

Utilização das imagens Landsat TM nas épocas de verão e inverno para classificação de diferentes formações vegetais.

Carlos Augusto Zangrado Toneli¹
Luis Marcelo Tavares de Carvalho¹
Fausto Weimar Acerbi Junior¹

¹ Universidade Federal de Lavras - UFLA
Caixa Postal 3037 - 37200-000 - Lavras - MG, Brasil
carlos-toneli@plantar.com.br
{passarinho, fausto}@ufla.br,

Abstract. The regions around the Peruaçu river are rich transition areas between Savanna and Seasonal Deciduous Forests. They are located in the San Francisco river basin and they form a mosaic of several different preservation units. This paper aims to describe a methodology for land cover classification using two Landsat TM images acquired in different dates: one image from summer time and the other from winter time. The pre-processing methods used were: geometric correction, radiometric calibration and radiometric correction of the summer image considering the winter image as a reference. The CUBIST algorithm was used in order to characterize the special patterns of different land cover types. The use of a non parametric model through a decision tree algorithm presented 92.4 % of global accuracy and 0.9145 of *Kappa* index for classification of the study area. The status of each protected area unit was reported as well as the classification of different forest formations. It was concluded that the temporal images, from summer and winter, are essential for discrimination between deciduous and semideciduous forest formation.

Palavras-chave: remote sensing, classification image, Savanna and Seasonal Deciduous Forests
sensoriamento remoto, classificação de imagens, Savana e Florestas Estacional Decidual.

1. Introdução

Atualmente, as pressões da sociedade sobre os ecossistemas vêm alterando as paisagens em grande escala. Em Minas Gerais, grande parte da biodiversidade, ainda não conhecida, vem anualmente sendo ameaçada pela forma de exploração descontrolada. Conforme Scolforo & Carvalho (2006), cerca de 1,7% ou 152 mil hectares da vegetação nativa no estado de Minas Gerais foi modificado entre os anos de 2003 a 2005.

De fato, é importante o monitoramento das diferentes paisagens no meio ambiente para tornar possível a análise da conservação destes ecossistemas. Um dos maiores desafios científicos deste século é desenvolver e implementar estratégias para evitar a perda de uma parcela importante da biodiversidade global devido, principalmente, ao aumento acelerado das atividades humanas no planeta (Silva, 2003).

Com os avanços da geotecnologia, em relação aos novos equipamentos de sensores remotos, novas metodologias de processamento de imagens digitais e análise em sistema de informação geográfica, têm-se adquirido maior robustez na análise e monitoramento da cobertura terrestre. As atuais tecnologias possibilitam a compreensão do comportamento na dinâmica da superfície terrestre em diferentes escalas, dando suporte à tomada de decisões mais confiáveis e eficientes.

1.1 Objetivo

- Objetivo principal deste estudo é caracterizar as diferentes formações vegetais dos domínios fitogeográficos presentes nas unidades de conservação do rio Peruaçu, tal como as formações de Savana, Florestas Estacionais Deciduais e suas áreas de transições.
- Objetivo específico desta metodologia é ajustar os parâmetros necessários para a classificação de imagens em épocas diferentes, por meio do algoritmo de árvore de decisão.

2. Metodologia do Trabalho

Para a realização deste trabalho foram utilizadas as imagens de 30 metros de resolução espacial com 6 bandas espectrais do sistema sensor TM a bordo do satélite Landsat 5 nas épocas de inverno 14/9/06 e verão 20/1/07, fornecidos pelo Diretório Geral de Imagens - DGI do INPE. Foram utilizados os softwares de processamento de imagens ENVI 4.3, algoritmo de árvore de decisão CUBIST e o sistemas de informações geográficas ArcGIS 9.2.

2.1 Área de estudo

A área de estudo está localizada na sub-bacia hidrográfica do rio Cochá presente na bacia hidrográfica do rio São Francisco. A Figura 1 mostra os limites dos principais rios que fazem parte da área de estudo: rio São Francisco, Peruaçu, Itacarambi, Japoré, riacho Borrachudo e da cruz.

