

COMPORTAMENTO DE SOFTWARES SIG NA CLASSIFICAÇÃO SUPERVISIONADA

Mikael Timóteo Rodrigues¹, Bruno Timóteo Rodrigues¹, Lincoln Gehring cardoso¹, Sérgio campos¹
Felipe de Souza Nogueira Tagliarini¹, Mateus de Campos Leme¹, Flávia Luíze Pereira de Souza¹

¹Faculdade de Ciências Agrônômicas – FCA/UNESP, Avenida Universitária, nº 3780 – CEP 18610-034, Altos do Paraíso, Botucatu-SP, mikaelgeo@gmail.com; brunogta21@hotmail.com; cardosolg@fca.unesp.br; felipe_tagliarini@hotmail.com; seca@fca.unesp.br; mateus.leme93@gmail.com; flavialuizesouza@hotmail.com.

RESUMO

O objetivo desse trabalho foi averiguar a atuação dos aplicativos de Sistemas de Informações Geográficas (SIG) TerraView 4.2.2 e IDRISI Selva desempenhando a classificação supervisionada por meio de padrão espectral em imagem Landsat 8, associada a comparação do uso da terra da bacia hidrográfica do rio Lavapés, inseridas no município de Botucatu/SP utilizando-se técnicas de sensoriamento remoto e geoprocessamento. As áreas de treinamento supervisionado foram definidas a partir de nove classes para bacia do Lavapés, fundamentais para o estudo e análise do uso e ocupação da terra, como mata, solo, culturas - agricultura, corpos d'água e malha urbana dentre outras classes encontradas. Foi evidente a diferença dos produtos derivados a partir da classificação gerada e posteriormente pós-classificados com o filtro majoritário (majority filter) do aplicativo TerraView, onde após a reclassificação a acurácia foi elevada, apresentado menos erros (em relação ao IDRISI e ao TerraView sem a utilização do filtro) de omissão e comissão nas matrizes e suavização dos mapas classificados, com a eliminação de classes de 10 e 75 pixels por região, o que abrandou consideravelmente a estética dos mapas e consequentemente a diminuição de erros.

Palavras-chave — Geoprocessamento, Sensoriamento Remoto, Processamento de Imagens, Uso do solo.

ABSTRACT

The objective of this work was to verify the performance of the Geographic Information Systems (GIS) applications TerraView 4.2.2 and IDRISI Selva performing the supervised classification using a spectral standard in Landsat 8 image, associated with a comparison of the land use of the hydrographic basin of the Lavapés watershed, inserted in the city of Botucatu/SP using techniques of remote sensing and geoprocessing. The supervised training areas were defined from nine classes for the Lavapés watershed, which are fundamental for the study and analysis of land use and

occupation, such as forest, soil, crops - agriculture, water bodies and urban fabric among other classes found. It was evident to the difference of the derived products from the generated classification and later post-classified with the majority filter of TerraView application, where after the reclassification the accuracy was high, presented less errors (in relation to IDRISI and TerraView without the use of the filter) of omission and commission in the matrices and smoothing of the classified maps, with the elimination of classes of 10 and 75 pixels per region, which considerably softened the aesthetics of the maps and consequently the reduction of errors.

Key words — Geoprocessing, Remote Sensing, Image Processing, Use of the soil.

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 [1]. 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 [2]. A forma mais usual e difundida entre os pesquisadores que trabalham com sensoriamento remoto para classificar grandes áreas é a classificação supervisionada. Por sua vez, essa técnica de classificação é adotada pelos softwares SIG comerciais e livres.

