

Comparação entre estimadores por expansão direta e por regressão no cálculo da área de milho e soja

Marcos Adami¹
Mauricio Alves Moreira²
Bernardo Friedrich Theodor Rudorff²
Corina da Costa Freitas²
Rogério Teixeira de Faria³

¹ Secretaria de Estado da Agricultura e do Abastecimento do Paraná,
Av. Minas Gerais, 1351 - 86300-000 - Cornélio Procópio – PR, Brasil
adami@seab.pr.gov.br

² Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12201-970 - São José dos Campos - SP, Brasil
{mauricio, bernardo}@ltid.inpe.br; corina@dpi.inpe.br

³ Instituto Agrônômico do Paraná - IAPAR
Caixa Postal 481- 86047-902 - Londrina - PR, Brasil
rtfaria@iapar.br

Abstract. The objective of this work was to compare the direct expansion method with the regression method to estimate soybean and corn crop areas in the municipality of Cornélio Procópio, Paraná State, Brazil. Thirty five regular segments of 1 km x 1 km were randomly selected from a sample panel stratified according to land use and land cover based on Landsat image classification procedure. The segments were used for data collection during fieldwork. The regression method produced the lowest coefficient of variation (CV) values for both soybean and corn crop area estimation, although the direct expansion method performed quite similar, provided fast estimates and is independent of cloud free images. The regression method was also more efficient than the direct expansion method but considering that only 23.5% of the study area was cloud free the efficiency of the regression method could have been much better if a cloud free image happened to be available.

Palavras-chave: regular segments, crop area estimation, agricultural statistics, segmentos regulares, estimativa de área de culturas, estatísticas agrícolas.

1. Introdução

É incontestável a importância econômica das culturas de milho e soja, para o estado do Paraná. Ambas representam em torno de 40% do valor bruto da produção agrícola do Estado (Paraná.SEAB/DERAL, 2004). Neste sentido, o conhecimento da área cultivada com estas culturas é de fundamental importância para o estabelecimento de políticas agrícolas.

Há várias maneiras de determinar a área cultivada, porém três abordagens são as mais empregadas: a) os censos que implicam em um gasto maior de tempo e custo; b) as amostragens subjetivas que são ágeis, mas não têm associado o erro; e c) as amostragens probabilísticas que têm embasamento científico e erro associado, podendo ser mais confiáveis.

O emprego da teoria da amostragem por área teve início na década de 60 do século passado e esteve amparada no uso de fotografias aéreas e materiais cartográficos (FAO, 1998). A partir do lançamento do primeiro satélite de sensoriamento remoto da série Landsat em 1972, essa teoria foi aprimorada, incorporando o uso de imagens de satélites. Os projetos *Crop Identification Technology Assessment by Remote Sensing* (CITARS), *Large Area Crop Inventory Experiment* (LACIE) e *Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing* (AGRISTARS) forneceram uma base sólida para os modelos

estatísticos de estimativa de área cultivada por amostragem (Chen, 1980). Nestes modelos, são empregados dois métodos: a) por expansão direta; e b) por regressão que visa reduzir a variância das estimativas para conferir maior confiabilidade à estimativa.

Segundo Ipoliti-Ramilo (1999) a probabilidade de se obter uma imagem livre de nuvens durante o período da safra de verão na região Centro-Sul é inferior a 20%. Isso que indica que para se realizar a estimativa de área por regressão, para culturas de verão, será necessário utilizar imagens com presença parcial de nuvens.

Neste sentido o presente trabalho tem por objetivo estimar a área cultivada com milho e soja por expansão direta e compará-la com a estimativa por regressão para a qual foi utilizada uma imagem parcialmente coberta com nuvens no município de Cornélio Procópio, PR.

2. Material e Método

A área teste selecionada para analisar a eficiência do estimador foi o município de Cornélio Procópio, Paraná, com uma área de 650 km² (**Figura 1**). Imagens dos sensores TM e ETM⁺ dos satélites Landsat-5 e 7, respectivamente, adquiridas nos anos 2001, 2002 e 2003 foram interpretadas para dividir o município em três estratos de uso do solo: 1) Estrato A – região agrícola com 50 a 80 % de área cultivada; 2) Estrato B – agricultura extensiva com 15% a 49% de área cultivada e com predominância de pastagens e 3) Estrato C – áreas não agrícolas que compreendem cidades, parques, reservas florestais, instalações militares, montanhas etc. (Adami, 2004).

