

## **AMBDATA: Variáveis ambientais para Modelos de Distribuição de Espécies (MDEs)**

Silvana Amaral  
Cristina Bestetti Costa  
Luciana Satiko Arasato  
Arimatéa de Carvalho Ximenes  
Camilo Daleles Rennó

Instituto Nacional de Pesquisas Espaciais - INPE  
Caixa Postal 515 - 12227-010 - São José dos Campos - SP, Brasil  
{silvana, cbestetti, luarasato, arimatea, camilo}@dpi.inpe.br

**Abstract.** Species distribution models (SDMs) provide the geographical distribution of a species based on the definition of its potential ecological niche, from the relation between species occurrences and environmental characteristics. As SDMs have been widely used for biodiversity studies at regional scale, there are increasing demands for different environmental data in a standardized format. In order to facilitate the manipulation of these geographical datasets, at this paper we present Ambdata: a website created to provide environmental variables from official data sources useful for SDMs purpose. Environmental variables related to climate (monthly minimum, maximum, and mean temperature; monthly precipitation; and bioclimatic data), topography (altitude, slope, aspect, drainage density, vertical height to the closest drainage), soil and vegetation characteristics, usually used for SDMs, were first acquired, produced and then organized in a geographical information system. Each variable was resampled to spatial resolution of 30 arc-seconds (approximately 1 km), converted to Lat-Long geographical projection, and clipped to the geographical boundaries of Brazil and Brazilian Amazon. Once the ascii-raster grid files were prepared, we organized a webpage called Ambdata that describes the dataset and provides free access to these data. This article contributes to present Ambdata webpage to the scientific community: a database available to assist SDM studies of Amazon and Brazilian species, improving the knowledge about our biodiversity.

**Palavras-chave:** environmental variables, geographical database, species distribution model (SDM), geoprocessing, variáveis ambientais, banco de dados geográficos, modelos de distribuição de espécies (MDE), geoprocessamento.

### **1. Introdução**

Modelos de distribuição de espécies (MDE) podem ser definidos como modelos que relacionam dados de ocorrência ou abundância de espécies em locais específicos (dados de distribuição) a informações das características ambientais daqueles locais (Elith & Leathwick, 2009; Cassini, 2011). Uma vez que se conhece a relação entre a distribuição e um conjunto de variáveis ambientais, os modelos permitem a previsão da distribuição potencial da espécie para outro espaço geográfico ou condições ambientais, possibilitando, por exemplo, estudos do efeito de invasão de espécies em novos ambientes, estudo de paleo-biogeografia, ou mesmo avaliar o efeito de simulações de mudanças climáticas futuras na distribuição das espécies.

MDEs têm sido amplamente utilizados para estimar a distribuição potencial de espécies dando suporte a questões científicas e aplicadas tais como: para identificar áreas de *hotspot* de biodiversidade ou de ocorrência de espécies ameaçadas; indicar prioridade de áreas para conservação, avaliar impacto de mudanças climáticas, entre outras (Guisan e Thuiller, 2005). São especialmente úteis no monitoramento da vida silvestre e planejamento de estratégias de conservação ambiental, ao fornecer critérios objetivos produzidos a partir de dados secundários (Cassini, 2011).

Nas últimas décadas, a produção científica envolvendo MDE para aplicações em biogeografia e conservação cresceu acentuadamente. Isto se explica pela crescente disponibilidade de dados de distribuição das espécies, decorrente da organização dos herbários e coleções digitais online, e principalmente pela facilidade de aquisição de dados

ambientais, resultado da recente democratização de dados de sensoriamento remoto. Esta disponibilidade de dados de escalas locais a globais tem impulsionado tanto as publicações quanto os estudos relativos ao procedimento de modelagem (Kamino et al., 2011).

Várias são as técnicas de modelagem disponíveis para explorar as relações entre as variáveis respostas, neste caso a ocorrência da espécie, e as variáveis preditoras, ou seja, as variáveis de caracterização ambiental (Guisan e Zimmerman, 1999). Neste processo, há várias fontes que disponibilizam dados ambientais, em diferentes formatos (Kamino et al., 2012). Contudo, o uso destes dados para modelagem normalmente requer a integração geográfica de modo a compatibilizar escala, recorte e projeção cartográfica, o que pode não ser trivial para usuários pouco familiarizados com técnicas de geoprocessamento.

