

## Desempenho e acurácia dos SIGs Terra View e Idrisi e seus respectivos classificadores supervisionados

Mikael Timóteo Rodrigues<sup>1</sup>  
Bruno Timóteo Rodrigues<sup>1</sup>  
Felipe de Souza Nogueira Tagliarini<sup>1</sup>  
Lincoln Gehring Cardoso<sup>1</sup>  
Sérgio Campos<sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP  
Faculdade de Ciências Agrônomicas – FCA, Campus de Botucatu, Departamento de  
Engenharia Rural  
Rua José Barbosa de Barros, nº 1780  
18.610-307 - Botucatu - SP, Brasil  
mikaelgeo@gmail.com; brunogta21@gmail.com; felipe\_tagliarini@hotmail.com;  
cardosolg@fca.unesp.br; seca@fca.unesp.br

**Abstract.** The main objective of this study is to investigate the performance of TerraView 4.2.2 and Idrisi Selva performing classification oversees through the spectral pattern on Landsat 5, associated with comparing the land use of the river Capivara watershed, inserted in the municipality of Botucatu, São Paulo, Brazil. The areas of supervised training were defined through seven land use classes, founded by the Manual Use of Technical IBGE Earth. In the region of Capivara watershed, they are practiced multiple types of management, which can be found planting crops from subsistence scale, through small and medium-sized farms, to major agro-industrial structures, thus providing a panorama of great complexity to mapped and subsequently patterned. An aggravating the methodology were the weed common in cultivated pastures and soils with various forms of culture, because they cause interference in the spectral pattern of land use classes, thus providing "noise" that changed the pure spectral response of crops inducing error digital classification. Post-classification also improved matrix realignment estimates for removal of pixel groups, reaching a higher order than 50% accuracy, increasing accuracy, allowing a lower inclusion of items of other classes, thus making it the best classification. Unlike the products derived from the supervised classification by maximum likelihood post classified with the majority filter, which after reclassification accuracy was high, presented fewer errors, as well as smoothing of classified maps.

**Key-words:** Digital classification, Remote Sensing, Land Truth, Thematic Mapping.

### 1. Introdução

O Sensoriamento Remoto tem papel significativo no ambiente SIG (Sistemas de Informação Geográfica), como uma das alternativas mais viáveis para extração de informações sobre os diversos usos do solo, assim, gerando grande número de informações e possibilitando programação de ações adequadas para tomada de decisões por constituir ferramentas de articulação e informações geoespaciais (FUSHITA, et al, 2013). A partir do conhecimento das características físicas dos alvos de uma dada região, e seu comportamento espectral (interação da radiação solar com os elementos que compõe o mosaico da superfície terrestre como os rios, a vegetação, o tipo de solo, a rocha) é possível fazer um monitoramento das condições ambientais da área em questão e descobrir possíveis modificações (AMORIM, et al., 2007).

Múltiplos trabalhos utilizando diversas ferramentas SIG têm sido desenvolvidos com a intenção de delimitar e classificar bacias hidrográficas e os respectivos manejos do uso da

terra nas mais variáveis áreas. Com base nisso, foram elaboradas as técnicas de classificação digital de imagens, que automatizam a metodologia de extração das imagens de satélite, minimizando a subjetividade da interpretação humana, bem como, reduzindo esforços e encurtando o tempo de trabalho do analista.

A forma mais usual e difundida entre os pesquisadores que trabalham com Sensoriamento Remoto e que foi adotada pelos softwares comerciais e livres que utilizam módulos de processamento de imagens, dividi os classificadores em supervisionados e não-supervisionados. A classificação supervisionada, que foi empregada no presente trabalho requer conhecimentos prévios das classes de alvos, a fim de classificar a imagem nos alvos de interesse pré-fixados pelo analista, onde o algoritmo precisa ser treinado para poder diferenciar as classes uma das outras. O presente trabalho teve por objetivo averiguar o comportamento de dois softwares SIG (TerraView 4.2.2 e Idrisi versão Selva) realizando a classificação supervisionada para mapear usos na bacia hidrográfica do rio Capivara.

