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

ADRIANA PAESE^{1*}

SIDNEY DORNELLES²

JOSÉ EDUARDO DOS SANTOS²

JOSÉ SALATIEL RODRIGUES PIRES²

¹ Conservação Internacional, Minas Gerais, Brasil.

² 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 m².

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 (x²), os quais são aplicados sequencialmente. 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.

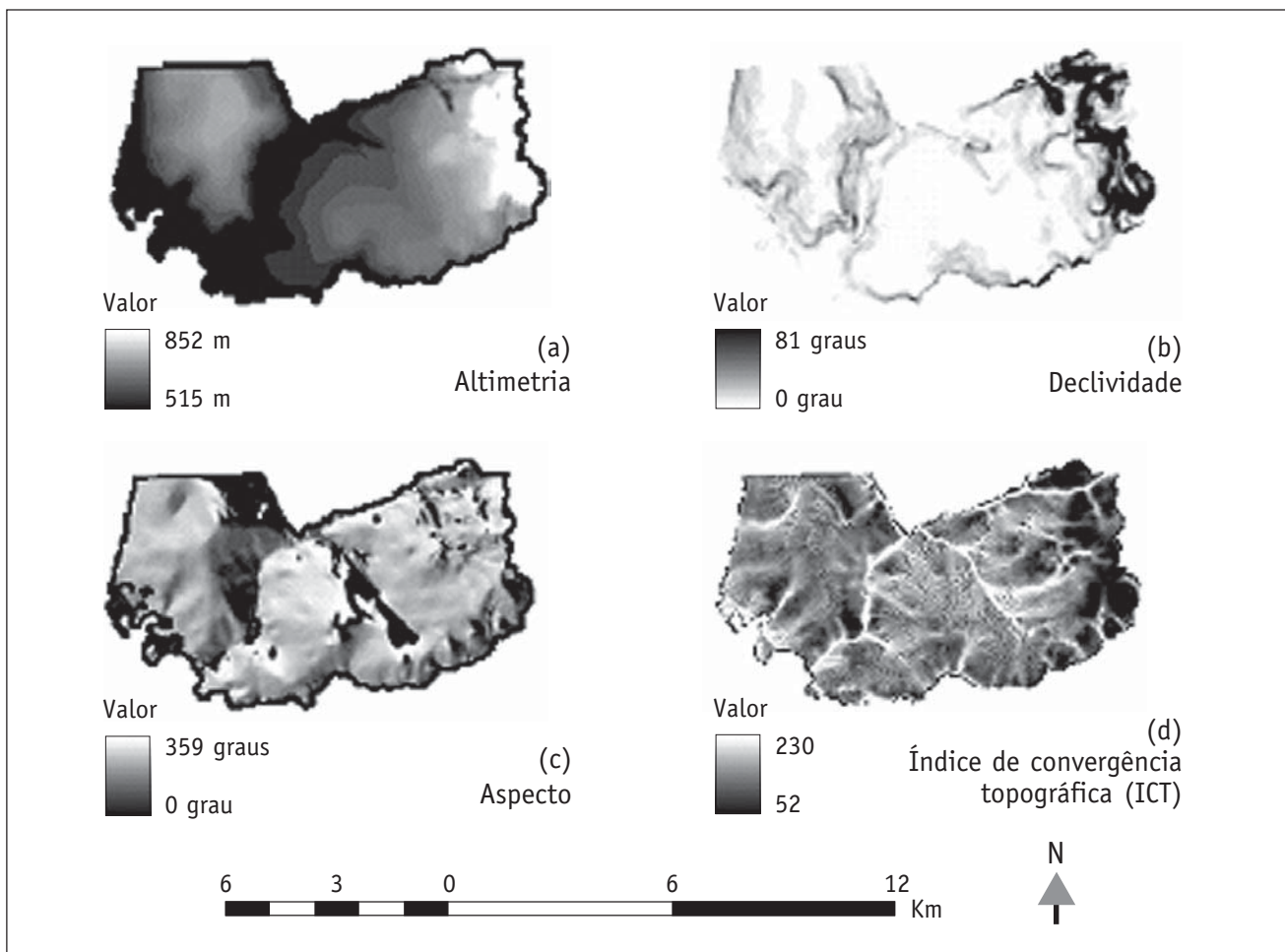


FIGURA 2 – Variáveis topográficas utilizadas nos modelos de classificação em árvore: (a) altimetria; (b) declividade; (c) aspecto; e (d) índice de convergência topográfica (ICT).

