

Uso de Maxent e Garp para modelagem de distribuição potencial espécies pioneiras no estado de Minas Gerais

Guilherme Leite Nunes Coelho¹
Luis Marcelo Tavares de Carvalho²
Nathália Silva de Carvalho¹

¹ Universidade Federal de Lavras - UFLA
37200-000 - Lavras - MG, Brasil
{guilhermeleite.bio, nathalia.bioufla}@gmail.com.br

² Universidade Federal de Lavras - UFLA
Caixa Postal: 3037 - 37200-000 - Lavras - MG, Brasil
passarinho@dcf.ufla.br

Abstract. Model the potential distribution can be a powerful tool for conservation of natural resources programs. The aim of this work is to evaluate the performance of this two algorithms Garp and Maxent in mapping environmental suitability of two pioneer species *Bowdichia virgilioides* and *Acosmium dasycarpum* in the state of Minas Gerais. Data were used from many different sources to build these four models. Fourteen variables were chosen to represent environmental conditions of places that species were found. Temperature, rainfall, altitude, slope, drainage density, distance drain and vegetation indices. Models produced by Maxent outperformed the results of Garp. Although both showed satisfactory and his maps are considered reliable. Modeled species showed potential distribution under the ecological conditions described in the literature. The results of MaxEnt identified areas with high environmental similarity to the conditions found in the records of occurrence of the species. We also found that vegetation index gave the highest contribution for two models and two species. Another important variable is solar radiation, because in both models gave substantial contribution and without this variable, the model will not be good. Mean diurnal temperature range gave an important contribution to *Acosmium dasycarpum* and actual evapotranspiration was more importante to *Bowdichia virgilioides*.

Palavras-chave: pioneer species, specie distribution modelling, Maxent, GARP, openModeller, potential distribution, espécies pioneiras, modelagem de distribuição de espécies, distribuição potencial.

1. Introdução

As políticas públicas atuais de conservação da biodiversidade têm crescido em escala mundial, resultante de estudos e um forte apelo popular é cada dia mais comum a aprovação de grandes projetos de conservação e recuperação de áreas naturais. Nesse cenário existe uma grande demanda de estudos que indiquem espécies ideais para recuperação de cada região do país.

O estado de Minas Gerais tem grande destaque nesse cenário, visto que abriga três importantes biomas, Cerrado, Mata Atlântica e Caatinga. Segundo Myers et al. (2000) Cerrado e Mata Atlântica são os únicos “hotspots” brasileiros, tornando imprescindíveis ações de conservação e recuperação desses biomas.

Na atualidade os modelos preditivos de distribuição de espécie são uma importante ferramenta para o gerenciamento de recursos naturais, e em muitas outras áreas como biogeografia, evolução e ecologia (Anderson et al, 2003, Mortara e Valeriano, 2001; Mortara, 2000). Diversos trabalhos com esses modelos têm identificado áreas de importância para biodiversidade e auxiliado na identificação de alterações temporais na distribuição de determinadas espécies (Peterson 2003). Segundo Siqueira (2005) a modelagem preditiva é eficiente na definição de áreas potenciais para re-introdução de espécies.

A modelagem preditiva de distribuição baseia se no conceito de nicho fundamental que caracteriza se pelas condições abióticas que os indivíduos necessitam para sobreviverem.

Esses fatores são temperatura, precipitação, topografia dentre outros. Durante o processo de modelagem baseia-se nesse conjunto de condições ecológicas onde espécie pode ser manter e locais que são semelhantes a essas condições ecológicas representando o potencial de distribuição da espécie. (Phillips et al. 2006) Dessa forma, os resultados finais do modelo são os possíveis locais de ocorrência da espécie de interesse.

Existem inúmeros tipos de algoritmos utilizados para realizar a modelagem de distribuição. O algoritmo genético conhecido como Genetic Algorithm for Rule-set Production (GARP) é bastante utilizado para prever a distribuição potencial das espécies (Peterson 2001, Stockwell e Noble 1992, Stockwell e Peterson 2002, Peterson e Cohoon 1999). Muito citado na literatura esse algoritmo apresentou bom desempenho em diversos tipos de estudo como análises biogeográficas, evolução, ecologia, conservação e análises de espécies invasoras. (Peterson e Vieglais 2001, Anderson e Martinez-Meyer 2004, Ortega-Huerta e Peterson 2004, Gaubert et al. 2006).

