

# O emprego da geoestatística na determinação do tamanho “ótimo” de amostras aleatórias com vistas à obtenção de estimativas dos volumes dos fustes de espécies florestais em Paragominas, estado do Pará

Paulo Cerqueira dos SANTOS<sup>1</sup>, Antônio Cordeiro de SANTANA<sup>2</sup>, Paulo Luiz Contente de BARROS<sup>3</sup>, Joaquim Carlos Barbosa QUEIROZ<sup>4</sup>, Tiago de Oliveira VIEIRA<sup>5</sup>

## RESUMO

O objetivo do trabalho foi determinar o tamanho adequado de amostra para estimar o volume de fustes de espécies florestais de uma população de árvores a serem cortadas no sistema de manejo florestal da empresa Cikel Brasil Verde Madeiras - Pará. Utilizaram-se as metodologias da amostragem sistemática e do estimador geoestatístico da krigagem ordinária com simulação sequencial, respectivamente para a escolha das amostras e estimação dos volumes dos fustes das árvores. Os resultados mostraram que os métodos podem ser utilizados no cálculo dos volumes de fustes de árvores. Entretanto, o método da krigagem apresenta um efeito de suavização, tendo como consequência uma subestimação dos volumes calculados. Neste caso, um fator de correção foi aplicado para minimizar o efeito da suavização. A simulação sequencial indicativa apresentou resultados mais precisos em relação à krigagem, uma vez que tal método apresentou algumas vantagens, tal como a não exigência de amostras com distribuições normais e ausência de efeito de suavização, característico dos métodos de interpolação.

**PALAVRAS- CHAVE:** Krigagem ordinária, Resíduos lenhosos, Simulação seqüencial indicativa, Floresta amazônica.

## The use of geostatistics to determine the appropriate sample size in order to obtain stem volume estimates of tropical wood species

### ABSTRACT

The objective of this study was to determine the appropriate size sample to estimate the stem volumes stems of tree species from a population of trees to be cut in the forest management system of the timber company Cikel Brasil Verde Madeiras – Pará State, Brazil. The methodologies of systematic sampling and geostatistical kriging with sequential simulation were used, respectively, for the choice of samples and estimation of stem volumes of trees. The results showed that the methods can be used to calculate the stem volumes of trees. However, the kriging method has a smoothing effect, which resulted in an underestimation of the volumes. In such case, a correction factor was applied to minimize the effect of smoothing. The sequential simulation indicative presented more accurate results compared to kriging, since this method has certain advantages over kriging, such as not requiring samples with normal distributions and no smoothing effect characteristic of the interpolation methods.

**KEYWORDS:** Amazon, Ordinary Kriging, Sequential Indicator Simulation, Trees Stems Volume.

<sup>1</sup> Universidade Federal do Pará. e-mail: cerquera@ufpa.br

<sup>2</sup> Universidade Federal Rural da Amazônia . e-mail: acsantana@superig.com.br

<sup>3</sup> Universidade Federal Rural da Amazônia. e-mail: paulo.contente@ufra.edu.br

<sup>4</sup> Universidade Federal do Pará. e-mail: joaquimqz@hotmail.com

<sup>5</sup> Universidade Federal do Pará. e-mail: tiagovieira7@yahoo.com.br

## INTRODUÇÃO

A determinação do tamanho de amostras geralmente está ligada a algum objetivo de interesse principal, pois em função dessa amostra deseja-se conhecer as características fundamentais da população em estudo. Dessa forma, o tamanho da amostra deve ser tal que conecte a precisão, o erro máximo de amostragem permitido, o tamanho da população e propicie a redução dos custos operacionais. Neste sentido, diversos autores, tais como Stefanel *et al.* (1996), Marodim *et al.* (2000) e Lorentz *et al.* (2004), utilizaram a metodologia proposta por Cochran (1977), a qual faz a conexão dos elementos citados para determinar tamanhos de amostras utilizadas em seus estudos.

Em contrapartida, Guedes (2008) utilizou-se de uma metodologia conhecida como algoritmo de recozimento simulado (*annealing simulation*) e genético para obter uma solução ótima (tamanho ótimo da amostra) quando do dimensionamento de um conjunto de 256 pontos amostrais de teor de potássio. No entanto, foi utilizada neste trabalho a metodologia sugerida por Cochran (1977), face à sua simplicidade operacional, comparado a outras técnicas mais sofisticadas existentes na literatura para determinação do tamanho de amostras. Neste contexto, Mello (2004) indica que um dos principais objetivos dos inventários florestais é estimar o volume total de madeira existente num povoamento florestal, sendo que tal estimativa está diretamente ligada a uma amostra, retirada desse povoamento.

No presente estudo, que também priorizada a característica dendrométrica volume, utilizou-se o procedimento da amostragem sistemática para a escolha das unidades amostrais, em face da estrutura das informações constantes no inventário florestal. Dessa forma, este trabalho teve como objetivo principal determinar um tamanho adequado dos volumes de uma população 668 árvores selecionadas para corte que permitisse minimizar as discrepâncias entre os valores calculados por métodos de interpolação e/ou simulação e o valor real cubado quando da exploração da área em estudo.

