Paula Alves Condé

# MODELOS TEÓRICOS DE DISTRIBUIÇÃO DE ABUNDÂNCIA DAS ESPÉCIES PARA CARACTERIZAÇÃO DA DIVERSIDADE BIOLÓGICA E A UTILIZAÇÃO DE DIFERENTES MEDIDAS DE ABUNDÂNCIA

## THEORETICAL MODELS OF SPECIES ABUNDANCE DISTRIBUTION TO CHARACTERIZE THE BIOLOGICAL DIVERSITY AND THE USE OF DIFFERENT MEASURES OF ABUNDANCE

São Paulo 2012 Paula Alves Condé

# MODELOS TEÓRICOS DE DISTRIBUIÇÃO DE ABUNDÂNCIA DAS ESPÉCIES PARA CARACTERIZAÇÃO DA DIVERSIDADE BIOLÓGICA E A UTILIZAÇÃO DE DIFERENTES MEDIDAS DE ABUNDÂNCIA

## THEORETICAL MODELS OF SPECIES ABUNDANCE DISTRIBUTION TO CHARACTERIZE THE BIOLOGICAL DIVERSITY AND THE USE OF DIFFERENT MEASURES OF ABUNDANCE

Dissertação apresentada ao Instituto de Biociências da Universidade de São Paulo, para a obtenção de Título de Mestre em Ecologia, na Área de Ecossistemas Terrestres.

Orientador(a): Prof. Dr. Paulo Inácio de Knegt López de Prado

Versão corrigida Original disponível no Instituto de Biociências da USP

São Paulo

2012

Condé, Paula Alves Modelos teóricos de distribuição de abundância das espécies para caracterização da diversidade biológica e a utilização de diferentes medidas de abundância 102 páginas

Dissertação (Mestrado) - Instituto de Biociências da Universidade de São Paulo. Departamento de Ecologia.

1. Distribuição de abundância das espécies 2. Medida de abundância 3. Amostragem I. Universidade de São Paulo. Instituto de Biociências. Departamento de Ecologia.

Comissão Julgadora:

Prof<sup>a</sup>. Dr<sup>a</sup>. Malva Isabel Medina Hernández

Prof. Dr. João Luís Ferreira Batista

Prof. Dr. Paulo Inácio de Knegt López de Prado Orientador

## Dedicatória

Aos meus queridos pais ...pelo carinho sempre constante, conforto e suporte. As palavras muitas vezes não são suficientes para expressar o sentimento de carinho e gratidão a tantas pessoas especiais e fundamentais. Mas começo agradecendo saudosamente à minha família. Em primeiro lugar agradeço às duas pessoas que conquistaram minha primeira admiração na vida, meu pai Paulo Afonso e minha mãe Léa. Agradeço não só ao cuidado no ninho e amor incondicional, mas ao jeito simples e carinhoso de levar a vida. Também às minhas primeiras amizades e companhias, meus queridos irmãos Camila, Gabriel e Clara que preencheram não apenas nossa casa com a alegria, mas também meu coração... e agradeço por proporcionarem outros grandes presentes: minhas sobrinhas Ana Maria e Luciana, borboletinhas alegres e coloridas, e meus cunhados Marcos Paulo e Renata, amigos valiosos. Agradeço aos meus avós Sebastião e Terezinha, plantadores de frutas, hortas e jardins, por manterem meu melhor refúgio de paz e alimentação no interior de Minas. Também aos tios e primos queridos que colorem ainda mais esse recanto familiar.

Agradeço principalmente ao meu orientador Paulo Inácio Prado pelas inúmeras atenções e confiança depositada. Por tornar possível a conquista desta etapa e orientar a execução desse trabalho. A segurança transmitida, reuniões e discussões dedicadas à admirável tarefa de orientação foram fundamentais em todo o processo e merece minha mais sincera gratidão. Agradeço ao seu jeito mineiro, conhecimento teórico, metodológico e enciclopédico, além da forma paciente de conduzir as dedicadas orientações.

A professora Dra. Malva Hernández tem meu reconhecimento, gratidão e admiração pelo importante apoio e incentivo no processo de ingresso na pósgraduação, me confiando inclusive dados de sua pesquisa. Agradeço pelos dados cedidos e pela marcante participação durante minha formação na graduação, com uma maneira sempre alegre, paciente e carinhosa de ser.

Agradeço aos professores membros da LAGE, Glauco Machado e Paulo Guimarães, pela importante participação nas atividades e seminários do laboratório, e pela valorosa contribuição em minha formação, principalmente através do curso de campo e convivência no lab. Agradeço também aos demais professores que contribuíram em conhecimentos teóricos e metodológicos através das disciplinas, e aos professores membros da minha banca de qualificação: Renata Pardini, Roberto Kraenkel e Roberto Shimizu.

Ao CNPq pela bolsa inicial de estudo e à FAPESP pelo financiamento do projeto.

Ao Maurício Barbosa pelo patrocínio de milhas na ponte aérea Rio(Branco)-São Paulo.

Ao retiro científico do LET, que além de proporcionar um clima de motivação, produtividade e gastronomia gourmet, foi um impulso fundamental para a execução das análises e códigos em R.

Agradeço a alguns colegas de trabalho que colaboraram de diferentes formas para possibilitar minha estadia junto à Universidade de São Paulo: Antônio Bastos Nunes (*in memorian*), Diogo Selhorst, Elaine Oliveira, Júlio Resende, Antônio Marreiro, Keide Bessa, Simone Alencar, Evandro Camara e Nara Pantoja.

Amigos... são refresco da vida, sorriso na chegada e conforto na partida. Agradeço muito a esses presentes que encontramos pela vida, os bons, velhos e novos amigos... por sempre compartilharem agradáveis momentos e festejarem os bons acontecimentos. Os dias não teriam tantas cores sem vocês.

A todos os membros da LAGE por manterem um ambiente de trabalho agradável e proporcionarem momentos de descontração, cafés, chás, mates, almoços coletivos... e marcantes amizades. Lembrando principalmente aos irmãos de orientador, amigos do LET: Camila Mandai, por me ensinar a importância de usar perneiras, pela demasiada sinceridade e mau-humor aleatório, pelas ervas de mate e dias de snake-car, e pela simpática companhia nos momentos finais de correção do trabalho. Ao Rodolpho Rodrigues, por ser um mano prestativo e da hora, pelos auxílios com o R, amizade e café colonial (mesmo que escasso). À Camila Castanho pela amizade, presença alegre e sempre animada e por ser nosso exemplo de atleta e dedicação. Ao Renato Lima pela amizade e por fazer eu me sentir normal. À querida amiga Sara Mortara por sua incrível habilidade ilustrativa, por ser uma pessoa gourmet imprescindível para o lab, pela amizade, bom humor e alegria contagiante. Ao André Chalom pelo humor inteligente com expressão facial inalterável, pela linguagem didática do mundo matemático, pelo seminário de SVN e síntese de hipercubos latinos e por fazer eu me sentir mais normal ainda.

Agradecimento muito especial às minhas queridas sisters de campo, Rachel Werneck e Babi Henning, pelo presente de suas amizades e agradáveis sisjapas, pela parceria e motivação nos momentos precisos, por me apresentarem uma pequena parte desta mega-cidade e compartilharem os melhores momentos da vida no Brasil. À Rachel Werneck agradeço particularmente pelo acolhimento e aquecimento em São Paulo... o cantinho pra dormir e aquecedor para o lab foram fundamentais para minha aclimatação neste país.

Às companheiras de república, pela companhia e bons momentos compartilhados, chás da madrugada e bulling coletivo com argentinas... Julia Beneti, agradeço pela presença alegre, pelos comentários aleatórios e repertório de cultura útil e inútil. À Nati Nery pela amizade e informações gerais da vida neste planeta e pelas jantas da madrugada. À Nara Leão, pela presença silenciosa e pensamentos malignos.

À Ruli Astegiano, por ser sempre mucho más engraçada, e me revelar o quanto vale a pena e é divertido ter amigos, mesmo que sejam argentinos. Agradeço pela amizade e companhia, pelas milanesas e sobremesas, e principalmente pela disponibilidade de bulling e lista de pérolas e mocassos compartilhados para tornar meus dias mais alegres. À Esther Sebastian, pela agradável e responsável companhia, pela amizade e cultura brasileira, e pela deliciosa galinhada com omelete (ops! Paella com tortillas).

Agradeço também aos amigos de longa data, que mantêm viva as memórias de cada boa fase da vida. Em especial: Fernanda Mura, Anna Catarina, Carolina Condé, Iara Silveira & Cia, Mariana Fagundes e Sara Melo.

E por fim, agradeço enormemente ao meu amigo e amor, Rafael Barbosa, pelo companheirismo e conforto de todas as horas, pela paciência e carinho, e por ter sempre muitos bons motivos para pensar em você.

À todos que fazem parte da minha amostra humana de diversidade, histórias e cores, sintam-se agradecidos, mesmo que agora possa ter me falhado a memória.

### SUMÁRIO

RESUMO	10
ABSTRACT	12
Apresentação	14
Referências	16

#### 1 O USO DOS MODELOS DE DISTRIBUIÇÃO DE ABUNDÂNCIA DE ESPÉCIES

1.1	Introdução
1.2	Breve histórico e considerações teóricas relevantes das SADs
1.3	Abordagem da verossimilhança para o ajuste e seleção de modelos
1.4	A utilização da biomassa como medida de abundância nos modelos de SAD 25
1.5	Relevância dos besouros escarabeíneos para estudos em ecologia
1.6	Considerações finais
Refe	erências

# **2** Avaliação da eficiência da medida de abundância nos modelos de distribuição de abundância das espécies: Um estudo empírico em comunidades de besouros scarabaeinae

2.1	Introd	ução
2.2	Mater	rial e Métodos
	2.2.1	Descrição dos dados
	2.2.2	Análises estatísticas
		Ajuste e seleção dos modelos
		Rarefação das amostras
		Medida da probabilidade de evidência enganadora – PME
		Estimativa do viés e precisão das estimativas dos parâmetros
2.3	Resul	tados
	2.3.1	SADs das amostras completas de cada área
		<i>Vale</i>
		<i>Vale Alterado</i>
		<i>Encosta Sul</i>
		<i>Encosta Norte</i>

<i>Topo Morro</i>	62
Eucaliptus	65
2.3.2 Probabilidade de evidência enganadora - PME	68
2.3.3 Viés da estimativa dos parâmetros	78
2.3.4 Coeficiente de variação da estimativa dos parâmetros	84
2.3.5 Comparação dos resultados	89
2.4 Discussão e Conclusões	90
Referências	99
Anexos 10	02

#### RESUMO

As distribuições de abundância das espécies (SADs) são consideradas uma maneira informativa e sintética de descrever a diversidade biológica, e revelam um dos padrões mais bem estabelecidos da ecologia, que é a predominância de espécies raras nas comunidades biológicas. A generalidade deste padrão o torna relevante para a análise de outros parâmetros das comunidades. Além disso, ele levou ao desenvolvimento dos modelos teóricos de SAD. Os modelos estatísticos de SAD, por sua vez, passaram a ocupar um papel central na biologia, como principio empírico que sustenta várias teorias ecológicas. Preston sugeriu em 1948 que as SADs das comunidades naturais teriam uma distribuição aproximadamente Lognormal, porém apareciam truncadas devido ao efeito do tamanho amostral, cuja forma real só seria revelada, portanto, em amostras grandes. Esta ideia vai ao encontro da propriedade estatística da consistência, que implica que o acúmulo de evidência causado pelo aumento do tamanho amostral favorece a aproximação ao modelo verdadeiro, bem como às estimativas de seus parâmetros. Além do efeito do tamanho amostral, os padrões de distribuição de abundância podem diferir dependendo da medida de abundância utilizada. Estudos sugerem que a biomassa seria uma medida de abundância mais adequada para revelar a distribuição subjacente de uma comunidade biológica em amostras ou escalas menores (differential veiling). Neste contexto, nosso objetivo principal neste trabalho foi investigar e discutir a vantagem de considerarmos a biomassa como medida de abundância nos modelos teóricos de distribuição de abundância das espécies. Avaliamos, então, o efeito do tamanho da amostra sobre a qualidade dos ajustes dos modelos sob duas diferentes medidas de abundância: biomassa e número de indivíduos. Simulamos amostras de diferentes tamanhos a partir de amostragens aleatórias de uma base de dados empírica e as ajustamos a diferentes modelos de SADs. Com isso, pudemos avaliar a eficiência das amostras com cada medida de abundância - em revelar o modelo correto, bem como a acurácia e precisão das estimativas dos parâmetros desses modelos. Para o presente estudo, utilizamos um grupo com reconhecido potencial indicador e relevância para estudos ecológicos, os besouros Scarabaeinae. Os resultados encontrados neste estudo mostram que a maior eficiência da biomassa para revelar a distribuição subjacente não é tão geral quanto sugerem os trabalhos anteriores. Os critérios de análise

utilizados não favoreceram consistentemente a biomassa como medida mais eficiente em revelar o modelo correto e apresentar maior acurácia e precisão nas estimativas de seus parâmetros. Assim, a afirmativa de que a SAD de biomassa não é oculta (*veiled*) em escalas e amostras menores não parece ser geral. Os resultados obtidos no presente estudo também indicam que o padrão de *differential veiling* entre as SADs de biomassa e número de indivíduos podem variar dependendo da escolha do modelo e do conjunto de dados avaliados. No entanto, se a causa do *differential veiling* entre as SADs de biomassa e número de indivíduos for principalmente devido ao efeito de escala, uma explicação alternativa é que a escala do presente estudo teria que ser ampliada para podermos demonstrá-lo. Considerando então, os efeitos de amostragem apontados pelos nossos resultados sobre a medida de abundância, destacamos que, apesar da importância do assunto, há uma escassez de estudos que investiguem o uso da biomassa como medida de abundância nas SADs e as implicações de diferentes efeitos sobre esta medida destacam a necessidade de estudos adicionais que possibilitem isolar o efeito de escala do efeito de tamanho amostral.

Palavras-chave: Distribuição de abundância das espécies, medidas de abundância, amostragem, linha de véu, Scarabaeinae, biomassa.

#### ABSTRACT

The species abundance distributions (SADs) are considered an informative way to describe the biological diversity revealling one of the most wellestablished patterns in ecology: the predominance of rare species in biological communities. The generality of such pattern made it relevant to analyze other biodiversity parameters and to induce the development of theoretical models of SAD. On the other hand, statistical models of SAD occupied a central role in biology as an empirical principle that underlies many ecological theories. Was suggested that SADs follow an approximately lognormal distribution, but that may appear truncated by the effect of sample size, being completely revealed only in large samples. This idea is associated with the statistical property of consistency, which implies that the accumulation of evidence - represented by the increase in sample size - leads the samples distribution to the true model, as well as the estimates of the parameters to their real values. Beyond the effect of sample size, the SADs may differ depending on the measure of abundance adopted. Studies suggest that biomass could be a more efficient measure of abundance to reveal the underlying distribution of biological communities in smaller samples or scales (differential veiling). In this context, our aim in this study was to investigate and discuss the advantage of considering biomass in theoretical models of SAD. Thus, we evaluated the effect of sample size on the quality of models fitness under two different measures of species abundance: biomass and number of individuals. We simulated samples of different sizes taken from an empirical database of Scarabaeinae beetles. Furthermore, we adjusted the simulated samples to different SAD models and evaluated the efficiency of each kind of abundance index to reveal the correct model, as well as the accuracy and precision of the parameters estimates. Our results show that the efficiency of biomass in revealing the underlying distribution is not as general as suggested by previous works. According to our analysis criteria, biomass was not far more efficient in revealing the correct model or in providing greater accuracy and precision in parameters estimates than the number of individuals was. Thus, the statement that the SAD biomass is not veiled on small scales or samples does not seems to be general. Moreover, our results also indicate that the effect of differential veiling on SADs using biomass and number of individuals depends on the choice of the evaluated model and data set. However, if the cause of the differential veiling is mainly due to scale effects, an alternative explanation is that the scale of this study was not wide enough to show it. Considering the sampling effects studied here in biodiversity evaluation we highlight the lack of studies investigating the use of biomass in SADs and implications of different effects on this measure indicate the urgent call for further studies on this subject which enable to isolate the effect of scale from the sample size.

Keywords: Species abundance distributions, measure of abundance, sampling, veil line, Scarabaeinae, biomass.

#### **Apresentação**

Além de entender como os conjuntos de espécies são distribuídos na natureza, outro grande desafio para os ecólogos é buscar métodos consistentes de análises que possibilitem extrair e extrapolar informações através dos padrões de distribuição encontrados (Gaston 2000). As distribuições de abundância das espécies (SAD<sup>1</sup>) são consideradas uma maneira informativa e sintética de descrever a diversidade biológica, e revelam um dos padrões mais bem estabelecidos da ecologia, que é a predominância de espécies raras nas comunidades naturais (McGill *et al.* 2007; Prado 2010).

A generalidade desse padrão elevou-o a princípio empírico que sustenta outras teorias, e também a atributo básico das comunidades, que pode ser usado para descrevê-las e compará-las. Dado isto, desenvolveram-se modelos matemáticos de distribuição de abundância das espécies, que posteriormente foram criticados por não explicitarem os mecanismos que explicam o surgimento destes padrões (Hubbell 2001). No entanto, a grande utilidade das SADs em descrever e comparar comunidades, independentemente da composição de suas espécies, representa uma vantagem metodológica do uso dos modelos como ferramenta de acesso e monitoramento da diversidade biológica (Dornelas *et al.* 2009) e sustentou o papel histórico que as SADs ocuparam no desenvolvimento da ecologia de comunidades (Hubbell 2001; McGill 2003).

Os estudos das SADs são realizados a partir de amostras da comunidade, tentando chegar a conclusões relacionadas com a quantidade e a distribuição dos organismos que fazem parte daquela comunidade a partir de um modelo teórico de

<sup>&</sup>lt;sup>1</sup> Utilizaremos a sigla tradicionalmente usada na literatura (SAD – *species abundance distributions*) para referir à distribuição de abundância das espécies.

distribuição de probabilidade (Magurran 2004). No entanto, a SAD é notoriamente afetada pelo tamanho da amostra, sendo que as amostras maiores tendem a se aproximar da distribuição real da comunidade (Preston 1948; Pielou 1977).

Além do efeito da amostragem, os padrões de distribuição de abundância podem diferir dependendo da medida de abundância utilizada (Connolly *et al.* 2005; Saint-Germain *et al.* 2007). A medida de abundância tradicionalmente utilizada para medir a importância relativa das espécies é o número de indivíduos, sendo o uso da biomassa e suas implicações nas SADs ainda escassos na literatura (Morlon *et al.* 2009). Os estudos de Connolly *et al.* (2005) e Morlon *et al.* (2009) encontraram padrões em que as SADs das comunidades de árvores, aves, corais e peixes foram melhor reveladas em escalas menores com o uso da biomassa como medida de abundância. Esses resultados, embora interessantes e inéditos, não foram amplamente investigados e compreendidos na literatura, mesmo apresentando importantes implicações teóricas e aplicadas (Williamson 2010).

