

VOLUME 3 | Nº 1-2 | DEZEMBRO 2007



# MEGADIVERSIDADE



**Modelagem  
ambiental e a  
conservação da  
biodiversidade**

Editor Geral  
**José Maria Cardoso da Silva**

Editores Convidados  
**Ana Luiza Kerti Mangabeira Albernaz**  
**Thaís Pacheco Kasecker**

Editores Associados  
**Adrian Antônio Garda**  
**Guilherme Fraga Dutra**  
**Luiz Paulo Pinto**  
**Paulo Gustavo Prado**  
**Ricardo Bomfim Machado**

Coordenação de Edição e Produção  
**Isabela Santos**

Staff de Produção  
**Regiane Avelar**

Conselho Editorial  
**Ana Rodrigues**  
**Angelo Machado**  
**Anthony Rylands**  
**Claude Gascon**  
**Francisco Barbosa**  
**Gustavo Fonseca**  
**Ima Célia Vieira**  
**Katrina Brandon**  
**Keith Alger**  
**Marcelo Tabarelli**  
**Roberto Cavalcanti**  
**Russell Mittermeier**  
**Thomas Lewinsohn**  
**Thomas Lacher**

**Conservação Internacional**  
Av. Getúlio Vargas 1300 – 7º andar  
30112-021 Belo Horizonte – MG  
Tel.: 55 31 3261-3889  
e-mail: [info@conservacao.org](mailto:info@conservacao.org)  
[www.conservacao.org](http://www.conservacao.org)



A Conservação Internacional é uma organização privada sem fins lucrativos, fundada em 1987, com o objetivo de conservar o patrimônio natural do planeta – nossa biodiversidade global – e demonstrar que as sociedades humanas são capazes de viver em harmonia com a natureza.

**Volume 3 | Nº 1-2 | Dezembro 2007**

Fotos da capa: (1) e (2) Adriano Jerozolimski, (3) Haroldo Castro, (4) Enrico Bernard,  
(foto maior) Adriano Gambarini.

Projeto e edição gráfica: Grupo de Design Gráfico Ltda.  
Tiragem: 1.000 exemplares

**Apoio**



# MEGADIVERSIDADE

---

Volume 3 | Nº 1-2 | Dezembro 2007

## Modelagem ambiental e conservação da biodiversidade

### SUMÁRIO

- 3 **Apresentação**  
JOSÉ MARIA CARDOSO SILVA, ANA LUIZA KERTI MANGABEIRA ALBERNAZ & THAÍS PACHECO KASECKER
- 5 **Modelos lineares como ferramentas para a modelagem da distribuição de espécies**  
GUILHERME MOURÃO & WILLIAM E. MAGNUSSON
- 13 **Modelos de classificação em árvore para a localização de áreas de ocorrência potencial de espécies**  
ADRIANA PAESE, SIDNEY DORNELLES, JOSÉ EDUARDO DOS SANTOS & JOSÉ SALATIEL RODRIGUES PIRES
- 25 **O uso de regressão logística para espacialização de probabilidades**  
EDUARDO M. VENTICINQUE, JULIANA STROPP CARNEIRO, MARCELO PAUSTEIN MOREIRA & LEANDRO FERREIRA
- 38 **Uso de modelos aditivos generalizados na estimativa da distribuição potencial de espécies**  
PAULO DE MARCO JÚNIOR
- 46 **Algoritmo Genético para Produção de Conjuntos de Regras (GARP)**  
RICARDO SCACHETTI PEREIRA & MARINEZ FERREIRA DE SIQUEIRA
- 56 **Modelos de metapopulação**  
DOUGLAS F. M. GHERARDI
- 64 **O uso de modelos em ecologia de paisagens**  
JEAN PAUL METZGER, MARINA ANTONGIOVANNI DA FONSECA, FRANCISCO JOSÉ BARBOSA DE OLIVEIRA FILHO & ALEXANDRE CAMARGO MARTENSENS
- 74 **Modelagem de dinâmica de paisagem: concepção e potencial de aplicação de modelos de simulação baseados em autômato celular**  
BRITALDO SILVEIRA SOARES FILHO, GUSTAVO COUTINHO CERQUEIRA, WILLIAM LEITE ARAÚJO & ELIANE VOLL
- 87 **Planejamento sistemático para a conservação na Amazônia brasileira – uma avaliação preliminar das áreas prioritárias de Macapá-99**  
ANA LUIZA KERTI MANGABEIRA ALBERNAZ & MANUELLA ANDRADE DE SOUZA



# APRESENTAÇÃO

Um dos maiores desafios da moderna ciência da biodiversidade é determinar quantas espécies existem no planeta, como elas se relacionam filogeneticamente e, por fim, como elas se distribuem. O conhecimento detalhado das distribuições geográficas das espécies é fundamental para o planejamento da conservação e para prever, por exemplo, possíveis conseqüências, sobre a biodiversidade global, das mudanças globais que afetam o planeta. Além disso, as distribuições geográficas são uma das informações mais importantes para programas de pesquisa que visam compreender os processos ecológicos e evolutivos que determinam os grandes padrões de biodiversidade global.

Infelizmente, a distribuição de grande maioria das espécies é conhecida somente a partir de informações muito limitadas, o que restringe significativamente muitas análises. Um bom exemplo diz respeito à tomada de decisões sobre a conservação de ambientes e espécies tropicais. Geralmente o processo é complexo e envolve a ponderação de aspectos ambientais, econômicos e sociais. A falta de informações detalhadas sobre a distribuição das espécies é um dos fatores que contribui para que as discussões sobre conservação e a criação de áreas protegidas percam seu foco sobre biodiversidade, dando mais peso a outros fatores.

Uma das estratégias encontradas por cientistas para minimizar o problema da falta de informações foi a de desenvolver modelos computacionais para gerar predições sobre as distribuições das espécies a partir da associação da ocorrência da espécie com fatores ambientais. Nos últimos anos, novas ferramentas computacionais especialmente voltadas para a modelagem ecológica foram criadas e tornadas disponíveis para uso geral. Tais ferramentas incluem modelos de distribuição de espécies, de metapopulações, de paisagens e de escolha de áreas prioritárias para a conservação. Embora várias destas ferramentas estejam presentes na literatura internacional há muito tempo, e o valor de suas aplicações seja amplamente reconhecido, elas ainda têm sido pouco utilizadas no Brasil, e particularmente na Amazônia. Foi esta constatação que levou o Programa Institucional de Biodiversidade do Museu Paraense Emilio Goeldi (MPEG), em parceria com a Conservação Internacional e o projeto GEOMA de Modelagem Ambiental na Amazônia, a organizar o seminário “Ferramentas para a modelagem da distribuição de espécies em ambientes tropicais”, realizado em Belém, no período de 10 a 14 de fevereiro de 2003.

Este número especial da Megadiversidade apresenta os resultados do seminário. A seleção dos temas e ordem de sua apresentação foi feita de forma a apresentar os modelos dos mais simples aos mais complexos. Os primeiros capítulos tratam exclusivamente da distribuição de espécies (modelos gerais lineares, regressão logística, árvores de decisão, modelos gerais aditivos e o algoritmo genético GARP). A seguir são apresentados modelos que incorporam componentes de paisagem: metapopulações, ecologia de paisagem, autômatos celulares e sistemas de apoio à escolha de áreas para a conservação.

Nem todos os modelos disponíveis em 2003 foram apresentados durante o seminário: a seleção dependeu principalmente de profissionais interessados em desenvolver trabalhos sobre o tema. Todos os textos e análises foram elaborados por pesquisadores que atuam em instituições nacionais, e os exemplos apresentados também são extraídos da nossa realidade.

Este número só se tornou possível pelo extraordinário esforço dos autores, que fizeram várias atualizações dos artigos, e do trabalho de editoração e acompanhamento de Renata Valente e Mônica Fonseca, ambas da equipe da CI-Brasil. Esperamos com isso estimular novos pesquisadores a seguir por esse caminho, contribuindo, assim, para o avanço desse tipo de pesquisa no Brasil.

José Maria Cardoso da Silva – CONSERVAÇÃO INTERNACIONAL

Ana Luiza Kerti Mangabeira Albernaz – MUSEU PARAENSE EMÍLIO GOELDI

Thaís Pacheco Kasecker – CONSERVAÇÃO INTERNACIONAL

EDITORES

---

# Modelos lineares como ferramentas para a modelagem da distribuição de espécies

GUILHERME MOURÃO<sup>1\*</sup>  
WILLIAM E. MAGNUSSON<sup>2</sup>

<sup>1</sup> Embrapa Pantanal, Mato Grosso do Sul, Brasil.

<sup>2</sup> Instituto Nacional de Pesquisas da Amazônia – INPA, Amazonas, Brasil.

\* e-mail: gui@cpap.embrapa.br

## RESUMO

Modelos lineares gerais (MLG) podem ser empregados para descrever relações ecológicas, como a relação entre fator ou fatores abióticos e densidades de organismos. Entretanto, os modelos lineares teoricamente só são apropriados em condições restritas, geralmente difíceis de serem encontradas no mundo real. Apesar desta limitação, modelos lineares vêm sendo amplamente empregados em estudos ecológicos, principalmente porque permitem investigar o efeito de fatores que podem estar sendo mascarados por outros. Neste capítulo apresentamos exemplos de aplicação de modelos lineares em alguns estudos que enfocam a distribuição de organismos na Amazônia.

## ABSTRACT

*General linear models (GLM) can be used to describe ecological relationships, such as the relationships between abiotic factors and densities of organisms. However, linear models are strictly applicable under restricted conditions that are unlikely to be encountered in the real world. Despite this limitation, linear models have been frequently used in ecological studies, mainly because they allow studies of the effects of factors that may be confounded by others. In this chapter we present examples of the use of linear models in studies of Amazonian organisms.*

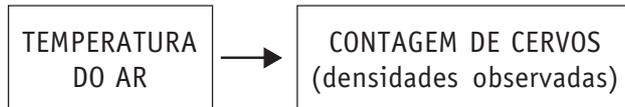
## MODELOS LINEARES SIMPLES

Na matemática, um modelo linear simples segue a fórmula  $y = a + b * x$  e a representação gráfica desta relação segue a forma geral apresentada na Figura 1a. Chamamos  $y$  e  $x$  de variáveis, porque assumem diferentes valores dentro de um mesmo modelo. Por convenção, chamamos  $y$  de variável dependente, *i.e.*,  $y$  varia em função de  $x$ , que é a variável independente.

Os valores denotados pelas letras  $a$  e  $b$  são chamados de parâmetros, porque são constantes e característicos de uma dada população de dados, embora possam variar entre modelos da mesma classe. Estimativas de parâmetros baseados em amostras da população são chamadas estatísticas. O parâmetro  $a$  se refere ao valor que  $y$  assume quando  $x=0$  e por isso é chamado de “intercepto” ou “elevação”. Observe que na Figura 1a o menor valor que aparece na escala do eixo  $x$  é 20 e o

valor de  $y$  onde  $x=0$  não é mostrado (neste exemplo,  $a=-13,9$ ). O parâmetro  $b$  indica quanto é o incremento (ou decréscimo) de  $y$  quando  $x$  aumenta em uma unidade, e por isso é chamado de inclinação.

Vamos considerar uma questão do mundo real, onde a temperatura do ar afeta a contagem de uma determinada espécie, digamos, cervos-do-pantanal. O fluxograma abaixo dá uma indicação adequada de nosso modelo conceitual:



A analogia do modelo matemático  $y = a + b * x$  com nosso modelo do mundo real se faz como  $\text{CONTAGEM} = \text{ELEVACÃO} + \text{INCLINAÇÃO} * \text{TEMPERATURA}$ , onde a  $\text{CONTAGEM}$  é a variável dependente e a  $\text{TEMPERATURA}$  é a variável independente ou, apropriadamente, o “fator” que afeta a variável dependente. Neste caso, a elevação apareceu

explicitamente no modelo, mas se tivéssemos razão para crer que a relação fosse linear e que não contaríamos cervo algum quando a temperatura do ar fosse zero (*i.e.*  $\text{ELEVACÃO}=0$ ) poderíamos simplesmente omiti-la. Entretanto, na maioria dos casos é mais seguro incluir a constante e deixar que os dados, em vez de nossos preconceitos, posicionem a reta.

No mundo real os modelos são sempre imperfeitos, e a relação que esperamos encontrar entre variáveis dependentes e fatores são mais semelhantes ao gráfico apresentado na Figura 1b. Podemos ver que, neste caso, a equação  $y = a + b * x$  já não é suficiente para descrever cada ponto do gráfico, e precisamos acrescentar mais um termo para ajustar a equação:  $y = a + b * x + e$ , ou  $\text{CONTAGEM} = \text{ELEVACÃO} + \text{TEMPERATURA} + \text{VARIACÃO RESIDUAL}$ , onde a variação residual represente o desvio da observação do nosso modelo.

Note que no gráfico da Figura 1b apareceu a linha que representa a relação linear entre as variáveis, mas não dissemos nada sobre o critério que escolhemos para definir onde a reta deveria passar. Obviamente desejamos traçar a linha na posição mais “próxima” possível

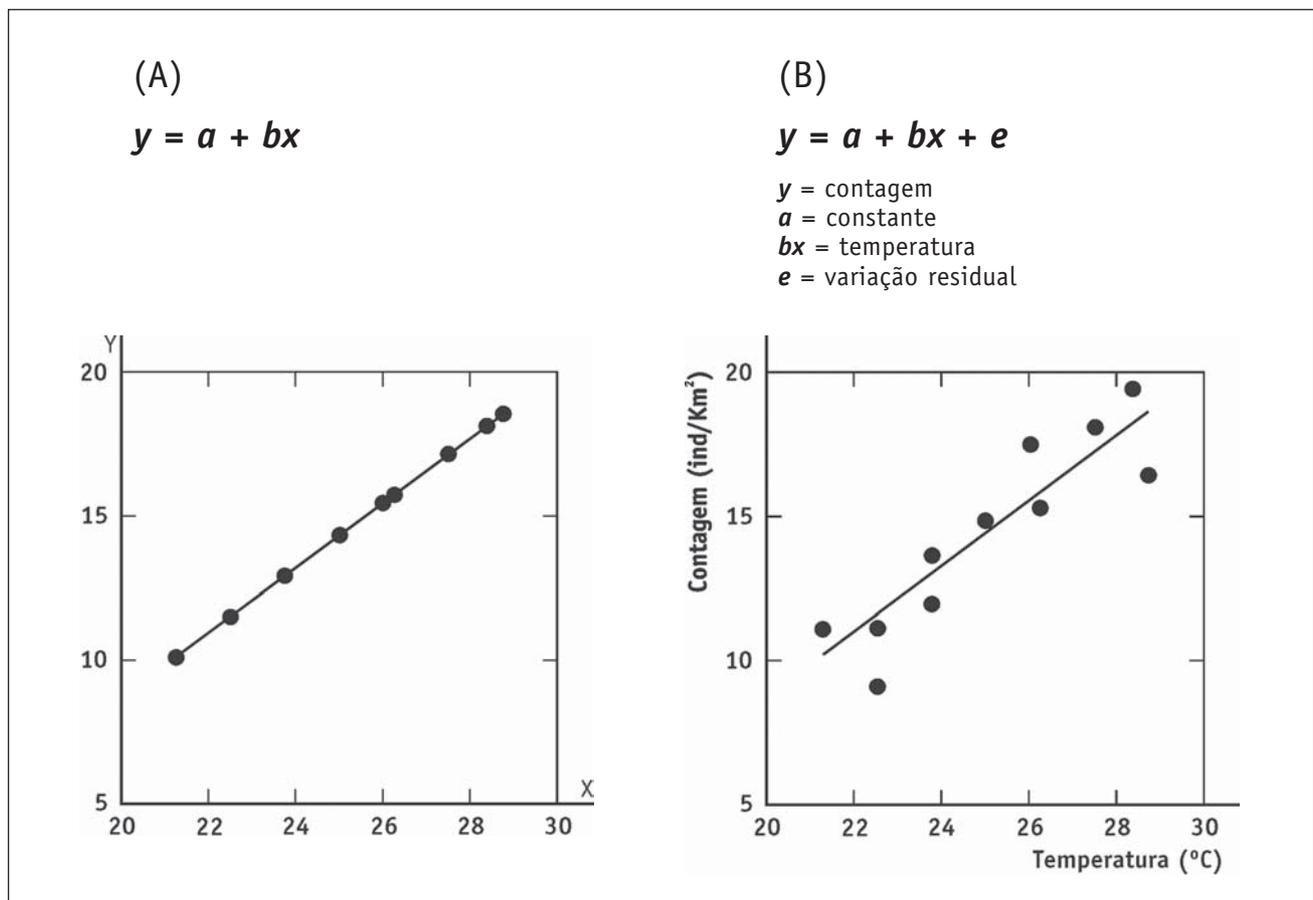
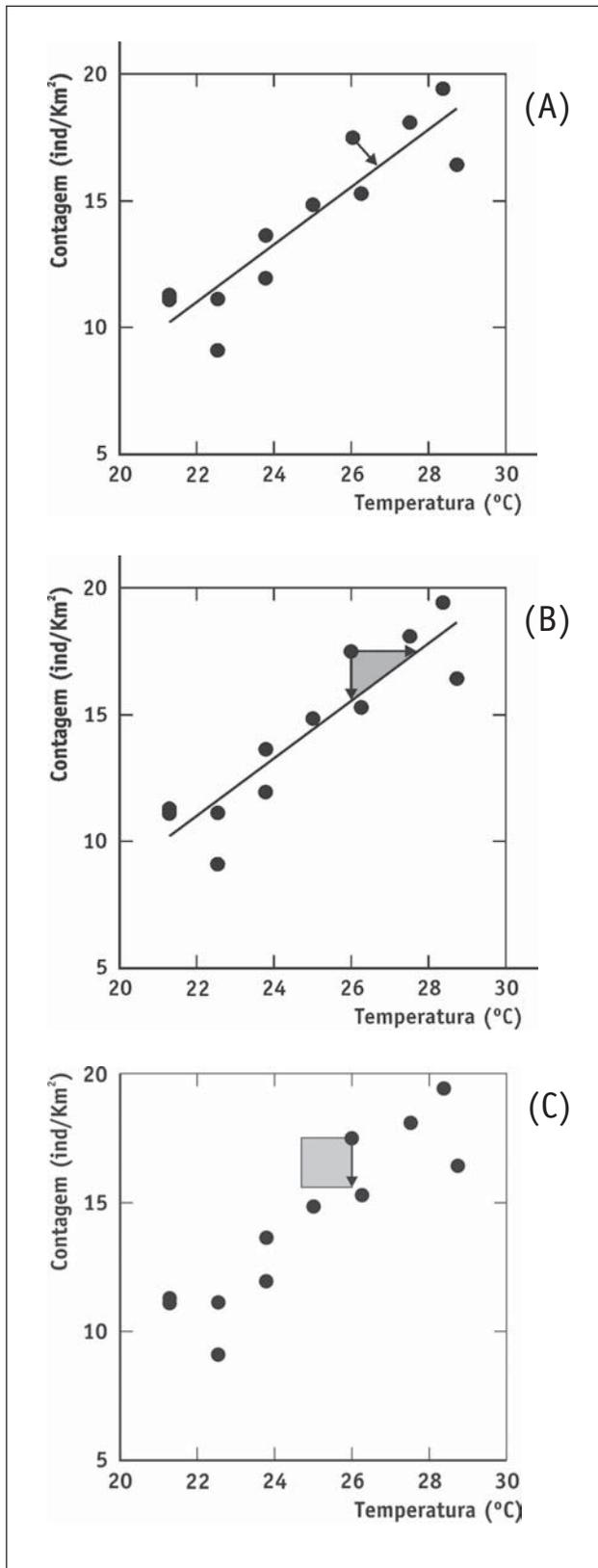


FIGURA 1 – (A) Representação matemática e gráfica de um modelo linear simples e (B) analogia com os modelos biológicos.



**FIGURA 2** – Diferentes alternativas de minimizar os resíduos dos pontos observados à reta de regressão: (A) distância perpendicular do ponto à linha; (B) simultaneamente as distâncias horizontal e vertical do ponto à linha; e (C) o quadrado da distância vertical do ponto à linha.

dos pontos em geral, o que significa minimizar as distâncias dos pontos até a linha, mas qual distância escolher? Poderíamos minimizar as distâncias perpendiculares à linha, como mostrado na Figura 2a, ou simultaneamente as distâncias horizontais e verticais de cada ponto até a linha (o que equivale a minimizar as áreas dos triângulos formados por estas distâncias e a linha), como na Figura 2b, ou ainda apenas a distância vertical dos pontos até a linha ou o quadrado desta distância (Figura 2c). Qualquer um destes métodos tem sua lógica e na verdade, deram origem a diferentes procedimentos estatísticos.

O procedimento de “regressão do maior eixo” minimiza a distância perpendicular do ponto à linha, como na Figura 2a, enquanto a regressão do “maior eixo reduzido” minimiza as áreas dos triângulos como mostrado na Figura 2b. Entretanto, o método de minimizar o quadrado das distâncias verticais dos pontos à linha, mostrado na Figura 2c, embora não pareça à primeira vista tão simples e intuitivo quanto o anterior, permitiu o desenvolvimento de uma grande variedade de análises complexas e geralmente as mais úteis em termos de previsão. Por razões óbvias, é chamada de “regressão dos mínimos quadrados”.

Mesmo quando trabalhamos com fatores categóricos, como nos modelos de análise de variância (ANOVA) convencional, estamos ajustando modelos lineares e de mínimos quadrados. De fato, internamente a maioria dos programas que computam ANOVA trabalham com códigos numéricos para as categorias que permitem uma análise de regressão.

A melhor estimativa do valor esperado para qualquer ponto em uma dada categoria é a média dos valores observados nesta categoria e o procedimento ANOVA minimiza os resíduos como o quadrado das distâncias verticais dos pontos até o valor esperado, de forma análoga à regressão linear.

## MODELOS LINEARES ADITIVOS (MULTIFATORIAIS)

Até agora, tratamos de modelos com somente um fator (variável independente) afetando a variável dependente. A mesma lógica pode ser estendida para modelos com mais que uma variável independente, como ilustrado no fluxograma abaixo:



A vantagem de se trabalhar com mais que uma variável por vez é que se pode investigar o efeito de um fator independente dos efeitos de outros fatores. Muitas vezes, uma variável esconde o efeito da outra (ver o exemplo a seguir), ou pode até criar a impressão de que a variável tem um efeito positivo quando em realidade seu efeito sobre a variável dependente é negativo (Magnusson & Mourão, 2003).

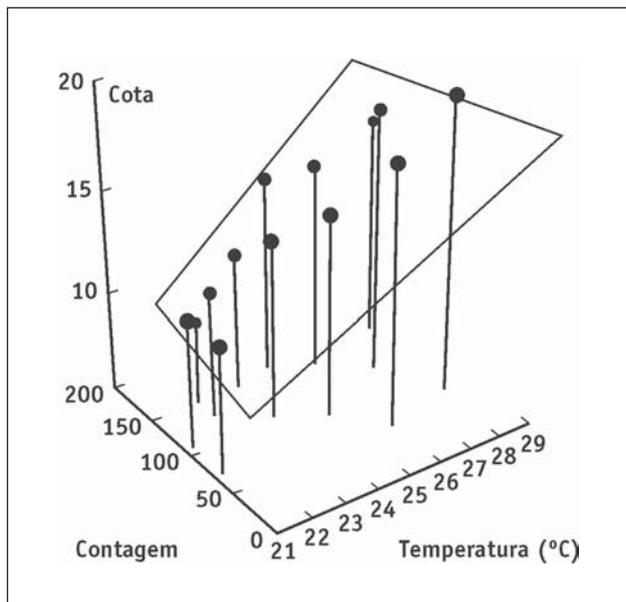


FIGURA 3 – Representação gráfica da superfície formada por um modelo linear de dois fatores.

O modelo de regressão múltipla (e os modelos lineares gerais [GLMs], em geral) é uma extensão do nosso modelo com um fator, mas tem a premissa de que os efeitos dos fatores são aditivos. O modelo conceitual é o seguinte:

$$y = a + b_1x_1 + b_2x_2 + \dots + b_ix_i + e$$

Em termos do nosso modelo específico, isto é expresso como:

$$\text{Contagem} = \text{elevação} + \text{inclinação}_1 * \text{temperatura} + \text{inclinação}_2 * \text{cota} + \text{resíduo}$$

Quando os efeitos das variáveis não são aditivos, é possível adicionar uma outra variável conceitual (chamada “interação”) para descrever este efeito (Magnusson & Mourão, 2003), mas estas variáveis conceituais têm pouco valor para modelagem de distribuições.

No caso de nosso exemplo com dois fatores, o modelo é uma superfície plana (Figura 3). Modelos com mais fatores produzem superfícies complexas que não podem ser apresentadas em três dimensões, mas cuja lógica matemática segue aquela de duas dimensões.

Para este exemplo, o exame dos gráficos bidimensionais simples indica um forte efeito da temperatura sobre o número de cervos contados, mas um efeito fraco, ou inexistente, da cota de inundação sobre o número de cervos contados (Figura 4).

O emprego de modelos multifatoriais (neste caso, regressão múltipla) permite o exame de efeitos “escondidos” por outros fatores. A análise produz gráficos que

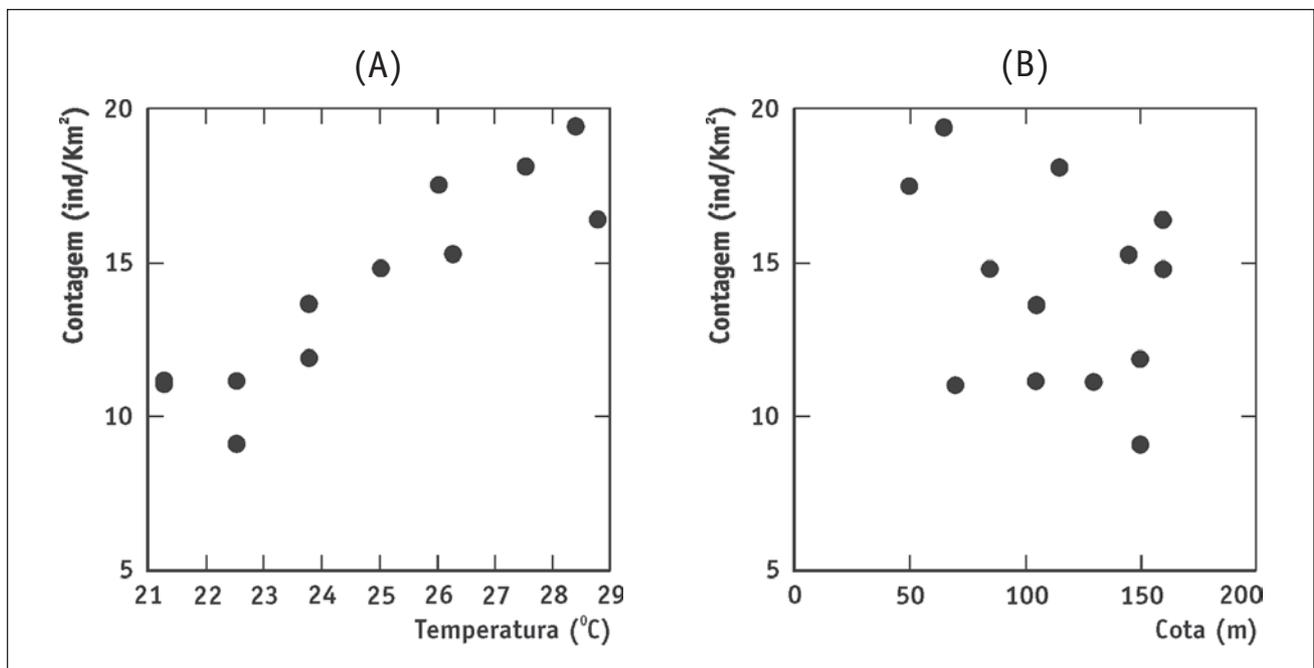
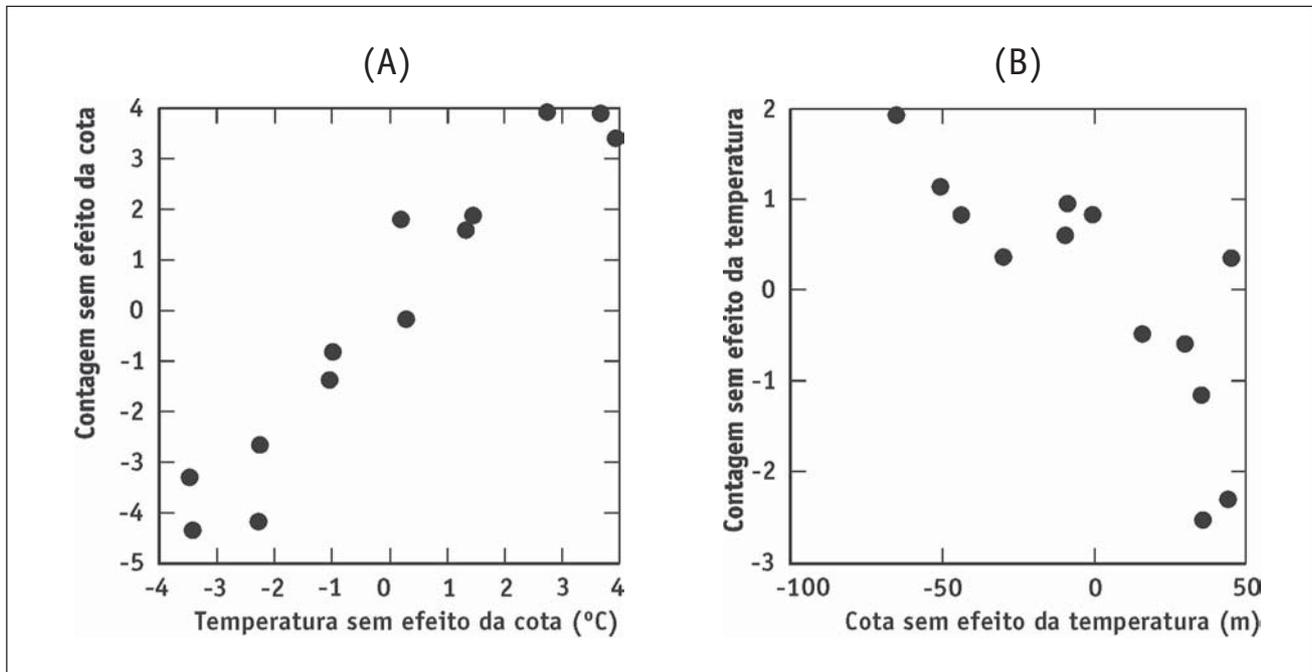


FIGURA 4 – Efeitos simples de (A) temperatura do ar e (B) da cota de inundação sobre o número de cervos contados.



**FIGURA 5** – Gráficos dos resíduos parciais, exemplificando como modelos multifatoriais podem ser usados para isolar os efeitos de fatores simultâneos (no caso, “cota” e “temperatura”) sobre uma variável dependente (“contagem”).

mostram os efeitos esperados caso as outras variáveis no modelo fossem mantidas constantes. Estes gráficos que representam as regressões parciais indicam um forte efeito de ambos, temperatura e cota (Figura 5). Isto pode ser importante, porque a temperatura provavelmente afeta apenas o comportamento do cervo de se esconder ou não à sombra e longe dos olhos dos observadores. A cota afeta o uso do hábitat e esta informação pode ser crítica para as previsões sobre a distribuição geográfica da espécie.

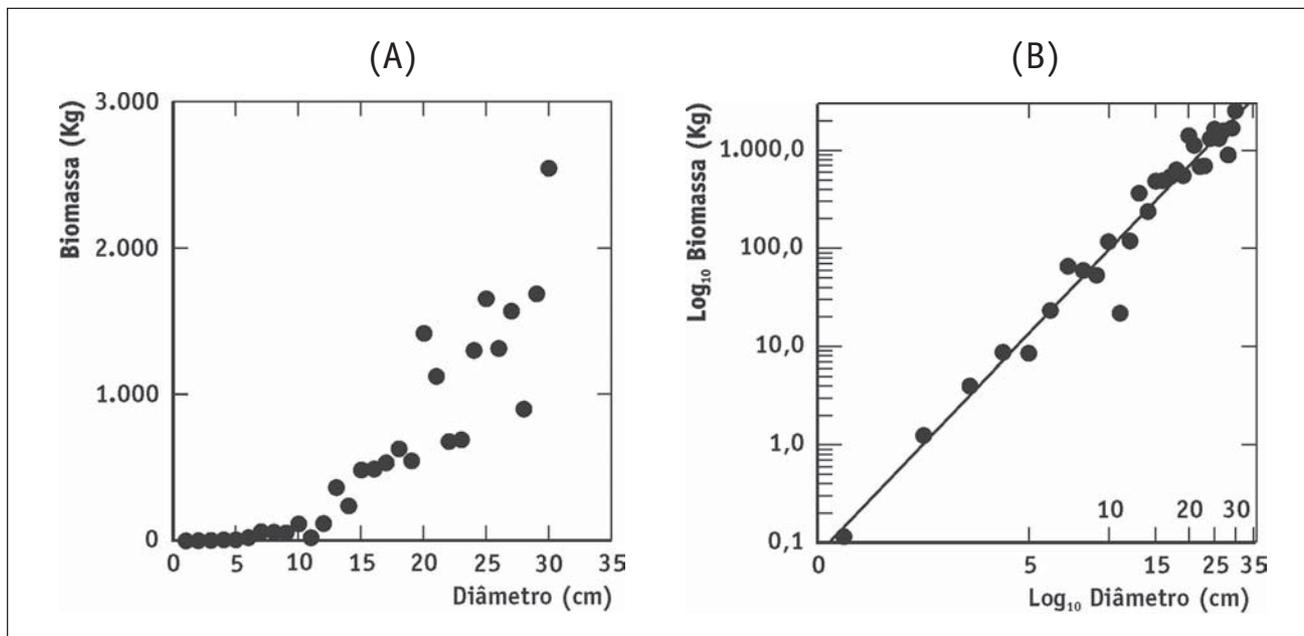
## TRANSFORMAÇÕES PARA OBTENÇÃO DE MODELOS LINEARES

Quando a relação entre a variável dependente e a variável independente não é linear, muitas vezes podemos transformar uma ou as duas variáveis para obtermos uma relação linear, como ilustrado na Figura 6, onde a transformação de potência (ambas as variáveis transformadas para o logaritmo dos valores originais) foi usada para se obter uma relação linear entre biomassa e diâmetro de árvores. Estas transformações são simples de se empregar com a ajuda de um computador, mas lembre-se que o emprego de uma transformação pode implicar em uma mudança da questão que será respondida (Noy-Mier *et al.*, 1975; Pielou, 1984; Johnson & Field, 1993).

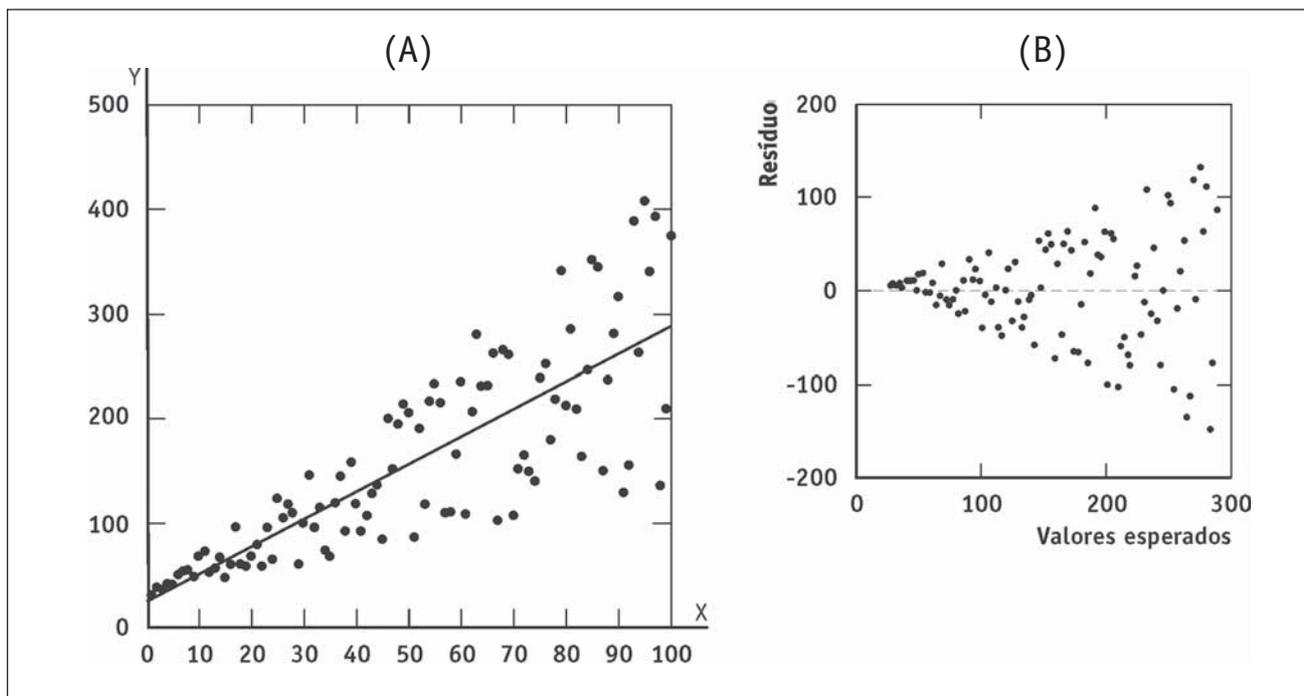
Transformações também podem ser usadas para atender outras premissas das análises, como a homogeneidade de variâncias. A Figura 6 representa um exemplo onde a premissa de normalidade ao longo da reta de regressão só foi atendida após uma transformação (Figuras 6A e B). Entretanto, a situação mais comum em sistemas biológicos é a de encontrarmos variâncias proporcionais à média, como mostrado na Figura 7.

Na realidade, os cálculos para se localizar a linha (estimar os parâmetros “*a*” e “*b*”) não dependem desta premissa de homogeneidade de variâncias, mas os testes estatísticos que seguem, para determinar se a inclinação da linha difere de zero (ou qualquer outro valor), têm esta premissa. Quando usarmos uma transformação para corrigir problemas da distribuição de resíduos, precisamos nos certificar de que não criamos problemas com outra premissa, como a linearidade, e que não introduzimos outras premissas na análise, como a exigência das relações serem multiplicativas, em vez de aditivas (Austin, 2002).

Até agora, nossas análises se basearam em modelos lineares, nos quais temos minimizado os desvios quadrados para estimarmos parâmetros. Estes métodos algébricos de mínimos quadrados são, teoricamente, apropriados somente sob um conjunto de condições muito restritivas. Eles assumem que as relações são lineares, que os efeitos dos fatores são aditivos, que



**FIGURA 6** – Exemplo de (A) uma relação não-linear entre o diâmetro e biomassa em árvores e (B) a relação linearizada por transformação logarítmica de ambas as variáveis.



**FIGURA 7** – Ao contrário do que acontece neste exemplo, modelos de mínimos quadrados requerem que a distribuição dos resíduos seja homogênea ao longo dos valores esperados.

os resíduos dos modelos têm distribuição normal, que não há erro estocástico na medida das variáveis independentes, que a variação da variável dependente é homoganeamente distribuída ao longo dos níveis da

variável independente e ainda outras condições improváveis. Quando estas condições não são satisfeitas, algum outro modelo será um melhor estimador dos parâmetros.

## EXEMPLOS PRÁTICOS

Quando nós não podemos cumprir as premissas da regressão linear de mínimos quadrados, muitas vezes podemos usar outras técnicas descritas neste volume. No entanto, estas técnicas são mais complicadas e o modelo linear de mínimos quadrados já se mostrou adequado em muitas situações. Por exemplo, Layme *et al.* (2004) usaram regressão múltipla para relacionar a distribuição do roedor *Bolomys lasiurus* com a ocorrência de queimadas, disponibilidade de alimento e estrutura de vegetação em savanas amazônicas na região de Alter do Chão, Pará. Neste trabalho, a questão principal foi verificar se a distribuição da espécie poderia ser prevista a partir de informações sobre a estrutura da vegetação e/ou presença de queimadas – fatores possíveis de serem determinados através de imagens de sensoriamento remoto – ou se seria necessário determinar a disponibilidade de alimento para o roedor, um fator muito mais oneroso e caro de ser medido. A densidade de *B. lasiurus* foi avaliada em 31 parcelas de 4 ha, distribuídas ao longo das savanas de Alter do Chão. Inicialmente, havia sido planejado amostrar 40 parcelas, mas queimadas imprevistas impediram a mensuração de todos os fatores em todas as parcelas. Os métodos multifatoriais requerem que todas as variáveis, dependentes e independentes, sejam medidas em todas as unidades amostrais. Portanto, se houver alguma unidade em que uma ou mais variáveis presentes no modelo não foi medida, esta unidade não pode entrar na análise. Portanto, um planejamento cuidadoso, com unidades de amostragem sobressalentes é sempre recomendável.

A estrutura de vegetação (cobertura de gramíneas e cobertura de arbustos) foi medida em 4 transecções de 250 m em cada parcela. Os ratos foram capturados em armadilhas tipo “Sherman” ao longo das mesmas transecções, e a disponibilidade de invertebrados foi estimada baseada na biomassa de invertebrados que caiu em armadilhas de queda (“pitfalls”). As variáveis potencialmente mensuráveis em imagens de satélite não previram a densidade de ratos com maior acerto do que era esperado para variáveis aleatórias (estrutura da vegetação –  $P = 0,13$ ; queimada no ano anterior –  $P = 0,99$ ). A disponibilidade de alimento explicou cerca de 50% da variação na densidade dos ratos ( $r^2 = 0,53$ ;  $P < 0,001$ ), mas não foi relacionada com a estrutura da vegetação ( $P = 0,71$ ).

Neste caso, como a extensão da área de estudo foi pequena (~100 km<sup>2</sup>), a premissa de que não houve autocorrelação espacial entre os valores da variável dependente e qualquer uma das variáveis independentes

(premissa de qualquer análise GLM – Legendre *et al.*, 2002) é suspeita de não ter sido atendida, mas análises subsequentes usando o teste de Mantel parcial feito por Ghizoni-Jr (2003) confirmaram o resultado. Como consequência do estudo, os pesquisadores mudaram o enfoque, deixando de tentar modelar a distribuição da espécie através de características da vegetação, para tentar entender como as flutuações temporais e espaciais na abundância de invertebrados afetam a espécie.

Regressões lineares podem ser combinadas com análises multivariadas para analisar as relações entre variáveis derivadas de técnicas de ordenação. Por exemplo, Hero *et al.* (1998) estudaram a relação entre a distribuição de comunidades de girinos, a distribuição de comunidades de predadores e as variáveis abióticas em corpos d’água na floresta amazônica. Uma ordenação foi usada para se obter uma dimensão que representasse a comunidades de girinos, uma dimensão para representar a comunidade de predadores (peixes e odonatas) e uma dimensão para representar os fatores físico-químicos. Dados de dois sítios independentes levaram aos mesmos resultados. Os predadores tinham efeitos significativos sobre as comunidades de girinos ( $P < 0,001$  em ambos os casos), mas as comunidades não eram significativamente relacionadas com os fatores físico-químicos ( $P = 0,539$  e  $P = 0,888$ ).

Um aspecto importante de se destacar é que as ordenações para se obter variáveis lineares apropriadas para GLM restringem as interpretações. Ordenações descrevem a similaridade entre comunidades, não a magnitude da comunidade. Uma comunidade não é mais ou menos que uma outra, elas somente são diferentes. O  $R^2$  da regressão reflete a magnitude do efeito, mas se o efeito é considerado positivo ou negativo é uma decisão puramente arbitrária.

Este estudo também ilustra uma limitação das análises GLM. Elas investigam somente os efeitos diretos e assumem que as variáveis independentes não afetam umas às outras. Na natureza, isto é improvável. Uma alternativa seria empregar “Análise de Equações Estruturais” (SEM) para estimar também os efeitos indiretos. Um exemplo desta abordagem pode ser encontrado em Hero *et al.* (1998), que usaram SEM para investigar os efeitos diretos e indiretos de peixes sobre a comunidade de girinos em riachos e poças na Amazônia Central. Estes autores concluíram que, embora os peixes tenham um efeito negativo direto sobre a densidade de determinadas espécies de girinos, eles podem propiciar o aumento da densidade destas espécies através de efeitos indiretos. Peixes comem girinos, mas peixes comem e controlam as populações de larvas de insetos, como

odonatas e besouros, que são predadores muito mais vorazes dos girinos. O efeito líquido favorece os girinos na presença de peixes. McCune & Grace (2002) explicam como as “Análises de Equações Estruturais” podem ser vistas como uma extensão dos modelos lineares (GLM).

## REFERÊNCIAS BIBLIOGRÁFICAS

- Austin, M.P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling* 157: 101-118.
- Ghizoni-Jr, I.R., Jr. 2003. Efeito das variações nas escalas temporais e espaciais na dinâmica de populações de *Bolomys lasiurus* (RODENTIA: MURIDAE) em uma savana na Amazônia Central. Dissertação de Mestrado. INPA/UFAM.
- Hero, J.M., C. Gascon & W.E. Magnusson. 1998. Direct and indirect effects of predation on tadpole community structure in an Amazonian rainforest. *Australian Journal of Ecology* 23: 474-482.
- Johnson, C.R. & C.A. Field. 1993. Using fixed-effects model multivariate analysis of variance in marine biology and ecology. *Oceanography and Marine Biology Annual Review* 31: 177-221.
- Layme, V.M.G., A.P. Lima & W.E. Magnusson. 2004. Effects of fire, food availability and vegetation on the distribution of the roedent *Bolomys lasiurus* in an Amazonian savanna. *Journal of Tropical Ecology* 20: 183-187.
- Legendre, P., M.R.T. Dale, M.J. Fortin, J. Gurevitch, M. Hohn & D. Myers. 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography* 25: 601-615.
- Magnusson, W.E. & G. Mourão. 2003. *Estatística sem Matemática*. Editora Planta, Londrina.
- McCune, B. & J.B. Grace. 2002. *Analysis of Ecological Communities*. MjM Software Designs, Gleneden Beach, EUA.
- Noy-Meir, I., D.Walker & W.T. Williams. 1975. Data transformations in ecological ordination II. On the meaning of data standardization. *Journal of Ecology* 63: 779-800.
- Pielou, E.C. 1984. *The interpretation of ecological data*. Wiley, New York.

---

# Modelos de classificação em árvore para a localização de áreas de ocorrência potencial de espécies

ADRIANA PAESE<sup>1\*</sup>

SIDNEY DORNELLES<sup>2</sup>

JOSÉ EDUARDO DOS SANTOS<sup>2</sup>

JOSÉ SALATIEL RODRIGUES PIRES<sup>2</sup>

<sup>1</sup> Conservação Internacional, Minas Gerais, Brasil.

<sup>2</sup> Universidade Federal de São Carlos – UFSCar, São Paulo, Brasil.

\* e-mail: a.paese@conservation.org.br

## RESUMO

Este trabalho apresenta um método para a localização de habitats de maior probabilidade de ocorrência de três espécies de primatas em áreas naturais de conservação da região nordeste do Estado de São Paulo, utilizando dados sobre a localização das espécies e de algumas variáveis ambientais analisadas com base em modelos de classificação por árvore. Os modelos resultantes mostram que o tipo de cobertura do solo foi a principal variável ambiental associada à presença das três espécies. A introdução dos resultados em ambiente SIG permitiu identificar outros locais com as mesmas combinações das variáveis ambientais em que as espécies foram localizadas. Embora seja necessário um número maior de observações para o desenvolvimento de modelos mais satisfatórios, ao menos para uma das espécies estudadas os resultados constituem modelos exploratórios importantes, que poderão subsidiar futuros trabalhos de campo, possibilitando, ainda, a inclusão de mais informações sobre a biologia das espécies e sobre as alterações ambientais locais.

## ABSTRACT

*This study describes a method of locating sites, in natural protected areas situated to the northeast region of São Paulo State, where the probability of occurrence of three primate species is highest. It analyses data on species occurrence and on landscape variables, developed in a Geographic Information System (GIS) using classification tree models. The resulting models showed that land cover is the landscape variable most strongly associated with the species. They also generated prediction rules for the species potential sites of occurrence, which are projected back onto the landscape, locating other places that have the same combination of environmental conditions. Although a great number of observations is necessary to develop better models for at least one of the species in the study, the models represent a valuable exploratory guide, which should facilitate prospective field surveys. They can be easily updated in a GIS to incorporate additional information on the species biology and on environmental change.*

## INTRODUÇÃO

Dados sobre a ocorrência de espécies são escassos, especialmente em regiões tropicais, onde a grande diversidade biológica dos ecossistemas torna esse tipo de inventário uma tarefa onerosa e inviável em curto prazo. Quando existentes, eles são freqüentemente apresentados como áreas de ocorrência potencial de espécies, estimadas com base na localização de barreiras geográficas e na relação entre a localização de populações e alguma variável climática. Raramente essas informações são atualizadas, não acompanhando as alterações antrópicas ao ambiente. Apesar da dificuldade na obtenção de dados precisos e atualizados sobre a ocorrência de espécies, existe uma grande demanda por informações detalhadas, que devem subsidiar o desenvolvimento de projetos conservacionistas em grandes áreas (Brooks *et al.*, 2004).

Os modelos de distribuição que relacionam a localização de espécies a conjuntos de variáveis ambientais, definindo o seu hábitat potencial ou sua área de ocorrência potencial são apresentados como uma alternativa para a dificuldade na obtenção de dados pontuais sobre a ocorrência de espécies. Modelos preditivos têm também sido empregados como modelos neutros para o entendimento dos efeitos da fragmentação (Milne *et al.*, 1989) e em trabalhos relacionados com a ocorrência e conservação de espécies (Davis & Goetz, 1990; Moore *et al.*, 1991; De'Ath & Fabricius, 2000).

Um exemplo dos modelos de distribuição potencial de espécies são as árvores de decisão, ou mais formalmente, os modelos de classificação e regressão em árvore (Moore *et al.*, 1990; Flick, 1998; Vayssières *et al.*, 2000; Urban *et al.*, 2002). Os modelos de classificação em árvore são técnicas ainda pouco utilizadas na ecologia e representam uma alternativa não paramétrica a outras técnicas estatísticas. Entre as vantagens dos modelos em árvore destacam-se: a flexibilidade para lidar com variáveis numéricas e categóricas, a facilidade de interpretação e a possibilidade de explorar dados não lineares e interações complexas entre variáveis (De'Ath *et al.*, 2000).

As árvores de decisão constituem modelos conceituais dos padrões ou processos na paisagem. Um benefício da utilização dessa estratégia é a formalização do modelo, que força o pesquisador a ser específico sobre as suas idéias, orientando-o na coleta de dados. A amostragem de uma árvore de decisão pode ser estruturada em um sistema de informações geográficas, por meio da transcrição da combinação das condições que definem o hábitat. Depois de mapeadas, essas condições

podem ser sub-amostradas por meio da utilização de outro método de amostragem (Urban, 2002).

O presente trabalho tem como objetivo demonstrar a utilização da classificação em árvore para a estimativa das áreas de ocorrência potencial de três espécies de primatas da área da Estação Ecológica de Jataí e da Estação Experimental de Luiz Antônio. O modelo estatístico foi desenvolvido procurando estabelecer as relações entre os pontos de localização das espécies e as combinações das variáveis ambientais disponibilizadas em um banco de dados georreferenciado, resultando no mapeamento, na paisagem, das relações ambientais estatisticamente definidas.

## MATERIAL E MÉTODOS

### Área de estudo

A área de estudo compreende a Estação Ecológica de Jataí (EEJ) e a Estação Experimental de Luiz Antônio (EELA), localizadas no município de Luiz Antônio, na região Nordeste do Estado de São Paulo, entre as coordenadas 21°30' e 21°40' de latitude Sul e 47°40' e 47°50' de longitude Oeste (Figura 1). Com a denominação de



**FIGURA 1** – Localização da Estação Ecológica de Jataí (EEJ) e da Estação Experimental de Luiz Antônio (EELA) (Modificado de Pires, 2000).

Parque Estadual de Jataí, esta área foi objeto de um plano de manejo direcionado à conservação da biodiversidade no contexto regional (Pires, 2000). A área se destaca por constituir um dos poucos remanescentes da vegetação nativa, suficientemente extensa para abrigar uma grande variedade de ecossistemas (Cavaleiro *et al.*, 1990) e uma grande diversidade de espécies, algumas das quais ameaçadas de extinção no Estado de São Paulo (Talamoni, 1996; Dalmolin, 1999; Dias, 2000; Dornelles, 2000, Talamoni *et al.*, 2000; Mantovani, 2001). Na área estudada, a vegetação nativa é composta por diferentes tipos fisionômicos de Cerrado e também pelas matas semidecídua, ciliar e de galeria. O histórico de perturbações antrópicas parece ser o principal fator causador dos padrões espaciais da vegetação na área.

### Procedimentos experimentais

**ESPÉCIES E VARIÁVEIS AMBIENTAIS** – As espécies de primatas usadas no desenvolvimento deste trabalho, o sauá, *Callicebus nigrifrons* (Spix, 1823), o bugio, *Alouatta caraya* (Humboldt, 1812), e o macaco-prego, *Cebus nigritus* (Goldfuss, 1809), foram anteriormente estudadas em termos da estimativa do tamanho populacional, da análise da utilização do hábitat e do mapeamento de sua distribuição para a área de estudo em questão (Dornelles, 2000).

Apesar de outros estudos terem relacionado a ocorrência de espécies a variáveis ambientais (variáveis da paisagem) na EEJ e EELA (Talamoni, 1996; Dalmolin, 1999; Dias, 2000; Motta-Junior, 2000; Motta-Junior & Alho, 2000; Talamoni *et al.*, 2000), somente o estudo de Dornelles (2000) incluiu informações explícitas sobre a distribuição das espécies, tendo sido, por esse motivo, utilizado como base para a realização deste estudo.

No presente trabalho, as seguintes variáveis ambientais foram utilizadas para modelar as exigências dessas espécies com relação ao ambiente: hipsometria, clinografia, aspecto (grau de exposição da rampa), índice de convergência topográfica (ICT), cobertura da terra, distância das áreas antrópicas, distância dos corpos d'água, distância das estradas principais, e densidade de estradas. Estas variáveis foram selecionadas por estarem disponibilizadas no acervo cartográfico do Laboratório de Análise e Planejamento Ambiental (LAPA) da UFSCar, bem como por serem consideradas potencialmente importantes como preditoras diretas da distribuição das espécies de primatas ou de espécies vegetais, sendo, neste caso, consideradas preditoras indiretas da distribuição das espécies de primatas.

No banco de dados digital do LAPA (BD-Jataí – Pires, 1994) foram obtidas as seguintes camadas de informação (layers): o modelo digital de elevação (MDE) e os mapas de hidrografia e de localização de estradas na área de estudo, digitalizados com base nas cartas topográficas do IBGE em escala 1:50.000 (Pires, 2000). Com base no modelo digital de elevação (MDE), foram derivadas outras quatro variáveis topográficas: hipsometria, clinografia, aspecto e índice de convergência topográfica. De acordo com Moore *et al.* (1991), a estrutura e a composição da vegetação estão frequentemente relacionadas a estas variáveis, as quais refletem as condições climáticas locais e as condições edáficas na paisagem.

A hipsometria é um determinante da distribuição da maioria das espécies, estando diretamente relacionada à temperatura e à precipitação, agindo como um substituto a estas duas variáveis, difíceis de serem estimadas (Moore *et al.*, 1991). A clinografia é um dos índices topográficos mais frequentemente calculados para a análise ambiental. Ela afeta o fluxo e o tempo de residência da água ou da umidade na paisagem. O aspecto é a direção da declividade e fornece informações sobre incidência solar, condições térmicas e exposição de diferentes locais. O índice de convergência topográfica (ICT) é um índice do acúmulo relativo da umidade na paisagem (Wolock, 1993), medindo a tendência da água de superfície acumular-se ou não em diferentes locais. O ICT calcula a área “morro acima” que contribui com o fluxo de água para determinado local e a declividade deste local. Seu cálculo é realizado pela expressão  $\ln(a/\tan b)$ , onde  $a$  é a área de contribuição “morro acima” e  $\tan b$  é a declividade local. Altos valores do ICT representam áreas para onde há grande probabilidade da água convergir (canais de córregos), enquanto que valores inferiores representam áreas em que a probabilidade da água convergir ou se acumular é baixa (topo de morros). O índice de convergência topográfica foi calculado usando o modelo digital de elevação e uma série de comandos do programa ARC/Info (ESRI, 2000) e programas escritos na linguagem FORTRAN de programação. O cálculo do ICT foi efetuado de acordo com os procedimentos definidos por Wolock (1993).

O mapa de cobertura da terra foi derivado de imagens do satélite Landsat TM5 datadas de 1997 e 1998. Estas imagens foram classificadas em 18 categorias: área antrópica (área contendo infra-estrutura rural), campo sujo, capoeira, capoeirão, cerrado alto, cerradão alto aberto, cerradão baixo, plantação de *Eucalyptus* com sub-bosque de vegetação de cerrado, plantação de *Eucalyptus* sp., mata ciliar, mata de galeria, mata de

galeria em regeneração, mata semidecídua, mata semidecídua em regeneração, mata de transição baixa, plantação de *Pinus* sp. e área de várzea e água, que inclui reservatórios, lagoas marginais, lagos, córregos e rios (Dornelles, 2000).

A distância das áreas antrópicas, a distância dos corpos d'água e a distância das estradas principais foi calculada no programa ARC/Info (ESRI, 2000), com base na condução de uma análise de proximidade. Para cada uma das características acima descritas foi derivada uma superfície contínua em que o valor de cada pixel representa a distância euclidiana da área antrópica, do corpo d'água ou da estrada mais próxima.