O algoritmo de Máxima Entropia – Maxent também se destaca na literatura (Phillips et al, 2006; Elith et al, 2007; Iwashita, 2008). Assim como o GARP necessita apenas de dados de presença da espécie o que se configura numa grande vantagem visto que dados de ausência são em geral raros e imprecisos.

Nesse presente estudo optou-se por realizar a modelagem preditiva com duas espécies pioneiras objetivando indicar espécies para colonização e recuperação de áreas degradadas. Segundo Lorenzi 1992 *Bowdichia virgilioides* é bastante utilizada em programas de reflorestamento e na recuperação de áreas degradadas de preservação permanente. Dessa forma a modelagem pode auxiliar a indicar áreas de potencial de colonização com maior precisão, diminuindo possíveis erros de planejamento.

2. Materiais e Métodos

Os mapas de potencial de distribuição foram elaborados através dos algoritmos GARP (algoritmo genético para produção de regras) e Maxent (algoritmo de máxima entropia). As ocorrências das espécies georreferenciadas foram extraídas do inventário florestal de Minas Gerais e posteriormente associadas a dados ambientais abióticos e bióticos. Assim para cada espécie foi avaliado a influência das variáveis ambientais na sua distribuição e qual algoritmo apresentou melhor desempenho para toda a área do estado de Minas Gerais.

O estado de Minas Gerais apresenta uma diversidade de fitofisionomias dos biomas Cerrado, Mata Atlântica e Caatinga, com a necessidade crescente de conhecimento sobre os biomas brasileiros devido a alta degradação. O estado torna-se uma importante área de estudo para gerar informações que possam auxiliar projetos de conservação.

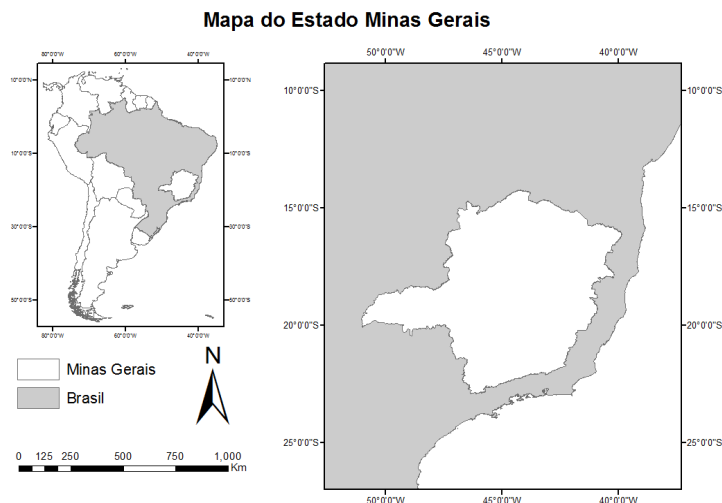


Figura 1. Mapa de localização do Estado de Minas Gerais

2.1 Espécies analisadas

A espécie *Bowdichia virgilioides* Kunt. é conhecida popularmente como sucupira-preta, sucupira-do-campo, dentre outros nomes. É uma espécie arbórea considerada pioneira adaptada a terrenos secos e pobres. (Brandão e Ferreira 1991, Lorenzi, 1992) Segundo Lorenzi, 1992 é bastante utilizada em programas de reflorestamento e recuperação de áreas degradadas. A espécie *Acosmium dasycarpum* (Vog) Yakovlev é conhecida popularmente por perobinha do campo, chapada, dentre outros nomes. É uma planta de características exclusivas de cerrado sensu stricto e cerradão, sua distribuição se dá na parte central do Brasil. Ocorre em estados brasileiros como Bahia, Minas Gerais, São Paulo, Mato Grosso e Goiás. (Lorenzi, 1998) Os registros das espécies foram extraídos do inventário florestal de Minas Gerais iniciado no ano de 2003 que consiste no mapeamento da flora nativa e dos reflorestamentos existentes no estado. O inventário foi realizado pela Universidade Federal de Lavras (UFLA) com objetivo de construir um banco de dados essencial para planejamento de gestão ambiental no estado. No presente estudo foi utilizado apenas dados de presença. Os registros das espécies foram extraídos dos fragmentos inventariados, em cada parcela que as espécies de interesse estiveram presentes foram considerados um ponto de ocorrência. Dessa forma foi obtido um total de 683 ocorrências para *Bowdichia virgilioides* Kunt. e 527 para a *Acosmium dasycarpum* (Vog) Yakovlev.

