

## **EFEITOS DA ANISOTROPIA NA IDENTIFICAÇÃO DE ESPÉCIES EM FLORESTA TROPICAL COM IMAGENS HIPERESPECTRAIS**

### **EFFECTS OF ANISOTROPY ON SPECIES DISCRIMINATION IN TROPICAL FOREST USING IMAGING SPECTROSCOPY**

Matheus Pinheiro Ferreira<sup>1</sup>, Yosio Edemir Shimabukuro<sup>1</sup>, Carlos Roberto de Souza Filho<sup>2</sup>

<sup>1</sup> Instituto Nacional de Pesquisas Espaciais - INPE, Av. dos Astronautas, 1758, 12.227-010 - São José dos Campos, SP, Brasil, mpf;yosio@dsr.inpe.br

<sup>2</sup> Universidade Estadual de Campinas - UNICAMP, R. João Pandiá Calógeras, 51, 13.083-870- Campinas, SP, Brasil, beto@ige.unicamp.br

#### **RESUMO**

Dados hiperespectrais obtidos por sensores a bordo de aeronaves e satélites tem demonstrado grande potencial para o estudo de florestas tropicais. Em particular, a identificação de espécies arbóreas com imagens de alta resolução espacial e espectral mostrou-se viável e pode contribuir para inventários fitossociológicos a nível local e regional. Neste trabalho, buscou-se avaliar o efeito da anisotropia das copas na identificação de nove espécies em uma floresta neotropical utilizando imagens hiperespectrais aerotransportadas. A distribuição de brilho dos pixels das espécies foi dividida nas classes iluminada e sombreada, por meio de um limiar calculado de forma automática. A performance de métodos paramétricos (Análise Discriminante Linear e Quadrática) e não-paramétricos (*Support Vector Machine* e *k*-vizinhos mais próximos) foi avaliada em termos de exatidão de classificação, tendo como dados de entrada pixels iluminados e sombreados simultânea e separadamente. A distribuição de brilho dos pixels de algumas espécies, principalmente as com copas de grandes dimensões, mostrou claro comportamento bimodal, o que afetou a performance de classificadores paramétricos de forma mais severa. De modo geral, a exatidão de classificação aumentou para todos os métodos testados ao se utilizar somente pixels iluminados. Os resultados apresentados neste trabalho demonstram a importância de se considerar os efeitos da anisotropia das copas na identificação de espécies arbóreas em ambientes tropicais com dados de sensoriamento remoto hiperespectral de alta resolução espacial.

*Palavras-chave:* Espectroscopia de imageamento, diversidade florística, efeitos de iluminação, mapeamento de espécies.

#### **ABSTRACT**

Airborne and spaceborne hyperspectral data provide important information for management and conservation of tropical forests. In particular, species identification using high-fidelity imaging spectroscopy have proven to be feasible and could contribute to floristic assessments at local to regional scales. This study aims to evaluate the effects of crown anisotropy for discrimination of nine species from a Neotropical forest using airborne hyperspectral data. Pixel brightness distribution was divided into sunlit and shaded, by means of an automatic thresholding method. The performance of parametric (linear and quadratic discriminant analysis) and non-parametric (*Support Vector Machine* and *k*-Nearest Neighbor) was evaluated in terms of classification accuracy, with sunlit and shaded pixels classified simultaneous and separately. Pixel brightness distribution of some species, particularly those with large crowns, showed a clear bimodal behavior, which affected the performance of parametric classifiers more severely. In general, classification accuracy increased for all tested methods when using only sunlit pixels. Results presented in this study demonstrate the importance of considering the effects of crown anisotropy on species discrimination in tropical environments using high spatial resolution hyperspectral data.

*Keywords:* Imaging spectroscopy, floristic diversity, illumination effects, tree species mapping.

