



SENSORIAMENTO REMOTO E GEOPROCESSAMENTO APLICADO AO USO DA TERRA PARA AVALIAÇÃO ENTRE CLASSIFICADORES A PARTIR DO ÍNDICE KAPPA

RODRIGUES, Mikael Timóteo¹; POLLO, Ronaldo Alberto Pollo¹;
RODRIGUES, Bruno Timóteo²; MALHEIROS, Jessica Moraes³; CAMPOS, Sérgio⁴

RESUMO – (SENSORIAMENTO REMOTO E GEOPROCESSAMENTO APLICADO AO USO DA TERRA PARA AVALIAÇÃO ENTRE CLASSIFICADORES A PARTIR DO ÍNDICE KAPPA). O presente artigo buscou comparar as técnicas de mapeamento do uso e ocupação do solo na bacia hidrográfica do Ribeirão Paraíso localizado no município de São Manuel – SP confrontando métodos de classificação supervisionada MAXVER e não supervisionada CLUSTER realizada em imagem do sensor TM (*Thematic Mapper*) do satélite Landsat-5 na combinação Falsa cor. A verdade terrestre do uso e ocupação das terras na bacia hidrográfica do Ribeirão Paraíso foi elaborado no software ArcGIS 10.1 e as amostras de cada uso do solo (MAXVER) foram processadas através do aplicativo MAXLIKE inserido no software IDRISI, bem como o mesmo número de classes para o CLUSTER. Por fim, as duas classificações foram comparadas com a verdade como parâmetro para identificar a melhor metodologia no estudo de classificação em bacias. Frente aos resultados obtidos, a classificação supervisionada MAXVER seria a mais indicada para se realizar análise do uso da terra por meio de classificação digital, devido sua maior capacidade de acurácia em relação à classificação não-supervisionada CLUSTER.

Palavras-chave: Geotecnologia; Imagem orbital; Bacia hidrográfica.

ABSTRACT – (REMOTE SENSING AND GEOPROCESSING APPLIED TO LAND FOR EVALUATION FROM BETWEEN CLASSIFIERS KAPPA INDEX). This article aimed to compare the techniques of mapping the use and occupation of land in the watershed of Ribeirão Paraíso located in the municipality of San Manuel - SP confronting methods of supervised classification MAXVER and unsupervised CLUSTER held on image sensor TM (*Thematic Mapper*) Landsat - 5 in combination False color. The ground truth of the use and occupation of land in the watershed of Ribeirão Paraíso was developed in ArcGIS 10.1 software and samples of each land use (MAXVER) were processed through MAXLIKE inserted in IDRISI software application as well as the same number of classes for CLUSTER. Finally, the two classifications were compared with the truth as a parameter to identify the best methodology in the study of classification basins. Given our results, the supervised classification MAXVER would be the most suitable to carry out analysis of land use by digital classification, due to its greater capacity for accuracy in relation to unsupervised classification CLUSTER

Keywords: Geotechnology; Orbital picture; Watershed.

¹Geógrafo, Doutorando em Agronomia (Energia na Agricultura) - Universidade Estadual Paulista - UNESP/FCA

² Geógrafo e Gestor Ambiental – Pesquisador Associado do Laboratório de Geoprocessamento Aplicado – LGA/UFAL

³ Zootecnista, Doutoranda em Melhoramento Genético Animal - Universidade Estadual Paulista - UNESP/FCAV

⁴Professor Titular - Universidade Estadual Paulista - UNESP/FCA

1. INTRODUÇÃO

Os sistemas de informações geográficas aliados as imagens de satélite são de grande importância para o estudo do meio ambiente, pois nos permitem analisar grandes quantidades de dados num menor período de tempo, possibilitando um melhor planejamento do uso do solo.

Na bacia hidrográfica do ribeirão Paraíso, São Manuel-SP, objeto deste estudo, observa-se a ação antrópica na modificação intensa da paisagem natural, ocasionado principalmente pelo uso intensivo das terras, onde se encontra a área urbana e porções da Área de Proteção Ambiental (APA) Perímetro Botucatu-SP, considerada uma área de recarga do Aquífero Guarani, além de várias nascentes que abastecem a população comércio e indústrias do município.