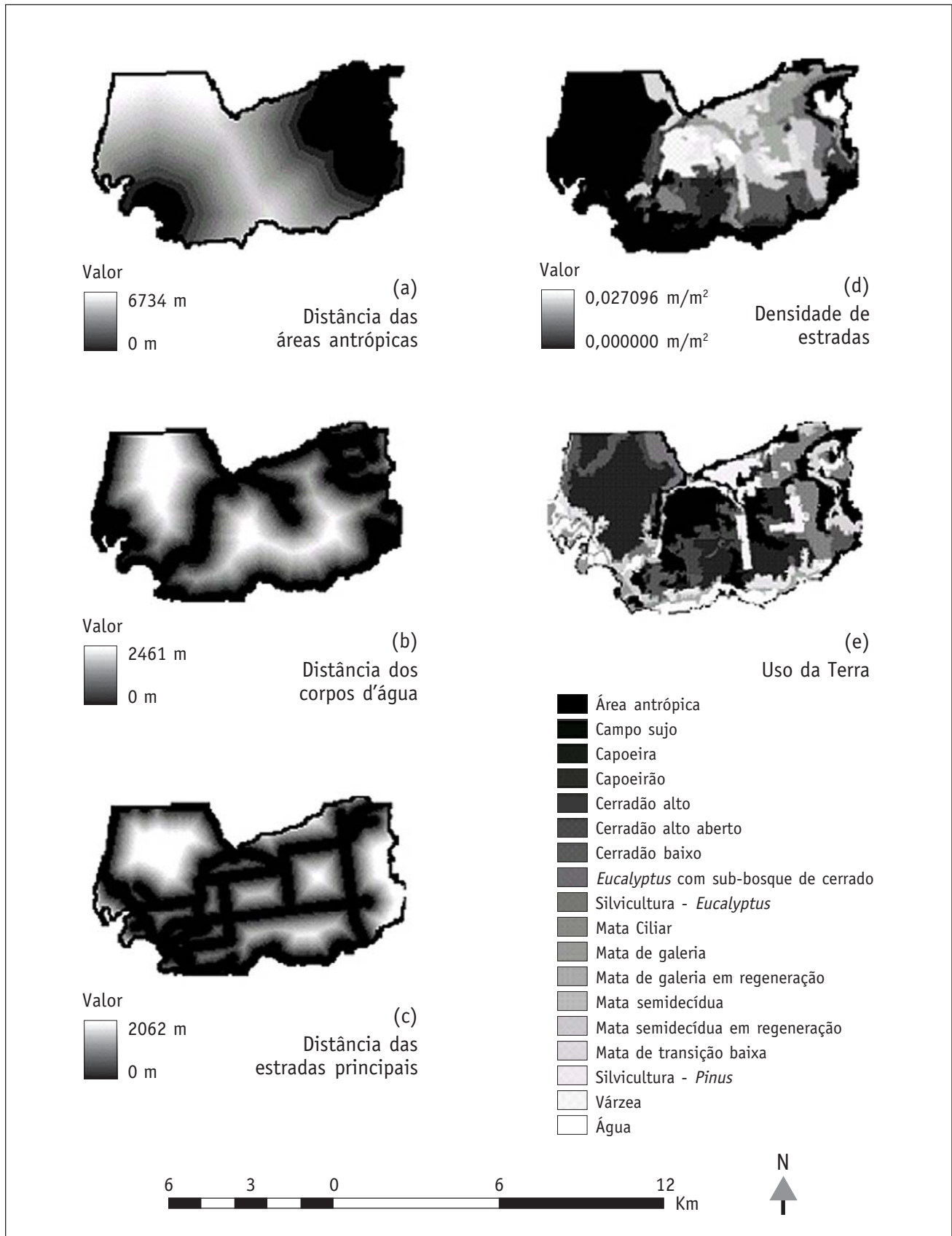


FIGURA 3 – Variáveis ambientais utilizadas nos modelos de classificação em árvore: (a) distância das áreas antrópicas; (b) distância dos corpos d'água; (c) distância das estradas principais; (d) densidade de estradas; e (e) uso da terra.

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/m² a 0,02796m/m². 