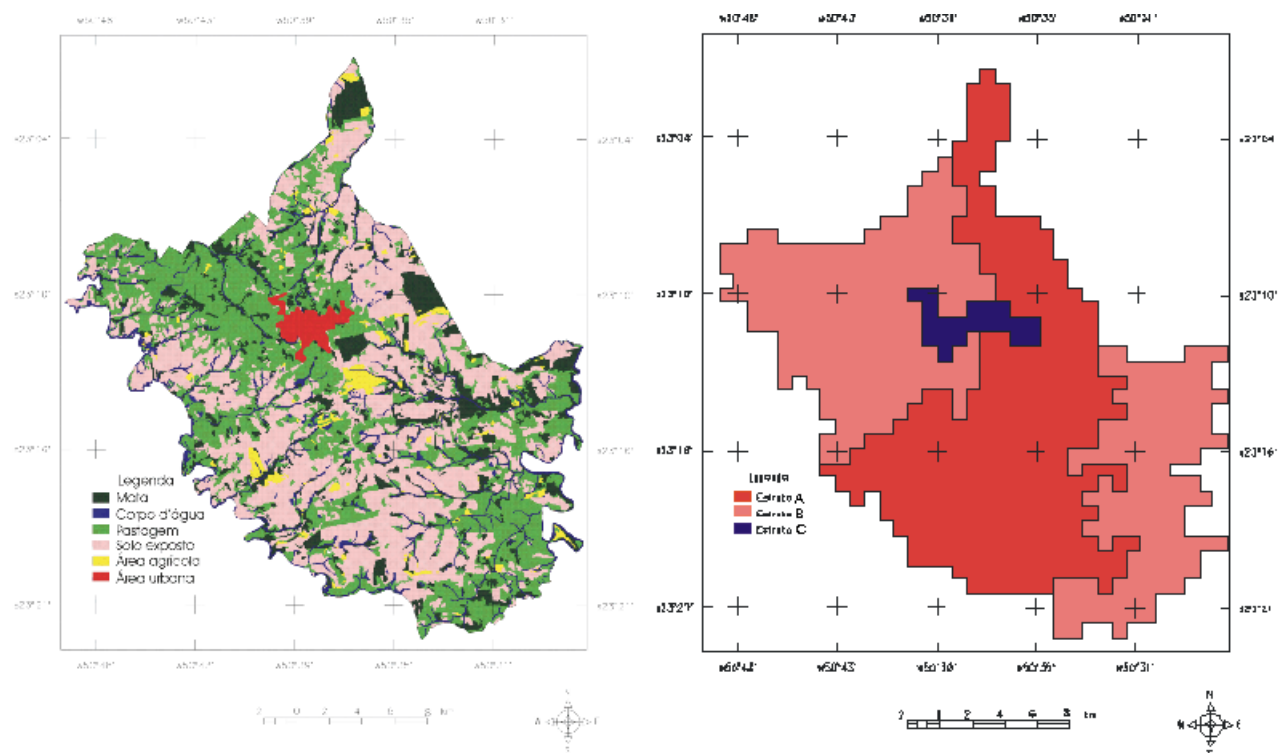


Figura 1 – Área de estudo e estratos de uso do solo.
Fonte: Adami (2004).

Outro fato que deve ser levado em consideração é a forma, o tamanho e a alocação dos segmentos da amostra na área de estudo. Gonzáles-Alonso et al. (1991) conduziram um experimento em Navarra, Espanha, cujo objetivo foi comparar a área estimada por expansão direta e por regressão com segmentos regulares (0,7 x 0,7 km) e irregulares (polígonos cadastrais do censo). O esquema de amostragem adotado foi o de amostra sistemática

