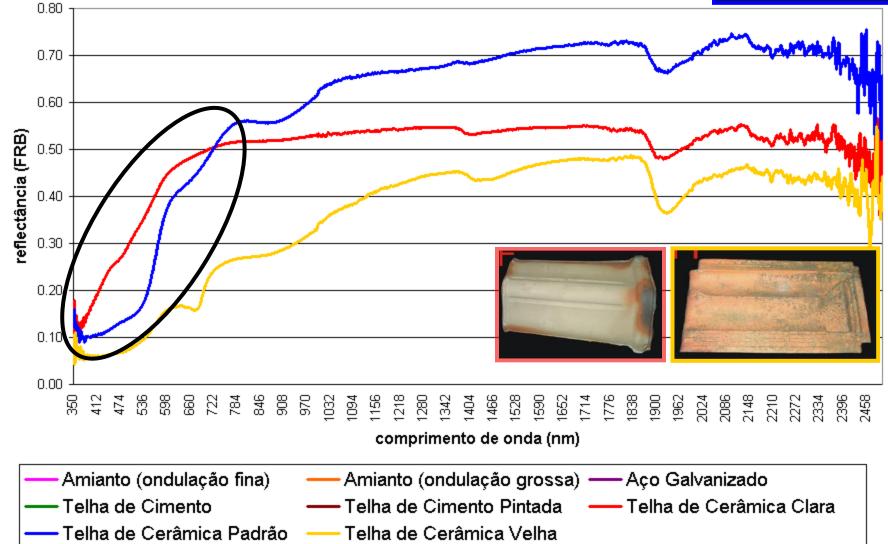






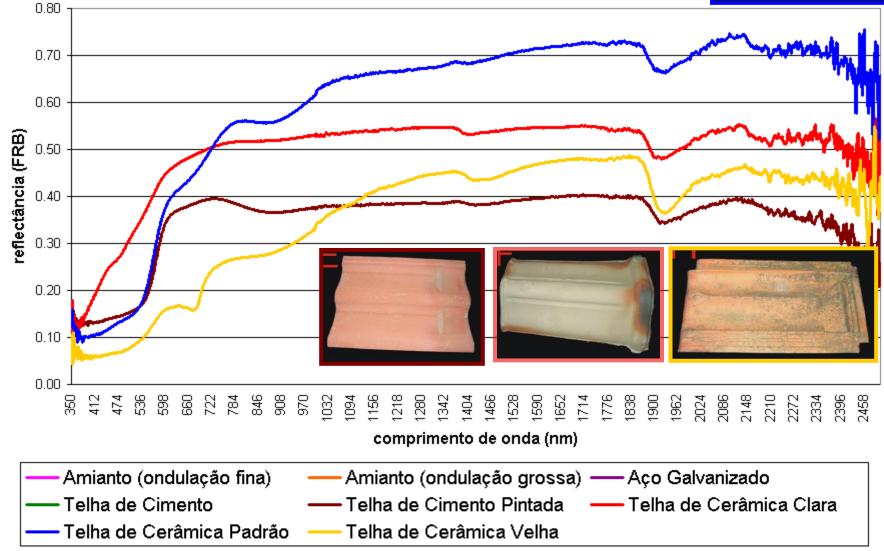










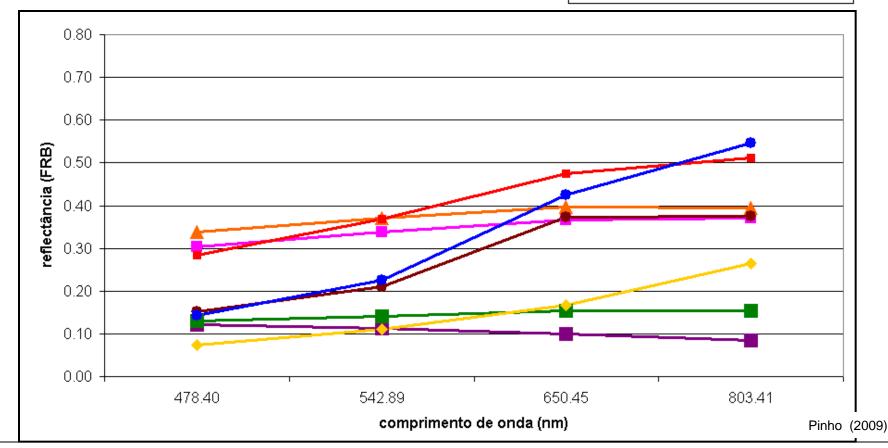


Espectros reamostrados para as bandas do sensor Quickbird (61cm - P / 2.4m - XS)

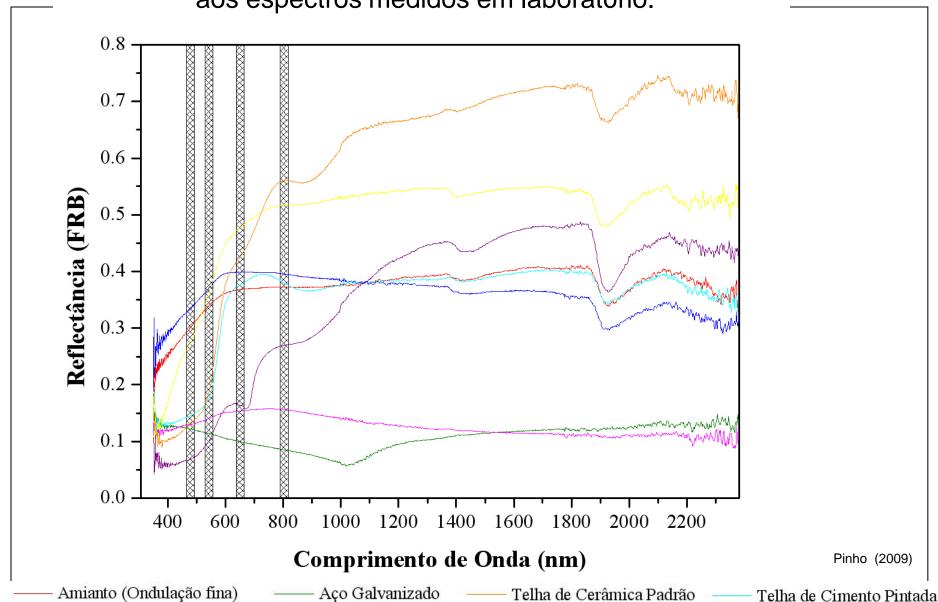
--- Telha de Cerâmica Clara

--- Telha de Cerâmica Padrão

🔶 Telha de Cerâmica Velha



Posicionamento de bandas do sensor Quickbird em relação aos espectros medidos em laboratório.



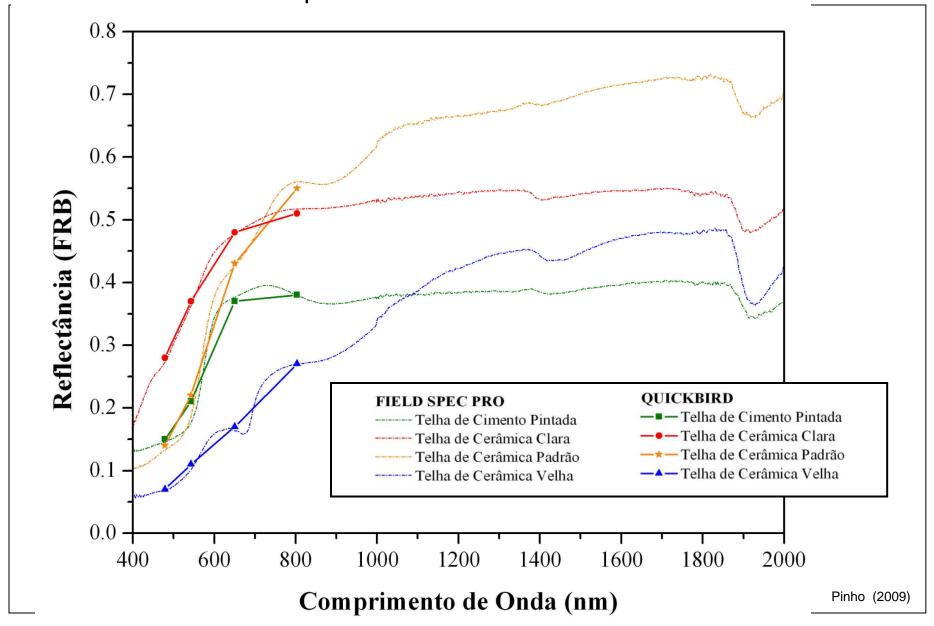
Telha de Cerâmica Velha

Telha de Cerâmica Clara

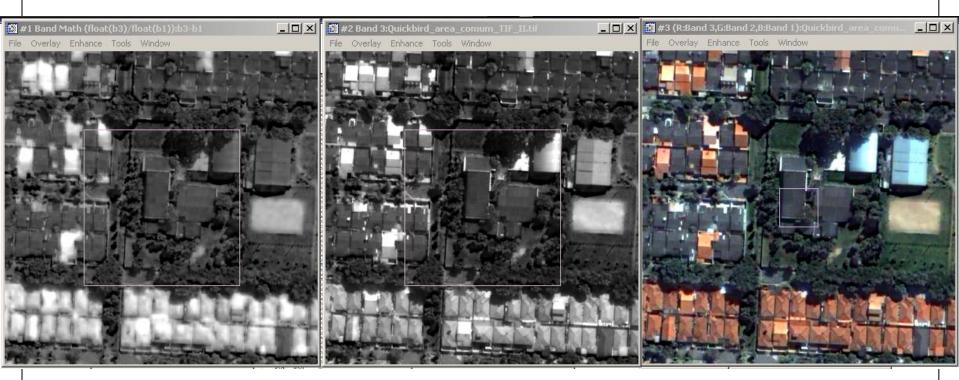
Telha de Cimento —

Amianto (Ondulação grossa)

Espectros de Telhas Cerâmicas obtidos em Laboratório e simulados para as bandas do Quickbird



Extração de Telhas Cerâmicas



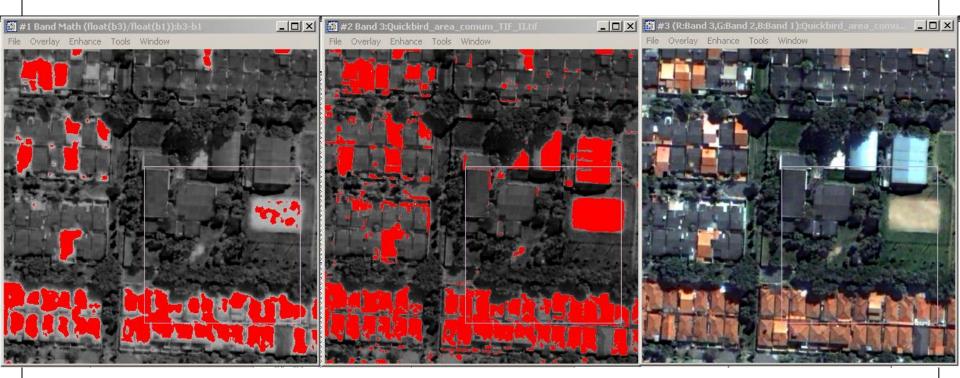
Razão de bandas

Banda 3

Composição colorida cor verdadeira

Extração de Telhas Cerâmicas

Pinho (2009)



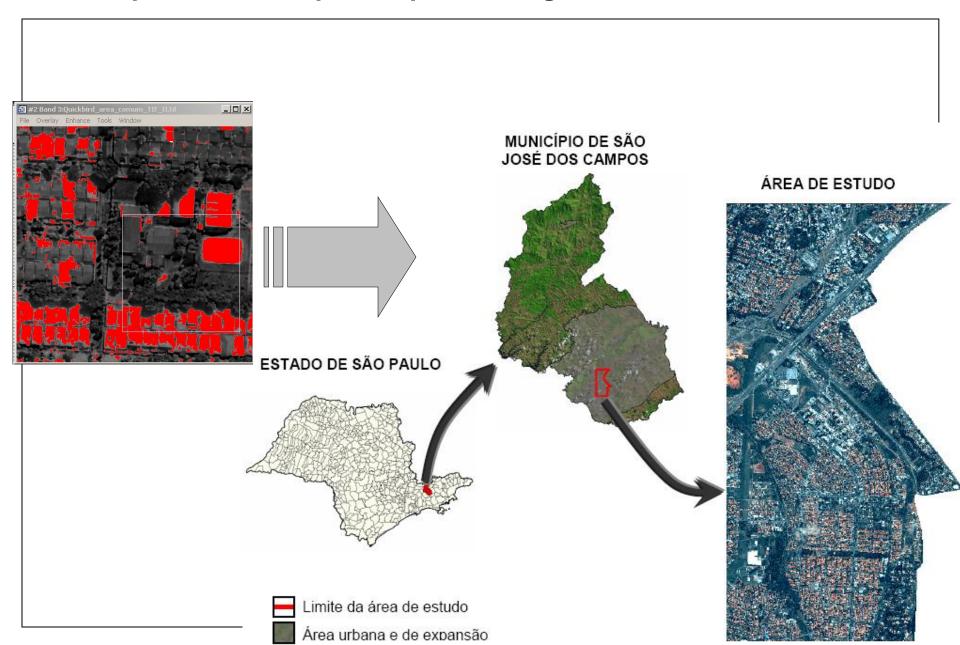
Razão de bandas

Banda 3

Composição colorida cor verdadeira

Detecção de classes de cobertura em imagens de alta resolução espacial pode ser obtida a partir da análise de medidas em laboratório.

Do comportamento espectral para a imagem classificada?

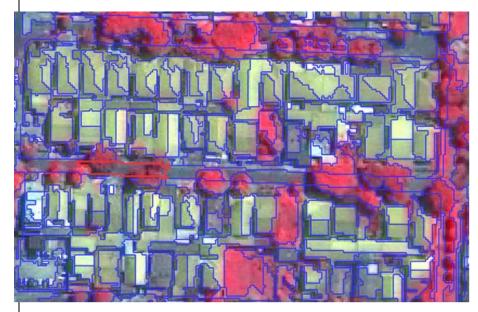


Classificação por objetos

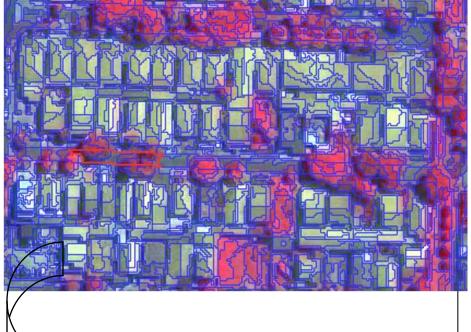
Segmentação

Identificar formas ?...

... ou informação espectral ??

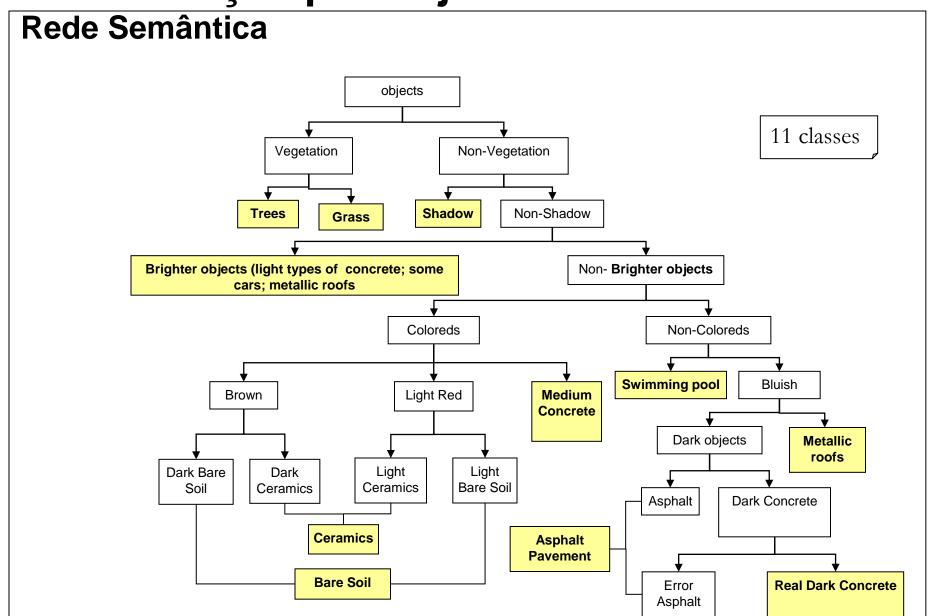


Quickbird (2004), PAN+XS, IHS, ENVI, E-Cognition



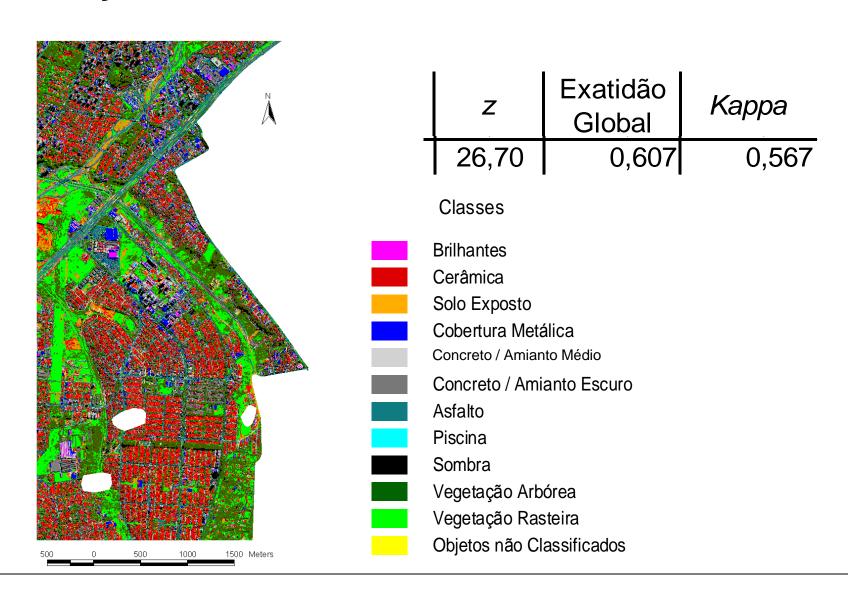
~ 400.000 objetos para classificar

Classificação por objetos

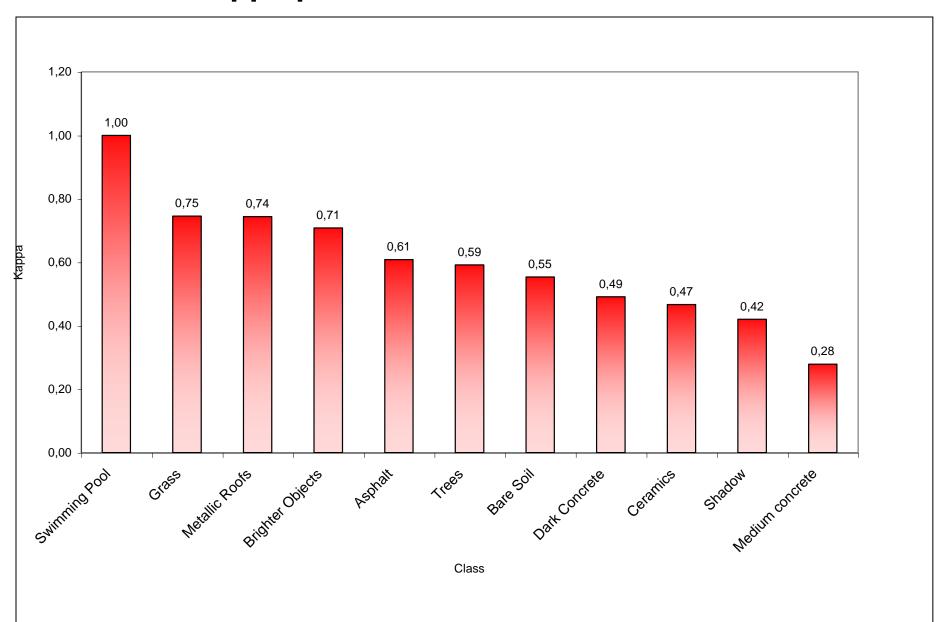


Classificação por objetos

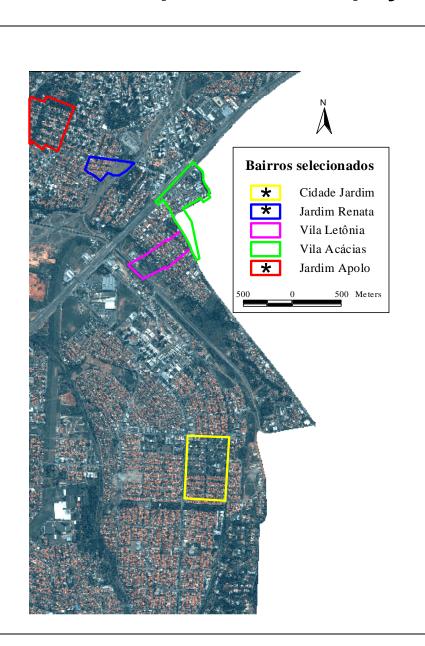
Classificação



Exatidão - Kappa por classe



Influência do padrão de ocupação



Pouco Organizados

Heteronegeindade em tamanho e tipo de material, objetos pequenos, cobertura do solo mais variada

Organizados (*)

Dispostos regularmente, homogeneidade em tamanho e tipo de material **Exemplos**

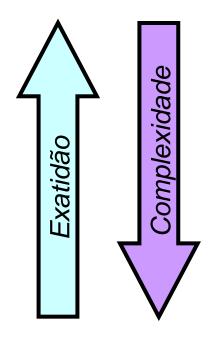




(*) Cidade Jardim

Conhecimento Local

Exatidão de mapeamento



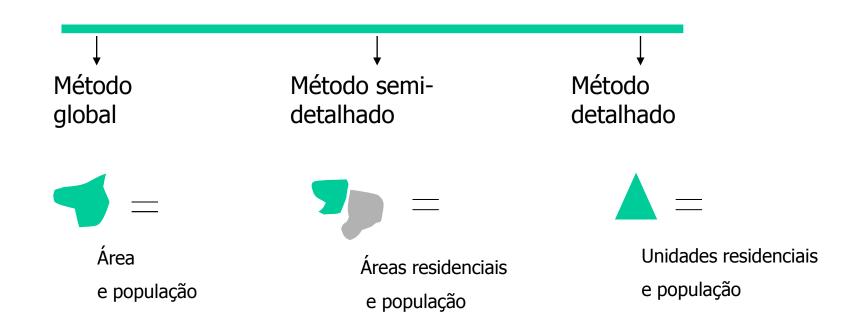
	Quickbird			
Bairros	Kappa	Exatidão Global		
Jardim Renata	0,76	0,79		
Cidade Jardim	0,74	0,76		
Jardim Apolo	0,63	0,68		
Vila Acácias	0,58	0,62		
Vila Letônia	0,50	0,55		



Da classificação à estimativa...

