

## Uso de análise de imagens baseada em objetos (OBIA) e classificação não-supervisionada para identificação de envelopes bioclimáticos no bioma amazônico.

Thiago Sanna Freire Silva<sup>1</sup>  
Carlos Leandro de Oliveira Cordeiro<sup>1</sup>

<sup>1</sup> Instituto de Geociências e Ciências Exatas, UNESP - Univ Estadual Paulista, Departamento de Geografia, Ecosystem Dynamics Observatory.  
Caixa postal 178, CEP13506-900. Rio Claro, SP, Brazil.  
tsfsilva@rc.unesp.br, carlosleandroc@gmail.com

**Abstract.** Mapping bioclimatic envelopes onto geographical space is an important process for predicting and understanding species range sizes and distribution limits. This mapping depends on the availability of relevant environmental datasets at the appropriate scale, several which are commonly derived from remote sensing sources. Object-based image analysis (OBIA) is a relatively new approach in remote sensing, and together with machine learning algorithms, has been used successfully to map land cover classes in highly heterogeneous satellite and aerial images. Its framework, however, is equally applicable to other problems involving the delineation and classification of spatially homogeneous regions. Therefore, we investigated the applicability of OBIA and unsupervised clustering to detect possible bioclimatic envelopes and support biogeographical studies. Our selected study area was a portion of the Amazon basin, corresponding to interfluvium between the Negro and Branco rivers. We combined topographic data from the SRTM mission, precipitation data from the CHIRPS dataset, radar backscattering from the ALOS/PALSAR sensor and spatially explicit estimates of canopy height and vegetation biomass, which was supplied as input data for image segmentation and posterior clustering, based on mean and standard deviation attributes. Overall, several bioclimatic envelopes could be mapped, some of them coinciding to land features usually associated with vicariance events, such as the Branco River channel. Our results emphasize the applicability of modern image analysis methods for identifying bioclimatic envelopes using spatially-explicit environmental data, and also the potential role of bioclimatic discontinuities, as well as vicariance, in explaining current distribution patterns for amazonian species.

**Palavras-chave:** remote sensing, biogeography, species distribution, sensoriamento remoto, biogeografia, distribuição de espécies.

### 1. Introdução

A modelagem de envelopes bioclimáticos é um processo que utiliza algoritmos para localizar espaços geográficos com características ambientais similares, usualmente do ponto de vista de uma ou mais espécies de organismos (Hirzel e Lay, 2008). O mapeamento da projeção destes envelopes bioclimáticos sobre o espaço geográfico (*E space* e *G space*, *sensu* Peterson et. al., 2011) se baseia principalmente em dados relativos a processos abióticos, observados em escalas mais amplas, como temperatura, precipitação, pedologia, hidrologia, e características geológicas e geomorfológicas; além de dados bióticos relacionados à estrutura de habitats, como altura ou estrutura do dossel, biomassa e fenologia da vegetação (He et al. 2015). Essa abordagem tem sido utilizada para a modelagem de distribuição de espécies (*species distribution modeling*, SDM) ou modelagem de nichos ambientais (*environmental niche modeling*, ENM), principalmente utilizando dados derivados de observações por

sensoriamento remoto orbital (Deblauwe et al. 2016; He et al. 2015), para representar as condições ambientais que refletem as dimensões do nicho ecológico grinnelliano (Peterson et al., 2011), representando limites e requisitos ambientais necessários para a sobrevivência das espécies (Mod et al. 2016).

O mapeamento da distribuição e extensão dos envelopes bioclimáticos sobre o espaço geográfico não é trivial, e depende de uma série de fatores como a disponibilidade de dados representativos dos processos ambientais, em uma escala adequada de representação; a complexidade dos modelos; e a disponibilidade de dados confiáveis de distribuição das espécies (Peterson et al., 2011). Outras dificuldades são as influências de vários fatores históricos, como vicariância e barreiras de dispersão, história geológica, e/ou interações bióticas, que aumentam a diferença entre o nicho realizado e o nicho fundamental (Mod et al., 2016).

