

Avaliação de métodos de classificação em imagens TM/Landsat e CCD/CBERS para o mapeamento do uso e cobertura da terra na região costeira do extremo sul da Bahia

Caroline Leão¹
Lilian Anne Krug¹
Milton Kampel¹
Leila Maria Garcia Fonseca¹

¹ Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12245-970 - São José dos Campos - SP, Brasil
{caroline, licakrug, milton}@dsr.inpe.br
leila@dpi.inpe.br

Abstract. The aim of this work is to evaluate the accuracy of three classification algorithms used to map land use and land cover in the coastal zone of the state of Bahia using TM/Landsat and CCD/CBERS images from 1996 and 2005. The accuracy test made use of the kappa index and statistics parameters. The use of CCD images showed relatively better statistical results when compared with the use of TM images. The reduction in the kappa index on thematic map regarding to the TM 2005 image was probably caused by the presence of clouds.

Palavras-chave: land use/cover, classification algorithms, multispectral images, uso e cobertura da terra, classificadores, imagens multiespectrais.

1. Introdução

A determinação do uso e cobertura da terra é fundamental na avaliação ambiental de uma região, permitindo um diagnóstico bem circunstanciado e confiável (Espinoza e Abraham, 2005). Estes diagnósticos são especialmente relevantes em áreas de grande importância ambiental e potencial turístico, como é o caso da região costeira do extremo sul da Bahia.

As imagens derivadas de produtos do sensoriamento remoto são excelentes fontes de dados para produzir mapas de uso e cobertura da terra. Estes mapas são gerados a partir de técnicas de processamento de imagens, como a classificação digital. Os produtos obtidos por meio de processamento digital de imagens devem ter sua acurácia avaliada a partir de uma verdade de campo a fim de conferir maior confiabilidade e facilitar o processo de tomada de decisão. Atualmente um dos principais parâmetros que devem ser fornecidos junto aos mapas é a avaliação do erro envolvido no produto (Powell et al., 2004).

O objetivo deste trabalho é avaliar a acurácia de três classificadores implementados no SPRING (Câmara et al., 1996), para a construção de mapas de uso e cobertura da terra com o uso de imagens TM/Landsat e CCD/CBERS. A área de estudo compreende a região costeira dos municípios de Alcobaça, Caravelas e Nova Viçosa, no extremo sul da Bahia. Além disso, o trabalho também visa comparar os resultados obtidos pelas classificações a partir das imagens TM e CCD.

1.1. Área de estudo

A área de estudo compreende a região costeira dos municípios de Alcobaça, Caravelas e Nova Viçosa, localizados no extremo Sul da Bahia, costa leste brasileira (**Figura 1**). Juntamente com os municípios litorâneos de Mucuri e Prado, eles constituem uma região conhecida como Costa das Baleias (Bahia-Brasil, 2006).