A utilização do Geoprocessamento como ferramenta é bastante eficaz na espacialização e quantificação dos recursos naturais de uma determinada área tanto a nível regional quanto a nível local, abrangendo o planejamento e emprego adequado da ocupação de determinado perímetro, assim se tornando uma alternativa viável na restrição das possíveis falhas e lacunas. Os Sistemas de Informações Geográficas (SIG) tem

colaborado para o mapeamento e zoneamento, atuando como método de fiscalização e planejamento de áreas agrícolas em determinadas bacias hidrográficas e seus respectivos espaços e uso da terra, uma vez que as mudanças ocorrem em uma pequena escala temporal (RODRIGUES, et al, 2013).

O sensoriamento remoto tem papel significativo no ambiente SIG, 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 objetos de uma dada região, e seu comportamento espectral (interação da radiação solar com os objetos 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).

A classificação de imagens diz respeito à interpretação com auxílio de aplicativos e computadores, atribuindo significado específico para cada pixel ou a

um conjunto de pixels através da criação de polígonos, em função de suas propriedades numéricas (ROSA, 2009). Esse processo de classificação se caracteriza pela atribuição aos conjuntos de pixels que possuem características espectrais comuns entre si, onde cada assinatura espectral profere ao referido conjunto de pixels representa um tipo específico de uso da terra, pastagem, mata nativa, área urbana dentre outros usos.

As técnicas de classificação de imagens podem ser divididas em não-supervisionada e supervisionada. A primeira não requer qualquer informação sobre classes de interesse, somente examinando sem auxílio do analista, os agrupamentos espectrais do alvo e dividindo-os em classes (TANGERINO & LOURENÇO, 2013). O CLUSTER ou Classificação Não-Supervisionada trabalha com uma a lógica conhecida como análise de agrupamento, sendo útil quando não se possui informações sobre a área imageada, sendo as classes definidas no algoritmo de classificação (ROSA, 2009).

A classificação supervisionada consiste na identificação prévia das classes de informação chamadas áreas de treinamento, que nada mais são do que representações do comportamento médio das classes que serão mapeadas automaticamente (TANGERINO &

LOURENÇO, 2013). A principal técnica utilizada para a realização de classificações supervisionadas é a MAXVER (máxima verossimilhança), que consiste na classificação da imagem ponto a ponto, a partir de classes fornecidas pelo usuário, (ROSA, 2009).

Este trabalho tem como objetivo comparar os métodos de classificação supervisionada e não supervisionada, através do padrão espectral em imagem Landsat 5, associada a comparação do uso da terra.

2. MATERIAL E MÉTODOS

A bacia hidrográfica do ribeirão Paraíso localiza-se no município de São Manuel e Areiópolis-SP, entre as coordenadas geográficas 48°32'49" a 48°37'44" de longitude W Gr. e 22°37'53" a 22°47'03" de latitude S, distando aproximadamente 270 Km da capital do estado.

Para o apoio cartográfico, foram utilizadas cartas planialtimétricas editadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em formato impresso e digital em escala 1:50.000 com equidistância das curvas de nível de 20 metros referente a Folha de São Manuel/SF-22-Z-B-V-2 e Pratânia/SF-22-Z-B-V-4 (IBGE, 1973).

Utilizou-se também imagem do sensor TM (*Thematic Mapper*) do satélite Landsat-5 cedida pelo Instituto Nacional de Pesquisas Espaciais (INPE) em formato digital, Figura 1, sendo posteriormente georreferenciada (Sistema de Projeção

UTM) nas bandas termais 3, 4 e 5 composição colorida (R-G-B), com resolução espacial de 30 metros, referentes à órbita 220, ponto 76, passagem de 19/11/2010.

