Camila Yumi Mandai

Aplicação de modelos teóricos de distribuição de abundância das espécies na avaliação de efeitos de fragmentação sobre as comunidades de aves da Mata Atlântica

São Paulo

-2010-

Camila Yumi Mandai

Aplicação de modelos teóricos de distribuição de abundância das espécies na avaliação de efeitos de fragmentação sobre as comunidades de aves da Mata Atlântica

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

Orientador: Prof.Dr. Paulo Inácio de Knégt López de Prado

São Paulo

-2010-

Mandai, Camila Yumi
Aplicação de modelos teóricos de distribuição de abundância das espécies na avaliação de efeitos de fragmentação sobre as comunidades de aves da Mata Atlântica
87 páginas

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

1. SAD 2. Série logarítmica 3. α de Fisher 4. Seleção de modelos 5. Fragmentação 6. Simulação 7. Probabilidades de captura
I. Universidade de São Paulo. Instituto de Biociências. Departamento de Ecologia.

Comissão Julgadora:

Prof(a). Dr(a).

Prof(a). Dr(a).

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

Dedicatória

Ao meu pai

Tetsuo Mandai (*in memorian*).

Agradecimentos

Agradeço primeiramente à minha família. Minha mãe Eliza pelo amor incondicional, por ser a pessoa mais compreensiva que conheço, mãe para todas as horas e em todos os sentidos. Ao meu irmão Kleber, por me fazer conhecer os limites da minha limitada paciência. À minha irmã, pela amizade, pelo carinho e pela minha formação. À minha cunhada Cristina por conceber meus desejos alimentícios nas festinhas de criança e por ter me dado como sobrinho o Vinícius, minha constante alegria e o Maurício que apesar de ainda não conhecer, já o amo. Ao meu cunhado Rafael por seu humor artístico e por ter me dado meu sobrinho Théo, que faz todas as minhas idas a Penápolis uma verdadeira descoberta.

Ao meu orientador Paulo Inácio de Knegt López de Prado pela dedicada orientação, pelas preciosas horas de ensino, pelo incentivo, e também, pela seriedade, calma e paciência ilimitadas comigo.

À FAPESP pela bolsa concedida para a realização deste trabalho.

Aos professores Dr. Jean Paul Metzger e Dra. Renata Pardini pela oportunidade de colaboração com o projeto "Conservação da biodiversidade em paisagens fragmentadas".

Ao Alexandre Martensen, pelos dados cedidos.

À Cristina Banks-Leite pelos dados cedidos, pelos conselhos, dicas, idéias, e principalmente pela dedicação e carinho extremos em tudo que faz.

Ao Rafael Pimentel por todo o trabalho e paciência que teve comigo, pela organização dos dados utilizados e pela extrema eficiência.

À professora Dra. Érica Hasui, pela paciência, disponibilidade, estímulo, e claro, pela simpatia incomparável.

Ao professor Dr. Gonçalo Ferraz, por todo o estímulo durante as aulas da disciplina de Ecologia de populações, que me abriram os olhos com suas críticas ácidas à aplicação de abundância relativa sem considerar as probabilidades de captura. Sem seus ensinamentos, este trabalho, literalmente, não seria metade do que é.

À Dalva, secretária da pós-graduação da Ecologia, pelo seu jeito de madrasta, mas coração de mãe.

Aos Drs. Alexandre Adalardo de Oliveira, Pérsio dos Santos e Rogério Rosa pela participação na banca de qualificação e dicas valiosas.

Não poderia jamais deixar de agradecer também ao pessoal do laboratório gigante, que se tornaram verdadeiros companheiros nesta minha jornada. À Camila Zatz por ser a pessoa mais surpreendente do mundo. Ao Muzgo, pela nerdice, por instalar pacientemente o Ubuntu para mim, pela paciência com os meus problemas (e reclamações) de informática e pelo ombro amigo. Ao Billy pelas tiradas " à la" chutes na cara aos fins de tarde, pelas discussões frenéticas e produtivas, e por me emprestar o computador (que aliás, estou usando para escrever os agradecimentos), meu eternos agradecimentos a você Billy Requena. À Marie(enta) Claire Chelini por ser uma criaturinha rabugenta, grossa, ranzinza e contraditoriamente meiga, que alegra e perfuma os dias no laboratório. À Rachel Werneck, pela vocação em lidar comigo, pelo companheirismo e coração de mãe, pelos bolos, brigadeiros, risotos que alegravam o meu dia e principalmente pelo melhor pão de queijo do mundo. À Flávia Marquitti pela energia que faz vibrar o ar, as cadeiras, as mesas, a sala, enfim, o universo. Ao Mathias por sua habilidade humorística, pelo suprimento de vídeos do youtube, e pela sua sensualidade feminina. À

Paulinha Lemos pelo sorriso maroto e pelo jeito direto. À Esther e ao Jean, pela simpatia, companhia, pelos treinos em espanhol e pela prometida paella. Ao Ernesto Schimdt, pelo entusiasmo e sorriso fácil, e claro, pelo chimarrão. Ao Roberto Mungía Steyer pelas discussões, disponibilidade e valiosas referências. Ao Paulo Guimarães Júnior, por toda a sua cultura útil, e inútil (nunca me esquecerei onde fica bairro do Limoeiro). Ao Glauco Machado, não pelo meu apelido, nem pelas cortadas homéricas, mas pela sua figura ilustre e nunca despercebida.

Ao pessoal do laboratório nave-mãe, o GET. À Camila Castanho por todas as conversas e conselhos sensatos, desabafos e terapias. À Marina Rizzo, por combinar humildade e conhecimentos matemáticos em uma só pessoa e por isso ser a esperança para as soluções dos meus problemas. À Paula Condé pelo humor indestrutível e misturas mágicas (e duvidosas) de chás. Ao Renato Lima pelo toque de classe e requinte que dá ao laboratório, e é claro, pela sua abertura constante para discussões. E finalmente ao Rodolpho Rodrigues pela amizade, companheirismo, discussões, conversas, café, humildade, pelo jeito fácil, pelas milhares de caronas, pela paciência, enfim, por toda a nossa convivência nesses dois anos.

À todos os amigos que eu fiz no Curso de Campo de Ecologia da Mata Atlântica e no IB, alunos e professores que compartilharam risadas, alegrias e muitos traumas comigo. Em especial, quero agradecer à Lucia Munari, amiga mais que compreensiva e de todas as horas, à Julia Stuart e Talita Zupo, pela aflição conjunta, à Flávia Borbo e Julia, pelas agradáveis idas ao bandejão, à Paula Valdujo, pela sensatez e nerdice com orgulho, Andrea Sendoda, a japa-chinesa louca , Denise Araújo Alves pela ajuda, companheirismo e horas de relax nesta última etapa, à Nicolle Sidney, pela amizade mais que fácil e intimidade mais que desnecessária, e é claro por ter me feito conhecer o Sanhudo, à Elaine Bornancin pelas conversas diárias sobre o tudo e o nada, à Eline pelo acompanhamento nos dias de sol e chuva trabalhando, à Laura Prado por salvar a minha vida me ajudando com o sumário, e tornar essas últimas horas de pré-entrega muito mais divertidas. São tantas pessoas que não tenho como colocar aqui, porém são pessoas que fizeram da minha vida no IB e no mestrado algo mais que alegre.

Aos meus amigos de Ribeirão Preto. Lucas Albrechet, amigo contraditório, companheiro de todas as horas e para toda a vida. À Marcia Pannuti, amiga mais que improvável, agradeço pelas viagens maravilhosas à Ilha e por ser uma irmã de coração. Às minhas eternas irmãs de república, Bianca Amaral e Katherine Foehringer, pela amizade por toda a vida, e por todo o carinho que eu aprendi a cultivar por vocês. Ao José Luíz Liberato e Lívia Godoy que estão aqui comigo nos últimos momentos, me acalmando e fazendo uma companhia mais que necessária. À Mayumi Yamada, Claire Jezéquèl, e Juliana Abumansur por vocês serem incríveis, lindas, mais que legais e muito mais insanas que eu!

Ao pessoal da escalada, minha terapia: Tiago Carrijo, Simone Policena, Paola, Tacio Philip, por ter me levado pela primeira vez para a rocha, Paulinha, Rodrigo Caderno e Nicolle.

Agradeço finalmente ao Carlos Eduardo Sanhudo, pelos momentos juntos, pelo companheirismo, por sua visão curiosa de mundo, e por me fazer uma pessoa mais feliz.

Neste momento, eu provavelmente me esqueci de muitas pessoas, pessoas não menos importantes do que aquelas que eu nomeei aqui. Sei que muitos nomes ainda me virão à mente, então meu muito obrigada.

Sumário

RESUMO 3
ABSTRACT 4
Apresentação 5
Caracterização geral do padrão de abundância de espécies e considerações teóricas 5
Parâmetro escalar (Probabilidade de captura) 9
Aplicação dos modelos de abundância 10
Ajuste e Seleção de modelos por máxima verossimilhança 11
A proposta desta tese 12
Bibliografia 13
Capítulo 1 – Efeitos da heterogeneidade interespecífica de probabilidade de captura sobre o viés de seleção e estimação dos modelos de distribuição de abundância das espécies 16
1. Introdução 16
2. Material e Métodos 18
2.1. Modelos de simulação 18
2.2. Simulações 20
2.3. Análises 22
3.Resultados 24
3.1. Viés de seleção de modelos 24
3.2.Vieses nas Estimativas dos parâmetros 24
4. Discussão 30
5. Bibliografia 34
Capítulo 2 - Aplicação dos modelos teóricos de distribuição de abundância das espécies na avaliação de efeitos ambientais quando a premissa dos modelos de igualdade de probabilidades de captura entre as espécies é desrespeitada 37
1. Introdução 37
2. Material e Método 38
2.1. Modelos de Simulação 39
2.2. Cenários de distribuição das componentes probabilísticas 42
2.3. Simulações 43
3. Resultados 46
3.1. Viés de seleção 46
3.2. Viés de estimação 46
4. Discussão 52
Capítulo 3 – Aplicação do modelo de distribuição de abundância série logarítmica na avaliação de efeitos de fragmentação sobre a comunidade de aves em paisagens fragmentadas de remanescentes da Mata Atlântica 58
1. Introdução 58

2. Material e Métodos 61
2.1 Áreas de coleta 61
2.2 Coleta de dados 63
2.3 Variável dependente: α de Fisher 63
2.4 Variáveis independentes: Área, Proximidade 65
2.5 Análise de dados 66
2.6. Variação nas abundâncias das espécies nos fragmentos e areas controle 67
3. Resultados 67
4. Discussão 78
5. Bibliografia 81
Considerações Finais 85
Bibliografia 87

RESUMO

As distribuições de abundância relativa das espécies tiveram um papel importante no desenvolvimento da ecologia de comunidades, revelando um dos padrões mais bem estabelecidos da ecologia, que é a alta dominância de algumas espécies nas comunidades biológicas. Este padrão provocou a criação de dezenas de modelos teóricos na tentativa de explicar quais mecanismos ecológicos poderiam gerá-lo. Os modelos teóricos de abundância relativa das espécies podem ser vistos como descritores das comunidades, e seus parâmetros, medidas sintéticas de dimensões da diversidade. Esses parâmetros podem ser utilizados não só como descritores biologicamente interpretáveis das comunidades, mas também como variáveis respostas a possíveis fatores ambientais que afetam as comunidades. Adotando então esta aplicação descritiva dos modelos, nosso objetivo foi comparar as comunidades de aves de áreas em um gradiente de fragmentação, utilizando como variável resposta os valores estimados do parâmetro do modelo série logarítmica, o α de Fisher. Como todos os modelos teóricos de abundância relativa propostos têm como premissa, a igualdade de probabilidade de captura entre as espécies, o que para comunidades de espécies de organismos móveis, como aves, parece pouco realista, neste trabalho investigamos também o grau de sensibilidade dos modelos quanto à quebra dessa premissa. Assim, por meio de simulações de comunidades, analisamos o viés de seleção e estimação, e revelamos que o aumento do grau de heterogeneidade entre as probabilidades de captura das espécies acarreta no incremento do viés de seleção do modelo real e também de estimação dos parâmetros. Porém, como o objetivo do estudo era identificar os fatores que influenciam a diversidade das comunidades, mesmo com o viés de estimação, talvez ainda fosse possível revelar o grau de influência sobre os valores dos parâmetros, quando ele existir. Assim, prosseguimos com mais uma etapa de simulações, em que geramos comunidades cujos valores de parâmetros tinham uma relação linear com a área dos fragmentos. O que encontramos é que independente da igualdade ou desigualdade de capturabilidade das espécies, quando o efeito existe, ele é sempre detectado, porém dependendo do grau de diferença de probabilidade de captura das espécies, o efeito pode ser subestimado. E, na ausência de efeito, ele pode ser falsamente detectado, dependendo do grau de heterogeneidade de probabilidades de captura entre as espécies, mas sempre com estimativas bem baixas para o efeito inexistente. Com esses resultados então, pudemos quantificar os tipos de efeitos da heterogeneidade de probabilidades de captura e prosseguir com as análises dos efeitos de fragmentação. O que nossos resultados mostraram é que na paisagem com 10% de cobertura vegetal, a área parece influenciar a diversidade dos fragmentos mais que o isolamento, e que na paisagem de 50% de cobertura vegetal, a variável de isolamento se torna mais importante que a área para explicar os dados. Porém, em uma interpretação mais parcimoniosa, consideramos as estimativas dos efeitos muito baixas para considerar que ele de fato existia. Com isso, concluímos que o processo de fragmentação provavelmente não tem efeito sobre a hierarquia de abundância das espécies, e é independente da porcentagem de cobertura vegetal da paisagem. Contudo, em uma descrição do número de capturas de cada espécie nos fragmentos, ponderada pelo número de capturas amostrado em áreas contínuas adjacentes, revelaram que o tamanho do fragmento pode ser importante na determinação de quais espécies serão extintas ou beneficiadas e que talvez a qualidade da matriz seja decisiva para a manutenção de espécies altamente sensíveis em fragmentos pequenos. Assim, demonstramos que, embora as SADs sejam pouco afetadas pela fragmentação, a posição das espécies na hierarquia de abundâncias pode mudar muito, o que reflete as diferenças de sensibilidade das espécies a área e isolamento dos fragmentos.

Palavras-chave: SAD, série logarítmica, α de Fisher, seleção de modelos, fragmentação, simulação, probabilidades de captura.

ABSTRACT

Species abundance distribution (SADs) had an important role in community ecology, revealing one of the most well established pattern in ecology, which is the high dominance by just a few species. This pattern stimulated the proposal of innumerous theoretical models in an attempt to explain the ecological mechanism which could generate it. However these models can also be a descriptor of the communities and their parameters synthetic measures of diversity. Such parameters can be used as response variables to environmental impact affecting communities. Adopting this approach our objective was to compare bird communities through areas of different levels of fragmentation, using as response variable the estimates of α, the parameter of Fisher's logseries. Considering the implicit assumption of equal capture probabilities among species in SAD models we also investigated the degree of sensibility of the models when this assumption is disrespected, once it seems so unrealistic. Thus simulating communities in which species had equal and different capture probabilities among them we found that increases in the degrees of heterogeneity in species catchability lead to a gain in biases on the model selection and parameters estimations. Additionally, since our goal in this study was identify some factors that may influence the diversity in communities, even with the biases, if they were constant, maybe it was still possible to test the relation. In this context we proceed to another stage of simulations, where we generate communities whose parameter values had a linear relationship with remnant area. What we find is that regardless of equal or unequal in catchability of species, when the effect exists, it is always detected, but depending on the degree of difference in probability of catching the species, the effect may be underestimated. Further, in the absence of effect, it can be falsely detected, depending on the degree of heterogeneity of capture probabilities among species, but always with very low estimates for the effect non-existent. With these results, we could quantify the types of effects of heterogeneity on capture probabilities and proceed with the analysis of the effects of fragmentation. What we showed is that the landscape with 10% vegetation cover, the fragment area appears to influence the diversity of the fragments rather than isolation, and landscape in 50% of plant cover, the isolation variable becomes more important than area to explain the data. But in a more parsimonious interpretation, we consider the estimated of the effects too low to consider that they actually exist. Therefore, we conclude that the fragmentation process probably has no effect on the hierarchy of species abundance. However, in a description of the number of captures of each species in the fragments, weighted by the number of catches sampled in continuous adjacent areas revealed that the fragment size may be important in determining which species will be extinct or benefit and that perhaps the quality of matrix is decisive for the maintenance of highly sensitive species in small fragments. Thus, we demonstrated that while the SAD are not significantly affected by fragmentation, the position in the hierarchy of species abundances can change a lot, which reflects the different sensitivity of species to area and isolation in the fragments.

Key-words: SAD, logseries, Fisher's α, model selection, fragmentation, simulation, capture probabilities.

Apresentação

Caracterização geral do padrão de abundância de espécies e considerações teóricas

Em qualquer ambiente, seja terrestre, aquático, temperado ou tropical, as abundâncias das espécies não se distribuem de forma uniforme, ao contrário, há um padrão de alta dominância por parte de uma minoria de espécies (Magurran, 2003). A recorrência deste padrão em diversos táxons, ambientes e escalas tornou-o uma das leis universais da ecologia de comunidades, consagrada pela forma côncava da curva de distribuição da abundância das espécies, conhecida na literatura por "hollow curve"(McGill et al. 2007). O padrão, então, assumiu um papel importante na biologia como principio empírico que sustenta várias teorias ecológicas (Hubbel 2001, McGill 2003). Desde os anos 30, vários modelos teóricos de distribuição de probabilidades foram propostos para descrever e explicar a forma das curvas côncavas das distribuições de abundância das espécies (SAD[1] ;McGill et al. 2007). Os modelos são classificados de maneira geral em duas categorias: os estatísticos, que ajustam uma distribuição estatística teórica *a priori*, normalmente com propriedades bem conhecidas e de fácil tratamento matemático; e os biológicos, que deduzem distribuições teóricas a partir de cenários ecológicos, mas que podem trazer dificuldades analíticas (Magurran, 2003).

Para alguns dos modelos estatísticos, propostos apenas para descrever e se ajustarem às curvas de abundância relativa, foi demonstrado que ser gerados por mecanismos biológicos simples (Bulmer 1974, May 1975, Caswell 1976). Esses modelos, juntamente com os modelos originalmente biológicos, são baseados em desdobramentos das "assembly rules", ou, "regras de montagem" (Hubbell 2001), cujo conceito fundamental é a competição pelo nicho. O primeiro modelo biológico a ser proposto foi em 1932, por Motomura, cuja lógica de partição sequencial do nicho serviu como base para o modelo *broken-stick* ("vara quebrada") de MacArthur (1957), que teve um papel importante na maneira como os ecólogos passaram a pensar sobre a diversidade ecológica das comunidades e o papel de modelos na ecologia (Kingsland 1995).Após do modelo *broken-stick,* e a partir dos anos 70, houve uma proliferação de propostas de modelos teóricos de SAD (McGill et al. 2007)

Efeito da amostragem na representação do padrão

Os modelos teóricos de abundância relativa preveem uma certa distribuição para as comunidades, mas muitas vezes não incorporam a maneira como a amostragem pode

[1] Utilizaremos essa sigla para designar as distribuições de abundâncias das espécies em uma comunidade, ou amostra dela tomada, em correspondência a SAD (species abundance distributions), que é o termo tradicionalmente utilizado na literatura.

influenciar a representação dessa distribuição. Assim, há dois tipos distintos de construção dos modelos de SADs, que Dewdney (1998) classifica como: a proposta de uma distribuição teórica particular para representar todos os padrões de abundância, e testá-la através de análises das observações disponíveis, ou um modelo que leve em consideração a relação matemática entre a SAD *g* numa comunidade e a distribuição *f* que se revela na amostra tomada desta comunidade. No segundo tipo de abordagem, as distribuições tornam-se equações compostas, que descrevem a probabilidade das espécies serem representadas na amostra por um certo número de indivíduos, dada sua abundância na comunidade. A inclusão do procedimento de amostragem na concepção dos modelos é dada então pela incorporação de uma teoria da amostragem. A teoria da amostragem é tradicionalmente aplicada na engenharia da comunicação e na teoria da informação, mas recentemente vem sendo utilizada em várias disciplinas, e pode ser aplicada sempre que as funções que descrevem um sistema precisam ser reconstruídas a partir de uma amostra (Zayed 1993).

Os modelos compostos foram primeiramente introduzidos na análise de distribuições de abundância quando Fisher (1943) propôs a série logarítmica. A partir dos dados de Corbet & Williams, ele assumiu que a comunidade fosse uma distribuição gama, que é uma distribuição bastante flexível. Sob a premissa que os indivíduos são tomados ao acaso e de maneira independente nas amostras, Fisher usou a distribuição Poisson para descrever o processo de amostragem. O modelo proposto, então, é uma distribuição de Poisson, em que o parâmetro lambda (número esperado de capturas de uma espécie) não é fixo, mas uma variável aleatória com distribuição Gama. A distribuição composta resultante é a distribuição binomial negativa. Como o processo amostral não permite conhecer a classe de abundância zero, que seriam as espécies não representadas na amostra, Fisher truncou a distribuição, e ainda assumiu que o número de espécies na comunidade fosse infinito, fazendo com que o parâmetro de forma da binomial negativa tendesse a zero (Hubbell 2001), resultando em um modelo com apenas um parâmetro, que é a série logarítmica. O parâmetro da série logarítmica (α), por sua vez, é utilizado como um índice de diversidade, e teoricamente, é independente do tamanho da amostra.

A distribuição série logarítmica se caracteriza por prever uma alta dominância, sendo a classe modal a de abundância um. O modelo série logarítmica parece se ajustar melhor em comunidades com baixa riqueza ou naquelas em que a amostra é muito pequena. Tradicionalmente, aceita-se que o aumento da amostra piora o ajuste a série logarítmica, e que uma distribuição lognormal passaria a ser um modelo mais apropriado para descrever as abundâncias (Krebs 1994, Magurran 2005). A lognormal, por sua vez, foi considerada pela

primeira vez para descrever abundâncias de espécies na comunidade em 1948, quando Preston plotou em um histograma as abundâncias relativas em escala logarítmica na base 2, que ele denominou "oitavas", e encontrou um padrão próximo ao da curva normal. Na sua proposta, Preston destacou que as populações tenderiam a crescer geometricamente e não aritmeticamente. Além disso, ele propôs que as curvas das amostras tomadas de uma comunidade apresentavam como classe modal as de abundância baixa, e que a série logarítmica era um artefato da amostragem, em que apenas uma pequena porção da comunidade era representada, e por isso a distribuição da amostra aparecia de forma parcial, ou truncada. Assim, as espécies raras na amostra não seriam necessariamente as mais raras na comunidade, e sim as de abundância modal, sendo a amostra uma representação incompleta das espécies da comunidade. Desta forma, a verdadeira distribuição só poderia ser revelada com o aumento do tamanho amostral.

Para incorporar o efeito da amostragem no modelo, assim como Fisher *et al.* (1943), Preston presumiu uma distribuição Poisson para representar o processo amostral, mas propos que uma boa aproximação seria a própria curva lognormal cortada na altura do que ele chamou de "linha de véu" (*veil line*), que seria a classe de abundância abaixo da qual as espécies não são incluídas na amostra. Essa linha seria deslocada com o aumento da porção da comunidade representada na amostra, ou seja, do tamanho amostral. A lognormal truncada desta maneira prevê uma comunidade com abundâncias mais homogênea que a série logarítmica. Um dos motivos pelos quais Preston rejeitou a série logarítmica de Fisher, e propôs a lognormal truncada, foi o melhor ajuste do seu modelo aos dados empíricos utilizados por ele.

Em 1969, Pielou generalizou e formalizou a aplicação da teoria da amostragem nos modelos de distribuição de abundância. Ela considerou que a variação de abundância de uma determinada espécie na amostra poderia ser representada também pela distribuição de Poisson, cuja média ou esperança, que corresponde ao único parâmetro da Poisson (λ) representaria as S observações independentes, correspondentes às espécies amostradas, de uma distribuição de probabilidades contínua com função de densidade $f(\lambda)$. A função $f(\lambda)$ da comunidade seria a escolhida *a posteriori* e não mais fixa como fez Fisher ao escolher a gama como representativa das abundâncias das espécies na comunidade.

Apesar de Preston ter ressaltado em seu trabalho que a *veil line* era só uma ferramenta de aproximação e não serve para cálculos exatos das estimativas das freqüências esperadas, Dewdney (1998) critica veemente a adoção da prática de truncagem nos modelos, e prefere a inclusão da teoria da amostragem tradicional. Segundo ele, a amostragem Poisson sempre

mantém a forma da distribuição de abundâncias da comunidade, o que não estaria de acordo com a *veil line* que pode chegar a cortar a curva ao meio, mudando radicalmente sua forma. Dewdney, assim como Pielou, explicita a relação entre a distribuição da amostra e da comunidade. Ele nota que o modelo correto para o processo amostral, caso ocorra sem reposição em uma comunidade finita, é a função hipergeométrica, mas que por ser de difícil tratamento pode ser aproximada pela função de Poisson, desde que a fração amostrada da comunidade seja pequena. A esse modelo de equações compostas, ele deu o nome de "transformação de Pielou". A partir da abordagem mais geral de Pielou, alguns pesquisadores incluíram a teoria da amostragem em seus modelos de SAD. Alguns mantiveram a Poisson para descrever a amostragem e variaram a distribuição f(λ) da comunidade, como a distribuição Poisson - lognormal de Bulmer (1975) e também a de Dewdney (1998) ao testar a lognormal misturada à Poisson e também à hipergeométrica.

Recentemente, a incorporação da teoria da amostragem nas funções de SAD tomou um lugar de destaque por estar presente na teoria neutra de Hubbel (2001), que fez ressurgir o interesse pelos modelos teóricos de SAD (e.g., Volkov *et al* 2003; Alonso & McKane 2004; Ettiene & Olff 2004, Etiene 2005, Etiene & Alonso 2005). Os trabalhos com teoria neutra utilizaram a teoria da amostragem para representar a heterogeneidade de amostragem, e assim modelar a probabilidade de uma espécie com limitação de dispersão e uma certa abundância estar na amostra dentro de uma comunidade neutra (Green & Plotkin 2007). Neste caso, o tipo de amostragem tanto foi a Poisson como a hipergeométrica. Desta forma, apesar de neste período, ter havido um desenvolvimento na aplicação da teoria da amostragem, os avanços só foram alcançados considerando comunidades neutras, excluindo as propriedades da amostra na presença de forças não-neutras (Green & Plotkin 2007). Em formulação mais recente, Green & Plotkin (2007) apresentaram a teoria aplicada às SADs em um contexto mais geral, e deixaram livre não só a função da distribuição de probabilidades da comunidade, mas também a da distribuição da amostra. A fórmula geral apresentada por eles, para se obter a distribuição de abundância amostrada $\Phi a(\gamma)$, foi:

$$\Phi a\gamma = \int_0^\infty \psi a(\gamma \mid n)\Phi(n)\mathrm{dn} \qquad (1)$$

Nesta equação, $\psi a(y|n)$ é uma função de probabilidade condicional, que descreve a probabilidade de que uma espécie da comunidade tenha abundância y na amostra, dada sua abundância na comunidade, n. A função de distribuição de abundâncias na comunidade, $\Phi(n)$,

representa a probabilidade de que uma espécie tenha abundância *n* na comunidade. O produto destas duas funções dá a probabilidade dos dois eventos ocorrerem, isto é, que a abundância na amostra seja *y*, dado que na comunidade seja *n*. Integrando-se este produto em *n,* e dividindo-o por uma constante de integração (não representada na fórmula acima), temos a probabilidade de um dado valor de abundância na amostra.

Além da fórmula geral, eles compilaram as soluções das equações compostas de Poisson com alguns modelos tradicionais de SADs (na maioria já deduzida por outros autores), que foram: série logarítmica (versão contínua), gama, exponencial e hiperbólica truncada. Na formulação destes autores, há um novo parâmetro, que seria a relação de escala entre a amostra e a comunidade, que pode ser representado pela fração de área ou de indivíduos da comunidade que foi incluída na amostra. Assim, ao se fazer uma amostra aleatória de uma comunidade, a distribuição de abundâncias resultante é reescalonada por esta proporção amostrada. No caso de animais móveis, esta fração corresponde à probabilidade de captura das espécies (Royle & Nichols 2003). Considerando, então, a abundância de uma espécie N_i e a probabilidade de captura *a,* o número esperado de capturas desta espécie na amostra é $E[x]=N_i \times a$. A equação composta geral da Poisson proposta por Green & Plotkin (2007) é:

$$\Phi \mathrm{a} y = \int_0^{\infty} e^{-\mathrm{an}} \frac{\mathrm{an}^y}{\mathrm{y!}} \Phi n \mathrm{dn} \qquad (2)$$

Os autores também propuseram soluções mais complexas, em que a amostragem continuasse ao acaso, mas as capturas correlacionadas, dependentes da distribuição dos indivíduos. Eles incluíram, então, o efeito de agregação na amostragem e o fizeram através de equações compostas com os mesmos modelos de distribuição citados acima para a comunidade, mas utilizando a binomial negativa como a função de amostragem, em que o parâmetro de dispersão (*k*) variasse de acordo com a agregação espacial dos indivíduos.

Parâmetro escalar (Probabilidade de captura)

Em resumo, para corrigir o efeito da amostragem e a representação imperfeita das abundâncias, os modelos de SAD passaram a incorporar um parâmetro escalar, representando a fração da comunidade que foi amostrada. Entretanto, a correção é composta de um único valor para todas as espécies, o que implica que a fração que a amostragem é capaz de representar de cada espécie é sempre a mesma. Assim, existe uma premissa implícita de que os indivíduos são igualmente capturáveis, independente da espécie, local, horário ou qualquer outra variação entre as sessões de capturas.

Entretanto, principalmente na ecologia de populações, em que é comum utilizar estimativas de abundâncias, como medidas para caracterizar as diferentes populações da mesma espécie em locais distintos, ou, de espécies diferentes na mesma área, é sabido como as diferenças nas probabilidades de captura de cada espécie podem alterar os valores estimados (Williams et al 2002). A relevância de se estimar a probabilidade de captura de cada população está relacionada a fatores da espécie ou do local que podem alterar a capturabilidade dos indivíduos e por isso tornar a contagem realizada de indivíduos resultado não só da abundância da espécie, mas também do quão visível ou capturável é a mesma (Nichols et al 1986).

Assim, uma vez que a capturabilidade distinta entre as espécies pode fazer com que o número de indivíduos capturados de cada espécie dependa de outros fatores além da abundância (Remsen & Good 1996), ela pode modificar as relações verdadeiras de abundância das espécies nas curvas de distribuições de abundância relativa, quando representadas por dados de contagem, na amostra. Entretanto, devido a dificuldades de se estimar as abundâncias na escala da comunidade (Mackenzie et al 2006), a magnitude do efeito da capturabilidade diferenciada entre as espécies no contexto dos estudos do padrão de abundância relativa nas comunidades é pouco explorado. O desconhecimento sobre os efeitos das probabilidades de captura heterogênea nos estudos de distribuição de abundância relativa tornam o uso indiscriminado de dados de contagem nessa escala pouco confiável, uma vez que ao não manterem as relações das abundâncias das espécies da comunidade, os dados podem levar vieses consideráveis nos resultados.

