

Revista Brasileira de Cartografia (2014) N<sup>o</sup> 66/2: 349-363  
Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto  
ISSN: 1808-0936

## **DESEMPENHO DE CLASSIFICADORES PARAMÉTRICO E NÃO PARAMÉTRICO NA CLASSIFICAÇÃO DA FISIONOMIA VEGETAL**

*Evaluating the Performance of Parametric and Non-parametric Classifiers for  
Identifying Vegetal Physiognomies*

**Alexandre Curvelo Andrade<sup>1</sup>, Cristiane Nunes Francisco<sup>2</sup> & Cláudia Maria de  
Almeida<sup>3</sup>**

**<sup>1</sup>AMS Kepler Engenharia de Sistemas  
Departamento de Geoprocessamento**

R. Alfredo Ignácio Nogueira Penido, 300/58 - Resid. Aquarius - São José dos Campos, SP, Brasil - 12246-000  
curvelo@amskepler.com

**<sup>2</sup>Universidade Federal Fluminense – UFF  
Instituto de Geociências – Departamento de Análise Geoambiental**  
Campus da Praia Vermelha – Boa Viagem – Niterói – RJ - Brasil - 24210-310  
crisnf@vm.uff.br

**<sup>3</sup>Instituto Nacional de Pesquisas Espaciais – INPE  
Coordenação de Observação da Terra – OBT/ Divisão de Sensoriamento Remoto – DSR**  
Av. dos Astronautas, 1758 - Jardim da Granja - São José dos Campos, SP - Brasil - 12227-010  
almeida@dsr.inpe.br

*Recebido em 31 de Março, 2013/ Aceito em 11 de Julho, 2013  
Received on March 31, 2013/ Accepted on July 11, 2013*

O objetivo do presente trabalho é o estudo comparativo entre dois algoritmos estatísticos e supervisionados, MAXVER (Máxima Verossimilhança) e SVM (Support Vector Machine), respectivamente paramétrico e não paramétrico, aplicados na classificação da fisionomia da vegetação em imagens de altíssima resolução espacial, com ênfase na avaliação do desempenho das classificações derivadas do incremento no número de atributos. A base de dados foi constituída pelas bandas do QuickBird fusionadas e por imagens geradas pelas transformações das bandas originais, além de dados de relevo do banco de dados TOPODATA. A área de estudo, com 16 km<sup>2</sup>, situa-se no município de Nova Friburgo, na região serrana do estado do Rio de Janeiro. No total, foram realizados quatro experimentos combinando o classificador com o número de atributos utilizados, considerando, em dois experimentos, apenas as quatro bandas espectrais corrigidas geométrica e radiometricamente e, nos outros dois, um espaço de atributos com 18 planos de informação. Os índices Kappa alcançados pelo classificador MAXVER variaram entre 0,64 e 0,66. Já os alcançados pelos experimentos conduzidos com o algoritmo SVM variaram entre 0,52 e 0,80. Considerando os índices alcançados, conclui-se que o número de planos de informação no espaço de atributos não aumenta significativamente a acurácia do desempenho do algoritmo MAXVER, mas possui grande influência no desempenho do SVM.

**Palavras-chave:** Sensoriamento Remoto, Classificação de Imagens, QuickBird, MAXVER, SVM, Coeficientes de Concordância, Índice Kappa.

## ABSTRACT

The present work is committed to conduct a comparative analysis between two supervised classifiers, Maximum Likelihood and Support Vector Machine, respectively parametric and non-parametric, for the classification of vegetal physiognomies using very high spatial resolution imagery, emphasizing the gain in performance with the accordingly increase in the number of attributes. The database consisted of pan-sharpened QuickBird images and transformed images derived from the original bands besides relief data obtained from the TOPODATA Project. The study area extends over a surface of 16 km<sup>2</sup> and is located within the municipality of Nova Friburgo, in the mountainous region of Rio de Janeiro state. In total, four experiments were accomplished, all of them combining the adopted classifier with a different number of attributes. In the first two experiments, only the four QuickBird spectral bands, previously subject to geometric and radiometric corrections, were used. In the remainder two experiments, eighteen input bands were employed. The Kappa indices obtained with the Maximum Likelihood classifier lied between 0.64 and 0.66, while those obtained for the Support Vector Machine ranged from 0.52 to 0.80. Considering the attained results, we concluded that the number of input bands does not meaningfully increase the accuracy of the Maximum Likelihood classifier, whereas this factor greatly influences the Support Vector Machine performance.

**Keywords:** Remote Sensing, Image Classification, QuickBird, Maximum Likelihood, Support Vector Machine, Agreement Coefficient, Kappa Coefficient.

## INTRODUÇÃO

A vegetação é um dos elementos da paisagem mais impactados pela atuação humana, e sendo assim, é de suma importância conhecer a sua distribuição espacial, com o intuito de adquirir informações necessárias para a adoção de práticas conservacionistas, bem como para que as intervenções antrópicas no espaço sejam previamente planejadas. Tradicionalmente, trabalhos de campo associados ao uso de documentos cartográficos vêm sendo o principal método de coleta de dados para fins de mapeamento da vegetação.

Com o avanço alcançado na obtenção de dados da superfície terrestre oriundos de sensores remotos, o uso de imagens tornou-se uma alternativa importante aos levantamentos de campo extensivos, diminuindo o tempo necessário dedicados a eles, bem como permitindo a coleta de informações em locais de difícil acesso.

As técnicas de processamento digital de imagens (PDI) têm sido massivamente testadas no mapeamento da vegetação. Dentre as técnicas mais difundidas, destaca-se a classificação que utiliza algoritmos computacionais para a identificação automática de alvos baseada

no reconhecimento de padrões. A partir da classificação dos dados digitais, os alvos que compõem a imagem ganham a designação de uma classe temática através da rotulação de cada *pixel*, ou região<sup>1</sup>, a uma classe de cobertura da terra que mais se assemelha a um determinado padrão gerado na fase de treinamento da classificação.

Atualmente, são muitos os métodos de classificação disponíveis, sendo que os mais usados nas diversas plataformas de PDI fundamentam-se na inferência estatística. Estes algoritmos, no caso da classificação *pixel a pixel*<sup>2</sup> em particular, procuram identificar a informação de cada *pixel* componente da cena, classificando-o de acordo com sua informação espectral. Os níveis de acurácia apresentados pelos diferentes métodos variam de acordo com o algoritmo de classificação e a configuração dos parâmetros do classificador (QUEIROZ *et al.*, 2004).

Os classificadores podem ser não-supervisionados ou supervisionados. Nos algoritmos do primeiro tipo, a classificação está apoiada em métodos estatísticos de análise de agrupamento e a identificação das classes ocorre sem a intervenção direta do

<sup>1</sup> No início da década de 1980, foram concebidos métodos de classificação de imagens baseados em regiões, isto é, agrupamentos rotulados de *pixels* contíguos, também denominados segmentos, segundo critérios de semelhança do valor digital e obedecendo a limiares de área.

<sup>2</sup> As técnicas tradicionais de classificação de imagens de sensoriamento remoto baseiam-se na análise *pixel a pixel* ou por regiões, enfocando as diferenças espectrais para a extração de informações temáticas. Nas técnicas de análise de imagens baseada em objeto (*object-based image analysis* - OBIA), os *pixels* são agrupados em objetos de acordo com as suas propriedades não somente espectrais, mas também espaciais, através da técnica de segmentação, e estes passam a ser utilizados como unidades primitivas para a classificação (NAVULUR, 2006).

intérprete. Os supervisionados fundamentam-se na definição das classes e na coleta das suas respectivas amostras, denominadas como áreas de treinamento, desenvolvidas, ambas as etapas, pelo usuário. Os algoritmos de classificação utilizam-se das amostras para gerar os padrões e, assim, rotular os *pixels* como pertencentes a uma determinada classe (NOVO, 2011). Estes classificadores, por sua vez, se dividem em paramétricos e não paramétricos.