A densidade de estradas foi calculada para cada classe de cobertura da terra pela intersecção do mapa de localização de estradas secundárias com o arquivo de uso e cobertura da terra. Desta forma, foi calculado o comprimento total das estradas em cada região, delimitada por um determinado tipo de uso ou cobertura da terra. A densidade de estradas foi calculada dividindo o comprimento total das estradas em cada região, por sua área em m<sup>2</sup>.

**A METODOLOGIA DA ANÁLISE DE CLASSIFICAÇÃO E REGRESSÃO POR ÁRVORE** – Para a obtenção do mapa da ocorrência potencial das espécies de primatas foi utilizado um modelo de classificação por árvore com o auxílio do pacote estatístico SPSS- Answer Tree e do SIG ARC/Info (ESRI, 2000). Diferentemente dos modelos lineares convencionais, que utilizam relações lineares entre as variáveis preditoras (x) e a variável resposta (y), os modelos de análise e regressão por árvore dividem o conjunto dos dados em grupos hierárquicos, os quais apresentam diferentes valores da variável resposta. Além disso, em modelos de regressão linear as exceções à regra preditora dos valores de y aparecem como resíduos, enquanto que nos modelos de regressão por árvore as exceções às regras preditoras são caracterizadas mais explicitamente (Breiman *et al.*, 1984).

Existem pelo menos três técnicas de classificação e regressão por árvore que diferem quanto ao método empregado para a divisão do conjunto de dados. Breiman *et al.* (1984) sugerem que, para testar efetivamente a precisão do modelo, mais de uma técnica deve ser utilizada. Neste estudo, foram aplicadas duas técnicas para a análise de classificação e regressão por árvore: a técnica CART (Classification And Regression Trees) e CHAID (Chi-Square Automatic Iterative Detection). A técnica CART divide os dados repetidamente e se-

qüencialmente, de forma que os subgrupos resultantes de cada divisão apresentam entre si a maior heterogeneidade possível e a maior homogeneidade interna. Essa técnica utiliza um modelo de classificação quando a variável resposta y é categórica e um modelo de regressão quando a variável resposta é numérica. Neste trabalho, a variável resposta foi categorizada como bugio, sauá, prego ou acaso. Portanto, foi utilizado um modelo de classificação por árvore. Diferentemente da técnica CART, que permite apenas partições binárias do conjunto de dados, a técnica CHAID permite a sua divisão em dois ou mais grupos por vez. Esta técnica é baseada em testes Qui-Quadrado (x<sup>2</sup>), os quais são aplicados seqüencialmente. Após cada aplicação, a variável preditora que apresenta a maior associação com a variável resposta é escolhida até que não ocorra mais uma associação significativa entre as variáveis preditora e resposta.

Modelos de distribuição de espécies geralmente utilizam dois conjuntos de dados, um contendo amostras dos locais onde a espécie foi observada (presença) e o outro contendo amostras de localizações onde a espécie não foi encontrada (ausência). Estes modelos representam o maior contraste das relações entre a espécie e o seu hábitat. A hipótese nula testada nestes casos é a de que os dois conjuntos de dados representam amostras equivalentes em um universo maior de habitats disponíveis.

Outra possibilidade é a utilização de conjuntos de dados com amostras dos locais onde a espécie foi observada (presença) e com pontos distribuídos ao acaso, representando amostras dos tipos de hábitat disponíveis. A hipótese nula testada nestes casos é a de que a espécie não apresenta preferências por tipos de hábitat distintos, ou seja, ela utiliza diferentes habitats na proporção em que eles são encontrados.

No presente trabalho, para a aplicação dos testes, foram amostradas no SIG ARC/Info todas as variáveis ambientais nos locais precisos onde as espécies foram observadas em campo. Para um número igual de localizações de cada animal foram amostrados pontos com localização determinada ao acaso. Nesta etapa, a variável resposta foi categorizada como espécie ou acaso. Esse procedimento permitiu diferenciar as condições ambientais nos pontos onde as espécies foram observadas daquelas dos pontos cuja distribuição é aleatória. Procurou-se assim, identificar preferências da espécie por diferentes tipos de hábitat.

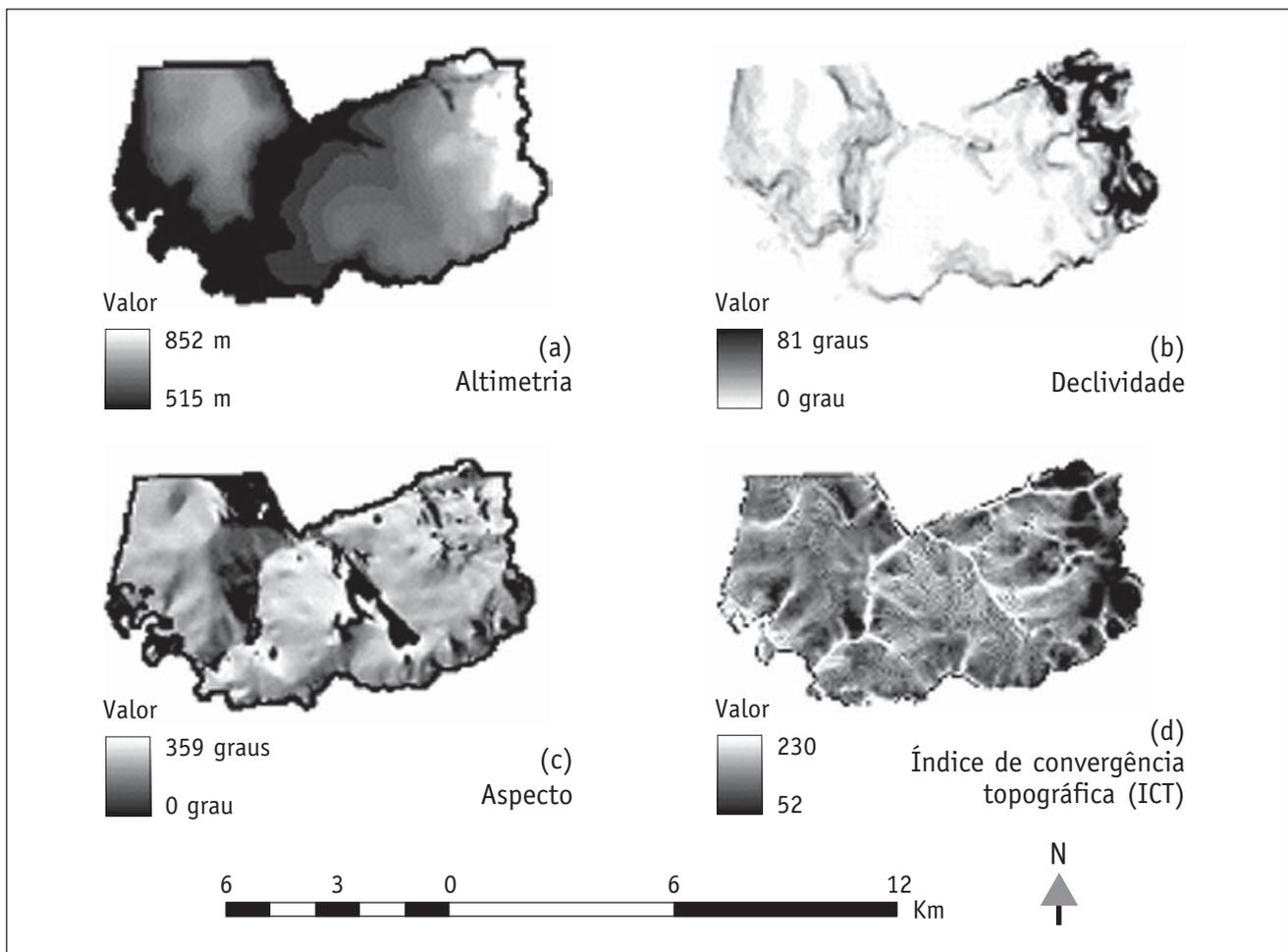
As observações referentes a cada espécie foram analisadas, separadamente, usando a técnica CART e CHAID.

Posteriormente às análises, foram aplicados testes de validação cruzada, com o objetivo de estimar o erro ou taxa de má-classificação (m.c.) de cada modelo. Os testes de validação cruzada dividem aleatoriamente o conjunto de dados em sub-amostras, utilizadas para verificar a probabilidade de que um de seus elementos seja classificado erroneamente. A média dessa probabilidade para todas as amostras é a taxa de má-classificação (Breiman *et al.*, 1984).

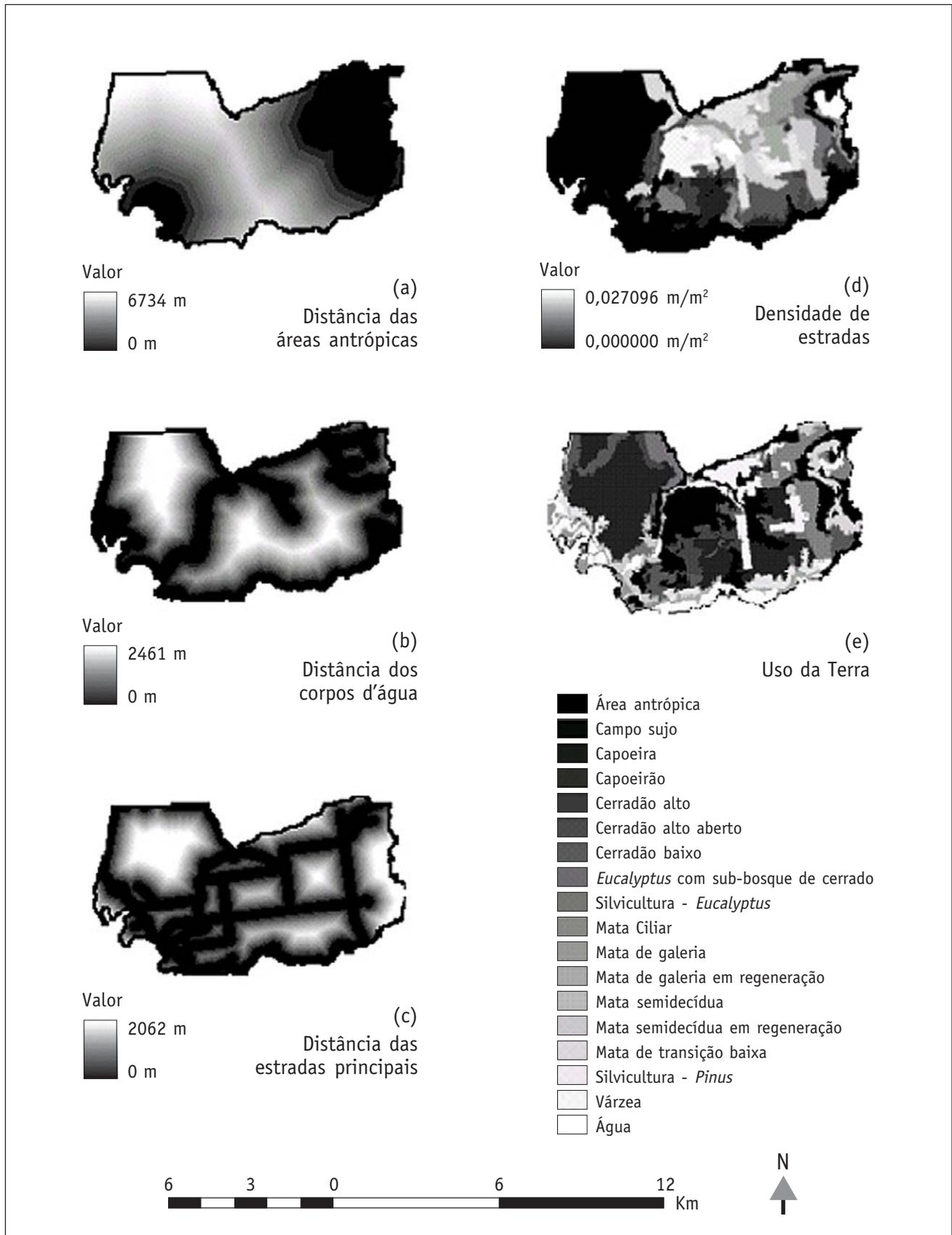
A leitura das árvores foi feita da raiz em direção aos nós terminais. As regras predictoras das relações entre as espécies e as variáveis ambientais que apresentaram as menores taxas de má-classificação foram projetadas na paisagem por meio de *scripts* escritos na linguagem macro do SIG ARC/Info. Este procedimento teve como objetivo localizar, na área de estudo, outros locais com as mesmas combinações das variáveis ambientais em que as espécies foram localizadas.

## RESULTADOS

Na área que compreende a EEJ e a EELA, a altitude varia entre 515 e 852 metros, estando as áreas mais altas localizadas a nordeste e as mais baixas no limite da EEJ com o rio Mogi-Guaçu (Figura 2a). As áreas de maior declividade atingem 81 graus e estão localizadas na direção nordeste, na Estação Experimental (Figura 2b). Os valores do aspecto (Figura 2c) indicam que a maior extensão do relevo na área de estudo está voltada para a direção sudoeste. Os valores do índice de convergência topográfica variaram de 52 a 230, com os valores mais altos localizados em regiões planas próximas aos córregos, e os mais baixos nas regiões mais elevadas e com maior declividade (Figura 2d). A área de estudo apresenta pequena variação altimétrica e, por esse motivo, os índices topográficos derivados do modelo digital de elevação não apresentam grande variação.



**FIGURA 2** – Variáveis topográficas utilizadas nos modelos de classificação em árvore: (a) altimetria; (b) declividade; (c) aspecto; e (d) índice de convergência topográfica (ICT).



**FIGURA 3** – Variáveis ambientais utilizadas nos modelos de classificação em árvore: (a) distância das áreas antrópicas; (b) distância dos corpos d'água; (c) distância das estradas principais; (d) densidade de estradas; e (e) uso da terra.

A distância de qualquer local para a área urbanizada mais próxima variou de 0 a 6734 m (Figura 3a). O ponto mais distante dos corpos d'água está localizado a 2461 m do córrego mais próximo (Figura 3b). As áreas mais isoladas (Figura 3c) estão localizadas a 2062 m de qualquer estrada principal. A densidade de estradas pode variar de 0m/m<sup>2</sup> a 0,02796m/m<sup>2</sup>. Área antrópica é o tipo de uso do solo em que se configura a maior densidade de estradas (Figura 3d). As diferentes categorias de cobertura da terra identificadas na área de estudo (Dornelles, 2000) estão representadas na Figura 3e.

Os modelos que apresentaram as menores taxas de má-classificação (m.c.) para as três espécies são resultados da análise do conjunto de dados pela técnica CART (Figuras 4, 5 e 6). Entre as espécies estudadas, o modelo resultante da análise de classificação em árvore para o bugio foi o que apresentou a menor taxa de má-classificação. A análise individual dos dados dessa espécie resultou em uma taxa de má-classificação de 27% para a técnica CART e 38% para a técnica CHAID. O bugio está mais associado aos ambientes de mata ciliar, mata semidecídua, mata semidecídua em regeneração, e mata de transição baixa (Figura 4).

Os modelos de classificação em árvore do saúá apresentaram taxas de má-classificação de 37% para a técnica CART e de 47% para a técnica CHAID. De acordo com as regras de predição da localização desta espécie, ela ocorre preferencialmente em áreas como campo sujo, cerradão alto, *Eucalyptus* com sub-bosque de cerrado em regeneração, mata ciliar, mata de galeria, mata de galeria em regeneração, mata semidecídua, ou mata de

transição baixa, e em áreas em que a distância aos corpos d'água é menor ou igual a 1598 m e a densidade de estradas é menor ou igual a 0,00022 m/m<sup>2</sup>. A espécie pode também ocorrer nas áreas ocupadas pelos mesmos tipos de cobertura da terra, cuja densidade de estradas, porém, é maior que 0,00022 m/m<sup>2</sup> se a altitude for maior que 604 m; ou em áreas ocupadas pelas mesmas coberturas da terra, se a distância aos corpos d'água for maior que 1830 m (Figura 5).

Considerando-se que a área de estudo apresenta uma pequena variação em sua altitude, a hipsometria não deveria ser um fator limitante na distribuição de nenhuma das espécies estudadas. Entretanto, ela aparece no modelo de classificação por árvore, provavelmente por estar relativamente correlacionada com outras variáveis que explicariam melhor a distribuição dos primatas, como a distância aos corpos d'água ( $r = 0,317$  e  $p < 0,01$ ) e a distância das estradas principais ( $r = 0,588$  e  $p < 0,01$ ).

As análises de classificação em árvore para o macaco-prego mostram que ele está mais associado com dois tipos de cobertura da terra: cerradão alto e mata semidecídua (Figura 6). Entretanto, estes resultados devem ser utilizados com cautela, uma vez que para esta espécie os resultados obtidos foram os menos satisfatórios. As taxas de má-classificação para os modelos em árvore do prego foram de 63% (técnica CART) e 78% (técnica CHAID). Esses resultados indicam que o número de observações para esta espécie foi insuficiente para determinar, por meio da análise de classificação em árvore, relações precisas e detalhadas entre as mesmas e o ambiente.

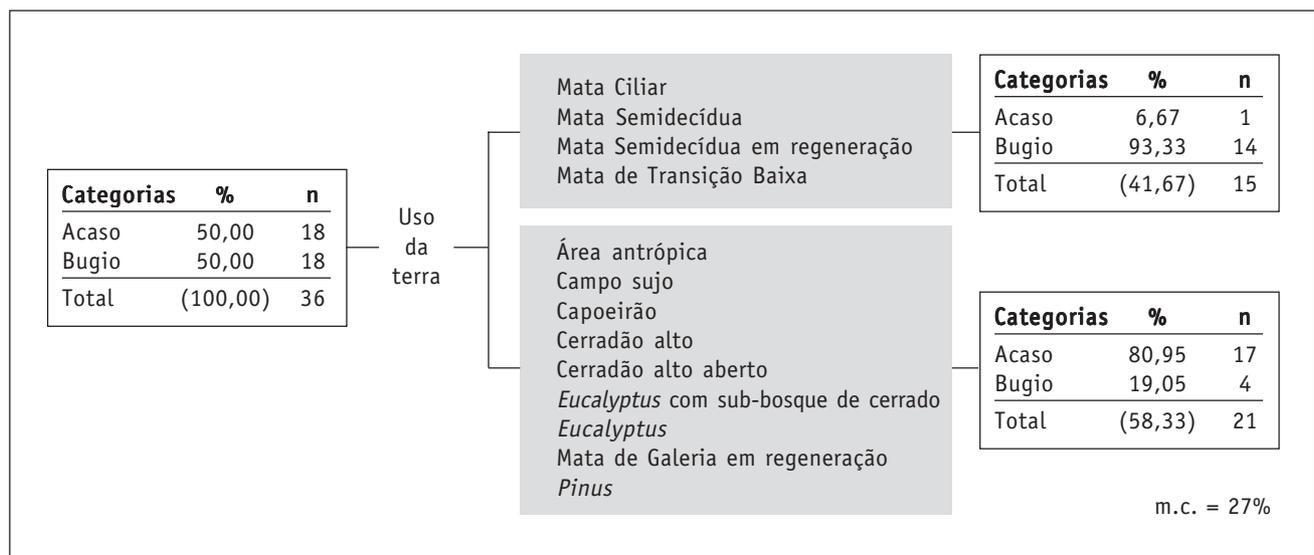


FIGURA 4 – Modelo de classificação em árvore da distribuição do bugio (*Alouatta caraya*) na área de estudo.

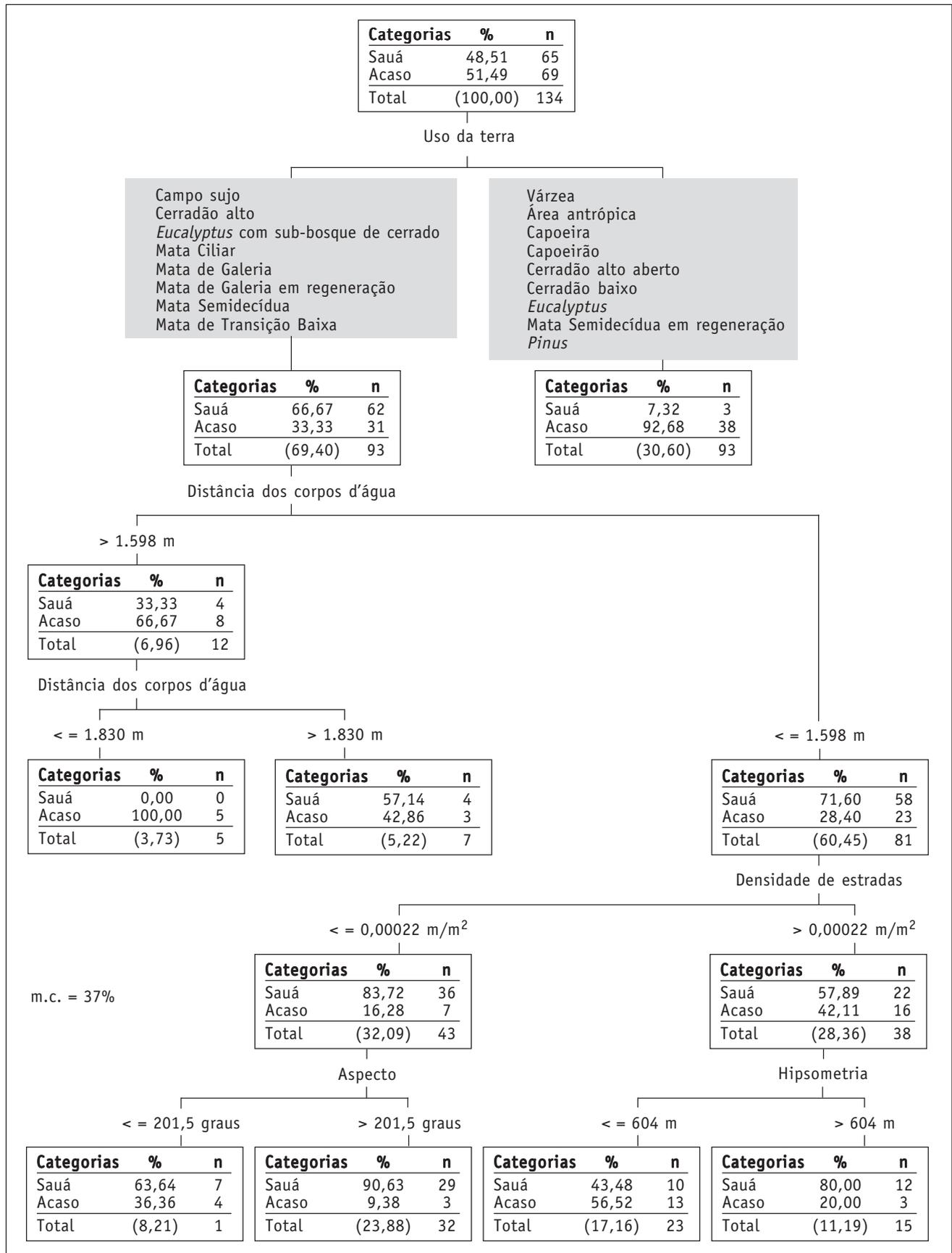
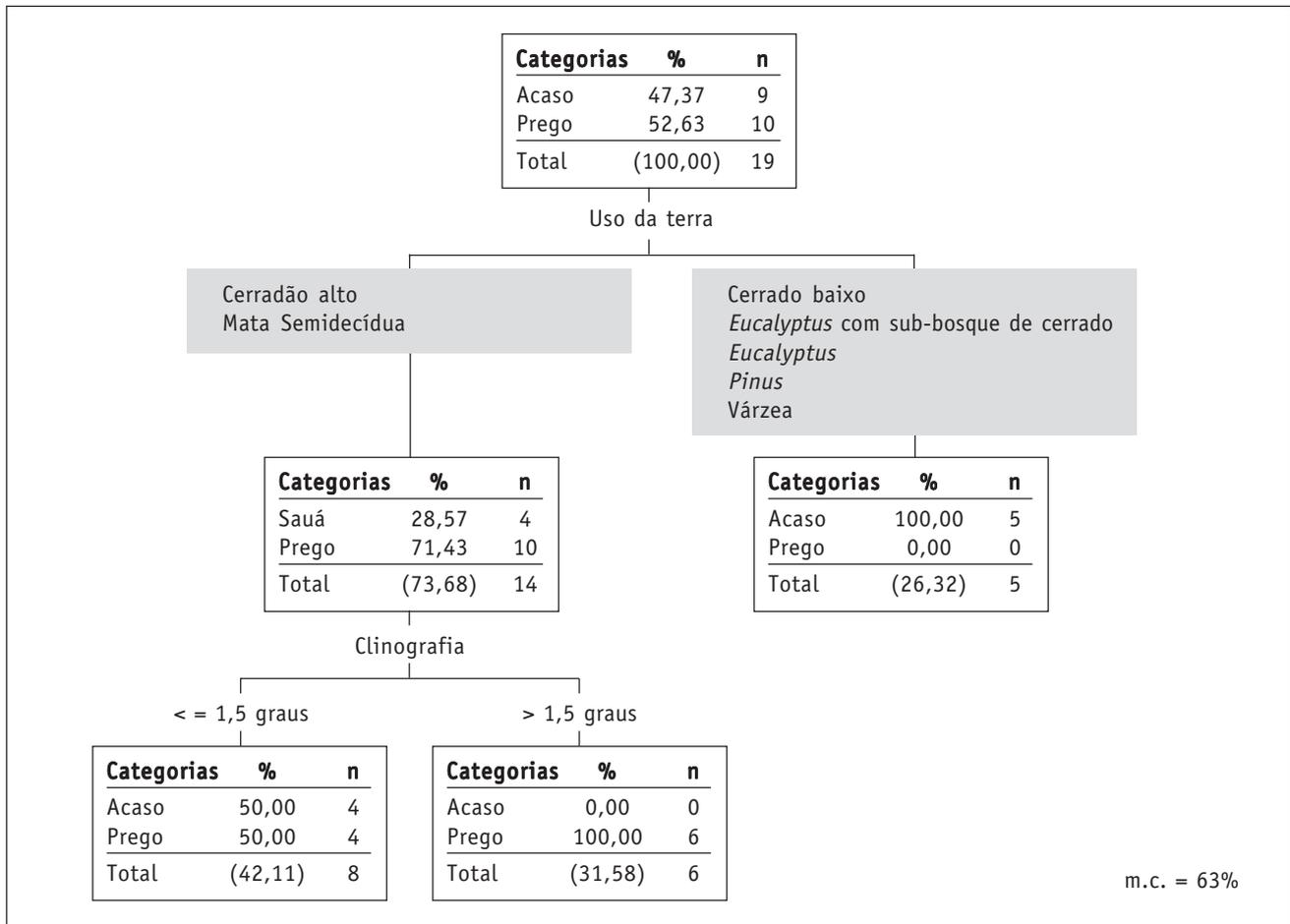


FIGURA 5 – Modelo de classificação em árvore para o sauá (*Callicebus nigrifrons*) na área de estudo.



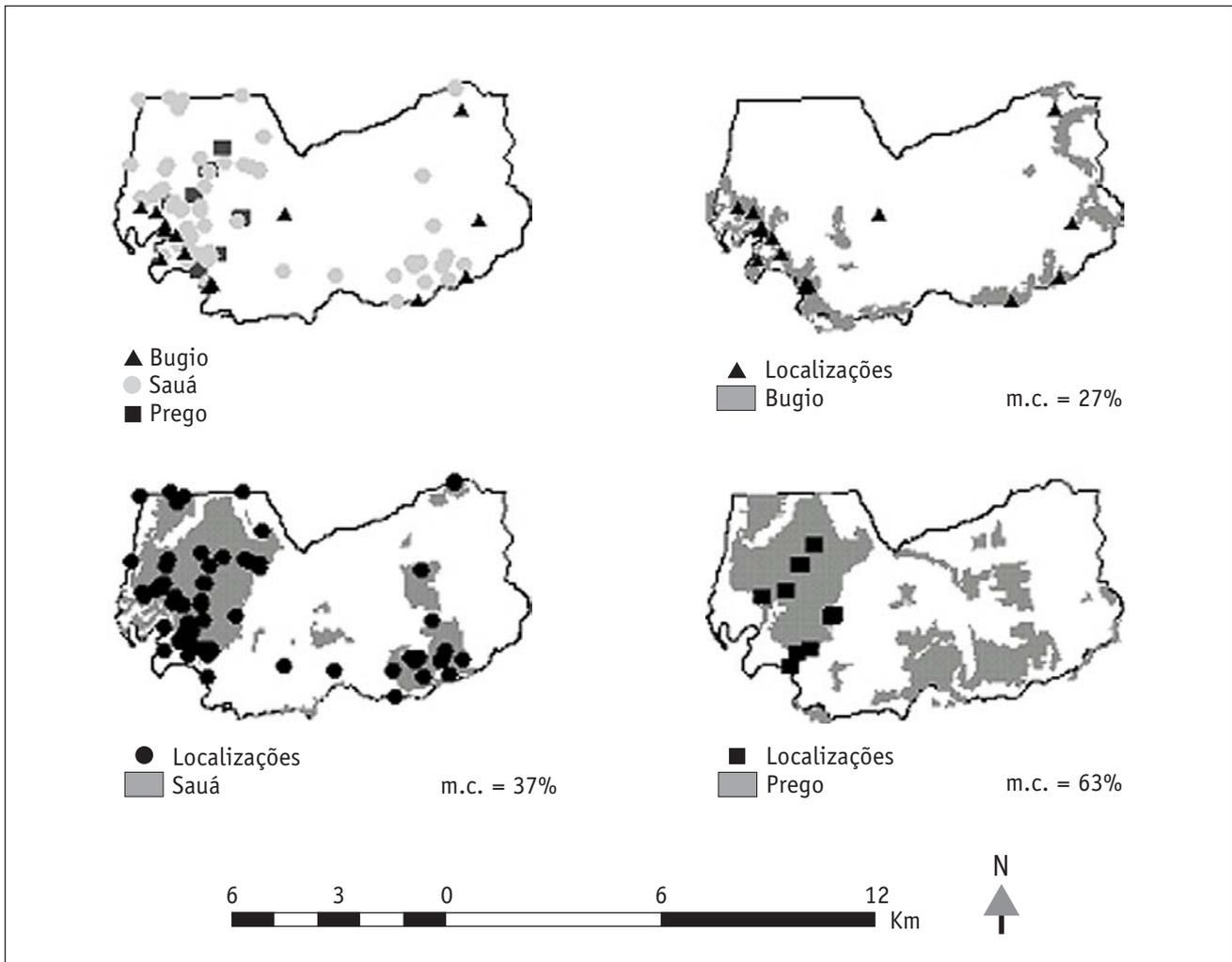
**FIGURA 6** – Modelo de classificação em árvore da distribuição do macaco-prego (*Cebus nigritus*) na área de estudo.

Outros locais que apresentam as mesmas combinações de condições ambientais definidas pelos modelos de classificação por árvore são mostrados na Figura 7. Estas áreas representam os locais em que há maior probabilidade das espécies de primatas serem encontradas: menor para o bugio e maior para o sauá. Comparando-se as áreas de maior probabilidade de ocorrência das espécies estudadas com o zoneamento proposto no plano de manejo da EEJ e da EELA (Pires, 2000), foi possível verificar que a área de maior probabilidade de ocorrência do macaco-prego (*Cebus apella*) está contemplada na zona intangível, enquanto que as áreas de maior probabilidade de ocorrência do sauá (*Callicebus personatus nigrifrons*) e do bugio (*Alouatta caraya*) estão contempladas nas zonas intangíveis e de recuperação. A provável ocorrência dessas duas espécies na zona de recuperação ressalta a necessidade da recuperação natural do ecossistema, por meio dos processos de sucessão ecológica ou de ações de recuperação, como propostas por Pires (2000).

## DISCUSSÃO

Ambas as técnicas utilizadas para o desenvolvimento do modelo em árvore mostraram-se pouco adequadas para a análise de conjuntos de dados cujo número de observações é pequeno, como ocorre neste trabalho: 65 para o sauá, 18 para o bugio e 10 para o prego. Na área estudada, o pequeno número de indivíduos destas espécies pode estar relacionado principalmente com as atividades de caça, com queimadas e com a fragmentação de habitats pela presença de um grande número de estradas (Dornelles, 2000).

De forma geral, o poder analítico dos modelos de distribuição de espécies está diretamente relacionado ao número de observações. A redução do número de áreas amostradas aumenta a probabilidade de que variáveis pouco significativas sejam incluídas e de que variáveis mais relevantes para explicar a ocorrência das espécies sejam excluídas do modelo (Vaughan & Ormerod, 2003; Miller *et al.*, 2004). Estudos consideram



**FIGURA 7** – Localização das três espécies de primatas (Dorneles, 2000), projeção das áreas de maior probabilidade de localização destas espécies na área de estudo e taxas de má classificação (m.c.) dos modelos de classificação em árvore.

satisfatório o número mínimo de 10 observações para cada variável preditora empregada (Vaughan & Ormerod, 2003).

Como resultado do pequeno número de observações, os modelos para o bugio, o sauá e o prego (Figuras 4, 5 e 6) detectaram um número pequeno de variáveis preditoras, sendo que a variável cobertura da terra foi a única incluída nos modelos do bugio e do prego. A cobertura da terra foi a variável mais importante para explicar as exigências das três espécies de macacos com relação ao ambiente. Entretanto, as espécies consideradas parecem ser pouco sensíveis a esta variável. Todas ocorrem em ambientes fragmentados e secundários e são generalistas quanto à utilização do habitat (Mittermeier *et al.*, 1988). Um maior detalhamento de aspectos estruturais e florísticos da vegetação poderiam facilitar a interpretação das relações entre os

diferentes tipos de cobertura do solo e as espécies de primatas.

O ICT, o aspecto, e a distância das áreas antrópicas não foram associados com a ocorrência das espécies em nenhum dos modelos. Como mencionado, provavelmente, isto se deve ao número reduzido de observações analisadas. Outro fator a ser considerado é que as variáveis ambientais utilizadas, em especial os índices topográficos, representam 'variáveis antropogênicas da paisagem', isto é, variáveis com baixa resolução, que podem ser facilmente mapeadas ou reconhecidas pelos seres humanos (Milne *et al.*, 1989), mas que não representam aspectos relevantes da biologia das espécies, como por exemplo, a localização de algum item importante da dieta alimentar das mesmas. A falta do mapeamento preciso de variáveis biologicamente significantes, as quais têm frequentemente extensões

muito pequenas para serem representadas em bancos de dados com grandes escalas, representa uma fonte potencial de erro nos modelos preditivos da distribuição de espécies (Lynn *et al.*, 1995). Além disso, como mediadores entre as condições edáficas e micro-climáticas, os índices topográficos provavelmente refletem mais satisfatoriamente as exigências das espécies vegetais do que as das espécies de primatas com relação ao ambiente. O uso de variáveis que apresentam relações indiretas com a distribuição de espécies dificulta a interpretação dos modelos de distribuição. Apesar disso, estas variáveis são comumente utilizadas por serem facilmente derivadas de imagens de sensoriamento remoto e por estarem prontamente disponíveis (Miller *et al.*, 2004).

Os modelos de distribuição extrapolam da ocorrência das espécies, com base no entendimento das suas relações com o ambiente. Existem disparidades entre as escalas refinadas em que os estudos ecológicos são conduzidos e as escalas mais amplas em que é realizado o planejamento para a conservação. Neste sentido, há uma grande expectativa de que os modelos de distribuição possam substituir os dados sobre a ocorrência de espécies, contornando as dificuldades logísticas da obtenção destes dados. Contudo, como demonstra o presente trabalho, aspectos metodológicos importantes devem ser considerados para o desenvolvimento de modelos satisfatórios.

Um grande número de programas de computador tem sido desenvolvido com a finalidade de gerar mapas da ocorrência potencial de espécies, com base em um conjunto de variáveis espacialmente representadas. Ao mesmo tempo, os modelos disponíveis e suas potenciais fontes de erro têm sido avaliados por diferentes autores (Loiselle *et al.*, 2003; Vaughan & Ormerod, 2003, entre outros). Mapas da presença/ausência de espécies geralmente são apresentados como resultados definitivos dos modelos de distribuição. Contudo, uma maior importância deve ser dada ao conjunto de hipóteses sobre as relações entre a espécie e o ambiente desvendado pelo modelo. Estas hipóteses podem ser testadas posteriormente, direcionando a amostragem da paisagem. O conjunto de hipóteses sobre a distribuição das espécies está explicitamente definido na estrutura do modelo de classificação em árvore, podendo ser diretamente mapeadas em um sistema de informações geográficas. Isto os torna especialmente úteis para o entendimento da combinação de condições que definem a distribuição de espécies em grandes áreas (Urban *et al.*, 2002).

## REFERÊNCIAS BIBLIOGRÁFICAS

- Breiman, L., J.H. Friedman, R.A. Olshen & C.J. Stone. 1984. Classification and regression trees. Wadsworth and Brooks/Cole, Monterey, CA.
- Brooks, T., G.B. da Fonseca & A.S.L. Rodrigues. 2004. Species data and conservation planning. *Conservation Biology* 18: 1682-1688.
- Cavalheiro, F., M.V.R. Ballester, A.V. Krusche, S.A. Melo, J.L. Waechter, C.J. Da Silva, M.C. Dárienzo, M. Suzuki, R.L. Bozelli, T.P. Jesus & J.E. Santos. 1990. Propostas preliminares referentes ao plano de zoneamento e manejo da Estação Ecológica de Jataí, Luiz Antônio, SP. *Acta Limnologica Brasiliensia* 3: 951-968.
- Dalmolin, P.C. 1999. Composição e história natural da comunidade de serpentes da Estação Ecológica de Jataí e município de Luiz Antônio, SP. Dissertação de Mestrado. Universidade Federal de São Carlos, São Carlos.
- Davis, F.W. & S. Goetz. 1990. Modeling vegetation pattern using digital terrain data. *Landscape Ecology* 4: 69-80.
- De'Ath, G. & K.E. Fabricius. 2000. Classification and Regression Trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81: 3178-3192.
- Dias, M.M. 2000. Avifauna das Estações Ecológica de Jataí e Experimental de Luiz Antônio, São Paulo, Brasil. In: J.E. Santos & J.S.R. Pires (eds). Estação Ecológica de Jataí. pp. 285-301. São Carlos: RIMA, v. 1.
- Dornelles, S.S. 2000. Censo e análise de habitat para conservação e manejo de primatas (Estações Ecológica de Jataí e Experimental de Luiz Antônio, SP). Dissertação de Mestrado. Universidade Federal de São Carlos, São Carlos.
- Environmental Systems Research Institute (ESRI). 2000. Inc. ArcGIS, Version 9.0. Redlands, CA.
- Flick, P. 1998. A multiple scale approach to reserve site selection. Dissertação de Mestrado. Nicholas School of the Environment, Duke University, Durham.
- Loiselle, B.A., C.A. Howell, C.H. Graham, J.M. Goerck, T. Brooks, K.G. Smith & P.H. Williams. 2003. Avoiding pitfalls of using species distribution models in conservation planning. *Conservation Biology* 17: 1591-1600.
- Lynn, H., C.L. Mohler, S.D. Degloria & C.E. McCulloch. 1995. Error assessment in decision-tree models applied to vegetation analysis. *Landscape Ecology* 10: 323-335.
- Mantovani, J.E. 2001. Telemetria convencional e via satélite na determinação das áreas de vida de três espécies de carnívoros da região nordeste do Estado de São Paulo. Tese de Doutorado. Universidade Federal de São Carlos, São Carlos.
- Miller, J.R., M.G. Turner, E. Smithwick, L. Dent & E.H. Stanley. 2004. Spatial Extrapolation: the science of predicting ecological patterns and processes. *Bioscience* 54: 310-320.
- Milne, B.T., K.M. Johnston & R.T.T. Forman. 1989. Scale-dependent proximity of wildlife habitat in a spatially-neutral Bayesian model. *Landscape Ecology* 2: 101-110.
- Mittermeier, R.A., A.B. Rylands, A.F. Coimbra-Filho & G.A.B. Fonseca. 1988. Ecology and Behavior of Neotropical Primates. World Wildlife Fund, Washington D.C.
- Moore, D.M., B.G. Lee & S.M. Davey. 1990. A new method for predicting vegetation distributions using decision tree analysis in a geographic information system. *Environmental Management* 15: 59-71.

- Moore, I.D., R.B. Gryson & A.R. Ladson. 1991. Digital terrain modeling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes* 5: 3-30.
- Motta-Junior, J.C. 2000. Variação temporal e seleção de presas na dieta do lobo-guará, *Chrysocyon brachyurus* (Mammalia: Canidae), na Estação Ecológica de Jataí, Luiz Antônio, SP. In: J.E. Santos & J.S.R. Pires (eds). Estação Ecológica de Jataí. pp. 331-346. São Carlos: RIMA, v. 1.
- Motta-Junior, J.C. & C.J.R. Alho, 2000. Ecologia alimentar de *Athene cunicularia* e *Tyto alba* (Aves: Strigiformes) nas Estações Ecológica de Jataí e Experimental de Luiz Antônio, SP. In: J.E. Santos & J.S.R. Pires (eds). Estação Ecológica de Jataí. pp. 303-315. São Carlos: RIMA, v. 1.
- Pires, A.M.Z.C.R. 2000. Diretrizes para a conservação da biodiversidade em planos de manejo de unidades de conservação. Caso de estudo: Estação Ecológica de Jataí e Estação Experimental de Luiz Antônio (Luiz Antônio - SP). Tese de Doutorado. Universidade Federal de São Carlos, São Carlos.
- Talamoni, S.A. 1996. Ecologia de uma comunidade de pequenos mamíferos da Estação Ecológica de Jataí, município de Luiz Antônio, SP. Tese de Doutorado. Universidade Federal de São Carlos, São Carlos.
- Talamoni, S.A., J.C. Motta-Junior & M.M. Dias. 2000. Fauna de mamíferos da Estação Ecológica de Jataí e da Estação Experimental de Luiz Antônio, São Paulo, Brasil. In: J.E. Santos & J.S.R. Pires (eds). Estação Ecológica de Jataí. pp. 317-327. São Carlos: RIMA, v. 1.
- Urban, D.L. 2002. Strategic monitoring of landscapes for natural resource management. In: J.L. Liu & W.W. Taylor (eds). Integrating landscape ecology into natural resource management. pp. 294-311. Cambridge University Press, Cambridge, Reino Unido.
- Urban, D.L., S. Goslee, K. Pierce & T. Lookingbill. 2002. Extending community ecology to landscapes. *Ecoscience* 9: 200-212.
- Vaughan, I.P. & S.J. Ormerod. 2003. Improving the quality of distribution models for conservation by addressing shortcomings in the field collection of training data. *Conservation Biology* 17: 1601-1611.
- Vayssieres, M.P., R.E. Plant & B.H. Allen-Diaz. 2000. Classification trees: an alternative non-parametric approach for predicting species distributions. *Journal of Vegetation Science* 11: 679-694.
- Wolock, D. 1993. Simulating the variable-source-area concept of streamflow generation with the watershed model TOPMODEL. Water-Resources Report, USGS.

---

# O uso de regressão logística para espacialização de probabilidades

EDUARDO M. VENTICINQUE<sup>1\*</sup>  
JULIANA STROPP CARNEIRO<sup>2</sup>  
MARCELO PAUSTEIN MOREIRA<sup>2</sup>  
LEANDRO FERREIRA<sup>3</sup>

<sup>1</sup> Wildlife Conservation Society, Programa de Conservação Andes-Amazônia, Amazonas, Brasil.

<sup>2</sup> Instituto Nacional de Pesquisas da Amazônia – INPA, Amazonas, Brasil.

<sup>3</sup> Museu Paraense Emílio Goeldi, Pará, Brasil.

\* e-mail: eventicinque@wcs.org

## RESUMO

Neste trabalho discute-se o uso de modelos de regressão logística em análises espaciais, fazendo uma breve introdução sobre regressões logísticas e usando estudos de casos da aplicação desta técnica em estudos ecológicos, utilizando aplicativos de Sistemas de Informação Geográfica.

## ABSTRACT

*In this chapter we discussed the use of logistic regression models in spatial analyses, doing a brief introduction on logistic regression and your application in some study cases related to ecology studies using with tools the Geographic Information System.*

## INTRODUÇÃO

A regressão logística vem sendo utilizada nas mais diversas áreas da ciência. Este método, assim como as regressões lineares e múltiplas, estuda a relação entre uma variável resposta e uma ou mais variáveis independentes. A diferença entre estas técnicas de regressão se deve ao fato de que na regressão logística as variáveis dependentes estão dispostas em categorias, enquanto na regressão linear estas variáveis são dados contínuos ou discretos. Outra diferença é que na regressão logística a resposta é expressa por meio de uma probabilidade de ocorrência, enquanto que na regressão simples obtém-se um valor numérico (Penha, 2002).

A estrutura do modelo logístico é apropriada para analisar o comportamento de uma variável dependente categórica. Geralmente, a regressão logística é realizada para dados binários (Cox, 1970), entretanto, também pode ser aplicada a dados multinominais. Tipicamente, a variável dependente é binária e codificada como 0 (ausência) ou 1 (presença); porém, pode ser multinominal, sendo codificada como um número inteiro, variando de 1 a  $k - 1$ , onde  $k$  é um número positivo qualquer. Embora a regressão logística possa ser aplicada a qualquer variável dependente categórica, ela é utilizada com maior frequência em análises de dados binários. Estes exemplos incluem a estimativa de probabilidade de ocorrência de uma espécie em

função de variações na altitude ou da quantidade de chuva, a estimativa da probabilidade de que uma área seja desflorestada em função de sua distância das estradas, rios ou sedes municipais, etc.

De forma sucinta, podemos dizer que existem três procedimentos distintos para manipular dados binários, ordinais e nominais em regressão logística. A escolha de qual método utilizar depende do número de categorias e das características da variável resposta, conforme mostra a Tabela 1.

**TABELA 1** – Tipos de variável resposta.

TIPO	NÚMERO DE CATEGORIAS	CARACTERÍSTICAS
Binária	2	Dois níveis
Ordinal*	3 ou +	Ordenação natural de níveis
Nominal*	3 ou +	Sem ordenação natural de níveis

(\*) São variações do estado multinominal ou politômico de uma variável (adaptado de Penha, 2002).

Uma variável binária é aquela que aceita apenas dois níveis de resposta, como sim ou não. Já uma variável ordinal segue uma ordenação natural dos fenômenos ou eventos, como pequeno, médio e grande, ou classificações como ruim, regular, bom, ou excelente (“ranks”). A nominal, por sua vez, pode ter mais de três níveis e não considera nenhuma ordenação. Um exemplo seria a classificação de algum objeto em azul, preto, amarelo e vermelho; ou a previsão do tempo como ensolarado, nublado e chuvoso (Penha, 2002).

Existem vários tipos de estudos que se pode analisar com modelos logísticos. Estes incluem bioensaios, epidemiologia, experimentos clínicos, pesquisa de mercado, distribuição de espécies, etc. Neste trabalho nós vamos nos ater às aplicações da regressão logística dentro de um Sistema de Informação Geográfica (SIG) com questões ligadas à ecologia e à conservação.

A Figura 1 compara o modelo linear com o logístico binário básico, utilizando os mesmos dados. Note que o modelo linear prediz valores de  $y$  contínuos infinitamente. Assim, se a predição é para compreensão das probabilidades, este modelo é claramente impróprio. Além disso, o modelo linear não se ajusta à média de  $x$  para qualquer um dos valores da resposta. Geralmente não se consegue ajustar estes dados satisfatoriamente. Assim, podemos dizer que o modelo linear não é apropriado para esta estrutura de dados. Já o modelo logístico é projetado para ajustar dados binários, quando é

assumido que  $y$  representa uma distribuição de probabilidades, ou quando é simplesmente expresso como uma medida binária que estamos tentando prever. Apesar da diferença entre os dois gráficos, o modelo linear e o logístico são variantes um ao outro. Assumindo a variável preditora ( $x$ ), o modelo linear é:

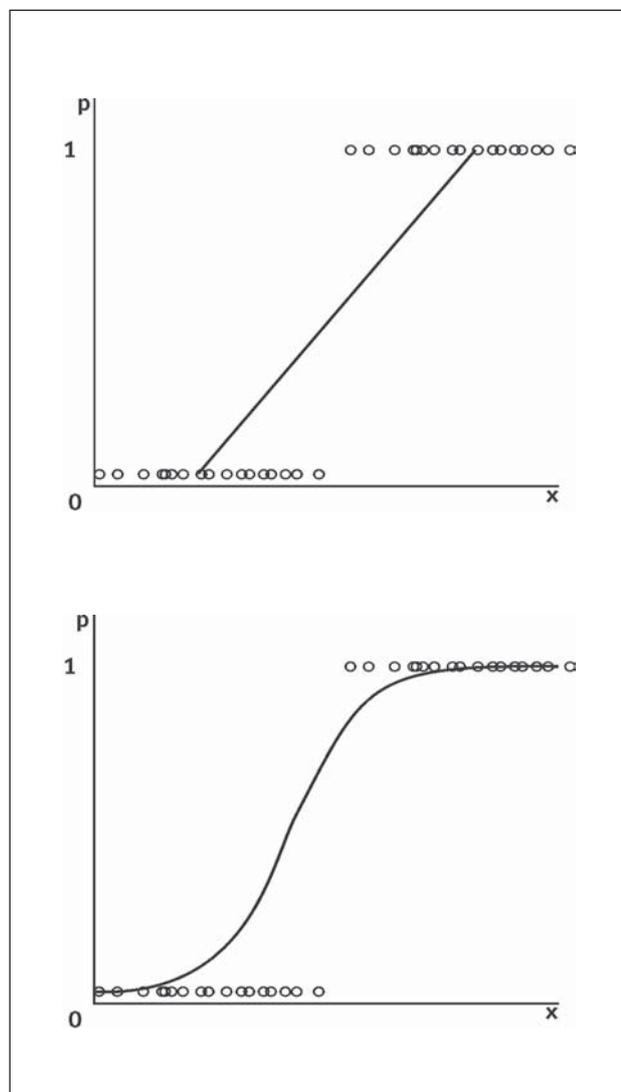
$$y = xb + e,$$

onde  $y$  é um vetor de observações,  $x$  é uma matriz dos preditores, e  $e$  é um vetor de erros.

Enquanto que o modelo logístico é:

$$y = \exp(xb + e) / [1 + \exp(xb + e)],$$

onde  $y$  é a probabilidade de ocorrência de um evento,  $x$  é uma matriz dos preditores, e  $e$  é um vetor de erros.



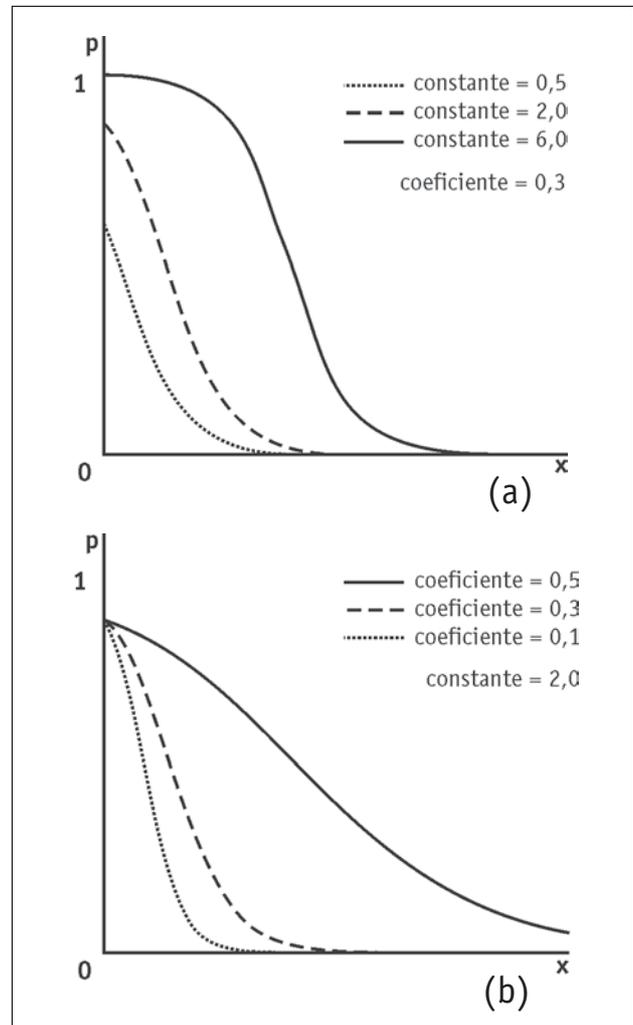
**FIGURA 1** – Modelos de regressão linear e logística com dados binários.

### Coeficientes e constantes

Podemos avaliar os coeficientes obtidos pela regressão logística de forma parecida com a que fazemos em uma regressão linear. No entanto, sua interpretação é diferente. O coeficiente da regressão logística indica o quanto aumenta a probabilidade de ocorrência de um evento para o aumento de uma unidade na variável independente. O coeficiente pode ser positivo ou negativo. No caso de um coeficiente positivo, quanto maior for seu valor, maior será o poder preditivo da variável independente sobre a probabilidade de ocorrência de um evento. No entanto, a probabilidade de 0 a 1 é resultado de uma função não linear da probabilidade de ocorrência de um evento.

É muito importante lembrar o que quer dizer, em termos de interpretação, uma função não linear. Na regressão linear o acréscimo (ou decréscimo) do valor de  $y$  em função do acréscimo de  $x$  é constante ao longo de toda escala de valores de  $x$ . Já na regressão logística isto não acontece, havendo áreas onde essa mudança é mais pronunciada e outras onde ela nem ocorre. As áreas onde pequenas variações nos valores de  $x$  causam grandes mudanças nos valores de  $y$  representam áreas de maior probabilidade de mudança de estado da variável  $y$  em função de  $x$ .

Na Figura 2 podemos visualizar o efeito da variação dos valores da constante e dos coeficientes sobre a curva de probabilidade estimada a partir de regressão logística. O gráfico da Figura 2a foi obtido somente com a troca dos valores da constante (intercepto) e podemos notar que as formas das curvas são exatamente as mesmas e a única mudança é sua localização no eixo  $x$ . Em outras palavras, todo modelo de regressão logística tem seus limites entre 0 e 1, só que muitas vezes estes limites estão fora do nosso intervalo de amostragem ou mesmo não são plausíveis de acontecer, por exemplo, como valores negativos de distância. Neste caso, não é possível visualizar em que intervalo de  $x$  as probabilidades alcançam valores próximos de 1. Já na Figura 2b temos uma situação distinta, onde a alteração dos coeficientes com uma constante fixa causa mudanças evidentes na distribuição da probabilidade de ocorrer um evento em função da mudança de valores no eixo  $x$ . Podemos notar que quanto maior o coeficiente, maior é a mudança na probabilidade estimada em função de mudanças no  $x$ . De forma simplificada, podemos dizer que o coeficiente modela a curva enquanto que a constante a localiza em função do  $x$ .



**FIGURA 2** – Modelos de regressão logística obtidos com alterações somente na constante (a) e no coeficiente (b).

### Razão de chances (*odds ratio*)

A razão de chances permite conhecer quais chances um evento tem de acontecer se, sob as mesmas condições, ele não acontecer. Ou seja, razão de chances é uma medida de associação e expressa a aproximação do quanto é mais provável (ou improvável) para o resultado estar presente entre aqueles com  $x = 1$  do que entre aqueles com  $x = 0$ . Por exemplo, se  $y$  denota a presença ou ausência de uma determinada espécie e  $x$  denota se a área tem ou não tem floresta, o *Odds* = 2 indica que a presença daquela espécie é duas vezes mais esperada em áreas com floresta do que em áreas sem floresta. Ou seja, a presença de floresta é muito importante para aumentar a chance de ocorrência daquela espécie. Outro exemplo, que talvez possa ser mais intuitivo, seria a razão de chances de ser atropelado toda vez que se atravessa uma avenida. Mesmo que você

atravesse a avenida e não seja atropelado, existia uma chance deste evento ocorrer, essa chance é a “razão de chances” ou “*odds ratio*”. A razão de chances de resposta é dada por  $p/(1-p)$  onde  $p$  é a probabilidade de resposta, e a razão de chances é o fator multiplicativo de mudança de estado de  $y$  quando a variável independente aumenta uma unidade. O livro de Hosmer & Lemeshow (1989) contém maiores explicações sobre a interpretação e forma de cálculo das razões de chance e de seus intervalos de confiança.

### Estatística de *Likelihood-Ratio*

Uma vez definido o modelo, é necessário testar a sua validade. Em regressão logística há uma série de gráficos, testes de ajuste, e outras medidas para assegurar a validade do modelo. Estas estatísticas permitem identificar as variáveis que não se ajustam bem, ou que têm forte influência sobre a estimativa dos parâmetros.

Uma das formas mais comuns de se avaliar o modelo como um todo, é por meio da estatística de *Likelihood-Ratio*. Esta estatística testa a hipótese de que todos os coeficientes, menos a constante, são iguais a 0. A significância da estatística de *Likelihood-Ratio* (LR) é testada utilizando a distribuição do  $X^2$  com os graus de liberdade iguais ao número de variáveis independentes no modelo, não incluindo a constante.