Um método recentemente empregado na distinção de elementos de cobertura da terra com baixa separabilidade em imagens espectrais é a análise de imagens baseada em objetos (*object-based image analysis*, OBIA) (Hay e Castilla 2008; Moffett e Gorelick 2013), a qual busca delimitar áreas homogêneas em um conjunto de imagens (chamadas de segmentos ou objetos), e que se tornam a unidade de análise para classificação e mapeamento de ambientes, substituindo a análise baseada em *pixels*. Objetos similares podem ser agrupados através de métodos de classificação, entre eles a análise de agrupamento (*clustering*). Um desses métodos é conhecido como UPGMA (*unweighted-pair group method using arithmetic mean*), popularmente utilizado em análises de agrupamento hierárquico e geralmente capaz de capturar bem a dissimilaridade estrutural dos dados (Legendre e Legendre 1998), especialmente dados de composição florestal (Tuomisto et al. 2003; Figueiredo et al. 2014).

Os métodos de OBIA e classificação não-supervisionada por aprendizado de máquina, contudo, tem sido pouco utilizado em áreas de aplicação fora do domínio de análise de imagens ou dados não-espaciais, e pode representar uma abordagem eficiente e robusta para o delineamento de envelopes ambientais a partir de conjuntos de camadas espaciais. Assim, o objetivo deste trabalho foi avaliar como o método de segmentação e análise de agrupamento pelo algoritmo UPGMA podem auxiliar na identificação de envelopes bioclimáticos em uma região no norte da Amazônia.

## 2. Metodologia de Trabalho

O estudo foi realizado na região do interflúvio dos rios Branco e Negro (Figura 1). Esse interflúvio está localizado nas partes baixas do escudo das Guianas, com algumas áreas mais elevadas e *inselbergs* com elevações acima de 2000 m. O clima da região é fortemente afetado pelos fenômenos climáticos El Niño e La Niña (Latrubesse e Ramonel 1994), com duas estações bem definidas, uma estação seca mais pronunciada entre junho a setembro e uma estação chuvosa mais pronunciada entre dezembro a março. Um gradiente de precipitação pode ser observado na região, com aproximadamente 1500 mm por ano ao norte e 2500 mm ao sul e uma variação média anual de temperatura fica entre 26°C e 33°C (Radambrasil 1976).

A maioria da região é coberta por floresta densa, montana e submontana, e áreas com florestas aluviais como igapós e várzeas, com árvores altas e alta biomassa (IBGE 2012). Em meio a essas áreas florestadas são encontradas porções de vegetação de areias brancas como campinas e campinaranas, encontradas principalmente nas áreas mais ao sul, próxima ao Rio

Negro (IBGE 2012) e uma grande porção de savanas no alto Rio Branco (IBGE 2012).

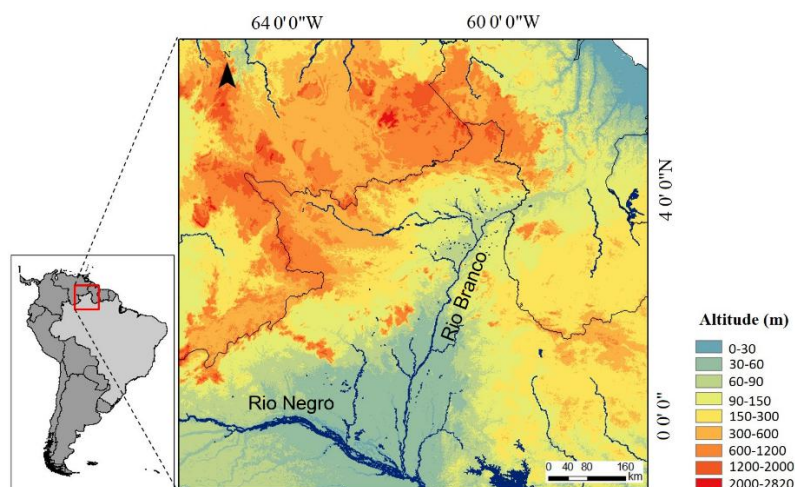


Figura 1. Localização da área de estudo no interflúvio dos rios Negro e Branco.

