

# ANÁLISE DA EFICIÊNCIA DOS ESTIMADORES DE EXPANSÃO DIRETA E DE REGRESSÃO PARA ÁREAS CULTIVADAS COM CAFÉ, MILHO E SOJA NO MUNICÍPIO DE CORNÉLIO PROCÓPIO, ESTADO DO PARANÁ<sup>1</sup>

Marcos Adami<sup>2</sup>  
Maurício Alves Moreira<sup>3</sup>  
Bernardo Friedrich Theodor Rudorff<sup>4</sup>  
Corina da Costa Freitas<sup>5</sup>

**RESUMO:** O objetivo deste trabalho é avaliar a eficiência dos métodos de expansão direta e de regressão, para estimar áreas de culturas agrícolas no município de Cornélio Procópio, Estado do 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 a ocupação do solo. O método de regressão apresentou os menores coeficientes de variação para as estimativas de área plantada com soja e milho, muito embora o método de expansão direta tenha tido bom desempenho, fornece estimativas de forma ágil e independe da disponibilidade de imagens. Já para a cultura do café, a disponibilidade de imagens livres de cobertura de nuvens foi boa por ser uma cultura perene e por permitir uma maior janela para aquisição das imagens. Nesse caso, o método de regressão mostrou-se bem mais eficiente do que o método de expansão direta.

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

## EFFICIENCY ANALYSIS OF DIRECT EXPANSION AND REGRESSION ESTIMATORS FOR COFFEE, CORN AND SOYBEAN CROP AREAS ESTIMATES IN THE MUNICIPALITY OF CORNÉLIO PROCÓPIO, PARANÁ STATE

**ABSTRACT:** The objective of this work is to evaluate the efficiency of the direct expansion and regression methods to estimate agricultural crop areas in the municipality of Cornélio Procópio, Paraná State. 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 produced the lowest coefficient of variation (CV) values for soybean and corn crop areas estimates, even though the direct expansion method had a good performance, provided fast estimates and is independent of cloud free images. For the coffee crop evaluation it was possible to acquire cloud free images once its perennial characteristic allows for a greater time window for image acquisition. In this case the regression method proved to be much more efficient than the direct expansion method.

**Key-words:** remote sensing, area estimate, sampling area, agricultural statistics.

**JEL Classification:** C8, C13, C42.

<sup>1</sup>Registrado no CCTC n. ASP-06/2004.

<sup>2</sup>Economista, Mestre, Departamento de Economia Rural da Secretaria de Estado da Agricultura e do Abastecimento do Paraná (e-mail: adami@seab.pr.gov.br).

<sup>3</sup>Engenheiro Agrônomo, Doutor, Instituto Nacional de Pesquisas Espaciais (INPE) (e-mail: mauricio@ltd.inpe.br).

<sup>4</sup>Engenheiro Agrônomo, Doutor, Instituto Nacional de Pesquisas Espaciais (INPE) (e-mail: bernardo@ltd.inpe.br).

<sup>5</sup>Estatística, Doutora, Instituto Nacional de Pesquisas Espaciais (INPE) (e-mail: corina@dpi.inpe.br).

## 1 - INTRODUÇÃO

Os programas de previsão de safras agrícolas, baseados em amostragem, são empregados em vários países há muitas décadas. Os Estados Unidos, por exemplo, utilizam modelos estatísticos para estimar a área e a produtividade agrícola desde 1938. Inicialmente, esses modelos tiveram por base listas do cadastro de propriedades rurais no Estado de Iowa, sendo que, em 1945, foram expandidos para quase todos os estados americanos (FAO, 1998).

A teoria da amostragem por área teve início na década de 1960, amparada por 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), que forneceram uma base sólida para os modelos estatísticos de estimativa de área cultivada por amostragem (CHEN, 1980).

