

COMPARAÇÃO ENTRE PROCEDIMENTOS DE AMOSTRAGEM PARA ESPÉCIES FLORESTAIS RARAS E PADRÃO DE DISTRIBUIÇÃO ESPACIAL AGREGADO¹

Carlos Pedro Boechat Soares², Christian Marques Rodello³, Agostinho Lopes de Souza², Helio Garcia Leite², Vicente Paulo Soares² e Gilson Fernandes da Silva⁴

RESUMO – O objetivo deste trabalho foi comparar procedimentos de amostragem para espécies florestais com populações raras e padrão de distribuição espacial agregado. Para isso, foi simulada uma população em uma floresta de 90 ha, subdividida em 100 unidades amostrais (N = 100) de 9.000 m² de área cada uma, apresentando número total de indivíduos igual a 44, a qual foi submetida a três procedimentos de amostragem: amostragem casual simples; amostragem sistemática; e amostragem adaptativa em *cluster*, com amostras iniciais selecionadas casual ou sistematicamente. Após as análises, verificou-se que a amostragem sistemática foi o melhor procedimento para estimar o número total de indivíduos da espécie em questão. Além disso, constatou-se a necessidade de investigar o efeito do tamanho e forma de parcelas, a escala de agregação e o tamanho da população, bem como suas combinações, sobre a eficiência dos estimadores da amostragem adaptativa em *cluster*.

Palavras-chave: *Amostragem sistemática, amostragem aleatória e amostragem adaptativa em cluster.*

COMPARISON BETWEEN SAMPLING PROCEDURES FOR RARE FOREST SPECIES AND AGGREGATED SPATIAL DISTRIBUTION PATTERN

ABSTRACT – *The objective of this work was to evaluate the procedure adaptive sampling procedure used in forest inventories for species with low population density and aggregated spatial distribution pattern, in relation to procedures normally used in forest inventories. Thus, a population was simulated in a 90 ha forest, subdivided into 100 units (N = 100) with areas of 9,000 m² each, presenting 44 individuals of a given species, and submitted to three sampling procedures: simple random sampling; systematic sampling and adaptive cluster sampling, with initial samples selected through simple random sampling and systematic sampling. After analyses, it was verified that the systematic sampling was the best procedure to estimate the total number of individuals of the studied species. Besides, it was observed the need of further investigation on the effect of the size and form of plots, the aggregation scale and size of the population, and their combinations, on the efficiency of the estimators of the adaptive cluster sampling.*

Keywords: Systematic sampling, random sampling and adaptive cluster sampling.

1. INTRODUÇÃO

O problema básico de qualquer procedimento de amostragem é a obtenção de estimativas fidedignas de alguma característica da população de interesse, tomando como base somente parte dessa população (SOUZA et al., 2006; THOMPSON e SEBER, 1996).

No contexto da Ciência Florestal, pode-se citar como exemplo a necessidade de obter a estimativa do

número de indivíduos de uma espécie que exibe um padrão de distribuição espacial altamente agregado e, ainda, que esses agregados se apresentam de forma esparsa em grande área na floresta, indicando raridade geográfica ou espacial dessa espécie (KALTON e ANDERSON, 1986; GOSH e INNES, 1996).

Um procedimento de amostragem que vem sendo estudado e testado em levantamentos de populações

¹ Recebido em 30.01.2007 e aceito para publicação em 24.04.2009.

² Departamento de Engenharia Florestal da Universidade Federal de Viçosa (UFV). E-mail: <csoares@ufv.br>; <alsouza@ufv.br>.

³ Programa de Pós-Graduação em Ciência Florestal da UFV. E-mail: <christianfloresta@gmail.com>.

⁴ Centro de Ciências Agrárias da Universidade Federal do Espírito Santo (UFES), Alegre-ES. E-mail: <gfsilva@cca.ufes.br>



de espécies vegetais raras e que exibe padrão de distribuição espacial agregado é a amostragem adaptativa e, entre os procedimentos de amostragem adaptativa, a amostragem adaptativa em *cluster* tem-se destacado nos levantamentos de populações que se conformam segundo o padrão descrito anteriormente devido à sua operacionalidade (THOMPSON e SEBER, 1996; BROWN, 2003).

A amostragem adaptativa em *cluster* refere-se ao procedimento de amostragem em que uma amostra inicial é selecionada de forma aleatória ou sistemática, e enquanto essas unidades de amostra oriundas dessa pré-seleção contemplam o atributo a ser mensurado, de acordo com um critério de inclusão estabelecido *a priori*, e novas unidades de amostra situadas na vizinhança dessas unidades de amostra pré-selecionadas serão incluídas na amostragem, formando redes de unidades de amostra (THOMPSON, 1990; THOMPSON e SEBER, 1996; BROWN, 2003).

Assim sendo, este estudo foi conduzido com o objetivo de comparar procedimentos de amostragem para espécies florestais raras e com padrão de distribuição espacial agregado.

2. MATERIALE MÉTODOS

2.1. Área de estudo

Para a realização deste trabalho, foi simulada uma população com 44 indivíduos em uma floresta com 90 ha, dividida em 100 unidades de amostra ($N = 100$) com 9.000 m² de área cada, onde uma dada espécie pudesse apresentar padrão de distribuição espacial agregado e raridade numérica. Os pontos presentes na Figura 1 representam os indivíduos dessa espécie, bem como caracterizam a sua distribuição na floresta.