Tendo em vista a larga utilização da classificação supervisionada, bem como dúvidas ligadas a eficiência dos softwares e seus filtros majoritários, o presente trabalho teve como objetivo analisar desempenho dos aplicativos SIG TerraView 4.2.2 e IDRISI Selva bastante utilizado no Brasil e em diversas partes do mundo para mapear grandes áreas que apresentam padrão espectral associado aos diversos usos da

terra, por meio do e a influência dos respectivos filtros no produto final da classificação

2. MATERIAIS E MÉTODOS

A área de estudo foi delimitada pela área que abrange a bacia hidrográfica do rio Lavapés, situada 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. Para a classificação supervisionada por máxima verossimilhança também conhecida como "MAXVER", foi utilizada uma imagem do satélite LANDSAT-8, atuando com sete bandas espectrais, de forma que no presente trabalho, será utilizado três, com a composição RGB, 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. A partir da imagem de satélite, elaborou-se um mapa da verdade terrestre do uso e ocupação das terras na bacia hidrográfica do rio Lavapés. Para auxiliar na interpretação visual das áreas de cobertura vegetal, onde foram identificadas 9 classes de uso e cobertura para bacia do rio Lavapés.

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. Também foi utilizado no TerraView, como forma de pós-classificação, a extensão Filtro Majoritário (majority filter) para a "limpeza" de pixels que possivelmente tenham sido classificados de forma equivocada pelo algoritmo do software, 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.

Para identificar a melhor classificação, foram cruzados os mapas de MAXVER derivados do TerraView e IDRISI 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)[3] 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 [4].

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 [5], 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 1) pode-se observar essa variação fenológica, da cana de açúcar, assim, subdividida em três subclasses.

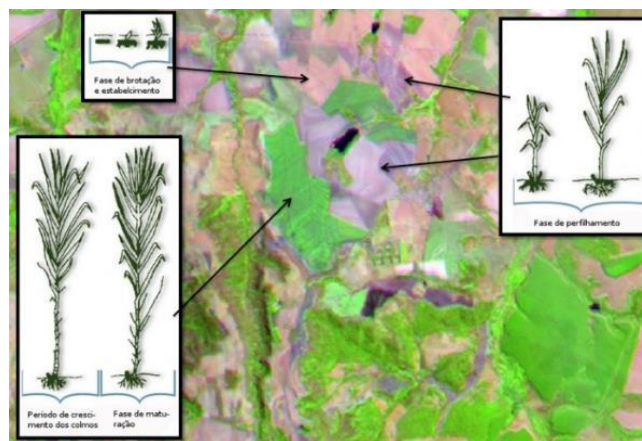


Figura 1. Imagem de composição RGB e subclasses da cana de açúcar com diferentes respostas espectrais.

Na região da bacia do Lavapés 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

difículdade para o mapeamento da região estudada, fazendo-se uso de imagens de orbitais é a ambiguidade 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 Lavapés (Figura 2A) 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 do TerraView e IDRISI.

A classificação supervisionada (Figura 2B) através do método máxima verossimilhança realizada no IDRISI Selva por meio da extensão MAXLIKE, apresentou resultado razoável com relação ao índice Kappa, com valor na ordem de 0.37. Tendo em vista as características de uso na bacia estudada e seu conflito espectral, o resultado encontrado foi positivo (pois existe um aceitável grau de acurácia nos dados classificados, onde cada uso teve maiores valores em suas respectivas classes), pois a imagem orbital da bacia analisada apresenta grande área com o cultivo da cana nos mais variados estágios de desenvolvimento e produção, bem como grandes áreas de pastagem e solo preparado para o manejo desta cultura, assim, emitindo diversas respostas espectrais de um mesmo uso, caracterizando um alto grau de confusão para o algoritmo na separação das classes.