A seleção dos atributos foi baseada em imagens de satélites e seus produtos, relacionados a geologia, topografia, clima e biota. A altitude, declividade e orientação da vertente (aspecto) foram derivados do modelo digital de elevação (MDE) da SRTM (*Shuttle Radar Topography Mission*), com resolução espacial de 90 m. Os dados de precipitação foram obtidos na forma do produto CHIRPS (*Climate Hazards group InfraRed Precipitation with Station data*), a partir de uma série de dados diários de precipitação, com resolução espacial de 5,5 km, disponível em <http://chg.geog.ucsb.edu/data/chirps/>. A partir dos dados de precipitação diários foram extraídos os extremos de precipitação, representados pelos percentis de 90% e 10% da série temporal correspondente ao intervalo de tempo de 1981 a 2016. As imagens do sensor orbital PALSAR (*Phased Array type L band Synthetic Aperture Radar*), à bordo do satélite *Advanced Land Observing Satellite* (ALOS) foram utilizadas diretamente, com resolução espacial de 100 m e polarização HH. Os dados da biota consistiram em informações de biomassa e altura do dossel extraídas de dados modelados e interpolados por Saatchi et al., (2007) e Simard et al., (2011), respectivamente, ambas com resolução espacial de 1km.

Todos os processamentos foram realizados nos softwares Python, QGIS e R. A seleção dos dados de entrada buscou representar fatores topográficos (DEM-SRTM), características bióticas (biomassa e altura do dossel), sazonalidade climática (precipitação) e umidade do solo e estrutura da vegetação (retroespalhamento PALSAR). A integração dos dados foi realizada através de reamostragem de todas as camadas para 90 m, a menor escala espacial utilizada. Todos os datasets foram então reprojatados e empilhados (*stacking*) e posteriormente foi realizada a segmentação, utilizando o software *The Remote Sensing and GIS Software Library* (RSGISLib, Bunting & Clewley, (2013) e o algoritmo de agrupamento K-Means, com número de classes = 10 e o número mínimo de pixels por objeto de 5000. Os objetos gerados foram então transformados em vetores.

Após a segmentação, os valores da média e desvio padrão dos pixels de cada camada de entrada foram calculados e extraídos para cada objeto. Estes atributos foram utilizados como dados de entrada para a análise de *agrupamentos* pelo método UPGMA, utilizando os pacotes *vegan*, *MASS* e *cluster* do software R. A seleção do número de classes a serem agrupadas foi

determinada através da análise visual do dendrograma de agrupamento. O procedimento de classificação foi hierárquico, identificando sucessivos grupos, e reaplicando o algoritmo UPGMA somente para os objetos inicialmente classificados como pertencendo a diferentes grupos (Figura 2).

Para a análise e avaliação dos agrupamentos foi realizada uma análise gráfica através de *boxplots*, observando-se a variação dos atributos bioclimáticos em cada classe e nível hierárquico. Um mapa final foi produzido distinguindo as diferentes classes.