O modelo paramétrico ou clássico emprega parâmetros, como média e matriz de covariância, para definir as regras do classificador para cada classe, o que demanda certa homogeneidade nas propriedades espectrais das classes (ANDREOLA *et al.*, 2009). O desempenho e a exatidão dos classificadores que usam a estratégia paramétrica estão diretamente ligados à distribuição normal dos dados. O problema reside nas imagens que apresentam comportamento mais heterogêneo, assim como nas feições anômalas (*outliers*) encontradas em imagens mais homogêneas.

Os classificadores contextuais ou não paramétricos diferenciam-se do primeiro na medida em que podem ser empregados com distribuições que não obedecem aos parâmetros da curva normal (SEMOLINI, 2002). As informações da vizinhança do *pixel* também são analisadas na classificação, a partir da aplicação de um fator de correção, denominado como contextual, que resguarda a probabilidade de rotular um *pixel* de forma duvidosa. Neste método, a classificação de cada *pixel* pode depender inclusive de todos os *pixels* da imagem (ABOUD-NETA *et al.*, 2009).

Dentro da abordagem da estatística paramétrica, a Máxima Verossimilhança (MAXVER) é um dos métodos mais usados pela comunidade de sensoriamento remoto. Este algoritmo estima, em uma abordagem quantitativa, a variância e a covariância dos padrões espectrais das amostras de treinamento. Primeiramente, pressupõe-se que os *pixels* que formam o conjunto de treinamento apresentam uma distribuição normal (Gaussiana). Desta forma, um vetor de médias e uma matriz de covariância podem descrever os padrões de respostas encontradas na imagem.

Segundo Centeno (2009), a probabilidade de se localizar um *pixel* com um vetor de estado  $x$  em uma imagem pode ser descrita por  $p(x)$

no método MAXVER. Como em uma imagem existem apenas alguns vetores  $x$ , o processo de classificação procura definir a classe mais provável para estes *pixels*. Dado um padrão de resposta espectral (classe de cobertura da terra), a probabilidade de encontrar um determinado *pixel* neste grupo pode ser descrita como  $(p(X | w_j))$ . Vale lembrar que o conhecimento dos valores exatos de probabilidade de cada classe não é o objetivo principal. O conhecimento destes valores se faz necessário apenas para viabilizar as comparações entre as classes. A regra mais importante refere-se à probabilidade de um *pixel* pertencer ou não a uma determinada classe, ou seja, dado um *pixel*  $x$ , ele pertence à classe  $w_j$  se a probabilidade que o associa à classe  $w_j$  é maior que a probabilidade que o associa a qualquer outra classe.

Apesar de ser um classificador tradicionalmente utilizado, o MAXVER não apresenta bom desempenho quando aplicado a imagens de alta resolução espacial. A definição precisa das áreas de treinamento, assim como a seleção de muitos *pixels* se faz necessária para produzir bons resultados com a utilização do classificador MAXVER (INPE, 2002).

O aumento na dimensionalidade dos dados eleva a acurácia da imagem classificada, na medida em que novas variáveis são acrescentadas. Entretanto, a partir de um determinado número de bandas incluídas no espaço dimensional, a acurácia decresce devido ao limite na geração de novas amostras de treinamento (QUEIROZ *et al.*, 2004; LICZBINSKI, 2007). Isto porque o número de parâmetros a serem estimados na matriz de covariância, particularmente, cresce rapidamente à medida que a dimensionalidade dos dados aumenta. Se o número das amostras de treinamento não aumenta na mesma proporção, a confiabilidade na estimação destes parâmetros decresce. Em outras palavras, quando o tamanho das amostras de treinamento permanece constante, o acréscimo de bandas espectrais adicionais resulta na incerteza da estimativa dos parâmetros e, conseqüentemente, em decréscimo na acurácia da imagem temática produzida pelo classificador (LICZBINSKI, 2007). Este efeito é conhecido como o fenômeno de *Hughes* ou a maldição da dimensionalidade (ERBERT, 2001 *apud* QUEIROZ, *et al.*, 2004).

Pesquisas apontam que a classificação

gerada pelo algoritmo MAXVER apresenta precisão prejudicada quando há ampla superposição entre os dados de treinamento no espaço  $n$ -dimensional (COSTA *et al.*, 1998). Pesquisadores, valendo-se da não-obrigação de pressupor a normalidade dos dados de treinamento, vêm buscando alternativas não paramétricas de classificação por *pixel*, com o intuito de obter melhores índices de acurácia na classificação de imagens (COSTA *et al.*, 1998).

Recentemente, uma nova família de algoritmos baseados em aprendizado de máquina (AM), área de pesquisa da Inteligência Computacional que estuda o desenvolvimento de técnicas capazes de extrair conhecimento a partir de amostras de dados, tem recebido grande atenção por parte de pesquisadores como técnica promissora para a classificação de imagens. Essas técnicas, baseadas em modelos matemáticos, têm a capacidade de “aprender” a partir de um número de amostras e generalizar o conhecimento gerado para toda a imagem (QUEIROZ *et al.*, 2004). Segundo Vapnik (1995) *apud* Nascimento *et al.* (2009), os algoritmos de aprendizagem de máquina têm como finalidade a construção de limites de decisão que possibilitem a maior separação entre classes, através da minimização dos erros.

Conforme relatam Lorena e Carvalho (2007), as técnicas de aprendizado de máquina são baseadas em um princípio de inferência estatística chamada indução, que consiste na obtenção de conclusões genéricas a partir de um conjunto particular de exemplos. No aprendizado indutivo supervisionado, o conhecimento do ambiente (das classes) é feito através de um conjunto de exemplos que seguem uma lógica simples, tal qual: entrada  $\rightarrow$  saída desejada, cujo objetivo é fazer com que as representações dos padrões presentes em uma imagem sejam capazes de produzir saídas corretas para novas entradas. Logo, a obtenção de um classificador por um algoritmo de aprendizado de máquina pode ser considerado como sendo um processo de investigação (LORENA; CARVALHO, 2007).

Entre essas técnicas estão as máquinas de vetores de suporte (*Support Vector Machines* – SVMs), conhecidas pelo seu alto grau de generalização e por embasamento matemático sólido apoiado em probabilidade

estatística. Estudos recentes demonstram que o classificador SVM tem proporcionado resultados superiores aos de outros métodos de classificação, como Redes Neurais e Árvores de Decisão, para imagens de sensoriamento remoto (NASCIMENTO *et al.* 2009). Dixon e Candade (2008) demonstraram que as SVMs têm obtido sucesso em várias aplicações de sensoriamento remoto, notadamente em classificações que apresentam problemas em vista da alta dimensionalidade.

As SVMs são apresentadas como uma classificação supervisionada derivada da teoria de aprendizagem estatística que, na maioria das vezes, gera bons resultados para dados complexos ou ruidosos (ENVI, 2009). Apresentam, basicamente, a abordagem dos problemas de dois modos. No primeiro, os erros de classificação não são considerados. No segundo, são instituídas variáveis de folga, de maneira que os padrões fora dos valores esperados da classe possam ser desconsiderados em função do limiar de aceitação usado, diminuindo, desta forma, a probabilidade de erros na classificação (ALIXANDRINI, 2009). De modo geral, os algoritmos baseados em uma SVM obtêm ótimos resultados no que se refere à generalização por apresentar baixo *overfitting*, ou seja, ajuste demasiado dos dados de treinamento (super-especialização), pois, além de extraírem os parâmetros gerais que permitem a generalização, armazenam os ruídos e as peculiaridades, tolerando o reconhecimento de padrões não observados durante a etapa de treinamento (ALIXANDRINI, 2009).