Assim, considerando a utilidade dos modelos de distribuição de abundância das espécies como descritores das comunidades naturais e a compreensão do uso da biomassa como medida de abundância nestes modelos, o objetivo deste estudo foi avaliar a eficiência da biomassa e número de indivíduos como medida de abundância para revelar a distribuição subjacente das comunidades e discutir as implicações do uso da biomassa nos estudos das SADs.

Para representar a eficiência das medidas de abundância em revelar a SAD das comunidades, utilizamos critérios de avaliação da consistência dos modelos (probabilidade de evidência enganadora) e da acurácia e precisão das estimativas de máxima verossimilhança de seus parâmetros, representadas pelo viés e coeficiente de variação das estimativas dos parâmetros, respectivamente. Esses critérios foram

15

avaliados em relação à redução do tamanho amostral. Para isto, modelos de SADs para abundância de biomassa e número de indivíduos foram ajustados a amostras de tamanhos crescentes, simuladas em computador por rarefação da amostra empírica das comunidades de besouros Scarabaeinae (Coleoptera: Scarabaeidae) de Mata Atlântica utilizadas para o estudo. A avaliação da eficiência da medida de abundância para descrever a SAD da comunidade foi feita através de uma abordagem comparativa entre as duas medidas. Desta forma, se a biomassa for mais eficiente em revelar a SAD da comunidade com a diminuição do tamanho amostral, é esperado que as medidas utilizadas para avaliar a consistência dos modelos e o viés e precisão das estimativas de seus parâmetros, tenham valores consistentemente menores quando comparados aos valores dos modelos de SAD com número de indivíduos, para os mesmos tamanhos de amostra.

Este trabalho está organizada em dois capítulos: o primeiro capítulo consiste em uma revisão sobre o uso dos modelos teóricos de distribuição de abundância das espécies para caracterização da diversidade biológica e o uso da biomassa como medida de abundância nestes estudos. O objetivo deste capítulo é introduzir as bases teóricas e metodológicas que fundamentam a abordagem do segundo capítulo. O segundo capítulo consiste na avaliação empírica comparando a eficiência das medidas de abundância (biomassa e número de indivíduos) em revelar a distribuição subjacente de comunidades de besouros Scarabaeinae de Mata Atlântica com a diminuição do tamanho amostral.

#### Referências

Connolly, S. R.; Hughes, T. P.; Bellwood, D. R. & Karlson, R. H. 2005. Community structure of corals and reef fishes at multiple scales. Science 309: 1363-1365.

- Dornelas, M.; Moonen, A. C.; Magurran, A. E. & Bàrberi, P. 2009. Species abundance distributions reveal environmental heterogeneity in modified landscapes. Journal of Applied Ecology 46: 666-672.
- Gaston, K. J. 2000. Global patterns in biodiversity. Nature 405: 220-227.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press.
- Magurran, A. E. 2004. Measuring Biological Diversity. Blackwell Publishing.
- McGill, B. J. 2003. Strong and weak tests of macroecological theory. Oikos 102: 679-685.
- McGill, B. J.; Etienne, R. S.; Gray, J. S.; Alonso, D.; Anderson, M. J.; Benecha, H. K.; Dornelas, M.; Enquist, B. J.; Green, J. L.; He, F.; Hurlbert, A. H.; Magurran, A. E.; Marquet, P. A.; Maurer, B. A.; Ostling, A.; Soykan, C. U.; Ugland, K. I. & White, E. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecology Letters 10: 995-1015.
- Morlon, H.; White, E. P.; Etienne, R. S.; Green, J. L.; Ostling, A.; Alonso, D.;
  Enquist, B. J.; He, F. L.; Hurlbert, A.; Magurran, A. E.; Maurer, B. A.;
  McGill, B. J.; Olff, H.; Storch, D. & Zillio, T. 2009. Taking species abundance distributions beyond individuals. Ecology Letters 12: 488-501.
- Prado, P. I. 2010. Distribuições de abundâncias de espécies: avanços analíticos para entender um padrão básico em ecologia. Ciência e Ambiente 39: 121-136.
- Preston, F. W. 1948. The commonness, and the rarity, of species. Ecology 29: 254-283.
- Pielou, E. C. 1977. Mathematical Ecology. John Wiley and Sons.
- Saint-Germain, M.; Buddle, C. M.; Larrivée, M.; Mercado, A.; Motchula, T.; Reichert, E.; Sackett, T. E.; Sylvain, Z. & Webb, A. 2007. Should biomass be considered more frequently as a currency in terrestrial arthropod community analysis? Journal of Applied Ecology 44: 330-339.
- Williamson, M. 2010. Why do species abundance distributions of individuals and biomass behave differently under sampling? Oikos 119: 1697-1699.

#### **1 O** USO DOS MODELOS DE DISTRIBUIÇÃO DE ABUNDÂNCIA DE ESPÉCIES

#### 1.1 Introdução

As observações de que as espécies não são uniformemente distribuídas na natureza podem ser encontradas na literatura desde Darwin (1859) e levaram a uma série de investigações sobre os padrões de distribuição de abundância das espécies. A distribuição de abundância de espécies (SAD) segue uma das leis mais antigas e universais da ecologia: a dominância numérica de poucas espécies e a presença de várias espécies raras nas comunidades biológicas e em suas amostras (McGill *et al.* 2007). Há pouquíssimas exceções a este padrão, o que o torna relevante para a análise de quaisquer outros parâmetros das comunidades (Prado 2010).

Assim, quando as distribuições das frequências relativas das espécies de uma comunidade são representadas em um histograma de abundância (Figura 1a), é revelado um padrão de curva côncava conhecida na literatura como "hollow curve" (Dewdney 2000). A generalidade desse padrão sugere que as SADs refletem princípios subjacentes comuns que estruturam as comunidades ecológicas e propulsionou o desenvolvimento dos modelos de distribuição de abundância das espécies na ecologia (Magurran 2004). Os modelos matemáticos de SAD ocuparam então um papel central na biologia, como principio empírico que sustenta várias teorias ecológicas. Além disso, as SADs permitem uma análise muito mais informativa de padrões de diversidade do que índices sintéticos de riqueza e diversidade, apresentando enorme potencial aplicado no monitoramento e acesso à diversidade biológica (McGill *et al.* 2007; Dornelas *et al.* 2009).

Uma abordagem amplamente utilizada para quantificar e explorar as SADs e descrever os padrões de diversidade nas comunidades são os métodos gráficos e visuais (ver Figura 1) (McGill 2011). Há vários métodos desenvolvidos para esta finalidade (e.g. Whittaker 1960; Kempton & Taylor 1978; Lambshead *et al.* 1983), e a forma mais comum de apresentar as SADs é por meio de um diagrama de rank-abundância ou diagrama de Whittaker (Figura 1b), onde os valores de log-abundância são plotados no eixo Y, e o rank das espécies de maior à menor abundância são ordenadas no eixo X (Magurran 2004).



**Figura 1.** Três diferentes formas de apresentar os dados de distribuição de abundância das espécies: (a) histograma representando as classes de frequência de abundância das espécies em escala aritmética; (b) clássico rank de distribuição de abundância das espécies (em escala logarítmica); (c) gráfico com a função de distribuição cumulativa empírica de abundância das espécies. Os dados são da abundância numérica de besouros Scarabaeinae de Mata Atlântica coletados entre Julho/2007 à Junho/2008 no Parque Municipal da Lagoa do Peri, Florianópolis, Santa Catarina.

#### 1.2 Breve histórico e considerações teóricas relevantes das SADs

Os modelos da série geométrica de Motomura (1932), da série logarítmica de Fisher *et al.* (1943) e da lognormal de Preston (1948) foram as primeiras tentativas de descrever matematicamente a relação entre o número de espécies e suas

abundâncias relativas. O modelo de Motomura (1932) propõe a distribuição de série geométrica, cuja ideia serviu de base para o modelo *broken-stick* de MacArthur (1957) e outros modelos orientados pelo conceito de divisão de um nicho comum limitado (Sugihara 1980; Tokeshi 1990; 1996), que assumem a divisão de nicho como processo dominante influenciando na abundância das espécies.

Já os estudos de Fisher et al. (1943) e Preston (1948), por diferentes formas, incorporaram pioneiramente a preocupação com o efeito da amostragem na representação das SADs das comunidades. Fisher et al. (1943) propuseram o modelo da série logarítmica para explicar a SAD obtida de amostras das comunidades biológicas. O modelo proposto foi derivado matematicamente de uma amostragem Poisson de uma distribuição do tipo gamma e, teoricamente, a proposição isola o efeito do tamanho amostral em um parâmetro da distribuição (Pielou 1977). Preston (1948) sugeriu uma distribuição lognormal como modelo mais adequado para descrever as SAD das comunidades naturais, que sob uma amostragem Poisson, manteria a forma de uma distribuição aproximadamente lognormal na SAD da amostra (Green & Plotkin 2007). Preston defendeu que o modelo da série logarítmica sugerido por Fisher et al. (1943) representaria a SAD para amostras insuficientes da comunidade (i.e. seria o equivalente à distribuição lognormal truncada), e que à medida que a amostra se torna suficiente, a distribuição lognormal seria revelada. Esta proposta originou a ideia de que o aumento da amostra revela a distribuição subjacente da comunidade (McGill et al. 2007).

A partir desses estudos pioneiros, vários modelos foram propostos para descrever a forma das curvas côncavas das SAD e interpretações alternativas proliferaram (veja revisão de McGill *et al.* 2007). Os estudos das SAD têm demonstrado relevância teórica e aplicada ao investigar a influência de importantes

20

processos, como diferenciação de nicho, dispersão, dependência de densidade, alterações ambientais, especiação e extinção, na estrutura e dinâmica de comunidades ecológicas (Tokeshi 1993; Hubbell 2001; McGill *et al.* 2007; Magurran & McGill 2011).

Os modelos de SAD existentes podem ser classificados em dois tipos principais (Magurran 2004): os modelos estatísticos, que se apresentam como descritores estatísticos das SADs, sem significado biológico proposto, e os modelos biológicos, que deduzem distribuições teóricas a partir de cenários ecológicos ou evolutivos. Os modelos do primeiro tipo, os estatísticos, invocam uma distribuição estatística teórica *a priori* para descrever as SADs. No entanto, seus parâmetros são muitas vezes difíceis de interpretar biologicamente (Tokeshi 1999). Esses modelos têm sido muito utilizados como modelos nulos para investigar como alguns fatores influenciam a SAD encontrada, assumindo que qualquer desvio nas distribuições pode ser causado por diferenças entre as espécies e os fatores ambientais avaliados (McGill *et al.* 2007). Os mais proeminentes dentre estes modelos são a distribuição de série logarítmica de Fisher *et al.* (1943) e a distribuição lognormal de Preston (1948).

O segundo tipo de modelo é representado principalmente pelos modelos orientados pelo conceito de espaço comum de nicho limitado. Os mais proeminentes entre esses modelos são: o modelo 'vara-quebrada' (*broken-stick*) de MacArthur (1957), o modelo de repartição sequencial (*sequential breakage*) de Sugihara (1980), e o modelo de repartição de nicho (*niche apportionment*) de Tokeshi (1990; 1996). De acordo com os modelos de Sugihara e de Tokeshi, um recurso limitado comum (ou espaço de nicho total) é dividido entre as espécies por dois processos básicos (Tokeshi 1990; 1996): primeiro o espaço de nicho é dividido em duas partes e o segundo passo aplica uma regra de escolha de uma dessas partes para futura subdivisão. Diferentes

distribuições de probabilidade para ambos os passos geraram uma série de modelos, mas a divisão de nicho é assumida como processo dominante influenciando na abundância das espécies.

A proliferação de tantos modelos teóricos para explicar um mesmo padrão empírico é um reconhecido embaraço para os ecólogos (Magurran 2004; McGill *et al.* 2007), já que a maioria dos modelos propostos deram bons ajustes apenas para alguns conjuntos de dados, mas nenhum foi geral o suficiente para abarcar todos estes conjuntos (McGill *et al.* 2007). Em princípio a ideia é que, encontrado o modelo que melhor se ajusta aos dados por meio de testes estatísticos de aderência, este poderia fornecer indícios dos processos subjacentes e permitir compreender mudanças no padrão das SADs (Magurran 2004).

Porém, o progresso tem sido limitado tanto devido aos problemas de interpretação quanto às dificuldades metodológicas em testar qual modelo melhor se ajusta aos dados (McGill 2003). Os tradicionais testes de aderência, como o teste de  $\chi^2$ e Kolmogorov-Smirnov, colocam o modelo de distribuição testado como hipótese nula ao comparar os dados observados e os esperados para diferentes modelos. Assim, infere-se que as abundâncias observadas ajustam-se aos valores previstos pelos modelos quando os testes não apresentam desvios significativos (Magurran 2004). No entanto, esta abordagem é um uso equivocado da lógica de teste de significância, pois faz com que sua aceitação dependa mais da força do teste que da qualidade do ajuste (Routledge 1980). Além disso, a impossibilidade de comparações simultâneas de modelos torna a aplicação desses testes de aderência inconclusivos, pois o ajuste pode acabar sendo aparentemente igual para mais de um tipo de distribuição teórica, já que estes são testados individualmente (Manly 1992). Uma solução promissora para este problema metodológico é o uso do princípio da verossimilhança estatística (Edwards 1992; Royall 2000) para criar protocolos de comparação simultânea entre hipóteses estatísticas concorrentes.

Os dados necessários para obter as SAD são amplamente coletados em ecologia e relativamente fáceis de obter em diversas comunidades biológicas, além de agregar informações fundamentais como riqueza, dominância e proporção de espécies raras (Magurran & McGill 2011). A grande quantidade de trabalhos que geraram esse tipo de dados aliado ao seu potencial de comparação entre diferentes comunidades representa um campo de pesquisa promissor dentro da ecologia e demanda estudos que fortaleçam o conhecimento empírico e teórico da área (McGill *et al.* 2007).

#### 1.3 Abordagem da verossimilhança para o ajuste e seleção de modelos

Dado um vetor de abundâncias de espécies em uma amostra, a verossimilhança é uma função das probabilidades previstas pelo modelo para cada abundância observada. O método da máxima verossimilhança para o ajuste de modelos se baseia na simples ideia de que, entre todos os valores de parâmetros possíveis de um modelo, há um que é o de maior plausibilidade, que é chamado de estimativa de máxima verossimilhança (Prado 2010). A máxima verossimilhança expressa, desta forma, a plausibilidade de cada modelo teórico, considerando os dados disponíveis, o que fornece um critério direto e objetivo para comparação (Royall 2000).

A estimativa de máxima verossimilhança (MLE – maximum likelihood estimate) é obtida a partir da função de verossimilhança ( $LL(\theta|y = yi)$ ) que representa a função das probabilidades geradas por um certo modelo teórico, dado um conjunto de observações (Edwards 1992). Para observações independentes, como assumido para as abundâncias que compõem as SADs, basta multiplicar as probabilidades atribuídas a cada valor de abundância pelo modelo. No entanto, em geral a função é transformada para obter a função de log-verossimilhança negativa, devido às vantagens de manipulação matemática e computacional (Prado 2010). Assim, a função de log-verossimilhança negativa de um conjunto de dados é definida como a soma do logaritmo das probabilidades atribuídas pela função de densidade probabilística de um modelo para cada valor da amostra (Edwards 1992):

$$LL(\theta|y = y_i) = -\sum_{i=1}^n \log f(\theta|y = y_i)$$

(1)

, em que  $y_i$  é a observação, y a variável aleatória que descreve os valores possíveis de  $y_i$  e que tem função de densidade probabilística  $f(y|\theta)$ , com um conjunto de parâmetros.

Encontrando então o mínimo da função, por dedução matemática em casos simples ou com aproximações numéricas, obtemos os valores de máxima verossimilhança do modelo e a estimativa de máxima verossimilhança dos parâmetros (Prado 2010; Bolker 2012).

Para selecionar entre os modelos concorrentes, é possível levar em consideração a complexidade de cada modelo, com medidas de verossimilhança que penalizem aqueles com maior número de parâmetros. Um dos critérios mais utilizados é o Critério de Informação de Akaike (AIC; Akaike 1973), que é uma medida de estimativa da distância relativa esperada de cada modelo ao modelo verdadeiro, baseada na distância de Kullback-Leibler e no máximo da função de log-verossimilhança (Burnham & Anderson 2002). Assim, dada a log-verossimilhança máxima de um modelo (L) e seu número de parâmetros (k), o AIC é dado por:

$$AIC = 2k - 2L$$

O modelo selecionado é aquele com menor valor de AIC, entretanto, quando a diferença entre o AIC dos modelos concorrentes for muito pequena, convencionalmente menor que dois, os modelos são considerados igualmente plausíveis (Prado 2010). Desta forma, o AIC ordena um conjunto de modelos de acordo com sua plausibilidade frente aos dados, levando em conta sua parcimônia.

Através desses passos, é possível encontrar os valores mais plausíveis para os parâmetros de cada modelo através do método de máxima verossimilhança, e, em seguida, comparar diferentes modelos pelo AIC. O uso crescente destes conceitos é considerada uma mudança no paradigma de análise de dados, em resposta às limitações que as abordagens tradicionais impõem a muitas áreas da biologia (Johnson & Omland 2004). Assim, a comparação simultânea de vários modelos - que podem representar hipóteses com graus crescentes de complexidade - é mais adequada para estudos observacionais, comuns em ecologia, já que os testes de significância tradicionais contemplam apenas um modelo por vez (a hipótese alternativa), e têm a premissa de que os dados provêm de experimentos controlados que podem ser repetidos (Royall 2000; Johnson & Omland 2004).

#### 1.4 A utilização da biomassa como medida de abundância nos modelos de SAD

A despeito da grande utilidade dos modelos matemáticos em descrever as variações nas SAD de maneira sintética, os modelos clássicos de SAD foram todos concebidos para tratar contagens de indivíduos como medida de abundância das espécies, sem contemplar a utilização de biomassa. Assim, um aspecto que ganha atenção recente nos estudos das SAD é a relação entre a forma da distribuição e a medida utilizada para medir a importância relativa das espécies de uma comunidade (Morlon *et al.* 2009). Alguns trabalhos têm apontado a importância de compreender as diferentes relações entre medidas alternativas de abundância para obter um melhor entendimento dos processos subjacentes estruturando as comunidades e as implicações nas interpretações dos modelos (e.g. Thibault *et al.* 2004; Connolly *et al.* 2005; Saint-Germain *et al.* 2007; Morlon *et al.* 2009; Henderson & Magurran 2010; Williamson 2010).

Um debate por trás da medida utilizada para representar a importância das espécies é sobre quando a utilização de diferentes medidas de abundância, como contagem numérica e biomassa, representam medidas equivalentes na forma como as espécies subdividem os recursos (Sugihara 1989; Pagel *et al.* 1991; Taper & Marquet 1996). Apesar da abundância das espécies ser geralmente expressa como número de indivíduos nos estudos ecológicos, que representam a forma como esses dados são tipicamente obtidos, o uso da biomassa é uma medida preferida por alguns ecólogos, que consideram a biomassa como uma medida mais direta da partição de recursos entre as espécies (Peters 1983; Magurran 2004; Saint-Germain *et al.* 2007).

