

Métodos de classificação e acompanhamento da dinâmica de alteração do uso da terra nos municípios de Analândia e Santa Cruz da Conceição/SP – 2001 a 2015

Antoniane Arantes de Oliveira Roque ¹

Roberto Barros dos Santos ²

Jansle Vieira Rocha ²

Gleyce Kelly Dantas Araujo Figueiredo ²

Rubens Augusto Camargo Lamparelli ⁽³⁾

¹ Núcleo de Estudos e Pesquisas Ambientais/Universidade Estadual de Campinas –
NEPAM/UNICAMP
Rua dos Flamboyants, nº155, Cid. Univ. Zeferino Vaz, CEP 13083-867, Campinas/SP, Brasil
antoniane@yahoo.com.br

² Faculdade de Engenharia Agrícola/Universidade Estadual de Campinas
FEAGRI/UNICAMP
Avenida Cândido Rondon, nº501, Barão Geraldo, CEP 13083-875, Campinas/SP, Brasil
ro.barrossan@gmail.com, jansle.rocha@feagri.unicamp.br, gleyce.figueiredo@feagri.unicamp.br

³ Unicamp, Núcleo Interdisciplinar de Planejamento Energético, Rua Cora Coralina, no
330, CEP 13083-896 Campinas, SP. E-mail: rubens.lamparelli@gmail.com

Abstract. The dynamics of the land-cover change, appears as the most important task at the present time, in which this activities have a direct influence on environmental resources available to society. The methods of the classification of remote sensing images are fundamental to understanding the dynamics of changing occupation of the territory, and are presented as a key tool for effective management of natural resources. This study aimed to define the best classifiers of images for the uses in the agricultural region of the center-east of Sao Paulo/Brazil, in the temporal cutouts 2001 and 2015, analyzing the ratings for each year, and between the two different years, using of error matrix and Kappa index. Was used images of Landsat (satellites 7 and 8), instruments ETM+ and OLI respectively, and processed in ENVI and ArcGIS. It was concluded that the classification supervised by distance of Mahalanobis should be used with caution in the event of clayey soils with high humidity, because the spectral signature of water is similar of soil wet, for this method of classification. In the analysis for the first year we obtained an overall accuracy of 85%, which is an good indicator of accuracy of the classifiers selected. In the comparative analysis between the years under review, the overall accuracy was 26.3% and the Kappa index of 0.13, thus indicating that there was a significant change in land-cover. It is emphasized wich to carry out the land use classification, it is necessary to use more than one classifier.

Palavras-chave: índice Kappa, sensoriamento remoto, processamento de imagens, Landsat 8, matriz de erro.

1. Introdução

Acompanhar a dinâmica da alteração do uso da terra pela sociedade se mostra como tarefa das mais importantes na atualidade, na qual as atividades exercem influência direta nos recursos ambientais disponíveis à sociedade.

Este acompanhamento, e seu consequente mapeamento, possuem ferramentas no atual nível tecnológico da evolução humana, que o permitem em escalas até então não possíveis de se trabalhar, associando acurácia e precisão aos diferentes objetivos pretendidos.

O levantamento do uso atual da terra, necessário para fins de planejamento, pode ser obtido a partir da utilização de dados multiespectrais, fornecidos por satélites de Sensoriamento Remoto, associados às técnicas de interpretação (Pereira et al., 1989).

Os produtos de Sensoriamento Remoto trazem insumos fundamentais aos atuais técnicos e especialistas da área da geomática, aliando elementos da cartografia clássica à possibilidade de análises de extensas porções da superfície.

A classificação digital de imagens associa cada pixel da imagem a um determinado rótulo, obtendo-se como resultado um determinado tema. Existem essencialmente duas abordagens na classificação de imagens multiespectrais de sensoriamento remoto: a classificação supervisionada e a não-supervisionada (Crosta, 1992).

Na primeira, classificação supervisionada, através de algoritmos pré-estabelecidos, o usuário procura identificar pontos na imagem pertencentes a classes desejadas, deixando ao programa de classificação a identificação de todos os demais pixels pertencentes àquelas classes. Já o processo de classificação não supervisionada consiste em atribuir cada pixel de uma imagem a um grupo de classe. Essa atribuição é feita com base em observações do pixel, em si e na sua vizinhança, bem como numa série de regras (Pereira, 1992).

