

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

Ilka Afonso Reis¹

¹ Programa de Pós-Graduação
Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12201-970 - São José dos Campos - SP, Brasil
ilka@dpi.inpe.br

Abstract. Small-area population counts were estimated for Brazilian census collection zones (SCs), using Landsat TM imagery. Two approaches were adopted: a SC-based model, which utilized spatial aggregations of the reflectance of pixels associated to each SC ; a pixel-based model, which utilized the reflectance of pixels classified as urban in iterated regression modeling. The performance of these approaches was evaluated at two level: the micro level (median relative error, calculated to the SC population estimates) and the macro level (total relative error, calculated to total urban population estimate). The first approach had a poor performance at both levels, especially at the micro level. The second approach has produced results comparable to those described in literature. At macro level, the pixel-based model results that can be considered very good (total relative error of -0.06% against -4.8% described in a recent work). Problems with overestimation at low densities and underestimation at high densities were found. Possible improvements in this model are discussed.

Palavras-chave: population estimation, reflectance, regression models, reflectância, modelos de regressão

1. Introdução

A informação sobre o número de habitantes de uma certa região é de grande importância para o desenvolvimento de estudos em várias áreas do conhecimento. No Brasil, o Instituto Brasileiro de Geografia e Estatística (IBGE) realiza um censo populacional a cada dez anos. Neste censo, vários tipos de dados sobre a população brasileira são coletados, além da simples contagem do número de habitantes. A cada cinco anos, no intervalo entre dois censos, o IBGE realiza uma contagem populacional, onde somente é registrado o número de habitantes por domicílio e algumas características demográficas, como sexo, idade e escolaridade.

Tanto o censo quanto a contagem populacional representam um grande esforço em termos humanos e financeiros. Este esforço deve ser feito, mas não pode ser repetido anualmente. Deste modo, a estimação do tamanho da população nos anos em que não há censo e nem contagem populacional é usualmente feita através de técnicas de projeção demográfica.

De algumas décadas até hoje, o desenvolvimento tecnológico e científico na área de Sensoriamento Remoto trouxe uma nova alternativa para a solução do problema de estimação de populações: o uso de imagens de satélite e fotografias aéreas.

As primeiras propostas de estimação de população através de Sensoriamento Remoto foram feitas utilizando-se fotografias aéreas e espaciais. O trabalho de Ogrosky (1975) é um dos exemplos destas tentativas. Neste trabalho, o autor utilizou fotografias aéreas com filme infravermelho na estimação da população de 18 cidades americanas. Variáveis como área urbana, número de estradas ligando cidades vizinhas, distância até a área urbana mais próxima e área urbana da cidade vizinha mais próxima foram medidas nas fotografias de cada cidade. Estas variáveis foram utilizadas como explicativas num modelo de regressão linear, no qual a variável dependente era o logaritmo da população da cidade.

Imagens de satélite também têm sido utilizadas na estimação de populações através de modelos de regressão. A variável a ser explicada (dependente) é a densidade populacional e as variáveis explicativas (independentes) são os valores médios de radiância espectral (ou de reflectância) associados aos pixels da imagem nas várias bandas de um sensor. Esta modelagem pode ser feita em diversos níveis de agregação, utilizando unidades (áreas) de

coleta de dados como setores censitários ou agrupamentos destes setores. Posteriormente, as densidades estimadas são multiplicadas pela área de sua respectiva unidade de coleta para produzir a estimativa de população naquela unidade. Outras variáveis também podem ser acrescentadas ao modelo, como, por exemplo, a porcentagem de pixels da unidade de coleta classificados como de uso residencial. Geralmente, o tamanho da unidade de coleta de dados é superior ao tamanho do pixel da imagem utilizada¹, ou seja, uma unidade de coleta contém vários pixels.

Utilizando imagens do satélite LANDSAT-MSS (Multi Spectral Scanner), Iisaka e Hegedus (1982) apresentaram um estudo no qual os valores médios de radiância espectral das quatro bandas foram utilizados para estimar a população de pequenas áreas numa região residencial de Tóquio, Japão. Lo (1995) utilizou imagens do satélite SPOT-HRV (resolução espacial de 20 m e três bandas espectrais) para estimar a população e o número de residências em uma parte da região metropolitana de Hong Kong, caracterizada por alta densidade populacional e uso misto do solo (prédios multifuncionais).