#### **INTRODUÇÃO**

A potencialidade do sensoriamento remoto para identificação de espécies de plantas foi verificada na década de 1980, quando as primeiras imagens de alta resolução espectral foram adquiridas (GOETZ e VANE,

1985). Desde então, estudos têm sido realizados visando obter informações sobre a possível utilidade de dados de reflectância para discriminação de espécies e fitofisionomias. Em ambientes tropicais, com alta diversidade florística, estudos desta natureza estão restritos à condição de que a variabilidade espectral intraespecífica seja significativamente menor do que a interespecífica. Esta condição foi verificada ao nível foliar (COCHRANE *et al.*, 2000; CASTRO-ESAU *et al.*, 2006; FERREIRA *et al.*, 2013) e, de modo mais restrito, ao nível de dossel (ZHANG *et al.*, 2006). Ao nível foliar, segundo Gates *et al.* (1965), a interação da radiação eletromagnética com o tecido vegetal é controlada basicamente pela concentração de pigmentos, estrutura e quantidade de água. Já ao nível de dossel, um número maior de variáveis se inter-relaciona. Dentre elas, além dos próprios elementos da vegetação (folhas, galhos, cascas, frutos, sementes, etc.), podemos citar as propriedades espectrais do substrato, os parâmetros biofísicos (Índice de Área Foliar (IAF); Distribuição Angular das Folhas (DAF)), fatores de natureza geométrica (ângulos de iluminação e visada) e anisotrópicos (PONZONI *et al.*, 2012).

Particularmente, em imagens de alta resolução espacial, o efeito da anisotropia das copas, afeta seu comportamento espectral, produzindo proporções distintas de pixels sombreados e iluminados. Os pixels iluminados geralmente possuem uma relação sinal/ruído mais alta em imagens hiperespectrais, o que melhora algumas aplicações como, por exemplo, a extração de atributos bioquímicos do dossel (MALENOVSKÝ *et al.*, 2013). No entanto, à identificação de espécies, esta questão foi pouco estudada, principalmente em ambientes tropicais. Nesse contexto, o objetivo deste trabalho foi avaliar o efeito da anisotropia das copas na discriminação de nove espécies ao nível de dossel, utilizando imagens hiperespectrais de alta resolução espacial. Para tanto, foi avaliada a performance de métodos paramétricos e não-paramétricos na classificação de espécies, tendo como dados de entrada pixels iluminados e sombreados simultânea e separadamente.

## MATERIAL E MÉTODOS

### Área de estudo

A área de estudo deste trabalho compreende a Reserva de Santa Genebra (22°48' - 22°50' S, 47°06' - 47°07' W), localizada no município de Campinas/SP (Figura 1). Esta área possui 251,8 hectares (GUARATINI *et al.*, 2008), onde a formação florestal predominante é a floresta estacional semi-decidual (VELOSO *et al.*, 1991). A Reserva de Santa Genebra é um remanescente bem preservado de Floresta Atlântica e abriga alta diversidade de espécies arbóreas.

### Imagens hiperespectrais

Foram utilizadas imagens hiperespectrais obtidas pelo sensor ProSpec-TIR-VS de um levantamento aéreo realizado no dia 07/06/2010. Estes dados possuem resolução espacial de 1 metro e 357 bandas espectrais no intervalo de 400 a 2.500 nm. Dez faixas foram necessárias para o imageamento completo da área de estudo. A alta resolução espacial e espectral das imagens permite visualizar com clareza a copa de árvores emergentes, bem como sua resposta espectral (Figura 1).

Primeiramente, foi realizada a correção atmosférica das imagens para converter os dados de radiância aparente em reflectância de superfície por meio do algoritmo *Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes* (FLAASH) (ITT Visual Information Solutions, 2009). Em seguida, as imagens foram georreferenciadas utilizando-se o arquivo *Internal Geometry Map* (IGM), gerado durante o levantamento aéreo de cada faixa e, então, mosaicadas, apresentando pouca variação de brilho entre as faixas (Figura 1). Após a eliminação dos canais com baixa relação sinal/ruído, localizados principalmente ao redor das feições da absorção de água atmosférica (1.400 nm e 1.900 nm), a imagem foi composta de 260 bandas espectrais.