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/m². 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/m² 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.

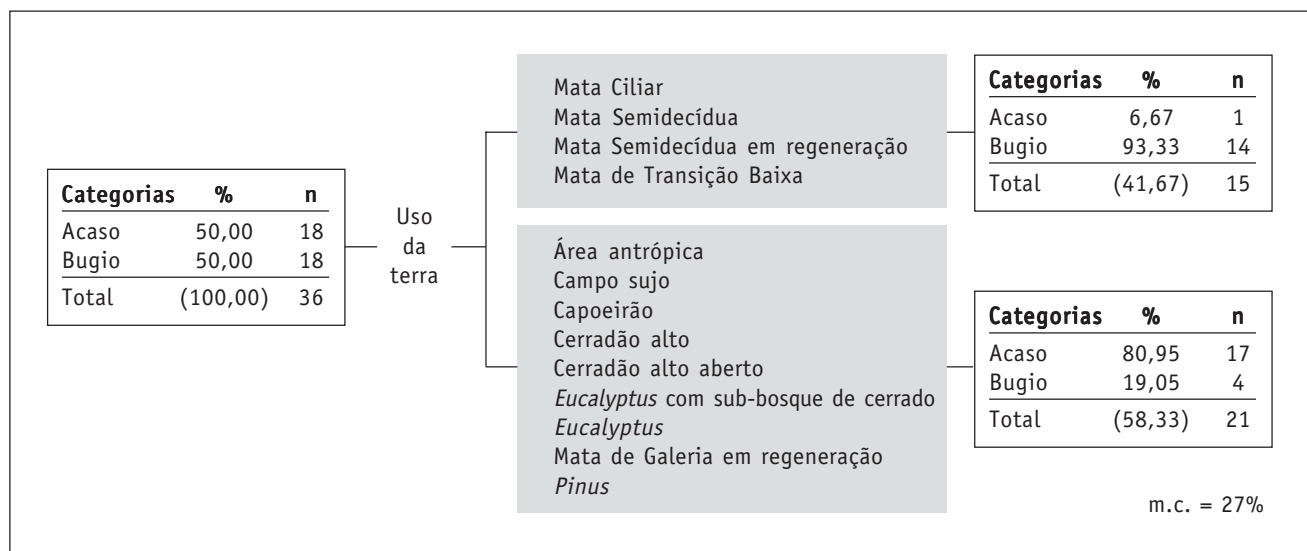


FIGURA 4 – Modelo de classificação em árvore da distribuição do bugio (*Alouatta caraya*) na área de estudo.