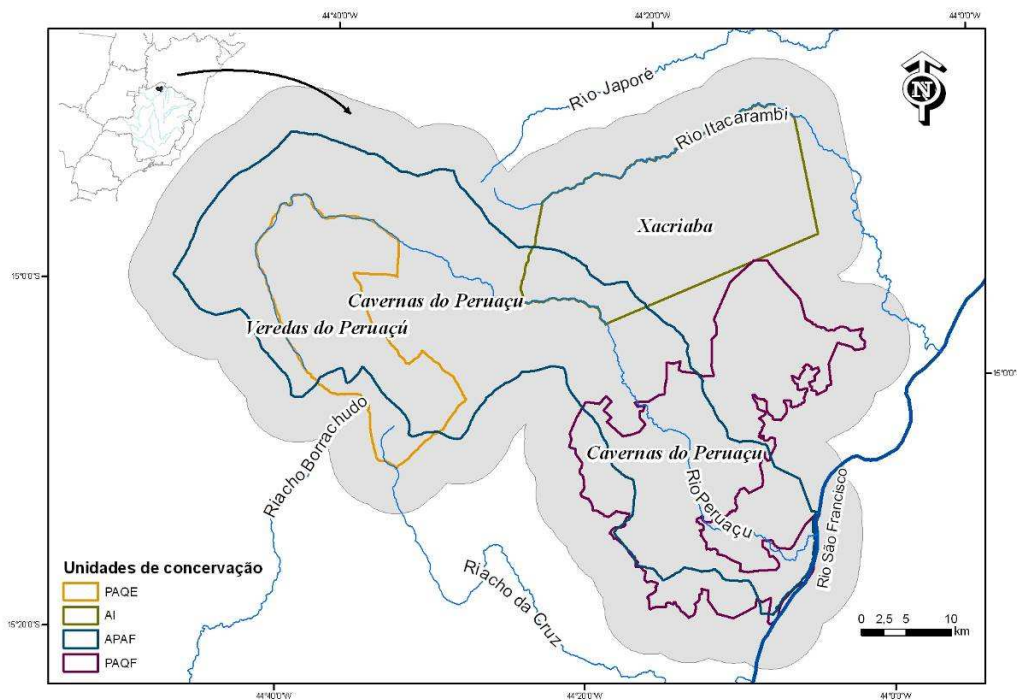


Figura 1. Localização da área de estudo.

Esta região encontra-se em uma zona de transição entre os domínios fitogeográficos de Savana e Florestas Estacionais Deciduais. É dotada de um extenso patrimônio natural, paisagístico, espeleológico e arqueológico. A Figura 2 mostra as dimensões que cada unidade representa em hectares.

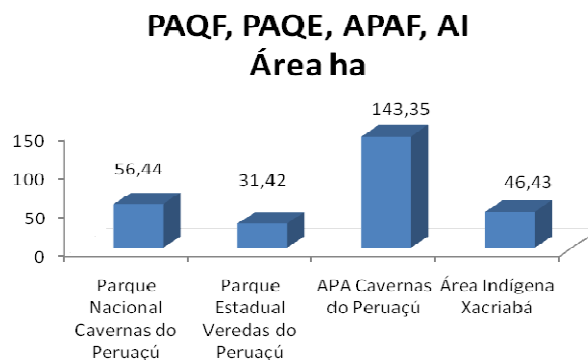


Figura 2. Proporções em hectares para cada unidade de conservação da área de estudo.

2.2 Pré-processamentos

O pré-processamento utilizado nesta metodologia consiste em: (a) correção geométrica das duas imagens, (b) calibração dos valores digitais para valores de reflectância e (c) retificação radiométrica dos valores de reflectância da imagem de verão utilizando a imagem de inverno como referência.

As correções geométricas das imagens Landsat TM, (inverno e verão) foram realizadas com o objetivo de obter a integridade geométrica de um mapa, sem as distorções causadas durante o processo de aquisição das imagens orbitais.

Os métodos utilizados para correções sistemáticas de ambas as imagens, foram o modelo polinomial do segundo grau e a reamostragem bilinear dos valores de cada *pixel*. A avaliação deste procedimento foi realizada utilizando e a raiz do erro quadrático médio (RMSE) das diferenças entre as posições dos pontos de controles da imagem referência e a posição estimada pelo modelo polinomial.