2.2. Análise do padrão de distribuição espacial e raridade numérica

De acordo com a metodologia proposta por Krebs (1999), para confirmação do padrão de distribuição espacial dos indivíduos da espécie, verificou-se se a frequência de unidades de amostra com diferentes números de indivíduos aderiria-se à distribuição de probabilidades de Poisson, que indica aleatoriedade, ou Binomial Negativa, que indica agregação (Tabela 1).

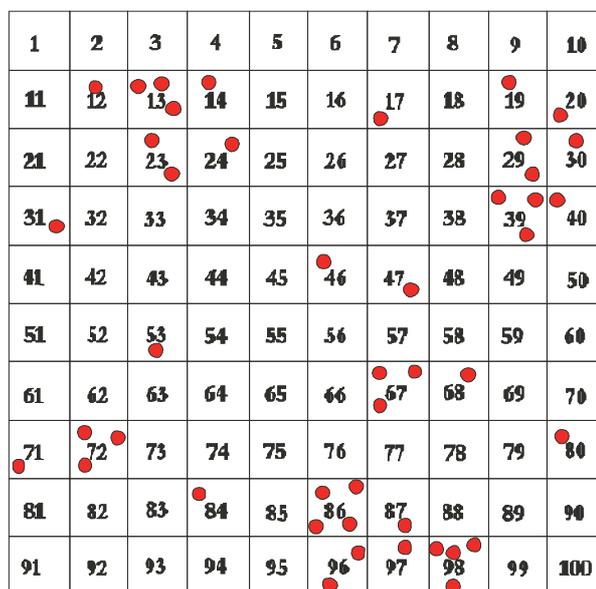


Figura 1 – Distribuição dos indivíduos de dada espécie em uma população composta por 100 unidades de amostra com 9.000 m² cada.

Figure 1 – Distribution of individuals of a given species in a population composed by 100 sapling units with 9000 m² each.

As probabilidades estimadas [$f(x)$] foram calculadas com o emprego das expressões apresentadas na Tabela 1, e as respectivas frequências estimadas (ou esperadas) de unidades de amostra que contêm x indivíduos ($x = 1, 2, 3, \dots, n$) na população foram obtidas ao multiplicar-se o número total de unidades de amostra na população ($N=100$) pelas respectivas probabilidades estimadas [$f(x)$].

A aderência das frequências observadas em relação às estimadas (esperadas) pelas funções de probabilidades de Poisson e Binomial Negativa foi verificada pelo teste de χ^2 (qui-quadrado). O valor de χ^2 calculado foi comparado com o valor de χ^2 tabelado, considerando-se um nível de probabilidade de 95% e (n-2) graus de liberdade para a função de probabilidade de Poisson e (n-3) graus de liberdade para a função de probabilidade Binomial Negativa (KREBS, 1999). Assim, se $\chi^2_{calculado}$ fosse maior que $\chi^2_{tabelado}$ a hipótese de nulidade seria rejeitada, ou seja, os dados não seguiriam a distribuição testada. Caso contrário, a hipótese de nulidade seria aceita, ou seja, os dados em questão seguiriam a distribuição testada.

Tabela 1 – Distribuições de Poisson e Binomial Negativa
Table 1 – Poisson and Negative Binomial Distributions

Distribuição	Expressão	Termos
Poisson	$f(x) = \begin{cases} e^{-\mu} * \left(\frac{\mu^x}{x!}\right); & \text{para } x = 0, 1, 2, 3, \dots, n. \\ 0, & \text{caso contrário} \end{cases}$	$f(x)$ = probabilidade de encontrar indivíduos em uma unidade de amostra e = base de logaritmo natural x = número de indivíduos (0, 1, 2, 3, ...) $x!$ = fatorial de x μ = média populacional
Binomial Negativa	$f(x) = \begin{cases} \left[\frac{\Gamma(k+x)}{x! \Gamma(k)} \right] * \left(\frac{\mu}{\mu+k} \right)^x * \left(\frac{k}{k+\mu} \right)^k; & \text{para } x = 0, 1, 2, 3, \dots, n \\ 0, & \text{caso contrário} \end{cases}$	$f(x)$ = probabilidade de se encontrar x indivíduos em uma unidade de amostra x = número de indivíduos (0, 1, 2, 3, ..., n) μ = valor verdadeiro da média populacional k = expoente da função Binomial Negativa Γ = função gamma

Como critério de raridade numérica, considerou-se que espécies com até um indivíduo por hectare seriam consideradas espécies raras (MARTINS, 1991; KAGEYAMA e GANDARA, 2000 *apud* RODRIGUES et al., 2003).

2.3. Procedimentos de amostragem

Foram utilizados três procedimentos de amostragem: Amostragem Casual Simples, sem reposição das unidades de amostra; Amostragem Sistemática; e a Amostragem Adaptativa em *Cluster*, com amostra inicial por amostragem casual simples e amostragem sistemática.

Na amostragem casual simples foram utilizadas amostras de tamanhos (n) iguais a 5, 10, 15, 20, 25 e 50 unidades de amostra. Em cada tamanho de amostra foram obtidas 30 amostras (repetições) por sorteio.

Na amostragem sistemática, foram utilizadas amostras de tamanho (n) igual a 5, 10, 20, 25 e 50 unidades de amostra, sendo que para cada tamanho de amostra e para cada repetição, a primeira unidade de amostra foi aleatorizada e as demais, distribuídas de forma equidistante. Em cada tamanho de amostra foi obtido o número de amostras (repetições) possíveis, ou seja, 20, 10, 5, 4 e 2 amostras.