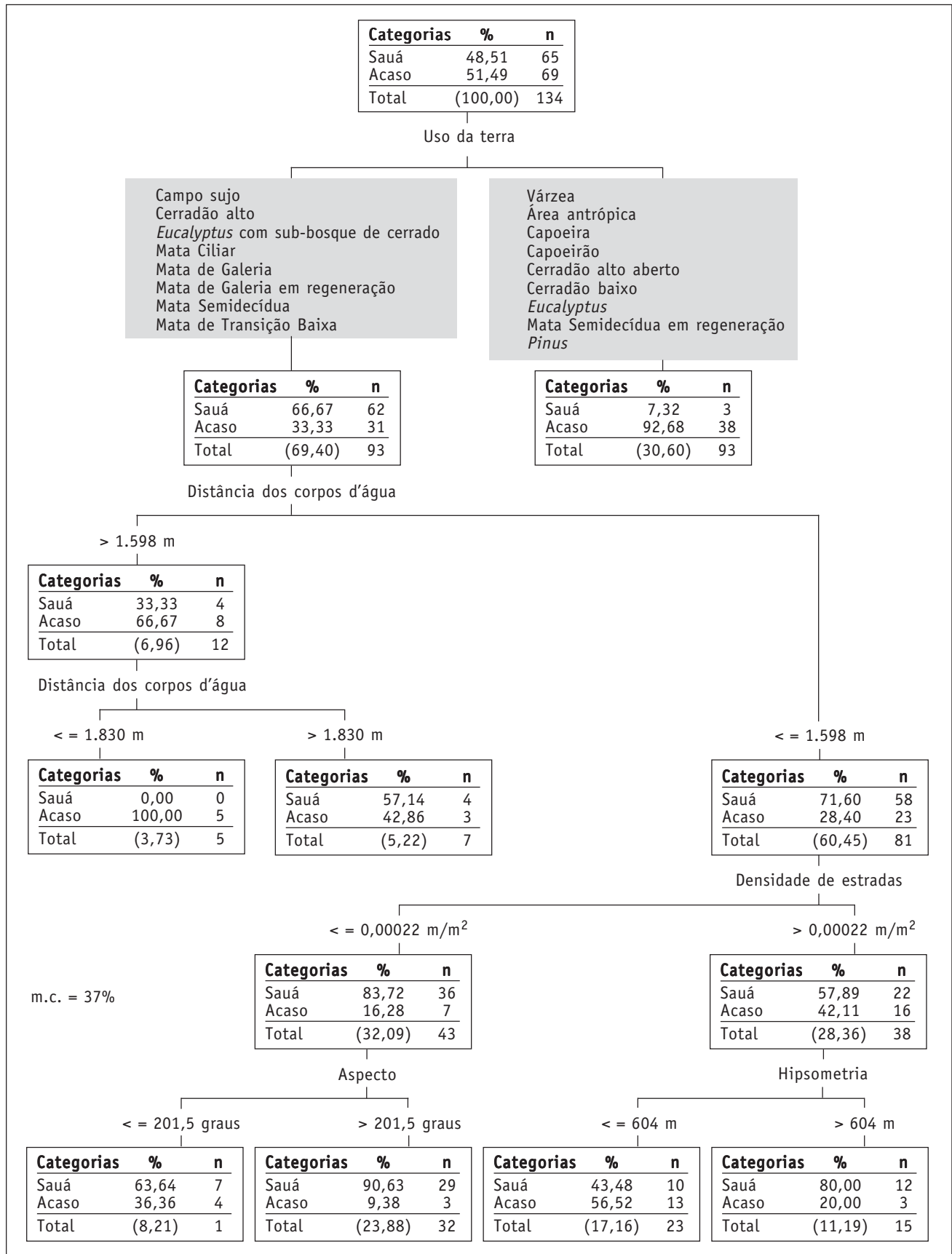


FIGURA 5 – Modelo de classificação em árvore para o sauá (*Callicebus nigrifrons*) na área de estudo.