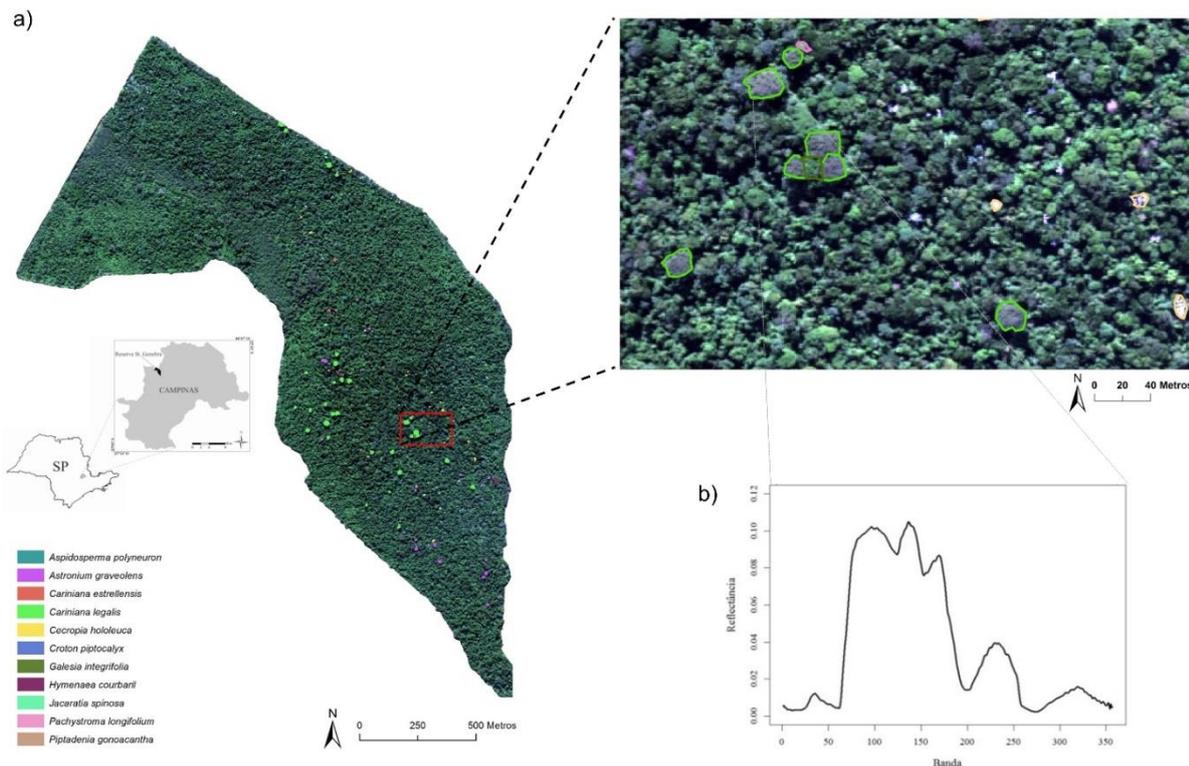


Figura 1. (a) Localização da área de estudo e distribuição espacial das espécies arbóreas identificadas em campo; (b) Resposta espectral média de uma copa de Cariniana legalis, localizada na porção central.

Figure 1. (a) Geographic location of the study area and spatial distribution of tree species identified at the field; (b) Mean spectral response a Cariniana legalis crown, located in the central portion.

Tabela 1. Número de pixels total, iluminados e sombreados por espécie, com destaque para as espécies que possuíram menos de 300 pixels e foram excluídas da análise

Table 1. Total, sunlit and shaded number of pixels per species, boldfaced lines indicate species with less than 300 pixels, which were excluded from the analysis.