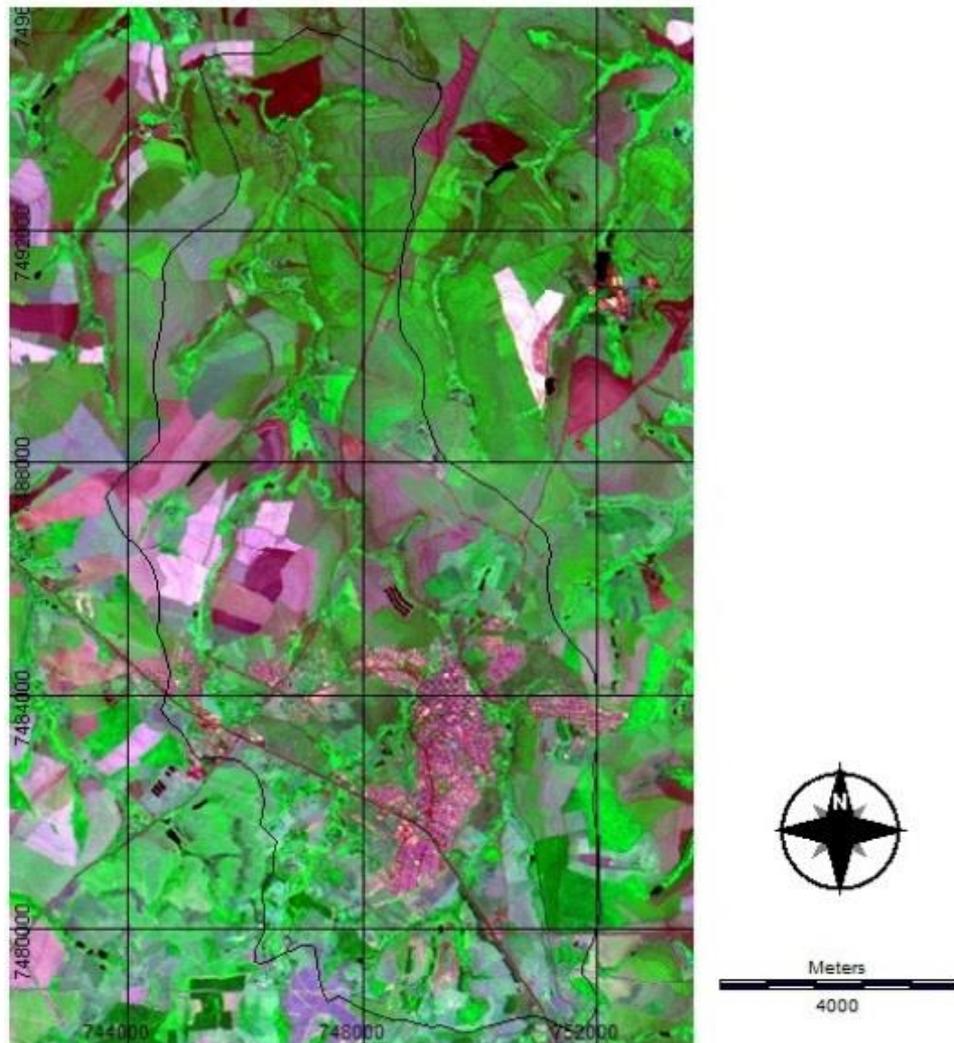


Figura 1. Área de estudo na Imagem de Satélite em 2010 pelo sensor TM do Landsat-5.

A partir das imagens, elaborou-se um mapa da verdade terrestre do uso e ocupação das terras na bacia hidrográfica do Ribeirão Paraíso utilizando-se do software ArcGIS 10.1. 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 12 classes de uso e cobertura para o ano de 2010 através da imagem de satélite. De posse do arquivo completo vetorizado, o mesmo foi exportado para o IDRISI.

Após a vetorização, realizou-se uma reclassificação da verdade terrestre para que a mesma tivesse precisamente o mesmo número de classes do uso da terra correspondente as classificações supervisionada e não-supervisionada, decorrendo assim seis classes: Pastagem, Mata, Cana-de-açúcar, Área urbana, Outras culturas e Solo exposto.