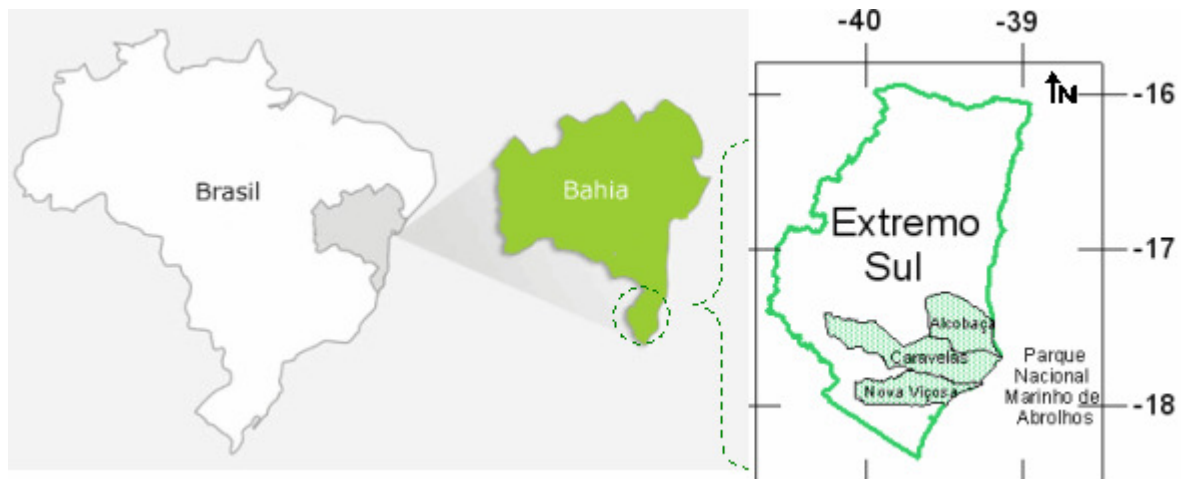


Figura 1: Mapa de localização dos municípios de Alcobaça, Caravelas e Nova Viçosa, no extremo sul da Bahia.

Fonte: Adaptado de FTC (2006).

A área de estudo abrange todo o ecossistema de manguezal em torno dos rios Caravelas, Peruípe e seus afluentes. Tem como principais atividades a pecuária, a agricultura de pequeno porte e a extração de eucalipto para obtenção de celulose. No litoral, destaca-se a atividade turística, sendo o município de Caravelas o principal porto de saída para visitantes do Parque Nacional Marinho de Abrolhos.

2. Materiais e Métodos

As imagens TM/Landsat 5 de 05/08/1996 e 13/07/2005 (órbita/ponto 215/72) e CCD/CBERS-2 de 18/07/2005 (órbita/ponto 148/120) foram obtidas no catálogo de imagens do INPE. A escolha das imagens foi baseada nas datas dos mapas de verdade de campo juntamente com a informação de cobertura de nuvens, devido à presença quase que constante de nebulosidade na região. Os mapas temáticos de 1996 e 2005 utilizados como verdade de campo foram extraídos de Santos (2006).

Foram utilizados os *softwares* SPRING 4.2 para o processamento das imagens, e Microsoft Excel para a avaliação estatística das classificações obtidas.

2.1 Procedimentos metodológicos

As imagens foram registradas, tendo como base imagens ortorretificadas do satélite Landsat-7, sensor ETM+ do ano de 2001, processadas pela *National Aeronautics and Space Administration* (NASA, 2006). Em seguida, foram recortadas para cobrirem somente a área de estudo que compreende este trabalho. A imagem CCD de 18/07/2005 com resolução espacial de 20 m foi reamostrada para 30 m para fins de cruzamento com os levantamentos de campo, cuja resolução era de 30 m. Esta reamostragem permitiu uma comparação adequada entre as classificações da imagem CCD e da imagem TM de 13/07/2005, também com resolução espacial de 30 m.

Para a classificação das imagens foram escolhidos os seguintes algoritmos (Jensen, 1996):
 - Máxima Verossimilhança (MAXVER), classificador supervisionado por pixel, onde amostras previamente selecionadas são utilizadas para o treinamento do classificador. Esta classificação considera a ponderação das distâncias entre médias dos níveis de cinza das classes.
 - *Bhattacharya*, classificador supervisionado por regiões. O Algoritmo utiliza a distância de *Bhattacharya* para medir a separabilidade estatística entre cada par de classe espectral. A

separabilidade é calculada através da distância média entre as distribuições de probabilidades de classes espectrais.

- IsoSeg, classificador não supervisionado por regiões. Trata-se de um algoritmo de agrupamento de dados não-supervisionado, aplicado sobre o conjunto de regiões, que por sua vez são caracterizadas por seus atributos estatísticos de média e matriz de covariância, e também pela área.

Para a classificação por regiões foi necessária a segmentação das imagens a partir de composições coloridas R3 G4 B2. O processo de segmentação de imagens exige a definição das bandas espectrais e dos limiares de similaridade e de área. Esses parâmetros são responsáveis pela qualidade da segmentação que servirá de base para a classificação temática posterior, sendo que neste trabalho foram escolhidos os valores 10 para similaridade e 20 para área de pixel.

Após a segmentação, foram realizados os treinamentos com aquisição de amostras das classes pré-determinadas para a classificação supervisionada *Bhattacharya* e o mapeamento para a classificação não supervisionada IsoSeg.