## MATERIAL E MÉTODOS

### Área de estudo e coleta dos dados

A área de estudo fica numa Unidade de Trabalho da Fazenda Rio Capim, com uma área de 140.658 ha, pertencente à empresa madeireira Cikel Brasil Verde Madeiras Ltda. na cidade de Paragominas e dista cerca de 320 km da capital Belém. A área de estudo possui basicamente as mesmas características apresentadas pelo município, em relação aos índices pluviométricos, topografia variando de plana a suavemente ondulada e florestas do tipo ombrófila densa submontana. A coleta de dados foi realizada na Unidade de

Trabalho nº 18 (UT 18), com 98 ha, localizada na Unidade de Produção Anual 10 (UPA 10) que é composta por 2.385 ha. Essa área era composta por 1161 árvores com DAP  $\geq$  45 cm, distribuídas em 54 espécies, sendo que 668 dessas árvores foram selecionadas para corte e 493 deixadas como estoque de reserva.

### Análises dos Dados

Dessas 668 árvores foram retiradas 17 amostras aleatórias (Tabela 1) para a variável volume dos fustes, sendo fixado previamente, para cada uma delas, um nível de significância 95% e erro amostral variando com o tamanho da amostra. O tamanho da amostra foi determinado por meio da expressão devido a Cochran (1977).

$$n = \frac{Nz^2\hat{p}\hat{q}}{(N-1)\varepsilon^2 + z^2\hat{p}\hat{q}} \quad (1)$$

Em que,

N = é o tamanho da população;

$z^2$  = abscissa da distribuição normal padrão para um nível de confiança previamente fixado (95%);

$\hat{p}$  = estimador de p, em que p representa a chance do elemento amostral pertencer ou não à amostra;

$\hat{q} = 1 - \hat{p}$ ;

$\varepsilon$  = erro amostral.

Tal expressão conecta o tamanho da população N, o nível de confiança e o erro máximo amostral permitido. Além disso, considerou-se que a população possui uma distribuição de Bernoulli com  $\hat{p}$  e  $\hat{q}$  máximos, ou seja,  $\hat{p} = \hat{q} = 0,5$ .

A partir dessas amostras, procurou-se investigar, através do estimador geoestatístico de krigagem ordinária (KO) e da simulação geoestatística indicativa, qual o tamanho de amostra produziria a melhor estimativa para o volume total dos fustes da área em estudo. Assim, os valores das estimativas (em função das amostras) para os volumes dos fustes foram comparados com o valor real cubado que foi de 3.408,6216 m<sup>3</sup>. Vale ressaltar que o volume final estimado, VFE, é formado por duas parcelas: a parcela devido à estimativa nos pontos não amostrados (VE) e a parcela formada pelo volume dos pontos amostrados (VA), ou seja, VFE=VE+VA.

Para a utilização das técnicas geoestatísticas, foi necessário avaliar o grau de dependência espacial (DE) para as amostras, em que essa dependência foi definida através da variação estruturada em relação ao patamar, sugerido por Biondi *et al.* (1994). Dessa forma, foram construídos para cada amostra os semivariogramas para as características dendrométricas Área Basal, Volume do Fuste, Altura Total e Diâmetro a Altura do Peito-DAP, a fim de se avaliar o grau de dependência espacial do povoamento florestal em questão.

**Tabela 1-** Amostras de volumes, segundo suas médias, erro padrão da média (EPM), coeficiente de variação (CV), volume e proporção em relação ao total.

Variável	Média das amostras	EPM	CV	Volume da Amostra (m3)	% da amostra em relação ao total
V267	4,848	0,240	80,93	1.294,314	40
V233	5,402	0,305	86,45	1.264,158	35
V200	4,970	0,281	80,07	994,0280	30
V187	4,991	0,242	66,41	933,3795	28
V174	5,034	0,256	67,09	875,8332	26
V160	4,913	0,279	71,91	786,1433	24
V147	4,925	0,335	82,48	723,9534	22
V134	5,094	0,309	70,15	682,6160	20
V120	5,187	0,373	78,84	622,4241	18
V107	4,657	0,285	63,35	498,2899	16
V94	5,846	0,651	107,93	549,5207	14
V80	5,351	0,385	64,33	428,0868	12
V67	4,592	0,369	65,79	307,6771	10
V53	4,837	0,594	89,45	256,3816	8
V40	4,408	0,400	57,44	176,3021	6
V27	4,561	0,482	54,86	123,1542	4
V13	4,549	0,709	56,23	59,13459	2

Fonte: Base de dados da CIKEL Brasil Verde Madeiras Ltda.

As principais técnicas utilizadas neste estudo são descritas a seguir, com base nas abordagens feitas por Mello (2004), Landim (1998) e Goovaerts (1997).