A classificação supervisionada MAXVER foi realizada com quatro amostras de cada uso do solo, e processadas através do aplicativo MAXLIKE inserido no software IDRISI Selva. Para classificação não supervisionada CLUSTER, foi processado o mesmo número de classes referentes ao mesmo uso da verdade

terrestre e da classificação supervisionada. Neste procedimento, o IDRISI Selva reconhece os padrões espectrais de cada grupo de pixels e os incorpora por aceção em classes distintas.

3. RESULTADOS E DISCUSSÃO

A verdade terrestre referente à bacia do Ribeirão Paraíso (Figura 2), foi obtida como resultado respectivo a real utilização do uso do solo na área de estudo, assim, realizado o índice para comparação com as classificações MAXVER e CLUSTER.

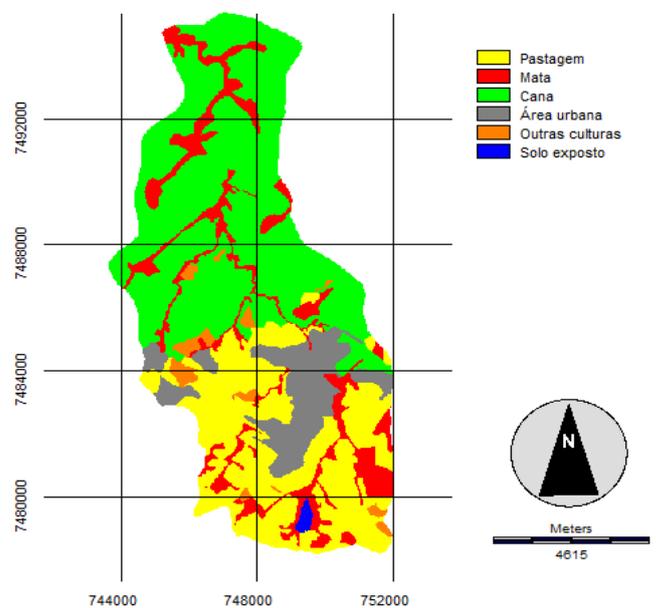


Figura 2. Verdade terrestre da bacia do Ribeirão Paraíso.

A classificação supervisionada (Figura 3) através do método MAXVER (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 0.24, tendo em vista as características de uso na bacia estudada, pois a mesma apresenta grande área com o cultivo da cana nos mais variados estágios de desenvolvimento e produção, bem como grandes áreas de 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.

Podemos observar de acordo com a matriz de confusão gerada a partir da classificação supervisionada (Tabela 1), que existe um aceitável grau de acurácia nos dados classificados, onde cada uso teve maiores valores em suas respectivas classes. Entretanto, como já foi mencionado anteriormente, a classe de Cana gerou muita confusão com as outras classes, o que já era esperado por se tratar de um uso

bastante diversificado no que se diz respeito às respostas espectrais.