Para o mapeamento temático foram utilizadas 8 classes derivadas dos mapas utilizados como verdade de campo (Santos, 2006):

Água – corresponde à rede de drenagem e ao oceano;

Floresta secundária em estágio médio-avançado de regeneração – floresta que apresenta formações variando de 10 – 25 m de altura;

Pastagem/agricultura de pequeno porte/áreas abertas – estas áreas foram caracterizadas pela sua uniformidade, campos abertos e altura relativamente baixa (menor que 2m);

Manguezal – ocupam espaços bem delimitados na desembocadura dos rios e rasgados de mar, e apresentam bosques homogêneos, com solos lamacentos;

Restinga – são caracterizadas por apresentarem alternâncias entre linhas de vegetação e faixas de exposição de areia;

Área urbana – compreende áreas de maior densidade populacional, incluindo núcleos urbanos;

Eucalipto – monocultura semiperene, apresenta plantas fibrosas de alto porte, formando bosques homogêneos de grande importância econômica como matéria prima da celulose;

Brejo – são planícies que passam grande parte do ano alagadas por sofrerem influência fluvial ou pluvial, acompanhando os cordões litorâneos;

Nuvem – presente somente no mapeamento para a imagem TM de 2005.

Para avaliar a acurácia dos classificadores na construção dos mapas temáticos foi utilizado o índice Kappa, juntamente com os seguintes parâmetros estatísticos: precisão geral do mapa, precisão de cada classe temática, precisão média das classes, erros de omissão e erros de inclusão (Remote Sensing Resources, 2006).

Os mapas de verdade e os mapas gerados a partir das classificações digitais foram cruzados (tabulação cruzada) e dessa forma, foram geradas matrizes de confusão. Estas matrizes foram exportadas do SPRING 4.2 para o Microsoft Excel, onde foram realizados os cálculos dos parâmetros estatísticos e do índice Kappa.

A precisão geral do mapa (P_{geral}) foi calculada a partir da razão entre o somatório do total de acertos da classificação digital das imagens (concordância entre os dois mapas), PX_{cc} , pelo total de pixels classificados (PX_t):

$$P_{geral} = \frac{\sum PX_{cc}}{PX_t} \quad (1)$$

A precisão de cada classe temática (P_{classe_X}) foi obtida pela razão entre o total de pixels corretamente classificados da respectiva classe (PX_{ccc}) pelo total de pixels da classe (total de pixels da classe obtido a partir do mapa verdade), PX_{tc} :

$$P_{classe_X} = \frac{\sum PX_{ccc}}{PX_{tc}} \quad (2)$$

A precisão média do mapa ($P_{média}$) foi dada pela razão entre o somatório da precisão de cada classe pelo número de classes (n):

$$P_{média} = \frac{\sum P_{classe_X}}{n} \quad (3)$$

Para calcular os erros de omissão ($E_{omissão}$) foi usada a precisão de cada classe:

$$E_{omissão} = 1 - P_{classe_X} \quad (4)$$

Os erros de omissão ocorrem quando determinada feição que deveria ser classificada em dada classe temática é classificada como pertencente à outra classe.

Para calcular o erro de inclusão primeiro foi calculado o erro de exatidão ($E_{exatidão}$), sendo ele a razão entre o total de pixels corretamente classificados na classe x (PX_{tpcc}) pelo total de pixels atribuídos à classe x no mapa a ser testado (PX_{tpce}):

$$E_{exatidão} = \frac{PX_{tpcc}}{PX_{tpce}} \quad (5)$$

Assim, de posse do erro de exatidão, foi calculado o erro de inclusão, que ocorre quando uma feição é classificada como uma classe temática que não corresponde à realidade:

$$E_{inclusão} = 1 - E_{exatidão} \quad (6)$$

O índice Kappa (k), que mede o grau de concordância da classificação digital com a verdade de campo, foi calculado conforme mostrado abaixo, sendo X o elemento considerado.

$$k = \frac{\sum X_{ij} - \sum X_i \cdot \sum X_j}{1 - \sum X_i \cdot X_j} \quad (7)$$

Para avaliar o índice Kappa foi adotada a **tabela 1**, adaptada de Galparsoro e Fernández (2001).

Tabela 1 – Valor para avaliar o grau de concordância a partir do índice Kappa.