Na amostragem adaptativa em *cluster*, foram utilizadas como amostras iniciais aquelas obtidas pela amostragem casual simples e pela amostragem sistemática, considerando-se os mesmos tamanhos de amostra propostos para esses dois procedimentos. Para este estudo, a condição de inclusão adotada foi a presença de pelo menos um indivíduo ($y_i \geq 1$) da espécie dentro da unidade de amostra. Desse modo, se houvesse um ou mais indivíduos em uma unidade de amostra inicial,

as unidades de amostra adjacentes a ela, isto é, que possuíam um lado em comum, seriam aderidas à amostra, formando as redes (THOMPSON e SEBER, 1996).

2.4. Análise da amostragem

Independentemente do procedimento de amostragem e da intensidade da amostragem, foram calculados para cada amostra (repetições), estimativas da média do número de indivíduos por unidade de amostra e variância da média. No cômputo da média e variância da média dos procedimentos tradicionais (amostragem casual simples e amostragem sistemática), foram empregados os estimadores usualmente utilizados.

Para a amostragem adaptativa em *cluster*, as estimativas da média e variância da média foram obtidas pelos estimadores de Hansen-Hurwitz modificados (HANSEN e HURWITZ, 1943; THOMPSON, 1990; THOMPSON e SEBER, 1996; BROWN e MANLY, 1998; BROWN, 2003), que foram calculados pelas seguintes expressões:

$$\bar{Y}_{HH^*} = \frac{1}{n_1} \sum_{i=1}^{n_1} w_i; \text{ sendo } w_i = \frac{\sum_{j=1}^{m_i} y_j}{m_i} \text{ e } S_{\bar{Y}_{HH^*}}^2 = \frac{N - n_1}{N n_1 (n_1 - 1)} \cdot \sum_{i=1}^{n_1} (w_i - \bar{Y}_{HH^*})^2$$

em que:

\bar{Y}_{HH^*} = estimador de Hansen-Hurwitz modificado para a média de indivíduos por unidade de amostra;

$S_{\bar{Y}_{HH^*}}^2$ = estimador de Hansen-Hurwitz modificado para a variância da média;



N = número total de unidades de amostra na população;

n_i = número inicial de unidades da amostra;

w_i = média das m_i observações na i -ésima rede; e

y_i = número de indivíduos da espécie na i -ésima unidade de amostra.

Em todos os procedimentos foram calculados o erro-padrão da média, erro de amostragem absoluto, erro de amostragem em porcentagem, exatidão e intervalo de confiança.

Foram elaborados gráficos para identificar visualmente a tendência de comportamento da precisão e a exatidão dos procedimentos de amostragem, bem como o tamanho final de unidades de amostras final no procedimento de amostragem adaptativa.

3. RESULTADOS E DISCUSSÃO

3.1. Padrão de distribuição espacial

Em termos comparativos, a distribuição Binomial Negativa obteve melhor ajuste aos dados observados (Tabela 2 e Figura 2).

Pelo teste de χ^2 (Tabela 3), verificou-se que os dados observados seguem a distribuição Binomial Negativa. Assim, confirma-se que a espécie em questão apresenta padrão de distribuição espacial agregado.

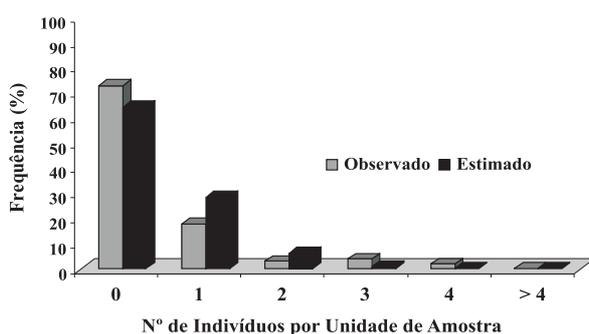
3.2. Análise dos procedimentos de amostragem

Observou-se (Figura 3) tendência de diminuição do erro de amostragem com o aumento do tamanho da amostra, em todos os procedimentos utilizados. Em pequenas amostras, a amostragem adaptativa em *cluster* com seleção sistemática da amostra inicial (Figura 3D) resulta em erros de amostragem menores. Em amostras maiores, os procedimentos são igualmente precisos.

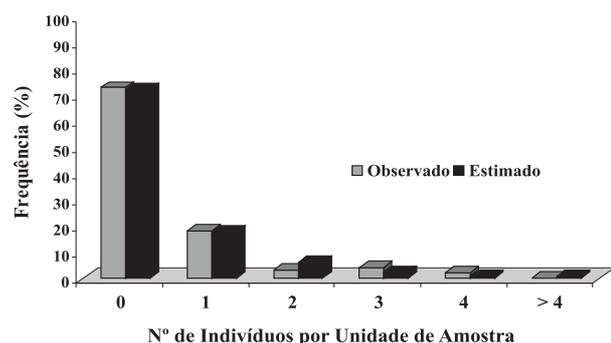
Tabela 2 – Número de indivíduos, frequência observada, probabilidade ($P(x)$) de se encontrar o número (x) de indivíduos em uma unidade de amostra e frequência esperada para distribuição de Poisson e Binomial Negativa