Aplicação dos modelos de abundância

Apesar do aumento de propostas de modelos teóricos para descrever o padrão de alta dominância nas comunidades, nenhum, até hoje, pareceu ser superior aos outros, uma vez que mesmo que o modelo falhe para um conjunto particular de dados, ele não pode ser descartado (Mcgill *et al.,* 2007). A impossibilidade de invalidar modelos é devido, principalmente, à deficiência de testes para analisar os processos em uma abordagem estatística (Alonso *et al* 2006), uma vez que os métodos estatísticos tradicionalmente utilizados são inadequados para o ajuste dos dados (McGill, 2003). Os tradicionais testes de aderência, como o teste de χ^2 e Kolmogorov – Smirnov, ao comparar os dados observados e os esperados para diferentes modelos, colocam o modelo de distribuição testado como hipótese nula. Assim, a aceitação do modelo ocorre quando não há desvios significativos e o objetivo do teste passa a ser a não rejeição da hipótese nula. Além disso, não é possível fazer comparações simultâneas de modelos, o que torna a aplicação desses testes de aderência inconclusivos, pois os dados

podem se ajustar a mais de um modelo, já que estes são testados individualmente. Uma solução promissora para este problema metodológico é o uso do princípio de verossimilhança estatística (Edwards 1972, Royall 2000) para criar protocolos de comparação simultânea de várias hipóteses concorrentes. Um dos mais simples é a seleção de modelos baseada em índices de informação (Burnham & Anderson, 2002), que ordena um conjunto de modelos de acordo com sua plausibilidade frente aos dados, levando em conta sua parcimônia. Na escala de McGill (2003), essa seria a categoria mais rigorosa de testes de ajustes dos dados a modelos teóricos.

No caso das SADs, os modelos concorrentes podem ser diferentes distribuições teóricas de probabilidade, ou a mesma distribuição, mas com parâmetros diferentes. Com isso é possível, por exemplo, avaliar se um gradiente ambiental afeta o tipo de distribuição de abundância, ou apenas a sua forma, através do uso dos valores dos parâmetros dos modelos. Os parâmetros, portanto, passam a ser utilizados como indicadores de atributos da comunidade e são considerados ferramentas poderosas de avaliação e comparação entre áreas (Pielou 1977).

Ajuste e Seleção de modelos por máxima verossimilhança

A função de log-verossimilhança negativa de um conjunto de dados é definida como a soma do logaritmo das probabilidades atribuídas pela equação de densidade probabilística de um modelo para cada valor da amostra:

$$\mathrm{LL}(\theta \mid \mathrm{y}=\mathrm{y}_i) = -\sum_{\mathrm{i}=1}^{n} \log f(\theta \mid \mathrm{y}=\mathrm{y}_i) \qquad (3)$$

(Edwards 1972), em que y_i é a observação, y a variável aleatória que descreve os valores que y_i pode assumir e que tem função de densidade probabilística $f(y|\theta)$, com um conjunto de parâmetros θ. A função $LL(\theta|y=y_i)$ expressa a plausibilidade de se ter uma certa observação, considerando as probabilidades previstas pelo modelo. Como o valor de verossimilhança é uma medida relativa de plausibilidade em relação a modelos concorrentes, na seleção de modelos com o AIC é necessário escolher mais que um modelo candidato, com diferentes funções de densidade probabilística, e deduzir suas funções de verossimilhança. No caso dos modelos de abundância relativa, as distribuições a serem inseridas nas funções de verossimilhança são truncadas na classe zero, uma vez que na amostra as espécies representadas por abundância zero são desconhecidas

Escolhidos os modelos concorrentes, segue-se o seguinte protocolo de ajuste de mode-

los por máxima verossimilhança:

1 – Usa-se as funções de densidade probabilística de cada modelo para aplicar as funções de log-verossimilhança (equação 3) e calcular as probabilidades atribuídas para cada conjunto de dados, dado certos valores atribuídos de parâmetro. Posteriormente, por meio das rotinas de otimização computacional buscam-se os valores de parâmetro que minimizem a função de log-verossimilhança.

2 – Com a minimização da função de log-verossimilhança da amostra, obtém-se os valores de máxima verossimilhança do modelo e as estimativas de máxima verossimilhança dos parâmetros, que são os valores dos parâmetros do modelo no ponto em que a função de log-verossimilhança atinge seu valor mínimo.

3- Para a seleção dos modelos, um dos critérios mais utilizados atualmente é o critério de informação de Akaike - AIC (Akaike 1973) que é uma medida de estimativa de distância relativa esperada de cada modelo, ao modelo verdadeiro. O AIC pode ser enunciado também como uma medida de plausibilidade do modelo penalizado pelo número de parâmetros do mesmo, pois é uma função da log-verossimilhança e do número de parâmetros do modelo. O cálculo de AIC é o dobro do valor de verossimilhança negativa do modelo somado ao dobro do número de parâmetros do mesmo.

4- O modelo selecionado é aquele com menor valor de AIC. Entretanto, quando a diferença entre AIC dos modelos concorrentes for muito pequeno, dependendo do critério, considera-se empate, e ambos os modelos são selecionados.

A proposta desta tese

Embora haja atualmente muitos modelos teóricos de SAD ainda não há muitos testes empíricos comparando-os sistematicamente, seja em situações ecológicas diversas ou então para diferentes grupos taxonômicos com numerosos conjuntos de dados (McGill, 2003). McGill et al. (2007) afirmam que analisar pequenas variações na natureza da "hollow curve" nunca levará a inferências fortes e testes decisivos para as teorias de distribuição de abundância relativa. A ideia apresentada aqui é que considerando que os modelos de SAD são acima de tudo ferramentas descritivas das relações de abundância das, ele pode ser escolhido por critérios práticos, desde que se ajuste a um amplo conjunto de dados, e cujo os parâmetros podem ser utilizados como medidas de comparação entre dados de comunidades diferentes (Pielou 1977).

Adotando então esta aplicação descritiva dos modelos, nosso objetivo foi avaliar os efeitos da fragmentação sobre a diversidade de comunidades de aves, utilizando como variá-

vel resposta os valores estimados dos parâmetros de um modelo teórico de SAD. Além do ajuste, o primeiro critério de escolha do modelo foi que ele fosse um modelo de distribuição de distribuições de abundancias na amostra, i.e. um modelo composto. Como ainda não se sabe qual o efeito da probabilidade de captura diferenciada entre as espécies, as análises foram precedidas por duas etapas de avaliação da robustez dos modelos de abundância relativa em relação à premissa de igualdade de probabilidades de captura.

A dissertação é composta de três capítulos e uma conclusão geral. O capítulo de número um investiga o viés de seleção e estimação quando a probabilidade de captura é diferente entre as espécies. O capítulo dois avança nesta investigação, avaliando a robustez dos parâmetros dos modelos como indicativo da sensibilidade da comunidade a uma variável de fragmentação, quando há a quebra da premissa de igualdade de capturabilidade entre as espécies. Por fim, o capítulo três discorre sobre a descrição, desenvolvimento, resultados e discussão do objetivo principal do projeto de mestrado, relativo ao efeito da fragmentação sobre as SADs das comunidades. Finalmente, a conclusão geral relatará brevemente sobre como cada etapa do projeto foi pensada e como todas elas foram essenciais para que pudéssemos atingir os objetivos propostos aqui.

Bibliografia

AKAIKE, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. In: B. N. PETROV & F. CSAKI, eds. Second International Symposium on Information Theory. Budapest: Akademiai Kiado, pp. 267−281.

ALONSO, D., ETIENNE, R. S. & MCKANE, A. J. 2006. The merits of neutral theory. Trends in Ecology & Evolution. 21: 451 − 457.

ALONSO, D. & MCKANE, A.J. 2004. Sampling Hubbell's neutral theory of biodiversity. Ecology Letters. 7: 901-910.

BULMER, M.G. 1975. On fitting the poisson lognormal distribution to species-abundance data. Biometrics 30: 101-110.

BURNHAM, K. P., AND D. R. ANDERSON. 2002. Model selection and multimodel inference: a practical information-theoretic approach. 2nd Edition. Springer-Verlag, New York, New York, USA. 488 pp.

DEWDNEY, A.K. 1998. A general theory of sampling process with applications to the "veil line". Theoretical Population Biology. 54: 294- 302.

EDWARDS, A. W. F. 1972. Likelihood − An Account of the Statistical Concept of Likelihood and its Application to Scientific Inference. New York, Cambridge University Press.

ETIENNE, R. S. 2005. A new sampling formula fro neutral biodiversity. Ecology Letters. 8: 253-260.

ETIENNE, R.S. & OLFF, H. 2004. A novel genealogical approach to neutral biodiversity theory. Ecology Letters. 7: 170-175.

ETIENNE, R. S. & ALONSO, D. 2005. A dispersal-limited sampling theory for species and alleles. Ecology Letters. 8: 1147- 1156.

FISHER, R. A., CORBET, A. S. & WILLIAMS, C.B. 1943. The relation betwen the number of species and the number of induviduals in a random sample of an animal population. The Journal of Animal Ecology. 12: 42-58.

GREEN, J. & PLOTKIN, J. 2007. A statistical theory for sampling species abundance. Ecology Letters., 10: 1037-1045.

HUBBELL, S.P. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton, Princeton University Press.

KINGSLAND, S. E. 1995. Modeling nature: episodes in the history of population ecology. Chicago, University of Chicago Press.

KREBS, C.J. 1994. Ecology: The Experimental Analysis of Distribution and Abundance, Harper Collins, New York.

MACARTHUR, R. H. 1957. On the relative abundance of bird species. *Proceedings of National Academy of Science*, 43, 293-295.

MACKENZIE, D.I., NICHOLS, J. D., ROYLE, J.A., POLLOCK, K.H., BAILEY, L.L. & HINES, J.E. 2006. Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence. Academic Press.

MAGURRAN, A.E. 2004. Measuring Biological Diversity . *Blackwell Science, Oxford* .

MAGURRAN, A.E. 2005. Species abundance distribution: Pattern or Process ? Functional Ecology. 19 : 177-181.

MAY, R. M. 1975. Patterns of Species Abundance and Diversity. *In:* M. L. CODY & J. M. DIAMOND (eds.), (pp. 81-120). Cambridge, Harvard University Press.

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 (in press).

MOTOMURA, I. 1932. A statistical treatment of associations, Jpn. J. Zool. 44: 379–383.

NICHOLS, J. D., R. E. TOMLINSON AND G. WAGGERMAN. 1986. Estimating nest detection probabilities for White-winged Dove nest transects in Tamaulipas, Mexico. Auk 103: 825-828.

PIELOU, E. C. 1977. *Mathematical Ecology*. New York: John Wiley and Sons.

PRESTON, F.W. 1948.The commonness, and the rarity, of species. Ecology. 29: 254 − 283.

REMSEN, J. V. & GOOD, D.A. 1996. Misuse of data from mist-net captures to assess relative abundance in bird populations. The awk. 113(2):381-398.

ROYALL, R. M. 2000. Statistical Evidence: A Likelihood Paradigm. London, Chapman and Hall.

ROYLE, J. A. AND J. D. NICHOLS. 2003. Estimating abundance from repeated presence-absence data or point counts. Ecology 84: 777 – 790.

VOLKOV, I. BANAVAR, J.R., HUBBEL, S.P. & MARITAN, A. 2003. Neutral theory and relative species abundance in ecology. Nature. 424: 1035-1037.

WILLIAMS, B. K., NICHOLS, J. D., & CONROY, M. J. 2001. Analysis and management of animal populations: modeling, estimation, and decision making. Academic Press.

ZAYED. A.I. 1993. Advances in Shannon's sampling theory. CRC Press.

Capítulo 1 – *Efeitos da heterogeneidade interespecífica de probabilidade de captura sobre o viés de seleção e estimação dos modelos de distribuição de abundância das espécies*

1. Introdução

Embora alvo de muitas críticas, estudos com abundância relativa das espécies em comunidades de organismos móveis geralmente utilizam dados de contagem de indivíduos como índices das abundâncias reais (Mackenzie et al 2006, Williams et al 2001). Os motivos de tais críticas estão relacionados à heterogeneidade de capturabilidade das espécies e a respectiva variação entre ambientes, uma vez que variações ambientais e atributos específicos, além da abundância da espécie, podem ter influência sobre a probabilidade de captura final das espécies (Remsem & Good 1996). A constatação dessa heterogeneidade interespecífica surgiu de observações em campo, que mostraram como características específicas (e.g. comportamento e área de forrageamento) influenciavam o sucesso de captura dos animais (MacArthur & MacArhtur 1974; Keyes & Grue 1982).

No caso de capturas de aves por rede de neblina, MacArthur & MacArthur (1974) e, posteriormente, Karr (1981) e Remsen & Good (1996) apontaram alguns atributos das espécies, do ambiente e do procedimento de amostragem que poderiam alterar a proporção de capturas de cada espécie em relação às suas abundâncias como: altura de forrageamento no estrato vegetal, tamanho corpóreo, frequência e amplitude de voo, tensão da rede e altura do dossel. Cada um destes fatores condiciona uma probabilidade de captura que em conjunto determinam a probabilidade de captura realizada. Se supomos que estas probabilidades componentes são independentes, a probabilidade de captura final é um produto destas várias probabilidades.

Há anos são estão disponíveis métodos de estimação de abundância das espécies que levam em conta a probabilidade de captura imperfeita (Dahl 1919, Lincoln 1930 Schnabel 1938, Pollock 1982). Os métodos mais usuais envolvem um delineamento amostral de marcação e recaptura, em que as populações são amostradas várias vezes, gerando um histórico de capturas sequenciais que permite a estimação das abundâncias correspondentes, ajustadas para probabilidades de captura (Nichols 1992; Williams et al. 2002). Entretanto, no nível de comunidade, as estimativas das abundâncias de cada espécie, em geral, não são viáveis, pois a maioria apresenta poucos indivíduos capturados e baixas taxas de recaptura (Jung et al 2000).

Entre as possíveis consequências de se utilizar dados de contagem para descrever as relações de abundâncias nas comunidades, sem o reconhecimento da distribuição de probabilidades de captura das espécies, é que independente das relações reais de abundância relativa das espécies, a distribuição obtida pode não ser a distribuição das abundâncias das espécies e sim das suas probabilidades de captura. Assim, as distribuições de contagem de capturas ou número de indivíduos capturados nas amostras podem ser radicalmente diferentes das distribuições de abundância das espécies nas comunidades (Williams et al 2001). Em um cenário mais pessimista, o padrão de abundância até então estabelecido poderia ser invalidado e os modelos teóricos propostos descartados. Alguns modelos teóricos de abundância relativa de espécies em comunidades incorporam o processo amostral, mas sob a premissa de que de igualdade de capturabilidade das espécies.

Considerando, então, a possibilidade do uso dos modelos de abundância como descritores das comunidades, com parâmetros considerados medidas comparativas entre comunidades (Pielou 1977), e a premissa dos mesmos de que as probabilidades de captura são equivalentes entre as espécies, é possível que muitos equívocos possam ser cometidos na utilização dos modelos teóricos de abundância como descritores de diversidade das comunidades. Neste contexto, nosso objetivo foi investigar com simulações computacionais se a probabilidade de captura diferenciada das espécies cria (1) viés de seleção, definido como mudanças nas chances do modelo correto ser selecionado e (2) viés de estimação, definido como a diferença entre a esperança dos valores dos coeficientes de cada modelo estimados e seu valor real. Para tanto, fizemos simulações em que comunidades com distribuição lognormal, série logarítmica e exponencial foram amostradas com e sem diferenças na probabilidade de captura das suas espécies.

2. Material e Métodos

2.1. Modelos de simulação

2.1.1 Comunidade

As distribuições probabilísticas de abundância das espécies escolhidas para representarem as comunidades foram a distribuição lognormal (equação 1), distribuição exponencial (equação 2) e a distribuição série logarítmica (equação 3), com as respectivas funções de densidade probabilística:

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

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

$$f\langle r|\theta\rangle = \theta e^{(-\theta r)}$$

(2)

, em que $f\,(r|\theta)$ é a probabilidade de uma espécie ter abundância r na comunidade, dados os valores do parâmetro θ, equivalente ao inverso da esperança do número de indivíduos de uma espécie.

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

(3)

, em que $f\,(r|\alpha)$ corresponde à probabilidade de uma espécie ter abundância r na comunidade, dado os valores do parâmetro α, que é proporcional ao número de espécies representadas por um só indivíduo na amostra e N o total de indivíduos (para mais detalhes, ver apêndice de Alonso 2006).

Optamos pelas distribuições descritas acima por serem modelos que se ajustam a um grande conjunto de dados empíricos, tornando-os de amplo uso em estudos de representação de comunidades biológicas (Hubbel 2001) e, portanto, familiares aos ecólogos.

2.1.2. Probabilidade de captura

Em nossas simulações, a captura depende de atributos das espécies e do método, sendo que cada atributo, individualmente, atribui uma certa probabilidade de que a captura ocorra. Por exemplo, a distribuição de alturas de voo de uma espécie de ave define a proporção do tempo em que os indivíduos voam à altura da rede, que é a probabilidade de captura condicionada por este fator. Supomos também que pode haver outros fatores que condicionam a captura, cada um com uma probabilidade de captura associada (e.g., tamanho corporal, frequência de voo, área de vida, comportamento territorial, sensibilidade à temperatura, e tempo e período de exposição da rede). Se assumimos que cada um destes fatores é independente dos demais, a probabilidade final de captura será o produto das probabilidades atribuídas por cada um dos fatores. Aqui chamaremos cada uma dessas probabilidades de componentes da probabilidade de captura da espécie. Portanto, dados n fatores e suas componentes p_j, a probabilidade final de captura (a) de um indivíduo de uma espécie i ser capturado foi calculada como o produto destas componentes probabilísticas:

$$a_i = \prod_{j=1}^{n} p_j \qquad (4)$$

Para representar estes componentes de probabilidade p_j associados a cada fator j, escolhemos a distribuição Beta, que descreve uma variável aleatória restrita entre zero a um, adequada para representar probabilidades (Bolker 2008). A distribuição Beta tem dois parâmetros, que podem ser combinados para resultar em formas de distribuição muito diferentes. Assim, é possível representar cenários em que os componentes atribuem alta probabilidade para a captura da maioria das espécies, o contrário, e mesmo uma distribuição bimodal (figura 1).

2.1.3. Amostragem

A distribuição de probabilidades escolhida para representar o número de indivíduos amostrados aleatoriamente de cada espécie da comunidade foi a distribuição Poisson. Ela foi escolhida por descrever um processo de amostral em que não há agregação dos indivíduos nas unidades amostrais, portanto em que as capturas das espécies são variáveis aleatórias independentes com taxa de captura constante (λ). A distribuição de densidade probabilística da Poisson é:

$$f(r \mid \lambda) = \frac{e^{\lambda} \lambda^{r}}{r!} \qquad (5)$$

, em que $f(r \mid \lambda)$ é a probabilidade de uma espécie ter r indivíduos na amostra, dado o valor de parâmetro de taxa de capturas λ. O parâmetro λ corresponde à esperança do número de capturas de indivíduos de uma espécie i, que é o produto da abundância da espécie na comunidade, N_i , com a probabilidade de captura correspondente, ai,:

$$\lambda_i = N_i \times a_i \qquad (6)$$

2.2. Simulações

Primeiro sorteamos 200 valores de uma das distribuições de probabilidades escolhidas para compor as abundâncias esperadas de uma comunidade fictícia com 200 espécies. Em seguida, sorteamos n valores de componentes da distribuição Beta para cada espécie, e calculamos o produto dos mesmos para ser a probabilidade de captura (a_i). Realizamos os sorteios de Beta e o cálculo do produto das componentes para cada uma das 200 espécies. Determinadas as abundâncias das espécies na comunidade e as probabilidades de captura correspondentes, simulamos a amostragem sob dois diferentes cenários: um em que cada espécie tem uma probabilidade diferente de ser capturada e outro em que as espécies são igualmente capturáveis. Para a amostragem no cenário de probabilidade de captura variável, (i.e., distinta entre as espécies), sorteamos 200 valores da distribuição Poisson em que o termo de probabilidade de captura a de λ (equação 6) foi determinado para cada espécie (equação 4). Já para o cenário de probabilidade de captura constante (i.e. igual para todas as espécies), sorteamos 200 valores da distribuição Poisson em que o termo a de λ foi o valor médio dos 200 valores de a de cada espécie.

2.2.1 Valores de parâmetros da simulação

Os valores escolhidos para os parâmetros μ e σ da lognormal foram 3,6 e 0,5, respectivamente. Já para os valores do parâmetro θ da distribuição exponencial e do parâmetro α da distribuição série logarítmica foram 0,02 e 35,44, respectivamente. Os valores dos parâmetros foram escolhidos por serem valores que geravam esperanças de 10000 indivíduos e 200 espécies nas comunidades, como forma de controlar o número de indivíduos e espécies em todas as comunidades.

No caso dos parâmetros da distribuição Beta, fizemos uma simulação para cada uma das cinco combinações de valores dos dois parâmetros (figura 1). Optamos por esta forma para abranger vários cenários de probabilidade de captura, do mais pessimista ao mais otimista de probabilidades de captura. Assim, quando assumimos valores de A = 1 e B = 5 , criamos um cenário em que os valores baixos das componentes probabilísticas são mais prováveis

(distribuição assimétrica à esquerda), ao contrário da combinação A= 5 e B =1, que simula o cenário mais otimista de alta eficiência de captura para a maioria das espécies, gerando uma distribuição assimétrica à direita das probabilidades componentes. Os cenários intermediários seriam quando assumimos A= 5 e B= 5 (distribuição unimodal) e A= 1 e B=1 (distribuição uniforme), em que no primeiro caso há um valor intermediário (0,5) mais provável e no segundo em que todos os valores de componentes são igualmente prováveis. Já quando atribuímos valores A= 0,5 e B = 0,5, criamos uma distribuição bimodal, com os valores mais baixos e mais altos da variável são os mais prováveis, simulando um método de captura altamente seletivo para um grupo de espécies.

Variamos também entre dois e cinco o número de componentes de probabilidade p_j da probabilidade de captura. No cenário mais pessimista (componentes de probabilidade com distribuição assimétrica à esquerda, figura 1), mesmo dois fatores resultaram em valores de probabilidade de captura muito baixos. Com isto, poucas espécies foram incluídas na amostra, em geral com um indivíduo. Assim, para o cenário assimétrico à esquerda, não foi possível fazer a seleção de modelos, nem estimativas de parâmetros para amostras em que as probabilidades de captura final eram compostas por mais que dois componentes na produtória (equação 4).

Ao total foram, então, 17 cenários de probabilidade de captura (cinco combinações de parâmetros de Beta e quatro variantes de números de componentes que determinam a probabilidade de captura realizada, com exceção de um dos cenários das componentes Beta) e dois cenários de amostragem (probabilidades de captura iguais e diferentes), totalizando 34 simulações de cenários de amostragem para cada um dos três tipos de comunidades. Cada simulação foi repetida 200 vezes.

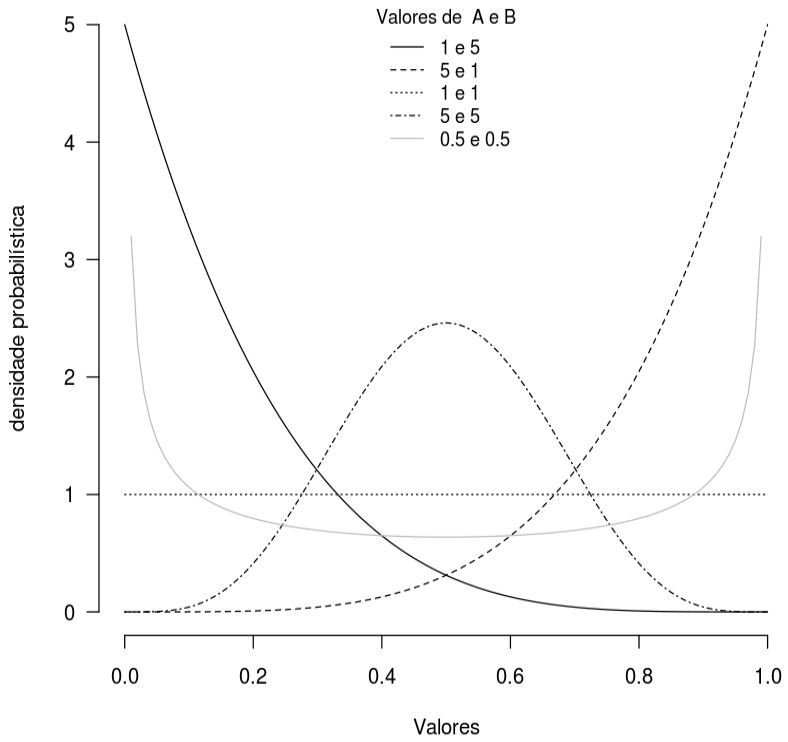

Figura 1. Distribuições Beta usadas para gerar os valores das componentes das probabilidades de captura, cujo produto é a probabilidade final de captura de cada espécie. Cada distribuição de valores das componentes é obtida com uma combinação de diferentes valores dos parâmetros (A e B) da distribuição Beta, e representam um cenário diferente para a variação das probabilidades destes fatores entre as espécies. Os cenários vão de iguais chances de ocorrência de todos os valores de probabilidades, à maior chance de valores extremos, ou médios (ver métodos).

2.3. Análises

2.3.1. Ajuste e Seleção de modelos

As funções de densidade probabilística utilizadas para o ajuste e seleção dos modelos das amostras foram funções compostas, que combinavam as funções de densidade das comunidades e a distribuição Poisson, como representativa do processo amostral. As funções de densidade resultantes utilizadas para o ajuste e seleção de modelos foram então: Poisson lognormal, série logarítmica e Poisson Exponencial. Com exceção da série logarítmica, que já é uma função truncada, as funções de densidade probabilística dos demais modelos foram truncadas na classe zero.

Para o cálculo dos vieses de estimação e seleção, então, precisávamos estimar os valores dos parâmetros das distribuições de abundâncias das comunidades com as amostras e analisar qual o modelo melhor se ajustava às mesmas. Assim, depois de geradas as amostras, ajustamos os modelos compostos correspondente a origem da comunidade que cada amostra era proveniente e obtivemos as estimativas de máxima verossimilhança dos parâmetros.

Depois prosseguimos com a seleção de modelos a fim de obtermos o modelo mais plausível para cada amostra. O critério de seleção que utilizamos foi o critério de informação de Akaike - AIC (Akaike 1973), em que o modelo de menor valor de AIC é selecionado, exceto quando a diferença entre AIC dos modelos concorrentes foi menor ou igual a dois, quando consideramos empate e ambos os modelos eram selecionados.

2.3.2. Viés de seleção e de estimador de parâmetros

Para o viés de seleção, como para todas as amostras sabíamos qual o modelo tinha gerado cada uma delas, a frequência com que ele não foi selecionado expressa o viés de seleção. Para estimar o viés das estimativas dos parâmetros, calculamos a diferença entre a média de todas as estimativas dos valores do parâmetro obtidas nas simulações e o valor do parâmetro. Esta diferença foi dividida pelo valor real dos parâmetros e multiplicada por 100, para expressar o viés em termos percentuais.

2.3.3. Viés do estimador de riqueza

Como as funções de densidade probabilística das distribuições Poisson logormal e Poisson exponencial, utilizadas nas funções de verossimilhanças inseridas na seleção de modelos, foram truncadas na classe zero, uma vez estimados os valores dos parâmetros é possível calcular a probabilidade da classe zero da função de densidade não truncadas. Assim, obtemos a proporção de espécies que não foram representadas na amostra, o que nos permite estimar o número de espécies na comunidade (Pielou 1977). Já para a distribuição série logarítmica, como parâmetro α é calculado a partir dos valores de número de espécies e indivíduos na amostra, e, é um parâmetro da comunidade, e por isso independente do tamanho da amostra, é possível com uma estimativa de total de indivíduos e o α estimado da amostra, obter o número de espécie em uma comunidade. A equação utilizada é:

$$S = \alpha \ln(1 + N/\alpha) \qquad (7)$$

, em que S é o número de espécies e N o número de indivíduos. O N utilizado para estimar o S da comunidade, na prática, seria um valor estimado a partir das amostras, porém utilizamos o valor real de 10000 indivíduos da comunidade simulada para representar um cenário ideal de estimativa do número de indivíduos e assim, não inserir nenhum componente adicional de erro além daquele que estávamos testando. Para a estimativa de viés, calculamos a diferença da riqueza estimada média obtida nas simulações e a riqueza real da comunidade.

Todas as simulações e análises foram executadas no programa R versão 2.11.1 (R

Development Core Team 2010). Além de das funções desenvolvidas especialmente para este estudo, contamos também com funções já implementadas em R, e disponível nos seguintes pacotes adicionais do R: poilog (Grøtan & Engen 2008), bbmle (Bolker 2010), vegan (Oksanen 2010) e untb (Hankin 2007).

3.Resultados

3.1. Viés de seleção de modelos

Com exceção do cenário de probabilidades componentes assimétrica à direita, que representa o cenário mais otimista de amostragem, a heterogeneidade de capturabilidade das espécies provocou vieses de seleção maiores que quando todas as espécies tinham a mesma probabilidade de captura. Embora em todos os tipos de comunidades, os vieses de seleção fossem maiores quando as probabilidades de captura eram diferentes entre as espécies, as comunidades com distribuição exponencial foram mais sensíveis aos cenários de heterogeneidade de capturabilidade e foram as que sofreram os maiores vieses. Além da heterogeneidade de capturabilidades, a forma da distribuição das componentes também provocou o aumento nos vieses de seleção. As formas que provocaram os maiores vieses foram unimodal e bimodal. E quando as probabilidades componentes tinham uma distribuição assimétrica à direita, não houve diferenciação entre os cenários com ou sem heterogeneidade de probabilidades de captura.

Assim, quando a premissa foi desrespeitada, o viés máximo chegou a 99,7% (figura 4b), e quando a premissa foi respeitada, o viés de seleção não ultrapassou 13,5% para nenhuma das distribuições (figuras 2a-4a). Este viés de seleção aumentou com a adição de componentes probabilísticas da capturabilidade. (figuras 2b-4b). A distribuição exponencial foi a que atingiu os maiores vieses, com um pico de 99,7% (figura 4b), seguida da lognormal, com valor máximo de 78,85% (figura 2b), e, a série logarítmica, com 37,45% (figura 3b).

3.2. Vieses nas Estimativas dos parâmetros

Para todas as combinações de cenários de amostragem, quando a probabilidade de captura variava, os vieses de estimativas dos parâmetros de todos os três modelos foram mais altos em relação aos vieses provocados pela probabilidade de captura equivalente (figuras 5-8). O parâmetro que se mostrou menos sensível à quebra da premissa foi o parâmetro de escala (μ) da lognormal (figura 5b), que sempre foi subestimado. Por outro lado, o parâmetro que apresentou o maior viés com a quebra da premissa foi o parâmetro de forma (σ), também da

lognormal (figura 6b), que em média foi superestimado.

O aumento do número de probabilidades componentes, com a exceção da estimativa de μ no cenário de distribuição bimodal das componentes (figura 5), causou um aumento do viés da estimativa de todos os parâmetros, quando as probabilidades de captura eram diferentes entre as espécies. A distribuição das probabilidades componentes também afetou o viés das estimativas. Os cenários mais otimistas de distribuição de componentes probabilísticas, (distribuições uniformes e assimétricas à direita dos componentes da capturabilidade), provocaram menor viés em todos os casos.

3.3.Estimativa de riqueza

Com exceção do cenário das componentes com distribuição assimétrica à direita, que é o cenário mais otimista, a probabilidade de captura diferenciada acarretou no aumento do viés de estimativa de riqueza, com a subestimação em relação ao valor real de 200 espécies (figuras 9-11). Além disso, o acréscimo de números de componentes da probabilidade de captura provocou aumento do viés para todos os cenários de amostragem, excluindo o cenário de componentes de probabilidades com distribuição assimétrica à direita.