O teste de *Likelihood-Ratio*, ou teste G, é calculado utilizando o valor da estatística de log likelihood do modelo saturado e do insaturado. Tipicamente, o modelo saturado contém o conjunto de variáveis analisadas e o modelo insaturado omite um subconjunto selecionado, embora outras restrições sejam possíveis. A estatística do teste é duas vezes a diferença do *Likelihood-Ratio* do modelo saturado para o insaturado e é testada com a distribuição do  $X^2$ , sendo o grau de liberdade igual ao número de restrições impostas. Se um modelo contém uma constante, podemos calcular um teste de *Likelihood-Ratio* da hipótese nula em que todos os coeficientes, exceto a constante, são iguais a 0. A fórmula da estatística G usada para testar o modelo é a seguinte:

$$G = 2*[LL(N)-LL(0)]$$

Onde:

LL(N) = log *likelihood* do modelo saturado  
(todas as variáveis inclusas)

LL(0) = log *likelihood* do modelo insaturado  
(somente a constante inclusa)

Quando fazemos essa subtração, estamos olhando, simplesmente, o quanto as variáveis estão causando mudanças nas probabilidades de ocorrência de um evento e se essas mudanças são maiores que esperadas ao acaso.

Para ilustrar o uso do teste de *Likelihood-Ratio*, considere o seguinte modelo:

Presença de uma espécie =  
CONSTANTE + altitude + chuva + temperatura (saturado)  
Presença de uma espécie =  
CONSTANTE + altitude + chuva (insaturado)

Podemos formular a hipótese nula de que a temperatura não contribui para explicar a variação do modelo e proceder ao teste desta forma. Suponha que para este exemplo os valores de G são 12,05 e 5,01, com 3 e 2 graus de liberdade para os modelos saturado e insaturado, respectivamente. Agora podemos entender a variação que é explicada pela temperatura, entendendo quanto perdemos de poder de explicação ao removermos essa variável do modelo. Isso pode ser realizado da seguinte forma:

Efeito da temperatura = G (insaturado) – G (saturado),  
com 3 – 2 graus de liberdade.

Essa expressão fica assim:

G = 12,05 – 5,01, com 1 grau de liberdade  
G = 7,04, gl = 1 e  $p < 0,05$ , rejeitando-se a hipótese nula de que a temperatura não tem influência sobre a probabilidade de ocorrência de uma determinada espécie.

## TESTES ESTATÍSTICOS DE AJUSTE DO MODELO ÀS OBSERVAÇÕES

### Rho<sup>2</sup> de McFadden

Rho<sup>2</sup> de McFadden é uma transformação da estatística de LR para imitar um R<sup>2</sup> da regressão linear. Seus valores estão sempre entre 0 e 1 e, quanto mais alto, melhor é o ajuste do modelo aos resultados. Entretanto, o Rho<sup>2</sup> de McFadden tende a ser muito mais baixo que R<sup>2</sup>. Porém, baixos valores não implicam, necessariamente, num ajuste pobre. Valores entre 0,2 e 0,4 são considerados satisfatórios (Hensher & Johnson, 1981).

### Pearson

Mede quão bem a observação é prevista pelo modelo. Observações que não se ajustam bem ao modelo têm um alto valor de Pearson.

### Hosmer-Lemeshow

Este teste avalia o modelo ajustado, comparando as frequências observadas e as esperadas. O teste associa os dados às suas probabilidades estimadas, da mais baixa à mais alta, e então faz um teste qui-quadrado para determinar se as frequências estimadas estão próximas das frequências observadas (Hosmer & Lemeshow, 1989).

### Diagnósticos de regressão

Na regressão logística, a representação gráfica permite visualizar vários testes de ajuste, sendo que há gráficos relacionados à probabilidade do evento e outros relacionados à alavancagem (que diz se uma observação é um ponto extremo e possui uma forte influência na determinação da reta de regressão, o que diminui a capacidade de análise do modelo). A inspeção gráfica é realizada com base nos pontos extremos de influência (ou *outliers*). Em alguns casos, o ponto que foi identificado como extremo deve ser excluído da amostra e, em seguida, deve ser novamente calculada a equação e o gráfico. Quando os coeficientes desta nova equação forem muito diferentes dos coeficientes da antiga, significa que aquele era um ponto de influência. Se o contrário ocorrer, significa que aquele ponto era apenas um ponto extremo. A decisão de se remover dados da amostra deve ser procedida com muito critério e cuidado. Geralmente, existem informações importantes nestes pontos discrepantes. Por exemplo, os gráficos do delta qui-quadrado (DELPSTAT) *versus* probabilidade do evento identificam os pontos que não se ajustam bem aos modelos.

Em diversos pacotes estatísticos pode-se criar um arquivo para elaborar diagnósticos da regressão logística (Pregibon, 1981; Cook & Weisberg, 1984; Steinberg & Colla, 1998). No caso do programa SYSTAT, o arquivo contém as variáveis apresentadas na Tabela 2.

Podemos entender a variável: LEVERAGE (1) como uma medida da influência de uma observação no ajuste do modelo, e a variável DELBETA (1) como uma medida da mudança no vetor do coeficiente devido àquela observação. Por exemplo, os gráficos de PEARSON, DEVIANCE, LEVERAGE (1), DELPSTAT, com o CASO, destacam pontos de dados diferenciados. Para discussão adicional e interpretação de gráficos de diagnóstico, veja o Capítulo 5 de Hosmer & Lemeshow (1989).

**TABELA 2** – Variáveis de diagnóstico da regressão logística geradas pelo programa estatístico SYSTAT. Uma descrição detalhada destas variáveis pode ser encontrada no manual do SYSTAT ou em Hosmer & Lemeshow (1989).

NOME NO ARQUIVO	LEGENDA
ACTUAL	Valor da variável dependente
PREDIGA	Valor predito (1 ou 0)
PROB	Probabilidade predita
LEVERAGE (1)	Elemento diagonal da matriz "chapéu" de Pregibon
LEVERAGE (2)	Componente de LEVERAGE (1)
PEARSON	Resíduo de observação de Pearson
VARIANCE	Variância de resíduo de Pearson
PADRÃO	Resíduo de Pearson padronizado
DEVIANCE	Desvios Residuais
DELPSTAT	Mudança no $\chi^2$ de Pearson
DELBETA (1)	Mudança padronizada em Beta
DELBETA (2)	Mudança padronizada em Beta
DELBETA (3)	Mudança padronizada em Beta

### Principais problemas e vantagens

#### PROBLEMAS

- Se o fenômeno de interesse não for monotônico e seu pico de frequência tiver valores intermediários, será difícil obter um bom ajuste do modelo. Esse tipo de problema pode ser identificado através da análise dos resíduos da regressão;
- Pontos discrepantes, principalmente nos limites da distribuição das variáveis explanatórias, podem causar resultados espúrios;
- Obter dados confiáveis e não viciados para alimentação do modelo;
- Autocorrelação espacial.

#### VANTAGENS

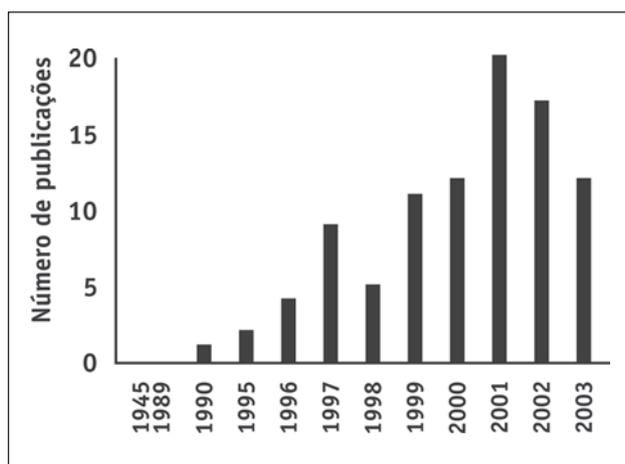
- O modelo logístico requer informações simples e, portanto pode ser alimentado com facilidade;
- Trabalha com  $N$  variáveis simultaneamente;
- Trabalha simultaneamente com  $N$  vetores direcionais de variação. Essa flexibilidade pode ser obtida quando construímos em um Sistema de Informação Geográfica uma camada de dados independente. No caso de se trabalhar com informações com formas irregulares, por exemplo, distância da estrada ou declividade, o sentido de variação espacial do fenômeno pode ocorrer em diferentes direções;
- As probabilidades obtidas podem ser espacializadas e então se aplicar um filtro para que os padrões possam ser mais facilmente visualizados;
- Facilidade de interpretação e uso dos resultados em probabilidades.

### Validação do modelo

Assim como a maioria dos métodos, a regressão logística necessita de novos dados (nova amostra) ou de uma amostra reservada dos dados para verificar se o mesmo modelo pode ser satisfatoriamente ajustado a estes novos dados. Ou seja, é preciso determinar se os coeficientes e os erros-padrão obtidos a partir dos dados utilizados para elaboração do modelo são similares aos obtidos para os dados de validação.

### O uso de regressão logística na análise espacial

Fizemos uma consulta bibliográfica ([www.webofscience.com](http://www.webofscience.com)) a partir das palavras-chave “regressão logística” e “Sistemas de Informação Geográfica” e obtivemos 93 trabalhos publicados ao longo de 58 anos. O uso das regressões logísticas associado a Sistemas de Informação Geográfica começou a ser mais praticado a partir de 1990 (Figura 3). De 1999 até 2003 foram publicados, em média, 14 trabalhos por ano.



**FIGURA 3** – Número de publicações encontradas no site [www.webofscience.com](http://www.webofscience.com), referente ao emprego de regressão logística em Sistemas de Informação Geográfica (SIG).

### EXEMPLOS DE APLICAÇÃO DA REGRESSÃO LOGÍSTICA

Aqui vamos citar três exemplos onde técnicas de regressão logística simples são utilizadas. O leitor poderá reparar que o uso pode ter variação na escala espacial empregada, podendo ser usada desde a detecção de respostas de Odonata à proporção de florestas circundando igarapés na Amazônia central até modelos de desflorestamento em função da malha viária na Amazônia Legal. Outra característica relevante diz respeito às diferentes formas com que estes

modelos podem ser usados. Há casos, como no exemplo da probabilidade de ocorrência de espécies de Odonata em função da quantidade de floresta, onde os resultados não são reprojatados, ou seja, não há espacialização das probabilidades, pois estávamos interessados simplesmente em saber se há influência da quantidade de floresta preservada ao longo dos pequenos cursos d’água sobre a ocorrência de determinadas espécies.

No estudo realizado com distribuição de árvores na Mil Madeireira Itacoatiara Ltda, as probabilidades de ocorrência de cada espécie foram projetadas no espaço, utilizando como base os mapas de altitude e declividade (variáveis independentes). Neste caso, foi utilizada uma regressão logística múltipla e o modelo pode ser considerado espacialmente explícito, pois podemos localizar, no espaço, todas as probabilidades.

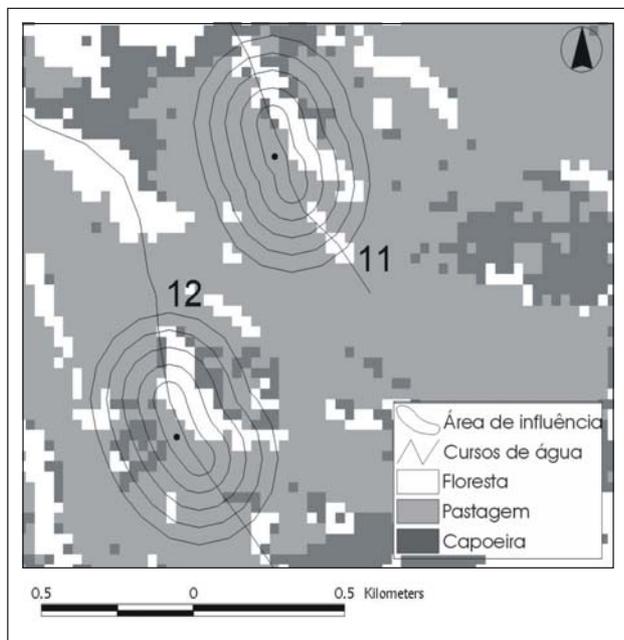
Outro exemplo em que o espaço continua implícito, mas as probabilidades não são projetadas no mapa, é o trabalho sobre a importância das unidades de conservação e terras indígenas, ajudando a conter o desmatamento na Amazônia brasileira. Nesse caso, a regressão logística foi utilizada para entender a probabilidade de uma área florestada ser convertida em área desflorestada, considerando se está localizada dentro ou fora de uma terra indígena ou unidade de conservação, e a distância que está da malha viária.

O que tentamos aqui, por meio destes exemplos, é oferecer ao leitor um panorama geral de alguns usos que podemos ter com regressão logística, e também chamar atenção para o uso da técnica com problemas espaciais.

### Mudanças na fauna de odonatas em igarapés amazônicos, em função de alterações na cobertura vegetal

Este estudo, realizado na Amazônia central, estima a probabilidade de ocorrência de espécies de odonatas em função da quantidade de floresta. Neste estudo, o pesquisador Dr. Paulo de Marco estava interessado em saber se existe influência da quantidade de floresta preservada ao longo dos pequenos cursos de água sobre a ocorrência de determinadas espécies de odonatas. O resultado ajuda a avaliar se a largura da mata ripária prevista no Código Florestal Brasileiro é suficiente para manter o conjunto de espécies de odonatas.

Probabilidade de ocorrência de uma espécie =  $\exp (\% \text{ de cobertura florestal} * \text{coeficiente} + \text{Intercepto} + \text{erro}) / [1 + \exp (\% \text{ de cobertura florestal} * \text{coeficiente} + \text{Intercepto} + \text{erro})]$



**FIGURA 4** – Exemplo de como a paisagem é tratada neste estudo. A análise é realizada utilizando o valor de cobertura florestal contida dentro das áreas de influência. No caso deste estudo, os anéis são distanciados 50 metros.

Podemos notar nos resultados (Tabela 3) que somente uma espécie respondeu às alterações na cobertura de forma negativa, ou seja, quanto maior era a porcentagem de cobertura florestal menor era a probabilidade de encontrar a espécie. As demais espécies não responderam à proporção de mata ao redor dos pontos amostrais, na área do estudo.

#### Uso de regressão logística para modelar a distribuição espacial de espécies arbóreas na Amazônia central

O presente trabalho é parte dos resultados apresentados na dissertação de mestrado de Juliana Stropp Carneiro, sob a orientação do Dr. Eduardo Venticinque

(Carneiro, 2004). Este trabalho teve como objetivo elaborar modelos preditivos de ocorrência de *Aniba roseodora* (pau-rosa), *Cariniana micrantha*, *Caryocar villosum*, *Dinizia excelsa*, *Dipteryx odorata*, *Goupia glabra*, *Manilkara bidentata*, *Manilkara huberi*, *Parkia multijuga*, *Parkia pendula*, *Peltogyne paniculata* e *Pseudopiptadenia psilostachya* em função da topografia. Estimamos a probabilidade de ocorrência dos indivíduos com regressão logística múltipla, sendo a variável dicotômica a presença e a ausência dos indivíduos, e as variáveis contínuas a altitude e a declividade do terreno. As informações sobre a ocorrência das árvores foram cedidas pela Mil Madeireira Itacoatiara Ltda. Os dados sobre a ocorrência dos indivíduos arbóreos foram coletados pela empresa durante a prospecção e o mapeamento das árvores com DAP  $\geq 40$  cm. As informações sobre a ocorrência das árvores consistem em um arquivo do tipo pontos, em formato *shapefile*, com a lista de espécies e as coordenadas da localização dos indivíduos em UTM. Convertimos esse arquivo para o formato matricial e obtivemos um arquivo do tipo GRID, com células de 93 m. Elaboramos o Modelo Digital do Terreno a partir dos dados do Shuttle Radar Topography Mission (SRTM) e adquirimos os dados sobre altitude no site <http://seamless.usg.gov>. Para a correção geométrica, utilizamos como base uma imagem Landsat TM 7 (órbita/ponto 230/62) georreferenciada (projeção UTM – zona 21; datum WGS 84). Posteriormente, coregistramos a imagem SRTM com a base dos igarapés da área de interesse digitalizada. Para obter os parâmetros da regressão logística, exportamos os dados do ArcView 3.2 e os analisamos em um pacote estatístico. Posteriormente, aplicamos as equações obtidas nos modelos logísticos aos temas de altitude e declividade e obtivemos os mapas de probabilidade de ocorrência de indivíduos (ver anexo). Deste modo, estes mapas expressam a probabilidade de ocorrência dos indivíduos em células de 93 m, dada a altitude e declividade daquela célula.

**TABELA 3** – Análise de regressão logística para a dependência da presença de algumas espécies de Odonata em relação à proporção de mata ao redor dos pontos amostrais, na área do Projeto Dinâmica Biológica de Fragmentos Florestais (PDBFF), Manaus, AM. Valores entre parênteses são os erros padrões dos parâmetros estimados.

ESPÉCIE	COEFICIENTE BO	% DE MATA	X <sup>2</sup> (VALOR DE p)
<i>Argia</i> sp.1	-1,086 (1,138)	2,305 (1,631)	2,145 (0,143)
<i>Argia</i> sp. 2	2,665 (1,461)	-3,233 (1,861)	3,656 (0,050)
<i>Chalcopteryx scintilans</i>	-0,782 (1,114)	1,490 (1,549)	0,954 (0,329)
<i>Dicterias atosanguinea</i>	-0,080 (1,094)	-0,181 (1,504)	0,014 (0,904)

Para determinar a capacidade preditiva do modelo, obtivemos a tabela de sucesso de predição para cada um dos modelos gerados. Esta tabela é composta pelas variáveis expressas abaixo:

$$resposta = \sum Pi,$$

onde  $Pi$  é a probabilidade estimada para as células de presença;

$$referência = \sum Pj,$$

onde  $Pj$  é a probabilidade estimada para as células de ausência;

$$\text{Índice de acerto de presença} = \frac{resposta}{Ni},$$

onde  $Ni$  é o número de células de presença;

$$\text{Índice de acerto de ausência} = \frac{referência}{Nj},$$

onde  $Nj$  é o número de células de ausência.

As variáveis Índice de acerto de presença, Índice de acerto de ausência e Índice de acerto total refletem a relação entre a distribuição observada e a esperada, indicando o nível de acerto do modelo.

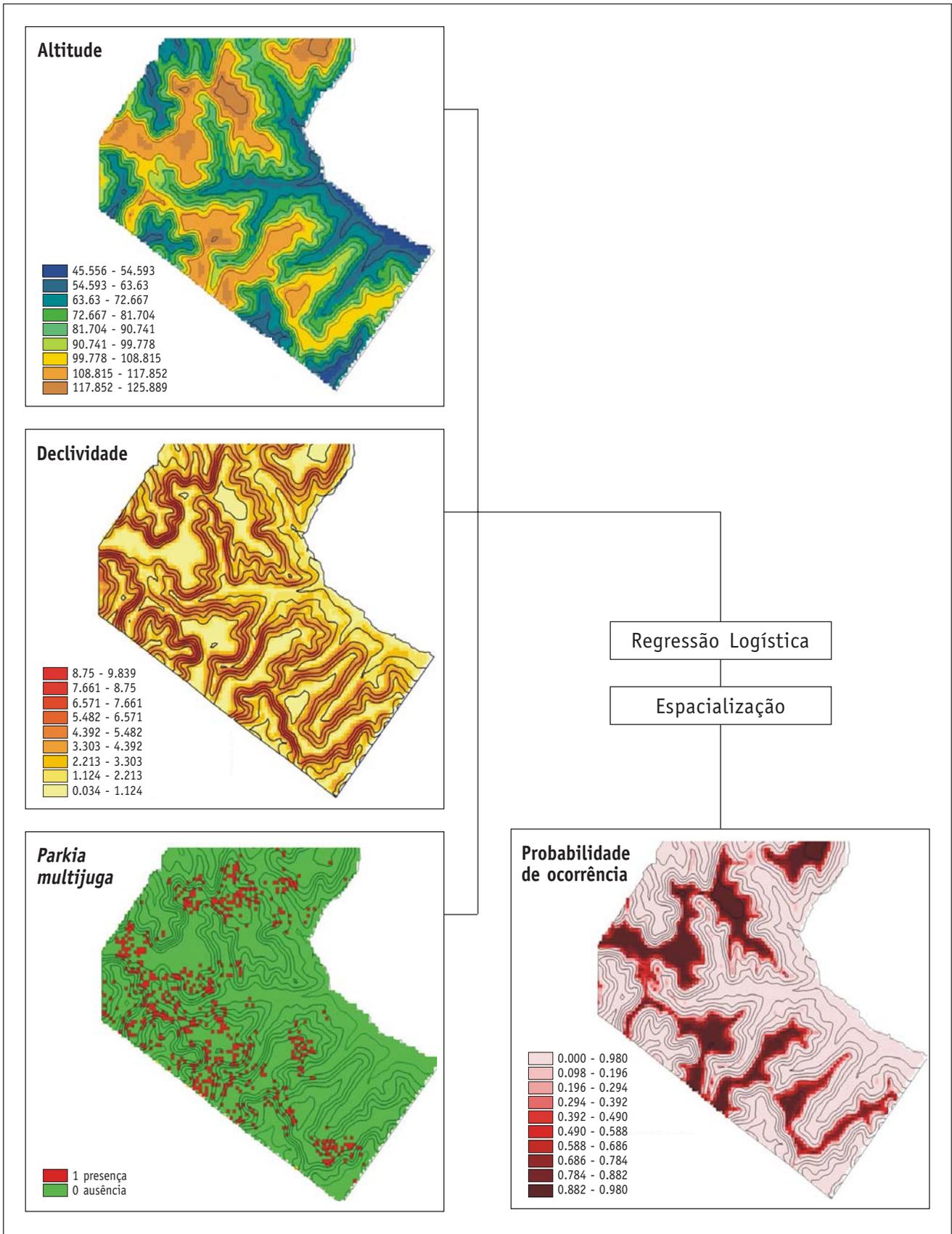
Ainda com o objetivo de determinar se as probabilidades geradas pelos modelos refletem aumento no acerto de ocorrência de um indivíduo, comparamos a probabilidade de acerto usando o modelo com a probabilidade de acerto ao acaso.

A análise de regressão logística indicou associação entre a ocorrência dos indivíduos e a topografia para 10 espécies. As espécies *D. excelsa*, *A. rosaeodora* e *C. villosum* tiveram o padrão de distribuição distinto das demais, em relação à topografia. *D. excelsa* teve relação positiva tanto com a declividade quanto com a altitude, sugerindo que a probabilidade de encontrar indivíduos dessa espécie é maior em lugares altos e íngremes, ou seja, no início dos platôs. Já a ocorrência de *C. villosum* teve relação positiva com a declividade e negativa com a altitude, indicando que esta espécie ocorre nas vertentes e em baixas altitudes. *A. rosaeodora* mostrou-se negativamente relacionada com as variáveis topográficas analisadas, ocorrendo em locais de altitude e declividade baixas, estando assim associada às regiões de baixo. Entretanto as espécies *C. micrantha*, *G. glabra*, *M. huberi*, *M. bidentata*, *P. multijuga*, *P. pendula* e

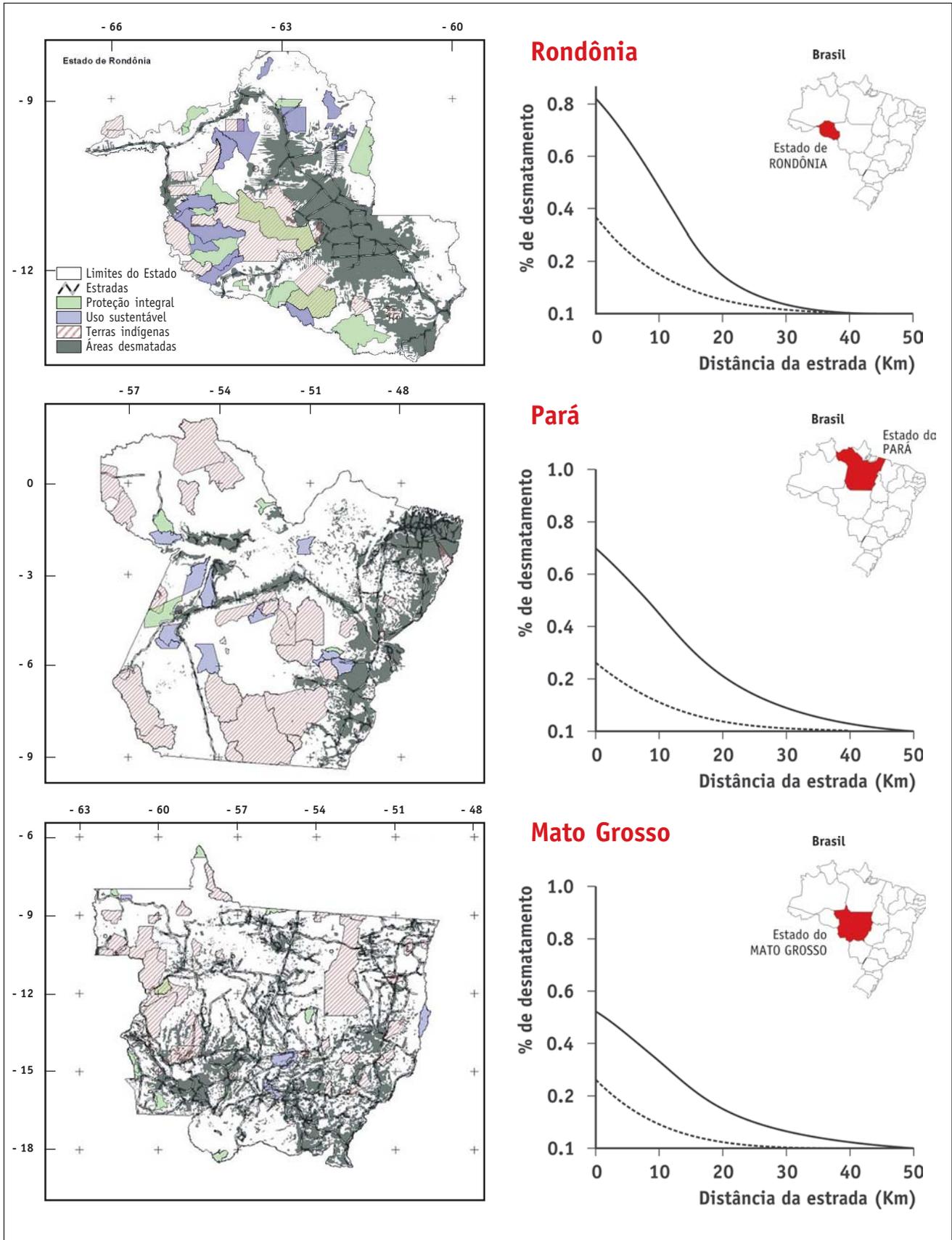
*P. psilostachya* evidenciaram semelhanças na maneira em que se distribuem ao longo da toposeqüência. Essas espécies mostraram-se associadas a locais de altitude elevada e baixa declividade, características que definem os ambientes de platô. As espécies *D. odorata* e *P. paniculata* não tiveram a distribuição estruturada pela topografia.

Os mapas de probabilidade de ocorrência das espécies estudadas representam a configuração espacial da distribuição prevista para cada espécie. Nos mapas das espécies que têm sua ocorrência influenciada pela topografia, é possível visualizar concordância entre as probabilidades mapeadas e a variável topográfica de maior influência sobre a distribuição de seus indivíduos. A Figura 5 ilustra o exemplo do mapa de probabilidade de ocorrência de *P. multijuga*. Neste caso, o índice de acerto de presença foi maior que a probabilidade de encontrar indivíduos ao acaso, indicando que o modelo é capaz de prever a distribuição desta espécie na área onde foi elaborado.

O modelo preditivo foi capaz de prever corretamente a ocorrência de *A. rosaeodora*, *C. micrantha*, *C. villosum*, *D. excelsa*, *M. huberi*, *M. bidentata*, *P. multijuga*, *P. pendula* e *P. psilostachya* na área onde foi elaborado. Esse resultado indica que, em determinados compartimentos da paisagem, a topografia pode condicionar a distribuição de algumas espécies. Em geral, as características edáficas das florestas de terra firme da Amazônia central se alteram ao longo do gradiente de altitude. Dessa forma, a topografia é uma medida indireta das mudanças do ambiente na paisagem (Tuomisto & Ruokolainen, 1994) e, portanto, as diferentes respostas das espécies frente às posições topográficas refletem a influência que a variação ambiental pode ter sobre a estrutura espacial dessas espécies. No caso do trabalho aqui apresentado, a liberação dos dados SRTM na Internet foi fundamental para alcançar os objetivos propostos. Vale lembrar que já existem imagens do SRTM com resolução de 30 metros disponíveis para os Estados Unidos da América e, em breve, é provável que tenhamos acesso a esta informação para a região Amazônica. Caso isso ocorra, poderemos fazer modelos ainda mais precisos, baseados na topografia. Outro fato importante é a ausência de custo para se trabalhar com estas imagens. Se conseguirmos gerar bons modelos preditivos sobre a distribuição de espécies arbóreas com dados provenientes do SRTM, passaremos a ter uma ferramenta eficiente e de baixo custo para modelar a distribuição destas espécies.



**FIGURA 5** – Mapa de probabilidade de ocorrência de *Parkia multijuga*, obtido a partir dos dados de ocorrência dos indivíduos, altitude e declividade, na área da Mil Madeireira Itacoatiara Ltda.



**FIGURA 6** – Proporção de área desmatada em função da distância das estradas, dentro (tracejado) e fora (contínua) de áreas protegidas, em Rondônia, no Pará e no Mato Grosso.

O aumento do poder de predição providenciado pelo modelo é influenciado pelos fatores que estruturam espacialmente as espécies. Assim, se o modelo contempla os fatores preponderantes na ocorrência das espécies, o poder de predição é maior. Deste modo, a compreensão dos fatores que interferem na distribuição espacial das espécies e sua incorporação aos modelos preditivos podem providenciar modelos mais próximos da realidade. Portanto, a incorporação de informações da variação ambiental, bem como estudos aprofundados da relação espécie-ambiente (Pitman *et al.*, 2001) e a análise da distribuição das árvores em escala regional, podem contribuir para a modelagem da distribuição espacial das árvores. Dado o contexto em que se insere a análise da configuração espacial da vegetação, a análise dos dados sobre variáveis ambientais relacionadas com a ocorrência de espécies pode ser proveitosa para a compreensão da distribuição da diversidade na Amazônia.

#### **Unidades de conservação e terras indígenas ajudam a conter desmatamento na Amazônia brasileira**

O objetivo deste tópico foi testar diferenças no nível de desmatamento dentro e fora de unidades de conservação (proteção integral e uso sustentável) e terras indígenas (denominadas aqui como áreas protegidas) em relação à distância das estradas, nos Estados de Rondônia, Pará e Mato Grosso, para ilustrar a importância de unidades de conservação como redutores do efeito do desmatamento na Amazônia. Estes estados foram escolhidos como estudos de caso devido a sua importância na participação do desmatamento da Amazônia, já que somam cerca de 70% do total da área desmatada nesta região entre 2000-2001 (INPE, 2003). Os Estados de Rondônia, Pará e Mato Grosso têm cerca de 29,2%, 20,4%, e 28,4% de sua área já desmatada, respectivamente.

A análise demonstra que a proporção total da área desmatada fora das áreas protegidas sempre foi significativamente mais elevada do que no interior destas. Uma diferença que pode variar de 9,8 a 19,6 vezes, dependendo do estado analisado. A regressão logística também permite demonstrar que a proporção do desmatamento decai exponencialmente em função da distância das estradas. Contudo, o desmatamento dentro das áreas protegidas é sempre menor do que fora delas nos três estados analisados, mesmo quando estas áreas situam-se próximas a estradas. Isso derruba a crença de que as áreas protegidas sofreriam menos desmatamento somente pelo fato de estarem situadas mais distantes das estradas (Figura 6).

#### **AGRADECIMENTOS**

Agradecemos à Mil Madeireira Itacoatiara Ltda. pela disponibilização da base de dados, ao Projeto Dinâmica Biológica de Fragmentos Florestais (PDBFF), ao WWF – Brasil, Projeto Experimento de Grande Escala da Biosfera-Atmosfera da Amazônia (LBA) e ao Projeto Geoma, pelo suporte financeiro (Bolsa de Marcelo Moreira). À Marina Antongiovanni, Amanda Mortati e Ana Albernaz pela revisão do texto. Ao Dr. Paulo De Marco por nos autorizar a usar seus dados em um dos exemplos.

#### **REFERÊNCIAS BIBLIOGRÁFICAS**

- Carneiro, J.S. 2004. Mapeamento preditivo da vegetação: uso de SIG para modelar a distribuição espacial de espécies arbóreas na Amazônia central. 2004. Dissertação de Mestrado. Instituto Nacional de Pesquisas da Amazônia (INPA), Manaus.
- Cook, D.R. & S. Weisberg. 1984. Residuals and influence in regression. Chapman and Hall, New York.
- Cox, D.R. 1970. The analysis of binary data. Methuen, Londres.
- Hensher, D. & L.W. Johnson. 1981. Applied discrete choice modelling. Croom Helm, London.
- Hosmer, D.W. & S. Lemeshow. 1989. Applied logistic regression. John Wiley & Sons, Inc., New York.
- INPE. 2003. Disponível em: <http://www.obt.inpe.br/prodes> (acessado em janeiro 2007).
- Penha, R.N. 2002. Um estudo sobre regressão logística binária. Disponível em: <http://www.iem.efei.br/dpr/td/producao2002/PDF/Renata.PDF> (acessado em novembro de 2003).
- Pitman, N.C.A., J. Terborgh, M.R. Silman, P.V. Núñez, D.A. Neill, C.E. Cerón, W.A. Palacios & M. Aulestia. 2001. Dominance and distribution of tree species in upper Amazonian terra firme forests. *Ecology* 82: 2101-2117.
- Pregibon, D. 1981. Logistic regression diagnostics. *Annals of Statistics* 9: 705-724.
- Steinberg, D. & P. Colla. 1998. Logistic regression. In: Wilkinson, L. (ed). SYSTAT 8.0 Statistics. pp. 517-584. Chicago.
- Tuomisto, H. & K. Ruokolainen. 1994. Distribution of Pteridophyta and Melastomataceae along an edaphic gradient in an Amazonian rain forest. *Journal of Vegetation Science* 5: 25-34.

#### **ANEXO**

##### **Como espacializar regressão logística no ArcView versão 3.2**

A espacialização da regressão logística no ArcView 3.2 se inicia pela determinação dos temas a serem relacionados. Para isso, define-se o tema que contém a variável dependente binária (1/0 – presença e ausência

do fenômeno de interesse) e o tema que representa a variável preditora contínua. No caso da regressão logística múltipla, é possível estabelecer dois ou mais temas referentes às variáveis preditoras. Após essa determinação, é feita a análise estatística dos dados, a fim de se obter os parâmetros da regressão logística. Finalmente, os parâmetros da regressão são incorporados ao ArcView 3.2 e então realizadas as operações para sua espacialização. Os tópicos abaixo descrevem detalhadamente esses três procedimentos:

### Obtenção dos temas referentes à variável dicotômica e contínua

Os temas (camadas digitais) deste tópico devem estar em formato GRID e apresentar a mesma resolução espacial, número de linhas e colunas. O arquivo GRID referente ao tema da variável binária deve ter os valores das células 0 e 1. Assim, se a representação do evento for do tipo ponto, linha ou polígono, é necessário converter o arquivo para GRID e atribuir o valor 1 e 0 às células correspondentes à presença e ausência do evento, respectivamente. A conversão para GRID e a associação do valor 1 às células de presença pode ser feita a partir do menu do ArcView 3.2 e a associação do valor 0 pode ser feita a partir da extensão Grid PIG Tolls (<http://arcscripsts.esri.com> ou <http://arcscripsts.esri.com/details.asp?dbid=11872>). Este tema será utilizado nas operações de obtenção dos valores a serem utilizados na análise estatística. Assim, é necessário que a tabela associada a ele tenha quatro campos, como na Figura 7.

Os campos *value* e *count* são criados automaticamente pelo ArcView 3.2 e indicam o valor numérico do pixel (*value*) e o respectivo número de pixels (*count*) com valor 0 e 1. Os campos *presença* e *ausência* são criados pelo usuário, sendo que o campo *presença* apresenta valor 1 para presença e 0 para ausência e o campo *ausência* valor 1 para ausência e 0 para presença. Até aqui definimos o tema referente à variável categórica. O próximo passo é determinar os temas com as variáveis contínuas. Isso é definido pela experiência e pela disponibilidade de dados do usuário.

### Obtenção dos dados para a análise estatística

Neste tópico será descrito como obter os valores das variáveis contínuas na área de estudo do evento de interesse. Para isso, realizaremos algumas operações matemáticas na função MAP CALCULATOR do ArcView 3.2, com os temas definidos anteriormente. O esquema das operações entre as camadas é mostrado na Figura 8.

- Obtenção dos valores da variável contínua nas células de presença do evento de interesse:  
[GRID variável contínua] ÷ [GRID variável categórica “campo presença = 1”]
- Obtenção dos valores da variável contínua nas células de ausência do evento de interesse:  
[GRID variável contínua] ÷ [GRID variável categórica “campo ausência = 1”]

Os GRIDs gerados por essas operações devem ser exportados no formato ASCII Raster (opção disponível no menu do programa). A planilha da primeira operação contém os valores das variáveis contínuas nos pixels referentes à presença e a da segunda, os valores referentes à ausência. O valor “-9999” é atribuído à ausência de dados.

As planilhas podem ser editadas no Excel. Sugerimos a elaboração de uma única planilha com duas colunas: uma contendo a variável contínua e outra a informação de presença e ausência. No caso da regressão logística múltipla, a planilha pode conter três ou mais colunas. A partir dessas planilhas é possível se obter os parâmetros necessários para espacialização da regressão logística em um pacote estatístico.

### Espacialização da regressão logística no ArcView 3.2

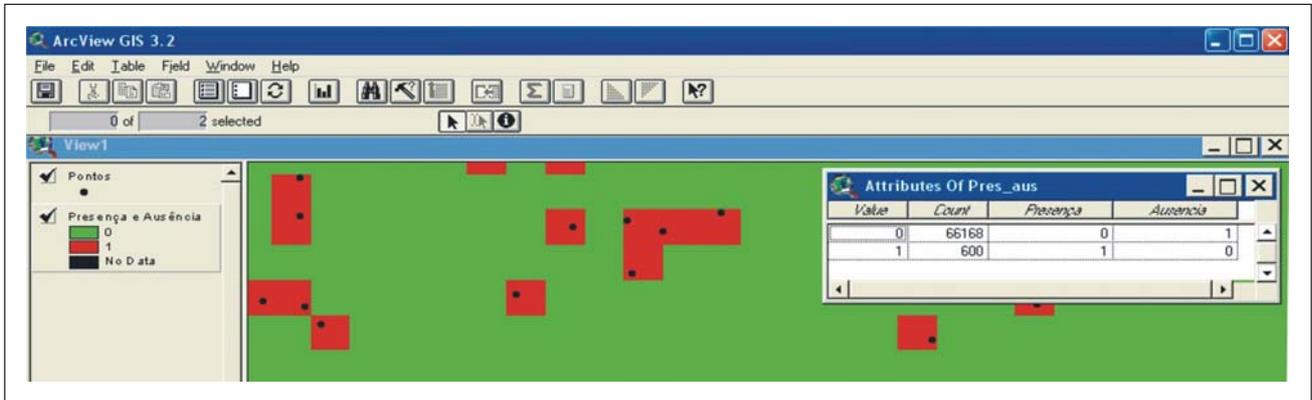
A equação da regressão logística simples pode ser espacializada no ArcView 3.2 a partir das seguintes operações:

- (([GRID variável contínua]) \* coeficiente) - > [GRID A]
- (([GRID A] + Constante) - > [GRID B]
- (([GRID B] .Exp) - > [GRID C]
- (([GRID C] + 1) - > [GRID D]
- (([GRID C] / [GRID D]) - > [GRID E]

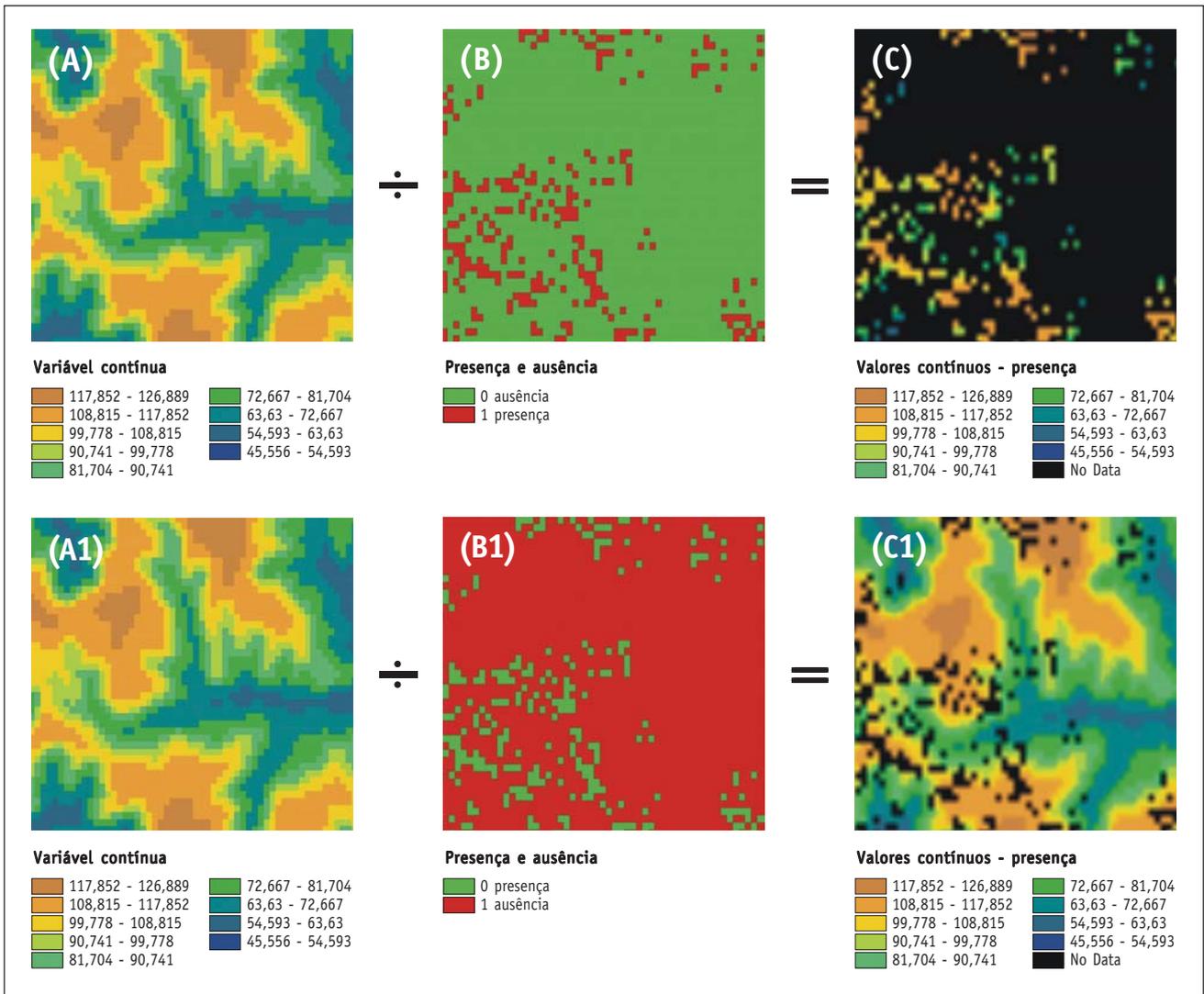
Já para a equação da regressão logística múltipla, sua espacialização é feita com as operações indicadas abaixo.

- (([GRID variável contínua<sub>1</sub>] \* (coeficiente<sub>1</sub>)) + ([GRID variável contínua<sub>2</sub>] \* (coeficiente<sub>2</sub>)) - > A
- (([GRID A] + Constante) - > [GRID B]
- (([GRID B] .Exp) - > [GRID C]
- (([GRID C] + 1) - > [GRID D]
- (([GRID C] / [GRID D]) - > [GRID E]

Todas essas operações podem ser realizadas a partir da função MAP CALCULATOR do módulo Spatial Analyst do ArcView 3.2.



**FIGURA 7** – Tabela de atributos do arquivo formato GRID da variável dependente. Os pontos eram um arquivo que estava em formato vetorial que foi transformado para GRID. Os pixels em vermelho correspondem a pelo menos uma presença e os pixels em verde às ausências.



**FIGURA 8** – Representação das operações para obtenção dos valores das variáveis contínuas nas células de ausência e presença do evento de interesse. A e A1 são variáveis contínuas; B é a variável dicotômica com valor 1 para presença; B1 é a variável dicotômica com 1 para ausência; C são os valores das células da camada digital da variável contínua com presença e C1 a mesma operação para as células com ausência.

---

# Uso de modelos aditivos generalizados na estimativa da distribuição potencial de espécies

PAULO DE MARCO JÚNIOR

Universidade Federal de Goiás, Goiás, Brasil.  
e-mail: pdemarco@icb.ufg.br

## RESUMO

Os modelos de distribuição potencial têm sido cada vez mais utilizados para prever a distribuição atual das espécies, discutir padrões de riqueza e modelar sua distribuição em relação às mudanças climáticas globais. Nesse trabalho, apresento algumas bases teóricas relevantes à compreensão desses modelos e à discussão sobre os mecanismos que determinam a distribuição das espécies. A partir disso, discuto a estrutura estatística dos modelos aditivos generalizados e seu potencial como métodos preditivos eficientes de modelagem. Um exemplo, usando espécies de planta do gênero *Inga*, é apresentado, mostrando a facilidade de execução dessa estratégia utilizando a metodologia GRASP (predição espacial por regressão generalizada) e algumas limitações e perspectivas para melhorar as técnicas de definição de pseudoausências nesses modelos.

## ABSTRACT

*Species distribution models are increasingly used to predict present day distribution of species, discuss their richness patterns and model its distribution in relation to global climatic change. Here, I present some theoretical basis to understand these models and the discussion about the mechanisms that determine species distribution. Based on this general background, I discuss the statistical structure of the generalized additive models and its potential use as efficient predictive modeling technique. A working example using plant species of *Inga* is presented, showing how easy is the execution of this strategy using GRASP (generalized regression and spatial prediction) methodology and some limitations and perspectives to improve the pseudo-absence use in these models.*

## COMO SE ESPERA DETERMINAR A DISTRIBUIÇÃO POTENCIAL DE UMA ESPÉCIE?

Dos distintos modelos que estão sendo apresentados neste conjunto de trabalhos e em muitos outros, que estão disponíveis na literatura (Peterson, 2003; Guisan & Thuiller, 2005; Phillips *et al.*, 2006), pode-se perceber

que a busca para determinar a distribuição potencial de espécies biológicas mantém um procedimento geral comum. Na maioria destes métodos, se assume que:

1. As espécies respondem a gradientes ambientais de uma forma previsível;
2. A forma da relação entre a presença ou a presença/ausência da espécie relatada para uma parte de sua

área de distribuição ou em uma amostra de sua distribuição reflete a forma desta relação em toda a área;

3. É possível estimar a forma desta distribuição e extrapolar este resultado para outras áreas, determinando áreas onde potencialmente esta espécie está ou deveria estar presente.

O item 1 é básico e implicitamente admitido em todos os modelos. O item 2 é um dos que têm atraído grande atenção por incluir o problema sobre a forma como os dados são relatados. Podem-se distinguir dois tipos de dados: i) dados onde só a presença é relatada, como os dados de museus e herbários; e ii) dados de presença e ausência. Espera-se que dados de presença e ausência sejam tomados de forma padronizada entre locais de coleta, evitando problemas de amostragem (Zaniewski *et al.*, 2002). Modelos como o GARP – “Genetic Algorithm for Rule Prediction” (Stockwell & Noble, 1992; Peterson, 2001) e o ENFA – “Ecologic Niche Factor Analysis” (Hirzel *et al.*, 2001, Brotons *et al.*, 2004) podem trabalhar com dados apenas de presença de espécies e são, nesse sentido, mais robustos.

A maior parte dos métodos se distingue pela forma como estimam a forma da curva (o item 3 no esquema acima), mas existem diferenças maiores em alguns deles. Em especial, os modelos aditivos generalizados vão diferir de uma forma radical neste aspecto, o que pode determinar que seu uso deva estar relacionado a diferentes propostas, como espero demonstrar no decorrer deste texto.

## QUE FATORES DETERMINAM A DISTRIBUIÇÃO DAS ESPÉCIES?

O padrão de distribuição de espécies sempre tomou conta de uma parte considerável da literatura ecológica (Andrewartha & Birch, 1954; Bock & Ricklefs, 1983; Brown, 1984; Arita *et al.*, 1990; Austin, 2007). A forma como determinada espécie se distribui ao longo de um gradiente ambiental tanto é usada como base para teorias gerais, incluindo sucessão ecológica (Reader, 1992) ou a teoria do rio-contínuo (Vannote *et al.*, 1980), como para explicações de sua distribuição em uma escala biogeográfica.

Muitas destas abordagens baseiam suas predições em uma forma específica de relação entre a probabilidade de ocorrência e o gradiente ambiental. Assim, as funções de incidência de Diamond (Diamond, 1975; Gilpin & Diamond, 1981) podem ser consideradas como parte deste sistema geral, distinguindo-se apenas pelo fato do gradiente sob análise ser a área de ilhas oceânicas.

A abordagem das funções de incidência é interessante por diferenciar modelos possíveis de distribuição em função de características bionômicas das espécies, em especial suas habilidades competitivas (Grant, 1966; Grant & Abbott, 1980), mas também suas habilidades de colonização (Cole, 1981; Berggren, 2001). No entanto, outras abordagens também buscam distinguir, em maior ou menor grau, a forma da curva de resposta da espécie em relação ao gradiente ambiental e suas características intrínsecas.

Evidentemente, a forma mais básica que se pode apresentar do que foi exposto acima são as curvas que determinam os limites de tolerância das espécies (Austin, 2007). Em geral, espera-se que a eficiência das espécies em sobreviver, crescer e reproduzir ocorra em função de sua eficiência em resposta ao gradiente ambiental. Os livros-texto de ecologia buscam expressar esta relação com uma função normal ou outra de forma semelhante (*e.g.* Begon *et al.*, 1996).

Se as respostas das espécies aos gradientes ambientais são, em parte, resultado de suas características intrínsecas, espera-se que haja diferentes tipos de formas das curvas de resposta, no mínimo algo mais diverso do que variações de curtose e assimetria da curva normal. Estes diferentes modelos deveriam resultar da forma como a espécie responde ao gradiente, seja ele simplesmente um fator abiótico que interage com o conjunto de adaptações morfológicas e fisiológicas da espécie, ou um recurso que determina complexos modelos de resposta competitiva e de interações intra- e interespecíficas, incluindo um aumento do desvio da simetria típica da curva normal e respostas não uni-modais (Oksanen & Minchin, 2002). Evidentemente, os modelos mais complexos deveriam advir da forma como alguns gradientes abióticos podem determinar a distribuição, o acesso ou a eficiência de uso de recursos no ambiente.

Neste ponto, podemos construir a seguinte cadeia lógica: características bionômicas → resposta a gradientes ambientais → forma da curva de resposta → padrão de distribuição biogeográfica. Partindo desta teoria, modelos estatísticos usados para estimar a forma da curva podem ser também utilizados para testar hipóteses sobre como se processa o efeito dos fatores ambientais, ou como determinadas espécies, que compartilham características ecológicas, devem apresentar o mesmo tipo de resposta ao gradiente. Modelos deste tipo, à semelhança dos modelos baseados na função de incidência, utilizados nos estudos de metapopulação e ecologia da paisagem (Taylor, 1991; Wahlberg *et al.*, 2002; Ovaskainen & Hanski, 2003), podem ser

utilizados não apenas para a estimativa da distribuição potencial de uma espécie, mas também para testar os processos que determinam esta distribuição.

Como qualquer modelo teórico que se presta a distinguir processos, estes métodos devem representar curvas simples. Bons exemplos de modelos com estas características seriam os modelos logísticos (ou Logit) e Probit. Ambos têm certa flexibilidade, ajustando-se a alguns poucos padrões possíveis de distribuição, e são estimados com o uso de poucos parâmetros (Manel *et al.*, 1999; Pearce & Ferrier, 2000). Alguns modelos são, portanto, mais interessantes para refletir essas relações, sendo desejável que nós otimizemos neles sua generalidade, enquanto em outros o que buscamos otimizar é uma predição prática acurada da distribuição potencial da espécie (Guisan & Zimmermann, 2000). Neste último grupo é que os métodos baseados em modelos aditivos generalizados estão agrupados.

## O QUE SÃO MODELOS ADITIVOS GENERALIZADOS?

Os modelos aditivos generalizados foram propostos por Hastie & Tibshirani (1990). Estes modelos são caracterizados pelo fato de assumirem que a média de uma variável resposta depende de um preditor aditivo, através de uma função de ligação. Uma característica importante destes métodos é que a função de distribuição de probabilidade desta variável dependente pode ser qualquer uma dentre aquelas que fizerem parte da família exponencial. Isto permite a construção de modelos aditivos para variáveis com distribuição normal (*e.g.* biomassa de espécies), Poisson (*e.g.* riqueza de espécies) ou binomial (*e.g.* função de incidência das espécies).

Suponha que  $y$  é uma variável resposta aleatória e que  $X_1 \dots X_p$  é um conjunto de variáveis preditoras. Em um procedimento de regressão, consideramos como a variável dependente  $Y$  pode depender de  $X_1 \dots X_p$ , e o valor esperado de  $Y$  assume a seguinte forma linear:

$$E(Y) = f(X_1, \dots, X_p) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$$

Equação 1

Dada uma amostra, os parâmetros  $b_1 \dots b_j$  podem ser estimados, normalmente pelo método dos quadrados mínimos. O modelo aditivo generaliza o modelo linear da seguinte maneira:

$$E(Y) = f(X_1, \dots, X_p) = s_0 + s_1(X_1) + \dots + s_p(X_p)$$

Equação 2

Onde  $s_i(\cdot)$  são as funções de ajuste. Já o modelo aditivo generalizado assume também a existência de uma função de ligação entre  $f(X_1, \dots, X_p)$ . É isto que permite que se possam utilizar diferentes funções de distribuições, além da distribuição normal, para modelar os resíduos. É claro que para cada tipo de distribuição haverá uma função de ligação apropriada.

É esta propriedade que expande a utilidade dos modelos generalizados. Usualmente os modelos baseados na função de distribuição normal não são apropriados para uma série de aplicações na área ecológica. Os dados de riqueza de espécies são, na maior parte, considerados discretos (sendo discreta uma variável em que a menor distância entre dois números é grande comparada com a amplitude de variação). Quando estudamos riqueza de espécies de mamíferos ou répteis, o modelo mais apropriado assumiria uma distribuição Poisson para os resíduos. É claro que se você estiver estudando número de espécies de Coleoptera no dossel de florestas Amazônicas, a curva normal poderá ser bastante apropriada.

Voltando ao modelo, a partir da Equação 2:

$$\eta = s_0 + \sum_{i=1}^p s_i(X_i)$$

Equação 3

onde  $\eta$  é considerado o termo aditivo. A relação entre a média da variável resposta (digamos  $\mu$ ) e o fator aditivo é definida pela função de ligação:  $g(E(Y)) = \eta$ .

O ajuste de um modelo generalizado é feito mantendo algumas características dos modelos aditivos que o precederam. Nestes modelos, cada função é ajustada separadamente, garantindo uma maior eficiência local, mas podendo perder em universalidade. Esta estratégia permite enfrentar o que foi chamado de “maldição da dimensionalidade” na qual o aumento de variáveis preditoras eleva de forma drástica as estimativas de variância. Além disto, o ajuste local é feito minimizando uma função de perda (normalmente um erro quadrático) pela escolha de funções, em vez de parâmetros individuais. Este procedimento poderia ser importante para uma maior eficiência quando os parâmetros individuais apresentam alta covariância dentro dos modelos.

Adicionalmente, Hastie & Tibshirani (1990) chamam muita atenção para o algoritmo de ajuste por “backfitting”, considerado essencial nesta aplicação. Neste mecanismo, cada interação do procedimento de ajuste busca minimizar a função de perda em relação a uma das funções (uma das variáveis preditoras)

independentemente, até a convergência. Hastie & Tibshirani (1990) provaram que este algoritmo atinge uma solução única independente de valores iniciais para funções de ajuste simétricas, como, por exemplo, as funções *spline*.

Por fim, o tipo de função de ajuste é uma característica essencial no modelo aditivo generalizado. Nestes modelos, funções não-paramétricas *spline* são utilizadas. Nestas funções, o número de parâmetros pode ser controlado e seu aumento (que leva um aumento do número de graus de liberdade do modelo) leva um aumento da qualidade de ajuste. Este comportamento pode ser comparado ao aumento do número de graus de um ajuste polinomial, que produz uma função mais flexível ao ajuste aos dados coletados com sacrifício de graus de liberdade.

Esse tipo de modelo foi utilizado para predição de distribuição de espécies em uma variedade de contextos nos últimos anos (Guisan *et al.*, 2002; Zaniwski *et al.*, 2002; Leathwick *et al.*, 2006), e faz parte de programas computacionais para análise de distribuição como o BIOMOD (Thuiller, 2003) e o GRASP (Lehmann *et al.*, 2002), o que tende a popularizar ainda mais seu uso.

#### **Padrões de distribuição de duas espécies do gênero *Inga***

Em um pequeno exercício sobre o funcionamento deste método, eu usei a distribuição de espécies do gênero *Inga* (Mimosaceae). As razões para isto são: a disponibilidade destes dados (cedidos em formato digital pelo Dr. Mike Hopkins) e a existência, neste conjunto de dados, de informações de espécies restritas e amplamente distribuídas, o que seria desejável para a avaliação do método. Este conjunto de dados refere-se também a uma classe importante de informações para estimar a distribuição potencial de espécies, que são espécies depositadas em herbários e museus. Estes dados não advêm de levantamentos completos de flora, mas do acúmulo de informações na literatura e em herbários. Esta fonte será extremamente utilizada na determinação de áreas de distribuição potencial de espécies de agora a um futuro próximo e é interessante avaliar seus resultados. Por fim, é evidente, mas sempre importante dizer, que estes dados são incompletos e resultam de um sem número de vieses de coletas, de pesquisadores e até mesmo de identificação de espécies. Talvez seja a hora de confiar mais no teorema da estatística que sugere que a soma de  $n$  variáveis aleatórias, independentemente distribuídas, tende a uma distribuição normal.