## 2. Metodologia de Trabalho

A área de estudo abrange a bacia hidrográfica do rio Capivara, situado no Município de Botucatu, região Centro-Oeste do Estado de São Paulo entre as coordenadas geográficas 22° 47' 10" a 22° 52' 38" de latitude S e 48° 22' 07" a 48° 26' 38" de longitude W e Coordenadas UTM variando de 761500 a 778000 metros a Leste e 7458000 a 7484000 metros a Sul dentro do Fuso 22 e Zona K (Figura 1).

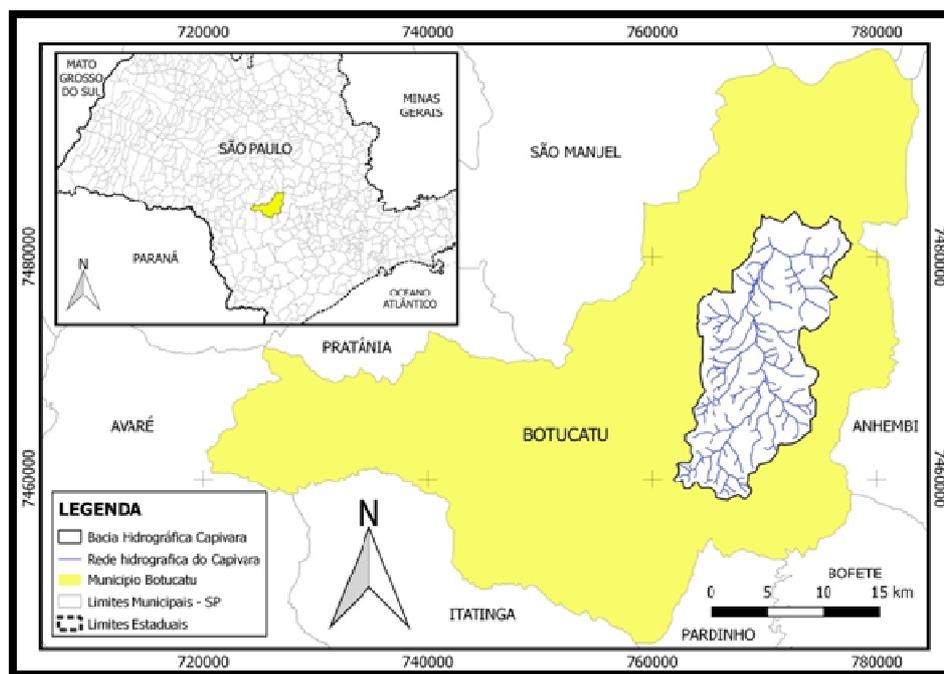


Figura 1: Localização da bacia rio Capivara, Botucatu-SP.

Para realização do georreferenciamento da área de interesse, foi obtido pelos softwares Idrisi versão Selva e o ArcGIS 10.1, utilizando como pontos de controle a base cartográfica da carta planialtimétrica da folha Botucatu (SF-22-R-IV-3), escala 1:50000 e equidistância vertical de 20m situada no fuso 22, Datum Córrego Alegre.

Para a classificação supervisionada por máxima verosimilhança também conhecida como "MAXVER", foram utilizados dados coletados pelo sensor Thematic Mapper de uma imagem do satélite Landsat-5, atuando com sete bandas espectrais, de forma que no presente trabalho, será utilizado três, com a composição RGB, composto pelas das bandas 3, 4 e 5 (TM3, TM4 e

TM5) com escala 1:50000, na órbita 220, ponto 76 inserida no mosaico de recobrimento sobre o estado de São Paulo, com resolução de 30 X 30 metros, com data de passagem em 19 de setembro de 2011.

A imagem orbital foi registrada no software ArcGIS 10.1 com auxílio da carta planialtimétrica da folha Botucatu SF-22-R-IV-3, onde a mesma foi exportada para o formato GEOTIFF.