Espécie	Nº de pixels	Nº de pixels iluminados	Nº de pixels sombreados
<i>Croton piptocalyx</i>	537	273	264
<i>Cecropia hololeuca</i>	686	443	243
<i>Astronium graveolens</i>	1.837	896	941
<i>Jacaratia spinosa</i>	300	155	145
<i>Hymenaea courbaril</i>	559	311	248
<i>Cariniana legalis</i>	7.805	3.355	4.450
<b><i>Cariniana estrellensis</i></b>	<b>278</b>	-	-
<i>Pachystroma longifolium</i>	320	95	225
<i>Galesia integrifolia</i>	405	221	184
<b><i>Piptadenia gonoacantha</i></b>	<b>158</b>	-	-
<i>Aspidosperma polyneuron</i>	701	412	289

### Seleção de amostras de treinamento

Inicialmente, as copas de 100 árvores emergentes foram manualmente delineadas, tendo como base uma composição colorida de cor verdadeira em uma escala fixa de 1:1.500. Em campo, foi possível identificar 92 destas árvores que se distribuíram entre 11 espécies (Figura 1, Tabela 1). Visto que a proporção de pixels entre as espécies foi desigual (Tabela 1), após extrair suas respostas espectrais, foram selecionados

aleatoriamente 300 pixels de cada uma delas. No entanto, apenas 9 espécies possuíam quantidade igual ou superior a esta e, assim, foi construído um conjunto de dados composto por 2.700 pixels. Deste conjunto, foram selecionadas ao acaso um total de 1.000 amostras para compor os grupos de treinamento e teste, que foram utilizados na avaliação da exatidão das classificações. Para separar os pixels em duas classes, ou seja, iluminados e sombreados, utilizou-se um limiar de brilho, escolhido de forma automática pelo método de Otsu (OTSU, 1979). O método de Otsu baseia-se no histograma da classe e busca encontrar um valor que divide a distribuição dos pixels em duas, de maneira a minimizar a variância intraespecífica (i.e. maximizar a variância interespecífica). Visando manter proporções equivalentes entre as classes, foram selecionados ao acaso 90 e 100 pixels, iluminados e sombreados para cada espécie, respectivamente.

## Classificações

A classificação de imagens hiperespectrais normalmente requer um procedimento de seleção de bandas, devido à elevada dimensionalidade deste tipo de dado. Os principais objetivos da seleção de atributos é evitar o efeito de Hughes (HUGHES, 1968), ou seja, redução da acurácia com o aumento da dimensionalidade, e diminuir o custo computacional. Neste trabalho, foi utilizado o método de regressão linear múltipla *stepwise* na seleção de bandas para classificação. O método *stepwise* verifica a importância das variáveis (bandas), incluindo ou excluindo-as em um modelo de regressão, com base em uma regra de decisão definida em termos de uma medida de significância estatística, no caso, o *p*-valor. Ao final do processo, são identificadas as bandas que foram incluídas no modelo, sendo estas utilizadas para classificar as espécies. O método *stepwise* foi aplicado separadamente para os três conjuntos de dados, ou seja, todos os pixels (iluminados+sombreados), apenas iluminados ou apenas sombreados.

Buscou-se avaliar a performance de métodos paramétricos e não-paramétricos na classificação das espécies ao nível de dossel, visto que a distribuição dos pixels das classes pode não seguir uma distribuição normal. Os métodos paramétricos utilizados foram: Análise Discriminante Linear (ADL) e Análise Discriminante Quadrática (ADQ). ADL e ADQ foram utilizados com sucesso na discriminação de espécies com dados de reflectância ao nível foliar (CASTRO-ESAU *et al.*, 2006; FERREIRA *et al.*, 2013), mas seu emprego com dados obtidos ao nível de dossel é controverso (ZHANG *et al.*, 2006). A análise discriminante é uma técnica da estatística multivariada que busca obter uma combinação linear de características observadas, ou seja, bandas, que apresente o melhor poder de discriminação entre classes. Esta combinação linear é conhecida como função discriminante e procura maximizar a variância interespecífica e minimizar a variância intraespecífica, assumindo que as classes possuem distribuição normal e compartilham da mesma matriz de covariância (como é o caso da ADL) ou não (como no caso ADQ, em que a covariância é estimada separadamente para cada classe).