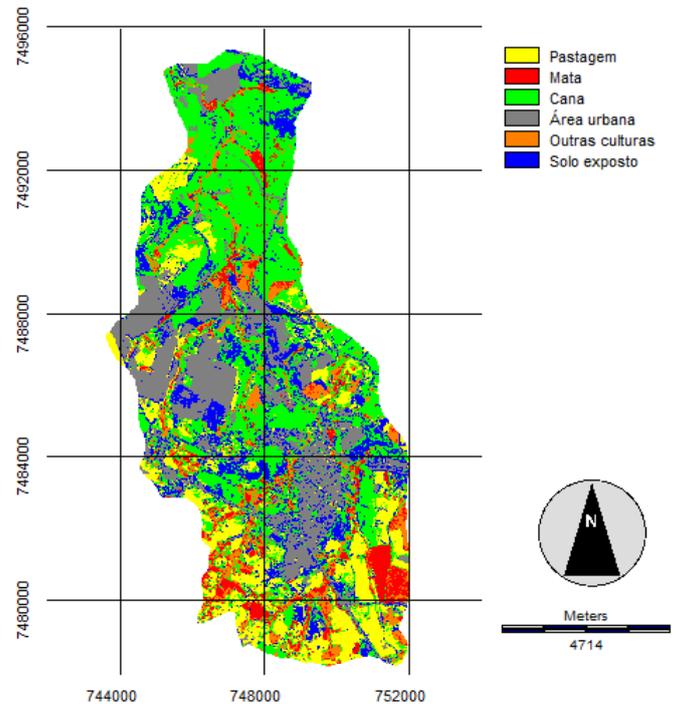


Figura 3. Classificação Supervisionada MAXVER.

Juntamente com a Cana, a Pastagem, Mata e Outras Culturas também tiveram um elevado grau de confusão com outras classes, bem como entre elas por se encontrarem mal representadas espectralmente (figura 4). Ainda com base na análise da matriz de confusão, o maior grau de acurácia foi da classe Cana, com 35,57% de precisão, seguida da classe Área Urbana, com acurácia na ordem de 20,8%.

Tabela 1. Tabela determinada pela Matriz de Confusão gerada a partir da Classificação Supervisionada

Classes	Pastagem	Mata	Cana	Área urbana	Outras culturas	Solo exposto	Total
Pastagem	9162	445	3938	336	252	67	14200
Mata	1510	4189	711	119	297	1	6827
Cana	7942	4065	20364	1228	647	52	34298
Área urbana	1785	574	11852	5769	35	46	20061
Outras culturas	1730	3770	1158	120	678	8	7464
Solo exposto	3464	748	6733	2145	279	197	13566
Total	25593	13791	44756	9717	2188	197	96416

O resultado da classificação poderia ser melhor se fosse confrontado uma verdade terrestre com mais classes do uso da terra, ou com subdivisões para algumas classes, como diferenciação dos grupos de Cana e seus respectivos solos preparados e subdivisões na classe Mata, dividindo-a

em Reflorestamento, Mata Nativa, Várzea, Mata Degradada, assim sendo, cruzada com uma classificação supervisionada com mais classes, ocorrendo uma melhor representação espectral da cena com menos interferências de dados amostrais de outras classes.