Os dois principais modelos de estimativa de áreas agrícolas que utilizam imagens de sensoriamento remoto são: expansão direta e expansão por regressão. Em ambos os métodos, imagens recentes ou de arquivo são utilizadas na construção do painel amostral. Uma vez selecionadas as amostras, imagens atuais são utilizadas para identificação e delimitação das culturas agrícolas. Caso não se disponha de imagens livres de cobertura de nuvens, a identificação e delimitação das culturas agrícolas nos segmentos selecionados é feita a campo com o auxílio de imagens de arquivo. Para o método de expansão por regressão, além dos dados coletados a campo, utiliza-se a classificação de imagens de satélites ou fotografias aéreas atuais para reduzir a variância da estimativa (KRUG e YANASSE, 1986; MUELLER; SILVA; VILLALOBOS, 1988; GONZÁLES-ALONSO; SORIA; GOZALO, 1991; GALLEGO, 1995; FAO, 1996). Por exemplo, Gon-

záles-Alonso et al. (1997) compararam as estimativas de áreas agrícolas obtidas pelo estimador de expansão direta e de regressão e verificaram que a eficiência relativa do estimador por regressão foi nove vezes superior à da expansão direta.

Utilizando imagens de sensoriamento remoto, técnicas de geoprocessamento e coleta de dados a campo, essa pesquisa tem como objetivo analisar a eficiência dos estimadores de expansão direta e de regressão, para estimar áreas cultivadas com café, milho e soja, em nível municipal.

## 2 - MATERIAL E MÉTODO

### 2.1 - Estratificação da Área

A área teste, selecionada para analisar a eficiência dos estimadores, foi o município de Cornélio Procopio, localizado ao norte do Estado do Paraná, com uma área de 650 km<sup>2</sup>. 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<sup>+</sup>, respectivamente, adquiridas nos anos 2001, 2002 e 2003 (ADAMI, 2004), 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.).

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; Soria; Gozalo (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 700m de lado) e irregulares (polígonos cadastrais utilizados para o censo). O esquema de amostragem adotado foi o de amostra siste-

mática 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. Dessa forma, os estratos A e B foram divididos em segmentos quadrados de 1km x 1km (100ha) de área (GONZÁLES-ALONSO; SORIA; GOZALO, 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 construção, 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 esses 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 das variáveis pesquisadas, 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 no Município de Cornélio Procópio, Estado do Paraná, 2004

Estrato	Número total de segmentos no estrato (a)	Número de segmentos sorteados (b)	Percentual da área amostrada (b/a)
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 café, milho e soja, por meio do estimador de regressão, foi necessário classificar as imagens do sensor ETM+ do Landsat-7 referentes à órbita 222 ponto 76 adquiridas nos dias 03/11/2002 e 07/02/2003. A imagem de 03/11/2002 foi empregada para obter a área classificada com café no município e dentro dos segmentos amostrados, uma vez que essa cultura, por ser perene, pode ser identificada na imagem nessa época do ano. A imagem de 07/02/2003 foi utilizada para classificar a área das culturas de soja e milho, que nesse período encontram-se em pleno desenvolvimento vegetativo. A classificação foi realizada por meio do sistema SPRING.

Para as culturas de soja e milho, a classificação da imagem foi realizada parcialmente no município, devido à presença de nuvens na área de estudo. Dessa forma, foram criados dois pós-estratos geográficos (com e sem nuvem). No pós-estrato geográfico com nuvens, a área de cada cultura foi estimada somente pelo estimador por expansão direta, e no pós-estrato sem nuvens foi aplicado o estimador por regressão. A área total do município foi obtida somando-se estas duas estimativas. Para a cultura do café, a imagem utilizada não apresentava cobertura de nuvens e, portanto, não houve a subdivisão em pós-estratos.

## 2.2 - Cálculo das Estimativas

A seguir são apresentadas as fórmulas dos estimadores por expansão direta e por regressão, assim como suas respectivas variâncias.