2.2 Variáveis ambientais

A seleção de variáveis foi dividida em algumas etapas. Em primeiro foi realizado uma busca na literatura para identificar comportamentos ecológicos da espécie e determinar possíveis variáveis que melhor explicassem a distribuição das plantas em estudo. Em seguida foi realizado uma análise de correlação utilizando uma ferramenta chamada ENMTools's core methods (Warren et al. 2008). Foram selecionadas ao todo 14 variáveis a fim de caracterizar as condições abióticas em que as espécies estavam submetidas. As variáveis selecionadas foram temperatura, precipitação, topografia, radiação de disponibilidade de água no solo. Todas as camadas ambientais foram re-amostradas para 270m.

As variáveis climáticas e bioclimáticas foram obtidas do a partir do projeto WorldClim (Hijmans et al, 2005) com resolução espacial de 30 arc-segundos (aproximadamente 1km). As variáveis selecionadas foram Variação Diurna de Temperatura (bio2), Isotermalidade (bio3), Sazonalidade da Temperatura (bio4), Precipitação do mês mais seco (bio14).

A distância vertical à drenagem mais próxima é uma variável disponibilizada pelo projeto Ambdata coordenado pelo Instituto de Pesquisas Espaciais. O algoritmo HAND (Height Above the Nearest Drainage) foi desenvolvido por Rennó et al. 2008 para calcular a distância vertical de cada ponto em relação à drenagem mais próxima o que indica disponibilidade de água no solo.

A densidade de drenagem deriva dos dados projeto HydroSHEDS (Lehner et al. 2006). A partir desses dados a densidade foi calculada utilizando o algoritmo de Kernel (Silverman, 1986, p.76, equação 4.5). O projeto HydroSHEDS gerou os dados de drenagem a partir do SRTM (Shuttle Radar Topographic Mission) de resolução espacial original de 15 arc-segundos, ou aproximadamente 500m.

O projeto Ambdata citado anteriormente disponibiliza uma variedade grande de dados ambientais. Os dados de altitude e declividade foram gerados a partir de dados do SRTM) de resolução espacial de 3 arc-segundos (~90m) e resolução vertical (altura) de 1 m. Ambos os dados são utilizados para caracterização topográfica da área estuda.

Evapotranspiração real é a quantidade de água que removida do solo através dos processos de evaporação e transpiração. Os dados foram obtidos do projeto Climond – Global

climatologies for bioclimatic modelling versão v 1.1 (Kriticos et.al, 2012). Nesse mesmo projeto foram extraídos dados de radiação solar como Radiação média anual (bio20), Mês com maior radiação (bio21), Mês com menor radiação (bio22), Sazonalidade da radiação (bio23).

O índice de vegetação NDVI: Normalized Difference Vegetation Index foi utilizado por ser considerado um indicador de variações de densidade de vegetação principalmente em estudos de áreas de grande extensão (Freitas et al. 2005; Kawashima et al. 2007). O NDVI é derivado de imagens de satélites provenientes do sensor AVHRR – Advanced Very High Resolution Radiometer (<<http://www.edu.usgs.gov/products/satellite/avhrr.html>>) na resolução espacial de 1 km². Foram utilizadas quatro imagens NDVI uma para cada estação, a fim de indicar o comportamento da vegetação ao longo de 2004.