Os métodos não-paramétricos empregados foram: *Support Vector Machines* (SVM) e *k*-vizinhos mais próximos (*k*-NN). SVM é uma técnica de aprendizado computacional baseada em *kernels*, que realizam a classificação construindo hiperplanos no espaço multidimensional para separar amostras de diferentes classes (VAPNIK, 1995). Este classificador tem sido utilizado com sucesso na discriminação de espécies arbóreas florestais (DALPONTE *et al.*, 2014; FÉRET e ASNER, 2013). Dentre os vários tipos de *kernels* existentes, este estudo aplicou somente o *kernel* linear, que possibilita apenas o ajuste do parâmetro de penalidade (C), estimado por um procedimento de busca exaustiva. *k*-NN é um classificador baseado em analogia, que procura *k* elementos do conjunto de treinamento que estejam mais próximos do elemento que se pretende classificar, de acordo com a menor distância euclidiana. O número ideal de vizinhos do algoritmo *k*-NN foi escolhido pelo método de validação cruzada *leave-one-out*. Todo processamento dos dados foi realizado em ambiente MatLab®.

## RESULTADOS E DISCUSSÃO

A distribuição de frequência do brilho dos pixels mostrou um comportamento claramente bimodal para algumas espécies, principalmente as com copas de elevado diâmetro. A Figura 2 apresenta a distribuição de frequência dos pixels, em termos de intensidade, para *Cariniana legalis*. A copa de uma árvore foi inserida

na figura para ilustrar o efeito da anisotropia nos valores de brilho. Esta espécie, conhecida popularmente como Jequitibá-Rosa destaca-se por atingir grandes dimensões. Devido ao tamanho e formato das suas copas, a resposta espectral ao nível de dossel é influenciada pela anisotropia, causada tanto pela própria estrutura da copa, quanto por fatores de natureza geométrica, como os ângulos de iluminação e visada. Isto causa o efeito bimodal apresentado no histograma da Figura 2, produzindo pixels sombreados e iluminados no interior da copa de uma mesma árvore. Por este motivo, assumir que as espécies possuem distribuição normal, como é feito por métodos paramétricos, pode não ser a melhor abordagem de classificação.