As vantagens de utilizar dados de sensoriamento remoto nos levantamentos do uso atual das terras, segundo Freitas Filho e Medeiros (1993), são atingir grandes áreas de difícil acesso e fazer o imageamento a altas altitudes, possibilitando uma visão sinóptica da superfície terrestre, com repetitividade, viabilizando as ações de monitoramento.

Os instrumentos computacionais do geoprocessamento chamados de SIG ou GIS (Geographic Information System) permitem a realização de análises complexas ao integrar dados de diversas fontes e ao criar bancos de dados georreferenciados por possibilitarem a automatização da produção de documentos cartográficos (Assad e Sano, 2003). Os SIGs são formados pelos módulos de interface com o usuário, entrada e integração de dados, funções de processamento gráfico e de imagem, visualização e plotagem e armazenamento e recuperação de dados.

Segundo Maximiliano (1996), os Sistemas de Informações Geográficas (SIGs), na análise integrada dos dados do meio físico, tem sido usado por vários setores que tratam da questão ambiental, como importante ferramenta para o planejamento ambiental.

Este trabalho pretende investigar a alteração do uso do solo nos municípios de Analândia e Santa Cruz da Conceição, estado de São Paulo/Brasil, por meio do uso de imagens de sensoriamento remoto e utilização de ferramentas de geoprocessamento, em sistemas de informação geográfica, visando-se acompanhar a dinâmica desta alteração, bem como avaliar a qualidade das classificações realizadas para os diferentes anos.

O objetivo geral é demonstrar os classificadores de imagens de sensoriamento remoto que melhor se adequam à localidade e às imagens utilizadas; sua respectiva análise da qualidade do procedimento realizado e a comparação entre duas datas de anos distantes entre si. Analisar as classificações de cada ano, e entre os dois diferentes anos por meio de matriz de erro ou de contingenciamento, utilizando-se a exatidão global e o Índice Kappa para tal feito.

2. Metodologia de Trabalho

A área de estudo está situada na região centro-leste do estado de São Paulo/Brasil e constituem uma região de expressiva importância no abastecimento de insumos agrícolas à região metropolitana de Campinas e São Paulo.

Analândia é um município categorizado como Estância Climática, e localiza-se a uma latitude 22,07° sul e a uma longitude 47,39° oeste. Possui uma área de 326,6 km².

Santa Cruz da Conceição localiza-se a uma latitude de 22,14° sul e a uma longitude de 47,45° oeste, estando a uma altitude de 645 metros. Possui uma área de 149,87 km².

As imagens das cidades de Analândia e Santa Cruz da Conceição foram obtidas através do site www.earthexplorer.usgs.gov, da U.S Geological Survey. Para a obtenção das imagens

foi utilizada a grade 220/75 do Landsat 7 ETM+ e Landsat 8 OLI, datadas de 23 de março de 2001 e 05 de maio de 2015, respectivamente.

O processamento das imagens, bem como sua classificação foi realizado por meio do uso dos *softwares* ENVI versão 4.5 e ENVI versão 5.1. A confecção dos mapeamentos e obtenção dos mapas finais foi realizado com o uso do *software* ArcGIS versão 10.2.2. Para visualização de padrões atuais da localidade, bem como suporte a entendimento da dinâmica da região utilizou-se o *software* GoogleEarth Pro versão 7.1.5.

3. Resultados e Discussão

Como o objetivo do presente estudo reside na diferenciação do uso da terra, optou-se pela composição das bandas 4, 5 e 3 do satélite Landsat 7, sensor ETM+, e das bandas 5, 6 e 4 do satélite Landsat 8, sensor OLI. Tal escolha residiu em constatações de significância, após testes realizados, corroborados com bibliografias consultadas de processamento de imagens digitais destes satélites.

Utilizou-se ainda a composição cor natural (RGB 321/Landsat7 e RGB 432/Landsat8) para auxílio nas análises. Efetuou-se ainda a aplicação de realce para melhoria de visualização em tela (Figura 1).