Valor do Kappa		Concordância
<	0,20	Pobre
0,21	- 0,40	Fraca
0,41	- 0,60	Moderada
0,61	- 0,80	Boa
0,81	- 1,00	Muito boa

Fonte: Adaptado de Galparsoro e Fernández (2001).

3. Resultados e Discussão

Os valores do índice Kappa para as imagens CCD de 2005 e TM de 2005 e 1996, para os três métodos de classificação, são apresentados na **tabela 2**.

Tabela 2 – Valores de índice Kappa para as classificações referentes às imagens CCD de 2005, TM de 2005 e TM de 1996.

Imagens	Classificadores		
	<i>Bhattacharya</i>	IsoSeg	MAXVER
CCD 2005	0,62	0,48	0,49
TM 2005	0,53	0,46	0,37
TM 1996	0,61	0,49	0,54

Comparando os valores obtidos com o grau de concordância da **tabela 1**, observa-se que os resultados para o índice Kappa foram, em sua grande maioria, de concordância moderada. Os melhores resultados obtidos, tanto para a CCD de 2005 quanto para as TM de 2005 e de 1996, foram com o uso do classificador *Bhattacharya* (**Figuras 2, 3 e 4**), que obtiveram uma concordância moderada para a TM de 2005 e boa para a CCD de 2005 e TM de 1996.

Os índices Kappa para a classificação sobre a imagem TM de 2005 foram relativamente os mais baixos, provavelmente devido à presença de nuvens na referida imagem.

Com base na classificação *Bhattacharya*, que foi a que obteve os melhores resultados de índice Kappa, calculou-se os valores de precisão geral dos mapas sendo de 70% para a classificação referente à CCD de 2005, 63% para a TM de 2005 e 69% para a TM de 1996, respectivamente. Os demais parâmetros estatísticos encontram-se na **tabela 3**.