Outra característica importante sobre as SVMs é a robustez no tratamento de dados de alta dimensionalidade, perante objetos de grandes dimensões, como são as imagens de satélites (LORENA; CARVALHO, 2007). Normalmente, outros classificadores encontram problemas com este tipo de dados, devido à grande ocorrência de *overfitting*. Além disto, as SVMs revelam grande capacidade de discernimento para dados com alta entropia, definida como a quantidade de incerteza associada a esta imagem (aleatoriedade), ou seja, quanto maior a entropia de uma imagem, mais atípica (fora de padrão) será a imagem analisada.

O objetivo do presente trabalho é comparar dois algoritmos estatísticos e supervisionados, MAXVER e SVM, aplicados na classificação da fisionomia da vegetação em imagens de altíssima

resolução espacial, com ênfase na avaliação do desempenho das classificações derivadas do incremento de novos atributos. A área de estudo, com aproximadamente 16 km<sup>2</sup>, está situada no município de Nova Friburgo, na região serrana do estado do Rio de Janeiro, e é coberta por formações vegetais com diferentes níveis de intervenção antrópica (Figura 1). A opção pela classificação das fitofisionomias deve-se às dificuldades para se realizar mapeamento dos estágios sucessionais de vegetação baseando-se apenas em métodos de classificação automática de imagens de sensoriamento remoto.

## MATERIAL E MÉTODOS

Para a classificação da fisionomia da cobertura vegetal da área de estudo, foram utilizadas as bandas fusionadas do satélite QuickBird, obtidas em julho de 2008. Essa plataforma possui dois sensores que atuam no modo pancromático (PAN) e multi-espectral (MS), com resolução radiométrica de 11 bits.

À base de dados, também foram incorporados o Modelo Digital de Elevação (*Digital Elevation Model - DEM*) e a grade de declividade, ambos com 30 m de resolução espacial, oriundos do projeto TOPODATA (VALERIANO, 2005). Esse projeto consistiu no refinamento de dados SRTM (*Shuttle Radar Topography Mission*), com 90 m de resolução, disponibilizados pelo USGS (*United States*

*Geological Survey*), por meio de processo de *krigagem* ou interpolação geoestatística (VALERIANO, 2005).

Para comparar os resultados entre algoritmos de classificação automática supervisionados, paramétrico e não paramétrico, assim como para avaliar a influência que o incremento de novos planos de informação adicionados ao espaço de atributos (*n*-dimensional) exerce no desempenho da classificação, foram selecionados os algoritmos MAXVER, paramétrico, e SVM, não paramétrico. A escolha foi pautada na revisão da literatura que revela o MAXVER como um dos algoritmos paramétricos mais usados (QUEIROZ *et al.*, 2004), enquanto o SVM surge, atualmente, como um dos algoritmos não paramétricos com melhor desempenho, notadamente em imagens ruidosas (DIXON; CANDADE, 2008).

As etapas gerais do trabalho, descritas a seguir, consistiram no pré-processamento das imagens, na geração de imagens derivadas de operações entre as bandas fusionadas, na classificação pelos métodos supervisionados não paramétrico e paramétrico e na validação dos resultados.

### Pré-processamento dos dados

As imagens obtidas por sensores remotos apresentam distorções geométricas e radiométricas, e a finalidade dessa etapa foi

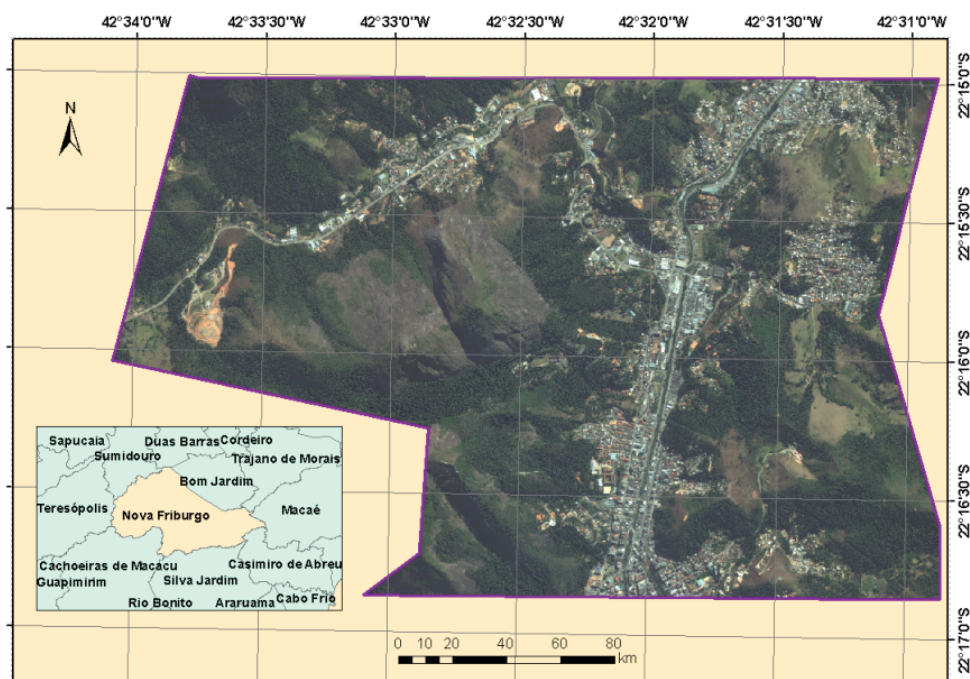


Fig. 1 – Localização da área de estudo, Centro de Nova Friburgo, RJ.

o tratamento dos dados em sua forma bruta, com objetivo de eliminar essas distorções, preparando-as para o processo de classificação (MOREIRA, 2005). Foram realizadas as operações de ortorretificação e de conversão dos níveis digitais para valores de radiância no topo da atmosfera.

Segundo TOUTIN (2004), as distorções geométricas podem ser causadas pela variação do movimento, ângulo e altitude da plataforma e do sensor, refração e turbulência atmosférica, curvatura e rotação da Terra, topografia, e diferenças entre o geoide e elipsoide. A finalidade da ortorretificação é recuperar a qualidade geométrica da cena, eliminando ou minimizando essas distorções, de tal forma que os dados possuam escala e projeção próprias de um mapa. No presente trabalho, a ortorretificação foi executada na plataforma ENVI 4.7 pelo método de Coeficientes Polinomiais Racionais (*Rational Polynomial Coefficients* - RPC).

A conversão dos níveis digitais para valores de radiância<sup>3</sup> é fundamental em estudos que comparam bandas ou imagens geradas por diferentes sensores. Embora representem o comportamento espectral dos alvos da cena, os níveis digitais não possuem uma unidade escalar que comporte a comparação entre imagens de diferentes fontes (PONZONI; SHIMABUKURO, 2009). Assim sendo, a radiância torna-se uma unidade radiométrica mais concisa para ser empregada em sensoriamento remoto (JENSEN, 2009).

Como este trabalho previa operações entre bandas, os parâmetros para a transformação dos níveis digitais para radiância de topo de atmosfera, obtidos em DIGITAL GLOBE (2003), foram aplicados às imagens. No entanto, a correção atmosférica em nível de superfície, ou mesmo aparente (topo da atmosfera), não se mostrou necessária, visto que não é o objetivo do presente estudo a análise de séries temporais de imagens.