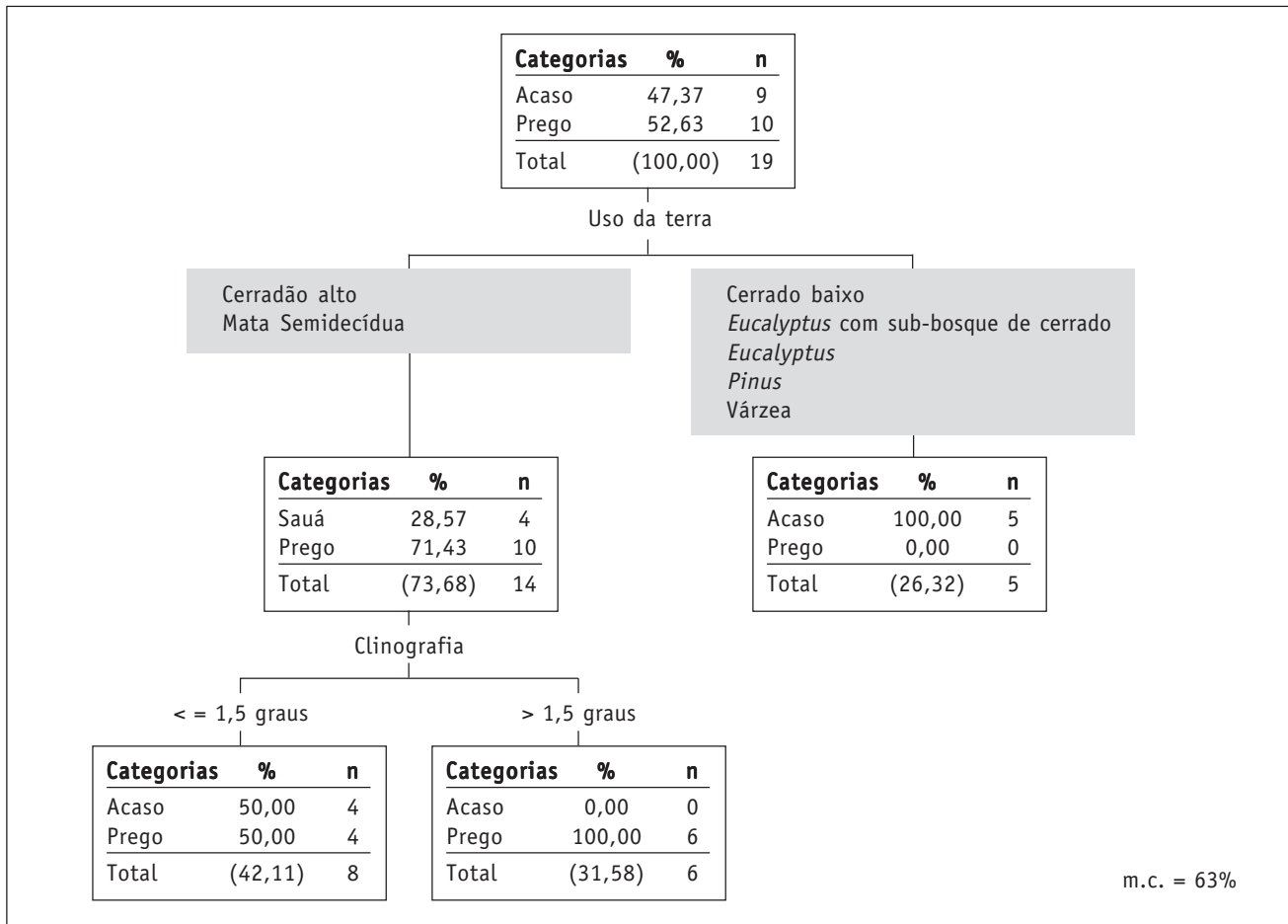


FIGURA 6 – Modelo de classificação em árvore da distribuição do macaco-prego (*Cebus nigritus*) na área de estudo.