Informações gerais sugerem que o gênero *Inga* é composto por espécies que se distribuem principalmente na Amazônica e na Mata Atlântica, com espécies que

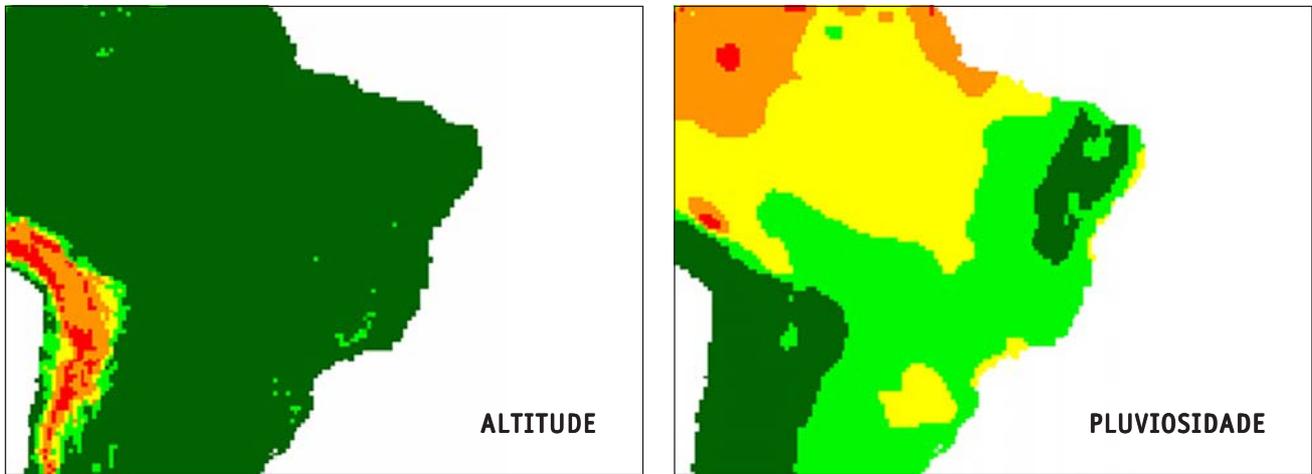
possivelmente estão associadas a regiões de maior pluviosidade e temperatura (M. Hopkins, com. pess.).

#### **AJUSTANDO MODELOS ADITIVOS GENERALIZADOS PELO PROCEDIMENTO GRASP**

Neste exercício usei o procedimento GRASP – “Generalized Regression Analysis for Spatial Prediction” de acordo com a descrição e sugestões presentes em Lehmann *et al.* (2002). O procedimento pode ser implementado através do programa estatístico R, de distribuição gratuita pela rede mundial de computadores. Informações detalhadas e aquisição dos programas necessários podem ser obtidas em <http://www.cscf.ch/grasp>.

O processo começa com a aquisição das variáveis preditoras em um formato apropriado para a análise. Usualmente estes dados são informações climáticas, de relevo ou de distribuição de alguma característica considerada importante e que está apresentada no formato de uma matriz, onde cada célula corresponde a um pixel da imagem base do estudo (na terminologia dos usuários de sistemas de informação geográfica: “ASCII RASTER”). Esta imagem representa o mapa da região em que, tanto os dados de distribuição foram obtidos, quanto para onde as predições do modelo serão produzidas. Na Figura 1 apresento dados de altitude e pluviosidade para uma porção da América Latina, os quais serão a base de dados ambiental neste exercício.

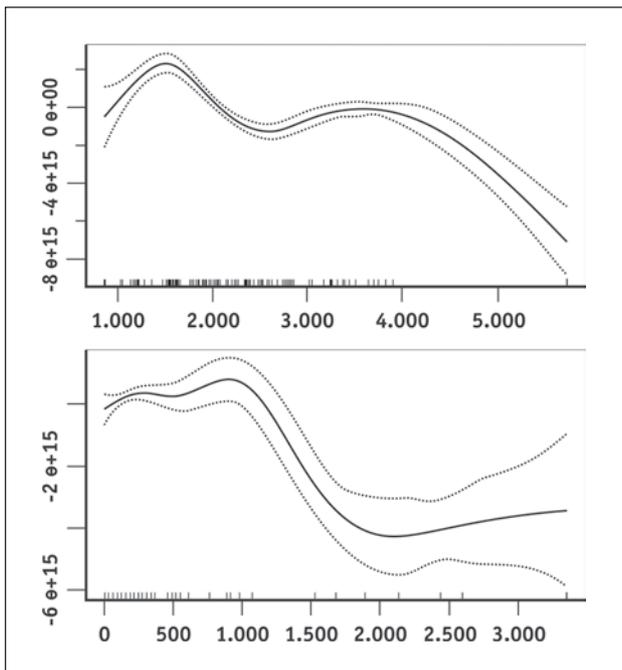
Para este exercício, utilizei os dados de *Inga virescens* Benth. e *I. macrophylla* Willd. No entanto, os dados utilizados provêm de material depositado em herbários onde apenas é relatada a informação da presença de determinada espécie. Zaniwski *et al.* (2002) discutem em detalhe as várias alternativas para se tratar dados de presença apenas, utilizando os modelos aditivos generalizados. O processo passa, necessariamente, por se estabelecer o que aqueles autores chamaram de “pseudoausências”. Ao invés de usar um critério estatístico (que discutirei mais além) para estabelecer as pseudoausências, utilizei um arrazoado prático. Considerarei que os coletores que estiveram em campo não tinham vícios sérios de amostragem de uma espécie de *Inga* em relação à outra e que, quando estava presente em uma localidade, ela sempre foi relatada. Assim, associei pseudoausências para cada espécie para os locais onde ela não foi coletada, mas a outra espécie foi. Este método é obviamente viciado, principalmente comparando-se espécies de ampla distribuição com espécies de distribuição restrita, mas foi utilizado aqui apenas como uma abordagem preliminar.



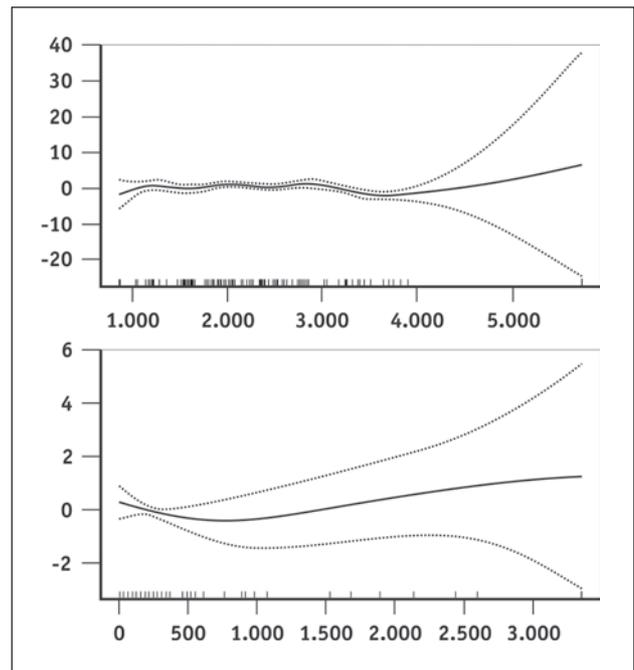
**FIGURA 1** – Distribuição de valores de altitude e pluviosidade em uma porção da América Latina que será a base deste exercício. As cores representam as classes de valores de cada variável em um gradiente, na seguinte ordem: verde escuro, verde claro, amarelo, alaranjado e vermelho escuro. Esta é apenas uma representação geral com os dados em classes para fácil visualização dos grandes padrões, porém os dados utilizados na análise são os dados brutos, sem classificação.

No ajuste pelo procedimento GRASP, o próximo passo é a escolha do modelo que será utilizado. Como os dados agora representam presença/ausência, escolhi uma ligação por uma função logística e, como função de distribuição dos resíduos, a distribuição Binomial. Outras escolhas podem ser admitidas ou testadas, mas essa é provavelmente a mais apropriada.

A partir disso, o programa exibe uma série de resultados que descrevem o ajuste do modelo e seu resultado geral. Dentre eles, é apresentado um gráfico que apresenta as previsões para cada variável e seus respectivos intervalos de confiança. Estes resultados são apresentados na Figura 2, para *I. virescens*, e na Figura 3, para *I. macrophylla*.



**FIGURA 2** – Apresentação da saída do programa GRASP para o ajuste das presença/ausência de *I. virescens* para pluviosidade (gráfico superior) e altitude (gráfico inferior). As linhas tracejadas representam o intervalo de confiança de 95%.



**FIGURA 3** – Apresentação da saída do programa GRASP para o ajuste das presença/ausência de *I. macrophylla* para pluviosidade (gráfico superior) e altitude (gráfico inferior). As linhas tracejadas representam o intervalo de confiança de 95%.

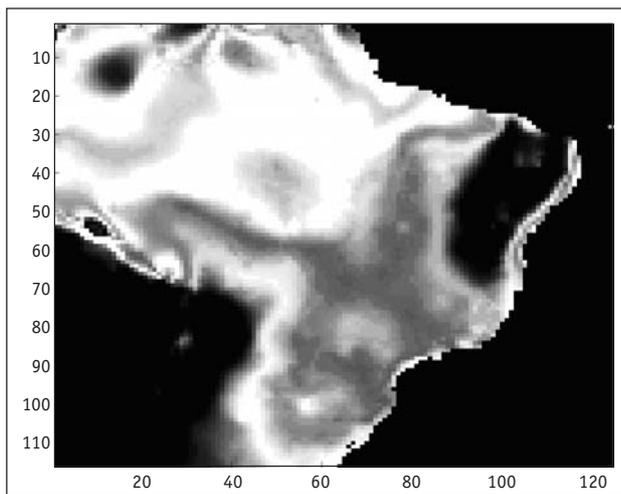
Os resultados sugerem um ajuste muito mais claro para *I. virescens* do que para *I. macrophylla*. Em especial, os maiores intervalos de confiança podem ser observados para o ajuste à altitude em *I. macrophylla*. Isto pode sugerir que a precipitação é mais importante que a altitude na determinação da ocorrência desta espécie. A interpretação do modelo é feita principalmente através da inspeção dos gráficos para “interpretar e entender esses modelos de regressão” (Lehmann *et al.*, 2002).

Na realidade, o ajuste de *I. macrophylla* é pobre ( $R^2=0,199$ ), mas o ajuste da função de precipitação [S(precipitação)] foi significativo ( $\chi^2=23,279$ , graus de liberdade estimado = 8,  $p=0,002$ ), enquanto o ajuste à altitude não foi ( $\chi^2=2,763$ , graus de liberdade estimado = 2,  $p=0,283$ ).

Ao final do processo, o programa produz arquivos de predição da ocorrência da espécie para a área da imagem de entrada das variáveis ambientais, que podem ser lidos e tratados nos sistemas de informação geográfica mais comuns (Figuras 4 e 5). Há uma diferenciação maior entre as áreas de ocorrência para *I. virescens*, com distribuição mais claramente associada a áreas de clima mais úmido, enquanto *I. macrophylla* apresenta uma distribuição que inclui as áreas de cerrado do Brasil central. A ocorrência de ambas não é predita para áreas muito secas do nordeste brasileiro.

#### CONSIDERAÇÕES E PERSPECTIVAS SOBRE O USO DE MODELOS ADITIVOS GENERALIZADOS

O ajuste apresentado nas Figuras 2 e 3 é muito bom para a natureza e quantidade de dados deste exercício. Em

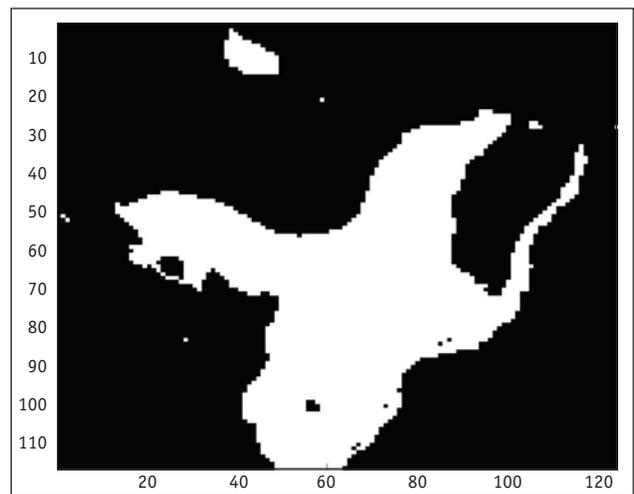


**FIGURA 4** – Predição de ocorrência para *I. virescens*, segundo o modelo GRASP. Áreas claras representam áreas de presença potencial da espécie.

grande parte isto é resultado da grande flexibilidade dos modelos aditivos generalizados, gerada pelo uso de funções não-paramétricas de ordem alta. Isto torna a interpretação dos ajustes como processos causais muito difíceis, podendo preferivelmente ser considerada uma análise exploratória das possíveis relações entre essas variáveis. Em muitas áreas de pesquisa os modelos aditivos generalizados são principalmente considerados como métodos de análise exploratória ou “*data mining*”.

A facilidade de acesso a programas de distribuição livre deve ser considerada uma vantagem dessa abordagem. Espera-se um desenvolvimento de novos aplicativos ou a adição de algumas novas propriedades em modelos como o GRASP, à medida que este ramo de aplicações se desenvolva (Lehmann *et al.*, 2003).

Como quaisquer outros modelos, estes também são sensíveis ao tamanho das amostras. Stockwell & Peterson (2002) demonstraram este efeito sobre vários métodos, incluindo a regressão logística e o GARP, e concluíram que este efeito é grande e que a escolha de quais variáveis ambientais utilizar tem efeito imprevisível sobre a acurácia das estimativas. Nada há que não sugira que o mesmo não ocorra nos modelos aditivos generalizados. Quanto menor a amostragem do padrão de distribuição de uma determinada espécie, maiores podem ser os vícios dos modelos gerados. Espera-se que estes vícios sejam no sentido de prever áreas de distribuição mais restritas, mas a escolha de uma variável incorreta pode superestimar a distribuição de uma espécie. Isto ocorrerá principalmente se outra variável capaz de restringir a distribuição da espécie em questão não for incluída no modelo.



**FIGURA 5** – Predição de ocorrência para *I. macrophylla*, segundo o modelo GRASP. Áreas claras representam áreas de presença potencial da espécie.

Uma das questões centrais é o uso de pseudoausências. Zaniewsky *et al.* (2002) utilizam duas formas distintas de produzir pseudoausências a partir de métodos estatísticos. Seus resultados são bastante razoáveis se comparados com o modelo baseado em dados reais de ausência (produzidos com coletas de campo com esforço padronizado). No entanto, o número de amostras daquele estudo está muito além da maioria dos bancos de dados atuais sobre sistemas amazônicos e, possivelmente, da maioria absoluta dos bancos de dados de distribuição de espécies brasileiras. Este permanece como um ponto sério na utilização destes métodos e sugere que a coleta de informações de campo com esforço padronizado ainda é uma prioridade para a utilização dessa técnica.

Por fim, é preciso salientar que a validação é um ponto crucial na avaliação do sucesso do uso de modelo. A coleta de dados de campo nas áreas onde o modelo prevê e em áreas onde ele não prevê a ocorrência da espécie é vital, tanto para o uso apropriado dos resultados do modelo no futuro, garantindo sua acurácia, quanto para encontrar falhas e, talvez, novas possibilidades de ajuste.

## AGRADECIMENTOS

Os dados utilizados neste trabalho foram gentilmente cedidos pelo Dr. Mike Hopkins, incluindo parte de uma de suas apresentações interativas em um curso de campo. Esse trabalho foi desenvolvido como parte do projeto “Escalas biogeográficas e conservação de espécies: Modelos preditivos para distribuição de espécies e de diversidade alfa e beta”, financiado pelo CNPq, e do projeto “Hotspots de Biodiversidade de Odonata”, financiado pela Fundação “O Boticário” e pelo Instituto de Pesquisas da Mata Atlântica (IPEMA).

## REFERÊNCIAS BIBLIOGRÁFICAS

- Andrewartha, H.G. & L.C. Birch. 1954. The distribution and abundance of animals. University of Chicago Press, Chicago.
- Arita, H.T., J.G. Robinson & K.H. Redford. 1990. Rarity in Neotropical forest mammals and its ecological correlates. *Conservation Biology* 4: 181-192.
- Austin, M. 2007. Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling* 200: 1-19.
- Begon, M., J.L. Harper & C.R. Townsend. 1996. *Ecology: individuals, populations and communities*, 3rd edition. Blackwell Science, Oxford.
- Berggren, A. 2001. Colonization success in Roesel's bush-cricket *Metroptera roeseli*: The effects of propagule size. *Ecology* 82: 274-280.
- Bock, C.E. & R.E. Ricklefs. 1983. Range size and local abundance of some North American songbirds: a positive correlation. *The American Naturalist* 122: 295-299.
- Brotos, L., W. Thuiller, M.B. Araujo & A.H. Hirzel. 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27: 437-448.
- Brown, J.H. 1984. On the relationship between abundance and distribution of species. *The American Naturalist* 124: 255-279.
- Cole, B.J. 1981. Colonizing abilities, island size, and the number of species on archipelagoes. *The American Naturalist* 117: 629-638.
- Diamond, J.M. 1975. Assembly of species communities. In: M.L. Cody & J.M.C.P. Diamond (eds). *Ecology and evolution in communities*. pp. 342-444. Harvard University Press, Cambridge.
- Gilpin, M.E. & J.M. Diamond. 1981. Immigration and extinction probabilities for individual species: relation to incidence functions and species colonization curves. *Proceedings of National Academy of Sciences of USA* 78: 392-396.
- Grant, P.R. 1966. Ecological compatibility of bird species on islands. *The American Naturalist* 100: 451-462.
- Grant, P.R. & I. Abbott. 1980. Interspecific competition, island biogeography and null hypothesis. *Evolution* 34: 332-341.
- Guisan, A., T.C. Edwards & T. Hastie. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling* 157: 89-100.
- Guisan, A. & W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8: 993-1009.
- Guisan, A. & N.E. Zimmermann. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147-186.
- Hastie T. & R. Tibshirani. 1990. *Generalized additive models*. Chapman & Hall, London.
- Hirzel, A.H., V. Helfer & F. Metral. 2001. Assessing habitat-suitability models with a virtual species. *Ecological Modelling* 145: 111-121.
- Leathwick, J.R., J. Elith & T. Hastie. 2006. Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecological Modelling* 199: 188-196.
- Lehmann, A., J.M. Overton & J.R. Leathwick. 2002. GRASP: generalized regression analysis and spatial prediction. *Ecological Modelling* 157: 189-207.
- Manel, S., J.M. Dias & S.J. Ormerod. 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecological Modelling* 120: 337-347.
- Oksanen, J. & P.R. Minchin. 2002. Continuum theory revisited: what shape are species responses along ecological gradients? *Ecological Modelling* 157: 119-129.
- Ovaskainen, O. & I. Hanski. 2003. The species-area relationship derived from species-specific incidence functions. *Ecology Letters* 6: 903-909.
- Pearce, J. & S. Ferrier. 2000. An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecological Modelling* 128: 127-147.

- Peterson, A.T. 2001. Predicting species' geographic distributions based on ecological niche modeling. *Condor* 103: 599-605.
- Peterson, A.T. 2003. Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology* 78: 419-433.
- Phillips, S.J., R.P. Anderson & R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231-259.
- Reader, R.J. 1992. Herbivory, competition, plant mortality and reproduction on a topographic gradient in an abandoned pasture. *Oikos* 65: 414-418.
- Stockwell, D.R.B. & I.R. Noble. 1992. Induction of sets of rules from animal distribution data - A robust and informative method of data-analysis. *Mathematics and Computers in Simulation* 33: 385-390.
- Taylor, B. 1991. Investigating species incidence over habitat fragments of different areas: A look at error estimation. *Biological Journal of the Linnean Society* 42: 177-192.
- Thuiller, W. 2003. BIOMOD - optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology* 9: 1353-1362.
- Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell & C.E. Cushing. 1980. The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences* 37: 130-137.
- Wahlberg, N., T. Klemetti & I. Hanski. 2002. Dynamic populations in a dynamic landscape: the metapopulation structure of the marsh fritillary butterfly. *Ecography* 25: 224-232.
- Zaniewski, A.E., A. Lehmann & J.M.C. Overton. 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling* 157: 261-280.

---

# Algoritmo Genético para Produção de Conjuntos de Regras (GARP)

RICARDO SCACHETTI PEREIRA<sup>1\*</sup>  
MARINEZ FERREIRA DE SIQUEIRA<sup>1</sup>

<sup>1</sup> Centro de Referência em Informação Ambiental – CRIA, São Paulo, Brasil.

\* e-mail: scachett@gmail.com

## RESUMO

Neste trabalho é apresentado o uso de algoritmo genético GARP (*Genetic Algorithm for Rule Set Production*), com o objetivo de modelar a distribuição geográfica potencial de espécies. Tais algoritmos tentam encontrar relações não-aleatórias entre os dados de ocorrência da espécie e os dados ecológico/ambientais relevantes para a espécie (tais como: temperatura, precipitação, topografia, tipo de solo, geologia, entre outros) no ponto onde a espécie foi registrada. Os conceitos em que os algoritmos genéticos se baseiam estão relacionados com os conceitos biológicos de evolução das espécies e de seleção natural. O GARP é um algoritmo genético projetado para criar e otimizar um conjunto de regras ou proposições lógicas. Este conjunto de regras é a base de inferência usada para modelar o nicho ecológico das espécies. A qualidade de cada regra presente no modelo é avaliada através de uma função de adaptação, que é calculada através da significância estatística obtida pela aplicação da regra ao conjunto de pontos de treinamento fornecidos ao algoritmo. Técnicas de modelagem de nicho ecológico são importantes, porque ainda é precário o nível de conhecimento existente sobre a distribuição geográfica de plantas tropicais, sendo a maioria das espécies representadas por poucos pontos de amostragem e, muitas vezes, esses pontos são georreferenciados de maneira muito imprecisa. Essas técnicas podem ser utilizadas para aumentar o conhecimento sobre a distribuição potencial dessas espécies, analisar possíveis rotas de distribuição de espécies invasoras e de vetores de doenças infecciosas, avaliar os melhores ambientes para re-introdução de espécies a seus hábitos naturais, entre outros. Neste trabalho foram utilizados dados de espécies nativas do Brasil para avaliar a eficácia dos algoritmos em prever a distribuição geográfica potencial de espécies tropicais.

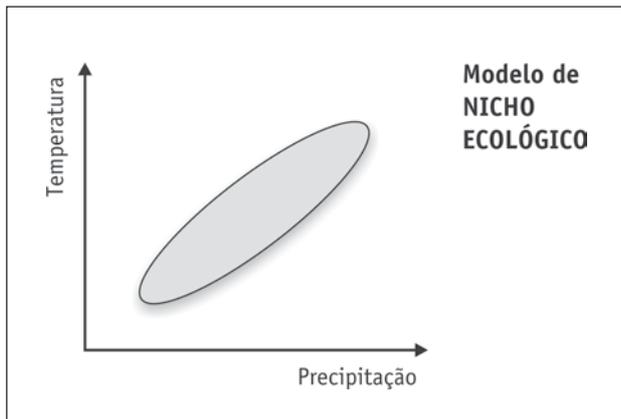
## ABSTRACT

*This work presents the use of the genetic algorithm GARP (Genetic Algorithm for Rule Set Production) in modeling the potential geographic distribution of species. Such algorithms tend to find non aleatoric relations between species occurrence data and ecological/environmental data relevant to the species (such as temperature, precipitation, topography, type of soil, geology, among others) at the point where the species was found. The concepts that genetic algorithms are related to species biological evolution and natural selection. GARP is a genetic algorithm projected to create and optimize a number of rules or logical propositions. This set of rules is the base of inference used to model species ecological niche. The quality of each rule present in the model is*

evaluated through an adaptation function that is calculated using statistical significance obtained through the application of the rule to the set of training points furnished by the algorithm. Ecological niche modeling techniques are important because the level of understanding about the geographic distribution of plants in tropics is still precarious, and most species are represented by very few sampling points that many times are georeferenced with little precision. Among many other applications, these techniques may be used to increase the knowledge about the potential distribution of species, analyze possible distribution routes of invasive species and of infectious disease vectors, and to evaluate the best environments to reintroduce species to their natural habitat. In this work, data on native species from Brazil was used to evaluate the algorithms' efficiency in predicting the potential geographic distribution of tropical species.

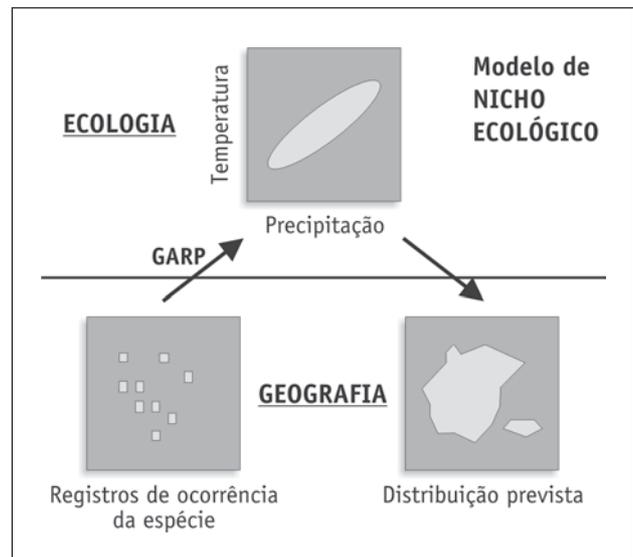
## MODELAGEM DE NICHU ECOLÓGICO

O processo de modelagem de nicho ecológico consiste em converter dados primários de ocorrência de espécies em mapas de distribuição geográfica, indicando a provável presença ou ausência da espécie, neste caso, através da aplicação de algoritmo genético. Estes modelos trabalham, na maioria dos casos, com o conceito de nicho ecológico fundamental da espécie. Este conceito foi definido por Hutchinson (1957) como sendo um conjunto de condições ecológicas com as quais as populações conseguem se manter, podendo ser representado por um espaço ecológico/ambiental multidimensional (Figura 1).

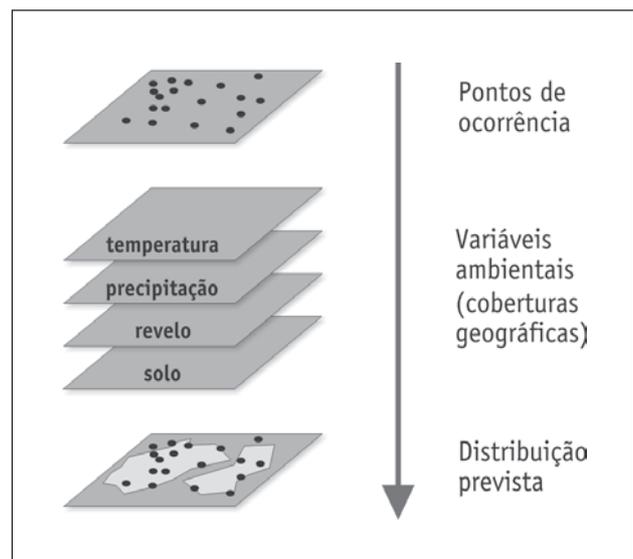


**FIGURA 1** – Exemplo de espaço bidimensional definido como nicho ecológico de espécie, baseado em dois parâmetros ambientais (temperatura e precipitação).

Tais algoritmos tentam encontrar relações não-aleatórias entre os dados de ocorrência da espécie e os dados ecológico/ambientais relevantes para a espécie (tais como: temperatura, precipitação, topografia, tipo de solo, geologia, entre outros) no ponto onde a espécie foi registrada (Figuras 2 e 3).



**FIGURA 2** – Esquema mostrando a relação entre os dados de campo e a previsão de distribuição geográfica, através do processo de modelagem de nicho ecológico.



**FIGURA 3** – Esquema mostrando a relação entre as variáveis ambientais e os pontos de ocorrência da espécie.

Existem vários métodos na literatura que podem ser aplicados para se obter modelos de nicho ecológico a partir dos pontos de ocorrência de espécies, conforme é mostrado na Figura 2, sendo que vários destes métodos são apresentados em outros artigos deste volume. No presente trabalho será apresentado em detalhe um destes métodos, chamado Algoritmo Genético para Produção de Conjuntos de Regras (GARP). Entretanto, antes de descrever o algoritmo GARP, é necessário apresentar uma breve descrição geral dos algoritmos genéticos.

## ALGORITMOS GENÉTICOS

Segundo Mitchel (1996), foi entre as décadas de 1950 e 60 que vários grupos de cientistas da computação começaram a estudar os sistemas evolucionários de maneira independente. Estes sistemas se baseavam na idéia de que a teoria da evolução das espécies (Darwin, 1859) poderia ser utilizada como ferramenta de otimização para problemas de engenharia. Todos esses sistemas seguiam a idéia central de que uma população de soluções candidatas para a resolução de um problema evoluía, e seus indivíduos eram melhorados através da aplicação de operadores inspirados na variação genética e na seleção natural. Ainda durante a década de 60, várias aplicações específicas foram desenvolvidas usando este conceito, em problemas de engenharia aeronáutica (Rechenberg, 1965), em problemas de ciência da computação (Fogel *et al.*, 1966), entre outras.

Foi no início da década de 70 que Holland e seu grupo de pesquisa começaram a estudar métodos para estudar o fenômeno da adaptação das espécies como ocorre na natureza, e desenvolver estratégias para a utilização desses conceitos em sistemas computacionais. Ao contrário dos outros pesquisadores da área que vinham aplicando a teoria da evolução a problemas específicos, Holland e seu grupo tinham como objetivo usar a teoria para desenvolver uma metodologia genérica de resolução de problemas. Como resultado de suas pesquisas, foram introduzidos os algoritmos genéticos (Holland, 1975). Nesta publicação, Holland apresentou o conceito dos algoritmos genéticos como uma abstração da evolução biológica e estabeleceu a base teórica para a adaptação de soluções através da variação genética e da seleção natural.

### Inspiração na teoria da evolução das espécies e terminologia

Grande parte dos conceitos em que os algoritmos genéticos se baseiam está relacionada com os conceitos

biológicos da evolução das espécies e da seleção natural. Para ilustrar como estes conceitos são empregados na resolução de um problema através de um algoritmo genético, considere o seguinte problema de minimização da função:

$$f(x, y, z) = x - 2y + 5z$$

no domínio dos números inteiros entre 0 e 255.

Uma solução válida para o problema é denominada um indivíduo. Um conjunto de indivíduos é considerado uma população. Neste caso, os pontos:  $p = (156, 18, 2)$  e  $q = (4, 8, 100)$  são dois indivíduos e o conjunto dos dois pontos é uma população. Cada indivíduo tem seu próprio genoma, que é representado por um ou mais cromossomos. No exemplo proposto, um cromossomo poderia ser representado pela concatenação das representações binárias das coordenadas dos pontos. Assim, os cromossomos  $P$  e  $Q$  (representando os pontos  $p$  e  $q$ ) teriam a seguinte forma:

$$P = 100111000001001000000010$$

$$Q = 000001000000100001100100$$

Cada cromossomo é subdividido em genes que se encontram disponíveis em um número de alelos, e cada indivíduo pode ter presente em seu genoma uma combinação arbitrária de alelos. No caso do exemplo proposto, os cromossomos seriam subdivididos em três genes, um para cada variável da função,  $x$ ,  $y$  e  $z$ . Os cromossomos  $P$  e  $Q$  teriam os alelos 156 e 4 para o gene  $x$ , respectivamente.

### Espaço de busca de soluções e as superfícies de otimização

A idéia de se buscar uma solução ótima para um problema dentre um conjunto de soluções candidatas é um conceito bastante comum na Ciência da Computação. Este conjunto que contém todas as possíveis soluções para um problema é chamado de espaço de busca. Cada solução pertencente ao espaço de busca está associada a um valor que define a qualidade da solução. Estes valores associados a cada solução, se expressados em um espaço multidimensional, formam o que se chama de função ou superfície de otimização. A Figura 4 mostra um exemplo de um espaço de soluções bidimensional em que as diferentes soluções são mostradas no eixo  $x$  e os respectivos valores da função de otimização  $f(x)$  são mostrados no eixo  $y$ .

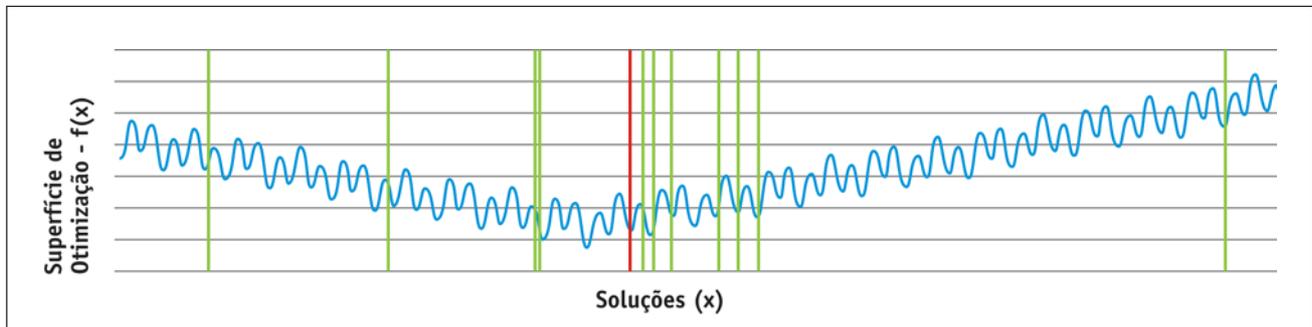


FIGURA 4 – Exemplo de superfície de otimização. Fonte: Obitko, 1998.

Na Figura 4, as barras verticais representam exemplos de soluções presentes no espaço de soluções para um problema arbitrário. A curva em azul representa a função de otimização. No exemplo da função de otimização proposta anteriormente, o espaço de busca seria a combinação de todos os valores possíveis das três variáveis em seus domínios e a superfície de otimização consistiria na superfície tridimensional definida pelo valor função de otimização em todos os pontos do espaço de busca.

### Operadores heurísticos

Uma vez definida a maneira com que as soluções de um dado problema são codificadas, pode-se, então, aplicar um conjunto de operadores heurísticos para modificar soluções existentes, obtendo novas soluções para o problema.

Este processo de modificação de soluções existentes com objetivo de obter novas soluções é o mecanismo básico usado pelos algoritmos genéticos para explorar o espaço de soluções de um problema.

Os operadores heurísticos mais empregados são a recombinação genética (*crossing-over*) e a mutação.

**Recombinação genética** – A operação de recombinação consiste na criação de novos cromossomos a partir da permutação de segmentos de dois outros cromossomos parentais. Veja um exemplo de permutação entre os cromossomos P e Q, definidos anteriormente na posição 12.

Antes da permutação tinha-se P e Q:

$P = 100111000001-001000000010$   $p = (156, 18, 2);$   
 $f(p) = 130$

$Q = 000001000000-100001100100$   $q = (4, 8, 100);$   
 $f(q) = 488$

Após a permutação tem-se P' e Q':

$P' = 100111000001-100001100100$   $p' = (156, 24, 100);$   
 $f(p') = 158$

$Q' = 000001000000-001000000010$   $q' = (4, 2, 2);$   
 $f(q') = 10$

Como resultado, têm-se dois novos pontos do espaço de soluções, com valores distintos na superfície de otimização.

**Mutação** – A mutação consiste em realizar uma alteração do valor da seqüência de um cromossomo numa posição aleatória. Veja o exemplo de uma mutação no cromossomo P, na posição 17.

Antes da mutação tinha-se:

$P = 100111000001<0>01000000010$   $p = (156, 18, 2);$   
 $f(p) = 130$

Após a mutação tem-se P':

$P = 100111000001<1>01000000010$   $p' = (156, 26, 2);$   
 $f(p') = 114$

Da mesma forma que ocorre na recombinação, a mutação proporciona um mecanismo para o algoritmo explorar o espaço de soluções do problema.

### Estrutura de execução dos algoritmos genéticos

De maneira geral, os algoritmos genéticos seguem a seguinte estrutura de execução:

1. Início: é gerada uma população inicial de n indivíduos;
2. Adaptação: avalia-se o grau de adaptação de cada indivíduo da população através da função de adaptação  $f(x)$ . Os indivíduos menos aptos, isto é, aqueles que têm valor de  $f(x)$  menor que um limiar  $v$ , morrem e são eliminados da população;
3. Reprodução: os indivíduos mais adaptados, isto é, que têm valor de  $f(x)$  maior que um limiar  $v$ , se reproduzem e criam um novo conjunto de indivíduos. A reprodução se dá da seguinte forma:
  - a. Seleção: dois indivíduos pais são selecionados a partir da população existente, de acordo com os valores de suas funções de adaptação  $f(x)$ ;
  - b. Recombinação: os cromossomos dos dois indivíduos são alinhados e recombinados para formar novos indivíduos;

- c. Mutação: cada novo indivíduo sofre mutação em posições aleatórias dos seus cromossomos, segundo uma probabilidade de mutação pré-definida;
4. Criação da nova população: os indivíduos gerados na fase de reprodução são incorporados à população;
5. Teste da condição de parada e início de nova interação: verifica-se se o número total de iterações chegou ao fim. Se sim, a execução do algoritmo é interrompida e o indivíduo mais apto da população atual é a solução do problema. Caso contrário, o algoritmo volta ao passo 2.

Os algoritmos genéticos processam as populações de cromossomos através das etapas descritas acima, substituindo uma população pela próxima. Os cromossomos são transferidos de uma geração para a próxima de acordo com o valor da função de adaptação, que consiste numa medida de quão boa é a solução representada pelo cromossomo (uma metáfora para o fenótipo). Para mais detalhes sobre os algoritmos genéticos, sugerem-se a leitura de Goldberg (1989), Davis (1996) e Mitchell (1996).

## GENETIC ALGORITHM FOR RULE-SET PRODUCTION (GARP)

O GARP é um algoritmo genético projetado para criar e otimizar um conjunto de regras ou proposições lógicas e foi desenvolvido originalmente por Stockwell e Noble (1992). Este conjunto de regras é a base de inferência usada para modelar o nicho ecológico das espécies.

### Tipos de regras

As regras usadas no algoritmo apresentam o seguinte formato:

Se **A** é verdadeiro, então **B** é verdadeiro.

Onde **A** é a pré-condição e **B** é o resultado ou previsão da regra. Neste algoritmo, a pré-condição é uma expressão lógica que define uma região do espaço ecológico (através das variáveis ambientais) e o resultado é a previsão de presença ou ausência da espécie.

No GARP as regras podem ser de quatro tipos diferentes: regras atômicas, de faixa, de faixa negada e logit.

As regras atômicas são as regras mais simples usadas pelo algoritmo. Elas usam um único valor de cada uma das variáveis ambientais na sua pré-condição. Um exemplo de regra atômica seria:

Se **Temperatura = 23°C** e **Elevação = 2000 m**, então o táxon está presente.

Ao serem criadas no modelo, cada variável ambiental recebe um valor aleatório.

As regras de faixa (ou *range rules*) são uma generalização das regras atômicas, sendo que as variáveis ambientais podem assumir valores dentro de uma faixa. Por exemplo:

Se **Temperatura está entre 23° e 28°C** e **Elevação está entre 2000 e 2700 m**, então o táxon está presente.

Ao serem incluídas no modelo, estas regras recebem valores das faixas associados a cada variável ambiental através do algoritmo de previsão de distribuição de espécies, chamado BIOCLIM (Nix, 1986).

As regras de faixa negadas consistem das regras de faixa em que a precondição da regra é negada. Por exemplo, uma regra de faixa negada teria a seguinte forma:

Se **Temperatura não está entre 23° e 28°C** e **Elevação não está entre 2000 e 2700 m**, então o táxon está presente.

As regras logit são uma adaptação dos modelos de regressão logística para o formato de regras lógicas. A regressão logística é uma forma de regressão em que o seu resultado é transformado em uma probabilidade. A regressão logística é definida pelas expressões abaixo.

$$p = 1 / [ 1 + e^{-y} ], y = c_0 + c_1x_1 + c_2x_2 + \dots + c_nx_n$$

onde  $p$  é a probabilidade que determina se a regra deve ser aplicada (equivalente à expressão lógica dos outros tipos de regras),  $y$  é a variável dependente (presença ou ausência da espécie),  $x_i$  é a  $i$ -ésima variável independente (variável ambiental),  $c_i$  é o coeficiente que multiplica a  $i$ -ésima variável independente e  $c_0$  é o termo constante ou de interceptação. A função de probabilidade  $p$  tem a forma de uma curva em S, como a apresentada a seguir.

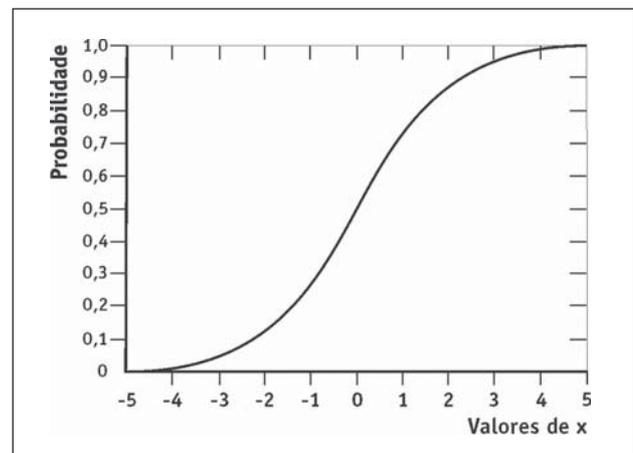


FIGURA 5 – Exemplo de função de probabilidade na regressão logística.

### Modelo de nicho ecológico no GARP

O algoritmo GARP define o modelo de nicho ecológico das espécies através de um conjunto de regras como as apresentadas anteriormente. As regras presentes num modelo de nicho ecológico são aplicadas hierarquicamente, de forma que as primeiras regras do modelo se sobrepõem às últimas. Um exemplo possível de modelo do GARP seria:

REGRA 1 – Se **Temperatura está entre 5° e 10°C** e **Elevação está entre 1000 e 2000 m**, então o táxon está presente (regra de faixa).

REGRA 2 – Se **Temperatura é de 18°C** e **Elevação é 1500 m**, então o táxon está ausente (regra atômica).

REGRA 3 – Se  $p = 1 / [1 + e^{-y}] > 0,75$ , onde  $y = 0,2 + \text{Temperatura} \times 0,45 + \text{Elevação} \times -0,12$ , então o táxon está presente (regra tipo logit).

### Codificação das regras no algoritmo genético

No GARP, cada regra é considerada como um indivíduo, e o conjunto de regras é considerado uma população, segundo a terminologia definida para os algoritmos genéticos. Internamente, as regras são codificadas através das faixas de valores ou coeficientes relativos às variáveis ambientais e também ao valor da previsão da

regra. As mesmas regras apresentadas acima seriam codificadas como cromossomos como na Tabela 1.

Os coeficientes das variáveis ambientais (representados nas colunas 3 a 6) correspondem aos genes que compõem os cromossomos. No caso das regras de faixa (e negadas) existem dois valores que definem a faixa para cada variável ambiental, e por isso cada variável é representada por dois genes. A previsão das regras (coluna “Presença ou Ausência”) também é codificada como um gene, podendo sofrer alterações durante a execução do algoritmo. Note que as regras do tipo logit e atômica não utilizam o segundo fator das variáveis. Estas posições (marcadas com um asterisco – \*) são ocupadas por valores irrelevantes para estas regras, mas que são importantes nas operações de mutação e recombinação.

A qualidade de cada regra presente no modelo é avaliada através de uma função de adaptação (última coluna da tabela), que é calculada através da significância estatística obtida pela aplicação da regra ao conjunto de pontos de treinamento fornecidos ao algoritmo.

### Aplicação dos operadores heurísticos sobre as regras

Durante a execução do algoritmo, as regras são modificadas aleatoriamente através de operadores heurís-

TABELA 1

Regra	Tipo da Regra	Temperatura: Primeiro Fator	Temperatura: Segundo Fator	Elevação: Primeiro Fator	Elevação: Segundo Fator	Presença ou Ausência	Função de Adaptação
1	Faixa	5	10	1000	2000	P	50%
2	Atômica	18	25*	1500	3000*	A	12%
3	Logit	0,45	0,17*	-0,12	0,02*	P	95%

TABELA 2

Situação	Regra	Temperatura: Primeiro Fator	Temperatura: Segundo Fator	Elevação: Primeiro Fator	Elevação: Segundo Fator	Presença ou Ausência	Função de Adaptação
Antes	1	5	10	1000	2000	P	50%
Depois	4	5	17	1000	2000	P	67%

TABELA 3

Situação	Regra	Temp: Primeiro Fator	Temp: Segundo Fator	Elevação: Primeiro Fator	Elevação: Segundo Fator	Presença ou Ausência	Função de Adaptação
Antes	1	5	10	1000	2000	P	50%
	2	18	25	1500	3000	A	12%
Depois	5	5	25	1500	3000	A	0%
	6	18	10	1000	2000	P	97%

ticos de recombinação e mutação. Por exemplo, uma operação de mutação sobre a regra 1, na posição do segundo gene (segundo fator da variável Temperatura), seria realizada como na Tabela 2.

Note que a mutação criou uma nova regra, que quando aplicada aos pontos de treinamento, obtém um valor maior na função de adaptação, devido à mudança realizada em um de seus genes.

Já uma operação de recombinação sobre a posição do segundo gene resultaria como na Tabela 3.

Note que os novos cromossomos gerados a partir da recombinação e da mutação são diferentes dos cromossomos originais e, por isso, apresentam valores de adaptação diferentes. Estes valores de adaptação resultantes podem ser tanto melhores quanto piores que os valores dos cromossomos originais.

### Seleção natural no GARP e início de uma nova iteração

Após a criação de novos cromossomos e inclusão destes na população existente, é executada uma operação de seleção natural. Nesta operação, aqueles cromossomos que têm valor da função de adaptação abaixo de um certo limiar pré-definido são eliminados da população.

O GARP processa os conjuntos de regras, gerando novos cromossomos, aplicando os operadores heurísticos a eles e em seguida executando a seleção natural, excluindo-se as regras menos eficientes sucessivamente. Quando um número predeterminado de iterações é atingido, o algoritmo é encerrado. É criado como resultado um conjunto de regras a partir dos indivíduos sobreviventes. Este conjunto de regras representa o modelo de nicho ecológico da espécie. Este modelo é aplicado de volta ao espaço geográfico, indicando as regiões onde provavelmente a espécie está presente ou ausente.

Diferentemente da concepção original dos algoritmos genéticos, em que a solução para o problema é representada apenas pelo indivíduo mais apto da população, o GARP considera toda a população (conjunto de regras) como solução para o problema de modelagem. Mais detalhes do processo utilizado pelo GARP está disponível na literatura (Stockwell, 1999; Stockwell & Peters, 1999).

Testes anteriores mostraram a habilidade do modelo em prever a distribuição de espécies em várias regiões (Peterson & Cohoon, 1999; Peterson *et al.*, 1999; Peterson, 2001; Peterson & Vieglas, 2001; Anderson *et al.*, 2002a, b; Fera & Peterson, 2002; Peterson *et al.*, 2002a, b, c; Stockwell & Peterson 2002a, b; Anderson *et al.*, 2003; Peterson *et al.*, submetido; Peterson *et al.*, 2003).

## APLICAÇÕES DO GARP NO BRASIL

O GARP vem sendo utilizado para estabelecer previsões sobre a distribuição geográfica de espécies para condições tropicais. A seguir, mostramos alguns exemplos do que vem sendo desenvolvido no Brasil, com base no uso do GARP.

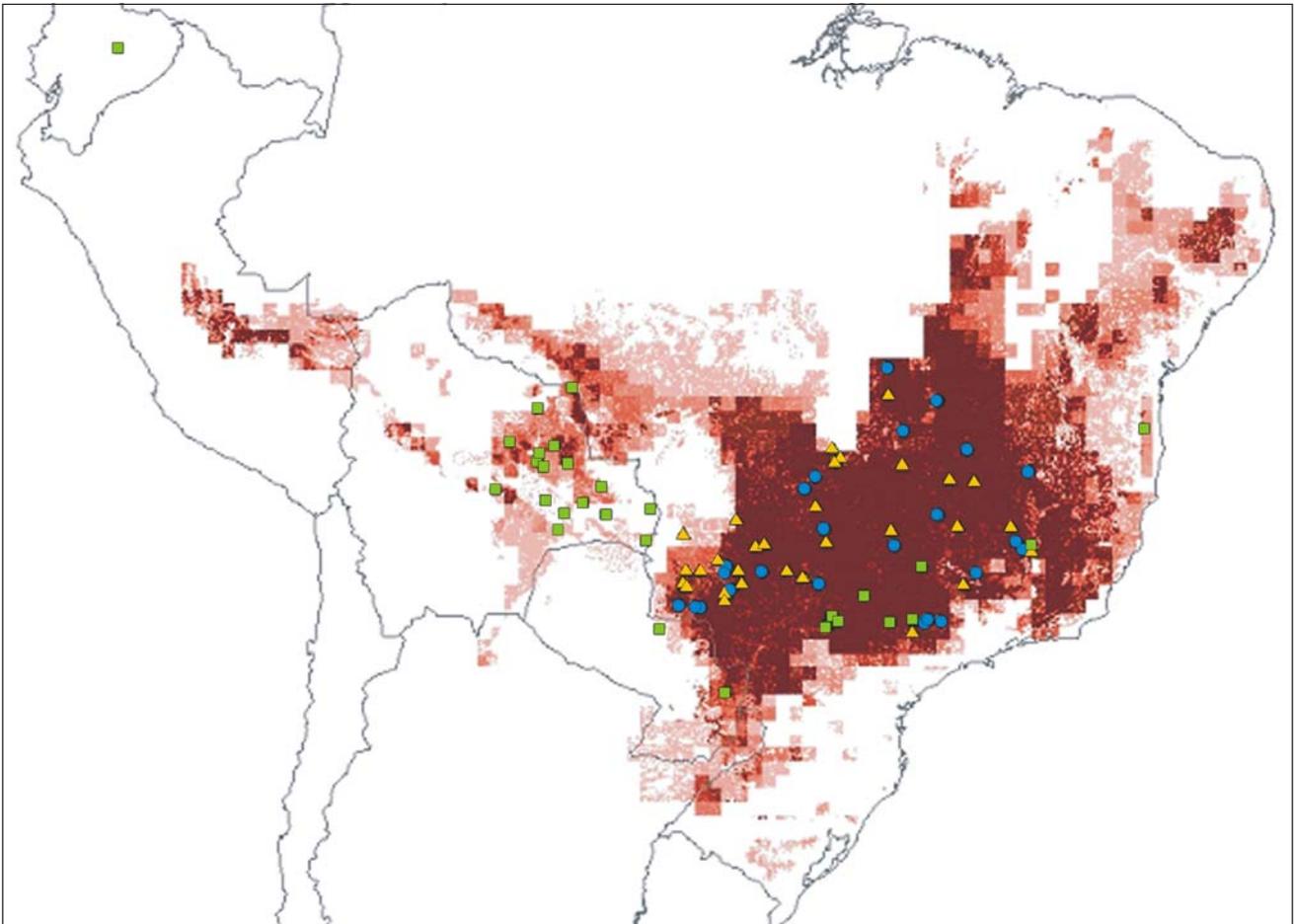
### Uso do GARP para prever a distribuição geográfica de espécies

Atualmente ainda é precário o nível de conhecimento existente sobre a distribuição geográfica de plantas tropicais, sendo a maioria das espécies representadas por poucos pontos de amostragem e, muitas vezes, esses pontos são georreferenciados de maneira muito imprecisa.

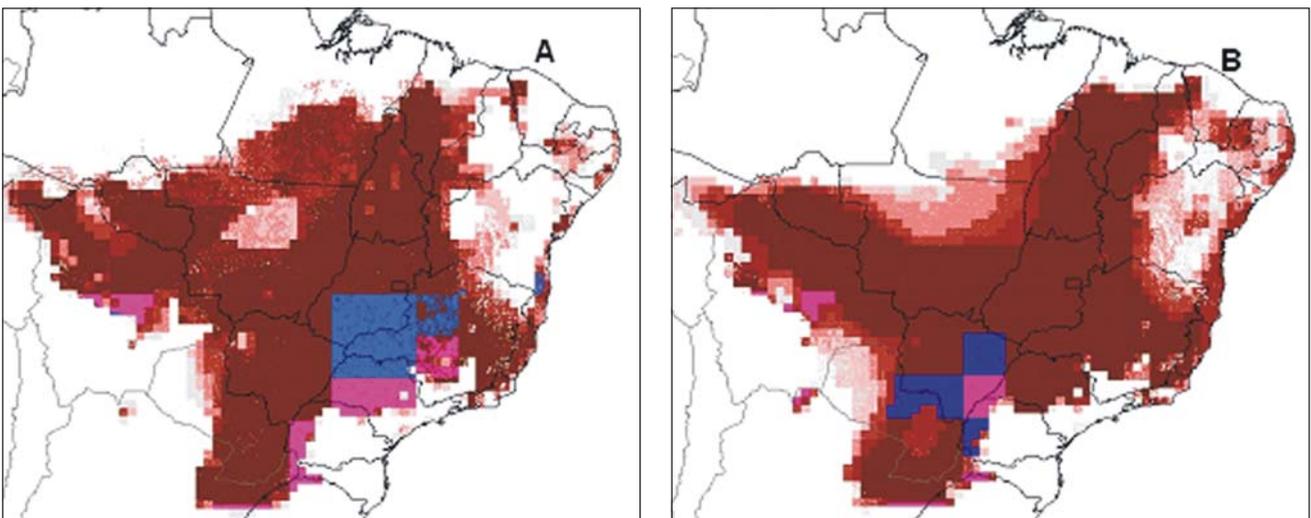
Como um primeiro teste para se avaliar a eficácia do GARP em prever a distribuição geográfica de espécies tropicais (Figura 6), foi selecionado um conjunto de dados, provenientes do Projeto de Cooperação Técnica Conservação e Manejo da Biodiversidade do Bioma Cerrado - EMBRAPA Cerrados - UnB - Ibama/DFID e RBGE/Reino Unido, para gerar os modelos (Koch, dados não publicados). Foi escolhido este conjunto de dados por se tratar de uma base de dados bastante completa e precisa, referente às espécies arbóreas de cerrado (Ratter *et al.*, 2001).

Este conjunto foi dividido em dois subconjuntos de dados independentes, um utilizado para construir os modelos e prever a distribuição geográfica das espécies (dados de treinamento), e outro utilizado para testar a previsão de distribuição (dados de teste). Posteriormente, devem-se empregar análises estatísticas para validar se o resultado obtido é significativo estatisticamente. Com base em análises deste tipo é possível avaliarmos se a modelagem de nicho ecológico é uma ferramenta robusta para auxiliar no entendimento da distribuição geográfica de espécies de plantas tropicais. A Figura 6 mostra um exemplo de resultado do uso de GARP para prever a distribuição geográfica de espécies arbóreas de cerrado.

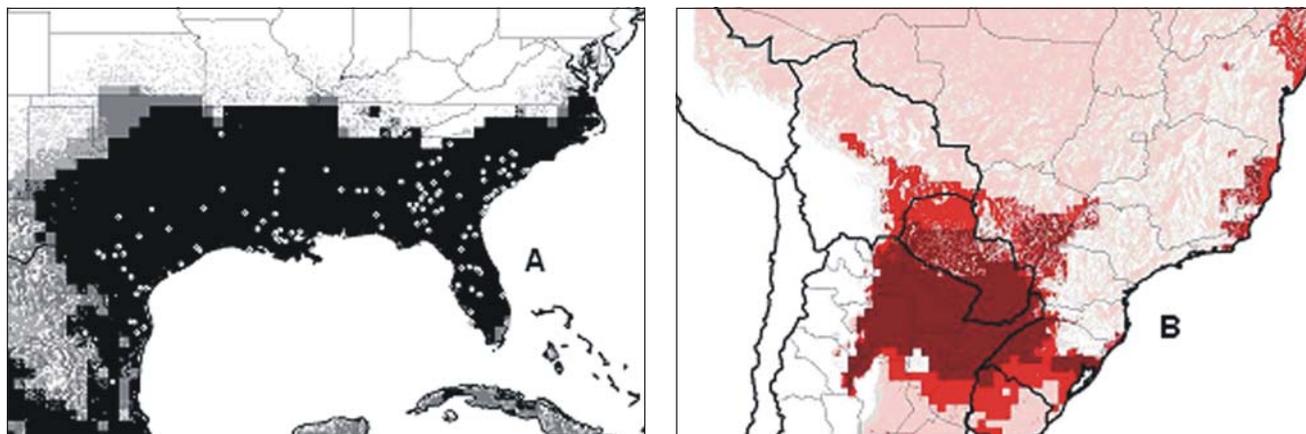
Os dados provenientes de outras bases de dados que não as disponíveis em Ratter *et al.* (2001), pontos representados por quadrados na Figura 6, não participaram da modelagem, mas ilustram, em alguns casos, a qualidade do modelo gerado pelo GARP. Em outros casos (como o ponto isolado no canto superior esquerdo da figura, localizado no Equador) mostram que o modelo não foi capaz de prever aquela região como de provável ocorrência da espécie. A causa disso, neste caso, muito provavelmente é resultante da ausência de pontos no conjunto de dados de treinamento que representassem as condições ambientais presentes naquela região.