## Processamento dos dados

Esta etapa consistiu na realização de operações entre as bandas com o objetivo de aumentar o contraste das imagens e, assim, inserir informações no banco de dados que possibilitassem a discriminação entre os alvos pelos algoritmos de classificação. Esse procedimento foi adotado, pois, em trabalho realizado por Francisco e Almeida (2012) sobre mineração de dados visando à classificação de imagens, com banco de dados constituído por 138 descritores relacionados a imagens derivadas e a bandas sem processamento radiométrico, as árvores de decisão foram constituídas, predominantemente, por descritores correspondentes às imagens derivadas, enquanto apenas um único nó correspondeu a uma banda sem qualquer tipo de processamento.

Foram aplicadas a Análise das Principais Componentes (*Principal Components Analysis* - PCA) e a conversão do sistema de cores aditivo RGB (*Red* ou Vermelho, *Green* ou Verde, *Blue* ou Azul) para o espaço IHS (*Intensity* ou Intensidade - I, *Hue* ou Matiz - H, *Saturation* ou Saturação - S).

Também foram realizadas operações para cálculo de índices de vegetação - o Índice de Vegetação por Diferença Normalizada (*Normalized Difference Vegetation Index* - NDVI) e o Índice de Vegetação por Razão Simples (*Simple Ratio* - SR) - com a finalidade de explorar as propriedades espectrais da vegetação, em especial nas regiões do vermelho (na faixa do espectro visível) e do infravermelho próximo. Estes índices fundamentam-se no comportamento oposto da reflectância da vegetação nessas duas regiões espectrais (PONZONI; SHIMABUKURO, 2009). Ressalta-se que no trabalho realizado por Francisco e Almeida (2012), citado no parágrafo anterior, as árvores de decisão foram constituídas por no mínimo 30% de descritores relacionados ao NDVI.

A PCA é uma transformação linear

<sup>3</sup> A radiância ( $L_\lambda$ ) consiste no fluxo radiante, refletido ou emitido, por unidade de área projetada e dentro de um ângulo sólido em uma direção específica. A radiação anotada pelo sensor, avaliada como radiância de topo de atmosfera (Top-of-Atmosphere Radiance), corresponde à radiância aparente, pois sofre influência atmosférica e, portanto, não representa a resposta espectral pura dos alvos contidos na superfície terrestre (JENSEN, 2009).

dos dados da imagem, de forma a rotacioná-los e deslocá-los para um novo sistema de coordenadas, elevando ao máximo a sua variância (CENTENO, 2009), e é baseada em transformações algébricas aplicadas sobre a matriz de covariância entre as bandas de uma imagem. A transformação PCA gera número de componentes (bandas) igual ao número de bandas da imagem de entrada. A primeira componente é a que possui maior variância do conjunto, e este valor decresce sucessivamente nas componentes seguintes (LIU, 2006). O principal propósito dessa transformação (ou análise) por principais componentes é reduzir a correlação entre as bandas da imagem utilizada. Essa técnica é útil para aumentar o volume da informação, isolar componentes de ruído e reduzir a dimensionalidade de ruído dos dados (CENTENO, 2009). Convém mencionar que no trabalho realizado por Francisco e Almeida (2012), já anteriormente citado, as árvores de decisão foram constituídas por pelo menos um descritor relacionado às principais componentes, inclusive a última componente principal foi selecionada para discriminar plantio de eucalipto da vegetação nativa..

A transformação RGB-IHS fundamenta-se na conversão das cores do espaço cúbico RGB (contendo os canais vermelho, verde e azul) para o espaço cônico IHS. Neste espaço, a intensidade representa o brilho dos alvos; a matiz, a cor através do comprimento de onda, e a saturação corresponde ao grau de pureza da cor. Neste trabalho, foram geradas as transformações relativas às composições 3(R)2(G)1(B) e 4(R)3(G)2(B), obtendo-se seis novas bandas transformadas.

Os índices de vegetação baseiam-se em operações aritméticas entre bandas, utilizados tanto para diagnosticar parâmetros biofísicos e de biomassa quanto parâmetros mais específicos, como densidade de cobertura vegetal, atividade fotossintética e produtividade (PONZONI; SHIMABUKURO, 2009). Comumente, são utilizadas as bandas do infravermelho próximo (entre 0,74  $\mu\text{m}$  e 0,90  $\mu\text{m}$ ) e do vermelho (entre 0,63 e 0,69  $\mu\text{m}$ ), devido ao comportamento

antagônico em relação à resposta espectral esperada da vegetação. Quanto maior a densidade da cobertura vegetal em uma determinada área, menor a reflectância na região do visível, devido à maior oferta de pigmentos fotossintetizantes. Por outro lado, maior é reflectância na região do infravermelho próximo, devido ao espalhamento múltiplo da radiação eletromagnética nas diferentes camadas de folhas (PONZONI; SHIMABUKURO, 2009).

O NDVI, índice de vegetação calculado neste trabalho, corresponde a uma “razão simples” e normalizada entre as bandas do infravermelho próximo e vermelho, com intervalo compreendido entre -1 a +1 (PONZONI; SHIMABUKURO, 2009). Huete *et al.* (2002) *apud* Jensen (2009) salientam que o NDVI pode saturar em condições de elevado Índice de Área Foliar (IAF)<sup>4</sup>, o que dificultaria a diferenciação entre as fisionomias em formações vegetais mais densas.

O índice SR é a razão entre as bandas do infravermelho próximo e do vermelho. Revela-se um bom indicador do crescimento da cultura, assim como de estimativa de biomassa, porque as folhas verdes absorvem a radiação na faixa do visível (principalmente no vermelho) e refletem na faixa do infravermelho próximo (LIU, 2006). Valiosas informações sobre o IAF e a biomassa da vegetação são fornecidas pelo índice SR (SCHLERF *et al.*, 2005 *apud* LIU, 2006), por ser especialmente sensível a variações de biomassa ou IAF em vegetação de grande biomassa, tais como florestas (HUETE *et al.*, 2002, *apud* JENSEN, 2009).

### **Classificação supervisionada**

O banco de dados para classificação apresentou, no total, dezoito planos de informação, composto por: quatro bandas multiespectrais fusionadas do sistema QuickBird convertidas para radiância de topo de atmosfera, quatro componentes principais, seis componentes IHS, dois índices de vegetação, um DEM e uma grade de declividade.

Apesar de o foco do trabalho ser o mapeamento fisionômico da vegetação, houve

<sup>4</sup> Corresponde a razão entre a área foliar do dossel e a unidade de superfície projetada no solo ( $\text{m}^2/\text{m}^2$ ) (PONZONI & SHIMABUKURO, 2009).

a necessidade de definir classes representativas dos alvos de não-vegetação presentes na imagem para executar a classificação automática, pois não é possível ao algoritmo classificar apenas os pixels “desejados”. Além das três classes de vegetação, foram criadas três de não-vegetação (sombra, antrópico e afloramento), devido à diferença espectral entre esses alvos. A classe antrópica abrangeu um conjunto de alvos distintos, no entanto, como apresentam respostas espectrais semelhantes, foram inseridos em uma única classe. A seguir, estão descritas as classes de mapeamento definidas.

**Vegetação arbórea** – fisionomia vegetal com dossel elevado e denso, predominantemente formada por árvores de diversas espécies (em conjunto ou isoladas) sob diversos níveis de influência antrópica. Apresenta grande heterogeneidade, tanto com relação à forma, quanto à resposta espectral, ocasionada pelas diferentes espécies vegetais componentes da cena e seus respectivos parâmetros biofísicos.

**Vegetação herbácea-arbustiva** - inclui fisionomia formada por vegetação rasteira (gramíneas), com presença de ervas e arbustos, sem ocorrência de árvores, caracterizada por dossel regular e contínuo, com menor influência do solo, além de rara presença de sombras.