estratificada. Os autores observaram que a precisão dos estimadores foi similar, independente da forma dos segmentos. No entanto, a eficiência relativa da regressão em relação à expansão direta variou de 1,66 a 13,05 nos segmentos regulares e de 1,44 a 26,29 nos segmentos irregulares. Assim, os autores concluíram que é mais indicado utilizar segmentos regulares para o estimador por regressão. Desta forma, para o presente trabalho os estratos A e B foram divididos em segmentos quadrados de 1 x 1 km (100 ha) de área (González-Alonso et al., 1991). Cada unidade foi codificada por um número de 1 a n_i ($i = A, B$) para o sorteio das amostras por estrato, empregando-se o critério de amostragem aleatória estratificada. Não foram selecionados segmentos no estrato C, pois, por definição, o mesmo não continha áreas agrícolas. As unidades amostrais ou segmentos selecionados foram identificados nas imagens na composição colorida das bandas 3 (vermelho), 4 (infravermelho próximo) e 5 (infravermelho médio). Para cada segmento selecionado foi gerado um produto analógico na escala de 1:25.000 chamado de módulo. Sobre estes módulos foram traçados os limites dos segmentos e indicadas as coordenadas dos vértices a fim de facilitar a localização no campo (Adami, 2004). Um papel transparente foi sobreposto aos segmentos para que a informação coletada a campo pudesse ser devidamente anotada sobre as imagens.

Devido ao desconhecimento do comportamento estatístico da variável pesquisada, o percentual médio amostrado foi de aproximadamente 5% (Rudorff e Krug, 1986). O número de segmentos sorteados por estrato pode ser observado na **Tabela 1** (Adami, 2004).

Tabela 1 – Número total de segmentos sorteados e percentuais da área amostrada

Estrato	Total de Segmentos a	Segmentos Sorteados b	Área Amostrada (%) (b/a)*100
A	304	25	8,22
B	336	10	2,97
Total	640	35	5,47

Fonte: Adami, 2004.

Para a orientação a campo, quanto à localização dos segmentos, foi utilizado um GPS de navegação XL12 e o software TRACKMAKER, que faz a interface entre o GPS e o microcomputador.

Para estimar a área de milho e soja por meio do estimador de regressão foi necessário classificar a imagem do sensor ETM⁺ do Landsat-7 referente à órbita 222 ponto 76 adquirida no dia 07/02/2003. Foi realizada uma classificação não-supervisionada por meio do classificador ISOSEG no aplicativo SPRING. Posteriormente foi feita uma interpretação visual por meio do módulo de edição matricial do SPRING para adequar os limites das classes. A estimativa das áreas de milho e soja foram feitas utilizando-se os estimadores por regressão e o por expansão direta. Para a área do município coberta com nuvens foi empregada apenas a expansão direta enquanto que para a área livre de nuvens foi utilizado o estimador por regressão (Gallego, 1995; FAO, 1998; Adami, 2004).

3. Resultados e discussão

Devido à alta incidência de nuvens durante o período de desenvolvimento das culturas de verão (milho e soja), a única imagem disponível com cobertura parcial de nuvens foi a do dia 07/02/2003. Desta forma, o modelo de regressão foi testado apenas em parte do estrato A, denominado A'. O resultado desta classificação foi utilizado para estimar a área por regressão para as culturas de milho e soja e pode ser observado tanto na **Tabela 2** quanto na **Figura 2**.

Tabela 2 - Resultado do mapeamento temático para o estrato A' (região sem nuvem)

Tema	Área (km ²) a	(a/total)*100 %
Corpos d' água	2,05	1,3
Café	3,55	2,3
Mata	20,31	13,1
Pastagem	26,92	17,3
Agricultura	20,21	13,0
Milho	11,42	7,4
Soja	70,75	45,6
Nuvem	496,76	-
Área urbana	-	-
Total sem nuvem	155,21	100,0