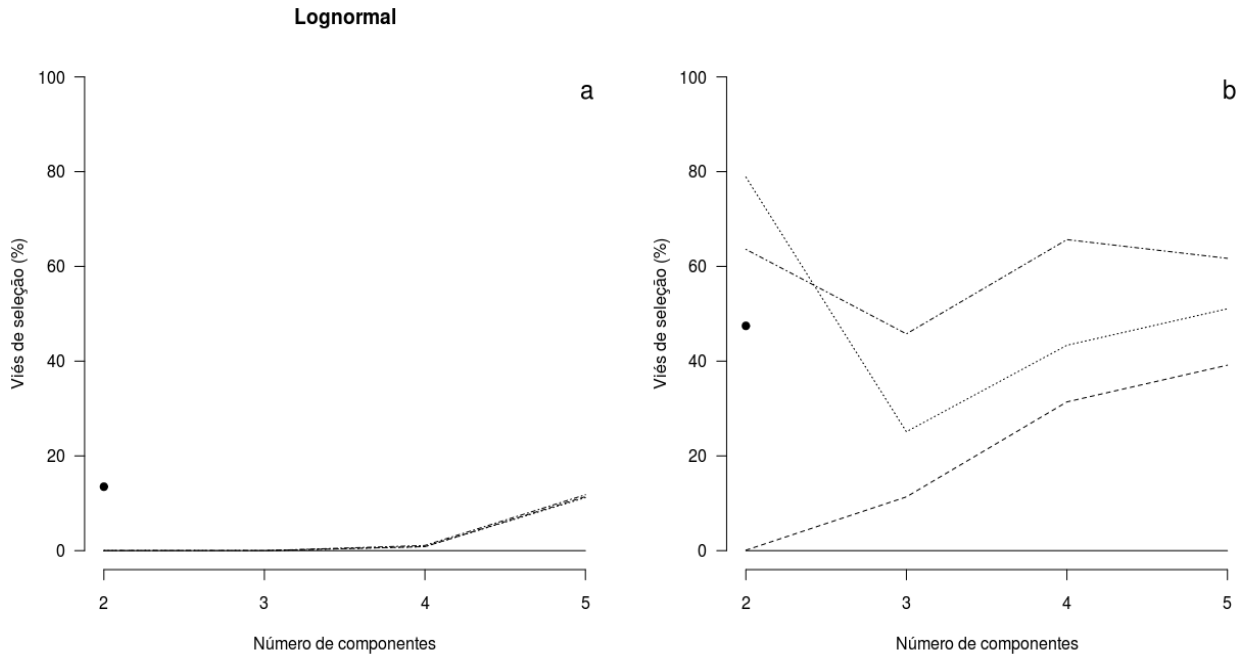

Figura 2. Porcentagem das simulações de amostras provenientes de comunidades Lognormal, em que as amostras não selecionaram o modelo Poisson-lognormal, e sim outro modelo concorrente (série logarítmica ou poisson-exponencial) quando probabilidade de captura é igual (a) e, quando varia entre as espécies (b) em função do número de componentes da probabilidade de captura. Cada linha representa um tipo de distribuição dos componentes de probabilidade (fig. 1): • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ····, unimodal; –·–, bimodal.

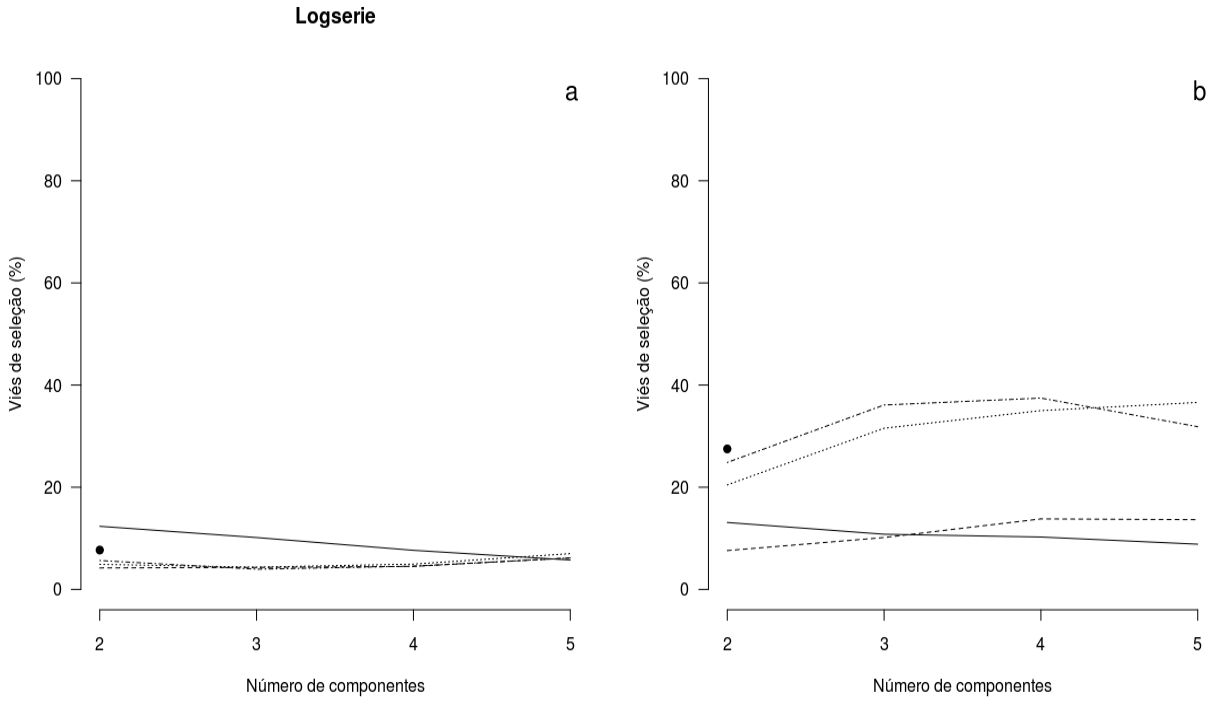

Figura 3. Porcentagem das simulações de amostras provenientes de comunidades série logarítmica, em que as amostras não selecionaram o modelo série logarítmica, e sim outro modelo concorrente (poisson-lognormal ou poisson-exponencial) quando probabilidade de captura é igual (a) e, quando varia entre as espécies (b) em função do número de componentes da probabilidade de captura. Cada linha representa um tipo de distribuição dos componentes de probabilidade (fig. 1): • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ····, unimodal; –·–, bimodal.

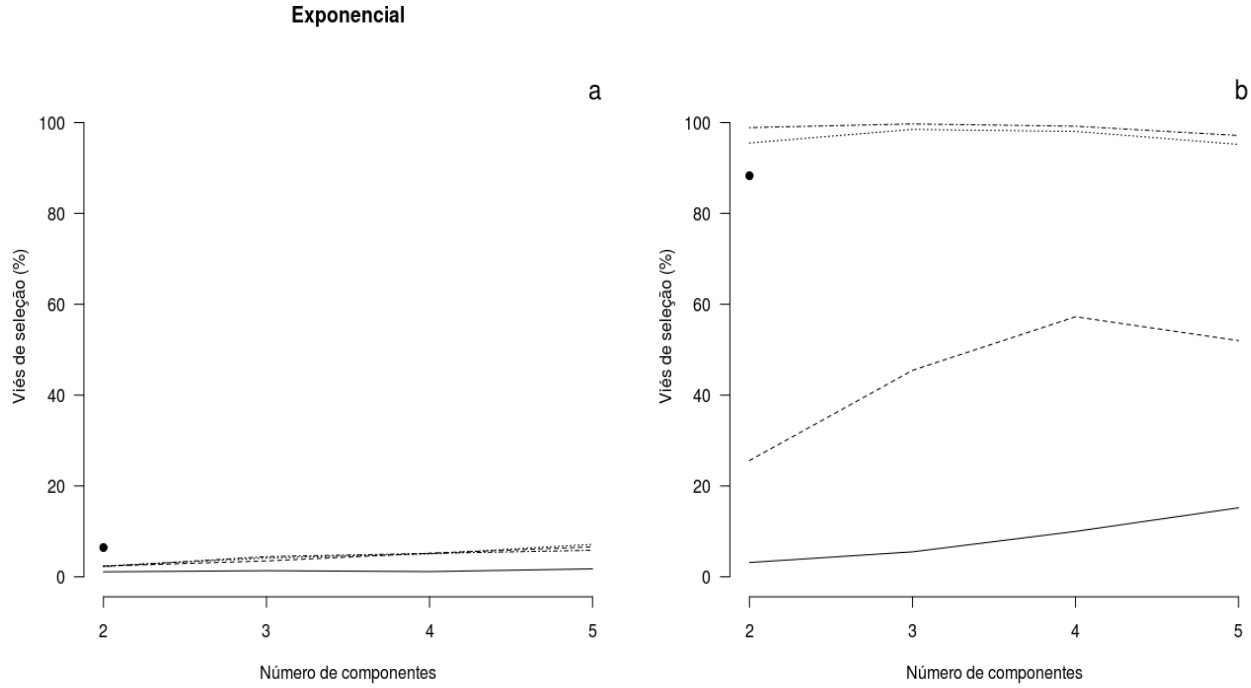

Figura 4. Porcentagem das simulações de amostras provenientes de comunidades Exponencial em que as amostras não selecionaram o modelo Poisson-exponencial, e sim outro modelo concorrente (série logarítmica ou poisson-lognormal) quando probabilidade de captura é igual (a) e, quando varia entre as espécies (b) em função do número de componentes da probabilidade de captura. Cada linha representa um tipo de distribuição dos componentes de probabilidade (fig. 1): • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ····, unimodal; –·–, bimodal.

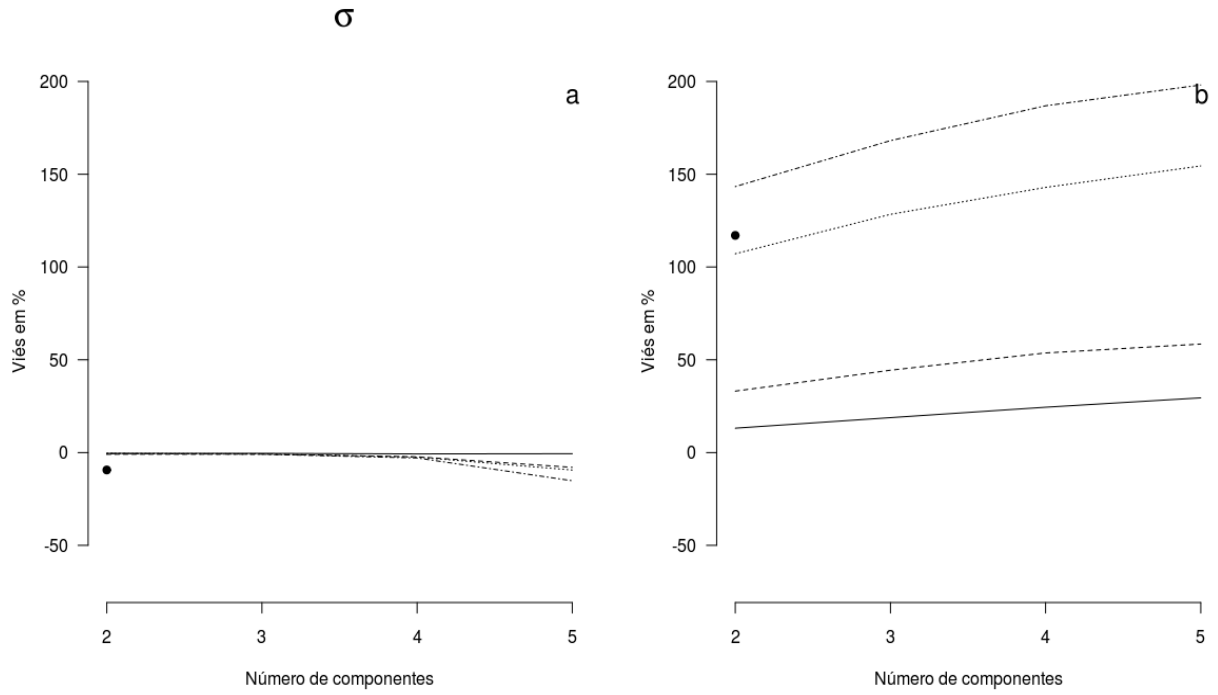

Figura 5. Viés percentual da média das estimativas do parâmetro σ da Lognormal em relação ao seu valor real, quando a probabilidade de captura varia (a) e, quando é igual para todas as espécies (b) de acordo com o número de probabilidades componentes. Cada linha representa um tipo de distribuição dos componentes de probabilidade (fig. 1): • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ····, unimodal; –·–, bimodal.

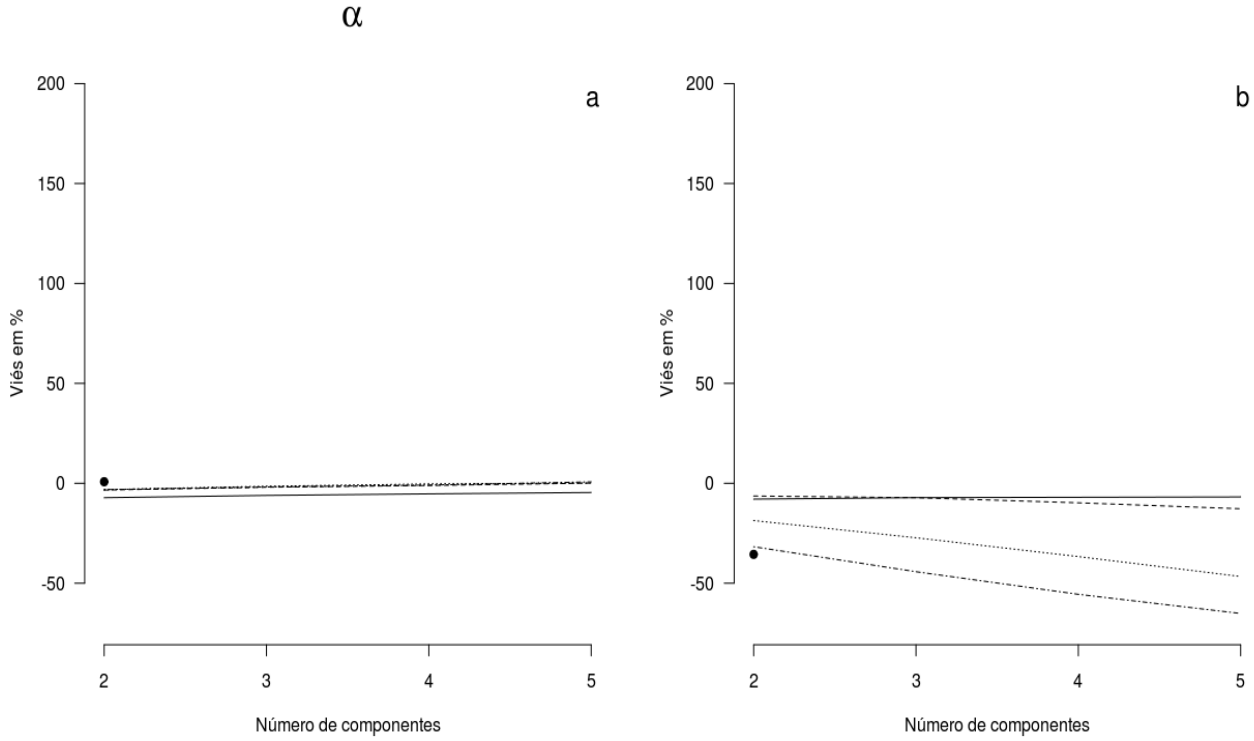

Figura 6. Viés percentual da médias das estimativas do parâmetro α da série logarítmica em relação ao seu valor real, quando a probabilidade de captura varia (a) e, quando é igual para todas as espécies (b) de acordo com o número de probabilidades componentes. Cada linha representa um tipo de distribuição dos componentes de probabilidade (fig. 1): • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ····, unimodal; –·–, bimodal.

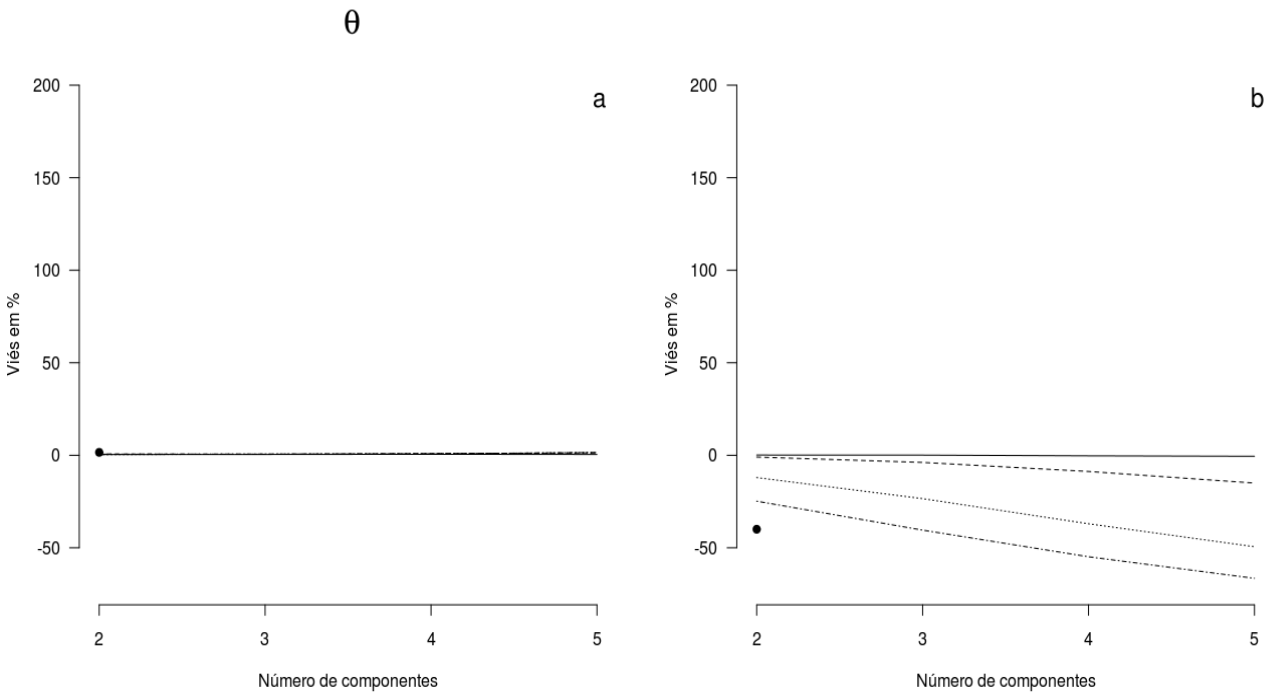

Figura 7. Viés percentual da média das estimativas do parâmetro θ da Exponencial em relação ao seu valor real, quando probabilidade de captura varia (a) e, quando é igual para todas as espécies (b) de acordo com o número de probabilidades componentes. Cada linha representa um tipo de distribuição dos componentes de probabilidade (fig. 1): • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ····, unimodal; –·–, bimodal.

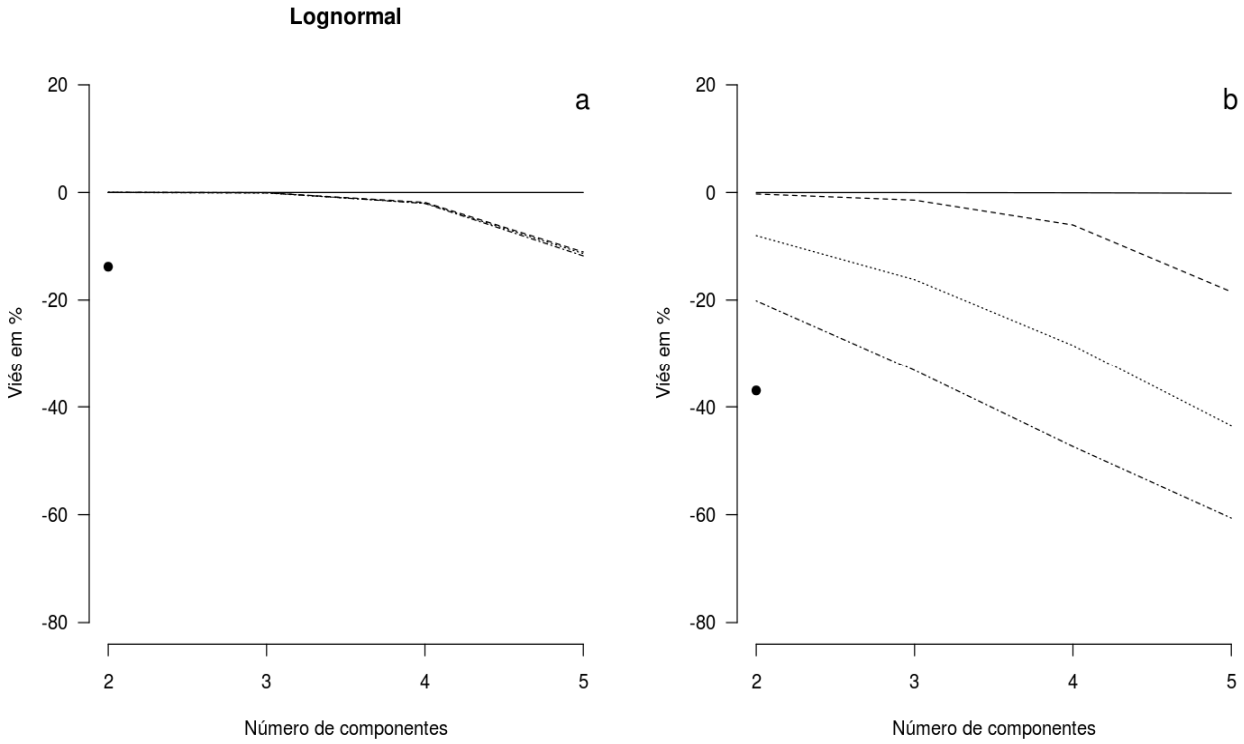

Figura 8. Desvio percentual da média das estimativas de riqueza, em relação ao número real de espécies nas comunidades, a partir do modelo Poisson Lognormal ajustado, quando probabilidade de captura varia (a) e, quando é igual para todas as espécies (b) de acordo com o número de probabilidades componentes. Cada linha representa um tipo de distribuição dos componentes de probabilidade (fig. 1): • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ····, unimodal; –·–, bimodal.

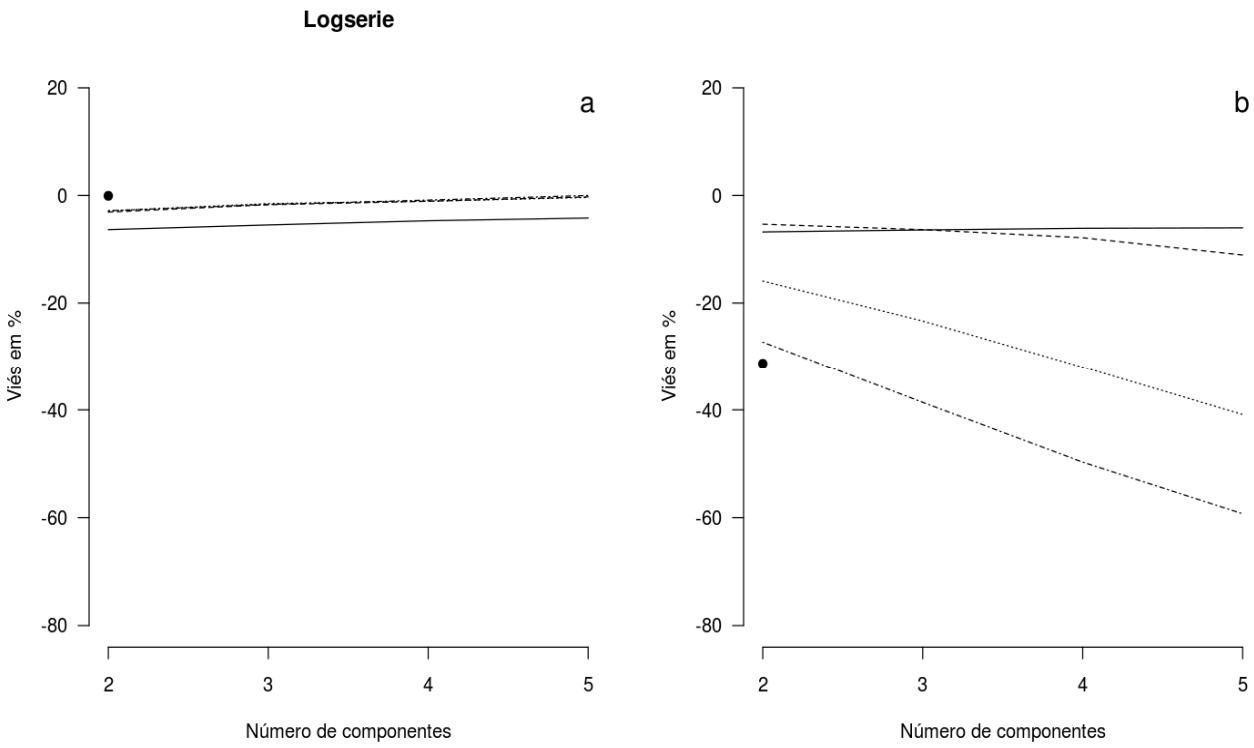

Figura 9. Desvio percentual da média das estimativas de riqueza, em relação ao número real de espécies nas comunidades, a partir do modelo série logarítmica ajustado, quando probabilidade de captura varia (a) e, quando é igual para todas as espécies (b) de acordo com o número de probabilidades componentes. Cada linha representa um tipo de distribuição dos componentes de probabilidade (fig. 1): • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ····, unimodal; –·–, bimodal.

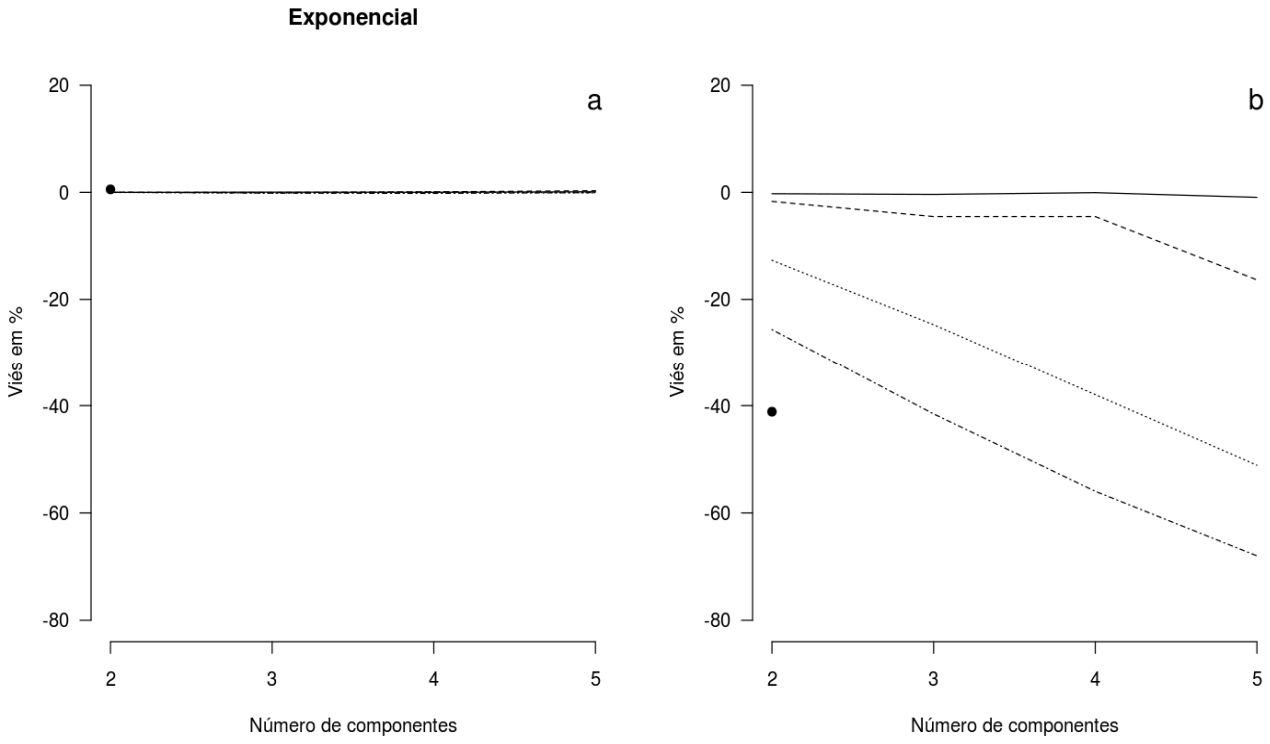

Figura 10. Desvio percentual da média das estimativas de riqueza, em relação ao número real de espécies nas comunidades, com o modelo Poisson Exponencial ajustado, quando probabilidade de captura varia (a) e, quando é igual para todas as espécies (b) de acordo com o número de probabilidades componentes; • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ····, unimodal; –·–, bimodal.

4. Discussão

Preston (1948) já havia notado que amostras cada vez menores poderiam representar a distribuição das abundâncias das espécies de uma comunidade de forma parcial, como se a verdadeira distribuição fosse truncada. Posteriormente, Dewdney (1998) reforçou a idéia, mas demonstrou que com o efeito da amostragem, não só a distribuição seria truncada, mas também deformada. No nosso trabalho, abordamos a existência de dois tipos possíveis e complementares de deformação da distribuição das abundâncias relativas causada pela amostragem: aquela causada por a amostragem ser sempre uma representação parcial da comunidade, e outra, causada não só pela amostragem ser uma fração da comunidade, mas por essa fração variar entre as espécies e estar relacionada a outros fatores além das diferenças de abundância entre elas. O primeiro tipo é incorporado em muitos modelos teóricos de distribuição de abundância, como as tradicionais série logarítmica de Fisher (1943), lognormal truncada de Preston (1948), e o recente modelo multinomial de soma zero de Hubbel (2001). Já o segundo tipo de deformação é um tópico pouco explorado nos modelos teóricos (Engen 1977, Alonso et al 2008), principalmente em uma abordagem de simulação como a que executamos aqui. Alonso et al. (2008) deduziram analiticamente o efeito da abundância e probabilidade de captura distinta entre as espécies no aprimoramento da descrição das comunidades, mas não fizeram uma avaliação deste efeito em comunidades reais ou simuladas. No nosso trabalho, portanto, não tentamos demonstrar se há ou não distorção na distribuição de abundância da comunidade, mas sim, quantificar a distorção quando uma das premissas dos modelos é desrespeitada.

O que encontramos é que quando existem diferenças de capturabilidade entre as espécies, dependendo da distribuição das probabilidades de captura das espécies e da distribuição de abundância que a comunidade segue, podemos não conseguir revelar a verdadeira distribuição através das amostras. Os modelos de distribuição de abundância são muito semelhantes entre si, uma vez que foram propostos para descrever um mesmo padrão (Connolly et al 2009), e dependendo das combinações de valores de parâmetros das distribuições, elas podem se tornar indistinguíveis. Magurran (2005) discute o efeito da variação de valores do parâmetro de dominância da binomial negativa, que derivou a série logarítmica de Fisher (1943), e concluiu que quando o parâmetro é alto (baixa dominância), a curva tende a uma lognormal completa e está geralmente associada a amostras grandes. Já quando assume valores baixos (alta dominância), a curva tende a uma série logarítmica de Fisher (1943) ou lognormal truncada de Preston (1948), e descreve bem abundâncias em amostras pequenas. A distribuição do

tipo lognormal truncada, por sua vez, dependendo de onde for a truncagem, aproxima-se muito de uma série logarítmica ou uma exponencial e por isso podem se tornar empiricamente indistinguíveis. Assim, os desvios nas estimativas dos parâmetros da distribuição original causados pela heterogeneidade de probabilidades de captura das espécies pode levar a seleção de outros modelos que não seja o de origem das amostras.