### 2.2.1 - Estimador por expansão direta

A estimativa de área ocupada por uma determinada cultura, por meio da expansão direta, é obtida apenas com dados coletados nos segmentos

amostrais para cada estrato de uso do solo. O somatório da área estimada com a cultura nos diferentes estratos fornece a área total da cultura de interesse na região de estudo. A estimativa total de uma cultura ( $\hat{Z}_c$ ), é (HANSEN; HURWITZ; MADOW, 1953; COCHRAN, 1977; KRUG e YANASSE, 1986; MUELLER; SILVA; VILLALOBOS, 1988; ALLEN, 1990; PRADHAN, 2001; DAY, 2002):

$$\hat{Z}_c = \sum_{i=1}^m e_i \sum_{k=1}^{n_i} z_{c,ik} \quad (1)$$

sendo  $c$  o índice para representar a cultura de interesse;  $i$  o índice para representar o estrato de uso do solo com  $i = 1, \dots, m$ ;  $k$  o índice para representar o segmento amostrado com  $k = 1, \dots, n_i$ ;  $m$  é o número de estratos;  $n_i$  o número de segmentos amostrados no  $i$ -ésimo estrato de uso do solo;  $N_i$  o número total de segmentos no  $i$ -ésimo estrato de uso do solo;  $e_i$  ( $n_i/N_i$ ), ou seja, fator de expansão ou inverso da probabilidade de que um segmento tem de estar na amostra escolhida no  $i$ -ésimo estrato de uso do solo; e  $z_{c,ik}$  a área da cultura  $c$ , no  $k$ -ésimo segmento do  $i$ -ésimo estrato de uso do solo.

A variância estimada de  $\hat{Z}_c$  é:

$$v(\hat{Z}_c) = \sum_{i=1}^m N_i(N_i - n_i) \frac{s_{c,i}^2}{n_i} \quad (2)$$

em que  $s_{c,i}^2$  é a variância estimada da cultura  $c$  no estrato  $i$ , dada por:

$$s_{c,i}^2 = \frac{\sum_{k=1}^{n_i} (z_{c,ik} - \bar{z}_{c,i})^2}{(n_i - 1)} \quad (3)$$

e  $\bar{z}_{c,i}$  é a área média da cultura  $c$  no  $i$ -ésimo estrato de uso do solo:

$$\bar{z}_{c,i} = \frac{\sum_{k=1}^{n_i} z_{c,ik}}{n_i} \quad (4)$$

## 2.2.2 - Estimador por regressão

No método de estimativa por regressão, além dos dados dos segmentos amostrais coletados no campo, é necessário obter os dados da classificação da imagem para toda a área de estudo, ou seja, um mapa com a distribuição espacial das culturas. Uma análise de correlação entre esses dados é realizada para os segmentos amostrais (MOREIRA, 1983; KRUG e YANASSE, 1986; MUELLER; SILVA; VILLALOBOS, 1988; ALLEN, 1990; DAY, 2002). A estimativa da área da cultura  $c$ , para o  $i$ -ésimo estrato de uso do solo, por meio do estimador por regressão é obtida por:

$$\hat{Y}_{c,i}(reg) = N'_i [\bar{y}_{c,i} + b_{c,i}(\bar{X}_{c,i} - \bar{x}_{c,i})]. \quad (5)$$

sendo  $N'_i$  o número total de segmentos no  $i$ -ésimo estrato de uso do solo, no pós-estrato sem cobertura de nuvens;  $\bar{y}_{c,i}$  é a área média da cultura  $c$  por segmento, utilizando os dados coletados em campo, no  $i$ -ésimo estrato de uso do solo, para o pós-estrato livre de nuvens obtido através de:

$$\bar{y}_{c,i} = \frac{\sum_{k=1}^{n_i} y_{c,ik}}{n_i} \quad (6)$$

em que  $y_{c,ik}$  é a área total da cultura  $c$  obtida em campo no  $k$ -ésimo segmento amostrado do  $i$ -ésimo estrato de uso do solo, para o pós-estrato sem cobertura de nuvens;  $n_i$  é o número de segmentos amostrados no  $i$ -ésimo estrato de uso do solo, no pós-estrato sem cobertura de nuvens;  $\bar{X}_{c,i}$  o número médio populacional de *pixels* classificados com a cultura  $c$ , por segmento, no  $i$ -ésimo estrato do uso do solo para o pós-estrato sem cobertura de nuvens, e é obtido por:

