

DETERMINAÇÃO DO USO DO SOLO POR AMOSTRAGEM DE PONTOS EM FOTOGRAFIA AÉREA

Eng.º Agr.º Salomão Schattan

1 — O PROBLEMA

A fotografia aérea da superfície da terra tem hoje inúmeras aplicações. Uma delas, ligada diretamente à agricultura, é a determinação do uso que se fazia de determinada área da terra no momento em que foi fotografada.

Esta determinação terá que ser feita usando diferentes técnicas conforme a área a ser estudada, seja pequena como um município ou grande como o Estado de São Paulo.

No caso de um município, pode-se proceder à montagem das fotografias em um mosaico e nele delimitar, por inspeção, as áreas que no município estejam alocadas aos diferentes usos agrícolas. A seguir, mede-se a área dedicada a cada um dos usos por meio de um planímetro ou um retículo de pontos. A cada ponto deste retículo corresponderá uma

área determinada que varia conforme a escala da fotografia e a distância entre os pontos do retículo.

As medições feitas desta forma deverão ser corrigidas para as variações de escala provocadas pela variação da altura do avião que tira as fotografias e as distorções de escala que ocorrem do centro para a periferia de uma mesma fotografia.

Como já dissemos, estas medições e correções são exequíveis quando se trata de pequenas áreas, entretanto esta mesma técnica para ser aplicada a grandes áreas exigiria numeroso pessoal especializado e grande equipamento tornando o trabalho além de moroso, caro de precisão por véses duvidosa.

Para o caso de grandes áreas foi idealizada a técnica de amostragem de pontos que usa

os princípios teóricos da amostragem de caracteres qualitativos.

2 — AMOSTRAGEM DE PONTOS

A amostragem de pontos na fotografia aérea pode ser levada a efeito usando a amostragem irrestrita aleatória, ou amostragem sistemática. Na amostragem irrestrita aleatória os pontos são alocados na fotografia por meio de coordenadas escolhidas com o auxílio de “tabela de números ao acaso”. Na amostragem sistemática usa-se o “retículo de pontos” que consiste, como diz o próprio nome, em sobrepor à fotografia da área a ser estudada um retículo que defina sobre a mesma, certo número de pontos igualmente distanciados entre si.

Se se trata por exemplo de estudar um município, do qual possuímos um mosaico devidamente montado, podemos definir um sistema de coordenadas e nele local o número de pontos desejado quando usamos amostragem irrestrita aleatória ou então, usando técnicas variadas lhe sobrepor um retículo.

Quando se prefere a amostragem sistemática e portanto o retículo, o primeiro passo consiste em definir o número de pontos desejados. A partir desse número pode-se calcular facilmente a distância entre as paralelas que formarão o retículo. O cruzamento de duas linhas, uma horizontal e outra vertical definirá sobre a foto-

grafia um ponto que constituirá uma “unidade de amostragem”. A seguir constrói-se o retículo de dimensões desejadas sobre a fotografia ou então se sobrepõe à fotografia o retículo desejado.

Qualquer um desses dois métodos da amostragem pode ser conjugado com a estratificação geográfica ou com a de estágios múltiplos. Havendo necessidade pode-se igualmente amostrar em uma segunda fase fazendo-se então em uma sub-amostra o trabalho de campo desejado.

3 — IDENTIFICAÇÃO E CLASSIFICAÇÃO DOS PONTOS

A importância do sistema de amostragem de pontos sobre a fotografia aérea reside no fato de se poder identificar, por simples inspeção, o uso que se fazia da área onde cai o ponto que é então considerado solidário com uma pequena área que lhe é imediatamente circunjacente.

Desta forma, agrupando os usos em classes definidas, pode-se alocar cada ponto a uma classe de usos.

O resto é simples. Cada ponto constitui uma “unidade de amostragem” e ele só pode pertencer a uma única classe. Se consideramos por exemplo as pastagens como uma das classes de uso, vemos que determinado ponto só pode pertencer a uma das duas classes C = com pastagens e C' = Sem pastagens, representando neste

caso a classe C' todos os usos excepto pastagens.

Está claro que, o estabelecimento das classes fica limitado pela possibilidade de identificação na fotografia.