Neste contexto, nossos resultados apontam que o viés de estimativa do parâmetro provocado pela heterogeneidade de capturabilidade pode ter levado à seleção de outro modelo, uma vez que os vieses de estimação aumentam nos mesmos cenários que os vieses de seleção são maiores . Para as comunidades com distribuição lognormal que simulamos, a probabilidade de captura diferenciada pode causar vieses, principalmente, nas estimativas de σ, que é o parâmetro de forma. Na lognormal, quanto maior o valor de σ, mais assimétrica fica a distribuição (Bolker 2008). O parâmetro de forma está associado à agregação dos valores por observação, que na distribuição de abundâncias de espécies corresponderia ao grau de dominância na comunidade (Magurran 2005), em que poucas espécies acumulam a maior parte dos indivíduos. Além do σ, o parâmetro de escala μ também foi sensível à probabilidade de captura diferenciada, uma vez que o valor real do parâmetro foi subestimado, quando estimado a partir de amostras que desrespeitaram a premissa de probabilidades iguais entre as espécies. O parâmetro de escala nas distribuições está relacionado à extensão da curva à direita e à sua altura, sendo que valores pequenos estão associados às curvas mais altas e caudas mais curtas (menos espécies).

Embora em menor grau que o parâmetro σ da lognormal, o parâmetro θ da exponencial também sofreu maior viés com a quebra da premissa de igual capturabilidade das espécies. Entretanto, apesar de também se tratar de um parâmetro indicativo de dominância, ao contrário de σ , com a quebra da premissa o θ foi subestimando, e portanto, a dominância subestimada. Ao introduzirmos a heterogeneidade das probabilidades de captura, permitimos que eventualmente algumas espécies raras possam ter probabilidades de captura mais altas que aquelas mais abundantes, o que torna os valores das esperanças das espécies na amostragem mais próximos, levando a amostras mais homogênea, resultando em subestimação da dominância. O mesmo não ocorreu com a lognormal, pois mesmo que ela tenha o mesmo tamanho de uma comunidade exponencial, ela apresenta maior dominância. Assim, na lognormal, as espécies raras têm menos indivíduos que as que as espécies mais raras de uma comunidade exponencial. Por isso, mesmo que na comunidade lognormal as espécies raras tenham probabilidades de captura altas, elas ainda permanecem com as esperanças bem mais baixas que as abundantes nas amostras. Nas amostras retiradas de cada tipo de comunidade, portanto, quan-

do a probabilidade de captura varia, aquelas provenientes da comunidade lognormal sofrem um viés de superestimação da dominância e a exponencial, o contrário. Entretanto, quando a probabilidade de captura não varia, mas apenas diminui, ambos os tipos de amostras sofrerão superestimação da dominância, pois não haverá a variação das probabilidades de captura para alterar as relações das esperanças das espécies nas amostras.

Finalmente, o parâmetro α da série logarítmica, que é amplamente utilizado como um índice de diversidade (Kempton & Taylor 1974; Kempton & Taylor 1976; Magurran 2004), também sofre maior viés de subestimação quando a premissa de homogeneidade de capturabilidade é desrespeitada, o que na prática levaria inferir menor diversidade que o real. Assim, todos os parâmetros dos modelos utilizados para representar a distribuição das abundâncias das espécies foram sensíveis à quebra da premissa testada, com resultados que poderiam levar a interpretações distintas. Como o desrespeito da premissa de probabilidade de captura equivalente entre as espécies afetou as estimativas dos parâmetros, a estimativa de riqueza também foi afetada, uma vez que ela é derivada dos parâmetros e por isso dependente do grau de dominância. A riqueza, portanto, também sofreu um viés de subestimação acentuado.

Não só a variação das probabilidades de captura, mas também o aumento do número de componentes contribuiu para o aumento de viés nas estimativas dos parâmetros e de riqueza. Como definimos a probabilidade de captura como uma produtória de probabilidades quando aumentamos o número de fatores na produtória, tendemos a probabilidade de captura para números extremamente pequenos. Logo, o acréscimo de fatores acelera a diminuição das chances das espécies serem capturadas. E, considerando o número de atributos nomeados por MacArthur & MacArthur (1974) e Remsen & Good (1996), como contribuintes da probabilidade de captura de aves por rede de neblina, é possível que a probabilidade de captura realizada seja composta por mais de uma componente e, por isso, tenda de fato a valores muito baixos. Foi por isso que o aumento do viés ocorreu para ambos os tipos de amostras, uma vez que ao aumentarmos o número de componentes, diminuímos as probabilidades de captura em geral, e a média delas também, levando a amostras cada vez menores e menos representativas das comunidades. Além disso, o aumento do número de componentes de diferentes valores acarreta também o aumento do coeficiente de variação da capturabilidade, uma vez que a média tende a zero mais rapidamente que a variância. O aumento da variação relativa à média implicaria em eventuais probabilidades de captura em valores extremos, resultando em uma ou algumas espécies com o número de capturas muito altos na amostra, o que por sua vez afetaria a dominância. Resumindo, o aumento da variação das probabilidades de captura, em combinação a amostras cada vez menores, podem potencializar as grandes distorções nas a-

bundâncias na amostra. Portanto, os resultados ressaltam a importância de se estimar o quão representativa a amostra pode não ser, dependendo das premissas que foram desrespeitadas.

Além do número de componentes, a distribuição das probabilidades componentes (figura 1) também afetou o viés das estimativas. Isso porque algumas distribuições das probabilidades componentes, inserem maior ou menor variância nas probabilidades de captura, resultando em maior ou menor viés. O cenário de distribuição de componentes bimodal, por exemplo, que foi o que causou os maiores vieses nos valores estimados dos parâmetros, estabelece maiores probabilidades a valores extremos da variável (i.e. zero e um), inserindo maior variância para as probabilidades de captura (Figura 1). Já os menores vieses ocorreram para as amostras a partir dos cenários de distribuição das componentes assimétrica à direita, que são os cenários mais otimistas das probabilidades das componentes, pois além de ser uma das distribuições que atribuem as menores variâncias às probabilidades de captura das espécies, é a que prevê os valores mais altos para as capturabilidades.

Como os parâmetros dos modelos analisados podem ser interpretados como descritores de dominância das espécies, e o estimador de riqueza, uma medida sintética resultante dos valores dos parâmetros, as estimativas que analisamos no trabalho são utilizados como medidas de diversidade. Assim, a subestimação de μ e α , a superestimação de σ, dos modelos série logarítmica e lognormal, e, por último, a subestimação da riqueza, causada pela heterogeneidade de probabilidade de captura das espécies, levariam a subestimação da diversidade. E, mesmo no caso do modelo exponencial, se por um lado, o parâmetro de decaimento θ indique menor dominância que o real, e por isso seria interpretado como alta diversidade, a riqueza continuou a ser subestimada, o que novamente levaria a interpretações de menor diversidade.

Os modelos utilizados, escolhidos por meio da seleção, e as estimativas de parâmetros desses modelos descrevem a forma e as relações de dominância ou raridade relativas das espécies percebidas podem ser utilizados para guiar critérios e decisões com objetivos conservacionistas. Entretanto nossos resultados indicam que dependendo da dimensão da variação de probabilidades de captura entre as espécies, o modelo selecionado e os valores estimados dos parâmetros a partir das amostras podem tornar tais resultados pouco confiáveis. Assim, essa sensibilidade dos estimadores à quebra de uma premissa que parece não ser muito realista pode inviabilizar o uso dos modelos como descritores da diversidade, bem como a sensibilidade ao viés de seleção inviabiliza o uso dos modelos como explanações teóricas dos mecanismos ecológicos estruturadores das comunidades.

Porém, o uso dos modelos como explicações alternativas de mecanismos ecológicos

estruturadores das comunidades têm outras falhas além desta apontada por nós. Essas falhas estão relacionadas à deficiência em se descartar os modelos teóricos por questões de limitações dos métodos estatísticos tradicionais, que não a rejeição dos modelos e não permite a comparação simultânea de vários modelos, e também pela própria abordagem padrão-processo, que é sempre questionada pelo fato de vários processos poderem dar origem a um mesmo padrão. Por esse motivo, McGill et al. (2007) no artigo "Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework" chamou a atenção para a quantidade de modelos teóricos que se acumularam, sem meios de descartá-los, e de como a aplicação dos modelos como explicações teóricas tem trazido poucos progressos para a área. E, assim como Pielou (1977), ele aconselha o uso dos modelos como ferramentas descritivas das comunidades, que podem ser aplicadas para a comparação entre as comunidades. Neste caso, é possível escolher um modelo, por critérios práticos como número de parâmetros, em que os valores dos parâmetros seriam utilizados para estabelecer as relações entre comunidades. Assim, embora tenhamos demonstrado que pode haver um aumento do viés de estimação e seleção quando as capturabilidades são diferentes entre as espécies, se o viés for constante, ou seja, independente dos valores dos parâmetros, o uso dos mesmos como variáveis resposta a efeitos ambientais ainda pode ser válido. Permanece em aberto então, a aplicação dos modelos como ferramentas comparativas entre comunidades, assunto que será abordado no próximo capítulo.

5. Bibliografia

AKAIKE, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. In: B. N. PETROV & F. CSAKI, eds. Second International Symposium on Information Theory. Budapest: Akademiai Kiado, pp. 267–281.

ALONSO, D., OSTILING, A. & ETIENNE, R.S. 2008. The implicit assumption of simmetry and species abundance distribution. Ecology Letters. 11: 93-105.

BOLKER, B. 2008. Ecological Models and Data in R. Princenton University Press.

BOLKER, B. 2010. bbmle:Tools for general maximum likelihood estimation. Pacote de R versão 0.9.5.1 baseado no pacote stats4 por R Development Core Team. URL http://CRAN.R-project.org/package=bbmle

CONNOLLY, S. R., DORNELAS, M., BELLWOOD, D. R. & HUGHES, T.P. Testing species

abundance models: a new bootstrap approach applied to Indo-Pacific coral reefs.Ecology, 90(11), 2009, pp. 3138–3149.

DAHL, K. 1919. Studies of trout and trout-waters in Norway. Salmon Trout Mag. 18:16-33.

DEWDNEY, A.K. 1998. A general theory of sampling process with applications to the "veil line". Theoretical Population Biology. 54: 294- 302.

ENGEN, S. 1977. Comments on two differents approaches to the analysis of species frequency data. Biometrics. 33: 205-213.

FISHER, R. A., CORBET, A. S. & WILLIAMS, C.B. 1943. The relation betwen the number of species and the number of induvidulas in a random sample of an animal population. The Journal of Animal Ecology. 12: 42-58.

GRØTAN, V & ENGEN, S. 2008. poilog: Poisson lognormal and bivariate Poisson lognormal distribution. Pacote do R versão 0.4.

HANKIN, R.K.S. 2007. Introducing untb, an Rpackage for simulating ecological drift under the unified neutral theory of biodiversity. Journal of Statistical Software., 22(12).

HUBBELL, S.P. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton, Princeton University Press.

JUNG, R.E.; DROEGE S.; SAUER, J. R.; LANDY, R. B. 2000. Evaluation of terrestrial and streamside salamander monitoring techniques at shenandoah national park. Environment Monitoring and Assessment, 63: 65-79.

KEMPTON, R.A. & TAYLOR, L.R. 1974. Log-series and log-normal parameters as diversity discriminants for the Lepidoptera. J. Anim. Ecol. 43: 381–399.

KEMPTON, R.A. & TAYLOR, L.R. 1976. Models and statistics for species diversity. Nature. 262: 818–820.

KEYES, B.E. & C.E. GRUE. 1982. Capturing birds with mist nets: a review. North American Bird

Bander 7: 2-14.

LINCOLN, F. C. 1930. Calculating water fowl abundance on the basis of banding returns. Circ. U. S .Dept. Agric., No. 118.

MACARTHUR, R. H. & MACARTHUR, A. T. 1974. On the use of mist nets for population studies of birds. Proc. Natl. Acad. Sci. USA 71:3230-3233.

MAGURRAN, A.E. 2004. Measuring Biological Diversity . *Blackwell Science, Oxford* .

MAGURRAN, A.E. 2005. Species abundance distribution: Pattern or Process ? Functional

Ecology. 19 : 177-181.

MACKENZIE, D.I., NICHOLS, J. D., ROYLE, J.A., POLLOCK, K.H., BAILEY, L.L. & HINES, J.E. 2006. Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence. Academic 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 (in press).

NICHOLS, J. D. 1992. Capture-recapture models: using marked animals to study population dynamics. Bio- science , 42:94–102.

OKSANEN, J.; BLANCHET, G.; KINDT, R.; LEGENDRE, P.; O'HARA, R.B.; SIMPSON G.L.; SOLYMOS, P.; STEVENS, M.H. & WAGNER, H. 2010. Vegan: Community Ecology Package. Pacote do R versão 1.17-2. URL http://CRAN.R-project.org/package=vegan

PIELOU, E. C. 1977. *Mathematical Ecology*. New York: John Wiley and Sons.

POLLOCK, K.H. 1982. A Capture-Recapture Design Robust to Unequal Probability of Capture . The Journal of Wildlife Management, 46 (3): 752-757 .

PRESTON, F.W. 1948.The commonness, and the rarity, of species. Ecology. 29: 254 – 283.

REMSEN, J. V. & GOOD, D.A. 1996. Misuse of data from mist-net captures to assess relative abundance in bird populations. The awk. 113(2):381-398.

R DEVELOPMENT CORE TEAM. 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org .

SCHNABEL, Z. E. 1938. The Estimation of Total Fish Population of a Lake . The American Mathematical Monthly. 45(6): 348-352.

WILLIAMS, B. K., NICHOLS, J. D., & CONROY, M. J. 2001. Analysis and management of animal populations: modeling, estimation, and decision making. Academic Press.

Capítulo 2 - *Aplicação dos modelos teóricos de distribuição de abundância das espécies na avaliação de efeitos ambientais quando a premissa dos modelos de igualdade de probabilidades de captura entre as espécies é desrespeitada*

1. Introdução

A diversidade biológica é geralmente decomposta em duas variáveis: riqueza, que é o número de espécies da comunidade, e dominância (ou uniformidade), que reflete as relações de abundância relativa das espécies (Magurran 2004). Essas variáveis são frequentemente utilizadas para caracterizar as comunidades e avaliar impactos antrópicos sobre a biodiversidade, com fins de manejo e preservação (e.g. Barbour et. al 1995). Uma das maneiras de representá-las sinteticamente é por meio das distribuições de abundância das espécies (SADs[1],McGill et al. 2007). As SADs, por sua vez, revelaram um dos padrões mais bem estabelecidos da ecologia de comunidades, de que independente do grupo taxonômico ou do ecossistema descrito, existe sempre uma alta dominância nas comunidades biológicas (Pielou 1977). Esse padrão deu origem a diversas propostas de modelos teóricos (e.g. Motomura 1932, Fisher 1943, Preston 1948, MacArthur 1957), cujo conteúdo e generalidade fez com que McGill (2010) incluísse o tópico na lista do que ele chamou de teorias unificadas da ecologia. Além do componente teórico, os modelos já foram destacados como ferramentas comparativas, cujos parâmetros são considerados medidas sintéticas de diversidade e utilizados para estabelecer relações entre comunidades (Pielou 1977), que não necessariamente têm as mesmas composições (McGill et al. 2007).

Como a realização de censos é incomum em estudos com comunidades biológicas (William et. al 2001), houve a inclusão de uma teoria da amostragem na construção dos modelos teóricos de SADs, como forma de representar o processo amostral imperfeito (e.g. Fisher 1943) . Ao inserir uma teoria da amostragem nos modelos, passa-se a considerar explícita (Green & Plotkin 2007) ou implicitamente (Fisher 1943) um parâmetro escalar relativo à fração da comunidade amostrada. Entretanto, esse parâmetro escalar é aplicado para a comunidade como um todo e não por espécie, o que insere uma premissa de que as espécies

1 Utilizaremos essa sigla para designar as distribuições de abundâncias das espécies em uma comunidade, ou amostra dela tomada, em correspondência a SAD (species abundance distributions), que é o termo tradicionalmente utilizado na literatura.

têm as mesmas probabilidades de captura. No entanto, quando se trata de amostras de organismos móveis, é possível que haja uma variação interespecífica acentuada de probabilidade de registro ou captura, relacionadas a fatores específicos e ambientais (Remsen & Good 1996), além da abundância da espécie (MacArthur &MacArthur 1974). Essa heterogeneidade torna as SADs das amostras muito diferentes das SADs reais das comunidades (William et al 2001). Essa falta de correspondência entre as SADs das amostras e das comunidades podem ocasionar vieses tanto na identificação do modelo correto como nas estimativas dos parâmetros destes modelos.

No capítulo anterior mostramos que o viés de estimativa dos parâmetros dos modelos de abundância, a partir de amostras, desconsiderando a heterogeneidade de probabilidade de captura das espécies, podem ser muito altos. Esses vieses impossibilitam a interpretação dos valores absolutos dos parâmetros, como medidas de diversidade. Entretanto ainda não sabemos se o viés é sempre o mesmo, independente dos valores do parâmetro da SAD na comunidade. Então, se existe uma variável ambiental e os parâmetros dos modelos de SAD variam em função dela, se o viés for sempre o mesmo, independente dos valores que os parâmetros assumirem, a utilização de suas estimativas como medida comparativa para detectar o efeito do gradiente ainda seria válida.

O objetivo deste estudo foi avaliar o efeito da variação das probabilidades de captura entre as espécies sobre a detecção da influência de uma variável ambiental, quando se utiliza as SADs para descrever a resposta ao efeito da mesma. Investigamos se o efeito permanece perceptível, bem como se a magnitude do efeito permanece a mesma. Para tanto, simulamos amostras tomadas de comunidades em fragmentos florestais de diferentes tamanhos, com probabilidades de captura iguais e diferentes entre as espécies.

2. Material e Método

Criamos um vetor com vários valores de uma variável ambiental hipotética. Tal variável se referia a um atributo das manchas de habitats nas quais as comunidades eram estabelecidas na simulação. Dentro de cada uma dessas paisagens simuladas, as comunidades sempre pertenciam a um mesmo modelo de SAD, e apenas a forma da distribuição poderia variar, que por sua vez depende dos valores de parâmetros. Assim, as comunidades foram geradas a partir de modelos teóricos de SAD, cujos parâmetros poderiam ser: (1) dependentes da variável ambiental; (2) constantes ao longo da variação da variável ambiental. Para cada comunidade foram retiradas dois tipos de amostras, uma em que as probabilidades de captura

das espécies eram diferentes e outra em que as probabilidades de captura eram iguais entre as espécies. Depois, ajustamos às amostras o o mesmo modelo de SAD usado para gerar a comunidade, e obtivemos os valores estimados dos parâmetros. Os parâmetros então foram utilizados como variáveis-resposta em uma análise para detectar a relação com a variável ambiental. A frequência com que as amostras indicaram corretamente a presença ou ausência de efeitos, e os vieses sofridos nas estimativas do efeito, quando ele estava presente, nos permitiram avaliar a possibilidade do uso de parâmetros de modelos de SAD como variáveis resposta a um efeito ambiental. Todas as simulações, e análises foram executadas no programa R versão 2.11.1 (R Development Core Team 2010). Além das funções que criamos, utilizamos funções de pacotes já implementados em R. Os pacotes utilizados foram: vegan (Oksanen et al. 2010), untb (Hankin 2007), bbmle (Bolker 2010) e poilog (Grøten & Engen 2008).

2.1. Modelos de Simulação

2.1.1 SAD das comunidades

Os modelos teóricos de SAD escolhidos para gerar as comunidades foram: exponencial, série logarítmica e lognormal (equações nos métodos do capítulo um). O parâmetro θ da exponencial expressa a dominância da comunidade. Já o parâmetro α da série logarítmica é um índice de diversidade, que expressa a proporção com que o aumento do número de indivíduos na amostra adiciona espécies na mesma. Finalmente, os parâmetros μ e σ da Lognormal são parâmetros de escala e forma, respectivamente, sendo σ interpretado como um parâmetro indicativo de dominância na comunidade.

2.1.2 Relação da variável ambiental e a diversidade

A variável ambiental hipotética escolhida foi a área de fragmento. O intuito foi simular o processo de fragmentação atuando sobre a diversidade das comunidades, em que nossa predição teórica foi que aumento da área do fragmento diminui a dominância (θ, e σ), e aumenta a diversidade (α). No caso do modelo lognormal, de dois parâmetros, decidimos, no cenário em que havia relação com a variável ambiental, atribuir a relação somente para o parâmetro σ, que tem uma interpretação biológica mais bem estabelecida. Assim, quando σ variava em função da área do fragmento, μ permanecia constante. A área dos fragmentos foi expressa em hectares, com transformação em logaritmo na base natural (ln). No cenário em que havia relação entre a variável ambiental e os parâmetros dos modelos de SAD, a relação estabelecida entre os valores dos parâmetros a serem designados às comunidades e o

logaritmo do tamanho da área foi uma relação linear. Os valores dos parâmetros de cada modelo de SAD foram definidos, então, como:

$$I + b.\ln(ÁREA) \qquad (1)$$

,em que I é o intercepto e *b* o coeficiente angular.

Os extremos de valores de tamanho dos fragmentos bem como os valores de intercepto I e coeficiente angular *b* para definição dos valores de parâmetros das SADs das comunidades foram definidos com base em valores empíricos de estudos realizados em paisagens fragmentadas em área de Mata Atlântica. Todos esses estudos fazem parte do projeto de pesquisa temático "**Conservação da biodiversidade em paisagens fragmentadas"**, coordenado pelos Drs. Jean Paul Metzger e Renata Pardini (IB-USP) financiado atualmente pelo Programa Mata Atlântica (CNPq/Alemanha), Auxílio Jovem Pesquisador FAPESP 05/56555-4, e anteriormente pelo Auxílio Temático FAPESP 99/05123-4. Neste projeto foram estudadas três áreas do Planalto Atlântico Paulista com condições abióticas similares, porém diferentes graus de fragmentação em um gradiente com três níveis de cobertura vegetal com o objetivo inicial de se testar questões relacionadas aos limiares teóricos de fragmentação. Foram inventariados os grupos de aves e pequenos mamíferos (Bueno 2008; Martensen 2008; Banks-Leite 2009). Assim baseados nos tamanhos de fragmentos encontrados nas três áreas estudadas no projeto citado acima, estabelecemos como valores extremos de áreas nas nossas simulações. Já os valores de I e *b* (tabela 1) foram estabelecidos baseados nos valores obtidos com regressões lineares ajustadas com os dados de amostragem da avifauna obtidos a partir dos estudos dentro do projeto citado acima.

Em cada simulação, primeiro ajustamos os três modelos de SAD utilizados aqui, às abundâncias das espécies nas amostras tomadas de cada fragmento, por meio de máxima verossimilhança (ver capítulo 1). Com isso, obtivemos as estimativas de máxima verossimilhança dos parâmetros. Em seguida usamos um modelo Gaussiano linear para descrever a relação dos parâmetros estimados para cada fragmento e atributos do fragmento. Este modelo foi comparado com o modelo de ausência de efeito, no qual o valor do parâmetro tem distribuição normal, mas esperança constante.

Tabela 1. Valores de intercepto e coeficiente angular estabelecidos para a relação linear entre os parâmetros das SADs das comunidades simuladas em cada fragmento e o logaritmo da área dos fragmentos.

Parâmetro	Valor de I	Valor de *b*
θ da Exponencial	0,005	-0,001
α de Fisher (série logarítimica)	10,020	3,000
σ da Lognormal	1,070	-0,100

Para o cenário sem relação dos valores dos parâmetros das distribuições e o gradiente ambiental, os valores dos parâmetros dos modelos de SADs foram constantes para todos os fragmentos, definidos para os parâmetros μ e σ da lognormal como 3,6 e 0,5, respectivamente, como 0,02 para o valor de θ da distribuição exponencial, e como 35,44 para o parâmetro α da distribuição série logarítmica. Esses valores foram os mesmos que os utilizados nas simulações do capítulo anterior, escolhidos por terem ordem de grandeza compatível com comunidades de aves que são analisadas no capitulo seguinte.

2.1.3 Processo amostral

A distribuição de probabilidades escolhida para representar o número de indivíduos amostrados aleatoriamente de cada espécie da comunidade foi a distribuição Poisson. Ela foi escolhida por descrever um processo de amostragem sem agregação, em que as capturas das espécies são variáveis aleatórias independentes com taxa de captura constante (λ). A distribuição de densidade probabilística da Poisson é:

$$f(r \mid \lambda) = \frac{e^{\lambda}\lambda^{r}}{r!} \qquad (2)$$

, em que $f(r \mid \lambda)$ é a probabilidade de uma espécie ter *r* indivíduos na amostra, dado o valor de parâmetro de taxa de capturas λ. O parâmetro λ corresponde à esperança do número de capturas de indivíduos de uma espécie *i*, que é o produto da abundância da espécie na comunidade, N_i , com a probabilidade de captura correspondente, a_i:

$$\lambda_i = N_i \times a_i \qquad (3)$$

Em nosso modelo, a captura depende de atributos das espécies e do método, sendo que

cada atributo, individualmente, define uma certa probabilidade de que a captura ocorra. A probabilidade final de captura, portanto, é o produto das probabilidades atribuídas por cada um dos fatores. Aqui chamaremos cada uma dessas probabilidades de componentes da probabilidade de captura da espécie. Para representar os componentes probabilísticos associados a cada fator, escolhemos a distribuição Beta, que descreve uma variável aleatória restrita entre zero a um, adequada para representar probabilidades (Bolker 2008). A distribuição Beta tem dois parâmetros, que podem ser combinados para resultar em formas de distribuição muito diferentes. Assim, é possível representar cenários em que os componentes atribuem alta probabilidade para a captura, o contrário, e mesmo uma distribuição bimodal (figura 1 do capítulo um).

Portanto, dados n fatores e suas componentes p obtidas de uma distribuição Beta, a probabilidade final (a) de um indivíduo de uma espécie i ser capturado foi calculada como o produto destas componentes probabilísticas p:

$$a_i = \prod_{j=1}^{n} p_j \qquad (4)$$

2.2. Cenários de distribuição das componentes probabilísticas

No caso dos parâmetros da distribuição Beta, fizemos uma simulação para cada uma das cinco combinações de valores dos dois parâmetros (figura 1 do capítulo um). Optamos por esta forma para abranger vários cenários de probabilidade de captura, do mais pessimista ao mais otimista de probabilidades de captura. Assim, quando assumimos valores de A = 1 e B = 5 , criamos um cenário em que os valores baixos das componentes probabilísticas são mais prováveis (distribuição assimétrica à esquerda), ao contrário da combinação A= 5 e B =1, que ilustra o cenário mais otimista de alta eficiência de captura para a maioria das espécies, gerando uma distribuição assimétrica à direita das probabilidades componentes. Os cenários intermediários seriam quando assumimos A= 5 e B= 5 (distribuição unimodal) e A= 1 e B=1 (distribuição uniforme), em que no primeiro caso há um valor intermediário (0,5) mais provável e no segundo em que todos os valores de componentes são igualmente prováveis. Já quando atribuímos valores A= 0,5 e B = 0,5, criamos uma distribuição bimodal, na qual os valores mais baixos e mais altos da variável são os mais prováveis, representando a mais alta heterogeneidade na capturabilidade das especies.

Variamos também entre dois e cinco o número de componentes de probabilidade p_j da probabilidade de captura. No cenário mais pessimista (componentes de probabilidade com

distribuição assimétrica à esquerda, figura 1 do capítulo um), mesmo dois fatores resultaram em valores de probabilidade de captura muito baixos. Com isto, poucas espécies foram incluídas na amostra, em geral com um indivíduo. Assim, para o cenário assimétrico à esquerda, não foi possível fazer a seleção de modelos, nem estimativas de parâmetros para amostras em que as probabilidades de captura final era composta por mais que dois componentes na produtória (equação 4).

Ao total foram, então, 17 cenários de probabilidade de captura: cinco combinações de parâmetros de Beta e quatro variantes de números de componentes que determinam a probabilidade de captura realizada, com exceção de um dos cenários das componentes Beta e dois cenários de amostragem (probabilidades de captura iguais e diferentes). No total foram 34 simulações de cenários de amostragem, para cada cenário de simulação de presença ou ausência de relação do parâmetro com o gradiente.

2.3. Simulações

2.3.1 Comunidades

Quando os parâmetros dos modelos de SADs são dependente da variável ambiental

Determinado os valores assumidos pela variável ambiental, calculamos os valores dos parâmetros para cada comunidade (equação 1), definidos pela relação linear com a área do fragmento. Para as comunidades com SADs Exponencial e Lognormal sorteamos 200 valores das distribuições de probabilidades correspondentes, para cada fragmento . Obtivemos assim, comunidades com 200 espécies, cujos valores de parâmetros variavam em função do tamanho dos fragmentos simulados. Já para as comunidades série logarítmica, não foi possível definir um número único de espécies nas comunidades, pois como o parâmetro α é determinado pelo número de indivíduos e espécies da comunidade, ao variá-lo alteramos a relação entre a riqueza e abundância total. Assim, se fixássemos a comunidade em 200 espécies, com a variação de α, o número de indivíduos total atingiria valores pouco realistas, em que as comunidades dos fragmentos pequenos teriam muito mais indivíduos que os fragmentos maiores. Para gerarmos comunidades série logarítmica mais realistas, definimos uma relação linear entre o número de indivíduos das comunidades e o tamanho do fragmento. Uma vez definidos o número de indivíduos e o α, a partir da relação linear estabelecida com o tamanho de fragmento (eq. 1), o número de espécies pôde ser definido e assim as comunidades geradas.

Quando os parâmetros dos modelos de SAD são constantes independentes dos valores da variável ambiental

Para comunidade cujo parâmetro de modelo de SAD não tinha relação com o tamanho do fragmento, foi adotado o mesmo procedimento descrito acima, mas com valores de parâmetros iguais para todas as comunidades na paisagem simulada.

2.3.2 Amostragem

Geradas as distribuições de abundâncias dos fragmentos a serem amostrados, sorteamos *n* valores de componentes da probabilidade de captura para cada espécie, e calculamos o produto dos mesmos para ser a probabilidade de captura (a_i). Com isso, simulamos a amostragem sob dois diferentes cenários: um em que cada espécie tinha uma probabilidade diferente de ser capturada e outro em que as espécies são igualmente capturáveis. Para a amostragem no cenário de probabilidade de captura variável, (i.e., distinta entre as espécies), sorteamos S valores da distribuição Poisson, sendo S o número de espécies da comunidade a ser amostrada, em que o termo de probabilidade de captura a_i de λ foram os valores definidos para cada espécie (equação 3). Já para o cenário de probabilidade de captura igual para todas as espécies, sorteamos S valores da distribuição Poisson em que o termo a_i de λ foi o valor médio dos S valores das probabilidade de captura definidas para cada espécie. Cada simulação de comunidades no gradiente foi repetida 1000 vezes. Como foram três tipos de distribuição (Exponencial, série logarítmica e Lognormal) e dois tipos de relação com o gradiente, foram 6000 amostras em cada ponto do gradiente, para cada um dos cenários de amostragem.

2.4. Análises

Primeiro ajustamos os modelos de SAD para cada amostra do fragmento para obtermos as estimativas de máxima verossimilhança dos parâmetros (detalhes deste procedimento de máxima verossimilhança no capítulo 1). Obtidas as estimativas de parâmetros de cada fragmento, utilizamos os valores como variáveis dependentes (Vd) no modelo de regressão linear simples. Ao utilizarmos o modelo de regressão linear, consideramos os parâmetros variáveis aleatórias gaussianas, cuja média era uma função linear do logaritmo (ln) da área ou uma constante:

Modelo 1: **E[Vd] ~ I + b*ln(AREA)**

Modelo 2: **E[Vd] ~ constante**

Decidimos utilizar o modelo Gaussiano baseado no fato de que as variáveis-resposta no modelo são estimativas de máxima verossimilhança dos parâmetros dos modelos de SAD e que os estimadores de máxima verossimilhança tendem a uma distribuição Gaussiana com o aumento do tamanho amostral (propriedade de normalidade assintótica, Royle & Dorazio 2008).