$$\bar{X}_{c,i} = \frac{X_{c,i}}{N'_i} \quad (7)$$

em que  $X_{c,i}$  é o número total de *pixels* classificados com a cultura  $c$ , no  $i$ -ésimo estrato do uso do solo

para o pós-estrato sem cobertura de nuvens;  $\bar{x}_{c,i}$  é o número médio amostral de *pixels* por segmento da cultura  $c$  no  $i$ -ésimo estrato de uso do solo, para o pós-estrato sem cobertura de nuvens, definido por:

$$\bar{x}_{c,i} = \frac{\sum_{k=1}^{n_i} x_{c,ik}}{n_i} \quad (8)$$

em que  $x_{c,ik}$  é o número total de *pixels* da cultura  $c$  no  $k$ -ésimo segmento amostrado do  $i$ -ésimo estrato de uso do solo, para o pós-estrato sem cobertura de nuvens;  $b_{c,i}$  é o estimador de mínimos quadrados do coeficiente de regressão para a cultura  $c$  no  $i$ -ésimo estrato de uso do solo no pós-estrato sem cobertura de nuvens, obtido por:

$$b_{c,i} = \frac{\sum_{k=1}^{n_i} (y_{c,ik} - \bar{y}_{c,i})(x_{c,ik} - \bar{x}_{c,i})}{\sum_{k=1}^{n_i} (x_{c,ik} - \bar{x}_{c,i})^2} \quad (9)$$

A estimativa da área da cultura  $c$ , para toda a região livre de nuvens é dada por:

$$\hat{Y}_c(reg) = \sum_{i=1}^m \hat{Y}_{c,i}(reg), \quad (10)$$

A variância de  $\hat{Y}_c(reg)$ , é dada por:

$$v(\hat{Y}_c(reg)) = \sum_{i=1}^m \frac{N_i(N_i - n_i)}{n_i} s_{c,i}^2 (1 - r_{c,i}^2) \quad (11)$$

em que:

$$s_{c,i}^2 = \frac{\sum_{k=1}^{n_i} (y_{c,ik} - \bar{y}_{c,i})^2}{n_i - 2} \quad (12)$$

e  $r_{c,i}^2$  é o estimador do coeficiente de correlação entre os dados de campo e os dados classificados por satélite, para a cultura  $c$ , no  $i$ -ésimo estrato de uso do solo, dado por:

$$r_{c,i}^2 = \frac{\sum_{k=1}^{n_i} (y_{c,ik} - \bar{y}_{c,i})(x_{c,ik} - \bar{x}_{c,i})}{\sum_{k=1}^{n_i} (y_{c,ik} - \bar{y}_{c,i})^2 \sum_{k=1}^{n_i} (x_{c,ik} - \bar{x}_{c,i})^2} \quad (13)$$

Na situação em que a área de estudo é parcialmente coberta por nuvens, como é o caso das culturas de milho e soja, a estimativa da área para a cultura é obtida através dos dois estimadores, isto é, na área coberta com nuvens, emprega-se o estimador por expansão direta e no restante da área o estimador por regressão.

O passo final é unir as estimativas obtidas pelo estimador por regressão com aquelas obtidas pelo estimador por expansão direta. Assim, para os pós-estratos com cobertura de nuvens, a estimativa da cultura  $c$  e sua respectiva variância são dadas por:

$$\hat{Y}_c(exdir) = \sum_{i=1}^m \frac{N_i''}{n_i''} \sum_{k=1}^{n_i''} y_{c,ik}'' \quad (14)$$

$$v(\hat{Y}_c(exdir)) = \sum_{i=1}^m \frac{N_i(N_i'' - n_i'')}{n_i''(n_i'' - 1)} \sum_{k=1}^{n_i''} (y_{c,ik}'' - \bar{y}_{c,i}'')^2 \quad (15)$$