**Vegetação de gramínea** – predomínio de gramíneas, sem formar uma superfície contínua e, desta forma, a resposta espectral recebe influência do solo, distinguindo-se das demais classes de vegetação.

**Afloramento Rochoso** – superfície rochosa localizada em regiões mais altas e íngremes. Pode apresentar uma fina camada de solo coberta por vegetação rupestre com respostas espectrais próximas às da classe de vegetação.

**Área antrópica** – composta por áreas urbanizadas, construções isoladas, vias pavimentadas, solos expostos e áreas recentemente queimadas.

**Sombra** - ausência total ou parcial da incidência solar, devido ao sombreamento provocado por algum obstáculo (dossel, construções, topografia etc.), com distribuição espacial em toda a imagem.

A etapa seguinte consistiu na coleta de amostras de treinamento, realizada no ambiente do ENVI, distribuídas na imagem e representativas da diversidade de respostas

espectrais de cada uma das classes. Foram coletados 724 *pixels* distribuídos entre as seguintes classes: vegetação arbórea – 134, vegetação herbácea-arbustiva – 98, vegetação de gramínea – 120, afloramento rochoso – 112, área antrópica – 164, e sombra – 134.

Foram feitos quatro experimentos de classificação, dois utilizando o algoritmo MAXVER, e dois o algoritmo SVM utilizando a função de base radial (FBR). Este último algoritmo possui, além da radial, outras três funções – linear, quadrática e polinomial, que são variações da FBR, sendo que esta última apresenta os melhores resultados na separação ótima de classes (NASCIMENTO *et al.*, 2009).

Duas das classificações, cada uma com um algoritmo, foram executadas usando apenas quatro bandas pré-processadas da imagem QuickBird, ou seja, corrigidas das distorções geométricas e radiométricas. Outras duas classificações foram executadas, cada uma com um algoritmo, utilizando o total de dezoito planos de informação correspondentes aos dados topográficos, às bandas originais corrigidas e às bandas derivadas por processamento digital.

### **Avaliação da acurácia das classificações**

Para a avaliação da acurácia das classificações, foram coletadas cerca de 300 amostras para cada experimento, estratificadas pela área esperada das seis classes, totalizando 1,2 mil amostras. Procurou-se obedecer ao número mínimo de 50 amostras por classe, definido por Congalton e Green (2009), para mapas cobrindo menos de um milhão de acres e com menos de 12 classes. No entanto, devido à reduzida área de algumas classes, não foi possível a observância deste número para todas elas.

A seguir, foram construídas as matrizes de erros, compostas, nas colunas, pelas amostras de referência, e nas linhas, pelas classes obtidas na classificação. Na matriz, são expressos os erros de omissão, ou seja, amostras que não foram classificadas de acordo com as classes de referência, e os erros de comissão, correspondentes a amostras de referência classificadas erroneamente como pertencentes a outras classes. A partir das matrizes, foram calculados os seguintes índices: (a) exatidão global - relação entre o número de amostras classificadas corretamente sobre o número



total de amostras de referência; (b) exatidão do produtor – relativa aos erros de omissão, a qual representa a relação entre o número de amostras classificadas corretamente da classe  $k$  e o número total de amostras de referência da classe  $k$ ; (c) exatidão do usuário - referente aos erros de comissão, a qual representa a relação entre o número de amostras classificadas corretamente da classe  $k$  e o número total de amostras classificadas da classe  $k$ ; (d) *Kappa* ( $K$ ) - analisa todos os elementos da matriz de erros (equação 1); e (e) *Kappa* condicional - (CONGALTON; GREEN, 2009).

$$K = \frac{\sum_{i=1}^k [n_{ii} - \sum_{i=1}^k (n_{i+}] n_{+i}]}{n^2 - \sum_{i=1}^k [(n_{i+}] n_{+i}]} \quad (1)$$

em que:

$n_{ii}$  – número total de amostras corretamente classificadas da classe  $k$ ;

$n_{i+}$  – número total de amostras classificadas da classe  $k$ ;

$n_{+i}$  – número total de amostras coletadas da classe  $k$ ; e

$n$  – número total de amostras.

O índice *Kappa* varia de -1 a 1, e quanto mais próximo de 1, maior a precisão da classificação. De acordo com Landis e Koch (1977), a qualidade da classificação pode ser definida em função dos seguintes intervalos de índice *Kappa*: [ $<0,00$ ]: muito ruim; [ $0,00 - < 0,20$ ]: ruim; [ $0,20 - < 0,40$ ]: aceitável; [ $0,40 - < 0,60$ ]: bom; [ $0,60 - < 0,80$ ]: muito bom; [ $0,80 - 1,00$ ]: excelente.

O índice *Kappa* condicional avalia a acurácia de cada classe e é calculado com base no mesmo princípio do índice *Kappa* utilizado para a avaliação global da classificação, assim como a interpretação dos resultados obedece aos mesmos princípios. O índice *Kappa* condicional do usuário avalia os erros de comissão (equação 2):

$$K_u = \frac{nn_{ii} - n_{i+} n_{+i}}{nn_{+i} - n_{+i} n_{i+}} \quad (2)$$

## RESULTADOS E DISCUSSÕES

A tabela 1 refere-se à matriz de erros correspondente à classificação do Experimento 1, ou seja, obtida pela aplicação do algoritmo MAXVER utilizando apenas as quatro bandas

pré-processadas do QuickBird. O índice de exatidão global foi de 70%, e o índice *Kappa*, 0,64, considerado muito bom de acordo com Landis e Koch (1977). Apenas a classe vegetação herbácea-arbustiva apresentou erros superiores a 50%, ou seja, a exatidão do produtor foi de 48%, e a exatidão do usuário foi de 37%, confundindo-se principalmente com a classe vegetação arbórea. Da mesma forma, os índices de exatidão da classe vegetação arbórea apresentaram grande confusão com herbácea-arbustiva, tanto por erros de omissão como de comissão, com a exatidão em torno de 60%. Entre as classes de vegetação, a que apresentou melhor desempenho foi a vegetação de gramínea, com exatidão do produtor de 64% e do usuário de 83%. As classes de não-vegetação – antrópica, sombra e afloramento rochoso – apresentaram menores erros do que as classes de vegetação, com a maioria dos índices de exatidão acima de 80%.

A tabela 2 corresponde à matriz de erros da classificação do Experimento 2, ou seja, obtida pelo algoritmo SVM, usando apenas as quatro bandas pré-processadas do QuickBird. O índice de exatidão global foi de 60%, e o índice *Kappa*, 0,52, considerado bom de acordo com Landis e Koch (1977). A classe vegetação herbácea-arbustiva apresentou resultado ruim, com exatidão do produtor correspondendo apenas a 6%, confundindo-se com a classe arbórea, e exatidão do usuário próxima a 20%, confundindo-se com a classe sombra. A vegetação de gramínea apresentou índices de exatidão próximos a 50%, apresentando maiores erros de omissão com a classe afloramento e de comissão com a classe vegetação arbórea. Entre as classes de vegetação, a arbórea apresentou melhor desempenho, com índices de exatidão acima de 50%. As classes de não-vegetação apresentaram erros menores do que as classes de vegetação, com a maioria dos índices de exatidão acima de 70%.

