

# ALEATORIZAÇÃO DE PONTOS NO TERRITÓRIO DE UM MUNICÍPIO, USANDO O SPRING, PARA A ESTIMATIVA DE ÁREA AGRÍCOLA POR AMOSTRAGEM

Alfredo José Barreto Luiz<sup>1</sup>  
Salette Gürtler<sup>2</sup>

## RESUMO

Esse trabalho descreve um procedimento para a geração de um conjunto de pontos amostrais aleatórios, definidos no espaço bidimensional pelas suas coordenadas geográficas, e determina como fazer a combinação desse conjunto com uma imagem obtida por satélite de sensoriamento remoto, usando o programa SPRING. O uso de aparelhos GPS auxilia a localização desses pontos no campo, permitindo a utilização do método em levantamentos de estatísticas agrícolas. É fornecido um exemplo de amostragem aleatória de 100, 200 ou 300 pontos dentro dos limites do município de Ipuã - SP, visando a estimativa de área das principais culturas agrícolas.

**PALAVRAS-CHAVE:** SIG, amostragem, estatísticas agrícolas, distribuição espacial.

## POINT RANDOMIZATION IN COUNTY TERRITORY , USING SPRING, TO ESTIMATE CROP AREA BY SAMPLING

## ABSTRACT

This paper describes a procedure to generate a set of random sampling points, well defined by their geographic coordinates, and establish how to do the combination between this set and an image obtained from remote sensing satellite, using SPRING software. The use of GPS equipment aids these points localization in the ground, allowing the method utilization in agricultural statistics surveys. An example is furnished with random sampling of 100, 200 or 300 points inside the limits of Ipuã municipality, in São Paulo State (Brazil), aiming to estimate the main crops area.

**KEYWORDS:** GIS, sampling, agricultural statistics, spatial distribution.

## 1. INTRODUÇÃO

No passado, o uso de amostragem aleatória por pontos em levantamentos estatísticos para estimativa de área plantada era quase impossível pela dificuldade de localização e pela inexistência de uma lista enumerável destes pontos. A disponibilização de três tecnologias: imagens de sensoriamento remoto, aparelhos GPS, e sistemas de informações geográficas (SIG), mudou esse quadro. A vantagem do uso das imagens na construção de painéis amostrais é que a sua aquisição produz uma divisão imaginária da superfície terrestre em elementos de cena (pixels) de mesmas dimensões que a recobrem totalmente e sem sobreposição. Se a área de um município é a região de interesse, existe um número finito  $N$  de pixels que representam na imagem a sua superfície real, sendo  $N = (\text{área do município})/(\text{área do pixel})$ . Dessa maneira obtém-se uma lista enumerável dos componentes da população, da qual pode-se retirar amostras aleatórias. A dificuldade em identificar no terreno elementos amostrais com limites imaginários (e não reais, como rios, matas, etc.), como o pixel, é contornada pela precisão e facilidade de operação dos aparelhos de GPS, disponíveis a baixo custo. Finalmente, para integrar os benefícios do uso das imagens e dos aparelhos GPS, estão disponíveis os SIG, que facilitam o tratamento da informação georreferenciada (Luiz et al., 2002).

---

<sup>1</sup> Eng°. Agrônomo, MSc. em Estatística, Dr. Sensoriamento Remoto – Embrapa Meio Ambiente/INPE

<sup>2</sup> Ecóloga, MSc. em Sensoriamento Remoto – INPE

---

O objetivo desse trabalho é descrever um procedimento para a geração de um conjunto de pontos amostrais aleatórios, definidos no espaço bidimensional pelas suas coordenadas geográficas, e determinar como fazer a combinação desse conjunto com uma imagem, através do programa SPRING. Será fornecido um exemplo de amostragem aleatória de 100, 200 ou 300 pontos dentro dos limites do município de Ipuã - SP, visando a estimativa de área das principais culturas agrícolas.