A utilização do número de indivíduos como variável resposta nos modelos orientados pela repartição de nicho, por exemplo, sugere que a abundância numérica das espécies reflete os processos pelos quais as espécies compartilham recursos dentro da comunidade (Saint-Germain *et al.* 2007). Desta forma, a abundância numérica é utilizada como uma medida de aproximação da utilização de recursos. Para esta caracterização ser válida, como ressalta MacArthur (1960), é necessário assumir que os indivíduos entre as espécies estudadas são comparáveis em seus tamanhos e demandas energéticas (i.e. espécies ecologicamente equivalentes). No entanto, estas

medidas não são necessariamente equivalentes (Taper & Marquet 1996; Saint-Germain *et al.* 2007) e Morlon *et al.* (2009) demonstraram que a forma da distribuição da SAD geralmente não é conservada na conversão estatística entre as diferentes medidas de abundância das espécies.

A medida de abundância utilizada exerce grande influência sobre as formas das SAD e interpretações dos modelos (Thibault *et al.* 2004; Connolly *et al.* 2005) mas estudos de casos que permitam construir descrições das mudanças da diversidade biológica em resposta aos fatores investigados utilizando simultaneamente dados de contagem numérica e biomassa para as SAD são extremamente raros na literatura (McGill *et al.* 2007).

A proposição do modelo lognormal associado ao efeito do tamanho amostral feita por Preston (1948) só contempla número de indivíduos como medida de abundância nas SADs, e foi apenas com os trabalhos recentes de Connolly *et al.* (2005) e Morlon *et al.* (2009) que o modelo sugerido por Preston foi relacionado com a utilização da biomassa como medida de abundância. Houve um longo período entre essas duas proposições diferentes: (1) A SAD com número de indivíduos segue uma distribuição lognormal, que seria revelada em amostras maiores (Preston 1948); (2) A SAD de biomassa e número de indivíduos seguem uma distribuição lognormal, no entanto, para escalas menores, as SAD de biomassa podem reter mais ou menos a distribuição lognormal da comunidade, enquanto as SAD de número de indivíduos das mesmas comunidades permanecem truncadas.

Considerando a impossibilidade de amostrar todos os indivíduos em comunidades naturais e o aspecto central da influência do processo amostral nos estudos das SAD desde a proposição dos primeiros modelos (Fisher *et al.* 1943; Preston 1948), os resultados encontrados pelos trabalhos de Connolly *et al.* (2005) e Morlon *et al.* (2009) suscitam investigações se o uso da biomassa seria mais eficiente em revelar a SAD subjacente da comunidade em amostras menores e outras implicações do uso da biomassa nas teorias e modelos de SAD existentes.

#### 1.5 Relevância dos besouros escarabeíneos para estudos em ecologia

Os besouros da subfamília Scarabaeinae (Coleoptera: Scarabaeidae) possuem uma série de vantagens para estudos ecológicos: é um grupo que possui taxonomia e biologia relativamente bem conhecida; são globalmente distribuídos e facilmente coletados; possuem importância ecológica em todos os níveis tróficos; apresentam sensibilidade à alterações no ambiente e alta diversidade e associação com florestas tropicais; e permitem, desta forma, o estudo de padrões da comunidade associados a diferentes fatores e condições ecológicas (Halffter & Matthews 1966; Hanski & Camberfort 1991; Gardner et al. 2008a). Além disso, em decorrência do hábito copro-necrófago e consequências do comportamento e ciclo de vida, esses besouros participam de importantes serviços ambientais (Nichols *et al.* 2008), que incluem: o aumento da taxa de ciclagem de nutrientes no ambiente (Nealis 1977); fertilização e aeração do solo (Mittal 1993); redução da transmissão de alguns parasitas de vertebrados por matar larvas e ovos (Bergstrom *et al.* 1976); dispersão secundária de sementes (Andresen 2002; Vullinec 2002) e contribuição de biomassa para níveis tróficos superiores (Halffter & Mathews 1966; Halffter & Edmonds 1982).

A estreita relação e dependência do meio em que vivem e da oferta de alimentos fazem com que os besouros Scarabaeinae apresentem alta sensibilidade às alterações no ambiente (Klein 1989; Halffter *et al.* 1992; Nichols *et al.* 2007). Diversos estudos demonstram que a composição e estrutura das comunidades de escarabeíneos são afetadas negativamente em áreas modificadas e deterioradas pela ação humana (Klein 1989; Halffter *et al.* 1992; Halffter & Favila 1993; Davis *et al.* 2001; Didham *et al.* 1998; Estrada & Estrada 2002; Andresen 2003; Scheffler 2005; Nichols *et al.* 2007; Gardner *et al.* 2008a). Por esses motivos, além de serem facilmente coletados e frequentemente afetados de forma mais rápida e forte que outros taxa, esses besouros vêm sendo reconhecidos como um importante grupo indicador de mudanças ambientais por diversos autores (Klein 1989; Halffter & Favila 1993; Didham *et al.* 1998; Davis *et al.* 2001; Gardner *et al.* 2008b). No grupo, a abundância numérica e biomassa das espécies são características facilmente mensuráveis e comumente utilizadas para acessar a resposta e efeito da estrutura das comunidades associadas a diferentes condições (Gardner *et al.* 2008a).

#### 1.6 Considerações finais

Considerando então, que os padrões de estrutura da comunidade podem diferir dependendo da medida de abundância utilizada (Connolly *et al.* 2005), a utilidade dos besouros Scarabaeinae como grupo focal indicador, e o uso das SAD como descritores das comunidades, no próximo capítulo investigamos como o tamanho amostral afeta as SADs construídas a partir de medidas de biomassa e número de indivíduos para descrever a abundância relativa das espécies na comunidade. Essa investigação combina relevância teórica e aplicada. Vimos que é possível escolher o modelo mais adequado ao conjunto de dados com cada medida de abundância através de critérios práticos e objetivos. Podemos avaliar a consistência do modelo e as medidas de acurácia e precisão das estimativas dos parâmetros para diferentes tamanhos de amostras, e, assim, obter uma alternativa para distinguir qual medida de abundância é mais eficiente em revelar a SAD da comunidade em amostras menores. Com isso discutiremos como as mudanças entre essas distribuições associadas ao processo amostral representam contribuições teóricas e aplicadas relevantes no uso da SAD como ferramenta para representar e avaliar a diversidade biológica.

#### Referências

- Andresen, E. 2002. Dung beetles in a Central Amazonian rainforest and their ecological role as secondary seed dispersers. Ecological Entomology 27: 257-270.
- Andresen, E. 2003. Effect of forest fragmentation on dung beetle communities and functional consequences for plant regeneration. Ecography 26: 87-97.
- Akaike, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. *In:* Second International Symposium on Information Theory (Eds. B. N. Petrov & F. Csaki). Akademiai Kiado.
- Bergstrom, R. C.; Maki, L. R. & Werner, B. A. 1976. Small dung beetles as biological control agents: laboratory studies of beetles action on Trichostrongylid eggs in sheep and cattle feces. Proceedings of the Helminthological Society of Washington 43(2): 171-175.
- Bolker, B. 2012. Package bbmle: Tools for general maximum likelihood estimation. (R package version 1.0.4.1, based on stats4 by the R Development Core Team). URL http://cran.r-project.org/.
- Burnham, K. P. & Anderson, D. R. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer.
- Connolly, S. R.; Hughes, T. P.; Bellwood, D. R. & Karlson, R. H. 2005. Community structure of corals and reef fishes at multiple scales. Science 309: 1363-1365.
- Darwin, C. 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray.
- Davis, A. J.; Holloway, J. D.; Huijbregts, H.; Krikken, J.; Kirk-Spriggs, A. & Sutton,S. L. 2001. Dung beetles as indicators of change in the forests of northernBorneo. Journal of Applied Ecology 38: 593-616.
- Dewdney, A. K. 2000. A dynamical model of communities and a new speciesabundance distribution. Biological Bulletin 54: 294-302.

- Didham, R. K.; Hammond, P. M., Lawton, J. H., Eggleton, P. & Stork, N. E. 1998. Beetles species responses to Tropical Forest fragmentation. Ecological Monographs 68(3): 295-323.
- Dornelas, M.; Moonen, A. C.; Magurran, A. E. & Bàrberi, P. 2009. Species abundance distributions reveal environmental heterogeneity in modified landscapes. Journal of Applied Ecology 46: 666-672.
- Edwards, A. W. F. 1992. Likelihood. The Johns Hopkins University Press.
- Estrada, A. & Estrada, R. C. 2002. Dung Beetles in continuous Forest, Forest fragments and in an agricultural mosaic habitat island at Los Tuxtlas, Mexico. Biodiversity and Conservation 11: 1903-1918.
- Fisher, R. A.; Corbet, A. S. & Williams, C. B. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. Journal of Animal Ecology 12: 42-58.
- Gardner, T. A.; Hernández, M. I. M.; Barlow, J. & Peres, C. A. 2008a. Understanding the biodiversity consequences of habitat change: the value of secondary and plantation forests for neotropical dung beetles. Journal of Applied Ecology 45:883-893.
- Gardner, T. A.; Barlow, J.; Araujo, I. S.; Pires, T. C. A.; Bonaldo, A. B.; Costa, J. E.;
  Eposito, M. C.; Ferreira, L. V.; Hawes, J.; Hernández, M. I. M.; Hoogmoed,
  M. S.; Leite, R. N.; Lo-Man-Hung, N. F.; Malcolm, J. R.; Martins, M. B.;
  Mestre, L. A. M.; Santos, R. M.; Overal, W. L.; Parry, L.; Peters, S. L.;
  Ribeiro-Junior, M. A.; Silva, M. N. F.; Motta, C. S. & Peres, C. A. 2008b.
  The cost-effectiveness of biodiversity surveys in tropical forests. Ecology
  Letters, 11: 139-150.
- Green, J. & Plotkin, J. 2007. A statistical theory for sampling species abundance. Ecology Letters 10: 1037-1045.
- Halffter, G. & Edmonds, W. D. 1982. The nesting behavior of dung beetles (Scarabaeinae): An ecological and evolutive approach. Man and the Biosphere Program - UNESCO.
- Halffter, G. & Favila, M. E. 1993. The Scarabaeinae (Insecta: Coleoptera): an animal group for analysing, inventorying and monitoring biodiversity in tropical rain forest and modified landscapes. Biology International 27: 15-21.

- Halffter, G. & Matthews, E. G. 1966. The natural history of dung beetles of the subfamily Scarabaeinae (Coleoptera, Scarabaeidae). Folia Entomologica Mexicana 12/14: 1-312.
- Halffter, G.; Favila, M. E. & Halffter, V. 1992. A comparative study of the structure of the Scarab guild in Mexican Tropical Rain Forests and derived ecosystems. Folia Entomologica Mexicana 84: 132-157.
- Hanski, I. & Cambefort, Y. 1991. Dung Beetle Ecology. Princeton University Press.
- Henderson, P. A. & Magurran, A. E. 2010. Linking species abundance distributions in numerical abundance and biomass through simple assumptions about community structure. Proceedings of The Royal Society 277: 1561-1570.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press.
- Johnson, J. B. & Omland, K. B. 2004. Model selection in ecology and evolution. Trends in Ecology and Evolution 19: 101-108.
- Kempton, R. A. & Taylor, L. R. 1978. The Q-statistic and the diversity of floras. Nature 275: 252-253.
- Klein, B. C. 1989. Effects of forest fragmentation on dung and carrion beetle communities in Central Amazonia. Ecology 70(6): 1715-1725.
- Lambshead, P. J. D.; Platt, H. M. & Shaw, K. M. 1983. The detection of differences among assemblages of marine benthic species based on an assessment of dominance and diversity. Journal of Natural History 17: 859-874.
- MacArthur, R. H. 1957. On the relative abundance of bird species. Proceedings of National Academy of Science 43: 293-295.
- MacArthur, R. 1960. On the relative abundance of species. American Naturalist 94: 25-36.
- Magurran, A. E. 2004. Measuring Biological Diversity. Blackwell Publishing.
- Magurran, A. E. & McGill, B. J. 2011. Biological Diversity. Oxford University Press.
- Manly, B. F. J. 1992. The design and analysis of research studies. Cambridge University Press.
- McGill, B. J. 2003. Strong and weak tests of macroecological theory. Oikos 102: 679-685.
- McGill, B. J. 2011. Species abundance distribution. *In:* A. E. Magurran & B. J. McGill (ed.) Biological Diversity. Oxford University Press.

- McGill, B. J.; Etienne, R. S.; Gray, J. S.; Alonso, D.; Anderson, M. J.; Benecha, H. K.; Dornelas, M.; Enquist, B. J.; Green, J. L.; He, F.; Hurlbert, A. H.; Magurran, A. E.; Marquet, P. A.; Maurer, B. A.; Ostling, A.; Soykan, C. U.; Ugland, K. I. & White, E. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecology Letters 10: 995-1015.
- Mittal, I. C. 1993. Natural manuring and soil conditioning by dung beetles. Tropical Ecology 34: 150-159.
- Morlon, H.; White, E. P.; Etienne, R. S.; Green, J. L.; Ostling, A.; Alonso, D.;
  Enquist, B. J.; He, F. L.; Hurlbert, A.; Magurran, A. E.; Maurer, B. A.;
  McGill, B. J.; Olff, H.; Storch, D. & Zillio, T. 2009. Taking species abundance distributions beyond individuals. Ecology Letters 12: 488-501.
- Motomura, I. 1932. A statistical treatment of associations. Japanese Journal of Zoology 44: 379-383.
- Nealis, V. G. 1977. Habitat association and community analysis of south Texas dung beetles (Coleoptera: Scarabaeidae). Canadian Journal of Zoology 55: 138-147.
- Nichols, E.; Larsen, T.; Spector, S.; Davis, A. L.; Escobar, F.; Favila, M. & Vulinec,
  K. 2007. Global dung beetle response to tropical forest modification and fragmentation: A quantitative literature review and meta-analysis. Biological Conservation 137: 1-19.
- Nichols, E.; Spector, S.; Louzada, J.; Larsen, T.; Amezquita, S. & Favila, M. E. 2008. Ecological functions and ecosystem services provided by Scarabaeinae dung beetles. Biological Conservation 141: 1461-1474.
- Pagel, M. D.; Harvey, P. H. & Godfray, H. C. J. 1991. Species-abundance, biomass and resource use distribution. American Naturalist 138: 836-850.
- Peters, R. H. 1983. The ecological implications of body size. Cambridge University Press.
- Pielou, E. C. 1977. Mathematical Ecology. John Wiley and Sons.
- Prado, P. I. 2010. Distribuições de abundâncias de espécies: avanços analíticos para entender um padrão básico em ecologia. Ciência e Ambiente 39: 121-136.
- Preston, F. W. 1948. The commonness, and the rarity, of species. Ecology 29: 254-283.

- Routledge, R. D. 1980. The form of species-abundance distributions. Journal of Theoretical Biology 82: 547-558.
- Royall, R. M. 2000. Statistical Evidence: A likelihood paradigm. Chapman & Hall/CRC.
- Saint-Germain, M.; Buddle, C. M.; Larrivée, M.; Mercado, A.; Motchula, T.; Reichert, E.; Sackett, T. E.; Sylvain, Z. & Webb, A. 2007. Should biomass be considered more frequently as a currency in terrestrial arthropod community analysis? Journal of Applied Ecology 44: 330-339.
- Scheffler, P. Y. 2005. Dung beetle (Coleoptera: Scarabaeinae) diversity and community structure across three disturbance regimes in eastern Amazonia. Journal of Tropical Ecology 21: 9-29.
- Sugihara, G. 1980. Minimal community structure: an explanation of speciesabundance patterns. American Naturalist 116: 770-787.
- Sugihara, G. 1989. How do species divide resources? American Naturalist 133: 770-787.
- Taper, M. L. & Marquet, P. A. 1996. How do species really divide resources? American Naturalist 147: 1072-1086.
- Thibault, K.; White, E. & Ernest, S. K. M. 2004. Temporal dynamics in the structure and composition of a desert rodent community. Ecology 85: 2649-2655.
- Tokeshi, M. 1990. Niche apportionment or random assortment: species abundance patterns revisited. Journal of Animal Ecology 59: 1129-1146.
- Tokeshi, M. 1993. Species abundance patterns and community structure. Advances in Ecological Research 24: 111-186.
- Tokeshi, M. 1996. Power fraction: a new explanation of relative abundance patterns in species-rich assemblages. Oikos 75: 543-550.
- Tokeshi, M. 1999. Species coexistence. Blackwell.
- Vullinec, K. 2002. Dung beetle communities and seed dispersal in primary forest and disturbed land in Amazonia. Biotropica 34(2): 297-309.
- Whittaker, R. H. 1960. Vegetation of the Siskiyou mountains, Oregon and California. Ecological Monographs 30: 279-338.
- Williamson, M. 2010. Why do species abundance distributions of individuals and biomass behave differently under sampling? Oikos 119: 1697-1699.

## 2 Avaliação da eficiência da medida de abundância nos modelos de distribuição de abundância das espécies: Um estudo empírico em comunidades de besouros Scarabaeinae

#### 2.1 Introdução

A mensuração da biodiversidade representa um grande desafio para os ecólogos e envolve a questão central do quão bem os dados amostrados representam a comunidade encontrada no ambiente natural. Agravado pelo cenário atual de grande perda de espécies e alterações no ambiente natural, a necessidade de produzir boas estimativas de parâmetros da biodiversidade com a menor amostra possível é cada vez mais urgente (Magurran & McGill 2011).

Alguns avanços tecnológicos representaram um salto nas possibilidades do uso de ferramentas estatísticas e de modelagem aplicados à ecologia. Neste sentido, a utilização de modelos teóricos de distribuição de abundância de espécies para caracterização da diversidade biológica se destaca como uma ferramenta de grande potencial teórico e aplicado para a ecologia e biologia da conservação, devido principalmente à sua grande utilidade em descrever e comparar diferentes comunidades naturais (McGill *et al.* 2007).

A distribuição de abundância de espécies (SAD) revelou o padrão mais bem estabelecido da ecologia de comunidades, que é a alta dominância de algumas espécies e a presença de muitas espécies raras (Magurran 2004). A generalidade deste padrão assumiu um importante papel na biologia como princípio empírico que sustenta várias teorias ecológicas (Hubbell 2001; McGill 2003) e levou à proposição de dezenas de modelos teóricos de distribuição de probabilidade para descrever e explicar a forma das curvas côncavas das SAD (McGill *et al.* 2007). Apesar desses modelos terem sofrido muitas críticas, principalmente por não explicitarem mecanismos que expliquem o surgimento destes padrões (Hubbell 2001), a grande disponibilidade de dados necessários para obter as SADs aliado ao seu potencial de comparações e sua natureza visual (Figura 1, Capítulo 1) são importantes vantagens metodológicas do uso dos modelos de SAD nos estudos em ecologia (McGill *et al.* 2007).