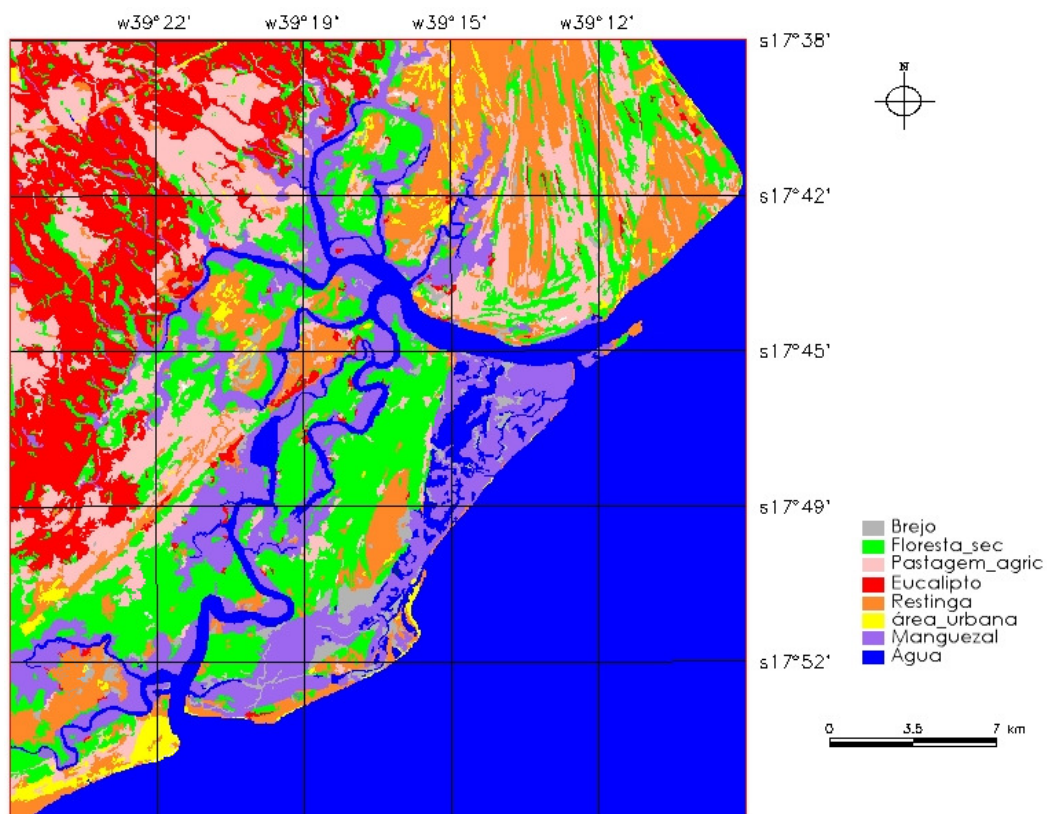


Figura 2: Mapa temático gerado pelo classificador *Bhattacharya* a partir da imagem TM de 05/08/1996.

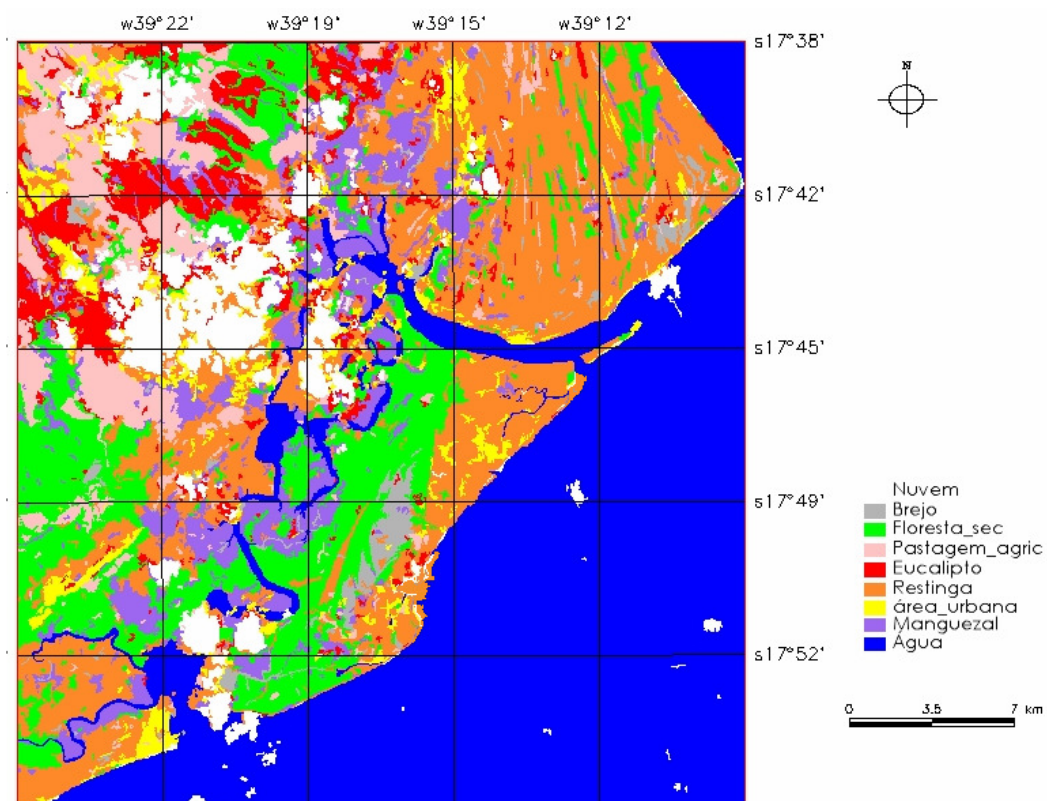


Figura 3: Mapa temático gerado pelo classificador *Bhattacharya* a partir da imagem TM de 13/07/2005.

