

UMA NOVA ABORDAGEM PARA DEFINIÇÃO DA SUFICIÊNCIA AMOSTRAL EM FRAGMENTOS FLORESTAIS NATIVOS

Lucas Rezende Gomide¹, José Roberto Soares Scolforo², Cláudio Roberto Thiersch³,
Antônio Donizette de Oliveira⁴

(recebido: 18 de agosto de 2005; aceito: 18 de outubro de 2005)

RESUMO: A suficiência amostral se caracteriza como um importante ponto na garantia do conhecimento da variável de interesse dentro de uma população. Entretanto, um problema comum encontrado diz respeito à subjetividade dos métodos para a sua determinação, proporcionando interpretações muitas vezes equivocadas da fisionomia estudada. Desta maneira objetivou-se com este estudo captar a estimativa da suficiência amostral pelo método da regressão linear de platô, auxiliada pelo teorema do limite central (nova abordagem) e compará-lo com o método tradicional de ajuste, tendo como área de estudo 20 fragmentos de vegetação nativa, dispersas ao longo da bacia do rio São Francisco - MG. O método de ajuste tradicional apresentou resultados duvidosos, ou ainda um possível “falso” ponto de platô nos fragmentos: 22, 27, 77, 100 e 112. Na nova abordagem, a caracterização do platô foi bem clara e sem deixar dúvidas sobre a suficiência amostral.

Palavras-chave: Regressão linear de platô, suficiência amostral, teorema do limite central.

A NEW APPROACH TO DETERMINE THE SAMPLING SUFFICIENCY IN NATIVE FOREST FRAGMENTS

ABSTRACT: *The sampling sufficiency is an important point to guarantee the knowledge of the variable of interest in a population. However, a common problem is the subjectivity of the methods used to determine the sampling sufficiency, leading to equivocated interpretations of the studied physiognomy. Thus, the objective of this study was to determine the sampling sufficiency applying a new approach combining the plateau response linear regression with the central limit theorem and to compare it with the traditional adjustment method, in 20 native fragments along the São Francisco River Watershed - MG. The traditional adjustment method presented some dubious results such as a “false” plateau point in fragments: 22, 27, 77, 100 and 112. Based on the new approach the plateau did not occur and clearly the plateau point left no doubt about the sampling sufficiency.*

Key words: plateau response; linear regression; sampling sufficiency; central limit theorem.

1 INTRODUÇÃO

Ao estudar a fisionomia de um ambiente é muito claro que a determinação de toda a variação das espécies na comunidade só será alcançada quando a amostragem representar toda a área (censo). Entretanto, seja pelo cronograma temporal, seja pelo orçamento elevado, a alternativa é fazer uso de procedimentos de amostragem. No caso de estudos florísticos a suficiência amostral pode ser obtida de métodos como a curva espécie-área, regressão linear de platô, regressão quadrática de platô. A regressão exponencial de platô pode ser testada ainda, para este tipo de análise.

Autores como Camargo (1997), Gomide (2004), Gomide et al. (2005), Lima (1997), Nappo et al. (1999), Vasconcelos (1992) e Volpato (1994), entre outros, utilizaram a regressão linear de platô em seus estudos florísticos para validar a suficiência amostral e obtiveram resultados positivos na inferência da amostragem. Mas a sua introdução no Brasil se deu na área de Ciência dos Solos, ao desenvolver um ensaio sobre fertilidade do solo, realizado por Braga (1983).

Apesar destes estudos um problema frequente ao analisar a composição florística de uma população é a intensidade amostral, mais

¹Engenheiro Florestal, M.S. – Rua Leon Jofre Avayou, 585 – Vera Cruz – 37.200-000 – Lavras, MG – lrgomide@hotmail.com

²Professores do Departamento de Ciências Florestais – Universidade Federal de Lavras/UFLA – Cx. P. 3037 – Lavras, MG – 37.200-000 – scolforo@ufla.br; donizete@ufla.br

³Engenheiro Florestal, M.S., Doutorando do Departamento de Ciências Florestais – Universidade Federal de Lavras/UFLA – Cx. P. 3037 – 37.200-000 – Lavras, MG – claudio.thiersch@vcp.com.br

precisamente no ponto de suficiência amostral. Esta é uma questão muito discutida pelos pesquisadores, os quais sempre afirmam a necessidade de que as espécies de uma comunidade estejam significativamente mensuradas na amostragem e que esta retrate seus detalhes ao longo de gradientes heterogêneos.

A deficiência amostral proporciona interpretações e conclusões muitas vezes equivocadas da fisionomia estudada, por não contemplar corretamente o ambiente em estudo. Isso impede a realização de comparações estruturais, florísticas e de diversidade influenciando decisivamente na análise da vegetação. Somente após a determinação da suficiência amostral, pode-se então proceder à quantificação de vários índices de diversidade e similaridade, e tirar conclusões sobre as peculiaridades da vegetação amostrada.

Desta maneira, objetivou-se com este trabalho captar a estimativa da suficiência amostral pelo método da regressão linear de *plateau*, auxiliada pelo teorema do limite central em 20 fragmentos de vegetação nativa dispersas ao longo da Bacia do Rio São Francisco, no Estado de Minas Gerais, além da comparação entre o método tradicional de ajuste, a partir de um único sorteio aleatório, e a metodologia proposta neste trabalho, com 30 sorteios aleatórios.

2 MATERIAL E MÉTODOS

2.1 Descrição da área de estudo

A área de estudo abrange áreas de preservação permanente associadas aos cursos de água da Bacia do rio São Francisco, no Estado de Minas Gerais, compreendida entre a Serra da Canastra e o município de Manga (Figura 1). Ela abrange as regiões do Alto e Médio São Francisco e, nesta faixa de terra, o rio apresenta 36 afluentes que são considerados os mais expressivos para supri-lo com água. Além disso, existem outros rios e riachos fundamentais para a bacia, porém com menor porte. Os rios mais expressivos são: Abaeté, Bambuí, Borrachuda, Carinhonha, Córrego São Miguel, Corrente, Indaiá, Jequitaiá, Pará, Pardo, Pandeiro, Paracatu, Paraopeba, Piauí, Peixe, Rio das Velhas, Urucuia, e Verde Grande.