Por meio de dados de Sensoriamento Remoto

Métodos de inferência de população



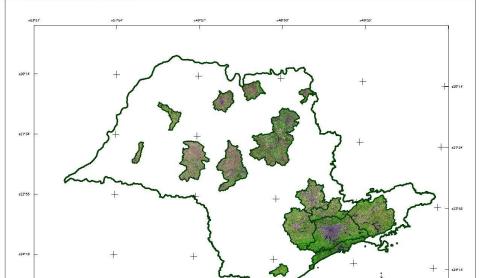
Liu e Clark (2002)

MÉTODO GLOBAL

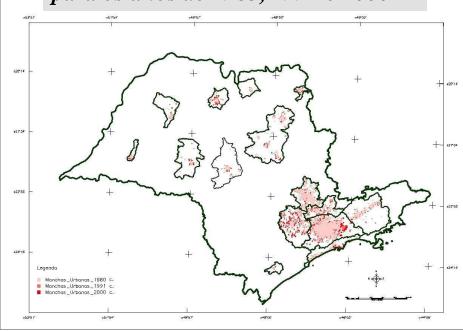
Urbanização Dispersa e Novas Formas de Tecido Urbano

Estudo de caso: Estado de São Paulo

Landsat TM − 30 m



Mapeamento das Áreas Urbanizadas para os anos de 1985, 1991 e 2000



- ✓ Levantamento das áreas das Manchas urbanas
 - ✓ Determinação da correlação entre População e área

Correlação dos municípios por Região de Estudo Correlação de todos os Municípios

MÉTODO GLOBAL

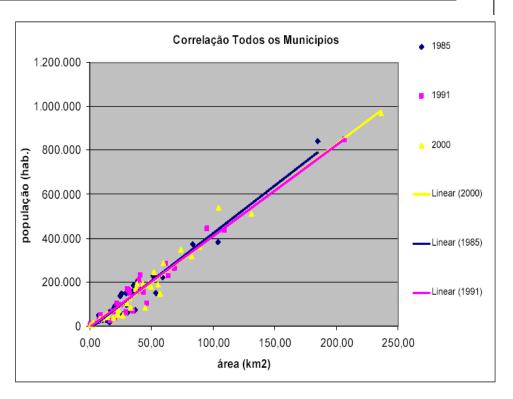
Correlação a partir do Mapeamento das

Áreas Urbanizadas

	Índices de Correlação População x Área							
	1985		1991		2000			
	r ²	equação	r ²	equação	r ²	equação		
Bauru	0,9997	y=3675,6x	0,9998	y=3759,4x	1	y=3826,8x		
Campinas	0,9787	y=4673,3x	0,9819	y=4200,4x	0,9747	y=4244,5x		
São José dos Campos	0,9901	y=4580,9x	0,9923	y=4717,4x	0,9902	y=5019,8x		
Todos os Municípios	0,9564	y=4339,2x	0,961	y=4141,3x	0,965	y=4203,9x		

Anos de 1985, 1991 e 2000

Com os dados Censitários IBGE



MÉTODO SEMIDETALHADO

Comparação das classificações a partir de dados de SR com diferentes resoluções espaciais

Gonçalves et al (2003)

Município de São José dos Campos



IKONOS 2 - referência terrestre (4m)

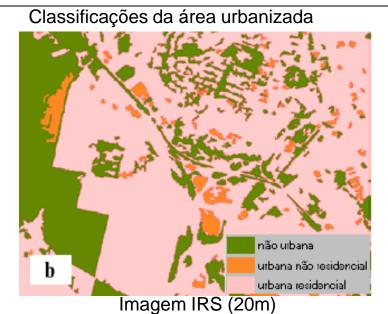
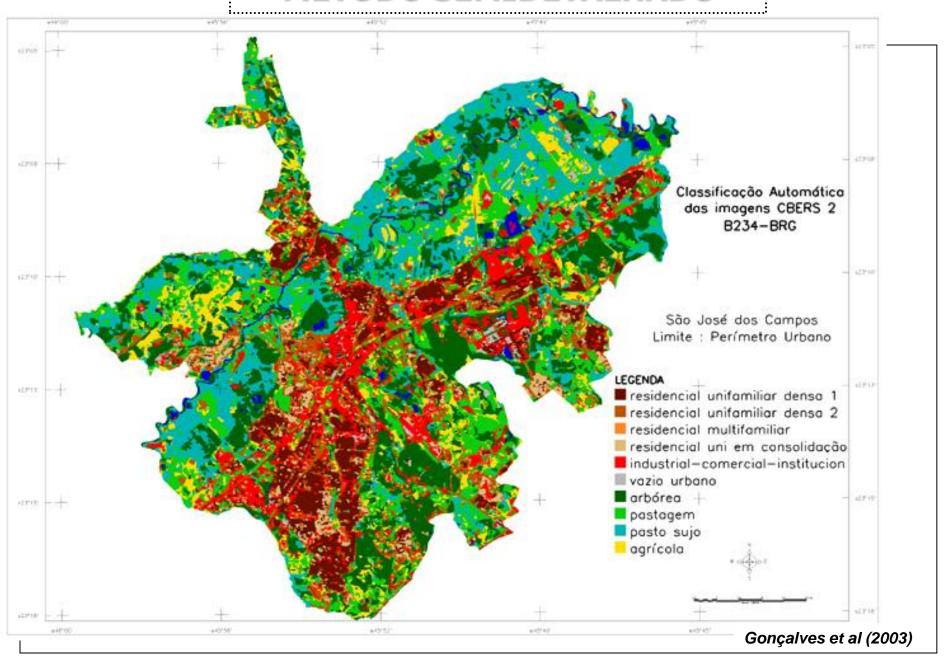




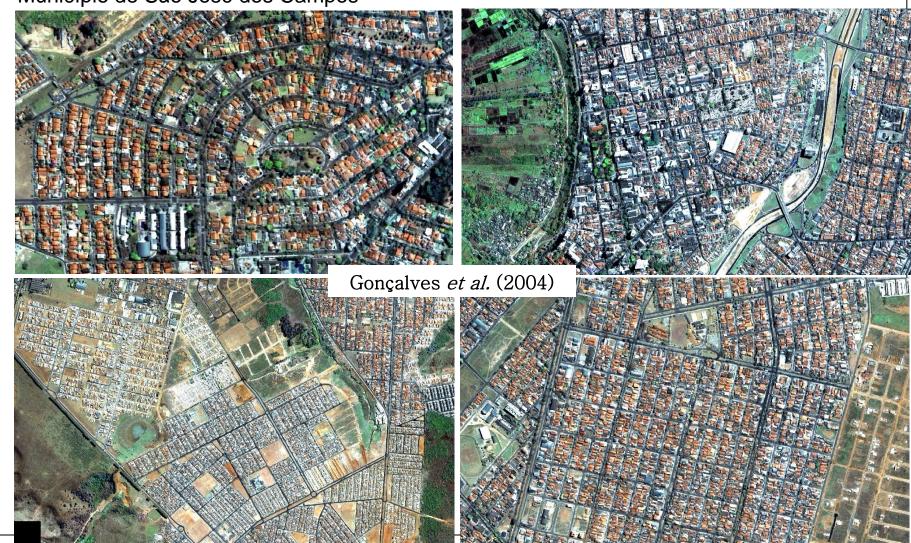
Imagem LANDSAT 2000 (30m)

MÉTODO SEMIDETALHADO



Texturas encontradas nas imagens de Alta resolução Espacial - Delimitação de Zonas Residenciais Homogêneas quanto ao espaço residencial construído

Município de São José dos Campos



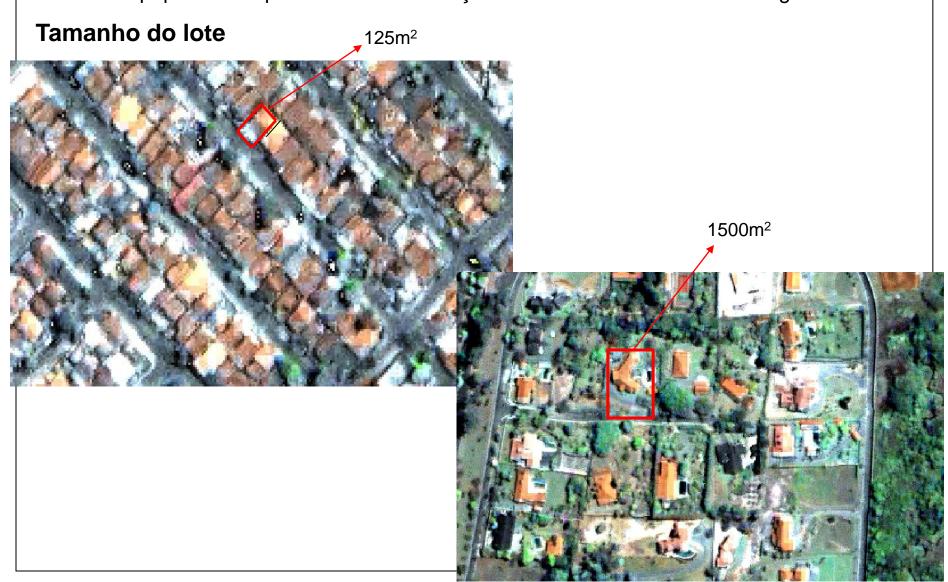
Contextualização

- Objetivo: avaliar a inferência de população a partir de dados de sensoriamento remoto orbital de alta resolução
- São José dos Campos 539.313 habitantes, 532.717 urbana, 6.596 rural (Censo 2000)
 - Crescimento populacional em função do crescimento econômico industrial
- Base de Dados:
 - Imagem IKONOS-2 (2000), fusão multi-espectral (4 m) e pancromática (1 m), fotos aéreas para correção geométrica
 - Base de informação dos setores censitários + atributos para 77% da área urbana (cobertura imagens)
 - Estratificação em zonas residenciais unifamiliares homogêneas (ZRUFs)

Contextualização

- Estratificação em zonas residenciais unifamiliares homogêneas (ZRUFs)
 - relação entre uso residencial unifamiliar e padrão construtivo –
 condições sócio-econômicas e padrões de textura nas imagens
 - Interpretação visual, escala 1:10.000, chaves de interpretação
 - ZRUFs associadas às classes sócio-econômicas
 - para área urbana consolidada: alta, média, baixa e favela,
 - para área urbana não consolidada: expansão

Estimativa populacional por meio de Delimitação de Zonas Residenciais Homogêneas



Estimativa populacional por meio de Delimitação de Zonas Residenciais Homogêneas **Organização das quadras e lotes**







Estimativa populacional por meio de Delimitação de Zonas Residenciais Homogêneas

Arborização das ruas e lotes



Estimativa populacional por meio de Delimitação de Zonas Residenciais Homogêneas

Densidade residencial



Estimativa populacional por meio de Delimitação de Zonas Residenciais Homogêneas

Homogeneidade dos materiais dos telhados







Estimativa populacional por meio de Delimitação de Zonas Residenciais Homogêneas

Classificação e subdivisões:



Estimativa populacional por meio de Delimitação de Zonas Residenciais Homogêneas

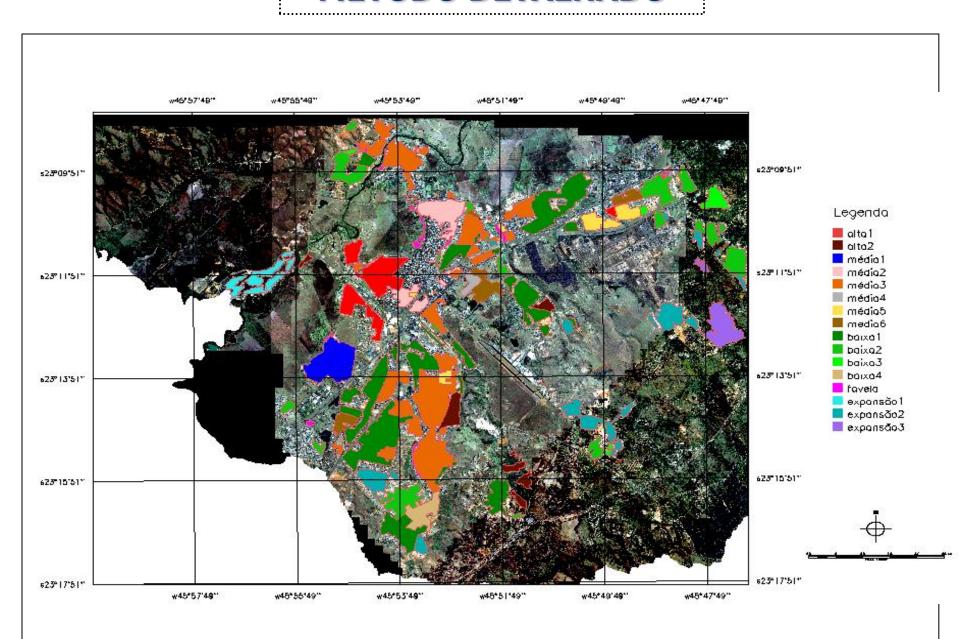
Classificação e subdivisões - exemplos:

Alta 1 - Próxima do centro / lotes e quadras bem definidos / arborização urbana de grande porte / presença de áreas verdes intralotes / lotes inteiros (12x30m ou maiores) / um domicílio por lote com construções secundárias, incluindo piscinas / recuos laterais e de fundos identificáveis na imagem / predomínio de cobertura cerâmica.

Média 1 - Localização específica em uma região da cidade / lotes e quadras padronizados / arborização urbana de pequeno porte / predominância de lotes inteiros (10x25m) / taxa de ocupação alta, sendo os lotes quase totalmente construídos, mas com predomínio de uma única unidade por lote / predominância de telhados cerâmicos (ocupação nova).

BAIXA 1-Lotes e quadras padronizados/em áreas de ocupação antiga há arborização urbana; nas mais recentes, há áreas criadas para tal;vegetação intralote ausente/porcentagens equivalentes de lotes inteiros (10x25m) e meios lotes (5x25m)/alta taxa de ocupação e ausência de recuos obrigatórios devido à presença de coabitação/coberturas predominantemente cerâmicas.

Expansão 1 -Área urbana não-consolidada, com características de ocupação tipo classe média alta. Lotes (¡Ý 12x30m) e quadras padronizados/vegetação urbana de pequeno porte ou inexistente (loteamentos novos)/um domicílio por lote com construções secundárias, incluindo piscinas/coberturas bastante variadas (grande diferença na resposta espectral)/classe presente em única localização da cidade (distante do centro e com relevo pouco acidentado).



Procedimentos

Base de dados amostrais:

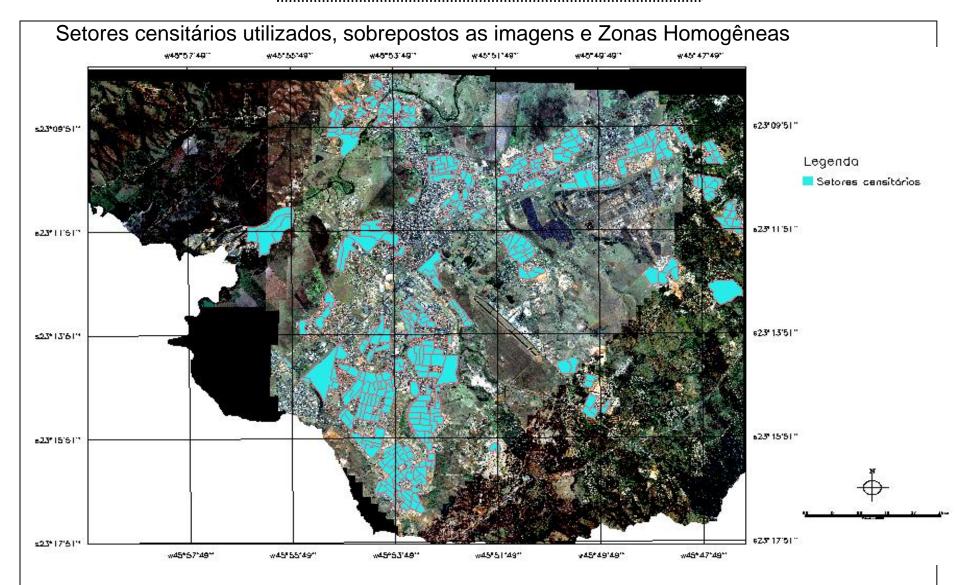
- Setores censitários pertencentes a uma única ZRUFs (sobreposição de mapas)
- 328 setores homogêneos (classe socioeconômica) de referência

Seleção das amostras:

- para inferência da população foram escolhidos 24 setores representativos do universo amostral,
- centro de massa das ZRUFs e atributos dos setores núm. méd. moradores/dom; renda média responsável; méd. anos estudo responsável e núm. méd.banheiros/dom.

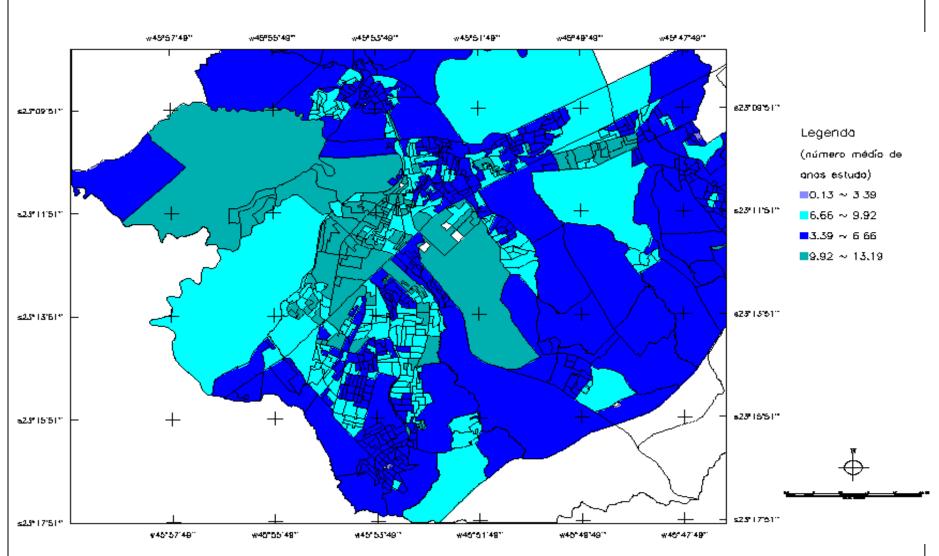
Chaves de interpretação unidades domiciliares

 definidas por critérios de localização, organização, áreas verdes, características de construção, etc.