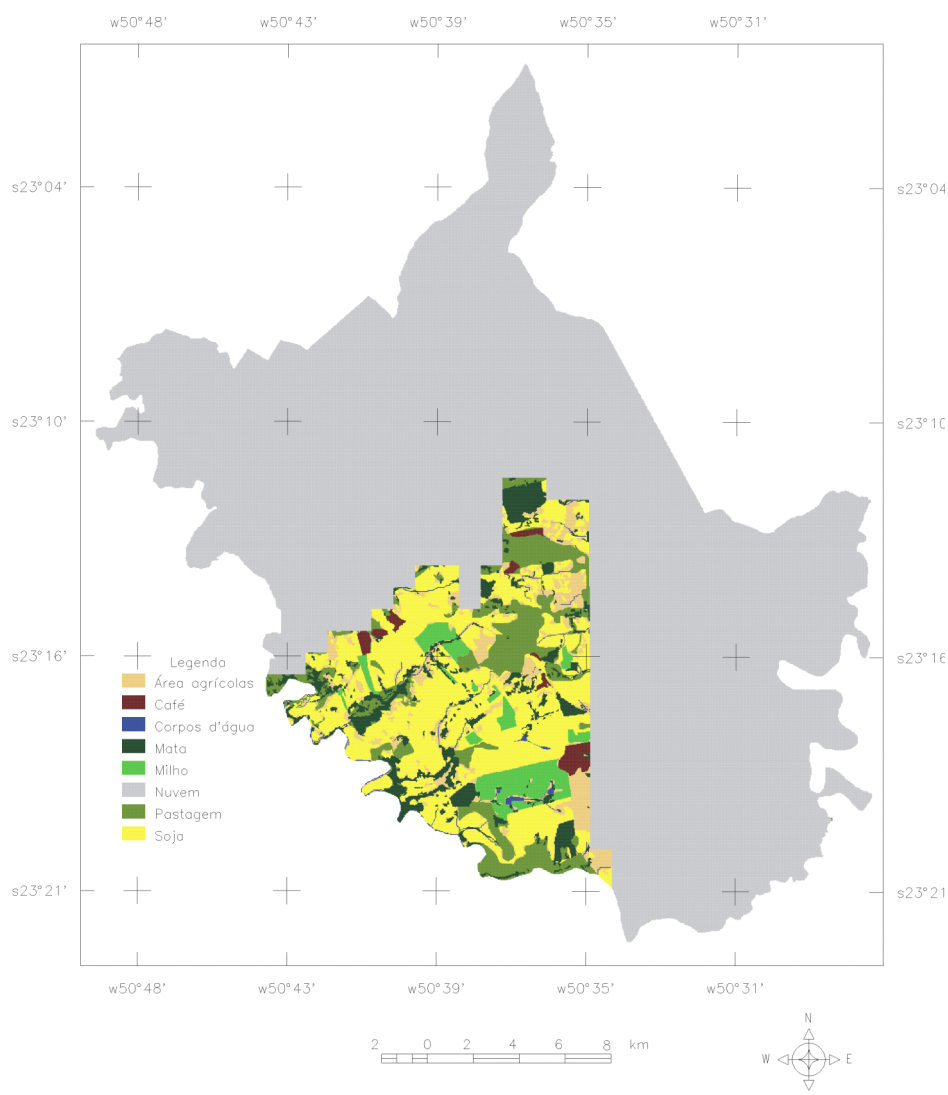


Figura 2 – Resultado do mapeamento temático para a área livre de nuvens.

Para estimar a área de milho e soja pelo o modelo de regressão linear empregaram-se os dados referentes à classificação temática e os resultados da pesquisa de campo. Para a estimativa da área de milho e soja no restante do município, ou seja, parte do estrato A e o estrato B onde ocorreu incidência de nuvem, usou-se a expansão direta, em que foram empregados os dados coletados nos segmentos por Adami (2004).

Os resultados da estimativa de área de soja e milho, obtidos pela combinação dos dois modelos (expansão direta e regressão) estão na **Tabela 3**. Nesta tabela pode-se observar que há grande variabilidade entre os dados estimados pelos dois modelos para a cultura do milho e também em relação à estimativa subjetiva do PARANÁ.SEAB/DERAL (2003). Contudo, devido à grande variabilidade de ocorrência de áreas plantadas com milho, nos segmentos da amostra, os resultados não diferem entre si (IC > 95%).

Tabela 3 - Estimativas de área das culturas de milho e soja obtidas por regressão, por expansão direta e pelo DERAL (subjetivo) além da variância e do coeficiente de variação (CV) para a regressão e expansão direta e a eficiência relativa (ER) da regressão em relação à expansão direta.

Estimadores	Cultura do milho			
	Área (km ²)	Variância (km ⁴)	CV (%)	ER
Regressão	32,68	94,62	29,76	2,27
Expansão Direta	45,12	215,47	32,54	-
Subjetivo (DERAL)*	40,00	-	-	-
	Cultura da soja			
Regressão	271,44	784,03	10,31	1,09
Expansão Direta	252,42	860,31	11,62	-
Subjetivo (DERAL)*	255,00	-	-	-

* Fonte: PARANÁ.SEAB/DERAL (2003)

Com relação à cultura de soja, pode-se observar na Tabela 3 que quando comparados os dados pontuais, a expansão direta teve maior aproximação com a estimativa subjetiva (PARANÁ.SEAB/DERAL, 2003). A área estimada pela regressão para a cultura da soja foi de 271,44 km² e não difere significativamente da expansão direta (95% de confiança).