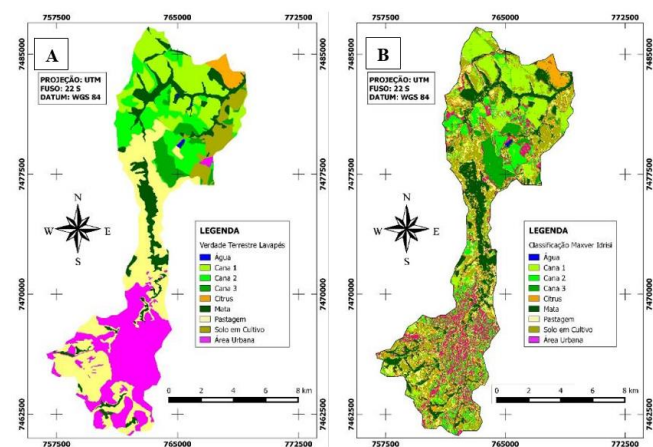


Figura 2. (A) Verdade terrestre da bacia do rio Lavapés. (B) Classificação supervisionada IDRISI Selva.

Tal acontecimento referente ao cultivo da cana nos mais variados estágios de desenvolvimento e produção (classe cana colhida no ano, classe cana de ano e meio), bem como grandes áreas de solo preparado emitindo diversas respostas espectrais de um mesmo uso causando “ruídos” na classificação supervisionada. A classificação supervisionada derivada do software TerraView (Figura 3) por meio do aplicativo MAXVER (modelagem do algoritmo da máxima verossimilhança), que considera a ponderação das distâncias entre as médias dos níveis do conjunto de pixels das amostras por classes, apresentou resultado razoável com relação ao índice Kappa, com valor na ordem de 0.21.

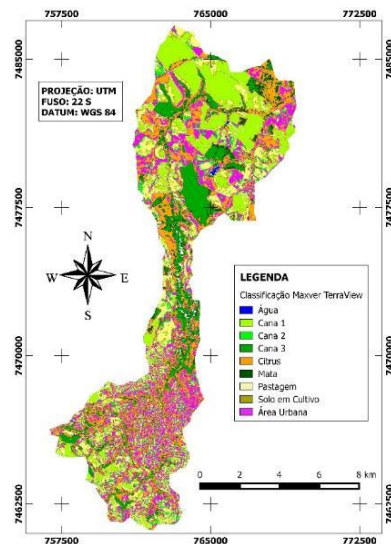


Figura 3. Classificação supervisionada TerraView.

O TerraView possui em sua extensão de Filtro Majoritário a flexibilidade da escolha referente à quantidade de pixels que pode ser eliminado por classes em sua pós-classificação, onde o usuário estabelece a quantidade de pixel por bloco. Por tal ferramenta apresentar esta mobilidade de inúmeras tentativas a pós-classificação com seu Filtro Majoritário, foram modeladas e consideradas para o presente resultado, as pós-classificações com retirada de grupos 10 pixels e grupos com 75 para redução de ruídos provocados por pixels fora de suas respectivas classes. O grupo de 10 pixels foi considerado por ser valor mínimo significativo de uma pós-classificação com o TerraView. Todavia, o grupo de 75 pixels foi considerado por apresentar o valor majoritário máximo referente a grupos de pixels removidos que causam suavização e melhor acurácia.

Os testes feitos com valores maiores que 75 (90, 150, 200 e 250) ocasionaram uma maior homogeneização das classes, porém uma acurácia menor. Isso ocorreu devido a um grande deslocamento de pixels, onde o algoritmo desloca os grupos de pixels ruidosos (que estavam classificados incorretamente), mas também os grupos classificados em suas classes corretas, desta forma a pós-classificação com valores acima de 75 pixels suaviza o mapa e modifica classes corretas. Assim sendo, a modelagem da pós-classificação com retirada de grupos 10 pixels por região reduziram de forma bastante significativa os ruídos de pixels isolados elevando a acurácia das primeiras classificações do TerraView, saindo de 0,21 (onde faziam parte do intervalo referente a “razoável” segundo Tabela 1, elevando para ordem de 0.45 (Figura 4A), onde o mapa foi suavizado e suas estimativas recalculadas, elevando a acurácia de classificação para o intervalo de “boa” segundo a escala Kappa. As classes mata, solo em cultivo, cana 2 e pastagem que vinham apresentando erros de omissão nas outras matrizes, apresentaram mais acertos nas suas respectivas classes, porém, a água passou a ter mais pixels confundidos com a mata. Com o uso do filtro (grupo de 10 pixels) os valores