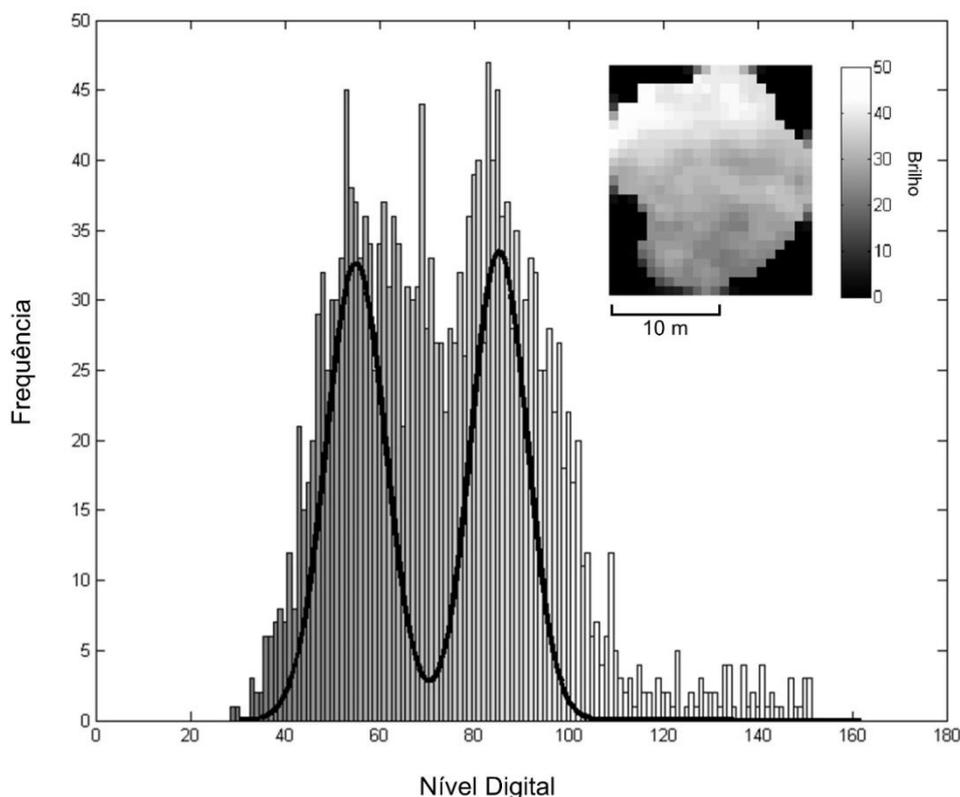


Figura 2. Distribuição de frequência dos valores de brilho dos pixels de *Cariniana legalis*, com a copa de uma árvore para ilustrar o efeito da anisotropia. Uma distribuição normal bimodal foi ajustada ao histograma (curva em preto), para representar pixels sombreados e iluminados.

Figure 2. Frequency distribution of pixel brightness values of *Cariniana legalis* with a tree crown to illustrate the effects of anisotropy. A bimodal normal distribution was fitted to the histogram (black line), in order to represent shaded and sunlit pixels.

Na realidade, após separar a distribuição dos pixels em sombreados e iluminados, observou-se um considerável ganho na exatidão de classificação, principalmente para os métodos paramétricos, como mostra a Tabela 2. Nota-se que LDA e QDA aumentaram sua performance em aproximadamente dez pontos percentuais. Dentre os métodos não-paramétricos, este aumento de performance foi maior para o SVM-linear do que para o  $k$ -NN. De modo geral, tanto para métodos paramétricos quanto não-paramétricos, a exatidão foi maior ao se utilizar apenas pixels iluminados. Acredita-se que estes pixels possuam uma relação sinal/ruído alta, o que contribuiu na caracterização da resposta espectral das espécies. A seleção de pixels iluminados provou melhorar algumas aplicações do sensoriamento remoto hiperespectral da vegetação, como, por exemplo, a extração de atributos bioquímicos do dossel (ZARCO-TERJADA *et al.*, 2004; MALENOVSKÝ *et al.*, 2013) e, inclusive, a classificação de espécies em florestas boreais (DALPONTE *et al.*, 2014). Os resultados apresentados neste trabalho, também demonstraram os benefícios da utilização de pixels iluminados na identificação de espécies arbóreas em florestas tropicais.

Tabela 2. Exatidão de classificação das espécies apresentadas na Tabela 1, utilizando classificadores paramétricos e não-paramétricos, bem como pixels iluminados e sombreados simultânea e separadamente.

Table 2. Classification accuracy of the species listed in Table 1, using parametric and non-parametric methods for classification of shaded and sunlit pixels simultaneous and separately.

Exatidão de classificação (%)					
Dados de entrada	Paramétrico		Não-paramétrico		Nº Bandas selecionadas
	LDA	QDA	SVM linear	k-NN	
Todos os pixels	74.3	83.3	82.1	89.8	30
Pixels iluminados	85.9	93	94.7	95	40
Pixels sombreados	82.5	94.1	90.2	91.3	26

## CONCLUSÕES

Neste trabalho, foi avaliado o efeito da anisotropia das copas na classificação de espécies arbóreas em floresta tropical utilizando dados de sensoriamento remoto hiperespectral de alta resolução espacial. A performance de métodos paramétricos e não-paramétricos foi avaliada em termos de exatidão de classificação, com a separação dos pixels de cada espécie em duas classes: iluminados e sombreados. A estrutura e dimensão das copas, bem como parâmetros de natureza geométrica (ângulos de iluminação e visada), contribuíram para o comportamento anisotrópico do dossel, principalmente de espécies de grande porte, influenciando sua resposta espectral. Métodos paramétricos (LDA e QDA) mostraram-se mais sensíveis ao caráter bimodal de distribuição dos pixels do que métodos não-paramétricos. Em todos os testes, a utilização de pixels iluminados mostrou-se mais adequada para classificar as espécies.

