

TAMANHO DA AMOSTRA DE REFERÊNCIA PARA VERIFICAÇÃO DE EXATIDÃO EM IMAGENS ORBITAIS

Thomaz Corrêa e Castro da Costa¹
Ricardo Seixas Brites¹

¹Depto Eng. Florestal-UFV, 36571-000 Viçosa-MG
tcorrea@alunos.ufv.br

Abstract. The sample size reference for accuracy images is defined with normal approximation of binomial function without pilot sample. Therefore accuracy and variance are not estimated, but fixed *a priori*. For large discrepancy between a *priori* and estimate variance from accuracy the sample error will deviate from the intended error. The other investigation in this work was the study of the spatial autocorrelation to help the definition of the reference sample size. Semivariograms were applied on bands of image and sills were estimates. The results show the sensible deviation between a *priori* and estimate accuracy do not cause a large tendency towards the sample error. The semivariograms estimate smaller sampled size for total image with concern distinct pattern image.

Keyword: Semivariogram, sample size, image accuracy

INTRODUÇÃO

A qualidade da classificação de imagens orbitais é avaliada por meio de índices de exatidão (Exatidão Global e Kappa, dentre os mais comuns), calculados a partir de matrizes de erros que expressam a concordância entre a imagem classificada e a verdade de campo, representada pela amostra de referência. Para construção desta amostra deve-se considerar o desenho de amostragem apropriado, o tamanho da amostra e a distância entre unidades de amostra.

Com relação ao desenho de amostragem, o único método inadequado para selecionar amostras de referência é a amostragem blocada, que superestima a exatidão de uma cena, devido ao problema da autocorrelação espacial. Os demais procedimentos, amostragem aleatória, aleatória estratificada, e amostragem sistemática não diferenciaram significativamente em termos de tendência na estimação da exatidão (BRITES, 1996).

O tamanho da amostra de referência pode ser calculado pela fórmula de dimensionamento para a distribuição binomial,

$$n = Z^2 pq / E^2,$$

sendo p=exatidão desejada;

$$q = 1 - p;$$

$Z \sim N(0, 1)$, o ponto crítico para um limite de confiança de inclusão para P, o parâmetro exatidão real; e

E=erro de amostragem.

O dimensionamento indica que é necessário n pontos independentes para se obter um intervalo de confiança com E pré estabelecido, que incluirá o verdadeiro parâmetro, P, ao nível

de probabilidade determinado por Z. Embora a solução analítica para o dimensionamento esteja correta, comete-se na prática um erro de aproximação. A estimativa usada para P não é obtida por uma amostra de referência piloto, mas fixada como exatidão desejada. Isto implica que ao obtermos uma exatidão estimada muito distante da fixada para dimensionamento, o erro de amostragem planejado, E, não será alcançado.

As Figuras 1 e 2 ilustram esta conseqüência. A Figura 1 representa o comportamento de n em função de p estimado. Note que para uma exatidão desejada de 0,90 dimensiona-se a amostra para 34 pontos, mas ao obter uma estimativa de exatidão da imagem de 0,60, a amostra de 34 pontos estará subdimensionada, aumentando conseqüentemente o erro de amostragem para 0,16 e não 0,1 como planejado (Figura 2).

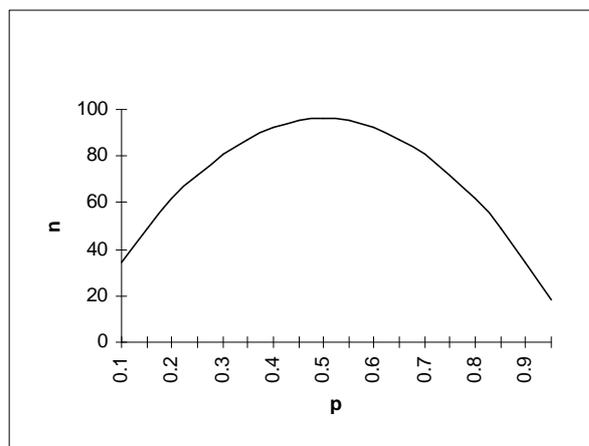


FIGURA 1-Tamanho da amostra (n) em função da exatidão estimada (p), para um intervalo de confiança com E=0,1 a 95% de probabilidade

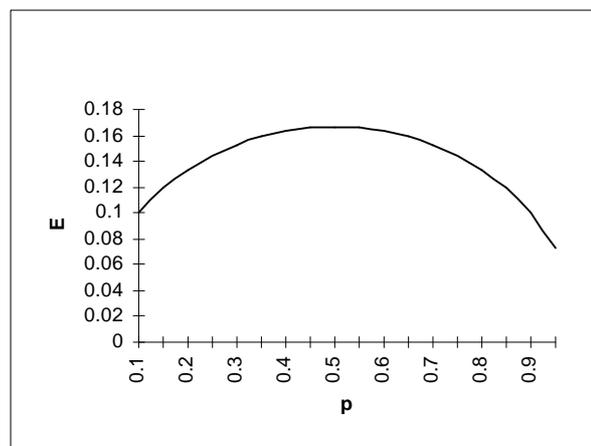


FIGURA 2-Erro de Amostragem (E) em função da exatidão estimada (p) a 95% de probabilidade para n=34