A tabela 3 refere-se à matriz de erros da classificação do Experimento 3, ou seja, obtida pelo algoritmo MAXVER, usando o total de 18 planos de informação, compostos pela bandas pré-processadas e processadas do QuickBird, além do DEM e grade de declividade. O índice de exatidão global foi de 73%, e o índice *Kappa*, 0,66, considerado muito bom de acordo com Landis e Koch (1977). A classe vegetação

herbácea-arbustiva apresentou pior desempenho entre as classes de vegetação, com exatidão do produtor correspondendo apenas a 31%, e do usuário próxima a 50%, confundindo-se principalmente com a classe arbórea. Por sua vez, a classe arbórea apresentou índices de exatidão superiores a 60%, confundindo-se com a herbácea-arbustiva. Entre as classes de vegetação, a gramínea apresentou melhor desempenho, com índices de exatidão acima de 80%. As classes de não-vegetação apresentaram erros menores do que as classes de vegetação, com a maioria dos índices de exatidão acima de 80%.

A tabela 4 refere-se à matriz de erros da classificação referente ao Experimento 4 (Figura 2), ou seja, obtida pelo algoritmo SVM, usando 18 planos de informação. O índice de exatidão global foi de 84%, e o índice *Kappa*, 0,80, considerado muito bom de acordo com Landis e Koch (1977). Neste experimento, os erros de omissão e comissão foram pequenos em todas as classes, com a maioria dos índices de exatidão superiores a 70%. A classe herbácea-arbustiva apresentou pior desempenho entre as classes de vegetação, com exatidão do produtor

correspondendo 74%, e do usuário próxima a 60%, confundindo-se principalmente com a classe arbórea. Esta, por sua vez, apresentou índices de exatidão variando entre 70 e 90%, confundindo-se principalmente com a classe herbácea-arbustiva. Entre as classes de vegetação, a gramínea apresentou melhor desempenho, com índices de exatidão acima de 90%. As classes de não-vegetação apresentaram erros semelhantes aos das classes de vegetação, com a maioria dos índices de exatidão acima de 80%.

A figura 3 apresenta o desempenho das classificações a partir do incremento do número de planos de informação no espaço de atributos. Verifica-se que o aumento do número de atributos melhorou sensivelmente o desempenho da classificação (Experimento 4) realizada com o algoritmo SVM, com índice *Kappa* passando de 0,52 para 0,80. No entanto, o desempenho da classificação com MAXVER (Experimento 3) não apresentou melhora significativa, com índice *Kappa* passando de 0,64 para 0,66.

A figura 4 apresenta o *Kappa* condicional de usuário das classes de cobertura da terra, confirmando a tendência apresentada nas matrizes de erros da classificação de que as classes de não-

Tabela 1 - Matriz de erros de classificação do MAXVER com 4 bandas (Experimento 1).

Classes		Amostras de referência						Total classificadas
		Afloramento	Árborea	Herbácea arbustiva	Gramínea	Sombra	Antrópico	
Classificadas	Afloramento	27	3		11	1	2	44
	Árborea		42	15	6	8	1	72
	Herbácea arbustiva		23	15	1	1	1	41
	Gramínea	3	1		35		3	42
	Sombra	1	2	1	1	43	2	50
	Antrópico				1		46	47
<b>Total coletadas</b>		31	71	31	55	53	55	296
Exatidão do produtor		87%	59%	48%	64%	81%	84%	70%
Exatidão do usuário		61%	58%	37%	83%	86%	98%	

Tabela 2 - Matriz de erros de classificação do SVM com 4 bandas (Experimento 2).

Classes		Amostras de referência						Total classificadas
		Afloramento	Árborea	Herbácea arbustiva	Gramínea	Sombra	Antrópico	
Classificadas	Afloramento	26	5	1	21	5	4	62
	Árborea		39	19	1	4		63
	Herbácea arbustiva		1	2		6		9
	Gramínea	3	20	6	29		1	59
	Sombra	2	7	3	2	37	1	52
	Antrópico				6	1	49	56
<b>Total coletadas</b>		31	72	31	59	53	55	301
Exatidão do produtor		84%	54%	6%	49%	70%	89%	60%
Exatidão do usuário		42%	62%	22%	49%	71%	88%	

Tabela 3 - Matriz de erros de classificação do MAXVER com 18 planos de informação (Experimento 3).

Classes		Amostras de referência						Total classificadas
		Afloramento	Árborea	Herbácea arbustiva	Gramínea	Sombra	Antrópico	
Classificadas	Afloramento	25		3			3	31
	Árborea		55	18	4	13	1	91
	Herbácea arbustiva		10	11		2		23
	Gramínea	4		1	43	2	1	51
	Sombra	2	6	1	3	34	3	49
	Antrópico			2	2		47	51
<b>Total coletadas</b>		31	71	36	52	51	55	296
Exatidão do produtor		81%	77%	31%	83%	67%	85%	73%
Exatidão do usuário		81%	60%	48%	84%	69%	92%	

Tabela 4 - Matriz de erros de classificação do SVM com 18 planos de informação (Experimento 4).

Classes		Amostras de referência						Total classificadas
		Afloramento	Árborea	Herbácea arbustiva	Gramínea	Sombra	Antrópico	
Classificadas	Afloramento	31			2	1	6	40
	Árborea		51	4		2		57
	Herbácea arbustiva		13	23		4		40
	Gramínea		1	3	50		1	55
	Sombra		6	1	3	45		55
	Antrópico					1	48	49
<b>Total coletadas</b>		31	71	31	55	53	55	296
Exatidão do produtor		100%	72%	74%	91%	85%	87%	84%
Exatidão do usuário		78%	89%	58%	91%	82%	98%	

vegetação e de gramínea obtiveram melhores índices de acurácia nos experimentos realizados, com valores acima de 0,60 na maioria das situações testadas. Neste grupo, destaca-se a classe antrópica, com *Kappa* superior a 0,85, valores considerados excelentes segundo Landis e Koch (1977). Entre as classes de vegetação, a herbácea-arbustiva apresenta o pior desempenho nos experimentos realizados, com valores de *Kappa* condicional abaixo de 0,60, alcançando 0,13 no Experimento 2. O melhor desempenho foi da gramínea, com índices entre 0,80 e 0,90, com exceção do Experimento 2. Verifica-se também que o Experimento 4 (algoritmo SVM e 18 bandas) apresentou melhor desempenho em quase todas as classes, com índices variando entre 0,75 e 0,98, exceção feita às classes afloramento rochoso e herbácea-arbustiva. No caso desta última, contudo, o índice, próximo a 0,5, foi o maior apresentado por esta classe nos experimentos realizados

A figura 5 mostra o desempenho do algoritmo SVM através da comparação do

*Kappa* condicional de usuário das classes os Experimentos 2 e 4. Verifica-se que, com o aumento do número de planos de informação no espaço de atributos, a classificação apresentou sensível melhora na discriminação das classes estabelecidas, com aumento substancial dos valores do *Kappa* condicional em todas as classes. Com destaque para a classe herbácea-arbustiva, que apesar de ter apresentado o pior desempenho nos outros três experimentos, obteve, no Experimento 4, o seu melhor desempenho.

No entanto, a figura 6 não evidenciou o mesmo comportamento nos experimentos que utilizaram o algoritmo MAXVER (Experimentos 1 e 3). Verificou-se que para as classes de vegetação e afloramento rochoso, os índices *Kappa* se elevaram com o incremento de atributos, com aumento entre 0,20 (afloramento) e 0,05 (arbórea), enquanto os índices das classes sombra e antrópica reduziram-se com o incremento de atributos, 0,25 e 0,05, respectivamente.