Neste contexto, este artigo tem por objetivo divulgar nossa iniciativa de organizar uma base de dados geográficos, usualmente utilizados para modelagem de distribuição de espécies em escala regional, e promover através de uma página na web, o acesso livre a esta base para usuários que queiram exercitar a modelagem tendo o Brasil e Amazônia Legal como recorte geográfico. A esta base, demos o nome de Ambdata – Variáveis Ambientais para Modelagem de Distribuição de Espécies (<http://www.dpi.inpe.br/Ambdata/>). A maioria dos dados foram obtidos a partir das fontes provedoras oficiais e apenas recortados/reamostrados para os limites geográficos definidos, com exceção das variáveis densidade de drenagem e distância vertical à drenagem mais próxima (HAND), que foram produzidas pelos autores. As variáveis disponíveis, fontes e procedimento de organização desta base de dados são descritos a seguir.

## 2. Metodologia

A partir de diferentes fontes de dados, as variáveis ambientais relacionadas a clima e caracterização topográfica, usualmente utilizadas para MDEs, foram adquiridas e organizadas em um sistema geográfico de informação. Cada variável foi então recortada para os limites geográficos definidos e reamostrada para uma mesma resolução espacial e projeção cartográfica. Uma vez que todas as variáveis estavam preparadas, organizou-se uma página na internet para apresentar as variáveis e proporcionar o livre acesso (download) a estes dados.

### 2.1 Extensão geográfica, escala e projeção cartográfica

Considerando potenciais usuários interessados na biodiversidade brasileira, duas extensões geográficas foram definidas para o recorte dos dados ambientais: Brasil e Amazônia Legal Brasileira (Tabela 1).

Tabela 1. Coordenadas geográficas para o recorte das variáveis ambientais para os limites do Brasil e da Amazônia Legal (UL - superior esquerda; LR - inferior direita).

	Latitude	Longitude
<b>Brasil</b>		
UL	05° 16' 16,5" N	73° 59' 16,5" W
LR	33° 45' 13,5" S	32° 23' 16,5" W
<b>Amazônia Legal</b>		
UL	05° 16' 11,65" N	74° 00' 14,98" W
LR	18° 02' 44,99" S	43° 59' 19,3" W

Os dados ambientais foram convertidos e são fornecidos como grades no formato ASCII-raster, projetados no sistema de coordenadas geodésicas de projeção “Lat-Long”, Datum WGS-84, com resolução espacial de 30 arcos de segundo (aproximadamente 1 km no Equador), interpolados pelo algoritmo Vizinho-mais-próximo.

## 2.2 Variáveis ambientais

Apresenta-se abaixo a descrição das variáveis ambientais sistematizadas no Ambdata.

As variáveis **climáticas** e bioclimáticas foram adquiridas a partir do WorldClim - Global Climate Data, (<http://www.worldclim.org/>) correspondem a dados de observação de 1950 a 2000 que foram interpolados para a resolução de 30 arc-segundos (~1km) (Hijmans et al. 2005). Os valores de temperatura são fornecidos em graus Celsius ( $^{\circ}\text{C}$ )\*10, e os valores de precipitação em mm. Para as variáveis precipitação total mensal (prec), temperatura máxima mensal (temp), temperatura mínima mensal (tmin), e temperatura média mensal (tmean), os dados estão disponíveis para os meses de janeiro a dezembro (mês 1 ao 12). As variáveis bioclimáticas, derivadas de temperatura e precipitação mensal representam tendências anuais, sazonalidade e fatores ambientais extremos ou limitantes, identificadas como bio1 a bio19: temperatura média anual (bio1); variação diurna média de temperatura média mensal (Tmax-Tmin) (bio2); isothermalidade ((bio2/bio7) (\* 100)) (bio3); sazonalidade da temperatura (desvio padrão \* 100) (bio4); temperatura máxima do mês mais quente (bio5); temperatura mínima do mês mais frio (bio6); amplitude térmica anual (bio5-bio6) (bio7); temperatura média do trimestre mais úmido (bio8); temperatura média do trimestre mais seco (bio9); temperatura média do trimestre mais quente (bio10); temperatura média do trimestre mais frio (bio11); precipitação Anual (bio12); precipitação do mês mais chuvoso (bio13); precipitação do mês mais seco (bio14); sazonalidade da precipitação (coeficiente de variação) (bio15); precipitação do trimestre mais chuvoso (bio16); precipitação do trimestre mais seco (bio17); precipitação do trimestre mais quente (bio18); precipitação do trimestre mais frio (bio19).