1.1. O trabalho de Harvey

Em seu primeiro trabalho de 2002, Harvey utiliza imagens do sensor TM (Thematic Mapper), a bordo do LANDSAT, para estimar a população dos CD's (Collect Districts) de duas regiões situadas na Austrália. Os valores médios das reflectâncias em seis bandas espectrais (bandas 1 a 5 e 7) foram utilizados como variáveis explicativas no modelo de regressão. Uma das regiões foi utilizada para estimar a equação de regressão e outra para validá-la (validação externa). Vários tipos de transformações e combinações das reflectâncias nas seis bandas do TM foram estudadas, tanto ao nível dos CD's (transformações/combinções de médias) quanto ao nível dos pixels de um CD (médias de transformações/combinções), gerando vários modelos de regressão. Transformações na variável dependente (densidade populacional do CD) também foram estudadas. A análise da acurácia das estimativas foi feita em separado para CD's urbanos e rurais, mas nenhum tipo de classificação dos pixels ou CD's foi incorporado aos modelos. Os melhores resultados foram obtidos no nível macro (população total), na área urbana, utilizando-se o modelo com transformações das reflectâncias ao nível do pixel (erro relativo de 0.6%). Os resultados na área rural foram pouco satisfatórios, onde populações de locais de baixa densidade populacional foram superestimadas em todos os casos. Isto mostra que o potencial desta metodologia é limitado pela heterogeneidade da cobertura do solo e densidade populacional dos CD's.

Partindo da observação de que as combinações das reflectâncias foram mais eficazes quando feitas ao nível dos pixels, o autor realizou um segundo estudo (Harvey, 2002b). Neste trabalho, Harvey utilizou os pixels da imagem como unidades de coleta de dados e comparou o desempenho dos modelos de regressão assim estimados com aqueles estimados ao nível dos CD's. Como não existem dados de população associados aos pixels, o autor utiliza uma versão do algoritmo EM (expectation-maximization, Dempster *et al.*, 1977) através de uma regressão iterada, uma idéia bastante interessante e muito bem ilustrada em seu artigo. Ao contrário de seu primeiro trabalho, Harvey classificou previamente os pixels em residenciais e não residenciais. Tanto modelos simples quanto modelos baseados em combinações das reflectâncias foram investigados. O modelo que obteve o melhor desempenho foi aquele baseado na combinação linear das reflectâncias nas seis bandas utilizadas. No momento da validação externa, algumas modificações foram feitas neste modelo, tendo sido obtidos resultados bastante satisfatórios no nível macro (erro relativo de 0.3% na área urbana) e

¹ Isto justifica o emprego de uma medida-resumo como a média na construção das variáveis explicativas do modelo de regressão (valores médios das radiâncias (ou reflectâncias) dos pixels).

razoáveis no nível micro, a população de cada setor (erro relativo mediano de 17.2%). Outra vez, os resultados na área rural não foram considerados satisfatórios. Comparando os dois tipos de modelos investigados (modelos complexos baseados em CD's e o modelo simples baseado em pixels), o autor observa que o desempenho do segundo tipo é melhor, no sentido de que produz estimativas mais acuradas. Isto ocorre mesmo no caso de CD's rurais com baixa densidade populacional, áreas que se revelaram um problema em outros estudos devido à grande superestimação de suas populações.

1.2. Objetivo

O objetivo deste trabalho é estudar as técnicas apresentadas por Harvey (2002a e 2002b) na estimação das populações dos setores censitários de Belo Horizonte (MG), no ano de 1996, com o auxílio da imagem do sensor Thematic Mapper, a bordo do satélite LANDSAT 5.

2. Materiais e métodos

Neste trabalho, foram necessários três tipos de dados: o cadastro urbano do município de Belo Horizonte com a divisão em setores censitários de 1996; a contagem populacional de 1996, por setor censitário; e as imagens das bandas 1 a 5 e 7 do LANDSAT-TM (órbita/ponto 218/74 de 31/01/1996). Estas imagens foram corrigidas atmosféricamente pelo modelo 6S (*Second Simulation of Satellite Signal in the Solar Spectrum*, Vermote et al., 1997).