Definido os modelos para cada conjunto de parâmetros estimados das amostras dos fragmentos, procedemos com o ajuste e seleção de modelos (Burnham & Anderson 2002). Para a seleção de modelos, então, utilizamos os valores de verossimilhança dos modelos concorrentes para o cálculo de AIC (Akaike 1973), que expressa a plausibilidade do modelo para a amostra, penalizado pelo número de parâmetros. Quanto menor o valor de AIC, mais plausível é o modelo. Na seleção de modelos então, fizemos uma classificação hierárquica de modelos de acordo com aos seus valores de AIC, sendo que diferenças de AIC do modelo mais plausível para os demais, menor que dois, foram considerados empates.

2.4.1. Viés de seleção

Como para todos os conjuntos de valores de parâmetros sabíamos qual a relação havia sido estabelecida com a variável de tamanho de fragmento, a frequência com que o modelo verdadeiro não era selecionado foi usado como a estimativa de viés de seleção.

2.4.2 Viés do estimador de parâmetros

O parâmetro de interesse a ser estimados na percepção da influência ou não do efeito da área dos fragmentos sobre as comunidades era o coeficiente angular *b,* obtido a partir do modelo 1. Para estimar o viés nos estimadores dos parâmetros, calculamos a diferença entre a média de todas as estimativas dos valores de *b* obtidos nas simulações e o valor do parâmetro estabelecido *a priori*. Esta diferença foi dividida pelo valor dos parâmetros e multiplicada por 100, para expressar o viés em termos percentuais. Para o cenário sem relação dos parâmetros com a variável de área de fragmento, como o valor real do coeficiente angular *b* era zero, o viés foi representado como os valores absolutos estimados, uma vez que não há como calcular porcentagens.

3. Resultados

3.1. Viés de seleção

Cenário em que há relação

Para as simulações em que os valores dos parâmetros das SADs eram uma função linear do logaritmo da área, todas as amostras, independente se as probabilidades de captura entre as espécies eram iguais ou diferentes, selecionaram corretamente o modelo que previa relação. Assim, o viés de seleção foi zero para todos os cenários.

Cenário em que não há relação

Para as simulações em que os valores dos parâmetros não tinham relação com a área dos fragmentos, houve viés, e com exceção das amostras tiradas de comunidades com uma SAD de série logarítmica, os vieses foram baixos, entre 3,3 % e 7,3%, independente de as amostras terem sido retiradas com probabilidades de captura distintas ou equivalente entre as espécies (Figuras 1-2). Já as amostras retiradas de comunidades com SAD de série logarítmica foram mais sensíveis à quebra da premissa de homogeneidade de capturabilidade. Nestes casos, foi selecionado com maior frequência o modelo errado, que previa relação, e o viés de estimativa do coeficiente angular aumentou em função do cenário de heterogeneidade de probabilidades de captura e também do aumento o número de probabilidades componentes (Figura 3). Assim, as amostras retiradas das comunidades série logarítmica quando a probabilidade de captura era diferente entre as espécies não selecionaram o modelo real, sem relação, em até 48% das vezes, enquanto que as amostras em que as espécies tinham a mesma probabilidade de captura, a porcentagem não passou de 9%.

3.2. Viés de estimação

Cenário em que há relação

No cenário em que havia relação entre os valores de parâmetros das comunidades e o tamanho dos fragmentos, o viés de estimação da magnitude do efeito, com exceção do cenário mais otimista de distribuição das probabilidades componentes (Beta assimétrica à direita), foi maior quando calculados a partir das amostras em que as probabilidades de captura eram diferentes entre as espécies (Figuras 4-6). O viés causado pela heterogeneidade de probabilidades de captura em questão foi sempre de subestimação e ocorreu para todos os

tipos de comunidades (Figuras 4b-6b), entretanto ele foi maior ou menor dependendo do cenário de probabilidades das componentes. O cenário de probabilidades componentes que provocou os maiores vieses para todos os tipos de comunidades foi o bimodal (Figura 1 do capítulo um), chegando a provocar o viés de subestimação máxima de 62,47% nos cenários de comunidades série logarítmica (Figura 6b). Além da inserção da variação de probabilidade de captura das espécies, o aumento do número de componentes multiplicativas, que diminui a média das probabilidades de captura e assim o tamanho amostral, também provocou acréscimos nos vieses de estimação do coeficiente angular. Essa sensibilidade ao tamanho amostral não ocorreu somente para as amostras em que as espécies tinham diferentes probabilidades de captura, mas também na estimativa da inclinação na relação entre o gradiente e os parâmetros μ e σ da lognormal, calculada das amostras em que as espécies eram igualmente capturáveis (Figuras 4b-6b).

Cenário em que não há relação

Quando não havia relação entre os parâmetros das comunidades e a variável ambiental, as estimativas de coeficiente angular, que na realidade era zero para todos os conjuntos de amostras, ficaram muito próximas de zero, independente se as amostras eram geradas com ou sem igualdade de capturabilidade entre as espécies (Figuras 7-9). Entretanto, apesar de terem sido pequenas, as estimativas a partir das amostras provenientes de cenários em que as probabilidades de captura eram diferentes entre as espécies apresentaram uma variação maior (Figuras 7b-9b).

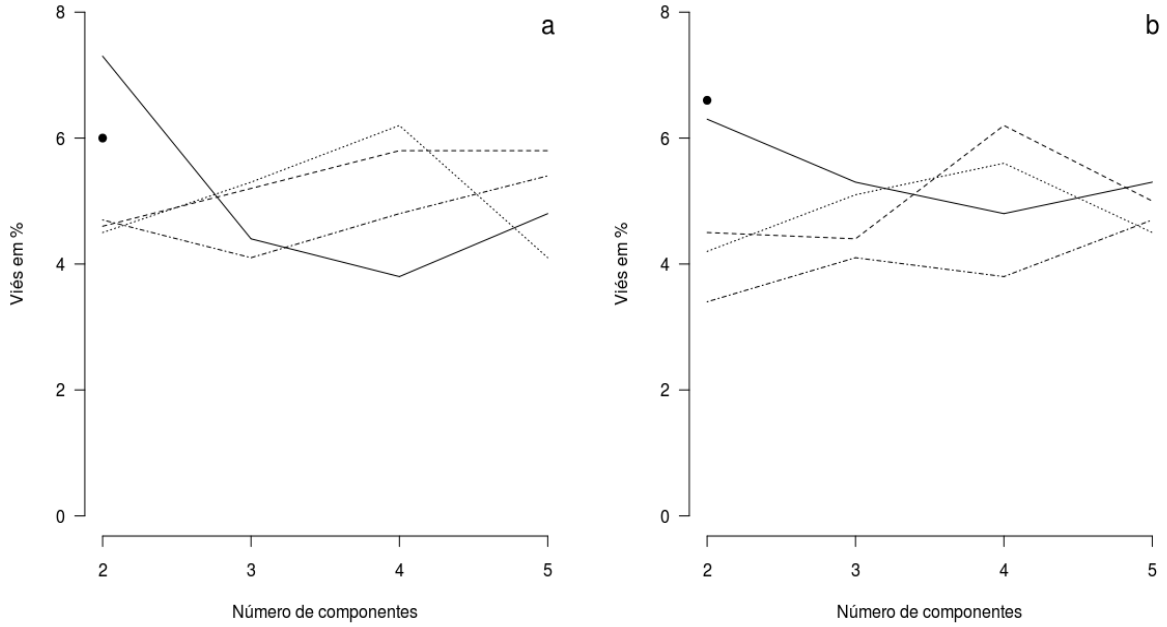

Figura 1. Viés de seleção na ausência de efeito da área dos fragmentos sobre as SADs das comunidades em cada fragmento, para a SAD exponencial, considerando que nos cenários em que o efeito estava presente o viés foi sempre zero. O viés de seleção é a porcentagem das simulações em que o modelo sem relação entre o parâmetro θ da SAD exponencial e o logaritmo da área do fragmento não foi selecionado, quando a probabilidade de captura é igual (a), ou diferente entre as espécies (b). O viés está apresentado em função do número de probabilidades que compõem as probabilidades de captura das espécies. As diferentes linhas representam cada cenário de distribuição de probabilidades componentes Beta: • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ···, unimodal; –·–, bimodal.

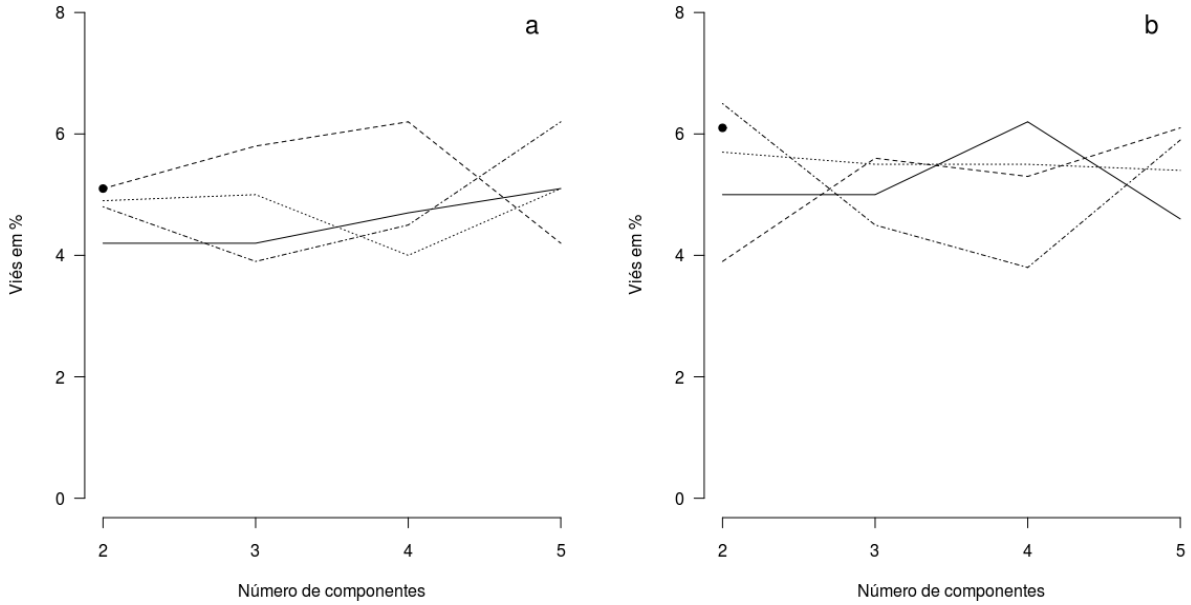

Figura 2. Viés de seleção na ausência de efeito da área dos fragmentos sobre as SADs das comunidades em cada fragmento, para a SAD lognormal, considerando que nos cenários em que o efeito estava presente o viés foi sempre zero. O viés de seleção é a porcentagem das simulações em que o modelo sem relação entre o parâmetro σ da SAD lognormal e o logarítmo da área do fragmento não foi selecionado, quando a probabilidade de captura é igual (a), ou diferente entre as espécies (b). O viés está apresentado em função do número de probabilidades que compõem as probabilidades de captura das espécies. As diferentes linhas representam cada cenário de distribuição de probabilidades componentes Beta: • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ···, unimodal; –·–, bimodal.

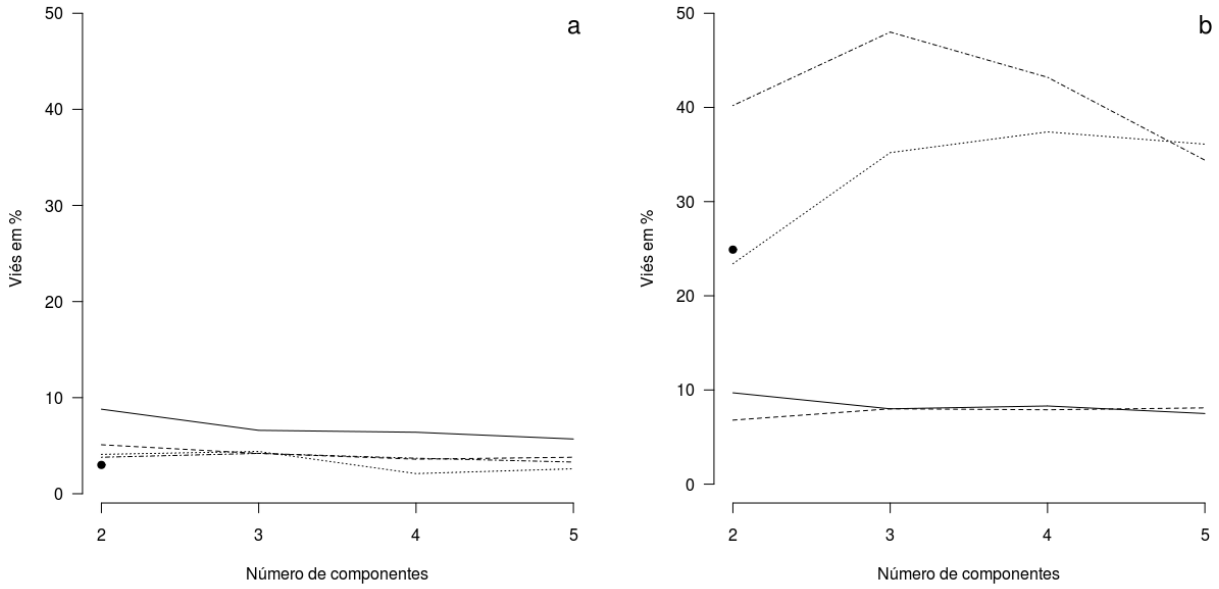

Figura 3. Viés de seleção na ausência de efeito da área dos fragmentos sobre as SADs das comunidades em cada fragmento, para a SAD série logarítimica, considerando que nos cenários em que o efeito estava presente o viés foi sempre zero. O viés de seleção é a porcentagem das simulações em que o modelo sem relação entre o parâmetro α da SAD série logarítimica e o logarítmo da área do fragmento não foi selecionado, quando a probabilidade de captura é igual (a), ou diferente entre as espécies (b). O viés está apresentado em função do número de probabilidades que compõem as probabilidades de captura das espécies. As diferentes linhas representam cada cenário de distribuição de probabilidades componentes Beta: • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ···, unimodal; –·–, bimodal.

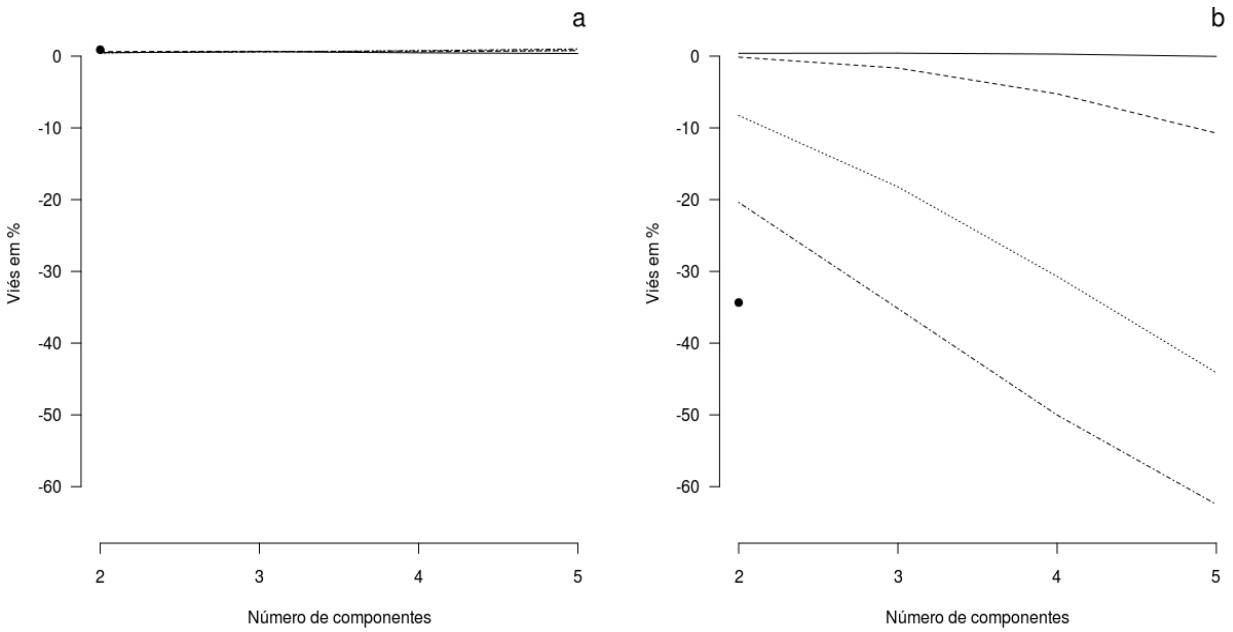

Figura 4. Viés do estimador de coeficiente angular da relação do logaritmo da área do fragmento sobre o parâmetro θ da SAD exponencial, quando há relação, quando a probabilidade de captura é igual (a), ou diferente entre as espécies (b), em função do número de probabilidades componentes e os cenários Beta das mesmas: • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ···, unimodal; –·–, bimodal.

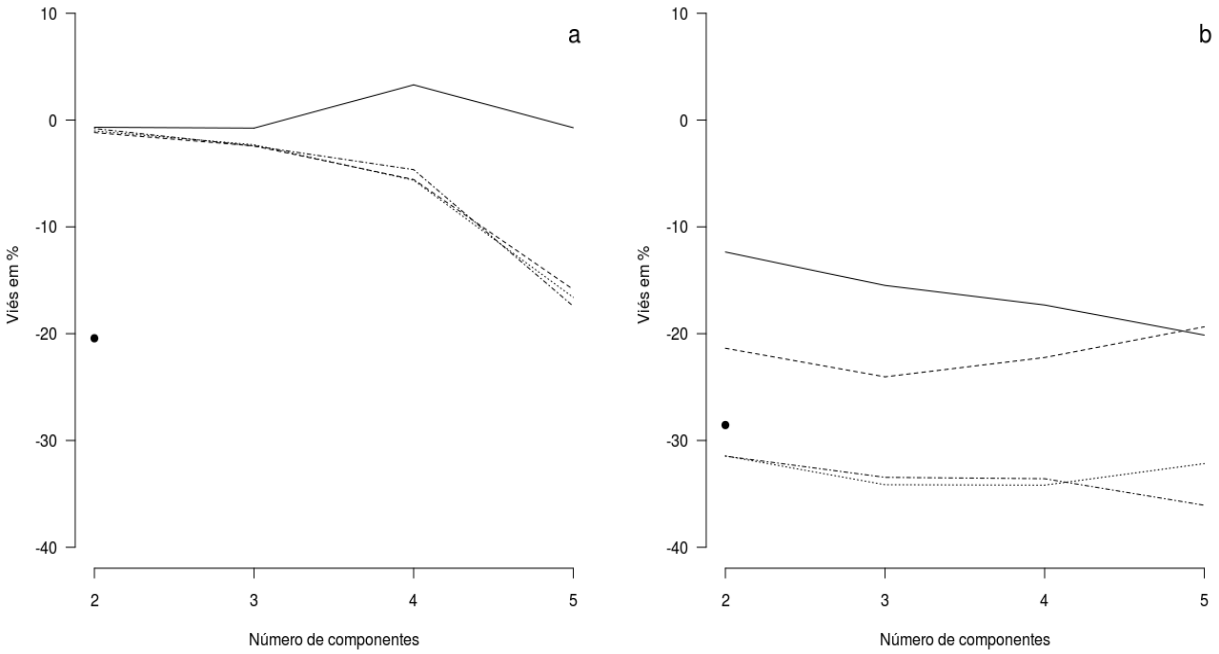

Figura 5. Viés do estimador de coeficiente angular da relação do logaritmo da área do fragmento sobre o parâmetro σ da SAD lognormal, quando há relação, quando a probabilidade de captura é igual (a), ou diferente entre as espécies (b), em função do número de probabilidades componentes e os cenários Beta das mesmas: • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ···, unimodal; –·–, bimodal.

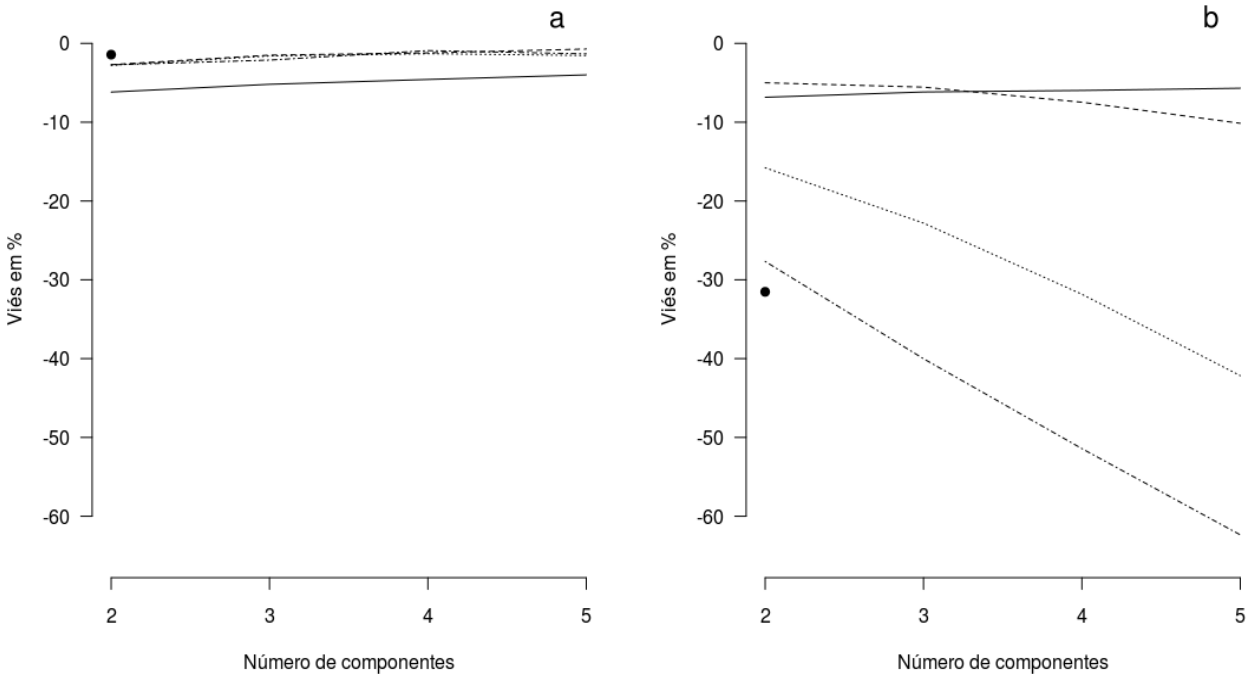

Figura 6. Viés do estimador de coeficiente angular da relação do logaritmo da área do fragmento sobre o parâmetro α da SAD série logarítmica, quando há relação, quando a probabilidade de captura é igual (a), ou diferente entre as espécies (b), em função do número de probabilidades componentes e os cenários Beta das mesmas: • , assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ···, unimodal; –·–, bimodal.

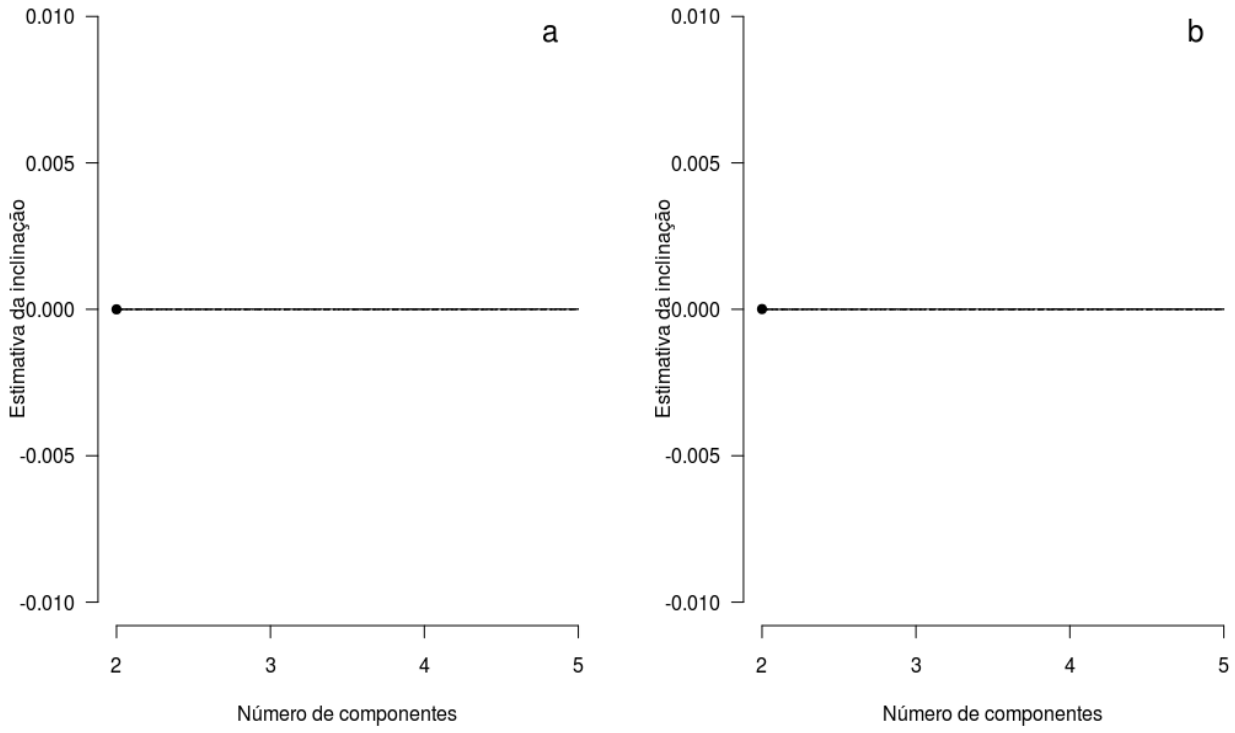

Figura 7. Média das **e**stimativas de coeficiente angular (inclinação) para a relação entre θ e o gradiente de fragmentação, quando não há efeito (o valor real da inclinação é zero). A linha visualizada é na verdade composta por cinco linhas sobrepostas que individualmente representa um tipo de distribuição dos componentes de probabilidade (fig. 1 do cap.1): assimétrica à esquerda; assimétrica à direita; uniforme; unimodal;bimodal.

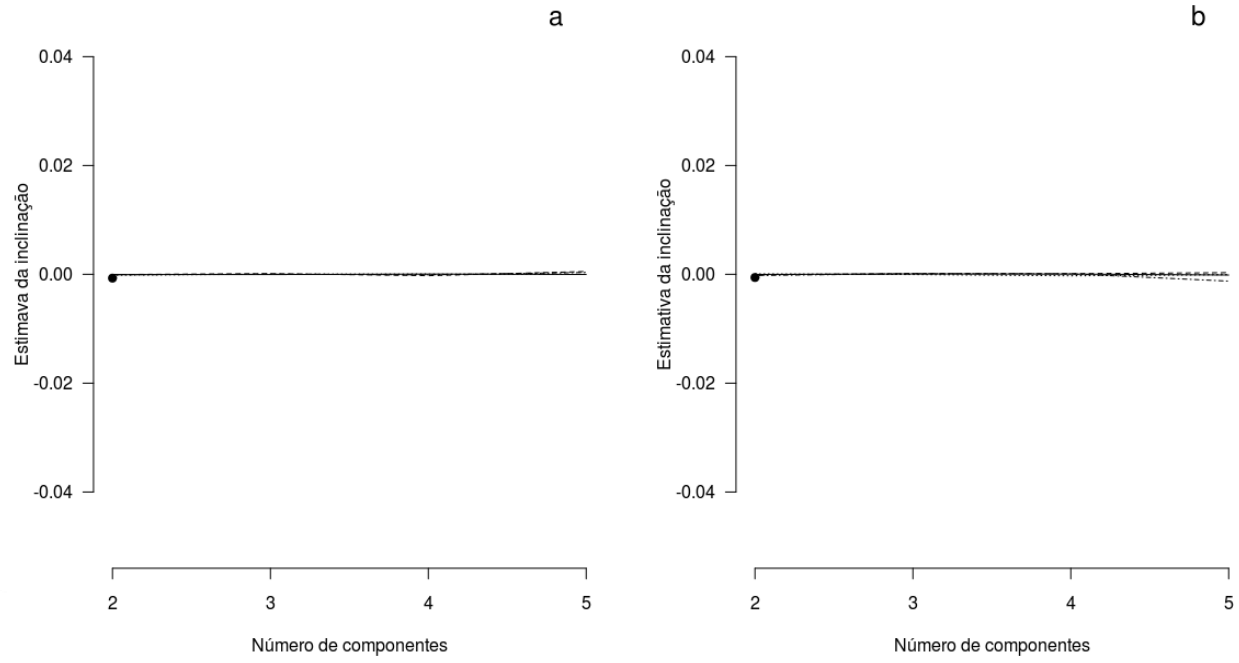

Figura 8. Média das **e**stimativas de coeficiente angular (inclinação) para a relação entre σ e o gradiente de fragmentação, cujo valor real é zero. A linha visualizada é na verdade composta por cinco linhas sobrepostas que individualmente representa um tipo de distribuição dos componentes de probabilidade (fig. 1 do cap.1): assimétrica à esquerda; assimétrica à direita; uniforme; unimodal;bimodal.

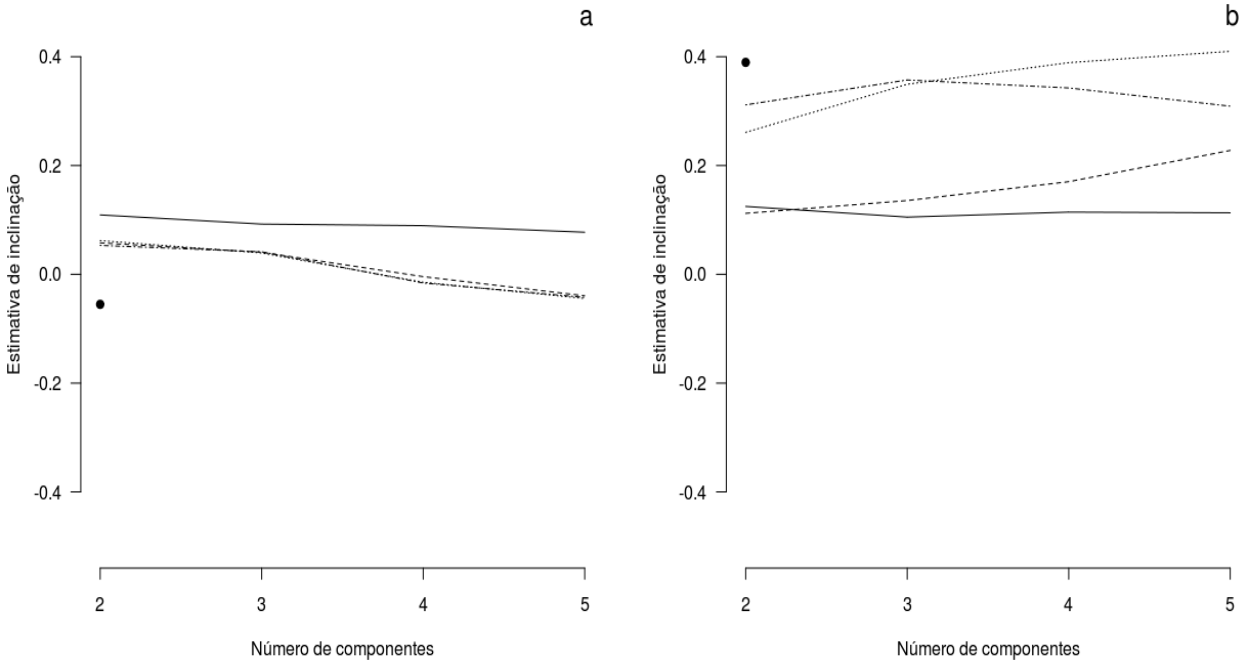

Figura 9. Média das **e**stimativas de coeficiente angular (inclinação) para a relação entre α e o gradiente de fragmentação, cujo valor real é zero. Cada linha representa um tipo de distribuição dos componentes de probabilidade (fig. 1 do cap.1): •, assimétrica à esquerda; — , assimétrica à direita; - - -, uniforme; ····, unimodal; –·–, bimodal.