A fim de atenuar os problemas da correlação entre os dados de diferentes satélites, os números digitais são convertidos para valores de radiância. Segundo Markham & Barker (1987) os valores em números digitais das imagens brutas são transformados para valores de radiância a partir de parâmetros orbitais do satélite e de calibração do sensor usando a seguinte equação:

$$L_{\lambda} = L_{\text{MIN}\lambda} + ((L_{\text{MAX}\lambda} - L_{\text{MIN}\lambda}) * \text{QCAL} / \text{QCAL}_{\text{MAX}}) \quad (1)$$

onde:

λ = comprimento de onda; L_{λ} = radiância espectral aparente; $L_{\text{MIN}\lambda}$ = radiância espectral mínima; $L_{\text{MAX}\lambda}$ = radiância espectral máxima; QCAL_{MAX} = número digital máximo; QCAL = número digital.

Após a transformação dos valores digitais em valores de radiância espectral aparente (L_{λ}), esses valores serão utilizados para o cálculo de reflectância aparente, segundo o método proposto por Markham e Barker (1987) através da seguinte equação:

$$\rho_a = \frac{\pi * L_{\lambda} * d^2}{E_{\text{sun}\lambda} * \cos \theta} \quad (2)$$

onde:

ρ_a = reflectância aparente; L_{λ} = radiância espectral ($\text{mW.cm}^{-2}.\text{ster}^{-1}.\mu\text{m}^{-1}$); d = distância sol-terra em unidades astronômicas; $E_{\text{sun}\lambda}$ = irradiância espectral média do sol no topo da atmosfera ($\text{mW.com}^{-2}.\mu\text{m}^{-1}$); θ = ângulo solar zenital.

A utilização de imagens multitemporais em sensoriamento remoto demanda etapas de pré-processamento que normalizem a radiometria devido as diferentes condições climáticas ou atmosféricas, decorrentes dos períodos sazonais. Em virtude da indisponibilidade de informações no momento de aquisição da imagem, Hall et.al.,(1991) propôs a normalização das condições atmosférica utilizando médias de objetos claros e escuros presentes nas diferentes imagens de satélites previamente calibradas para os valores de reflectância, conforme a (equação 3).

$$T_i = m_i * x_i + b_i, \quad (3)$$

onde:

T_i = número digital da imagem retificada; x_i = número digital da imagem original; $m_i = (B_{ri} - D_{ri}) / (B_{si} - D_{si})$; $b_i = (D_{ri} * B_{si} - D_{si} * B_{ri}) / (B_{si} - D_{si})$

e onde:

Bri = média do conjunto de referência clara; Dri = média do conjunto de referência escura;
Bsi = média do conjunto claro a ser retificado; Dsi = média do conjunto escuro a ser retificado; e i = bandas do TM (1 a 5 e 7).

2.2 Processamento de dados digitais

Para o processamento de dados digitais utilizam-se algoritmos que categorizam, rotulam ou classificam os valores dos atributos espectrais em padrões estatísticos ou computacionais. Os algoritmos de árvore de classificação, de maneira geral, são formados por seqüência de procedimentos bem definidos que, quando seguidas, levam a solução de uma determinada classe de problema. Este classificador, quando aplicado aos dados multiespectrais, segundo Tso & Mather (2001), separa cada classe com base no conhecimento das propriedades espectrais e na relação parental entre as classes vizinhas.

Considerado um classificador não paramétrico que se baseia no aprendizado indutivo, onde o algoritmo aprende as regras para separar as classes de acordo com amostras de treinamento (Quilan, 1986). A partir das amostras de treinamento, o algoritmo separa os atributos do conjunto de amostras por regras de decisões binárias em subconjuntos mais homogêneos. Este procedimento é submetido a novas regras, a fim de gerar conjuntos puros determinando as classes finais.

A árvore de decisão é composta de um nó radical (**t1**), conjunto de nós interiores (**t2**), que contem as decisões, ($x_1 \leq 0.7$ e $x_2 \leq 0.5$) das repartições em nós terminais, folhas (**t3**, **t4** e **t5**), os quais contêm as classes temáticas finais exemplificadas na Figura 3.