sendo  $y_{c,ik}''$  a área total da cultura  $c$  obtida em campo no  $k$ -ésimo segmento do  $i$ -ésimo estrato de uso do solo, no pós-estrato com cobertura de nuvens;  $N_i''$  o número total de segmentos no  $i$ -ésimo estrato de uso do solo, no pós-estrato com cobertura de nuvens;  $n_i''$  o número de segmentos amostrados no  $i$ -ésimo estrato de uso do solo, no pós-estrato com cobertura de nuvens;  $\bar{y}_{c,i}''$  é a área média da cultura  $c$  por segmento, utilizando os dados coletados em campo, no  $i$ -ésimo estrato de uso do solo, no pós-estrato com cobertura de nuvens obtido através de:

$$\bar{y}_{c,i}'' = \frac{\sum_{k=1}^{n_i''} y_{c,ik}''}{n_i''} \quad (16)$$

A estimativa total da área da cultura  $c$  em toda a região de estudo será:

$$\hat{Y}_c(tot) = \hat{Y}_c(reg) + \hat{Y}_c(exdir) \quad (17)$$

em que  $\hat{Y}_c(reg)$  é dado por (10) e  $\hat{Y}_c(exdir)$  é dado por (14).

A variância estimada de  $\hat{Y}_c(tot)$  é dada por:

$$v(\hat{Y}_c(tot)) = v(\hat{Y}_c(reg)) + v(\hat{Y}_c(exdir)) \quad (18)$$

onde  $v(\hat{Y}_c(reg))$  é dada por (11) e  $v(\hat{Y}_c(exdir))$  é dada por (15).

A avaliação do estimador por regressão em relação ao estimador por expansão direta, também conhecida como eficiência relativa (ER) (KRUG e YANASSE, 1986 e GONZÁLES-ALONSO; SORIA; GOZALO, 1991), é realizada pela seguinte equação:

$$ER = \frac{v(\hat{Z}_c)}{v(\hat{Y}_c(tot))}, \quad (19)$$

sendo que  $v(\hat{Z}_c)$  é definida em (2) e  $v(\hat{Y}_c(tot))$  é definida em (18). ER indica quantas vezes deve-se aumentar a amostra para que a expansão direta tenha a mesma precisão da estimativa por regressão.

### 3 - RESULTADOS E DISCUSSÃO

O resultado de área plantada com café, milho e soja, 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, 2003), referente ao ano-safra 2002/03, é apresentado na tabela 2. Nota-se nessa tabela que o menor CV foi obtido para a estimativa da área de soja (10,31%) pelo estimador de regressão. Os altos valores do CV obtidos para as culturas de milho e café podem estar relacionados com a menor expressividade dessas culturas na área de estudo, principalmente, quando comparados à cultura da soja. Outra causa pode estar relacionada com o critério adotado na estratificação, ou seja, se a estratificação fosse

baseada apenas na porcentagem da área agrícola as estimativas apresentariam diferentes graus de precisão, em função da área ocupada. Isso sugere que a estratificação deve ser cautelosa quando se busca a estimativa de culturas agrícolas presentes em diferentes proporções, tal como a soja em relação ao milho e café. Uma alternativa para aumentar a precisão da estimativa é aumentar o número de segmentos da amostra.

**Tabela 2** - Área Estimada, Variância, Coeficiente de Variação (CV) e Eficiência Relativa (ER) para as Culturas de Café, Milho e Soja Obtida pelos Estimadores de Expansão Direta e de Regressão, além da Estimativa Subjetiva do DERAL

Método de estimativa	Área (km <sup>2</sup> )	Variância (km <sup>4</sup> )	CV (%)	ER
Cultura do café				
Regressão	16,8	4,7	12,9	7,99
Expansão direta	17,5	37,6	34,9	-
Subjetivo (DERAL)	17,5	-	-	-
Cultura do milho				
Regressão + expansão	32,7	94,6	29,8	2,27
Expansão direta	45,1	215,4	32,5	-
Subjetivo (DERAL)	40,0	-	-	-
Cultura da soja				
Regressão + expansão	271,4	784,0	10,3	1,09
Expansão direta	252,4	860,3	11,6	-
Subjetivo (DERAL)	255,0	-	-	-

Fonte: Dados da pesquisa e Paraná (2003).