A terceira questão com relação a amostra de referência é definir a distância ideal entre os pontos de referência. Imagens de sensoriamento remoto apresentam forte correlação espacial, pois pixels mais próximos apresentam valores de brilho similares (demarcam uma feição), e a medida que se aumenta a distância entre pixels, o desvio entre valores de brilho tende a aumentar, até que a variação não sofra mais a influência da distância. Esta variação representa a dependência espacial entre valores vizinhos. Podemos então medir o raio de influência de um ponto na imagem por meio de uma função de correlação espacial entre valores vizinhos, afim de evitar a marcação de pontos que sejam dependentes entre si.

Funções de correlação espacial como os semivariogramas vem sendo utilizadas em sensoriamento remoto para identificar padrões espaciais de feições (CURRAN, 1988; RAMSTEIN, 1989; COHEN et al., 1990; LACAZE et al., 1994), para estimar erros provocados por ruídos em imagens de sensoriamento remoto (ATKINSON, 1996; ATKINSON, 1997), para classificar imagens de radar (MIRANDA et al., 1992; 1996) entre outros.

Em estudos de amostragem, McBRATNEY e WEBSTER (1983) demonstraram um método para determinar o tamanho de amostra, considerando a dependência espacial determinada pelo semivariograma. A variância total pode ser minimizada para um dado tamanho de amostra se a

posição das observações formam uma grade sistemática. Exemplos mostraram que a relação entre erro padrão e tamanho de amostra pode ser alcançada para qualquer precisão desejada. Outro estudo é apresentado em ENTZ e CHANG (1991), que avaliaram delineamentos de amostragem usando análise geoestatística (semivariograma e Kriging).

Neste estudo seguimos a sugestão de CURRAN (1988), que propôs o uso da geoestatística na determinação da amostra de referência. Aplicamos semivariogramas para determinar o raio de influência, ou alcance entre pontos, nas bandas utilizadas para classificação, com o intuito de auxiliar na construção da amostra de referência. Semivariogramas indicariam a distância ideal para marcação de pontos independentes, exigidos pela teoria de amostragem para definição de intervalos de confiança.

O objetivo deste trabalho foi investigar o efeito da exatidão *a priori* para dimensionar amostras de referência, e verificar se o conhecimento da correlação espacial em uma imagem auxilia no dimensionamento de amostras de referência.

O Método Tradicional

A pressuposição de distribuição binomial para a concordância entre pixels de referência e pixels classificados é explicada em RICHARDS (1986) conforme segue: suponha uma variável aleatória $X=\{0,1,2,\dots, n\}$ para representar a quantidade de pixels corretamente classificados de uma amostra de referência de n pixels (provas), que assume valor 1, se o pixel é corretamente classificado (sucesso), e 0, para classificação incorreta (fracasso). Sendo p a probabilidade de sucesso, ou exatidão esperada, a probabilidade, de X pixels serem corretamente classificados em uma amostra aleatória de n pixels é dada pela função binomial:

$$p(X=x)=C_x^n p^x (1-p)^{n-x}$$

O intervalo de confiança para o parâmetro P (exatidão real) da distribuição binomial é construído a partir da aproximação normal descrito em MEYER (1984). Considerando n suficientemente grande em dados de referência, o intervalo de confiança pode ser calculado aproximadamente como:

$$IC \cong p \pm Z \sqrt{\frac{pq}{n}}, \quad \text{pois a variância de } p \text{ é } V(p) = \frac{pq}{n}$$

A fórmula de dimensionamento é facilmente obtida a partir do erro de amostragem, E , isolando-se n , conforme abaixo:

$$E = Z \sqrt{\frac{pq}{n}} \therefore E^2 = Z^2 \frac{pq}{n} \therefore n = \frac{Z^2 pq}{E^2}$$

Geoestatística

A Geoestatística, também chamada de “Teoria de Variáveis Regionalizadas” foi desenvolvida por Matheron (1963, 1971), citado por VIEIRA et al. (1983). Sendo $Z(\mathbf{x})$ o valor de uma variável regionalizada, Z , no ponto \mathbf{x} , é possível descrever a variabilidade da função $Z(\mathbf{x})$ no espaço \mathbf{x} (coordenadas) variando dependentemente do local da amostragem (Grossi sad, 1986, citado por HAMAKAWA, 1991). Seu comportamento pode ainda variar com a direção, ou seja, refletir a anisotropia do fenômeno. Pelo fato de utilizarmos apenas uma amostra de um processo aleatório,

duas pressuposições são necessárias para sua consistência, que são Estacionaridade e Ergodicidade (VIEIRA et al., 1983).