2.3 Algoritmos

O primeiro algoritmo utilizado no estudo foi MAXENT – Maximum entropy modelling (Phillips et al. 2004; Phillips et al. 2006), trabalho realizados em modelagem indicam este algoritmo apresenta melhor desempenho em estudos com dados somente de presença. O algoritmo é implementado em um software de mesmo nome. A probabilidade de ocorrência se dá pela distribuição mais próxima da distribuição uniforme, considerando um conjunto de restrições que são as informações (variáveis ambientais) referentes ao alvo (espécies) (Dudík et al. 2004; Phillips et al. 2004; Phillips et al. 2006).

Neste estudo foi utilizada a versão 3.3.3k do software, foram mantidos os valores padrões do programa. No entanto o limite de corte foi alterado para o valor fixo acumulado de 10, técnica que aceita 10% de taxa de omissão dos pontos de presença. É uma metodologia arbitrária que identifica regiões de alta similaridade ambiental com o conjunto de treino. (Kamino, 2009)

O segundo algoritmo é o GARP – Genetic Algorithm for Rule-Set Prediction (Stockwell & Noble 1992; Stockwell & Peters 1999; Muñoz et al. 2009) foi empregado utilizando a plataforma Openmodeller versão 1.1.0 (<http://openmodeller.cria.org.br/>), na categorial GARP with best subsets – DesktopGARP implementation. Foi mantido todos os parâmetros padrões do programa. Este algoritmo tem como base a inteligência artificial em um processo iterativo é associado os pontos de ocorrência e as camadas ambientais, gerando um conjunto de regras que caracteriza se na área de potencial de ocupação da espécie (Peterson et al. 1999)

Durante as análises do maxent foi avaliado a importância de cada variável para a distribuição potencial da espécie. Essa avaliação é feita por um teste jackknife sobre os dados de testes e de treinamento e vem implementada no software no maxent. A curva ROC (receiver operating characteristic) foi utilizada para avaliar a qualidade do modelo, nesses casos de modelos que com dados de presença, é avaliado se a classificação é melhor que uma classificação aleatória (Phillips et al, 2006).

3. Resultados s Discussão

O algoritmo Maxent apresentou bons resultados. Com base no teste jackknife implementado no software Maxent foi possível mensurar a contribuição de cada variável para o modelo. No experimento realizado com a espécie *Bowdichia virgilioides* o NDVI apresentou maior contribuição ao modelo na distribuição da espécie com 26,9%, seguido das variáveis, Radiação média anual (bio20) com 18,9% e evapotranspiração real com 11,5%. Assim a omissão dessas prejudicaria os resultados finais. Considerando a distribuição da espécie *Acosmium dasycarpum* as três variáveis que apresentaram maiores contribuições foram NDVI, Radiação média anual (bio20), Variação Diurna de Temperatura (bio2), com 19,2%, 18,2% e 15,2% respectivamente.

Acurácia do modelo foi considerada muito boa. Para o modelo da *Bowdichia virgilioides* o valor da AUC foi de 0,95 para os dados de treino e 0,97 para a espécie *Acosmium dasycarpum*. Os mapas da (Figura 2a) da (Figura 2b) mostram os locais no estado de Minas Gerais onde existe adequabilidade ambiental para as duas espécies. Distribuição que esta de acordo com a literatura. Segundo Lorenzi, 1998 a espécie *Acosmium dasycarpum* é exclusiva de cerrado e cerradões, observando o mapa da Figura 2a nota se que a distribuição potencial ocorre exatamente no bioma cerrado. A espécie *Bowdichia virgilioides* é citada na literatura com uma espécie adaptada a terrenos secos e pobres (Lorenzi, 1992). Observando o mapa da espécie de adequabilidade ambiental na Figura 2b nota se que os maiores valores de adequabilidade foram encontrados em regiões com baixos valores de precipitação e altos índices de evapotranspiração condições de locais secos.

Nas Figuras (2c) e (2d) pode ser observado os mapas gerados pelo algoritmo GARP para as espécies *Acosmium dasycarpum* e *Bowdichia virgilioides* respectivamente. Nota se uma grande diferença entre os mapas gerados pelo Maxent. Em geral as áreas de alta adequabilidade ambiental para as espécies são maiores. Isso deve se a técnica utilizada no processamento do Maxent foi adotada um valor fixo acumulado de 10, assim apenas regiões de alta similaridade ambiental foram mapeadas. (Kamino, 2009).