Dos 2110 setores censitários de Belo Horizonte em 1996, 99.4% eram classificados como urbanos e somente este tipo de setor foi utilizado neste trabalho. Destes setores, 35 foram retirados: 13, por serem considerados especiais (asilos, orfanatos e presídios), e 22, por possuírem baixa densidade populacional (menos de 1 habitante por 10 m²), o que poderia prejudicar a análise (Harvey, 2002a e 2002b). Para os 2060 setores censitários restantes, estava disponível também a classificação do IBGE para o tipo de ocupação: normal e aglomerado sub-normal (favelas).

Os bancos de dados foram construídos com o *software* SPRING 4.0 (INPE, 2003) e analisados no programa R 1.8.1 (R Development Core Team, 2003). Para a construção do banco de dados de setores, a média da reflectância em cada banda foi calculada para todos os setores censitários. Na construção do banco de dados de pixels, foi necessário classificar a imagem para retirar os pixels não-urbanos. O tipo de classificação utilizada foi a supervisionada, por pixel, através do método da máxima verossimilhança (Mather, 1999). As bandas utilizadas na classificação foram a 3, 4 e 5. Ao final, restaram 228957 pixels classificados como urbanos e cada um deles foi associado a um dos setores censitários urbanos.

2.1. O modelo de regressão iterada

Os modelos de regressão utilizados na modelagem dos bancos de dados de setores e de pixels têm a mesma forma do modelo de regressão usual (equação 1)

$$p_i = \beta_0 + \sum_{j=1}^k \beta_j r_{ij} + \varepsilon_i \quad (1)$$

Na abordagem dos setores censitários (Harvey, 2002a), p_i representa a densidade populacional do setor censitário i , r_{ij} é a média da reflectância dos pixels do setor i na j -ésima banda do sensor, β_0 e β_j , $j = 1, 2, \dots, k$, são os parâmetros a serem estimados e ε_i representa a parte da densidade populacional dos setores que não é explicada pelo modelo de regressão.

Já na abordagem dos pixels (Harvey, 2002b), r_{ij} é a reflectância do pixel i na j -ésima banda do sensor e p_i representa a população do pixel. Assim, pode-se notar que a modelagem

dos dados ao nível dos pixels esbarra em uma dificuldade: as contagens populacionais disponíveis estão associadas aos setores censitários e não aos pixels.

Para contornar esta dificuldade, Harvey (2002b) utilizou o que ele denominou **regressão iterada**. Depois de classificar os pixels como residenciais e não-residenciais, uma estimativa inicial da população em cada pixel foi feita e, então, estas estimativas foram refinadas iterativamente. A estimativa inicial de população em cada pixel, p_i , é dada pela divisão da população do setor pelo número de pixels do setor. A seguir, a equação de regressão em (1) é estimada e os valores preditos (estimados) para p_i são ajustados de modo que o total da população do setor censitário depois do ajuste seja mantido como o total populacional conhecido do setor. A população ajustada do pixel i é dada pela soma da população estimada para o pixel i e da média dos resíduos do setor censitário ao qual pertence o pixel.

Na iteração seguinte, os p_i 's ajustados na iteração anterior substituem as estimativas iniciais de população na variável dependente e a equação de regressão é novamente estimada. Novos p_i 's ajustados são calculados e substituem os p_i 's atuais na iteração seguinte. As iterações continuam e o critério de parada pode ser definido em função de alguma medida de qualidade de ajuste como o coeficiente de determinação (R^2) ou o quadrado médio dos resíduos, por exemplo. As iterações terminariam quando uma destas medidas não tivesse mudanças consideradas relevantes de uma iteração para outra.