A **Semivariância** é a função de autocorrelação mais utilizada para descrever a variabilidade de $Z(\mathbf{x})$, pois atende tanto a hipótese de forte quanto fraca estacionaridade. Semivariância é a média entre variâncias de pares de dados separados por um vetor distância \mathbf{h} ou *lag*. Considere um par de valores $[Z(\mathbf{x}), Z(\mathbf{x}+\mathbf{h})]$, separados pela distância \mathbf{h} . A variância S^2 para este par de valores é:

$$S^2 = \frac{[Z(\mathbf{x}) - \bar{Z}]^2 + [Z(\mathbf{x} + \mathbf{h}) - \bar{Z}]^2}{2 - 1} = [Z(\mathbf{x}) - \bar{Z}]^2 + [Z(\mathbf{x} + \mathbf{h}) - \bar{Z}]^2$$

Sendo $\bar{Z} = \frac{Z(\mathbf{x}) + Z(\mathbf{x} + \mathbf{h})}{2}$, e substituindo-se em S^2 , temos que:

$$S^2 = \frac{[Z(\mathbf{x}) - Z(\mathbf{x} + \mathbf{h})]^2}{2} = \text{SemiVar},$$

que é a Semivariância para este par de valores. Se considerarmos \mathbf{k} pares de valores separados por um mesmo \mathbf{h} de \mathbf{n} valores, a semivariância média será:

$$\text{SemiVar}(\mathbf{h}) = \frac{\mathbf{E}[Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})]^2}{2} = \frac{1}{2\mathbf{k}} \sum_{i=1}^{\mathbf{k}} [Z(\mathbf{x}_i + \mathbf{h}) - Z(\mathbf{x}_i)]^2$$

sendo $\mathbf{k}=\mathbf{n}-\mathbf{h}$.

O gráfico gerado por esta função é chamado **Semivariograma**. A curva do Semivariograma, aumenta a medida que \mathbf{h} cresce atingindo, teoricamente, um patamar (*sill*), quando a semivariância é aproximadamente igual a variância da população (Clark, 1979, citado por PREVEDELLO, 1987). O início deste patamar define o alcance ou raio de influência (*range*), ou seja, pares de dados a uma distância maior que o raio de influência são independentes.

No exemplo apresentado na Figura 3 o semivariograma para a população P2E, formada por quatro estratos de 25 elementos arranjados de forma aleatória, segue este comportamento, indicando um alcance de 25 elementos. Este comportamento padrão pode variar em algumas populações, não sendo este patamar tão fácil de definir. Podem ocorrer curvas parabólicas com o ponto de máximo indicando o *lag*, conforme observado o semivariograma para a população P1E, na Figura 3. Para esta população, formada de três estratos com correlação linear crescente, o primeiro com 25 elementos, o segundo com 50 elementos, e o terceiro com 25 elementos, verifica-se que a curva do semivariograma apresenta um ponto de máximo na distância de 25 m, tornando a decair com o aumento do lag.

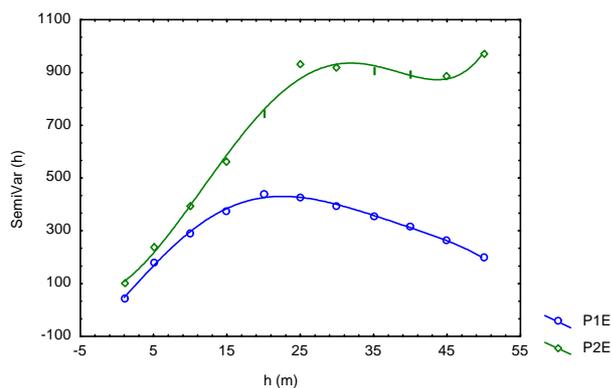


FIGURA 3-Semivariogramas para duas populações estratificadas simuladas (P1E e P2E)

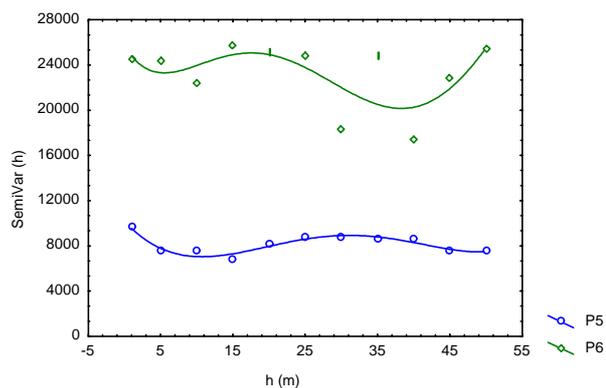


FIGURA 4-Semivariogramas para duas populações aleatórias simuladas (P5 e P6)

É esperado também em semivariogramas que, para $h=0$, $SemiVar(h)$ também igual a zero, que pode não ocorrer na prática. As curvas geralmente não passam pela origem, e sim por um valor positivo. Este fenômeno é conhecido como efeito pepita (*nugget*), que é explicado por flutuações aleatórias que ocorrem a distâncias mais curtas que o primeiro *lag* considerado na análise (Burgess e Webster, 1980 citados por PREVEDELLO, 1987). O efeito nugget pode também ocorrer independentemente da distância, fenômeno típico de populações aleatórias, conforme pode ser observado para as populações P5 e P6, na Figura 4.