## 2. MATERIAL E MÉTODOS

O local do estudo foi o município de Ipuã, SP, no qual as principais culturas são a soja e a cana-de-açúcar (Epiphany et al., 2001). Foram feitas duas visitas ao campo, em agosto de 2001 e fevereiro de 2002, nas quais foi realizado, com o auxílio de imagens do sensor ETM+ do satélite Landsat 7, órbita/ponto 220/74, o mapeamento do uso do solo no município, na escala de 1:60.000, sobre cartas preparadas conforme Luiz et al. (2002). Os mapas foram digitalizados como “Planos de Informação” (PIs) temáticos no SPRING (INPE/DPI, 2003).

Os aplicativos computacionais utilizados foram: o SPRING, para o tratamento das imagens, informações georreferenciadas e mapas; e o EXCEL, para a geração dos números aleatórios, gráficos e cálculos. As imagens foram registradas e importadas para um projeto criado no SPRING, na projeção UTM - CA, onde o retângulo envolvente era o menor a conter os limites do município. Para a delimitação do município, foi utilizada a grade municipal digital do IBGE, ajustada utilizando-se a carta de Ipuã, escala 1:50.000 (IBGE, 1972). Foram calculados no SPRING os valores das áreas: do retângulo envolvente, do município de Ipuã e das classes de uso do solo nas duas datas.

Devido às características irregulares das divisas municipais, para distribuir aleatoriamente os pontos amostrais na área do município realizou-se primeiro um sorteio dentro do retângulo envolvente, sobre a imagem armazenada no SPRING, fazendo a equivalência entre ponto e pixel, descartando-se depois os pontos que caíram fora dos limites municipais.

Trabalhando-se sobre a representação matricial da imagem, a posição de cada pixel é definida pelo binômio linha e coluna. Chame-se de  $x$  o número da coluna e  $y$  o valor da linha em uma imagem (note-se que  $x$  e  $y$  podem ser multiplicados pela dimensão do pixel, transformando sua unidade em metros, ou convertidos em coordenadas de latitude e longitude). O retângulo envolvente é formado por  $r$  linhas e  $s$  colunas e absolutamente definido pelos pares ordenados das posições de apenas dois elementos de cena, um no canto inferior esquerdo ( $x_1, y_1$ ) e o outro no canto superior direito ( $x_2, y_2$ ). Observa-se que  $r = (|x_2 - x_1| + 1)$  e que  $s = (|y_2 - y_1| + 1)$  e que o número total de pixels do retângulo é igual ao produto entre  $r$  e  $s$ , que será chamado de  $T$ . Para permitir a realização da escolha aleatória de  $m$  candidatos a pontos amostrais entre os  $T$  pixels é necessário que se gerem  $m$  pares ordenados ( $x_i, y_i$ ). Se o gerador retorna um número aleatório  $z$  maior ou igual a 0 e menor que 1, distribuído uniformemente, basta aplicar  $m$  vezes a transformação:  $x_i = (z \times r)$  e  $y_i = (z \times s)$ .

Esses pontos são considerados candidatos pois, como nem todo o retângulo é ocupado pela área do município, uma fração destes  $m$  elementos estará localizada fora dos limites da região de interesse e não serão amostrados. Apenas  $n$  elementos sorteados, cuja localização esteja dentro dos limites determinados, serão considerados pontos amostrais. Como a escolha dos pontos foi aleatória, espera-se que, em média, a relação entre  $n$  e  $m$  seja a mesma existente entre a área do município ( $A$ ) e a área do retângulo envolvente ( $R$ ). Desta relação é possível determinar a melhor estimativa de  $m$  que produza um determinado valor de  $n$ , que pode ser calculada da seguinte forma:  $m = \{1 + \text{INT}[n \times (R/A)]\}$ , onde  $\text{INT}$  significa a parte inteira da expressão entre colchetes.

Considerando a área de Ipuã como 46.837,4 ha (IBGE, 2003), e a do retângulo envolvente como 86.265 ha; para selecionar amostras de tamanho  $n_1 = 100$ ,  $n_2 = 200$  e  $n_3 = 300$ , seria necessário sortear um total de:  $m_1 = \{1 + \text{INT}[100 \times (86.265/46.837,4)]\} = 185$ ,  $m_2 = 369$ , e  $m_3 = 553$ ; ou seja, gerados 185 pontos dentro do retângulo envolvente, a esperança matemática é de que 100 deles estejam localizados dentro dos

limites do município, e assim, da mesma forma, para  $m_2$  e  $m_3$ . Para aumentar a probabilidade de conseguir o número desejado de pontos dentro da região de interesse, e como o custo de geração desses números aleatórios é computacionalmente muito baixo, gerou-se o dobro desses pontos, ou seja, 370 para garantir 100, 738 para 200, e 1106 para 300.