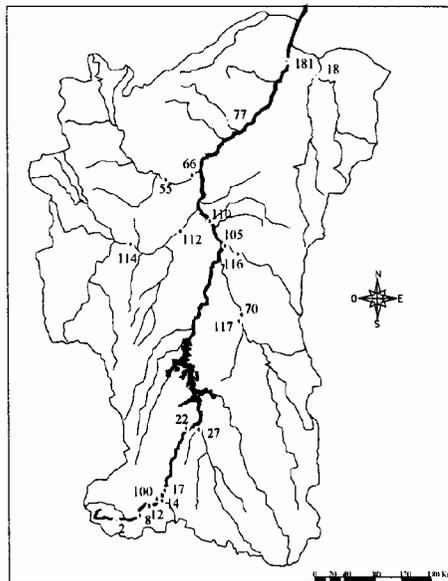


Figura 1 – Mapa da Bacia do rio São Francisco em Minas Gerais, e a localização dos 20 fragmentos estudados.

Figure 1 – Map of the São Francisco River Watershed in Minas Gerais, and location of the studied 20 fragments.

Os principais afluentes do rio São Francisco situam-se no Estado de Minas Gerais e fornecem cerca de 70% da água (SATO & GODINHO, 1999). O percurso do rio São Francisco em Minas Gerais, acompanhando sistematicamente suas margens, apresenta 1400 km de extensão, com área de drenagem de 235,207 km² correspondente a 40,07% da área do Estado, constituindo-se o maior rio totalmente brasileiro.

2.2 Amostragem e coleta dos dados

O método de amostragem utilizado foi o conglomerado em estágio único, com sistematização das unidades amostrais na área. Cada conglomerado foi composto por três subunidades de 10 x 25 m, distantes 25 metros uma da outra. Ao longo do fragmento foram estabelecidos transectos, distantes 100 metros entre conglomerados. Já ao longo de um transecto, os conglomerados ficaram a uma distância de 50 metros um do outro (Figura 2). O número de conglomerados variou de acordo com a área de cada fragmento (Tabela 1).

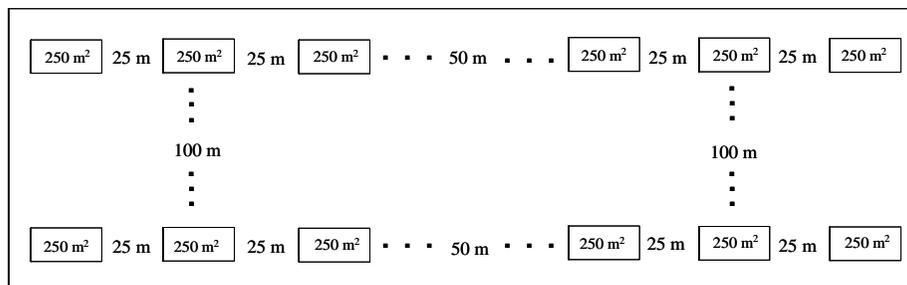


Figura 2 – Representação da amostragem em conglomerado, aplicada aos fragmentos da bacia do rio São Francisco, MG.

Figure 2 – Sampling cluster scheme applied to each fragment of the São Francisco river watershed - MG.

Tabela 1 – Dados da amostragem no levantamento florístico para cada fragmento.

Table 1 – Sampling data of the floristic inventory for each fragment.

Fragmento	Município	Área do fragmento (ha)	Área amostral (ha)	Número de subunidades
2	Vargem Bonita	21,83	0,600	24
8	Bambuí	34,95	0,650	26
100	Iguatama	5,17	0,450	18
12	Iguatama	45,12	0,750	30
14	Iguatama	43,40	1,050	42
17	Lagoa da Prata	165,26	1,500	60
22	Abaeté	2,88	0,375	15
27	Martinho Campos	15,31	0,475	19
117	Lassance	17,79	0,850	34
70	Lassance	41,08	0,950	38
116	Jequitaiá	107,64	1,250	50
105	Ibiaí	15,15	0,750	30
110	Ibiaí	47,16	0,900	36
112	Santa Fé de Minas	40,00	0,500	20
114	Brasilândia	38,25	1,000	40
66	São Romão	40,57	1,250	50
55	Urucuia	9,34	0,625	25
77	Januária	4,74	0,475	19
181	Manga	24,93	1,050	42
18	Matias Cardoso	85,83	0,850	34
Total		806,40	16,30	652,00

Embora as subunidades maiores que 250 m² (10 x 25 m) sejam desejáveis para estudos de comunidades, a característica ciliar dos fragmentos encontrados (a largura das áreas de preservação permanente) e, principalmente, a ausência de grandes fragmentos impediu a obtenção de subunidades amostrais com áreas superiores, o que tornou como medida possível de subunidade no conglomerado o

tamanho 10 x 25 metros.

Cochran (1963) argumentou que o conglomerado é a maneira mais fácil e econômica de obter inferências populacionais, considerando pequenas unidades amostrais. A unidade amostral do conglomerado é constituída por um grupo ou conglomerado de unidades menores chamadas de elementos.

Desta forma, foram amostrados 20 fragmentos florestais distribuídos ao longo das áreas de preservação permanente associada aos cursos de água da Bacia do rio São Francisco (Figura 1), com o propósito de representar as diferentes fisionomias ali encontradas. A caracterização dos fragmentos indicando a localização e suas descrições ambientais e fisionômicas encontra-se na Tabela 1. Nesta pode se perceber que a área dos fragmentos selecionados totalizou 806,40 ha, e que a amostragem correspondeu a 16,3 ha, advinda de 625 subunidades dos conglomerados.

Nessas subunidades, as variáveis dendrométricas quantificadas foram a altura total e a CAP, para todos os indivíduos que apresentavam a circunferência a 1,30 m de altura (CAP) igual ou superior a 15,7 cm, ou diâmetro igual ou superior a cinco cm. Os indivíduos perfilhados foram incluídos quando pelo menos uma das ramificações obedecia ao diâmetro mínimo de inclusão.