Segundo Harvey (2002b), a multicolinearidade entre as reflectâncias nas bandas do TM pode causar problemas de convergência deste processo de regressão iterada. Outro problema lógico é o das estimativas negativas para as populações associadas aos pixels, já que o modelo de regressão linear utilizado não possui restrições. Na tentativa de solucionar este problema, um procedimento alternativo foi utilizado: a cada iteração, as estimativas negativas eram transformadas em zero e ajustes nas estimativas dos outros pixels eram feitos de modo a manter constante o total populacional do CD.

3. Resultados e Discussão

Estimados os modelos de regressão, os resultados foram avaliados através de medidas de erro. O erro relativo para cada setor foi calculado pela diferença absoluta entre o valor observado da população e o valor estimado pelo modelo, dividida pelo valor observado. O Erro Relativo Mediano é a mediana dos valores absolutos dos erros relativos e representa o *erro no nível micro* (setores). O Erro Relativo Total foi calculado como a diferença entre o total populacional estimado (todos os setores) e o total populacional observado, dividida pelo total observado. O erro relativo total representa a variação do total estimado em relação ao total observado, sendo chamado *erro no nível macro* (área urbana).

3.1 Banco de dados de setores

Considerando a regressão ao nível dos setores, foram ajustados três modelos: completo (todas as bandas) e completo com transformações na variável dependente (raiz quadrada e logaritmo natural). No entanto, o desempenho destes modelos não foi considerado satisfatório (erros relativos muito altos). A análise de resíduos do modelo com a transformação logarítmica apresentou-se melhor do que a dos outros modelos. Levando em conta o fato de que a distribuição da densidade populacional dos setores é bastante assimétrica, este modelo foi refinado na tentativa de melhorar seu desempenho, resultando nos modelos com indicador de favela e porcentagem de área urbana.

A incorporação do indicador de favela e da porcentagem de área urbana melhorou um pouco o desempenho do modelo, mas o erro nos dois níveis (macro e micro) ainda não pode ser considerado satisfatório. No entanto, o modelo com indicador de favela e porcentagem de área urbana conseguiu resultados similares aos do modelo completo ajustado por Harvey

(2002a) no nível micro (erro relativo mediano de 27.81% contra os 28.5% de Harvey) e melhores no nível macro na área urbana (erro relativo total de 8.61% contra os 25.6% de Harvey).

3.2 Banco de dados de pixels

Como colocado em Harvey (2002b), os bancos de dados de pixels podem ser muito grandes, significando muito esforço computacional nas análises. Para solucionar este problema, Harvey sugere que se trabalhe com uma amostra dos pixels. No presente trabalho, além de solucionar a questão do esforço computacional, a amostragem permitiu que se fizesse a validação interna do modelo estimado com os pixels amostrados, já que não se dispunha de uma segunda imagem, como era o caso em Harvey (2002b). No processo de amostragem, 25% dos pixels de cada setor censitário era selecionado. Este valor de fração amostral é arbitrário. Harvey (2002b), por exemplo, utiliza 2% dos pixels de cada CD. A opção por uma fração amostral bem superior deve-se ao fato de que os setores censitários de Belo Horizonte são menores do que os CD's de Harvey. O número de pixels nos setores censitários urbanos varia entre 4 e 6091. Para muitos setores, uma fração tão pequena quanto aquela utilizada por Harvey implicaria em uma amostra de tamanho nulo. Depois de amostrados os pixels de cada setor, uma fração de 25% da população do setor era igualmente distribuída entre os pixels selecionados. Esta é a população inicial a ser redistribuída através da regressão iterada.

A correlação linear entre as reflectâncias das bandas 1, 2 e 3 é bastante grande (coeficiente de correlação linear de Pearson de 0.87 a 0.96), assim como a correlação entre as bandas 5 e 7 (0.82). A banda 4 é a que possui as menores correlações com as outras bandas (0.18 a 0.55). Estes relacionamentos entre as variáveis independentes podem causar problemas de multicolinearidade, afetando a estimação dos coeficientes do modelo (Draper e Smith, 1998). Deste modo, optou-se por ajustar os modelos sem a banda 2, mas com a banda 3, e o modelo sem as bandas 2 e 3. Além destes dois modelos, o modelo completo foi ajustado para possibilitar a comparação com os resultados obtidos por Harvey (2002b). Variações destes modelos também foram ajustadas, como o modelo sem o intercepto, o modelo com a transformação logaritmo natural da população (variável dependente) e o modelo sem os ajustes para populações negativas. Os resultados são apresentados no **Quadro 1**.