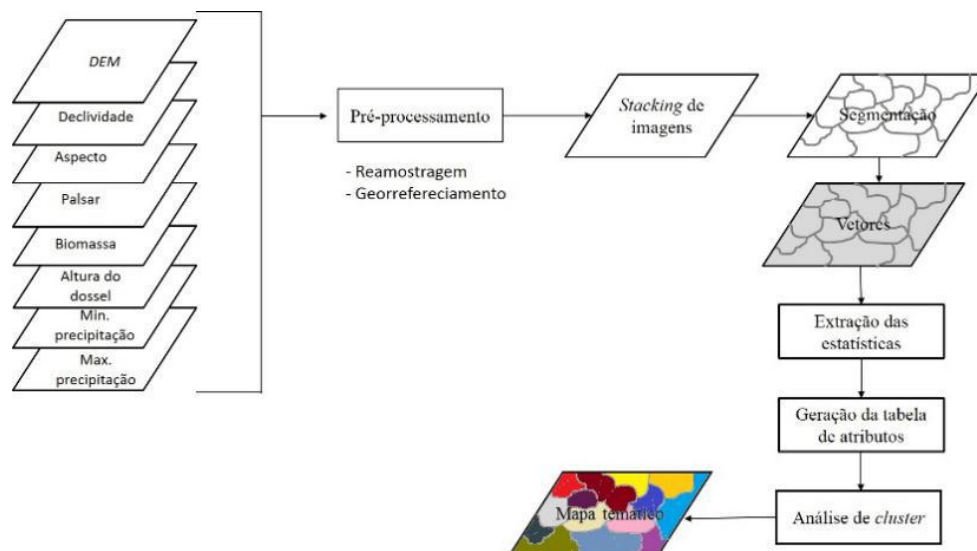


Figura 2. Fluxograma com as etapas das análises dos dados para a classificação de espaços geográficos no interflúvio dos rios Negro e Branco, utilizando o método de segmentação e de agrupamento.

### 3. Resultados e discussão

De acordo com dendrograma inicial produzido pela análise UPGMA, foram realizados quatro processos hierárquicos de classificação. Na primeira classificação foram separados quatro grupos (Fig. 3A), revelando um padrão com duas classes maiores dominando a área de estudo e duas classes menores ocupando pequenas áreas, uma destas restrita à parte nordeste e outra distribuída em diferentes partes da área de estudo (Fig. 3B-a). Uma segunda classificação foi realizada no agrupamento 2, o qual foi separado dois grupos (Fig. 3B-b); a classe 6 foi distribuída em diferentes partes, enquanto a classe 5 ocupou uma maior porção da área de estudo. Essa classe foi dividida em mais três classes (Fig. 3B-c), e nessa separação a grande savana de Roraima (classe 7), ao norte da bacia do rio Branco, foi distinguida dos demais ambientes. No último passo de classificação, o agrupamento 3 foi separado em cinco classes, sendo duas principais que ocuparam diferentes porções da área de estudo. A classe 12 dominou a porção oeste do rio Branco e a classe 10 a porção à leste deste rio (Fig. 3B-d).

A dominância de 3 classes principais no mapeamento (classes 9, 10 e 12) caracterizou a existência de envelopes bioclimáticos distintos entre as duas margens do rio Branco. Na margem direita o ambiente pode ser caracterizado por áreas de altitude moderada, com uma vegetação predominantemente arbórea de florestas de planície. A margem esquerda é



caracterizada por ambientes mais altos e íngremes, dominados por vegetação de floresta montana e submontana. Por fim, a classe 9, recorrente na margem esquerda, pode ser caracterizada por terras baixas, dominadas por florestas ombrófilas (Rede Rio Negro 2015) (Fig. 3A-d). Nossos resultados mostram similaridade com o mapeamento ambiental realizado pela Rede Rio Negro (2015) para essa região, porém avança em considerar quais informações ambientais foram mais importantes para a caracterização de cada um dos ambientes encontrados neste interflúvio.