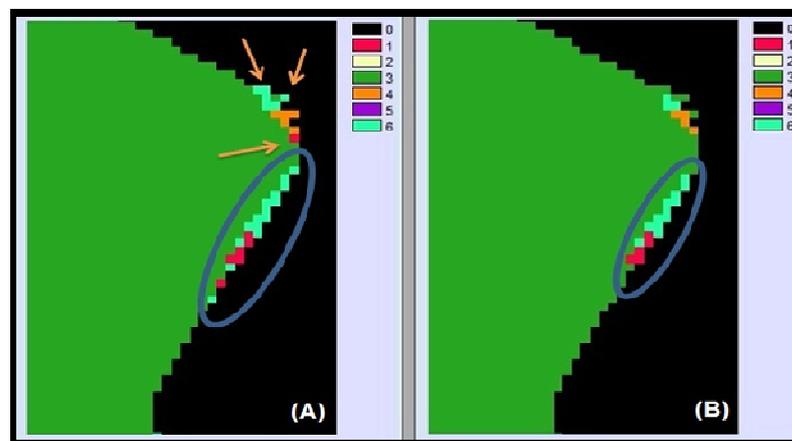
A partir da imagem de satélite, elaborou-se um mapa da verdade terrestre do uso e ocupação da terra na bacia hidrográfica dos rio Capivara, fazendo-se uso do software QGIS 2.2.0. Para auxiliar na interpretação visual das áreas de cobertura vegetal, foi utilizado o Google Earth como ferramenta de apoio, permitindo delimitar com maiores detalhes as ocupações, onde foram identificadas 7 classes para bacia do rio Capivara no ano de 2011 por meio da imagem derivada do sensor TM a bordo do Landsat 5. De posse do arquivo completo vetorizado, o mesmo foi exportado em formato GEOTIFF para o Idrisi Selva edição 17.0 (EASTMAN, 2012).

As áreas de treinamento supervisionado foram definidas por meio de polígonos que representaram as respectivas classes de uso e ocupação da terra, considerando a cor, brilho e textura emitida por cada pixel da imagem. Assim, cada polígono supervisionado, teve um conjunto de pixel com o mesmo contexto, iguais feições e padrões, sendo consolidado o mesmo objeto irradiante.

Área de treinamento (*training sites*) é uma amostra homogênea da classe com toda a variabilidade dos níveis de cinza. É recomendado que o usuário adquira mais de uma área de treinamento (CROSTA, 1992), utilizando o maior número de informações disponíveis, como trabalhos de campo, mapas, dentre outras informações. O número de "pixels" de treinamento de uma classe aumenta com a complexidade da área a ser classificada.

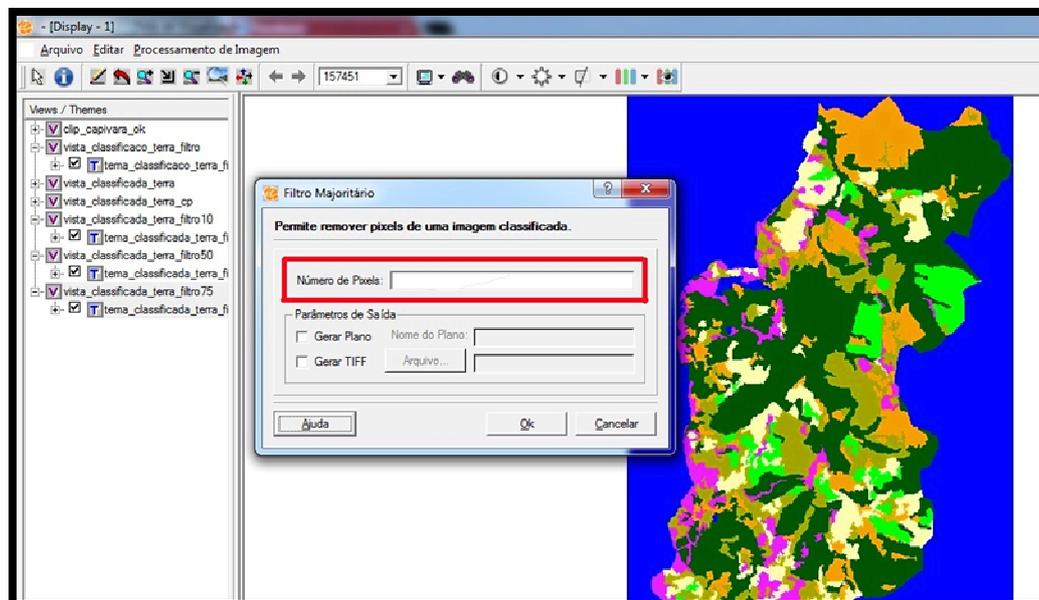
Assim sendo, a classificação supervisionada MAXVER (máxima verossimilhança) foi realizada por meio dos softwares TerraView e Idrisi, onde foram definidas amostras de treinamento, tomando como base o Manual Técnico de Uso da Terra do IBGE (2006), por meio de polígonos representados nos dois aplicativos.