2.3 Coleta e identificação do material botânico

A identificação botânica das espécies arbóreo-arbustivas se deu a partir da coleta do material botânico de todos os indivíduos e encaminhados ao Herbário ESAL da Universidade Federal de Lavras, para consultas a especialistas, literatura e/ou comparações com espécimes lá existentes. A classificação dos indivíduos amostrados foi de família, gênero e espécie, tendo como referência *Angiosperm Phylogeny Group* (JUDD et al., 1999).

2.4 Suficiência amostral

A suficiência amostral foi estimada e avaliada utilizando-se o procedimento da regressão linear com resposta em platô (REGRELRP) do Sistema para Análises Estatísticas (SAEG, 1997). Assim, a REGRELRP permitiu determinar a área mínima de amostragem, para representação florística das áreas.

2.4.1 Arquivo de dados

Para realizar o ajuste da regressão e calcular a suficiência amostral, as subunidades dentro de cada fragmento foram sorteadas aleatoriamente 30 vezes. Em cada sorteio, calculava-se a frequência acumulada (FA) desta combinação. No final dos sorteios, extraía-se a média de (FA) e calculava-se ainda a área acumulada referente às subunidades do levantamento florestal. A partir deste ponto aplicava-se a REGRELRP, obtendo-se seus parâmetros e o ponto de encontro entre as duas regressões. Ao todo foram realizados 600 sorteios distribuídos nas 20 fisionomias estudadas.

2.4.2 Teorema do limite central

O teorema considera que se uma variável qualquer, com uma distribuição muito diferente da normal (pode até mesmo ser discreta), poderá se tornar uma curva normal, desde que desta população seja sorteado um grande número de amostras (BUSSAB, 1988). Esse número é considerado suficientemente grande quando $n \geq 30$, como mostra a Figura 3.

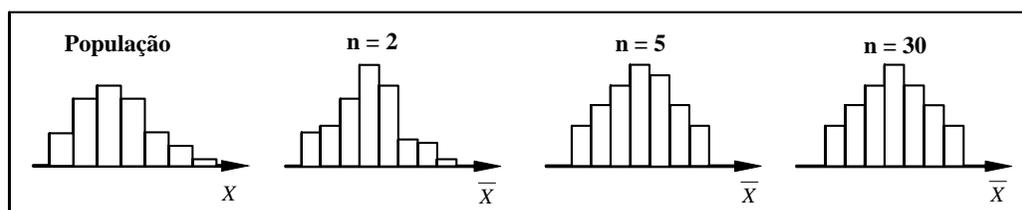


Figure 3 – Histograma de frequência de uma população, seguindo uma seqüência de sorteios aleatórios.

Figure 3 – Frequency histogram for a population, following the random sampling sequence.

O princípio desta teoria parte da aleatorização de unidades amostrais, em que novas combinações amostrais são geradas provenientes da amostragem original. Para Wallace (1988), estes novos conjuntos amostrais são considerados independente entre si e tem covariância igual a zero. Considere-se então X_1, X_2, \dots, X_n como variáveis aleatórias independentes com média μ e variância σ^2 distintas, e provenientes de cada aleatorização. Como todos os X_i são gerados a partir de N amostras da mesma população, pode-se considerar que:

$$\mu_{\bar{X}} = \frac{\mu_{X_1}}{N} + \frac{\mu_{X_2}}{N} + \dots + \frac{\mu_{X_n}}{N} = \mu \quad e$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma_{X_1}^2}{N} + \frac{\sigma_{X_2}^2}{N} + \dots + \frac{\sigma_{X_n}^2}{N} = \frac{\sigma^2}{N}$$

Logo, assume-se pelo teorema que X tende a uma distribuição Normal:

$$\bar{X} \rightarrow N \left[\mu_X, \frac{\sigma_X^2}{n} \right]$$

Portanto com esta regra estabelecida, uma dada variável que não obedeça a distribuição normal, pode perfeitamente tornar-se uma aproximação da distribuição normal.

2.4.3 Método da regressão linear de platô

A Regressão linear de Plateau foi desenvolvida para a análise de modelos matemáticos descontínuos, sendo assim classificado como uma regressão segmentada. A segmentação do modelo se dá em um único ponto, o que proporciona a formação de um bi-segmento. A primeira parte do modelo, antes da divisão do segmento é representada por um modelo linear simples ($Y_i = \beta_0 + \beta_1 X_i$), já na segunda parte aparece o modelo de platô ($Y_i = P$). O modelo de platô apresenta apenas a constante (P) como parâmetro, o que proporciona a formação de uma reta contínua (platô), sem a influência de β_1 que confere a inclinação da reta, por isso forma-se um platô.

O modelo geral possui um segmento de reta antes do ponto de junção (X_0) com o platô, como mencionado, e o uso de uma variável binária (*Dummy*)

é empregado para unir os dois modelos, sendo dado por:

$$Y_i = (\beta_0 + \beta_1 X_i)Z_i + P(1-Z_i)$$

Em que:

β_0, β_1 - parâmetros a serem estimados na equação da reta;

P - parâmetros a serem estimados na equação de Platô;

Y_i - variável dependente;

X_i - variável independente;

Z - variável *Dummy*.

Segundo Ferreira (2005), este modelo apresenta 3 parâmetros (β_0, β_1 e P), entretanto P não pode ser expresso em função dos demais. Apesar das variáveis parciais não dependerem dos parâmetros, este é um modelo não-linear uma vez que a matriz Jacobiana depende de X_0 para ser construída.

O mecanismo que aciona cada modelo é em função de Z . Logo, para o valor de $Z = 1$, ou seja ($X_i < X_0$), o modelo linear é ativado. Com o valor de $Z = 0$, isto é ($X_i \geq X_0$) o modelo de platô é acionado.

O cálculo do ponto de junção (X_0) é dado a seguir:

$$Y = \beta_0 + \beta_1 X_0 \quad (1)$$

$$Y = P \quad (2)$$

Igualando os dois modelos (1) e (2) gera uma continuidade entre as retas, tornando-as unidas por X_0 , e assim tem-se:

$$\beta_0 + \beta_1 X_0 = P \quad (3)$$

Rearranjando a equação (3) em função de X_0 , tem-se:

$$X_0 = \frac{P - \beta_0}{\beta_1}$$

No processo de ajuste, a base de dados é rearranjada em vários conjuntos de dados, começando por 2 e $n-2$ dados, respectivamente, até $n-1$ e 1, em que realiza-se uma série de ajustes de regressão até a definição do ponto de platô, sendo 'n' o número de unidades amostrais (ALVAREZ, 1985).