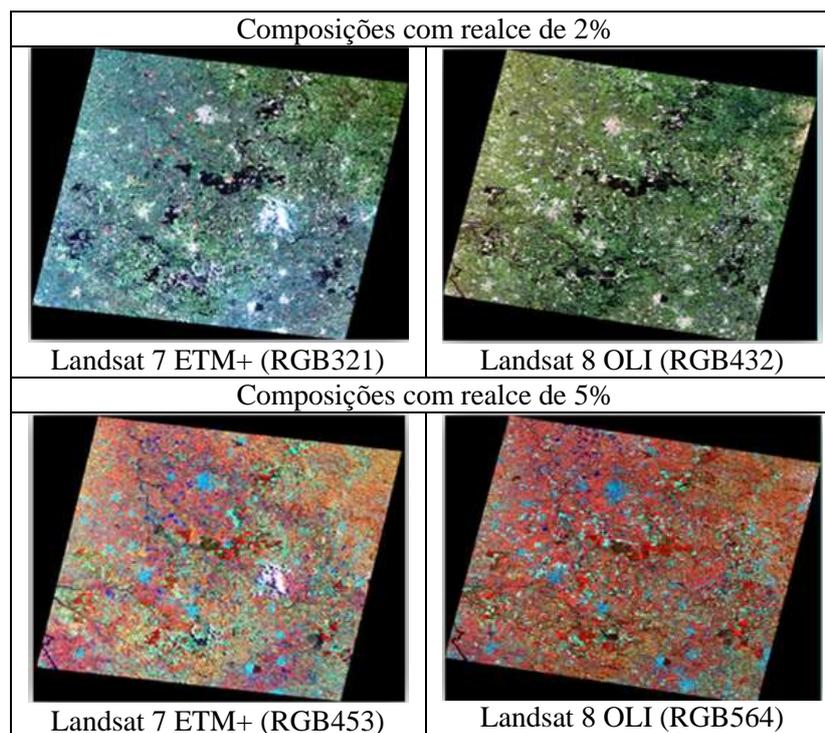


Figura 1 - Imagens das composições utilizadas no estudo.

Passo importante no uso das imagens de sensores remotos reside na conversão dos valores de níveis de cinza de seus pixels, para valores de radiância, realizada banda a banda.

Procedeu-se então à conversão para valores de reflectância e correção atmosférica destas, utilizando-se o módulo do ENVI 5.1 FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes), sendo este um módulo para correção atmosférica de diversos sensores utilizando os algoritmos mais avançados (MODTRAN). A nova versão do módulo FLAASH proporciona uma melhor capacidade para a correção atmosférica. Os parâmetros utilizados foram: 1.Elevação do solo = 650 m; 2.Data do vôo = de acordo com a data de cada imagem;

3.Modelo Atmosférico = Tropical; 4.Modelo de aerossol = Rural; 5.Recuperação de aerossol = Nenhum; 6.Visibilidade Inicial = 40 km.

Visando-se melhoria de desempenho da máquina, objetivando-se o trabalho baseado apenas na região de interesse efetuou-se o recorte para somente o foco do estudo (Figura 2).

Efetuuou-se a determinação de áreas para classificação supervisionada para os dois diferentes anos, mantendo-se pelo menos 1.500 pixels por classe.

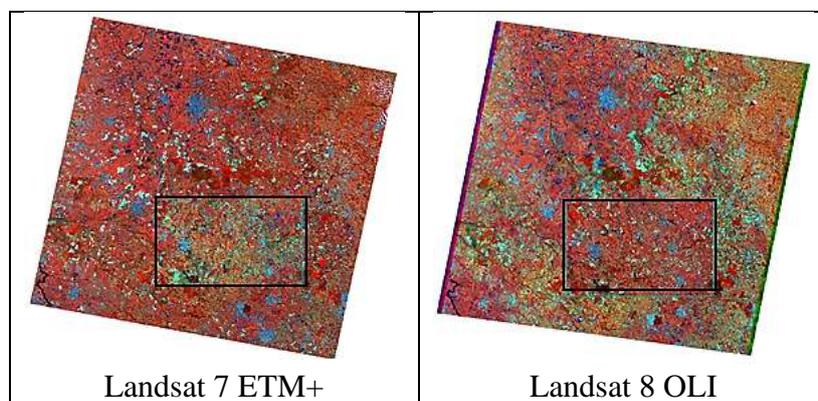


Figura 2 - Recorte das imagens para a região de estudo.