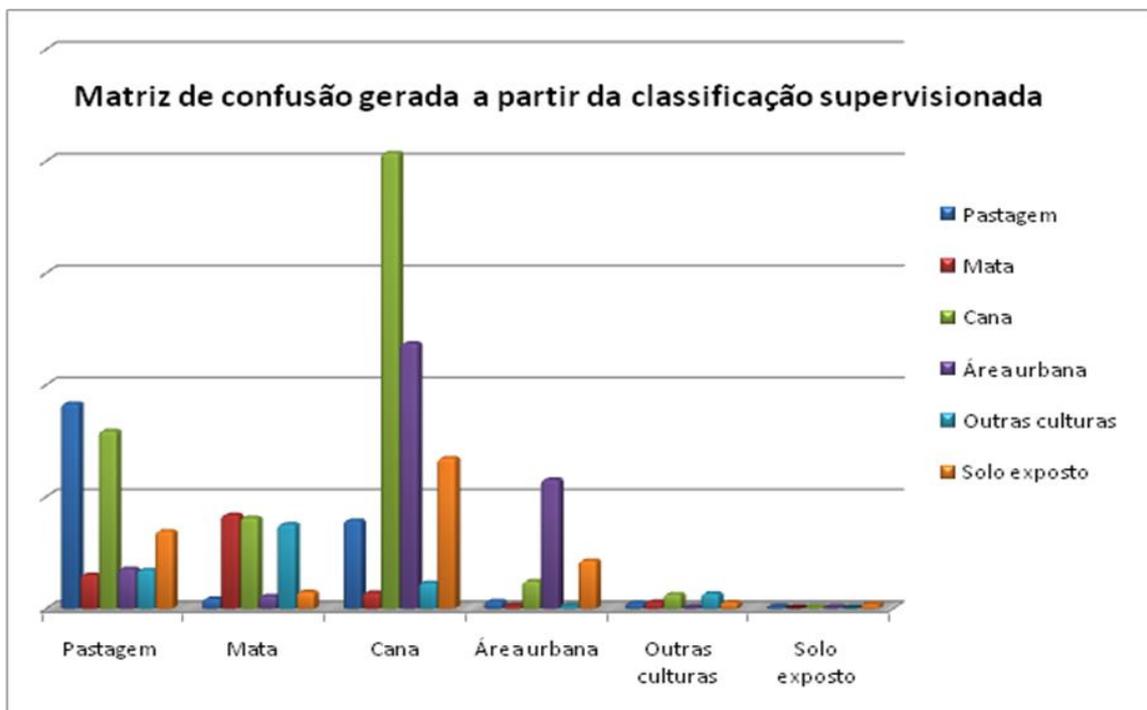


Figura 4. Matriz de Confusão Gerada a partir da Classificação Supervisionada.

Para a classificação não-supervisionada, realizada através do método CLUSTER, onde o algoritmo define as classes por meio de grupos espectrais (Análise de agrupamento) sem qualquer influência do analista, selecionando pixels e correlacionando com grupos mais próximos dos seus valores espectrais, foi encontrado resultados insatisfatórios, onde o índice Kappa atingiu valor na ordem de 0.12, o que caracteriza, segundo Landis e Koch (1977) adaptado por Piroli, (2010), a classificação como ruim, onde tal classificação não representa os verdadeiros usos da terra encontrado na área de estudo, como pode ser observado na Figura 5.

O algoritmo apresentou grande confusão nas separações das classes, tendo em vista que para esse tipo de classificação é usada uma técnica de seleção pixels pelos picos dos histogramas, e quando os picos foram identificados, todos os possíveis valores são associados ao pico mais próximo, e as classes tendem a cair no ponto médio entre os picos (EASTMAN, 2006).

Portanto, o algoritmo desalinhou automaticamente as respostas espectrais de algumas classes (Tabela 2), assim reduzindo o grau de

acurácia nos dados classificados, onde alguns usos tiveram seus valores subestimados e o seu grau de acurácia comprometido com erros de inserção de pixel que não pertencem a determinadas classes a elas atribuídas.

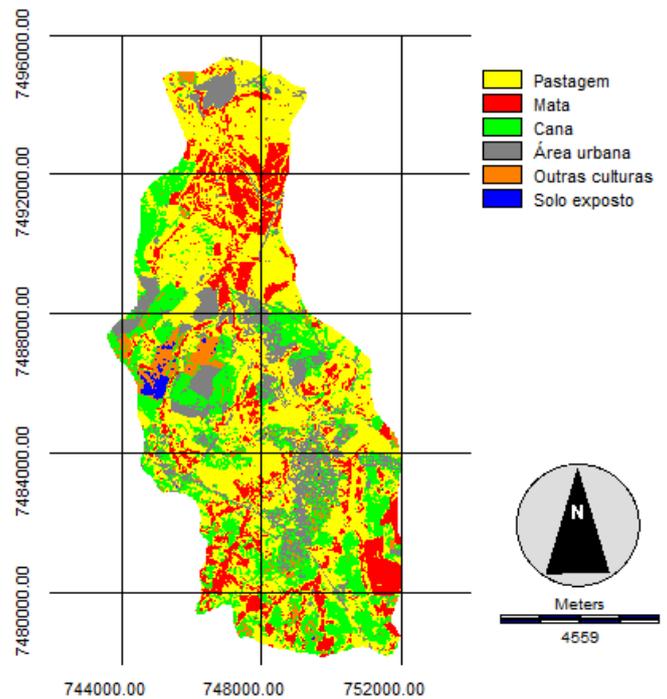


Figura 5. Classificação Não-Supervisionada CLUSTER.