Modelos para Ajuste dos Semivariogramas

Os modelos mais utilizados para ajuste da curva da função de semivariância são (PREVEDELLO, 1987):

Modelo linear : $SemiVar(h) = Co + Ch$

Modelo esférico: $SemiVar(h) = Co + C\left\{\left[\frac{3}{2}\right] \cdot \left(\frac{h}{a}\right) - \left[\frac{1}{2}\right] \cdot \left(\frac{h}{a}\right)^3\right\}$, para $0 \leq h \leq a$
 $Co + C$, para $h > a$

Modelo exponencial: $SemiVar(h) = Co + C[1 - \exp(-h/a)]$

Modelo de Gauss: $SemiVar(h) = Co + C\{1 - \exp[-(h/a)^3]\}$

Todos estes modelos pressupõe o efeito pepita e a ocorrência de um patamar (“sill”): $h > a$

MATERIAL E MÉTODOS

Área de Estudo

A área está situada na microrregião de Viçosa-MG, com predominância neste Município. Foi utilizada parte de uma imagem, VIC, de 15 x 15 km (22.500 ha), extraída da imagem Landsat5-TM, Órbita 217 Ponto 74, Quadrante Sul, de 10/10/94, composta de 3 bandas espectrais: TM3, TM4 e TM5. A correção geométrica da imagem foi feita com pontos de controle coletados em

Ortofotocartas da CEMIG, resolução 1:10.000, utilizando-se a função de mapeamento quadrática, com reamostragem pelo vizinho mais próximo, atingindo um RMS (raiz do erro médio quadrático) de 0,2028 metro.

Semivariogramas foram aplicados a imagem VIC de 15 x 15 km e a três subcenas de 3 x 3 km da imagem de acordo com o seguinte critério: a primeira com um grande fragmento de vegetação (F17); a segunda com predominância de zona urbana e vegetação (F29); e a terceira com predominância de zona rural (pastagens e pequenos fragmentos de vegetação-F39). Veja Figura 5.

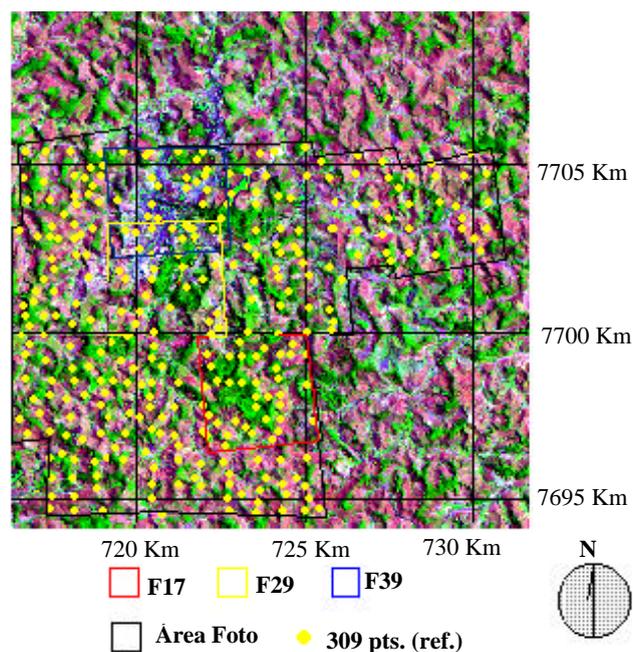


FIGURA 5-Composição colorida 345 da imagem VIC, área das fotos referentes às subcenas F17, F29 e F39, área coberta por fotos aéreas, e os 309 pontos da amostra de referência

O motivo de utilizar estas subcenas foi o de estudar a correlação espacial em diferentes padrões da imagem. A função semivariância foi aplicada em quatro transectos: Norte-Sul (NS); Oeste-Leste (OL); Noroeste-Sudeste (NOSE) e Nordeste-Sudoeste (NESO), e a média entre os quatro transectos. Os lags utilizados para cada transecto foram: $h=1,2,3,\dots,50$ pixels. Os transectos foram aplicados as bandas 3, 5 e a imagem gerada pelo índice de vegetação TVI. O modelo utilizado para ajuste da função semivariância não foi um dos citados na literatura, mas o polinomial de quarto grau

$$\text{SemiVar}(h) = ah^4 + bh^3 + ch^2 + dh + C_0$$

por melhor se ajustar ao comportamento dos dados, pois alguns transectos não apresentam estabilização da variância. Os alcances foram determinados pela média entre alcances estabelecidos nas bandas 3, 5 e TVI.

O cálculo dos n pontos independentes a partir de alcances médios obedeceu ao seguinte procedimento: Divide-se o número de colunas e linhas pelo alcance (distância em pixels). A

multiplicação dos resultados fornece o número de pontos independentes para toda a cena. O fator de 0,502181 corrige o número de pontos para a área coberta pelas fotos aéreas.