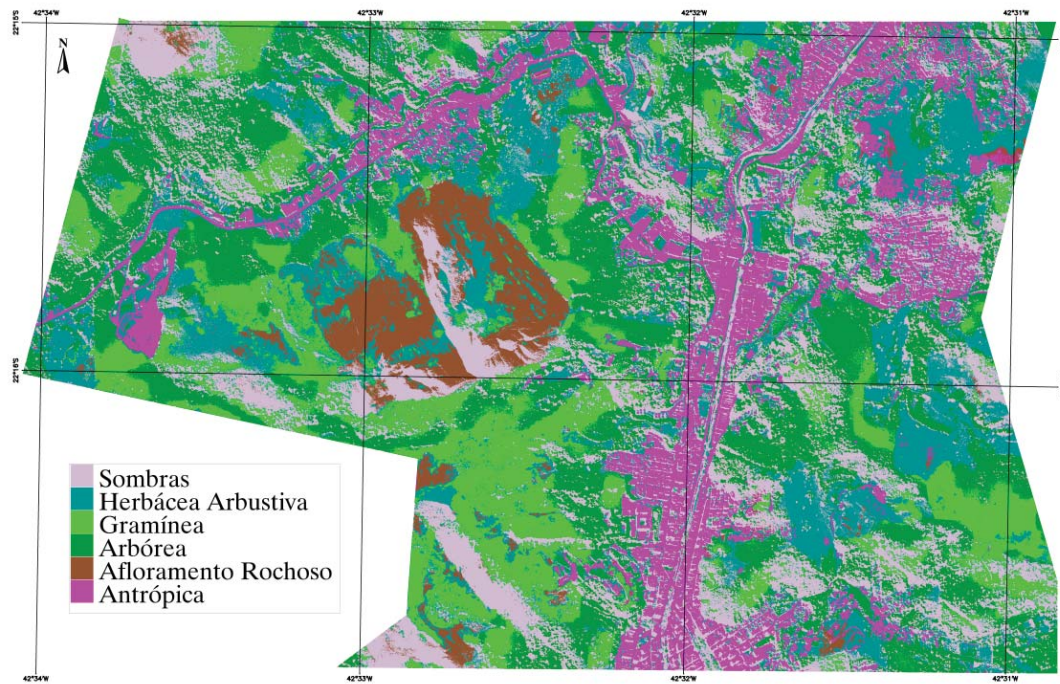


Fig. 2 – Experimento 4: classificação de cobertura da terra com algoritmo SVM e base de dados constituída por 18 planos de informação.

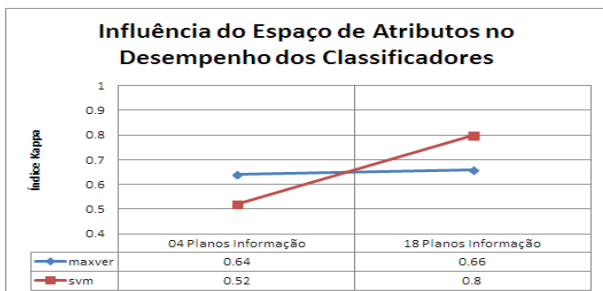


Fig. 3 – *Kappa* da classificação de acordo com o número de planos de informação no espaço de atributos.

## CONCLUSÕES

Tendo em vista que o objetivo deste estudo foi analisar o desempenho entre um algoritmo não paramétrico e um paramétrico na classificação de fisionomias vegetais em imagens de alta resolução, pode-se afirmar que o classificador não paramétrico, SVM, apresentou melhor desempenho do que o paramétrico, MAXVER, quando utilizou maior número de planos de informação no espaço de atributos de classificação. A classificação correspondente ao algoritmo SVM com incremento de planos de informação atingiu *Kappa* de 0,80, valor considerado excelente segundo Landis e Koch (1977).

Nos experimentos que aproveitaram apenas as quatro bandas pré-processadas

do QuickBird, o desempenho do algoritmo paramétrico MAXVER mostrou-se superior tanto na avaliação global quanto individualmente para as classes de vegetação, com exceção da classe vegetação arbórea, que apresentou melhor desempenho pelo algoritmo SVM em ambos os cenários (com 4 e 18 bandas).

Considerando o experimento de melhor desempenho (SVM com 18 planos de informação), as classes referentes às fisionomias vegetais alcançaram excelente desempenho individual, com *Kappa* condicional próximo a 0,90, exceto a classe herbácea-arbustiva, que obteve desempenho próximo a 0,50, considerado bom, segundo Landis e Koch (1977), no entanto foi a que teve o segundo maior aumento do *Kappa* condicional. O aumento do número de atributos revelou-se de fundamental importância para o desempenho individual das classes alcançado pelo algoritmo SVM.

Deve ser destacado que, considerando-se todos os experimentos realizados, a vegetação herbácea-arbustiva obteve os menores índices de acurácia entre todas as classes, ao passo que a gramínea obteve os maiores índices entre as classes de vegetação. Este resultado está relacionado à resposta espectral semelhante entre as classes vegetação herbácea-arbustiva e arbórea, o que causa muita confusão na classificação. Já

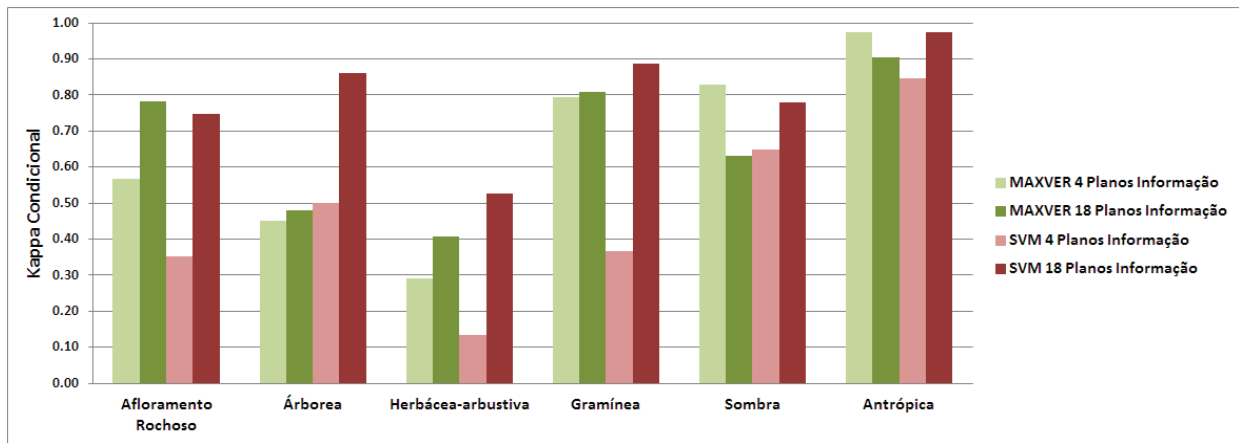


Fig. 4 – *Kappa* condicional do usuário das classes de cobertura da terra.

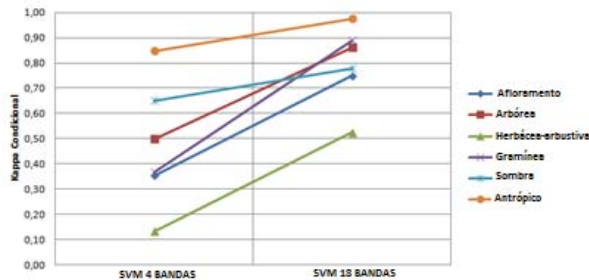


Fig. 5 - *Kappa* condicional do usuário do algoritmo SVM.

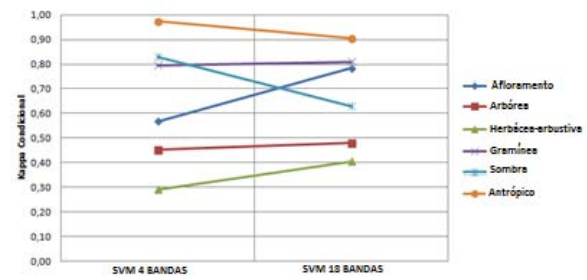


Fig. 6 - *Kappa* condicional do usuário do algoritmo MAXVER.

a gramínea se difere espectralmente das outras classes de vegetação devido à influência do solo na resposta espectral. Resultados semelhantes foram obtidos por Francisco e Almeida (2012).