A análise de variância é montada para o modelo linear, calculando-se a Soma de Quadrado Total (SQT), Soma de Quadrado da Regressão (SQR) e a Soma de Quadrado do Desvio (SQD) a partir das fórmulas:

$$SQT = \sum_{i=1}^n Y_i^2 - \frac{\left(\sum_{i=1}^n Y_i\right)^2}{n};$$

$$SQR = \beta_1 \left(\sum_{i=1}^n X_i Y_i - \frac{\sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n} \right) e$$

$$SQD = SQT - SQR$$

Em que: Y_i e X_i - definidos anteriormente.

Para o modelo de Platô, apenas a SQT é calculada, devido a inexistência do parâmetro β_1 . Entretanto esta soma de quadrado é considerada como a SQD. Isto ocorre devido ao platô, apresentar um comportamento constante, após o ponto X_0 . Logo, a SQT é a mesma que a SQD, para platô. E na seleção das equações (reta e platô) que apresentarem a menor SQD é eleita o melhor conjunto de equações, para o conjunto de dados. Em que a Soma de Quadrado do Desvio da combinação é composta pelo somatório das SQD (reta) e SQT (platô).

2.4.4 Tradicional x Nova abordagem

Inúmeros trabalhos que empregam a técnica da Regressão linear de platô utilizam este procedimento, para o cálculo da suficiência amostral de forma única ou tradicional. Ou seja, partem para o ajuste da regressão, sem a utilização de um mecanismo de base matemática, que sirva de balizamento na confiabilidade do conjunto de dados. Aquele conjunto de dados observados na *frequência acumulada X área acumulada*, deve ser o espelho do que acontece na fisionomia estudada. Pois quando se trata de combinações, algumas delas podem não expressar o caráter de uma população, deixando a desejar e prejudicando futuras interpretações dos dados. A nova abordagem mescla o uso da

REGRELRP com o teorema do limite central. E neste estudo, essa nova abordagem foi confrontada com a forma tradicional de ajuste, Na qual se realiza um único sorteio, contra a média de 30 sorteios da nova proposta.

3 RESULTADOS E DISCUSSÃO

3.1 Tradicional

Realizado os ajustes a partir de um único sorteio, para cada fragmento estudado, verificou-se, que houve a formação do platô em todos os casos, conforme observado na Tabela 2.

Entretanto, alguns fragmentos como: 22, 27, 77 e 100 demonstraram que a formação do platô se deu próximo ao limite máximo de área amostrada, nos respectivos fragmentos, em que os valores de X_0 foram: 3630,54 m² (22), 4708,60 m² (27), 4581,75 m² (77) e 4410,99 m² (100). Observe-se que a área total amostrada pelo levantamento, nos fragmentos em questão foram: 3750 m² (22), 4750 m² (27), 4750 m² (77), 4500 m² (100). Arredondando-se os valores de X_0 em relação a unidade amostral (250 m²), percebe-se que a suficiência amostral é atingida exatamente sobre o último ponto, pois em se tratando de inventário florestal não há como mensurar frações de uma parcela.

Este é um ponto interessante para o surgimento de um possível “falso” platô, no qual o platô é formado na posição do último par ordenado observado.

O exemplo hipotético a seguir, exemplifica melhor este fato. Considere-se a hipótese do ajuste de um modelo linear “perfeito”, no qual se espera que a Soma de Quadrado do Desvio e Total da regressão sejam nulos, pois a reta ajustada passa exatamente sobre os pontos observados. A tabulação dos dados segue na Tabela 3, em que cada valor de FA é exatamente 10% do valor da área acumulada com uma taxa de crescimento, sempre crescente e constante. Este ponto foi promovido para supor a não formação do platô, uma vez que o incremento de espécies é crescente e não forma pontos de estabilidade do crescimento.

Após o ajuste deste conjunto de dados, observou-se o que a hipótese testada foi alcançada, como pode-se notar a seguir, os valores dos parâmetros ajustados, $\beta_0 = 0$, $\beta_1 = 0,1$ e $P = 175$.

Tabela 2 – Parâmetros calculados dos modelos e parte da Análise de Variância utilizada para a seleção do modelo, pelo método tradicional de ajuste.

Table 2 – Results of regression parameters in a traditional adjustment method and part of the Analysis of Variance used to select the model.