**FIGURA 6** – Distribuição prevista para *Terminalia argentea*. Os círculos azuis representam os pontos usados na confecção do modelo (dados de treinamento); os triângulos amarelos são os pontos usados para testar o modelo (dados de teste); e os quadrados verdes são registros dessa espécie provenientes de outros conjuntos de dados (New York Botanical Garden, Biota/Fapesp e Missouri Botanical Garden).



**FIGURA 7** – Exemplo de previsão de distribuição atual (1961-1990) em vermelho escuro e futura (2055), em azul, para o cenário HHGDX50, e em rosa para o cenário HHGAX50, para espécies arbóreas de cerrado: A - *Rudgea viburnoides* (Cham.) Benth. (Rubiaceae) e B - *Qualea parviflora* Mart. (Vochoysiaceae).



**FIGURA 8** – Previsão de invasão de *Homalodisca coagulata* na América do Sul. A - previsão de distribuição nativa da espécie no sudoeste dos Estados Unidos. B - projeção do nicho ecológico potencial de ocupação da espécie na América do Sul (modificado de Peterson *et. al.*, 2003).

#### Uso do GARP para prever algumas conseqüências de alterações climáticas futuras na distribuição geográfica de espécies

A Figura 7 mostra as conseqüências de alterações climáticas para três espécies arbóreas de cerrado, baseadas em dois cenários futuros de alteração climática (Siqueira & Peterson, 2003). Em vermelho, temos a área prevista de distribuição da espécie no presente. Em azul, temos a área habitável em 2055, baseada no cenário HHGSDX50 (cenário mais otimista). Em rosa temos a área habitável em 2055, baseada no cenário HHGGAX50 (cenário mais pessimista).

Os modelos de circulação usados foram provenientes do HadCM2 (Carson, 1999). Foi utilizado um cenário mais otimista e outro mais pessimista em relação ao modo como o clima poderia se alterar nos próximos 50 anos, usando cenários Hadley HHGSDX50 e HHGGAX50, respectivamente. O cenário HHGSDX50 baseia-se em um aumento de 0,5%/ano de CO<sub>2</sub> e o cenário HHGGAX50 baseia-se em um aumento de 1%/ano de CO<sub>2</sub> na atmosfera, prevendo um aquecimento global e redução de precipitação em várias regiões. Como um exemplo sobre o que representam essas alterações climáticas futuras, foi feito um cálculo para o México, cuja conclusão foi que essas alterações ocasionariam um aumento de 1,6° a 2,5°C na temperatura média e uma diminuição de 70 a 130 mm na precipitação média (Peterson *et al.*, 2002d). Como resultado, podemos observar uma redução de área na distribuição geográfica, que varia de intensidade dependendo da espécie, para algumas espécies arbóreas de cerrado.

#### Uso do GARP para prever o potencial de invasão de espécies exóticas

A possibilidade de invasão de espécies exóticas em outras regiões é uma questão de relevância conservacionista e econômica. É possível utilizar o GARP para se tentar prever o comportamento de uma espécie potencialmente invasora do seguinte modo: modela-se o nicho ecológico de uma espécie na sua área de distribuição geográfica nativa e posteriormente projeta-se essa modelagem, através da projeção do nicho potencial de ocupação dessa espécie, em outras áreas de interesse. O exemplo a seguir (Peterson *et al.*, 2003) ilustra a possibilidade de invasão da América do Sul por *Homalodisca coagulata* (inseto vetor de transmissão da bactéria *Xylella fastidiosa*), dado o potencial de seu impacto como um vetor para agentes fitopatológicos. Foram desenvolvidos modelos de nicho ecológico para a espécie na sua área de distribuição geográfica nativa, no sudeste dos Estados Unidos (Figura 8a).

A capacidade de previsão dos modelos foi testada tanto na área de distribuição nativa quanto em projeções no estado da Califórnia, onde a espécie tem estado presente por muito tempo como uma espécie invasora. Em ambos os casos, os testes indicaram alta significância estatística para as previsões. Projeções dos mesmos modelos para a América do Sul indicaram poucas chances de invasão ao sudeste do Brasil, local em que doenças em culturas de cítricos são preocupantes. Entretanto, todos os modelos são concordantes ao prever alto risco de estabelecimento da espécie na região onde se encontram as vinícolas do norte da Argentina e no extremo sul do Brasil (Figura 8b).

## REFERÊNCIAS BIBLIOGRÁFICAS

- Anderson, R.P., M. Laverde & A.T. Peterson. 2002a. Geographical distributions of spiny pocket mice in South America: Insights from predictive models. *Global Ecology and Biogeography* 11: 131-141.
- Anderson, R.P., M. Laverde & A.T. Peterson. 2002b. Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos* 93: 3-16.
- Anderson, R.P., D. Lew & A.T. Peterson. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling* 162: 211-232.
- Carson, D.J. 1999. Climate modelling: Achievements and prospects. *Quarterly Journal of the Royal Meteorological Society* 125: 1-27.
- Darwin, C.R. 1859. *On the origin of species by means of natural selection*. John Murray, London.
- Davis, L. 1996. *Handbook of genetic algorithms*. Thomson Computer Press, Boston, Massachusetts.
- Feria, T.P. & A.T. Peterson, 2002. Using point occurrence data and inferential algorithms to predict local communities of birds. *Diversity and Distributions* 8: 49-56.
- Fogel, L.J., A.J. Owens & M.J. Walsh. 1966. *Artificial intelligence through simulated evolution*. Wiley, New York.
- Goldberg, D. 1989. *Genetic algorithms in search, optimization and machine learning*. Addison. Wesley Publishing Company, Inc., Reading, Massachusetts.
- Hutchinson, R.E. 1957. Concluding remarks. *Cold Spring Harbor. Symposia on Quantitative Biology*. 22: 415-427.
- Holland, J. 1975. *Adaptation in natural and artificial systems*. The University of Michigan Press, Ann Arbor, Michigan.
- Mitchell, M. 1996. *An introduction to genetic algorithms*. MIT Press, Cambridge, MA.
- Nix, H.A. 1986. A biogeographic analysis of Australian Elapid snakes. In: R. Longmore (ed). *Atlas of Australian Elapid Snakes*. Australian Flora and Fauna Series 8: 4-15.
- Obitko, M. 1998. *Genetic Algorithms*. Disponível em <http://www.professor.webizu.org/ga/> (acessado em agosto de 2007).
- Peterson, A.T. 2001. Predicting species' geographic distributions based on ecological niche modeling. *Condor* 103: 599-605.
- Peterson, A.T. & K.C. Cohoon. 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecological Modelling* 117: 159-164.
- Peterson, A.T. & D. A. Vieglais. 2001. Predicting species invasions using ecological niche modeling. *BioScience* 51:363-371.
- Peterson, A.T., J. Soberon & V. Sanchez-Cordero. 1999. Conservatism of ecological niches in evolutionary time. *Science* 285: 1265-1267.
- Peterson, A.T., L.G. Ball & K.C. Cohoon. 2002a. Predicting distributions of tropical birds. *Ibis* 144: 27-32.
- Peterson, A.T., V. Sanchez-Cordero, C.B. Beard & J.M. Ramsey. 2002b. Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. *Emerging Infectious Diseases* 8: 662-667.
- Peterson, A.T., D.R.B. Stockwell & D.A. Kluza. 2002c. Distributional prediction based on ecological niche modeling of primary occurrence data. In: M. Scott (ed). *Predicting species occurrences: Issues of scale and accuracy*. pp. 617-623. Island Press, Washington, D.C.
- Peterson, A.T., M.A. Ortega-Huerta, J. Bartley, V. Sanchez-Cordero, J. Soberon, R.H. Buddemeier & D.R.B. Stockwell. 2002d. Future projections for Mexican faunas under global climate change scenarios. *Nature* 416: 626-629.
- Peterson, A.T., M. Papes & D.A. Kluza. 2003. Predicting the potential invasive distributions of four alien plant species in North America. *Weed Science*. 51: 863-868.
- Peterson, A.T., R. Scachetti-Pereira & D.A. Kluza. 2003. Assessment of invasive potential of the glassy-winged sharpshooter *Homalodisca coagulata* in California and South America. *Biota Neotropica* 3. Disponível em: <http://www.biotaneotropica.org.br/v3n1/pt/abstract?article+BN00703012003> (acessado em julho de 2007).
- Ratter, J.A., S. Bridgewater & J.F. Ribeiro. 2001. Espécies lenhosas da fitofisionomia cerrado sentido amplo em 170 localidades do bioma cerrado. *Boletim do Herbário Ezechias Paulo Heringer, Brasília* 7: 5-112.
- Rechenberg, I. 1965. *Cybernetic solution path of an experimental problem*. Ministry of Aviation, Royal Aircraft Establishment, United Kingdom.
- Siqueira, M.F. & A.T. Peterson. 2003. Consequences of global climate change for geographic distributions of cerrado tree species. *Biota Neotropica* 3: 1-14.
- Stockwell, D.R.B. 1999. Genetic algorithms II. In: A. H. Fielding (ed). *Machine learning methods for ecological applications*. pp. 123-144. Kluwer Academic Publishers, Boston.
- Stockwell, D.R.B. & I.R. Noble, 1992. Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. *Mathematics and Computers in Simulation* 33: 385-390.
- Stockwell, D.R.B. & A.T. Peterson, 2002a. Controlling bias in biodiversity data. In: J. M. Scott (ed). *Predicting species occurrences: Issues of scale and accuracy*. pp. 537-546. Island Press, Washington, D.C.
- Stockwell, D.R.B. & A.T. Peterson, 2002b. Effects of sample size on accuracy of species distribution models. *Ecological Modelling* 148: 1-13.
- Stockwell, D.R.B. & D.P. Peters, 1999. The GARP modelling system: Problems and solutions to automated spatial prediction. *International Journal of Geographic Information Systems* 13: 143-158.

---

# Modelos de metapopulação

DOUGLAS F. M. GHERARDI

Instituto Nacional de Pesquisas Espaciais – INPE, Divisão de Sensoriamento Remoto, São Paulo, Brasil.  
e-mail: douglas@ltd.inpe.br

## RESUMO

O conceito de metapopulação é definido como uma rede formada por populações espacialmente estruturadas em agrupamentos, cujos indivíduos se reproduzem localmente e onde a migração entre as populações pode influenciar a dinâmica local. São apresentados o modelo clássico de Levins (1969), seguido pelo modelo de metapopulação espacialmente implícito de Roughgarden (1997) e pelo modelo de metapopulação espacialmente estruturado de Ovaskainen & Hanski (2001).

## ABSTRACT

*Metapopulation is defined as a network of spatially structured local populations that reproduce locally and where migration of individuals influence population dynamics. Three different models are presented: the classic Levins model (1969), the spatially implicit metapopulation model proposed by Roughgarden (1997), and the spatially structured metapopulation model proposed by Ovaskainen & Hanski (2001).*

## INTRODUÇÃO

O conceito de metapopulação tem feito parte de publicações especializadas desde o final da década de 80, muito embora as idéias centrais ao conceito já fossem apresentadas na forma de biogeografia de ilhas. Uma metapopulação pode ser grosseiramente definida como uma rede formada por populações espacialmente estruturadas em agrupamentos, cujos indivíduos se reproduzem localmente e onde a migração entre as populações pode influenciar a dinâmica local, incluindo a possibilidade de uma população se restabelecer após uma extinção (Hanski & Simberloff, 1997).

No centro do conceito de metapopulação está a idéia de que o espaço é discreto e composto por manchas

de habitat apropriadas ao desenvolvimento de uma espécie alvo. O espaço restante é geralmente chamado de matriz. As manchas de habitat devem ser grandes o suficiente para acomodar populações locais vivendo em pan-mixia, porém, não muito grandes. Estas premissas estão de acordo com o conceito original desenvolvido por Levins (1969), sendo que para uma metapopulação ideal devemos acrescentar mais três características: 1) as manchas devem ter áreas e isolamento iguais; 2) as populações locais devem ter dinâmica totalmente independente, ou não-correlacionadas; e 3) a taxa de troca de indivíduos entre as populações locais deve ser tão baixa a ponto de não influir na dinâmica local, isto é, a dinâmica local ocorre em uma escala de tempo mais rápida do que a dinâmica da metapopulação.

A prática tem demonstrado que não é possível encontrar metapopulações ideais na vida real, mas os conceitos apresentados acima podem ser aplicados com sucesso, afrouxando-se um pouco aquelas premissas. O que de fato importa para que o conceito possa ser aplicado de modo a gerar resultados cientificamente interpretáveis, é a noção de populações discretas que se reproduzem e são conectadas por migração. De qualquer forma, alguns cuidados devem ser tomados antes que se decida empregar essa teoria, caso contrário alguma outra abordagem deve ser utilizada que não a de metapopulação.

### Condições para a aplicação dos conceitos de metapopulação

Mais importante do que tentar classificar as populações vivendo em ambientes fragmentados, é encontrar maneiras de compreender e prever suas dinâmicas. Isto porque as populações exibem variações em suas estruturas espaciais de uma maneira contínua. Por isso, sugere-se que o sistema seja avaliado de forma a satisfazer quatro condições indicativas de uma aplicação informativa dos conceitos de metapopulação (Hanski, 1997): (1) o hábitat considerado adequado deve ocorrer em manchas discretas e ser ocupado por populações locais que se reproduzam; (2) mesmo a maior população deve ter um risco de extinção significativo, caso contrário a metapopulação persistiria unicamente por causa desta população, se transformando em uma metapopulação do tipo continente-ilha; (3) as manchas de hábitat não devem estar muito isoladas, de forma a impedir a recolonização. Isto levaria uma metapopulação fora do equilíbrio, se encaminhando para uma extinção global; e (4) as populações locais não podem ter dinâmicas completamente sincronizadas. Isto levaria a metapopulação à extinção no espaço de tempo daquela população local com a menor probabilidade de extinção.

## MODELOS DE METAPOPLAÇÃO

A modelagem de populações, que assume o espaço como entidade discreta, pode ser dividida em duas classes, desde que o modelo trate de duas populações da mesma espécie conectadas por migração, ou da interação de várias populações locais (Hanski & Simberloff, 1997). A primeira classe se aplica a estudos cujo interesse se concentra no efeito da migração sobre a dinâmica local, com ausência de extinção local. A segunda classe, mais indicada para o estudo de metapopulações, se aplica ao estudo de várias manchas de hábitat e populações locais, podendo ocorrer uma renovação das

populações. Dentro desta segunda classe, pode-se distinguir as abordagens espacialmente implícitas, as espacialmente explícitas e as espacialmente realistas. A seguir serão delineadas estas abordagens de maneira resumida para que, em seguida, se apresente algumas formulações específicas.

### Abordagens espacialmente implícitas

Encarar a modelagem de populações, a partir de uma simplificação crítica de um problema naturalmente complexo, pode permitir que se intuem aspectos importantes do problema atacado. Assim, em vez de se tentar modelar a dinâmica de várias populações conectadas por migração, pode-se modelar a mudança no número de manchas ocupadas, sem se preocupar com o que acontece em cada mancha, ou com a sua localização. Esta é a essência do modelo proposto por Levins (1969), que não considera a posição de cada mancha e assume o espaço de maneira implícita. Será visto mais adiante que esse tipo de abordagem pode revelar aspectos esclarecedores sobre a ecologia de populações. A vantagem desses modelos é que facilitam a análise conceitual e matemática, e a desvantagem é que só podem investigar uma parcela reduzida de aspectos relevantes à dinâmica de metapopulações.

As restrições impostas por modelos espacialmente implícitos levaram pesquisadores a levantar questões importantes. Em particular, sobre quais seriam as consequências do posicionamento realista das populações locais. Poderia uma taxa de migração ser alta o suficiente para resgatar populações locais da extinção? Se os tamanhos das manchas apresentarem diferenças muito grandes, isto poderia influenciar nas probabilidades de extinção local? De que maneira? Estas questões, que demandam a incorporação de informações sobre tamanho e posição relativa das manchas, têm sido investigadas com o uso de modelos espacialmente explícitos.

### Abordagens espacialmente explícitas

A incorporação explícita do espaço aos modelos produziu algumas estruturas distintas de modelagem, como o autômato celular (Caswell & Etter, 1993) e os modelos baseados em mapas de grades acopladas (Hassel *et al.*, 1991). Nestes casos, assume-se que as populações locais (ou indivíduos) estão organizados em células dispostas em uma grade regular e seus tamanhos são modelados como variáveis discretas ou contínuas.

Esta categoria de modelos permite que se defina um tamanho e espaçamentos regulares das células, facilitando a construção de regras matemáticas mais simples que governam o comportamento local. É este comportamento local que separa as abordagens implícitas

tas e explícitas, uma vez que aqui as populações interagem apenas com aquelas de células vizinhas. Esta condição gera conseqüências dinâmicas interessantes, como a necessidade de um intervalo de tempo longo para que uma metapopulação alcance um estado estacionário, e a ocorrência de dinâmica espacial caótica.

A principal desvantagem é computacional, uma vez que se torna necessário o emprego de todo o vetor de presenças e ausências, ao invés da simples proporção de células ocupadas. Entretanto, os modelos baseados em grades e aqueles baseados em descrições do ambiente em formato raster dentro de um sistema de informações geográficas (*cf.* Akçakaya, 1994), como os utilizados em ecologia de paisagem, guardam muitas semelhanças entre si. Há um vasto campo de investigação a se explorar, buscando acoplar as características da paisagem às regras dinâmicas que regem populações vivendo em agrupamentos de “células” localizadas em um ambiente matricial (*raster*).

### Abordagens espacialmente realistas

Os modelos realistas estão preocupados em associar a dinâmica de metapopulações reais a partir da incorporação das características geométricas das manchas pertencentes a uma rede. Neste ponto, os modelos buscam em geral fazer previsões sobre a dinâmica das manchas frente às alterações de forma, número e localização, dependendo fortemente de dados coletados *in situ*. Talvez o modelo mais simples desta categoria seja aquele baseado na função de incidência (IF model, Hanski 1994, 1997), que se assemelha a uma cadeia de Markov simples. Ele é derivado do modelo de Levins, com a diferença de não ser determinístico, mas estocástico. Assim, as manchas de hábitat podem variar em área e possuir posições específicas. Isto permite que se associem diferentes probabilidades de extinção local a diferenças de área das manchas, e que diferenças nas probabilidades de recolonização sejam associadas às distâncias entre as manchas.

## OS MODELOS

Serão apresentadas as linhas gerais de três modelos de metapopulação, seguindo uma sequência cronológica e estrutural de desenvolvimento. O primeiro a ser apresentado é o modelo clássico de Levins (1969), que ao longo dos anos foi expandido para investigar importantes processos ecológicos, como dispersão (Ruxton *et al.*, 1997; Parvinen, 1999, Silva *et al.*, 2001), competi-

ção (Garret & Dixon, 1997; Lehman & Tillman, 1997; Buttel *et al.*, 2002), estabilidade (Neuhauser, 2002), conservação (Man *et al.*, 1995), destruição de hábitat e extinção (Tilman & Lehman, 1997). Em seguida será apresentado o modelo de metapopulação espacialmente implícito para a função de produção de invertebrados sésseis proposto por Roughgarden (1997). Este modelo será ilustrado por um experimento numérico simulando um ecossistema recifal costeiro brasileiro. Finalmente, será apresentado o modelo de metapopulação espacialmente estruturado proposto por Ovaskainen & Hanski (2001), que propõe algumas medidas da capacidade de uma rede de manchas de hábitat suportar uma metapopulação viável.

Antes, entretanto, cabe uma nota de esclarecimento quanto ao propósito e validade do modelo ecológico simples, em geral mal compreendido (Hanski, 1997). Seu propósito não é o de recriar tantos detalhes das populações reais quanto possível. Seu propósito é o de isolar algum aspecto de interesse, para um estudo teórico. As diferenças entre a vida real e um modelo simplificado podem não ter maiores conseqüências sobre a sua capacidade de incorporar variáveis críticas e processos que afetem o fenômeno em estudo.

### O modelo de Levins (1969)

Este modelo assume a existência de um grande número de manchas de hábitat discretas, preferencialmente de mesmo tamanho, e todas conectadas umas às outras via migração. O fato de na vida real haverem restrições à migração de indivíduos não compromete os resultados do modelo quando em estado estacionário, a menos que a heterogeneidade seja muito acentuada. As condições ideais de aplicação do modelo são aquelas em que a dinâmica da população local ocorre em uma escala de tempo menor do que a da metapopulação. Isto pode ocorrer se as manchas de hábitat forem menores e as populações locais alcançarem a capacidade de suporte rapidamente, ou se a taxa de colonização for lenta. As populações devem apresentar a mesma probabilidade de extinção, a taxa de colonização  $c$  deve ser proporcional ao número de manchas ocupadas  $p$  (fonte de colonizadores), dada por  $cp$ , e à fração de manchas não ocupadas,  $1-p$  (alvo da colonização). A taxa de extinção local deve ser proporcional ao número de manchas ocupadas, dada por  $ep$ , obtendo-se, assim, a dinâmica de ocupação no tempo contínuo,

$$\frac{dp}{dt} = cp(1-p) - ep \quad \text{Equação 1}$$

onde  $c$  é a taxa de colonização e  $e$  é a taxa de extinção. Quando  $c > e$ , a proporção de manchas ocupadas aproxima-se do equilíbrio,  $dp/dt = 0$ , obtendo-se:

$$\hat{p} = 1 - \frac{e}{c} \quad \text{Equação 2}$$

Este equilíbrio tem estabilidade global, ou seja,  $p$  irá se aproximar de  $\hat{p}$  para qualquer densidade inicial e qualquer perturbação, desde que  $p > 0$ . O modelo de ocupação apresenta algumas características importantes (Tilman *et al.*, 1997). O crescimento populacional tem o comportamento de uma curva logística, sendo que uma população persiste em um hábitat se  $c > e$ . Um dos aspectos mais interessantes deste modelo é que nenhuma espécie é capaz de preencher completamente seu hábitat em equilíbrio. É possível demonstrar facilmente que, sendo  $\hat{s}$  a proporção de todas as manchas de hábitat viáveis não ocupadas por uma única espécie em equilíbrio, então,

$$\hat{s} = 1 - \hat{p} = \frac{e}{c} \quad \text{Equação 3}$$

O resultado inevitável de se viver tendo o espaço como uma variável importante é que uma proporção das manchas sempre estará vazia, uma vez que, em termos ecológicos,  $e$  não pode ser zero. Quanto maior a taxa de mortalidade de uma espécie relativamente à sua taxa de colonização, maior será a proporção de espaço não ocupado.

Para que a metapopulação persista, é necessário que a recolonização ocorra a uma taxa suficientemente alta para compensar as extinções e para permitir que pequenas metapopulações se desenvolvam. Em outras palavras,  $e/c < 1$  ou  $c/e > 1$ , isto é, uma população cercada de manchas de hábitat não ocupadas deve ser capaz de estabelecer pelo menos uma nova população ( $1/e$ ) durante sua existência.

A simplicidade matemática e a facilidade de se manipular analiticamente este modelo vêm do fato da dispersão ser tratada considerando-se que todos os propágulos estão dispersos aleatoriamente ao longo de todo o hábitat. Isto elimina os efeitos da dispersão local sobre a dinâmica da metapopulação, causada pela heterogeneidade espacial (mas veja Doebeli, 1995; Ruxton *et al.*, 1997).

### Modelo de metapopulação espacialmente implícito (Roughgarden, 1997)

O modelo de metapopulação desenvolvido por Roughgarden (1997) tem sua formulação baseada em uma função de produção biológica de invertebrados limita-

dos espacialmente, como é o caso dos invertebrados bentônicos. Ele é um modelo de metapopulação de uma espécie que apresenta um ciclo de vida consistindo de ovo, larva pelágica e adulto de vida bentônica. Portanto, é uma formulação adequada para a modelagem da produção natural de populações de corais e da influência de diferentes parâmetros na performance da população.

A formulação assume que as larvas estão contidas em um reservatório de larvas não estruturado e que os adultos residem em um substrato bentônico. As larvas deverão recrutar sobre um substrato não ocupado. Assume-se, também, que o reservatório de larvas se mistura rapidamente no sentido da costa para o mar aberto (note a semelhança com o modelo anterior, do tipo Levins). O espaço livre na formulação de metapopulação é definido como a área não ocupada por organismos,

$$F = A - aB \quad \text{Equação 4}$$

onde  $A$  é a área do hábitat bentônico,  $a$  é a área basal de um indivíduo e  $B$  é o número de organismos adultos. O assentamento de larvas (recrutamento) sobre o espaço não ocupado é dado pelo termo  $cLF$ , incluído na dinâmica do ponto de assentamento em questão. Assim, a dinâmica no hábitat bentônico é:

$$\frac{dB}{dt} = cLF - \mu B \quad \text{Equação 5}$$

onde  $c$  é o coeficiente de assentamento de larvas,  $L$  é o número de larvas no reservatório de larvas, e  $\mu$  é a taxa de mortalidade no hábitat bentônico. Assume-se que o recrutamento seja proporcional à quantidade de espaço vazio e ao número de larvas. A dinâmica do reservatório de larvas responde à reprodução dos indivíduos adultos, à perda por mortalidade larval e ao recrutamento. Para se obter a função de produção para os adultos, deve-se assumir que a população de larvas alcança equilíbrio relativamente ao número de indivíduos adultos (o estoque). Isto leva à dinâmica de indivíduos adultos expressa como:

$$\frac{dB}{dt} = \left[ \frac{m}{1 + \nu/c (A - aB)} - \mu \right] B \quad \text{Equação 6}$$

onde  $m$  é a taxa de fecundidade dos adultos e  $\nu$  a taxa de mortalidade do reservatório de larvas.

As equações para  $r_0$ , taxa intrínseca de crescimento, e  $r$ , o crescimento quando a população alcança a capacidade de suporte ( $K$ ) são:

$$r_0 = \frac{cA(m - \mu) - \mu v}{cA + v} \quad \text{Equação 7}$$

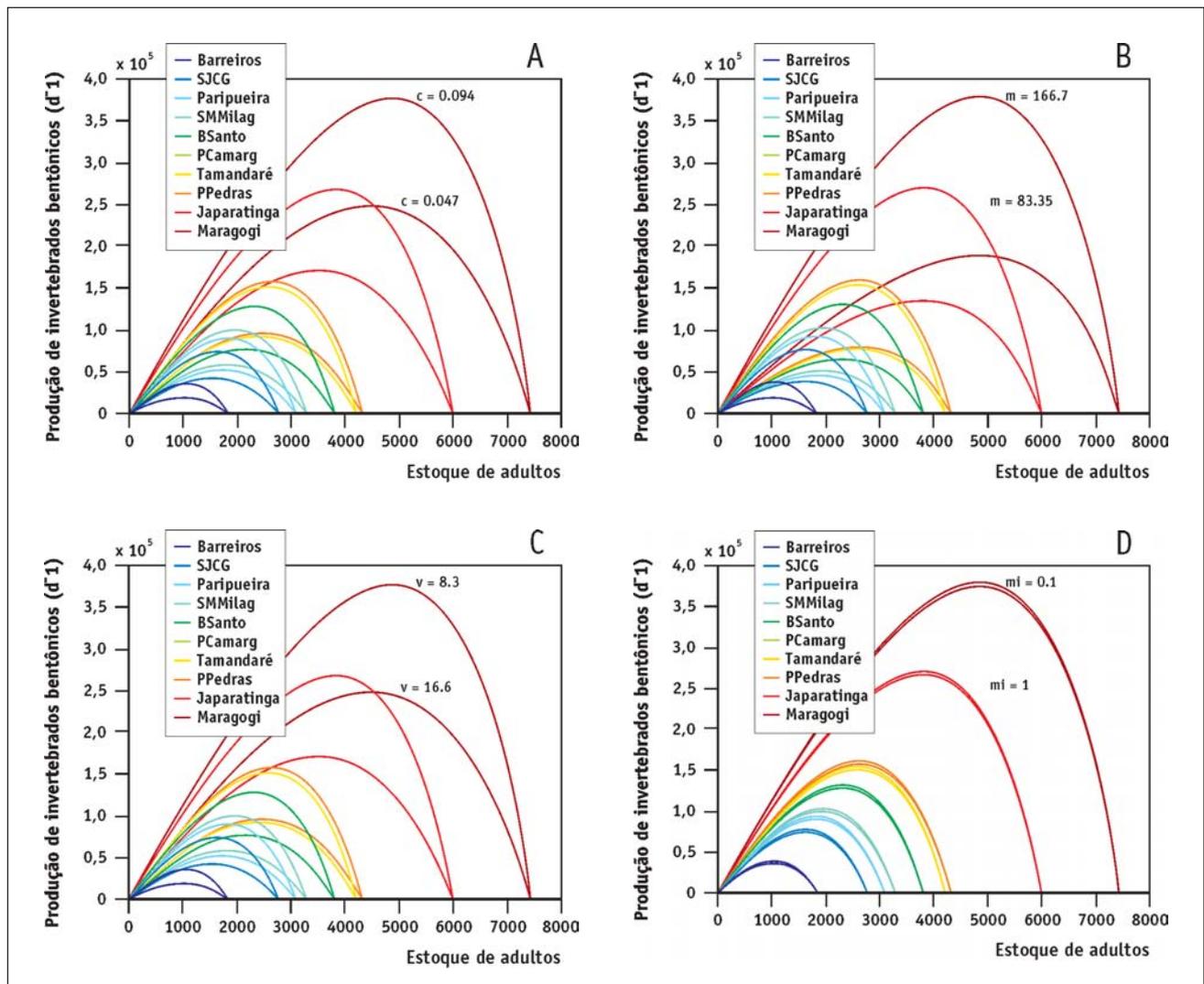
$$r = \frac{(m - \mu)[cA(m - \mu) - \mu v]}{mv} \quad \text{Equação 8}$$

**EXEMPLO DE APLICAÇÃO**

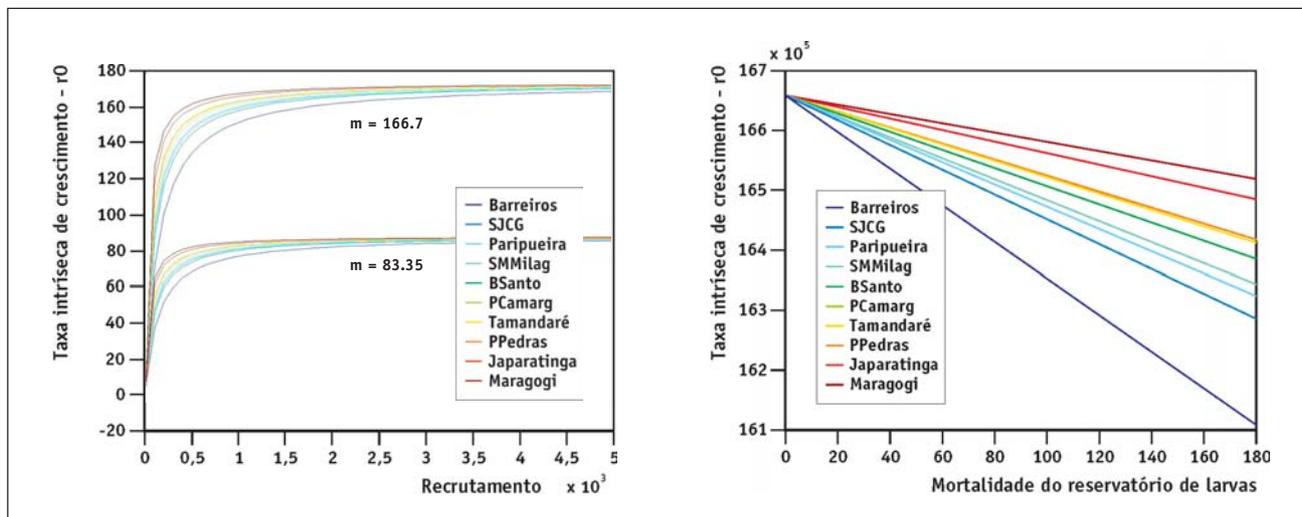
A seguir são apresentados alguns resultados de experimentos numéricos, usando como exemplo os recifes mapeados da área de proteção ambiental (APA) Costa dos Corais, localizada entre os estados de Alagoas e Pernambuco. Apenas alguns dos parâmetros empregados pelo modelo podem ser encontrados na literatura

que trata dos ambientes recifais brasileiros (calculados com base em Maida & Ferreira, 1995). Para suprir esta falta, serão utilizados dados obtidos na literatura internacional, coletados por pesquisadores trabalhando em outros recifes. Como exemplos destes parâmetros podemos citar a fecundidade dos corais (*m*) e a taxa de mortalidade no reservatório de larvas (*v*).

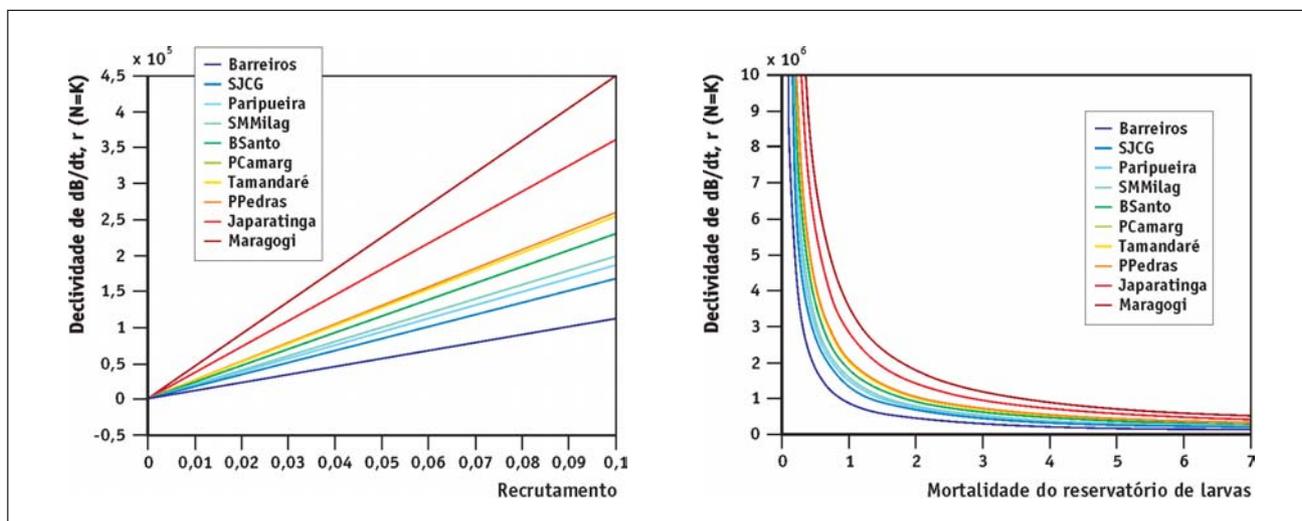
Resultados dos experimentos numéricos executados mostram o comportamento da produção de indivíduos de uma espécie hipotética de coral para recifes agrupados em dez municípios (Figuras 1, 2 e 3). Um cenário padrão é comparado a outro, onde fecundidade e recrutamento foram reduzidos em 50%, e a mortalidade de adultos e larvas foi aumentado na mesma proporção.



**FIGURA 1** – Comportamento da função de produção comparando os cálculos feitos com os dados da literatura contra um cenário ambiental negativo. Cores quentes correspondem a áreas maiores de recife (o significado de cada parâmetro pode ser encontrado no texto). *m*=166,7; *c*=0,094; *n*=8,3; *a*=0,03; *m*=0,1.



**FIGURA 2** – Resultados dos experimentos numéricos usando a taxa intrínseca de crescimento. Note a diferença nos limites de estabilização para  $r_0$  quando a fecundidade é reduzida (gráfico esquerdo);  $r_0$  mostra uma tendência linear negativa em relação à taxa de mortalidade das larvas (gráfico direito).



**FIGURA 3** – Contrastando com os resultados mostrados na Figura 2,  $r$  aumenta linearmente com o aumento no recrutamento e cai exponencialmente com pequenos incrementos na mortalidade larval.

As mudanças mais marcantes na produção biológica foram observadas quando o recrutamento (Figura 1A) e a fecundidade (Figura 1B) foram reduzidos e a mortalidade de larvas aumentada (Figura 1C). A morte de adultos parece causar um impacto relativamente menor sobre a produção de indivíduos (Figura 1D). De todos os parâmetros, o recrutamento parece ser o mais crítico, e quanto maior for a área de hábitat bentônico, maior será a perda absoluta da produção natural. Assim, o modelo sugere que, nos esforços de conservação, deva ser dada atenção especial ao assentamento de larvas, juntamente com aqueles fatores que possam contribuir para a redução da acessibilidade do substrato às larvas.

### Modelo de metapopulação espacialmente estruturado: capacidade de metapopulação (Ovaskainen & Hanski, 2001)

O modelo espacialmente realista (estruturado) é capaz de incorporar, de maneira explícita, as áreas e a localização das manchas de hábitat, sendo isto tão mais importante quanto mais fragmentado for o hábitat. Esta seção se limitará a apresentar a versão baseada no modelo determinístico de Levins, onde o tempo é uma variável contínua. Esta escolha foi feita porque a formulação apresentada por Ovaskainen & Hanski (2001) para o tempo contínuo (modelo IF) exige que as estimativas de parâmetros relativos a cada mancha sejam mais arbitrárias.

A partir da versão determinística deste modelo, são derivadas medidas locais que dão a importância das manchas de hábitat, individualmente, para a persistência da metapopulação. Essas medidas podem ser estendidas para se ponderar a respeito da influência que a adição de novas manchas, em locais específicos, pode exercer sobre a persistência da metapopulação.

O ponto central da análise é  $p_i(t)$ , a probabilidade da mancha  $i$  estar ocupada no tempo  $t$ . Sendo então  $p(t) = \{x \in \mathbb{R}^n \mid 0 \leq x_i \leq 1 \forall i\}$  o vetor com componentes  $p_i(t)$ , se buscará modelar a mudança de  $p(t)$  no tempo contínuo (ou discreto) como

$$p_i(t) = C_i(p(t)) (1-p_i(t)) - E_i(p(t)) p_i(t) \quad \text{Equação 9}$$

Neste caso,  $C_i(p(t))$  dá a probabilidade de colonização da mancha  $i$  quando ela está desocupada e  $E_i(p(t))$  a probabilidade de extinção quando a mancha  $i$  está ocupada. Para incorporar os efeitos da área da mancha de hábitat sobre a extinção e da conectividade sobre a colonização, Ovaskainen & Hanski (2001) modificaram a Equação 9,

$$\frac{dp_i(t)}{dt} = C_i(p(t))(1-p_i(t)) - E_i p_i(t), \quad \text{Equação 10}$$

onde  $C_i(p(t))$  dá as taxas de colonização das manchas desocupadas e  $E_i$  as taxas de extinção das populações restantes, agora independente do vetor  $p(t)$ . Pode-se, então, assumir que  $E_i = e/A_i$ , uma vez que manchas maiores ( $A$ ) tendem a possuir populações maiores com menor risco de extinção ( $e$ ). A contribuição das populações restantes para a conectividade da mancha  $i$  pode ser relacionada às áreas das manchas e suas distâncias da mancha  $i$  na forma

$$S_i(p(t)) = \sum_{j \neq i} A_j e^{-\alpha d_{ij}} p_j(t), \quad \text{Equação 11}$$

onde  $1/\alpha$  é a distância média de migração,  $d_{ij}$  é a distância entre as manchas  $i$  e  $j$ , e  $p_j$  a probabilidade de  $j$  estar ocupada em  $t$ . O raciocínio associado à equação 11 é que a taxa de emigração da mancha  $j$ , quando ocupada, é proporcional à área da mancha, e a contribuição da mancha  $j$  para a imigração em direção à mancha  $i$  decresce exponencialmente com a distância  $d_{ij}$ . A taxa de colonização é então calculada por  $C_i(p(t)) = cS_i(p(t))$ . É importante observar que a notação matricial é  $g_i(p) = G((Mp)_i)$ , onde  $M$  é uma matriz irredutível e não negativa composta dos elementos  $m_{ij} = A_i A_j e^{-\alpha d_{ij}}$  para  $i \neq j$  e  $m_{ii} = 0$ . O importante aqui é

notar que  $M$  incorpora a influência da extinção local em  $i$ , com o termo  $A_i$ , e a influência de  $j$  na colonização de  $i$ , com o termo  $A_j e^{-\alpha d_{ij}}$ .

É possível agora definir a capacidade de persistência ( $\lambda_M$ ) e de invasão ( $\lambda_i$ ) da metapopulação. A medida  $\lambda_M$  dá a capacidade de suporte, ou o tamanho de uma população quando o sistema está em equilíbrio. Em outras palavras, é a possibilidade de uma metapopulação persistir no longo prazo, dada pela existência de um estado de equilíbrio estável não-trivial. A medida  $\lambda_i$  define a condição de invasão de uma rede de manchas de habitats vazios a partir de uma população local pequena quando  $\lambda_i > e/c$ . No modelo determinístico, o valor de  $\lambda_M = \lambda_i = \lambda_i(M)$  onde  $\lambda_i(M)$  é o primeiro autovalor da matriz  $M$ . Ovaskainen & Hanski (2001) também definiram uma medida para estimar a variação na importância de  $\lambda_M$  frente à remoção de uma mancha  $q$ , dada por:

$$V_q^r = \frac{A_q^2 [S_q(x)]^2}{\lambda_M^2} \quad \text{Equação 12}$$

onde  $A_q$  é a área de  $q$ ,  $S$  é a conectividade e  $x$  é o autovetor direito de  $M$ .

A formulação desenvolvida para este modelo espacialmente estruturado mostra que a principal variável de controle ambiental não é a quantidade de habitats, mas uma medida que integra os efeitos das áreas das manchas e suas conectividades na capacidade da rede de manchas de hábitat suportar uma metapopulação. Talvez a qualidade mais relevante deste modelo espacialmente estruturado é o seu potencial de aplicação na análise de paisagem para espécies ou grupo de espécies com disponibilidade de informação biológica limitada.

A unificação das vertentes de modelagem, quais sejam, a de ecologia de metapopulação (ou de baixa cobertura espacial) e a de ecologia de paisagem (ou de alta cobertura espacial) é a fronteira a ser ultrapassada para que se possa avançar na direção da efetiva gestão de mosaicos.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Akçakaya, H.R. 1994. RAMAS/GIS. Linking landscape data with population viability analysis. Applied Biomathematics, Setauket, New York.
- Buttel L.A., R. Durrett & S.A. Levin. 2002. Competition and species packing in patchy environments. Theoretical Population Biology 61: 265-276.
- Caswell, H. & R.J. Etter. 1993. Ecological interactions in patchy environments: from patch-occupancy models to cellular automata. In: S.A. Levin, T.M. Powell & J.H. Steele (eds). Patch Dynamics. pp 93-109. Springer Verlag, Berlin.

- Doebeli, M. 1995. Dispersal and dynamics. *Theoretical Population Biology* 47: 82-106.
- Garrett, K.A. & P.M. Dixon. 1997. Environmental pseudo-interaction: the effects of ignoring the scale of environmental heterogeneity in competition studies. *Theoretical Population Biology* 51: 37-48.
- Hanski, I.A. 1994. A practical model of metapopulation dynamics. *Journal of Animal Ecology* 63: 151-162.
- Hanski, I.A. 1997. Metapopulation dynamics: from concepts and observations to predictive models. In: I.A. Hanski & M.E. Gilpin (eds). *Metapopulation biology: Ecology, genetics, and evolution*. pp 69-91. Academic Press, San Diego, California.
- Hanski, I.A. & D. Simberloff. 1997. The metapopulation approach, its history, conceptual domain, and application to conservation. In: I.A. Hanski & M.E. Gilpin (eds). *Metapopulation biology: Ecology, genetics, and evolution*. pp 5-26. Academic Press, San Diego, California.
- Hassell, M.P., H.N. Comins & R.M. May. 1991. Spatial structure and chaos in insect population dynamics. *Nature* 353: 255-258.
- Lehman, C. L. & D. Tilman. 1997. Competition in spatial habitats. In: D. Tilman & P. Kareiva (eds). *Spatial Ecology: the role of space in population dynamics and interspecific interactions*. Princeton University Press, Princeton, N.J.
- Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America* 15: 237-240.
- Maida, M. & B.P. Ferreira. 1995. Estudo preliminar sobre o assentamento de corais em um recife na Baía de Tamandaré - PE. *Boletim Técnico e Científico do CEPENE/IBAMA* 3: 23-36.
- Man, A., R. Law & N.V.C. Polunin. 1995. Role of marine reserves in recruitment to reef fisheries: a metapopulation model. *Biological Conservation* 71: 197-204.
- Neuhauser, C. 2002. Effects of local interactions and local migration on stability. *Theoretical Population Biology* 62: 297-308.
- Ovaskainen, O. & I. Hanski. 2001. Spatially structured metapopulation models: global and local assessment of metapopulation capacity. *Theoretical Population Biology* 60: 281-302.
- Parvinen, K. 1999. Evolution of migration in a metapopulation. *Bulletin of Mathematical Biology* 61: 531-550.
- Roughgarden, J. 1997. Production function from ecological populations: a survey with emphasis on spatially implicit models. In: D. Tilman & P. Kareiva (eds). *Spatial Ecology*. pp 296-317. *Monographs in Population Biology* 30, Princeton University Press, Princeton.
- Ruxton, G.D., J.L. Gonzalez-Andujar & J.N. Perry. 1997. Mortality during dispersal and the stability of a metapopulation. *Journal of Theoretical Biology* 186: 389-396.
- Silva, J.A.L., M.L. Castro & D.A.R. Justo. 2001. Stability in a metapopulation model with density-dependent dispersal. *Bulletin of Mathematical Biology* 63: 485-505.
- Tilman, D. & C.L. Lehman. 1997. Habitat destruction and species extinctions. In: D. Tilman & P. Kareiva (eds). *Spatial Ecology*. pp 233-249. *Monographs in Population Biology* 30, Princeton University Press, Princeton.
- Tilman, D., C.L. Lehman & P. Kareiva. 1997. Population dynamics in spatial habitats. In: D. Tilman & P. Kareiva (eds). *Spatial Ecology*. pp. 3-20. *Monographs in Population Biology* 30, Princeton University Press, Princeton.

---

# O uso de modelos em ecologia de paisagens

JEAN PAUL METZGER<sup>1\*</sup>  
MARINA ANTONGIOVANNI DA FONSECA<sup>2</sup>  
FRANCISCO JOSÉ BARBOSA DE OLIVEIRA FILHO<sup>3</sup>  
ALEXANDRE CAMARGO MARTENSEN<sup>1</sup>

<sup>1</sup> Universidade de São Paulo – USP, Instituto de Biociências, São Paulo, Brasil.

<sup>2</sup> Instituto Socioambiental – ISA, Amazonas, Brasil.

<sup>3</sup> Instituto Brasileiro do Meio Ambiente – IBAMA, Direção de Proteção Ambiental, Coordenação de Zoneamento Ambiental, Distrito Federal, Brasil.

\* e-mail: [jpm@ib.usp.br](mailto:jpm@ib.usp.br)

## RESUMO

Por lidar com unidades complexas e muitas vezes espacialmente extensas, a Ecologia de Paisagens utiliza usualmente representações abstratas e simplificadas das paisagens reais, *i.e.* modelos. Este artigo faz uma análise de alguns destes modelos e apresenta três estudos de caso, onde diferentes modelos são utilizados para: i) entender os efeitos da fragmentação sobre comunidades biológicas; ii) simular a percepção da paisagem por espécies de aves; e iii) inferir sobre os riscos de extinção relacionados a um processo de desmatamento. A Ecologia de Paisagens utiliza muitos modelos desenvolvidos em outras disciplinas, e os adapta para o estudo das relações entre padrões espaciais e processos ecológicos. Todos estes modelos caracterizam-se por considerarem a heterogeneidade do espaço. O principal desafio é encontrar um equilíbrio entre a complexidade destes modelos, que tende a ser cada vez maior, e a capacidade deles permitirem generalizações, contribuindo assim para o fortalecimento de uma teoria em Ecologia de Paisagens.

## ABSTRACT

*Since Landscape Ecology works with complex and usually spatially large areas, it often makes use of theoretical and simplified representations of real landscapes, i.e. models. This article makes a set of analyses of some commonly used models, and presents three case studies where different models are utilized for: i) understanding the effects of fragmentation in the biological communities; ii) simulate landscape perception for bird species; and iii) make inferences about extinction risks related to deforestation processes. Landscape Ecology uses models that were developed in different disciplines, and adapt them to analyze the relation between the spatial patterns and the ecological processes. All these models present as a common feature different ways to represent spatial heterogeneity. The main challenge is to reach the fine balance between adding complexity to those models, which tends to be bigger and bigger, and their capacity to allow generalizations, thus contributing to the enforcement of a Landscape Ecology theory.*

## MODELOS E ECOLOGIA DE PAISAGENS

### Ecologia de Paisagens

A paisagem pode ser definida como “um mosaico heterogêneo formado por unidades interativas, sendo esta heterogeneidade existente para pelo menos um fator, segundo um observador e numa determinada escala de observação” (Metzger, 2001). Numa visão antropocêntrica, a paisagem é vista como um conjunto de ecossistemas ou unidades de uso e cobertura do território, definidos por critérios fisionômicos, composicionais ou de utilidade para uso humano. Numa visão ecológica, a paisagem pode ser vista como um mosaico de diferentes tipos de habitats, submetidos a diferentes regimes de perturbação (naturais ou antrópicos), e de menor ou maior qualidade para uma determinada espécie. Fica claro que, dependendo do observador, as definições dos limites e dos constituintes da paisagem vão se modificar. Em particular, podemos esperar que espécies de menor ou maior capacidade de dispersão utilizem espaços mais ou menos restritos, respectivamente, ou então que a especificidade do habitat ou o requerimento por diferentes características abióticas modifique a definição das manchas.

Numa abordagem “ecológica”, a Ecologia de Paisagens vem sendo considerada como uma ecologia espacial, interessada em analisar a influência do arranjo espacial de suas unidades nos processos ecológicos (Turner, 1989). Muito frequentemente, a Ecologia de Paisagens tem sido considerada uma ecologia de macro-escalas, interessada nas interações de padrões e processos ecológicos em mosaicos de vários quilômetros quadrados (Forman, 1995). Contudo, em alguns casos, as perguntas da Ecologia de Paisagens podem ser desenvolvidas em espaços de apenas algumas dezenas de metros quadrados, na medida em que a heterogeneidade espacial é percebida nesta escala por alguns organismos, como a micro-fauna de musgos (Gonzales *et al.*, 1998) ou artrópodes em agroecossistemas (With *et al.*, 1999).

Como em outros ramos da ciência, a obtenção de dados em Ecologia de Paisagens pode ocorrer de diversas formas (Metzger, 2003a). A experimentação, por meio da manipulação da paisagem, e da comparação do estado do sistema antes e depois da intervenção, é sem dúvida umas das formas mais eficientes para se estabelecer relações causais (Debinski & Holt, 2000, McGarigal & Cushman, 2002). No entanto, são raras as ocasiões em que esse tipo de manipulação é possível e eticamente correto quando se trabalha com paisagens, em particular quando esta se estende por vários quilômetros quadrados (Tabela 1).

Para superar esta limitação, outras estratégias têm sido procuradas, como a comparação de processos ecológicos em paisagens que apresentam diferentes padrões espaciais. A principal vantagem destes “experimentos observacionais” (*sensu* McGarigal & Cushman, 2002) é que eles permitem considerar as paisagens em toda a sua complexidade, ganhando, assim, grande poder de generalização (Tabela 1). Por outro lado, esta vantagem é também uma desvantagem, na medida em que a complexidade dos fatores envolvidos nas relações entre padrões espaciais e processos ecológicos não permite categoricamente estabelecer relações causais simples, como geralmente se espera de um procedimento cartesiano de pesquisa. Diversos outros fatores, além do padrão espacial considerado na escolha das paisagens, podem interferir nas relações, em particular o histórico de perturbação e a heterogeneidade do micro-habitat, o que também torna quase impossível encontrar réplicas e controles verdadeiros para estudos comparativos (Tabela 1).

Um outro método utilizado para superar as limitações da manipulação de amplas paisagens são os experimentos em micro-paisagens. Por exemplo, o efeito de corredores na distribuição e abundância de animais associados a musgos foi estudado em áreas restritas de rochas recobertas por musgos (Gonzales *et al.*, 1998). Estas eram fragmentadas, raspando-se as rochas e

**TABELA 1** – Comparação de alguns atributos de experimentação em escalas pontuais e globais (adaptado de Bissonette, 1997).

ATRIBUTOS	ESCALAS PONTUAIS	MACRO-ESCALAS ( <i>e.g.</i> paisagem)
Manipulação experimental	possível	difícil
Réplicas e controles	possível	difícil
Rigor	alto	baixo
Testabilidade de hipóteses	alta	baixa
Problemas devido à falhas na amostragem	graves	menos graves
Generalizações	baixas	altas

mantendo-se “fragmentos de musgo” isolados, em alguns casos, ou unidos por corredores, em outros casos. Estes experimentos permitem um maior controle das variáveis, o uso de réplicas e controles verdadeiros, logo, um conhecimento mais preciso de relações causais. O grande problema deste tipo de experimentação é o “efeito de transmutação”, *i.e.*, a mudança nas propriedades em função da escala ou do nível hierárquico (Bissonette, 1997). Em outras palavras, não é possível assegurar que as respostas observadas para a fauna de musgos em relação à presença ou ausência de corredores possa ser aplicada para outros organismos que percebem a heterogeneidade (a paisagem) em outra escala espaço-temporal.

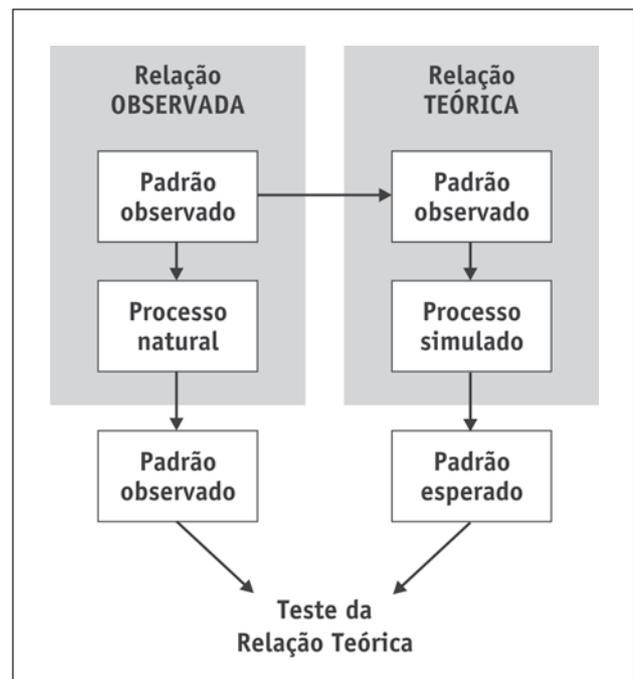
Outra alternativa seria o monitoramento de uma paisagem ao longo de um processo de perturbação, seja ele natural (*e.g.*, propagação de fogo, tornados, doenças) ou antrópico (*e.g.*, criação de estradas, canalizações de rios, grandes projetos madeireiros). Desta forma, a possibilidade de se fazer observações antes e depois destas alterações na estrutura da paisagem pode permitir inferir como estas estão alterando um determinado processo ecológico. Apesar da limitada possibilidade do pesquisador controlar este tipo de “experimento”, tanto as perturbações naturais como as grandes intervenções humanas na paisagem geram oportunidades singulares de experimentação em macro-escalas e deveriam ser utilizadas mais frequentemente pelos ecólogos.

### O uso de modelos em Ecologia de Paisagens

De forma complementar às alternativas apresentadas acima, a modelagem vem se firmando como uma excelente aliada para a obtenção de conhecimento e geração de hipóteses em Ecologia de Paisagens. Numa revisão de trabalhos publicados na revista *Landscape Ecology*, foi possível observar que a porcentagem de artigos que utilizam modelos subiu de *ca.* 15%, entre 1987 e 1991, para *ca.* 25%, entre 1991 e 1995 (Hobbs, 1996), e para mais de 35% a partir de 2000 (Metzger, 2006). Questões populacionais, incluindo dinâmica de metapopulações, efeitos de fragmentação, importância de corredores e processos de dispersão ou invasão, estão entre os temas mais abordados com modelos em Ecologia de Paisagens (37% dos 101 trabalhos publicados entre 1994 e 1999 analisados por Turner *et al.*, 2001). Em seguida, aparecem temas ligados a efeitos de perturbação (17%), dinâmica de vegetação (15%) e fatores abióticos (11%).

Os modelos são idealizações ou caricaturas da realidade. Eles são utilizados nas mais diversas disciplinas

como um meio para formalizar, de modo simplificado, nosso entendimento sobre um determinado assunto e testar, empiricamente, a validade das premissas ou hipóteses utilizadas em sua construção. Em Ecologia de Paisagens, o principal objetivo é analisar as relações entre padrões espaciais e processos ecológicos. A partir do conhecimento adquirido empiricamente (por experimentação ou observação), modelos podem ser utilizados para formalizar estas relações, de acordo com algumas premissas, e simular, a partir de um padrão espacial conhecido, as alterações esperadas. Estas, por sua vez, podem ser consideradas hipóteses a serem testadas empiricamente, de forma a sabermos se nosso modelo está ou não adequadamente representando as relações entre padrões e processos (Figura 1). Ou seja, é da estreita relação entre a formalização de modelos e o teste empírico para sua validação que se espera uma sedimentação do conhecimento. A modelagem é assim totalmente complementar à experimentação e à observação na geração do conhecimento.



**FIGURA 1** – O uso de modelos em Ecologia de Paisagens como instrumento para gerar hipóteses a serem testadas sobre a relação entre padrões espaciais e processos ecológicos.