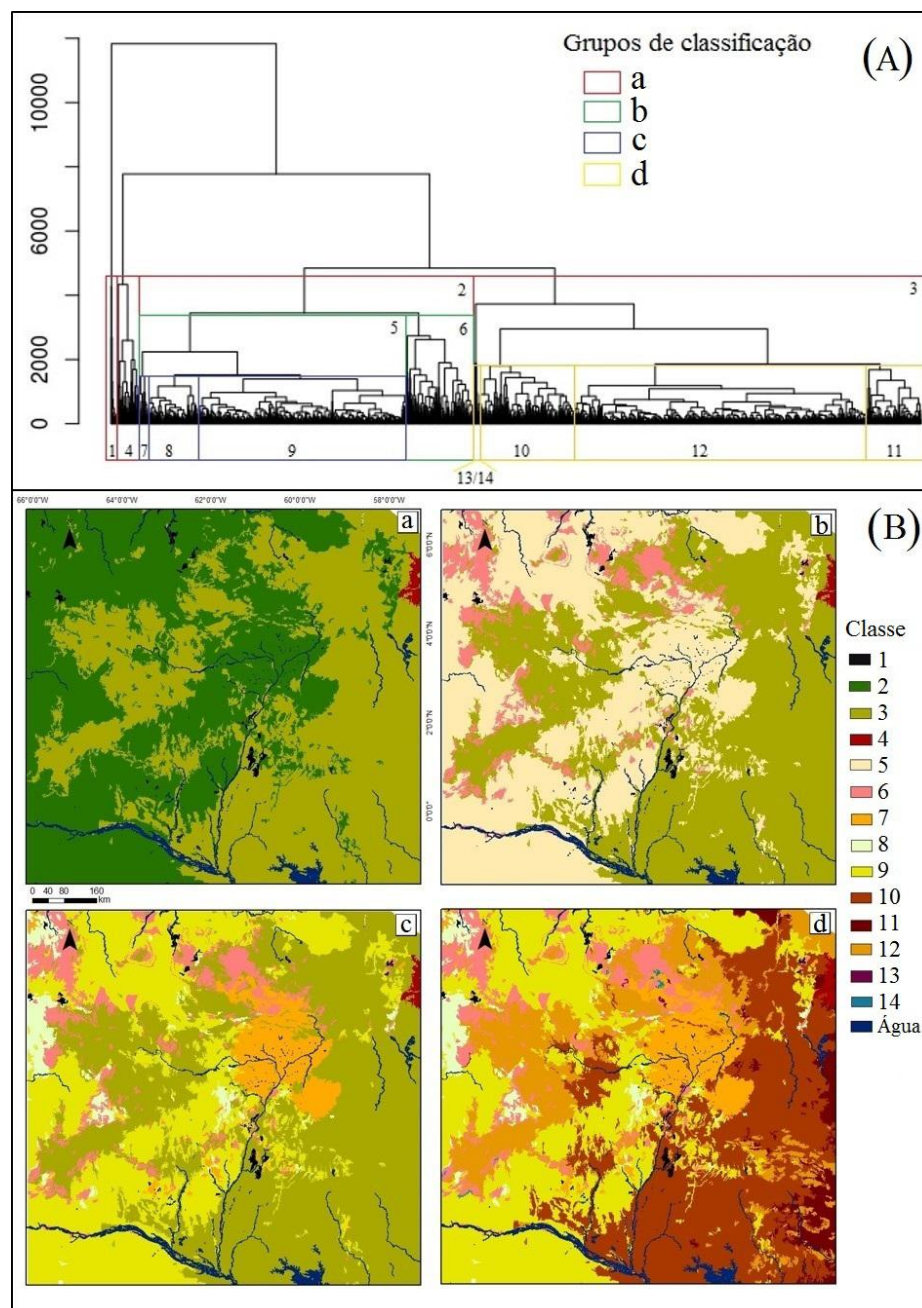


Figura 3. Dendrograma da classificação UPGMA, considerando 14 classes para o mapeamento (A). Foram realizadas 4 grupos de classificação (a, b, c e d), onde as cores representam cada um dos agrupamentos classificados. (a) primeira classificação; (b) segunda classificação; (c) terceira classificação; e (d) quarta classificação, mostrando o mapa final da distribuição dos envelopes bioclimáticos no interflúvio dos rios Negro e Branco.

As informações de biomassa, altura do dossel e retroespalhamento (PALSAR), e derivadas do DEM (declividade e altitude) foram os dados mais importantes na separação e definição dos espaços geográficos (Fig. 4). Os dados de precipitação foram os que menos influenciaram na separação dos grupos, apresentando baixa diferença de valores entre as classes ao longo do processo de classificação. Contudo, espera-se que padrões de precipitação variem em escalas regionais e continentais, maiores que a extensão da área de estudo selecionada.