Também foi utilizado no TerraView, como forma de pós-classificação, a extensão Filtro Majoritário para a "limpeza" de pixels que possivelmente tenham sido classificados de forma equivocada pelo algoritmo do software (Figura 2), aparecendo como pequenos pontos dentro de classes maiores, uniformizando as classes de uso e eliminando pontos isolados, classificados diferentemente de grupos vizinhos. Contudo, é gerada uma segunda classificação do mesmo SIG (TerraView) com menos ruídos, por sua vez, com menor erro agregado.



**Figura 2.** Classificação supervisionada sem filtro (A); Pós-classificação com uso Filtro Majoritário (B).

O software livre Terraview possui em sua extensão de Filtro Majoritário a flexibilidade da escolha referente à quantidade de pixels que o usuário pode eliminar por classes em sua pós-classificação (Figura 3), ou seja, reduzir os ruídos de pixels classificados em classes não pertencentes ao mesmo. Com essa opção, o usuário pode fazer inúmeras tentativas de pós-classificação até chegar a uma acurácia próxima da desejada.



**Figura 3.** Extensão do filtro majoritário no aplicativo TerraView e sua opção de escolher a quantidade de pixel que o usuário pode eliminar por classes.

Foram elaborados vários testes de pós-classificação com números de exclusão de pixels baixo, moderado e altos, porém, a acurácia tende a cair quando se eleva bastante o número de exclusão de ruídos, pois com altos grupos de pixels migrados de uma classe para outra, o algoritmo tende mover não só os pixels que causam ruídos, mas também pixel que já se encontra em sua classe verdadeira, assim, fazendo o sentido inverso do que seria uma satisfatória pós-classificação.

Para identificar a melhor classificação, foram cruzados os mapas de MAXVER derivados do TerraView com a verdade terrestre, onde a mesma se caracteriza como a real utilização do solo, apontando a exatidão (acurácia) de cada classificação, cruzando matrizes de pixel ou conjuntos de pixels. O índice Kappa (Tabela 1) é o mais recomendado por utilizar todas as células da matriz ao invés de somente os elementos diagonais, diferentemente de métodos que fazem uso apenas da Exatidão Global, utilizando dados derivados dos cruzamentos das classes, onde os mesmos se situam na diagonal principal da matriz de confusão, o que garante ao índice Kappa uma maior precisão da acurácia em relação aos outros métodos devido à medição de um pixel estar corretamente classificado, em relação à probabilidade de estar incorretamente classificado, avaliando a acurácia temática por ser mais sensível as variações de erros de omissão e comissão, também conhecido na literatura como erro de inclusão. (DEMARCHI et. al., 2011).

Desta forma, o erro de omissão representa os pixels “perdidos” para outras classes, bem como o erro de comissão reflete os pixels “recebidos” de outras classes. Na matriz de erro, o erro de omissão indica a probabilidade de uma unidade amostral de referência estar corretamente classificada e o erro de inclusão representa a probabilidade que um pixel classificado no mapa, represente a categoria no terreno.

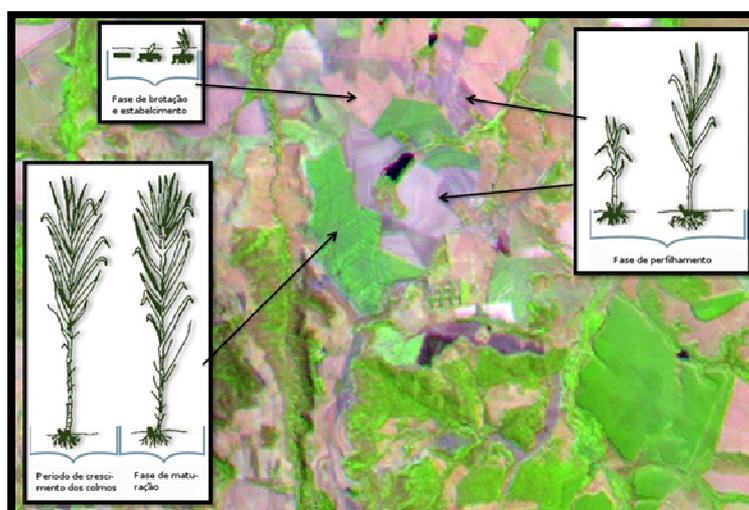
**Tabela 1.** Intervalos de caracterização da acurácia em relação a verdade terrestre.