Table 2 – Number of individuals, observed frequency, probability [$P(x)$] of the number (x) of individuals in a sampling unit and expected frequency for the distribution of Poisson and Negative Binomial

Amostra		Distribuição de Poisson		Distribuição Binomial Negativa	
Nº de Indivíduos (x)	Frequência Observada (%)	P(x)	Frequência Esperada (%)	P(x)	Frequência Esperada (%)
0	73	0,644	64,40	0,723	72,30
1	18	0,283	28,34	0,178	17,80
2	3	0,062	6,23	0,061	6,10
3	4	0,009	0,92	0,023	2,30
4	2	0,001	0,10	0,009	0,90
> 4	0	0,0001	0,01	0,006	0,60
Total	100	1,000	100,00	1,000	100,00



(a)



(b)

Figura 2 – Frequências observadas e esperadas do número de unidades de amostra, considerando-se a distribuição de Poisson (a) e Binomial Negativa (b).

Figure 2 – Observed and expected frequencies of the number of sampling units, considering the distribution of Poisson (a) and Negative Binomial (b).

Tabela 3 – Resultados do teste χ^2 (Qui-quadrado) das distribuições testadas

Table 3 – Results of χ^2 test for the distributions tested

Distribuição	$\chi^2_{\text{calc.}}$	χ^2_{tab}	Conclusão
Poisson	52,89	9,49	Rejeita-se H_0 *
Binomial Negativa	4,82	7,81	Aceita-se H_0 *

* H_0 : Os dados em estudo seguem a distribuição testada.

Em todos os procedimentos de amostragem, quanto maior o tamanho da amostra, mais exatas são as estimativas do número de indivíduos (Figura 4). No entanto, a amostragem adaptativa em *cluster* (Figura 4CD) mostrou tendência de subestimação do número total de indivíduos com o aumento do tamanho da amostra inicial (n_i). Esse fato pode ser explicado pelo reduzido tamanho das redes presentes na região de amostragem e pelo baixo número de indivíduos presentes em cada unidade de amostra, gerando pesos (w_i) pequenos,

conferindo esse comportamento ao estimador de Hansen-Hurwitz modificado para a média.

Na amostragem casual simples (Figura 5) e na amostragem adaptativa em *cluster* com seleção casual da amostra inicial (Figura 7), observaram-se oito amostras com zero indivíduo, para o tamanho de amostra $n = 5$ e para $n = 15$, respectivamente. Na amostragem sistemática (Figura 6), observaram-se quatro casos para $n = 5$. Na amostragem adaptativa em *cluster* com seleção sistemática da amostra inicial (Figura 8), observaram-se quatro casos para $n = 5$ e um caso para $n = 10$, respectivamente.

Em todos os procedimentos, à medida que aumenta o tamanho da amostra (Figuras 5, 6, 7 e 8), observa-se menor amplitude do intervalo de confiança, em virtude da redução dos erros de amostragem.

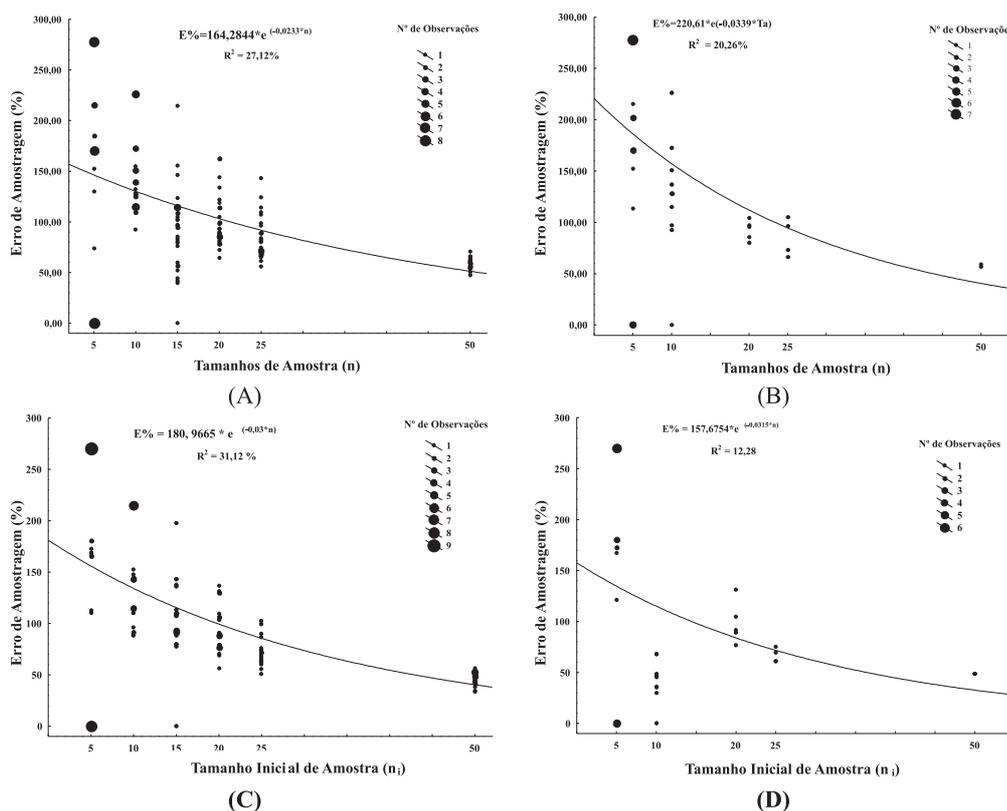


Figura 3 – Relação entre o erro de amostragem e o tamanho de amostra. Em (A) para amostragem casual simples, em (B) para sistemática, em (C) para a adaptativa em *cluster* com seleção casual da amostra inicial e (D) para a adaptativa em *cluster* com seleção sistemática da amostra inicial.