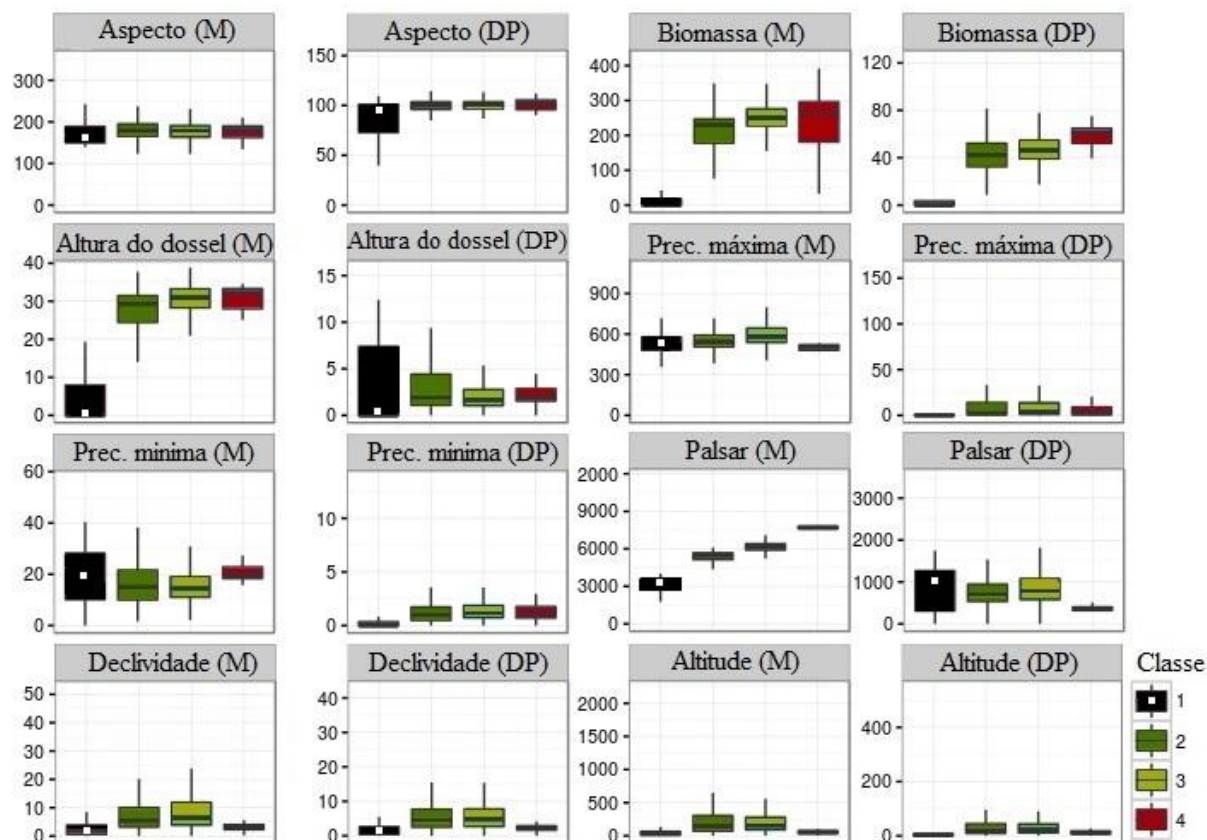


Figura 4. *Boxplots* mostrando a distribuição de valores observados para cada classe de envelope bioclimático, na primeira etapa de classificação, ilustrando a variação dos dados para as quatro classes geradas.

#### 4. Conclusões

O presente trabalho demonstra o potencial dos métodos de análise de imagens baseada em objetos e de classificação não-supervisionada a partir de aprendizado de máquina, tradicionalmente utilizados para a classificação de cobertura da terra em imagens espectrais, para a identificação e mapeamento de diferentes envelopes bioclimáticos em escala regional.

A região do interflúvio Rio Branco - Rio negro apresentou-se como um grande mosaico ambiental, representado por um gradiente topográfico e de vegetação bastante pronunciado e os *layers* apontados como os mais importantes na separação dos espaços geográficos refletiram esses dois gradientes. Por outro lado, a precipitação não se mostrou importante nas análises, provavelmente por ser um processo que varia em escalas maiores que as analisadas neste trabalho.

Os resultados obtidos enfatizam a heterogeneidade ambiental observada mesmo dentro de um bioma tipicamente tropical, comumente considerado como “homogêneo” em termos de sua composição ambiental e pressões seletivas associadas, relegando a explicação de limites de distribuição das espécies à eventos pretéritos de vicariância (e.g. migração e formação dos grandes rios e soerguimento do terreno através de processos tectônicos). Os métodos utilizados indicam a potencial existência de ecoclinas dentro do bioma amazônico, muitas vezes coincidindo com outros elementos da paisagem, as quais podem também explicar as áreas e limites de distribuição atuais de diversas espécies amazônicas.