A outra etapa foi construir um arquivo de referência com 309 pontos (Figura 5), que foi sistematicamente reduzido resultando nos arquivos 309, 292, 282, 272, 262,..., 12 pontos. Para cada arquivo foi calculado o índice de exatidão global (G), e sua variância. Para testar a diferença entre os índices de exatidão (G_1 e G_2), o teste estatístico utilizado foi o de Z, conforme expressão abaixo (MA e REDMOND, 1995):

$$Z = \frac{G_2 - G_1}{\sqrt{V(G_2) + V(G_1)}}$$

sendo

$$G = p = \frac{1}{n} \sum_{i=1}^k n_{ii} \quad \text{e} \quad V(G) = \frac{pq}{n}$$

RESULTADOS E DISCUSSÃO

Alguns semivariogramas aplicados às subcenas estão representados nas Figuras 6 a 9. O comportamento diferenciado das curvas de diferentes transectos comprovam o efeito anisotrópico. Assim, a análise em apenas uma direção não fornece resultados seguros de alcance.

Analisando a curva média entre transectos para cada banda, obtivemos alcances aproximados de 10 a 30 pixels como descrito no Quadro 1. O alcance médio por subcena variou, provocando uma grande diferença no tamanho da amostra (n). Para F17 e F39 são necessários no máximo 472 pontos igualmente espaçados para independência. Para F29, n foi de apenas 163 pontos, indicando maior correlação espacial nesta subárea da imagem. E para VIC, a subcena total, n ficou em 738 pontos. É provável que menores alcances para toda a cena sejam devidos ao tamanho do transecto, proporcionando um maior número de pares de dados.

A média geral do alcance entre as subcenas, incluindo a imagem VIC, foi de 19 pixels, resultando em um n de 372 pontos de referência, o que significa que amostras de referência menores ou iguais a 372 pontos sistematicamente afastados a distâncias maiores ou iguais a 19 pixels garantem pontos, em sua maioria, independentes.

Estes resultados demonstram a instabilidade do método proposto por CURRAN (1988), aplicado a imagens de sensoriamento remoto, porque se a análise fosse realizada apenas na imagem VIC teríamos o resultado de até 738 pontos independentes, sendo que para algumas áreas da imagem, representando padrões distintos, o número de pontos necessários foi bem menor.

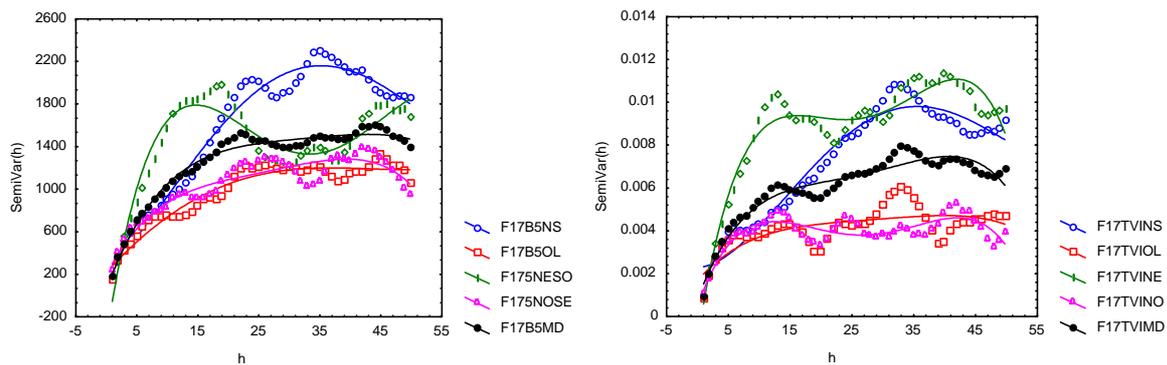


FIGURA 6-Semivariogramas para os transectos NS (Norte-Sul); OL (Oeste-Leste); NESO (Nordeste-Sudoeste); NOSE (Noroeste-Sudeste); e MD (Média entre Transectos) para as bandas 5 e TVI da Subcena F17

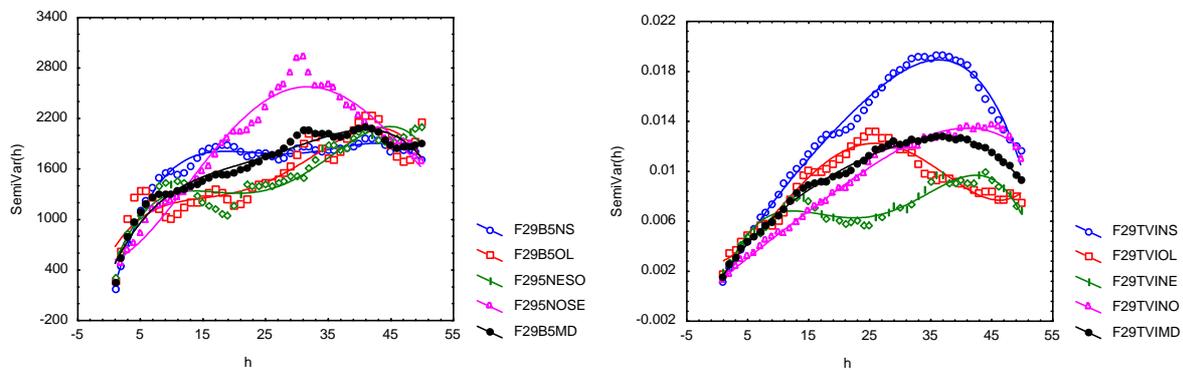


FIGURA 7-Semivariogramas para os transectos NS (Norte-Sul); OL (Oeste-Leste); NESO (Nordeste-Sudoeste); NOSE (Noroeste-Sudeste); e MD (Média entre Transectos) para a bandas 5 e TVI da Subcena F29