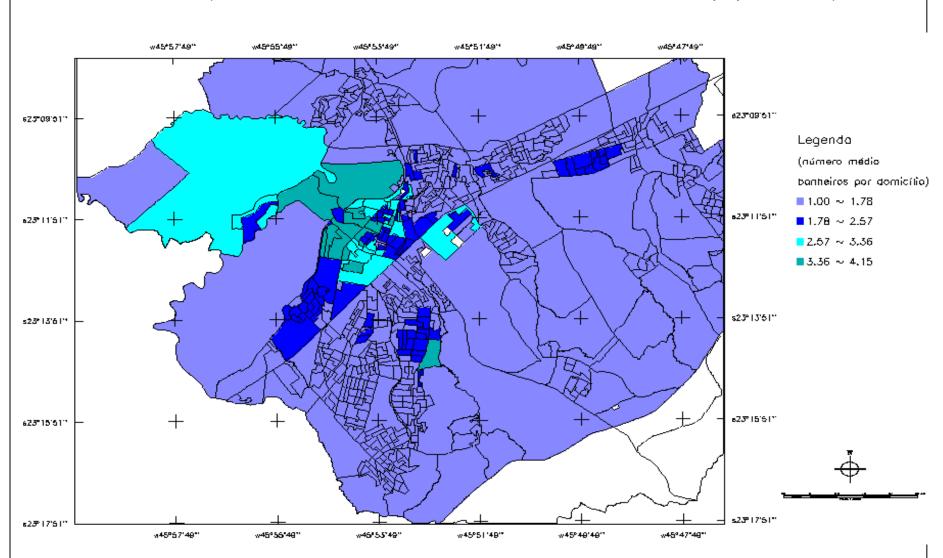


Base de dados amostrais: Setores censitários pertencentes a uma única ZRUFs (sobreposição) – 328 setores homogêneos (classe socioeconômica) de referência

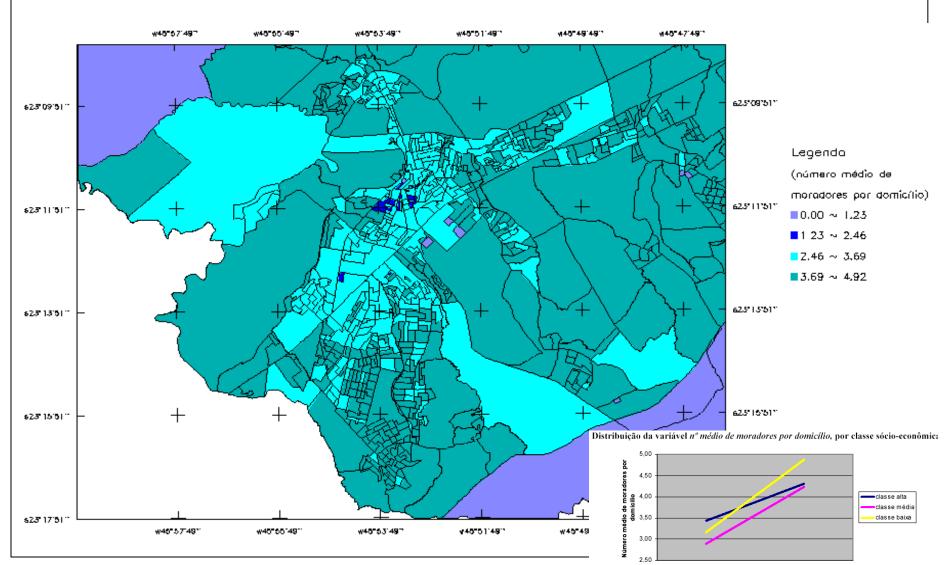
Distribuição espacial dos dados socioeconômicos a fim de se realizar a seleção das amostras (Setores censitários a serem utilizados na inferência populacional)

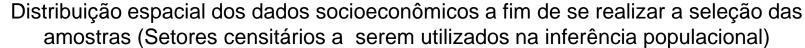


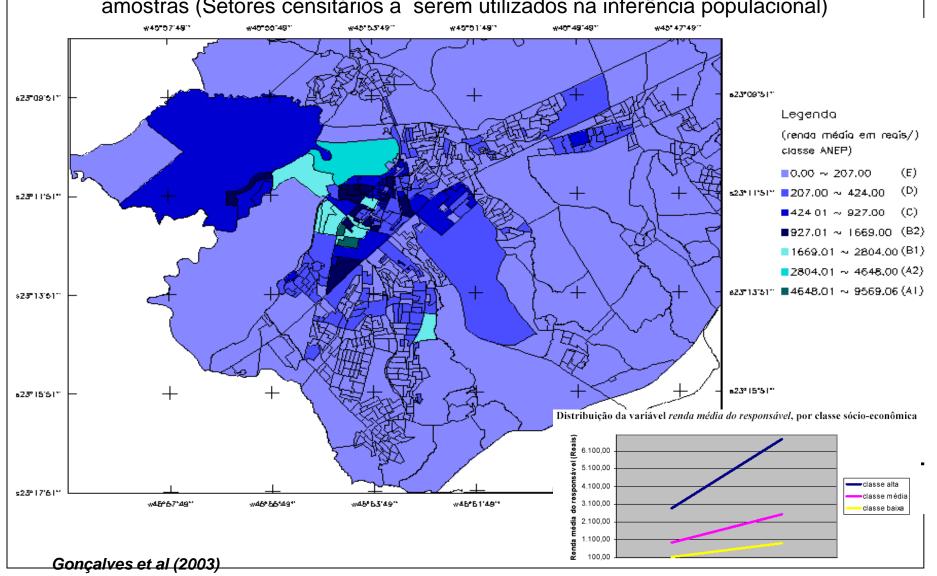
Distribuição espacial dos dados socioeconômicos a fim de se realizar a seleção das amostras (Setores censitários a serem utilizados na inferência populacional)



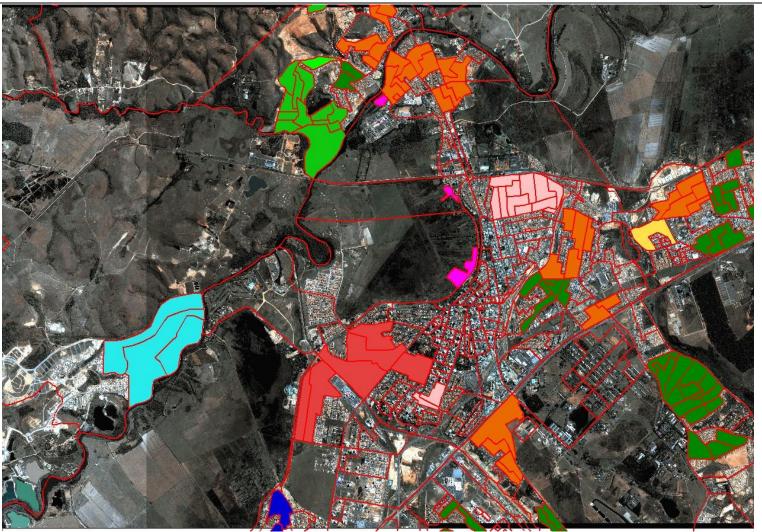
Distribuição espacial dos dados socioeconômicos a fim de se realizar a seleção das amostras (Setores censitários a serem utilizados na inferência populacional)







Setores Censitários Amostrais selecionados



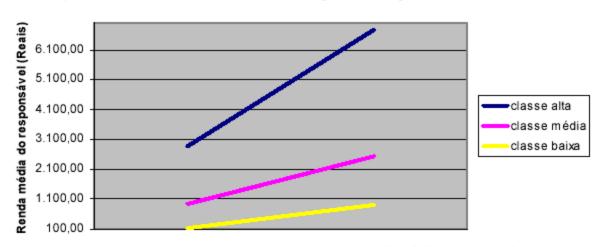
Seleção das amostras: para inferência da população foram escolhidos 24 setores representativos do universo amostral, centro de massa das ZRUFs e atributos dos setores – núm. méd. moradores/dom; renda média responsável; méd. anos estudo responsável e núm. méd.banheiros/dom.

Inferência e Análise

- Identificação das unidades domiciliares:
 - para todos os 24 setores selecionados contagem de domicílios ocupados
- Inferência de população:
 - População setor = n. de domicílios * n. médio de residentes por domicílio
- Precisão da estimativa de população: correlação entre população IBGE 2000 e a estimativa IKONOS. Erros de omissão e inclusão:
 - Erro de Estimativa = (População IBGE População IKONOS 2)
 População IBGE

Resultados

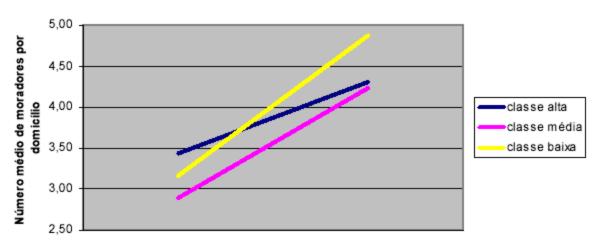
Distribuição da variável renda média do responsável, por classe sócio-econômica



 Correlação entre média de salários e padrão de ocupação (SR)

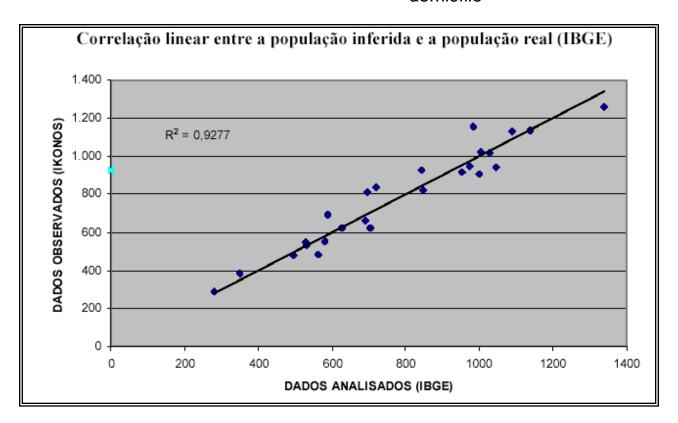
Distribuição da variável nº médio de moradores por domicílio, por classe sócio-econômica

 Maior superposição de classes - número médio de moradores por domicílio



Inferência de população:

População setor = n. de domicílios * n. médio de residentes por domicílio

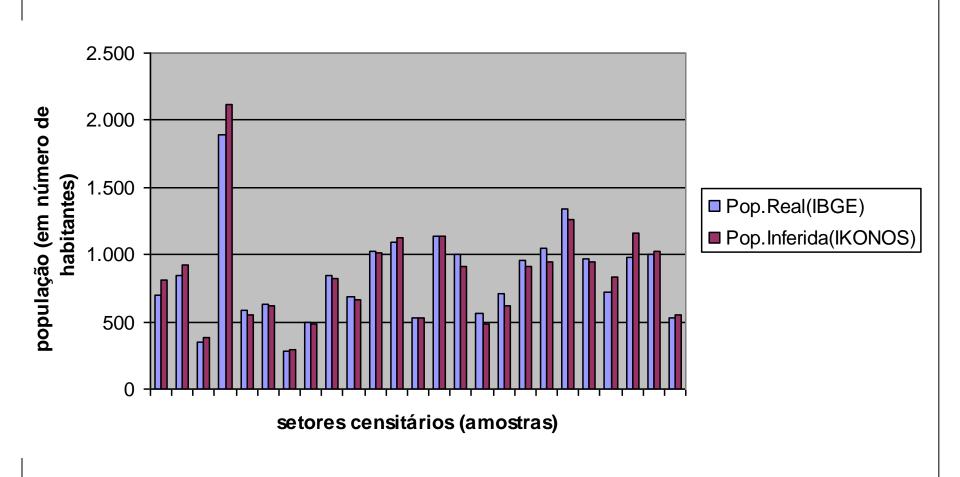


Erro médio de inferência da população: intervalo de 0,41 ± 4,51 %, com 99% de confiança

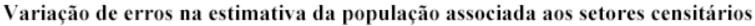
24 setores censitários, erro médio de inferência populacional de +0,41%, e desvio padrão de 8,58

Dados populacionais inferidos semelhantes aos dados populacionais do Censo Demográfico ("t" de Student, 99% d.f.)

Relação entre a população inferida e a referência (IBGE)



Resultados



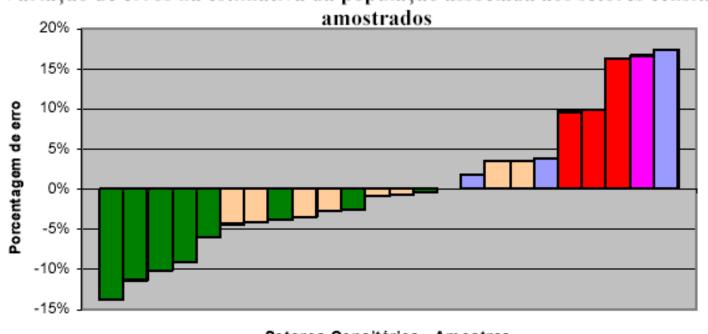
Classe Alta

Classe Média

Classe Baixa

Expansão

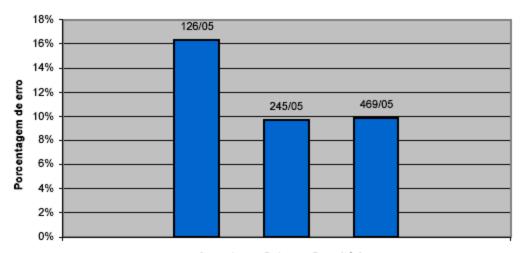
Favela



Setores Censitários - Amostras

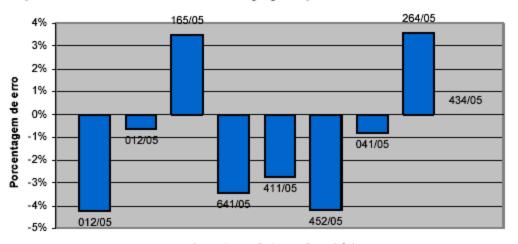
Resultados - Erros/amostra

Variação de erros na estimativa da população associada à Classe Alta



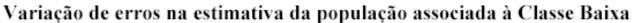
Amostras - Setores Censitários

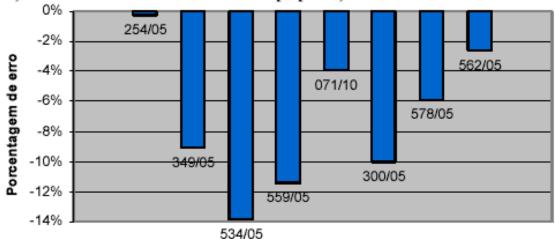
Variação de erros na estimativa da população associada à Classe Média



Amostras - Setores Censitários

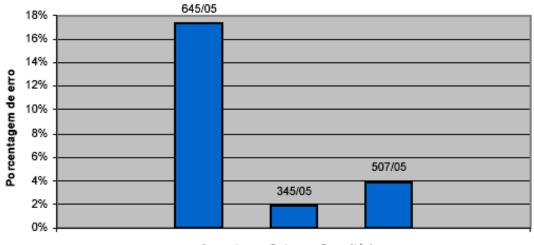
Resultados - Erros/amostra





Amostras - Setores Censitários

Variação de erros de estimativa da população associada a áreas não consolidadas



Amostras - Setores Censitários

Considerações

- Erro médio de inferência da população intervalo de 0,41 ± 4,51
 %, com 99% de confiança
 - 24 setores censitários, erro médio de inferência populacional de +0,41%, e desvio padrão de 8,58
- Dados populacionais inferidos <u>semelhantes</u> aos dados populacionais do Censo Demográfico (teste "t" de Student, 99% d.f.)
- <u>Viabilidade</u> do uso de imagens orbitais de alta resolução espacial para a estimativa de dados populacionais a partir do método detalhado, em períodos intercensitários.
- Dados <u>úteis para auxiliar</u> a delimitação e definição dos setores censitários para a realização dos censos demográficos decenais, o que aumentaria a precisão dos dados levantados pelos recenseadores.



Delimitação de Zonas Residenciais Homogêneas

Automatizado - regras de classificação

Classes de uso nas quadras





Souza et. al, 2009

Inferência Populacional - Estimativas

Regressão Iterada Regressão Estatística Espacial

Estimando Populações

Estimação da população dos setores censitários de Belo Horizonte usando imagens de satélite

- Censo Demográfico (IBGE)
 - a cada 10 anos.
- Contagem Populacional :
 - entre dois censos.
- Projeções demográficas.



Estimando Populações

Modelo de regressão linear

Reflectâncias nas bandas do sensor

$$P=eta_0+eta_1\dot{R_1}+eta_2R_2+...+eta_7R_7+arepsilon$$
 População

Setores Censitários (Harvey, 2002a)

Duas abordagens

Pixels (Harvey, 2002b)

Modelo de Regressão por Pixels

• Problema: não existem dados de população associados aos pixels e sim aos setores censitários.

- Solução (Harvey, 2002b): regressão iterada
- Idéia da regressão iterada : redistribuir a população do setor em seus pixels através de sucessivas regressões.

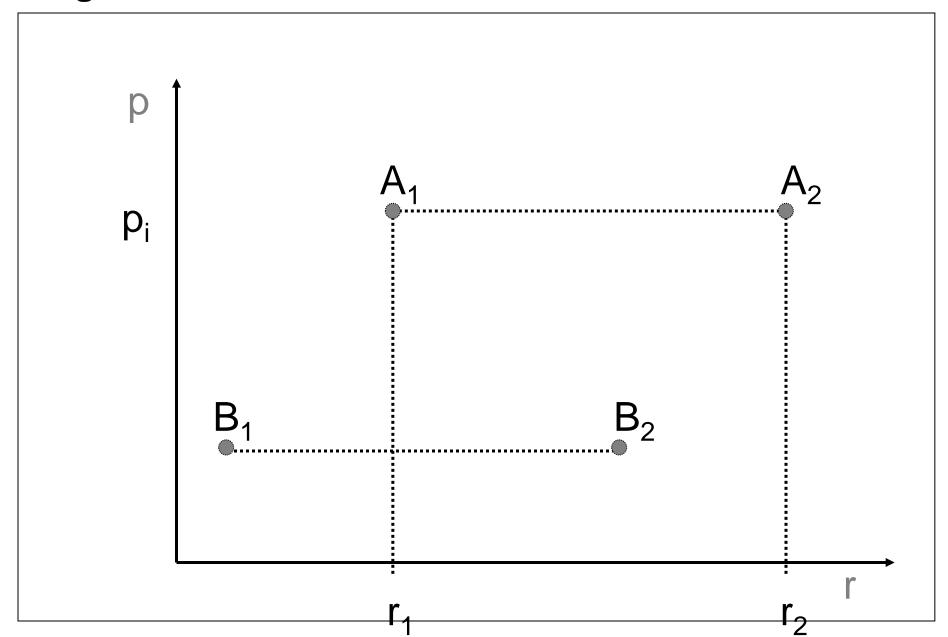
População redistribuída igualmente no setor censitário:

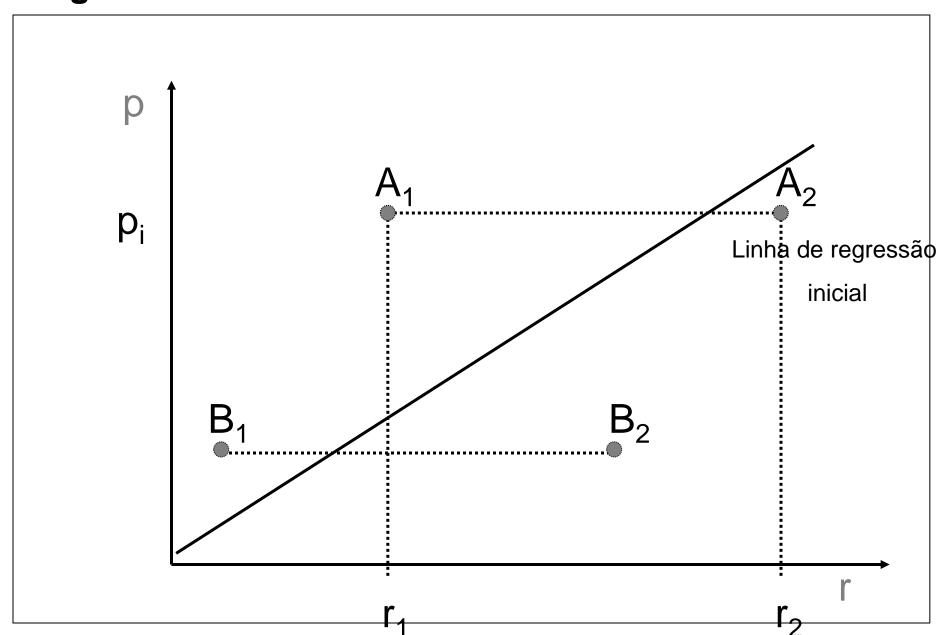
- População pixel i = Pop setor / n. pixels
- Modelo de regressão ajustado e as estimativas de população são calculadas (Pi – estimado);
- Para cada setor, as estimativas são corrigidas de modo que a população total do setor se mantenha igual ao valor conhecido $p_{i(\text{ajust})} = p_{i(\text{estim})} + \bar{r}$