Além de gerar hipóteses, os modelos, uma vez testados, permitem: i) comparar diferentes cenários de alteração da paisagem; ii) extrapolar os resultados de um caso particular para diferentes escalas espaciais ou temporais ou para condições diferentes das observadas

atualmente (por exemplo, para uma ampla gama de porcentagens de hábitat na paisagem; diferentes configurações espaciais; diferentes condições climáticas); e iii) testar a importância relativa de cada parâmetro (em modelos, é possível controlar as variáveis) ou a sensibilidade do modelo a um determinado parâmetro (dando assim indicações de que parâmetros devem ser observados ou monitorados no campo). Em termos práticos, bons modelos podem ser valiosos instrumentos de decisão (enquanto modelos falsos podem ser extremamente perniciosos).

Uma ampla variedade de modelos é utilizada em Ecologia de Paisagens (Turner *et al.*, 2001), desde os mais simples, como modelos determinísticos baseados em manchas, até os extremamente complexos, como os modelos espacialmente explícitos baseados em indivíduos (os *Individual-Based Models*, IBM) (Tabela 2). Apesar desta ampla gama de possibilidades, os modelos utilizados em Ecologia de Paisagens caracterizam-se por considerarem as relações “padrões-processos” e a heterogeneidade do espaço, de forma implícita, explícita ou realista<sup>1</sup>.

**TABELA 2** – Principais características de modelos com menor ou maior complexidade que podem ser usados em Ecologia de Paisagens (baseado em Turner *et al.*, 2001 e em comunicação pessoal de Jane Jepsen).

PARÂMETRO	MODELOS SIMPLES	MODELOS COMPLEXOS
Formalização	Analíticos	Simulações
Estocasticidade	Determinísticos	Estocásticos
Tempo	Discretos	Contínuos
Espaço	Não espaciais	Espaciais
Unidade	Manchas ou populações	Indivíduos
Requerimento de dados	Baixo	Alto
Custo	Baixo	Alto

Os modelos em Ecologia de Paisagens são utilizados para as mais diversas finalidades, em particular para: (1) descrição da estrutura da paisagem, como por exemplo, através da teoria dos grafos (Keitt *et al.*, 1997; Urban & Keitt, 2001) ou da teoria da percolação (Stauffer, 1985), utilizando-se ou não modelos neutros (Gardner *et al.*, 1987); (2) análise da dinâmica da paisagem, através de modelos de dinâmica de manchas (Pickett & Thompson, 1978) ou de matrizes de transição (*e.g.*, Acevedo *et al.*, 1995), geralmente enfatizando os principais agentes causadores desta dinâmica, através, por exemplo, de regressões logísticas, modelos de vizinhança, ou de análises multivariadas; (3) estudos espacialmente explícitos de dinâmica de populações (revisados em Dunning *et al.*, 1995; ver também revisão de IBM, em Grimm, 1999) ou metapopulações (por exemplo, as Funções de Incidência; Hanski, 1992, 1994); e (4) análise da relação entre a estrutura da paisagem e os mais diversos processos ecológicos (por exemplo, propagação de perturbação, deslocamento de indivíduos, taxa de mortalidade) ou padrões biológicos (*e.g.*, diversidade ou riqueza de espécies, presença ou abundância de espécies), através de regressões, análises de correspondência, testes de Mantel, entre outros. Na próxima seção, procura-se exemplificar algumas das possíveis aplicações de modelos em Ecologia de Paisagens.

## EXEMPLOS DE APLICAÇÃO

### Modelos que enfatizam a relação entre estrutura e diversidade da paisagem

O estudo das relações entre a estrutura da paisagem e a composição e riqueza de espécies é sem dúvida um dos temas de pesquisa mais freqüente em Ecologia de Paisagens. Para tanto, diversos modelos podem ser utilizados em dois principais momentos: na quantificação da estrutura da paisagem; e no estabelecimento das relações entre paisagem e riqueza de espécies.

<sup>1</sup> **Modelos espacialmente implícitos** incorporam parâmetros espaciais, porém sem considerar a posição exata dos componentes no espaço. Por exemplo, a Teoria da Biogeografia de Ilhas considera o tamanho e a distância das ilhas ao continente (ambos, parâmetros espaciais), porém, uma vez obtidos esses parâmetros, não é necessário ter a localização exata das ilhas num mapa. Trata-se tipicamente de um modelo espacialmente implícito.

Em **modelos espacialmente explícitos** é necessário ter a posição exata dos componentes do sistema no espaço, pois as propriedades deste componente dependem do contexto no qual ele se insere. Estes modelos simplificam o espaço em unidades (células) de tamanho e forma idênticos, com regras de mudança de estado (regras de transição) que dependem da vizinhança.

**Modelos espacialmente realistas** também usam a posição exata dos componentes do sistema no espaço, porém não procuram simplificar esse espaço em unidades (células) semelhantes. Estes modelos lidam com objetos (*e.g.*, manchas) de tamanhos e formas diferentes, situados em matrizes heterogêneas. Em geral, estes são os modelos mais complexos.

Um exemplo deste tipo de pesquisa pode ser fornecido pelo projeto *Biodiversity conservation in fragmented landscapes at the Atlantic Plateau of São Paulo* (Metzger, 1999a). Este projeto tem por objetivo entender a influência do tamanho e do grau de conectividade<sup>2</sup> de fragmentos de florestas ombrófilas densas da Mata Atlântica na distribuição de espécies com diferentes sensibilidades à modificação do hábitat, na composição e riqueza de comunidades de diferentes táxons (árvores, mamíferos, aves, répteis, anfíbios, aranhas e borboletas), na variabilidade genética de algumas populações (em particular, de anfíbios e pequenos mamíferos) e em alguns processos ecológicos (mortalidade de plântulas, polinização, dispersão de sementes e ciclagem de nutrientes). Para tanto, estão sendo estudadas duas paisagens de 100 km<sup>2</sup> no Planalto de Ibiúna (cerca de 50 km a sudoeste da cidade de São Paulo): uma paisagem florestal, a Reserva do Morro Grande, considerada como área controle, não-fragmentada; e uma paisagem fragmentada, na região de Caucaia do Alto, onde as áreas de floresta encontram-se reduzidas a fragmentos de diferentes tamanhos e graus de conectividade, circundadas por áreas de agricultura anual e silvicultura. Estão sendo estudados 21 fragmentos, sendo cinco grandes (50-275 ha), oito médios (10-45 ha) e oito pequenos (< 5 ha).

A análise da estrutura da paisagem foi baseada numa abordagem categórica, onde unidades discretas são identificadas, mapeadas e posteriormente caracterizadas quantitativamente por uma série de índices espaciais (Metzger, 2003b). Esta abordagem é consistente com o modelo conceitual de “mancha-corredor-matriz”<sup>3</sup>, o mais empregado para se definir os elementos da paisagem em função de sua disposição espacial. Com exceção das paisagens nas quais a heterogeneidade se apresenta sob forma de gradientes ambientais, praticamente toda paisagem pode ser representada a partir de manchas, corredores e matriz. Desta forma, utilizou-se um levantamento aerofotogramétrico recente,

de abril de 2000, na escala de 1/10.000, e foi feito, por foto-interpretção, o mapeamento de 17 unidades da paisagem, incluindo vegetação natural em sete estádios de sucessão, quatro tipos de campos antrópicos e áreas agrícolas, reflorestamentos, instalações rurais, núcleos urbanos, corpos d’água e mineração. Para a análise quantitativa, as áreas de vegetação natural foram agrupadas em três classes: floresta em estádios inicial/médio a avançado (fisionomia predominantemente arbórea); floresta em estádios iniciais (fisionomias arbustivas); e vegetação natural pioneira (fisionomias predominantemente arbustivas/herbáceas).

Para responder a pergunta central do projeto, duas características estruturais da paisagem foram analisadas: o tamanho e o grau de conectividade dos fragmentos. Os limites dos fragmentos foram definidos pela predominância de fisionomias arbóreas naturais, o que permitiu acessar as extensões de cada fragmento. No caso da conectividade, o cálculo foi baseado na *correlation length* (Keitt *et al.*, 1997), um índice derivado da teoria dos grafos (Gross & Yellen, 1999). Esta teoria tem sido aplicada para descrever diferentes tipos de redes, como as formadas por estradas, computadores, ou por cadeias hierárquicas em empresas. Em análises espaciais em Ecologia de Paisagens, os grafos podem ser entendidos como conjunto de fragmentos espacialmente isolados, porém funcionalmente unidos, de forma similar ao que ocorre numa metapopulação. Considera-se que dois fragmentos estão unidos quando a distância entre eles ou o tipo de matriz que os separa permite fluxos biológicos. Uma vez definidos todos os grafos de uma paisagem, é possível medir a conectividade de cada um deles, em função do número de fragmentos unidos ou de sua extensão espacial (*gyrate*, ou área de giro), e de uma média da conectividade de todos os grafos da paisagem, *i.e.* a conectividade da paisagem (Keitt *et al.*, 1997). Quatro índices de conectividade foram assim calculados para cada fragmento, considerando duas capacidades de deslocamento das espécies (não se

<sup>2</sup> A conectividade é a capacidade da paisagem de facilitar os fluxos biológicos. Essa conectividade depende muito de características das espécies, em particular da capacidade de deslocamento em diferentes tipos de hábitat. Entretanto, ela também pode ser medida em termos estruturais, em função da proximidade dos fragmentos de hábitat, do grau de conexão espacial entre eles, através de corredores ou *stepping stones* (pontos de ligação; Metzger, 1999b), ou, ainda, em função da permeabilidade da matriz inter-hábitat.

<sup>3</sup> As manchas são áreas homogêneas, numa determinada escala espacial, que se distinguem das unidades vizinhas e apresentam extensões espaciais reduzidas e não-lineares. Os corredores são também áreas homogêneas, que se distinguem das unidades vizinhas e que apresentam disposição espacial linear. E a matriz é definida, segundo Forman (1995), como a unidade da paisagem que controla a dinâmica da paisagem. Em geral, essa unidade pode ser reconhecida por recobrir a maior parte da paisagem (*i.e.*, sendo a unidade dominante em termos de recobrimento espacial), ou por ter um maior grau de conexão de sua área (*i.e.*, um menor grau de fragmentação).

deslocam fora do hábitat; podem se deslocar até 50 m pela matriz) e diferentes cenários de permeabilidade das unidades da paisagem (conectividade apenas por áreas florestais e conectividade por toda a vegetação natural, independente de sua fisionomia). Todos os índices foram calculados no programa FRAGSTATS™ versão 3.3, numa imagem matricial de 5 metros de resolução.

Uma vez obtidos os índices de estrutura da paisagem e feitos os levantamentos de composição e riqueza das espécies dos diferentes grupos taxonômicos estudados nos fragmentos, procurou-se relacionar os padrões espaciais com os padrões de distribuição das espécies através de regressões lineares. No caso de aves de sub-bosque amostradas por redes de neblina (ca. 540 horas/rede por fragmento; Martensen *et al.*, 2007), a riqueza de espécies está marginalmente relacionada com a área dos fragmentos, porém está fortemente associada à proporção de mata numa vizinhança de 800m e à conectividade florestal do fragmento. Estes dados sugerem que os limites dos fragmentos, tal qual definidos pela foto-interpretação, não correspondem aos limites funcionais (ou biologicamente significativos) para o conjunto da avifauna. Ou seja, as espécies respondem mais a características da paisagem (proporção de mata, conectividade) do que ao tamanho do fragmento, indicando que a permanência de algumas espécies nos fragmentos depende do tipo de entorno. Estes resultados por si só não permitem estabelecer relações causais, porém, ao indicarem possíveis interações, têm como benefício o direcionamento (afunilamento) da pesquisa em questões mais pontuais e funcionais, como por exemplo, em experimentos de deslocamento pela paisagem.

Este tipo de análise vem sendo comumente utilizado em estudos de fragmentação, onde variáveis biológicas são relacionadas a características de manchas de hábitat (grande maioria dos estudos), ou a características da paisagem no entorno destas manchas (McGarigal & Cushman, 2002). Estes trabalhos devem ser considerados como uma etapa exploratória no estabelecimento de relações causais entre padrões espaciais e processos ecológicos (Metzger, 2003a). McGarigal & Cushman (2002) sugerem ainda que, como a fragmentação é um processo que ocorre no nível da paisagem, é necessário mudar o foco da pesquisa, analisando não apenas fragmentos com diferentes características (*e.g.*, tamanho e isolamento) inseridos numa única paisagem, mas também comparando paisagens com diferentes graus de fragmentação.

### Modelos de simulação da paisagem

A importância do arranjo espacial do hábitat sobre a persistência das espécies é outro foco de estudos em Ecologia de Paisagens. A idéia é entender, a partir das ocorrências e das propriedades das populações atualmente presentes, em que condições estas populações tenderiam a se extinguir ou, inversamente, a se perpetuar. Para tanto, é necessário extrapolar os padrões atuais de presença para outras condições espaciais ou ao longo do tempo, o que muitas vezes só pode ser feito através de modelagem. Vários modelos sofisticados têm sido utilizados para responder tais questões, entre eles os Modelos de Viabilidade Populacional (Akçakaya & Ferson, 1992; Possingham *et al.*, 1992) ou modelos de dinâmica de populações espacialmente explícitos (Dunning *et al.*, 1995). Quando estes modelos incorporam parâmetros espaciais para analisar características populacionais, obrigatoriamente eles assumem uma perspectiva de Ecologia de Paisagens.

Um dos grandes desafios destes modelos é parametrizar a importância dos diferentes tipos de unidades da paisagem nos processos demográficos e no deslocamento das espécies, o que só pode ser atingido ao se considerar o espaço sob a ótica das espécies estudadas (Vos *et al.*, 2001). Uma estratégia para se evoluir neste sentido é considerar que a estrutura da paisagem biologicamente significativa (na ótica das espécies) é aquela que melhor se correlaciona com os padrões biológicos observados. Ou seja, inversamente ao que se fez em Caucaia, onde se impunha um padrão espacial numa visão antropocêntrica e procurava-se relacioná-lo com as propriedades biológicas (no caso, a riqueza de espécies), uma outra estratégia de pesquisa é de simular diferentes padrões espaciais (ou “visões” de uma mesma paisagem) e analisar qual destes padrões espaciais melhor se correlaciona com os padrões biológicos.

Esta foi a estratégia adotada para se testar a importância da permeabilidade da matriz inter-hábitat na migração de aves da floresta contínua para fragmentos florestais na Amazônia (Antongiovanni, 2001; Antongiovanni & Metzger, 2005). Este trabalho foi desenvolvido nas áreas do Projeto de Dinâmica Biológica de Fragmentos Florestais (PDBFF), situadas a cerca de 80 km ao Norte de Manaus, AM. Foram estudados nove fragmentos, de 1 e 10 ha, isolados entre 1980 e 1984, e que se encontram envoltos por três diferentes tipos de matrizes: capoeiras dominadas por *Vismia* spp, em áreas que sofreram corte e queima; capoeiras dominadas por *Cecropia* spp, em áreas que sofreram apenas corte; e áreas de pastagem ainda ativas. A comunidade de aves

de sub-bosque foi amostrada nos fragmentos com redes de neblina, desde antes do isolamento dos fragmentos até 1992 (Bierregaard Jr. & Lovejoy, 1989; Stouffer & Bierregaard, 1995). Onze espécies insetívoras foram analisadas pormenorizadamente, por estarem entre as espécies mais comuns antes da fragmentação e por serem facilmente capturadas em redes de neblina. Nove destas espécies sofreram extinções locais após o isolamento dos fragmentos, porém, com o abandono das pastagens e conseqüente desenvolvimento das capoeiras de *Cecropia* spp e *Vismia* spp, houve recolonização (Stouffer & Bierregaard, 1995). Segundo estes autores, as recolonizações foram mais freqüentes quando os fragmentos estavam envoltos por capoeiras de *Cecropia* spp, indicando que estas capoeiras eram mais permeáveis aos fluxos biológicos. O trabalho de Antongiovanni (2001) teve justamente por objetivo testar essa afirmação. Desta forma, para cada fragmento e ano analisado (1985 a 1992), foram calculados índices de permeabilidade da matriz que levam em conta: (1) a *distância* a ser percorrida no caminho de deslocamento entre o fragmento e a área de mata primária contínua mais próxima (considerada como única fonte de indivíduos, uma vez que estudos anteriores mostram que há pouco movimento de indivíduos entre fragmentos); e (2) a *resistência* oferecida pelos diferentes tipos de matriz. Os valores de permeabilidade da matriz foram relacionados, com o uso de regressões logísticas, com a ocorrência de migração de indivíduos de cada uma das 11 espécies estudadas, por fragmento e para cada ano em estudo ( $n = 60$ ). Para o cálculo da permeabilidade da matriz, assumiu-se como premissa que o indivíduo iria percorrer o caminho mais curto entre a fonte e o fragmento. As relações explicitadas acima foram realizadas para cada uma das espécies e quatro conjuntos de valores de permeabilidade da matriz, obtidos por meio da simulação de quatro cenários possíveis onde: (1) capoeiras de *Cecropia* spp são mais permeáveis; (2) capoeiras de *Vismia* spp são mais permeáveis; (3) capoeiras de *Vismia* spp e *Cecropia* spp são igualmente permeáveis, porém capoeiras mais velhas são mais permeáveis que as mais jovens; e (4) não há diferença na permeabilidade das unidades da matriz, *i.e.*, apenas é considerada a distância entre a floresta contínua e o fragmento. Cada um destes cenários corresponde a uma possível “visão” da estrutura da paisagem, sendo que o objetivo é justamente ver qual destas visões melhor se ajusta aos padrões observados de migração das espécies.

As regressões logísticas mostram que o cenário em que apenas o isolamento é considerado não se ajusta aos dados de ocorrência de migrações para nenhuma

espécie. Por outro lado, os cenários em que *Cecropia* spp e capoeiras mais velhas são mais permeáveis foram significativamente ajustados à ocorrência de migração de sete espécies, confirmando a hipótese de Stouffer & Bierregaard (1995). Para outras quatro espécies, nenhum dos cenários propostos explicou significativamente a migração de indivíduos para os fragmentos, o que pode significar que estas espécies são indiferentes à matriz, ou que nenhum dos cenários se aproxima da visão da matriz destas espécies. Este trabalho mostra que o uso de modelos de simulação para criar paisagens segundo a percepção das espécies é uma estratégia útil da análise da paisagem. A aplicação de modelos de permeabilidade de matriz no estudo da migração de indivíduos em paisagens fragmentadas deve ter um papel essencial no entendimento de mosaicos antrópicos e na conservação da diversidade biológica nestas paisagens (Antongiovanni & Metzger, 2005).

#### **Modelos de dinâmica de paisagem e suas conseqüências em termos de biodiversidade**

Outro tópico relevante nas pesquisas em Ecologia de Paisagens é o estudo da influência de dinâmica da paisagem sobre os processos biológicos. Sabe-se que diferentes padrões de modificação do hábitat resultam em diferentes estruturas espaciais, o que, por sua vez, implica em diferentes influências sobre os processos biológicos (Gustafson, 1999). Modelos de dinâmica da paisagem têm sido utilizados para analisar as mais variadas conseqüências das mudanças da paisagem, como na emissão de gases de efeito estufa, na produção madeireira, na sustentabilidade social de uma região, ou ainda na manutenção da biodiversidade (Dale *et al.*, 1993a, b; 1994a, b; Gustafson, 1999; entre muitos outros). Dentre estes tópicos, um tem despertado especial atenção: a influência de padrões de desmatamento sobre os riscos de extinção. Para abordar esta problemática, é interessante analisar qual seria o padrão esperado de perda de espécies ao longo de um processo de desmatamento. O uso de princípios da teoria da percolação em modelos neutros tem sido muito útil neste sentido (Gardner *et al.*, 1987; Gardner & O'Neill, 1991).

Um modelo neutro consiste num conjunto de regras “simples” usadas para criar um padrão espacial, em geral bidimensional. O modelo neutro permite criar e analisar uma estrutura na ausência de fatores geradores de padrão (*i.e.*, de ordem), e, ao compararmos este padrão ao observado, é possível ter uma idéia de quanto a paisagem real se ajusta às regras do modelo neutro. Em geral, testa-se inicialmente o modelo mais simples de todos: o padrão neutro totalmente aleatório.

A teoria da percolação foi desenvolvida inicialmente na física para responder perguntas relacionadas à conectividade em espaços bidimensionais (Stauffer, 1985), tais como: quanto metal é necessário acrescentar de forma que a eletricidade possa passar num determinado material? A idéia seria encontrar a quantidade mínima necessária para ter condutividade, pois o acréscimo de mais metal aumentaria os custos. Nesta teoria, há apenas dois estados em que um material pode se encontrar: há percolação (no caso, passagem de eletricidade), pois o metal condutor atravessa o material de ponta a ponta; e não há percolação, quando metal condutor encontra-se fragmentado e não permite unir duas pontas do material estudado. Ecólogos estão interessados em questões similares às dos físicos, porém, ao invés de percolação elétrica, o interesse está na percolação de indivíduos de uma espécie ou de perturbações: quanto hábitat é necessário para permitir que uma determinada espécie possa se locomover de um dado ponto a outro de uma paisagem? Que quantidade de mata, com alta probabilidade de pegar fogo, pode permitir (ou impedir) a propagação de uma queimada na paisagem? Devido à similaridade das perguntas de físicos e ecólogos, a teoria da percolação vem sendo constantemente aplicada em Ecologia de Paisagens (Gardner *et al.*, 1987; Turner *et al.*, 1989; Andren, 1994, With *et al.*, 1997).

Quando se analisa a percolação de hábitat em paisagens neutras, observa-se que ocorrem mudanças bruscas na estrutura da paisagem no limiar de percolação, quando há ruptura da continuidade do hábitat original, aumento brusco do grau de fragmentação e redução abrupta do tamanho dos fragmentos (With & King, 1999). Todas estas mudanças sugerem um aumento dos riscos locais de extinção e uma diminuição nas possibilidades de recolonização, o que pode significar um rápido aumento dos riscos de extinção no conjunto da paisagem (Metzger & Decamps, 1997).

Para testar esta hipótese do aumento brusco dos riscos de extinção em paisagens reais, modelos de simulação de perda de espécies foram utilizados para se analisar três padrões comuns de desmatamento na Amazônia brasileira (Oliveira-Filho, 2001; Oliveira-Filho & Metzger, 2006): (1) *espinha de peixe*, onde lotes de mesmo tamanho (*e.g.*, 50 ha) são distribuídos regularmente ao longo de estradas, como ocorre em projetos de colonização governamentais; (2) *desordenado*, formado por lotes pequenos, de tamanhos variados e irregularmente distribuídos no espaço, geralmente resultante de colonização não-induzida; e (3) *grandes proprietários*, *e.g.* fazendeiros que promovem pecuária extensiva em lotes de mais de 1000 ha.

Estes padrões foram estudados na região de Alta Floresta, no norte do estado de Mato Grosso. Para cada um destes padrões, foram selecionadas três áreas (réplicas), de 8 por 8 km, onde em 1984 havia mais de 90% de mata e em 1998 havia menos de 30% de mata. As alterações na estrutura da paisagem foram seguidas a partir de imagens de satélite Landsat-TM, classificadas em mata e não-mata, numa série bianual de 1984 a 1998. Um padrão aleatório de desmatamento foi gerado pelo programa RULE (Gardner, 1999), contendo entre 95% e 15% de floresta de forma a simular oito momentos ao longo de um processo de desmatamento. Para simular a perda de espécies ao longo do processo de desmatamento, foram criados 60 grupos funcionais, resultantes da combinação de características de: (1) deslocamento pela matriz (5 classes: não se desloca pela matriz ou se desloca 60, 120, 240 ou 480 metros matriz adentro); (2) área de vida (6 classes: 125, 250, 500, 1000, 2000 e 4000 ha); e (3) tolerância ou não aos efeitos de borda, numa faixa de 90 m a partir do contato entre mata e não-mata (duas classes). Todos os grupos funcionais foram inseridos nas paisagens iniciais (1984 para as paisagens reais e 95% de hábitat para a simulação) e foram mantidos na paisagem enquanto os requerimentos de área de vida fossem respeitados, considerando-se a capacidade de deslocamento e a sensibilidade às bordas. O modelo utilizado é determinista, na medida em que para cada estrutura há apenas uma condição possível para cada grupo funcional (presente ou ausente, dependendo do caso). Ademais, esse modelo não considera o tempo de latência entre as modificações da paisagem e a extinção do grupo funcional.

Os resultados da dinâmica da paisagem mostram que a paisagem simulada (aleatória) apresenta um padrão não-linear de alteração de sua estrutura para todos os índices considerados (número de fragmentos, tamanho médio dos fragmentos, isolamento e conectividade). Por outro lado, os padrões reais apresentaram tanto variações lineares (em particular, no caso do número de fragmentos) quanto não-lineares (para os demais índices), sendo que três momentos de mudança brusca ocorreram: no início do processo de desmatamento, quando há uma perda rápida no tamanho médio dos fragmentos; num momento intermediário, quando há perda de conectividade ou aumento do isolamento; e no final do processo de desmatamento, quando há um aumento brusco no grau de isolamento. Os momentos de mudança brusca variaram pouco entre os padrões de desmatamento reais, contrariamente ao esperado. Surpreendentemente, e apesar da ocorrência destes limiares de mudança brusca da paisagem, o padrão de perda dos grupos funcionais foi linear para todos os

padrões de desmatamento, tanto os reais quanto o aleatório. As principais diferenças foram observadas entre grupos funcionais com diferentes características, em particular, o grupo de espécies de interior se extinguiu mais rapidamente que o grupo de espécies de borda, ou ainda os grupos funcionais com menor capacidade de deslocamento e/ou maior área de vida tenderam também a se extinguir mais cedo no processo de desmatamento.

Desta forma, foi possível distinguir quatro grandes conjuntos de grupos funcionais: (1) aqueles que são muito sensíveis ao desmatamento e que se extinguem constantemente no início do processo, quando há uma queda brusca no tamanho médio dos fragmentos; (2) aqueles que são sensíveis ao padrão de desmatamento e ao arranjo espacial da floresta remanescente, tendendo a se extinguir mais tarde quando o padrão é mais agregado; esse grupo é sensível aos limiares dos índices de isolamento e conectividade; (3) os grupos funcionais que são apenas sensíveis ao desmatamento, mas não ao arranjo espacial dos remanescentes, tendendo a se extinguir progressivamente (linearmente) à medida que ocorre o desmatamento; e (4) finalmente, os grupos funcionais pouco sensíveis, que não se extinguem nas condições estudadas. Apesar da simplicidade e das limitações do modelo utilizado (não considera o tempo de latência; é determinista; não considera nenhuma característica relacionada à demografia das espécies), ele permite evidenciar padrões de perda de espécies que só poderiam ser observados através do monitoramento biológico de paisagens em processo de desmatamento, algo quase impossível de ser realizado. Ademais, esses modelos, se corretamente validados, podem dar importantes indicações sobre que padrão de desmatamento é menos prejudicial em termos biológicos, sendo assim de grande valia no planejamento da expansão da fronteira agrícola na Amazônia e na otimização da conservação da biodiversidade em paisagens antropizadas.

## CONSIDERAÇÕES FINAIS

Devido à complexidade das questões abordadas em Ecologia de Paisagens e à dificuldade de se realizar experimentos com paisagens, a aquisição de conhecimento através do uso de modelos revela-se uma estratégia útil e complementar à experimentação ou à comparação de paisagens. Modelos podem ser úteis para criar hipóteses, avaliar a influência de um determinado fator, ou mesmo para extrapolar os padrões observados ao longo do tempo ou para condições espaciais de difícil observação no campo. Há uma grande diversidade

de modelos utilizados em Ecologia de Paisagens, sendo que estes foram, em geral, inicialmente desenvolvidos em outras disciplinas. Há, assim, poucos modelos de Ecologia de Paisagens, mas muitos modelos usados ou adaptados para o estudo de problemáticas próprias da Ecologia de Paisagens, *i.e.*, relacionados com a investigação das relações entre padrões espaciais e processos ecológicos, ou pelo menos na análise de um destes componentes. Por incorporarem parte da heterogeneidade da paisagem, estes modelos são relativamente complexos e, quase que obrigatoriamente, são espacialmente explícitos ou realistas. Um desafio futuro é encontrar um balanço entre a complexidade destes modelos, que acabam se tornando cada vez mais específicos a um tipo de situação, e a capacidade de generalização, essencial para o avanço de uma teoria em Ecologia de Paisagens.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Acevedo, M.F., D.L. Urban & M. Aflan. 1995. Transition and Gap Models of Forest Dynamics. *Ecological Applications* 5: 1040-1055.
- Akçakaya, H.R. & S. Ferson. 1992. RAMAS/Space user manual: spatially structured population models for conservation biology. Applied Biomathematics, New York.
- Andren, H. 1994. Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat - a review. *Oikos* 71: 355-366.
- Antongiovanni, M.A. 2001. Efeito da matriz sobre a avifauna de sub-bosque de fragmentos florestais na Amazônia (AM). Dissertação de Mestrado. Instituto de Biociências, Universidade de São Paulo, São Paulo.
- Antongiovanni, M.A. & J.P. Metzger. 2005. Matrix permeability effect on inter-habitat migration of insectivorous bird species. *Biological Conservation*. 122: 441-451.
- Bierregaard Jr, R.O. & T.E. Lovejoy. 1989. Effects of forest fragmentation on Amazonian understory bird communities. *Acta Amazonica* 19: 215-241.
- Bissonette, J.A. 1997. *Wildlife and landscape ecology: effects of pattern and scale*. Springer-Verlag, New York.
- Bissonette, J.A. & I. Storch, 2003. *Landscape ecology and resource management*. Island Press, Washington.
- Dale, V.H., R.V. Oneill, M. Pedlowski & F. Southworth. 1993a. Causes and effects of land-use change in central Rondonia, Brazil. *Photogrammetric Engineering and Remote Sensing* 59: 997-1005.
- Dale, V.H., F. Southworth, R.V. O'Neill, A. Rosen & R. Frohn. 1993b. Simulating spatial patterns of land-use change in Rondônia, Brazil. *Lectures on Mathematics in the Life Sciences* 23: 29-55.
- Dale V.H., R.V. Oneill, F. Southworth & M. Pedlowski. 1994a. Modeling effects of land management in the Brazilian Amazonian settlement of Rondonia. *Conservation Biology* 8: 196-206.
- Dale V.H., S.M. Pearson, H.L. Offerman & R.V. Oneill. 1994b. Relating patterns of land-use change to faunal biodiversity in the Central Amazon. *Conservation Biology* 8: 1027-1036.

- Debinski, D.M. & R.D. Holt. 2000. A survey and overview of habitat fragmentation experiments. *Conservation Biology* 14: 342-355.
- Dunning, J.B., D.J. Stewart, B.J. Danielson, B.R. Noon, T.L. Root, R.H. Lamberson & E.E. Stevens. 1995. Spatially Explicit Population-Models - Current Forms and Future Uses. *Ecological Applications* 5: 3-11.
- Forman, R.T.T. 1995. *Land mosaics: the ecology of landscapes and regions*. Cambridge University Press, Cambridge.
- Gardner, R.H. 1999. RULE: map generation and a spatial analysis program. In: J.M. Klopatek & R.H. Gardner (eds). *Landscape ecological analysis: issues and applications*. pp. 280-303. Springer, New York.
- Gardner, R.H. & R.V. O'Neill, 1991. Pattern, process and predictability: the use of neutral model for landscape analysis. In: M.G. Turner & R.H. Gardner (eds). *Quantitative methods in landscape ecology - The analysis and interpretation of landscape heterogeneity*. pp. 77-103. Springer-Verlag, New York.
- Gardner, R.H., B.T. Milne, M.G. Turner, & R.V. O'Neill. 1987. Neutral models for the analysis of broad-scale landscape pattern. *Landscape Ecology* 1: 19-28.
- Gonzales, A., J.H. Lawton, F.S. Gilbert, T.M. Blackburn & I. Evans-Freke. 1998. Metapopulation dynamics, abundance and distribution in micro ecosystems. *Science* 281: 2045-2047.
- Grimm, V. 1999. Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? *Ecological Modelling* 115: 129-148.
- Gross, J. & J. Yellen. 1999. *Graph theory and its applications*. CRC Press, Florida.
- Gustafson, E.J. 1999. HARVEST: a timber harvest allocation model for simulating management alternatives. In: J.M. Klopatek & R.H. Gardner (eds). *Landscape ecological analysis: issues and applications*. pp. 109-124. Springer, New York.
- Gutzwiller, K.J. 2002. *Applying landscape ecology in biological conservation*. Springer, New York.
- Hanski, I. 1992. Inferences from Ecological Incidence Functions. *American Naturalist* 139: 657-662.
- Hanski, I. 1994. Patch-occupancy dynamics in fragmented landscapes. *Trends in Ecology & Evolution* 9: 131-135.
- Hobbs, R.J. 1996. Future landscapes and the future of landscape ecology. *Landscape & Urban Planning* 37: 1-9.
- Keitt, T.H., D.L. Urban & B.T. Milne. 1997. Detecting critical scales in fragmented landscapes. *Conservation Ecology* [on line] 1: 4.
- Martensen, A.C., P.F. Develey & J.P. Metzger. 2007. Effects of forest fragmentation on the understory bird community of the Atlantic Rain Forest of Southern Brazil. *Ornitologia Neotropical, Canadá*, v press.
- McGarigal, K. & S.A. Cushman. 2002. Comparative evaluation of experimental approaches to the study of habitat fragmentation effects. *Ecological Applications* 12: 35-45.
- Metzger, J.P. 1999a. Biodiversity conservation in fragmented landscapes at the Atlantic Plateau of São Paulo (Brazil). Projeto Fapesp 99/05123-4. Disponível em <http://eco.ib.usp.br/lepac/biota-caucaia> (acessado em julho de 2007).
- Metzger, J.P. 1999b. Estrutura da paisagem e fragmentação: análise bibliográfica. *Anais da Academia Brasileira de Ciências* 71: 445-463.
- Metzger, J.P. 2001. O que é ecologia de paisagens? *Biota Neotropica* 1: 1-9.
- Metzger, J.P. 2003a. Delineamento de experimentos numa perspectiva de ecologia da paisagem. In: L. Cullen Jr., R Rhudram & C. Valladares-Padua (eds). *Métodos e Técnicas na Biologia da Conservação e no Manejo da Vida Silvestre*. Universidade Federal do Paraná, Curitiba.
- Metzger, J.P. 2003b. Quantificação da estrutura da paisagem: o uso adequado de métricas. In: L. Cullen Jr., R Rhudram & C. Valladares-Padua (eds). *Métodos e técnicas na biologia da conservação e no manejo da vida silvestre*. Universidade Federal do Paraná, Curitiba.
- Metzger, J.P. 2006. *Ecologia de paisagens fragmentadas*. Tese de Livre Docência. Instituto de Biociências, Universidade de São Paulo.
- Metzger, J.P. & H. Decamps. 1997. The structural connectivity threshold: an hypothesis in conservation biology at the landscape scale. *Acta Oecologica - International Journal of Ecology* 18: 1-12.
- Oliveira-Filho, F.J.B. 2001. Padrão de desmatamento e evolução da estrutura da paisagem em Alta Floresta (MT). Dissertação de Mestrado. Instituto de Biociências, Universidade de São Paulo, São Paulo.
- Oliveira Filho, F.J.B. & J. P. Metzger. 2006. Thresholds in landscape structure for three common deforestation patterns in the Brazilian Amazon. *Landscape Ecology* 21: 1061-1073.
- Pickett, S.T.A. & J.N. Thompson. 1978. Patch dynamics and design of nature reserves. *Biological Conservation* 13: 27-37.
- Possingham, H.P., I. Davies, I.R. Noble & T.A. Norton. 1992. A metapopulation simulation model for assessing the likelihood of plan and animal extinctions. *Mathematics and Computers in Simulation* 33: 367-372.
- Stauffer, D. 1985. *Introduction to percolation theory*. Taylor and Francis, London.
- Stouffer, P.C. & R.O. Bierregaard. 1995. Use of amazonian forest fragments by understory insectivorous birds. *Ecology* 76: 2429-2445.
- Turner, M.G. 1989. Landscape Ecology - the effect of pattern on process. *Annual Review of Ecology and Systematics* 20: 171-197.
- Turner, M.G., R.H. Gardner, V.H. Dale & R.V. Oneill. 1989. Predicting the spread of disturbance across heterogeneous landscapes. *Oikos* 55: 121-129.
- Turner, M.G., R.H. Gardner & R.V. O'Neill. 2001. *Landscape ecology in theory and practice: pattern and process*. Springer, New York.
- Urban, D. & T. Keitt. 2001. Landscape connectivity: a graph-theoretic perspective. *Ecology* 82: 1205-1218.
- Vos, C., J. Verboom, P. Opdam, & C.J.F. Braak, 2001. Toward ecologically scaled landscape indices. *American Naturalist* 158: 24-41.
- With, K.A. & A. W. King. 1999. Dispersal success on fractal landscapes: a consequence of lacunarity thresholds. *Landscape Ecology* 14: 73-82.
- With, K.A., S.J. Cadaret & C. Davis. 1999. Movement responses to patch structure in experimental fractal landscapes. *Ecology* 80: 1340-1253.
- With, K.A., R.H. Gardner & M.G. Turner. 1997. Landscape connectivity and population distributions in heterogeneous environments. *Oikos* 78: 151-169.
- Wu, J. & R. Hobbs. 2002. Key issues and research priorities in landscape ecology: an idiosyncratic synthesis. *Landscape Ecology* 17: 355-365.

---

# Modelagem de dinâmica de paisagem: concepção e potencial de aplicação de modelos de simulação baseados em autômato celular

BRITALDO SILVEIRA SOARES FILHO\*  
GUSTAVO COUTINHO CERQUEIRA  
WILLIAM LEITE ARAÚJO  
ELIANE VOLL

Universidade Federal de Minas Gerais, Minas Gerais, Brasil.  
\* e-mail: britaldo@csr.ufmg.br

## RESUMO

Modelos de simulação recentemente se tornaram campo de pesquisa promissor, recebendo atenção por parte de pesquisadores de diversas áreas. Uma classe especial de modelos de simulação é representada pelos modelos espaciais ou modelos de paisagem, que simulam mudanças dos atributos do meio ambiente através do território geográfico. O uso de tais modelos visa auxiliar o entendimento dos mecanismos causais e dos processos de desenvolvimento de sistemas ambientais, e assim determinar como eles evoluem diante de um conjunto de circunstâncias. Neste artigo, são delineados os principais passos de desenvolvimento de um modelo de simulação espacial. A ênfase é dada à aplicação de autômatos celulares para replicação dos padrões espaciais de mudanças. Como exemplo, é apresentado o modelo DINAMICA, enfocando a suas funções de transição. Como potencial de aplicação, são descritos três estudos de caso, a saber: 1) a avaliação da fragmentação da paisagem florestal em função da arquitetura de projetos de colonização; 2) desenho de corredores de conservação; e 3) replicação de padrões de processos de difusão. O uso do DINAMICA em tais modelos visa demonstrar o seu potencial de adaptação para simular diversos fenômenos espaciais. Espera-se, para um futuro próximo, que modelos de simulação espacial deixem de ser apenas um objeto de pesquisa acadêmica, para se tornarem de fato um instrumento indispensável ao suporte de decisões sobre ordenamento territorial.

## ABSTRACT

*Simulation models have become a promising field of research, drawing the attention of a diversity of researchers. Spatially explicit simulation models represent a particular class of models that attempt to replicate the changes of ecological attributes across the landscape. These models are useful to help us understand the ways an environmental system evolves, and thereby determine the possible system trajectory under a set of circumstances. In this paper, we outline the main*

*steps towards the development of a simulation model. Emphasis is given to the application of cellular automata to replicate the spatial patterns of change. As an example of cellular automata, we introduce the DINAMICA model, focusing on the design of its transition functions. To highlight DINAMICA's potential, we describe its application in three case studies: 1) assessment of forest fragmentation as a function of types of road network; 2) design of conservation corridors; and 3) replication of patterns generated from diffusion processes. These different applications aim to demonstrate the potential of DINAMICA to be easily adapted to simulate various dynamic spatial phenomena. As a result, we expect for the near future that simulation models will become, not only a research topic, but also an essential tool for environmental planning.*

## POR QUE MODELOS DE SIMULAÇÃO?

Em um processo de simulação, é modelada a dinâmica de um sistema, reproduzindo-se, em ambiente computacional, a complexidade de seu mecanismo de desenvolvimento, que opera através de processos de troca de materiais, energia, informação e espécies/estados entre os componentes ou elementos do sistema.

Modelos de simulação se tornaram recentemente um campo de pesquisa promissor, recebendo especial atenção por parte de pesquisadores de diversas áreas. A importância da simulação advém de sua capacidade de multiplicar a nossa imaginação individual, permitindo que grupos de pessoas compartilhem, através de uma experiência comum, modelos mentais de uma certa realidade, independente de sua complexidade. A simulação é vista como um instrumento de aquisição de conhecimento, a partir da integração da informação sobre sistemas dinâmicos. Neste ponto, Lévy (1998) destaca que, dentre os gêneros aportados pela cibercultura, a simulação ocupa um lugar central, constituindo um importante dispositivo heurístico, que só se tornou viável graças ao recente desenvolvimento científico, tanto na área de informática, como também na modelagem de sistemas.

Uma classe especial de modelos de simulação é representada pelos modelos espaciais ou modelos de paisagem, que simulam mudanças dos atributos do meio ambiente através do território geográfico. O uso de tais modelos visa auxiliar o entendimento dos mecanismos causais e processos de desenvolvimento de sistemas ambientais, e assim determinar como eles evoluem diante de um conjunto de circunstâncias, que são as condições circunjacentes ou de contorno e representam cenários traduzidos por diferentes quadros socioeconômicos, políticos e ambientais. À luz do resultado do modelo, podem-se testar hipóteses sobre possíveis trajetórias do sistema e suas implicações ambientais.

Hoje, já se encontra disponível uma miríade de modelos espacialmente explícitos, aplicáveis a uma diversidade de áreas, tais como difusão de epidemias, dinâmica populacional, mudanças de uso do solo, dinâmica florestal e propagação de fogo. Para uma revisão, veja Baker (1989), Sklar & Constanza (1991) e Soares-Filho *et al.* (2002).

Dentro dos seus diversos propósitos, modelos espaciais de simulação têm se tornado um importante instrumento de auxílio ao ordenamento territorial, considerando que qualquer planejamento regional ou estratégia de conservação ambiental que não leve em conta o fator tempo está fadado ao insucesso. Além disso, modelos de simulação podem ser usados para promover idéias comuns, como no exemplo das mudanças globais e suas conseqüências ao Planeta. Portanto, modelos espaciais de simulação, que realisticamente reproduzem padrões espaciais de mudanças, são, hoje, requisito para o entendimento e avaliação de complexas questões ambientais em escalas local, regional e global.

## CONCEPÇÃO E DESENVOLVIMENTO DE MODELOS DE SIMULAÇÃO DE MUDANÇAS DA PAISAGEM

A comunidade científica vem cada vez mais buscando respostas e extrapolações dos resultados de suas pesquisas através do espaço geográfico. Um raciocínio lógico nessa linha de pesquisa baseia-se na assertiva de que os processos naturais ou antrópicos, quer sejam eles físicos, ecológicos, técnico-culturais ou de outro tipo qualquer, são, em certo grau, controlados pela organização espacial de seu conjunto ambiental. Tal organização espacial é chamada de paisagem. O objeto de pesquisa em paisagem é o estudo de sua trama espacial e seu papel em relação aos processos ambientais (Gulinck *et al.*, 1993). Nesses termos, uma

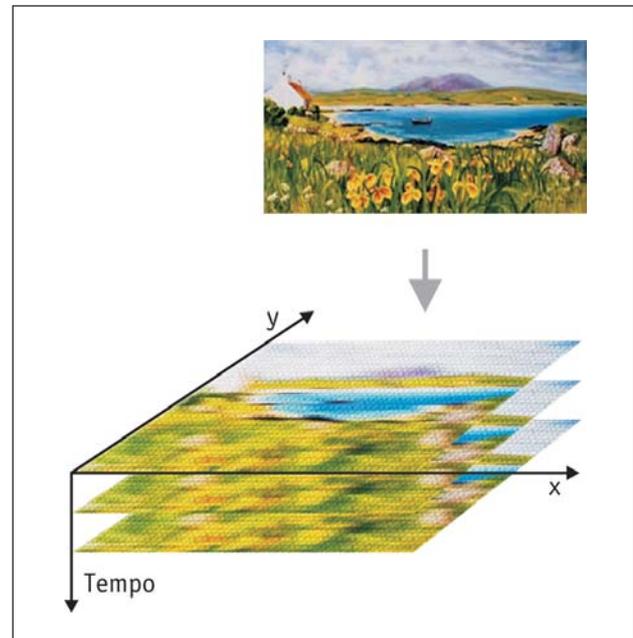
paisagem se encontra em perpétua mutação, podendo a sua estrutura e composição mudar drasticamente através do tempo e implicar em importantes ramificações quanto ao arranjo e interação de seus ecossistemas (Forman & Godron, 1986). Em efeito, o estudo de Paisagem substanciou o desenvolvimento de uma nova classe de modelos de simulação, os modelos espacialmente explícitos ou de paisagem, que abrangem, em geral, quadros espaciais na ordem de dezenas de quilômetros ou mais, como extensões de municípios, e intervalos temporais na ordem de anos a dezenas de anos.

Uma maneira conveniente de se pensar sobre um modelo espacial consiste em imaginar uma paisagem composta por uma variedade de arranjos de elementos de paisagem – os quais representam ecossistemas ou classes de cobertura e uso do solo – sobrepostos a uma grade ou matriz, sistema de representação 2D conhecido como *raster* ou matricial. Elementos de paisagem passam a ser representados por *pixels* ou células com o mesmo valor em uma única camada de informação (Figura 1).

Uma questão pertinente a esta representação refere-se a escolha do tamanho da célula ou *pixel* (grão da paisagem) em função da escala de observação. Para tanto, busca-se uma solução de compromisso, definindo-se um tamanho que atenda uma certa exatidão cartográfica, que capture a frequência espacial do fenômeno modelado e ainda permita um desejável desempenho computacional. Na representação matricial, a forma, configuração e tamanho dos elementos de paisagens não são modelados explicitamente, mas sim construídos para cada mancha da paisagem, que se faz representar por um conjunto de células contíguas de mesmo valor ou atributo.

Há algumas vantagens em representar uma paisagem, usando-se de uma construção matricial ou *raster*. A utilização de um SIG (Sistema de Informações Geográficas) matricial possibilita uma maior flexibilidade na representação e tratamento do contínuo espacial; a modelagem dos processos ambientais se torna mais direta através de um modelo de dados em camadas (cf. Burrough, 1991); os SIGs matriciais são inteiramente compatíveis com imagem de satélite, grande fonte de dados para a modelagem de mudança de paisagem; existem várias pesquisas e programas desenvolvidos que trabalham com dados no formato *raster* para análise dos padrões da estrutura da paisagem (cf. Baker & Cai, 1992; McGarigal & Marks, 1995); por último, a representação matricial para um processo de simulação é mais fácil de lidar matematicamente, porquanto cada célula unitária do modelo possui o mesmo tamanho e forma,

o que resulta numa maior capacidade analítica – por exemplo, usando-se de álgebra cartográfica (cf. Tomlin, 1990) – além de tender a ser mais rápida para avaliações de múltiplas camadas cartográficas.



**FIGURA 1** – Abstração da paisagem em um modelo espacial de mudanças, usando-se a estrutura *raster*.

O desenvolvimento de um modelo de simulação de mudanças da paisagem passa pela solução da equação 1, a qual representa a função que descreve as mudanças nos padrões espaciais de um tempo  $t$  para um novo padrão espacial no tempo  $t+v$ :

$$X_{t+v} = f(X_t, Y_t) \quad \text{Equação 1}$$

onde  $X_t$  é o padrão espacial no tempo  $t$ , e  $Y_t$  é o vetor ou conjunto escalar de variáveis que afetam a transição, no caso, variáveis com coordenadas espaciais.

Desse modo, os componentes básicos de qualquer modelo de mudança da paisagem são: (1) configuração inicial; (2) função de mudança; e (3) configuração de saída. A configuração inicial consiste em um mapa da paisagem, o qual pode ser derivado de uma série de dados, incluindo dados históricos de uso do solo, obtidos via sensoriamento remoto.

A função de mudança pode usar ambas as matemáticas contínuas ou discretas, envolvendo desde uma equação linear diferencial, no caso do tempo ser considerado como uma variável contínua, ou, ao contrário, uma

equação de diferença, até exemplos mais complexos, como a utilização de seqüências de regras de decisão, conjunto complexo de equações não-lineares, com interações recursivas ou não (Baker, 1989). Quanto ao espaço de estados, muito embora ele possa ser contínuo, em geral, assume-se uma representação discreta. Assim, à medida que o número de estados aumenta, a capacidade computacional e a necessidade de dados para esses modelos crescem consideravelmente.

A construção de uma função de mudança não é trivial, consistindo sempre na parte mais complexa do modelo. Na verdade, ela pode ser decomposta em duas funções, uma de quantificação das mudanças e uma outra dedicada à alocação espacial das mudanças, lembrando sempre que as mudanças em uma paisagem não se dão ao acaso, mas sim condicionadas a uma estrutura espacial.

Seguindo este raciocínio, um primeiro passo na construção de um modelo de simulação de mudanças passa pela definição de um modelo conceitual, no qual são definidos os elementos ou estados de uma paisagem e suas transições possíveis (Figura 2). Uma maneira simples de representar esse modelo de estado e transições dá-se pelo uso da matriz de transição (Equação 2), que, por si só, não consiste em um modelo espacial, mas na distribuição da quantidade de mudanças por toda a paisagem, ou seja, o primeiro componente da função de mudança.

$$\begin{bmatrix} 1 \\ 2 \\ \cdot \\ j \end{bmatrix}_{t=v} = \begin{bmatrix} P_{11} & P_{21} & P_{\cdot 1} & P_{j1} \\ P_{12} & P_{22} & P_{\cdot 2} & P_{j2} \\ P_{1\cdot} & P_{2\cdot} & P_{\cdot\cdot} & P_{j\cdot} \\ P_{1j} & P_{2j} & P_{\cdot j} & P_{jj} \end{bmatrix}^v * \begin{bmatrix} 1 \\ 2 \\ \cdot \\ j \end{bmatrix}_{t=0} \quad \text{Equação 2}$$

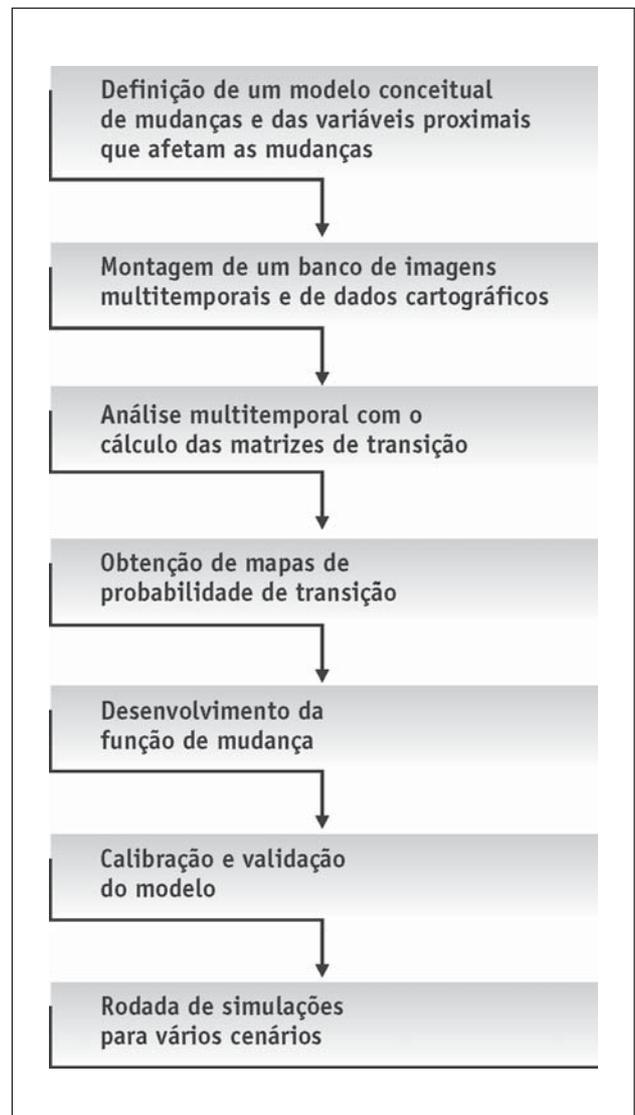
A matriz  $P_{ij}$  é conhecida como matriz de transição, na qual as colunas representam as probabilidades de um determinado estado  $i$  permanecer no mesmo estado ou mudar para o estado  $j$  durante o intervalo de tempo de  $t \rightarrow t + v$ , de modo que:

$$\sum_{i=1}^n P_{ij} = 1, i = 1, 2, \dots, n. \quad \text{Equação 3}$$

sendo o estimador de  $P_{ij}$  dado pela Equação 4, onde  $n$  é o número de estados na paisagem.

$$\hat{P}_{ij} = \frac{n_{ij}}{\sum_{j=1}^n n_{ij}} \quad \text{Equação 4}$$

Um modelo de transição, como definido anteriormente, requer apenas que seja especificado um número finito de estados e que sejam conhecidas as probabilidades, ou melhor dizendo, as taxas de transição de um estado para outro em um intervalo de tempo discreto. Um caso específico de modelo de transição, no qual as taxas de transição são estáticas através do tempo, é conhecido como modelo Markoviano. Tal tipo de modelo não é realístico, sendo limitado o seu uso para simular a dinâmica de fenômenos ambientais, a não ser por curto intervalo de tempo. Portanto, modelos atuais tendem a usar matrizes dinâmicas, nas quais as probabilidades de transição são recalculadas a cada nova iteração por submodelos acoplados (cf. Soares-Filho *et al.*, 2004).



**FIGURA 2** – Fluxograma de desenvolvimento e aplicação de um modelo de simulação.

De acordo com a Equação 1, a solução de modelos espaciais dinâmicos requer a definição dos processos no tempo  $t$ , que controlam o desenvolvimento do padrão espacial para o tempo  $t+v$ . Nesses termos, a construção de um modelo de simulação visa obter medidas de alguns aspectos do estado atual da paisagem que possam fornecer indicações do seu potencial para se tornar algo diferente, diante de diversos conjuntos de condições. Portanto, um primeiro passo na construção de um modelo de dinâmica de paisagem consiste no mapeamento dos padrões mutáveis de uso e cobertura do solo da região escolhida para estudo. Para tanto, são necessários métodos de levantamento que possibilitem identificar as mudanças na paisagem. Para este propósito, podem ser utilizados, como fonte de dados, levantamentos de campo, mapas previamente existentes, fotografias aéreas, dados publicados por censos e sensoriamento remoto orbital. Dentre esses métodos, o desenvolvimento e aplicação de técnicas de sensoriamento remoto orbital para o monitoramento de fenômenos e processos ambientais têm recebido uma especial atenção. O sensoriamento remoto entra como a principal fonte de dados para o desenvolvimento, calibração e refinamento de modelos de dinâmica de paisagem.

A aplicação do sensoriamento remoto envolve técnicas de interpretação e classificação de imagens multitemporais e a subsequente análise dos mapas resultantes, através de tabulação cruzada, visando a detecção e quantificação de mudanças. Um dos resultados desse processo consiste na matriz de transição para o período analisado. Como algumas vezes pretende-se rodar o modelo com uma maior resolução temporal, por exemplo, em passos anuais, a matriz do período deve ser ainda anualizada através da seguinte propriedade de matrizes ergódicas (Equação 5), as quais convergem para uma distribuição estacionária.

$$P^{1/t} = HV^{1/t}H^{-1} \quad \text{Equação 5}$$

onde  $P$  é a matriz de transição,  $H$  e  $V$  são seus autovetores e autovalores e  $1/t$  é a fração desejada do intervalo de tempo. No exemplo de cinco anos, a fração anual corresponderia a  $1/t = 0,2$

Modelos de simulação requerem como entrada, substanciais dados cartográficos, os quais serão armazenados em conjunto com os mapas multitemporais em um Sistema de Informações Geográficas. Nesse ambiente, técnicas de análise espaciais serão usadas para quantificar os efeitos espaciais das variáveis que afetam as mudanças.

Nesse caso, deve-se diferenciar entre variáveis proximais e causais. As variáveis proximais explicam as influências regionais na configuração espacial dos padrões de mudanças, no exemplo de distâncias às estradas e às cidades, fisiografia e divisão político-administrativa. A análise do efeito dessas variáveis resulta no mapa de favorabilidade ou probabilidade de mudança. Já as variáveis causais explicam as forças subjacentes ou motivadoras das mudanças, sendo muitas vezes consideradas como variáveis de contorno ou exógenas ao modelo espacial, mas que podem ser eventualmente integradas ao modelo espacial através de um modelo gerador de cenários, o qual processa a interação de subsistemas encadeados – *e.g.*, modelos climático, demográfico, econômico e político-econômico – e seus efeitos na dinâmica do sistema estudado (Soares-Filho *et al.*, 2004). Nesse aspecto, modelos de paisagem têm se tornado objeto de estudo multidisciplinar, envolvendo freqüentemente a participação de especialistas em diversas áreas, haja vista que eles devem considerar as interações entre o complexo de atividades humanas e os sistemas ecológicos.

A discretização de modelos de simulação de paisagem em subunidades de área passa pela seleção de variáveis proximais e pelo cálculo dos seus efeitos nas mudanças. A saída dessa análise consiste no mapa de favorabilidade ou probabilidade de mudança. Diversos métodos têm sido aplicados para tal propósito, sendo, em geral, caracterizados como técnicas dirigidas aos dados, notadamente regressão logística (Ludeke *et al.*, 1990; Turner *et al.*, 1996a, b; Mertens & Lambin, 2000; Soares-Filho *et al.*, 2001) e pesos de evidência (Almeida *et al.*, 2002, 2003; Soares-Filho *et al.*, 2004).