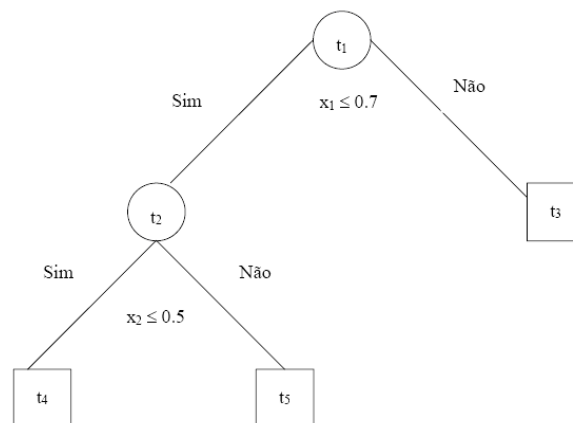


Figura 3. Modelo de árvore de decisão.

A amostragem para o mapeamento das imagens Landsat TM foi realizada por interpretações visuais, utilizando como referência o mapeamento do estado de Minas Gerais (Scolforo & Carvalho, 2006) e adotado o sistema de classificação vegetal proposta por Veloso (1991). As nomenclaturas e legendas associadas ao sistema são: Floresta Estacional Decidual- **FED**; Savana Arborizada -**As**; Pecuária- **Ap**; Contato Savana / Floresta Estacional- **SNc**; Floresta Estacional Semidecidual Aluvial- **Fa**; Formações com influência fluvial e/ou lacustre- **Pab**; Savana Gramíneo Lenhosa- **Sgs**; Agricultura- **Ac**; Afloramento Rochoso- **Afl. Roc.**; Solo Exposto- **Solo exp.** e Água.

Foi necessário o agrupamento das duas imagens em um único conjunto de dados, formando um conjunto temporal com total de 14 bandas espectrais. Foram utilizadas 20% do total de *pixels* da imagem para amostragem do classificador por árvore de classificação. Deste conjunto, metade foi selecionada para amostra de treinamento e a outra metade para amostra

de validação. Durante a checagem da classificação, foram realizadas inspeções de campo para o reconhecimento das classes presentes na região de estudo.

3. Resultados e Discussão

3.1 Correção Geométrica das imagens

Para a imagem de inverno, foram utilizados 58 pontos de controles identificados na imagem referência, com aproximação de 0,67 *pixels* de erro, o equivalente a 20,1 metros em campo. Para imagem verão foram utilizados 60 pontos de controle, com aproximação de 0,78 *pixels* de erro, o equivalente a 23,4 metros de erros em campo.

Os métodos de correções geométricas resultaram em imagens com erros aceitáveis para o mapeamento da cobertura vegetal na escala semidetalhada de 1:50.000, uma vez que os erros aceitáveis nesta escala sejam menores que 25 metros ou 0,83 *pixels* segundo referencia de D'Alge (1999).

3.2 Calibração para valores de reflectância

No intuito de reduzir o volume de dados, os valores de reflectância aparente de 32 bits foram truncados em 0,6375 e posteriormente multiplicados por 400 para produzir dados em 8 bits novamente (Scolforo & Carvalho, 2006). Nos gráficos na Figura 4 é facilmente compreensível a diferença nas bandas espectrais 1 e 5. Contudo, para uma efetiva caracterização dos objetos em diferentes sensores e em diferentes épocas do ano, deve ser considerada a atenuação de efeitos causados pela atmosfera.

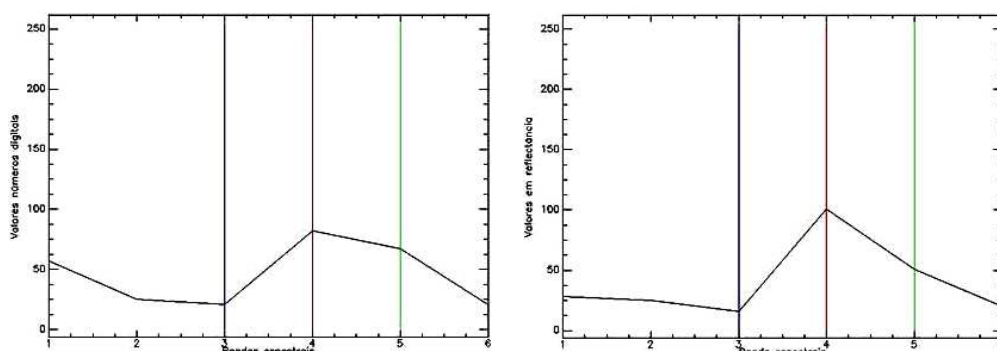


Figura 4. a.) Imagem sem calibração dos números digitais. (b.) mesma área em uma imagem com valores calibrados para reflectância.