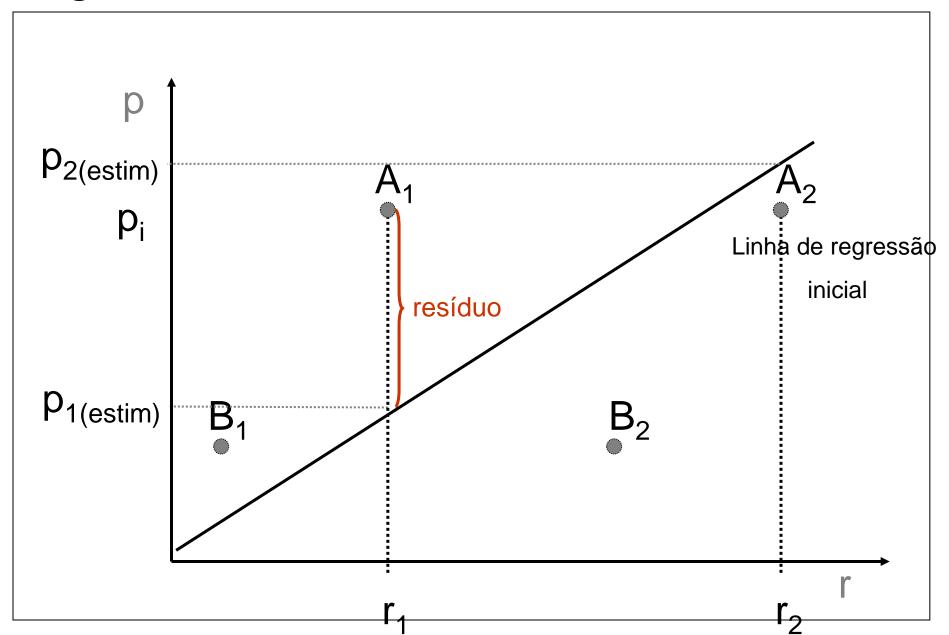
$$\bar{r} = \frac{\sum_{i=1}^{n} (p_i - p_{i(\text{estim})})}{n}$$

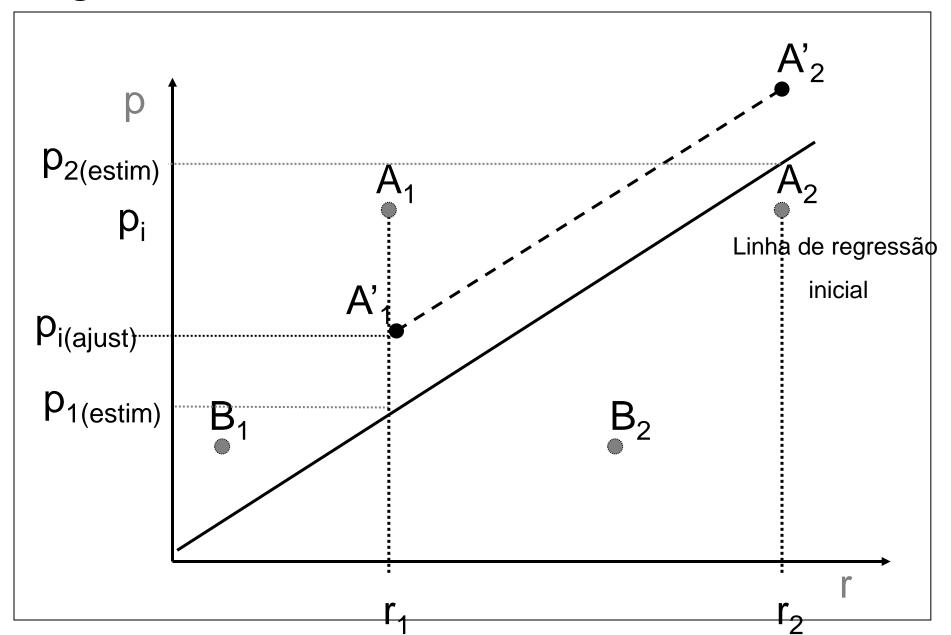
r é a média dos resíduos em cada setor.

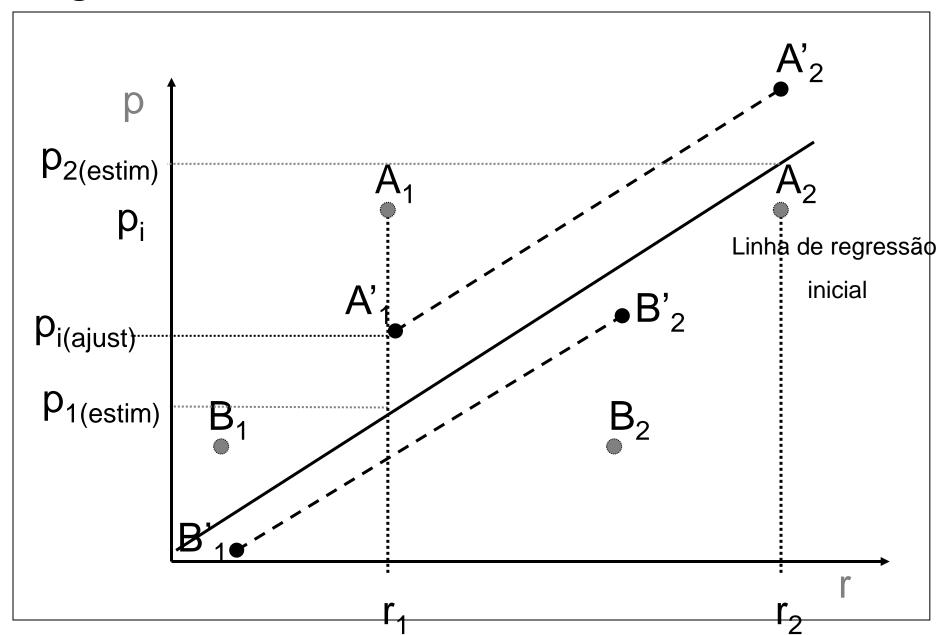
- Ou seja, a população ajustada para o píxel i é:
 - População ajustada pixel i = Pop est pixel i + média dos resíduos do setor.
- Na próxima iteração, os valores ajustados substituem as estimativas iniciais de população, e a regressão roda novamente
- O procedimento pára quando o incremento no valor de coeficiente de determinação (R2) passa a ser muito pequeno;
- A última equação é utilizada para fazer as estimativas de população.

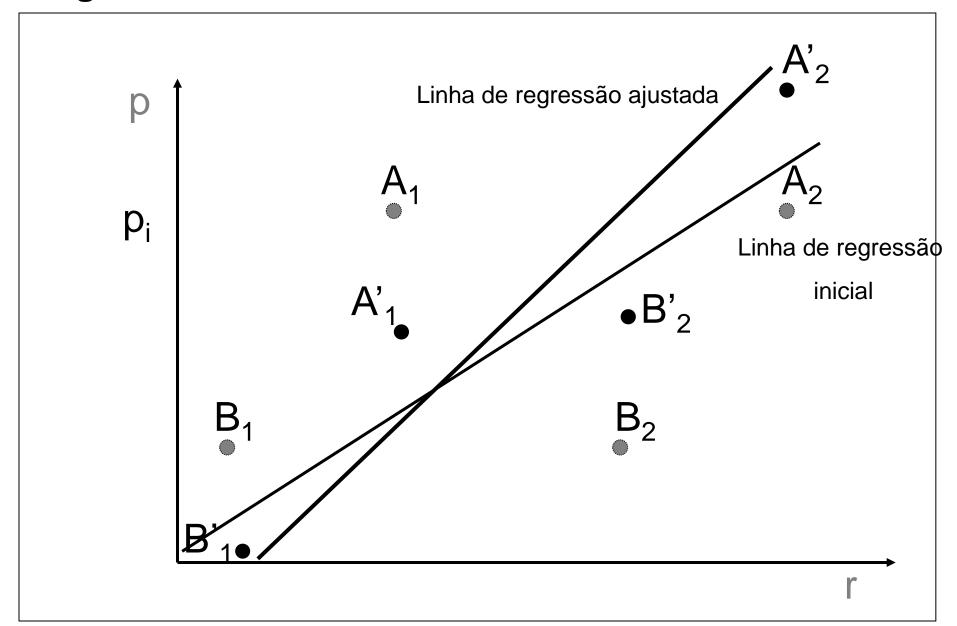












TM3 (R) TM4 (G) TM5 (B)

Aplicação

Regressão por pixels para estimação da população dos setores censitários de Belo Horizonte (MG) com imagens TM/LANDSAT

- Landsat5-TM 218/74, 31/01/1996
- Bandas 1 a 5 e 7
- Setores censitários
 Contagem IBGE 1996

Procedimento

- 1. Transformação das imagens ND para reflectância de superfície
- 2. Classificação dos pixels em urbanos e não-urbanos;
- 3. Associação dos pixels urbanos aos setores censitários;
- 4. Regressão iterada (amostra de pixels = 25% dos pixels dos setores);
- 5. Aplicação do modelo estimado a todos os pixels;
- 6. Combinação das estimativas de população para os setores

Avaliação dos Modelos

• Nível micro: erro na estimação da população do setor

$$Erro relativo = \frac{observado - estimado}{observado}$$

• Nível macro: erro na estimação da população total

$$Erro\ relativo\ total = \frac{total\ observado\ -\ total\ estimado}{total\ observado}$$

Resultados (modelo completo)

Afonso (2005)

Harvey (2002b)*

0.257	R ² Iterada (amostra)	0.855
0.193	R ² "back" (total)	0.822
30.43	Erro Relativo Mediano (%) (pop. Setor)	14.0
0.47	Erro Relativo Total (%)	-4.8
Negativo	Sinal do coeficiente da banda 4	Negativo
Negativo	Sinal do coeficiente da banda 4	Negativo
Positivo	Sinal do coeficiente da banda 4	Positivo

^{*} Área urbana

Resultados

Modelo escolhido :

$$\hat{p} = 13.243 + 0.011 \, \mathbf{r}_1 - 0.020 \, \mathbf{r}_4 - 0.102 \, \mathbf{r}_5 + 0.132 \, \mathbf{r}_7$$

- Erro relativo total: -0.06% (Harvey, -4.8%)
- Erro relativo mediano : 31.57% (Harvey, 14.0%)
- Tendências : (também observadas por Harvey)
 - Superestimação em setores pouco densos
 - Subestimação em setores muito densos

Considerações

- Os modelos de estimação de população usando imagens de satélite via modelos de regressão conseguem a bons resultados no nível macro, mas os resultados no nível micro são pobres.
- A densidade populacional do setor parece ter grande influência na super ou subestimação de sua população.
- O modelo foi validado apenas internamente.

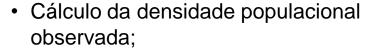
Considerações

Alternativas para melhoria dos resultados no nível micro:

- Incorporar indicadores indiretos de densidade populacional, como zonas de ocupação e uso do solo;
- Agrupar as estimativas de população em um nível superior ao dos setores (agrupamentos deles). Ex: UP (Unidade de Planejamento) no caso de Belo Horizonte.

Estimativa baseada em regressão espacial

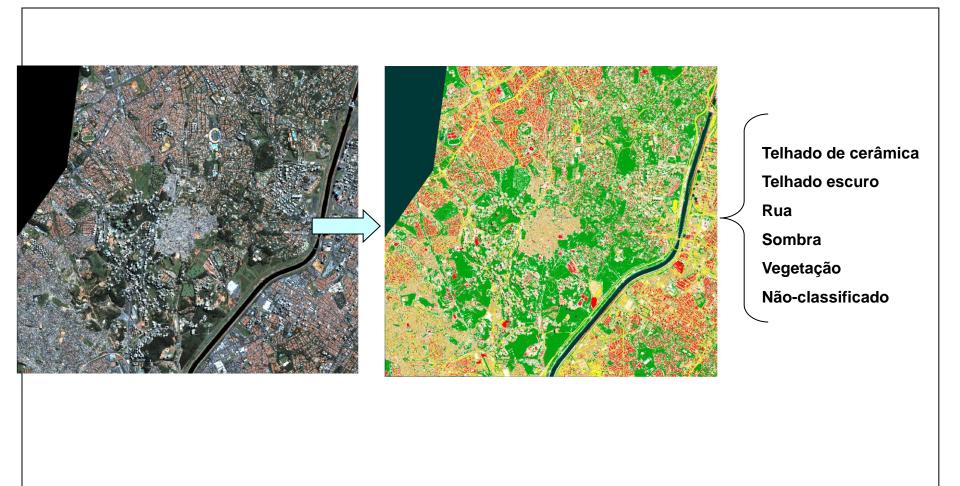
Estimar a densidade populacional agregada por setor censitário utilizando técnicas de regressão espacial multivariada e métricas espaciais obtidas sobre imagem de alta resolução espacial (QuickBird II).



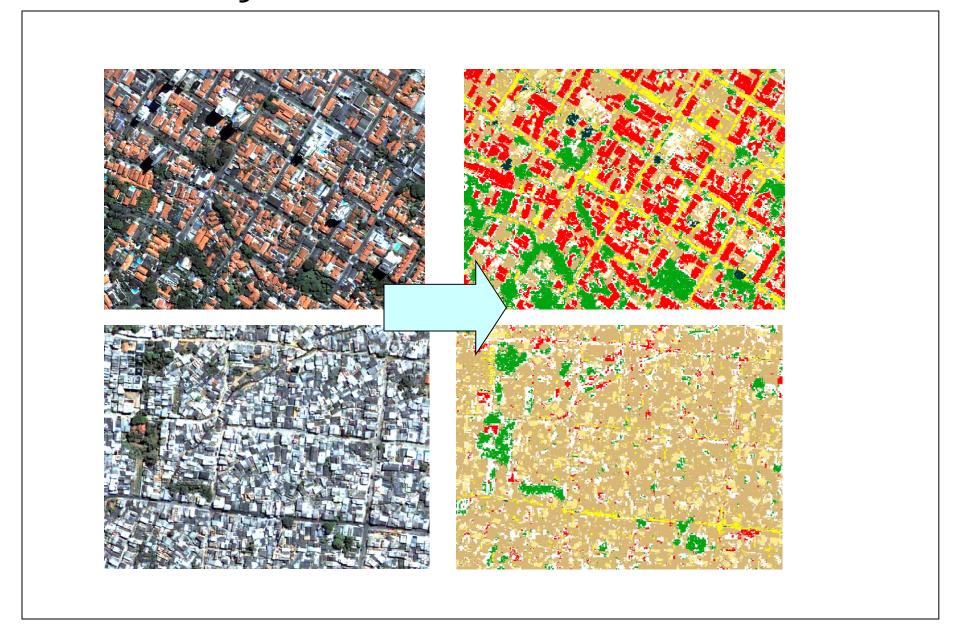
- Classificação supervisionada da imagem;
- Obtenção das variáveis;
- Construção de modelos de regressão linear;
- Análise da dependência espacial dos resíduos;
- Construção de modelos de regressão espacial;
- Análise dos resultados e escolha do modelo final.

Pinho (2009) / Novack et al. (2008)

Classificação

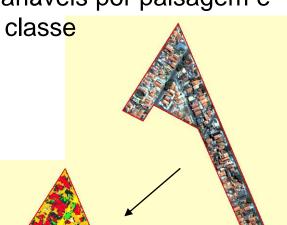


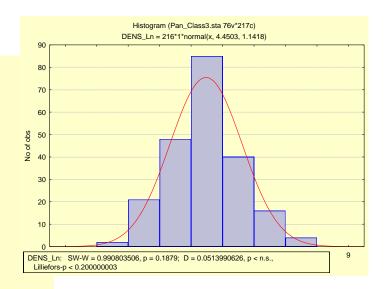
Classificação



Estatística/setor censitário

- Variáveis por paisagem e





	DENS_Ln
Variable	
PD_1	-0.38
NP_2	0.75
COHESION_2	0.57
Al_3	0.56
PLAND 5	-0.57

Modelo com 4 variáveis:

- 1. no de polígonos de telhado de cerâmica:
- 2. porcentagem de telhado escuro;
- 3. agregation index de rua; e
- 4. patch density de vegetação.

R2: 0.7516

Modelo com 2 variáveis:

- 1. nº de polígonos de telhado de cerâmica e
- 2. porcentagem de telhado escuro

R2: 0.6524

Modelos

Modelo com 4 variáveis:

Matriz de Distância - OLS		Matriz de Distância - Spatial Lag			
R-squared	0.7434	R-squared	0.7551		
Akaike	366	Akaike	359		
Moran	0.042	Moran	-0.017		
Breusch-Pagan	11.75	Breusch-Pagan	8.39		

Modelo com 2 variáveis:

Matriz de Distâr	ncia - OLS	Matriz de Distância - OLS - Regimes Locais			
R-squared	0.6583	R-squared	0.7763		
Akaike	423	Akaike	339		
Moran	0.0825	Moran	0.01		
Breusch-Pagan	14.13	Breusch-Pagan	17.64		

	4 •		4 •	
\Box	\ †!/	na	+11	10
<u> </u>		110	111	1 1
	<i>,</i>	114		<i>,</i>

Variável	Descrição
N_VIZINHOS	Número de setores vizinhos ao setor em questão
S_5	Área em m² da classe Sombra de Vegetação
S_1/10000	Área em m² dividida por 10.000 da classe Área Construída
NUM_2	Número de Polígonos da Classe Asfalto dentro das quadras
S_3	Área em m² da classe Solo Exposto
MQ_4	Área Média dos polígono da Classe Sombra de Prédio dentro das Quadras
NDVIM_1	NDVI médio dos Polígonos da classe Área Construída
REL_6	Porcentagem da área da classe Vegetação sobre a Área do setor

Depois de concluída a validação, construiu-se um modelo final com todos os parâmetros re-estimados baseados no conjunto total de dados.

	Regression Summary for Dependent Var R= .87607515 R ² = .76750766 Adjusted I F(3,77)=84.731 p<0.0000 Std.Error of es									
	Beta	Beta Std.Err. B Std.Err.								
N=81		of Beta of B								
Intercept			10.1687	0.08200						
REL_6	-0.37541	0.06105	-1.81609	0.29535						
S_1/10.000	-0.52337	0.08232	-0.12660	0.01991						
NUM_2	-0.19463	0.07627	-0.0019	0.00076						

Resultado do modelo da regressão final.

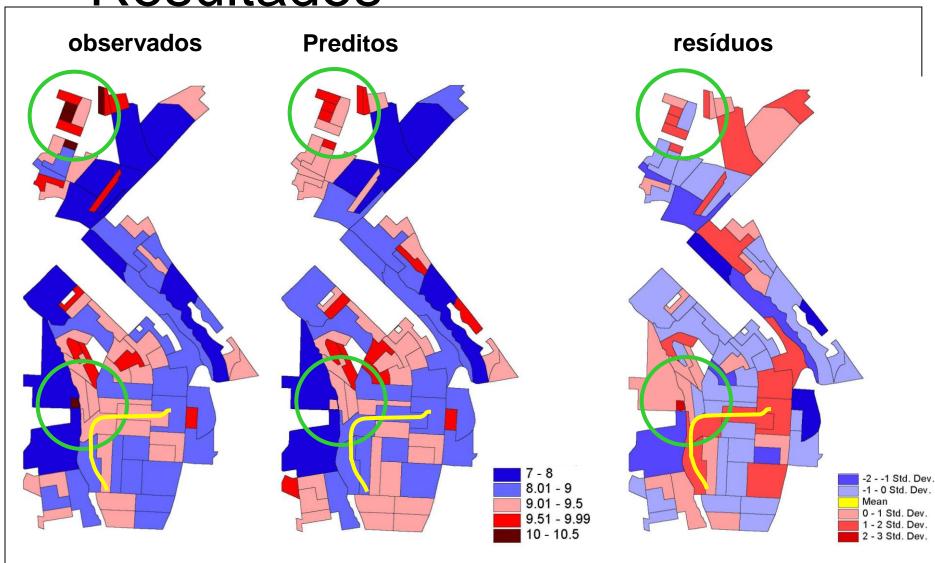
$$LnY = 10.16872 + (-0.37X_1) + (-0.53X_2) + (-0.19X_3)$$

 $X_1 = REL_6$

$$X_2 = S_1/10.000$$

$$X_3 = NUM_2$$

Resultados



Índice de Moran Global = 0.0550556 pcrítico = 0.05 e pvalue = 0.31 (não correlacionado)

Novack et al. (2008)

População – Imagens luzes noturnas









Sensoriamento remoto de luzes noturnas para estimativas populacionais em escalas regional e local: os casos do Distrito Florestal Sustentável da BR163 (PA) e da Região Metropolitana de São Paulo

Vinicius Etchebeur Medeiros Dória

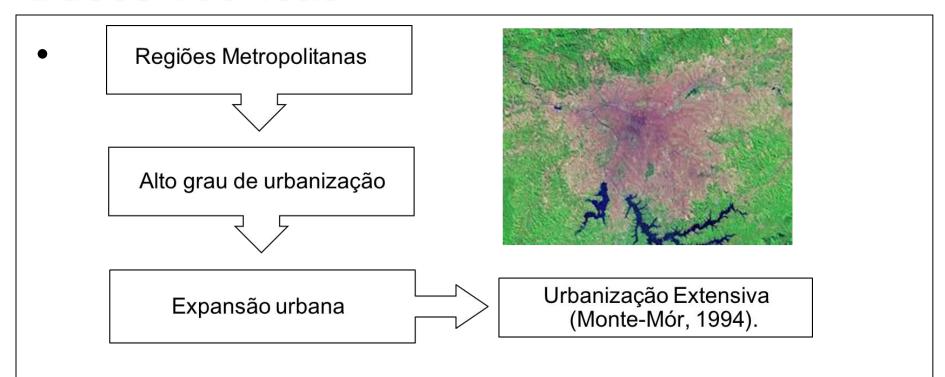
viniciusdoria@dsr.inpe.br

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

Motivação

 É possível utilizar as imagens de luzes noturnas como informação auxiliar nos estudos populacionais?