Valor de Kappa	Qualidade da Classificação
<0,00	Péssima
0,00 – 0,20	Ruim
0,20 – 0,40	Razoável
0,40 – 0,60	Boa
0,60 – 0,80	Muito boa
0,80 – 1,00	Excelente

Fonte: Landis e Koch (1977).

### 3. Resultados e Discussão

Em decorrência da variação temporal das culturas encontradas na área estudada, é que durante a o processo de escolha de amostras de treinamento (*training sites*) e a elaboração da verdade terrestre, o usuário tem que determinar várias subclasses da mesma cultura, como elaborado por Rodrigues et al. (2014), ou avaliar todas as amostras em uma mesma classe, onde desta forma irá ocorrer um grande desvio padrão e por sua vez uma confusão entre classes e diminuição significativa da acurácia. Na imagem de composição RGB (Figura 4) podemos observar essa variação fenológica, da cana-de-açúcar, assim, subdividida em três subclasses.



**Figura 4.** Imagem de composição RGB e subclasses da cana-de-açúcar com diferentes respostas espectrais.

Na região da bacia do Capivara são praticados múltiplos tipos de manejo, onde pode ser encontrado plantios de culturas desde a escala de subsistência, passando por pequenas e médias propriedades rurais, até maiores estruturas agroindustriais, assim, constituindo um panorama de grande complexidade a ser mapeado e posteriormente modelado, onde a lista de ocupação e uso das terras é muito diversificado.

Um agravante à metodologia são as plantas daninhas, comuns em pastagens e solos cultivados com diversas formas de cultura, pois elas ocasionam interferência no padrão espectral das referidas classes, assim, proporcionando “ruído” que altera a resposta espectral pura das culturas induzindo a erros de classificação digital.

Outra dificuldade para o mapeamento da área estudada, fazendo-se uso de imagens de orbitais, é a ambigüidade espectral entre classes de mesma cultura. Tais erros podem ser

inferidos pela apreciação de índices de exatidão específicos para cada classe da modelagem, para tanto, erros de omissão e comissão.

A verdade terrestre referente à bacia do rio Capivara (Figura 5 A) foi obtida como resultado da real utilização do uso do solo na área de estudo, assim, realizado o índice para comparação com as classificações MAXVER provenientes dos softwares SIG TerraView e Idrisi.

A classificação supervisionada (Figura 5 B) realizada no Idrisi apresentou bom resultado com relação ao índice Kappa, atingindo a ordem de 0.47, o que segundo Landis e Koch (1977) indica uma classificação de boa qualidade (0,40 a 0,60).

O melhor nível de conservação da classe “solo em cultivo”, em função de cuidados aderidos pelos respectivos agricultores utilizando insumos e outras formas adequadas de manejo do solo, apresentou um melhor padrão vegetal e conseqüentemente espectral, teve uma menor porcentagem de pixels errados em relação à classe “pastagem” que comumente se encontram pastos abandonados e cortados por estradas de solo exposto ou cobertos por vegetação daninha, proporcionando uma propriedade irregular referente a padrão, textura e espectro eletromagnético, o que ocasiona confusão em pontos dessa classe, principalmente com a área urbana que possui cor rosa escuro, textura grossa.