Fragmento	Modelo	Dados	β_0	β_1	P	SQD	SQD (total)	F	R ² (%)
2	Reta	16	35,05	0,0275	-	925,1132	1306,6132	242,78	94,54
	Platô	8	-	-	148,75	381,5000	-	-	-
8	Reta	20	30,821	0,0142	-	749,3620	896,6950	201,35	91,79
	Platô	6	-	-	103,67	147,3330	-	-	-
12	Reta	27	11,972	0,0049	-	215,7640	216,4307	286,9	91,98
	Platô	3	-	-	45,6667	0,6667	-	-	-
14	Reta	12	7,0303	0,0075	-	39,1422	322,5088	127,36	92,72
	Platô	30	-	-	30,4333	283,3666	-	-	-
17	Reta	32	35,794	0,0141	-	3217,3300	5316,1900	315,87	91,32
	Platô	28	-	-	150,429	2098,8600	-	-	-
18	Reta	15	13,248	0,0076	-	91,5012	263,2912	145,23	91,78
	Platô	19	-	-	42,8947	171,7900	-	-	-
22	Reta	14	22,0769	0,0132	-	181,1252	181,1252	163,59	93,17
	Platô	1	-	-	70	0	-	-	-
27	Reta	18	27,026	0,0121	-	576,9050	576,9050	122,61	88,45
	Platô	1	-	-	84	0	-	-	-
55	Reta	10	8,7333	0,0147	-	63,7333	211,4663	139,23	94,56
	Platô	15	-	-	47,1333	147,7330	-	-	-
66	Reta	19	11,018	0,0051	-	67,5719	83,2862	230,08	93,11
	Platô	7	-	-	35,4286	15,7143	-	-	-
70	Reta	24	28,917	0,0123	-	1048,3200	1309,1770	227,46	91,18
	Platô	14	-	-	104,714	260,8570	-	-	-
77	Reta	18	15,346	0,008	-	126,0770	126,0770	235,65	93,64
	Platô	1	-	-	52	0	-	-	-
100	Reta	17	19,0956	0,0095	-	114,8603	114,8603	300,55	95,25
	Platô	1	-	-	61	0	-	-	-
105	Reta	13	8,1538	0,0095	-	110,7690	233,2400	102,77	90,33
	Platô	17	-	-	40,1765	122,4710	-	-	-
110	Reta	24	13,569	0,0078	-	337,5130	449,1800	281,49	92,75
	Platô	12	-	-	60,8333	111,6670	-	-	-
112	Reta	19	7,4561	0,0026	-	17,4930	17,4930	229,64	93,1
	Platô	1	-	-	24	0	-	-	-
114	Reta	17	2,8456	0,0055	-	42,5858	119,6767	269,77	94,73
	Platô	22	-	-	26,6364	77,0909	-	-	-
116	Reta	32	33,655	0,0068	-	2188,8200	2476,5980	107,68	78,21
	Platô	18	-	-	88,8889	287,7780	-	-	-
117	Reta	22	10,584	0,0016	-	31,7527	44,6694	89,14	81,67
	Platô	12	-	-	19,5833	12,9167	-	-	-
181	Reta	22	9,1558	0,0037	-	128,0750	247,0250	118,87	85,59
	Platô	20	-	-	30,05	118,9500	-	-	-

Tabela 3 – Base de dados hipotética para mostrar o caso de um “falso” platô.

Table 3 – Hypothetic data base to show a “false” plateau point.

N	Área Acumulada (m ²)	Frequência Acumulada (número de espécies)
1	250	25
2	500	50
3	750	75
4	1000	100
5	1250	125
6	1500	150
7	1750	175

A Soma de quadrado do Desvio para ambos os modelos foi 0, e conseqüentemente para a SQD do somatório entre os modelos igual a zero. Dos sete dados presentes neste ajuste, seis dados pertenceram a regressão linear e 1 a regressão de platô.

O encontro entre as duas retas (linear e platô) ocorreu na posição 1750 m² de área acumulada e 175 espécies (FA), ou seja no último par ordenado observado para este conjunto de dados. A seleção deste ponto (X_0) pode ser observada na escolha da menor SQD, apresentada na Tabela 4. Esse ponto indica um “falso” platô, assim chamado, por representar de forma errônea um suposto platô que não deveria ocorrer, devido a não estabilidade da reta, que apresenta um crescimento constante e crescente. Isto implica no erro de uma “falsa” suficiência amostral, ao se aplicar a regressão linear de platô, com um único sorteio. Como a seleção do

melhor conjunto de equações é a partir, única e exclusivamente sobre a menor SQD do somatório entre a reta e platô. O caso de ocorrer no último dado, indica que o ponto selecionado sempre será a menor SQD, e para o último valor sempre será zero.

A combinação de dados “6” (reta) e “1” (platô) foi a selecionada na formação do platô, por apresentar a menor Soma de Quadrado do Desvio, entretanto, a combinação de dados “7” (reta) e “1” (platô) apresentou a mesma SQD = 0, e não pode ser selecionada, devido a incapacidade de dividir o mesmo ponto, entre as duas regressões.

O fragmento 112 é um caso especial, pois pelo ajuste realizado, indica que o ponto de encontro entre os segmentos de reta (linear e platô) é $X_0 = 6363,04$ (m²) e $P = 24$ (FA), mas observando melhor a área total amostrada, percebe-se que foram mensurados apenas 5000 m², com as mesmas 24 espécies. Logo, o platô é formado após o limite da base de dados, e teoricamente não há a formação do platô.

3.2 Nova abordagem

A aplicação da regressão linear de platô permitiu compreender o comportamento da amostragem, ao longo de todos os levantamentos realizados, observando se a quantificação da variável espécie foi suficiente. Na Tabela 5 são mostrados os parâmetros estimados para o modelo linear e o platô, na qual o coeficiente de determinação variou de 89,15% (F_{117}) a 96,55% (F_{77}), os quais reproduziram boas estimativas, mesmo analisando-se apenas esta medida.

Tabela 4 – Relação das combinações dos dados pertencentes a cada modelo, com suas respectivas Soma de Quadrado do Desvio (SQD) e a soma de ambas.

Table 4 – Relation of data combination of each model and their sum of square error and a total sum.

Número de pontos na equação da Reta	Número de pontos na equação de Platô	SQD		
		(Platô)	(Reta)	(Total)
2	6	10937,50	0	10937,50
2	5	6250,00	0	6250,00
3	5	6250,00	0	6250,00
3	4	3125,00	0	3125,00
4	4	3125,00	0	3125,00
4	3	1250,00	0	1250,00
5	3	1250,00	0	1250,00
5	2	312,50	0	312,50
6	2	312,50	0	312,50
6	1	0	0	0
7	1	0	0	0

Tabela 5 – Parâmetros calculados dos modelos e parte da Análise de Variância utilizada para a seleção do modelo, pelo novo método de ajuste.

Table 5 – Regression model parameters of the new approach and part of the Analysis of Variance used to select the model.