#### a) Semivariograma

O semivariograma mostra o grau de dependência entre amostras ao longo de um suporte específico para as chamadas variáveis regionalizadas, representadas por  $Z(\mathbf{u})$  em que  $\mathbf{u}$  indica a localização da amostra, coletadas em diversos pontos que podem estar regularmente distribuídos ou não. Caso haja a dependência espacial, o valor de cada ponto deverá apresentar um relacionamento, de algum modo, com valores obtidos no entorno deste ponto e a influência dos mesmos será tanto maior quanto menor forem as distâncias que os separam. Assim, considerando-se o vetor de distâncias  $\Delta\mathbf{h}$  (para uma distância  $h$ ) com uma direção específica, o grau de relação entre os pontos podem ser expressos pela função covariância ou semivariograma. Para análise da dependência espacial, a função semivariograma é mais utilizada em geoestatística devido apresentar algumas vantagens em relação à covariância. Dessa forma, o semivariograma para valores encontrados ao longo da distância  $\mathbf{h}$ , separados por  $\Delta\mathbf{h}$  é dado por:

$$K(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} [z(\mathbf{u}_{\alpha}) - z(\mathbf{u}_{\alpha} + \mathbf{h})]^2 \quad (2)$$

Para Queiroz (2003), os semivariogramas expressam o comportamento das variáveis regionalizadas ou de seus resíduos, e mostram ainda:

- o tamanho da zona de influência em torno da amostra;
- a anisotropia, quando o comportamento dos semivariogramas se mostram diferentes para várias direções da linha de amostragem;
- a continuidade, pela forma do variograma, que para  $h = 0$ ,  $\gamma(h) = \frac{1}{2n} \sum \{Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})\}^2$ , já apresenta algum valor. Tal situação é conhecida como *efeito pepita (nugget)* e é representada por  $C_0$ .

Para a modelagem do semivariograma, realizado depois de construído o semivariograma experimental, é necessário que a hipótese de estacionariedade seja considerada. Esta hipótese estabelece, em resumo, que os dois primeiros momentos (média e variância) da diferença  $[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]$  sejam independentes da localização  $\mathbf{u}$  e função somente do vetor  $\mathbf{h}$ . O segundo momento dessa diferença corresponde ao variograma, ou seja:

$$2g(\mathbf{h}) = E\{[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]^2\} \text{ é independente de } \mathbf{u} \hat{=} \text{ a área } \mathcal{A} \quad (3)$$

Os modelos básicos permissíveis de semivariogramas mais utilizados em geoestatística são apresentados a seguir:

#### Modelo esférico

$$\gamma(h) = C \left\{ \left( \frac{3}{2} \right) \left( \frac{h}{a} \right) - \left( \frac{1}{2} \right) \left( \frac{h^3}{a^3} \right) \right\}, \text{ para } h < a \quad (4)$$

$$\gamma(h) = C, \text{ para } h \geq a$$

Este modelo é o mais comum e, por esse motivo, é comparado à função de distribuição normal da estatística clássica (Landim 1998).

#### Modelo exponencial

$$\gamma(h) = C \left( 1 - e^{-\frac{h}{a}} \right) \quad (5)$$

#### Modelo Gaussiano

$$\gamma(h) = C \left( 1 - e^{-\frac{h^2}{a^2}} \right) \quad (6)$$

Onde o parâmetro  $C$  é conhecido como *patamar* ou “*sill total*” e representa o nível de variabilidade do semivariograma até a sua estabilização, enquanto o parâmetro  $a$  é conhecido como *alcance* e representa a distância observada até o nível onde a variabilidade se estabiliza. O parâmetro  $a$  indica ainda a distância que as amostras estão correlacionadas espacialmente.  $C_0$  é o que se conhece como *efeito pepita* e está relacionado a erros de amostragem, medidas, etc.  $C_1$  representa o sill parcial, de modo que  $C = C_0 + C_1$ .

Neste estudo foram testados os modelos teóricos esférico, exponencial e gaussiano para se avaliar qual deles melhor se ajustava aos dados de volumes dos fustes obtidos através das amostras para, posteriormente, ser utilizado como estimador do volume dos pontos não amostrados. A avaliação do desempenho de cada um destes modelos se deu através do critério de Akaike (Akaike’s Information Criterion - AIC), determinado através da função lik.GRF do software ‘geoR’ v. 1.6-25. Este critério é dado pela seguinte expressão:

$$AIC = -2\log L + 2K \quad (7)$$

onde  $L$  é a Verossimilhança maximizada pelo modelo candidato e  $K$  o número de parâmetros deste modelo.

#### b) Krigagem

Segundo Mello (2004) a krigagem é um método de inferência espacial que estima informações em pontos não amostrados a partir de informações situadas em pontos amostrados, considerando a estrutura de dependência espacial do fenômeno. Existem vários tipos de krigagem (Isaak e Shrivastava 1989). Dentre eles pode-se citar a krigagem Simples (KS), Ordinária (KO), Universal, Blocos, dentre outros.

Neste estudo foi utilizado a Krigagem Ordinária (KO) por levar em consideração a variação local da média, limitada

ao domínio de estacionariedade da mesma à vizinhança local  $W(u)$ , centrada sobre a localização  $u$  a ser estimada. Em outras palavras, a krigagem ordinária utiliza médias locais ou tendências locais estimadas a partir dos elementos amostrais vizinhos, ao invés de uma única média estacionária, como o faz o algoritmo de interpolação simples (Imai *et al.* 2003).