Bases Teóricas



- Monte-Mór baseado em Lefebvre (1999) → Urbanização Extensiva.
- Unidades Espaciais de Ocupação Humana (Dal'Asta et al., 2012) → áreas urbanizadas de sedes municipais, vilas e comunidades, madeireiras, serrarias, etc.









Sensores de Luzes Noturnas

- Defense Meteorological Satellites Program (DMSP)
- Operational Linescan System (OLS)
- OLS desde 1976
 - Whisk broom
 - Diurno e Noturno
 - Tubo Fotomultiplicador
 - Resolução espacial de 2,7 km
 - Resolução radiométrica de 6 bit
 - Produto anual de luzes estáveis de resolução espacial de 1 km



Sensores de Luzes Noturnas

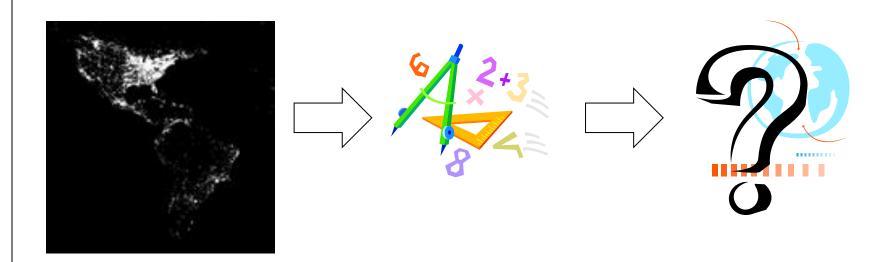
Visible Infrared Imaging Radiometer Suite (VIIRS)

- Joint Polar Satellite System (JPSS)
- Suomi National Polar-orbiting Partnership (SNPP)
- VIIRS desde out/2011
 - Whisk broom
 - 22 bandas
 - Day/Night Band (DNB)
 - 4 sensores Charged Coupled Device (CCD)
 - Resolução espacial de 742 m
 - Resolução radiométrica de 14 bit
 - Mosaico global livre de nuvens de 2012 com resolução espacial de 500 m
 - Calibração a bordo da plataforma.

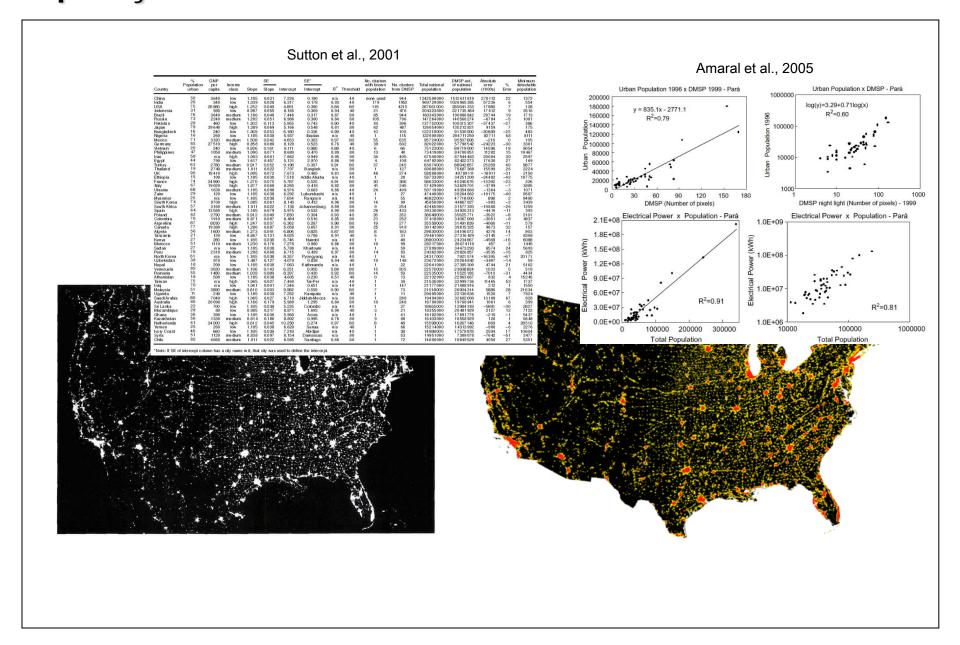


Sensoriamento Remoto de Luzes Noturnas e Estudos Populacionais

É possível utilizar as imagens de luzes noturnas como informação auxiliar nos estudos populacionais?

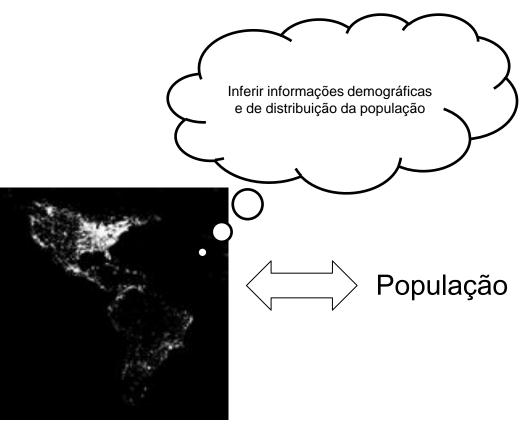


Aplicações do Sensoriamento Remoto de Luzes Noturnas



Objetivo Geral

 O objetivo geral deste trabalho é identificar as contribuições de imagens de luzes noturnas para estimativas de população em duas escalas e regiões distintas: a Amazônia Brasileira e a Região Metropolitana de São Paulo.



Objetivos Específicos

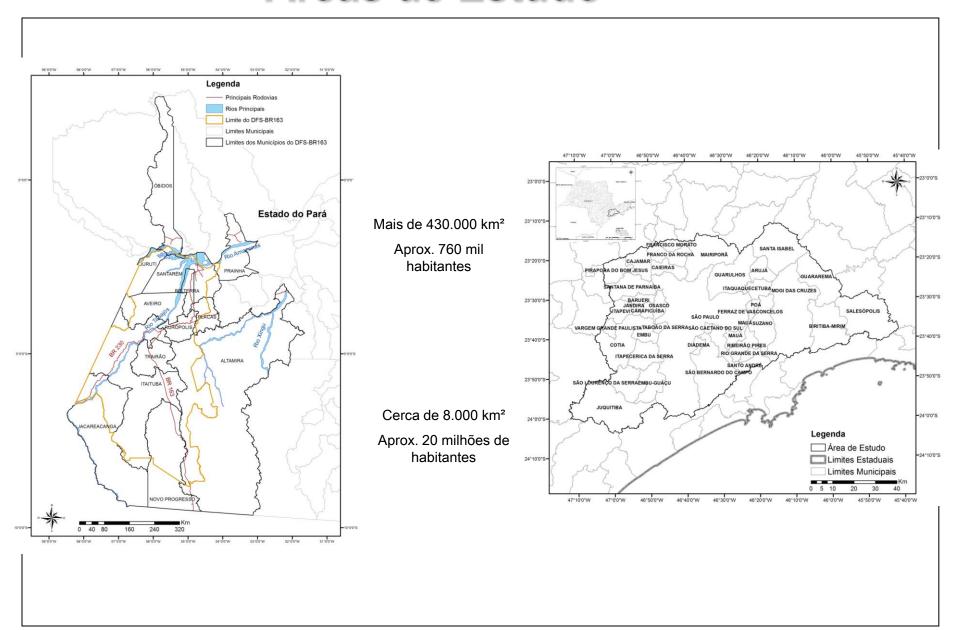
Melhoria dos dados de luzes noturnas

- Intercalibração das imagens mosaico do OLS e organização da série temporal;
- Melhoria do mosaico de luzes do VIIRS;

Estudo das relações com estimativa e distribuição de população

- Escala regional: qual a relação entre as imagens de luzes noturnas e as informações demográficas nesta escala?
- Escala local: qual a importância do espaço nas análises?

Áreas de Estudo



Material



Origem	Dados	1999	2000	2001 a 2006	2007	2008	2009	2010	2011	2012	2013
S. R.	Luzes Noturnas	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS e VIIRS	
	Censo Populacional										
Demográficos	Contagem da População										
	Setores Censitários										
	Localidades IBGE*									-	
Mapeamento sedes Secundários Mapeamento sedes municipais e localidades											
Securidarios	-limites**										
Pontos de Queimadas***				•							
	Pontos de Campo										

Diagrama Metodológico

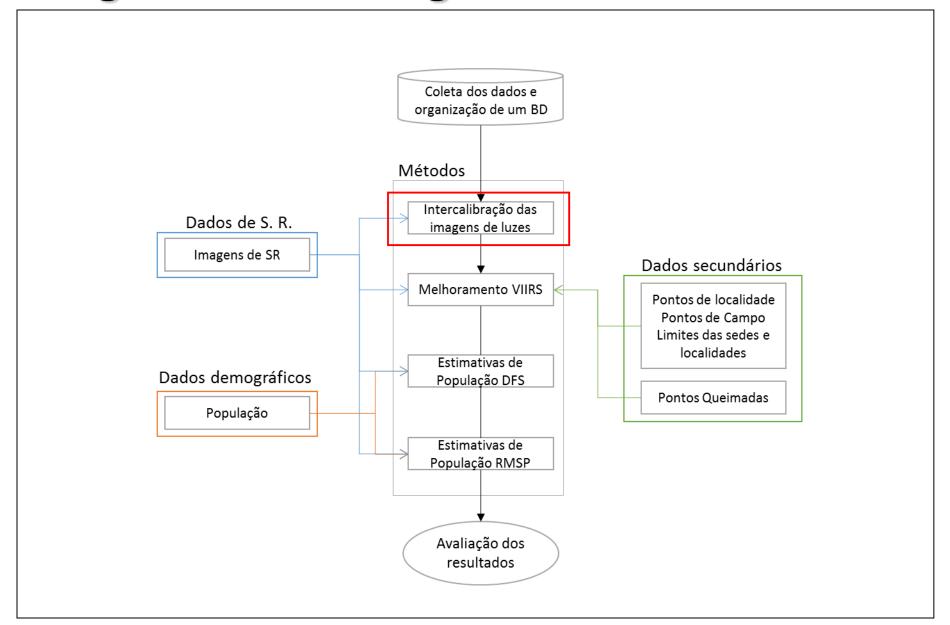
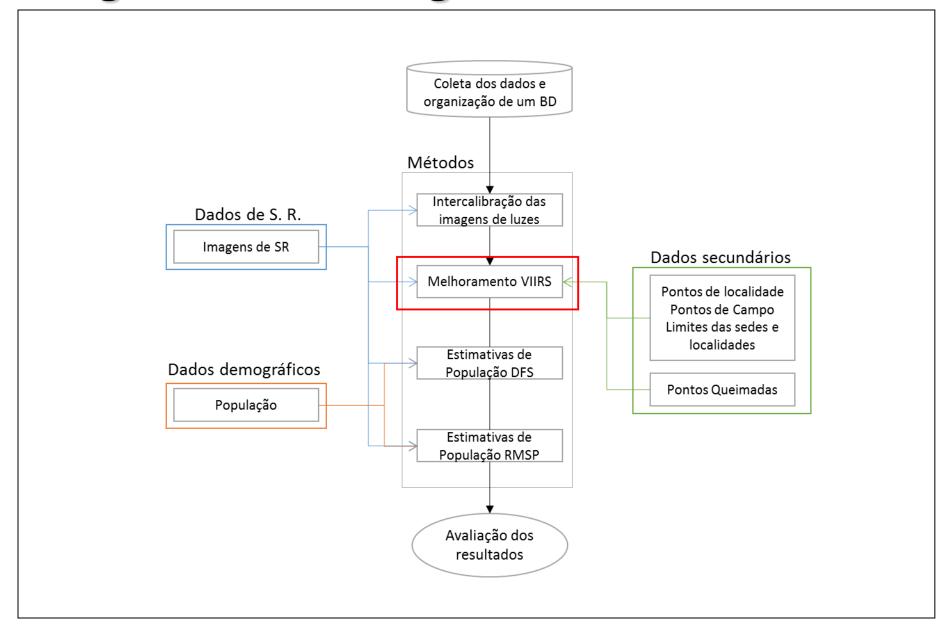
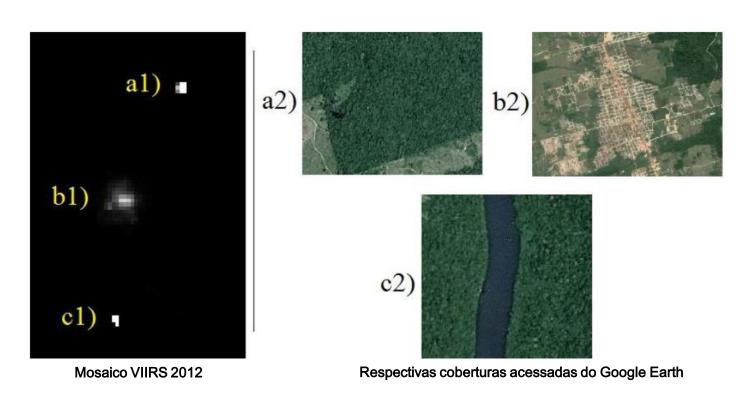


Diagrama Metodológico



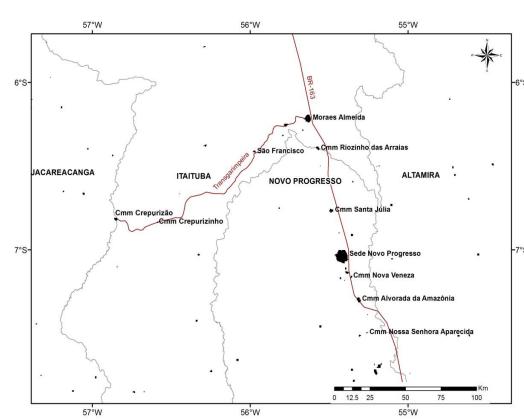
Melhoramento do mosaico de luzes noturnas do VIIRS

 Exemplos de detecção das luzes do sensor VIIRS



Melhoramento do mosaico de luzes noturnas do VIIRS

- Classificação nãosupervisionada
 ISODATA da imagem mosaico VIIRS 2012:
 - Máximo de 3
 iterações; Máximo
 de 15 classes (foram
 geradas 14); Limiar
 de mudança de
 classe de 50%;
 Mínimo de um pixel
 em cada classe;
 - Distância de 0.3 (unidade é ND) na média de cada classe;
 - Duas classes: luzes e background.



Exemplo da classificação do mosaico VIIRS com identificação de algumas comunidades visitadas em campo.

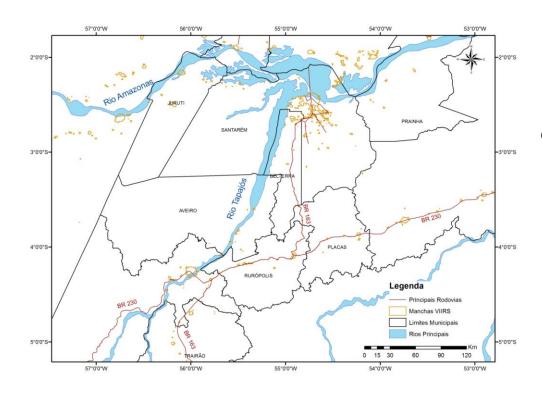
(baseado em Li et al., 2013)

Melhoramento do mosaico de luzes noturnas do VIIRS

Intersecção:

- Classificação OLS
- Pontos de Localidade
- Pontos de Campo
- Setores
 Censitários
 Urbanos 2010
- Queimadas INPE

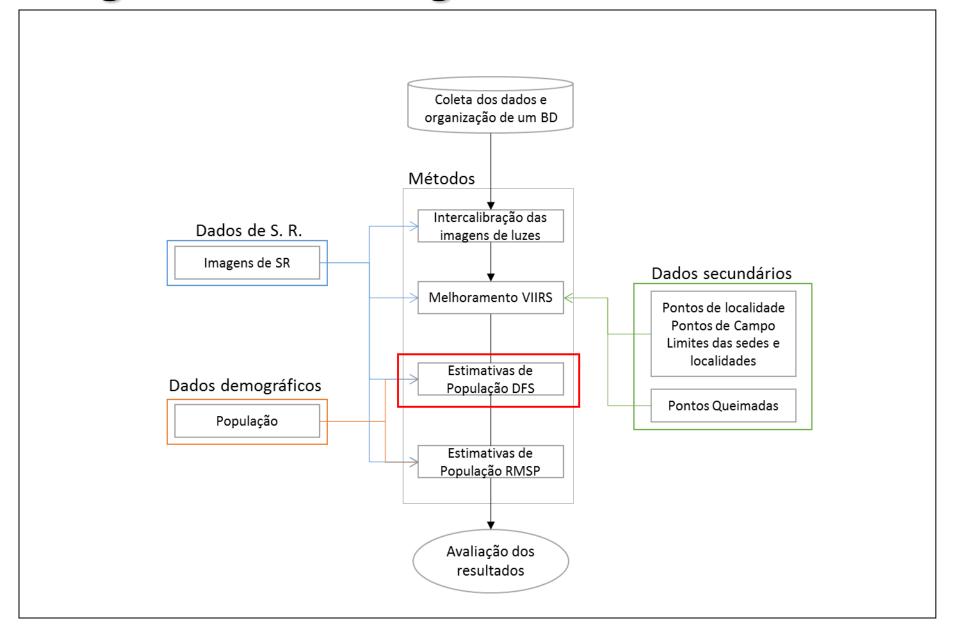
(baseado em Li et al., 2013)



Class. VIIRS

Class. VIIRS pós execução da metodologia

Diagrama Metodológico

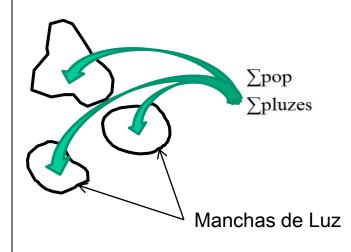


Variável população:

- •Identificação dos setores
 - •Soma da população

Variável luzes:

•Soma dos NDs dos pixels das manchas



Vila do Porto Rico, Jacareacanga

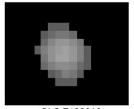


OLS F152007



Google Earth de maio/2007

Bela Vista do Caracol, Trairão



OLS F182010



Google Earth de setembro/2013

Jardim do Ouro, Itaituba



VIIRS 2012



Google Earth de dezembro/2014

Santa Júlia, Novo Progresso



VIIRS 2012



Google Earth de outubro/2008

Período considerado: 2000, 2007 e 2010



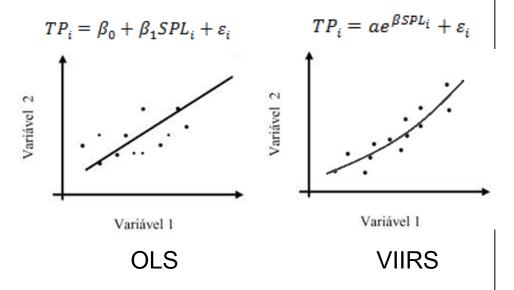
Síntese das manchas classificadas como luzes nas imagens OLS e VIIRS que foram selecionadas e descartadas, e da natureza dos setores censitários selecionados para compor o quantitativo populacional para cada sensor/ano.