De acordo com os resultados contidos na tabela 2, pode-se notar que, para a cultura do café, o uso do estimador por regressão foi quase oito vezes (7,99) mais eficiente do que a expansão direta e, ao mesmo tempo, o CV passou de 34,9% (expansão direta) para 12,9% (regressão). No caso da cultura do milho, em que houve a presença de nuvem em parte do município, a eficiência da regressão foi pouco mais de duas vezes. Esse fato é atribuído à maneira como foi estimada a área da cultura e o seu percentual no município. Nota-se também que o CV foi relativamente alto para ambas as estimativas, ou

seja, a regressão não melhorou a precisão da estimativa em relação à expansão direta. Para a cultura da soja, que é bastante expressiva no município (255km<sup>2</sup>), a ER foi praticamente a mesma com valores de CV relativamente baixos e similares, indicando que a expansão direta fornece boa estimativa. Além disso, ela é ágil e independe de imagens livre de cobertura de nuvens.

Observou-se que os CVs apresentados na tabela 2 estão dentro dos intervalos apresentados por diversos autores (MUELLER; SILVA; VILLALOBOS, 1988; VILLALOBOS et al., 1988; HILL e MÉGIER, 1988; GALLEGO, 1995; GONZÁLES-ALONSO; SORIA; GOZALO, 1991; TSILIGIRIDES, 1998; GONZÁLES-ALONSO et al., 1997; GÓMEZ e GALLEGO, 2000; e NEMATZADEH, 2001). Entretanto, deve-se salientar que as escalas de amostragem para os trabalhos realizados por esses autores foram regionais e estaduais, enquanto, neste caso, a escala de amostragem foi municipal. 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, pois neste estudo ela foi realizada com base na região administrativa e não no município (ADAMI, 2004).

Com relação à agilidade na obtenção dos resultados, salienta-se que a expansão direta fornece a estimativa tão logo termine o trabalho a campo. Já para a regressão, é necessário obter as imagens livres ou parcialmente livres de nuvens no período de pleno desenvolvimento vegetativo das culturas, para só então realizar a classificação e obter a estimativa por esse método, limitando seu uso para estimativas com previsão embora tenham uma melhor precisão. Neste trabalho os resultados da expansão direta foram obtidos 3 dias após o término da coleta de dados no campo (10/12/2002), enquanto para a regressão, os resultados só foram obtidos no final de fevereiro de 2003.

A comparação do resultado de estimativa de área obtida pela expansão direta e pela regressão com aquela obtida pelo levantamento subjetivo do DERAL (PARANÁ, 2003), para o ano de 2003, é apresentada na tabela 3. Observa-se que as estimativas

pelos métodos de regressão e de expansão direta foram muito próximas às do método subjetivo, exceto para a cultura do milho, cuja diferença em relação ao método subjetivo foi de 18,3% para a estimativa por regressão e de 12,8% para a estimativa por expansão direta.

**Tabela 3** - Diferença Relativa (DR) da Área Estimada pela Amostragem por Regressão e por Expansão com a Área Estimada pelo DERAL

Cultura	Estimativa (km <sup>2</sup> ) regressão (a)	Estimativa (km <sup>2</sup> ) expansão direta (b)
Café	16,8	17,5
Milho	32,7	45,1
Soja	271,4	252,4
Total	320,9	315,1

  

Cultura	Estimativa (km <sup>2</sup> ) DERAL (c) <sup>1</sup>	DR (%) (c-a)/c	DR (%) (c-b)/c
Café	17,5	4,6	-0,2
Milho	40,0	18,3	-12,8
Soja	255,0	-6,4	1,0
Total	312,5	-2,7	-0,8

Fonte: Dados da pesquisa e Paraná (2003).

Na tabela 3 pode ser observado que há boa diferença entre os dados estimados pelos diversos modelos para a cultura do milho. Entretanto, devido à grande variabilidade de ocorrência de áreas com milho nos segmentos amostrados, foi observado que essas estimativas não diferiram entre si a 95% de confiança.