#### c) Simulação geoestatística

A partir do modelo de dependência espacial pode-se direcionar o problema da estimação, usando-se krigagem, ou fazendo a simulação de um valor desconhecido a partir dos valores disponíveis. A simulação condicional é uma aproximação estocástica para predição que tenta reproduzir as condições reais em vez de estimá-las em termos de médias e variabilidades (Goovaerts 1997). As aproximações produzidas pela simulação condicional exibem as mesmas características amostrais apresentadas pelo histograma e semivariogramas amostrais. Entretanto, no caso da simulação condicional, ao contrário da krigagem, que é um método de interpolação baseado em médias ponderadas e que, por isso, apresenta um efeito de suavização, o grau de variabilidade dos valores simulados é bem próximo do medido pelos valores amostrais, geralmente resultando em maiores flutuações entre os valores adjacentes, ou seja, sem o efeito da suavização.

Em aplicações ambientais onde, com muita frequência se encontram distribuições que apresentam fortes assimetrias, deve-se procurar fugir da exigência de hipóteses Gaussianas (distribuições normais) na análise dos dados. Entre os principais algoritmos de simulação que não são limitados ao formalismo Gaussiano temos a Simulação Sequencial Indicativa e a Simulação *Annealing*. Neste estudo, pelas características apresentadas pelos conjuntos de dados em estudo, forte assimetria, conduziram à escolha do método da simulação indicativa, descrita a seguir.

A simulação indicativa é uma técnica de simulação não Gaussiana bastante utilizada em geoestatística. Neste caso, o valor da função de distribuição acumulada condicional, FDAC,  $F(\mathbf{u}; z_k | (n))$ , é determinado por uma série de valores de corte  $z_k$  discretizando a faixa de variação de  $z$ , ou seja,

$$F(\mathbf{u}; z_k | (n)) = \Pr \{ Z(\mathbf{u}) \leq z_k | (n) \}, \quad (8)$$

$$k = 1, \dots, K$$

onde  $\mathbf{u}$ , representa as coordenadas das localizações não amostradas,  $z$ , é o valor estimador da variável em estudo e  $n$  é o tamanho da amostra.

Esse formalismo é utilizado para modelar a seqüência de FDACs a partir das quais valores simulados são obtidos. O procedimento usual da simulação sequencial é o seguinte (Goovaerts 1997):

i) Inicialmente é feita a discretização da variável utilizando-se  $(K + 1)$  classes para  $K$  valores de corte  $z_k$ , selecionados de

acordo com os valores das variáveis. É comum a utilização de decis ou quartis para os pontos de corte. Então são obtidas variáveis indicativas, definidas como:

$$i(u_\alpha, z_k) = \begin{cases} 1 & \text{se } z(u_\alpha) \leq z_k \\ 0 & \text{caso contrário} \end{cases} \quad (9)$$

ii) É construída uma malha na área de estudo onde os valores a serem simulados se localizam nos cruzamentos das linhas da malha (nó da malha). Então é definido um caminho aleatório para visitar cada nó da malha de cada vez

iii) Em cada cruzamento

- Determinar os  $K$  valores da FDAC utilizando-se qualquer algoritmo de krigagem indicativa.
- Corrigir desvios e completar o modelo da FDAC, utilizando-se algoritmos de interpolação/extrapolação para as partes extremas (caudas) da função.
- Selecionar um valor simulado  $z(\mathbf{u})$  daquela FDAC.
- Adicionar o valor simulado ao conjunto de dados condicionantes (condicionados aos dados amostrais).
- Ir para o próximo nó da malha ao longo do caminho aleatório e repetir os passos acima (de *a* a *d*).

iv) Repetir o procedimento com um caminho aleatório diferente para gerar outra realização ou simulação. Isso é feito utilizando-se diferentes números, ou *seeds*.

As simulações foram realizadas utilizando-se o programa *sisim* da biblioteca GSLIB (Deusch e Journel 1996). Uma medida quantitativa e visual da incerteza espacial é fornecida pela geração de muitas simulações que se ajustem razoavelmente às mesmas amostras estatísticas (histograma e variograma) e dados condicionantes.

## RESULTADOS E DISCUSSÃO

Este estudo se inicia com a análise descritiva dos dados com a forma de visualizar o comportamento das amostras individualmente. O pressuposto é que as amostras, individualmente, guardem as mesmas características da população que as originaram.

### Análise Exploratória dos Dados

Foram utilizadas 17 amostras aleatórias de diversos tamanhos, selecionadas segundo um procedimento sistemático. Esse procedimento amostral permitiu uma boa representação espacial dos conjuntos amostrais para a área de estudo. O *Box plot* conjunto das 17 amostras mostrou a presença de valores atípicos (ou “outliers”) em todas as amostras, fato este que pode ser uma característica das florestas tropicais (Figura 1). Segundo Mello (2004) estes valores influenciam negativamente o semivariograma experimental, no sentido de produzir estimativas irreais. Deste modo, para a aplicação

da krigagem, optou-se por se fazer uso da transformação logarítmica para minorar os efeitos da presença de tais valores.