Figure 3 – Relation between the sampling error and the sample size. In (A) for the simple random sampling, in (B) for the systematic sampling, in (C) for the adaptive cluster sampling with initial selection by simple random sampling, and (D) for the adaptive cluster sampling with initial selection by systematic sampling.



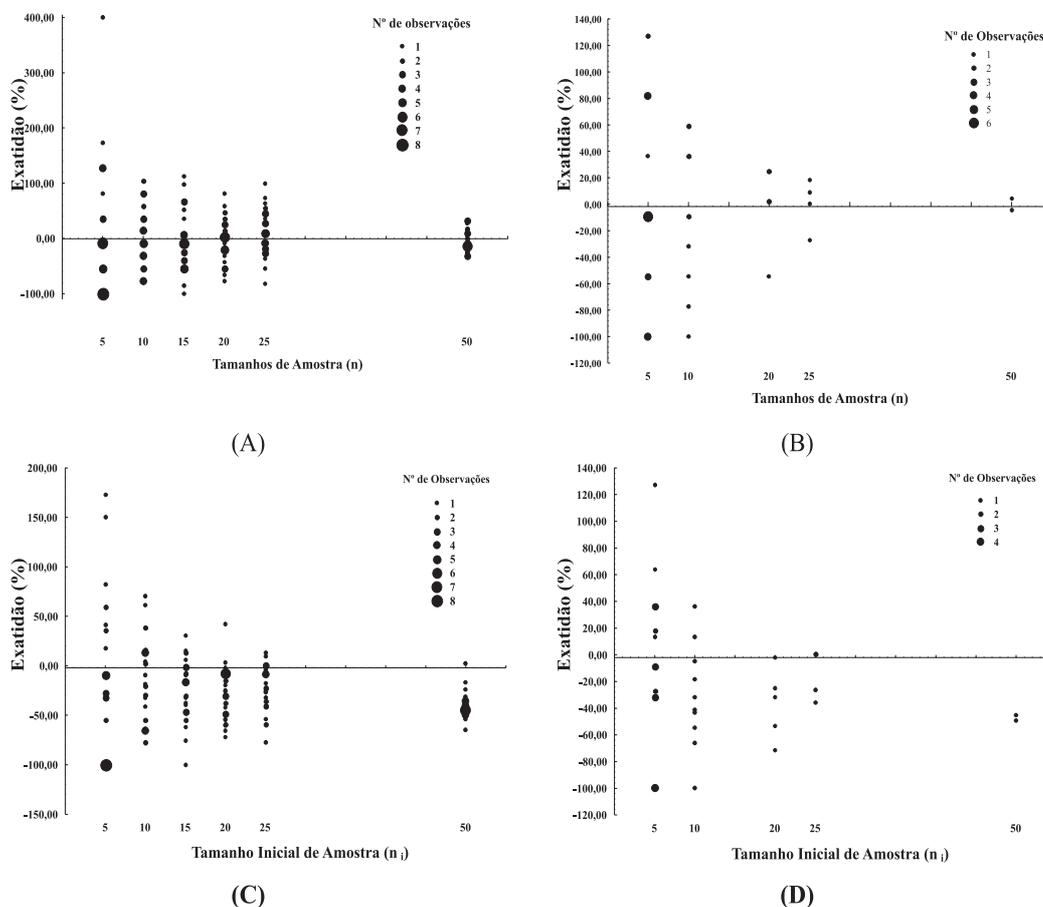


Figura 4 – Relação entre a exatidão e o tamanho de amostra. Em (A) para amostragem casual simples, em (B) para sistemática, em (C) para a adaptativa em *cluster* com seleção casual da amostra inicial e (D) para a adaptativa em *cluster* com seleção sistemática da amostra inicial.

Figure 4 – Relation between the accuracy and the sample size. In (A) for the simple random sampling, in (B) for the systematic sampling, in (C) for the adaptive cluster sampling with initial selection by simple random sampling, and (D) for the adaptive cluster sampling with initial selection by systematic sampling.

Para a amostragem casual simples, amostragem sistemática e amostragem adaptativa em *cluster* com seleção sistemática da amostra inicial (Figuras 5, 6 e 8), quanto menor o tamanho da amostra, maior o número de casos em que o verdadeiro valor do número total de indivíduos (44) esteve fora do intervalo de confiança, em virtude de algumas amostras terem apresentado estimativas para o total do número de indivíduos iguais a zero.

A amostragem adaptativa em *cluster* com seleção casual da amostra inicial mostrou comportamento inverso (Figura 7). Quanto maior o tamanho da amostra, maior o número de casos em que o verdadeiro valor do total

do número de indivíduos esteve fora do intervalo de confiança, chegando ao extremo de 28 casos em 30, no tamanho de amostra $n = 50$. Isso aconteceu devido ao reduzido tamanho das redes formadas e dos respectivos pesos (w_i), aliado ao grande tamanho da amostra.