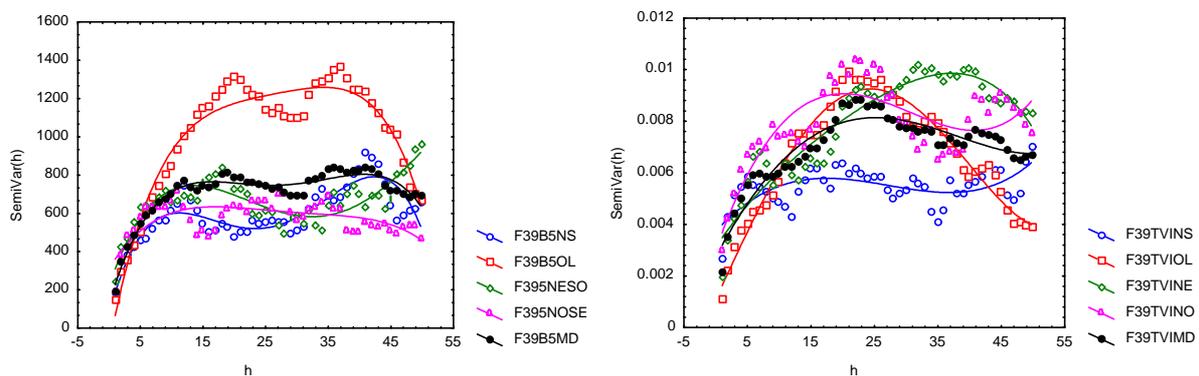


FIGURA 8-Semivariogramas para os transectos NS (Norte-Sul); OL (Oeste-Leste); NESO (Nordeste-Sudoeste); NOSE (Noroeste-Sudeste); e MD (Média entre Transectos) para as bandas 5 e TVI da Subcena F39

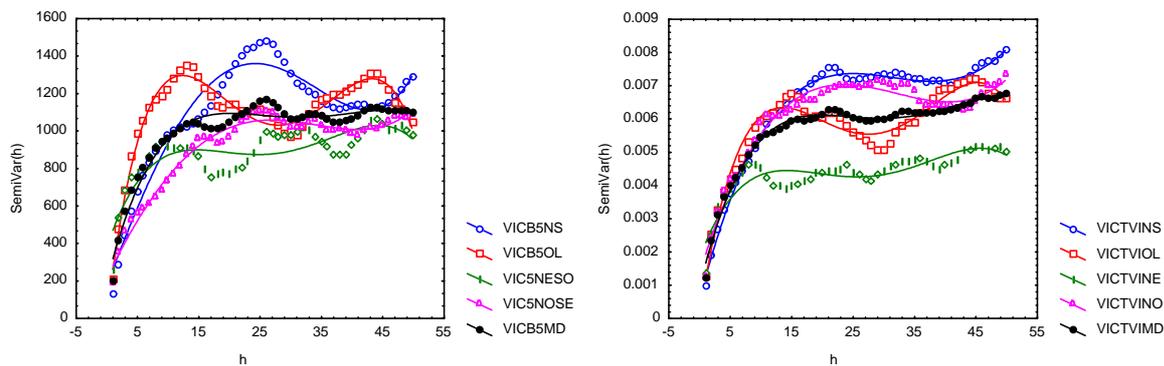


FIGURA 9-Semivariogramas para os transectos NS (Norte-Sul); OL (Oeste-Leste); NESO (Nordeste-Sudoeste); NOSE (Noroeste-Sudeste); e MD (Média entre Transectos) para as bandas 5 e TVI da Imagem VIC

QUADRO 1- Alcances para as bandas B3, B5 e TVI para as subcenas F17, F29 e F39 e VIC, Dimensionamento de Pontos Independentes (n) para o alcance médio em cada subcena, e para média geral de Alcances

Subcena	Banda	Alcance	n
F17	B3	15	
	B5	20	
	TVI	15	
	Média	17	472
F29	B3	25	
	B5	30	
	TVI	30	
	Média	28	163
F39	B3	20	
	B5	10	
	TVI	20	
	Média	17	472
VIC	B3	10	
	B5	15	
	TVI	15	
	Média	13	738
Média geral		19	373

O Quadro 2 mostra o procedimento piloto aplicado as amostras de 12 a 309 pontos, para dimensionar a amostra de referência com erro de amostragem fixado em 0,1. O uso de vários tamanhos de referência (12,22,..., 309 pontos) foi para verificar a estabilidade do dimensionamento, que variou de 59 a 80 pontos, com média de 71 pontos.