A Tabela 1 mostra o comportamento das 17 amostras em função das suas médias, erro padrão da média (EPM), coeficiente de variação (CV), volume dos fustes e proporção de cada uma delas em relação ao total das árvores (668) em estudo. Tais amostras apresentam uma amplitude que varia de 2% a 40% da população, sendo que as três maiores (V267, V233 e V200), possuem uma taxa de decaimento de 5% e, para as demais, a taxa de redução é 2% até o limite em que a amostra representa 2% da população (V13). Também é importante destacar que a maior amostra com 1.294,314 m<sup>3</sup>, concentra 22 vezes mais volume de madeira do que a menor amostra com 59,3816m<sup>3</sup>. Além do mais, o coeficiente de variação foi substancialmente mais baixo (56,23% da menor amostra contra 80,93% da maior amostra).

### Os Modelos Semivariográficos Estimados para a Krigagem

Foram ajustados os modelos esférico, exponencial e gaussiano aos semivariogramas experimentais das 17 amostras, através dos três métodos de ajuste: Mínimos Quadrados Ordinários (MQO), Mínimos Quadrados Ponderados (MQP) e Máxima Verossimilhança (MV). O critério de Akaike foi utilizado para a escolha do melhor modelo, sendo que o modelo teórico exponencial foi o que apresentou o melhor ajuste com 58,82% dos casos, seguido do modelo esférico com 23,53%. O modelo gaussiano foi ajustado em apenas 17,65% dos casos. Em contrapartida, o método de ajuste dos mínimos quadrados ponderados (65% dos casos) e mínimos quadrados ordinários (35% dos casos) foram os métodos mais consistentes no ajuste dos dados, enquanto o método da máxima verossimilhança se mostrou inadequado para ajustar esses dados aos modelos teóricos, em 100% dos casos. É importante frisar que a aplicação deste método requer a normalidade dos dados e, assim sendo, poderá ser a principal causa da inadequação do método, uma vez que nenhuma das amostras apresentou distribuição normal.

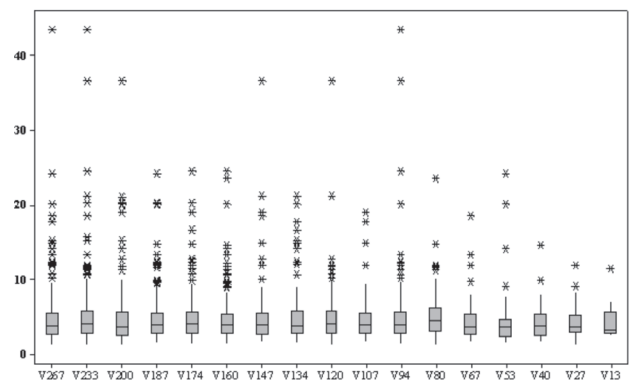


Figura 1 - Box plot conjunto mostrando a presença de “outliers” para as 17 amostras.



A Tabela 2 mostra os modelos teóricos escolhidos e a estimativa dos parâmetros efeito pepita ( $C_0$ ), variação estruturada ou *Sill* parcial ( $C_1$ ), alcance ( $a$ ), patamar ou *Sill* total ( $C$ ) e dependência espacial ( $DE$ ). A verificação do grau de dependência espacial ( $DE\%$ ) das amostras foi realizada através da razão entre a variação estrutura e o patamar. Essa razão, proposta por Cambardella *et al.* (1994) quando de estudos realizados com atributos de solo, preconiza que variáveis com níveis de dependência espacial inferiores a 30%, sejam considerada fraca, enquanto que variáveis com níveis de dependência entre 30 e 70% sejam consideradas dependentes espacial, moderadamente. Por outro lado, as variáveis com  $DE\%$  maiores ou iguais a 75%, sejam consideradas fortemente dependentes espaciais. Por esse critério, os resultados deste estudo mostraram que 52,94% (9) das amostras apresentaram baixa dependência espacial enquanto as restantes 47,06% (8) apresentaram dependência espacial variando de mediana a alta. O critério devido a Cambardella *et al.* (1994) também foi relatado por Biondi *et al.* (1994), quando estudaram o comportamento espacial do diâmetro a altura do peito ( $DAP$ ) e área basal em floresta natural de *Pinus ponderosa*.

É importante destacar que o grau de dependência espacial observado nas amostras é compatível com o grau dependência observado na população como um todo, sendo que, geralmente, as amostras com os mais altos níveis de dependência espacial correspondem àquelas amostras com tamanhos que representam menos de 14% da população investigada. Ressalte-se ainda que duas das amostras, V120 e V13, apresentaram  $DE\%$  iguais a zero e, dessa forma, sugerem a inconsistência da aplicação do estimador geoestatístico de krigagem ordinária (KO) aos dados. No entanto, mesmo considerando essa limitação, os valores finais das estimativas realizadas sob tal condição, diferem muito pouco daqueles realizados com amostras com forte dependência espacial ( $DE\%$ ).

Foram ajustados os semivariogramas para cada amostra, segundo o método mais adequado de ajustamento onde se constatou que os semivariogramas correspondentes às amostras de tamanhos variando entre 107 e 267 fustes, provieram dependência espacial ( $DE\%$ ) considerada de baixa a moderada, quando atingiu um valor máximo de 58%. No entanto, quando o tamanho das amostras vai se reduzindo

Tabela 2 - Parâmetros estimados para cada amostra, segundo o método de ajuste.