Sensor/Ano	Man	chas	Setores Co Selecio	
	Selecionadas	Descartadas	Urbanos	Rurais
OLS F142000	28	7	296	40
OLS F152000	27	4	296	42
OLS F152007	28	6	347	30
OLS F162007	25	6	347	32
OLS F182010	43	6	564	75
VIIRS2012	61	84	566	116

INPE

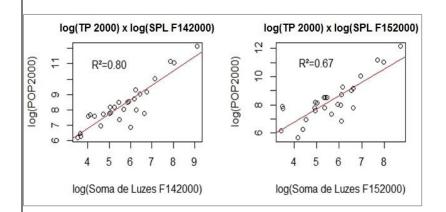
 Análise estatística de regressão das variáveis População (TP) e Luzes Noturnas (SPL)

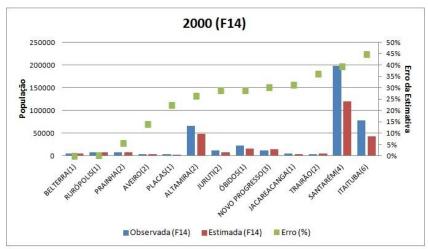
Regressão						
	Variável					
Período	Resposta	Variável Explicativa				
2000	População	Soma de Luzes OLS				
2007	População	Soma de Luzes OLS				
2010	População	Soma de Luzes OLS				
		Soma de Luzes VIIRS				
2010	População	2012				

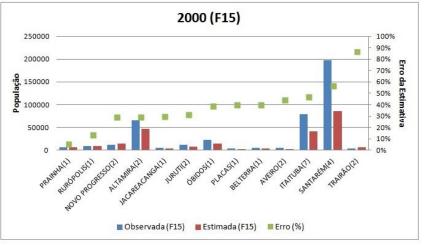




Estimativas DFS-BR163

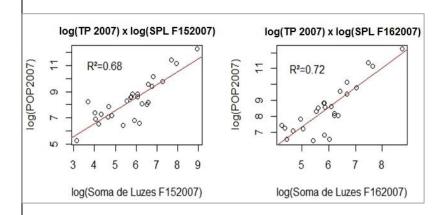


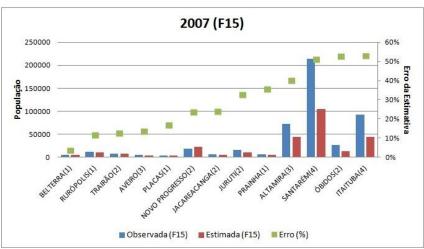


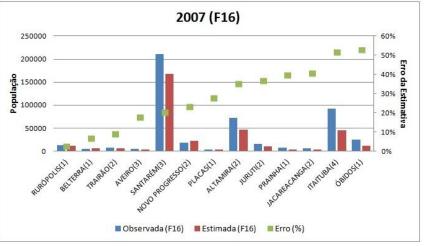




Estimativas DFS-BR163

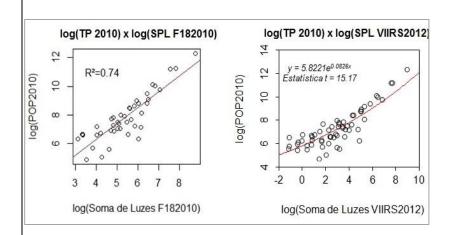


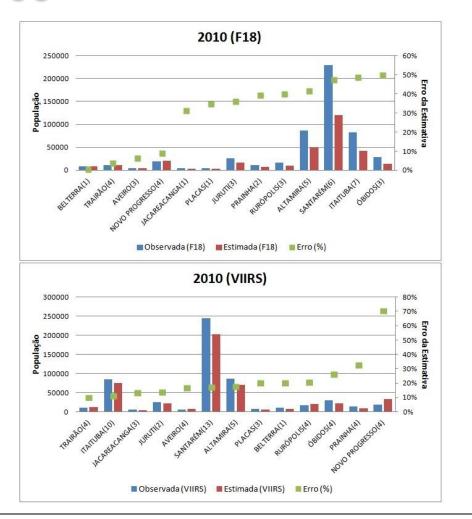






Estimativas DFS-BR163







Erros (porcentagem) das estimativas de população das manchas de luz por município.

Município	200	2000)7	201	-	
Município	F14	F15	F15	F16	F18	VIIRS	
Itaituba	-44.81	-46.93	-52.76	-51.98	-48.68	-11.14	
Santarém	-39.42	-56.88	-50.84	-20.67	-47.48	-17.36	
Jacareacanga	-31.27	-30.08	-23.73	-40.98	-31.39	-13.47	
Óbidos	-28.86	-39.25	-52.52	-53.04	-49.88	-26.25	
Juruti	-28.79	-31.36	-32.41	-37.17	-36.27	-13.70	
Altamira	-26.63	-29.48	-39.92	-35.32	-41.52	-17.83	
Placas	-22.55	-40.05	-16.72	-27.85	-34.91	-20.44	
Aveiro	-14.03	-44.55	-13.52	-18.07	-6.48	16.98	%
Prainha	-5.80	5.49	-35.53	-40.08	-39.41	-32.67	
Rurópolis	-0.49	13.44	-11.44	-2.79	-39.88	20.86	
Belterra	0.17	-40.11	3.48	6.93	-0.56	-20.48	
Novo Progresso	30.30	29.26	23.51	23.34	8.91	70.49	
Trairão	36.25	86.68	12.64	-9.15	3.88	9.99	
EMAP	23.80	37.97	28.39	28.26	29.94	22.44	
MEAP	28.79	39.25	23.73	27.85	36.27	17.83	

Métodos e Resultados



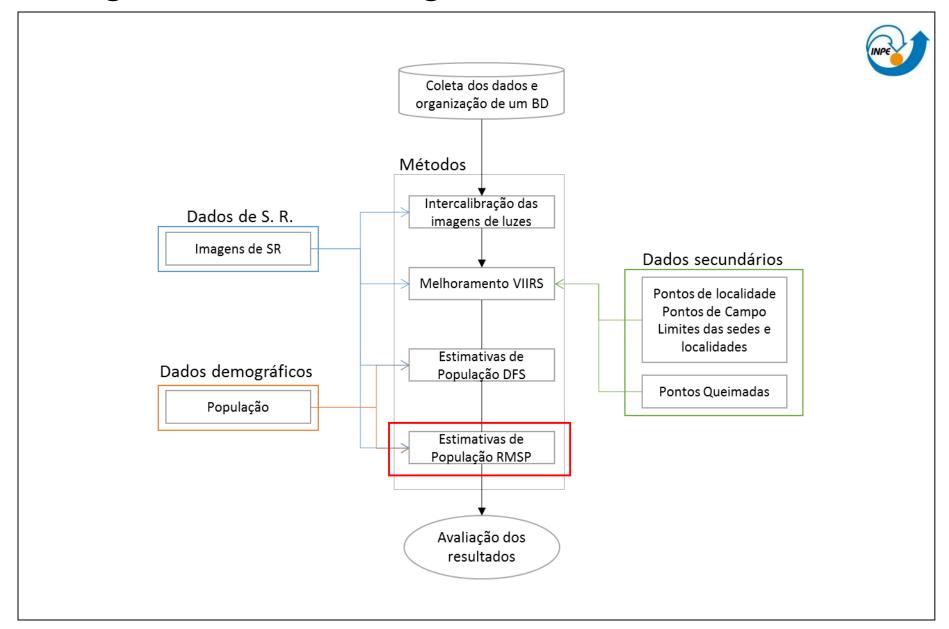


Complexos de madeireiras na sede de Novo Progresso, ao longo da BR 163. Imagens de satélite obtidas do Google Earth.



Localização de exemplo de madeireira no entorno da sede de Novo Progresso, ao longo da BR 163. Imagens de satélite obtidas do Google Earth.

Diagrama Metodológico



Estimativa e Distribuição de População na RMSP



- Áreas de Ponderação Urbanas;
- População (Censo 2010);
- Imagem de luzes noturnas do VIIRS 2012.
- Regressão Linear (TP x SLP)

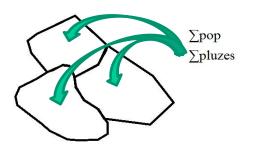
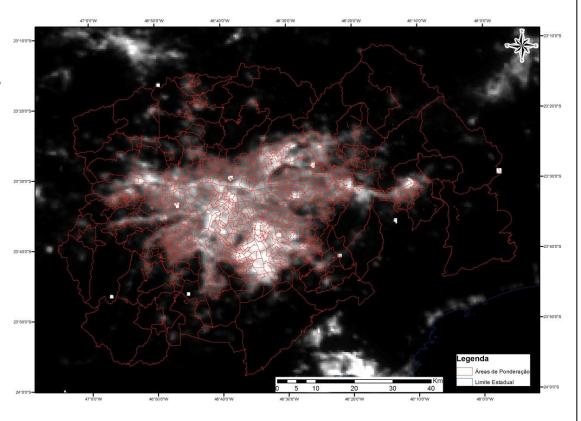
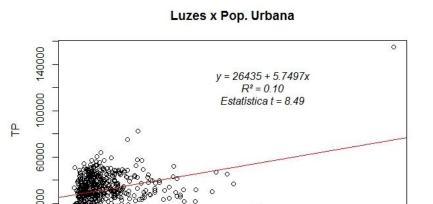


Imagem mosaico de luzes do VIIRS 2012 e áreas de ponderação (vermelho) do censo IBGE 2010 da RMSP





Regressão Linear



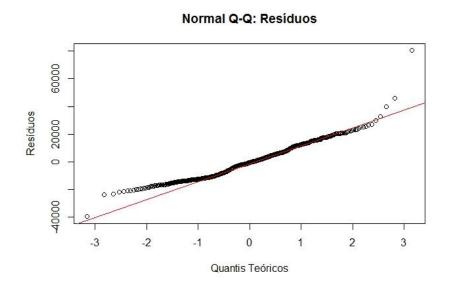
4000

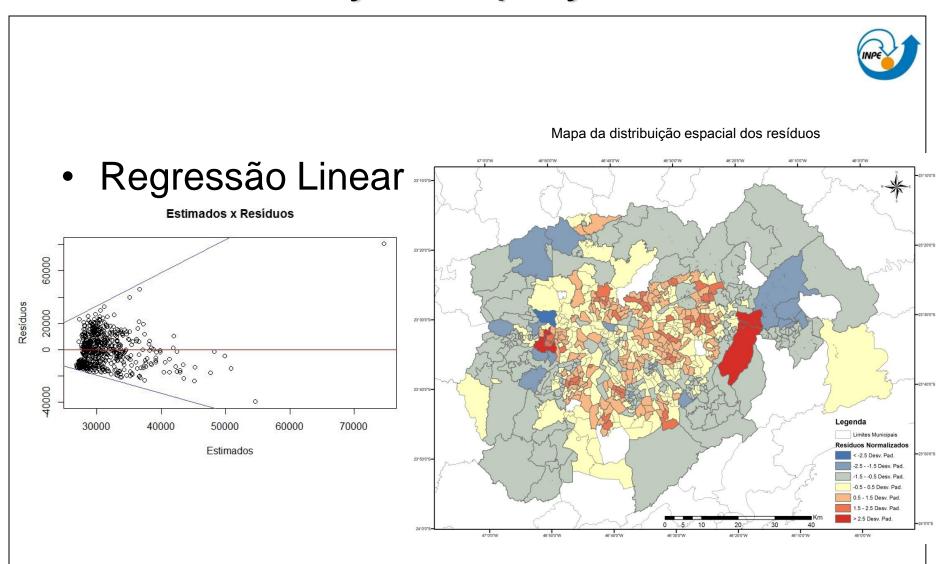
SPL

6000

8000

2000



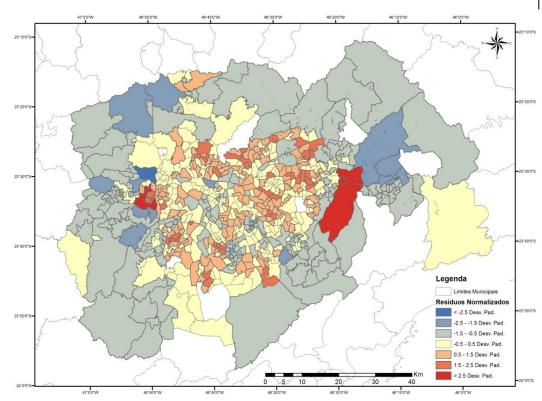




Índice de Moran Global: 0,45

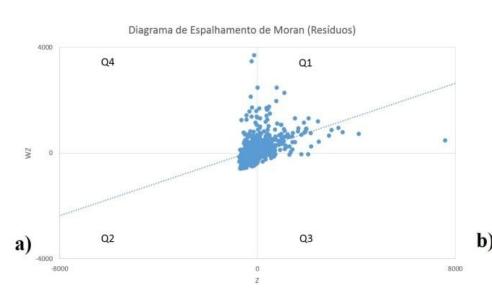
P-valor = 0.01

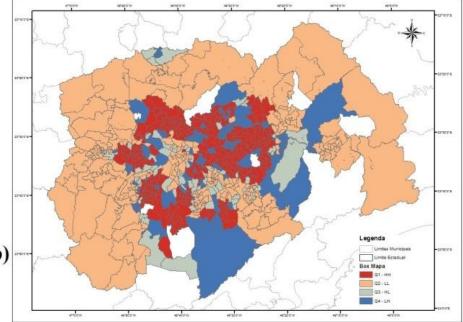
Mapa da distribuição espacial dos resíduos



Global - o valor deste índice varia de -1 a +1, onde -1 indica uma autocorrelação espacial negativa ou inversa, 0 indica aleatoriedade, isto é, quando não há autocorrelação espacial, e +1 indica autocorrelação espacial positiva ou direta.



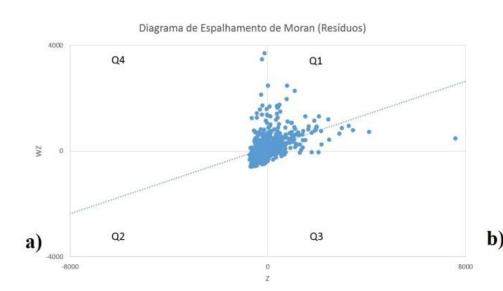




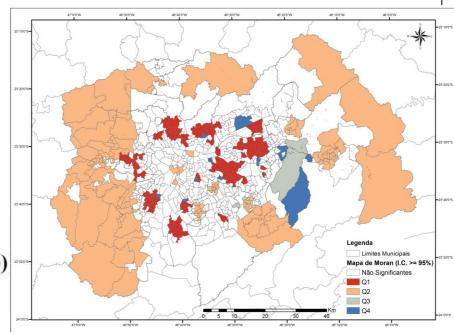
Local – decomposição do índice global e produz um valor para cada área (polígono) da região de estudo

Box Mapa dos Resíduos da regressão linear entre a soma dos pixels de luzes e população total das áreas de ponderação da RMSP.





Local – decomposição do índice global e produz um valor para cada área (polígono) da região de estudo

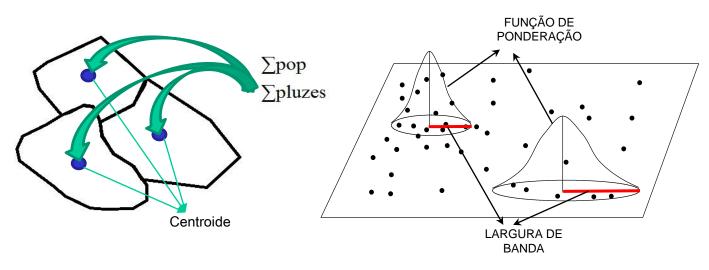


Box Mapa dos Resíduos da regressão linear entre a soma dos pixels de luzes e população total das áreas de ponderação da RMSP, considerando 5% de significância.



Geographically Weighted Regression – GWR (local)

$$TP_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)SPL_i + \varepsilon_i$$



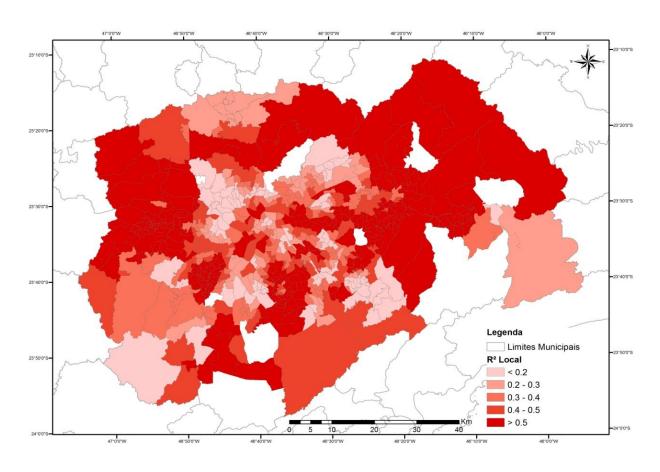
Função: Gaussiana

Banda adaptativa: 25 vizinhos

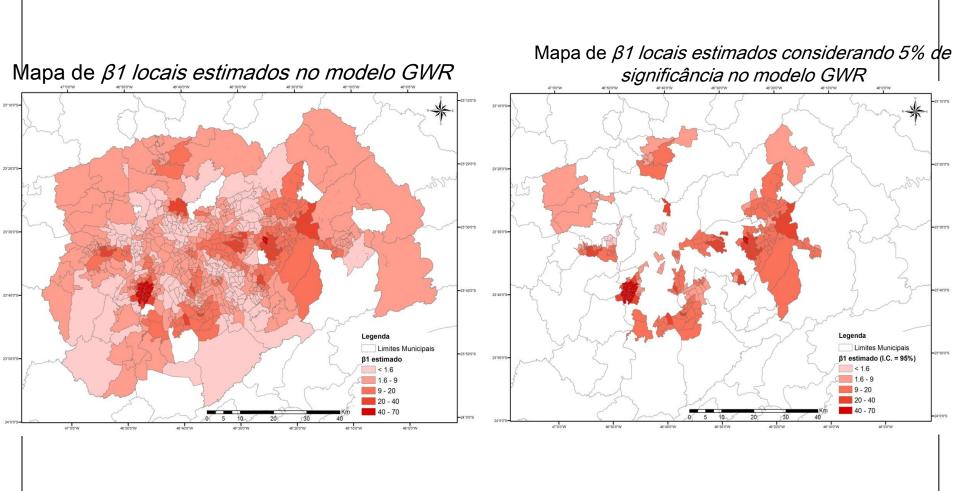


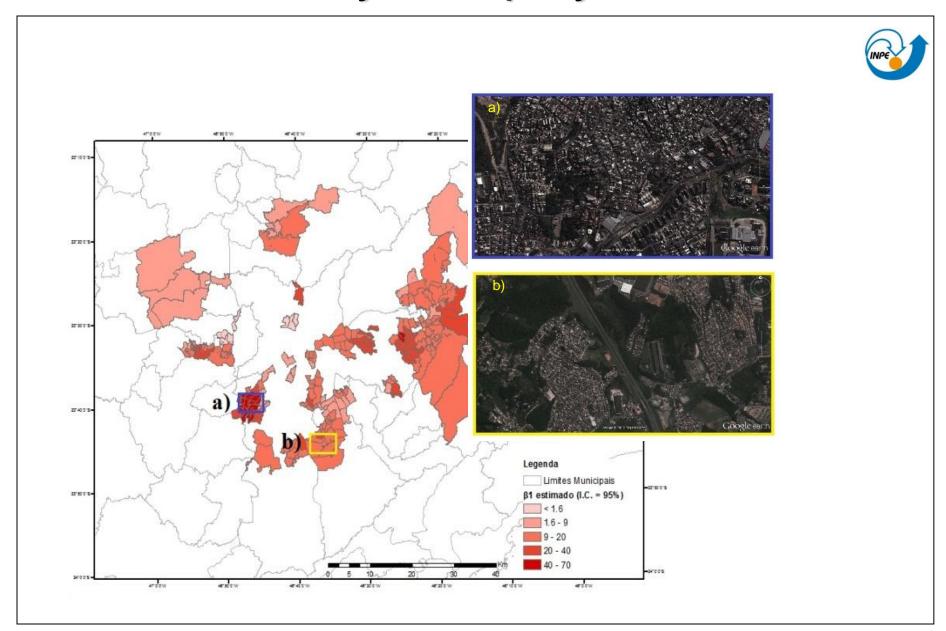
Mapa dos coeficientes de determinação locais do modelo GWR













Níveis de exatidão da população estimada usando o modelo GWR

Exatidão do erro relativo de população estimada por GWR		Áreas de ponderação (%)	
Exatidão Alta	(0 - 25%)	77,87	
Exatidão Moderada (25 - 50%)		18,09	
Inexato	(> 50%)	4,04	

(baseado em Li et al., 2013)





Francisco Morato - Estas áreas urbanas são, em sua maioria, caracterizadas como de baixa renda e carentes de investimento público em infraestrutura, equipamentos urbanos e projetos sociais (MORAES, 2007).



Cotia - Essa superestimativa pode ser atribuída ao fato de que na região destas duas áreas de ponderação existirem diversos empreendimentos industriais que emitem luz durante a noite, contribuindo para um valor alto na estimativa, sem contudo possuírem população residente.

Considerações Finais



- Intercalibração das imagens multitemporais do sensor OLS;
- Melhoria do mosaico VIIRS;
- Estimativa de população para a região do DFS-BR163;
- Capacidade do sensor VIIRS em identificar e individualizar áreas de ocupação humana:
 - Dado auxiliar para distribuição espacial da população;
 - Pode auxiliar órgãos governamentais e, por exemplo, na desagregação dos dados de população da Grade Estatística (IBGE).