3.3 Retificação radiométrica da imagem de verão

Foram ajustadas as regressões lineares aos valores da imagem de verão. A Tabela 1 apresenta os valores dos coeficientes de retificação (m_1 e b_2), os valores descritivos do grau de associação (R^2) entre a imagem referente (inverno) e a imagem estimada, e o erro quadrado médio (EQM) para cada banda espectral.

Tabela 1 – Coeficientes dos modelos lineares para cada banda espectral.

Coeficiente de Retificação	Banda 1	Banda 2	Banda 3	Banda 4	Banda 5	Banda 7
m_1	0.8350	0.9431	1.03803	0.76939	1.02706	0.97217
b_2	4.9517	2.8252	4.94016	9.84455	5.49778	8.44524
R^2	92.33	95.91	95.76	96.80	97.02	95.96
EQM	1.7292	3.7056	5.6126	7.3946	9.0658	8.0437

Estudo desenvolvido por Shimabukuro et.al.(2002) utilizando o método de retificação apresentaram resultados próximos aos coeficientes encontrados neste estudo. Deste modo, a imagem de verão obteve uma boa atenuação dos efeitos da atmosfera e pode ser comparada com a imagem de inverno.

3.4 Acurácia da classificação das imagens Landsat TM

A classificação encontrado na Figura 5, obteve 92,4 % de acurácia global e 0,9145 de coeficiente *Kappa*. A Tabela 2 mostra a matriz de erro utilizando 1000 pontos aleatorizados das amostras de teste para validação da classificação.