4. Discussão

Embora a possibilidade do uso dos parâmetros dos modelos de SAD como variáveis resposta a efeitos ambientais seja aconselhada por alguns autores (Pielou 1977, McGill et al.2007), os resultados obtidos aqui indicam que é preciso ter cautela com este procedimento, quando as probabilidades de captura das espécies são desconhecidas. Em ambos os cenários simulados, com e sem relação entre os parâmetros dos modelos de SAD e o efeito de área de fragmento, os resultados da análise para estimar as relações a partir de amostras sofreram alguma forma de aumento de viés com o incremento da heterogeneidade de capturabilidade das espécies. Porém a magnitude dos vieses variou dependendo do grau de heterogeneidade de capturabilidade. Nos cenários de menor heterogeneidade, compostos por todas as variações de cenários de probabilidades componentes com distribuição assimétrica à direita, os vieses de seleção e estimação foram praticamente equivalentes aos cenários de igualdade de capturabilidade. Assim, o efeito da heterogeneidade de probabilidades de captura entre as espécies é importante apenas quando aumentamos seu coeficiente de variação, que depende do número de componentes e da distribuição de cada componente. O cenário de distribuição de probabilidades componentes que provocou os maiores vieses foi o bimodal, que é o que em

média insere a maior variância das probabilidades de captura, uma vez que este cenário prevê que os valores extremos têm maior chance de ocorrer.

Vamos discorrer, então, sobre os possíveis resultados que podem ser obtidos ao aplicar os modelos de SAD para detectar o efeito de tamanho de fragmento e as limitações de interpretação de cada um. Se o efeito não for detectado é porque provavelmente ele não existe, uma vez que encontramos que quando existe a relação com o tamanho de fragmento, ela sempre foi detectada. Por outro lado, detectar um efeito não garante que ele de fato exista. Se a estimativa da magnitude do efeito for muito baixa, há uma chance, ainda que pequena, e independente de as probabilidades de captura serem iguais ou diferentes, de na análise termos detectado um efeito que não existe. Para o modelo série logarítmica em especial, a probabilidade de selecionar erroneamente o modelo com efeito, quando ele não existe, aumenta à medida que o grau de heterogeneidade de capturabilidade também aumenta. Entretanto nesses casos em que há a possibilidade de um efeito inexistente ser detectado, a estimativa de coeficiente angular foi sempre muito baixa. Consideramos as estimativas baixas porque nas simulações, os valores máximos atingidos previam uma variação muito baixa dos parâmetros entre as comunidades dos diferentes fragmentos simulados. No caso do parâmetro θ da distribuição Exponencial, o modelo de relação previa uma diferença na ordem de 10^{-5} desde a comunidade no fragmento de menor tamanho até o de maior tamanho. Para o parâmetro σ do modelo Lognormal, as diferenças dos valores entre as comunidades do maior e menor fragmento não ultrapassariam 0,006 . Finalmente para o parâmetro α da série logarítmica, o modelo de relação previu uma diferença máxima entre comunidades de 1,9. Considerando que os tamanhos dos fragmentos variaram entre 2 e 175 hectares, a variação nos valores de parâmetros prevista pelos modelos de relação, provocaria alterações, na prática, desprezíveis na diversidade. Neste contexto, concluímos que só podemos considerar que de fato o efeito existe quando, além do modelo de relação ter sido selecionado, que o coeficiente angular estimado seja alto, sendo ao menos maior do que os valores estimados nas nossas simulações de cenários sem relação.

É imprescindível, portanto, considerar a heterogeneidade de probabilidades de captura entre as espécies na interpretação de resultados que estabeleçam relação entre os parâmetros dos modelos de SAD e uma variável ambiental. Em termos práticos, se a heterogeneidade de capturabilidade entre as espécies for totalmente ignorada, o viés de subestimação do efeito de um gradiente ambiental sobre os parâmetros dos modelos de SAD das comunidades ou a detecção de efeitos inexistentes pode trazer consequências relevantes em estudos que relacionam diversidade com gradientes ambientais. Muitos desses estudos têm como objetivo

identificar os fatores ambientais aos quais tais comunidades são mais sensíveis, essenciais na definição de critérios para guiar práticas de manejo e preservação. Assim, ao aplicar os modelos teóricos de SAD para identificar os efeitos de uma variável ambiental, mesmo que não se saiba como é a distribuição das probabilidades de captura das espécies, é preciso levar em consideração os tipos de vieses a que podem estar sujeitos os resultados.

Entretanto, o ideal é que se tenha uma idéia do grau de variação das probabilidades de captura das espécies coletadas, pois a partir da avaliação do grau de heterogeneidade de probabilidades de captura e a comparação com os cenários de heterogeneidade de capturabilidade considerados aqui, é possível ter uma estimativa da magnitude do viés a que os resultados estão sujeitos. Porém, estimativas de probabilidades de captura no nível da comunidade podem ser pouco viáveis. Os métodos para estimar probabilidades de captura mais recomendados são aqueles baseados em dados de marcação e recaptura (William et al 2002), em que para a estimativa são utilizados dados de histórico de capturas sequenciais, geralmente aplicados a uma só população de cada vez (Nichols 2000). E, uma vez que o padrão que se encontra nas amostras de contagem de indivíduos das espécies das comunidades é de que a maioria são poucas vezes capturadas, com muitas espécies apresentando apenas um ou dois indivíduos ("singleton" e "doubletons"), não é possível obter dados suficientes que sustentem as premissas dos modelos (Pollock et al., 1990) e as estimativas dos parâmetros. Mesmo utilizando modelos de ocupação, que visam estimar abundâncias médias, e que são mais simples e talvez menos precisos, muitas vezes não é possível calcular as estimativas dos parâmetros do modelo para todas as espécies da comunidade. Mesmo em estudos de longa duração de comunidades pequenas esta é uma limitação importante. Este é o caso do estudo de uma comunidade de salamandras conduzido por Dodd e Dorazio (2004), em que eles utilizaram um modelo simples adaptado de Royle (2004), com o uso de contagens de indivíduos em locais de amostragens fixos em um período de alguns anos, sem necessidade de marcação. Foram encontradas 13 espécies, que possivelmente são todas as que ocorrem no local e, ainda assim, só foi possível estimar as abundâncias médias de seis delas.

Diante das dificuldades amostrais de se estimar as probabilidades de captura das espécies utilizando os modelos com dados de marcação e recaptura para no nível da comunidade, talvez fosse mais eficiente investir esforços no delineamento amostral. A adequação do procedimento amostral, a escolha do grupo animal a ser amostrado e a exclusão de espécies coletadas com probabilidades de captura muito baixas, por exemplo, podem auxiliar na diminuição da variância das probabilidades de captura das espécies coletadas.

Mesmo que tais estratégias tornem o universo amostral cada vez mais restrito, a adoção das mesmas pode diminuir o efeito da heterogeneidade de probabilidades de captura e diminuir os vieses, uma vez que a estimativa da magnitude do efeito é mais robusta a medida que o grau heterogeneidade de probabilidades de captura diminui. Se o delineamento amostral for planejado para diminuir o grau de heterogeneidade das probabilidades de captura das espécies, os resultados serão mais indicativos e as interpretações mais conclusivas, ainda que as espécies não tenham probabilidades de captura exatamente iguais. Porém é necessário destacar que a aplicação de medidas como essas requer cautela e muito conhecimento sobre o sistema de estudo, as histórias naturais das espécies da comunidade e os procedimentos amostrais disponíveis. Uma vez que tais medidas também não possam ser tomadas (e.g. dados já coletados e comunidades altamente diversas), ao utilizar os modelos de SAD como indicadores do efeito de fragmentação em comunidades de organismos móveis é imprescindível que se leve em conta todos os possíveis cenários de heterogeneidade de capturabilidade. Isso porque mesmo que não seja possível estimar as probabilidades de captura das espécies ou avaliar o grau de heterogeneidade, ao considerar a possibilidade dela existir pode aumentar o poder de inferência dos resultados.

Como pudemos ver neste trabalho, quando a relação dos parâmetros dos modelos de SAD e a variável ambiental existe, ela é sempre detectada, mesmo que o efeito real seja subestimado. Assim, ao não detectarmos o efeito, é porque ele provavelmente não existe. Além disso, mesmo quando o efeito é detectado, dependendo da estimativa do efeito, é possível tecer alguns cenários de heterogeneidade de capturabilidade e descartar as hipóteses menos plausíveis e chegar a interpretações mais conclusivas e cautelosas dos resultados. Entre os modelos utilizados aqui, o indicado seria utilizar aqueles em que está inserido o processo amostral e que tenha pelo menos um parâmetro independente do esforço amostral, isto é, que seja uma medida da diversidade da comunidade sem a necessidade de ser reescalonada, como é o caso dos parâmetros α e σ das distribuições série logarítmica e lognormal, respectivamente.

5. *Bibliografia*

AKAIKE, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. In: B. N. PETROV & F. CSAKI, eds. Second International Symposium on Information Theory. Budapest: Akademiai Kiado, pp. 267–281.

BANKS-LEITE, C. Conservação da comunidade de aves de sub-bosque em paisagens fragmentadas da Floresta Atlântica. Tese (Doutorado). Instituto de Biociências da Universidade de São Paulo. 227p.

BOLKER, B. 2008. Ecological Models and Data in R. Princenton University Press.

BOLKER, B. 2010. bbmle:Tools for general maximum likelihood estimation. Pacote de R versão 0.9.5.1 baseado no pacote stats4 por R Development Core Team. URL http://CRAN.R-project.org/package=bbmle

BUENO, A. A. 2008. Pequenos mamíferos da Mata Atlântica do Planalto Atlântico Paulista: uma avaliação da ameaça de extinção e da resposta a alteração no contexto e tamanho dos remanescentes. Tese (Doutorado). Instituto de Biociências da Universidade de São Paulo, São Paulo. 124p.

BURNHAM, K. P., AND D. R. ANDERSON. 2002. Model selection and multimodel inference: a practical information-theoretic approach. 2nd Edition. Springer-Verlag, New York, New York, USA. 488 pp.

DODD, C.K. & DORAZIO, R.M. Using counts to simultaneously estimate abundance and detection probabilities in a salamander community. Herpetologica. 60 (4): 468-478.

FISHER, R. A., CORBET, A. S. & WILLIAMS, C.B. 1943. The relation betwen the number of species and the number of induviduals in a random sample of an animal population. The Journal of Animal Ecology. 12: 42-58.

GREEN, J. & PLOTKIN, J. 2007. A statistical theory for sampling species abundance. Ecology Letters., 10: 1037-1045.

GRØTAN, V & ENGEN, S. 2008. poilog: Poisson lognormal and bivariate Poisson lognormal distribution. Pacote do R versão 0.4.

HANKIN, R.K.S. 2007. Introducing untb, an Rpackage for simulating ecological drift under the unified neutral theory of biodiversity. Journal of Statistical Software., 22(12).

MACARTHUR, R. H. 1957. On the relative abundance of bird species. *Proceedings of National Academy of Science*, 43, 293-295.

MACARTHUR, R. H. & MACARTHUR, A. T. 1974. On the use of mist nets for population studies of birds. Proc. Natl. Acad. Sci. USA 71:3230-3233.

MAGURRAN, A.E. 2004. Measuring Biological Diversity . *Blackwell Science, Oxford.*

MARTENSEN, A. C. Conservação de aves de sub-bosque em paisagens fragmentadas: Importância da cobertura e da configuração do habitat. Dissertação (Mestrado). Instituto de Biociências da Universidade de São Paulo. 160p.

MCGILL, B.J. 2010. Towards a unification of unified theories of biodiversity. Ecology Letters. 13 (5): 627-642.

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 (in press).

MOTOMURA, I. 1932. A statistical treatment of associations, Jpn. J. Zool. 44: 379–383.

OKSANEN, J.; BLANCHET, G.; KINDT, R.; LEGENDRE, P.; O'HARA, R.B.; SIMPSON G.L.; SOLYMOS, P.; STEVENS, M.H. & WAGNER, H. 2010. Vegan: Community Ecology Package. Pacote do R versão 1.17-2. URL http://CRAN.R-project.org/package=vegan

PIELOU, E. C. 1977. *Mathematical Ecology*. New York: John Wiley and Sons.

PRESTON, F.W. 1948.The commonness, and the rarity, of species. Ecology. 29: 254 – 283.

R DEVELOPMENT CORE TEAM. 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org .

REMSEN, J. V. & GOOD, D.A. 1996. Misuse of data from mist-net captures to assess relative abundance in bird populations. The awk. 113(2):381-398.

ROYLE, J. A. 2004. N-mixture models for estimating population size from spatially replicated counts. Biometrics 60:108–115.

ROYLE, J. A. & DORAZIO, R. M. 2008. Hierarchical modeling and inference in ecology:The analysis of data from populations, metapopulations and communiteis.Academic Press. San Diego, CA. 444 p.

POLLOCK, K. H., J. D. NICHOLS, C. BROWNIE & J. E. HINES. 1990. Statistical inference for capture-recapture experiments. Wildlife Monographs 107:1–97.

WILLIAMS, B. K., NICHOLS, J. D., & CONROY, M. J. 2001. Analysis and management of animal populations: modeling, estimation, and decision making. Academic Press.

Capítulo 3 *– Aplicação do modelo de distribuição de abundância série logarítmica na avaliação de efeitos de fragmentação sobre a comunidade de aves em paisagens fragmentadas de remanescentes da Mata Atlântica*

1. Introdução

O processo de desmatamento se dá pela perda e fragmentação do habitat, em que a paisagem de mata contínua dá lugar a uma paisagem formada por trechos pequenos e isolados de florestas (With 1997; Fahrig 2003). Os estudos dos efeitos da fragmentação estão fundamentados na teoria de Biogeografia de Ilhas de MacArthur e Wilson (1967), para a qual o tamanho das ilhas e a distância do continente são as principais variáveis preditoras da riqueza de espécies (Ewers & Didham 2006). Assim, em uma correspondência direta entre as ilhas terrestres (fragmentos) e as ilhas oceânicas, a principio a área e, mais tarde, a conectividade passaram a ser as principais grandezas para indicar o grau de fragmentação, e por isso fundamentais em diagnósticos de efeito da fragmentação sobre a presença das espécies (Hanski & Simberloff 1997 ; Metzger 2000). Nos estudos de efeitos de fragmentação, portanto, a redução do tamanho dos fragmentos acaba por ser associada à diminuição do tamanho de algumas populações, e por isso relacionada às probabilidades de extinção das espécies (Jules 1998). Aliado ao tamanho, a conectividade, ou grau de isolamento dos fragmentos, é considerada uma medida de permeabilidade do fragmento focal a indivíduos de fora (Bowman et al. 2002) e por isso associada às probabilidades de colonização e recolonização dos fragmentos (Hanski & Simberloff, 1997).

Entretanto com o acúmulo de estudos empíricos dos efeitos de fragmentação, algumas falhas foram encontradas na lógica da correspondência entre fragmentos e ilhas oceânicas (Watling & Donnelly 2006). A matriz em geral não é tão hostil nem tão homogênea como um oceano para espécies terrestres, e por isso ela permite maior ou menor trânsito de indivíduos dependendo da sua qualidade e complexidade, e também dos atributos das espécies componentes dos fragmentos (Ricketts 2001). Considerando, então, as atuais críticas sobre o uso isolado de métricas de tamanho de área e isolamento como preditores da fragmentação florestal (Fisher & Lindenmayer 2007), foram propostos os limiares teóricos de porcentagem de habitat como uma maneira de inserir o efeito da mudança conjunta destas duas variáveis, que ocorre no processo de fragmentação (Fahrig 2001). A proposta dos limiares teóricos estabelece que abaixo de aproximadamente 20-30% de habitat nativo na paisagem, a área dos fragmentos é a variável mais influente para a persistência das espécies da comunidade, e acima de 30%, a conectividade dos fragmentos passa a ser o melhor

preditora (Andrén 1994; Fahrig 2003).

Se, por um lado, o tamanho dos fragmentos e a conectividade foram constantemente utilizados como variáveis preditoras dos efeitos do processo de fragmentação sobre a biota da paisagem, por outro, a riqueza se tornou a variável resposta mais utilizada para representar a diversidade (e.g. Boulinier et al. 2001, Pidgeon et al. 2007, Martensen 2008). Entretanto, a diversidade e uma grandeza teórica também composta pela dominância, que pode ser descrita pelas diferenças nas abundância das espécies (Magurran 2004). Uma maneira mais completa de representar as duas dimensões da diversidade é por meio da distribuição de abundância das espécies (SAD[2], William et al. 2001). Tal representação revelou um dos padrões mais bem estabelecidos da ecologia de comunidades que é a alta dominância, portanto, que a maioria das espécies é rara (Magurran 2004). A recorrência deste padrão nas comunidades de diferentes grupos taxonômicos em diferentes tipos de ambientes estimulou a proposição de vários modelos teóricos na tentativa de encontrar os mecanismos ecológicos subjacentes (Pielou 1977).

No entanto, os modelos de SAD além de representarem hipóteses alternativas dos possíveis mecanismos que estruturam as comunidades, também são descritores matemáticos da hierarquia de abundâncias das espécies que uma comunidade pode ter. Em uma abordagem fenomenológica, os modelos podem ser utilizados como medidas comparativas para estabelecer relações entre comunidades, independentes dos mecanismos que os geraram (Pielou 1977). Usados desta maneira, os modelos de SADs são úteis para detectar mudanças nas comunidades relacionadas a impactos ambientais tais como poluição (Gray 1979) e fragmentação (Mouquet & Loureau 2003).

Nesta mesma linha fenomenológica, embora se possa investigar o modelo que melhor se ajusta a cada comunidade, é possível também escolher somente um só modelo e avaliar as mudanças de forma do mesmo em diferentes comunidades (Pielou 1977). O que define a forma das distribuições previstas pelos modelos são os seus parâmetros. Alguns desses parâmetros podem ser interpretados como medidas de dominância ou índices de diversidade, e por isso podem ser utilizados como variáveis-resposta pra representar a diversidade nos estudos de fragmentação (e.g. Cayuela et al. 2006; Yaacobi et al. 2007). Um dos modelos melhor estabelecidos e mais amplamente utilizados na ecologia é a série logarítmica de Fisher (Hubbel 2001; Magurran 2004) originalmente proposto para descrever a SADs em amostras de comunidades abertas. A série logarítmica tem um único parâmetro, chamado α de Fisher,

[2] Utilizaremos essa sigla para designar as distribuições de abundâncias das espécies em uma comunidade, ou amostra dela tomada, em correspondência a SAD (species abundance distributions), que é o termo tradicionalmente utilizado na literatura.

que define a taxa de acréscimo de espécies à medida que há um incremento de indivíduos na amostra (Fisher 1943). Por isso, o α de Fisher e é amplamente aceito como um índice de diversidade (Southwood 1984). O α é reconhecido como um bom índice de diversidade por considerar a distribuição de abundância das espécies (Magurran 2004), e ser teoricamente independente do tamanho amostral (Taylor 1978) A independência do esforço amostral de dá porque o modelo série logarítmica já inclui o efeito da amostragem na sua concepção, cuja solução, ao contrário dos demais modelos de SAD que também incorporaram o processo amostral, não requer o conhecimento da fração amostrada da comunidade. Apesar do α ser o parâmetro de um modelo de SAD, há evidências de que pode ser utilizado para medir a diversidade de espécies mesmo quando a série logarítmica não é o modelo que melhor descreve os dados, (Taylor 1978; Hayek & Buzas 1997).

Há duas maneiras, portanto, de se utilizar os modelos de SAD para detectar os efeitos de impactos ambientais, uma em que se avalia a variação dos modelos ajustados entre comunidades e outra que se escolhe um modelo, por critérios práticos, como ter poucos parâmetros ou com parâmetros que tenham interpretação biológica simples, para analisar as variações dos valores estimados dos mesmos. Em uma abordagem de seleção de modelos (Burnham & Anderson 2002), os modelos concorrentes podem ser de diferentes distribuições teóricas de probabilidade, ou a mesma distribuição, mas com valores de parâmetros diferentes. Entretanto, a aplicação dos modelos de SAD para avaliar efeitos de impactos ambientais não considera a identidade das espécies. Se por um lado essa característica é vantajosa quando se quer comparar comunidades com composições muito distintas, como marinhas e terrestres (McGill et al. 2007), por outro ela leva à perda da informação da variação de posição de cada espécies no ranking de abundância em uma comunidade em diferentes momentos (Mac Nally 2007) ou então entre comunidades similares. Uma maneira de se considerar a identidade das espécies, mantendo a sinteticidade da abordagem do uso de ranking de abundância, é, além de investigar variações na forma das SADs, acompanhar a variação da posição de cada espécie na hierarquia de abundância. Mac Nally (2007) defende que esse acompanhamento da posição das espécies é essencial em investigações de efeitos de mudanças de origem antrópica em paisagens. A aplicação desta abordagem se dá, de maneira geral, ordenando as espécies de maior abundância até a menos abundante em uma comunidade de referência para fixar a posição das espécies no ranking e a partir daí acompanhar mudanças nas suas abundâncias sob a influência do efeito estudado. A comunidade de referência pode ser da comunidade de uma área antes de sofrer um impacto a fim de avaliar seus efeitos, ou então, de uma comunidade controle.

O nosso objetivo neste trabalho foi identificar o efeito da fragmentação sobre a diversidade de aves e sobre as posições das espécies na hierarquia de abundâncias em paisagens fragmentadas de remanescentes da Mata Atlântica. Para isso, usamos as capturas de aves por procedimentos padronizados feitas em fragmentos e áreas controle de floresta contínua, com características semelhantes distribuídos em três paisagens com graus crescentes de perda de habitat na região do Planalto Atlântico paulista. Para representar a diversidade escolhemos o parâmetro do modelo de distribuição de abundância relativa série logarítmica, cujos valores estimados foram comparados entre fragmentos de diferentes tamanhos e graus de isolamento. Como o parâmetro da série logarítmica é o índice de diversidade α, se a fragmentação tem um efeito negativo sobre a diversidade, esperamos que com a redução da área e o aumento do isolamento dos fragmentos, os valores de α sejam gradativamente menores. Para avaliar as conseqüências da fragmentação sobre a posição das espécies na hierarquia de abundâncias, procedemos com uma descrição gráfica dos dados. Nessa descrição, posicionamos as espécies em ordem decrescente de número de capturas nas amostras de áreas de floresta contínua e observamos se as posições das mesmas se mantinham nas amostras dos fragmentos de diferentes tamanhos.

2. Material e Métodos

2.1 Áreas de coleta

Os dados utilizados para as análises dos efeitos de fragmentação são provenientes do Projeto "**Conservação da biodiversidade em paisagens fragmentadas**" , coordenado pelos Drs. Jean Paul Metzger e Renata Pardini (IB-USP), financiado atualmente pelo Programa Mata Atlântica (CNPq/Alemanha), Auxílio Jovem Pesquisador FAPESP 05/56555-4, e anteriormente pelo Auxílio Temático FAPESP 99/05123-4. As coletas foram realizadas em três áreas do Planalto Atlântico Paulista (Figura 1) e foram inventariados além de aves, outros grupos taxonômicos, com o objetivo inicial de se testar questões relacionadas aos limiares teóricos de fragmentação.

As áreas de estudo apresentam, portanto, condições abióticas similares, porém diferentes graus de fragmentação, em um gradiente com três níveis de cobertura vegetal da paisagem, se estendendo abaixo, no limiar e acima do limiar teórico de fragmentação, que é de 30% de cobertura vegetal. (Metzger 2006). As paisagens escolhidas foram áreas rurais nos municípios de Ribeirão Grande (24° 07' S, 48° 24' W), com 10% de cobertura vegetal;

Caucaia (23° 47' S, 47° 07'W) com 30% e Tapiraí (23° 50' S, 47° 20'W) com 50% de cobertura vegetal (Figura 1). Além dos fragmentos foram amostradas também, nas três paisagens, áreas de floresta contínua, adjacentes aos fragmentos como representativas de situações controle ausentes dos efeitos da fragmentação. As áreas de floresta contínua estão dentro de uma propriedade privada, uma Reserva Florestal e um Parque Estadual e representam a situação controle para as paisagens de 10%, 30% e 50% de cobertura vegetal, respectivamente. Para as análises dos efeitos de fragmentação sobre a diversidade foram consideradas apenas as amostras feitas nos fragmentos, excluindo as coletas nas áreas de floresta contínua.

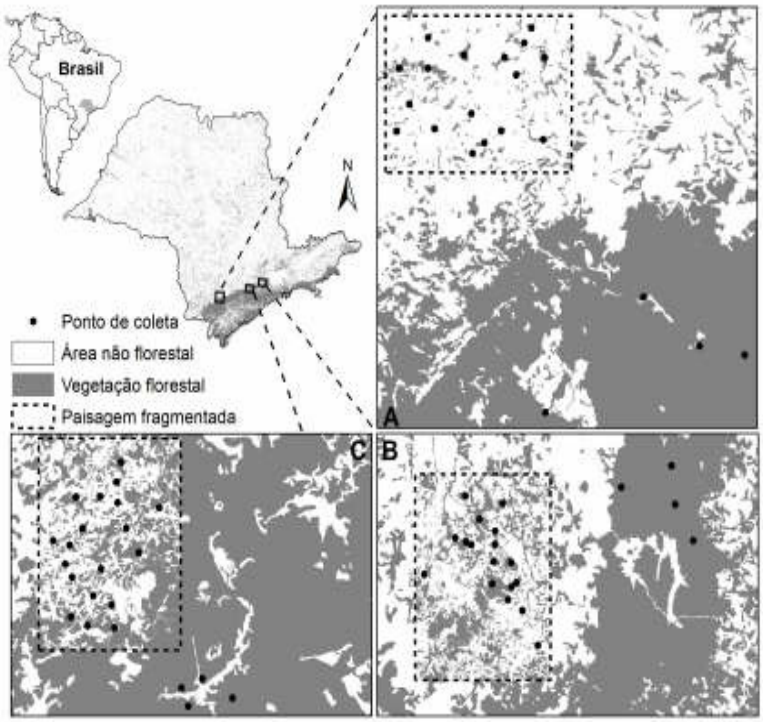

Figura 1. Localização das paisagens estudadas em três níveis de porcentagens de cobertura vegetal: (A) Ribeirão Grande com 10%; (B) Caucaia com 30% e (C) Tapiraí com 50% de cobertura vegetal. A área dos retângulos corresponde às áreas amostradas e os pontos de coleta correspondentes. As áreas cinza contínuas com quatro pontos de coleta, fora dos retângulos, são áreas de floresta contínua amostradas para representar a situação controle, sem os efeitos de fragmentação.

A vegetação original da região das três áreas de estudo é classificada como floresta ombrófila densa montana (Veloso *et al.*, 1991). Atualmente, a vegetação nativa é formada por um mosaico de manchas em estádios avançados e intermediários de sucessão. As altitudes das

três paisagens variam entre 800 e 1100 metros acima do nível do mar, com temperaturas entre 15° e 25°C e precipitações médias entre 1221 e 1808 mm ao ano (Centro de Pesquisas Meteorológicas e Climáticas Aplicadas a Agricultura: www.cpa.unicamp.br).

2.2 Coleta de dados

A coleta dos dados ocorreu sob a responsabilidade dos pesquisadores Cristina Banks-Leite e Alexandre Camargo Martensen durante seus projetos de doutorado e mestrado, respectivamente, que resultaram nas teses finais: **"Importância da cobertura e da configuração do hábitat"** (Martensen 2008) e **"Conservação das aves de sub-bosque em áreas fragmentadas da Floresta Atlântica"** (Banks-Leite 2009). As coletas ocorreram durante os anos de 2001 e 2007 com duas campanhas por semestre: duas vezes na estação seca e duas vezes na estação chuvosa.

As amostragens das aves foram feitas pelo método de rede de neblina, com 53 pontos amostrais distribuídos por fragmentos florestais das três paisagens. Na paisagem de Ribeirão Grande foram amostrados 17 pontos localizados em fragmentos florestais em três níveis de tamanho: pequenos (4 a 9 ha), medianos (11 a 37 ha) e grandes (43 a 92 ha). Na paisagem de Caucaia foram amostrados 17 pontos também em fragmentos pequenos (2 a 9 ha), medianos (14 a 52 ha) e grandes (56 a 158 ha) e, por fim, na paisagem de Tapiraí, foram amostrados 19 pontos nos fragmentos florestais em fragmentos de tamanho pequeno (2 a 5 ha), mediano (18 a 41 ha) e grandes (78 a 156 ha).

Em cada ponto amostral foram colocadas 10 redes de neblina (12 m de comprimento, 2,5 de altura e 31 mm de malha, cada rede), dispostas em linha reta. A média de horas/rede foi de 637h/rede (Desv. Pad.= 76,7). As aves capturadas foram marcadas com anilhas cedidas pelo Centro Nacional de Pesquisa para a Conservação de Aves Silvestres - CEMAVE (registro junto ao CEMAVE: 522117), sendo as recapturas registradas.

2.3 Variável dependente: α de Fisher

Para comparar as SADs de cada fragmento, utilizamos o modelo série logarítmica. Escolhemos o modelo série logarítmica porque ele: (a) considera o processo amostral que opera sobre a SAD das comunidades para gerar as abundâncias observadas; (b) possui apenas um parâmetro, que teoricamente é independente do tamanho da amostra (Fisher *et al.* 1943); (c) ao contrário dos demais modelos que consideram o processo amostral, não exige o conhecimento do parâmetro de fração da comunidade amostrada, cuja estimativa não é viável

com o conjunto de dados utilizados. Embora, a série logarítmica, como os demais modelos de SAD, assuma a premissa de que as espécies são igualmente capturáveis; como apresentado nos dois capítulos anteriores, ela ainda pode ser uma indicadora da sensibilidade da diversidade ao processo de fragmentação. Como mostramos no capítulo dois desta tese, mesmo que as capturabilidade das espécies sejam diferentes, se existe o efeito da fragmentação sobre a forma da série logarítmica, ele será detectado. E mesmo que sob capturabilidades distintas entre as espécies possam levar a falsa detecção do efeito, o efeito estimado é sempre de baixa intensidade, o que torna possível avaliar pelos valores estimados dos efeitos se houve ou não. Assim, é possível, mesmo desconhecendo as capturabilidades das espécies, avaliar o poder de inferência dos resultados e tirar conclusões sobre a presença ou não dos efeitos da fragmentação.