Ainda para a caracterização climática, o índice de Walsh (Walsh, 1996) traduz a intensidade e duração da estação seca que pode ser limitante para espécies Amazônicas. O início da estação seca ocorre quando pela primeira vez no ano chove menos de 100 mm no mês. A partir dos dados de pluviosidade acumulada mensal do WorldClim, Nelson (comunicação pessoal) gerou o índice de Walsh para Amazônia Brasileira.

Para caracterização de **relevo**, as variáveis altitude (em metros), declividade ou gradiente (em porcentagem), e exposição ou orientação de vertentes (em graus a partir do Norte) foram geradas a partir de dados do SRTM (*Shuttle Radar Topographic Mission*), de resolução horizontal (resolução espacial) de 3 arc-segundos (~90m) e resolução vertical (altura) de aproximadamente 1 m. O dado SRTM original foi reamostrado para a resolução espacial de ~1km, de modo a garantir a compatibilidade com os demais dados disponíveis no Ambdata.

A variável de distância vertical à **drenagem** mais próxima foi obtida com o uso do algoritmo HAND (*Height Above the Nearest Drainage*) a partir da topografia proveniente do Modelo Digital de Elevação (MDE) do SRTM (Rennó et al., 2008). Deve-se definir um limiar referente à área de contribuição mínima a partir do qual inicia-se a drenagem (nascentes) que será utilizada como referência. Quanto maior o limiar, menor a densidade de drenagem, ou seja, serão considerados apenas os cursos d'água principais (maior ordem). Limiares menores incluirão as drenagens menores, aumentando o detalhamento da rede de drenagem considerada. As grades de distância vertical geradas pelo HAND foram produzidas a partir dos dados SRTM, usando três limiares (50, 100 e 500 pixels), sendo os valores expressos em metros.

A variável densidade de drenagem foi produzida aplicando-se o algoritmo de interpolação espacial Kernel sobre os dados de rede de drenagem do projeto HydroSHEDS (Lehner et al., 2006), derivado das imagens SRTM. O dado de densidade de drenagem disponível foi gerado a partir do raio de influência de 0,27 graus na superfície e resolução espacial final de 1 km (Ximenes, 2008).

Para caracterização da **cobertura florestal**, duas variáveis foram incluídas: a cobertura florestal (*Percent Tree Cover*) proveniente do sensor MODIS (*MODerate-resolution Imaging Spectroradiometer*) e o mapa de vegetação do Brasil (IBGE, 1992).

O “Percent Tree Cover” representa a percentagem da cobertura do dossel florestal anual (Hansen et al., 2003), derivado de dados globais adquiridos entre 2000 e 2001 (<http://glcf.umiacs.umd.edu/data/vcf/>), gerado a partir da composição mensal das bandas do sensor MODIS, a bordo do satélite Terra da NASA, de 500 m de resolução.

O mapa de vegetação do Brasil reconstitui a situação da vegetação no território brasileiro com origem nas cartas-imagens de radar produzidas pelo Projeto RADAMBRASIL na escala 1:250.000 e foi obtido do IBGE (1992). O sistema de classificação do RADAMBRASIL reflete a chave aberta de Veloso et al. (1991). Para gerar a grade de vegetação, apenas o atributo LEG\_UVEG do mapa vetorial de vegetação foi selecionado. Para o recorte da Amazônia, apresenta-se o mapa a partir dos dados do RADAMBRASIL, com as atualizações feitas pelo Projeto SIVAM, publicado em formato digital, na escala original de 1:250.000 (PROJETO SIVAM, 2002).