Na Figura 3 são apresentadas as classificações utilizadas para o ano inicial, sendo estas de máxima verossimilhança (Maximum Likelihood Classification) e a distância de Mahalanobis (Mahalanobis Distance Classification).

A distância de Mahalanobis é uma medida de distância introduzida pelo matemático indiano Prasanta Chandra Mahalanobis em 1936. É baseada nas correlações entre variáveis com as quais distintos padrões podem ser identificados e analisados. É uma estatística útil para determinar a similaridade entre uma amostra desconhecida e uma conhecida.

A estimativa por máxima verossimilhança é um método para estimar os parâmetros de um modelo estatístico. Assim, a partir de um conjunto de dados, e dado um modelo estatístico, a estimativa por máxima verossimilhança estima valores para os diferentes parâmetros do modelo.

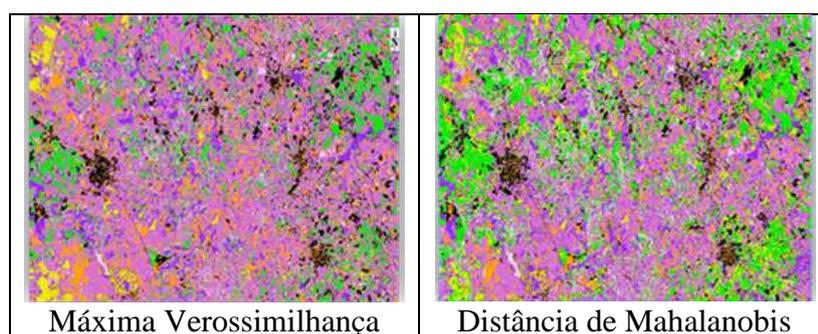


Figura 3 – Classificações supervisionadas sobre Landsat 7 ETM+.

Efetuuaram-se os testes iniciais utilizando-se da classificação não supervisionada por Isodata e pela classificação supervisionada pelo método do paralelepípedo (Figura 4).

O método do paralelepípedo é um método de classificação supervisionada que define áreas quadradas ou em paralelepípedo limitadas pelo maior e pelo menor valor de pixels contidos em agrupamentos pré-escolhidos, determinados pelas amostras de treinamento, que representarão uma determinada classe presente na imagem para a classificação.

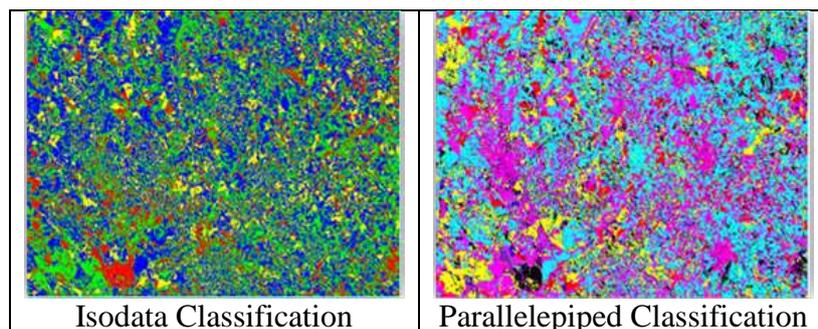


Figura 4 – Classificação não supervisionada e supervisionada sobre Landsat 7 ETM+.

As mesmas classificações adotadas para o ano inicial foram utilizadas para o ano final.

Visando-se a obtenção de áreas homogêneas, que dessem subsídio ao planejamento a nível regional, optou-se pela aplicação de um filtro de média 5x5 sobre todas as imagens classificadas, apresentando-se como exemplo a Figura 5.

