

USO DO ESTIMADOR DE REGRESSÃO PARA AUFERIR MAIOR CONFIABILIDADE À ESTIMATIVA DA ÁREA CULTIVADA COM CAFÉ NO MUNICÍPIO DE CORNÉLIO PROCÓPIO-PR.

Marcos Adami¹
Mauricio Alves Moreira²
Bernardo Friedrich Theodor Rudorff²

¹ Secretaria de Estado da Agricultura e do Abastecimento do Paraná – Departamento de Economia Rural – SEAB/DERAL.
Av. Minas Gerais, 1351
CEP 86.300-000 – Cornélio Procópio – Pr.
adami@seab.pr.gov.br

² Instituto Nacional de Pesquisas Espaciais – Divisão de Sensoriamento Remoto – INPE/DSR
Avenida dos Astronautas, 1.758 - Jd. Granja
CEP 12227-010 – São José dos Campos – SP.
{mauricio, bernardo @ltid.inpe.br}

RESUMO: Informação precisa sobre a estimativa da área plantada com café é importante para o estabelecimento de políticas adequadas em vários setores do agronegócio do café. Uma das alternativas viáveis para obtenção desta estimativa é a amostragem probabilística. O objetivo deste trabalho é avaliar a eficiência dos métodos de expansão direta e de regressão para estimativa de área de café no município de Cornélio Procópio, Paraná. Para a coleta de dados a campo foram selecionados aleatoriamente 35 segmentos de 1 km x 1 km, de um painel amostral estratificado segundo o uso e ocupação do solo. O método de regressão foi oito vezes mais eficiente do que o método de expansão direta e também teve coeficiente de variação bem menor (2,7 vezes), muito embora ambos os métodos produziram estimativas muito similares à estimativa subjetiva.

Palavras-chave: sensoriamento remoto, estimativa de área, amostragem, estatísticas agrícolas.

ABSTRACT: Precise information about coffee crop planted area is important to establish appropriate policies in several sector of the coffee agribusiness. One of the viable alternatives to obtain this estimate is the probabilistic sampling. The objective of this work is to evaluate the efficiency of the direct expansion and regression methods to estimate coffee crop areas in the municipality of Cornélio Procópio, Paraná State, Brazil. Thirty five segments of 1 km x 1 km were randomly selected from a sample panel stratified according to land use and cover. These segments were used for data collection during field work. The regression method was eight times more efficient than the direct expansion method and also had a much lower (2.7 times) coefficient of variation, although both methods produced estimates that were quite similar to the subjective estimate.

Key-words: remote sensing, area estimation, sampling area, agricultural statistics

1 - INTRODUÇÃO

O conhecimento preciso sobre a estimativa da área plantada com café é importante para o estabelecimento de políticas adequadas em vários setores da cafeicultura. Uma das alternativas viáveis para obtenção desta estimativa é a teoria da amostragem por área que teve início na década de 60 do século passado e está amparada em 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 na metodologia o uso de imagens de satélites como fonte de dados para pesquisa e previsão de safras nos 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). Estes projetos forneceram uma base sólida para os modelos estatísticos de estimativa de área cultivada por amostragem (Chen, 1980).

Auxiliado pela teoria da amostragem e pela metodologia para estimativa de área por amostragem utilizando imagens de sensoriamento remoto, técnicas de geoprocessamento e coleta de dados a campo, a presente pesquisa tem como objetivo verificar a eficiência do estimador de regressão para estimativa da área cultivada com café em relação ao método de expansão direta no município de Cornélio Procópio, PR.

2 - MATERIAL E MÉTODO

2.1 - Estratificação da Área

A área de estudo compreende o município de Cornélio Procópio localizado ao norte do estado do Paraná, com uma área de 650 km² (Figura 1). O município foi inicialmente dividido em três estratos de uso do solo utilizando-se imagens dos satélites Landsat-5 e 7 dos sensores TM e ETM⁺, respectivamente, adquiridas nos anos 2001, 2002 e 2003 (Adami, 2003), conforme mostrado a seguir:

Estrato A – região agrícola com 50 a 80 % de área cultivada;

Estrato B – agricultura extensiva com 15% a 49% de área cultivada e com predominância de pastagens;

Estrato C – áreas não agrícolas (cidades, parques, reservas florestais, instalações militares, montanhas etc.).