Uma das vantagens que a amostragem adaptativa em *cluster* pode apresentar em relação à amostragem casual simples ou sistemática, em procedimentos de inventários florestais é a redução de tempo e de custos, uma vez que há maior eficiência na coleta de dados, visto que o esforço de amostragem é mais concentrado na área de ocorrência do recurso objeto de investigação.

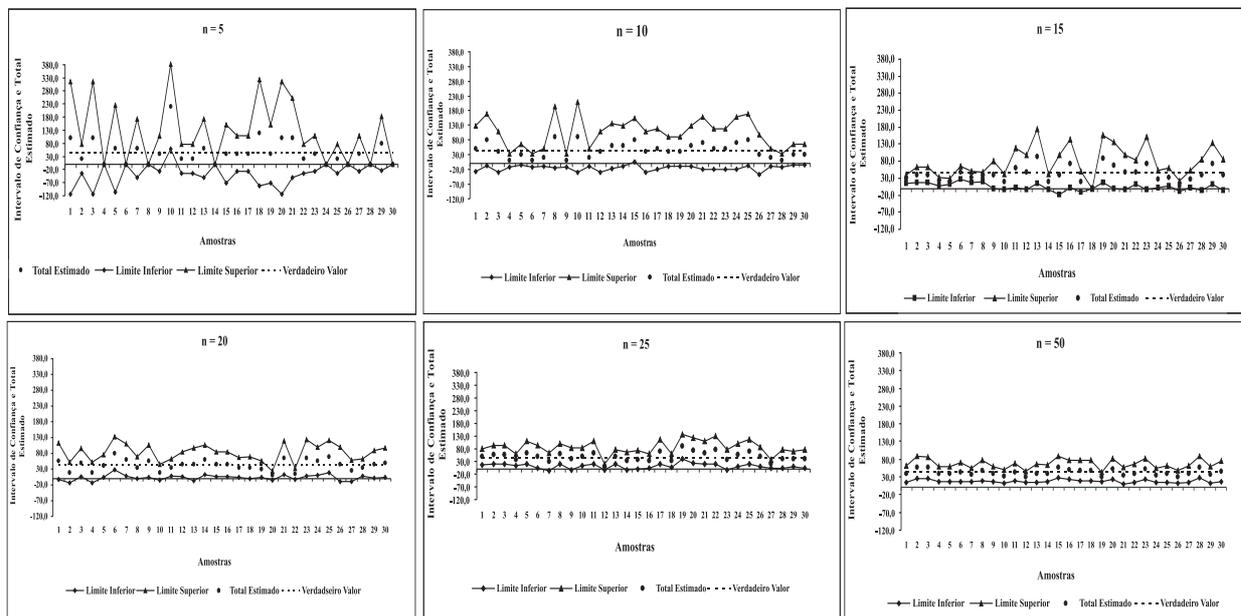


Figura 5 – Estimativa do total do número de indivíduos e limites inferior e superior do intervalo de confiança dos tamanhos de amostra $n = 5, 10, 15, 20, 25$ e 50 com 30 repetições cada da amostragem casual simples. A linha hachurada representa o verdadeiro total populacional de 44 indivíduos.

Figure 5 – Estimation of the total number of individuals and the upper and lower limits of the confidence interval for the sample sizes $n = 5, 10, 15, 20, 25$ and 50 with 30 repetitions each for simple random sampling. The dotted line represents the real population total of 44 individuals.

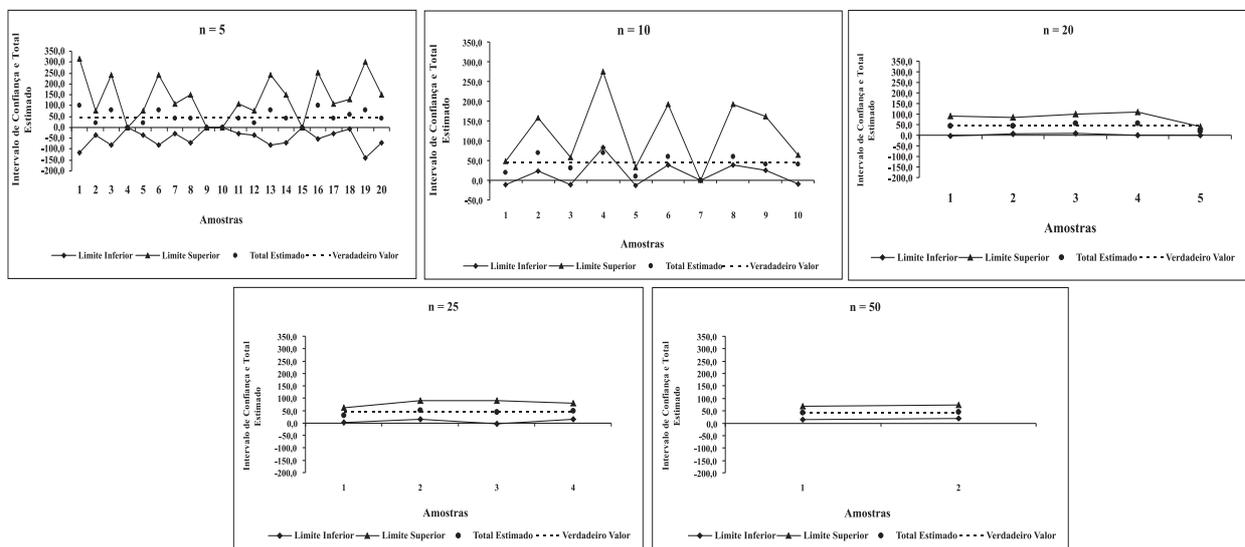


Figura 6 – Estimativa do total do número de indivíduos e limites inferior e superior do intervalo de confiança dos tamanhos de amostra $n = 5, 10, 20, 25$ e 50 com 30 repetições cada da amostragem sistemática. A linha hachurada representa o verdadeiro total populacional de 44 indivíduos.