## Agradecimentos

Este projeto recebeu apoio financeiro do projeto# 2012/50260-6, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). CLO Cordeiro é bolsista de pós-doutorado FAPESP, projeto #2016/08685-0, e TSF Silva é bolsista de produtividade do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq 310144/2015-9).

## Referências bibliográficas

Bunting, P.J.; Clewley, D.; Lucas, R.M.; Gillingham, S. The Remote Sensing and GIS Software Library (RSGISLib). **Computers & Geosciences**, v. 60, p. 216-226, 2014.

Deblauwe, V.; Droissart, V.; Bose, R.; Sonké, B.; Blach-Overgaard, A.; Svenning, J. C.; Wieringa, J.J.; Ramesh, B.R.; Stévant, T.; Couvreur, T.L.P. Remotely sensed temperature and precipitation data improve species distribution modelling in the tropics. **Global Ecology and Biogeography**, v. 25, n. 4, p. 443–454, 2016.

Figueiredo, F.O.G.; Costa, F.R.C.; Nelson, B.W.; pimentel, T.P. Validating forest types based on geological and land-form features in central Amazonia. **Journal of Vegetation Science**, v. 25, n. 1, p. 198–212, 2014.

Hay, G. J.; Castilla, G. Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline. In: Blaschke, T.; Lang, S.; Hay G. (Org.). **Object Based Image Analysis**, New York: Springer, 2008. cap. 1.4, p. 93–112.

He, K.S.; Bradley, B.A.; Cord, A.F; Rocchini, D.; Tuanmu, M; Schmidtlein, S.; Turner, W; Wegmann, M; Pettorelli, N. Will remote sensing shape the next generation of species distribution models? **Remote Sensing in Ecology and Conservation**, v. 1, n.1, p. 4-18, 2015.

Hirzel, A. H.; Lay, G.L. Habitat suitability modelling and niche theory. **Journal of Applied Ecology**, v. 45, p. 1372–1381, 2008.

Instituto Brasileiro de Geografia e Estatística – IBGE. **Manual Técnico da Vegetação Brasileira. Sistema fitogeográfico Inventário das formações florestais e campestres Técnicas e manejo de coleções botânicas Procedimentos para mapeamentos**. Rio de Janeiro: Diretoria de Geociências Coordenação de Recursos Naturais e Estudos Ambientais. 2012. 271 p. (ISBN 978-85-240-4272-0)

Latrubesse, E.M.; Ramonell, C. A climatic model for Southwestern Amazonia at Last Glacial times. **Quaternary International**, v. 21, p. 163-169, 1994.

Legendre, P.; Legendre, L. **Numerical ecology**. Amsterdam: Elsevier, 1998. 852 p.

Mod, H.K.; Scherrer, D.; Luoto, M.; Guisan, A. What we use is not what we know: environmental predictors in plant distribution models. **Journal of Vegetation Science**. doi:10.1111/jvs.12444, 2016.

Moffett, K.B.; Gorelick, S.M. Distinguishing wetland vegetation and channel features with object-based image segmentation. **International Journal of Remote Sensing**, v. 34, n. 4, p. 1332-1354, 2013.

RADAMBRASIL. Folha SC. **19-Rio Branco**. , Rio Janeiro: Ministério das Minas e Energia, Departamento Nacional da Produção Mineral, Projeto RADAM BRASIL, 1976.

Rede Rio Negro, 2015. Bacia do Rio Negro uma visão socioambiental. Disponível em:  
[http://www.rionegro.org.br/sites/rionegro.org.br/files/mapa\\_port\\_fim.pdf](http://www.rionegro.org.br/sites/rionegro.org.br/files/mapa_port_fim.pdf)

Saatchi, S.S.; Houghton, R.A.; Dos Santos A.R.C.; Soares, J.V.; Yu, Y. Distribution of aboveground live biomass in the Amazon basin. **Global Change Biology**, v. 13, n.4, p. 816–837, 2007.

Simard, M.; Pinto, N.; Fisher, J.B.; Baccini, A. Mapping forest canopy height globally with spaceborne lidar. **Journal of Geophysical Research**, v. 116, n. G04, p. 1-12. 2011.