O mapa de **solos**, adquirido junto a EMBRAPA, foi elaborado com base no novo sistema brasileiro de classificação de solos (EMBRAPA, 1999) e publicado pelo IBGE (2001). Para gerar a grade de solos, apenas o atributo CLASSE do mapa vetorial de solo foi selecionado. Da mesma forma que o mapa de vegetação, o mapa de tipos de solo é um arquivo ASCII-raster categórico, ou seja, os números da grade são apenas referências das classes temáticas associadas. Nem todos os algoritmos para modelagem de distribuição de espécies usam dados categóricos como variáveis ambientais, e portanto o uso das variáveis dos mapas de vegetação e solo limita-se a algoritmos que façam distinção de variáveis categóricas.

Por esta natureza, além das grades compatíveis com todos os outros dados disponíveis no Ambdata, os mapas de solo e vegetação podem também ser acessados no seu formato vetorial original (*shapefile*) para verificação das classes e atributos.

Para facilitar a visualização dos dados, os limites de território para o Brasil e a Amazônia Legal Brasileira foram obtidos a partir do IBGE (2010), e estão disponíveis em formato vetorial (*shapefile*), incluindo-se os limites das Unidades de Federação.

### **3. Resultados – dados ambientais disponíveis na internet**

Após a aquisição de todos os dados e tê-los recortado para os limites do Brasil e Amazônia Legal, organizou-se uma página na internet, hospedada no INPE/DPI para facilitar o livre acesso aos dados ambientais customizados para qualquer usuário interessado em MDEs. O Ambdata encontra-se hospedado nas versões em português e inglês no seguinte endereço eletrônico: <http://www.dpi.inpe.br/Ambdata/index.php> (Figura 1). A partir da página principal (índice), pode-se conhecer o conteúdo da base de dados, e selecionar variáveis de interesse para download.



Figura 1. Página na internet de apresentação do Ambdata (<http://www.dpi.inpe.br/Ambdata/index.php>).

O menu lateral apresenta e dá acesso às páginas que descrevem todas as variáveis disponíveis, como por exemplo os dados climáticos (Figura 2a). Cada página específica descreve o dado original, indicando a fonte e o procedimento adotado para geração do produto, se for o caso. Uma figura apresenta um exemplo da variável espacializada e um link direciona o usuário para a página de download. O item “Download” do menu principal direciona a qualquer momento o usuário para a página de download, onde uma tabela (Figura 2b) permite o download dos dados em um arquivo único, ou um arquivo para cada variável individual, compactados (formato zip).