O ponto de vista de alguns especialistas em foto-interpretção é que em fotografia de 1 : 25000 pode-se identificar com segurança as seguintes classes:

- 1 — Mata:
 - 1.1 — Floresta + Cerradão
 - 1.2 — Talhadia : eucaliptus + pinheiros
 - 1.3 — Áreas recentemente desmatadas
- 2 — Cerrado:
 - 2.1 — Cerrado
 - 2.2 — Campo natural
- 3 — Pasto:
 - 3.1 — Pasto limpo
 - 3.2 — Pasto sujo
- 4 — Lavouras:
 - 4.1 — Culturas permanentes
 - 4.2 — Culturas anuais
 - 4.3 — Áreas de culturas abandonadas.
- 5 — Rios - Lagos - Reprêas - Estradas - Áreas inaproveitadas
- 6 — Áreas urbanas
- 7 — Áreas em construção - na zona rural (inclui terreiros).

A classificação acima permite separar 13 usos diferentes do solo e cada ponto (unidade de amostragem) tendo um uso inequívoco terá que ser alocado a uma e só uma destas 13 classes.

4 — AMOSTRAGEM AO ACASO PARA PROPORÇÕES

4.1 — *Classificação dos pontos*

Temos interesse na determinação da porcentagem ou área que, em determinada região, é dedicada a um uso específico. Queremos por exemplo saber, qual a área com pastagens no Estado de São Paulo. Teoricamente a área ocupada em pastagem no Estado de São Paulo pode variar de zero ao total da área do Estado, constituindo portanto uma variável contínua. Entretanto, por meio de uma pequena transformação faremos com que ela possa ser tratada como se fôra variável descontínua.

Em vez de fazer meditações na fotografia aérea o que seria evidentemente possível, faremos a contagem das unidades de amostragem (pontos) que caem nas áreas cobertas de pastagens, classificando cada um dos pontos que constituem a amostra em "ponto caído sobre pastagem" e em "ponto caído sobre área não de pastagem". Pode-se evidentemente levantar um sem número de discussões e problemas sobre casos hipotéticos; entretanto na prática a técnica é, não só exequível como, fácil e rápida, não suscitando no geral qualquer problema a identificação da classe de uso e que deve ser alocado o ponto.

Chamemos de C a classe cujos elementos apresentam a característica "caído sobre pastagem" e C' a outra classe

cujos elementos não apresentam a referida característica não é pastagem).

Adotaremos a seguinte notação:

Convenhamos antes de mais nada em que é infinito o número de pontos contido na área de pastagem, isto é $N = \infty$. Chamaremos de n o número total de pontos na amostra. Teremos além disso

Número de unidades em C
 Na população | Na amostra
 A | a

Proporção de unidades em C
 Na população | Na amostra
 A | a

(Nos textos de estatística usa-se geralmente P e não Q , porém quando se escreve à máquina de datilografia fica mais fácil distinguir entre amostra e população usando Q).

As fórmulas para o cálculo da variância de Q são facilmente deduzidas das fórmulas correspondentes da amostragem irrestrita aleatória se atribuímos aos elementos da classe C valor 1 isto é ($x_1 = x_2 = \dots = 1$) e aos elementos da classe C' o valor 0 ($x_1' = x_2' \dots = 0$).

Em nosso caso, tratando-se realmente de uma variável contínua, é igualmente infinito o número de pontos que contém a classe C na população; entretanto dado o artifício que usamos, este fato não traz inconvenientes já que é finito o número de ponto em C na amostra e é igual a a .

4.2 — Cálculo da proporção

Desta forma, a proporção de elementos que na amostra, pertence á classe C é dada por:

$$q = \frac{\sum_i^n x_i}{n} = \frac{a}{n}$$

Está claro que q da amostra é a estimativa não visada de Q da população. Consequentemente, o problema da estimativa de Q pode ser olhada como

análogo ao cálculo da média de uma população onde a variável x_i só pode tomar os valor zero ou um.

4.3 — Cálculo da variância

$$s_x^2 = \frac{\sum_i^n (x_i - q)^2}{n - 1} = \frac{\sum_i^n x_i^2 - \frac{(\sum x_i)^2}{n}}{n - 1} = \frac{1}{n - 1} (nq - nq^2)$$

$$s_x^2 = \frac{n}{n - 1} (q \cdot p)$$

onde $p = 1 - q$

Da fórmula de s_x^2 vemos que a variabilidade de um tributo das unidades de amostragem depende unicamente da proporção de unidades que na po-

pulação possuem o atributo.