globais em diagonal foram remodelados, assim, elevando significativamente a acurácia do mapeamento. Melhorando ainda mais a suavização e as estimativas de realinhamento de matriz e acurácia (Figura 4B), a modelagem da pós-classificação com retirada de grupos 75 pixels por região atingiu a ordem de 0,523, observando-se um grande aumento na acurácia permitindo uma menor inclusão de pontos de outras classes, assim, tornando-se a melhor classificação supervisionada da bacia do Lavapés, com mais de 50% de acerto. Pode ser notado de forma clara, que o desempenho do classificador com a retirada de grupos 75 pixels por região foi bastante semelhante à retirada do grupo de 10, uma vez que se passou 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 do mapa final ao qual não pertence, enquanto deveriam estar em sua classe de origem (outra classe diferente da que foi classificada). A classe de água continuou sendo interferida pela mata, porém, com uma pequena melhoria no valor global. O citrus apresentou melhor ajuste ao filtro 75, confundindo apenas 74 pixels com a classe mata.

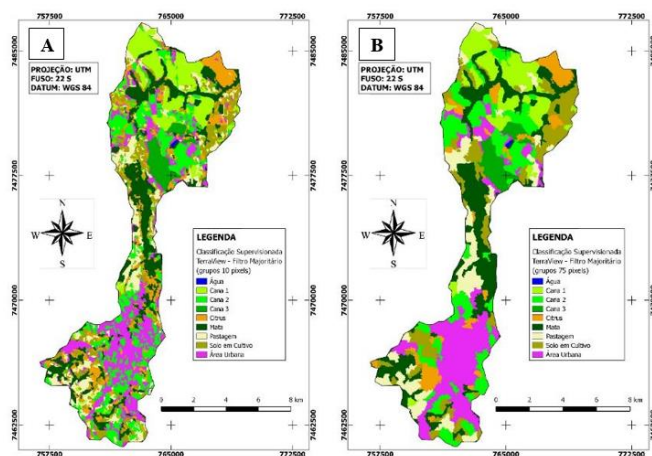


Figura 4. (A) Classificação Supervisionada TerraView com Filtro Majoritário (grupos 10 pixels). (B) Classificação Supervisionada TerraView com Filtro Majoritário (grupos 75 pixels).

5. CONCLUSÕES

O TerraView foi o SIG que apresentou maior variação nas classificações, ora apresentando resultados mais distantes da verdade terrestre (menor acurácia), ora com os melhores valores, provenientes da pós-classificação com o uso do seu filtro majoritário, onde o mesmo suprimiu grupos de pixels erroneamente classificados por classes e região, o que abrandou consideravelmente a estética dos mapas e consequentemente a diminuição de erros.

É importante salientar que, provavelmente se incorporados subdivisões espectrais de usos que apresentam tempos fenológicos distintos, bem como imagens com melhor resolução espacial e aplicação dos métodos em uma área com mais homogeneidade, com superfície menos rugosa

e espectralmente menos ruidosa (sombras e outros aspectos que provocam dispersão de ondas eletromagnéticas e/ou escalas maiores com nível de detalhamento menor) os resultados de acurácia dos aplicativos SIG utilizados, seriam possivelmente mais elevados e confiáveis, pois tais ferramentas apresentam um alto poder de processamento de dados digitais.

6. REFERÊNCIAS

- [1] 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.
- [2] 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.
- [3] LANDIS, R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v.33, n.1, p.159-174, Mar. 1977.
- [4] 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.
- [5] RODRIGUES, M. T.; RODRIGUES, B. T.; TAGLIARINE, 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.