Muitos autores defendem que as SADs em amostras de comunidades naturais bem representadas (ou extensivamente amostradas) apresentam uma distribuição da abundância de indivíduos aproximadamente lognormal (e.g. Preston 1948; Connolly et al. 2005; Williamson & Gaston 2005; Ulrich & Zalewski 2006; Sizling et al. 2009). Preston (1948), em seu estudo pioneiro, sugeriu que a ausência das espécies raras decorrentes da insuficiência amostral torna a distribuição lognormal truncada e, à medida que a amostra se torna suficiente, a distribuição lognormal se revela. Os estudos de Connolly et al. (2005) e Morlon et al. (2009) demonstraram que muitas SADs de biomossa se aproximam de uma distribuição lognormal mesmo em amostras em escalas menores, enquanto as SADs de indivíduos das mesmas comunidades permanecem truncadas, e só se revelam agregando-se amostras de vários sítios ou épocas, ou seja, em escalas maiores. Esses resultados sugerem que a biomassa seria uma medida mais efetiva para revelar a SAD subjacente da comunidade, em amostras pequenas ou de pequena abrangência. No entanto, a medida de abundância tradicionalmente utilizada para medir a importância relativa das espécies é o número de indivíduos, e o uso da biomassa e suas implicações é muito menos estudada nas SADs (Williamson 2010). Assim, a notável escassez desses estudos indica que há ainda a necessidade de mais investigações sobre o uso da biomassa como medida de abundância utilizada nas SADs e o efeito do tamanho amostral (McGill et al. 2007).

36
A lei da verossimilhança, amplamente aceita pelos estatísticos e que baseia a abordagem utilizada neste estudo para o ajuste e seleção dos modelos, postula que: se uma hipótese A implica que a probabilidade de uma variável aleatória X assumir o valor x é  $p_A(x)$ , enquanto a hipótese B implica que a probabilidade é  $p_B(x)$ , a observação X = x é uma evidência que suporta A sobre B apenas se  $p_A(x) > p_B(x)$ , e a razão de verossimilhança,  $p_A(x)/p_B(x)$ , mede a força dessa evidência. Edwards (1992) demonstrou que o uso da lei da verossimilhança satisfaz a propriedade de *consistência* requerida para uma boa medida de suporte de evidência. Isto quer dizer que se consideramos um modelo como verdadeiro, é esperado que o acumulo de evidência representado pelo aumento do tamanho amostral, aumente o suporte de evidência, favorecendo a aproximação ao modelo verdadeiro bem como as estimativas de seus parâmetros (Edwards 1992; Royall 2000). Se a hipótese de que a biomassa é mais eficiente em revelar a SAD é verdadeira, os modelos de SADs e seus parâmetros devem convergir mais rapidamente para o modelo real da comunidade com o aumento do tamanho da amostra. Assim, o aumento do tamanho amostral deve aproximar a distribuição amostral dos dados reais, e com isso, medidas de suporte de evidência dos modelos de SAD (e.g. Probability of Misleading Evidence - PME), e de acurácia e precisão das estimativas dos seus parâmetros tendem a zero com o aumento da amostra. Desta forma, um modelo selecionado com um dado conjunto amostral deve mostrar uma queda monotônica dessas medidas de análise com subamostras de tamanhos crescentes. Neste estudo, a expectativa é que as estimativas de PME dos modelos de SAD e as medidas de acurácia e precisão das estimativas de seus parâmetros apresentem uma queda mais rápida com o tamanho das subamostras com biomassa, em comparação com número de indivíduos.

O objetivo principal deste trabalho é investigar e discutir a vantagem de considerar a biomassa nos modelos teóricos de distribuição de abundância das espécies avaliando o efeito do tamanho da amostra sobre a qualidade dos ajustes dos modelos, usando as duas medidas de abundância: biomassa e número de indivíduos. Amostras de diferentes tamanhos foram simuladas por rarefação de uma base de dados empírica e diferentes modelos de SADs foram ajustados a estas amostras simuladas. Com isso, foi possível avaliar a eficiência das amostras com cada medida de abundância em revelar o modelo correto, bem como a acurácia e precisão das estimativas dos parâmetros desses modelos. Para o presente estudo, utilizamos um grupo com reconhecido potencial indicador e relevância para estudos ecológicos, os besouros Scarabaeinae, que também é um dos poucos grupos terrestres para o qual é usual medir as abundâncias tanto numérica como em biomassa.

## 2.2 Material e Métodos

#### 2.2.1. Descrição dos dados

Os dados das comunidades de besouros Scarabaeinae utilizados no estudo são provenientes do Projeto "Morfometria de besouros escarabeídeos (Coleoptera: Scarabaeidae) de uma comunidade de Floresta Atlântica", executado pela Dra. Malva Isabel Medina Hernández durante seu Projeto de Doutorado na Universidade Estadual Paulista (UNESP - Rio Claro).

As coletas para obtenção dos dados foram realizadas em seis diferentes áreas de Floresta Atlântica que compõem parte do mosaico florestal do Parque Estadual da Serra do Japi, São Paulo (localizado entre 23°12' e 23°22' S, e 46°53' e 47°03' W). Cada uma das áreas de coleta diferem conforme características descritas em Hernández & Vaz-de-Mello (2009): a área 1 é um vale (com 1.050 m de altitude) (Vale); a área 2 também é um vale (1.000 m), porém visivelmente mais degradado que o primeiro (Vale Alterado); a área 3 é uma encosta voltada ao sul (1.000 m) e influenciada pela umidade do mar (Encosta Sul); a área 4 é uma encosta voltada para o norte (1.025 m) e seca (Encosta Norte); a área 5 é um topo de colina (1.150 m) (Topo Morro); e a área 6 é uma floresta secundária que regenerou abaixo de uma plantação de *Eucaliptus* sp. com cerca de 30 anos (1.150 m) (Eucaliptus).

A coleta de dados foi realizada mensalmente durante o período de setembro de 1997 a agosto de 1998. As amostragens foram feitas com armadilhas de queda com iscas de atração, conforme descrito em Hernández *et al.* (2011). As armadilhas foram dispostas ao longo de uma linha de amostragem em intervalos de 50 metros. Para cada área foram utilizadas quatro armadilhas que permaneceram iscadas durante o período de 48 horas consecutivas em cada mês. Assim, o total de unidades amostrais obtidas ao longo do ano de estudo para representar a comunidade de cada área foi de 48 armadilhas. Daqui em diante denominaremos 'amostra' o conjunto de unidades amostrais de cada área (total de 48 unidades) e 'subamostra' o subconjunto tomado destas 48 unidades.

Os indivíduos coletados nas armadilhas foram secos em estufa a 45°C durante 48 horas e pesados em balança analítica com 0,001g de precisão para obtenção do peso seco, utilizado neste estudo como medida de biomassa. A partir destes dados, obtivemos um vetor de abundância numérica e de biomassa das espécies coletadas em cada unidade amostral de cada área do estudo para gerar as SADs das comunidades.

### 2.2.2. Análises estatísticas

Para avaliar o efeito do tamanho amostral sobre a qualidade dos ajustes, modelos teóricos de SAD foram ajustados a subamostras simuladas por rarefação do conjunto amostral de cada área de estudo, que foi o total de armadilhas em cada área (48 unidades amostrais). Estes modelos ajustados às subamostras foram então comparados com os modelos ajustados às amostras completas de 48 unidades de cada medida de abundância. Repetindo-se o procedimento muitas vezes, foi possível estimar a probabilidade do modelo correto não ser selecionado e também o viés e precisão das estimativas dos parâmetros de cada modelo. As seções seguintes descrevem os procedimentos de cada passo das análises, sendo: (1) ajuste e seleção dos modelos, (2) rarefação das amostras, (3) medida da probabilidade de evidência enganadora – PME, (4) estimativa do viés e precisão das estimativas dos parâmetros.

#### (1) Ajuste e seleção dos modelos

Em primeiro lugar, identificamos os modelos teóricos de SAD que melhor descreviam o conjunto de dados das duas medidas de abundância obtidos em cada área. Para isso, os modelos teóricos utilizados no estudo foram ajustados pelo método de máxima verossimilhança estatística e selecionados através do Critério de Informação de Akaike (AIC).

O ajuste dos modelos foi feito pelo método da máxima verossimilhança estatística. Para isso, obtivemos a estimativa de máxima verossimilhança dos parâmetros (MLE - *maximum likelihood estimate*) por meio da minimização da função de log-verossimilhança negativa (descrita no capítulo 1), utilizando otimizadores computacionais (Bolker 2012).

Os modelos teóricos (representados pelas respectivas distribuições probabilísticas) utilizados para representar as abundâncias das espécies de besouros Scarabaeinae foram: Lognormal, Poisson-Lognormal, Logserie, Gamma e Weibull. Essas distribuições foram escolhidas no presente estudo por apresentarem amplo uso nos estudos de representação de comunidades biológicas (Hubbell 2001; McGill 2011).

Para os dados de abundância numérica das espécies, as distribuições probabilísticas utilizadas foram a distribuição Lognormal (equação 3), distribuição Poisson-Lognormal (equação 4) e distribuição Série Logarítmica (ou Logserie) (equação 5), com as respectivas funções de densidade probabilística:

(3)

$$f(r|\mu,\sigma) = \frac{1}{r\sigma\sqrt{2\pi}} e^{\frac{(-\ln(r)-\mu)^2}{2\sigma^2}}$$

, em que  $f(r|\mu, \sigma)$  é a probabilidade de uma espécie ter abundância em um dado intervalo r na comunidade, dados os valores  $\mu e \sigma$ , que são os parâmetros de escala e forma, respectivamente.

(4)  
$$f(r|\mu,\sigma^2) = \frac{1}{\sigma\sqrt{(2\pi)x!}} \int_0^\infty (-\theta)\theta^{x-1} exp\left(-\frac{(\log\theta-\mu)^2}{2\sigma^2}\right) d\theta$$

, em que  $f(r|\mu, \sigma^2)$  corresponde à probabilidade de uma espécie ter abundância r na comunidade, dados os valores  $\mu \in \sigma^2$ , que também são parâmetros de escala e forma, respectivamente.

$$f(r|\alpha) = \frac{N^r}{r(N+\alpha)^r \log\left(\frac{N+\alpha}{N}\right)}$$

, em que  $f(r|\alpha)$  corresponde à probabilidade de uma espécie ter abundância r na comunidade, dado os valores do parâmetro  $\alpha$ , que é proporcional ao número de espécies representadas por um só indivíduo na amostra e N o total de indivíduos.

Para os modelos Poisson-Lognormal e Logserie, as funções de densidade probabilística foram truncadas na classe zero.

Já para os dados de biomassa das espécies, as distribuições probabilísticas utilizadas foram a distribuição Lognormal (equação 3), distribuição Gamma (equação 6) e distribuição Weibull (equação 7), com as respectivas funções de densidade probabilística (equação 3 já demonstrada acima):

(6)

$$f(r|s,a) = \frac{1}{a^s \prod(s)} r^{s-1} e^{\frac{-r}{a}}$$

, em que f(r|s, a) é a probabilidade de uma espécie ter abundância em um dado intervalo r na comunidade, dados os valores  $s \in a$ , que são os parâmetros de forma e escala, respectivamente.

(7)

$$f(r|a,s) = \frac{s}{a} \left(\frac{r}{a}\right)^{s-1} e^{(r/a)^s}$$

, em que f(r|a, s) corresponde à probabilidade de uma espécie ter abundância em um dado intervalo r na comunidade, dados os valores a e s, que são os parâmetros de escala e forma, respectivamente.

Após obtidos os valores de máxima verossimilhança dos modelos e a estimativa de máxima verossimilhança dos parâmetros, selecionamos entre os modelos concorrentes através do Critério de Informação de Akaike – AIC o modelo mais plausível (Akaike 1973). O AIC, conforme descrito no capítulo 1, ordena um conjunto de modelos de acordo com sua plausibilidade frente aos dados, levando em conta sua parsimônia. Considerando então a fórmula para calcular o AIC (equação 2), o modelo que melhor descreve cada conjunto de dados é aquele com menor valor de AIC. Entretanto, quando a diferença entre o menor AIC dos modelos concorrentes foi menor que dois, consideramos empate e ambos os modelos concorrentes são selecionados.

Desta forma, obtivemos os modelos que melhor descrevem cada conjunto amostral (48 unidades amostrais) de cada área (Vale, Vale Alterado, Encosta Sul, Encosta Norte, Topo Morro e Eucaliptus), e as estimativas de seus parâmetros. Para os propósitos deste trabalho, estes modelos foram tomados como as descrições "reais" de cada conjunto de dados. Assim, os modelos ajustados às subamostras de cada conjunto devem convergir para o modelo e estimativas obtidas com a amostra completa dos dados de cada área. As seções seguintes detalham como as subamostras foram simuladas, como foram ajustados os modelos a elas, e como esses modelos foram avaliados.

O ajuste e seleção de modelos pelo método da máxima verossimilhança e critério do AIC foram realizados a partir de funções do pacote *bbmle* (Bolker 2012) e

*sads* (Prado 2012) no ambiente estatístico R versão 2.14.2 (R Development Core Team 2012).

# (2) Rarefação das amostras

Uma vez identificados os modelos mais plausíveis de SADs para as amostras de cada área, geramos subamostras de menor tamanho e repetimos os ajustes dos modelos para cada tamanho das subamostras simuladas.

A unidade amostral tomada ao acaso para compor as amostras rarefeitas (subamostras) foram os dados de cada armadilha de queda do estudo em um dado mês. Assim, o sorteio para compor as subamostras foi feito selecionando ao acaso, entre o conjunto das 48 unidades amostrais de cada área de estudo, um número de unidades amostrais sem reposição. O número de unidades amostrais tomadas ao acaso sem reposição correspondem aos tamanhos das subamostras, e foram: 42, 36, 30, 24, 18, 12 e 6; representando a redução gradativa de 6 unidades amostrais a partir do conjunto de 48 unidades amostrais de cada área (rarefação da amostra).

Uma vez compostas as subamostras simuladas, as abundâncias numéricas ou de biomassa de cada espécie em cada armadilha foram somadas para obter a abundância total por espécie na subamostra. Assim, os modelos de SAD foram ajustados a cada subamostra simulada através do método de máxima verossimilhança descrito na seção anterior, obtendo-se os valores de máxima verossimilhança dos modelos e as MLEs dos parâmetros. O procedimento foi repetido 1000 vezes para cada tamanho de amostra, área e medida de abundância.

As simulações das reamostragens foram realizadas através de função específica criada no ambiente estatístico R (R Development Core Team 2012).

44

Obtivemos desta forma 1.000 vetores de abundância (número de indivíduos e biomassa) das espécies para cada área e tamanho de subamostra simulada.

#### (3) Medida da probabilidade de evidência enganadora - PME

Considerando que o AIC é uma medida de distância relativa esperada ao modelo teoricamente verdadeiro (Burnham & Anderson 2002), o AIC obtido de uma amostra é uma medida sujeito a erro. Assim, algumas amostras podem não revelar o melhor modelo que representa os dados originais. Royall (2000) chamou a isto de evidência enganadora (*misleading evidence*). A probabilidade de evidência enganadora (PME – *probability of misleading evidence*) pode ser estimada repetindo-se a amostragem e seleção de modelos de uma mesma população estatística cujo melhor modelo é conhecido. A proporção dessas repetições em que o modelo correto não é selecionado estima a PME.

Portanto, para estimar a PME em nosso estudo, utilizamos os valores de máxima verossimilhança obtidos pelo ajuste dos modelos a cada subamostra simulada para calcular os valores de AIC dos mesmos. Posteriormente, calculamos a PME contando o número de vezes que o modelo correto não foi selecionado pelo critério do AIC. Isso corresponde a quantas vezes a diferença entre os valores de AIC do modelo correto ao modelo de menor AIC ( $\Delta$ AIC) foi maior que dois. Após obtido o número de vezes que o modelo correto não foi selecionado, dividimos o valor pelo número de subamostras (1000) e multiplicamos por 100 para estimar o valor em percentual. Assim, obtivemos as estimativas de PME dos modelos para todas as subamostras simuladas de cada área, e para ambas as medidas de abundância. Considerando que um modelo selecionado com uma dada amostra deve mostrar uma queda monotônica de PME com subamostras de tamanhos crescentes, a expectativa é que a PME caia

mais rápido com o tamanho das subamostras com biomassa, em comparação com número de indivíduos.

#### (4) Estimativa do viés e precisão das estimativas dos parâmetros

As estimativas de máxima verossimilhança (MLEs) dos parâmetros obtidos pelo ajuste dos modelos de SAD a cada tamanho de subamostra foram utilizados para avaliar o viés e precisão (coeficiente de variação) destas estimativas. Os parâmetros usados para essas avaliações foram os parâmetros de seus respectivos modelos que possuem interpretação biológica comum em análises de diversidade, por serem associados à dominância das espécies nas SADs (McGill 2011). Esses parâmetros foram: os parâmetros de *forma* das distribuições Lognormal, Poisson-Lognormal, Gamma e Weibull, e o parâmetro único alfa da Logserie.

Para obter o viés das estimativas dos parâmetros com cada medida de abundância e em cada tamanho de amostra, calculamos a diferença entre a média das estimativas dos valores dos parâmetros obtidos nas subamostras de um dado tamanho simuladas e o valor do parâmetro ajustado aos dados da amostra completa (de 48 unidades amostrais) de cada área pela estimativa de máxima verossimilhança. Esta diferença foi dividida pelo valor real dos parâmetros e multiplicada por 100 para expressar o viés em termos percentuais e, com isso, facilitar a comparação entre os diferentes modelos e medidas de abundância.

O coeficiente de variação das estimativas dos parâmetros foi usado como uma medida de precisão dos MLEs obtidos de cada subamostra simulada. O coeficiente de variação das estimativas dos parâmetros para cada subamostra foi calculado dividindo o valor do desvio padrão dos parâmetros estimados nas 1.000 repetições de cada subamostra pelo valor da média desses parâmetros estimados. Considerando então que os parâmetros dos modelos ajustados às subamostras de tamanhos crescentes tendem a convergir para os parâmetros de cada conjunto amostral, esperamos que as MLEs dos parâmetros com as subamostras de biomassa convirjam mais rapidamente ao valor do parâmetro para o conjunto amostral de cada área que com número de indivíduos, e apresente dessa forma menor valor de viés e coeficiente de variação das estimativas dos parâmetros, com o aumento do tamanho amostral.

## 2.3 Resultados

## 2.3.1 SADs das amostras completas de cada área

As amostras coletadas mensalmente durante o período de 12 meses na Serra do Japi, São Paulo, registraram o total de 3.524 indivíduos pertencentes a 39 espécies de besouros da subfamília Scarabaeinae, cuja soma total de biomassa seca foi de 100,05g (Anexos 1 e 2). A espécie mais dominante numericamente foi *Canthidium trinodosum*, representada por 45,8% de todos os indivíduos do estudo. No entanto, para a abundância de biomassa, essa espécie representou apenas 11,34% da biomassa, sendo a terceira espécie com maior biomassa total.