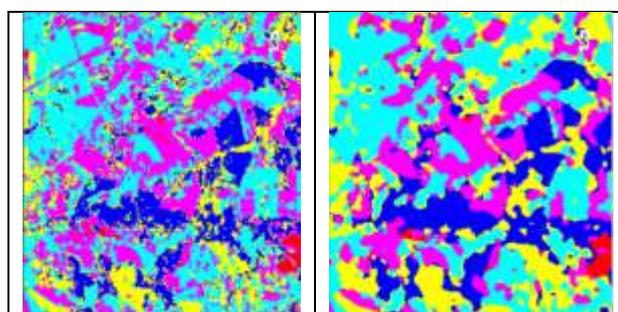


Figura 5 – Imagem classificada pela distância de Mahalanobis (Landsat 8 OLI), e após filtragem de média 5x5.

De posse das imagens em rasters, efetuou-se rotina de vetorização destas, obtendo-se assim diversos polígonos em formato vetorial.

Analisando-se as respostas de cada classificador, selecionaram-se as classes que melhor corresponderam às feições conforme Tabela 1.

Tabela 1 - Seleção das classes de cada classificação.

Classificação	Landsat 7 ETM+	Landsat 8 OLI
MaxVer	Pasto, Mata, Eucalipto e Cana-de-açúcar	Pasto, Mata, Eucalipto e Cana-de-açúcar
Mahalanobis	Água, Solo Exposto, Solo com palha e Laranja	Água e Solo exposto

Realizou-se a junção dos polígonos de cada classe num shapefile único, obtendo-se assim um arquivo para cada ano em análise.

Utilizado como base vetorial municipal, os limites obtidos em arquivo vetorial provenientes da base de dados da Infraestrutura Nacional de Dados Espaciais. A vetorização da área urbana foi realizada sobre imagem georreferenciada do Google Earth Pro e verificada sua não alteração ao longo dos anos em estudo.

Efetuoou-se então a junção dos polígonos de mesma classe, e a operação de continuidade vetorial, com posterior cálculo das áreas, com unidades em hectares (Figura 6), salientando-se que tal procedimento só foi possível, pois os vetores estavam em projeção plana.

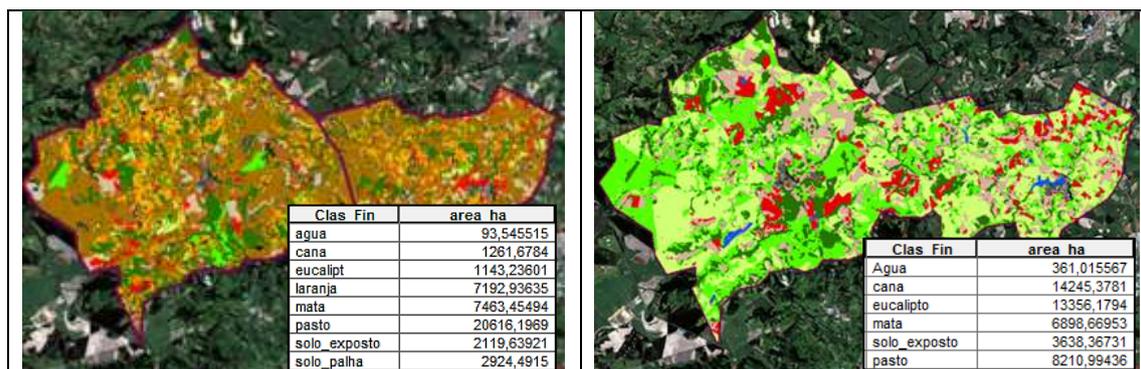


Figura 6 - Classificação Final para o ano inicial e final com as respectivas áreas de cada uso do solo.

Ressalta-se nesta etapa, que houve uma elevada confusão no classificador distância de Mahalanobis no Landsat 8 OLI, com relação a classificação de regiões com solo exposto e elevada umidade, classificando-as como água.

Para o acompanhamento da qualidade da classificação, gerou-se pontos em cada classe, utilizando-se o método aleatório. Optou-se por gerar 10 pontos por classe, sendo gerados no total 60 pontos para a classificação inicial e 80 pontos para a classificação final, devido a diferença de classes entre os dois.

Após a classificação das imagens, utilizou-se a estatística Kappa para testar a concordância entre os resultados observados e os classificados em uma tabela de contingência (matriz de erro).