QUADRO 2-Exatidão Global (G), sua Variância, V(G), Teste Z para Sucessivos pares de G: 22/12, 32/22,..., 309/292, Intervalo de Confiança (IC), Erro de Amostragem (E), e Redimensionamento (n), para as Amostras de Referência (Refer)

Refer	G	V(G)	Z	Zcal>1.96	IC	E	n	
12	0.75	0.015625	0.40	ns	0.51	1.00	0.25	72
22	0.81	0.006995	0.02	ns	0.65	0.97	0.16	59
32	0.81	0.004761	-0.29	ns	0.68	0.95	0.14	59
42	0.79	0.004009	-0.19	ns	0.66	0.91	0.12	65
52	0.77	0.003414	-0.14	ns	0.65	0.88	0.11	68
62	0.76	0.002958	-0.11	ns	0.65	0.86	0.11	70
72	0.75	0.002604	0.26	ns	0.65	0.85	0.10	72
82	0.77	0.002171	-0.97	ns	0.68	0.86	0.09	68
92	0.70	0.002267	1.29	ns	0.61	0.80	0.09	80
102	0.78	0.001659	-0.59	ns	0.70	0.86	0.08	65
112	0.75	0.001674	0.07	ns	0.67	0.83	0.08	72
122	0.75	0.001520	-0.08	ns	0.68	0.83	0.08	71
132	0.75	0.001420	0.19	ns	0.68	0.82	0.07	72
142	0.76	0.001285	0.06	ns	0.69	0.83	0.07	70
152	0.76	0.001189	-0.08	ns	0.70	0.83	0.07	69
162	0.76	0.001128	-0.10	ns	0.69	0.83	0.07	70
172	0.75	0.001077	0.05	ns	0.69	0.82	0.06	71
182	0.76	0.001011	-0.50	ns	0.69	0.82	0.06	71
192	0.73	0.001016	-0.04	ns	0.67	0.80	0.06	75
202	0.73	0.000970	0.18	ns	0.67	0.79	0.06	75
212	0.74	0.000906	-0.15	ns	0.68	0.80	0.06	74
222	0.73	0.000879	-0.03	ns	0.68	0.79	0.06	75
232	0.73	0.000844	-0.13	ns	0.68	0.79	0.06	75
242	0.73	0.000820	-0.03	ns	0.67	0.78	0.06	76
252	0.73	0.000789	-0.03	ns	0.67	0.78	0.06	76
262	0.73	0.000761	0.17	ns	0.67	0.78	0.05	77
272	0.73	0.000722	0.25	ns	0.68	0.78	0.05	75
282	0.74	0.000680	0.34	ns	0.69	0.79	0.05	74
292	0.75	0.000636	0.29	ns	0.70	0.80	0.05	71
309	0.76	0.000584			0.72	0.81	0.05	69
Média	0.75							71

A estatística Z que testa a significância entre pares sucessivos de exatidão global (G) foi aplicada apenas para provar que o tamanho da amostra não introduz tendência na estimativa de G. A comprovação está na não significância entre G's sucessivos a 5%, e G's extremos com testes não incluídos no Quadro 2.

Para uma amostra de referência piloto de 12 pontos, o erro de amostragem foi de 0,25. Para dimensioná-lo em 0,1 são necessários 72 pontos, como pode ser constatado pela nossa amostra de 72 pontos. O segundo ensaio piloto com amostra de 22 pontos redimensionou o n para 59 pontos. Pegando a amostra mais próxima de 62 pontos, obtivemos um E de 0,11.

Da mesma maneira uma amostra de 52 pontos, próxima a 49 pontos (dimensionamento para E=0,1, a 95% de probabilidade e exatidão pré-fixada "a priori" p=0,85), gerou o mesmo erro de

amostragem, de 0,11. Mas se a exatidão pré fixada for muito otimista, $p=0,90$, por exemplo, que dimensiona o n para 34 pontos, o Erro passa para 0,14 ($n=32$). Já o procedimento piloto apresentou desvios em E no máximo em 0,01, variando de 0,09 ($n=80$, aproximando para amostra de 82 pontos) a 0,11 ($n=59$, aproximando para a amostra 62 pontos).

Constata-se portanto que mesmo sendo o procedimento piloto a estratégia correta para definir o tamanho da amostra de referência, o uso de p pré-fixado de 0,85 não induziu a grandes desvios em E neste estudo, em que G estimado variou de 0,73 a 0,81, com exatidão pré-fixada de 0,85.

Comparando-se o dimensionamento tradicional com o dimensionamento por meio de semivariogramas verifica-se que o número de pontos independentes indicados pelos alcances foram bem superiores ao dimensionamento tradicional de 71 pontos a 95% de probab. com $E=0,1$. O Quadro 3 apresenta os resultados finais obtidos pelos três procedimentos.