Além da estratificação, outros aspectos a serem considerados são: a forma, o tamanho e a alocação dos segmentos da amostra. Gonzáles-Alonso et al. (1991) conduziram um experimento em Navarra na Espanha, cujo objetivo foi comparar a área estimada por expansão direta e pela regressão. A coleta de dados em campo foi realizada em segmentos regulares (quadrados de 700 m de lado) e irregulares (polígonos cadastrais utilizados para o 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, os estratos A e B foram divididos em segmentos quadrados de 1 km x 1 km (100 ha) de área (Gonzáles-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 por ele não conter 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, 2003). Um papel transparente foi sobreposto aos segmentos para que a informação coletada a campo pudesse ser devidamente anotada sobre as imagens.

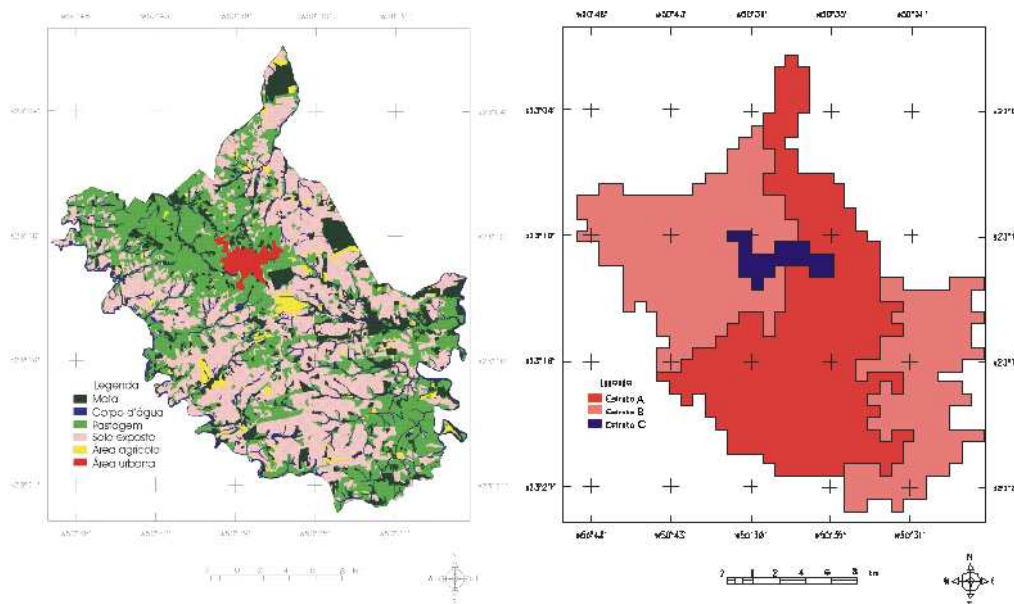


Figura 1 – Área de estudo e estratos de uso do solo.

Fonte: Adami, 2003.

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, 2003).

Tabela 1 – Número total de segmentos sorteados e área amostrada (%).

Estrato	Nº total de segmentos (a)	Nº de 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, 2003.

Para a localização dos segmentos no campo 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 café, por meio do estimador de regressão, foi necessário classificar a imagem do sensor ETM⁺ do Landsat-7 referentes à órbita 222 ponto 76 adquirida no dia 03/11/2002. A classificação foi realizada através do sistema SPRING, utilizando-se o

classificador ISOSEG com posterior edição matricial para adequar os limites das classes, exceto para as áreas ocupadas com café que foram obtidas por interpretação visual.

Para o cálculo das estimativas foram utilizados os estimadores por expansão direta e por regressão, assim como os estimadores de suas respectivas variâncias, descritos por Hansen et al. (1953), Cochran (1977), Moreira (1983), Krug e Yanasse (1986) Mueller et al. (1988), Allen (1990), Pradhan (2001), Day, 2002; e Adami (2003).

A avaliação do estimador por regressão em relação ao estimador por expansão direta, também conhecida como eficiência relativa (ER) é descrita por Krug e Yanasse (1986), Gonzáles-Alonso et al. (1991) e Adami (2003).

3 - RESULTADOS E DISCUSSÃO

Pode-se observar na Figura 2 o mapa temático gerado para obtenção da área de café no município de Cornélio Procópio. Foi verificada uma grande variabilidade espectral para as lavouras de café. Houve casos em que as lavouras foram plantadas no sistema tradicional e se apresentavam com baixo vigor vegetativo de tal forma que a resposta espectral da lavoura de café foi muito próxima da resposta espectral do solo. Já as plantadas no sistema adensado apresentaram alto vigor vegetativo com forte resposta espectral na banda 4 (infravermelho próximo) o que dificultou a classificação destas lavouras.