Fragmento	Modelo	Dados	β_0	β_1	P	SQD	SQD (total)	F	R ² (%)
2	Reta	17	32,4264	0,0267	-	884,7197	1156,0117	308,60	95,36
	Platô	7	-	-	149,1857	271,2920	-	-	
8	Reta	18	20,6349	0,0183	-	499,4294	677,8314	323,23	95,28
	Platô	8	-	-	104,965	178,4020	-	-	
12	Reta	19	11,4774	0,0064	-	99,1215	142,6300	253,71	93,72
	Platô	11	-	-	43,0091	43,5085	-	-	
14	Reta	25	12,5043	0,0037	-	117,5723	171,5435	216,72	90,41
	Platô	17	-	-	36,1	53,9712	-	-	
17	Reta	37	45,7024	0,0118	-	3028,9080	3929,1429	424,12	92,37
	Platô	23	-	-	155,8696	900,2349	-	-	
18	Reta	21	14,6705	0,0062	-	197,4340	252,8391	175,85	90,24
	Platô	13	-	-	47,6954	55,4051	-	-	
22	Reta	11	25,1743	0,0167	-	102,8535	125,8332	168,3	94,92
	Platô	4	-	-	66,8575	22,9797	-	-	
27	Reta	13	17,2865	0,0182	-	203,2670	273,3760	204,62	94,9
	Platô	6	-	-	79,2333	70,1090	-	-	
55	Reta	18	11,1652	0,0087	-	85,2734	115,5222	433,58	96,44
	Platô	7	-	-	51,82	30,2488	-	-	
66	Reta	29	11,8755	0,004	-	204,3085	283,8544	265,5	90,77
	Platô	21	-	-	41,1586	79,5459	-	-	
70	Reta	24	30,633	0,0123	-	788,5630	1109,0750	305,3	93,27
	Platô	14	-	-	106,4486	320,5120	-	-	
77	Reta	13	10,0065	0,0115	-	53,5996	86,1479	307,36	96,55
	Platô	6	-	-	48,4833	32,5483	-	-	
100	Reta	11	12,3764	0,0156	-	119,0136	166,5837	127,14	92,92
	Platô	7	-	-	57,0057	47,5701	-	-	
105	Reta	13	11,4262	0,0087	-	64,9500	146,1816	144,45	92,92
	Platô	17	-	-	40,7288	81,2316	-	-	
110	Reta	23	13,1841	0,0083	-	240,8400	345,7745	376,64	94,72
	Platô	13	-	-	61,594	104,9345	-	-	
112	Reta	13	6,015	0,0049	-	20,1950	26,0734	150,72	93,2
	Platô	7	-	-	22,5857	5,8784	-	-	
114	Reta	24	14,9429	0,0037	-	74,6393	102,5828	294,48	93,05
	Platô	15	-	-	29,7427	27,9435	-	-	
116	Reta	28	24,955	0,0091	-	789,3495	1135,7662	314,65	92,37
	Platô	22	-	-	90,00089	346,4167	-	-	
117	Reta	19	8,8287	0,0025	-	27,0506	40,2428	139,69	89,15
	Platô	15	-	-	20,5787	13,1922	-	-	
181	Reta	21	9,4314	0,0044	-	105,8960	184,3187	164,95	89,67
	Platô	21	-	-	32,9352	78,4227	-	-	

A distribuição entre os pontos para cada modelo ficou de forma bem harmônica, ou seja não foi encontrado nenhum fragmento, com a ocorrência de platô no limite máximo da área amostrada.

Verificou-se que, para representar a variável espécie, houve suficiência amostral para os 20 fragmentos amostrados. O número de conglomerados lançados a mais por fragmento, indicou que a amostragem foi bem representativa da vegetação (Tabela 6). Considerando toda a bacia inventariada, a média foi de 4,1 conglomerados a mais por fragmento, ou 12,3 subunidades, representando uma área de 3.075 m². O fragmento que alcançou uma suficiência amostral mais cedo, ou a uma menor área até a formação do platô, foi o 22 (2.500 m²). Porém, o fragmento com maior número de parcelas lançadas a mais, considerando um valor relativo por

área, foi o fragmento 114, que abrangeu cerca de 5.750 m² (58,7%), indicando uma alta amostragem. O crescimento acelerado e contínuo da curva de frequência acumulada, a partir do aumento da amostragem, indica uma diversidade específica de cada parcela. Esta diversidade determina o ponto ideal para se atingir a suficiência amostral, podendo ainda realizar comparações de riqueza de espécies entre os fragmentos, tendo como referência uma dada área amostral.

Na Figura 4, como critério de exemplificação, são demonstradas graficamente, o comportamento da regressão linear de platô, apenas para os fragmentos: 2, 8, 12, 14, 17 e 18. Em que FA *est* indica os valores estimados de FA e FA *real* os dados observados em campo, após o cálculo da distribuição de frequência.

Tabela 6 – Relação da amostragem total realizada por fragmento inventariado, discriminando a suficiência amostral após platô, e a variação no número de conglomerados lançados a mais.

Table 6 – Relation of the sampling size for each inventoried fragment, showing the plateau point, sampling sufficiency after plateau and the variation in the number of the extra clusters.

Fragmento	Amostragem Total	Platô (m ²)	Amostragem necessária		Amostragem realizada além do necessário		
			Subunidade	Conglomerado	(%)	Subunidade	Conglomerado
2	6000	4500	18	6,00	25,00	6	2,00
8	6500	4750	19	6,33	26,92	7	2,33
12	7500	5000	20	6,67	33,33	10	3,33
14	10500	6500	26	8,67	38,10	16	5,33
17	15000	9500	38	12,67	36,67	22	7,33
18	8500	5500	22	7,33	35,29	12	4,00
22	3750	2500	10	3,33	33,33	5	1,67
27	4750	3500	14	4,67	26,32	5	1,67
55	6250	4750	19	6,33	24,00	6	2,00
66	12500	7500	30	10,00	40,00	20	6,67
70	9500	6250	25	8,33	34,21	13	4,33
77	4750	3500	14	4,67	26,32	5	1,67
100	4500	3000	12	4,00	33,33	6	2,00
105	7500	3500	14	4,67	53,33	16	5,33
110	9000	6000	24	8,00	33,33	12	4,00
112	5000	3500	14	4,67	30,00	6	2,00
114	9750	4000	16	5,33	58,97	23	7,67
116	12500	7250	29	9,67	42,00	21	7,00
117	8500	4750	19	6,33	44,12	15	5,00
181	10500	5500	22	7,33	47,62	20	6,67