Amostra	Método de ajuste	Esférico				Exponencial				Gaussiano			
		Co	C1	a	DE(%)	Co	C1	a	DE(%)	Co	C1	a	DE(%)
V267	MQO					0,202	0,107	50	34				
V233	MQO									0,308	0,006	35	2
V200	MQP					0,224	0,147	121,7	40				
V187	MQP									0,129	0,177	46,4	58
V174	MQP	0,27	0,0356	258,74	12								
V160	MQO					0,129	0,078	40	26				
V147	MQP					0,184	0,113	35	38				
V134	MQP					0,282	0,061	93,65	18				
V120	MQP					0,302	0	14,99	0,0				
V107	MQP									0,260	0,019	62	7
V94	MQP					0,386	17,31	158	97,8				
V80	MQP	0,24	0,058	175	20								
V67	MQO	0	0,2763	32,61	100								
V53	MQO					0	0,372	30,78	100				
V40	MQO	5,84	0,9627	208,52	14								
V27	MQP					0,15	13,11	200	98,9				
V13	MQP					0,236	0	68,89	0				

Fonte: Levantamento de campo – UT18-UPA10 – Rio Capim.

(amostras de tamanhos variando de 13 a 94 fustes), já se percebe graficamente a prevalência de amostras (V94, V67, V53 e V27) com DE% satisfatória. Mesmo assim, as amostras de tamanhos 13 (V13), 40 (V40) e 80 (V80), apresentaram baixa dependência espacial. É importante frisar que este grupo contém 02 (duas) das três amostras que geraram as melhores estimativas do volume total dos fustes da área em estudo.

### Estimativas do Volume dos Fustes pela Krigagem

Os volumes dos fustes foram estimados através do estimador geoestatístico de krigagem ordinária (KO), em que os valores amostrais fornecidos pelas 17 amostras, funcionaram como os pontos amostrais conhecidos e possibilitou o cálculo das estimativas nos pontos não amostrados. Percebe-se que os valores das estimativas dos volumes a partir das amostras, tendem a subestimar o valor do volume real cubado. Esta subestimação pode ser em função de uma das características do estimador geoestatístico de krigagem ordinária, uma vez que o mesmo pode ser entendido como uma média dos pontos a serem estimados e dessa forma podem ser influenciados por valores extremos, grandes ou pequenos. Tal característica leva a uma superestimação de valores pequenos e uma subestimação dos valores grandes.

Todas as amostras, exceto a amostra V53, produziram estimativas acima de 80% do volume real. Em particular as amostras V233, V94 e V80, apresentaram estimativas onde as discrepâncias entre os valores estimados e o valor real medido não superam 8%.

Diante dos resultados obtidos, verificou-se a necessidade de se introduzir um fator de correção, como forma de amenizar as discrepâncias vislumbradas entre os valores estimados e o valor real cubado. Este fator de correção foi definido com base na variabilidade observada para cada variável, após a junção dos valores dos volumes correspondente à amostra e os valores estimados (fora da mostra). Assim, o Volume Final Estimado (VFE) pode ser expresso em função do Valor Estimado, acrescido desse Fator de Correção, como a seguir:

$$VFE = VE + 100 \sigma \text{ ou } VFE = VE + \bar{X} * CV \quad (10)$$

em que,

VFE= Volume Final Estimado;

VE= Volume estimado através da krigagem Ordinária;

$\sigma$ = Desvio padrão dos dados aglutinados (amostrados e estimados);

$\bar{X}$  = Média aritmética;

CV= Coeficiente de variação.

A Figura 2 compara, em termos de valores absolutos, os resultados obtidos para os valores cubados, estimados sem correção e estimados com correção. Observa-se que a

introdução do fator de correção proposto neste trabalho, provocou uma redução significativa dos níveis de discrepâncias entre os valores estimados e os valores cubados, notadamente naquelas amostras que já haviam apresentado as melhores estimativas (V233, V94 e V80). No caso de V233, a redução na discrepância foi de 114,242 m<sup>3</sup> enquanto que V94 e V80 de 245,622 m<sup>3</sup> e 129 m<sup>3</sup>, respectivamente.

Por fim, o estudo mostrou ser possível determinar o tamanho de amostras através do estimador geoestatístico de krigagem ordinária (KO), uma vez que através das estimativas de volumes obtidas, pode-se determinar um intervalo de confiança que contenha o melhor tamanho de amostra “n” que seja capaz de conectar precisão com a redução dos custos operacionais.