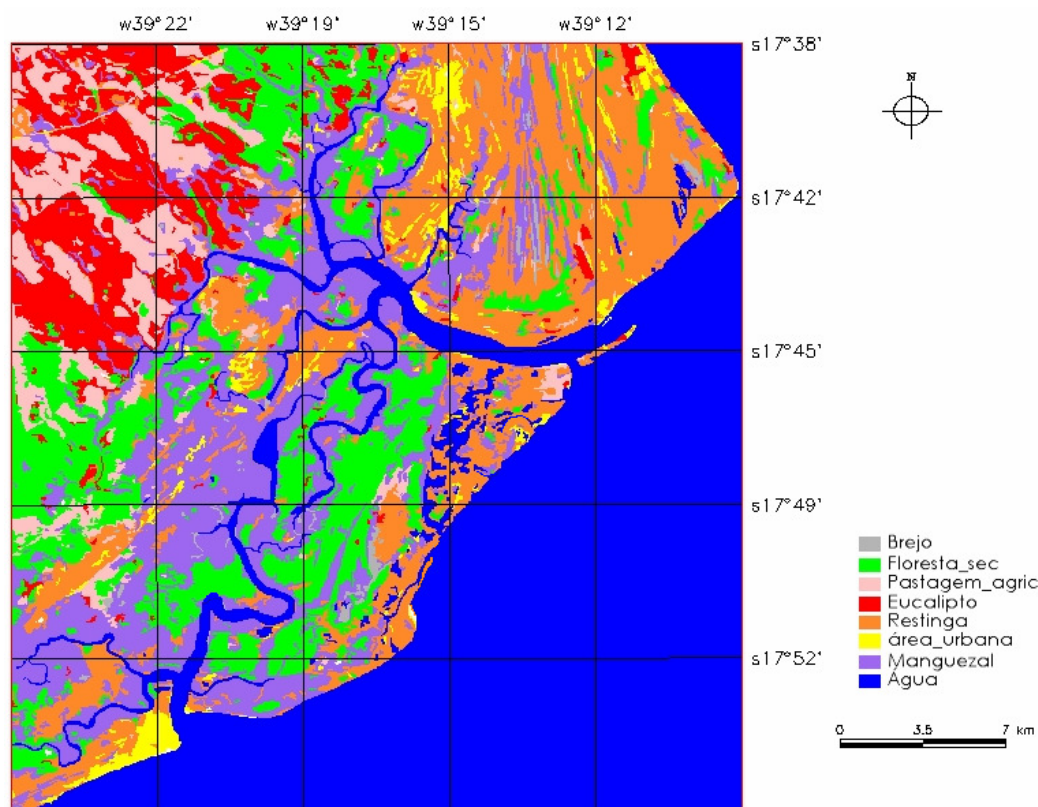


Figura 4: Mapa temático gerado pelo classificador *Bhattacharya* a partir da imagem CCD de 18/07/2005.

Tabela 3 – Parâmetros estatísticos obtidos a partir das classificações *Bhattacharya* referentes às imagens CCD de 2005, TM de 2005 e TM de 1996. Os valores estão em porcentagem (%).

CLASSES	CCD 2005			TM 2005			TM 1996		
	Precisão	Erros de inclusão	Erros de omissão	Precisão	Erros de inclusão	Erros de omissão	Precisão	Erros de inclusão	Erros de omissão
Água	96,40	5,11	3,60	92,45	2,95	7,55	94,31	6,11	5,69
Área_urbana	74,14	80,23	25,86	55,68	92,48	44,32	55,41	82,31	44,59
Restinga	59,40	22,33	40,60	64,50	34,74	35,50	41,53	29,43	58,47
Eucalipto	73,18	53,27	26,82	48,26	57,18	51,74	80,44	37,00	19,56
Pastagem	25,88	20,00	74,12	21,58	39,24	78,42	65,58	74,25	34,42
Floresta_sec	66,70	43,59	33,30	55,67	53,29	44,33	46,54	42,98	53,46
Brejo	33,13	84,92	66,87	60,15	90,87	39,85	51,80	91,31	48,20
Manguezal	77,97	68,52	22,03	31,99	69,69	68,01	59,54	34,28	40,46
Precisão média	63,35			47,81			61,89		

Observa-se que a classificação referente à imagem CCD de 2005 foi a que obteve o melhor resultado dentre as classificações, com a melhor precisão geral (70%), melhor precisão média das classes (63%) e os menores valores médios nos dois tipos de erro, inclusão e omissão. A classe **Água** foi a que obteve os menores erros, tanto de inclusão quanto de omissão, sendo, portanto, a classe menos confundida entre as classificações. As classes **Área_urbana** e **Brejo** foram as que obtiveram os maiores erros de inclusão para as três classificações, provavelmente devido às áreas serem pequenas e por isso poucas amostras foram coletadas relativamente às aquisições dos outros temas. As classes com maiores erros

de omissão foram: **Pastagem** e **Brejo** para a classificação referente à imagem CCD de 2005, **Pastagem** e **Manguezal** para a TM de 2005 e **Floresta_sec** e **Restinga** para a TM de 1996.