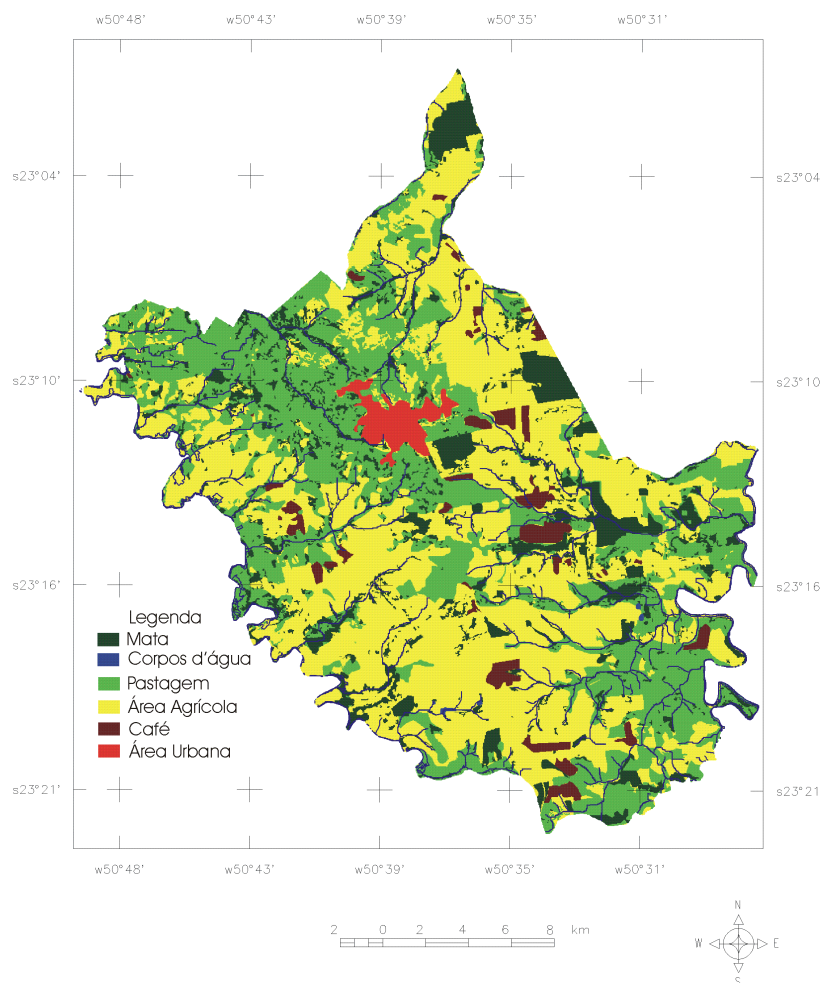


Figura 2 - Resultado do mapeamento temático utilizado no cálculo da regressão para estimar a área da cultura do café.

A área classificada com café e demais temas foi obtida de um mapa temático resultante da classificação da imagem Landsat-7/ETM⁺ da passagem de 03/11/2002 (Figura 2), cujos resultados são mostrados na Tabela 2.

Tabela 2 – Resultado do mapeamento temático para o município, utilizado para a expansão por regressão, para a cultura do café.

Tema	Área (km ²)	%
Corpos d' água	7,30	1,1
Café	18,74	2,9
Mata	91,51	14,0
Pastagem	220,12	33,8
Agricultura	304,03	46,6
Área urbana	10,40	1,6
Total	652,10	100,0

O resultado de área plantada com café obtido por meio dos estimadores de regressão e de expansão direta e pela estimativa subjetiva realizada pelo Departamento de Economia Rural (DERAL) da Secretaria de Estado da Agricultura e do Abastecimento (SEAB) do estado do Paraná (PARANÁ. SEAB/DERAL, 2003) referente ao ano safra 2002/03 é apresentado na Tabela 3. Os valores elevados do CV para a cultura do café podem estar relacionados com a baixa expressividade dessa cultura na área de estudo em relação a toda área agrícola a qual é predominantemente ocupada com a cultura da soja. Outra causa pode estar relacionada com o critério adotado na estratificação, ou seja, se a estratificação fosse baseada apenas na percentagem da área agrícola as estimativas apresentariam diferentes graus de precisão, em função da área ocupada. Isto sugere que a estratificação deve ser cautelosa quando se busca a estimativa de culturas agrícolas. Uma alternativa para aumentar a precisão da estimativa é aumentar o número de segmentos da amostra.

Tabela 3 – Área estimada, variância, coeficiente de variação (CV) e eficiência relativa (ER) para a cultura de café obtida pelos estimadores de expansão direta e de regressão, além da estimativa subjetiva do DERAL*.