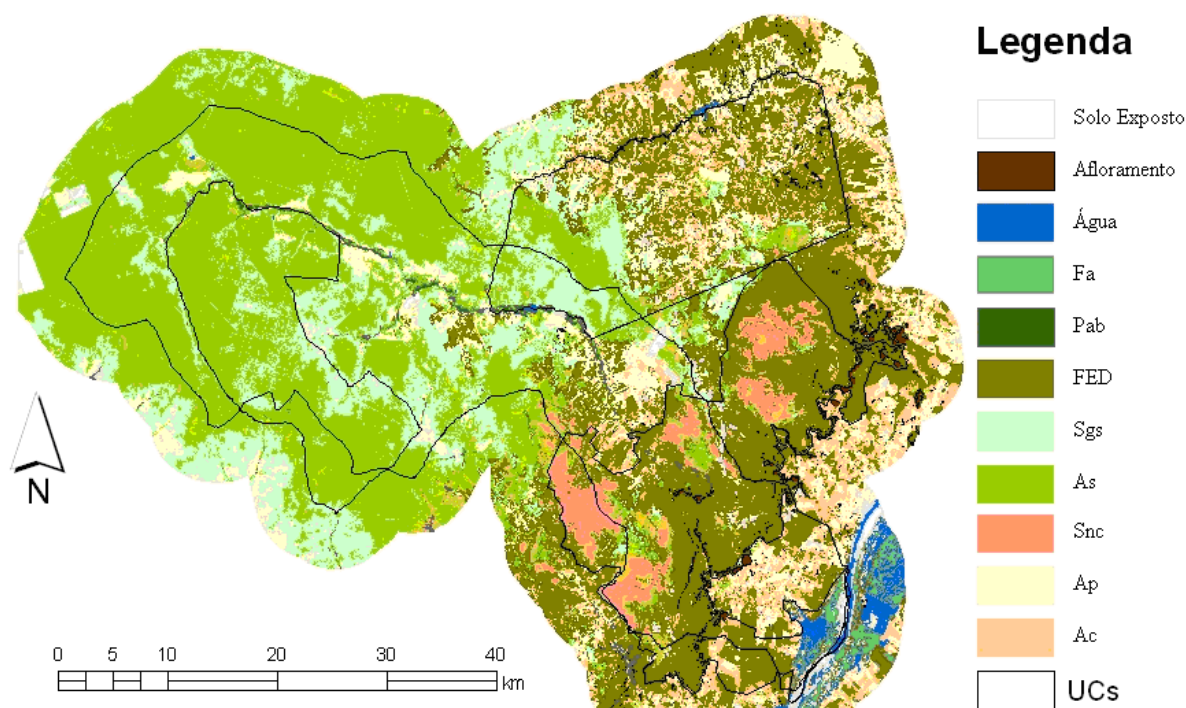


Figura 5. Classificação das imagens Landsat TM verão e inverno.

Tabela 2 - Matriz de erro para classificação das fisionomias vegetais utilizando a imagem TM.

	Água	FED	As	Ap	SNC	Afl. Roc.	Solo exp.	Fa	Pab	Sgs	Ac	Total
Água	64	0	0	0	0	0	0	1	0	0	0	65
FED	0	144	3	2	3	0	0	2	2	1	1	158
As	0	2	169	0	15	2	0	0	8	1	0	197
Ap	0	1	0	92	0	0	4	0	0	10	1	108
SNC	0	0	0	0	56	0	0	0	1	0	0	57
Afl. Roc.	0	0	0	0	0	98	0	0	0	0	0	98
Solo exp.	1	0	0	0	0	0	39	0	0	0	0	40
Fa	0	0	0	0	0	0	0	49	0	0	0	49
Pab	0	0	0	0	0	0	0	0	56	0	0	56
Sgs	0	0	1	7	0	0	0	0	0	93	0	101
Ac	0	2	0	5	0	0	0	0	0	0	64	71
Total	65	149	173	106	74	100	43	52	67	105	66	1000

A Figura 6 apresenta os resultados das proporções das classes da cobertura terrestre mapeadas para cada unidade. O Parque Nacional Cavernas do Peruaçu - **PAQF** apresentou maior proporção de Floresta Estacional Decidual - FED e Encrave - SNc. APA Cavernas do Peruaçu - **APAF** apresentou maior quantidade de área coberta por Savana arborizada - As e FED O Parque Estadual Veredas do Peruaçu - **PAQE** já apresentou predominância de As. A Área Indígena Xacriabá - **AI** apresentou a maioria de sua área coberta por As e FED com expressiva ocorrência de pastagens Ap em relação as unidades de conservação.