Os resultados gerados pelo algoritmo GARP foram considerados bons, para a espécie *Acosmium dasycarpum* a AUC obteve 0,82 já para outra espécie obteve se o valor de 0,74. Apesar de o algoritmo ter predito uma área potencial bem maior que o Maxent ainda sim houve coerência com a literatura. A espécie *Acosmium dasycarpum* que é exclusiva de cerrado obteve a maior parte da sua área predita no bioma de cerrado. Já a *Bowdichia virgilioides* que apresenta uma distribuição mais ampla, apresentou uma área potencial também mais abrangente.

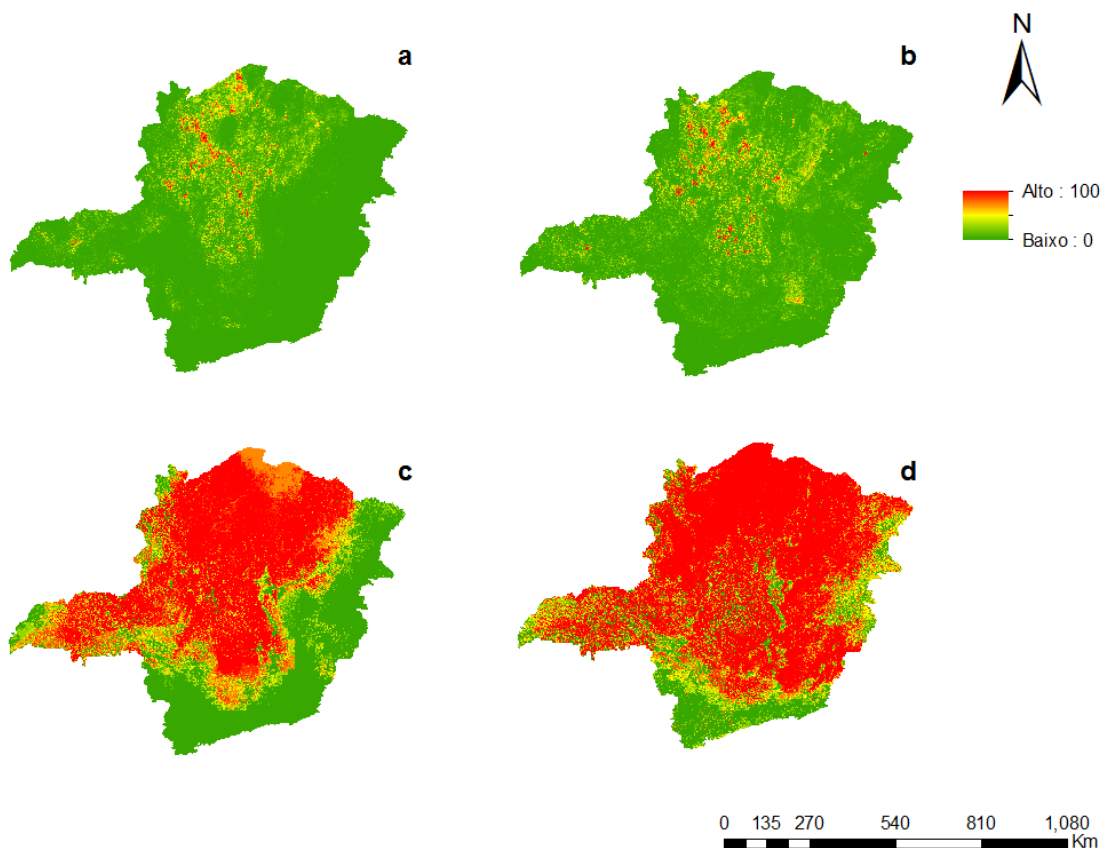


Figura 2. Mapas de predição potencial de distribuição geográfica resultantes da modelagem com os algoritmos Maxent e GARP. (a) Mapa de distribuição potencial para *Acosmium dasycarpum* utilizando o algoritmo Maxent. (b) Mapa de distribuição potencial para *Bowdichia virgilioides* utilizando o algoritmo Maxent. (c) Mapa de distribuição potencial para *Acosmium dasycarpum* utilizando o algoritmo GARP. (d) Mapa de distribuição potencial para *Bowdichia virgilioides* utilizando o algoritmo GARP.