As demais espécies de maior dominância numérica (>120 indivíduos) foram: *Eurysternus cyanescens*, *Uroxys kratochvili*, *Scybalocanthon nigriceps*, *Uroxys lata*, *Canthonella* sp. e *Dichotomius assifer* (Anexo 1). Já as espécies mais raras presentes no estudo, que apresentaram abundância numérica total menor que 10 indivíduos foram: *Canthidium dispar*, *Canthidium sulcatum*, *Trichillum* sp.2, *Dichotomius depressicollis*, *Dichotomius* sp.2, *Dichotomius carbonarius*, *Ateuchus histrio*, *Eurysternus parallelus*, *Canthon rutilans*, *Dichotomius bechynei*, *Phanaeus*  splendidulus, Dichotomius mormon, Coprophanaeus saphirinus, Canthidium sp.1, Uroxys sp.2, Ontherus azteca e Canthon sp. (Anexo 1).

No geral, a abundância de biomassa apresentou menor dominância relativa das espécies que a abundância numérica, e as espécies numericamente dominantes não foram necessariamente as espécies com maior abundância de biomassa. Assim, as espécies com menor abundância não foram necessariamente as mesmas com as duas medidas de abundância (Anexos 1 e 2). As espécies com maior abundância de biomassa no estudo (>10g) foram: *Dichotomius assifer*, com 22,74% da biomassa total, *Deltochilum furcatum* com 15,21%, *Canthidium trinodosum* com 11,34%, *Eurysternus cyanescens* com 11,21% e *Deltochilum brasiliense* com 10,52% (Anexo 2). Já as espécies com menor abundância de biomassa (<0,05g) foram: *Trichillum* sp.2, *Uroxys* sp.2, *Canthidium sulcatum*, *Uroxys* sp.1, *Ateuchus histrio, Canthidium* sp.1, *Canthidium dispar, Trichillum* sp.1 e *Paracanthon pereirai* (Anexo 2).

Houve uma grande variação de abundância das espécies entre as áreas, tanto para abundância numérica quanto pra biomassa. A área Vale Alterado teve a maior abundância numérica, com 1380 indivíduos, representados a maior parte pela espécie *Canthidium trinodosum*, que apresentou 1.031 indivíduos; e a área Topo Morro foi a de menor abundância numérica, com 285 indivíduos (Anexo 1). Já para a abundância de biomassa, a área Encosta Sul foi a que apresentou maior abundância, com o total de 25,72g, e a menor abundância de biomassa (assim como para a abundância numérica) ocorreu na área Topo Morro, com 4,39g (Anexo 2).

A riqueza de espécies não teve grande variação entre as áreas. A Encosta Sul apresentou a maior riqueza, com 30 espécies, e a área Eucaliptus foi a de menor riqueza, com 22 espécies (Anexo 1).

48

Em geral, as espécies mais dominantes em abundância de indivíduos e biomassa (e.g. >50 indivíduos; >5g de biomassa) foram distribuídas em todas as áreas, e as espécies com menor abundância relativa de ambas as medidas (número de indivíduos e biomassa) mantiveram o padrão de raridade em todas as áreas (Anexos 1 e 2).

Todos os modelos concorrentes para o ajuste das SADs com as diferentes medidas de abundância foram selecionados para o conjunto amostral de pelo menos duas áreas de estudo. Portanto, não houve um único modelo que se destacasse para descrever os dados de todas as amostras. A descrição das SADs de cada área com os modelos que melhor descrevem cada conjunto de dados é apresentado a seguir. O padrão mais claro que se repete para todas as áreas do estudo é que há mais dominância numérica do que por biomassa: a queda da linha nos gráficos de rankabundância é sempre mais acentuada para indivíduos, em geral por causa de 1 a 3 espécies numericamente dominantes (Figuras 2 a 7).

## Comunidade 1: Vale

Para a SAD de indivíduos na área Vale, todos os modelos concorrentes foram selecionados pelo AIC (Tabela 1), no entanto, o modelo da Logserie foi descartado devido ao grande viés apresentado na análise de PME (Figura 8; seção 2.3.2). Para a abundância de biomassa, todos os modelos concorrentes foram selecionados como bons descritores da SAD desta área de estudo (Tabela 2).

Os modelos Poisson-Lognormal e Logserie para a SAD de indivíduos diferiram principalmente na previsão das abundâncias das espécies raras, que pode ser observado no gráfico de rank-abundância através da maior divergência das caudas dos modelos de SADs do que o início da curva (Figura 2a). Já para a SAD de biomassa, a

Lognormal previu melhor a abundância das espécies raras (rank 15 em diante), enquanto as distribuições Gamma e Weibull previram melhor as espécies mais abundantes (rank até 15) e muito raras (rank 23 em diante) (Figura 2b). Assim, o empate entre os modelos concorrentes ocorrido para as SADs de biomassa e número de indivíduos provavelmente ocorreu devido ao fato de que cada modelo prevê melhor as abundâncias de um grupo de espécies (Figura 2).

No rank de abundância numérica, duas espécies se destacaram como mais dominantes, *Eurysternus cyanescens* com 65 indivíduos e *Uroxys lata* com 64. As três espécies seguintes apresentaram abundâncias de 43-25 indivíduos (*Canthidium trinodosum, Deltochilum furcatum* e *Canthon latipes*) e as demais espécies apresentaram abundâncias que diminuíram gradativamente de 16 indivíduos (rank 6) até as espécies mais raras, que foram: *Eurysternus parallelus, Phanaeus splendidulus, Trichillum* sp.1 e *Trichillum* sp.2 (Figura 2a). A espécie mais dominante numericamente, *Eurysternus cyanescens*, apresentou a terceira maior abundância de biomassa. Já a espécie *Uroxys lata*, que teve a segunda maior abundância, ficou na posição 13 do rank de biomassa. No rank de abundância de biomassa, a espécie de maior abundância foi *Deltochilum furcatum*, com 5,22g. Depois vê-se um grupo de três espécies com grande abundância de biomassa, entre 2,77g e 2,29g, e as demais espécies apresentaram biomassa menores que 1g até as espécies mais raras, que assim como para número de indivíduos, foram *Trichillum* sp.2 e *Trichillum* sp.1 (Figura 2b).

**Tabela 1.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância numérica** da comunidade de besouros Scarabaeinae coletadas na área **Vale**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes.

MODELO	AIC	ΔΑΙC
Poisson-Lognormal	189,48	0,00
Logserie	190,26	0,78
Lognormal	190,42	0,94

**Tabela 2.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância de biomassa** da comunidade de besouros Scarabaeinae coletadas na área **Vale**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes. Aqui utilizamos o **AIC** corrigido para pequenas amostras (**AICc**; Burnham & Anderson 2002).

MODELO	AICc	ΔΑΙΟ	
Weibull	16,50	0,00	
Gamma	17,10	0,60	
Lognormal	18,00	1,50	

-



**Figura 2.** Diagramas de rank-abundância para número de indivíduos (a) e biomassa (b) da comunidade de besouros Scarabaeinae coletada entre setembro/1997 a agosto/1998 na área Vale, Serra do Japi, São Paulo. Os pontos em preto são os dados empíricos e as linhas coloridas são os valores previstos pelos modelos teóricos, que são distribuições de densidade probabilística. O gráfico é apresentado em escala logarítmica.

# Comunidade 2: Vale Alterado

Na área Vale Alterado, os modelos selecionados pelo método do AIC que melhor descrevem a SAD de indivíduos foram Poisson-Lognormal e Logserie (Tabela 3), no entanto, o modelo da Logserie foi descartado devido ao grande viés apresentado na análise de PME (Figura 9; seção 2.3.2). Para a abundância de biomassa, a Lognormal foi o modelo que melhor descreveu a SAD nessa amostra (Tabela 4).

O modelo Poisson-Lognormal conseguiu aproximar a previsão de abundância da espécie mais dominante e das espécies mais raras (rank 18 em diante), no entanto, para as espécies de abundância intermediária o modelo apresentou um desvio grande em relação aos valores observados (Figura 3a). Já a SAD de biomassa foi mais equitativa, e o modelo Lognormal acertou relativamente bem a abundância de todas as espécies (Figura 3b).

A grande dominância numérica da espécie *Canthidium trinodosum* ocorrida nesta área de estudo (1031 indivíduos) ficou bastante evidente no gráfico de rank-abundância (Figura 3a). Vê-se também que as cinco espécies seguintes apresentaram abundâncias de apenas 80-40 indivíduos. Depois há um grupo de espécies de 10-5 indivíduos, e um grande número de espécies com abundância menor que cinco na amostra desta área (Figura 3a). As espécies mais raras numericamente foram: *Canthon latipes, Coprophanaeus saphirinus, Deltochilum rubripenne, Dichotomius mormon, Eurysternus hirtellus, Eurysternus sp., Paracanthon pereirai* e *Sylvicanthon foveiventre*. Para o rank de abundância de biomassa, a espécie mais dominante também foi *Canthidium trinodosum*, com 7,76g. As quatro espécies seguintes (*Dichotomius assifer, Eurysternus cyanescens, Deltochilum brasiliense* e *Deltochilum furcatum*) apresentaram biomassa entre 5-1,5g (Figura 3b). As demais

espécies apresentaram biomassa de 0,5g a 0,01g, e as duas espécies de menor biomassa, *Paracanthon pereirai* e *Uroxys* sp.1, com 0,004g e 0,002g, respectivamente (Figura 3b).

**Tabela 3.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância numérica** da comunidade de besouros Scarabaeinae coletadas na área **Vale Alterado**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes.

MODELO	AIC	ΔΑΙC	
Poisson-Lognorm	al 177,33	0,00	
Logserie	178,95	1,61	
Lognormal	185,43	8,10	

**Tabela 4.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância de biomassa** da comunidade de besouros Scarabaeinae coletadas na área **Vale Alterado**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes. Aqui utilizamos o **AIC** corrigido para pequenas amostras (**AICc**; Burnham & Anderson 2002).

MODELO	AICc	ΔΑΙC	
Lognormal	15,10	0,00	
Weibull	17,80	2,70	
Gamma	20,90	5,80	



**Figura 3.** Diagramas de rank-abundância para número de indivíduos (a) e biomassa (b) da comunidade de besouros Scarabaeinae coletada entre setembro/1997 a agosto/1998 na área Vale Alterado, Serra do Japi, São Paulo. Os pontos em preto são os dados empíricos e as linhas coloridas são os valores previstos pelos modelos teóricos de distribuição probabilística. O gráfico é apresentado em escala logarítmica.

# Comunidade 3: Encosta Sul

O melhor modelo para descrever a SAD de indivíduos com os dados da amostra completa da área Encosta Sul foi a Logserie (Tabela 5). Para a abundância de biomassa os modelos Lognormal e Weibull foram selecionados pelo AIC como bons descritores da SAD (Tabela 6), no entanto, o modelo Weibull foi descartado devido ao grande viés apresentado na análise de PME (Figura 10; seção 2.3.2).

O modelo Logserie para a SAD de indivíduo acertou melhor a previsão das espécies mais abundantes (rank até 5) e mais raras (rank 25 em diante), e nas abundâncias intermediárias houve algumas divergências nas previsões das abundâncias numéricas (Figura 4a). Já o modelo Lognormal apresentou uma previsão acurada da abundância de biomassa para praticamente todas as classes de abundância (Figura 4b). Também fica evidente nos ranks de abundância a maior equitabilidade da SAD de biomassa, bem como a maior quantidade de espécies muito raras (rank 22 em diante) (Figura 4).

No rank de abundância numérica, as duas espécies que se destacaram como mais dominantes foram *Canthidium trinodosum*, com 174 indivíduos e *Uroxys kratochvili*, com 158 indivíduos (Figura 4a). Vê-se em seguida um grupo de três espécies com abundâncias de 75-57 indivíduos (*Uroxys lata, Eurysternus cyanescens* e *Dichotomius assifer*) e depois um grupo grande de espécies com abundâncias menores que 18 indivíduos, sendo que nove destas espécies foram representadas por um único indivíduo na amostra desta área (Figura 4a), que foram: *Ateuchus histrio, Canthidium sulcatum, Coprophanaeus saphirinus, Deltochilum morbillosum, Dichotomius depressicollis, Ontherus azteca, Paracanthon pereirai, Sylvicanthon foveiventre* e *Uroxys aterrima*. No rank de abundância de biomassa, a espécie *Dichotomius assifer* se destacou por apresentar uma grande dominância, com 10,16g

(Figura 4b). A espécie seguinte apresentou biomassa de 3,3g (*Deltochilum furcatum*), com uma diminuição gradativa de biomassa para as demais espécies, até as espécies mais raras, que apresentaram biomassa de 0,035g (*Paracanthon pereirai*), 0,028g (*Uroxys* sp.1) e 0,017g (*Canthonella* sp.) (Figura 4b).

**Tabela 5.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância numérica** da comunidade de besouros Scarabaeinae coletadas na área **Encosta Sul**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes.

1			
MODELO	AIC	ΔΑΙC	
Logserie	215,39	0,00	
Poisson-Lognorm	al 218,45	3,06	
Lognormal	224,70	9,31	

**Tabela 6.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância de biomassa** da comunidade de besouros Scarabaeinae coletadas na área **Vale Alterado**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes. Aqui utilizamos o **AIC** corrigido para pequenas amostras (**AICc**; Burnham & Anderson 2002).

/			
MODELO	AICc	ΔΑΙC	
Lognormal	12,00	0,00	
Weibull	13,90	1,90	
Gamma	17,70	5,70	



**Figura 4.** Diagramas de rank-abundância para número de indivíduos (a) e biomassa (b) da comunidade de besouros Scarabaeinae coletada entre setembro/1997 a agosto/1998 na área Encosta Sul, Serra do Japi, São Paulo. Os pontos em preto são os dados empíricos e as linhas coloridas são os valores previstos pelos modelos teóricos de distribuição probabilística. O gráfico é apresentado em escala logarítmica.

## Comunidade 4: Encosta Norte

Na área Encosta Norte, os modelos selecionados que melhor descreveram a SAD de indivíduos foram a Poisson-Lognormal e Lognormal (Tabela 7). Para a abundância de biomassa, todos os modelos concorrentes foram selecionados como bons descritores das SAD (Tabela 8).

Os modelos Poisson-Lognormal e Lognormal se apresentaram praticamente como descritores equivalentes das abundâncias numéricas, diferindo um pouco apenas na previsão das espécies menos abundantes (rank 18 em diante) (Figura 5a). Já para a SAD de biomassa, os modelos Gamma e Weibull foram praticamente equivalentes e acertaram melhor a previsão das espécies mais dominantes (rank até 7), enquanto a Lognormal acertou melhor a previsão das espécies de menor abundância (rank 13 em diante) (Figura 5b). Assim, o empate entre a Lognormal e os modelos Gamma e Weibull para descrever a SAD de biomassa ocorreu provavelmente pelo fato de que cada modelo prevê melhor as abundâncias de um grupo de espécies (Figura 5b).

Na SAD de número de indivíduos fica evidente a presença de duas espécies de maior dominância, que são: *Canthidium trinodosum*, com 164 indivíduos e *Scybalocanthon nigriceps*, com 107 indivíduos (Figura 5a). Vê-se também que as espécies seguintes apresentaram abundâncias menores que 30 indivíduos, e três espécies que foram representadas por apenas um indivíduo na amostra dessa área (Figura 5a), que foram: *Canthidium dispar, Dichotomius bechynei* e *Dichotomius mormon*. Para o rank de abundância de biomassa, as duas espécies mais dominantes, diferente do rank de número de indivíduos, foram: *Deltochilum furcatum* com 3,48g e *Dichotomius assifer* com 3,15g (Figura 5b). As demais espécies apresentaram

biomassa (Figura 5b): *Uroxys* sp.1, que apresentou abundância numérica de seis indivíduos.

**Tabela 7.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância numérica** da comunidade de besouros Scarabaeinae coletadas na área **Encosta Norte**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes.

MODELO	AIC	ΔΑΙC
Poisson-Lognormal	189,71	0,00
Lognormal	190,43	0,72
Logserie	192,36	2,65

**Tabela 8.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância de biomassa** da comunidade de besouros Scarabaeinae coletadas na área **Encosta Norte**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes. Aqui utilizamos o **AIC** corrigido para pequenas amostras (**AICc**; Burnham & Anderson 2002).

MODELO	AICc	ΔΑΙC
Weibull	22,00	0,00
Gamma	22,10	0,10
Lognormal	22,50	0,50



**Figura 5.** Diagramas de rank-abundância para número de indivíduos (a) e biomassa (b) da comunidade de besouros Scarabaeinae coletada entre setembro/1997 a agosto/1998 na área Encosta Norte, Serra do Japi, São Paulo. Os pontos em preto são os dados empíricos e as linhas coloridas são os valores previstos pelos modelos teóricos de distribuição probabilística. O gráfico é apresentado em escala logarítmica.

#### Comunidade 5: Topo Morro

Para a SAD de indivíduos na área Topo Morro, os modelos selecionados pelo AIC foram a Poisson-Lognormal e a Logserie (Tabela 9). Para a abundância de biomassa, todos os modelos concorrentes foram selecionados pelo AIC (Tabela 10), no entanto, os modelos Weibull e Gamma foram descartados devido ao grande viés apresentado na análise de PME (Figura 11; seção 2.3.2).

Os modelos Poisson-Lognormal e Logserie para a SAD de indivíduos diferiram principalmente na previsão das abundâncias intermediárias das espécies (rank 3 à 13), que pode ser observado no gráfico de rank-abundância através da maior divergência das linhas nessas classes de abundância (Figura 6a). Já para as previsões das duas espécies mais dominantes e das espécies mais raras (rank 16 em diante) a previsão entre esses dois modelos foi praticamente equivalente, o que provavelmente gerou o empate entre os modelos, mesmo com uma previsão mais acurada do modelo Poisson-Lognormal que a Logserie (Figura 6a). O modelo Lognormal para SAD de biomassa acertou melhor a abundância das espécies com menor abundância e algumas intermediárias (rank 11 em diante), e fica evidente que a SAD de biomassa apresentou menor dominância de espécies que a SAD de número de indivíduos, devido à menor inclinação da linha dos dados (Figura 6b).

No rank de abundância numérica, a espécie *Canthidium trinodosum* se destacou pela maior dominância, com 128 indivíduos, seguida pela espécie *Canthonella* sp., com 50 indivíduos (Figura 6a). O grupo seguinte de espécies apresentaram abundância de 21-5 indivíduos, e depois destacou-se um grande número de espécies com abundância menor que quatro indivíduos (Figura 6a). As espécies *Ateuchus histrio, Deltochilum furcatum, Dichotomius* sp.1, *Paracanthon pereirai* e *Uroxys* sp.2 foram representados por um único indivíduo na amostra desta área. No rank de abundância de biomassa, todas as espécies apresentaram biomassa menor que

1g, sendo Canthidium trinodosum a espécie de maior biomassa (0,93g) e Uroxys sp.2

e Paracanthon pereirai as espécies com menor biomassa (cerca de 0,02g) (Figura 6b).