Quando Fisher propôs a série logarítmica, ele a descreveu como a distribuição de amostras retiradas das comunidades, representada por uma distribuição Poisson, cujo parâmetro de média segue uma distribuição Gama, resultando na equação de uma binomial negativa:

$$n_r = \frac{(k+r-1)!}{(k-1)!\,r!} \; X \; \frac{p^r}{(1+p)^{k+r}} \quad \mathrm{i} \geq 1 \qquad (1)$$

, que representa a probabilidade de uma espécie ter *r* indivíduos na amostra, sendo *k* o parâmetro de agregação e *p*, o parâmetro de representabilidade da amostra. Quando a diferença entre as abundâncias das espécies é muito alta, k diminui e se aproxima de zero. Em comunidades altamente diversas, k tende a zeroUtilizando os dados de capturas de borboletas de Corbet, Fisher percebeu que havia sempre uma grande dominância de poucas espécies, e que ajustando a binomial negativa aos dados, *k* torna-se indistinguível de zero (Fisher et al. 1943). Assim, Fisher tendeu o parâmetro k a zero e formulou sua série logarítmica como uma forma limite da distribuição binomial negativa para comunidades com muitas espécies. Já o termo da binomial negativa dependente do parâmetro p , $p/(p+1)$, foi redefinido como uma nova constante, X, e o termo (*k*-1)! como o parâmetro α, ambos solúveis a partir dos valores do número de indivíduos e de espécies na amostra (para mais detalhes ver parte III de Fisher et al 1943). Tais soluções originalmente deduzidas por Fisher ,constituídas por um sistema de duas equações (eq. 2 e eq.3) resolvidas numericamente, foi demonstrado mais tarde por Alonso et al. (2008) como sendo a solução de máxima verossimilhança para α.

$$X = \frac{N}{N+\alpha} \qquad (2)$$

, em que N é o número de indivíduos na amostra;

$$\alpha = \frac{-S}{\log(1-X)} \qquad (3)$$

, em que S é o número de espécies amostradas.

.A partir daí, Alonso et al. (2008) também deduziram a distribuição de densidade probabilística da série logarítmica (equação 3 do capítulo um). A série logarítmica estabelece uma relação entre número de indivíduos e numero de espécies::

$$S = \alpha \ln\left(1 + \frac{N}{\alpha}\right) \qquad (4)$$

No modelo série logarítmica, não há um limite ao número de espécies, que depende apenas do parâmetro α e do número de indivíduos. Portanto, é um modelo adequado para descrever comunidades abertas, que sempre revelam mais espécies à medida que incluímos mais indivíduos na amostra. Considerando, então, que a solução de Fisher para o cálculo de α resulta em uma estimativa de máxima verossimilhança, e que as estimativas de máxima verossimilhança tem como propriedade a normalidade assintótica (ver Royle & Dorazio 2008), utilizamos os valores estimados de α das amostras de cada fragmento como variáveis dependentes em modelos de regressão linear simples para avaliar os efeitos de fragmentação.

2.4 Variáveis independentes: Área, Proximidade

As variáveis preditoras do valor de α para cada ponto foram as métricas de área e isolamento, AREA e PROX, respectivamente, dos fragmentos amostrados, disponíveis no banco de dados do projeto "**Conservação da biodiversidade em paisagens fragmentadas**", calculados para todos os fragmentos amostrados e definidas como:

- Área (AREA): area total do fragmento em hectares.
- Proximidade (PROX): é uma medida de isolamento, calculado com a somatória das razões entre as áreas dos fragmentos vizinhos contidos, integral ou parcialmente, em um determinado raio ao redor do fragmento focal, e o quadrado das distâncias do fragmento focal aos vizinhos. Escolhemos como valor de raio ao redor do fragmento, 50 metros.

As duas métricas foram escolhidas para representar o processo de fragmentação em

termos de tamanho e isolamento de habitat dentro do que se acredita ser escala de percepção da paisagem pelas aves de sub-bosque.

2.5 Análise de dados

Para avaliar se há variação da diversidade em função da perda crescente de área e aumento do isolamento, construímos três modelos de regressão linear simples: um modelo nulo sem efeito e dois com efeitos das métricas do fragmento (ver abaixo). Assim, nos modelos, α é uma variável aleatória gaussiana, cuja esperança pode variar em função de uma das variáveis preditoras (modelos com efeito), ou então, ter um valor esperado constante e igual entre os fragmentos (modelo sem efeito). Nos modelos com efeito as variáveis preditoras foram transformadas em logaritmos em base natural (ln). Para a variável preditora de isolamento (PROX), adicionamos uma unidade nos valores da métrica antes da transformação logarítmica, uma vez que a variável pode assumir o valor zero. Para cada paisagem, os modelos foram ajustados minimizando-se a soma dos quadrados dos desvios, que é a solução de máxima verossimilhança para modelos lineares Gaussianos. Uma vez ajustados, os modelos foram comparados pelo protocolo de seleção de modelos de Burnham & Anderson (2002). Os modelos candidatos para as análises em cada paisagem foram os seguintes:

- ✓ $E[\alpha] = a$
- ✓ $E[\alpha] = a + b*\ln(\text{AREA})$
- ✓ $E[\alpha] = a + b*\ln(\text{PROX} + 1)$

, sendo *a* e *b* parâmetros estimados correspondentes ao intercepto e coeficiente angular, respectivamente.

Na seleção de modelos é feita uma classificação hierárquica dos modelos de acordo com os critérios de informação de Akaike, AIC (Akaike 1973). Para o cálculo de AIC, primeiro construímos a função de log-verossimilhança negativa do modelo (equação 1 do capítulo de Apresentação) e atribuímos valores variados para os parâmetros a fim de minimizar o valor da função. Quando a função atinge seu valor mínimo obtemos os valores de verossimilhança dos modelos, e também as estimativas de máxima verossimilhança negativa dos parâmetros (intercepto *a* e inclinação *b*). A partir dos valores de verossimilhança dos modelos podemos calcular o AIC que é:

$$AIC = 2LL(\theta) + 2\text{k} \qquad (6)$$

, em que LL(θ) é o valor de log-verossimilhança negativa do modelo, θ as estimativas de máxima verossimilhanças dos parâmetros, e k o número de parâmetros. O AIC pode ser convenientemente descrito como uma medida de plausibilidade do modelo penalizado pelo número de parâmetros. Quanto menor o valor de AIC, mais plausível é o modelo. Quando a diferença de valores de AIC dos modelos de menor AIC e de outro modelo (Δ AIC) era menor que dois, consideramos empate entre os mesmos (para mais detalhes de ajuste e seleção de modelos, ver capítulo de Apresentação).

2.6. Variação nas abundâncias das espécies nos fragmentos e areas controle

Para avaliar mudanças nas posições das espécies na hierarquia de abundância, utilizamos o procedimento gráfico proposto por Mac Nally (2007), designado pelo autor como "abundance spectrum". Para tanto, para cada paisagem, posicionamos todas as espécies amostradas em ordem decrescente de número de capturas presentes nas amostras da área de floresta contínua, que é a área controle. Determinadas as posições das espécies nessa ordem, plotamos, para cada fragmento e área controle, a proporção de número de captura de cada espécie. Além das informações de proporção de capturas de cada espécie, diferenciamos as espécies de acordo com o grau de sensibilidade a perturbação, seguindo a classificação de Stotz et al. (1996).

Todas as rotinas de otimização, análises e gráficos foram executadas no programa R versão 2.11.1 (R Development Core Team 2010). Além das funções implementadas por nós, utilizamos funções do pacote vegan (Oksanen et al. 2010) e bbmle (Bolker 2010), já implementadas em R.

3. Resultados

Descrição dos dados

Os conjuntos de dados registram 4818 indivíduos, sendo 1952 deles capturados na paisagem de 10% (Ribeirão Grande), 1293 na paisagem de 30% (Caucaia do Alto) e 1573 na paisagem de 50% (Tapiraí) de cobertura vegetal. Ao total foram amostradas 117 espécies: 70 na paisagem de 10%, 62 na de 30% e 87 na paisagem de 50% de cobertura vegetal. Entre as espécies, 55 espécies (aproximadamente 47%) foram capturadas em apenas uma das paisagens, sendo 18 exclusivas da paisagem de 10%, 8 exclusivas da paisagem de 30% e 29

exclusivas da paisagem de 50%. A média de espécies capturadas por fragmento na paisagem de 10% de cobertura vegetal foi de 26,29 (Desv. pad. = 4,5); já na paisagem com 30% foi de 25,12 (Desv. pad= 4,75) e finalmente na paisagem de 50% a média de espécies amostradas foi de 29,37 (Desv. pad= 9,21). O valor médio em α na paisagem de 10% de cobertura vegetal foi de 10,84 com valores entre 8,3 e 15,83. Na paisagem de 30%, a média dos α foi de 13,83 e teve valores entre 7,34 e 20,45. Por fim, na paisagem de 50% de cobertura vegetal, os α tiveram média de 16,35 e variaram entre 6,88 e 30,84.

Efeitos do tamanho e isolamento sobre a diversidade

Para a paisagem mais fragmentada (10% de cobertura vegetal), o modelo mais plausível foi aquele com efeito de área sobre o valor esperado do parâmetro α de Fisher, sendo o efeito esperado da área sobre a diversidade, positivo (tabela 1). Já para a paisagem com nível intermediário de cobertura vegetal (30% de cobertura vegetal), os três modelos foram igualmente plausíveis, sendo o modelo sem efeito das métricas o de menor AIC, sempre com um efeito positivo das métricas sobre a diversidade (tabela 2). Para a paisagem de maior cobertura vegetal (50%), embora o modelo com efeito de isolamento tivesse o menor valor de AIC, ele empatou com o modelo sem efeito das métricas. Nesta última paisagem, os efeitos estimados das métricas de isolamento e área sobre a diversidade dos fragmentos, foram positivo e negativo, respectivamente (tabela 3).

De acordo com o modelo com efeito de área para a paisagem de 10% α de cobertura vegetal, a diferença dos valores de α previstos para o maior e menor fragmentos da paisagem foi de 3,4. Isto indica que o efeito da área foi fraco, como mostra a similaridade das distribuições de abundâncias previstas por este modelo e pelo modelo com um único valor médio de α para todos os fragmentos (Figura 2).

Já o coeficiente angular previstopelo modelo de relação de menor AIC (modelo de isolamento), para a paisagem de 30% de cobertura vegetal, foi menor que 0,2 por unidade de ln (PROX +1) (tabela 1), resultando em uma variação de aproximadamente 1,05 em α entre os fragmentos de maior e menor isolamento. Essa baixa variação prevista torna todos os valores de α esperados pelo modelo com efeito muito próximos ao valor médio previsto pelo modelo sem efeito. Assim, as curvas de ajuste da série logarítmica com os valores de α previstos por ambos os modelos na distribuição de rank-abundância das amostras ficam praticamente sobrepostas (Figura 3).

Por fim, de acordo com o modelo com efeito de isolamento para a paisagem de 50% de cobertura vegetal, a diferença entre α dos fragmentos de menor e maior isolamento foi de

7,9. O modelo de isolamento, além de ser o que obteve menor valor de AIC para a paisagem, previu uma variação mais alta nos valores de α do que os modelos com o efeito de isolamento nas demais paisagens. Isso porque além de ter estimado os maiores valores da magnitude do efeito de isolamento, os fragmentos amostrados nessa paisagem apresentam a maior amplitude nas medidas de isolamento, permitindo que as estimativas de α atingissem valores mais altos. A maior amplitude prevista para os valores de α nos fragmentos, foram suficientes para diminuir a sobreposição das curvas de ajuste da série logarítmica com os valores de α previstos pelos modelos com efeito de isolamento, e da série logarítmica com um valor de α médio (Figura 4).

Alterações nas abundâncias das espécies entre fragmentos e áreas contínuas

Na paisagem de 10% de cobertura vegetal, a maior parte das espécies consideradas altamente sensíveis registradas na área controle, que no total foram seis, não foram amostradas nos fragmentos, sendo apenas três delas encontradas nos fragmentos maiores e uma nos fragmentos menores. Das quatro espécies de baixa sensibilidade amostradas na área controle, três foram encontradas em pelo menos 13 dos 17 fragmentos. Adicionalmente, todas as espécies que não registradas na área controle mas presentes nas amostras dos fragmentos são pouco ou medianamente sensíveis a perturbações (Figura 6).

Na paisagem de 30% de cobertura vegetal, das seis espécies altamente sensíveis, três foram amostradas nas áreas fragmentadas. Das espécies amostradas exclusivamente nos fragmentos, nenhuma é considerada altamente sensível. A espécie com o maior número de capturas na área controle, considerada pouco sensível a perturbações, foi amostrada em todos os fragmentos, sendo que na maior parte das vezes ela estava entre as com maior número de capturas.

Finalmente na paisagem de 50% de cobertura vegetal, das 16 espécies altamente sensíveis a perturbações encontradas nas áreas controle, somente cinco foram amostradas nas áreas fragmentadas. Essa paisagem foi a única entre as três paisagens que apresentou espécies altamente sensíveis presentes exclusivamente nas amostras dos fragmentos.

Tabela 1. Tabela de seleção de modelos para a paisagem com 10% de cobertura vegetal (Ribeirão Grande). Os modelos estão em ordem crescente de valores de AIC. **ΔAIC** é a diferença de valor de AIC entre o modelo considerado e o de menor AIC; ***w*AIC** é o peso de evidência do modelo, considerando o conjunto de modelos comparados, dois com efeito: um de área (AREA) e outro de isolamento(PROX); e um sem efeito (CONSTANTE). O coeficiente angular é a estimativa da magnitude do efeito previsto pelo modelo linear gaussiano que estabeleceu uma relação entre a diversidade (α de Fisher) e as métricas de fragmentação

MODELO	**AIC**	**ΔAIC**	***w*AIC**	**Coeficiente angular**
AREA	71,60	0	0,72	1,02
CONSTANTE	74,28	2,7	0,19	-
PROX	75,87	4,3	0,09	0,12

Tabela 2. Tabela de seleção de modelos para a paisagem com 30% de cobertura vegetal (Caucaia). Os modelos estão em ordem crescente de valores de AIC. **ΔAIC** é a diferença de valor de AIC entre o modelo considerado e o de menor AIC; ***w*AIC** é o peso de evidência do modelo, considerando o conjunto de modelos comparados, dois com efeito: um de área (AREA) e outro de isolamento(PROX); e um sem efeito (CONSTANTE). O coeficiente angular é a estimativa da magnitude do efeito previsto pelo modelo linear gaussiano que estabeleceu uma relação entre a diversidade (α de Fisher) e as métricas de fragmentação

MODELO	**AIC**	**ΔAIC**	**Peso de evidência**	**Coeficiente angular**
CONSTANTE	91,16	0	0,56	-
PROX	92,96	1,8	0,23	0,13
AREA	93,08	1,9	0,21	0,15

Tabela 3. Tabela de seleção de modelos para a paisagem com 50% de cobertura vegetal (Tapiraí). Os modelos estão em ordem crescente de valores de AIC. **ΔAIC** é a diferença de valor de AIC entre o modelo considerado e o de menor AIC; ***w*AIC** é o peso de evidência do modelo, considerando o conjunto de modelos comparados, dois com efeito: um de área (AREA) e outro de isolamento(PROX); e um sem efeito (CONSTANTE). O coeficiente angular é a estimativa da magnitude do efeito previsto pelo modelo linear gaussiano que estabeleceu uma relação entre a diversidade (α de Fisher) e as métricas de fragmentação

MODELO	**AIC**	**ΔAIC**	**Peso de evidência**	**Coeficiente angular**
PROX	123,11	0	0,64	0,75
CONSTANTE	124,94	1,8	0,25	-
AREA	126,61	3,5	0,11	-0,57

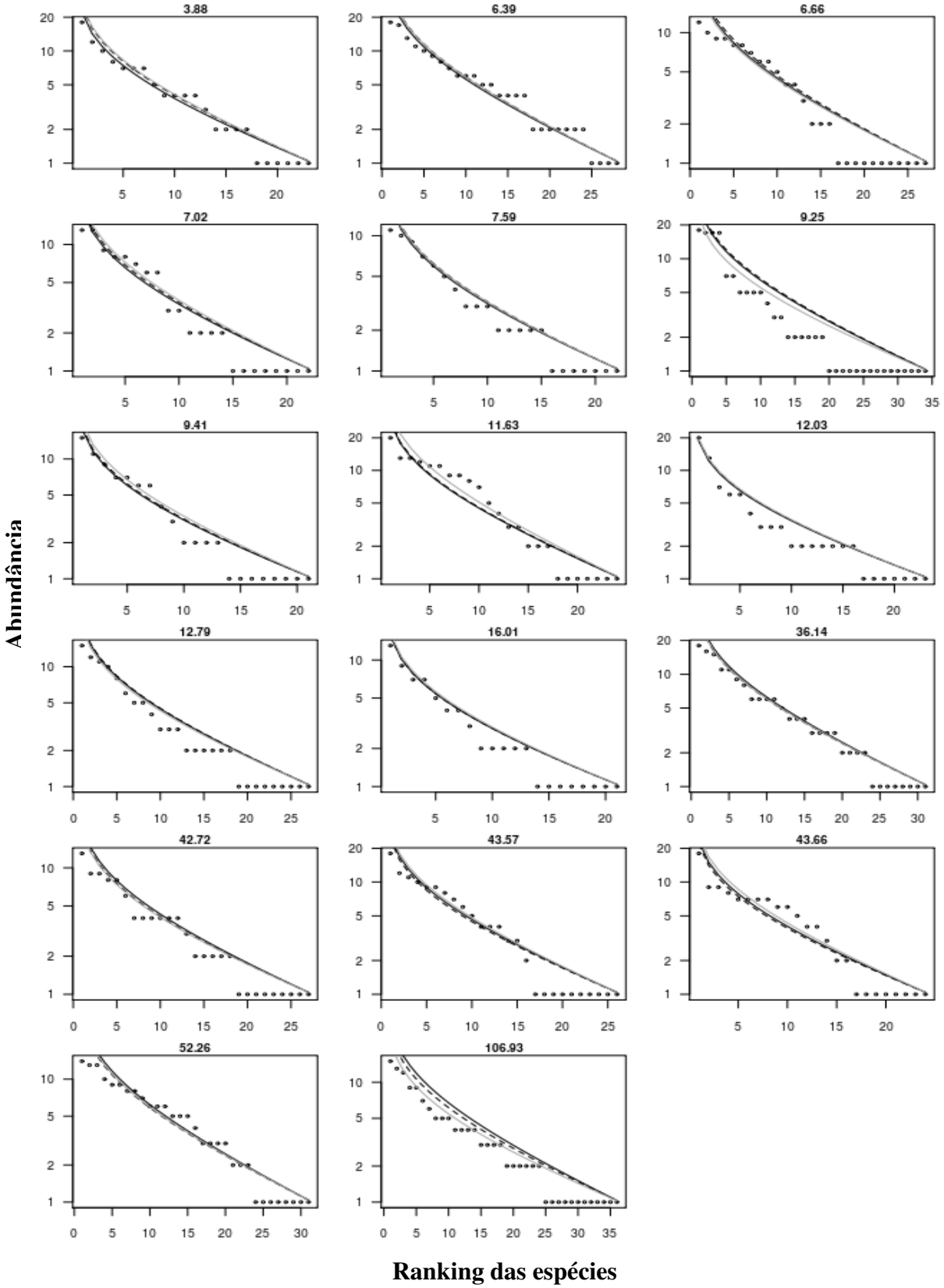

Figura 2. Diagramas de rank-abundância dos 17 fragmentos da paisagem com 10% de cobertura vegetal (Ribeirão Grande), dispostos em ordem crescente de area em hectares, que está indicada acima de cada diagrama. As linhas são as curvas de ajuste da distribuição série logarítmica com os valores de α previstos pelo modelo série logarítmica com valor de α : único para todos os fragmentos (—); variando em função do tamanho do fragmento (- - -); estimado para cada fragmento independentemente (—).

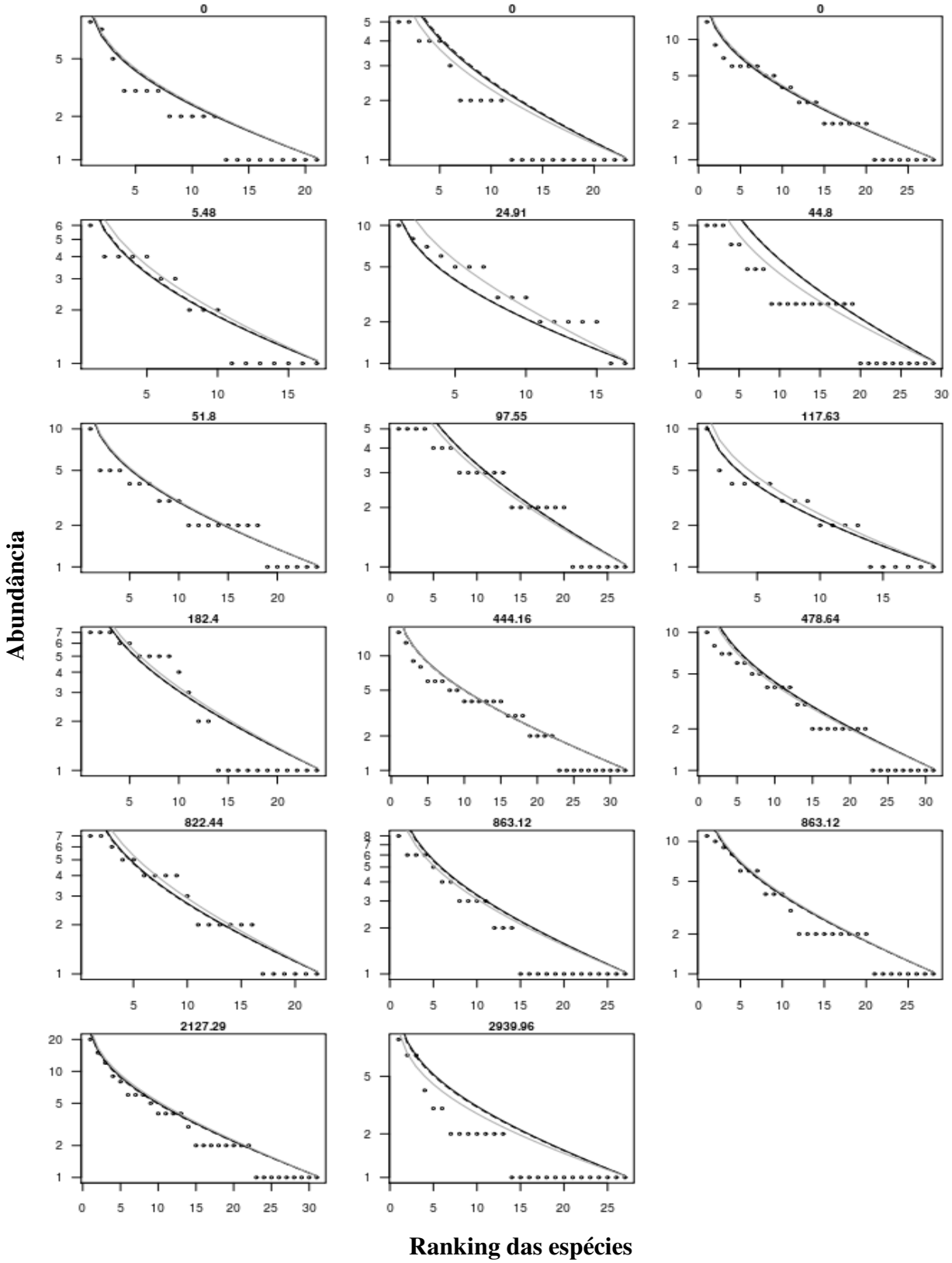

Figura 3. Diagramas de rank-abundância dos 17 fragmentos da paisagem com 30% de cobertura vegetal (Caucaia), dispostos em ordem decrescente de isolamento, que está indicado acima de cada diagrama. As linhas são as curvas de ajuste da distribuição série logarítmica com os valores de α previstos pelo modelo série logarítmica com valor de α : único para todos os fragmentos (—); variando em função do isolamento dos fragmentos (- - -); estimado para cada fragmento independentemente (—).

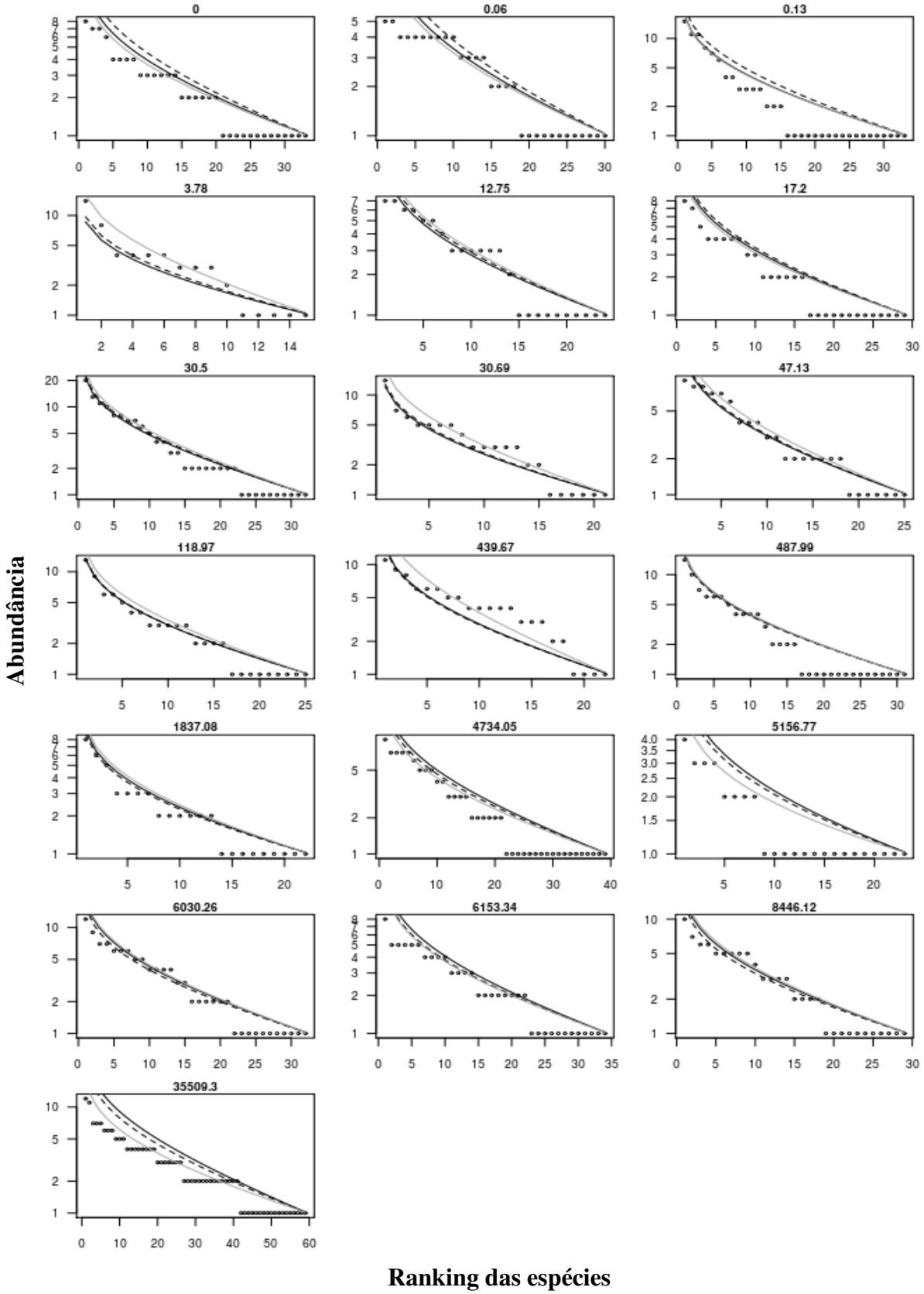

Figura 4. Diagramas de rank-abundância dos 19 fragmentos da paisagem com 50% de cobertura vegetal (Tapiraí), dispostos em ordem decrescente de isolamento, indicado acima de cada diagrama. As linhas são as curvas de ajuste da distribuição série logarítmica com os valores de α previstos pelo modelo série logarítmica com valor de α : único para todos os fragmentos (—); variando em função do isolamento dos fragmentos (- - -); estimado para cada fragmento independentemente (—).

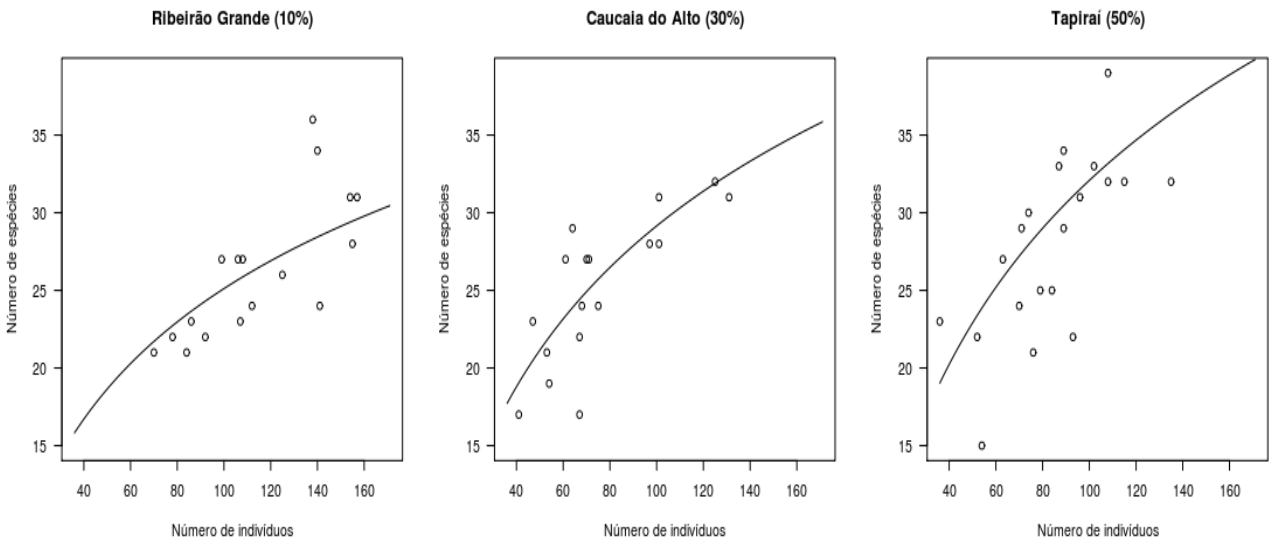

Figura 5. Curva prevista pelo modelo série logarítmica com o valor médio dos α estimados dos fragmentos para a relação de acréscimo do número de espécies em função do número de indivíduos nas amostras para cada uma das paisagens analisadas.