Os modelos que apresentam os melhores desempenhos no nível macro são os modelos (3.3) e (3.5), ajustados com as bandas 1, 4, 5 e 7, com intercepto, sem os ajustes para populações negativas na etapa da regressão iterada e, no caso do modelo (3.5), com indicador de aglomerados urbanos (favelas). A incorporação deste indicador ao modelo (3.3) não trouxe grande melhoria no desempenho das estimativas. Termos de interação entre as reflectâncias e o indicador de aglomerados foram introduzidos no modelo (3.5), sem sucesso. Por questão de simplicidade, o modelo escolhido foi o (3.3), que tem sua equação de regressão dada por

$$\hat{p} = 13.243 + 0.011 r_1 - 0.020 r_4 - 0.102 r_5 + 0.132 r_7$$

Os coeficientes estimados para as bandas 4, 5 e 7 pelo modelo (3.3) são bastante similares aos obtidos por Harvey (2002b) em seu modelo escolhido (0.014, 0.084 e 0.142, respectivamente). O coeficiente da banda 1 e o intercepto são bem diferentes (0.120 e 2.019, nesta ordem), provavelmente devido à presença das bandas 2 e 3 no modelo de Harvey.

O erro relativo total foi de -0.06%, bem inferior ao erro relativo total obtido por Harvey (2002b) na área urbana (-4.8%) com o modelo escolhido pelo autor. No entanto, no nível micro, o modelo escolhido por Harvey tem desempenho melhor (erro relativo mediano de 14.0% contra 31.57%).

Quadro 1 - Algumas medidas para avaliação dos modelos de regressão ajustados para o banco de dados de pixels.

	R² Iterada (amostra)*	R² “back” (total)**	Erro Relativo Mediano (%)	Erro Relativo Total (%)***
1.1) Completo	0.257	0.193	30.43	0.47
1.2) Completo, sem intercepto	-----	0.231	31.56	-0.67
1.3) Completo, sem ajustes	0.397	0.196	30.82	-0.09
1.4) Completo, log. natural	0.634	0.198	37.8	4.81
2.1) Sem banda 2	0.257	0.193	30.48	0.47
2.2) Sem banda 2, sem intercepto	-----	0.231	31.55	-0.69
2.2) Sem banda 2, sem ajustes	0.383	0.194	31.06	-0.09
2.3) Sem banda 2, log. natural	0.625	0.175	37.81	4.13
3.1) Sem bandas 2 e 3	0.241	0.186	31.36	1.66
3.2) Sem bandas 2 e 3, sem intercepto	-----	0.246	33.40	-0.44
3.3) Sem bandas 2 e 3, sem ajustes	0.393	0.190	31.57	-0.06
3.4) Sem bandas 2 e 3, log. natural	0.613	0.206	37.78	4.10
3.5) Modelo (3.3), indicador de favela	0.481	0.191	30.25	-0.06

* Valor correspondente à última iteração .

** Valores correspondentes ao quadrado do coeficiente de correlação entre as estimativas de população para os setores e os valores populacionais reais destes setores.

*** Valores positivos indicam superestimação e valores negativos indicam subestimação

Harvey (2002b) aponta problemas de superestimação em setores pouco densos e subestimação em setores muito densos. A **Figura 1** mostra o relacionamento entre a densidade populacional dos setores e o erro relativo na estimação de sua população (as escalas de ambos os eixos foram transformadas para melhorar a visualização). Considerando como erro tolerável de sub ou superestimação um valor menor do que 10%, somente 15% dos setores estiveram dentro destes limites (pontos entre as linhas horizontais pontilhadas). A linha vertical pontilhada representa a densidade de 1 hab/m². De maneira geral, pode-se notar que setores menos densos tendem a ter sua população superestimada. Por outro lado, em setores mais densos, tende a ocorrer a subestimação da população.