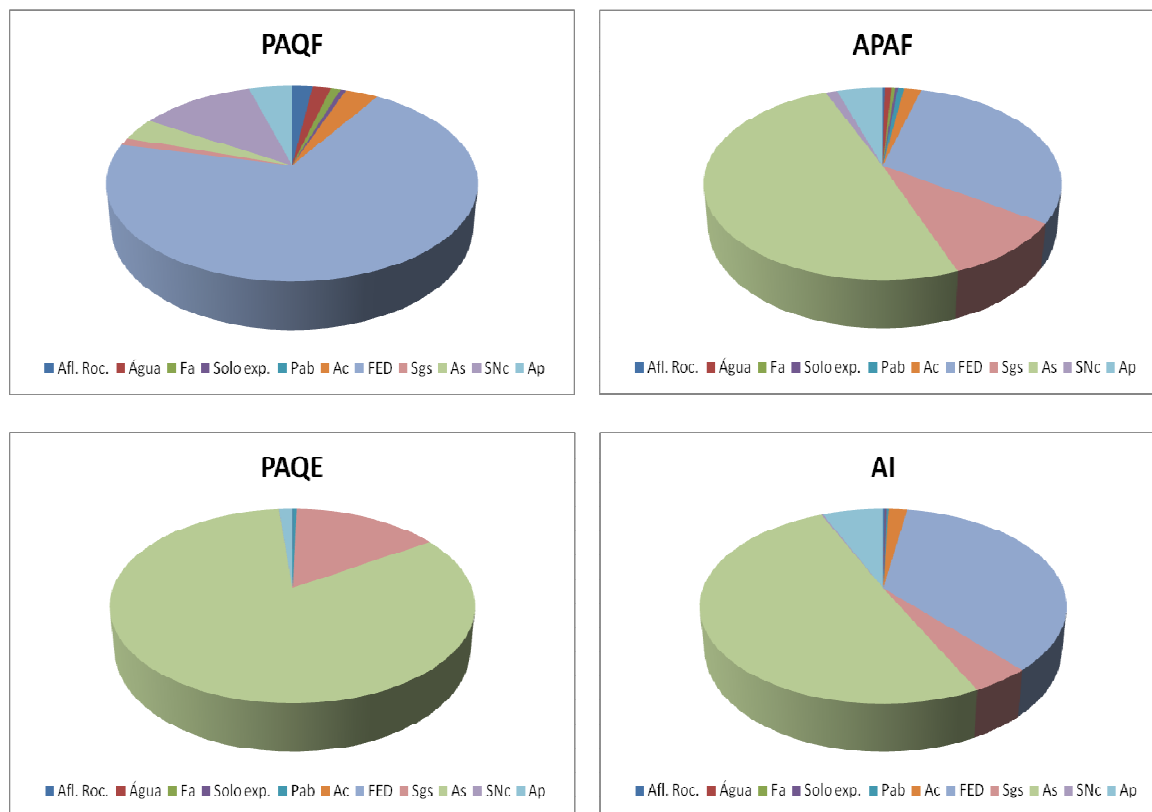


Figura 6. Proporções das classes para cada unidade de conservação.

4. Conclusões

De acordo com o objetivo proposto neste estudo foi concluído o mapeamento da cobertura vegetal utilizando o conceito de classes temáticas para diferenciar os domínios fitogeográficos que ocorrem nas unidades de conservação próximas ao rio Peruaçu.

As aplicações de modelos ambientais que utilizam dados temporais de imagens Landsat TM necessitam de importantes pré-processamentos, principalmente as correções geométricas, calibração e retificação radiométrica. Ainda assim, outros pré-processamentos são necessários, tal como a remoção de nuvens.

As imagens temporais, em diferentes épocas do ano, são essenciais para classificar o padrão de formações florestais deciduais e semideciduais, uma vez que a vegetação apresenta características específicas em determinadas épocas do ano. De fato, quanto melhor a resolução temporal das imagens ao longo do ano, devidamente pré-processada, melhor a condição de caracterização da vegetação utilizando algoritmos de árvore de decisão.

5. Referências Bibliográficas

D'ALGE, J.C.L. **Precisão de registro de imagens digitais**. (Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 1999). Comunicação pessoal.

HALL, F.G.; STREBEL, D.E.; NICKESON, J.E.; GOETZ, S.J. Radiometric rectification: toward a common radiometric response among multitemporal, multisensor images. **In Remote Sensing of Environment**, v.35, p.11-27, 1991.

QUILAN, J. R. Induction of decision trees. **Machine Learning**, Dordrecht, v.1, p. 81-106,1986.

MARKHAM, B.L.; BARKER, J.L. Radiometric properties of U.S. processes Landsat MSS data. **Remote Sensing of Environment**, v.17, p.39-71, 1987.

SILVA, A.de B. **Sistema de informações Geo-referenciadas: conceitos e fundamentos**. Campinas, SP: Unicamp, 2003. 236p.

SCOLFORO, J. R. ; CARVALHO, L. M. T. **Mapeamento e inventário da flora nativa e dos reflorestamentos de Minas Gerais**. Lavras: Universidade Federal de Lavras, 2006. 288 p.

TSO, B.; MATHER, P. M. **Classification Methods for remotely sensed data**. New York: Taylor & Francis, 2001. 332 p.

SHIMABUKURO, Y. E.; NOVO, E. M.; MERTES, L. K. Amazon River mainstem floodplain Landsat TM digital mosaic. **International Journal of Remote Sensing**, v. 23, n. 1, p. 57-69, 2002

VELOSO, H.P., RANGEL-FILHO, A.L.R. & LIMA, J.C.A. **Classificação da vegetação brasileira adaptada a um sistema universal**. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística, 1991. 124p.