De posse dos pontos de amostragem, efetuou-se a verificação para o ano inicial, obtendo-se como resultado os dados constantes na Tabela 2.

Na análise para o ano de 2001 (Tabela 2), obteve-se uma exatidão global de 85%, sendo este um indicador de concordância do que foi classificado, para com a referência.

O índice Kappa obtido foi de 0,828. O Kappa é uma medida de concordância inter observador e mede o grau de concordância além do que seria esperado tão somente pelo acaso. Esta medida de concordância tem como valor máximo o 1, sendo o valor 1 o qual representa total concordância, e os valores próximos e até abaixo de 0, indicam nenhuma concordância, ou a concordância foi exatamente a esperada pelo acaso.

Landis e Koch (1977) associam valores de Kappa à qualidade da classificação obtida, considerando valores superiores a 0,81 como excelentes.

Os maiores erros de inclusão estiveram na classificação das áreas de laranja, sendo esta uma classe de difícil identificação. Não foram obtidos erros de omissão para as classes de laranja e água, indicando que estas foram em sua totalidade classificadas.

Tabela 2 - Matriz de erro para o ano de 2001.

		Landsat 7								Soma linhas	Erro de inclusão
		Referência									
		Pasto	Mata	Eucalpto	Cana-de-açúcar	Solo_Palha	Solo_Exposto	Água	Laranja		
Classificação	Pasto	10	1	1						12	0,167
	Mata		11							11	0,000
	Eucalpto			9	1					10	0,100
	Cana-de-açúcar				2	8				10	0,200
	Solo_Palha	1				10				11	0,091
	Solo_Exposto	1					8			9	0,111
	Água							1	6	7	0,143
	Laranja	1		1		1	1		6	10	0,400
Soma colunas		13	14	11	9	11	10	6	6	80	
Acertos		0,769	0,786	0,818	0,889	0,909	0,800	1,000	1,000		Soma Diagonal 68
Erro omissão		0,231	0,214	0,182	0,111	0,091	0,200	0,000	0,000		Exat. Global 0,850 85,0 %
											Σ Produto Col*Lin. 823
											Índice Kappa 0,828 82,8 %

Na análise para o ano de 2015 (Tabela 3), a exatidão global ficou em 81,7% e o índice Kappa em 0,78, valor este dentro da faixa de classificação considerada como muito boa, segundo os critérios de Landis e Koch (1977).

A classe da laranja não foi identificada neste ano, nem a de solo palha, sendo portanto a análise destas linhas e colunas com valores não consideráveis.

Os maiores erros de inclusão (40%) estiveram na classe do eucalipto, tendo ocorrido a classificação de áreas de matas como de eucalipto.

A classe da água, apesar de ter obtido uma taxa de acerto de 100%, obteve uma confusão com as áreas de solo exposto (erro de inclusão de 30%), devido a elevada umidade destes solos, bem como sua característica argilosa.

Tabela 3 - Matriz de erro para o ano de 2015.

		Landsat 8								Soma linhas	Erro de inclusão
		Referência									
Classificação	Pasto	8					2			10	0,200
	Mata		10							10	0,000
	Eucalipto		4	6						10	0,400
	Cana-de-açúcar		1	1	8					10	0,200
	Solo Palha									0	#DIV/0!
	Solo Exposto						10			10	0,000
	Água						3	7		10	0,300
	Laranja									0	#DIV/0!
Soma colunas		8	15	7	8	0	15	7	0	60	
Acertos		1,000	0,667	0,857	1,000	#DIV/0!	0,667	1,000	#DIV/0!		Soma Diagonal 49
Erro omissão		0,000	0,333	0,143	0,000	#DIV/0!	0,333	0,000	#DIV/0!		Exat. Global 0,817 81,7 %
											E Produto Col*Lin. 600
											Índice Kappa 0,780 78,0 %

Na análise comparativa entre os anos (Tabela 4), a exatidão global ficou em 26,3% e o índice Kappa em 0,13, indicando, portanto que houve expressiva alteração do uso do solo.