**Tabela 9.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância numérica** da comunidade de besouros Scarabaeinae coletadas na área **Topo Morro**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes.

MODELO	AIC	ΔΑΙC	
Poisson-Lognormal	149,09	0,00	
Logserie	149,76	0,67	
Lognormal	153,98	4,89	

**Tabela 10.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância de biomassa** da comunidade de besouros Scarabaeinae coletadas na área **Topo Morro**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes. Aqui utilizamos o **AIC** corrigido para pequenas amostras (**AICc**; Burnham & Anderson 2002).

MODELO	AICc	ΔΑΙC	
Lognormal	40,90	0,00	
Weibull	41,20	0,30	
Gamma	41,90	1,00	



**Figura 6.** Diagramas de rank-abundância para número de indivíduos (a) e biomassa (b) da comunidade de besouros Scarabaeinae coletada entre setembro/1997 a agosto/1998 na área Topo Morro, Serra do Japi, São Paulo. Os pontos em preto são os dados empíricos e as linhas coloridas são os valores previstos pelos modelos teóricos de distribuição probabilística. O gráfico é apresentado em escala logarítmica.

# Comunidade 6: Eucaliptus

O melhor modelo para descrever a SAD de indivíduos com os dados da amostra completa da área Eucaliptus foi a Logserie (Tabela 11). Para a abundância de biomassa os modelos Gamma e Weibull foram igualmente plausíveis para descrever a SAD (Tabela 12).

O modelo Logserie para a SAD de indivíduos conseguiu acertar as previsões de abundância das espécies relativamente bem (Figura 7a), bem como os modelos Gamma e Weibull para SAD de biomassa (Figura 7b). Os modelos Gamma e Weibull se apresentaram equivalentes em suas previsões de abundância (Figura 7b). Nos ranks de abundância, também ficou evidente, através da queda menos acentuada da curva, uma grande equitabilidade da biomassa, bem como uma grande diferença de abundância das espécies raras entre as duas medidas de abundância (Figura 7).

No rank de número de indivíduos, *Canthidium trinodosum* foi a espécie que apresentou a maior dominância, com 74 indivíduos (Figura 7a). As demais espécies com maior abundância numérica foram: *Scybalocanthon nigriceps* com 52 indivíduos, *Eurysternus cyanescens* com 41 indivíduos e *Canthonella* sp. com 29 indivíduos. As espécies seguintes apresentaram abundâncias menores que 18 indivíduos, e seis espécies foram representados por um único indivíduo na amostra desta área (*Canthidium* sp.2, *Deltochilum morbillosum, Dichotomius mormon, Ontherus azteca, Onthophagus* sp. e *Uroxys* sp.1) (Figura 7a). Para o rank de abundância de biomassa, a espécie *Uroxys* sp.1 se destaca com a menor biomassa da amostra (0,0005g), sendo também a menor biomassa encontrada para uma espécie entre todas as áreas de estudo (Figura 7b; Tabela 15). Já a espécie com maior abundância de biomassa foi *Deltochilum furcatum*, com 1,65g, e as demais espécies apresentaram abundâncias entre 1,500-0,005g (Figura 7b).

**Tabela 11.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância numérica** da comunidade de besouros Scarabaeinae coletadas na área **Eucaliptus**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes.

MODELO	AIC	ΔΑΙC
Logserie	150,81	0,00
Poisson-Lognormal	155,02	4,20
Lognormal	157,69	6,87

**Tabela 12.** Comparação do ajuste de três modelos teóricos de SAD com dados de **abundância de biomassa** da comunidade de besouros Scarabaeinae coletadas na área **Eucaliptus**. Os modelos são comparados com o Critério de Informação de Akaike (AIC), que é uma função da log-verossimilhança e do número de parâmetros de cada modelo. O modelo mais plausível é o de menor **AIC**. Modelos com uma diferença de **AIC** igual ou menor que 2 são considerados igualmente plausíveis. O  $\Delta$ **AIC** indica esta diferença em relação ao modelo mais plausível, diagnosticando modelos equivalentes. Aqui utilizamos o **AIC** corrigido para pequenas amostras (**AICc**; Burnham & Anderson 2002).

MODELO	AICc	ΔΑΙC	
Gamma	3,00	0,00	
Weibull	3,70	0,70	
Lognormal	6,30	3,30	



**Figura 7.** Diagramas de rank-abundância para número de indivíduos (a) e biomassa (b) da comunidade de besouros Scarabaeinae coletada entre setembro/1997 a agosto/1998 na área Eucaliptus, Serra do Japi, São Paulo. Os pontos em preto são os dados empíricos e as linhas coloridas são os valores previstos pelos modelos teóricos de distribuição probabilística. O gráfico é apresentado em escala logarítmica.

## 2.3.2 Probabilidade de Evidência Enganadora - PME

Considerando o princípio estatístico geral de que o aumento do tamanho amostral deve aproximar a distribuição amostral dos dados reais, e a hipótese de que o uso da biomassa como medida de abundância é mais eficiente para revelar a distribuição subjacente, esperamos que: (1) a PME deve cair com o aumento da amostra, tendendo a zero; e (2) esta queda deve ser mais rápida para modelos ajustados à biomassa.

Como esperado pela teoria, na maioria dos casos a PME caiu com o aumento do tamanho amostral, ou seja, a rarefação das amostras tende a aumentar a PME. No entanto, em alguns modelos a PME não seguiu este padrão, pois apresentaram valores altos mesmo para as maiores subamostras. Assim, mesmo que tenham sido selecionados pelo critério do AIC, descartamos esses modelos que não seguiram o comportamento teoricamente esperado da PME para as demais análises, foram eles: modelo Logserie para dados de indivíduos do Vale (Figura 8); modelo Logserie para dados de indivíduos do Vale Alterado (Figura 9); modelo Weibull para dados de biomassa da Encosta Sul (Figura 10); e modelos Gamma e Weibull para dados de biomassa para Topo Morro (Figura 11). No modelo Logserie ajustado aos dados de número de indivíduos do Vale, a PME aumentou monotonicamente de 3,6% na subamostra de tamanho 6 até 30% na de tamanho 30, e depois diminuiu levemente, ficando com valor da PME em 27,6% na maior subamostra composta por 42 unidades amostrais (Figura 8). Para a área Vale Alterado, o modelo Logserie ajustado aos dados de número de indivíduos apresentou valores de PME que aumentaram gradativamente de 12,9% à 26,2% entre os tamanhos das subamostras (Figura 9). Já na área Encosta Sul, o ajuste do modelo Weibull com dados de biomassa apresentou PME com um padrão de parábola invertida, com valores menores para as subamostras

de tamanho intermediário, cerca de 4%, e maior PME para as menores e maiores subamostras, cerca de 12% (Figura 10). Na área Topo Morro, os modelos Gamma e Weibull ajustados aos dados de biomassa também apresentaram PME com um padrão de parábola invertida. No modelo Gamma a PME variou de cerca de 6% para a menor subamostra, alcançou o maior valor de 30% nas subamostras de tamanhos 18 e 24, e depois atingiu o valor de 16,6% para a maior subamostra de 42 unidades amostrais (Figura 11). E no modelo Weibull a PME variou entre cerca de 17% nas subamostras de tamanho intermediário e cerca de 2,5% nas amostras de menor e maior tamanho (Figura 11).



**Figura 8.** Valores da PME (*probability of misleading evidence*) do modelo Logserie ajustado para subamostras simuladas por rarefação dos dados de abundância numérica de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Vale, Serra do Japi, São Paulo. No eixo x temos os tamanhos das subamostras tomadas sem reposição de um conjunto de 48 unidades amostrais (armadilhas/mês). No eixo y temos a proporção de 1000 dessas subamostras simuladas para as quais o modelo não foi selecionado pelo critério do AIC. Embora o modelo tenha sido selecionado pelo critério do AIC como bom descritor do conjunto total da amostra, o aumento da PME com o tamanho das subamostras foi utilizado como critério adicional para seleção de modelos, e dessa forma, o modelo foi descartado.



**Figura 9.** Valores da PME do modelo Logserie ajustado para subamostras simuladas por rarefação dos dados de abundância numérica de besouros Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Vale Alterado, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 8.



**Figura 10.** Valores da PME do modelo Weibull ajustado para subamostras simuladas por rarefação dos dados de abundância de biomassa de besouros Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Encosta Sul, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 8.



**Figura 11.** Valores da PME do modelo Weibull ajustado para subamostras simuladas por rarefação dos dados de abundância de biomassa de besouros Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Topo Morro, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 8.

Os modelos em que a PME caiu com o aumento do tamanho amostral, como esperado pela teoria, foram então utilizados para as demais análises do estudo. A comparação da PME entre os modelos ajustados para as duas medidas de abundância não apresentou um padrão claro e consistente de maior ou menor PME quando comparadas entre si. Em alguns casos os modelos ajustados aos dados de biomassa apresentaram tendência de menor PME comparado aos dados de número de indivíduos, como nas áreas Vale (Figura 12) e Eucaliptus (Figura 17), onde todas as curvas dos valores de PME ajustados aos modelos com dados de biomassa se apresentaram abaixo das curvas de PME com dados de número de indivíduos. Já em outros casos, os modelos ajustados aos dados de número de indivíduos foram os que apresentaram tendência de menor PME comparado aos dados de biomassa, como nas áreas Vale Alterado e Encosta Sul (Figuras 13 e 14). Para a área Vale, o modelo Weibull com dados de biomassa foi o que apresentou menor PME, mantendo cerca de 0% para todos os tamanhos de subamostra (Figura 12). Já o modelo com maior valor de PME foi a Lognormal com dados de indivíduos, que apresentou 70% de PME para o menor tamanho de subamostra, e este valor diminuiu até cerca de 10% para o maior tamanho de subamostra (Figura 12). O modelo Poisson-Lognormal com dados de número de indivíduos apresentou o segundo maior valor de PME para a menor subamostra, sendo de 45%, seguido pelos modelos Lognormal e Gamma com dados de biomassa, com valores de PME de 21% e 11%, respectivamente (Figura 12).



**Figura 12.** Valores da PME (*probability of misleading evidence*) dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triângulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Vale, Serra do Japi, São Paulo. No eixo x temos os tamanhos das subamostras tomadas sem reposição de um conjunto de 48 unidades amostrais (armadilhas/mês). No eixo y temos a proporção de 1000 dessas subamostras simuladas para as quais cada modelo não foi selecionado pelo critério do AIC, embora seja um modelo válido para descrever o conjunto total da amostra.

Na área Vale Alterado, o único modelo ajustado às subamostras com número de indivíduos, Poisson-Lognormal, foi o que apresentou valores menores de
PME, sendo cerca de 5% para a menor subamostra, e cerca de 0% para os demais tamanhos crescentes de subamostra (Figura 13). Já para o único modelo ajustado às subamostras com biomassa (Lognormal), os valores de PME foram maiores: para o menor tamanho de subamostra, a PME foi de 13%, e diminuiu monotônicamente até cerca de 0% para a subamostra de tamanho 24, e manteve este valor para as subamostras seguintes (Figura 13).



**Figura 13.** Valores da PME (*probability of misleading evidence*) dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Vale Alterado, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 12.

Na área Encosta Sul, o modelo Logserie apresentou os menores valores de PME para os ajustes das subamostras com número de indivíduos, sendo cerca de 2% para a menor subamostra, e 0% para os demais tamanhos (Figura 14). Já para o modelo Lognormal com biomassa, os valores foram maiores e apresentaram PME de cerca de 29% para as subamostras de tamanho 6 e 12, e a PME diminuiu gradativamente até tender a 0% nas subamostras de tamanho 36 e 42 (Figura 14).



**Figura 14.** Valores da PME (*probability of misleading evidence*) dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Encosta Sul, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 12.

Para a área Encosta Norte, o modelo que apresentou menor PME foi a Weibull ajustada aos dados de biomassa, com cerca de 1% de PME para as duas menores subamostras e de 0% para as subamostras dos demais tamanhos (Figura 15). O modelo seguinte de menor PME foi a Gamma, também ajustada aos dados de biomassa, com cerca de 4% para as duas menores subamostras, e tendendo a 0% nas demais subamostras (de tamanho 24 em diante) (Figura 15). Já o modelo com maior PME foi a Lognormal ajustada às subamostras com número de indivíduos, que apresentou 83% de PME para o menor tamanho de subamostra, seguido pela Lognormal com biomassa, que apresentou cerca de 40% de PME para as três menores subamostras, e Poisson-Lognormal com número de indivíduos, que teve PME de 30% na menor subamostra (Figura 15).



**Figura 15.** Valores da PME (*probability of misleading evidence*) dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Encosta Norte, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 12.

Na área Topo Morro, não foi possível calcular a PME dos modelos Poisson-Lognormal e Logserie ajustados às subamostras de tamanho 6 com número de indivíduos (Figura 16). Os modelos Poisson-Lognormal com número de indivíduos e Lognormal com biomassa apresentaram PME menor para as subamostras de tamanho maior que 12, com valores semelhantes entre si (Figura 16). Já para esses mesmos tamanhos de subamostra (>12), o modelo Logserie apresentou PME maior, com a curva formando uma parábola com valores de 10% nas subamostras de tamanho 18 e 24, e 0% na maior subamostra (Figura 16). A maior PME ocorre no modelo Lognormal ajustado aos dados de biomassa da subamostra de tamanho 6, que é de 20% (Figura 16).



**Figura 16.** Valores da PME (*probability of misleading evidence*) dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Topo Morro, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 12. Os modelos Poisson-Lognormal e Logserie ajustados aos dados de abundância numérica apresentaram problemas no otimizador para a estimativa de máxima verossimilhança com as subamostras de tamanho 6.

Na área Eucaliptus, os valores de PME foram baixos para o ajuste de todos os modelos, no entanto, a PME dos modelos ajustados aos dados de biomassa apresentaram valores menores (Figura 17). A PME do modelo Weibull com dados de biomassa foi de 0% para todos os tamanhos de subamostra (Figura 17). E o modelo Gamma, também ajustado aos dados de biomassa, apresentou PME de 0,5% para a menor subamostra e de 0% para os demais tamanhos de subamostra (Figura 17). O único modelo ajustado às subamostras com número de indivíduos, Logserie, foi o que apresentou maior PME, sendo de 3,3% para a menor subamostra, e cerca de 0% para os demais tamanhos crescentes de subamostra (Figura 17).



**Figura 17.** Valores da PME (*probability of misleading evidence*) dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Eucaliptus, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 12.

Avaliando a PME para o ajuste de cada modelo, apenas os modelos Gamma e Weibull apresentaram uma tendência de menor valor da PME quando comparado ao ajuste dos demais modelos, pois para todos os conjuntos de subamostras em que esses modelos foram ajustados, as PME apresentaram valores menores (e.g. Figuras 12, 15 e 17). Para a avaliação da PME dos demais modelos, não houve um padrão consistente de menor ou maior valores de PME com o ajuste dos dados. Por exemplo, o ajuste do modelo Logserie para os dados de número de indivíduos da área Encosta Sul apresentou valores pequenos de PME (Figura 14), enquanto o ajuste do modelo para os dados da área Eucaliptus já apresentou valores maiores de PME (Figura 17).

### 2.3.3 Viés da estimativa dos parâmetros

Considerando que em subamostras maiores o modelo ajustado, bem como as estimativas de seus parâmetros, tendem a convergir para o da amostra completa, a acurácia dessas estimativas dos parâmetros deve aumentar com o tamanho da amostra (i.e. maior viés de estimativa dos parâmetros para amostras menores). Se a hipótese deste trabalho estiver correta, este aumento da acurácia das estimativas dos parâmetros deve ocorrer mais rápido para os modelos ajustados aos dados de biomassa.

Os valores de viés das estimativas de máxima verossimilhança (MLEs) dos parâmetros obtidos pelo ajuste dos modelos de SAD às subamostras, de maneira geral, caiu com o aumento do tamanho amostral, como esperado teoricamente (Figuras 18 a 23). No entanto, houve casos em que o comportamento do viés não seguiu claramente esse padrão. No viés de estimativa do parâmetro do modelo Poisson-Lognormal para dados de abundância de indivíduos da área Vale, por exemplo, o comportamento dos valores de viés apresentou uma leve parábola ao longo dos tamanhos de rarefação da amostra, houve no entanto, pouca variação desses valores, de 0% à 5% aproximadamente (Figura 18). Para os dados de abundância de biomassa do modelo Lognormal da área Encosta Sul, o viés de estimação do parâmetro foi menor para a subamostra de tamanho 18 (aproximadamente 0%), enquanto para as subamostras de tamanho maiores (24 em diante) os valores de viés se mantiveram próximos a 2% e para as subamostras menores (de tamanho 12 e 6) o viés triplicou com a redução do tamanho da subamostra (Figura 20). No viés de estimativa do parâmetro do modelo Lognormal para os dados de abundância de biomassa da área Encosta Norte, o comportamento dos valores de viés também apresentou uma leve parábola ao longo dos tamanhos de rarefação da amostra, variando aproximadamente entre 1% à 6% (Figura 21). Já para o modelo Poisson-Lognormal, os valores de estimativa dos parâmetros para os dados de abundância de indivíduos na área Topo Morro se apresentou como uma forte parábola invertida, variando aproximadamente entre 1% e 14% (Figura 22).

Se uma medida de abundância gerasse estimativas de parâmetro pelo ajuste dos modelos com maior acurácia, a curva de viés em função do tamanho da subamostra para os modelos de biomassa deveria estar sempre acima das curvas para os modelos ajustados com número de indivíduos. Entretanto, nenhuma das duas medidas de abundância apresentou um padrão claro de valores consistentemente menores de viés de estimativa dos parâmetros com o ajuste dos modelos, para todos os locais ou modelos usados. Os modelos com abundância de biomassa apresentaram menor valores de viés de estimativa dos parâmetros para as áreas Vale Alterado e Encosta Norte (Figuras 19 e 21). Esse resultado ficou mais evidente para a área Vale Alterado, para o qual a curva de viés de estimativa dos MLEs de modelos para número de indivíduos esteve sempre acima das curvas para biomassa, chegando a ter valores até cinco vezes maiores de viés (Figura 19). Já na Encosta Norte, os menores valores de viés de estimativa dos parâmetros dos modelos ajustados aos dados de biomassa ficou mais evidente nos valores a partir da redução de 50% do tamanho das subamostras, onde as curvas de viés dos modelos com número de indivíduos ficaram bem acima das curvas dos modelos para biomassa (Figura 21). Já para a área Eucaliptus, os modelos para abundância de indivíduos apresentaram menores valores de viés de estimativa dos parâmetros, sendo que as curvas para esses modelos mostraram vieses, em geral, três vezes menores (Figura 23). Para a área Encosta Sul, os valores de viés para as duas medidas foram equivalentes, exceto para a amostra de tamanho 18, em que o modelo ajustado aos dados de biomassa apresentou valor de

viés da estimativa do parâmetro quatro vezes menor que o modelo com dados de número de indivíduos (Figura 20). Para as áreas Vale e Topo Morro, os valores de menor viés dependeram do modelo comparado entre as medidas de abundância (Figuras 18 e 22). Por exemplo, para o ajuste do modelo Poisson-Lognormal com dados de número de indivíduos da área Vale para comparar o viés de estimativa dos parâmetros com o modelo Gamma para dados de biomassa, este último apresentou valores de menor viés (Figura 18). No entanto, para o modelo Lognormal para comparar o viés de estimativa dos parâmetros ajustados aos mesmos conjuntos de dados, o modelo Lognormal ajustado aos dados de biomassa apresentou valores de viés menores que o mesmo modelo com dados de número de indivíduos, sendo até três vezes menor (Figura 18).