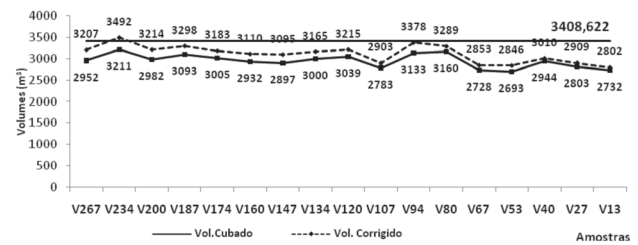


Figura 2 - Estimativas do volume total dos fustes com e sem correção. Fonte: Base de dados da CIKEL Brasil Verde Madeiras Ltda – UT 18

No presente estudo, os resultados mostraram que esse intervalo de confiança pode variar de um tamanho mínimo de “n” da ordem de 4% até um valor máximo de 11,9958% do total da população, passando por um valor médio da ordem de 8%. Acima destes valores, apesar do aumento da acurácia, certamente aumentará os custos operacionais e, dessa forma, não é aconselhável a utilização de valores de “n” que ultrapassem 12% da população. Outro fato relevante a ser destacado, são os resultados preliminares que já foram obtidos através de estudos realizados em três outras UTs (Unidades de Trabalho): UT01, UT07 e UT31 (Figura 4), localizadas na mesma área de estudo. Os resultados corroboram com as evidências observadas neste trabalho.

### Simulação sequencial Indicativa do Volume dos fustes

Na simulação indicativa do volume dos fustes, para o modelamento do padrão espacial foram utilizados semivariogramas indicativos, em que foram considerados como pontos de corte, os valores da variável em todas as amostras nos percentis  $Z_{15}$ ,  $Z_{25}$ ,  $Z_{50}$ ,  $Z_{75}$  e  $Z_{85}$ . Foram utilizadas somente 11 amostras (Z267 a Z94) na simulação sequencial indicativa. Em todos os casos, utilizou-se o modelo esférico. Todavia, em várias situações, houve a necessidade de ajustar dois modelos esféricos, indicados pelos alcances  $a_1$  e  $a_2$  e

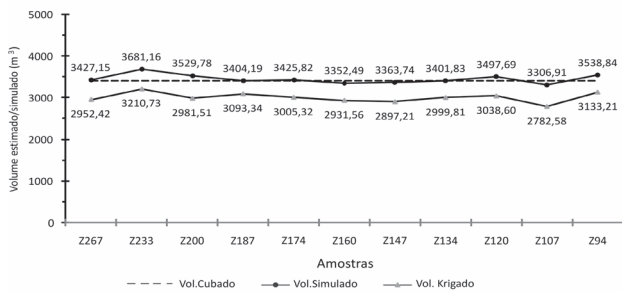


Figura 3 - Volume real dos fustes cubados, estimados por krigagem e simulados. Fonte: Base de dados da CIKEL Brasil Verde Madeiras Ltda – UT 18.

patamares  $C_1$  e  $C_2$ . Observa-se, por exemplo, para o ponto de corte nos centis  $Z_{15}$  uma correlação espacial que varia aproximadamente entre 25 a 140 m, dependendo do tamanho da amostra. O maior alcance (195 m) foi observado na amostra Z147 para o ponto de corte em  $Z_{75}$ .

As simulações foram realizadas utilizando-se o algoritmo de simulação seqüencial indicativa, SISIM (Deutsch e Journel 1998). As simulações foram condicionais às amostras selecionadas e reproduziram satisfatoriamente as estatísticas de primeira e segunda ordem, ou seja, histograma e variograma amostrais.

A Figura 3 mostra os resultados obtidos para os conjuntos de amostras utilizados. Pode-se observar que, as simulações não sofrem o efeito da suavização observada na krigagem. As simulações para as amostras de tamanhos entre 187 e 134, cujos erros amostrais ficaram entre 0,534 e 0,636, respectivamente, apresentaram valores muito próximos do valor cubado de 3408,22 m<sup>3</sup>. Observa-se ainda que o comportamento dos valores simulados e estimados por krigagem é praticamente o mesmo, confirmado pela forte correlação entre os valores simulados e estimados por krigagem, de 0,837.

## CONCLUSÕES

O grau dependência espacial (DE%) verificado nas amostras sugere um padrão que lembra muito um padrão aleatório, face aos valores apresentados. Tais valores indicaram baixa correlação espacial entre as amostras o que pode indicar que tais indivíduos são distribuídos aleatoriamente na floresta. Este fato influencia negativamente a precisão das estimativas.

As estimativas realizadas através da Krigagem Ordinária (KO) necessitam de um fator de correção como forma de amenizar os efeitos da subestimação verificada quando da aplicação dessa técnica.

Os resultados mostraram que a utilização do estimador geoestatístico de krigagem ordinária (KO) foi adequada para