Destaca-se nesta análise a significativa alteração das áreas de laranja para cana-de-açúcar, pasto e eucalipto. Nota-se novamente o problema obtido para a classificação da água, ressaltando-se que para 2015 áreas de ocupação do solo foram classificadas como água.

Os valores elevados na diagonal da matriz referente a cultura do eucalipto, indicam que 69,2% das áreas classificadas em 2001, permaneceram com o mesmo uso.

A cultura da cana-de-açúcar adentrou áreas antes ocupadas por pasto, laranja, solo exposto (classe esta eu já indicava a presença de cana, sendo apenas um estágio da cultura).

Outra cultura com incremento significativo de área foi a do eucalipto, entrando em áreas antes ocupadas por pasto, mata (provável confusão antes identificada aliada a desmate de vegetação nativa) e laranja.

Tabela 4 - Matriz de erro para o comparativo entre os anos de 2001 e 2015.

		2001 - 2015								Soma linhas	Erro de inclusão
		2001									
2015	Pasto	1		1	2		5		3	12	0,917
	Mata	2	1					1	1	5	0,800
	Eucalipto	5	5	9	2		1	4	2	28	0,679
	Cana-de-açúcar	7	1	2	5		3		5	23	0,783
	Solo Palha									0	#DIV/0!
	Solo Exposto	4		1	1					6	1,000
	Água						1	5		6	0,167
	Laranja									0	#DIV/0!
Soma colunas		19	7	13	10	0	10	10	11	80	
Acertos		0,053	0,143	0,692	0,500	#DIV/0!	0,000	0,500	0,000		Soma Diagonal 21
Erro omissão		0,947	0,857	0,308	0,500	#DIV/0!	1,000	0,500	1,000		Exat. Global 0,263 26,3 %
											E Produto Col*Lin. 977
											Índice Kappa 0,130 13,0 %

4. Conclusões

Houve uma substituição total da cultura da laranja por outros usos da terra, fato este explicado pelo surgimento do greening (doença causada por bactérias) na região, informação esta constatada junto aos moradores locais e bibliografias de estudo da região.

O incremento de aproximadamente 12.000 ha de cana-de-açúcar e também de aproximadamente 12.000 ha de eucalipto demonstram a utilização crescente destas culturas na ocupação da paisagem de Analândia e Santa Cruz da Conceição.

A classificação supervisionada por distância de Mahalanobis deve ser utilizada com cautela quando da ocorrência de solos argilosos e com alta umidade, pois possui assinatura espectral próxima a água para este método de classificação.

Constatou-se que a confecção de mapas provenientes de classificação de imagens de sensoriamento remoto traz resultados bastante satisfatórios para o planejamento municipal e regional e traduz numericamente as mudanças ocorridas.

Para a realização de classificação de uso da terra, faz-se necessário a utilização de mais de um classificador.

3. Referências Bibliográficas

ASSAD, E. D.; SANO E. E. **Sistemas de informações geográficas: Aplicações na agricultura**. 2^a ed., Brasília-DF: EMBRAPA-SPI / EMBRAPA-CPAC, 2003.

CROSTA, A. P. **Processamento digital de imagens de sensoriamento remoto**. Campinas: IG/UNICAMP. 1992. 170p.

FREITAS FILHO, M.R.; MEDEIROS, J.S. Análise multitemporal da cobertura vegetal em parte da Chapada do Araripe- CE, utilizando técnicas de Sensoriamento Remoto e Geoprocessamento. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 7, 1993, São José dos Campos. **Anais...** São José dos Campos: INPE, 1993. p.73-80.

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

MAXIMINIANO, G. A. **Bacia do Rio Pato Branco: ensaio cartográfico para análise da fragilidade do meio físico com uso de geoprocessamento**. Dissertação de Mestrado. São Paulo: USP, 1996.

PEREIRA, M. N.; KURKDJIAN, M. L. N. O DE; FORESTI, C. **Cobertura e uso da terra através de Sensoriamento Remoto**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 1989. 118p.

PEREIRA, R. S. Interpretação automática de imagens aplicada ao planejamento rural. In: CONGRESSO FLORESTAL ESTADUAL, 7, 1992, Nova Prata, RS. **Anais...** 2 v. p. 1191-1210.