Assim, a diferença das medidas de viés de estimativa dos parâmetros também não constituiu um padrão claro favorecendo uma medida de abundância em favor da outra. Também não houve um melhor modelo que apresentasse consistentemente os menores valores desse viés, em todas as áreas de estudo.



**Figura 18.** Valores do viés da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Vale, Serra do Japi, São Paulo. No eixo x temos os tamanhos das subamostras tomadas sem reposição de um conjunto de 48 unidades amostrais (armadilhas/mês). No eixo y temos a proporção do cálculo das diferenças entre a média das estimativas dos valores dos parâmetros obtidos nas 1000 subamostras simuladas e o valor do parâmetro ajustado aos dados da amostra completa (de 48 unidades amostrais) pela estimativa de máxima verossimilhança. Os parâmetros utilizados foram os parâmetros de forma dos modelos Lognormal, Poisson-Lognormal, Gamma e Weibull, e o parâmetro alfa da Logserie.



**Figura 19.** Valores do viés da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Vale Alterado, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 18.



**Figura 20.** Valores do viés da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Encosta Sul, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 18.



**Figura 21.** Valores do viés da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Encosta Norte, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 18.



**Figura 22.** Valores do viés da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Topo Morro, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 18.



**Figura 23.** Valores do viés da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Eucaliptus, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 18.

#### 2.3.4 Coeficiente de Variação da estimativa dos parâmetros

O coeficiente de variação das estimativas dos parâmetros foi usado como uma medida de precisão dos MLEs obtidos com o ajuste dos modelos a cada subamostra simulada. A precisão das estimativas dos parâmetros, assim como a acurácia, deve aumentar com o tamanho da amostra (i.e. valores menores de coeficiente de variação das estimativas dos parâmetros em amostras maiores). Assim, se a hipótese deste trabalho estiver correta, o aumento da precisão das estimativas dos parâmetros também deve ocorrer mais rápido para os modelos ajustados aos dados de biomassa. Todos os modelos comparados apresentaram consistentemente, valores maiores de coeficiente de variação das estimativas dos parâmetros com a diminuição do tamanho da amostra. Ou seja, como esperado em teoria, a precisão das estimativas dos parâmetros (MLEs) aumentou com o tamanho da amostra.

Os valores do coeficiente de variação das estimativas dos parâmetros nos modelos ajustados à abundância de biomassa foi menor para três das seis áreas de estudo (i.e. mais precisa): Vale Alterado, Encosta Sul e Topo Morro; como podemos observar pela sobreposição da linha do coeficiente de variação com dados de indivíduos em relação a linha com dados de biomassa (Figuras 25, 26 e 28). Nesses casos, os valores do coeficiente de variação das estimativas dos parâmetros ajustados aos dados de número de indivíduos foram cerca de duas vezes maiores que para os dados de biomassa (Figuras 25, 26 e 28). Para as demais áreas (Vale, Encosta Norte e Eucaliptus), a comparação do coeficiente de variação dos parâmetros com as diferentes medidas de abundância dependeu dos modelos comparados, que apresentaram menor ou maior precisão dependendo da escolha do modelo (Figuras 24 e 27). Por exemplo, se comparamos os valores do coeficiente de variação das estimativas dos parâmetros para os modelos Weibull e Lognormal com dados de biomassa, e para o modelo Lognormal com dados de número de indivíduos na área Encosta Norte, a precisão entre os modelos se demonstrou praticamente equivalente, enquanto se considerarmos o modelo Poisson-Lognormal para número de indivíduos ou Gamma para biomassa da mesma área, os valores do coeficiente de variação das estimativas dos parâmetros foram maiores que a dos modelos citados anteriormente (Figura 24). Também para os modelos Lognormal, Gamma e Weibull com dados de biomassa e Lognormal com número de indivíduos na área Vale, os resultados do coeficiente de variação foram equivalentes, enquanto os valores chegaram a ser três vezes maiores para o modelo Poisson-Lognormal com número de indivíduos da mesma área (Figura 27). Para a área Eucaliptus, o resultado da análise do coeficiente de variação das estimativas dos parâmetros apresentou valores praticamente equivalentes dessa medida de precisão entre os modelos ajustados aos dados de ambas as medidas de abundância (Figura 29).



**Figura 24.** Valores do coeficiente de variação da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Vale, Serra do Japi, São Paulo. No eixo x temos os tamanhos das subamostras tomadas sem reposição de um conjunto de 48 unidades amostrais (armadilhas/mês). No eixo y temos a proporção do cálculo do coeficiente de variação da estimativa dos parâmetros obtidos com o ajuste dos modelos às 1000 subamostras simuladas. O coeficiente de variação corresponde ao valor do desvio padrão desses parâmetros estimados dividido pela média dos mesmos parâmetros. Os parâmetros utilizados foram os parâmetros de forma dos modelos Lognormal, Poisson-Lognormal, Gamma e Weibull, e o parâmetro alfa da Logserie.



**Figura 25.** Valores do coeficiente de variação da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Vale Alterado, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 24.



**Figura 26.** Valores do coeficiente de variação da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Encosta Sul, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 24.



**Figura 27.** Valores do coeficiente de variação da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Encosta Norte, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 24.



**Figura 28.** Valores do coeficiente de variação da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Topo Morro, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 24.



**Figura 29.** Valores do coeficiente de variação da estimativa dos parâmetros dos modelos ajustados para subamostras simuladas por rarefação dos dados de abundância numérica (triangulos) e de biomassa (pontos) de besouros da subfamília Scarabaeinae coletados entre setembro/1997 a agosto/1998 na área Eucaliptus, Serra do Japi, São Paulo. Mais detalhes na legenda da Figura 24.

### 2.3.5 Comparação dos resultados

A Tabela 13 apresenta uma síntese dos resultados das análises que compararam a eficiência da medida de abundância para descrever as amostras das comunidades com os modelos de SAD. As estimativas da PME, que representam uma medida de suporte de evidência dos modelos de SAD com as diferentes medidas de abundância, e o viés e coeficiente de variação das estimativas dos parâmetros, que representam, respectivamente, medidas de avaliação da acurácia e precisão desses modelos, não apresentaram um padrão consistente. Em resumo, não há uma tendência clara de que os modelos de SADs ajustados a dados de biomassa convirjam para os modelos verdadeiros mais rapidamente, pelo menos pelos critérios de PME ou de acurácia e precisão das estimativas dos parâmetros.

**Tabela 13.** Síntese dos resultados das análises utilizadas para avaliar qual medida de abundância foi melhor para revelar a distribuição original da comunidade de besouros Scarabaeinae de diferentes áreas da Serra do Japi, São Paulo, com a diminuição da amostra. As células da tabela indicam a medida de abundância cujos modelos se adequaram melhor quanto a cada critério de convergência.

Critério de ineficiência	Vale	Vale Alterado	Encosta Sul	Encosta Norte	Topo Morro	Eucaliptus
РМЕ	Biomassa	Indivíduos	Indivíduos	Depende do modelo	Depende do modelo	Biomassa
Viés parâmetro	Depende do modelo	Biomassa	Depende do modelo	Biomassa	Depende do modelo	Indivíduos
CV parâmetro	Depende do modelo	Biomassa	Biomassa	Depende do modelo	Biomassa	Depende do modelo

## 2.4 Discussão e Conclusões

A primeira vez que o conceito de veil line ("linha de véu") foi considerado para descrever o efeito da amostragem nas SADs, foi através do trabalho de Preston (1948), junto com a proposição do modelo Lognormal para descrever as SADs de número de indivíduos nas comunidades biológicas. Este conceito surgiu da ideia de que as amostras insuficientes das comunidades naturais, ou de pequeno tamanho, geravam uma distribuição que se apresentava parcialmente revelada (Logserie ou Lognormal truncada) por excluírem as espécies mais raras da amostra. Assim, a veil line seria a classe de abundância abaixo da qual as espécies não seriam incluídas na amostra, e essa linha seria deslocada com o aumento do tamanho amostral revelando (unveiling) a distribuição original da comunidade. Esta ideia de unveiling, associada por Preston (1948) ao efeito do tamanho amostral, pode ser interpretada através do princípio estatístico de consistência, que implica que o acumulo de evidência, representado pelo aumento do tamanho amostral, favorece a aproximação ao modelo verdadeiro bem como as estimativas precisas e acuradas de seus parâmetros (Edwards 1992; Royall 2000). Desta forma, considerando que as estimativas de máxima verossimilhança - MLEs satisfazem o princípio estatístico de consistência (Edwards

1992), as MLEs representam uma forma robusta de investigar o unveiling dos modelos de SAD. Apesar do conceito de unveiling e o princípio estatístico de consistência serem conhecidos há tempos na literatura (Preston 1948; Edwards  $1992^2$ ), este trabalho é o primeiro a combiná-los para avaliar as SADs. A inovação do presente estudo para investigar o unveiling das SADs com as duas medidas de abundância avaliadas (biomassa e número de indivíduos) foi transformar estes conceitos em três critérios de convergência assintótica dos modelos: (1) redução de PME; (2) diminuição do viés da estimativa dos parâmetros; (3) aumento da precisão da estimativa de parâmetros com o aumento da subamostra obtidas por rarefação (amostragem sem reposição). Os critérios de análise utilizados apresentam a grande vantagem de serem aplicáveis e comparáveis para todos os modelos, além de sólida base teórica amplamente aceita e utilizada na teoria estatística, que são as abordagens da verossimilhança estatística para o ajuste de modelos e o critério do AIC para seleção de modelos concorrentes (Johnson & Omland 2004). Dessa forma os critérios de análise utilizados se demonstraram práticos, objetivos e adequados para a investigação proposta.

Alguns modelos utilizados para o ajuste do conjunto de dados das SAD dos besouros Scarabaeinae do estudo não apresentaram a maior convergência esperada pela teoria com o aumento do tamanho amostral. O fato de que para alguns modelos a PME não tenha caído monotonicamente como esperado pela teoria (ver Figuras 8 a 11, e resultados na seção 2.3.2) ofereceu um critério que parece interessante para o desempate de modelos. Levantamos a hipótese de que se o incremento de evidência com o aumento da amostra não aproxima a distribuição da amostra à sua distribuição original, isso indicaria falta de consistência do modelo

<sup>&</sup>lt;sup>2</sup> A primeira edição do livro "Likelihood", em que Edwards demonstrou que a lei da verossimilhança estatística satisfaz a propriedade de *consistência*, foi publicada em 1935.

(Royall 2000), o que pode ser utilizado como critério adicional da seleção de modelos. Então, quando o comportamento da PME em relação ao tamanho amostral não é o esperado pela teoria, contrariando um princípio estatístico bem estabelecido, isso indica que na verdade o modelo não representa um bom descritor dos dados. Assim, esse critério posterior ao uso do AIC para avaliar a qualidade do modelo parece lógico e interessante, além de ser inédito na literatura, e pode ser melhor explorado em estudos futuros. O uso de simulações computacionais oferece um interessante caminho para isso, pois é possível uma ampla investigação do comportamento da PME com diferentes modelos, onde as SADs podem ser simuladas a partir de modelos pré-estabelecidos. Com isso, a PME de cada modelo pode ser melhor avaliada através da rarefação de uma amostra com distribuição previamente conhecida, diminuindo desta forma as dúvidas de qual é a distribuição original dos dados.

Apesar do conceito de *unveiling* ser antigo na literatura, só recentemente ele foi relacionado com o uso de diferentes medidas de abundância nas SADs. O trabalho pioneiro que relacionou o conceito de *unveiling* para o uso de diferentes medidas de abundância para representar a importância relativa das espécies foi o de Connolly *et al.* (2005). Esses autores investigaram o efeito da escala de amostragem com a utilização da biomassa e número de indivíduos como medida de abundância nas SADs das comunidades marinhas de peixes e corais. O estudo sugeriu que nas comunidades estudadas, as SADs com as duas medidas de abundância seguem uma distribuição aproximadamente Lognormal, que, no entanto, emergem em escalas marcadamente diferentes. A SAD de biomassa se apresentou como uma Lognormal em escalas locais de amostragem, enquanto na SAD de número de indivíduos, a distribuição Lognormal foi revelada (*unveiling*) apenas com o aumento da escala espacial de amostragem (Connolly et al. 2005). O aumento de escala no estudo de Connolly et al. foi obtido agrupando as parcelas dos vários locais amostrados. Dessa forma, há dois efeitos misturados: o aumento da escala e o aumento do tamanho amostral. Posteriormente, Morlon et al. (2009) também encontraram resultados similares para investigação semelhante do efeito da escala de amostragem e da medida de abundância utilizada na SAD com comunidades de aves, peixes de água doce e árvores. No estudo de Morlon e colaboradores as SADs de biomassa também se aproximam mais rapidamente de uma Lognormal em escala espacial menor que as SADs de número de indivíduos, e os efeitos de aumento da escala e aumento do tamanho amostral também são misturados. Dessa forma, os trabalhos de Connolly et al. (2005) e Morlon et al. (2009) sugeriram que a Lognormal representa a distribuição subjacente mais comum para as comunidades naturais, tanto para a SAD de indivíduos quanto para a SAD de biomassa, enquanto outros autores propuseram o mesmo modelo apenas para SAD de número de indivíduos (e.g. Preston 1948; Williamson & Gaston 2005; Ulrich & Zalewski 2006; Sizling et al. 2009). Os resultados similares encontrados por Connolly et al. (2005) e Morlon et al. (2009) para diferentes grupos taxonômicos, em que a biomassa revelou a SAD mais rapidamente para escalas espaciais menores de amostragem, introduziu o que Williamson (2010) denominou de differential veiling ("revelamento diferencial"), e remete à ideia de que a biomassa seria uma medida de abundância mais eficiente em revelar a SAD da comunidade. No entanto, apesar de Williamson (2010) utilizar o termo *differential veiling* para se referir aos resultados encontrados pelos trabalhos de Connolly et al. (2005) e Morlon et al. (2009), que remete à ideia introduzida por Preston (1948) relacionada apenas ao efeito do tamanho amostral, os trabalhos desses autores abordam, na verdade, o efeito misturado do tamanho amostral e da escala espacial. Williamson (2010) em sua discussão sobre o tema relaciona diretamente os resultados encontrados por Connolly *et al.* e Morlon *et al.* ao efeito do tamanho amostral. No entanto, apesar de haver uma confusão entre estes dois efeitos, por estarem frequentemente associados, um aumento da amostra não implica necessariamente em aumento de escala. Por exemplo, se consideramos apenas um aumento na intensidade amostral (e.g. aumento do número de armadilhas) em uma mesma escala espacial e temporal, as amostras resultantes serão maiores sem representar uma escala espacial ou temporal maior. Da mesma forma, se consideramos um aumento da escala amostrada, em que as amostras entre todas as escalas apresentem o mesmo esforço amostral, teremos amostras de mesmo tamanho mas com escalas diferentes.

No presente estudo, a metodologia de rarefação das amostras com as diferentes medidas de abundância também misturou o efeito do tamanho amostral com o efeito de escala, devido à redução da representação dos meses e unidades amostrais nas subamostras, representando simultaneamente ambos os efeitos. Assim, os resultados encontrados neste estudo mostram que a maior eficiência da biomassa para revelar a distribuição subjacente não é tão geral quanto sugerem os trabalhos anteriores. Ou seja, a afirmativa de que a SAD de biomassa não é oculta (*veiled*) não parece ser geral. No entanto, se a causa do *differential veiling* entre as SADs de biomassa e número de indivíduos for principalmente a escala, uma explicação alternativa é que o presente estudo não abarcou uma escala tão ampla quanto os estudos de Connolly *et al.* e Morlon *et al.*, e, por isso não demonstrou o efeito do *differential veiling* devido à escala. No trabalho de Connolly *et al.* (2005), por exemplo, as escalas foram representadas por parcelas de dezenas de metros, que agrupadas abarcaram áreas de dezenas de quilômetros para representar as escalas

maiores. Já no presente estudo, a distância entre as unidades amostrais de cada área não ultrapassa a ordem de dezena de metros. No entanto a amostragem e metodologia de rarefação de nosso estudo também incorporam o efeito da escala temporal, pois a amostragem ocorreu mensalmente ao longo do período de um ano, enquanto os demais estudos não mencionam que as coletas abranjam escala temporal.

Os resultados obtidos no presente estudo também indicam que o padrão de differential veiling entre as SADs de biomassa e número de indivíduos podem variar dependendo da escolha do modelo e do conjunto de dados avaliados. Por exemplo, em alguns casos, os modelos ajustados aos dados de biomassa apresentaram tendência de menor PME comparado aos dados de número de indivíduos, como nas áreas Vale e Eucaliptus (Figuras 12 e 17); e em outros casos os modelos ajustados aos dados de número de indivíduos apresentaram tendência de menor PME comparado aos dados de biomassa, como nas áreas Vale Alterado e Encosta Sul (Figuras 13 e 14). Assim, a comparação da PME entre os modelos ajustados com as diferentes medidas de abundância neste estudo dependeu dos modelos comparados. Para a área da Encosta Norte, por exemplo, se comparamos o ajuste das distribuições Gamma ou Weibull para os dados de biomassa, ambos os modelos tiveram valores menores de PME se comparados ao ajuste do modelo Poisson-Lognormal para dados de número de indivíduos (Figura 15). No entanto, se comparamos o ajuste da Lognormal para os mesmos dados de biomassa, a PME foi maior quando comparada ao modelo Poisson-Lognormal ajustado aos dados de número de indivíduos (Figura 15). Já para a área Topo Morro, se comparamos os ajustes da Lognormal com dados de biomassa e da Poisson-Lognormal para dados de número de indivíduos, os valores da PME foram praticamente equivalentes, enquanto se ao invés da Poisson-Lognormal considerarmos os ajustes do modelo Logserie para os mesmos dados de número de indivíduos, a PME foi maior no ajuste de quase todos os tamanhos de subamostra (Figura 16). O fato da PME para o ajuste dos modelos Gamma e Weibull terem apresentado tendência de menor valor da PME para todos os conjuntos de subamostras em que esses modelos foram ajustados, quando comparado ao ajuste dos demais modelos (e.g. Figuras 12, 15 e 17), ocorreu provavelmente devido ao fato desses dois modelos (Gamma e Weibull) serem bastante flexíveis e apresentarem geralmente bons ajustes às SADs de vários conjuntos de dados. Essa variação do modelo de maior consistência entre as áreas e medidas de abundância do estudo também ocorreu para as análises de viés e precisão das estimativas dos parâmetros desses modelos (ver Figuras 18 a 29, e resultados nas seções 2.3.3 e 2.3.4). Assim, a comparação dos critérios de análise utilizados para avaliar a convergência dos modelos ajustados para as duas medidas de abundância (PME, viés e coeficiente de variação das estimativas dos parâmetros) não apresentou um padrão claro e consistente favorecendo uma medida de abundância em detrimento da outra (ver síntese dos resultados na Tabela 13). Apesar de haver diferença de desempenho entre os modelos, o padrão não é consistente a favor da medida de biomassa, mesmo havendo casos em que o uso da biomassa tenha proporcionado um melhor desempenho do modelo. Portanto, a hipótese do estudo, de que a biomassa seria uma medida mais eficiente em revelar a SAD subjacente não foi corroborada. Concluímos, desta forma, a partir dos critérios utilizados, que a melhor medida de abundância para revelar a SAD vai depender do conjunto de dados e dos modelos utilizados.