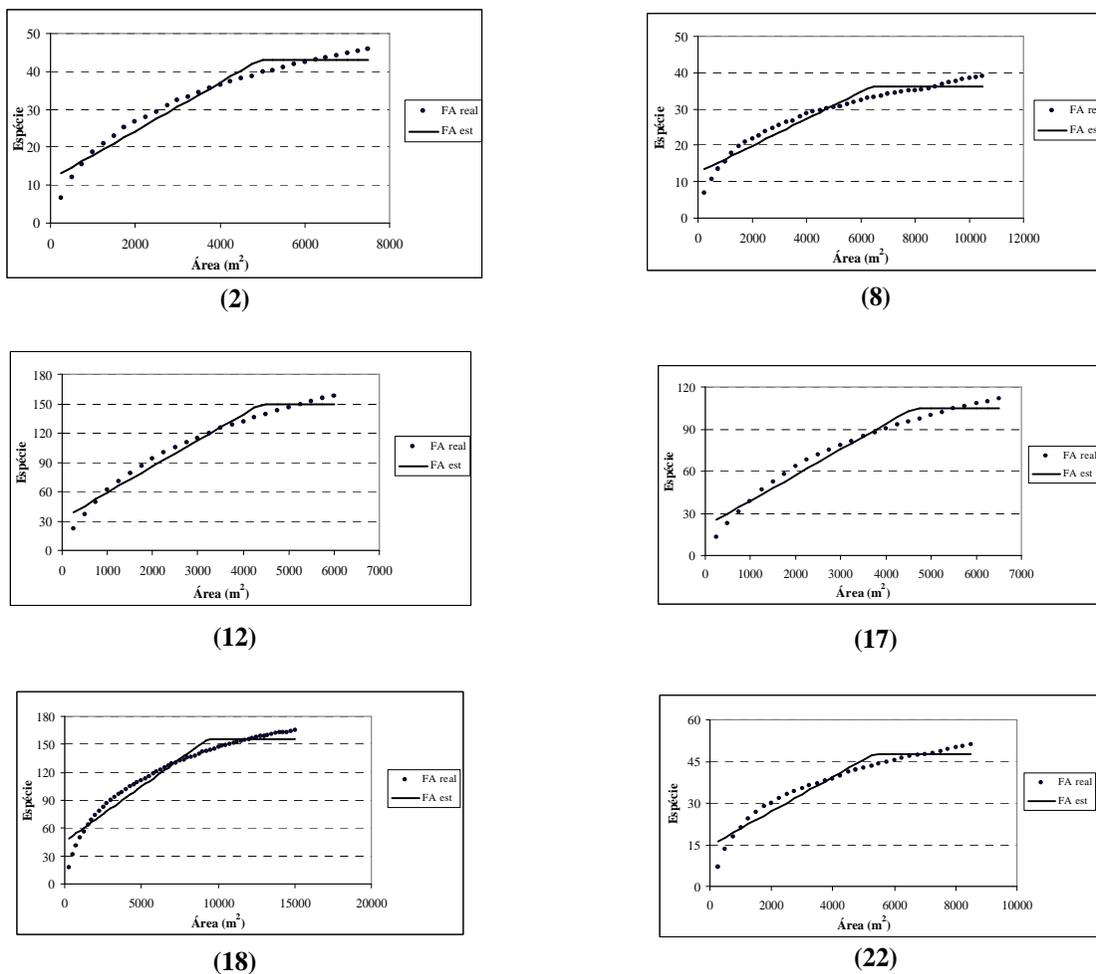


Figura 4 – Gráficos representativos do comportamento da amostragem, indicando o ponto de suficiência amostral (X_0), após o encontro da equação linear com a de platô.

Figure 4 – Graphic representation of the sampling behaviors, indicating the plateau point (X_0), after the crossing of the linear regression and the plateau regression.

A determinação desta suficiência garante uma análise a posteriori confiável, sobre padrões de diversidade, similaridade, equabilidade e conhecimento da estrutura da comunidade arbórea, sendo primordial quando se pretende correlacionar vegetações de forma geral, permitindo ainda embasamento nas análises da vegetação.

A união das retas gerada pelo modelo linear, em conjunto com a regressão de platô, expressa a suficiência amostral. Assim, deste ponto (X_0) em diante, a amostragem se faz suficiente, não havendo mais necessidade de lançar novas unidades amostrais.

Uma vez que a maioria das espécies recorrentes de determinado ambiente foi amostrada

e se atingiu a suficiência amostral, o comportamento da curva de frequência acumulada, para qualquer fisionomia, tende ao comportamento quase assintótico, ou seja, o aumento da área amostral reproduz ganhos de espécies continuamente, até o completo levantamento (censo) da riqueza total de espécie da área. Contudo, após a formação do platô, o aumento de informações de espécie/parcela é reduzido e economicamente desinteressante.

Em se tratando da diversidade real de uma vegetação, somente a realização de um censo na área reproduziria esta variável, dispensando, neste momento, os princípios da aplicação da amostragem. Porém, de acordo com Kenkel et al. (1989) e Orlóci (1993), a alta diversidade gera um grande número de interações entre indivíduos e ambiente, o que torna complexo amostrar fisionomias. O ponto de maior reflexo na amostragem talvez seja a formação de arranjos não-aleatórios espacialmente distribuídos na área pelas espécies, ou a distribuição espacial das espécies.

3.3 Tradicional x Nova abordagem

Em levantamentos ecológicos, as soluções clássicas para a sua determinação não são adequadas, pois estas assumem um universo mais heterogêneo (PILLAR, 1998). Negreiros (1982) e Scolforo & Mello (1997) comentam que a suficiência amostral, ou área mínima de amostragem, influi decisivamente na análise da vegetação.

O emprego de técnicas como a Regressão linear de Platô é uma alternativa viável, para se encontrar a suficiência amostral, porém deve-se utilizar com critérios. Ao se comparar o método tradicional de ajuste, com a proposta deste trabalho, percebeu-se que variações de resultados foram encontradas nos fragmentos 22, 27, 77, 100 e 112, indicando que o método tradicional pode deixar dúvidas, caso o comportamento dos dados apresentarem um ganho contínuo no número de espécies, ao longo da área amostral. Em alguns pontos pode-se chegar até a uma “falsa” suficiência, fato não observado ao se realizar ajustes com a nova abordagem.