Para transformar os pares ordenados sorteados  $(x_i, y_i)$  em valores compatíveis com as coordenadas planas  $(X_i, Y_i)$  do retângulo envolvente definido no SPRING, no caso de uma imagem com resolução espacial de 25 metros, utiliza-se a seguinte transformação:  $X_i = X_1 + 12,5 + [\text{INT}(x_i \times 25)]$  e  $Y_i = Y_1 + 12,5 + [\text{INT}(y_i \times 25)]$ , na qual  $(X_1, Y_1)$  define a posição da extremidade sudoeste do retângulo envolvente, no caso do hemisfério Sul. Esse procedimento garante que cada um dos os pixels da imagem tenha a mesma chance de ser escolhido para a amostra, e que a escolha se dará pelo centro do pixel.

As coordenadas dos pontos gerados aleatoriamente, conforme descrito, foram inseridas no projeto do SPRING via importação, na forma de arquivos de pontos em formato ASCII-SPRING, o que gerou um “Plano de Informação” (PI) para cada conjunto de pontos. Para eliminar os pontos que não estavam dentro do polígono que corresponde ao limite municipal de Ipuã, efetuou-se o recorte desses “PIs”, utilizando uma máscara com o limite. Como os “PIs” resultantes continham um número de pontos superior ao desejado, realizou-se a exportação dos “PIs” para possibilitar a abertura em uma planilha Excel. Os pontos foram, então, classificados em ordem crescente e destas listagens foi selecionado um conjunto com o número de pontos desejado para cada arquivo, ou seja, os primeiros 100, os primeiros 200 e os primeiros 300 pontos. Os demais pontos foram excluídos dos arquivos, que foram novamente importados para o SPRING, concluindo a etapa.

### 3. RESULTADOS E DISCUSSÃO

Foram geradas dez amostras tamanho 100, dez tamanho 200 e dez tamanho 300. Cada uma delas teve seus pontos cruzados automaticamente com o mapa temático de uso do solo, produzindo as estimativas de área ocupada com cada uso, com os respectivos desvios. Os dados assim obtidos foram comparados com a área calculada a partir do mapa, aqui considerada como “verdade”. A tabela 1 apresenta os valores verdadeiros e estimados (na média de dez repetições dos três tamanhos de amostra) para os diversos tipo de uso do solo, em agosto de 2001 e janeiro de 2002.

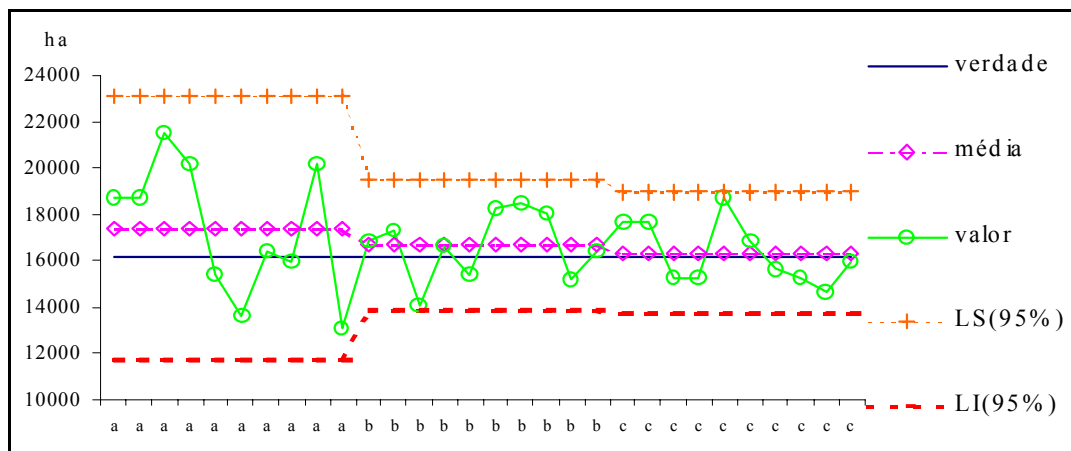
**TABELA 1: Área ocupada com diferentes tipos de uso do solo, em hectares, observada e estimada com três tamanhos de amostra (média de dez repetições), no município de Ipuã - SP, em duas datas.**