A interpretação dos resultados deve, no entanto, ser feita com cautela, já que uma questão importante na escolha da medida de abundância é também o que ela pretende investigar. Quando consideramos questões em que as interações tróficas ou partilha de recursos podem ser processos estruturadores importantes, a biomassa deve prover uma medida mais adequada desses processos por ser mais diretamente relacionada com a utilização de recursos que a abundância de indivíduos (Saint-Germain *et al.* 2007). Já se a questão do estudo for relacionada, por exemplo, aos fatores demográficos da população, como frequência de interação entre indivíduos (e.g. funções desempenhadas por polinizadores e vetores de patógenos), o uso do número de indivíduos como medida de abundância é mais informativo (Saint-Germain *et al.* 2007). Nesses casos, a escolha da biomassa ou do número de indivíduos como medida de abundância se daria pela adequação da medida em representar a variável de interesse, e não por revelar melhor a SAD subjacente.

Para os dados analisados, a distribuição de biomassa mostrou-se sempre mais equitativa que a distribuição de número de indivíduos, o que revela que as SADs das duas medidas de abundância não são equivalentes. A maior equitabilidade da biomassa se deve, provavelmente, à tendência de algumas espécies de menor tamanho apresentarem maiores abundâncias numéricas e as espécies de maior tamanho terem abundâncias numéricas menores. Alguns trabalhos sugerem uma grande semelhanca entre a distribuição de indivíduos, biomassa ou utilização de recursos (Sugihara 1989; Taper & Marquet 1996), enquanto outros sugerem importantes diferenças (Tokeshi 1990; Pagel et al. 1991; Connolly et al. 2005; Saint-Germain et al. 2007; Morlon et al. 2009; Henderson & Magurran 2010). Assim, é esperada uma observação de padrões de SAD similares na utilização de diferentes medidas de abundância das espécies se a abundância numérica, a biomassa e a utilização de recursos são proporcionais uns aos outros (Morlon et al. 2009). Esta condição pode ser encontrada se todos os indivíduos podem ser considerados ecologicamente equivalentes, ou seja, todas as espécies, independente de seu tamanho, possuem a mesma capacidade de utilização de recursos e taxas demográficas (Hubbell 2001). No entanto, para os casos

em que a relação entre medidas de abundância alternativas apresentam importantes diferenças, espera-se, por exemplo, que modelos baseados na repartição de nicho sejam mais diretamente explicados pela distribuição de biomassa e utilização de recursos que pela distribuição da abundância numérica (Tokeshi 1993; Thibault *et al.* 2004; Connolly et al. 2005).

Considerando que o tamanho da amostra representa uma importante fonte de viés na representação da SAD da comunidade e que a maioria dos dados de SAD são baseados em amostras limitadas (Pielou 1977; Green & Plotkin 2007), a proposição do presente estudo combina relevância teórica e aplicada para compreender como os padrões de abundância das amostras se relacionam com a SAD subjacente da comunidade com o uso de diferentes medidas de abundância. McGill et al. (2007) consideram as investigações de medidas alternativas de abundância e suas implicações como uma das prioridades no estudo das SADs. A metodologia proposta neste estudo para avaliar o differential veiling associado ao tamanho amostral através da convergência do ajuste dos modelos de SAD segue critérios simples e práticos para investigar a questão e também contribui para o fortalecimento metodológico para o estudo das SADs em ecologia de comunidades. No entanto, ainda há necessidade de futuras investigações do differential veiling entre as medidas de abundância nas SAD. Os efeitos da escala e tamanho amostral podem ser investigados separadamente em estudos futuros, combinando a análise dos critérios de convergência propostos neste estudo com simulação computacional. Com isso é possível investigar o differential veiling sistematicamente, desacoplando o efeito de escala e tamanho amostral através da simulação de SAD a partir de uma distribuição probabilística já conhecida, em que é possível simular a variação de escala e tamanho amostral separadamente, e posteriormente utilizar os critérios de convergência aqui propostos para as análises.

## Referências

- Akaike, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. *In:* Second International Symposium on Information Theory (Eds. B. N. Petrov & F. Csaki). Akademiai Kiado.
- Bolker, B. 2012. Package bbmle: Tools for general maximum likelihood estimation. (R package version 1.0.4.1, based on stats4 by the R Development Core Team). URL http://cran.r-project.org/.
- Burnham, K. P. & Anderson, D. R. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer.
- Connolly, S. R.; Hughes, T. P.; Bellwood, D. R. & Karlson, R. H. 2005. Community structure of corals and reef fishes at multiple scales. Science 309: 1363-1365.

Edwards, A. W. F. 1992. Likelihood. The Johns Hopkins University Press.

- Green, J. & Plotkin, J. 2007. A statistical theory for sampling species abundance. Ecology Letters 10: 1037-1045.
- Henderson, P. A. & Magurran, A. E. 2010. Linking species abundance distributions in numerical abundance and biomass through simple assumptions about community structure. Proceedings of The Royal Society 277: 1561-1570.
- Hernández, M. I. M.; Monteiro, L. R. & Favila, M. E. 2011. The role of body size and shape in understanding competitive interactions within a community of Neotropical dung beetles. Journal of Insect Science 11(13): 1-14.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press.
- Johnson, J. B. & Omland, K. B. 2004. Model selection in ecology and evolution. Trends in Ecology and Evolution 19: 101-108.
- Magurran, A. E. 2004. Measuring Biological Diversity. Blackwell Publishing.
- Magurran, A. E. & McGill, B. J. 2011. Biological Diversity. Oxford University Press.
- McGill, B. J. 2003. Strong and weak tests of macroecological theory. Oikos 102: 679-685.
- McGill, B. J. 2011. Species abundance distribution. *In:* A. E. Magurran & B. J. McGill (ed.) Biological Diversity. Oxford University Press.

- McGill, B. J.; Etienne, R. S.; Gray, J. S.; Alonso, D.; Anderson, M. J.; Benecha, H. K.; Dornelas, M.; Enquist, B. J.; Green, J. L.; He, F.; Hurlbert, A. H.; Magurran, A. E.; Marquet, P. A.; Maurer, B. A.; Ostling, A.; Soykan, C. U.; Ugland, K. I. & White, E. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecology Letters 10: 995-1015.
- Morlon, H.; White, E. P.; Etienne, R. S.; Green, J. L.; Ostling, A.; Alonso, D.;
  Enquist, B. J.; He, F. L.; Hurlbert, A.; Magurran, A. E.; Maurer, B. A.;
  McGill, B. J.; Olff, H.; Storch, D. & Zillio, T. 2009. Taking species abundance distributions beyond individuals. Ecology Letters 12: 488-501.
- Pagel, M. D.; Harvey, P. H. & Godfray, H. C. J. 1991. Species-abundance, biomass and resource use distribution. American Naturalist 138: 836-850.
- Pielou, E. C. 1977. Mathematical Ecology. John Wiley and Sons.
- Prado, P. I. 2012. Package sads: Statistical modeling of species abundance distributions. (R package version 0.2.04, work in progress with R version 2.14.2 by the R Development Core Team). Available in provisory URL http://143.107.246.244/let/doku.php?id=engl:tutorials:rcode.
- Preston, F. W. 1948. The commonness, and the rarity, of species. Ecology 29: 254-283.
- R Development Core Team 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing. ISBN 3-900051-07-0, URL http://www.R-project.org.
- Royall, R. M. 2000. Statistical Evidence: A likelihood paradigm. Chapman & Hall/CRC.
- Saint-Germain, M.; Buddle, C. M.; Larrivée, M.; Mercado, A.; Motchula, T.; Reichert, E.; Sackett, T. E.; Sylvain, Z. & Webb, A. 2007. Should biomass be considered more frequently as a currency in terrestrial arthropod community analysis? Journal of Applied Ecology 44: 330-339.
- Sizling, A. L.; Storch, D.; Sizlingová, E.; Reif, J. & Gaston, K. J. 2009. Species abundance distributions results from a spatial analogy of central limit theorem. Proceedings of the National Academy of Sciences of the United States of America 106(16): 6691-6695.

- Sugihara, G. 1989. How do species divide resources? American Naturalist 133: 770-787.
- Taper, M. L. & Marquet, P. A. 1996. How do species really divide resources? American Naturalist 147: 1072-1086.
- Thibault, K.; White, E. & Ernest, S. K. M. 2004. Temporal dynamics in the structure and composition of a desert rodent community. Ecology 85: 2649-2655.
- Tokeshi, M. 1990. Niche apportionment or random assortment: species abundance patterns revisited. Journal of Animal Ecology 59: 1129-1146.
- Tokeshi, M. Species abundance patterns and community structure. 1993. Advances in Ecological Research 24: 111-186.
- Ulrich, W. & Zalewski, M. 2006. Abundance and co-occurrence patterns of core and satellite species of ground beetles on small lake islands. Oikos 114: 338-348.
- Williamson, M. 2010. Why do species abundance distributions of individuals and biomass behave differently under sampling? Oikos 119: 1697-1699.
- Williamson, M. & Gaston, K. J. 2005. The appropriate null hypothesis for the speciesabundance distribution. Journal of Animal Ecology 74: 409-422.

# ANEXOS

Anexo 1. Abundância de indivíduos das espécies de besouros da subfamília Scarabaeinae coletados mensalmente entre setembro/1997 a agosto/1998 em seis diferentes áreas da Serra do Japi, São Paulo, Brasil. V=Vale; VA=Vale Alterado; ES=Encosta Sul; EN=Encosta Norte; TM=Topo Morro; EU=Eucaliptus. Dados extraídos de Hernández *et al.* (2011).

TRIBO	ESPÉCIE	V	VA	ES	EN	ТМ	EU	Total
Ateuchini	Ateuchus prox. histrio (Balthasar, 1939)	0	0	1	0	1	0	2
	Canthidium dispar Harold, 1867	0	0	0	1	0	0	1
	Canthidium prox. sulcatum (Perty, 1830)	0	0	1	0	0	0	1
	Canthidium trinodosum (Boheman, 1858)	43	1031	174	164	128	74	1614
	Canthidium sp. 1	0	0	0	0	5	0	5
	Canthidium sp. 2	6	46	17	18	2	1	90
	Trichillum sp. 1	1	8	4	4	2	0	19
	Trichillum sp. 2	1	0	0	0	0	0	1
	Uroxys aterrima Harold, 1867	7	4	1	9	3	0	24
	Uroxys kratochvili Balthasar, 1940	8	62	158	7	2	18	255
	Uroxys lata Arrow, 1933	64	38	75	5	11	3	196
	Uroxys sp. 1	15	5	4	6	0	1	31
	Uroxys sp. 2	0	0	5	0	1	0	6
Coprini	Dichotomius assifer (Eschscholtz, 1822)	16	31	57	18	2	9	133
	Dichotomius bechynei Martínez, 1973	0	0	3	1	0	0	4
	Dichotomius carbonarius (Mannerheim, 1929)	0	0	0	0	2	0	2
	Dichotomius depressicollis (Harold, 1867)	0	0	1	0	0	0	1
	Dichotomius mormon (Ljungh, 1799)	2	1	0	1	0	1	5
	Dichotomius sp. 1	10	2	6	23	1	0	42
	Dichotomius sp. 2	0	0	0	0	0	2	2
	Ontherus azteca Harold, 1869	0	6	1	0	0	1	8
Eurysternini	Eurysternus cyanescens Balthasar, 1939	65	79	59	26	21	41	291
	Eurysternus hirtellus Dalman, 1824	0	1	10	3	3	13	30
	Eurysternus parallelus Castelnau, 1840	1	0	0	0	2	0	3
	Eurysternus sp.	9	1	12	6	8	14	50
Phanaeini	Coprophanaeus saphirinus (Sturm, 1826)	3	1	1	0	0	0	5
	Phanaeus splendidulus (Fabricius, 1781)	1	0	3	0	0	0	4
Onthophagini	Onthophagus sp.	0	0	4	0	8	1	13
Canthonini	Canthon latipes Blanchard, 1845	25	1	5	19	9	5	64
	Canthon rutilans Castelnau, 1840	0	0	0	0	3	0	3
	<i>Canthon</i> sp.	4	0	0	4	0	0	8
	Canthonella sp.	12	34	6	30	50	29	161
	Deltochilum brasiliense (Castelnau, 1840)	8	7	8	5	0	3	31
	Deltochilum furcatum (Castelnau, 1840)	32	9	17	23	1	9	91
	Deltochilum morbillosum Burmeister, 1848	5	10	1	16	0	1	33
	Deltochilum rubripenne (Gory, 1831)	11	1	18	11	3	16	60
	Paracanthon pereirai d'Andretta & Martínez, 1957	4	1	1	0	1	8	15
	Scybalocanthon nigriceps (Harold, 1868)	12	0	18	107	14	52	203
	Sylvicanthon foveiventre (Schmidt, 1920)	3	1	1	7	2	3	17
	Numero de espècies	26	23	30	24	25	22	39
	Numero de indivíduos	368	1380	6/2	514	285	305	3524

**Anexo 2.** Abundância de biomassa das espécies de besouros da subfamília Scarabaeinae coletados mensalmente entre setembro/1997 a agosto/1998 em seis diferentes áreas da Serra do Japi, São Paulo, Brasil. V=Vale; VA=Vale Alterado; ES=Encosta Sul; EN=Encosta Norte; TM=Topo Morro; EU=Eucaliptus. Dados extraídos de Hernández *et al.* (2011).

TRIBO	ESPÉCIE	V	VA	ES	EN	TM	EU	Total
Ateuchini	Ateuchus near histrio (Balthasar, 1939)	0	0	0,0119	0	0,0118	0	0,0237
	Canthidium dispar Harold, 1867	0	0	0	0,0362	0	0	0,0362
	Canthidium near sulcatum (Perty, 1830)	0	0	0,0189	0	0	0	0,0189
	Canthidium trinodosum (Boheman, 1858)	0,2282	7,7638	1,0487	0,8463	0,9391	0,5196	11,3457
	Canthidium sp. 1	0	0	0	0	0,0337	0	0,0337
	Canthidium sp. 2	0,0301	0,2283	0,0741	0,1029	0,0089	0,0052	0,4495
	Trichillum sp. 1	0,0011	0,0156	0,0080	0,0079	0,0041	0	0,0367
	Trichillum sp. 2	0,0023	0	0	0	0	0	0,0023
	Uroxys aterrima Harold, 1867	0,0811	0,0654	0,0194	0,1433	0,0489	0	0,3581
	Uroxys kratochvili Balthasar, 1940	0,0211	0,1364	0,4542	0,0161	0,0065	0,473	0,6816
	Uroxys lata Arrow, 1933	0,2648	0,1458	0,3206	0,0137	0,0514	0,0115	0,8078
	Uroxys sp. 1	0,0099	0,0029	0,0028	0,0036	0	0,0005	0,0197
	Uroxys sp. 2	0	0	0,0120	0	0,0026	0	0,0146
Coprini	Dichotomius assifer (Eschscholtz, 1822)	2,7701	4,9831	10,1580	3,1486	0,2806	1,4174	22,7578
	Dichotomius bechynei Martínez, 1973	0	0	0,2424	0,0766	0	0	0,3190
	Dichotomius carbonarius (Mannerheim, 1929)	0	0	0	0	0,2978	0	0,2978
	Dichotomius depressicollis (Harold, 1867)	0	0	0,1900	0	0	0	0,1900
	Dichotomius mormon (Ljungh, 1799)	0,7378	0,4279	0	0,2975	0	0,4728	1,9360
	Dichotomius sp. 1	0,5846	0,1148	0,3152	1,3333	0,0308	0	2,3787
	Dichotomius sp. 2	0	0	0	0	0	0,1589	0,1589
	Ontherus azteca Harold, 1869	0	0,3186	0,0695	0	0	0,0689	0,4570
Eurysternini	Eurysternus cyanescens Balthasar, 1939	2,3888	3,1712	2,3220	1,0001	0,7848	1,5544	11,2213
	Eurysternus hirtellus Dalman, 1824	0	0,0112	0,0999	0,0304	0,0325	0,1484	0,3224
	Eurysternus parallelus Castelnau, 1840	0,0517	0	0	0	0,0366	0	0,0883
	Eurysternus sp.	0,4984	0,0631	0,9157	0,3230	0,5090	1,0229	3,3321
Phanaeini	Coprophanaeus saphirinus (Sturm, 1826)	0,9208	0,3068	0,1403	0	0	0	1,3679
	Phanaeus splendidulus (Fabricius, 1781)	0,2144	0	0,5663	0	0	0	0,7807
Onthophagini	Onthophagus sp.	0	0	0,0387	0	0,2392	0,0095	0,2874
Canthonini	Canthon latipes Blanchard, 1845	1,0120	0,0387	0,1610	0,7324	0,2709	0,1720	2,3870
	Canthon rutilans Castelnau, 1840	0	0	0	0	0,1576	0	0,1576
	<i>Canthon</i> sp.	0,0260	0	0	0,0304	0	0	0,0564
	Canthonella sp.	0,0048	0,0122	0,0017	0,0097	0,0231	0,0110	0,0625
	Deltochilum brasiliense (Castelnau, 1840)	2,2935	2,7742	2,8158	1,7601	0	0,8871	10,5307
	Deltochilum furcatum (Castelnau, 1840)	5,2219	1,4636	3,3371	3,4815	0,0619	1,6543	15,2203
	Deltochilum morbillosum Burmeister, 1848	0,3008	0,6112	0,0600	0,8758	0	0,0569	1,9047
	Deltochilum rubripenne (Gory, 1831)	0,9915	0,0623	1,6335	0,8918	0,2138	1,2141	5,0070
	Paracanthon pereirai d'Andretta & Martínez, 1957	0,0102	0,0041	0,0035	0	0,0021	0,0224	0,0423
	Scybalocanthon nigriceps (Harold, 1868)	0,2946	0	0,6727	2,3028	0,3233	1,1687	4,7621
	Sylvicanthon foveiventre (Schmidt, 1920)	0,0331	0,0131	0,0071	0,0724	0,0269	0,0451	0,1977
	Numero de especies	26	23	30	24	25	22	39
	Biomassa total (g)	18,9936	22,7343	25,7210	17,5364	4,3979	10,6689	100,0521