Pesos de evidência (*weights of evidence*) consistem em um método bayesiano, tradicionalmente usado por geólogos para indicar áreas favoráveis para algum fenômeno geológico, ex. mineralização ou sismicidade (Goodacre *et al.*, 1993; Bonham-Carter, 1994). O peso de evidência representa a influência de cada categoria (faixa de valores) de certa variável nas probabilidades espaciais de uma transição  $i \Rightarrow j$ , sendo calculado pelas equações 6 e 7:

$$O\{D/B\} = O\{D\} \frac{P\{B/D\}}{P\{B/\bar{D}\}} \quad \text{Equação 6}$$

$$\log\{D/B\} = \log\{D\} + W^+ \quad \text{Equação 7}$$

onde  $O\{D\}$  e  $O\{D/B\}$  são as razões de chances, respectivamente, de ocorrer *a priori* o evento  $D$  e de ocorrer  $D$

dado um padrão espacial  $B$ , no exemplo de um determinado tipo de solo ou faixa de distância à estrada.  $W^+$  é o peso de evidência de ocorrer o evento  $D$ , dado um padrão espacial  $B$ . A probabilidade *a posteriori* de uma transição  $i \Rightarrow j$ , dado um conjunto de dados espaciais e considerando que  $O\{D\}=1$ , pois esta já é passada ao modelo via matriz de transição, é expressa pela seguinte equação:

$$P(i \Rightarrow j(x, y)/V) = \frac{e^{\sum_k W_{kn_i \Rightarrow j(V), xy}}}{1 + \sum_{ij} e^{\sum_k W_{kn_i \Rightarrow j(V), xy}}} \quad \text{Equação 8}$$

onde  $V$  representa um vetor de  $k$  variáveis espaciais, medidas nas localidades  $x, y$  e representadas por seus pesos  $W^+_{k1xy}, W^+_{k2xy}, \dots, W^+_{knxy}$ , sendo  $n$  o número de categorias de cada variável  $k$ .

Como vantagem do método de pesos de evidência, cita-se que ele não é restringido pelas clássicas suposições dos métodos estatísticos paramétricos, os quais os dados espaciais freqüentemente violam. Além disso, é um método simples de se calcular, pois usa somente o resultado da tabulação cruzada entre o mapa de mudanças – produto do cruzamento de mapas multitemporais – e os mapas das variáveis proximais para alimentar fórmulas implementadas em planilhas eletrônicas (Figura 3). O efeito de cada variável pode ser calculado independentemente de uma solução conjunta, tendo como premissa apenas que as variáveis de entrada sejam espacialmente independentes, o que pode ser testado por diversos métodos, como coeficiente de  $V$  de Crammer (Bonham-Carter, 1994) ou o teste da “Incerteza da Informação Conjunta” – *Joint Information Uncertainty* (Almeida *et al.*, 2002). Variáveis correlacionadas são excluídas do modelo ou combinadas em uma terceira. Como o método usa dados categóricos, os pesos são também independentes de unidades de medida. No entanto, isto impõe a necessidade de um método de categorização que mantenha a estrutura original de variação dos dados.

Pesos de evidência são então assinalados para as categorias das variáveis representadas por suas camadas de informação no SIG (Figura 3). Os mapas resultantes da integração dos pesos assinalam as áreas mais favoráveis para cada tipo de mudança, podendo, portanto, ser interpretados como mapas de probabilidades espaciais de transição.

Em síntese, embora existam diversos métodos para se obter o mapa de favorabilidade de mudanças, sendo

aqui apresentado apenas o método de pesos de evidência, deve-se ressaltar que, mais do que a matemática utilizada, um método deve se pautar por ser robusto (independente de pressupostos), simples de ser implementado e transportado, sobretudo por fornecer uma explicação plausível consoante ao conhecimento do fenômeno analisado.

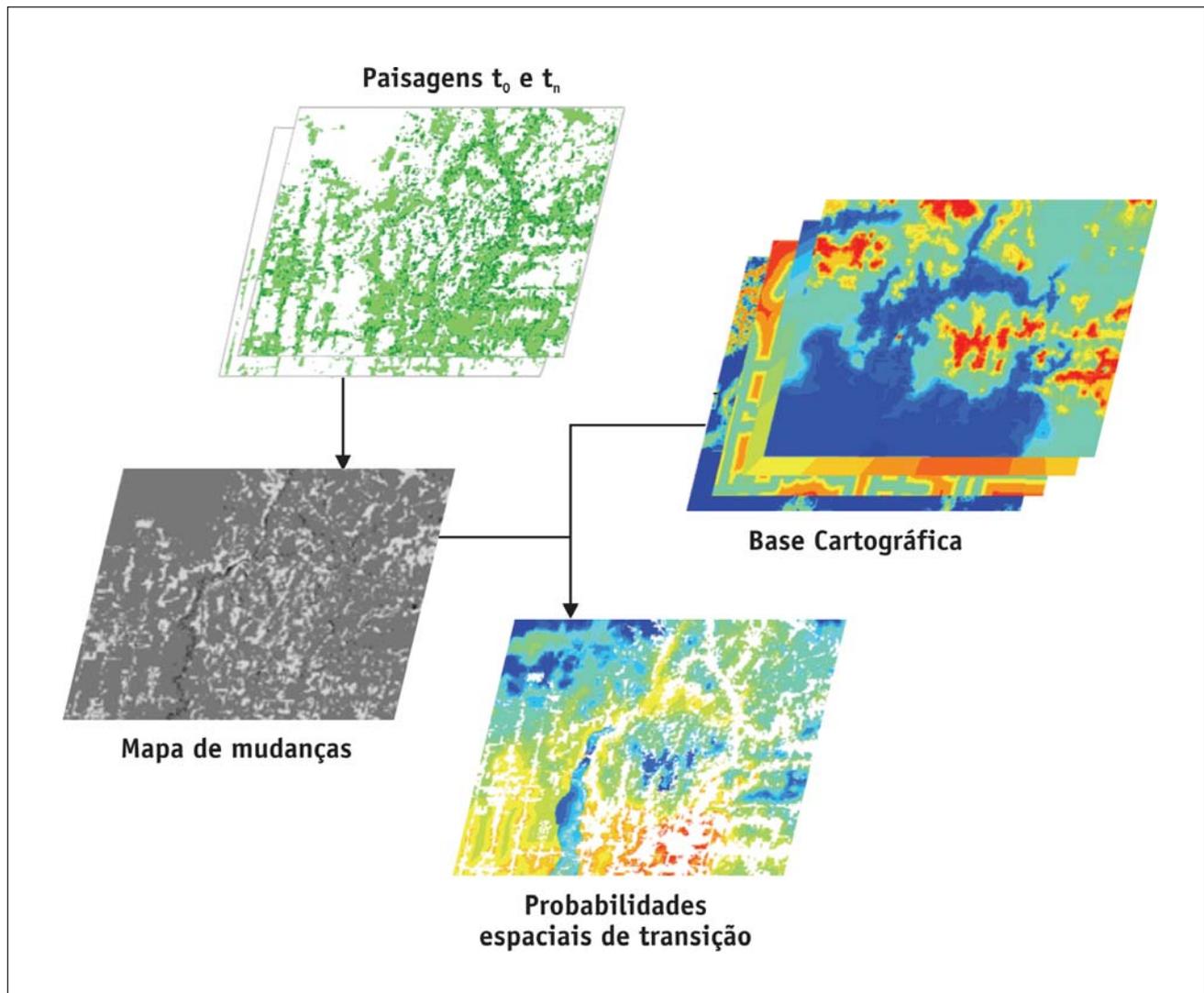
Em seqüência, o segundo componente da função de mudança irá operar sobre os mapas de probabilidades, buscando alocar as quantidades desejadas de mudança, através do ordenamento e sorteio das células mais prováveis. Como as mudanças na paisagem são fortemente influenciadas pela vizinhança, tendendo a formar padrões espaciais com distinto arranjo no tocante ao tamanho, forma e orientação das manchas, a função de alocação das mudanças deverá envolver algoritmos de interação da célula sorteada com as vizinhas. Isto será mostrado adiante pelo uso de técnicas de autômatos celulares.

Por fim, o processo de calibração deverá levar em conta o ajuste do modelo em relação a dois aspectos: (1) configuração espacial e (2) estrutura da paisagem. Para a configuração, são sugeridos métodos de comparação de imagens a múltiplas resoluções (*e.g.*, Hagen, 2003), tendo em vista que modelos com baixo ajuste numa resolução mais fina podem convergir dentro de uma certa vizinhança. Já para a comparação da estrutura da paisagem, são sugeridos os índices de dimensão fractal e medidas de conectividade da paisagem, tais como o índice de coesão de manchas – *patch cohesion index* (Mcgarigal & Marks, 1995).

O modelo de simulação, obtendo um razoável ajuste – algo acima de 75%, considerando o caráter falível das predições ambientais – estará validado para o caso de estudo específico. Por conseguinte, ele poderá ser extrapolado para outras áreas e casos semelhantes, ou mesmo ser usado para ilustrar ou testar diferentes hipóteses de cenário, alternando-se os parâmetros das variáveis de contorno.

## AUTÔMATOS CELULARES

Embora modelos de simulação espaciais estejam se tornando por demais híbridos para serem classificados em uma única categoria, a não ser por seu propósito de aplicação, uma abordagem de modelagem, em geral, fundamenta-se em três paradigmas: (1) baseado em indivíduos (*e.g.*, Liu & Ashton, 1998); (2) baseado em processos (*e.g.*, Voinov *et al.*, 1999); e (3) abordagem orientada ao espaço de autômatos celulares.



**FIGURA 3** – A seleção das variáveis proximais e o cálculo de seus efeitos nas mudanças são realizados através do cruzamento de variáveis armazenadas no SIG com o mapa de mudanças, produto da tabulação cruzada de mapas multitemporais.

A terceira categoria de modelos, conhecida como autômatos celulares, corresponde a engenhos compostos por conjuntos de células que imprimem determinados movimentos, assemelhando-se a seres animados. Em geral, o mecanismo de autômatos celulares consiste num arranjo  $n$ -dimensional de células, no qual o estado de cada célula depende de seu estado prévio e de um conjunto de regras de transição, de acordo com um arranjo específico de uma certa vizinhança, sendo todas as células atualizadas simultaneamente a passos discretos de tempo. O algoritmo que faz as células interagir localmente é conhecido como regra local do autômato celular (Sirakoulis *et al.*, 2000).

Autômatos celulares são, portanto, considerados como sistema dinâmico, tendo sido desenvolvidos para aplicação em diversas áreas, como arquitetura, engenharia,

matemática, simulação e jogos. Talvez o exemplo mais simples e característico de autômato celular seja o “jogo da vida”. Tendo sido desenvolvido pelo Matemático John Conway, em 1970, foi assim denominado devido à analogia com ascensão, queda e alternância de sociedades de organismos. John Conway escolheu, como regras locais de transição, arranjos espaciais que evitassem que colônias de células morressem ou expandissem rapidamente. Dessa forma, ele pertence à classe de jogos de simulação que se assemelham a processos da vida real, sendo um belo exemplo de sistemas auto-organizáveis.

Modelos de simulação espacial baseados em autômatos celulares estão se tornando populares, sendo mesmo incluídos em softwares comerciais de Sistemas de Informações Geográficas, *e.g.*, IDRISI (Eastman, 2001),

devido à sua facilidade de implementação, habilidade em mimetizar formas e capacidade de serem readaptados para reproduzir vários tipos de fenômenos espaciais, como espalhamento de fogo (Karafyllidis & Thanailakis, 1997; Hargrove *et al.*, 2000), difusão de epidemia (Sirakoulis *et al.*, 2000), dinâmica florestal (Lett *et al.*, 1999), crescimento urbano (Clarke & Gaydos, 1998; White *et al.*, 2000b; Almeida *et al.*, 2003), e mudanças de uso e cobertura do solo (Wu, 1998; White *et al.*, 2000a; Soares-Filho *et al.*, 2002; Soares-Filho *et al.*, 2004).

Como exemplo de modelo de simulação espacial do tipo autômato celular é apresentado o software DINAMICA (Soares-Filho *et al.*, 2002, Soares-Filho *et al.*, 2003). DINAMICA usa como entrada um conjunto de mapas, a saber: um mapa da paisagem inicial – por exemplo, um mapa de uso e cobertura do solo, obtido a partir da classificação de imagens de satélite –; um mapa do tempo de permanência de cada célula no seu estado atual; e um conjunto de variáveis cartográficas, que se dividem em dois tipos, estáticas e dinâmicas, sendo as últimas recalculadas em cada iteração do programa. As variáveis cartográficas – *e.g.*, solo, vegetação, distância às estradas, altitude, declividade, etc. – são combinadas, através da definição de seus pesos de evidência, para gerar os mapas de probabilidades de transição. Após cada iteração, DINAMICA produz um novo mapa da paisagem, os mapas de probabilidades de transição e mapas das variáveis dinâmicas. Seu modelo de transição pode ser ainda acoplado a um módulo

construtor de estradas, que passa ao programa mapas dinâmicos da rede viária, e a um gerador de cenários que produz matrizes de transição dinâmicas e as passa ao DINAMICA, usando-se de sua ligação com o software de modelagem VENSIM (Ventana, 2002) (Figura 4).

Como regras locais, DINAMICA usa um engenho de alocação de mudanças, composto por dois processos de transição complementares, denominados, respectivamente, de *Expander* (função expansora) e *Patcher* (função formadora de manchas), além de um módulo de difusão. A função expansora se dedica unicamente à expansão ou contração de manchas de uma determinada classe, já o segundo processo é responsável por formar novas manchas. Ambos os processos usam um mecanismo nucleador de manchas, que opera sobre o mapa de probabilidades de transição e tem como parâmetros de entrada a isometria, a variância e o tamanho médio das manchas. A quantidade definida de células a serem mudadas para uma determinada transição é repartida, de acordo com o especificado pelo usuário, entre as duas funções. Já o módulo de difusão leva um processo de transição  $i \Rightarrow j$  a migrar para novas áreas, em função de uma saturação assintótica de células  $j$  dentro de uma vizinhança especificada.

A combinação dos dois processos de transição, mais o módulo de difusão, apresenta numerosas possibilidades no tocante à geração e desenvolvimento de padrões espaciais de mudança, como será demonstrado pelo seu potencial de aplicação (*cf.* Soares-Filho *et al.*, 2004).

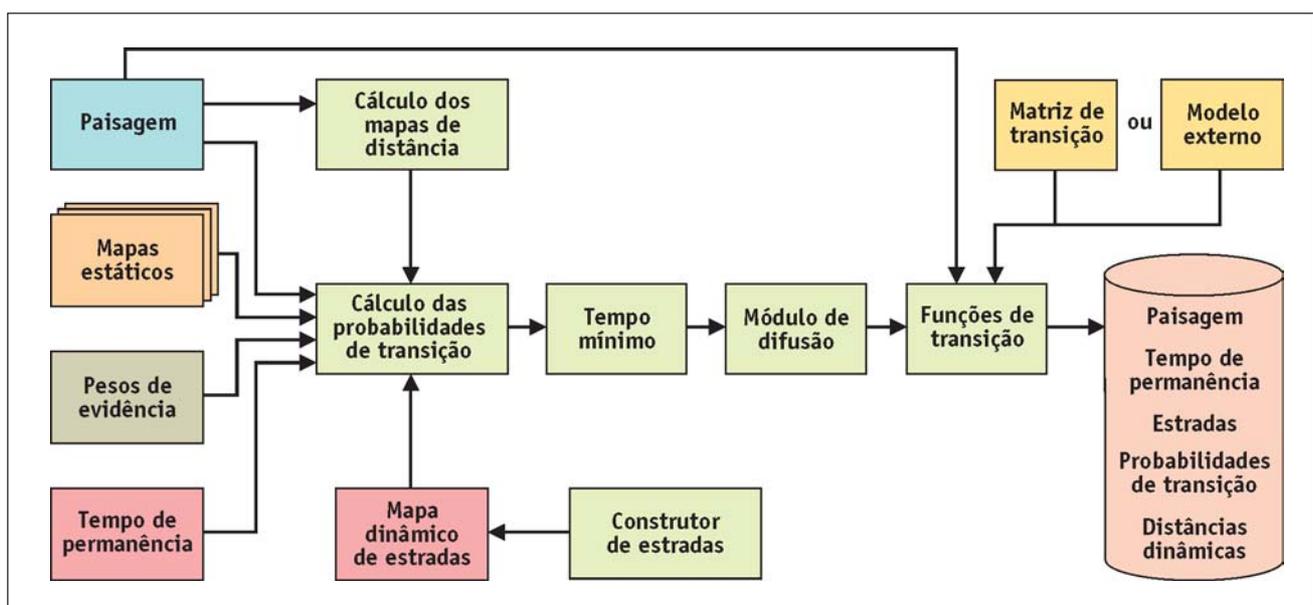


FIGURA 4 – Arquitetura do software DINAMICA.

## EXEMPLOS DE APLICAÇÃO

DINAMICA tem sido desenvolvido como um instrumento de investigação de trajetórias de paisagem e de dinâmica de fenômenos espaciais. Por ser um modelo genérico de mudanças, DINAMICA tem sido aplicado para modelar vários fenômenos dinâmicos, como desflorestamento (Soares-Filho *et al.*, 2002) e dinâmica urbana (Almeida *et al.*, 2002, 2003). Recentes avanços no seu desenvolvimento permitiram também a sua aplicação para modelagem de extensas regiões geográficas, como no exemplo de uma faixa de centenas de quilômetros de largura e de extensão, ao longo da BR-163, conhecida como corredor Cuiabá-Santarém (Soares-Filho *et al.*, 2004).

Como exemplos de seu potencial de aplicação, são ilustrados três estudos de caso, abordando: 1) avaliação da fragmentação da paisagem florestal em função da arquitetura de projetos de colonização; 2) desenho de corredores de conservação; e 3) replicação de padrões de processos de difusão.

O primeiro estudo consiste numa comparação da evolução da fragmentação da floresta remanescente em função do desenho da malha viária de projetos de colonização. A exemplo dos trabalhos de Batistella & Soares-Filho (1999) e Batistella (2001), são comparadas as regiões do projeto Machadinho d'Oeste e do Vale do Anari, em Rondônia. A primeira região apresenta uma

arquitetura orgânica, com as estradas acompanhando os interflúvios, enquanto a segunda caracteriza a típica arquitetura de “espinha de peixe”, freqüentemente encontrada através da Amazônia. Como as regiões desses projetos diferem também no tocante a outros aspectos, como história de ocupação, organização regional, fisiografia, taxa de desmatamento, tipo predominante e densidade de colonização, a simulação foi aplicada com o intuito de excluir da análise essas outras variações, criando-se uma situação “ideal” – a exemplo dos modelos neutros (Gardner *et al.*, 1987) –, na qual o único parâmetro variante é o arranjo geométrico da malha viária. Assim, na simulação, foram escolhidas para as duas áreas a mesma extensão geográfica (350 km<sup>2</sup>), densidade viária equivalente e taxa de desmatamento idêntica, de 0,03 ao ano. A alocação do desmatamento ocorreu em ambas as áreas apenas em função da distância às estradas; buscou-se reproduzir um tamanho médio de 10 hectares para as novas clareiras anuais e o modelo foi rodado por um intervalo de 10 anos.

Como DINAMICA é um modelo estocástico, a simulação foi rodada dez vezes, e sobre as paisagens resultantes foi calculada a métrica do *índice de mancha mais larga* (*Largest Patch Index*). A média obtida para esse índice, para a classe de floresta, foi de 71% para a região tipo Machadinho d'Oeste, enquanto para a região de espinha de peixe, tipo Vale do Anari, foi de 35% (Figura 5).

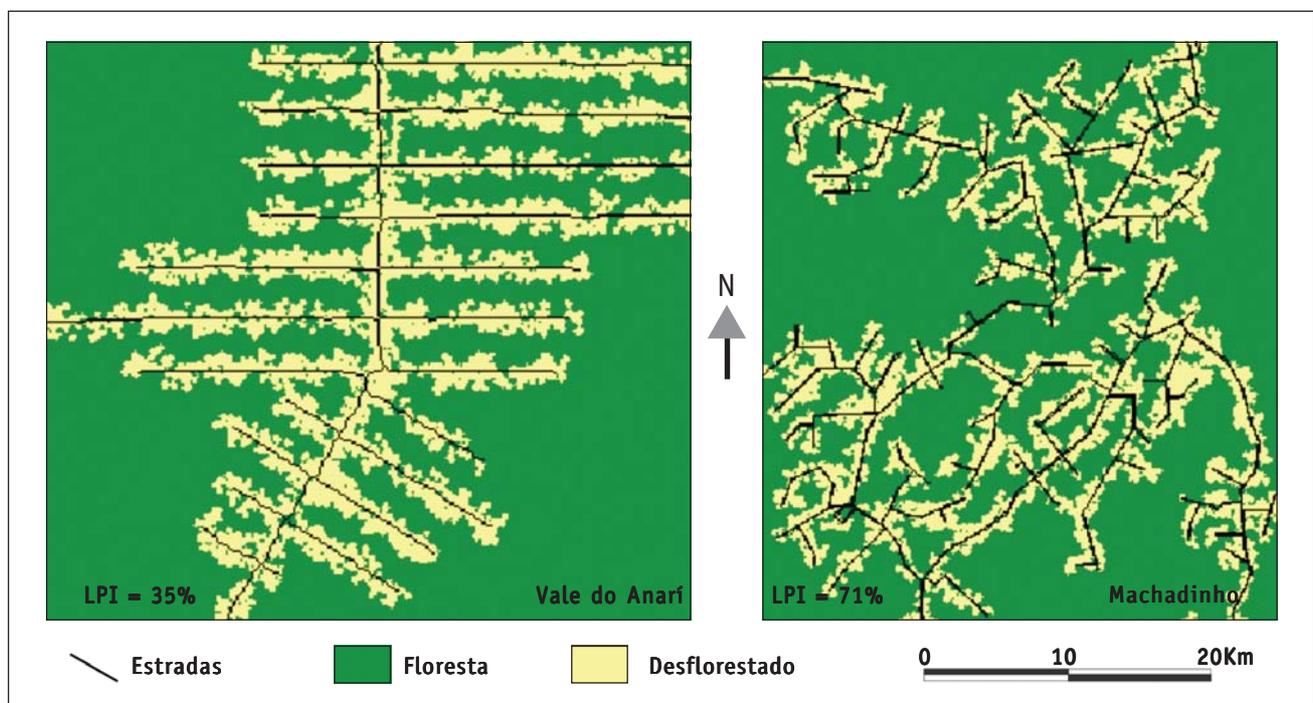


FIGURA 5 – Avaliação da fragmentação da paisagem em função da arquitetura de projetos de colonização.

Segundo Batistella (2001), o *índice de mancha mais larga* é uma potencial medida para avaliação da fragmentação da paisagem, sendo a fragmentação inversamente proporcional a esse índice. Em conclusão, o resultado da simulação demonstra uma clara vantagem da arquitetura tipo Machadinho d'Oeste para fins de melhor preservação da continuidade dos habitats florestais.

O segundo exemplo usa um recorte da região do município de Guarantã do Norte, Mato Grosso, para demonstrar o potencial da simulação para fins de seleção de áreas para restauração da paisagem florestal. Como ponto de partida, a simulação emprega um mapa de paisagem, obtido a partir da classificação de imagem TM/Landsat-5 de 1994. A paisagem simplificada é representada por quatro classes: floresta, desmatado, regeneração e corpo d'água. Na simulação, são especi-

ficadas taxas anuais de transição para que a dinâmica de paisagem convirja para um baixíssimo desmatamento, enquanto ocorre uma intensa regeneração florestal. O modelo busca então responder, à luz da importância da recuperação do ecossistema regional, qual será a configuração provável da paisagem em um tempo  $t+n$ , se a tendência atual for mantida; quais áreas desmatadas vão chegar a formar florestas secundárias; e, por conseguinte, quais serão as implicações ecológicas desse provável arranjo espacial.

O mapa de favorabilidade de regeneração florestal é construído a partir de estudos de campo e análise de dados de imagens multitemporais, a exemplo do trabalho de Soares-Filho *et al.*, (2001). Desse modo, as variáveis altitude, declividade, distância à rede hidrográfica, distância às estradas e distância à floresta, são selecionadas como sendo os fatores de controle espacial do

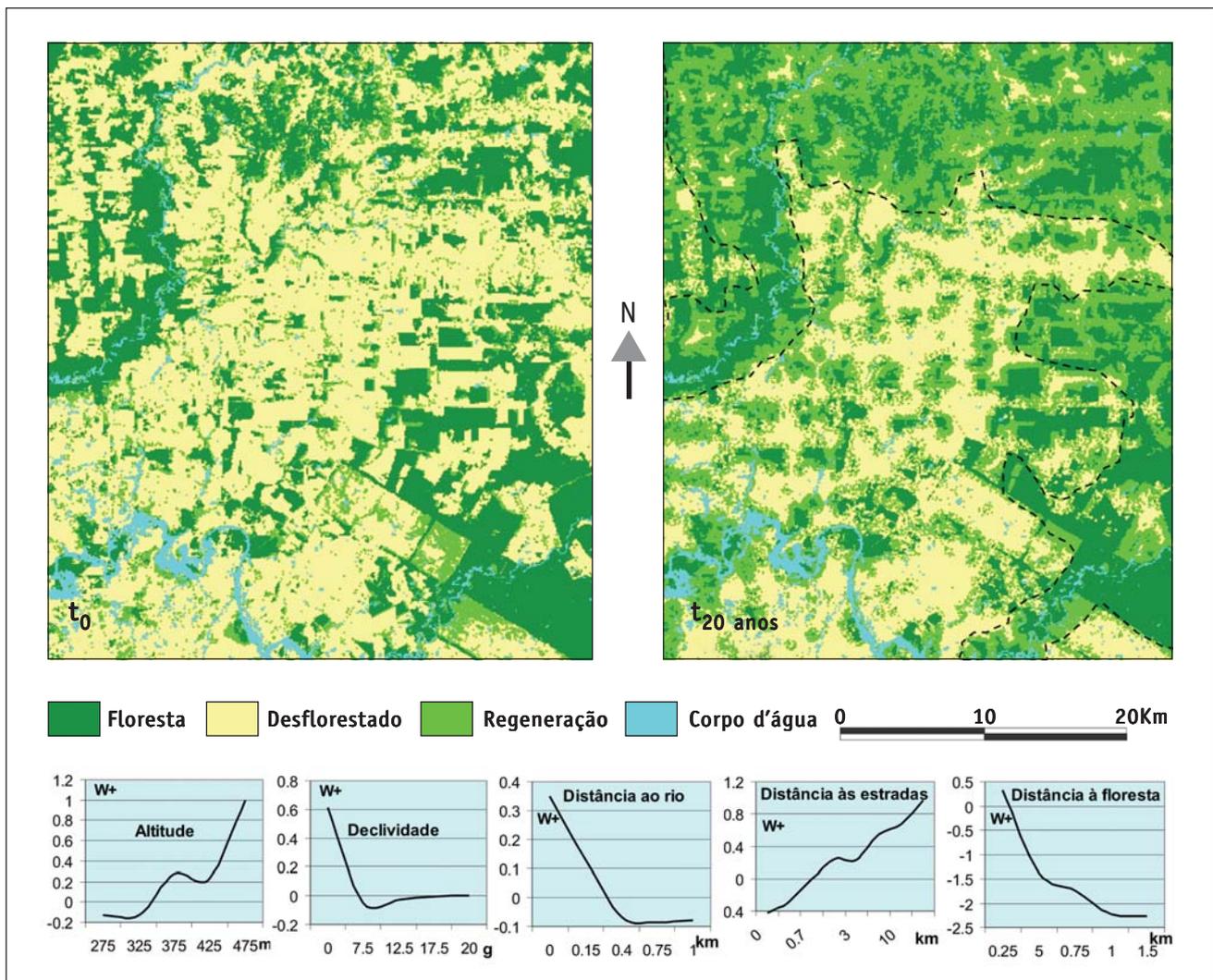


FIGURA 6 – Desenho de corredor de conservação.

processo de regeneração. Seus pesos de evidência são obtidos através do cruzamento de mapas de mudanças com os mapas dessas variáveis, conforme metodologia descrita na seção 2, e em seguida integrados para gerar o mapa de probabilidade de transição para a mudança “desmatado-regeneração”. O modelo é rodado em 20 passos anuais (Figura 6).

Como resultado, verifica-se que a porção nordeste da região apresenta o maior potencial para restauração florestal, gerando um arranjo de florestas secundárias, em torno das florestas remanescentes, que otimizam a conectividade dos habitats florestais. Em efeito, tal tipo de estudo poderia ser aplicado para seleção de áreas para a elaboração de corredores de conservação e restauração ambiental, como no exemplo de Landau (2001).

O terceiro caso de estudo ilustra o potencial do software DINAMICA em produzir padrões de difusão, que poderiam ser causados por processos, por exemplo, de propagação de fogo, epidemias ou migração de espécies. A Figura 7 mostra a geração de padrões de difusão a partir de uma paisagem sintética (a). O mapa (b) é produto de uma difusão por contágio (uso apenas do processo de expansão) em uma superfície isotrópica. O mapa (e) reproduz o mesmo processo, mas exibindo o efeito de saturação local, o qual inibe a continuação da transição  $i \Rightarrow j$  quando a presença de  $j$  atinge um certo limiar. O mapa (d) mostra processo análogo ao anterior, mas agora condicionado a um corredor de maior permeabilidade (c). Por fim, o mapa (f) mostra o resultado de dois processos de difusão competitivos, sendo que as células cinzas, com maior probabilidade, se superpõem às pretas.

Como ilustrado nos exemplos apresentados, a combinação das regras locais de transição do modelo DINAMICA apresenta numerosas possibilidades em relação à gênese e desenvolvimento de padrões espaciais de mudanças, o que o torna um potencial instrumento para a modelagem de dinâmica de paisagem.

## CONSIDERAÇÕES FINAIS

Modelos de simulação espacial já são uma realidade, embora exista ainda um longo caminho a percorrer antes que eles se tornem de fato mais realísticos e fáceis de operar. Uma grande limitação atual dessa classe de modelos refere-se à sua calibração, um processo extremamente laborioso, devido à quantidade de parâmetros a serem ajustados. Já o desempenho computacional vem crescendo exponencialmente, mais em função do aumento da velocidade dos processadores

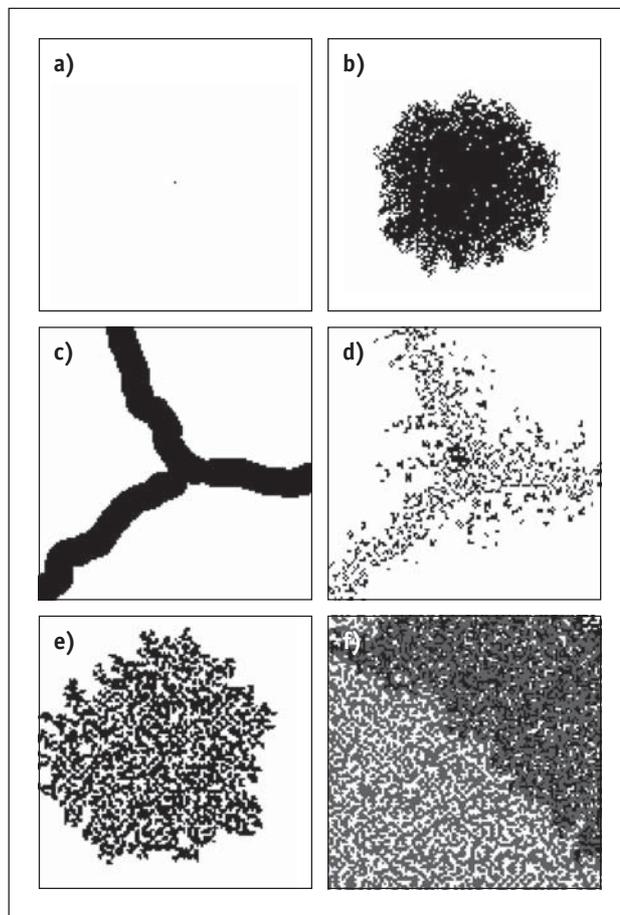


FIGURA 7 – Simulação de padrões de processos de difusão.

do que da otimização dos códigos. Este desempenho computacional já nos permite rodar modelos complexos, envolvendo arranjos matriciais acima de 100 megabytes, em computadores pessoais. Isto facilita em muito sua operação, pois evita a necessidade de supercomputadores, sistemas exóticos e difíceis de programar. A tendência será a incorporação, em breve, em SIGs comerciais, de módulos de simulação espacial, materializando o conceito do SIG dinâmico. Nesse sentido, a plataforma do DINAMICA oferece uma oportunidade única, pois seu caráter genérico, voltado à replicação de padrões espaciais, o torna fácil de ser adaptado para a simulação de diversos fenômenos ambientais, bastando apenas usar criatividade para abstrair o modelo conceitual para o ambiente do DINAMICA. Nesse sentido, deve-se ressaltar que qualquer modelo de simulação requer não só dados substanciais, mas uma investigação profunda e *in loco* dos processos controladores da dinâmica da paisagem em questão, que quase sempre envolve estudos multidisciplinares.

A arquitetura do DINAMICA possui também a versatilidade de operar em múltiplas escalas espaciais, tornando-o fácil de ser calibrado, pois seus modelos podem ser desenvolvidos usando-se de crescentes resoluções, haja vista que suas funções operam com valores em hectares, que são internamente transformados em números de células. Isto economiza um enorme tempo de computação, pois ele cresce geometricamente à dimensão da matriz de paisagem.

Espera-se que, com a maior difusão do geoprocessamento, modelos de simulação deixem de ser apenas um objeto de pesquisa acadêmica, para se tornarem de fato um instrumento indispensável ao suporte de decisões ao ordenamento territorial. É essa a derradeira meta da nossa pesquisa em modelagem de processos ambientais; por isso DINAMICA é um software de domínio público ([www.csr.ufmg.br/dinamica](http://www.csr.ufmg.br/dinamica)), havendo interesse por parte da nossa equipe em colaborar com potenciais usuários.

## AGRADECIMENTOS

Os autores agradecem a CABS/CI (Center for Applied Biodiversity Studies at Conservation International) e FAPEMIG (Fundação de Apoio à Pesquisa de Minas Gerais) – CRA2463/98 – pelo financiamento desta pesquisa. Primeiro autor é também bolsista da CAPES (Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) – BEX0438/02-2.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Almeida, C., A.M. Monteiro, G. Camara, B.S. Soares-Filho, G. Cerqueira, C. Pennachin & C. Batty. 2002. GIS and remote sensing as tools for simulation of urban land use change. In: Third International Symposium Remote Sensing of Urban Areas. D. Maktav & C. Jürgens (orgs). pp. 387-394. Istanbul Technical University and Regensburg University, Istanbul.
- Almeida, C., C. Batty, A.M. Monteiro, G. Camara, B.S. Soares-Filho, G. Cerqueira & C. Pennachin. 2003. Stochastic cellular automata modelling of urban land use dynamics: Empirical development and estimation. *Computers, Environment and Urban Systems* 27: 481-509.
- Baker, W.L. 1989. A review of models of landscape change. *Landscape Ecology* 2: 111-133.
- Baker, W.L. & Y. Cai. 1992. The r.le programs for multiscale analysis of landscape structure using GRASS geographical information system. *Landscape Ecology* 4: 291-302.
- Batistella, M. 2001. Landscape change and land-use/land-cover dynamics in Rondônia, Brazilian Amazon. Tese de Doutorado. School of Public and Environmental Affairs, Indiana University. 318p.
- Batistella, M. & B.S. Soares-Filho. 1999. Ensaio comparativo da fragmentação da paisagem em função de modelos arquitetônicos de assentamento rural em Rondônia. In: GIS-Brasil 1999. FatorGis, Cd-Rom, Salvador. Disponível em: <http://www.gisbrasil.com.br>.
- Bonham-Carter, G. 1994. *Geographic information systems for geoscientists: modelling with GIS*, Pergamon, New York. 414p.
- Burrough, P.A. 1991. *Principles of Geographical Information Systems for Land Resources Assessment*. 3rd edition. Clarendon Press, Oxford.
- Clarke, K.C. & L.J. Gaydos. 1998. Long term urban growth prediction using a cellular automaton model and GIS: Applications in San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science* 12: 699-714.
- Eastman, J. R. 2001. IDRISI 32.2 – guide to GIS and image processing. Clark Labs, Clark University, Worcester.
- Forman, R.T. & M. Godron. 1986. *Landscape Ecology*. John Wiley & Sons, New York.
- Gardner, R., H. Mime, T. Bruce, M. Turner & R. O'Neill. 1987. Neutral models for the analysis of broad-scale landscape patterns. *Landscape Ecology* 1: 19-28.
- Goodacre, C.M., G.F. Bonham-Carter, F.P. Agterberg & D.F. Wright. 1993. A statistical analysis of spatial association of seismicity with drainage patterns and magnetic anomalies in western Quebec. *Tectonophysics* 217: 205-305.
- Gulinck, H., O. Walpot & P. Janssens. 1993. Landscape structural analysis of central Belgium using Spot data. In: R. Haines-Young, D.R. Green & S. Cousins (eds.). *Landscape ecology and spatial information systems*. pp. 129-140. Taylor and Francis, Bristol.
- Hagen, A. 2003. Fuzzy set approach to assessing similarity of categorical maps. *International Journal of Geographical Information Science* 17: 235-249.
- Hargrove, W.W., R.H. Gardner, M.G. Turner, W.H. Romme & D.G. Despain. 2000. Simulating fire patterns in heterogeneous landscapes. *Ecological Modelling* 135: 243-263.
- Karafyllidis, I. & A. Thanailakis. 1997. A model for predicting forest fire spreading using cellular automata. *Ecological Modelling* 99: 87-89.
- Landau, E. C. 2001. Corredores ecológicos como paradigma inovador para a conservação da diversidade biológica; estudo de caso na mata atlântica do sul da Bahia. Tese de Doutorado. Ecologia, Conservação e Manejo da Vida Silvestre, UFMG.
- Lett, C., C. Silber & N. Barret. 1999. Comparison of a cellular automata network and an individual-based model for the simulation of forest dynamics. *Ecological Modelling* 121: 277-293.
- Lévy, P. 1998. *Cyberculture*. Odile Jacob, France.
- Liu, J. & P. Ashton. 1998. FORMOSAIC: an individual-based spatially explicit model for simulating forest dynamics in landscape mosaics. *Ecological Modelling* 106: 177-200.
- Ludeke A., R.C. Maggio & L.M. Reid. 1990. An analysis of anthropogenic deforestation using logistic regression and GIS. *Journal of Environmental Management* 31: 247-59.
- McGarigal, K. & B.J. Marks. 1995. Fragstats: spatial pattern analysis program for quantifying landscape structure. Gen. Tech. Report PNW-GTR-351, Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland.

- Mertens, B. & E. F. Lambin. 2000. Land-cover change trajectories in Southern Cameroon. *ANNA* 205: 467-494.
- Sirakoulis, G. C., I. Karafyllidis & A. Thanailakis. 2000. A cellular automaton model for the effects of population movement and vaccination on epidemic propagation. *Ecological Modelling* 133: 209-223.
- Sklar, F.H. & R. Costanza. 1991. The development of dynamic spatial models for landscape ecology: a review and prognosis. In: M.G. Turner & R.H. Gardner (eds.). *Quantitative methods in landscape ecology: the analyses and interpretation of landscape heterogeneity*. pp. 239-288. Springer Verlag, New York.
- Soares-Filho, B.S., R.M. Assunção & A. Pantuzzo. 2001. Modeling the spatial transition probabilities of landscape dynamics in an Amazonian colonization frontier. *BioScience* 51: 1039-1046.
- Soares-Filho, B.S., C. Pennachin & G. Cerqueira. 2002. DINAMICA – a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecological Modelling* 154: 217-235.
- Soares-Filho, B.S., L. Corradi, G. C. Cerqueira & W. Araújo. 2003. Simulating the spatial patterns of change through the use of the dinamica model. In: XI Simpósio Brasileiro de Sensoriamento Remoto, Cd-Rom. J.C.N. Epiphanyo & D.M. Valeriano (orgs). INPE, Belo Horizonte.
- Soares-Filho, B., A. Alencar, D. Nepstad, G.C. Cerqueira, M. Vera Diaz, S. Rivero, L. Solórzano & E. Voll. 2004. Simulating the response of land-cover changes to road paving and governance along a major Amazon highway: The Santarém-Cuiabá corridor. *Global Change Biology* 10: 745-764.
- Tomlin, C. D. 1990. *Geographic Information Systems and Cartographic Modeling*. Prentice-Hall Inc., New Jersey.
- Turner, M.G., D.N. Wear & R.O. Flamm. 1996a. Land ownership and land-cover change in southern Appalachian highlands and the Olympic Peninsula. *Ecological Applications* 6: 1150-1172.
- Turner, M.G., D.N. Wear & R.O. Flamm. 1996b. Ecosystem management with multiple owners: Landscape dynamics in a southern Appalachian watershed. *Ecological Applications* 6: 1173-1188.
- Ventana. 2002. Vensim software – linking systems thinking to powerful dynamic models. < <http://www.vensim.com/software.html> > (Novembro 2002).
- Voinov, A., R. Costanza, L. Wainger, R. Boumans, F. Villa, T. Maxwell & H. Voinov. 1999. Patuxent landscape model: integrated ecological economic modeling of a watershed. *Environmental Modelling & Software* 14: 473-491.
- White, R., G. Engelen & I. Uljee. 2000a. Modelling Land use change with linked cellular automata and socio-economic models: A tool for exploring the impact of climate change on the Island of St. Lucia. In: M. Hill & R. Aspinall (eds.). *Spatial Information for Land Use Management*. pp. 189-204. Gordon and Breach, Amsterdam.
- White, R., G. Engelen, I. Uljee, C. Lavalle & D. Ehrlich. 2000b. Developing an urban land use simulator for European cities. In: K. Fullerton (ed.). *5th EC GIS Workshop: GIS of Tomorrow*. pp. 179-190. European Commission Joint Research Centre.
- Wu, F. 1998. SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rule. *International Journal of Geographical Information Science* 12: 63-82.

---

# Planejamento sistemático para a conservação na Amazônia brasileira – uma avaliação preliminar das áreas prioritárias de Macapá-99

ANA LUIZA KERTI MANGABEIRA ALBERNAZ<sup>1\*</sup>  
MANUELLA ANDRADE DE SOUZA<sup>2</sup>

<sup>1</sup> Museu Paraense Emílio Goeldi – MPEG, Pará, Brasil.

<sup>2</sup> Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis – IBAMA, Pará, Brasil.

\* e-mail: anakma@museu-goeldi.br

## RESUMO

É feita uma descrição do método de planejamento sistemático para a conservação, incluindo suas principais etapas e os diferentes algoritmos desenvolvidos. As etapas incluem o levantamento de informações, a definição das metas, a análise de representatividade do sistema existente ou análise de lacunas, a seleção de novas áreas, a implementação e o monitoramento das áreas. Os algoritmos descritos são o MINSET, o de insubstituibilidade, programação linear, “*simulated annealing*” e “*environmental diversity*” (ED). Ao fim, é realizado um exemplo de aplicação comparando os resultados obtidos usando o algoritmo de insubstituibilidade (usando metas de 10% sobre ambientes gerados pelo cruzamento de informações de tipos de vegetação e ecorregiões) com as áreas prioritárias indicadas para a criação de unidades de conservação, entre as definidas em Macapá-1999.

## ABSTRACT

*A description of the systematic conservation planning is presented, including its main steps and the different types of algorithms developed. Steps include compilation of information, definition of targets, gap analysis based on existing reserve system, selection of new areas, implementation of the system, and monitoring. Described algorithms are MINSET, irreplaceability, linear programming, simulated annealing, and environmental diversity (ED). At the end, an example of application is presented, comparing results of irreplaceability analysis (using targets of 10%, based on the intersection of vegetation and ecoregion maps) with those areas indicated for reserve creation amongst the priority areas defined in Macapá-1999.*

## INTRODUÇÃO

Por muito tempo, a escolha de áreas para a criação de unidades de conservação foi baseada principalmente em critérios estéticos. Áreas de rara beleza cênica, como o Pico das Agulhas Negras (Parque Nacional de Itatiaia, criado em 1937) e as cataratas do Iguaçu (Parque Nacional do Iguaçu, criado em 1939) estão entre as unidades de conservação mais antigas do Brasil, acontecendo o mesmo em vários países do mundo. Por volta da década de 70, começou-se a discutir critérios mais científicos, como a seleção de áreas com maior riqueza de espécies ou maior quantidade de espécies endêmicas ou raras, os quais são até hoje utilizados na priorização de áreas para a conservação (Dinerstein & Wikramanayake, 1993; Prendergast *et al.*, 1993; Kershaw *et al.*, 1995; Caldecott *et al.*, 1996; Kerr, 1997; Mittermeier *et al.*, 1998).

No início da década de 80, foi proposto o critério de complementaridade para a escolha de novas áreas para a conservação (Kirkpatrick, 1983). Este autor sugeriu que, em vez de se considerar apenas a riqueza ou endemismo como critérios para a escolha de novas áreas a serem protegidas, deveria ser priorizada a proteção de áreas com elementos diferentes daqueles já representados em unidades de conservação. Tais elementos poderiam corresponder a qualquer nível de diversidade, fosse ambiental, específica ou genética. Hoje em dia, a complementaridade é a base de todas as técnicas para a seleção de áreas protegidas (Margules & Pressey, 2000).

A partir deste marco teórico é que se desenvolvem as principais técnicas para a seleção de áreas para a conservação. Todas elas são baseadas no estabelecimento de metas explícitas para a proteção, bem como na aplicação de algoritmos que buscam formas otimizadas de atingir as metas. O processo que envolve a aplicação destas técnicas é chamado genericamente de “planejamento sistemático” e é considerado a forma mais eficiente de ampliar a representatividade das redes de áreas protegidas (Pressey, 1994; Camm *et al.*, 1996; Nantel *et al.*, 1998; Araújo, 1999). A eficiência, neste sentido, é definida como a capacidade de atingir a maior parte das metas de conservação com um menor custo em termos de área (Pressey & Nichols, 1989).

O processo de planejamento sistemático para a conservação envolve as seguintes etapas: (1) levantamento das informações existentes na área de interesse; (2) estabelecimento de metas para a conservação; (3) análise da representatividade do sistema atual; (4) seleção de novas áreas; (5) implementação; e (6) monitoramento, controle e revisão do sistema (Margules & Pressey, 2000).

## Levantamento de informações

O levantamento das informações é uma etapa fundamental do trabalho. Para a maioria dos lugares do mundo, os dados de ocorrência de espécies são escassos e mal-distribuídos. Assim, embora espécies endêmicas e raras sejam de alto interesse para a conservação, este tipo de informação quase nunca é disponível. Uma alternativa encontrada foi a utilização de mapas de condicionantes ambientais – como solos, relevo e pluviosidade – na previsão de distribuição da espécie. A principal dificuldade com esta opção é selecionar quais informações são realmente relevantes para a distribuição da diversidade em uma dada região, e qual a escala espacial adequada para representar os padrões existentes (Margules & Redhead, 1995). É preciso ter em mente, também, que nem sempre o acréscimo de informações significa uma melhoria na resposta. Para uma área bem estudada na Austrália, por exemplo, foi verificado que o mapa de vegetação, sozinho, previa melhor a distribuição de diversos grupos de espécies que o mapa de domínios ambientais, no qual unidades homogêneas de paisagem foram criadas a partir de uma fusão de variáveis climáticas, geológicas e físicas (Ferrier & Watson, 1997). Dessa forma, esta etapa envolve não só o levantamento das informações disponíveis, mas também a seleção daquelas que têm relevância para a distribuição das espécies na área de interesse e que, por isso, deverão ser utilizadas nas demais etapas do processo.

## Estabelecimento de metas

A definição de metas para a conservação é provavelmente a etapa mais difícil e para a qual se tem menos parâmetros. As metas dizem respeito à quantificação de ambientes (e espécies, em alguns casos) que se pretende conservar. Embora a Convenção da Biodiversidade estabeleça uma meta de 10% para a proteção de ambientes e espécies, esta meta é considerada insuficiente, seja para incluir a maior parte dos ambientes e espécies em área protegida, seja para assegurar a manutenção dos processos ecológicos e evolutivos (Soulé & Sanjayan, 1998; Cowling *et al.*, 1999). A proporção de área necessária pode variar entre regiões, dependendo do número de espécies incluídas entre os alvos, do nível de endemismo das espécies, ou do grau e do tipo de ameaças a que a área de interesse está sujeita (Rodrigues & Gaston, 2001; Pressey *et al.*, 2003). Para as áreas tropicais, acredita-se que os alvos devam estar bem acima dos 10%, mas os conhecimentos atuais não permitem uma estimativa precisa da quantidade de área necessária para que um sistema de áreas de conservação

seja capaz de manter a maior parte das espécies e dos processos ligados à manutenção da biodiversidade. Esta é uma das razões pelas quais o sistema deve estar sempre sendo monitorado e revisto: as metas estabelecidas podem ter sido inadequadas ou podem mudar ao longo do tempo (Pressey *et al.*, 2003). No entanto, o estabelecimento de metas explícitas cria uma base sólida para discussões e negociações entre os diversos grupos de interesse, além de possibilitar o uso de sistemas computacionais de suporte à decisão (Pressey, 1998).

#### **Análise de representatividade e seleção de novas áreas**

A análise de representatividade e a seleção de novas áreas dependem diretamente das informações selecionadas e das metas estabelecidas. Estas etapas têm sido realizadas dentro dos programas computacionais de apoio. Diferentes algoritmos vêm sendo desenvolvidos desde a década de 1980; alguns com melhor solução matemática (Underhill, 1994; Rodrigues *et al.*, 2000), outros com soluções consideradas sub-ótimas, mas com maior possibilidade de incluir outras funções, como conectividade e custos (Pressey *et al.*, 1996a; Possingham *et al.*, 2000).

A análise de representatividade é também conhecida como análise de lacunas e, por meio dela, busca-se avaliar quais dentre as metas definidas já estão contempladas no sistema de unidades de conservação existente. Os elementos já incluídos em áreas protegidas são então excluídos das metas. Por meio do algoritmo, a prioridade relativa entre as áreas candidatas é recalculada com base nas metas ainda não contempladas.

#### **Os algoritmos de suporte à decisão**

Diversos algoritmos de suporte à decisão têm sido desenvolvidos para a seleção de novas áreas. Tais algoritmos consistem basicamente numa seqüência de regras que vão definindo as prioridades. Por exemplo, o algoritmo MINSET (Belbin, 1995) seleciona o mínimo de áreas a serem protegidas, baseado na inclusão de todas as espécies e respectivas quantidades definidas nas metas, incluídas em uma matriz espécie  $\times$  área. A seqüência de seleção inclui a área de maior riqueza como primeira prioridade, seguida por aquela que tem mais espécies diferentes da área selecionada, e assim sucessivamente, por meio de operações iterativas com a matriz. As iterações param quando as metas são atingidas. Alguns algoritmos mais sofisticados integram valores humanos (arqueológicos, arquitetônicos, históricos, recreacionais, religiosos, econômicos) aos naturais (geologia, clima, fitofisionomias, distribuição de

comunidades, espécies ou subespécies) na hierarquização dos valores para conservação (Keisler & Sundell, 1997). Embora nem sempre seja possível quantificar todos os valores envolvidos no processo e algumas vezes seja necessário atribuir índices arbitrários a alguns deles, uma das principais vantagens dos algoritmos é tornar todo o sistema de decisão explícito, de forma que seja possível a todos os interessados ter acesso aos princípios considerados nas escolhas e discutir de forma construtiva sobre as possibilidades ou necessidades de alteração.

Outro algoritmo desenvolvido para apoiar a seleção de áreas protegidas é a análise de insubstituibilidade (Ferrier *et al.*, 2000). O valor da insubstituibilidade (“*irreplaceability*”) atribuído a uma área corresponde à sua contribuição para que os alvos sejam atingidos e à redução da probabilidade de atingir os alvos se a área não for selecionada (Pressey, 1999). Uma das grandes vantagens deste algoritmo é permitir incluir na análise informações de diferentes níveis, como listas de espécies e tipos de vegetação, grau de ameaça e custos. Sua utilização é interativa, funcionando como uma extensão do SIG ArcView, chamada C-Plan (NPWS-NSW 2001). A extensão possibilita a execução do algoritmo diretamente sobre um mapa no monitor de um computador e a recalculagem instantânea dos valores de insubstituibilidade após a seleção de uma área ou a retirada da mesma do conjunto de áreas protegidas. Esta característica torna sua utilização adequada para decisões no mundo real, que freqüentemente envolvem negociação entre diversos grupos de interesse. Desenvolvido pelo Serviço Nacional de Parques e Vida Silvestre de New South Wales (NPWS-NSW 2001), Austrália, o C-Plan foi usado com sucesso em negociações para o aumento de área protegida envolvendo setores do governo, madeireiras e ambientalistas (Pressey, 1998).

Os algoritmos baseados em programação linear são matematicamente mais elegantes e são mais eficientes em promover a otimização do processo de seleção, no sentido de incluir todas as metas com um menor número total de áreas (Underhill, 1994). Conforme este autor, os algoritmos heurísticos não são capazes de encontrar o melhor conjunto de áreas porque selecionam áreas de forma seqüencial, o que aumenta a redundância entre os elementos selecionados. Outros autores também atestaram a performance superior desta técnica quando comparada aos algoritmos iterativos (Camm, 1996; Church, 1996; Csuti, 1997). No entanto, a programação linear não é muito eficiente quando o conjunto de áreas candidatas é muito grande (maior que 20 ou 30), o que torna o processamento

muito lento; além disso, produz uma solução única (Possingham *et al.*, 2000). A flexibilidade – que é a capacidade de gerar diferentes combinações de áreas que possibilitem atingir as mesmas metas, é considerada uma das principais vantagens dos algoritmos iterativos (Pressey *et al.*, 1996a).

Mais recentemente, foi desenvolvido um algoritmo para a seleção de reservas baseado em “*simulated annealing*”. Este método de minimização é baseado no processo de enrijecimento de metais e vidros. Ele inicialmente gera um sistema de reservas completamente randômico. Depois, explora diversas tentativas de solução, fazendo modificações randômicas sucessivas no sistema inicial. A cada passo, a nova solução é comparada com a solução prévia e a melhor delas é aceita. A vantagem deste método é que ele potencialmente pode evitar os ótimos locais. Ele permite que o sistema de reservas se mova temporariamente por todo o espaço de soluções sub-ótimas e, assim, aumenta o número de possibilidades de se atingir um mínimo global. Em princípio, o sistema aceita todas as mudanças, quer ela aumente ou diminua seu valor (custo). À medida que o processo avança, o sistema vai ficando cada vez mais seletivo sobre as mudanças que aceita, rejeitando aquelas mudanças que aumentam muito o valor do sistema. Ao final do processo, apenas mudanças que melhoram (diminuem) o valor do sistema são aceitas (Possingham *et al.*, 2000). O algoritmo para a seleção de reservas, desenvolvido por uma equipe da Universidade de Queensland, Austrália, se chama MARXAN (Ball & Possingham, 2000). Assim como C-Plan, ele pode funcionar de maneira integrada com o Arc-View e, inclusive, com o próprio C-Plan. Uma das suas grandes vantagens é que este algoritmo permite a incorporação de funções de custo (como distância mínima as outras unidades de conservação ou as áreas desmatadas ou urbanas) e de configuração espacial (como distância mínima entre áreas e minimização de perímetros).

Paralelamente ao desenvolvimento destes algoritmos, uma abordagem diferente foi proposta por Faith & Walker (1996a). A técnica que eles propuseram, chamada “*Environmental Diversity*” (ED), partia de trabalhos anteriores desenvolvidos no âmbito do projeto BioRap (Margules & Readhead, 1995), no qual análises de ordenação e classificação eram usadas como base para a diferenciação de ambientes. A principal vantagem da técnica é que ela permite aplicar a seleção sobre variáveis contínuas, obtidas a partir de análises de ordenação de dados ambientais ou biológicos. Conforme os autores, a ordenação das áreas, baseada em seus atributos, pode revelar padrões sutis de variação que

estão além da lista de elementos, podendo, por isso, constituir uma boa base para a seleção de áreas a serem protegidas (Faith & Walker, 1996b). As áreas a serem priorizadas são aquelas que apresentam a menor soma das distâncias aos seus vizinhos mais próximos, ou seja, as mais representativas de cada grupo (Faith & Walker, 1996a). As metas, neste caso, são relacionadas ao número de áreas a ser selecionado ou à área total a ser protegida, com uma possibilidade de ponderação para o tamanho das áreas candidatas. Esta maneira de estabelecer as metas é menos explícita e de difícil comparação com os demais métodos em termos de performance, e pouca atenção foi dada à técnica. No entanto, a possibilidade de trabalhar com variáveis contínuas, mais próximas à realidade ambiental, ainda mantém algum interesse sobre o método (veja Araújo & Williams, 2004; Faith *et al.*, 2004).