As classes de não-vegetação apresentaram desempenho maior do que as de vegetação em todos os experimentos, com destaque para a classe área antrópica. Desta forma, em relação a estas classes, pode-se afirmar que a utilização do algoritmo SVM e o incremento no número de atributos não melhoraram significativamente o desempenho da classificação quando comparada com outros experimentos. No caso específico das classes área antrópica e sombra, os desempenhos do classificador MAXVER com 4 bandas e do SVM com 18 bandas foram muito semelhantes. Como a classe área antrópica é composta por vários tipos de cobertura e uso, o melhor desempenho por esta classe em todos os experimentos pode estar relacionado à resposta espectral mais diversificada em relação às demais classes consideradas, o que reduz consideravelmente a confusão.

Como quase todas as bandas que incrementaram os dados de entrada nos experimentos realizados foram produzidas a

partir de operações que envolviam as próprias bandas que compõem a imagem QuickBird original (exceção feita aos planos de informação topográficos), é possível afirmar que o método de classificação não paramétrico SVM apresenta-se como uma alternativa viável ao mapeamento de fisionomias vegetais em imagens de alta resolução. Ainda assim, com o intuito de aperfeiçoar o método e aumentar a eficácia, outros estudos devem ser conduzidos, especialmente aqueles que procurem entender a importância que cada atributo apresenta no resultado final.

#### REFERÊNCIAS BIBLIOGRÁFICAS

ABOUD-NETA, S. R.; DUTRA, L. V.; SCOFIELD, G. B.; NEGRI, R. G.; FREITAS, C. C.; SILVA, D. L. A. Comparação entre classificação contextual e classificação por regiões para mapeamento de uso e cobertura da terra na região da Floresta Nacional de Tapajós - PA (FLONA): utilizando imagens polarimétricas em banda L. *In: Simpósio Brasileiro de Sensoriamento Remoto*, 14, 2009, Natal, Brasil. **Anais...** São José dos Campos: INPE, p. 7749-7756. 2009.

ALIXANDRINI, M. Classificação de imagens

- Landsat 7 ETM multiespectrais utilizando SVM. *In: Simpósio Brasileiro de Sensoriamento Remoto*, 14, 2009, Natal, Brasil. **Anais... São José dos Campos: INPE**, p. 7757-7764. 2009.
- ANDREOLA, R.; HAERTEL, V. Support Vector Machines na Classificação de Imagens Hiperespectrais. *In: Simpósio Brasileiro de Sensoriamento Remoto*, 14, 2009, Natal, Brasil. **Anais... São José dos Campos: INPE**, p. 6757-6764. 2009.
- CENTENO, J. A. S. **Sensoriamento remoto e Processamento de Imagens Orbitais**. Curitiba: Curso de Pós Graduação em Ciências Geodésicas, UFPR. Curitiba, PR. 2009.
- CONGALTON, R.; GREEN, K. **Assessing the Accuracy of Remotely Sensed Data: Principles and Practices**. Boca Raton: CRC/ Taylor & Francis. 2009. 183p.
- COSTA, T. C. C.; JÚNIOR, P. M.; BRITES, R. S.; SOARES, V. P. O Classificador NPVIC Comparado ao Não Paramétrico de Skidmore/ Turner e ao Classificador Maxver. *In: Simpósio Brasileiro de Sensoriamento Remoto*, 9, 1998, Santos, Brasil. **Anais... São José dos Campos: INPE**. p. 945-958.
- DIGITAL GLOBE. **Radiance Conversion of QuickBird Data**. Technical Note. 2003.
- DIXON, B.; CANDADE, N. Multispectral landuse classification using neural networks and support vector machines: one or the other, or both?. **International Journal of Remote Sensing**, v. 29, n. 4, p. 1185–1206, 2008.
- ENVI. **ENVI User's Guide - ENVI Version 4.7**. ITT Visual Information Solutions. 2009.
- FRANCISCO, C. N.; ALMEIDA, C. M. Avaliação de Desempenho de Atributos Estatísticos e Texturais em uma Classificação de Cobertura da Terra Baseada em Objeto. **Boletim de Ciências Geodésicas**, v. 18, n. 2, p. 302-326, 2012.
- INPE. **Tutorial SPRING**. São José dos Campos, SP: INPE. 2002. Disponível em: <<http://www.dpi.inpe.br/spring/portugues/tutorial/index.html>>. Acesso em: 10 de outubro de 2012.
- JENSEN, J.R. **Sensoriamento remoto do ambiente: Uma perspectiva em recursos terrestres**. São José dos Campos, SP: Parêntese Editora. 2009. 598p.
- LANDIS, J.; KOCH, G. The measurement of observer agreement for categorical data. **Biometrics**, v. 33, p. 159 – 174, 1977.
- LICZBINSKI, C. A. **Classificação de Dados Imagem em Alta Dimensionalidade, empregando Amostras Semi-Rotuladas e Estimadores para as Probabilidades a Priori**. 2007. 152p. Dissertação (Mestrado em Sensoriamento Remoto) – UFRGS . Porto Alegre, RS, 2007.
- LIU, W. T. H. **Aplicações de Sensoriamento Remoto**. Campo Grande, MS: UNIDERP. 2006.
- LORENA, A. C.; CARVALHO, A. C. P. L. F. Uma Introdução às Support Vectors Machines - Tutorial. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43-67, 2007.
- MOREIRA, M. A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**. Viçosa, MG: UFV. 2005. 307p.
- NASCIMENTO, R. F. F.; ALCÂNTARA, E. H.; KAMPEL, M.; STECH, J. L.; NOVO, E. M. L. M.; FONSECA, L. M. G. O Algoritmo Support Vector Machines (SVM): Avaliação da Separação Ótima de Classes em Imagens CCD-CBERS-2. *In: Simpósio Brasileiro de Sensoriamento Remoto*, 14, 2009, Natal, Brasil. **Anais... São José dos Campos: INPE**. p. 2079-2086.2009
- NOVO, E. M. L. M. **Sensoriamento remoto: Princípios e aplicações**. São Paulo, SP: Edgard Blücher, 2011. 388 p.
- PONZONI, F. J.; SHIMABUKURO, Y. E. **Sensoriamento remoto no estudo da vegetação**. São José dos Campos, SP: Parêntese. 2009. 127p.
- QUEIROZ, R. B. ; RODRIGUES, A. G. ; GÓMEZ, A. T. . Estudo Comparativo entre as Técnicas Máxima Verossimilhança Gaussiana e Redes Neurais na Classificação de Imagens IR-MSS CBERS 1. *In: I WorkComp Sul*, 1, 2004, Palhoça, SC, Brasil. **Anais...** Disponível em: <http://inf.unisul.br/~ines/workcomp/cd/pdfs/2307.pdf>. Acesso em: 05/08/2012.
- SEMOLINI, R. **Support Vector Machines**,

**Inferência Transdutiva e o Problema de Classificação.** Dissertação (Mestrado em Engenharia de Computação) - UNICAMP. Campinas, SP, 2002.

TOUTIN, T. Geometric Processing of Remote Sensing Images: Models, Algorithms and Methods. **International Journal of Remote**

**Sensing**, v. 25, n. 10, p. 1893–1924, 2004.

VALERIANO, M. M. Modelo digital de Variáveis Morfométricas com dados SRTM para o Território Nacional: o Projeto TOPODATA. *In: Simpósio Brasileiro de Sensoriamento Remoto*, 12, 2005, Goiânia, Brasil. **Anais... São José dos Campos: INPE**. p. 1-8. 2005