Para a cultura da soja, pode-se observar na tabela 3 que a estimativa pela expansão direta foi próxima à estimativa subjetiva do DERAL e também não diferiu da estimativa por regressão a 95% de confiança.

#### 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) para a cultura do café o estimador de regressão teve eficiência relativa (ER)

bem maior do que o estimador por expansão direta, devido à disponibilidade de imagens livres de cobertura de nuvens, além disso, o coeficiente de variação (CV) passou de 34,9% na expansão direta para 12,9% na regressão; b) para a cultura do milho, que tem pouca expressão na área de estudo, a ER do estimador de regressão, em conjunto com a expansão direta, foi apenas duas vezes maior, em relação à expansão direta apenas, no entanto, ambos os modelos apresentaram CVs altos, cabendo ressaltar que a disponibilidade de imagens livres de nuvens poderá acarretar uma significativa melhora na ER, tal como foi observado para a cultura do café que também tem pouca expressão em termos de área cultivada no município; c) para a cultura da soja, que é bastante expressiva em termos de área cultivada, não se observou melhora nas estimativas de área da cultura pelo estimador de regressão, em conjunto com a expansão direta, em relação à expansão direta apenas, ou seja, nessas condições ela fornece boa estimativa e tem a vantagem de ser obtida de forma ágil e independe de imagens livres de cobertura de nuvens; d) o *software* SPRING mostrou-se adequado para realizar os procedimentos descritos nesta pesquisa, pois abrange tratamento digital de imagens, funções matriciais e vetoriais, opera com banco de dados, etc. Recomenda-se que a estratificação seja realizada em função do percentual de área plantada com cada cultura, 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, 2004. 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 estimation 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.
- \_\_\_\_\_. **Multiple frame agricultural surveys: current survey based on area and list sampling methods**. Rome, 1996. v. 1.
- GALLEGO, F. J. **Sampling frames of square segments**. Luxembourg: Joint Research Centre, 1995. 72p.
- GÓMEZ, S. N.; GALLEGO, J. Crop area estimation pilot project: Zimbabwe. In: INTERNATIONAL SYMPOSIUM AN REMOTE SENSING OF ENVIRONMENT, 28., 2000, Cape Town. **Proceedings...** Cape Town: ICRSE, 2000. p. 4.
- 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.
- \_\_\_\_\_. et al. 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.
- 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.
- HILL, J.; MÉGIER, J. Regional land cover and agricultural area statistics and mapping in The Département Ardèche, France, by use Thematic Mapper data. *International Journal of Remote Sensing*, v. 9, n. 10-11, p. 1573-1595, 1988.
- 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 sub-estratos 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.
- NEMATZADEH, M. Estimation of area with area frame sampling in Hamadan province of Iran. In: IGARSS'01. Sydney, 2001. **Proceedings...** IEEE, 2001. v. 7, p. 2970-2972.
- PARANÁ. Secretaria de Estado da Agricultura e do Abas-



tecimento/Departamento de Economia Rural (SEAB/DERAL). **Dados da Região de Cornélio Procópio**. [mensagem pessoal]. Mensagem recebida por: [ccpseab@pr.gov.br](mailto:ccpseab@pr.gov.br) em fev. 2003.

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: SIMPÓSIO LATINO-AMERICANO DE SENSORIAMENTO REMOTO, 6., Gramado, 1986. **Anais...** São

José dos Campos: INPE, 1986. v. 1, p. 380-384.

TSILIGIRIDES, T. A. Remote sensing as a tool for agricultural statistics: a case study of area frame sampling methodology in Hellas. **Computers and Electronics in Agriculture**, n. 20, p. 45-77, 1998.

VILLALOBOS, A. G. et al. Desenvolvimento do sistema de informações agropecuárias. Características e resultados do modelo no Distrito Federal. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 5. Natal, 1988. **Anais...** Natal: Instituto Nacional de Pesquisas Espaciais & Sociedad de Especialistas Latinoamericanos en Percepción Remota, 1988. v. 1, p. 16.

---

Recebido em 13/05/2004. Liberado para publicação em 11/08/2004.