Uso	Agosto de 2001				Janeiro de 2002			
	verdade	média100	média200	média300	verdade	Média100	média200	média300
Cana	13004,5	12786,6	12833,5	13239,4	13420,1	13442,3	13348,7	13692,1
cana nova	2091,3	2529,2	1873,5	2014,0	1291,1	1264,6	1428,5	1139,7
Soja	0,0	0,0	0,0	0,0	16154,8	17376,7	16674,1	16299,4
Milho	193,8	93,7	117,1	187,3	1546,5	1639,3	1498,8	1857,9
Outras culturas	817,0	1264,6	819,7	843,1	111,8	187,3	70,3	140,5
vegetada	8905,6	8290,2	9086,5	8664,9	8891,3	8571,2	8992,8	8446,3
não vegetada	716,0	515,2	889,9	765,0	722,3	468,4	866,5	811,8
Palha	8115,3	8383,9	8243,4	7837,5	969,5	562,0	936,7	983,6
solo exposto	12993,9	12974,0	12974,0	13286,2	3729,9	3325,5	3021,0	3466,0

Pode-se observar dos dados apresentados que para os tipos de uso que ocupam uma significativa fração da área do município, não há grande ganho com o aumento do tamanho da amostra. Já para os usos pouco frequentes, esse ganho é bastante perceptível.

Na figura 1 está representado o comportamento das estimativas de área de soja, mostrando os valores obtidos em cada uma das 30 amostras, as médias das dez amostras com 100, 200 e 300 pontos, bem como os limites superior (LS) e inferior (LI) do intervalo de confiança para essas médias com 95% de probabilidade, todos comparados à “verdade” obtida do mapeamento realizado.

Observa-se que, com o aumento do tamanho da amostra, há uma consistente convergência da média em direção ao valor verdadeiro e uma redução da amplitude de variação das estimativas. Em todos os casos observados as estimativas sempre estiveram dentro do intervalo de confiança de 95%.



**FIGURA 1: Área estimada para a cultura da soja, em amostras de 100 (a), 200 (b) e 300 (c) pontos, distribuídos aleatoriamente no município de Ipuã, SP, em janeiro de 2002.**

#### 4. CONCLUSÕES

A combinação de imagens obtidas por satélite com o processamento digital de dados espaciais por meio de um SIG e com o uso do aparelho GPS para localização no campo permite a aplicação do método de amostragem por pontos em levantamentos estatísticos para a obtenção de estimativas agrícolas.

O método de amostragem por pontos permite gerar estimativas não tendenciosas de área ocupada por diferentes tipos de uso do solo. A precisão das estimativas é facilmente controlada através da determinação do tamanho da amostra.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

EPIPHANIO, J.C.N.; BARROS NETO, O.O.; LUIZ, A.J.B.; FORMAGGIO, A.R. Sistema de amostragem em imagem como base para estimativa de áreas agrícolas no município de Ipuã - SP. [CD-ROM]. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 10., 2001, Foz do Iguaçu. **Anais...** São José dos Campos: INPE, 2001. 8 p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **CIDADES**. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 20 maio 2003.

\_\_\_\_\_. **Ipuã: Folha SF-22-X-B-III-4**. [Rio de Janeiro], 1972. Carta do Brasil – Escala 1:50.000.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS/Divisão de Processamento de Imagens (INPE/DPI). **Manual do Spring**. Disponível em: <[http://www.dpi.inpe.br/spring/usuario/menu\\_spr.htm](http://www.dpi.inpe.br/spring/usuario/menu_spr.htm)>. Acesso em: 12 de junho de 2003.

LUIZ, A.J.B; OLIVEIRA, J.C; EPIPHANIO, J.C.N; FORMAGGIO, A.R. Auxílio das imagens de satélite aos levantamentos por amostragem em agricultura. **Agricultura em São Paulo**, São Paulo, v. 49, n. 1, p.41-54, 2002.