A imagem Landsat 5 foi trabalhada na composição RGB na faixa do visível, referente as bandas 3, 4 e 5 do sensor TM, sem fusão com a banda Pancromática, o que melhoraria a resolução espectral em 15 metros e auxiliaria o algoritmo no cálculo da matriz de confusão, bem como nas resposta da acurácia. A restrição da quantidade de classes (máximo de 6 classes) na utilização do método

CLUSTER também contribuiu para resultados de baixa acurácia. Na figura 6 referente a matriz de confusão, pode-se destacar o conflito gerado pela classe Pastagem em meio a alguns usos,

principalmente na classe Cana e Outras Culturas, onde as mesmas tiveram seus respectivos graus de acurácia reduzidos significativamente.

Tabela 2. Tabela determinada pela Matriz de Confusão gerada a partir da Classificação não - Supervisionada

Classes	Pastagem	Mata	Cana	Área urbana	Outras culturas	Solo exposto	Total
Pastagem	12946	5387	21600	2059	1047	91	43130
Mata	2618	7607	4914	191	828	3	16161
Cana	8291	412	8367	2761	218	187	20236
Área urbana	1032	303	7092	4119	89	28	12663
Outras culturas	700	82	2150	565	6	62	3565
Solo exposto	6	0	538	22	0	0	566
Total	25593	13791	44756	9717	2188	371	96416

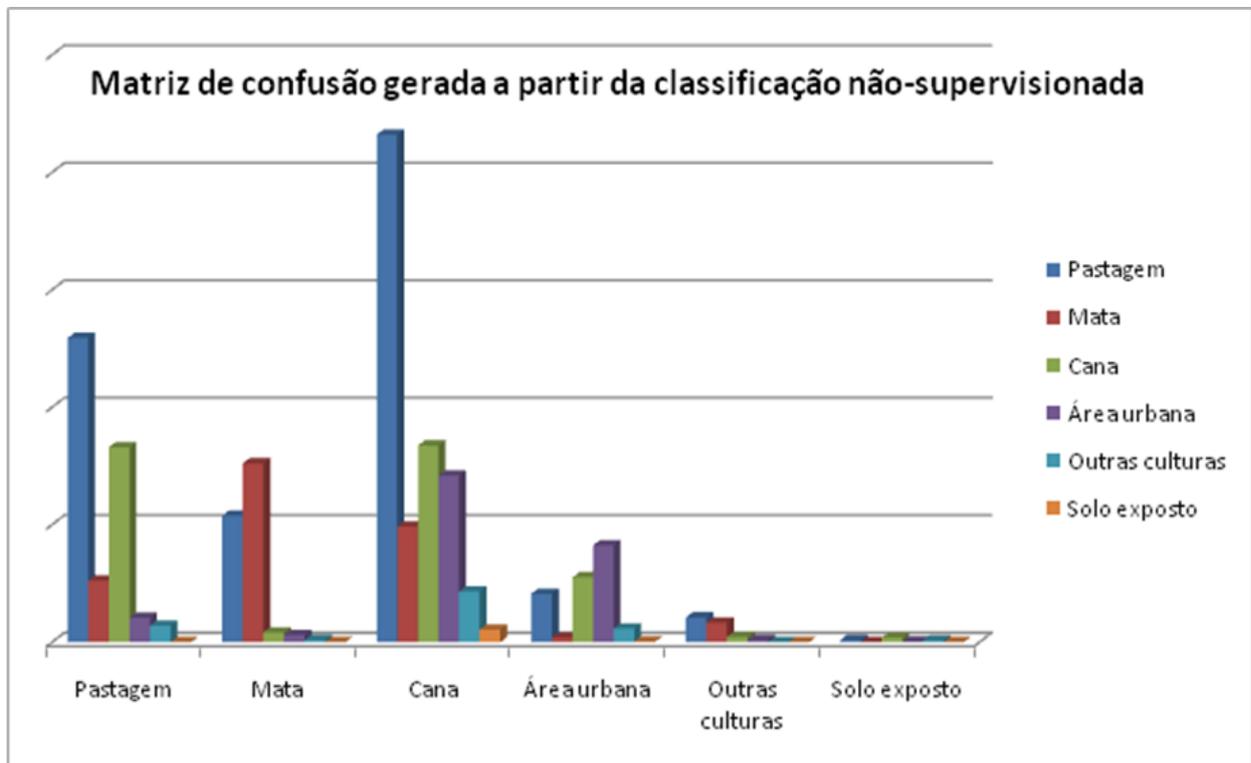


Figura 6. Matriz de Confusão Gerada a partir da Classificação Não-Supervisionada.