Considerações Finais



- Estimativa e distribuição espacial de população para a RMSP:
 - Os diferentes regimes espaciais apresentados pelos coeficientes do GWR indicam padrões diferentes de ocupação urbana;
 - Em função de poucos coeficientes terem sido significativos, talvez a escala de análise esteja muito grande;
 - A população estimada pelo GWR se apresentou bastante consistente com os dados oficiais, no entanto poucos coeficientes foram significantes;

Lidar 4 Pop estimation...

International Journal of Remote Sensing Vol. 31, No. 21, 10 November 2010, 5571–5586



Evaluation of small-area population estimation using LiDAR, Landsat TM and parcel data

PINLIANG DONG*, SATHYA RAMESH and ANJEEV NEPALI
Department of Geography, University of North Texas / 1155 Union Circle #305279,
Denton, TX 76203, USA

International Journal of Remote Sensing Vol. 31, No. 21, 10 November 2010, 5587–5604



Population estimation based on multi-sensor data fusion

ZHENYU LU†, JUNGHO IM*†, LINDI QUACKENBUSH† and KERRY HALLIGAN‡

International Journal of Remote Sensing Vol. 31, No. 21, 10 November 2010, 5605–5634



Assessing fine-spatial-resolution remote sensing for small-area population estimation

JOSE L. SILVÁN-CÁRDENAS†, LE WANG*†, PETER ROGERSON†, CHANGSHAN WU‡, TIANTIAN FENG† and BENJAMIN D. KAMPHAUS§

Lidar 4 Pop estimation...

International Journal of Remote Sensing Vol. 31, No. 21, 10 November 2010, 5571–5586



Evaluation of small-area population estimation using LiDAR, Landsat TM and parcel data

PINLIANG DONG*, SATHYA RAMESH and ANJEEV NEPALI

Department of Geography, University of North Texas / 1155 Union Circle #305279,

Denton, TX 76203, USA

- Lidar modelo digital de superfície e modelo digital de elevação → modelo digital de superfície normalizado
- Dados de parcelas máscara das áreas não residenciais e não comerciais
- Classificação TM-Landsat → refinamento das páreas residenciais
- Unidades do censo ("census block" contínuos e aleatórios) → Cálculo de área e volume das construções (morfologia matemática, estatísticas zonais, GIS)

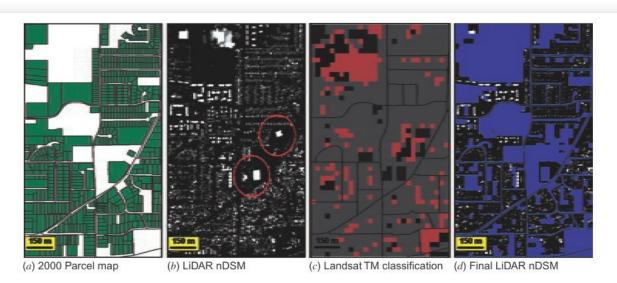


Figure 6. Samples of input data layers and final LiDAR nDSM. Several large buildings in the red circles in (b) are classified as CIT in (c), where grey is residential buildings, red is CIT, yellow is soil, and green is vegetation, and removed from the final LiDAR nDSM in (d), where blue areas represent a non-residential mask.

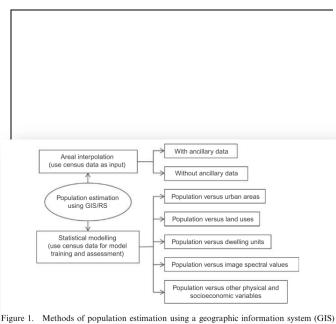
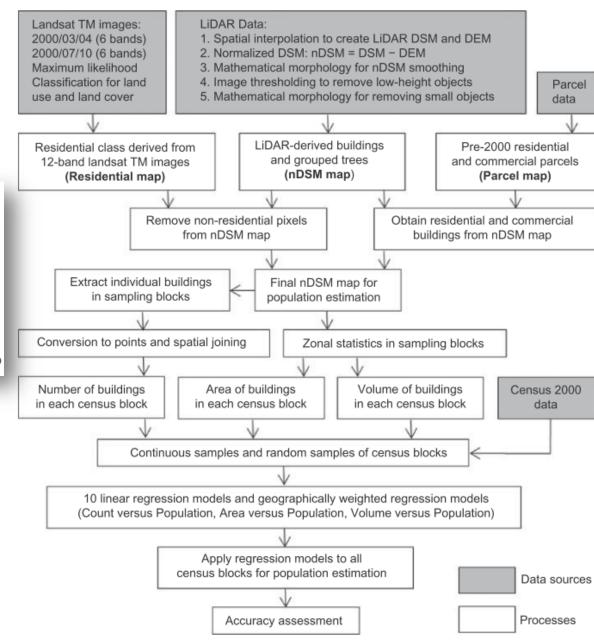


Figure 1. Methods of population estimation using a geographic information system (GIS) and remote sensing (RS).



- Para cálculo da população total:
 Modelos de regressão por mínimos
 quadrados (OLS) e
 geograficamente ponderada
 (GWR)
- A partir de variáveis independentes de contagem, área e volume de construções obtidos do modelo de superf normalizado (Lidar)
- Modelo aplicado aos censos blocks
- Análise de acurácia

Figure 7. Regression models derived from sampling blocks: (a), (b) and (c) are frou ous sampling blocks shown in figure 4(a), while (d), (e) and (f) are from random blocks shown in figure 4(b).

Table 1. Summary of continuous sampling model results.

Independent variable	Regression model	R^2	MRE (%)	MARE (%)
nDSM building count	y = 2.948x + 7.4341		53.46	29.43
nDSM building area	y = 0.0073x + 3.3004		52.84	41.75
nDSM building volume	y = 0.0005x + 6.0216		55.77	47.12

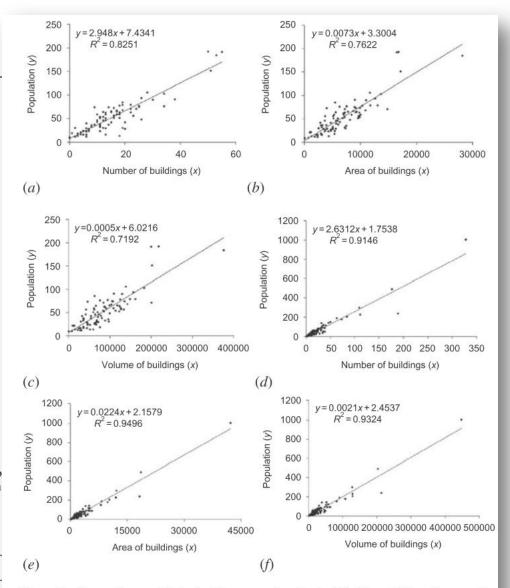


Figure 7. Regression models derived from sampling blocks: (a), (b) and (c) are from continuous sampling blocks shown in figure 4(a), while (d), (e) and (f) are from random sampling blocks shown in figure 4(b).

MRE: mean relative error for the overall dataset; MARE: median absolute relative error for the overall dataset; TE: total population estimation error (%) based on the overall dataset in the study area; R^2 : coefficient of determination.

- GWR melhores resultados que OLS – por melhor representar a heterogeneidade espacial
- Acurácia total é controlada pelo Sinal e magnitude dos erros relativos de estimação no nível dos census blocks
 - Substimação nas áreas de alta densidade populacional
 → total de pop é subestimado (23%)
- Sugere uso de Lidar com melhor resolução espacial

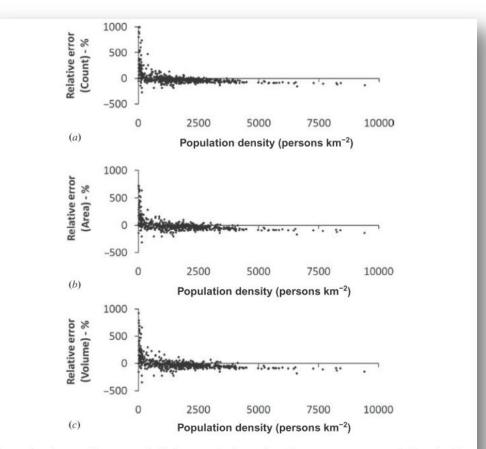


Figure 8. Scatter diagrams of relative population-estimation error versus population density at the census-block level. The relative errors were obtained from OLS models built from random samples using (a) building count as the independent variable, (b) building area as the independent variable and (c) building volume as the independent variable.

Concluindo

Ferramentas de Geoinformação para Estimativas de População

- Potencial de SR + metodologias
- Não há produção sistemática/ operacional específica para estimativa de população
- Há dados disponíveis, metodologia aplicável, é necessário competência e empenho...









Referências

- AFONSO, I. Estimação da população dos setores censitários de Belo Horizonte usando imagens de satélite. In: **XII Simpósio Brasileiro de Sensoriamento Remoto**, Proceedings. Goiânia (GO), 2005. p.2741-2748.
- FEITOSA, F.; MONTEIRO, A. M. V.; CÂMARA, G. Compatibilização de dados censitários para análises temporais com o auxílio de imagens Landsat. In: **XII Simpósio Brasileiro de Sensoriamento Remoto**, Proceedings. Goiânia (GO), 2005. p.2657-2664.
- GONÇALVES, C. D. A. B., et al. Análise do ambiente residencial urbano visando a inferência populacional a partir do uso de dados de sensoriamento remoto orbital de alta resolução. In: XIV Encontro Nacional de Estudos Populacionais ABEP, Proceedings. Caxambu (MG), 2004.
- NOVACK, T.; KUX,H. J. H.; MONTEIRO, A. M. V.; PINHO, C. M. D. . Estimation of Population Density Using High Resolution Remote Sensing Data and Spatial Regression Techniques: A Case Study in São Paulo City (Brazil). In: Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação, 2008, Recife. Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação, 2008
- PINHO, C. M. D.; UMMUS, M. E.; NOVACK, T. . Simulação do comportamento espectral de alvos urbanos em sensores multiespectrais. In: XIV SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 2009, Natal. XIV SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO XIV SBSR. São José dos Campos : INPE, 2009. p. 803-810.
- SOUZA, I. M. E.; DURAND, C.; ALMEIDA, Claudia Maria de; PINHO, C. M. D. . Uso de imagens de alta resolução espacial e análise orientada a objeto para caracterização socioeconômica do espaço residencial construído. In: XIV SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 2009, Natal. XIV SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO XIV SBSR. São José dos Campos: INPE, 2009. p. 875-882

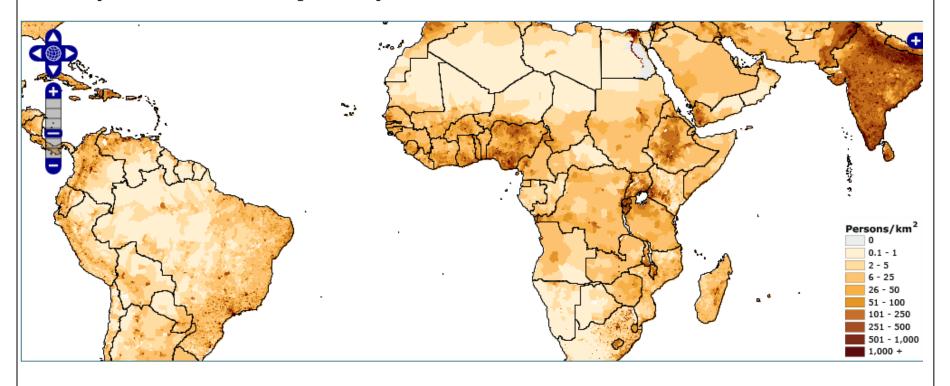


Superfícies de população para diferentes aplicações

GPW and GRUMP - Começou com Tobler...

Gridded Population of the World and the Global Rural - Urban Mapping Project

Base Layer: GPWv3 2005 Pop Density



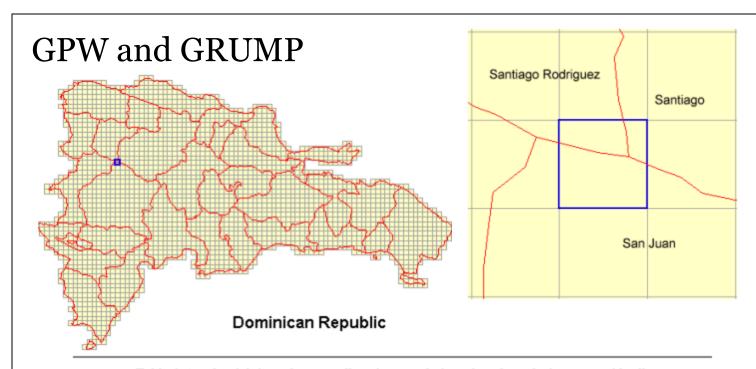


Table 6. Areal weighting scheme to allocation population whose boundaries cross grid cells

Administrative unit name	Administrative unit density (persons / sq km)	Area of overlap (sq km)	Population estimate for grid cell
Santiago Rodriguez	64.2	5.3	340
Santiago	246.5	2.2	542
San Juan	75.9	12.8	972
Total for cell	91.3	20.3	1854

http://sedac.ciesin.columbia.edu/gpw/

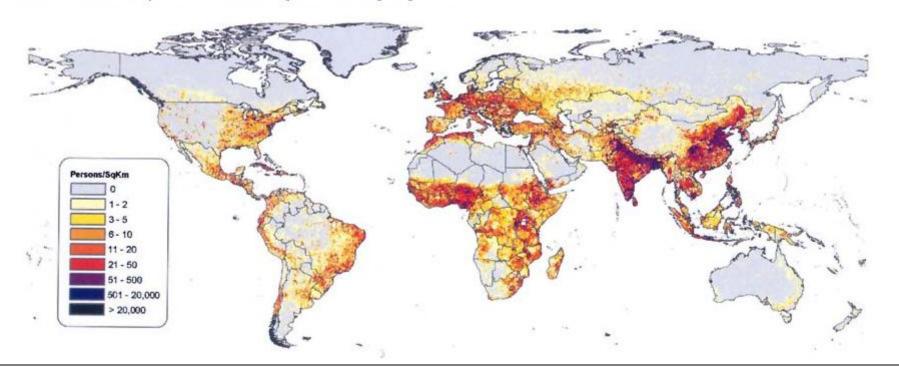
POPULATION AT RISK POPULATION AT

LANDSCAN 2000

- · Roads, weighted by distance from major roads.
- · Elevation, weighted by favourability of slope categories.
- · Land cover, weighted by type with exclusions for certain types.
- · Nighttime Lights of the World, weighted by frequency.

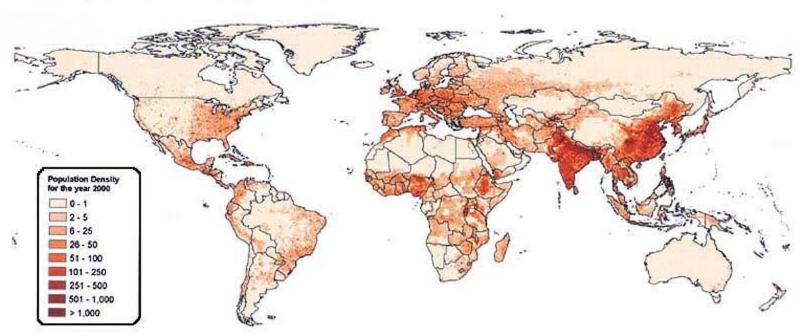
The resulting coefficients are weighted values, independent of census data, which can then be used to apportion shares of actual population counts within any particular area of interest. Coefficients vary considerably from country to country even within different regions of the same country.

MAP 3.2 LandScan Global Population Database, adjusted to UN figure year 2000



GRUMP 2000

MAP 3.3 Population density in 2000 from GRUMP adjusted to UN totals



Source: Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Inst. (IPFRI), the World Bank and Centro Internacional de Agricultura Tropical (CIAT)

http://www.fao.org/docrep/009/a0310e/A0310E06.htm#ch3

- Elvidge et al. A global poverty map derived from satellite data.
 Computers & Geosciences 35 (2009) 1652–1660
 - Mapa de pobreza baseado em estimativas populacionais globais
 - LANDSCAN 2004 x Luzes Noturnas 2003
 - Resolução 30 arsec ~1km

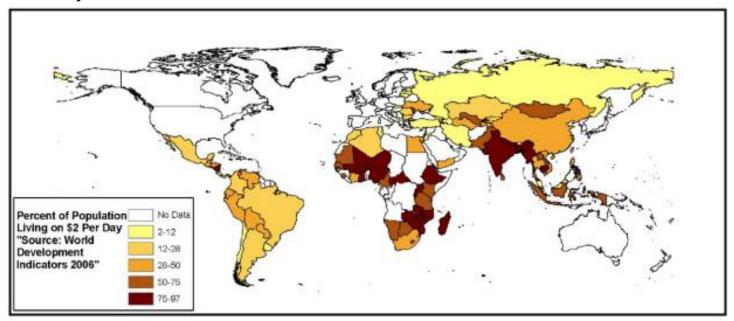
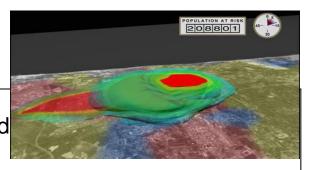


Fig. 1. Map of poverty levels for countries reporting international poverty line data (percent of population living on \$2 per day or less) from World Development Indicators 2006. Note that a number of countries have no data reported and that \$2 per day poverty line is not applicable to developed countries.

LandScan – alocação espacial de censo, baseado em mod de desagregação de dados



- population count is used instead of population density, baseado em residência.
- population count ou ambient pop quer representar pop baseado em hora/pessoa
 - NASA MODIS land cover (Friedl et al., 2002),
 - topographic data (SRTM) (Rodriguez et al., 2005),
 - high-resolution land cover data of the Controlled Image Base (CIB) from the US National Geospatial Intelligence Agency (NGA)

Multi-layered, <u>dasymetric</u>, <u>spatial modeling</u> approach - "smart interpolation" technique.

- 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.
- weighted for the possible occurrence of population during a day
- the population distribution model calculates a "**likelihood**" **coefficient** for each cell and applies the coefficients to the census counts
- The resultant population count is an ambient or average day/night population count

•http://www.ornl.gov/sci/landscan/landscan documentation.shtml

LandScan – alocação espacial de censo, baseado em mod de desagregação de da tead of g population count is u seado em residência. population count ou à baseado hora/pessoa "Our pricing structure is topog dependent upon the high-res the US National number of users who will have access to the polation" Multi-lav Dataset and the number of technio uses su different sites occupied by oads, slope, urban areas, • and those users" all of which are weighted for on during a alculates a kelihood" coefficient for each cell and applies the the population coefficients to census counts • The resultant population count is an ambient or average day/night population count

http://www.ornl.gov/sci/landscan/landscan documentation.shtml

Índice de Pobreza:

count POP LandScan

2004

DN médio das luzes

Maiores valores de IP ocorrem em áreas com muita população e pouca ou nenhuma luz detectada pelo DMSP!

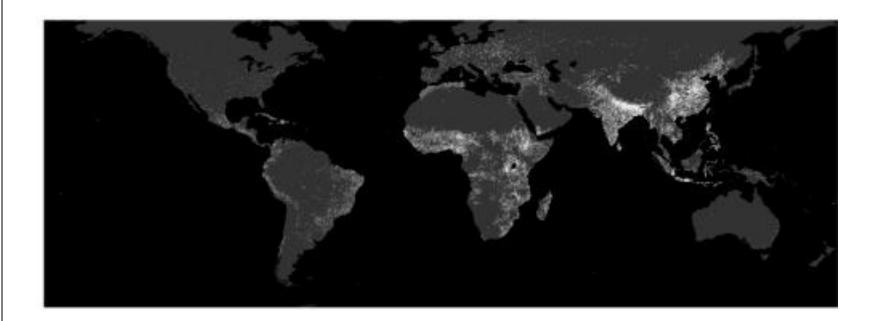


Fig. 2. Poverty index calculated by dividing LandScan 2004 population count by average digital number of DMSP satellite F15 nighttime lights from 2003.

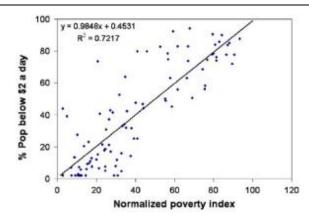


Fig. 3. Calibration of normalized poverty index (NPI) for estimation of poverty levels.

Table 2

Decadal classes of national poverty level estimates for 81 countries having populations in excess of 10 M.