Observa-se também que houve uma redução percentual do CV obtido pela regressão. A eficiência relativa da regressão quando comparada com a expansão direta, para a estimativa da área de milho foi de 2,27. Esse resultado está situado no limite inferior dos resultados obtidos por Gonzáles-Alonso et al. (1991), Gonzáles-Alonso e Cuevas (1993) e Gonzáles-Alonso et al. (1997).

No caso da cultura da soja, observa-se que a precisão das estimativas de área por meio dos dois modelos foi muito parecida. No caso, o CV obtido pela regressão foi de 10,31% e para a expansão direta foi de 11,62%. A eficiência relativa do estimador de regressão para a cultura da soja, quando comparado à expansão direta, foi de 1,09 vezes. Esta eficiência relativa (1,09) também foi menor que as descritas nos trabalhos de Gonzáles-Alonso et al. (1991), Gonzáles-Alonso e Cuevas (1993) e Gonzáles-Alonso et al. (1997).

4. Conclusões

Tendo em vista os objetivos propostos e os resultados encontrados, apresentam-se as seguintes conclusões: 1) os resultados da expansão direta, apesar de apresentarem um maior CV foram obtidos de modo mais rápido e não dependem da disponibilidade de imagens durante o período de desenvolvimento da cultura; 2) existe a possibilidade de se aplicar o estimador por regressão em áreas parcialmente cobertas com nuvens; 3) mesmo com a

cobertura de nuvem, o estimador por regressão teve uma eficiência relativa 2,27 vezes maior do que a expansão direta para a cultura do milho.

5. Referências

Adami, M. **Estimativa de áreas agrícolas por meio de técnica de sensoriamento remoto, geoprocessamento e amostragem**. 2004. 185p. (INPE-10235-TDI/900). Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2004.

Chen, S. C. **Precisão de classificação do mapa de distribuição de culturas, obtido a partir de imagens Landsat**. São José dos Campos: INPE, 1980. 7 p. (INPE-1889-RPE/232)

Food and Agriculture Organization of the United Nations (FAO). **Multiple frame agricultural surveys: agricultural survey programmes based on area frame or dual frame (area and list) sample designs**. Roma, 1998. v. 2.

Gallego, F. J. **Sampling frames of square segments**. Luxembourg: Joint Research Centre, 1995. 72p.

González-Alonso, F.; Cuevas, J. M. Remote sensing and agricultural statistics: crop area estimation through regression estimators and confusion matrices. **International Journal of Remote Sensing**, v. 14, n. 6, p. 1215-19, 1993.

González-Alonso, F.; Cuevas, J. M.; Arbiol, R.; Baulies, X. Remote sensing and agricultural statistics: crop area estimation in north-eastern Spain through diachronic Landsat TM and ground sample data. **International Journal of Remote Sensing**, v. 18, n. 2, p. 467-70, 1997.

González-Alonso, F.; Soria, S. L.; Gozalo, J. M. C. Comparing two methodologies for crop area estimation in Spain using Landsat TM images and ground-gathered data. **Remote Sensing of Environment**, v.35, n.1, p 29-35, Jan. 1991.

Ipoliti-Ramilo G.A. **Imagens TM/Landsat-5 da época de pré-plantio para a previsão da área de culturas de verão**. 1999. 183p. (INPE-7116-TDI/668). Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 1999.

Paraná. Secretaria de Estado da Agricultura e do Abastecimento – Departamento de Economia Rural (SEAB/DERAL). **Valor Bruto da Produção Agropecuária**. Disponível em: <<http://www.pr.gov.br/seab>> . Out. 2004.

Paraná. Secretaria de Estado da Agricultura e do Abastecimento – Departamento de Economia Rural (SEAB/DERAL). Dados da Região de Cornélio Procópio. (ccpseab@pr.gov.br, Fev. 2003). Comunicação Pessoal.

Rudorff, B. F. T.; Krug, T. Sensoriamento remoto na estimativa da área plantada com feijão, milho e mamona no município de Irece-BA. In: Simposio Latino-Americano de Sensoriamento Remoto, 6., Gramado. **Anais**. São José dos Campos: INPE, 1986. v. 1, p. 380-4.