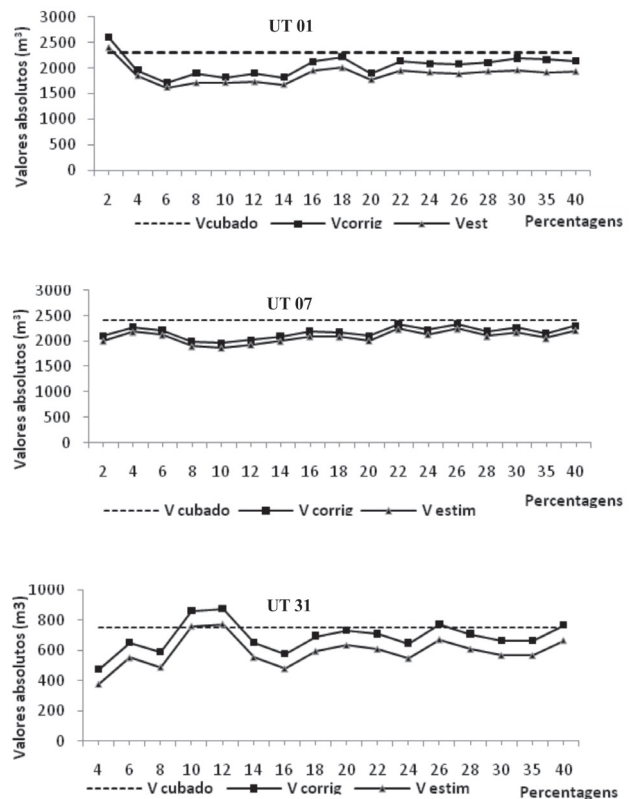


Figura 4 - Comparativo entre os volumes estimados/corrigidos com o volume cubado, segundo as UT 01, 07 e 31. Fonte: Base de dados da CIKEL Brasil Verde Madeiras Ltda.

fazer estimativas a partir de pequenas amostras. As estimativas obtidas através da maior amostra (V267) e da menor amostra (V13) divergem em apenas 6% ao mesmo tempo em que V13 é 20 vezes menor que V267.

Este estudo mostrou ainda ser plausível a adoção de tamanhos de amostras variando entre 4% e 12% da população, mesmo com a presença dos valores atípicos. O uso da simulação indicativa mostrou-se superior à krigagem para o cálculo do volume dos fustes das árvores. Isso ocorreu, provavelmente devido às características da simulação indicativa que pode ser utilizada para amostras com qualquer distribuição.

Finalmente, sugere-se que os resultados obtidos para UT 18 da UPA 10 sejam confrontados com outros resultados provenientes de outras UTs da mesma área (UPA), como no caso das UTs 01, 07 e 31, ou de UPAs diferentes, tendo em vista que o intuito é consolidar o método, diante do pioneirismo de sua aplicação em florestas da Amazônia e das características dos resultados encontrados com grande número de valores atípicos nos diversos tamanhos de amostras.



## BIBLIOGRAFIA CITADA

- Biondi, F.; Myers, D.E.; Avery, C.C. 1994. Geostatistically modeling stem size and increment in an old-growth forest. *Canadian Journal of Forest Research*, 24(7): 1354-1368.
- Cambardella, C.A.; Moorman, T.B.; Novak, J.M.; Parkin, T.B.; Karlen, D.L.; Turco, R.F.; Konopka, A.E. 1994. Field scale variability of soil properties in central Iowa soils. *Soil Science Society of America Journal*, 58(5): 1501-1511.
- Cochran, W.G. 1977. Sampling techniques. 3.ed. New York: Wiley. 555 pp.
- Deustch, C.V.; Journel, A.G. 1998. GSLIB: *Geostatistical Software Library and User's Guide*. Oxford University Press, New York. 369 pp.
- Estefanel, V.; Guedes, J.V.C.; Manara, V. 1996. Sample size for evaluation of yield components in common beans. *Ciência Rural*, 26(3): 32-39.
- Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York. 483 pp.
- Guedes, L.P.C.; Ribeiro Júnior, P.J.; Piedade, S.M.S.; Opazo, M.A.U. Optimization of spatial sampling using hybrid genetic algorithm. 2008 ([www.scholar.google.com.br](http://www.scholar.google.com.br)). Acesso em 1/09/2009.
- Imai, N.N.; Tachibana, V.M.; Vicente, J.; Lima, D.L.; Silva, E.A.; Voll, E.; Oliveira, H.E.S. 2003. Comparative analysis of kriging ordinary and kriging indicator in the case of weeds in soybean. Universidade Estadual Paulista (UNESP). Faculdade de Ciências e Tecnologia – Departamento de Cartografia, 2003 ([www.cartografia.org.br](http://www.cartografia.org.br)). Acesso em 19/05/2009.
- Isaaks, E.H.; Shrivastava, R.M. 1989. *An introduction to applied geostatistics*. Oxford University Press, New York.
- Landim, P.M.M. 1998. *Statistical analysis of geological data*. São Paulo: Fundação Editora da UNESP, (Ciência e Tecnologia) (In portuguese).
- Lorentz, L.H.; Lúcio, A.D.; Storck, L.; Lopes, S.J.; Boligon, A.A.; Carpes, R.H. 2004. Temporal variation of the sample size for the experiment in plastic greenhouses. *Ciência Rural*, 34(4): 1043-1049.
- Marodim, V.S.; Storck, L.; Lopes, S.J.; Santos, O.S.; Schmidt, D. 2000. Experimental design and sample size for lettuce grown in hydroponics. *Ciência Rural*, 30(5): 56-66. (In Portuguese).
- Mello, J.M. 2004. *Geostatistics applied to forest inventory*. Tese de Doutorado, Escola Superior Luiz de Queiroz, Piracicaba-SP. 111 pp. (In portuguese).
- Queiroz, J.C.B. 2003. *Use of geostatistics to quantify the risk of contamination by heavy metals in the port Santana-Amapá*. Tese de Doutorado, Universidade Estadual Paulista. Instituto de Ciências Exatas, Rio Claro, São Paulo. 200 pp. (In portuguese).

Recebido em 03/03/2010

Aceito em 06/07/2010