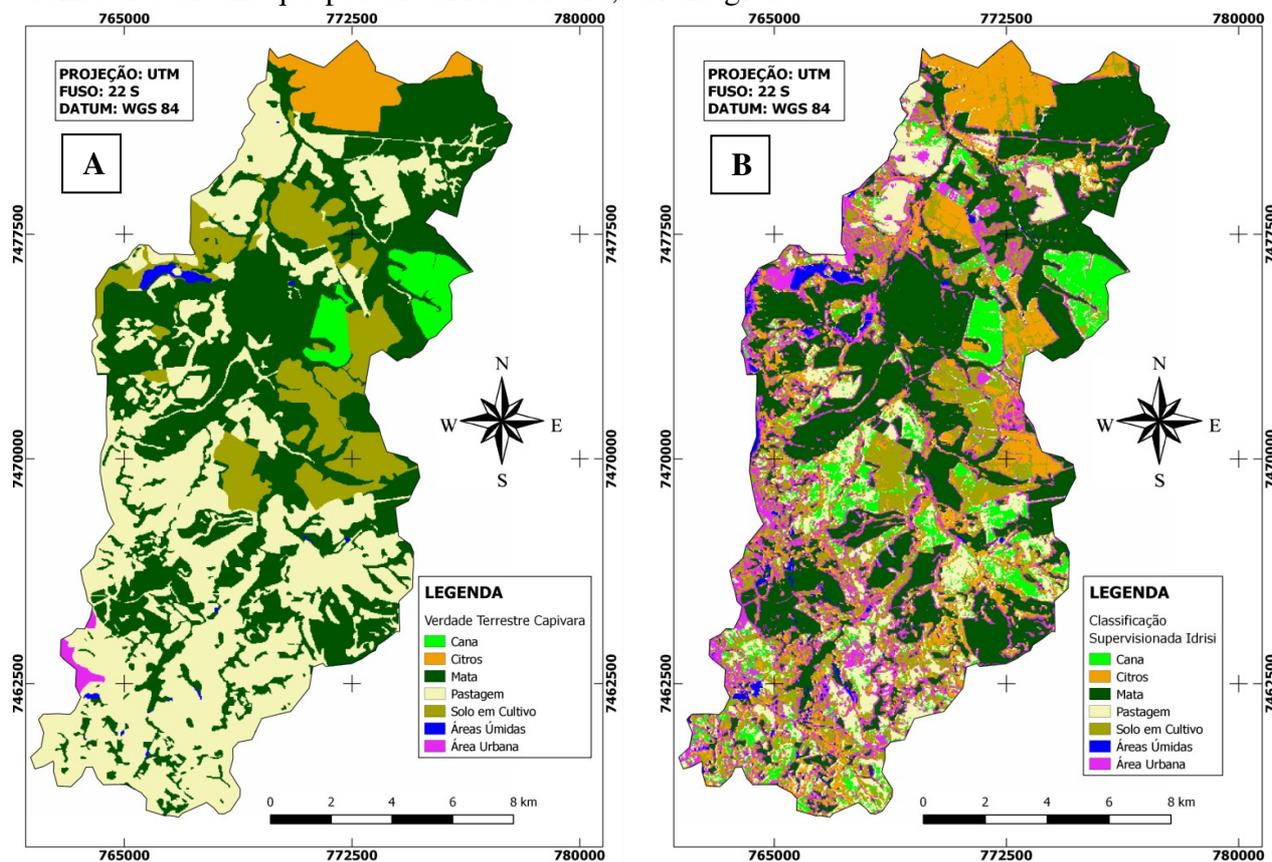
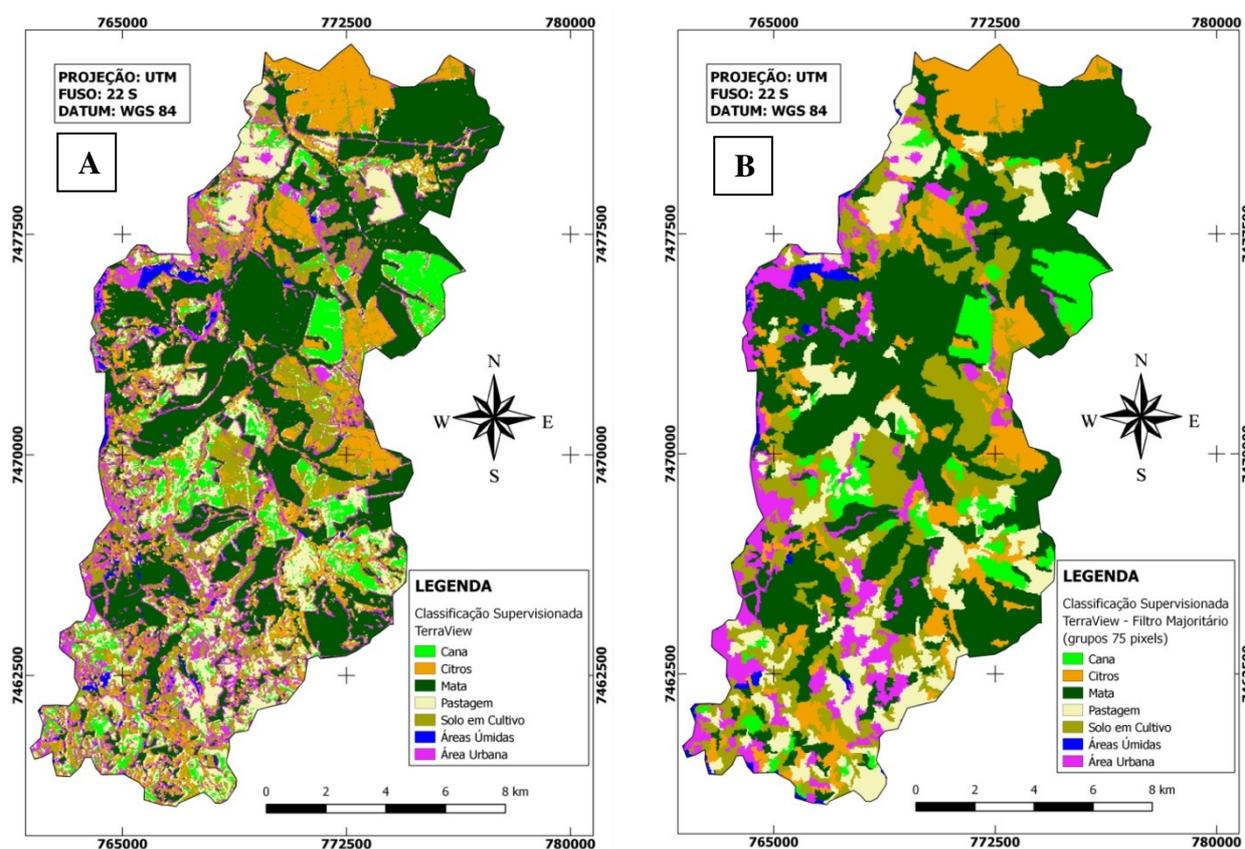