4. Conclusões

A estimação de populações utilizando imagens de satélite abre uma nova área dentro do Sensoriamento Remoto. Segundo Harvey (2002a), existem dois importantes aspectos que diferenciam e tornam a estimação de populações uma tarefa mais complexa do que as outras aplicações dos dados vindos de satélites. Primeiramente, estes dados estão relacionados às características da superfície da Terra e sua cobertura, embora, em muitas aplicações, este relacionamento seja indireto. Porém, no caso de populações, este relacionamento é ainda mais tênue. Por exemplo, as reflectâncias (ou radiâncias) captam pouco da informação existente sobre estruturas urbanas, como altura e tipo, o que causa problemas na estimação da

população em áreas de alta densidade populacional. O segundo aspecto diz respeito ao fato de que estimar uma variável quantitativa como tamanho de população através da dimensão espacial de uma imagem é mais ambicioso do que as aplicações mais usuais com objetivos qualitativos como classificação ou segmentação de áreas. Ainda assim, tentativas têm sido feitas nesta direção.

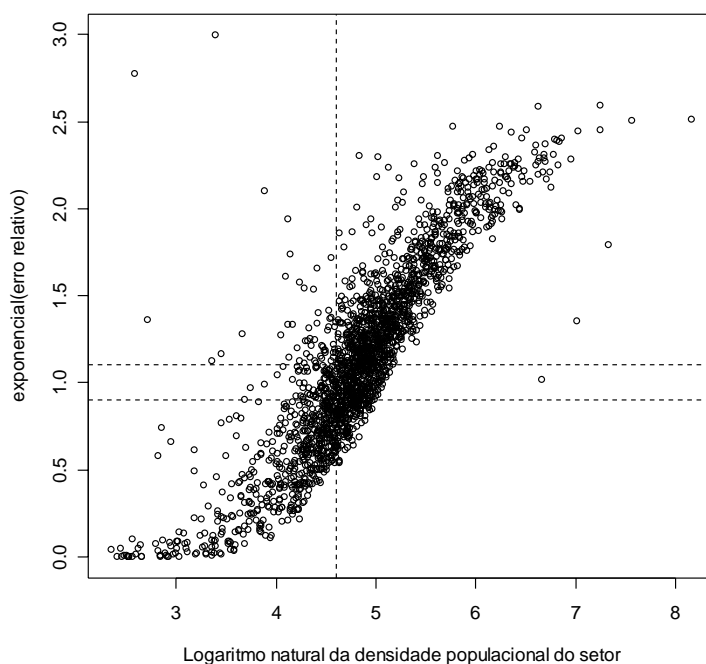


Figura 1 – Relacionamento entre a densidade populacional dos setores censitrios e o erro relativo na estimaco da populao (modelo 3.3)

No presente trabalho, o ajuste do modelo escolhido por Harvey (2002b) aos pixels classificados como urbanos (modelo completo) teve um desempenho pior no nvel micro (erro relativo mediano de 30.4% contra 14.0% de Harvey), mas bem superior no nvel macro (erro relativo total de 0.47% contra -4.8% de Harvey).

Em comparaco com o modelo completo, o modelo ajustado sem as bandas 2 e 3 e sem os ajustes para estimativas negativas na etapa da regresso iterada teve um desempenho semelhante no nvel micro (erro relativo mediano de 31.6%) e bem superior no nvel macro (erro relativo total de -0.06%).

Assim como descrito em Harvey (2002b), os problemas de superestimaco e subestimaco da populao dos setores (nvel micro) tendem se concentrar nos setores menos e mais densos, respectivamente. O modelo baseado nos pixels no leva em conta a densidade populacional, uma vez que os todos pixels tm a mesma rea (900 m², no caso do LANDSAT-TM). A incorporaco de uma varivel auxiliar como indicao de densidade populacional permitiria que os modelos fossem ajustados separadamente para setores muito densos e pouco densos. Este procedimento poderia melhorar as estimativas. Neste trabalho, no entanto, o indicador de aglomerados (favela) no se mostrou muito eficiente nesta melhoria. O cadastro de domiclios do municpio tambm poderia servir com um indicador de densidade populacional, desde que esteja atualizado, o que  particularmente difcil nas reas de favela e ocupaes recentes.