Figure 6 – Estimation of the total number of individuals and the upper and lower limits of the confidence interval for the sample sizes $n = 5, 10, 20, 25$ and 50 with 30 repetitions each for systematic sampling. The dotted line represents the real population total of 44 individuals.

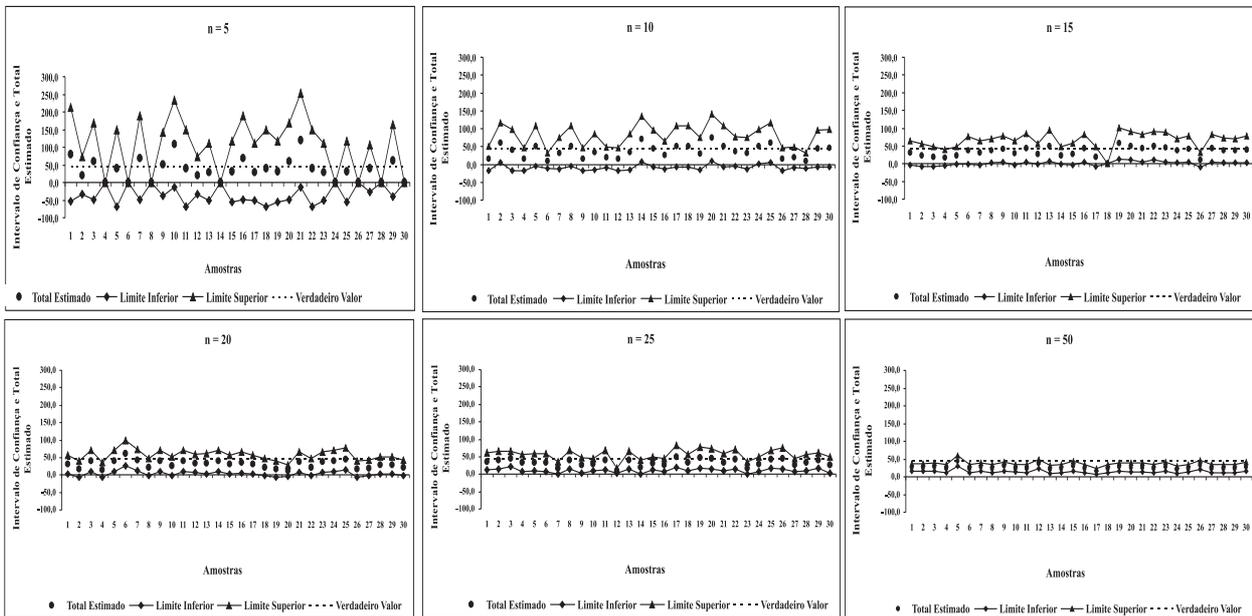


Figura 7 – Estimativa do total do número de indivíduos e limites inferior e superior do intervalo de confiança dos tamanhos de amostra $n = 5, 10, 15, 20, 25$ e 50 com 30 repetições cada da amostragem adaptativa em *cluster* com seleção casual da amostra inicial. A linha hachurada representa o verdadeiro total populacional de 44 indivíduos.

Figure 7 – Estimation of the total number of individuals and the upper and lower limits of the confidence interval for the sample sizes $n = 5, 10, 15, 20, 25$ and 50 with 30 repetitions each for the adaptive cluster sampling with initial selection by simple random sampling. The dotted line represents the real population total of 44 individuals.