Figura 5. Verdade terrestre da bacia do rio Capivara (A); Classificação supervisionada IDRISI Selva (B).

Nas classificações processadas pelo TerraView, foram observadas que as classes com maiores erros de comissão foram solo em cultivo, área urbana e a pastagem. Alguns maiores erros de omissão foram observados para a cana, a mata e novamente a pastagem, por apresentarem as maiores classes área dentro da bacia, conseqüentemente, com vários pontos de ambiguidade espectral confundindo o processamento do algoritmo, o que deu origem a uma classificação (Figura 6A) com um valor Kappa na ordem de 0,476, sendo considerada uma boa modelagem segundo os intervalos de acurácia da Tabela 1.

Além do fator ambigüidade espectral, a modelagem também é afetada pelo relevo acidentado, pois no sensoriamento remoto existe a interação macroscópica (relação da radiação eletromagnética com os objetos), assim, a imagem orbital utilizada no presente trabalho é detentora também da variável textura, que representa a rugosidade topográfica da superfície, podendo ter suas estimativas relativamente transformadas devido à variação vertical de altura da superfície nos vales das bacias, que por sua vez influencia nos tipos de dosséis da vegetação.

De modo geral, com todas as variáveis que podem causar tendência ao erro, pode-se dizer que os resultados da modelagem derivada da pós-classificação do TerraView fazendo-se uso do filtro majoritário foi mais coerente, atingindo um coeficiente Kappa na ordem de 0,515, alcançando mais de 50% de acerto referente a verdade terrestre. Esta função provocou um novo delineamento e suavização das bordas de todas as classes de uso, tornando o mapa aparentemente mais agradável, com um visual livre de ruídos (Figura 6B).



**Figura 6.** Classificação Supervisionada TerraView (A); Pós-classificação supervisionada TerraView com o uso do filtro majoritário (B).

A reclassificação fazendo-se uso no filtro majoritário excluiu dos grupos de pixels classificados erroneamente, passando a existir menos erros de omissão, bem como menos erros de comissão (que determinam grupos pixels ou pixel isolado que são atribuídos a uma classe qual não pertence), provocando um novo delineamento e suavização das bordas de todas as classes de uso, eliminando os ruídos e abrandando as classes, desta forma, tornando o mapa final aparentemente mais confortável ao visualizá-lo. Contudo, grandes grupos de classes classificados pelo algoritmo de forma equivocada não serão eliminados ou realocados pelo filtro, pois este apenas elimina grupos ruidosos.