Sendo q o correspondente da média em amostragem irrestrita aleatória temos

$$s_q^2 = \frac{q \cdot p}{n} \quad \text{e} \quad s_q = \sqrt{\frac{q \cdot p}{n}}$$

4.4 — Cálculo do intervalo de confiança

aproximação normal para este caso podemos passar ao cálculo do intervalo de confiança dado pela seguinte expressão:

Admitindo que seja válida a

$$q \pm t \sqrt{\frac{q \cdot p}{n}}$$

onde t é o desvio normal correspondente ao nível de confiança desejado.

faça uma correção para a continuidade; tal correção aumenta o intervalo de confiança e temos então a seguinte expressão

Cochran recomenda que se

$$q \pm t \left(\sqrt{\frac{q \cdot p}{n}} + \frac{1}{2n} \right)$$

Quando é grande o número de pontos na amostra a correção se aproxima de zero e voltamos à expressão anterior.

mente de C e C' às áreas cobertas pelas classes temos:

$$T = C + C' = \text{Área total}$$

4.5 — Tamanho da Amostra para Obter Precisão Desejada

$$Q \text{ é a proporção } \frac{C}{C + C'}$$

Para se determinar o tamanho necessário da amostra é preciso estabelecer previamente o que se espera da amostra em termos de erro, bem como o risco que desejamos correr de que o erro seja na realidade maior do que aquêle previamente fixado.

É fixada inicialmente uma margem de erro d na estimação da proporção q das unidades pertencentes à classe C e um risco que estamos dispostos a correr de que o verdadeiro valor divirja do valor estimado por uma quantidade maior do que .

Considerando as duas classes C e C' e chamando igual-

Tudo isso pode ser resumido na seguinte equação

$$\Pr \left\{ |q - Q| \geq d \right\} = \alpha$$

Admitindo que q tenha distribuição normal e ignorando a correção para população finita concluímos imediatamente

$$d = t \sqrt{\frac{q \cdot p}{n}} \quad \therefore \quad d^2 = t^2 \frac{q \cdot p}{n} \quad \therefore \quad n = \frac{t^2 q p}{d^2}$$

V — EXEMPLO PRÁTICO

5.1 — Daremos a seguir como exemplo o estudo feito em Tietê.

Foram definidas inicialmente sete classes de uso a saber:

- a) — Pasto
- b) — Culturas
- c) — Capoeira
- d) — Matas
- e) — Construções

f) — Estradas

g) — Rios e lagos

Sobre o mosaico do município foram alocados 1380 pontos por meio de um reticulado de pontos.

A seguir, os pontos foram inspecionados e classificados sem qualquer dificuldade ou ambiguidade em uma das sete classes acima mencionadas.

No quadro a seguir é dada a classificação dos 1380 pontos.

CLASSIFICAÇÃO DOS PONTOS NO MOSAICO DE TIETÊ

Porcentagem sobre o total e estimativa da área ocupada por cada classe

Classes	N.º de pontos	% sobre o total	Área em alqueires
Pasto	824	59,71	10 150
Cultura	340	24,64	4 189
Capoeira	86	6,23	1 059
Mato	41	2,97	505
Casa	23	1,67	284
Estrada	49	3,55	604
Rios e lagos	17	1,23	209
TOTAL	1 380	100,00	17 000

A área coberta por cada uma das classes de uso é estimada diretamente como porcentagem da área total do município que é conhecida e igual a 17 000 alqueires. É evidente que a percentagem de pontos dá igualmente a percentagem de área.

5.2 — Erro de amostragem

A fim de exemplificar o cálculo do erro de amostragem tomemos a classe “Capoeira” com 86 pontos correspondente a 6,23% do número total de pontos bem como da área do Município.

$$s^2_{6,23} = \frac{6,23 \times 93,77}{1.380} = \frac{584,19}{1.380} = 0,423$$

$$s^2_{6,23} = 0,65$$

5.3 — Intervalo de confiança

Admitindo um nível de con-

fiança igual a 95 teremos que o intervalo de confiança será dado por

$$6,23 \pm 1,96 \times 0,65 = 6,23 \pm 1,28$$

o que representa uma baixa precisão. A forma de aumentar a precisão da estimativa da área coberta com rios e lagos no município de Tietê está no aumento do número de pontos inspecionados.

5.4 — Tamanho necessário da amostra

O tamanho da amostra depende evidentemente da precisão desejada. Fixada a preci-

são calcula-se facilmente o tamanho necessário da amostra.

Se desejamos que a área ocupada com capoeiras no Município de Tietê seja calculada com erro menor igual a 0,62% e que o verdadeiro valor tenha 95% de probabilidade de se encontrar no intervalo $6,23 \pm 0,62$ então a amostra terá que ser constituída de um número de pontos dado pela fórmula abaixo.

$$\text{Número de pontos necessários} = \frac{(100 - 6,23) \times 38.400}{6,23 \times 10^2} = 5.780$$

Portanto, se desejamos um nível de confiança igual a 95% e se o erro de amostragem de-

ve ser igual a 0.311 então, a amostra terá que se constituir de 5.780 pontos