Partindo da observao de que a estimaco da populao de agrupamento de setores (nvel macro) mostrou-se bastante acurada, uma alternativa para melhorar o desempenho no

nível micro seria usar a população dos setores nos modelos, porém calcular as estimativas de população não para os setores, mas para um agrupamento deles. No caso de Belo Horizonte, este agrupamento poderia ser a Unidade de Planejamento (UP), que é, na verdade, a unidade utilizada para planejamento de políticas públicas na cidade.

Mesmo que esta abordagem alternativa melhore em muito a estimação no nível micro, seria ambicioso demais pensar que um mesmo modelo de estimação de populações, desenvolvido a partir de imagens produzidas por determinado sensor, pudesse ser aplicado a qualquer região. O caminho mais sensato seria desenvolver um modelo para cada região de interesse, incorporando a ele os dados disponíveis sobre a região, sem esquecer a necessidade de calibrá-lo sempre que os dados de contagem populacional fossem atualizados.

Embora haja opiniões um tanto quanto radicais sobre a estimação de populações via Sensoriamento Remoto (Brugioni, 1983), em seu estado da arte atual, ela seria uma tecnologia auxiliar na obtenção de informações sobre o número de habitantes de uma região, não substituindo a realização das contagens populacionais nem tampouco dos censos demográficos.

5. Agradecimentos

À Dra. Silvana Amaral Kampel, pesquisadora da Divisão de Processamento de Imagens do INPE, pela preciosa ajuda na parte inicial do trabalho, especialmente com a correção atmosférica e iniciação ao SPRING.

6. Referências

- Brugioni, D.A. The census: it can be done more accurately with space-age technology, **Photogrammetric Engineering and Remote Sensing**, vol. 49, p. 1337-1339, 1983.
- Dempster, A. N.; Laird, N. M.; Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of Royal Statistics Society**, Series B, vol. 39, p. 1-38, 1977.
- Draper, N. R. ; Smith, H. **Applied Regression Analysis**, 3a. edição. John Wiley and Sons, EUA, 706 p, 1998.
- Harvey, J. T. Estimating census district populations from satellite imagery: some approaches and limitations. **International Journal of Remote Sensing**, vol. 23, n. 10, p. 2071-2095, 2002a.
- _____. Population estimation models based on Individuals TM Pixels. **Photogrammetric Engineering and Remote Sensing**, vol. 68, n. 11, p. 1181-1192, 2002b.
- Iisaka J.; Hegedus, E. Population estimation from Landsat Imagery. **Remote Sensing of Environment**, vol. 12, p. 259- 272, 1982.
- INPE – Divisão de Processamento de Imagens. **Sistema de Processamento de Informações Georreferenciadas (SPRING)**. 2003. São José dos Campos - SP, URL <http://www.dpi.inpe.br/spring/>
- Langford, M. ; Maguire, D. J. ; Unwin, D. J. The aerial interpolation problem: estimating population using remote sensing within a GIS framework. In: **Handling Geographical Information: Methodology and Potencial Applications**, edited by I. Masser and M. Blakemore. London: Longman, 1991, p. 55-77.
- Lo, C. P. Automated population and dwelling unit estimation from high-resolution satellite images: a GIS approach. **International Journal of Remote Sensing**, vol. 16, n.1, p. 2071-2095, 1995.
- Mather, P.M. **Computer Processing of Remotely-Sensed Images: An Introduction**. John Wiley & Sons, 1999
- Ogrosky, C. E. Population estimates from satellite imagery. **Photogrammetric Engineering and Remote Sensing**, vol. 41, p. 707-712, 1975.
- R Development Core Team. R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, 2003, Vienna, Austria. ISBN 3-900051-00-3, URL <http://www.R-project.org>.
- Vermote, E.F. ; Tanre, D. ; Deuzé, J.L. ; Herman, M., and Morcrette, J.J. Second simulation of the satellite signal in the solar spectrum, 6S: An overview. **IEEE Trans. Geosc. Remote Sens**, vol. 35, n. 3, p. 675-686, 1997.