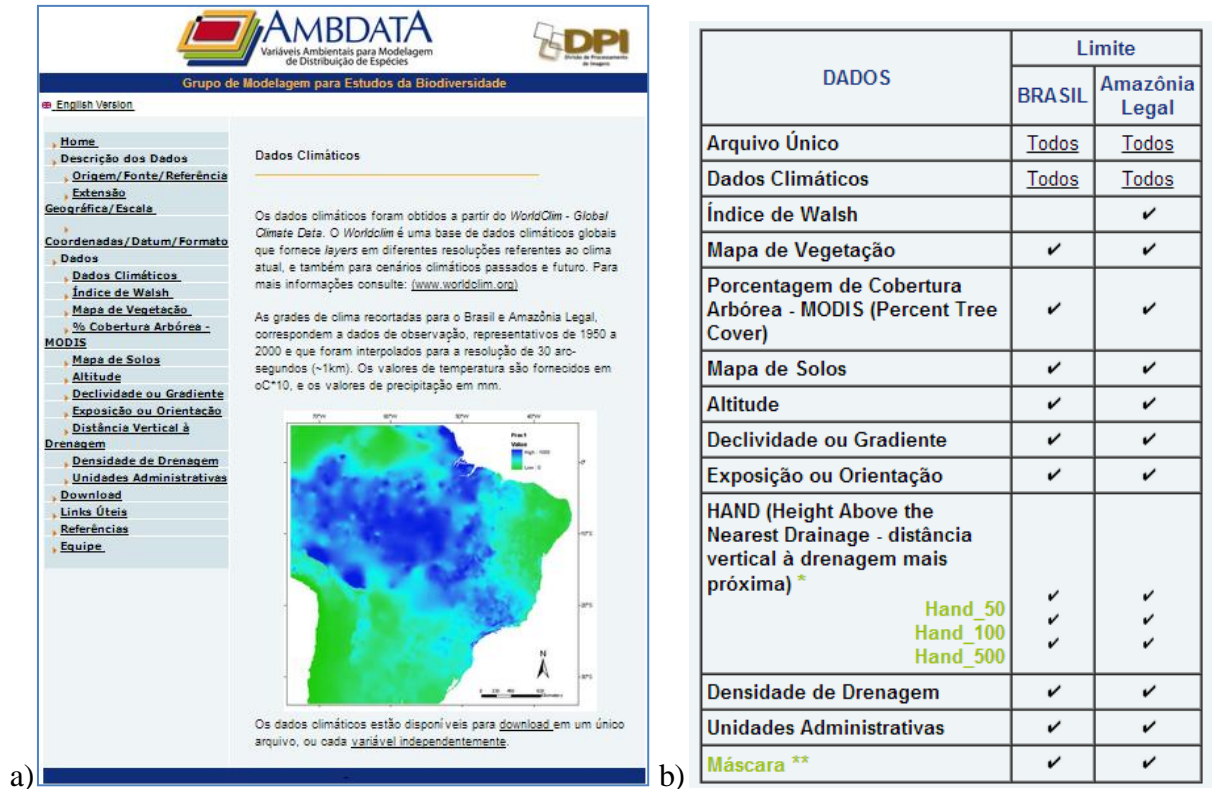


Figura 2. Exemplo da página que descreve as variáveis climáticas (a); e o detalhe da tabela de download para acessar arquivos compactados únicos ou conjunto de variáveis para o Brasil ou Amazônia Legal (b).

Adicionalmente, algumas indicações de dados ambientais e imagens disponíveis para GIS, software livre para MDE e SIG são apresentados no item de menu “Links Úteis”, e todas as referências estão listadas no item “Referências”.

#### 4. Conclusões

Este artigo apresenta a base de dados e a página da internet Ambdata, organizada para facilitar o acesso e uso das variáveis ambientais, provenientes de sensoriamento remoto e mapeamento cartográfico oficial para atividades de modelagem de distribuição de espécies.

Além de oferecer algumas variáveis como a distância vertical a drenagem mais próxima e a densidade de drenagem, produzidas a partir do SRTM, a partir do Ambdata, as variáveis ambientais encontram-se todas padronizadas quanto à resolução espacial e recortes geográficos para o Brasil e Amazônia Legal Brasileira. Estes simples procedimentos viabilizam, por exemplo, o uso direto dos dados no aplicativo Maxent (Phillips et al., 2011) que requer que todos os dados de entrada estejam no formato ASCII-raster, recortados para um único limite e resolução espacial. Desta forma, ecólogos e biogeógrafos podem se concentrar nos critérios de seleção de variáveis ambientais, o que é essencial no processo de modelagem, sem a necessidade de organizar os dados de entrada.

Este artigo contribui então para divulgar o Ambdata junto à comunidade científica: uma ferramenta que se propõe a auxiliar os estudos da distribuição de espécies Amazônicas e Brasileiras, facilitando a ampliação do conhecimento e a discussão a respeito de nossa biodiversidade.

#### Agradecimentos

Este trabalho foi motivado pela interação com os pesquisadores do Projeto openModeller (Fapesp process: 04/11012-0) e teve suporte do Projeto Cenários para Amazônica - Clima, Biodiversidade e Uso da Terra. Os autores agradecem à colega Hilcéa Ferreira pelo design das

páginas web, e à Divisão de Processamento de Imagem (DPI) do INPE por hospedar o Ambdata.