Os fragmentos 22, 27 e 77, por sua vez, atingiram uma suficiência amostral próxima do número lançado na área, entretanto, com um excedente em área após a suficiência amostral, que permaneceu em 1.250 m² ou 5 subunidades de 250 m². Para o método

tradicional, a performance da regressão ficou mais próxima do limite máximo de área amostral, não havendo sobra em subunidades amostrais.

O teorema do limite central auxilia na padronização das combinações selecionadas, ao se selecionar a média dos sorteios realizados, fato não esperado ao se realizar apenas um único sorteio, e ser possível de se obter uma combinação amostral não representativa da população, gerando dúvidas sobre o ponto X_0 .

4 CONCLUSÕES

Um possível “falso” platô pode ser obtido, ao se utilizar o método tradicional em futuros trabalhos.

A aplicação do teorema do limite central contribuiu na diluição de possíveis tendenciosidades geradas pela combinação dos dados, ou dúvidas referentes à determinação do ponto X_0 , o que de alguma forma garante a confiabilidade da informação adquirida.

Devido a não interferência do operador nos resultados, esta proposta metodológica pode ser utilizada como uma alternativa na determinação da suficiência amostral, frente ao método tradicional.

A amostragem realizada foi suficiente para todos os fragmentos inventariados, pelo método proposto. Porém, o mesmo não foi encontrado pelo método tradicional, deixando dúvidas nos fragmentos: 22, 27, 77, 100 e 112.

5 REFERÊNCIAS BIBLIOGRÁFICAS

- ALVAREZ, V. H. V. *Avaliação da fertilidade do solo: superfície de resposta, modelos aproximativos para expressar a relação fator-resposta*. Viçosa: UFV, 1985. 75 p.
- BRAGA, J. M. *Avaliação da fertilidade do solo. Ensaio de campo*. Viçosa: UFV, 1983. 101 p.
- BUSSAB, W. O. *Análise de variância e de regressão: uma introdução*. 2. ed. São Paulo: Atual, 1988. 147 p.
- CAMARGO, F. M. *Caracterização da vegetação lenhosa e dos solos de um mosaico de cerrado, floresta semidecídua, floresta decídua em Bocaiúva-MG*. 1997. 55 f. Dissertação (Mestrado em Engenharia Florestal) – Universidade Federal de Lavras, Lavras, 1997.

- COCHRAN, W. G. **Técnicas de amostragem**. Rio de Janeiro: USAID, 1963. 555 p.
- FERREIRA, D. F. **Uso de recursos computacionais**. Lavras: UFLA, 2005. 123 f. Apostila.
- GOMIDE, L. R. **Um modelo fitogeográfico para a bacia do rio São Francisco, em Minas Gerais**. 2004. 268 p. Dissertação (Mestrado em Florestas de Produção) - Universidade Federal de Lavras, Lavras, 2004.
- GOMIDE, L. R.; SCOLFORO, J. R. S.; OLIVEIRA, A. D.; THIERSCH, C. R. Suficiência amostral para fragmentos florestais da bacia do rio São Francisco. In: SCOLFORO, J. R. S.; OLIVEIRA, A. D. **Um modelo fitogeográfico para as áreas de preservação permanente da bacia do rio São Francisco, em Minas Gerais**. Lavras: UFLA, 2005. p. 128-136.
- JUDD, W. S.; CAMPBELL, C. S.; KELLOG, E. A.; STEAVENS, P. F. **Plant systematics: a phylogenetic approach**. Massachusetts: Sinauer, 1999.
- KENKEL, N. C.; JUNÁSZ-NAGY, P.; PODANI, J. On sampling procedure in population and community ecology. **Vegetatio**, Dordrecht, v. 83, n. 1/2, p. 195-207, Oct. 1989.
- LIMA, C. S. A. **Desenvolvimento de um modelo para manejo sustentado do cerrado**. 1997. 159 f. Dissertação (Mestrado em Engenharia Florestal) - Universidade Federal de Lavras, Lavras, 1997.
- NAPPO, M. E.; FONTES, M. A. L.; OLIVEIRA-FILHO, A. T. Suficiência amostral e análise do tamanho de parcelas para o estudo da regeneração natural do sub-bosque de povoamentos homogêneos de *Mimosa scabrella* Benth., em área minerada em Poços de Caldas, MG. **Revista Árvore**, Viçosa, v. 23, n. 4, p. 443-453, 1999.
- NEGREIROS, O. C. **Características fitossociológicas de uma comunidade de florestas latifoliada pluviosa tropical visando ao manejo do palmito, *Euterpe edulis* mart.** 1982. 104 f. Dissertação (Mestrado) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 1982.
- ORLÓCI, L. The complexities and scenarios of ecosystem analysis. In: PATIAL, G. P.; RAO, C. R. (Eds.). **Multivariate environmental statistics**. Amsterdam: North-Holland, 1993. p. 423-432.
- PILLAR, V. D. Sampling sufficiency in ecological surveys. **Abstract Botanica**, Budapest, v. 22, p. 37-48, 1998.
- SATO, Y.; GODINHO, H. P. Peixes da bacia do rio São Francisco. In: LOWEMcCONNEL (Ed.). **Estudos ecológicos de comunidades de peixes tropicais**. Tradução de A. E. A. M. Vazzoler, A. A. Agostinho e P. T. M. Cunhingham. São Paulo: EDUSP, 1999. p. 401-413.
- SCOLFORO, J. R. S.; MELLO, J. M. de. **Inventário florestal**. Lavras: UFLA/FAEPE, 1997. 341 p.
- SISTEMA PARA ANÁLISE ESTATÍSTICA E GENÉTICA. **Manual de uso**. Viçosa: UFV-Funarbe, 1997.
- VASCONCELOS, P. C. S. **Fitossociologia de uma vegetação em sucessão secundária, no Vale do Paraíba, São Paulo**. 1992. 116 f. Dissertação (Mestrado) - Universidade Federal de Viçosa, Viçosa, 1992.
- VOLPATO, M. M. L. **Regeneração natural em uma floresta secundária no domínio de mata Atlântica: uma análise fitossociológica**. 1994. 123 f. Dissertação (Mestrado em Engenharia Florestal) - Universidade Federal de Viçosa, Viçosa, 1994.
- WALLACE, T. D. **Econometrics: an introduction**. Mass: Addison-Wesley, 1988. 440 p.