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

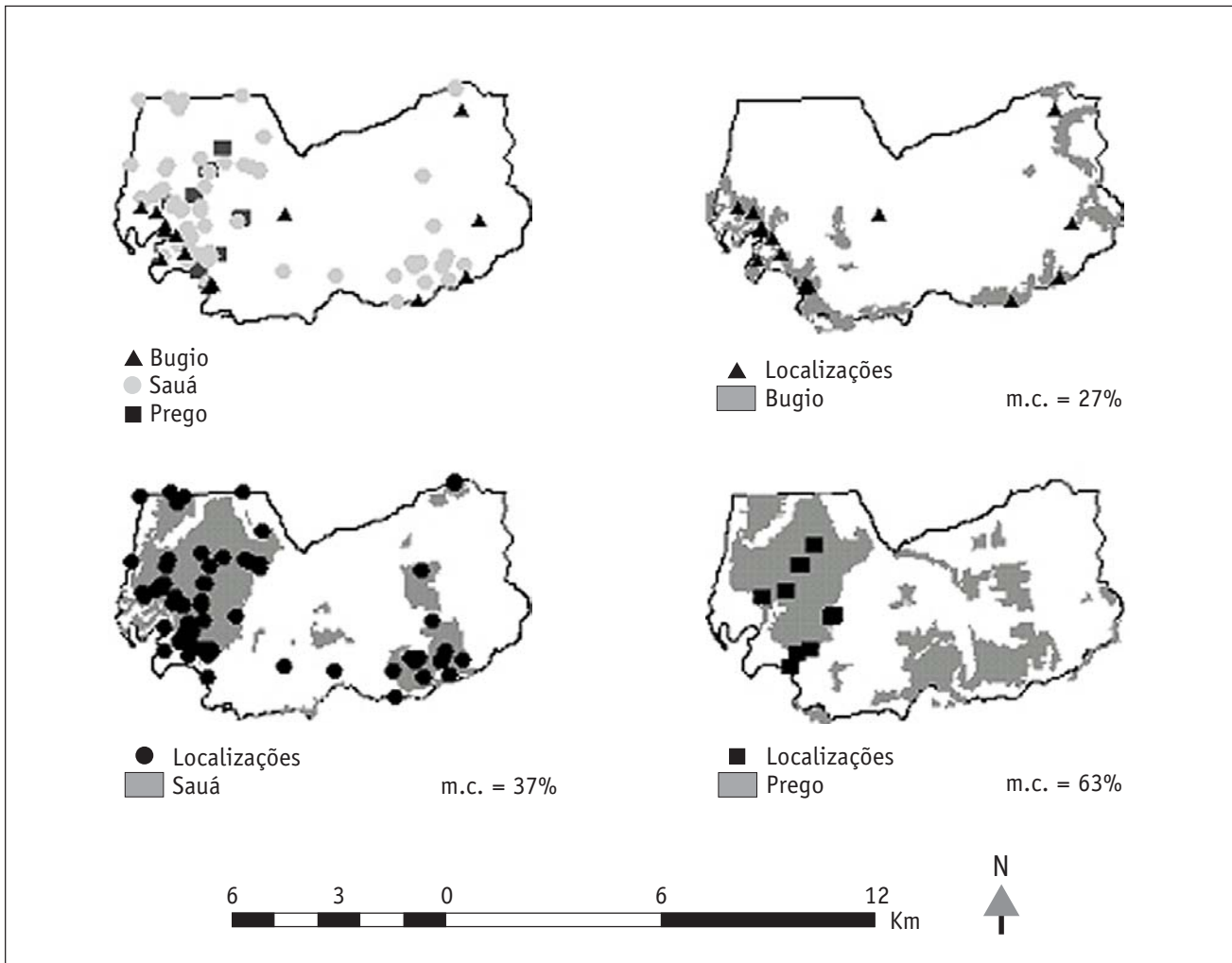


FIGURA 7 – Localização das três espécies de primatas (Dorneles, 2000), projeção das áreas de maior probabilidade de localização destas espécies na área de estudo e taxas de má classificação (m.c.) dos modelos de classificação em árvore.

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.
- Vayssières, 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.