#### 4. Conclusões

Devido os mais variados estádios de desenvolvimento e produção da cana-de-açúcar e de outras culturas presentes na bacia, bem como grandes áreas de solo preparado para o manejo, onde ora emitem diversas respostas espectrais de um mesmo uso, ora classes distintas apresentaram assinaturas espectrais idênticas ou muito próximas, caracterizou um alto grau de confusão na interpretação do algoritmo, influenciando diretamente na acurácia das classificações. Ficou presente à diferença do resultado da classificação supervisionada por pós-classificados com o filtro majoritário do SIG TerraView, onde após a reclassificação a acurácia foi elevada, apresentado menos erros, bem como suavização dos mapas classificados.

É importante salientar que, provavelmente, se incorporados subdivisões espectrais de usos que apresentam tempos fenológicos distintos, bem como imagens com melhor resolução espacial, aplicação dos métodos em uma área com mais homogeneidade e supressão de aspectos que provocam dispersão de ondas eletromagnéticas, os resultados de acurácia para o TerraView e o Idrisi, serão possivelmente mais elevados e confiáveis, pois tais ferramentas apresentam um alto poder de processamento de dados digitais.

#### Referências

Amorim, R. F.; Almeida, S. A. S.; Cuellar, M. Z.; Costa, A. M. B.; Gomes, C. Mapeamento de Uso e Ocupação do solo na Bacia Hidrográfica Piranhas/Açu, utilizando imagens CBERS e técnicas de classificação supervisionada. XIII Simpósio Brasileiro de Sensoriamento Remoto - SBSR.. 2007, Florianópolis, SC. **Anais...** São José dos Campos: INPE, 2007, p. 3710.

Crosta, A. P. **Processamento Digital de Imagens de Sensoriamento Remoto**. IG/UNICAMP, Campinas – SP, v 1, p 193, 1992.

Demarchi, J. C.; Piroli, E. L.; Zimback, C. R. L.. Análise temporal do uso do solo e comparação entre os índices de vegetação NDVI e SAVI no município de Santa Cruz do Rio Pardo - SP usando imagens Landsat-5. **Ra'e ga** (UFPR), v. 21, p. 234-271, 2011.

Eastman, J. R. **IDRISI Selva** – GIS and Image Processing Software – version 17.0. Worcester-MA/USA: Clark Labs, 2012.

Fushita, A. T.; Reis, R. R.; Faresin, L.; Santos, J. E. Desempenho da classificação supervisionada em diferentes programas: comparação por meio do uso da terra e do índice de naturalidade da paisagem. XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR. 2013, Foz do Iguaçu, PR. **Anais...** São José dos Campos: INPE, 2013, p. 6465.

Landis, R.; Koch, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v.33, n.1, p.159-174, Mar. 1977.

IBGE. Instituto Brasileiro de Geografia e Estatística. **Carta topográfica**: folha de Botucatu (SF-22-R-IV-3). Serviço gráfico do IBGE, 1969. Escala 1:50.000.

IBGE. Instituto Brasileiro de Geografia e Estatística. **Manual Técnico de Uso da Terra**. 2ª ed., Rio de Janeiro: IBGE, 2006.

INPE. Instituto Nacional de Pesquisas Espaciais - Divisão de Geração de Imagens (DGI-INPE). **Banco de Dados Geográficos**. Disponível em:<<http://www.dgi.inpe.br/>>. Acesso em: 08 jul. 2015.

INPE. Instituto Nacional de Pesquisas Espaciais - Divisão de Geração de Imagens (DGI-INPE). **Catálogo de Imagens**. Disponível em:<<http://www.dgi.inpe.br/>>. Acesso em: 15 abr. 2013.

Rodrigues, M. T.; Rodrigues, B. T.; Tagliarini, F. S. N. Comparação do desempenho de Sistemas de Informação Geográfica (IDRISI Selva e ArcGIS®) por meio de processamento digital de imagem. **Revista Fórum Ambiental da Alta Paulista**, Tupã, v. 10, n. 2, p. 265-280, 2014.