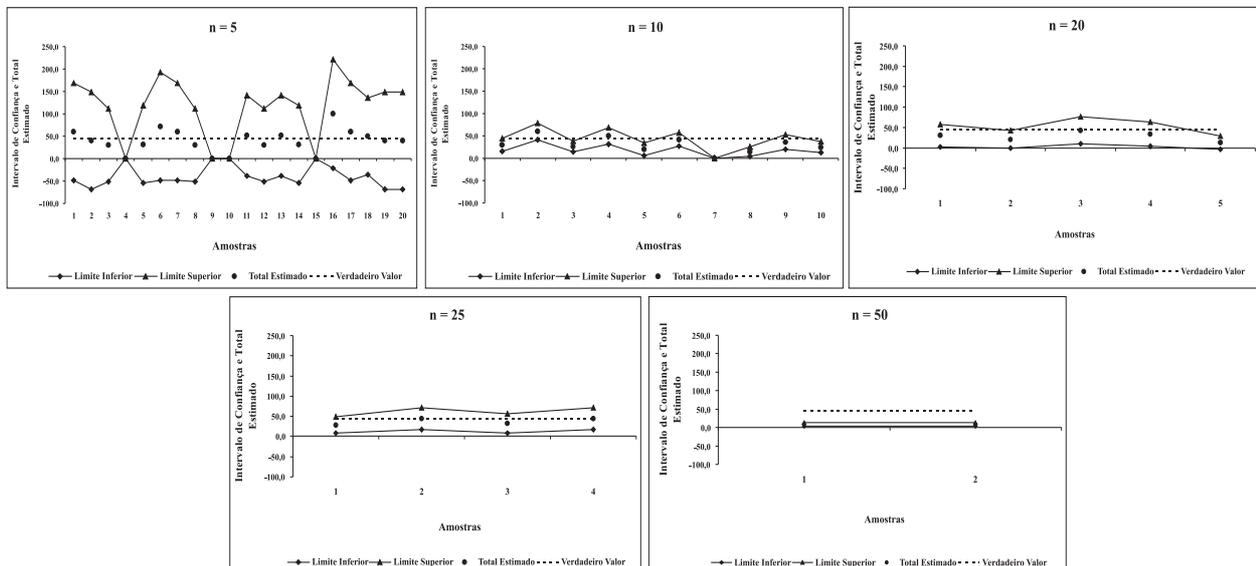


Figura 8 – Estimativa do total do número de indivíduos e limites inferior e superior do intervalo de confiança dos tamanhos de amostra $n = 5, 10, 20, 25$ e 50 com 30 repetições cada da amostragem adaptativa em *cluster* com seleção sistemática da amostra inicial. A linha hachurada representa o verdadeiro total populacional de 44 indivíduos.

Figure 8 – Estimation of the total number of individuals and the upper and lower limits of the confidence interval for the sample sizes $n = 5, 10, 20, 25$ and 50 with 30 repetitions each for adaptive cluster sampling with initial selection by systematic sampling. The dotted line represents the real population total of 44 individuals.

A eficiência da amostragem adaptativa em *cluster* pode ser aumentada em função da escolha da definição da vizinhança e do critério de inclusão. Segundo Brown (2003), se o critério de inclusão for estabelecido com valor muito baixo, o tamanho da amostra final poderá ser muito grande, pois as redes que serão formadas poderão abarcar muitas unidades amostrais. Do contrário, se o valor do critério de inclusão for muito alto, poucas unidades de amostra serão aderidas ao procedimento de amostragem. Além disso, outros fatores podem interferir na eficiência dos estimadores da amostragem adaptativa em *cluster*, por exemplo o tamanho e a forma de parcelas, o tamanho populacional e as diferentes escalas de padrão de agregação (KREBS, 1999).

A amostragem adaptativa em *cluster* possui potencial de utilização em procedimentos de inventário de florestas naturais brasileiras. Estudos desenvolvidos sobre padrão de distribuição espacial indicam que muitas espécies da Amazônia Legal ocorrem de forma agregada (SILVA e LOPES, 1996 OU 1982).

4. CONCLUSÃO

Após a realização deste trabalho, pôde-se concluir que, para pequenos tamanhos de amostra, a amostragem adaptativa em *cluster* com seleção sistemática da amostra inicial apresentou menores erros de amostragem. Em amostras maiores, os procedimentos têm a mesma precisão.

No entanto, considerando a exatidão das estimativas, bem como a tendência de sua distribuição em relação ao tamanho da amostra, pode-se concluir que a amostragem sistemática foi o melhor procedimento para estimar o número total de indivíduos da espécie em questão.

Conclui-se também que são necessários esforços adicionais no intuito de investigar o efeito do tamanho e forma de parcelas, da escala de agregação e do tamanho da população, e suas combinações, sobre a eficiência dos estimadores da amostragem adaptativa em *cluster*.

5. REFERÊNCIAS

BROWN, J.A.; MANLY, B. F. J. Restricted adaptive cluster sampling. **Environmental and Ecological Statistics**. v.5, n.1, p.49-63, 1998.

BROWN, J. A. Designing an efficient adaptive cluster sample. **Environmental and Ecological Statistics**, v.10, n.1, p.95-105, 2003.

GOSH, S.; INNES, J. L. Comparing strategies in forests monitoring programs. **Forest Ecology and Management**, v.82, p.231-238, 1996.

HANSEN, M.M.; HURWITZ, W.N. On the theory of sampling from finite populations. **Annals of Mathematical Statistics**. v.14, p. 333-362. 1943.

KALTON, G.; ANDERSON, D. W. Sampling rare populations. **Journal of Royal Statistics Society Series A**, v.149, p. 65-82, 1986.

KREBS, C. J. **Ecological methodology**. 2.ed. New York: Benjamin/Cummings, 1999. 581p.

MARTINS, F. R. **Estrutura de uma floresta mesófila**. Campinas: Unicamp, 1991. 246p.

RODRIGUES, L. A. et al. Florística e estrutura da comunidade arbórea de um fragmento florestal em Luminárias, MG. **Acta Botânica Brasílica**, v.17, n.1, p.71-87, 2003.

SILVA, J. N. M.; LOPES, J. C. A. **Distribuição espacial de árvores na Floresta Nacional do Tapajós**. Belém: Embrapa-CPATU, 1982. 14p. (Circular Técnica, 26).

SOUZA, D. R. et al. Análise estrutural em Floresta Ombrófila densa de terra firme não-explorada, Amazônia Oriental. **Revista Árvore**, v.30, n.1, p.75-87, 2006.

THOMPSON, S. K. Adaptive cluster sampling. **Journal of the American Statistical Association**, v.85, n.412, p.1050-1059, 1990.

THOMPSON, S. K.; SEBER, G. A. F. **Adaptive sampling**. New York: John Wiley & Sons, 1996. 265p.