QUADRO 3-Procedimentos para Dimensionamento de Amostras de Referência

	n	G	E
Dimens. c/ p a priori (p pré-fixado)	49	0,77	0,11
Dimens. c/ proced. piloto (p estimado)	71	0,75	0,10
Dimens. por semivariograma (> alcance)	163	0,76	0,07

O dimensionamento por semivariograma é obviamente mais trabalhoso, e não permite prever o erro de amostragem, mas uma vez determinado para uma área, nos informa o número mínimo de pontos necessários para cobrirmos a imagem de forma representativa, desde que estejam arranjados em grade sistemática.

CONCLUSÃO

A correlação espacial na imagem é anisotrópica, sendo necessária a aplicação de semivariogramas em diversas direções para obter alcance preciso.

O número de pontos independentes determinado por semivariogramas é maior para toda a cena comparado a padrões distintos da imagem

O dimensionamento de 71 pontos obtido pelo procedimento piloto resulta em pontos independentes na imagem VIC, desde que afastados no mínimo em 28 pixels.

O dimensionamento com p pré-fixado não provoca grandes tendências no Erro de amostragem quando a exatidão *priori* não se afastada muito da exatidão estimada pela amostragem piloto, mas o procedimento coerente com a teoria de amostragem é o uso de amostra piloto para dimensionar o tamanho da amostra.

AGRADECIMENTOS

Os autores agradecem ao Prof. Eupídio I. Fernandes Filho do Depto de Solos da Universidade Federal de Viçosa por introduzir o tema Geoestatística em nossa pesquisa, e ao Instituto Nacional de Pesquisas Espaciais - INPE, pela doação da cena utilizada neste estudo.

BIBLIOGRAFIA

- ATKINSON, P. M. On estimating measurement error in remotely-sensed images with the variogram. *International Journal Remote Sensing*, 18 (14), p.3075-3084, 1997.
- ATKINSON, P. M.; DUNN, R.; HARRISON, A. R. Measurement error in reflectance data and its implications for regularizing the variogram. *International Journal Remote Sensing*, 17 (18), p. 3735-3750, 1996.
- BRITES, R. S. *Verificação de exatidão em classificação de imagens digitais orbitais: efeitos de diferentes estratégias de amostragens e avaliação de índices de exatidão*. Viçosa: UFV, 101p. 1996.
- COHEN, W. B.; SPIES, T. A. ; BRADSHAW, G. A. . Semivariograms of digital imagery for analysis of conifer canopy structure. *Remote Sensing of Environment*, 34, p.167-178, 1990.
- CURRAN, P.J. The semivariogram in remote sensing: an introduction. *Remote Sensing of Environment*, 24, p.493-507, 1988.
- ENTZ, T.; CHANG, C. Evaluation of soil sampling schemes for geostatistical analyses: a case study for soil bulk density. *Canadian Journal Soil Science*, 71, 165-176, 1991.
- HAMAKAWA, P.J. *Variabilidade espacial de alguns componentes de produção de uma cultura de feijoeiro*. Piracicaba, 114p. 1991.
- LACAZE, B.; RAMBAL, S.; WINKEL, T. Identifying spatial patterns of Mediterranean landscapes from geostatistical analysis of remotely-sensed data. *International Journal Remote Sensing*, 15 (12) p.2437-2450, 1994.
- MA, Z.; REDMOND, R. L. Tau coefficients for accuracy assessment of classification of remote sensing data. *Photogrammetric Engineer and Remote Sensing*, 61 (4) p.453-439, 1995.
- McBRATNEY, A.B.; WEBSTER, R. How many observations are needed for regional estimation of soil properties? *Soil Science*, 135 (3), 177-183, 1983.
- MEYER, P. L. *Probabilidade: aplicações à estatística*, 2 ed. Rio de Janeiro: LTC-Livros Técnicos e Científicos Editora S. A., 426p. 1984.
- MIRANDA, F.P.; FONSECA, L. E. N.; CARR, J. R.; TARANIK, J. V. Analysis of JERS-1 (Fuyo-1) SAR data for vegetation discrimination in northwestern Brazil using the semivariogram textural classifier (STC). *International Journal Remote Sensing*, 17 (17), p.3523-3529, 1996.
- MIRANDA, F.P.; MACDONALD, J.A.; CARR, J. R. Application of the semivariogram textural classifier (STC) for vegetation discrimination using SIR-B data of Borneo. *International Journal Remote Sensing*, 13 (12), p.2349-2354, 1992.
- PREVEDELLO, B.M.S. *Variabilidade espacial de parâmetros de solo e planta*. Piracicaba, 166p. 1987.
- RAMSTEIN, G; RAFFY, M. Analysis of the structure of radiometric remotely-sensed images. *International Journal of Remote Sensing*, 10, p.1049-1073, 1989.

RICHARDS, J. A . *Remote sensing digital image analysis: an introduction*. Springer-Verlag Berlin Heidelberg, 281p. 1986.

VIEIRA, S. R.; HATFIELD, J. L.; NIELSEN, D.R.; BIGGAR, J.W. Geostatistical theory and application to variability of some agronomical properties, *Hilgardia*, 51 (3), p.1-75, 1983.