4. Conclusão

Em linhas gerais ambos os algoritmos apresentaram um bom desempenho. Pode se afirmar que o Maxent apresentou melhores resultados que o GARP devido aos valores de AUC que foram maiores do que do algoritmo genético. Assim indica se o uso do Maxent para buscar áreas com alta similaridade ambiental com áreas onde ocorre a espécie de interesse. O que pode diminuir chance de erro na tentativa de introduzir a espécie no ambiente ou buscar novas populações. O GARP não deve ser descartado já que apresentou bons resultados e é capaz de indicar regiões potencialmente importantes.

Agradecimentos

Os autores agradecem o apoio financeiro promovido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Gostaríamos de agradecer também a Universidade Federal de Lavras, o LEMAF - Laboratório de Estudos e Projetos em Manejo Florestal, CNPQ e a FAPEMIG.

5. Referências Bibliográficas

- Anderson, R. P.; Lew, D.; Peterson, A. T. Evaluating predictive models of species' distributions: criteria for selecting optimal models. **Ecological Modelling**, vol. 162, p. 211-232, 2003.
- Anderson, R. P.; Martinez-Meyer, E. Modeling species' geographic distributions for preliminary conservation assessments: an implementation with the spiny pocket mice (Heteromys) of Ecuador. **Biological conservation**, vol 116,p.167-179, 2004.
- Brandão, M.; Ferreira, P. B. D. **Flora apícola do cerrado**. Informe Agropecuário, Belo Horizonte, v. 15, n.168, p. 4-8, 1991.
- Dudík, M. et al. 2004. Performance guarantees for regularized maximum entropy density estimation. In: Proceedings of the Seventeenth Annual Conference on Computational Learning Theory. **ACM Press**, New York, pp. 655-662.
- Elith, J., Graham, C.H.; NCEAS Modeling Group. Novel methods improve prediction of species? Distributions from occurrence data. **Ecography**, v.29, n. 2, p. 129-151, 2006.
- Freitas, S. M.; De Mello, M. C. S.; Cruz, C. B. M. Relações entre maturidade estrutural da floresta e índices de vegetação na Mata Atlântica. **Anais XII Simpósio Brasileiro de Sensoriamento Remoto**, Goiânia, Brasil, 16-21 abril 2005, INPE, p. 1537-1544. 2005.
- Gaubert P.; Papes, M.; Peterson, A. T. Natural history collections and the conservation of poorly known taxa: Ecological niche modeling in central African rainforest genets (*Genetta* spp.). **Biological Conservation**, vol. 130, p.106 –117, 2006.

Hijmans, R. J.; Cameron, S.; Eparra, J.L.; Jones, P.G; Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. **International Journal of Climatology**, vol. 25, p. 1965-1978, 2005.

Iwashita, F. **Sensibilidade de modelos de distribuição de espécies a erros de posicionamento de dados de coleta**. 2007. 103 p. (INPE-15174-TDI/1291). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2007. Disponível em: <<http://urlib.net/sid.inpe.br/mtcm17@80/2007/06.13.12.04>>. Acesso em: 03 nov. 2008.

Kamino, L. H. Y. **Modelos de distribuição geográfica potencial: aplicação com plantas ameaçadas de extinção da floresta atlântica**. Tese de Doutorado. Belo Horizonte, MG: Universidade Federal de Minas Gerais, 2009.

Kriticos, D.J., Webber, B.L., Leriche, A., Ota, N., Macadam, I., Bathols, J. & Scott, J.K. (2012) CliMond: global high resolution historical and future scenario climate surfaces for bioclimatic modelling. **Methods in Ecology and Evolution** 3: 53-64. DOI: 10.1111/j.2041-210X.2011.00134.x

Lehner, B.; Verdini, K.; Jarvis, A. **HydroSHEDS – Technical Documentation**. Versão 1. 2006, p.27. Disponível em: <http://gisdata.usgs.net/HydroSHEDS/downloads/HydroSHEDS_TechDoc_v10.pdf>. Acesso em: 20 de out de 2008.