4. CONCLUSÃO

Devido os mais variados estágios de desenvolvimento e produção da Cana, bem como grandes áreas de solo preparado para o manejo desta cultura, emitiram diversas respostas espectrais de um mesmo uso, o que caracterizou um alto grau de confusão da reflectância, bem como e a restrição da quantidade de classes da verdade terrestre contribuíram para um resultado razoável na classificação MAXVER e um resultado insatisfatório na classificação CLUSTER.

Frente aos resultados obtidos, a classificação supervisionada MAXVER é a mais indicada para se realizar análise do uso da terra por meio de classificação digital, devido sua maior capacidade de acurácia em relação à classificação não-supervisionada CLUSTER.

A classificação não-supervisionada pode ser favorável, quando não se possui informações sobre a área de interesse ou até mesmo quando não pode ser definida visualmente pelo analista, devido a elementos mal definidos, onde não é possível atribuir amostras de treinamento (*training sites*), sendo assim, as classes definidas no algoritmo de classificação, a serviço

apenas da reflectância, sujeita a confusões referentes aos valores de pixel e sua respectiva acurácia que são influenciados pela resolução da imagem, bem como pelo uso da terra.

5. REFERÊNCIAS

ANDRADE, E. L.; SILVA, U. M. Avaliação da classificação não-supervisionada em imagens de alta resolução, utilizando a lógica fuzzy. XV Simpósio Brasileiro de Sensoriamento Remoto - SBSR. 2011, Curitiba, PR. **Anais...** São José dos Campos: INPE, 2011, p. 7418.

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.

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 15: The Andes Edition**. Worcester, MA: Clark University, 2006.

_____. **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.

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: 03 mar. 2013.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Carta topográfica**: folhas São Manuel - SF-22-Z-B-V-2 e Pratânia - SF-22-Z-B-V-4. Serviço gráfico do IBGE, 1973. Escala 1:50.000.

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

NASA. The Landsat program. Disponível em: <<http://landsat.gsfc.nasa.gov/>>. Acesso em: 17 ago. 2012.

PIROLI, E. L.. **Introdução ao geoprocessamento**. Ourinhos: Unesp/Campus Experimental de Ourinhos, 2010a. 46p.

RODRIGUES, M. T.; CERIGNONE, F. J.; RODRIGUES, B. T.; MALHEIROS, J. M.; MONTE, B. E. O.; CARDOSO, L. G. Utilização do geoprocessamento na classificação do uso da terra em áreas conflitantes da bacia hidrográfica do rio capivara no município de botucatu-sp. IV Workshop Internacional de Planejamento em Bacias Hidrográficas. 2013, Presidente Prudente. **Anais...** Presidente Prudente, SP: UNESP/FCT, 2013.

ROSA, R.. **Introdução ao sensoriamento remoto**. Uberlândia: EDUFU, 2009.

SÃO PAULO. Decreto Estadual Nº 20.960, de 8 de junho de 1983. **Área de Proteção Ambiental**: São Paulo. Diário Oficial v.96, n.002,1983. Disponível em:<<http://www.ambiente.sp.gov.br/apa/s/20960.htm>>. Acesso em: 23 ago.2012.

TANGERINO, D. F.; LOURENÇO, R. T. Comparação da exatidão de métodos de classificação supervisionada e não supervisionada a partir do índice kappa na microbacia do Ribeirão Duas Águas em Botucatu/SP. XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR. 2013, Foz do Iguaçu, PR. **Anais...** São José dos Campos: INPE, 2013, p. 4093.