## Referências Bibliográficas

- Cassini, M.H. Ecological principles of species distribution models: the habitat matching rule. **Journal of Biogeography**, v. 38, p. 2057–2065, 2011.
- Elith, J.; Leathwick, J.R. Species distribution models: ecological explanation and prediction across space and time. **Annual Review of Ecology, Evolution and Systematics**, v.40, p. 677–697, 2009.
- EMBRAPA. Centro Nacional de Pesquisa de Solos. **Sistema Brasileiro de Classificação de Solos**. Rio de Janeiro, 412 p., 1999
- Guisan, A.; Thuiller, W. Predicting species distribution: offering more than simple habitat models. **Ecology Letters**, v. 8, p. 993–1009. , 2005.
- Guisan, A.; Zimmermann, N. E. Predictive habitat distribution models in ecology. **Ecological Modelling**, v.135, p. 147-186, 2000.
- Hansen, M.; DeFries, R.S.; Townshend, J.R.G.; Carroll, M.; Dimiceli, C.; Sohlberg, R.A. Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm. **Earth Interactions**, v. 7, n.10, p.1-15, 2003.
- Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G.; Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. **International Journal of Climatology**, v.25, p.1965-78, 2005.
- Instituto Brasileiro de Geografia e Estatística – IBGE. Mapa de vegetação do Brasil. Rio de Janeiro. 1992. Escala 1:5.000.000.
- Instituto Brasileiro de Geografia e Estatística – IBGE; EMBRAPA - Mapa de Solos do Brasil. Rio de Janeiro. 2001 - Escala 1:5.000.000.
- Instituto Brasileiro de Geografia e Estatística – IBGE. Malha municipal digital - limites das Unidades de Federação, Censo 2010. 2010 - Escala 1:250.000.
- Kamino, L.H.Y.; Stehmann, J.R.; Amaral, S.; Marco Jr, P.; Rangel, T.F.; Siqueira, M.F.; Giovanni, R.; Hortal, J. Challenges and perspectives for species distribution modelling in the neotropics. **Biology Letters** (Print), v. 1, p., 2011.
- Kamino, L.H.Y.; Marco Jr, P.; Rangel, T.F., Amaral, S.; Siqueira, M.F.; Giovanni, R.; Hortal, J. Workshop summary: The application of species distribution models in the megadiverse Neotropics poses a renewed set of research questions. **Frontiers of Biogeography**, v. 4, p. 7-10, 2012.
- Lehner, B.; Verdin K.; Jarvis, A.. 2006. HydroSHEDS – Technical Documentation. Versão 1. p. 27. Disponível em: [http://gisdata.usgs.net/HydroSHEDS/downloads/HydroSHEDS\\_TechDoc\\_v10.pdf](http://gisdata.usgs.net/HydroSHEDS/downloads/HydroSHEDS_TechDoc_v10.pdf). Acesso em: 20 de out de 2008.
- Phillips, S. J.; Anderson, R. P.; et al. Maximum entropy modeling of species geographic distributions. **Ecological Modelling**, v.190, n.3-4, p.231-259, 2006.
- PROJETO SIVAM. 2002. **Relatório metodológico de trabalho. Revisão 3**. Diretoria de Geociência. Rio de Janeiro, 330 pp.
- 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 Amazonia. **Remote Sensing of Environment**, v. 112, p.3469-81, 2008.

Veloso, H. P.; Rangel Filho, A. L.; Lima, J. C. A. **Classificação da Vegetação Brasileira, adaptada a um sistema universal**. Rio de Janeiro, IBGE, Departamento de Recursos Naturais e Estudos Ambientais, 1991. 124p.

Walsh, R.P.D. The climate. In: P.W. Richards (ed). **The Tropical Rain Forest: an ecological study**. Cambridge University Press. , 1996, p.159-255.

Ximenes, A.C. Mapas auto-organizáveis para a identificação de ecorregiões do interflúvio Madeira-Purus: uma abordagem da biogeografia ecológica. 2008. 155 pp. (INPE-15332-TDI/1372). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2008. Disponível em: <http://urlib.net/sid.inpe.br/mtc-m18@80/2008/08.18.14.02>. Acesso em: 27 set. 2010.