Lorenzi H 1998. **Árvores Brasileiras: Manual de Identificação e Cultivo de Plantas Arbóreas do Brasil**. São Paulo: Nova Odessa:Plantarum.

Lorenzi, H. 1992. **Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas nativas no Brasil**. Nova Odessa: Plantarum,. 368 p.

Mortara, M. O. **Modelagem da distribuição potencial do palmitero (euterpe edulis Martius)em região montanhosa a partir de variáveis topográficas**. 2000. 157 p. (INPE-8857-TDI/805). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2000.

Mortara, M. O.; Valeriano, D. M. Modelagem da distribuição potencial do palmitero (Euterpe edulis Martius)a partir de variáveis topográficas. In: Simpósio Brasileiro De Sensoriamento Remoto, 10., 2001, Foz do Iguaçu. **Anais...** São José dos Campos: INPE, 2001. p. 459-471. CD-ROM, On-line. ISBN 85-17-00016-1. (INPE-8292-PRE/4082). Disponível em: <<http://urlib.net/dpi.inpe.br/lise/2001/09.14.15.54>>. Acesso em: 30 out. 2008.

Myers, N.; Mittermeier, R. A.; Mittermeier, C. G.; Fonseca, G. A. B.; Kent, J. Biodiversity hotspots for conservation priorities. **Nature**, v. 403, n. 1, p. 124-32, 2000.

Ortega-Huerta, M. A.; Peterson A. T. Modelling spatial patterns of biodiversity for conservation prioritization in North-eastern Mexico. **Diversity and Distributions**, vol. 10, p. 39-54, 2004.

Peterson, A. T.; Holt, R.D. Niche differentiation in Mexican birds: using point occurrences to detect ecological innovation. **Ecology Letters**, vol. 6, p. 774-782, 2003.

Peterson, A. T. Predicting species' geographic distributions based on ecological niche modeling. **The Condor**, vol. 103, p.599–605, 2001.

Peterson, A. T; Cohoon, K. P. Sensitivity of distributional prediction algorithms to geographic data completeness. **Ecological Modelling**, vol. 117, p. 159–164, 1999.

Peterson, A. T; Vieglais, D. A. Predicting species invasions using ecological niche modeling: new approaches from bioinformatics attack a pressing problem. **Bioscience**, vol 51 , nº 5, p. 363- 371, 2001.

Peterson, A. T. ; Soberon, J.; Sanchez-cordero,V. Conservatism of ecological niches in evolutionary time. **Science**, vol. 285, p. 1265–1267, 1999.

Phillips, S. J.; Anderson, R. P.; Schapire, R. E. Maximum entropy modeling of species geographic distributions. **Ecological Modelling**, vol. 190, p 231-259, 2006.

Philip, M. F. de. Desmatamento da Amazônia. 3º **Simpósio** Brasileiro de Pós-Graduação em Engenharia Florestal /1º Encontro Amazônico de Ciências Florestais, Manaus. 2004.

Rennó, C.D.; Nobre, A.D.; Cuartas, L.A; Soares, J.V.; Hodnett, M.G.; Tomasella, J.; Waterloo, M.J. HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazônia. **Remote Sensing of Environment**, v.112, pp. 3469-3481, 2008.

Silverman, B.W. Density Estimation for Statistics and Data Analysis. **New York: Chapman and Hall**, 1986. 176p.

Siqueira, M. F. **Uso de modelagem de nicho fundamental na avaliação do padrão de distribuição geográfica de espécies vegetais**. 2005. Universidade de São Paulo, São Carlos, SP. Tese de Doutorado, 2005.

Stockwell, D. R.B. ; Peterson, A.T. Effects of sample size on accuracy of species distribution models. **Ecological Modelling**, vol. 148, p. 1–13, 2002.

Stockwell, D. B.; Noble, I. R. Introduction of sets of rules from animal distribution data: A robust and informative method of analysis. **Mathematics and Computers in Simulation**, vol. 33, p. 385-390, 1992.

Warren, D.L., R. E. Glor, and M. Turelli. 2008. **Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution**. **Evolution** 62:2868-2883.