Poverty level	Countries
1-10	Taiwan, S. Korea, Egypt, Saudi Arabia, Japan, Belgium, Netherlands, Italy, United Kingdom, USA, Canada, Czech Republic, Germany, Greece, Spain, Hungary, France
11-20	Venezuela, Portugal, Iran, Poland, Australia, Mexico, Syria, Uzbekistan, Iraq, Serbia, Malaysia, Argentina, Chilo, Romania, Russia
21-30	Algeria, Pakistan, Colombia, Ukraine, Turke <mark>, Brazil, E</mark> tuador, Sri Lanka
31-40	Byelarus, Indonesia, Kazakhstan, Cuba, South Africa, Guatemala, Vietnam, Morocco, Peru, Thailand, Philippines
41-50	India, China
51-60	Bangladesh, Cote d'Ivoire, Ghana, Senegal, Yemen, Nigeria
61-70	North Korea, Zimbabwe, Zambia, Cameroon, Sudan
71-80	Nepal, Kenya, Angola, Myanmar, Congo DRC, Mali, Mozambique, Afghanistan
81-90	Malawi, Niger, Tanzania, Uganda, Cambodia, Madagascar, Burkina Faso, Ethiopia

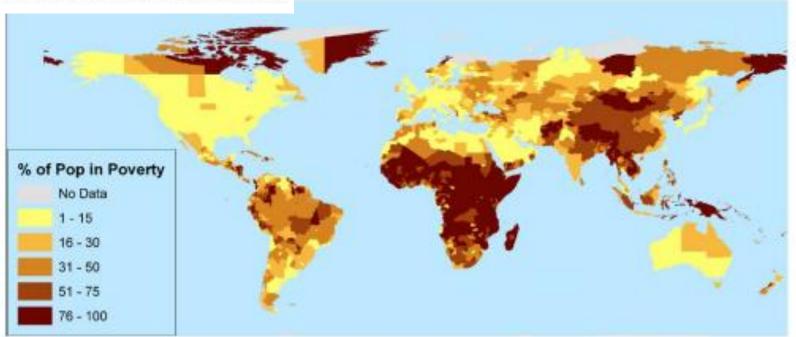


Fig. 4. Map of poverty levels for 2543 sub-national administrative units estimated based on satellite data-derived poverty index.

População por SR – Média resolução

Compatibilização de dados censitários para análises temporais com o auxílio de imagens Landsat.

População por SR – Média resolução

- Problema compatibilização das geometrias dos setores censitários e seus dados para viabilizar análise multi-temporal.
- Dados censitários de 1991 e 2000 relativos à área urbana do município de São José dos Campos – SP.
- Área de estudo passou de 347 para 739 setores entre 1991-2000
 - Censos demográficos 1991 e 2000
 - Tabelas de compatibilidade entre os setores 1991-1996 e 1996-2000
 - Imagens Landsat-5/TM (1990) e Landsat-7/TM (2000) bandas 3, 4
 e 5
 - Mosaico digital orto-retificado (1:30.000)
 - Dados complementares, sistema viário, quadras (1:2.000)

População por SR – média resolução

- Registro das imagens referência mosaico ortorretificado de 2000
- Interpretação visual das imagens composições coloridas para extrair as áreas de ocupação urbana

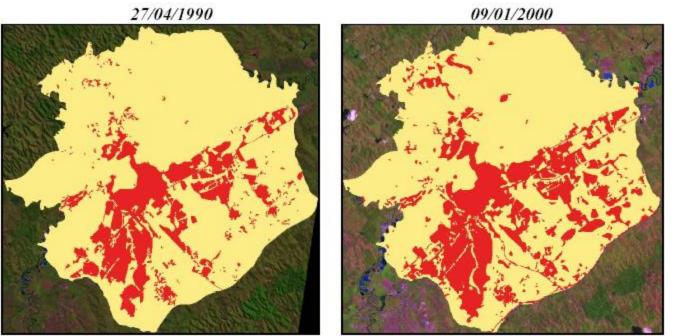


Figura 2 - Delimitação das manchas urbanas (em vermelho) sobre as imagens Landsat-5/TM (27/04/90) e Landsat-7/ETM+ (09/01/00).

Fonte: INPE (1990, 2000).

População por SR – média resolução

- Compatibilização das bases geográficas dos setores censitários
 - 1991=> áreas urbanas consolidadas
 - Demais áreas => setores 2000
 - Edição vetorial

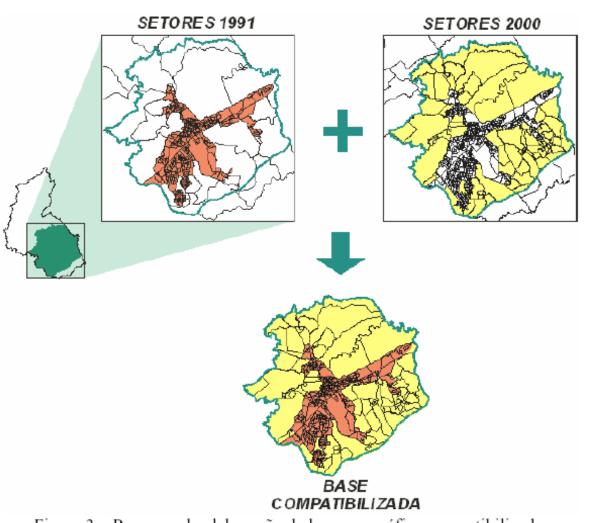


Figura 3 – Processo de elaboração da base geográfica compatibilizada.

População por SR - média resolução

- Tabelas de equivalência entre os polígonos da base compatibilizad a e os setores originais
 - Pesos proporcionai s às áreas

ID9100	ID 1991	PES 091	ID2000	PESO00
9100001	35499005000001	1	354990405000001	1
9100003	35499005000003	1	354990405000003	1
9100003	35499005000003	1	354990405000004	1
9100292	35499005000045	0,43	354990405000061	1

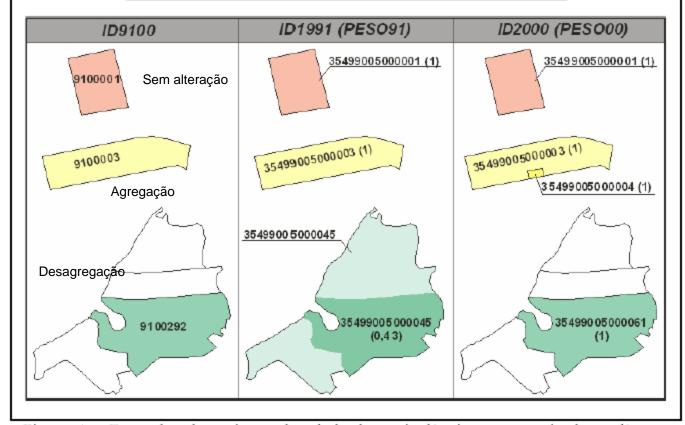


Figura 4 – Exemplos de registros da tabela de equivalência e geometria dos polígonos

População por SR – média resolução

- 421 polígonos na base compatibilizada
- Resultado: tabela com ID dos polígonos, dados censitários para 1991 e 2000 associada a base compatibilizada

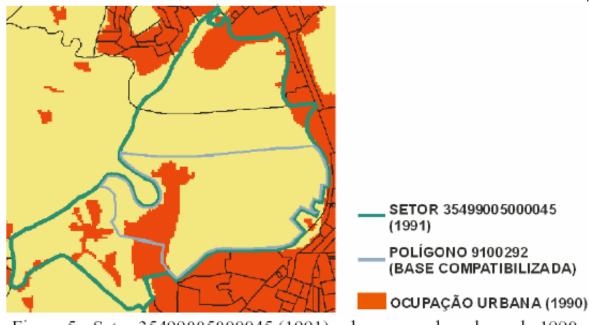


Figura 5 - Setor 35499005000045 (1991) sobre a mancha urbana de 1990.

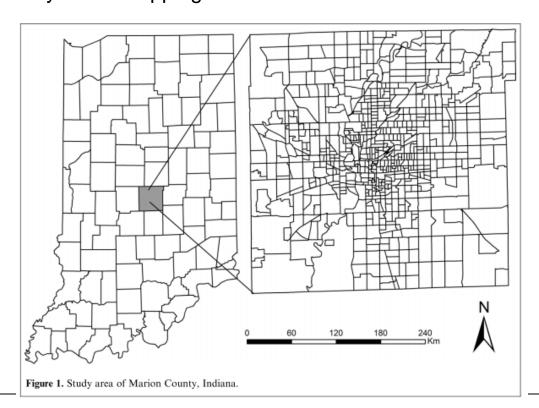
População por SR – média resolução

Considerações:

- O procedimento de Desagregação
 - considerou apenas área ocupada por usos urbanos em geral e os dados censitários referem-se apenas <u>às áreas com</u> <u>domicílios</u>;
 - Foi definida em função da área ocupada, desconsiderando diferenças de densidade.
- Ideal seria incorporar <u>dados auxiliares</u> como mapas de uso do solo identificando áreas residenciais e cadastro de imóveis com identificação de domicílios
- O procedimento minimiza o efeito das geometrias para estudos multi-temporais mas <u>não resolve problema de MAUP</u>, inerente ao processo de aquisição do dado (setores censitários)

3 Objectives:

- (i) examine the utility of <u>land surface temperature</u> and <u>fraction images</u> derived from a subpixel classifier for population estimation;
- (ii) (ii) compare the effectiveness of the <u>spectral response</u> based and the land use based methods <u>for population estimation</u> of US census block groups;
- (iii) (iii) produce a more accurate presentation of population distribution by combining dasymetric mapping with the land use based method.



Marion County:

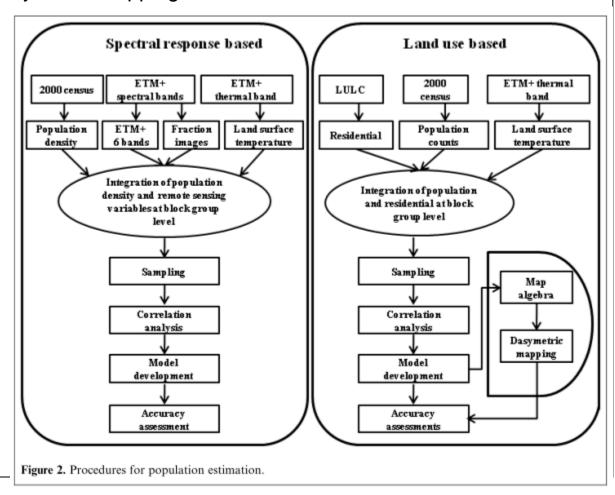
- total area of 1044 km2
- population of 860,454
- -density of 838 people/km²,and 387 183 housing units.
- Indianapolis ranked as the twelfth largest city in the US in 2000 (US Census Bureau, Census 2000 summary file)
- -Marion County has **658** block groups, with an average area of 1.59 km² and an average population density of 1587 people/km²

Methods:

- (i) the combination of spectral bands, land surface temperature, vegetation abundance, and impervious surface; and
- (ii) residential land use and dasymetric mapping based on land use.

Four steps were involved in population estimation:

- (i) extraction of population from the census data,
- (ii) development of remote sensing variables
- (iii) development of population estimation models by the integration of population and remote sensing derived variables, and
- (iv) accuracy assessment



- Two population parameters were: total population and population density.
- Linear spectral mixture analysis (LSMA) ETM+
- Land surface temperature (LST) from Landsat ETM+ thermal band (TIR)
- Urban LULC classification Maximum likelihood classifier
- high-density residential areas = units > 1300 people/km²;
 low-density < 400 people/km²; in between = medium-density
- a decision tree based on housing density at the block level.
- Fifty samples for each LULC type were randomly selected and compared to references collected from high spatial resolution aerial photographs.

Table 1. Accuracy assessment of land use and land cover classification results.

LULC type	Producer's accuracy (%)	User's accuracy (%)
Water	97.9	95.9
Urban	86.6	80.6
High-density residential	51.6	72.7
Medium-density residential	81.4	74.5
Low-density residential	76.3	63.0
Vegetation	91.9	97.0
Overall accuracy (%)	86.0	
Overall kappa	0.81	

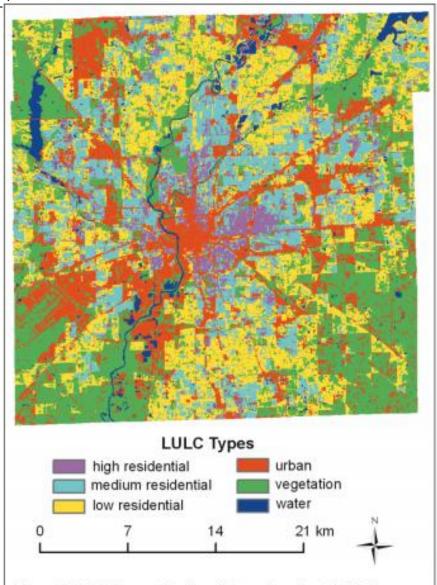


Figure 3. LULC map developed from Landsat ETM+ imagery with the maximum likelihood classifier.

Data integration and model development

- Remote sensing derived variables (e.g., spectral bands, fraction images, and land surface temperature) were aggregated to the block group level, and the mean values of these variables for the block group were then computed
- LULC types were summarized as the count of pixels falling within a block group.
- In addition, the sum of temperature values for residential land use types within a block group was also calculated.
- <u>Stepwise regression analysis</u> (identify suitable variables for development of population estimation models).
- 1) ??? the relationship between average population density in each census unit and scale-invariant indicators derived from remote sensing imagery:
 - independent variable: spectral response based population estimation for six Landsat ETM+ spectral bands, fraction images, and land surface temperature.
 - dependent variable: Population density
- 2) ??? relationship between population counts in each census unit and a number of scale-dependent indicators (e.g., pixel counts) from remote sensing data for the corresponding census unit.
 - independent variable: the areas of each type of residential land
 - dependent variable: Population count
- Mean relative error (RE) all block groups were used to assess population estimation models.

Results

Table 2. Correlation between population density and the mean of Landsat ETM+ spectral bands and derived variables.

	B1	B2	В3	B4	B5	В7	GV	IMP	TEMP
PD	0.144	0.083	0.088	-0.253*	-0.281*	0.006	-0.178	0.458*	0.513*

Note: B1-B5 and B7, Landsat ETM+ bands 1-5 and 7; GV, vegetation fraction; IMP, impervious fraction; PD, population density; TEMP, land surface temperature.

Table 3. Best models based on spectral data.

Model	Potential variables	Explanatory variables	R^2	RE (%)
1	Landsat ETM+ bands	B1 mean; B5 mean	0.18	
2	Landsat ETM+ bands; fractions; land surface temperature	TEMP mean; B7 mean	0.48	237

Note: For model 1, PD = 4083.082 - 61.875(B5 mean) + 36.677(B1 mean); for model 2, PD = -112.582.438 + 388.179(TEMP mean) - 61.511(B7 mean).

Best equation

$$PD = -112582.438 + 388.179(TEMP) - 61.511(B7)$$

The block groups with extremely high population density (especially greater than 3000 people/km²) were highly underestimated, and those with extremely low density (less than 400 people/km²) were largely overestimated, leading to high relative error.

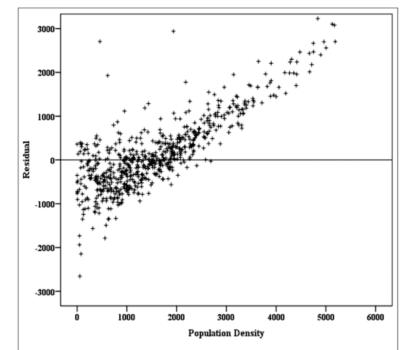


Figure 4. Residual distribution with population density (number of people per square kilometre) based on model 2.

^{*}Significant at the 0.01 level (two-tailed test).

Results - Analysis of population estimation with the land use based method

Table 4. Correlation between population counts and residential area and temperature.

	Low-density residential	Medium-density residential	High-density residential	TEMP sum
Population	0.392*	0.758*	0.824*	0.905*

^{*}Significant at the 0.01 level (two-tailed test).

the land use based method probetter estimation than the spectral response based n in terms of RE.

Model 3, which was based on residential land uses, provi best result forthis study.

Therefore, it was chosen to est population for the study are

Table 5. Models developed based on land use data.

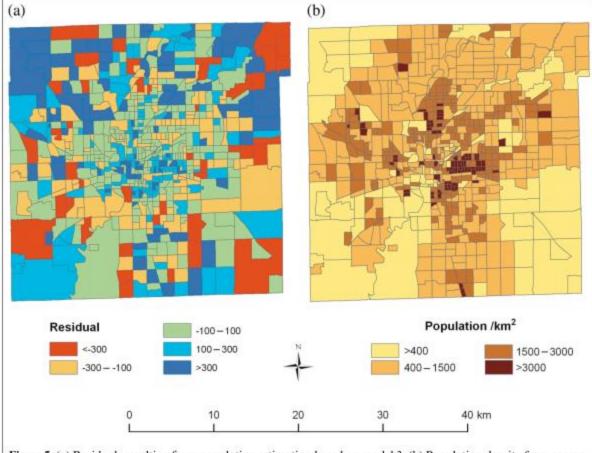


Figure 5. (a) Residuals resulting from population estimation based on model 3. (b) Population density from census.

Results - Analysis of population estimation with the land use based method

Model Explanatory variables R^2 RE (%)

3 Low-, medium-, and high-density residential 0.98 21.4

4 High-density residential; land surface temperature 0.98 21.4

Model 3 estimated population for the study area.

The estimated population based on model 3 for each block group was redistributed within residential classes using the dasymetric method.

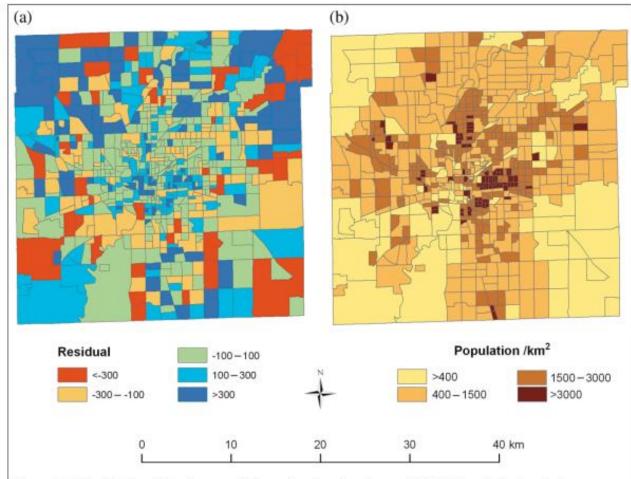


Figure 5. (a) Residuals resulting from population estimation based on model 3. (b) Population density from census.

Results - Analysis of population estimation with the land use based method

The dasymetric map shows the real population distribution pattern, on which LULC types such as agriculture, forest, and water had no people (Figure 6a),

the presentation of population distribution by a choropleth map (Figure 6b) did not conform to the population distribution patterns in the real world.

Model	Explanatory variables	R^2	RE (%)
3	Low-, medium-, and high-density residential	0.98	21.4
4	High-density residential; land surface temperature	0.98	21.4

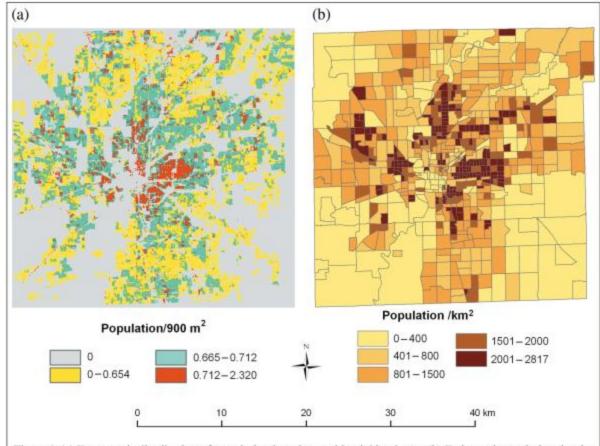


Figure 6. (a) Dasymetric distribution of population based on residential land uses. (b) Estimated population density based on model 3.

Remarks

- the direct use of mean values of spectral radiance could produce significant errors in population and population density.
 - variables may be highly correlated, thus multicollinearity may exist in multiple regression
- It appears that the residential impervious surface image fraction is an excellent population indicator because of its stability and underlying relationship with population
 - high-quality impervious surface data and the differentiation of nonresidential from residential impervious surfaces warrant further studies.
 - the same amounts of impervious surface may have significantly different population densities because of different patterns of residential use.
- → LULC data, especially those with high categorical resolution, have irreplaceable advantages for population estimation.

This procedure has been demonstrated to be capable of providing a good estimation result.

- They are more stable predictors compared with spectral responses and can be obtained from different remote sensing data sources.
- low-density residential areas scattered in forest and agricultural areas can have very similar spectral responses, making them difficult to separate
- One major uncertainty with population estimation comes from low-density residential areas with overestimation and from high-density residential areas with underestimation.
- This study has shown that the use of residential impervious surface, different densities of residential LULC classes, and land surface temperature can partially solve this problem.
- The incorporation of building height information seems to be another useful approach LIDAR