4. Conclusões

A avaliação do índice Kappa mostrou que o uso do classificador *Bhattacharya* foi o que obteve o melhor resultado entre as classificações digitais das imagens TM e CDD utilizadas neste trabalho. A classificação temática da imagem CCD de 2005, gerada com o uso deste classificador foi a que apresentou valores mais altos dos parâmetros estatísticos e índice Kappa.

Os resultados obtidos pela classificação da imagem TM de 2005 não foram próximos aos resultados obtidos com a imagem CCD do mesmo ano. A presença de nuvens na imagem TM pode ter ocasionado uma maior confusão entre as classes, influenciando nos resultados. Porém, de modo geral, observou-se que os valores do índice Kappa, da precisão geral do mapa e da precisão média das classes foram similares entre os mapas temáticos referentes às imagens CCD de 2005 e TM de 1996, respectivamente.

Isto pode indicar que as classificações de imagens TM e CCD com resolução espacial similar e mesma composição de bandas podem fornecer resultados semelhantes. Com isso, conclui-se que os produtos gerados a partir de imagens CCD podem simular produtos gerados por imagens TM, desde que pré-processadas e atendendo aos requisitos citados anteriormente.

5. Referências

- Bahia-Brasil. Extremo Sul. Disponível em: <<http://www.bahiabrasil.com.br/>>. Acesso em: 26 set 2006.
- Câmara, G.; Souza, R. C. M.; Freitas, U. M.; Garrido, J. C. P. SPRING: Integrating Remote Sensing and GIS with Object-Oriented Data Modelling. **Computers and Graphics**, v. 20, n. 3, p. 395-403, May-Jun 1996.
- Espinoza, H. F.; Abraham, A. M. Aplicação de técnicas de sensoriamento remoto e geoprocessamento para o estudo dos recursos hídricos em regiões costeiras. In: Simpósio Brasileiro de Sensoriamento Remoto (SBSR), 12. 2005. Goiânia. **Anais...** São José dos Campos: INPE, 2005. Artigos, p. 2487-2494. CD-ROM, On-line. ISBN 85-17-00018-8. Disponível em: <<http://marte.dpi.inpe.br/col/ltid.inpe.br/sbsr/2004/11.21.20.24.27/doc/2487.pdf>>. Acesso em: 15 ago. 2006.
- FTC. Faculdade de Tecnologias e Ciências. Disponível em: <<http://www.ftc.br/Galery.asp>> . Acesso em: 25 set 2006.
- Galparsoro, L. U. Fernández, S. P. **Medidas de concordância: el índice Kappa**. 2001. Disponível em: <<http://www.fisterra.com/mbe/investiga/kappa/kappa.htm>>. Acesso em: 21 jul. 2006.
- Jensen, J. R. **Introductory digital image processing: a remote sensing perspective**. 2ª. ed. New Jersey: Prentice Hall, 316 p. 1996.
- National Aeronautics and Space Administration. NASA. Disponível em: <<https://zulu.ssc.nasa.gov/mrsid/>>. Acesso em: 26 ago. 2006.
- Powell, R. L.; Matzke, N.; Souza Jr. C.; Clark, M.; Numata, I.; Hess, L. L.; Roberts, D. A. Sources of error in accuracy assessment of thematic land-cover maps in the Brazilian Amazon. **Remote Sensing of Environment**, n. 90, p. 221-233. 2004.
- Remote Sensing Resources. Remote Sensing Guides - Image Interpretation and Classification. Disponível em: <http://geospatial.amnh.org/remote_sensing/guides/image_interp/accuracy_assessment.html>. Acesso em: 10 set. 2006.
- Santos, P. S. **Técnicas de sensoriamento remoto voltadas à análise espaço-temporal da cobertura vegetal na região extremo sul da Bahia-Brasil**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais. 21 p. 2006 (No prelo).