### Unidades de planejamento

Uma característica comum a todos estes sistemas é que eles requerem a criação de unidades de planejamento. Unidades de planejamento são polígonos que representam as áreas candidatas à conservação. A cada unidade é atribuído um valor para conservação, que é uma função dos elementos que contém em relação aos alvos determinados. As unidades de seleção podem ser definidas artificialmente, como é o caso das grades regulares, utilizadas com frequência em estudos de conservação (*e.g.*, Kress *et al.*, 1998, para a Amazônia; Williams *et al.*, 1996; Araújo, 1999) ou podem ser baseadas em características naturais, como as bacias hidrográficas (Nix, 1997). Cabe lembrar que sua definição afeta o resultado final: em geral, quanto maior a unidade de planejamento, maior o requerimento absoluto em área para incluir a totalidade das metas (Pressey & Logan, 1998; Rodrigues & Gaston, 2001; Warman *et al.*, 2004). A utilização de unidades “artificiais”, como grades e hexágonos, requer a revisão dos limites após a seleção das áreas, mas tem a vantagem de não priorizar a seleção de determinadas unidades por ponderar seu tamanho.

### Persistência

Embora menos enfatizada na literatura pertinente, a chance de persistência de longo prazo dos elementos que se quer proteger é um aspecto tão importante do planejamento de áreas para a conservação quanto à representatividade. Algumas características que favorecem a persistência são: o tamanho, que deve incluir área suficiente para conter populações viáveis das espécies representadas (Burgmann & Lindenmayer, 1998; Berghout *et al.*, 1999); a conectividade com outras

áreas ecologicamente íntegras, para permitir a continuidade do fluxo gênico (Cowling *et al.*, 1999); a inclusão de unidades funcionais ecossistêmicas, como bacias hidrográficas, que possibilitem a manutenção de processos ecológicos (Cowling *et al.*, 1999); a defensibilidade natural da área (Peres & Terborgh, 1994); e a aceitação das estratégias de conservação pelas populações da área e entorno (Hyden, 1998). Para algumas destas características, como tamanho e conectividade, já podem ser atribuídas metas específicas, que podem também ser incorporadas como funções em alguns dos algoritmos já desenvolvidos (Ball & Possingham, 2000). Análises de ameaças e funções prevendo impactos presentes e futuros também têm sido incorporadas ao processo de escolha de áreas para a conservação, sendo, quando possível, desejável minimizar seus efeitos sobre as áreas escolhidas (Faith & Walker, 1996a). No entanto, áreas com características únicas, mas vulneráveis à ação antrópica, devem ser priorizadas nas ações de conservação (Pressey *et al.*, 1996b).

As etapas de implementação, monitoramento e revisão são efetuadas posteriormente. Tais etapas dependem de competências que vão muito além daquelas de um estudo acadêmico, e requerem a incorporação de informações de diversas outras naturezas, incluindo as demográficas, econômicas, sociais e de oportunidades.

## EXEMPLO DE APLICAÇÃO

O Brasil, como signatário da Convenção da Biodiversidade, tem o compromisso de ampliar o sistema de áreas integralmente protegidas na Amazônia, dos seus atuais cerca de 3% para os 10% acordados mundialmente. Até 2006, a principal orientação para a escolha de novas áreas era baseada nos resultados do Workshop “Avaliação e Identificação de Ações Prioritárias para a Conservação, Utilização Sustentável e Repartição dos Benefícios da Biodiversidade na Amazônia Brasileira”, realizado no âmbito do PROBIO, em Macapá, em 1999. Os resultados desta ação foram baseados na sobreposição de áreas indicadas por especialistas dos diversos grupos de trabalho (mamíferos, aves, anfíbios, etc.), em que maior importância relativa foi dada às áreas com maior número de sobreposições (MMA, 2002). Este documento foi a base para a definição das áreas onde deveriam ser criadas novas unidades de conservação nos vários níveis governamentais.

A escolha de áreas para a conservação por especialistas tem como vantagem que, características como a chance de persistência de espécies e as condições para

a implementação e manejo das áreas protegidas, são levadas em consideração, o que dificilmente ocorre na seleção feita por meio de algoritmos. Por outro lado, o julgamento de especialistas inevitavelmente envolve vieses associados ao conhecimento desigual entre diferentes regiões (Cowling *et al.*, 2003). Comparando as escolhas feitas por especialistas e algoritmos de suporte à decisão, estes autores chegaram à conclusão que os algoritmos seriam mais eficientes para detectar um padrão em escala ampla, enquanto os especialistas ajudam a especificar a alocação em escala fina, bem como as características de desenho que devem ser levadas em conta. Por isso, eles sugerem que estas duas abordagens devam ser mais integradas no planejamento para a conservação. As técnicas de planejamento sistemático para a conservação foram utilizadas na atualização das Áreas Prioritárias para a Conservação, realizada em 2006, para todos os biomas brasileiros.

Como exercício para a aplicação de uma das técnicas de planejamento sistemático, aplicamos o algoritmo “*irreplaceability*” (contido no C-Plan, gentilmente disponibilizado pelo NSW-NPWS) sobre os dados de vegetação e ecorregiões da Amazônia Legal, e comparamos as áreas de maior valor de insubstituibilidade com os polígonos indicados em Macapá-99 para a criação de unidades de conservação.

## MÉTODO

### Informações utilizadas

**Vegetação** – A vegetação de uma área é quase sempre reflexo de suas condicionantes ambientais, principalmente solo e clima (Richards, 1952; Aber & Melillo, 1991), e em geral é um bom preditor da distribuição de outros grupos de espécies (Ferrier & Watson, 1997). A base de dados digital de vegetação utilizada foi a da série POESIA-AM, produzida pelo IBAMA-CSR (2001), com 6.176 polígonos. Nesta base, as classes de vegetação seguem o sistema proposto por Veloso *et al.* (1991): este é um sistema hierárquico em seis níveis, em que o primeiro reflete as estruturas/formas de vida (savana, floresta, pioneiras, etc.); o segundo reflete clima/déficit hídrico (estacional, ombrófila, etc.); o terceiro os hábitos (arbórea, arbustiva, etc.); o quarto reflete o relevo (aluvial, terras baixas, submontana, etc.), e o quinto reflete alguns detalhes da fitofisionomia (dossel emergente, com palmeiras, com bambus, etc.). A documentação para o sexto nível é insuficiente para compreender esta subdivisão (Nelson & Oliveira, 2000). Por isso, foram utilizadas as 89 fitofisionomias que representam a

classificação até o quinto nível. Como o mapeamento já tem mais de 30 anos, não foram excluídas aquelas fitofisionomias antrópicas, como vegetação secundária, reflorestamentos e áreas de cultivo. É difícil saber qual a situação atual destas áreas e consideramos que florestas secundárias mais velhas, embora quase sempre revelem um empobrecimento em florística, podem ter um papel importante na manutenção de serviços ambientais e grupos faunísticos.

**Ecorregiões** – O uso de ecorregiões para subsidiar decisões sobre conservação foi proposto por Olson & Dinnerstein (1998). Uma ecorregião seria idealmente uma unidade biogeográfica capaz de representar os diferentes habitats e biotas (Wikramanayake *et al.*, 2002). As 23 ecorregiões definidas para a Amazônia Brasileira (WWF, 2001) têm suas delimitações baseadas principalmente nos interflúvios dos grandes rios, cuja atuação como barreira para a distribuição de aves e primatas já é bem conhecida (Ayres & Clutton-Brock, 1992; Haffer, 1992). Assim, embora sua eficiência para prever a distribuição de outros grupos ainda não tenha sido propriamente investigada, a subdivisão da região amazônica é interessante para as decisões de conservação, porque serve como um substituto categórico para a substituição de espécies com a distância, que é descrita ou esperada para a maioria dos grupos biológicos (Tuomisto *et al.*, 2003; Magnusson, 2004). Neste estudo, o mapa de ecorregiões foi cruzado com o de vegetação, criando um mapa de unidades ambientais.

### Metas

Foram utilizadas metas uniformes de 10% para cada unidade ambiental gerada a partir do cruzamento de tipos de vegetação e ecorregiões.

### Unidades de Planejamento

Foi criado um *layer* de unidades de seleção hexagonais, que emprestam maior conectividade ao desenho. A área de cada unidade foi de 1.200 km<sup>2</sup>, ou 120 mil hectares. Nas unidades de conservação de proteção integral os limites dos hexágonos foram dissolvidos, de forma que cada UC correspondesse a apenas uma unidade de planejamento. A área correspondente aos rios foi retirada do *layer*, uma vez que devem permanecer com trânsito livre.

### Importação para o C-Plan

O *software* C-Plan de suporte à decisão foi desenvolvido pelo Serviço Nacional de Parques e Vida Silvestre de

Nova Gales do Sul (NSW-NPWS), Austrália, e é distribuído sem custo para profissionais ligados ao planejamento em conservação, mediante cadastro junto à Instituição. O *software* opera com o algoritmo “*irreplaceability*” (Ferrier *et al.*, 2000), em que maior valor para a conservação é atribuído às unidades de planejamento que (1) contribuem com maior proporção de metas para atingir o total das metas, e (2) se não incluídas no sistema de áreas protegidas, acarretam maior dificuldade em atingir as metas globais do planejamento.

A construção de uma base de dados para aplicação no *software* envolve a montagem de uma matriz de locais, uma base de locais por feições e uma base de feições e metas. A base de locais é derivada do *layer* de unidades de planejamento e deve conter o código de cada unidade, sua área, e, quando pertinente, o tipo de propriedade (por exemplo, se a área já é uma unidade de conservação). Neste estudo, o campo de propriedade inclui as três categorias amplas de áreas protegidas, que são as unidades de conservação de proteção integral, as de uso sustentável e as terras indígenas. Tal inclusão nos permite avaliar o impacto de cada uma destas categorias na aproximação das metas. A cada uma destas categorias, na importação para o C-Plan, já pode ser atribuído um “*status*” específico, entre disponível, reservado e excluído. Na base confeccionada, as unidades de proteção integral foram consideradas já reservadas, mas as de uso sustentável e as terras indígenas, não. Isso permite avaliar sua importância relativa para a conservação e recomendar medidas de manejo, quando pertinente.

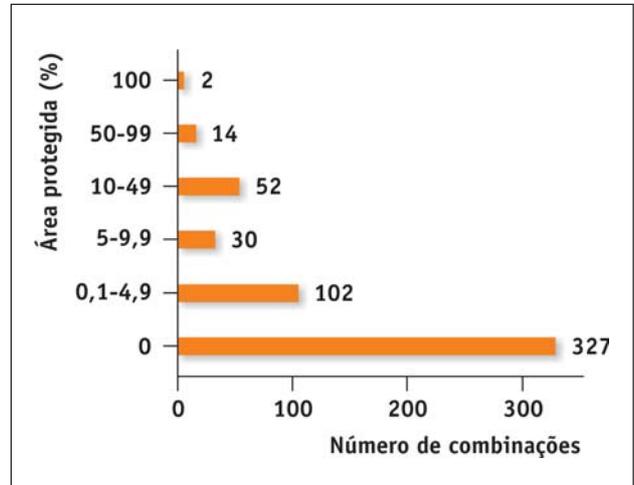
A matriz de locais por feições contém as quantidades de cada feição em cada local (unidades de planejamento). Na base construída, contém a área de cada tipo de vegetação em cada unidade de planejamento. A base de feições e metas contém os nomes ou códigos de cada fitofisionomia (iguais aos da tabela de locais por feições) e as metas, representadas como a área que deverá ser protegida para cada feição.

## RESULTADOS

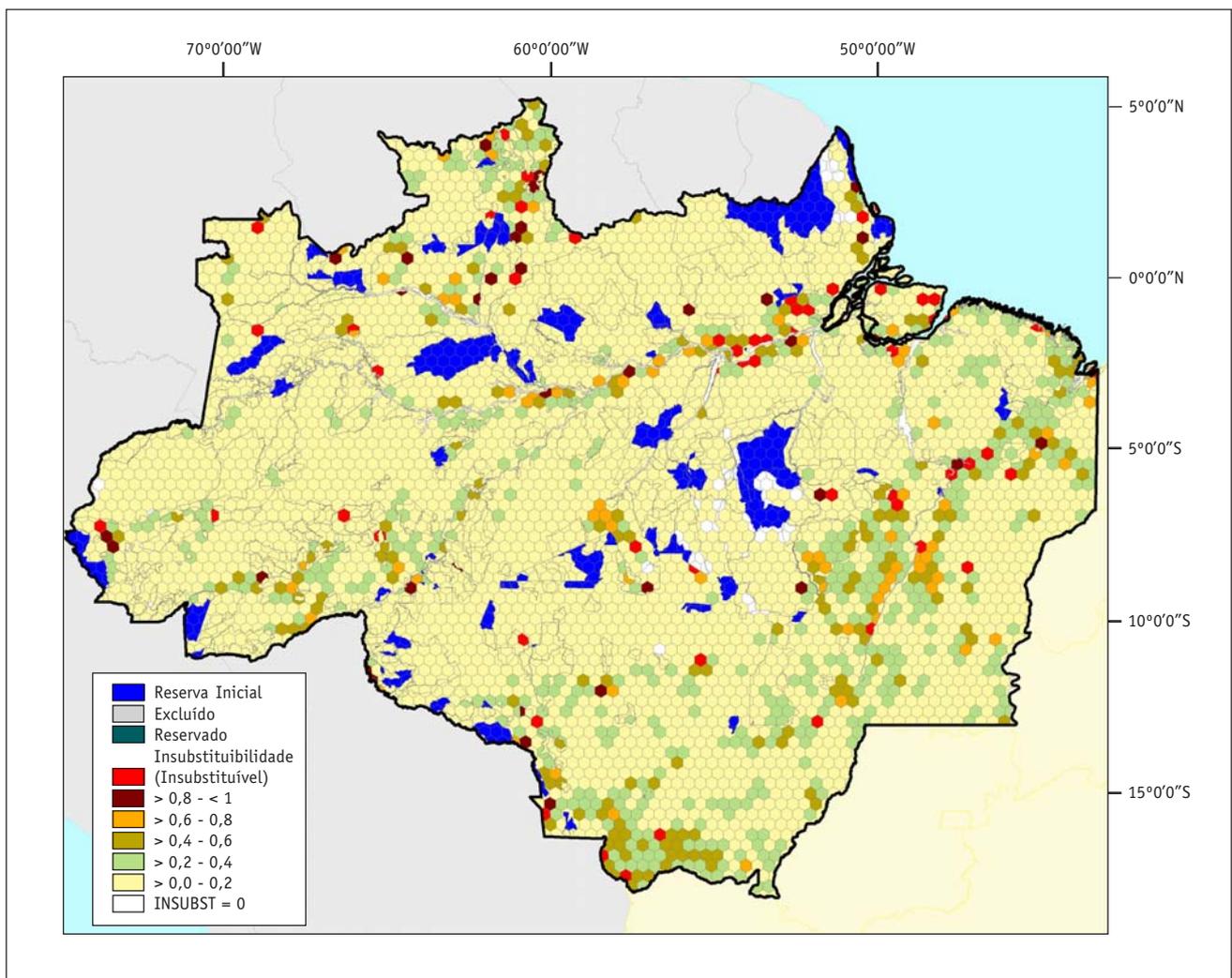
### Representatividade

Como já demonstrado em estudos anteriores sobre a representatividade do sistema de Unidades de Conservação na Amazônia Brasileira (Fearnside & Ferraz, 1995; Ferreira, 2001), a análise de insubstituibilidade usando tipos de vegetação e ecorregiões também revelou haver grande discrepância entre as metas atingidas. Considerando-se apenas o sistema atual de Unidades de

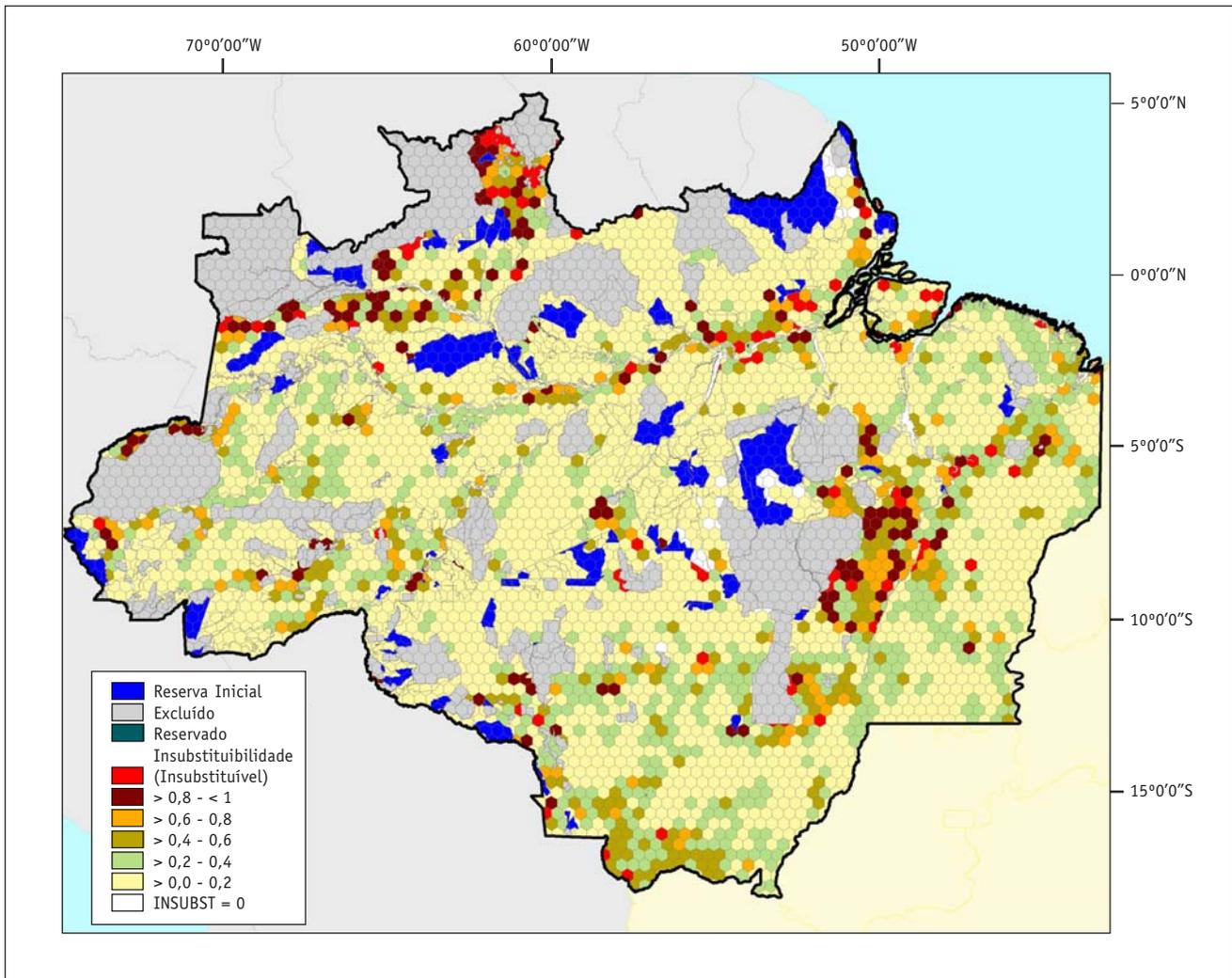
Conservação (UC) de proteção integral, das 527 combinações de vegetação e ecossistemas, duas estão 100% representadas no sistema, o que equivale a 10 vezes a meta atribuída. Outras 66 combinações também estão super-representadas com mais de 10% de sua área em unidades de conservação, enquanto 327 combinações não possuem representação alguma no sistema (Figura 1). Neste cenário, aglomerações de áreas com valores mais altos para a conservação situam-se em Roraima, na várzea do rio Amazonas, no interflúvio Tocantins-Xingu, no sul e no leste da Amazônia Legal (Figura 2). As áreas totalmente insubstituíveis (*irreplaceability* = 1) foram muito poucas e muito esparsas, mas é preciso considerar que nesta base de dados não houve nenhuma informação direta de distribuição de espécies – todas as informações utilizadas são consideradas de baixa resolução.



**FIGURA 1** – Proporção de área de cada uma das 527 combinações de tipos de vegetação e ecorregiões incluídas em áreas sob proteção integral.



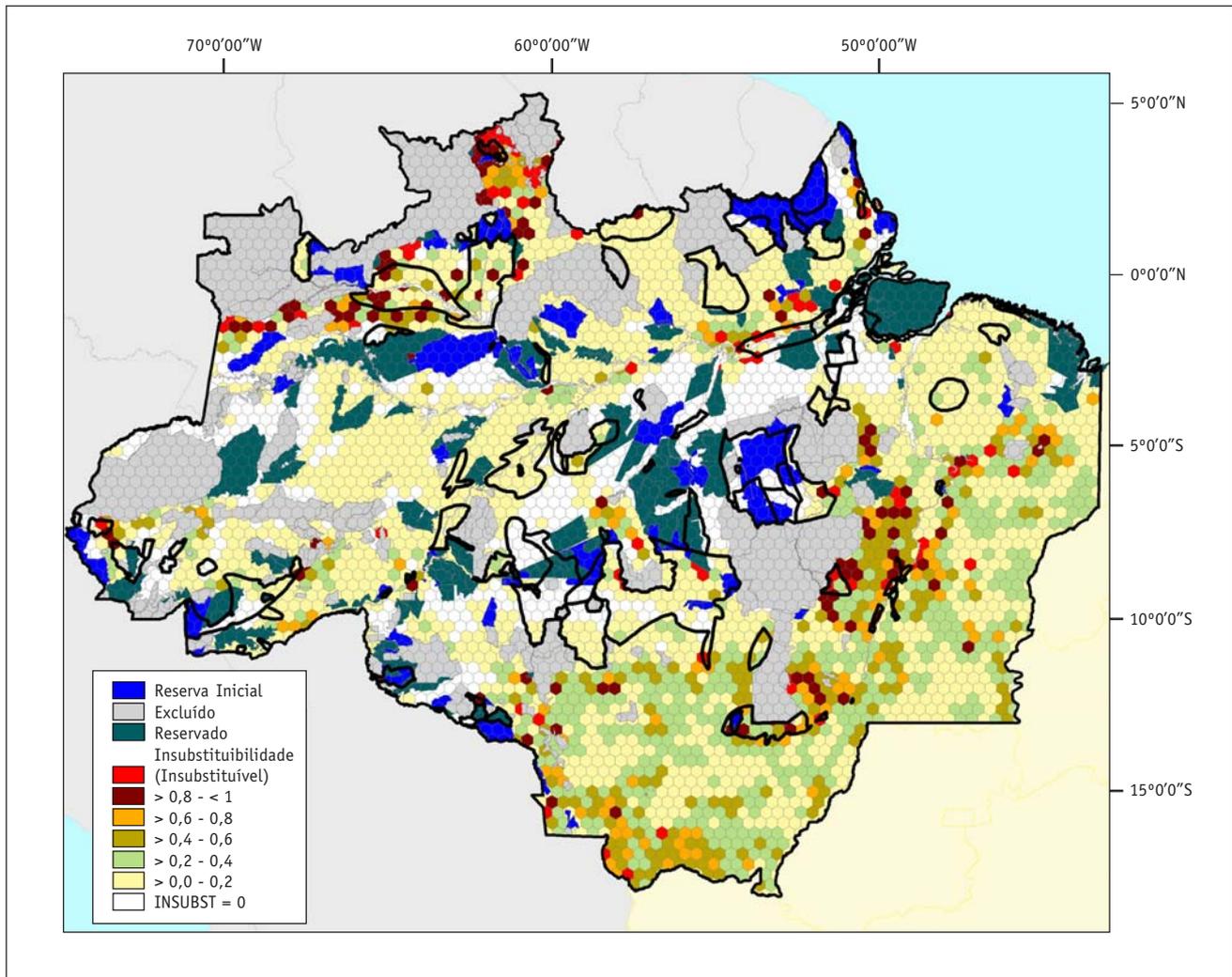
**FIGURA 2** – Resultado da análise de insubstituibilidade (*“irreplaceability”*) utilizando Cplan, com a reserva inicial de unidades de conservação de proteção integral.



**FIGURA 3** – Resultado da análise de insubstituibilidade (“irreplaceability”) utilizando C-Plan, com as terras indígenas (TIs) consideradas como não-disponíveis para o sistema de áreas para a conservação.

Embora a CDB inclua apenas áreas de proteção integral entre suas metas, há uma ampla discussão no Brasil sobre a inclusão de Terras Indígenas e áreas de uso sustentável como áreas que contribuem para atingir metas conservacionistas. A discussão sobre se estas áreas devem ser incluídas ou excluídas neste contexto, ou de que forma devem contribuir para as metas de conservação, está muito além do escopo deste exercício, mas nós fizemos algumas simulações para ver como se comportam os valores de insubstituibilidade mediante diferentes considerações quanto a essas categorias de áreas protegidas. Assim, por exemplo, se considerarmos que as terras indígenas (TI) não têm como função primordial a conservação da natureza e que são soberanas em relação às demais categorias de áreas protegidas, avaliamos como se comportariam os valores de

insubstituibilidade ao se excluir estas áreas do sistema, considerando-as como não-reservadas e não disponíveis para a conservação. O resultado reforçou ainda mais a importância de Roraima e de algumas áreas ao sul e ao leste da região amazônica, e revelou ainda como áreas importantes para a conservação parte do Amapá e algumas áreas ao redor do alto rio Negro (Figura 3). Ao se manter as áreas indígenas excluídas, mas considerando as UCs de uso sustentável como contribuindo para as metas, vê-se que grande parte das terras baixas da Amazônia encontram níveis satisfatórios de representação em unidades de conservação. No entanto, os valores das áreas ao sul, ao leste e noroeste da Amazônia são relativamente ainda mais realçados, mostrando que estes ambientes também não estão contemplados nas UCs de uso sustentável. A maioria destes ambientes



**FIGURA 4** – Resultado da análise de insubstituibilidade (*“irreplaceability”*) utilizando C-Plan, considerando as TIs excluídas, as UCs de uso sustentável reservadas e com os polígonos com propostas para a criação de UCs em Macapá-99 (linha preta).

não estavam incluídos nos polígonos propostos para a criação de UCs em Macapá-99 (Figura 4), que guiou as ações do Programa ARPA em sua primeira fase.

## CONSIDERAÇÕES FINAIS

Mesmo sendo uma extrema simplificação do processo, este exercício mostra como um modelo deste tipo é eficaz em auxiliar a tomada de decisões relativas à conservação de ambientes, podendo simular rapidamente a situação sob diferentes cenários. Ao processo podem ser incorporadas informações sobre a distribuição de espécies, idealmente por meio de mapas de sua distribuição (que podem ser gerados por alguma das ferramentas contidas neste volume).

Por fim, cabe ressaltar que este é apenas um exercício preliminar: o planejamento real deverá incluir as mencionadas funções de custo e configuração espacial, metas relacionadas aos processos ecológicos e evolutivos e uma atribuição de metas diferenciadas entre diferentes níveis hierárquicos da vegetação. Além disso, a análise seria muito mais eficiente, no sentido de minimizar a área total requerida para atingir as metas, se as unidades de planejamento fossem de tamanho menor.

## AGRADECIMENTOS

Este estudo foi apoiado pelo Projeto GEOMA/MCT. Manuella Andrade de Souza recebeu bolsa do CNPq/GEOMA.

**ANEXO** – Classes de vegetação até o quinto nível incluídas na análise de insubstituibilidade, com as respectivas metas atingidas pelas atuais unidades de conservação de proteção integral.

<b>CLASSE</b>	<b>NÍVEL 1</b>	<b>NÍVEL 2</b>	<b>NÍVEL 3</b>
Dm	FLORESTA	OMBRÓFILA	DENSA
Pfh	PIONEIRA	INFLUÊNCIA FLUVIOMARINHA	HERBÁCEA
rmb	REFUGIO		ARBUSTIVO
Pab	PIONEIRA	INFLUÊNCIA FLUVIAL	BURITIZAL
Sgs	SAVANA	ESTACIONAL	GRAMÍNEO-LENHOSA
Dsu	FLORESTA	OMBRÓFILA	DENSA
Pfm	PIONEIRA	INFLUÊNCIA FLUVIOMARINHA	ARBÓREA
LO	CAMPINARANA/FLORESTA OMBRÓFILA		
La	CAMPINARANA	OMBRÓFILA	ARBORIZADA
Pah	PIONEIRA	INFLUÊNCIA FLUVIAL	HERBÁCEA
Dbu	FLORESTA	OMBRÓFILA	DENSA
Sgf	SAVANA	ESTACIONAL	GRAMÍNEO-LENHOSA
Lap	CAMPINARANA	OMBRÓFILA	ARBORIZADA
Fbu	FLORESTA	ESTACIONAL	SEMIDECIDUAL
Ld	CAMPINARANA	OMBRÓFILA	FLORESTADA
Asb	FLORESTA	OMBRÓFILA	ABERTA
Dau	FLORESTA	OMBRÓFILA	DENSA
Ldp	CAMPINARANA	OMBRÓFILA	FLORESTADA
Paa	PIONEIRA	INFLUÊNCIA FLUVIAL	ARBUSTIVA
Dbe	FLORESTA	OMBRÓFILA	DENSA
Ma	MASSA DE ÁGUA	MASSA DE ÁGUA	MASSA DE ÁGUA
Abp	FLORESTA	OMBRÓFILA	ABERTA
Dse	FLORESTA	OMBRÓFILA	DENSA
Spf	SAVANA	ESTACIONAL	PARQUE
Sps	SAVANA	ESTACIONAL	PARQUE
Asp	FLORESTA	OMBRÓFILA	ABERTA
SO	SAVANA/FLORESTA OMBRÓFILA		
Dae	FLORESTA	OMBRÓFILA	DENSA
Dmu	FLORESTA	OMBRÓFILA	DENSA
Db	FLORESTA	OMBRÓFILA	DENSA
Sp	SAVANA	ESTACIONAL	PARQUE
Pas	PIONEIRA	INFLUÊNCIA FLUVIAL	ARBUSTIVA
Ds	FLORESTA	OMBRÓFILA	DENSA
Fsu	FLORESTA	ESTACIONAL	SEMIDECIDUAL
Saf	SAVANA	ESTACIONAL	ARBORIZADA
Sd	SAVANA	ESTACIONAL	FLORESTADA
rm	REFÚGIO		
Sas	SAVANA	ESTACIONAL	ARBORIZADA
Phs	PIONEIRA	INFLUÊNCIA FLUVIAL	HERBÁCEA
Da	FLORESTA	OMBRÓFILA	DENSA
Asb	FLORESTA	OMBRÓFILA	ABERTA
Fse	FLORESTA	ESTACIONAL	SEMIDECIDUAL
Fs	FLORESTA	ESTACIONAL	SEMIDECIDUAL
Asc	FLORESTA	OMBRÓFILA	ABERTA
Sa	SAVANA	ESTACIONAL	ARBORIZADA
Sg	SAVANA	ESTACIONAL	GRAMÍNEO-LENHOSA
Fae	FLORESTA	ESTACIONAL	SEMIDECIDUAL
Ap	AGROPECUÁRIA		
SN	SAVANA/FLORESTA ESTACIONAL		

<b>NÍVEL 4</b>	<b>NÍVEL 5</b>	<b>PROPORÇÃO METAS ATINGIDAS</b>
MONTANA		725,308
		434,286
MONTANO		208,634
ALUVIAL		166,529
	SEM FLORESTA-DE-GALERIA	151,098
SUBMONTANA	COM DOSSEL UNIFORME	149,459
		120,667
		116,868
		113,743
ALUVIAL		111,447
DE TERRAS BAIXAS	COM DOSSEL UNIFORME	90,965
	COM FLORESTA-DE-GALERIA	88,194
	COM PALMEIRAS	82,456
DE TERRAS BAIXAS	COM DOSSEL UNIFORME	81,634
		66,218
SUBMONTANA	COM BAMBUS	58,774
ALUVIAL	COM DOSSEL UNIFORME	51,716
	COM PALMEIRAS	49,128
ALUVIAL		43,352
DE TERRAS BAIXAS	COM DOSSEL EMERGENTE	40,594
MASSA DE ÁGUA	MASSA DE ÁGUA	39,306
DE TERRAS BAIXAS	COM PALMEIRAS	37,426
SUBMONTANA	COM DOSSEL EMERGENTE	36,088
	COM FLORESTA-DE-GALERIA	33,911
	SEM FLORESTA-DE-GALERIA	29,223
SUBMONTANA	COM PALMEIRAS	25,955
		23,755
ALUVIAL	COM DOSSEL EMERGENTE	22,095
MONTANA	COM DOSSEL UNIFORME	19,495
DE TERRAS BAIXAS		17,958
		11,366
ALUVIAL	SEM PALMEIRAS	10,150
SUBMONTANA		9,984
SUBMONTANA	COM DOSSEL UNIFORME	9,816
	COM FLORESTA-DE-GALERIA	8,700
		8,313
MONTANO		7,905
	SEM FLORESTA-DE-GALERIA	7,753
ALUVIAL	SEM PALMEIRAS	7,388
ALUVIAL		7,378
SUBMONTANA	DE BAMBUS	7,023
SUBMONTANA	COM DOSSEL EMERGENTE	6,792
SUBMONTANA		5,686
SUBMONTANA	COM CIPÓS	4,108
		2,397
		1,516
ALUVIAL	COM DOSSEL EMERGENTE	1,353
	PECUÁRIA	0,561
		0,071

continua...

continuação do anexo

CLASSE	NÍVEL 1	NÍVEL 2	NÍVEL 3
Abc	FLORESTA	OMBRÓFILA	ABERTA
Ab	FLORESTA	OMBRÓFILA	ABERTA
Acc	AGROPECUÁRIA		
Am	FLORESTA	OMBRÓFILA	ABERTA
Amc	FLORESTA	OMBRÓFILA	ABERTA
Amp	FLORESTA	OMBRÓFILA	ABERTA
As	FLORESTA	OMBRÓFILA	ABERTA
Ass	FLORESTA	OMBRÓFILA	ABERTA
Ca	FLORESTA	ESTACIONAL	DECIDUAL
Cau	FLORESTA	ESTACIONAL	DECIDUAL
Cb	FLORESTA	ESTACIONAL	DECIDUAL
Cbe	FLORESTA	ESTACIONAL	DECIDUAL
Cm	FLORESTA	ESTACIONAL	DECIDUAL
Cme	FLORESTA	ESTACIONAL	DECIDUAL
Cs	FLORESTA	ESTACIONAL	DECIDUAL
Cse	FLORESTA	ESTACIONAL	DECIDUAL
Dme	FLORESTA	OMBRÓFILA	DENSA
Fa	FLORESTA	ESTACIONAL	SEMIDECIDUAL
Fm	FLORESTA	ESTACIONAL	SEMIDECIDUAL
Fme	FLORESTA	ESTACIONAL	SEMIDECIDUAL
Las	CAMPINARANA	OMBRÓFILA	ARBORIZADA
Lg	CAMPINARANA	OMBRÓFILA	GRAMÍNEO-LENHOSA
ON	FLORESTA OMBRÓFILA/FLORESTA ESTACIONAL		
Pap	PIONEIRA	INFLUÊNCIA FLUVIAL	ARBUSTIVA
Pmb	PIONEIRA	INFLUÊNCIA MARINHA	ARBUSTIVA
Pmh	PIONEIRA	INFLUÊNCIA MARINHA	HERBÁCEA
Re	REFLORESTAMENTO		
Rf	REFLORESTAMENTO		
rlh	REFÚGIO		HERBÁCEO
rmh	REFÚGIO		HERBÁCEO
Rp	REFLORESTAMENTO		
ST	SAVANA/SAVANA ESTEPICA		
Ta	SAVANA-ESTEPICA	ESTACIONAL	ARBORIZADA
Td	SAVANA-ESTEPICA	ESTACIONAL	FLORESTADA
Tgs	SAVANA-ESTEPICA	ESTACIONAL	GRAMÍNEO-LENHOSA
Tp	SAVANA-ESTEPICA	ESTACIONAL	PARQUE
Vs	VEGETAÇÃO SECUNDÁRIA		
Vsp	VEGETAÇÃO SECUNDÁRIA		
Vss	VEGETAÇÃO SECUNDÁRIA		

<b>NÍVEL 4</b>	<b>NÍVEL 5</b>	<b>PROPORÇÃO METAS ATINGIDAS</b>
DE TERRAS BAIXAS	COM CIPÓS	0,017
DE TERRAS BAIXAS		0,000
	CULTURA CÍCLICA	0,000
MONTANA		0,000
MONTANA	COM CIPÓS	0,000
MONTANA	COM PALMEIRAS	0,000
SUBMONTANA		0,000
SUBMONTANA	COM SOROROCAS	0,000
ALUVIAL		0,000
ALUVIAL	COM DOSSEL UNIFORME	0,000
DE TERRAS BAIXAS		0,000
DE TERRAS BAIXAS	COM DOSSEL EMERGENTE	0,000
MONTANA		0,000
MONTANA	COM DOSSEL EMERGENTE	0,000
SUBMONTANA		0,000
SUBMONTANA	COM DOSSEL EMERGENTE	0,000
MONTANA	COM DOSSEL EMERGENTE	0,000
ALUVIAL		0,000
MONTANA		0,000
MONTANA	COM DOSSEL EMERGENTE	0,000
	SEM PALMEIRAS	0,000
		0,000
		0,000
ALUVIAL	COM PALMEIRAS	0,000
		0,000
		0,000
	<i>EUCALIPTUS</i>	0,000
	FRUTÍFERAS	0,000
ALTO-MONTANO		0,000
MONTANO		0,000
	<i>PINUS</i>	0,000
		0,000
		0,000
		0,000
	SEM FLORESTA-DE-GALERIA	0,000
		0,000
		0,000
	COM PALMEIRAS	0,000
	SEM PALMEIRAS	0,000

## REFERÊNCIAS BIBLIOGRÁFICAS

- Aber, J.D. & J.M. Melillo. 1991. *Terrestrial Ecosystems*. Saunders College Publishing, Orlando FL.
- Araújo, M.B. 1999. Distribution patterns of biodiversity and the design of a representative reserve network in Portugal. *Diversity and Distributions* 5: 151-163.
- Araújo, M.B., P.J. Densham & P.H. Williams. 2004. Representing species in reserves from patterns of assemblage diversity. *Journal of Biogeography* 31: 1037-1050.
- Ayres, J.M.C. & T.H. Clutton Brock. 1992. River boundaries and species range size in Amazonian primates. *American Naturalist* 140: 531-537.
- Ball, I. & H. Possingham. 2000. *Marxan (Marine Reserve Design using Spatially Explicit Annealing)* V 1.8.2.
- Belbin, L. 1995. PATN - Pattern Analysis Package. Technical Reference Guide. CSIRO, Division of Wildlife and Ecology, Australia.
- Berghout, M., J. Passioura & S. May 1999. Response to disturbance of forest species in comprehensive regional assessments regions in New South Wales- Southern Region. Sydney, NSW, Environment Australia.
- Burgmann, M.A. & D.B. Lindenmayer. 1998. *Conservation Biology for the Australian Environment*. Surrey Beatty & Sons, NSW.
- Caldecott, J.O., M.D. Jenkins, T.H. Johnson & B. Goombridge 1996. Priorities for conserving global species richness and endemism. *Biodiversity and Conservation* 5: 699-727.
- Camm, J.D., S. Polasky, A. Solow & B. Csuti, 1996. A note on optimal algorithms for reserve site selection. *Biological Conservation* 78: 353-355.
- Church, R.L., D.M. Stoms & F.W. Davis 1996. Reserve selection as a maximal covering location problem. *Biological Conservation* 76: 105-112.
- Cowling, R.M., R.L. Pressey, A.T. Lombard, P.G. Desmet & A.G. Ellis. 1999. From representation to persistence: requirements for a sustainable system of conservation areas in the species-rich Mediterranean-climate desert of southern Africa. *Diversity and Distributions* 5: 51-71.
- Cowling, R., R.L. Pressey, R. Sims-Castley, A. le Roux, E. Baard, C.J. Burgers & G. Palmer, 2003. The expert or the algorithm? Comparison of priority conservation areas in the Cape Floristic Region identified by park managers and reserve selection software. *Biological Conservation* 112: 147-167.
- Csuti, B.P., S. Polasky, P.H. Williams, R.L. Pressey, J.D. Camm, M. Kershaw, A.R. Kiestler, B. Downs, R. Hamilton, M. Huso & K. Sahr. 1997. A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biological Conservation* 80: 83-97.
- Dinerstein, E. & E.D. Wikramanayake. 1993. Beyond "hotspots": how to prioritize investments to conserve biodiversity in the Indo-Pacific Region. *Conservation Biology* 7: 53-65.
- Faith, D. & P. Walker. 1996a. Integrating conservation and development: effective trade-offs between biodiversity and cost in the selection of protected areas. *Biodiversity and Conservation* 5: 431-446.
- Faith, D. & P. Walker. 1996b. Environmental diversity: on the best possible use of surrogate data for assessing the relative biodiversity of sets of areas. *Biodiversity and Conservation* 5: 399-415.
- Faith, D.P., S. Ferrier & P.A. Walker. 2004. The ED strategy: how species-level surrogates indicate general biodiversity patterns through an 'environmental diversity' perspective. *Journal of Biogeography* 31: 1207-1217.
- Fearnside, P.M. & J. Ferraz. 1995. A conservation gap analysis of Brazil's Amazonian vegetation. *Conservation Biology* 9: 1134-1147.
- Ferreira, L.V. 2001. A Representação das unidades de conservação no Brasil e a identificação de áreas prioritárias para a conservação da biodiversidade nas ecorregiões do bioma Amazônia. Tese de Doutorado. INPA/UFAM.
- Ferrier, S. & G. Watson. 1997. An evaluation of the effectiveness of environmental surrogates and modelling techniques in predicting the distribution of biological diversity. Environment Australia, Canberra.
- Ferrier, S., R.L. Pressey & T.W. Barret. 2000. A new predictor of the irreplaceability of areas for achieving a conservation goal, its application to real-world planning, and a research agenda for further refinement. *Biological Conservation* 93: 303-325.
- Haffer, J. 1992. On the "river effect" in some forest birds of Southern Amazonia. *Boletim do Museu Paraense Emílio Goeldi, série Zoologia* 8: 217-245.
- Hyden, G. 1998. Governance in conservation and development. Disponível em: <http://www.cdf.ufl.edu/cdf/Gov-Hyden.htm> (acessado em janeiro de 2000).
- IBAMA-CSR. 2001. Base de dados Digital-Série Poesia-AM. CD-Rom.
- Keisler, J.M. & R.C. Sundell. 1997. Park boundary decisions: applying a multi-attribute utility. In: J.J. Pigram & R.C. Sundell (eds). *National Parks and protected areas: Selection, delimitation, and management*. Centre for Water Policy Research, University of New England, Australia.
- Kerr, J.T. 1997. Species richness, endemism, and the choice of areas for conservation. *Conservation Biology* 11: 1094-1100.
- Kershaw, M., G.M. Mace & P.H. Williams. 1995. Threatened status, rarity, and diversity as alternative selection measures for protected areas: a test using Afrotropical antelopes. *Conservation Biology* 9: 324-334.
- Kirkpatrick, J.B. 1983. An iterative method for establishing priorities for the selection of nature reserves: an example from Tasmania. *Biological Conservation* 25: 127-134.
- Kress, W.J., W.R. Heyer, P. Acevedo, J. Coddington, D. Cole, T.L. Erwin, B.J. Meggers, M. Pogue, R.W. Thorington, R.P. Vari, M.J. Weitzman & S.H. Weitzman, 1998. Amazonian biodiversity: assessing conservation priorities with taxonomic data. *Biodiversity and Conservation* 7: 1577-1587.
- Magnusson, W.E. 2004. Ecoregion as a pragmatic tool. *Conservation Biology (letters)* 18: 4-5.
- Margules, C.R. & R.L. Pressey. 2000. Systematic conservation planning. *Nature* 450: 243-253.
- Margules, C.R. & T.D. Redhead. 1995. *Guidelines for using the BioRap Methodology Tools*. Fyshwick, ACT, CSIRO.
- Mittermeier, R.A., N. Myers, J.B. Thomsen, G.A.B. da Fonseca & S. Olivieri. 1998. Biodiversity hotspots and major tropical wilderness areas: approaches to setting conservation priorities. *Conservation Biology* 12: 516-520.

- MMA - Ministério do Meio Ambiente. 2002. Biodiversidade Brasileira - Avaliação e identificação de áreas e ações prioritárias para a conservação, utilização sustentável e repartição dos benefícios da biodiversidade nos biomas brasileiros. Brasília, DF.
- Nantel, P., A. Bouchard, L. Brouillet, & S. Hay 1998. Selection of areas for protecting rare plants with integration of land use conflicts: a case study for the west coast of Newfoundland, Canada. *Biological Conservation* 84: 223-234.
- Nelson, B. & A.A. Oliveira. 2000. Área Botânica. In: A. Veríssimo, A. Moreira, D. Sawyer, I. dos Santos, L.P. Pinto & J.P.R. Capobianco (eds). Biodiversidade na Amazônia brasileira: avaliação e ações prioritárias para a conservação, uso sustentável e repartição de benefícios. pp. 132-176. Estação Liberdade, Instituto Sócio Ambiental, São Paulo.
- Nix, H.A. 1997. Management of parks and reserves for the conservation of biological diversity. In: J.J. Pigrum & R.C. Sundell (eds). National Parks and protected areas: Selection, delimitation and management. pp. 11-36. Centre for Water Policy Research, Armidale.
- Olson, D. & E. Dinnerstein. 1998. The global 200: a representation approach to conserve Earth's most biologically valuable ecoregions. *Conservation Biology* 12: 502-515.
- Peres, C.A. & J.W. Terborgh. 1994. Amazonian nature reserves: an analysis of the defensibility status of existing conservation units and design criteria for the future. *Conservation Biology* 9: 34-46.
- Possingham, H., I. Ball & S. Andelman. Mathematical methods for identifying representative reserve networks. 2000. In: S. Ferson & M. Burgman (eds). Quantitative methods for conservation biology. pp. 291-305. Springer-Verlag, New York.
- Prendergast, J.R., R.M. Quinn, J.H. Lawton, B.C. Eversham & D.W. Gibbons. 1993. Rare species, the coincidence of diversity hotspots and conservation strategy. *Nature* 365: 335-337.
- Pressey, R.L. 1994. Ad hoc reservations: forward and backward steps in developing representative reserve systems? *Conservation Biology* 8: 662-668.
- Pressey, R.L. 1998. Algorithms, politics and timber: an example of the role of science in a public, political negotiation process over new conservation areas in production forests. In: R.T. Wills & R.J. Hobbs (eds). Ecology for everyone: Communicating ecology to scientists, the public and the politicians. pp.73-87. Surrey Beatty and Sons, Chipping Norton, NSW.
- Pressey, R.L.B. 1999. Applications of irreplaceability analysis to planning and management problems. *Parks* 9: 42-51.
- Pressey, R.L. & A.O. Nicholls. 1989. Efficiency in conservation evaluation: scoring versus iterative approaches. *Biological Conservation* 50: 199-218.
- Pressey, R.L. & V.S. Logan. 1998. Size of selection units for future reserves and its influence on actual vs targeted representation of features: a case study in western New South Wales. *Biological Conservation* 85: 305-319.
- Pressey, R.L., H.L. Possingham & C.R. Margules. 1996a. Optimality in reserve selection algorithms: when does it matter and how much? *Biological Conservation* 76: 259-267.
- Pressey, R.L., S. Ferrier, T.C. Hager, C.A. Woods, S.L. Tully & K.M. Weinman. 1996b. How well protected are the forests of north-eastern New South Wales? Analyses of forest environments in relation to formal protection measures, land tenure, and vulnerability to clearing. *Forest Ecology and Management* 85: 311-333.
- Pressey, R.L., R.M. Cowling & M. Rouget. 2003. Formulating conservation targets for biodiversity pattern and process in the Cape Floristic Region, South Africa. *Biological Conservation* 112: 99-127.
- Richards, P.W. 1952. *The Tropical Rainforest: An ecological study*. Cambridge University Press, London.
- Rodrigues, A.S.L. & K.J. Gaston. 2001. How large do reserve networks need to be? *Ecology Letters* 4: 602-609.
- Rodrigues, A.S.L., J.C. Cerdeira & K.J. Gaston. 2000. Flexibility, efficiency, and accountability: adapting reserve selection algorithms to more complex conservation problems. *Ecography* 23: 565-574.
- Soulé, M.E. & M.A. Sanjayan. 1998. Conservation targets: do they help? *Science* 279: 2060-2061.
- Tuomisto, H., K. Ruokolainen & M. Yli-Halla. 2003. Dispersal, environment, and floristic variation of western Amazonian forests. *Science* 299: 241-244.
- Underhill, L.G. 1994. Optimal and suboptimal reserve selection algorithms. *Biological Conservation* 70: 85-87.
- Veloso, H.P., A.L.R. Rangel Filho & J.C.A. Alves. 1991. Classificação da vegetação brasileira, adaptada a um sistema universal. IBGE.
- Warman, L.D., A.R.E. Sinclair, G.G.E. Scudder, B. Klinkenberg & R.L. Pressey 2004. Sensitivity of systematic reserve selection to decisions about scale, biological data, and targets: case study from southern British Columbia. *Conservation Biology* 18: 655-666.
- Wikramanayake, E., E. Dinnerstein, C. Loucks, D. Olson, J. Morrison, J. Lamoreux, M. McKnight & P. Hedao. 2002. Ecoregions in ascendance: reply to Jepson and Whittaker. *Conservation Biology* 16: 238-243.
- Williams, P.H., D. Gibbons, C. Margules, A. Rebelo, C. Humphries & R. Pressey 1996. A comparison of richness hotspots, rarity hotspots, and complementary areas for conserving biodiversity of British birds. *Conservation Biology* 10: 155-174.



# INSTRUÇÕES PARA PUBLICAÇÃO NA REVISTA “MEGADIVERSIDADE”

Megadiversidade é uma publicação semestral editada pela Conservação Internacional e tem como objetivo principal publicar artigos relacionados com a conservação da biodiversidade no Brasil e no mundo. Cada número da revista trata de temas específicos, previamente selecionados pelo Conselho Editorial. De modo geral, os artigos serão convidados, mas todos passarão por um processo de revisão tanto por revisores externos como pelo Conselho Editorial.

## Orientação para a preparação dos manuscritos

**Título:** deve ser conciso e informativo. O mais curto possível. Deve ser em negrito e centralizado.

**Autores:** nomes completos dos autores alinhados a esquerda. No final de cada capítulo, cada autor deve apresentar um resumo (até 6 linhas) do seu curriculum-vitae, enfatizando local de nascimento, titulação, posição quando escreveu o capítulo, instituição e experiência. Acrescentar também o endereço institucional.

**Texto:** o(s) autor(es) devem organizar os artigos de forma mais apropriada, estabelecendo seções compatíveis com o desenvolvimento do texto. As seções não devem ser enumeradas. O texto deve ser corrido e as tabelas e figuras mostradas depois dele. Todo o texto deve ser escrito em Times New Roman 12, alinhado à esquerda e com espaçamento duplo. Usar o papel A4. O título das seções (introdução, material e métodos, etc...) deve ser em negrito e alinhado a esquerda. As sub-seções (quando for o caso!) devem ser em itálico e alinhadas à esquerda.

**Resumo:** esta seção deve conter de forma resumida os principais objetivos, metodologia, resultados e conclusões do estudo. Não exceder 300 palavras.

**Abstract:** esta seção deve conter de forma resumida e em inglês os principais objetivos, resultados e conclusões do capítulo. Não exceder 300 palavras.

**Referências bibliográficas:** a acurácia das referências é de responsabilidade dos autores. As referências deverão ser checadas no texto para assegurar (a) a correta grafia dos nomes dos autores e as datas, (b) que todos autores mencionados no texto são dados na lista de referências e vice versa. O título inteiro da referência deve ser citado assim como o nome completo do periódico. Os títulos dos livros devem ser seguidos da editora e do local de publicação. No texto, as referências devem ser organizadas cronologicamente por autor seguido da data. Deve ser utilizado a, b ou c após o ano para distinguir entre publicações do mesmo autor no mesmo ano.

No decorrer do texto a forma de citação da referência deve ser:

i – Dois autores: usar ambos os nomes seguidos do ano

ii – Três ou mais autores: Fornecer o nome do primeiro autor seguido de et al. e data.

Na lista, as referências devem ser organizadas, primeiro, em ordem alfabética dos autores, e, em seguida, por ordem cronológica para autores que são citados repetidamente.

Exemplos:

Andrew, D. 2001. Post fire vertebrate fauna survey: Royal and Heathcote national parks Garawarra State Recreation Area. New South Wales National Parks and Wildlife Service, Hurstville, Austrália.

Baker, J. R. 2000. The Eastern Bristlebird: cover dependant and fire sensitive. *Emu* 100: 286-298.

Keith, D. A., W. L. McCaw & R.J. Whelan. 2001. Fire regimes in Australian heathlands and their effects on plants and animals. In: R.A. Bradstock, J. Willians & A.M. Gills (eds). *Flammable Australia: the fire regimes and biodiversity of a continent*. pp 199-237. Cambridge University Press, Cambridge, Reino Unido.

**Tabelas e figuras:** os manuscritos podem conter figuras e tabelas. No corpo do manuscrito cada figura deve ser apresentada em página separada com os créditos e a indicação do número da figura. O local de inserção deve estar indicado no decorrer do texto. As figuras deverão ser preferencialmente em preto e branco. Serão publicadas em cores somente quando a compreensão da informação estiver comprometida. Iniciar cada tabela e figura em uma nova página. As legendas devem ser claras e auto-explicativas, ou seja, devem conter todas as informações necessárias para o leitor entendê-las sem necessitar do texto. Primeiro são mostradas todas as tabelas, depois as figuras. As legendas das tabelas são em cima e das figuras embaixo. Fotos e ilustrações são consideradas figuras. Não mostrar o mesmo dado em forma de tabela e figura. Tanto as tabelas como as figuras devem ser numeradas por números arábicos (e.g., Tabela 1, Figura 1, etc.).

**Apêndices:** só utilizar no caso de tabelas muito grandes, com informações relevantes mais não necessárias para o entendimento do capítulo.

**Nomes científicos:** dar o nome latino completo de cada espécie, seguido da autoridade que a descreveu (quando a tradição no respectivo grupo taxonômico assim exigir) e a família a que pertence na primeira citação. Citar o nome popular da espécie, quando existir, na primeira vez que a espécie é mencionada.

**Unidades, símbolos e números:** usar o sistema internacional. Em expressões matemáticas usar símbolos e não abreviações. Escrever os números de um a nove por extenso e de 10 em diante usar algarismos.

**Provas:** um conjunto de provas será enviado ao primeiro autor, que deverá retorná-las em um prazo de 7 dias. As correções deverão se limitar àquelas sugeridas pelos revisores e a erros tipográficos.

**Direitos autorais:** é condição para a publicação que o autor assine de o termo de cessão de direitos autorais à Conservação Internacional. Ao assinar o termo de cessão o autor tem o direito de usar seu próprio material indicado que a revista é o local original de publicação do artigo. A Conservação Internacional se comprometerá a ceder 25 separatas de cada artigo, as quais serão enviadas somente ao primeiro autor. Cópias adicionais serão enviadas com um custo adicional. Os autores também receberão uma cópia do artigo no formato pdf.

**Submissão:** os artigos deverão ser submetidos à Megadiversidade no seguinte endereço: Av. Getúlio Vargas, 1300 – 7º andar, Belo Horizonte – MG – Cep: 30112-021. E-mail: megadiversidade@conservacao.org. Os autores deverão enviar 3 cópias em papel A4 e uma em formato eletrônico (de preferência Word 2000) do manuscrito em português ou inglês, em espaçamento duplo. Os manuscritos serão submetidos a revisores selecionados pelo Conselho Editorial.



**CONSERVAÇÃO**  
INTERNACIONAL

BRASIL



# MEGADIVERSIDADE

Volume 3 | Nº 1-2 | Dezembro 2007

## Modelagem ambiental e conservação da biodiversidade

### SUMÁRIO

- 3 **Apresentação**  
JOSÉ MARIA CARDOSO DA SILVA, ANA LUIZA KERTI MANGABEIRA ALBERNAZ & THAÍS PACHECO KASECKER
- 5 **Modelos lineares como ferramentas para a modelagem da distribuição de espécies**  
GUILHERME MOURÃO & WILLIAM E. MAGNUSSON
- 13 **Modelos de classificação em árvore para a localização de áreas de ocorrência potencial de espécies**  
ADRIANA PAESE, SIDNEY DORNELLES, JOSÉ EDUARDO DOS SANTOS & JOSÉ SALATIEL RODRIGUES PIRES
- 25 **O uso de regressão logística para espacialização de probabilidades**  
EDUARDO M. VENTICINQUE, JULIANA STROPP CARNEIRO, MARCELO PAUSTEIN MOREIRA & LEANDRO FERREIRA
- 38 **Uso de modelos aditivos generalizados na estimativa da distribuição potencial de espécies**  
PAULO DE MARCO JÚNIOR
- 46 **Algoritmo Genético para Produção de Conjuntos de Regras (GARP)**  
RICARDO SCACHETTI PEREIRA & MARINEZ FERREIRA DE SIQUEIRA
- 56 **Modelos de metapopulação**  
DOUGLAS F. M. GHERARDI
- 64 **O uso de modelos em ecologia de paisagens**  
JEAN PAUL METZGER, MARINA ANTONGIOVANNI DA FONSECA, FRANCISCO JOSÉ BARBOSA DE OLIVEIRA FILHO & ALEXANDRE CAMARGO MARTENSEN
- 74 **Modelagem de dinâmica de paisagem: concepção e potencial de aplicação de modelos de simulação baseados em autômato celular**  
BRITALDO SILVEIRA SOARES FILHO, GUSTAVO COUTINHO CERQUEIRA, WILLIAM LEITE ARAÚJO & ELIANE VOLL
- 87 **Planejamento sistemático para a conservação na Amazônia brasileira – uma avaliação preliminar das áreas prioritárias de Macapá-99**  
ANA LUIZA KERTI MANGABEIRA ALBERNAZ & MANUELLA ANDRADE DE SOUZA

ISSN 1808-3773



9 771808 377755