Método de Estimativa	Área (km ²)	Variância (km ⁴)	CV (%)	ER
Regressão	16,8	4,7	12,9	7,99
Expansão Direta	17,5	37,6	34,9	-
Subjetivo (DERAL)*	17,5	-	-	-

Fonte: Dados da pesquisa, * PARANÁ. SEAB/DERAL, 2003.

De acordo com os resultados contidos na Tabela 3 pode-se notar que o uso do estimador por regressão foi quase oito vezes (7,99) mais eficiente do que a expansão direta, e o CV passou de 34,9% (expansão direta) para 12,9% (regressão).

Cabe ressaltar ainda que para melhorar a precisão dos resultados obtidos nessa pesquisa é necessário fazer um ajuste no painel de amostra, principalmente no que se refere à estratificação que foi realizada com base na região administrativa e não com base no município (Adami, 2003).

A comparação entre a estimativa do levantamento subjetivo do DERAL (PARANÁ. SEAB/DERAL, 2003) com a estimativa obtida pela expansão direta e pela regressão mostrou que a expansão direta superestimou a área em 0,8% e a regressão subestimou em 4,6%. As estimativas por regressão e expansão direta não diferiram da estimativa subjetiva do DERAL ao nível de 95% de confiança. Contudo o CV da estimativa por regressão foi 2,7 vezes menor do que o CV da expansão direta.

4 - CONCLUSÕES E RECOMENDAÇÕES

Considerando que a estimativa de áreas agrícolas, por meio de amostragem probabilística, é uma metodologia objetiva e fornece resultados de forma ágil pode-se concluir que: a) devido à disponibilidade de imagens livres de cobertura de nuvens foi possível obter a estimativa por regressão a qual teve eficiência relativa (ER) bem maior do que o estimador por expansão direta, além disso, o coeficiente de variação (CV) passou de 34,9% na expansão direta para 12,9 % na regressão; b) o software SPRING se mostrou adequado para realizar os procedimentos descritos nesta pesquisa, pois abrange tratamento digital de imagens, funções matriciais e vetoriais, operar com banco de dados etc. Recomenda-se que a estratificação seja realizada em função do percentual de área plantada com a cultura do café a fim de diminuir valores do coeficiente de variação das estimativas.

LITERATURA CITADA

ADAMI, M. Estimativa de áreas agrícolas por meio de técnicas de sensoriamento remoto, geoprocessamento e amostragem. São José dos Campos, 2003. 183 p. (INPE-10235-TDI/900).

ALLEN, J. D. A look at the Remote Sensing Applications Program of the National Agricultural Statistics Service. Journal of Official Statistics, v. 6, n. 4, p. 393-409, 1990.

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

COCHRAN, W. G. Técnicas de Amostragem 2 ed. Rio de Janeiro, RJ: Fundo de Cultura, 1977. 555p.

DAY, C. A compilation of PEDITOR estimaton formulas. Washington, DC: National Agricultural Statistics Service, 2002. 19 p. (RDD-02-03)

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). Multiple frame agricultural surveys: agricultural survey program based on area frame or dual frame (area and list) sample designs. Rome, 1998. v. 2.

GONZÁLES-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, 1991.

HANSEN, M. H.; HURWITZ, W. N.; MADOW, W.G. Sample Survey Methods an Theory - Theory. New York: John Wiley & Sons, 1953. v. 2, 332 p.

KRUG, T.; YANASSE, C. C. F. Estimativa de safras agrícolas utilizando dados coletados por satélites de sensoriamento remoto e dados terrestres, através de amostras de substratos geográficos. São José dos Campos, 1986. 51 p. (INPE-4102-RPE/534).

MOREIRA, M. A. Sistema de amostragem para estimar a área da cultura do trigo (*Triticum aestivum* L.) através de dados do LANDSAT. São José dos Campos, 1983. 86 p. (INPE-2941-TDL/150).

MUELLER, C. C.; SILVA, G.; VILLALOBOS, A. G. Pesquisa Agropecuária do Paraná - Safra 1986/87 (Programa de Aperfeiçoamento das Estatísticas Agropecuárias). Revista Brasileira de Estatística, v. 49, n. 191, p. 55-84, 1988.

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. 2003b). Comunicação Pessoal.

PRADHAN, S. Crop area estimation using GIS, remote sensing and area frame sampling. International Journal of Applied Earth Observation and Geoinformation, v. 3, n. 1, p. 86-92, 2001.

RUDORFF, B. F. T.; KRUG, T. Sensoriamento remoto na estimativa da área plantada com feijão, milho e mamona no município de Irecê-BA. In: SIMPOSIO LATINO-AMERICANO DE SENSORIAMENTO REMOTO, 6., Gramado, 1986. Anais... São José dos Campos: INPE, 1986. v. 1, p. 380-384.