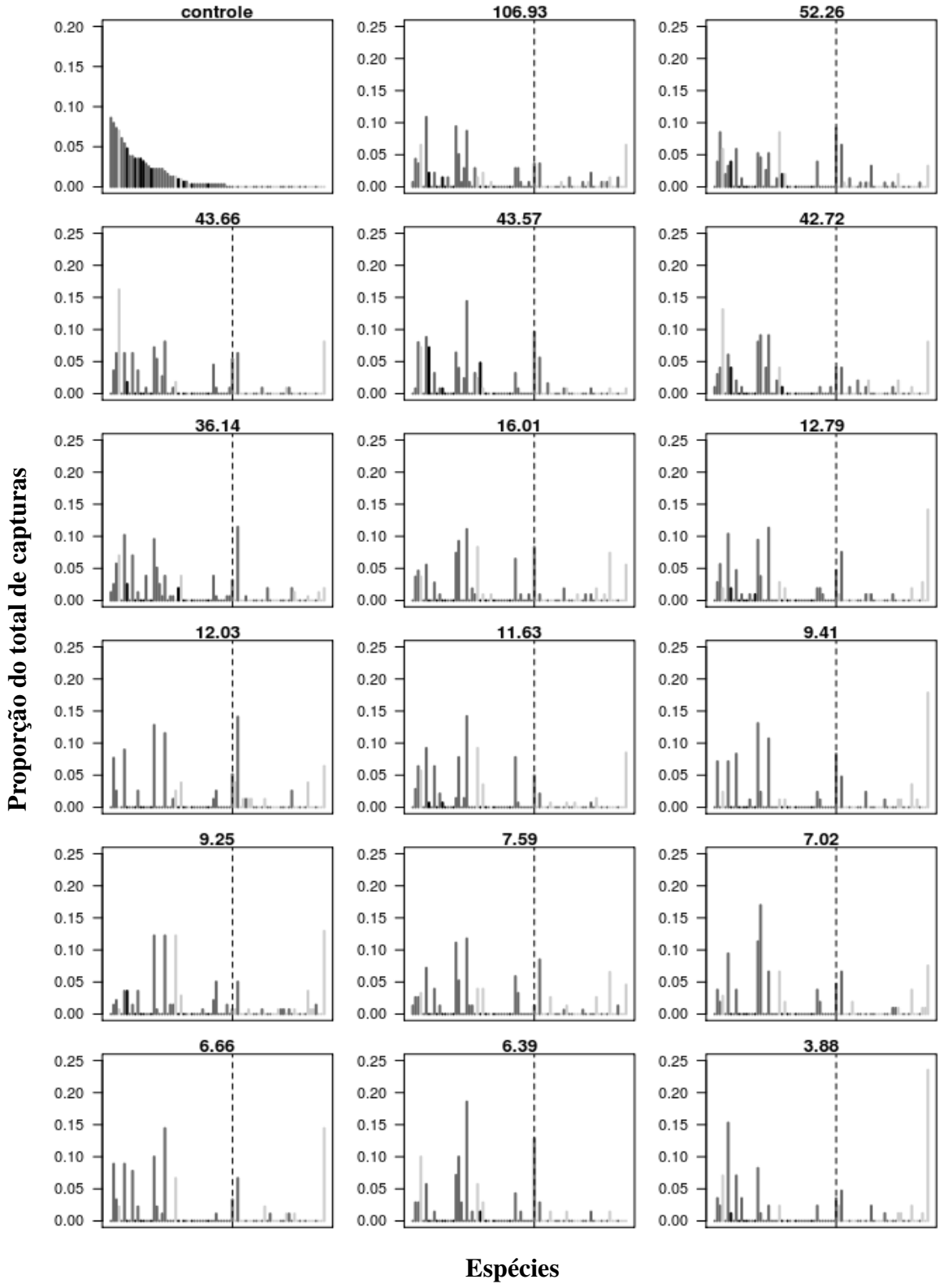

Figura 6. Distribuição do total de capturas relativas em função das espécies amostradas na paisagem de 10% de cobertura vegetal, ordenadas em ordem decrescente de número de capturas obtido nas amostras da área controle. Os gráficos estão ordenados em ordem decrescente de áreade fragmento em hectares, indicada acima de cada diagrama, e as espécies foram classificadas de acordo com o nível de sensibilidade de perturbações: alta ( ); média ( ) e baixa ( )

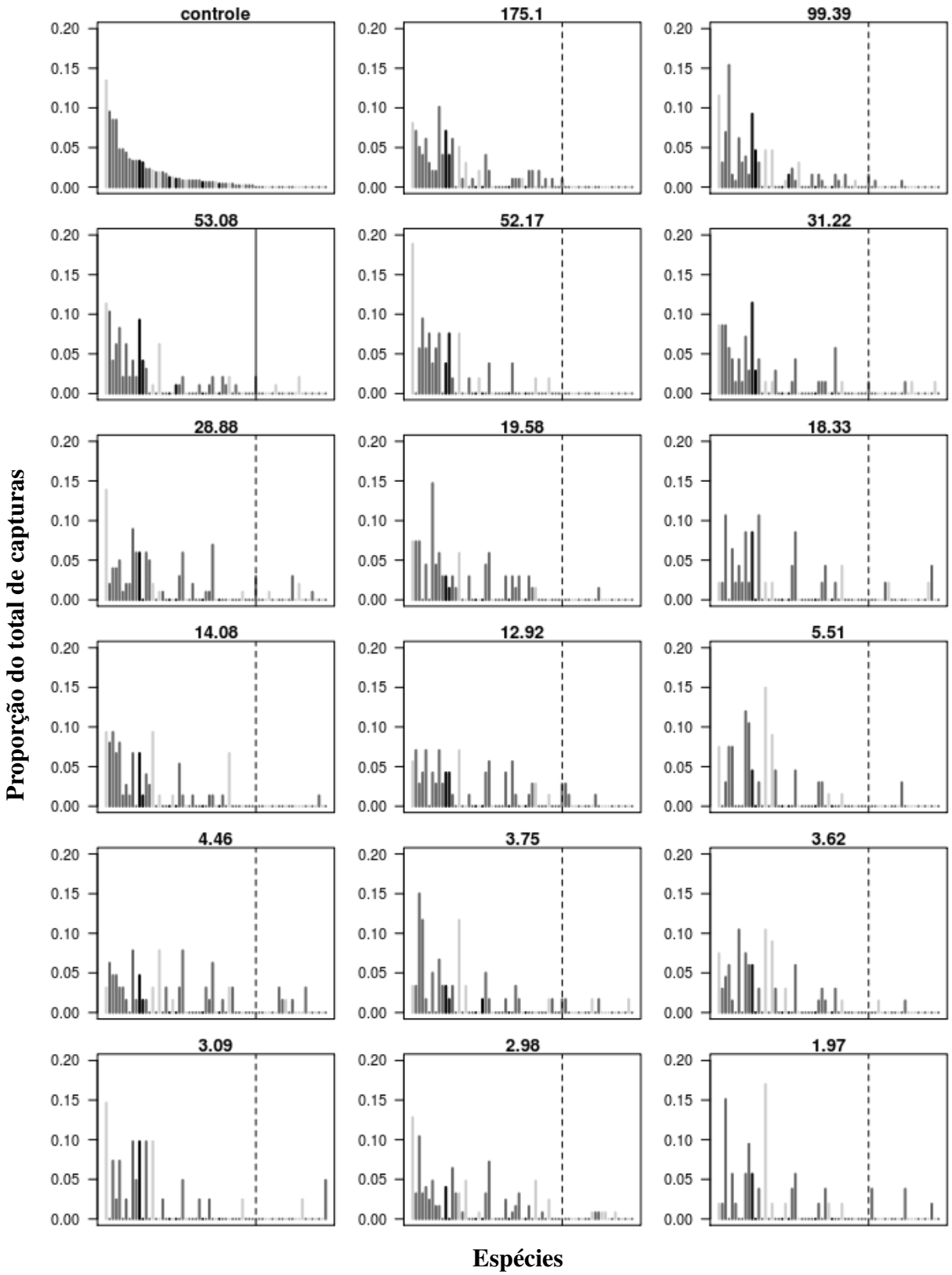

Figura 7. Distribuição do total de capturas relativas em função das espécies amostradas na paisagem de 30% de cobertura vegetal, ordenadas em ordem decrescente de número de capturas obtido nas amostras da área controle. Os gráficos estão ordenados em ordem decrescente de área de fragmento em hectares, indicada acima de cada diagrama, e as espécies foram classificadas de acordo com o nível de sensibilidade de perturbações: alta (▮); média (▯▮)e baixa ▯▯

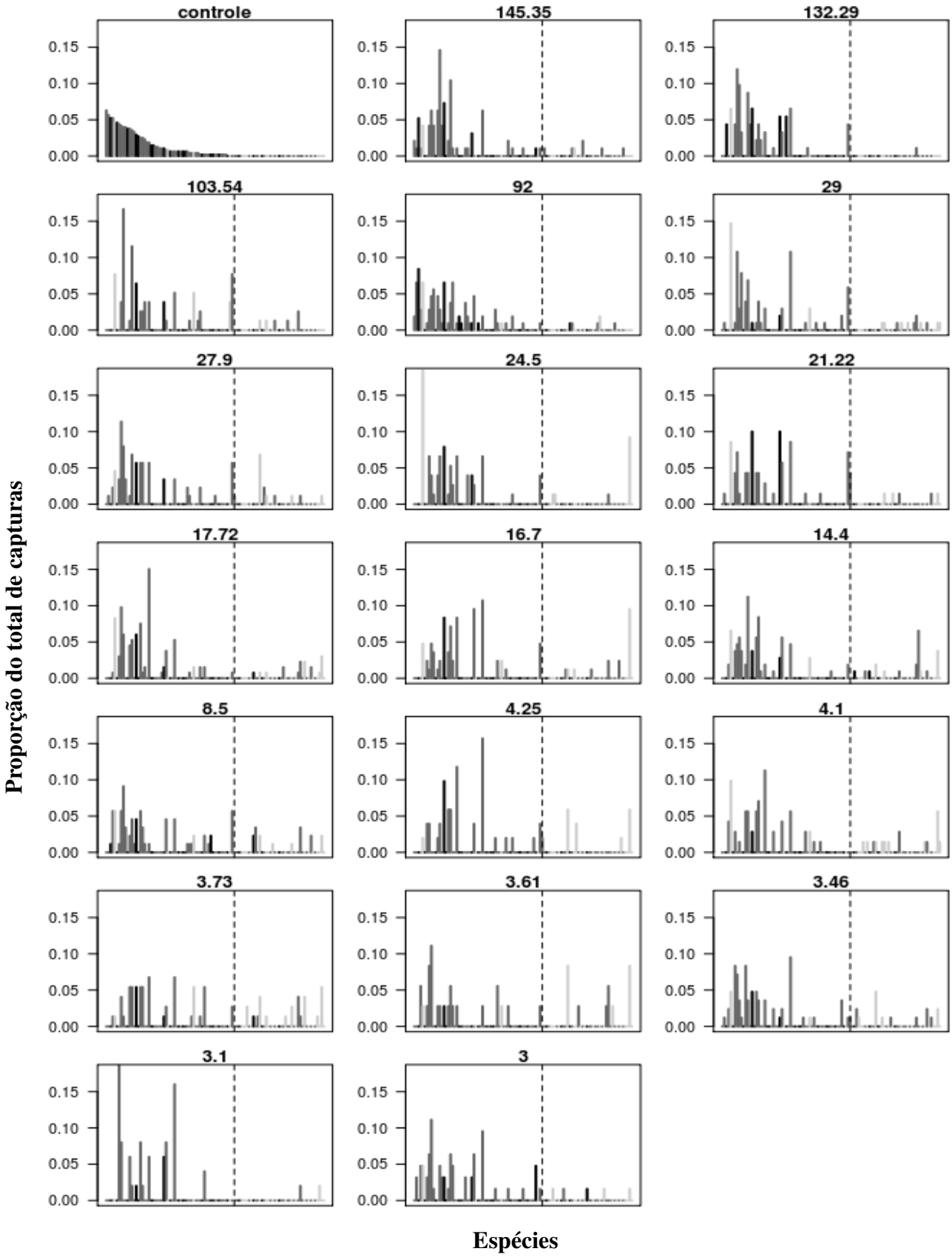

Figura 8. Distribuição do total de capturas relativas em função das espécies amostradas na paisagem de 50% de cobertura vegetal, ordenadas em ordem decrescente de número de capturas obtido nas amostras da área controle. Os gráficos estão ordenados em ordem decrescente de tamanho de fragmento em hectares, e as espécies foram classificadas de acordo com o nível de sensibilidade de perturbações: alta (▮); média (▮) e baixa (▯).

4. Discussão

Nos nossos resultados encontramos que na paisagem abaixo do limiar teórico, de 10% de cobertura vegetal, a área parece influenciar a diversidade dos fragmentos mais que o isolamento, e que na paisagem acima do limiar teórico (50%), a variável de isolamento se torna mais importante que a área para explicar os dados. Apesar dos resultados estarem de acordo com o padrão esperado pela hipótese dos limiares de fragmentação, é preciso ter cautela, pois não sabemos qual o grau de heterogeneidade de probabilidades de captura das espécies para este conjunto de dados. Embora sob altos graus de diferenciação de capturabilidade das espécies a seleção de modelos seja eficiente em detectar a relação entre o α e o gradiente de área quando ela existe, na ausência de relação ela pode ser falsamente detectada (Capítulo 2). No caso da série logarítmica que utilizamos aqui, a chance de haver uma falsa detecção da relação aumenta quando a diferença entre as probabilidades de captura das espécies são maiores. Quando esse viés ocorre, os coeficientes angulares estimados são sempre muito baixos. Sendo conservadores, para considerar que a relação existe, então, não basta detectar o efeito de uma variável ambiental sobre α, é preciso também que a estimativa de coeficiente angular seja alta.

Isso posto, vamos prosseguir com a interpretação dos resultados sob as restrições impostas quando não sabemos o grau de heterogeneidade de probabilidades de captura entre as espécies. Nas análises das três paisagens, embora os modelos que preveem o efeito de uma das duas variáveis de fragmentação analisadas estivessem entre os selecionados, os valores estimados de coeficientes angulares foram sempre baixos. Então, considerando os diferentes cenários de amostragem testados no capítulo anterior, temos algumas possibilidades de interpretação dos nossos resultados: (a) existe a relação e o coeficiente angular é realmente muito baixo; (b) existe a relação, o efeito é importante, porém a heterogeneidade de probabilidades de captura entre as espécies é alta e isso provocou um viés de subestimação acentuado; (c) não existe a relação, porém sob o efeito de diferenças grandes nas capturabilidades das espécies, foi selecionado o modelo com efeito com baixas estimativas de coeficiente angular.

Adotando uma interpretação mais conservadora e parcimoniosa, propomos que a inclusão dos modelos com efeitos entre os selecionados deveu-se a uma falsa detecção de efeitos em função da heterogeneidade de probabilidades de captura. Além das baixas estimativas de coeficiente angular obtidas, há as evidências empíricas apontadas por outros autores de que as probabilidades de captura provavelmente têm alta variabilidade entre as espécies de aves (MacArthur & MacArthur 1974, Keyes & Grue 1982, Remsem & Good

1996), e são definidas por vários fatores (Remsem & Good 1996). Além disto, mesmo que houvesse o efeito e ele estivesse subestimado, ele ainda seria muito baixo. A paisagem que nos deixaria com a maior dúvida quanto à presença ou não do efeito é a de 10% de cobertura vegetal, para a qual foi selecionado o modelo com efeito de área, com o maior valor de coeficiente angular estimado. Porém, mesmo assumindo que nesta paisagem existe uma influência do tamanho do fragmento sobre a diversidade, e que ele foi subestimado em função do alto grau de heterogeneidade de capturabilidade entre as espécies; corrigindo a nossa estimativa pelo viés máximo obtido no capítulo 2 (62,47%), obteríamos um coeficiente angular de no máximo de 2,7. Considerando que α é proporcional ao número de espécies raras na comunidade, um coeficiente angular de 2,7 seria algo na ordem de acréscimo de três espécies a mais que quando o α é constante, a cada vez que os fragmentos dobram de tamanho. Assim, mesmo que nas paisagens existam os efeitos das variáveis de fragmentação, eles são provavelmente de pequena magnitude, e assumir então um único valor médio de α para todos os fragmentos parece ser uma boa aproximação. A pequena variação em α, prevista pelo valor de coeficiente angular estimado pelo modelos também pode ser visualizado quando comparamos as curvas das distribuições de abundâncias previstas, com e sem efeito, que estão sempre muito próximas, mesmo nos fragmentos de maior e menor área ou isolamento (figuras 2-4).

Assim, apesar do desconhecimento das distribuições das probabilidades de captura das espécies nas três paisagens, a escolha de um parâmetro de um modelo de SAD, o α de Fisher, como variável resposta para testar os efeitos de fragmentação ainda pode ser bastante informativa. Os outros estudos realizados no local (Bueno 2008; Martensen 2008), um deles utilizando os mesmos conjuntos de dados (Martensen 2008), encontraram o efeito de algumas métricas sobre a riqueza e abundância total, sendo que a importância de cada métrica variava de acordo com a paisagem. Como o α expressa a proporção com que o aumento do número de indivíduos na amostra adiciona espécies na mesma (eq. 3), ao encontrarmos que ele não é fortemente influenciado pelas métricas de fragmentação, podemos dizer que a variação da riqueza entre fragmentos com diferentes atributos se deve apenas à variação do número de indivíduos capturados em cada fragmento. De fato, a relação entre o número de espécies em função de indivíduos capturados nos fragmentos parece seguir o que foi previsto pelo valor médio de α, estimado para todos os fragmentos juntos, para as três paisagens (Figura 5). Assim, a perda de espécies não está sendo acelerada pelo processo de fragmentação, mas é apenas uma decorrência da diminuição do número de indivíduos.

O número de indivíduos, por sua vez, está relacionado à capacidade de suporte do

local, que deve aumentar com a área dos fragmentos. Se α reduzisse em função da fragmentação, esperaríamos que a dominância ficasse maior, pela perda de algumas espécies mais sensíveis e aumento de outras espécies beneficiadas pela fragmentação. Entretanto, quando encontramos que α não se altera ou então varia muito pouco em função da fragmentação, embora haja uma diminuição de indivíduos e espécies, não há alteração na hierarquia entre as abundâncias, indicando que a forma da distribuição mantém-se a mesma. Neste cenário uma hipótese atraente seria que, com a perda de habitat pela fragmentação, as espécies perdem indivíduos e por isso as mais raras na paisagem original seriam aquelas mais propensas a desaparecer, com reduções proporcionais nas abundâncias daquelas que permanecem (Caugley 1994).

Porém, encontramos que nem sempre são as espécies raras que se extinguem, parecendo haver uma alta taxa de substituição, entre as espécies que eram as mais raras e as mais abundantes nas áreas contínuas (figuras 6-8). Mac Nally (2007), em um estudo que também analisou mudanças nas abundâncias das espécies com a fragmentação, chegou a encontrar um padrão em que há uma inversão de dominância entre as espécies que são originalmente mais abundantes e as mais raras. Embora não tenhamos encontrado um padrão de substituição tão radical, é evidente que a redução da riqueza não se dá por amostragem passiva, que tenderia a eliminar as mais raras (Coleman et. al 1982). Com nossa analise é possível perceber, sobretudo nas paisagens de 10 e 30% de cobertura vegetal, que nas áreas controle adjacentes ou até mesmo nos fragmentos maiores há um número maior de espécies consideradas altamente sensíveis a perturbações. Além disso, nas duas paisagens de menor porcentagem de cobertura vegetal, todas as espécies ausentes nas áreas controle, que apareceram nos fragmentos são compostas de espécies pouco sensíveis a impactos.

Assim, embora a relação entre as abundâncias das espécies não se altere com o processo de fragmentação, a perda de espécies parece não estar acontecendo pela diminuição proporcionalmente equivalente para todas as populações. Algumas espécies parecem ser mais prejudicadas com a perda e isolamento de habitat, o que por si só pode beneficiar o crescimento das espécies existentes e também a invasão de espécies de fora. E talvez a matriz com maior porcentagem de cobertura vegetal seja importante para a manutenção de algumas populações mais sensíveis, uma vez a paisagem com maior porcentagem de cobertura vegetal foi a única onde os fragmentos menores ainda apresentam espécies consideradas altamente sensíveis. O processo de fragmentação, portanto, embora tenha um efeito sobre a diversidade dos fragmentos apenas em termos de riqueza e não de dominância, independente da porcentagem de cobertura vegetal na paisagem, não atua de maneira homogênea sobre todas

as espécies. Algumas espécies parecem ser realmente mais sensíveis enquanto outras podem acabar se beneficiando com o processo de fragmentação, o que causa mudanças na sua posição na hierarquia de abundancias, embora a hierarquia em si não se altere muito.

5. Bibliografia

AKAIKE, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. In: B. N. PETROV & F. CSAKI, eds. Second International Symposium on Information Theory. Budapest: Akademiai Kiado, pp. 267–281.

ALONSO, D., OSTILING, A. & ETIENNE, R.S. 2008. The implicit assumption of simmetry and species abundance distribution. Ecology Letters. 11: 93-105.

ANDRÉN, H., 1994. Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat - A Review. Oikos 71(3): 355-366.

BANKS-LEITE, C. Conservação da comunidade de aves de sub-bosque em paisagens fragmentadas da Floresta Atlântica. Tese (Doutorado). Instituto de Biociências da Universidade de São Paulo. 227p.

BOLKER, B. 2010. bbmle:Tools for general maximum likelihood estimation. Pacote de R versão 0.9.5.1 baseado no pacote stats4 por R Development Core Team. URL http://CRAN.R-project.org/package=bbmle

BOWMAN, J., CAPPUCCINO, N. & FAHRIG, L., 2002. Patch size and population density: the effect of immigration behavior. Conservation Ecology 6(1): 9 .

BURNHAM, K. P., AND D. R. ANDERSON. 2002. Model selection and multimodel inference: a practical information-theoretic approach. 2nd Edition. Springer-Verlag, New York, New York, USA. 488 pp.

BUENO, A. B. 2008. Pequenos mamíferos da Mata Atlântica do Planalto Atlântico Paulista: uma valiação da ameaça de extinção e da resposta a alterações no contexto e tamanho dos remanescentes. Instituto de Biociências – Universidade de São Paulo. Tese, Departamento de Zoologia. 124 p.

CAYUELA, L., GOLICHER, D. J., BENAYAS, J. M. R. GONZÁLEZ-ESPINOSA, M. & RAMÍREZ-MARCIAL, N. Fragmentation, disturbance and tree diversity conservation in tropical montane forest. Journal of applied Ecology. 43: 1172-1181.

CAUGLEY, G. 1994. Directions in conservation biology. Jounal of animal Ecology. 63: 215-

244.

COLEMAN, B.D. 1982. Randomness, area, and species richness. Ecology. 63(4): 1121-1131.

EWERS, R. M. & DIDHAM, R. K. 2006. Confounding factors in the detection of species responses to habitat fragmentation. Biological Reviews. 81: 117-142.

FAHRIG, L. 2001. How much habitat is enough? Biological Conservation, 100: 65-74.

FAHRIG, L. 2003. Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution and Systematics 34: 487-515.

FISHER, R. A., CORBET, A. S. & WILLIAMS, C.B. 1943. The relation betwen the number of species and the number of induvidulas in a random sample of an animal population. The Journal of Animal Ecology. 12: 42-58.

FISHER, J. & LINDENMAYER, D. B. 2007. Landscape modification and habitat fragmentation: a synthesis. Global Ecology and Biogeography. 16 (3): 265-280.

GRAY, J. S. 1979. Pollution-induced changes in populations. Philosophical Transactions of the Royal Society B: Biological Sciences
. 286:545–557.

HANSKI, I. & SIMBERLOFF, D. 1997. The metapopulation approach, its history, conceptual domain, and application to conservation, In: Hanski, I., Gilpin, M. (Eds.), Metapopulation dynamics: empirical and theorethical investigations. Academic Press, London, UK, pp. 5-26.

HUBBELL, S.P. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton, Princeton University Press.

JULES, E.S. 1998. Habitat fragmentation and demographic change for a common plant: trillium in old-growth forest. Ecology 69: 1645-1656.

KEYES, B.E. & C.E. GRUE. 1982. Capturing birds with mist nets: a review. North American Bird

Bander 7: 2-14.

MACARTHUR, R. H. & MACARTHUR, A. T. 1974. On the use of mist nets for population studies of birds. Proc. Natl. Acad. Sci. USA 71:3230-3233.

MACARTHUR, R. H. & WILSON, E. O. 1967. The theory of island biogeography. Princeton University Press, Princeton.

MAC NALLY, R. 2007. Use of the abundance spectrum and relative-abundance distributions to analyze assemblage change in massively altered. Landscape. 170(3): 319-330.

MAGURRAN, A.E. 2004. Measuring Biological Diversity . Blackwell Science, Oxford.

MARTENSEN, A. C. Conservação de aves de sub-bosque em paisagens fragmentadas: Importância da cobertura e da configuração do habitat. Dissertação (Mestrado). Instituto de Biociências da Universidade de São Paulo. 160p.

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 (in press).

METZGER, J.P. 2000. Tree functional group richness and landscape structure in a Brazilian tropical fragmented landscape. Ecological Applications 10: 1147–1161.

METZGER, J. P. 2006. Conservação da Biodiversidade em Paisagens Fragmentadas no Planalto Atlântico de São Paulo - II. Projeto de Pesquisa, CNPq, Brasília.

MOUQUET, N. & M. LOREAU. 2003. Community patterns in source-sink metacommunities. American Naturalist 162:544–557.

OKSANEN, J.; BLANCHET, G.; KINDT, R.; LEGENDRE, P.; O'HARA, R.B.; SIMPSON G.L.; SOLYMOS, P.; STEVENS, M.H. & WAGNER, H. 2010. Vegan: Community Ecology Package. Pacote do R versão 1.17-2. URL http://CRAN.R-project.org/package=vegan

PIELOU, E. C. 1977. *Mathematical Ecology*. New York: John Wiley and Sons.

R DEVELOPMENT CORE TEAM. 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org .

REMSEN, J. V. & GOOD, D.A. 1996. Misuse of data from mist-net captures to assess relative abundance in bird populations. The awk. 113(2):381-398.

RICKETTS, T.H. 2001. The matrix matters: effective isolation in fragmented landscapes. American Naturalist. 158(1):87-99.

ROYLE, J. A. & DORAZIO, R. M. 2008. Hierarchical modeling and inference in ecology:The analysis of data from populations, metapopulations and communiteis.Academic Press. San Diego, CA. 444 p.

TAYLOR, L.R., 1978. Bates, Williams, Hutchinson—a variety of

diversities. In: Mound, L.A., Waloff, N. (Eds.), The diversity of Insect Faunas. Blackwell, Oxford.

VELOSO, H. P., RANGEL-FILHO, A. L. R., & LIMA, J. C. A. 1991. Classificação da vegetação brasileira adaptada a um sistema universal. IBGE, Rio de Janeiro, Brasil.

WATLING, J. I. & M. A. DONNELLY. 2006. Fragments as islands: a synthesis of faunal responses to habitat patchiness. Conservation Biology. 20:1016—1025.

WILLIAMS, B. K., NICHOLS, J. D., & CONROY, M. J. 2001. Analysis and management of animal populations: modeling, estimation, and decision making. Academic Press.

WITH, K.A., 1997. The application of neutral landscape models in conservation biology. Conservation Biology 11, 1069-1080.

YAACOBI, G., YARON, Z. , ROSENZWEIG, M. L. 2007. Habitat fragmentation may not matter to species diversity. Proceedings of the Royal Society B. 274: 2409-2412.

Considerações Finais

Iniciamos este trabalho com a proposta de aplicar os modelos teóricos de distribuição de abundância para testar efeitos de variáveis ambientais sobre a diversidade de comunidades biológicas. A proposta de se usar modelos em uma abordagem fenomenológica já foi levantada por alguns autores (Pielou 1977; McGill et al. 2007), porém ela permanece pouco usual (McGill et al.2007). Assim, a execução do projeto foi planejada não somente para responder as perguntas sobre efeitos ambientais, mas também para avaliar o quanto tal abordagem poderia ser frutífera.

Iniciamos o projeto pelo estudo dos diferentes modelos de distribuição de abundância relativa para escolhermos aqueles que fossem mais adequados para serem utilizados nas análises. No decorrer do estudo, percebemos a importância que a teoria da amostragem teve nas construções das propostas dos modelos e decidimos optar pelo uso exclusivo das alternativas que contemplassem o processo amostral na sua concepção. Porém, tais alternativas resultam em modelos compostos que representam a distribuição das capturas das espécies nas amostras, dada a distribuição das abundâncias das espécies na comunidade, cuja formulação em geral tem um parâmetro adicional: a fração da comunidade amostrada, equivalente a probabilidade de captura das espécies.

Durante a busca pelos métodos mais adequados para a estimação das probabilidades de captura, encontramos alguns trabalhos com evidências empíricas de que as probabilidades de captura das espécies são provavelmente muito diferentes entre elas (eg. MacArthur & MacArthur 1974; Remsen & Good 1996; Dodd & Dorazio 2004). Para isto foram decisivas a disciplina de Ecologia de Populações, do professor Gonçalo Ferraz, que nos chamou a atenção para a importância deste problema, e nos apontou a literatura pertinente. Foi desta maneira, então, que nos vimos obrigados a perguntar quais eram os riscos de se assumir um único valor de parâmetro de probabilidade de captura para todas as espécies, ou seja, o quanto a variação nas probabilidades de captura distorciam a representação das relações de abundância das espécies na amostra. O problema poderia ser contornado se substituíssemos as abundâncias observadas na amostra por estimativas das abundancias das espécies na comunidade, obtidas com métodos que levassem em conta a capturabilidade (Williams et al. 2002). Tentamos, então estimar as abundâncias de cada espécie nos conjuntos de dados de marcação e recaptura de aves que tínhamos disponíveis. Porém, pela pequena quantidade de capturas e, principalmente, recapturas de muitas espécies, não foi possível estimar as abundâncias para a maior parte delas, mesmo com os modelos mais simples de estimação. Tal problema parece ser recorrente em estudos que utilizam dados de contagem no nível da

comunidade (Jung et al. 2000, Dodd & Dorazio 2004). A primeira pergunta a ser respondida, portanto era se a variação interespecífica nas probabilidades de captura poderia enviesar nossos resultados; e se o viés acontecesse, em que situações ele nos impediria de responder a pergunta do projeto. A partir daí, decidimos prosseguir com simulações dos cenários de amostragem em que as espécies tinham probabilidades diferentes de serem capturadas (capítulos um e dois)

Apesar de não ter sido inicialmente planejada, a inclusão das duas etapas de simulação no projeto foi essencial para que pudéssemos responder nossas perguntas mais adequadamente. A maneira como analisamos os dados não mudou muito do que já tínhamos planejado, porém a interpretação dos resultados das análises foi diferente do que teria sido se não tivéssemos feito as simulações. Podemos dizer que a interpretação que fizemos dos resultados das análises só pôde ser mais conclusiva devido aos resultados descritos nos dois primeiros capítulos desta dissertação. Saber a que tipo de distorção as distribuições do número de capturas relativa as amostras podem estar expostas possibilita a delimitação do espaço de possíveis explicações a serem exploradas para entender o sistema de estudo.

O que podemos concluir nesta dissertação, portanto, é que a variação das probabilidades de captura pode alterar muito a distribuição de abundância das espécies pela amostra, mas que talvez seja possível atenuar seu efeito pela adequação do delineamento amostral, do método de captura ou pela conhecimento da distribuição das probabilidades de captura entre as espécies. Adotando medidas de cautela sobre os possíveis cenários de amostragens e tipos de vieses, concluímos também que o processo de fragmentação, em termos de tamanho e isolamento de fragmentos, tem um efeito muito fraco sobre a hierarquia de abundância das espécies, independente da porcentagem de cobertura vegetal da paisagem. Porém o tamanho do fragmento pode ser importante na determinação de quais espécies serão extintas ou beneficiadas e que talvez a qualidade da matriz seja decisiva para a manutenção de espécies altamente sensíveis em fragmentos pequenos.

Além de sua contribuição para a teoria de amostragem aplicada a modelos de SADs e a aplicação destes modelos para elucidar efeitos da fragmentação, esta tese ilustra os ciclos de aperfeiçoamento de um modelo, e o enorme potencial que há para isto na combinação de soluções analíticas, simulação computacional, e testes empíricos. A pré-condição para termos chegado até aqui é a convicção de que modelos são meios para compreender melhor uma situação, e não fins em si. Isto porque nesta concepção os modelos são representações de processos biológicos que determinam as abundâncias das espécies nas comunidades, considerando que há um processo amostral na representação das mesmas.

Bibliografia

DODD, C.K. & DORAZIO, R.M. Using counts to simultaneously estimate abundance and detection probabilities in a salamander community. Herpetologica. 60 (4): 468-478.

MACARTHUR, R. H. & MACARTHUR, A. T. 1974. On the use of mist nets for population studies of birds. Proc. Natl. Acad. Sci. USA 71:3230-3233.

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 (in press).

PIELOU, E. C. 1977. *Mathematical Ecology*. New York: John Wiley and Sons.

REMSEN, J. V. & GOOD, D.A. 1996. Misuse of data from mist-net captures to assess relative abundance in bird populations. The awk. 113(2):381-398.

ROYLE, J. A. & DORAZIO, R. M. 2008. Hierarchical modeling and inference in ecology:The analysis of data from populations, metapopulations and communiteis.Academic Press. San Diego, CA. 444 p.