Os resultados apresentados neste trabalho mostram a importância de se considerar a anisotropia das copas na classificação de espécies arbóreas tropicais com imagens de alta resolução espacial e espectral. Esta questão deve ser mais profundamente estudada, visando o aperfeiçoamento de técnicas de processamento de imagens e a elaboração de metodologias de mapeamento de espécies com dados hiperespectrais. Visto que, em um futuro próximo, a disponibilidade destes dados irá aumentar, tanto pelo surgimento de missões hiperespectrais em nível orbital, quanto pela crescente popularização dos veículos aéreos não tripulados (VANTs) e sensores de alta resolução espectral especialmente projetados para estes veículos.

## AGRADECIMENTOS

Agradecemos a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela concessão de auxílio financeiro para o desenvolvimento desta pesquisa, através do projeto nº 2013/11.589-5, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro para realização do trabalho de campo e a empresa FotoTerra por gentilmente ceder as imagens hiperespectrais.

## REFERÊNCIAS

- CASTRO-ESAU, K.L. *et al.* Variability in leaf optical properties of Mesoamerican trees and the potential for species classification. *American Journal of Botany*, v. 93, n. 4, p. 517-530, 2006.
- COCHRANE, M.A. Using vegetation reflectance variability for species level classification of hyperspectral data. *International Journal of Remote Sensing*, v. 21, n. 10, p. 2075-2087, 2000.
- DALPONTE, M. *et al.* Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. *Remote sensing of environment*, v. 140, p. 306-317, 2014.
- FÉRET, J.B.; ASNER, G.P. Tree Species Discrimination in Tropical Forests Using Airborne Imaging Spectroscopy. *IEEE Geoscience and Remote Sensing*, v. 51, n. 1, p. 73-84, 2013.

- FERREIRA, M.P. *et al.* Analyzing the spectral variability of tropical tree species using hyperspectral feature selection and leaf optical modeling. *Journal of Applied Remote Sensing*, v. 7, 073502-1-073502-13, 2013.
- GATES, D. M. *et al.* Spectral Properties of Plants. *Applied Optics*, v. 4, n. 1, 1965.
- GOETZ, A.; VANE, G. Imaging spectrometry for earth remote sensing. *Science*, v. 228, n. 4704, p. 1147-1153, 1985.
- GUARATINI, M.T.G. *et al.* Composição florística da Reserva Municipal de Santa Genebra, Campinas, SP. *Revista Brasileira de Botânica*, v. 31, n. 2, p. 323-337, 2008.
- HUGHES, G.F. On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, vol. IT-14, n. 1, p. 55-63, 1968.
- ITT Visual Information Solutions, 2009. Atmospheric Correction Module: QUAC and FLAASH User's Guide. Version 4.7. *ITT Visual Information Solutions*, Boulder, CO.
- MALENOVSKÝ, Z. *et al.* Retrieval of spruce leaf chlorophyll content from airborne image data using continuum removal and radiative transfer. *Remote Sensing of Environment*, v. 131, p. 85-102, 2013.
- OTSU, N. A threshold selection method from gray-level histograms. *IEEE transactions on Systems, Man, and Cybernetics*, v. 9, n. 1, p. 62-66, 1979.
- PONZONI, F.J. *et al.* *Sensoriamento remoto da vegetação*, 2012. 160 p.
- VAPNIK, V.N. *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995.
- VELOSO, H.P. *et al.* *Classificação da vegetação brasileira, adaptada a um sistema universal*. IBGE, Rio de Janeiro, 1991.
- ZARCO-TEJADA, P.J. *et al.* Needle chlorophyll content estimation through model inversion using hyperspectral data from boreal conifer forest canopies. *Remote sensing of environment*, v. 89, n. 2, p. 189-199, 2004.
- ZHANG, J. *et al.* Intra and interclass spectral variability of tropical tree species at La Selva, Costa Rica: Implications for species identification using HYDICE imagery. *Remote Sensing of Environment*, v. 105, n. 2, p. 129-141, 2006.