

## Uso da técnica de mistura na obtenção dos futuros dados de treinamento a serem empregados no modelo de monitoramento da cobertura terrestre na Amazônia

Marcelo Lopes Latorre<sup>1,2</sup>  
Yosio Edemir Shimabukuro<sup>1</sup>  
Osmar Abílio de Carvalho Júnior<sup>3</sup>

<sup>1</sup>INPE – Instituto Nacional de Pesquisas Espaciais  
Caixa Postal 515 – 12201-970 – São José dos campos – SP, Brasil)  
(yosio, latav) @ltid.inpe.br

<sup>2</sup>CTA–Comando-Geral de Tecnologia da Aeronáutica  
Praça Marechal Eduardo Gomes Nº 50, São José dos Campos, SP, Brasil

<sup>3</sup>Universidade de Brasília – Departamento de Geografia – Campus Universitário Darcy  
Ribeiro - ICC Norte, 70910-900, Brasília, DF, Brasil.

**Abstract:** This work had as objective to establish a preliminary analysis of the viability of training data use, through the mixture technique, in the elaborated model for monitoring the vegetation cover dynamic in Amazonia. This model is based on the technique denominated Vegetation Continuous Fields (VCF), where a larger emphasis is in the integration of different sensor data with different spatial resolutions. As the study area, it was chosen one region in Mato Grosso State, for presenting a great representativity of the Brazilian Amazonia in terms of land cover characteristics. The methodological process elaboration occurred, jointly, with the application of several tests, using different training data (ETM+ and TM scenes) obtained through the same process. Through its preliminary results and its comparison with the data obtained from PRODES Project (year of 2002), the feasibility of the mixture application was verified. It is expected that with this methodology the monitoring of vegetation cover dynamics contribute and complement projects like PRODES.

**Palavras-Chave:**Campos Contínuos de Vegetação, Sistema de Monitoramento, Modelo de Mistura, IDL.

### 1. INTRODUÇÃO

Este trabalho foi motivado pela necessidade de se estabelecer uma análise da viabilidade do uso de dados de treinamento, obtidos através de diferentes técnicas, neste caso a de mistura, em um modelo elaborado para o monitoramento da dinâmica da cobertura terrestre na Amazônia (Latorre et al, 2005).

Este modelo enfatiza a integração de dados de diferentes sensores e resoluções (Terra/MODIS, Landsat5/TM e ETM+ (*Enhanced Thematic Mapper*)), sendo baseado na técnica desenvolvida por Hansen et al. (2002, 2003) e denominada *Vegetation Continuous Fields (VCF)*.

O modelo proposto e desenvolvido apresenta inúmeras fases, podendo ser resumidas da seguinte forma:

1) Criação das variáveis independentes (as métricas) e dependentes (dados de treinamentos gerados a partir de cenas de melhor resolução espacial: ETM+ e TM).

2) Transformação das variáveis em arquivos .TXT para o emprego no programa *S-PLUS*, que realiza uma regressão baseada em árvore.

3) A elaboração de uma função que permita a leitura dos dados gerados pelo programa *S-PLUS* e posterior expansão a toda cena MODIS, obtendo ao final um mapa de cobertura terrestre da região proposta.

Tendo em vista o foco deste trabalho, serão abordadas nas seções posteriores apenas as descrições referentes à obtenção dos dados de treinamento e os resultados obtidos ao

empregá-los, usando para isto da técnica de mistura. A aplicação desta técnica fez parte de uma das etapas de desenvolvimento do processo metodológico que teve como objetivo de verificar a eficiência da aplicação de diferentes sensores e técnicas na elaboração do modelo proposto (Latorre et al, 2005).

O processo é finalizado através de uma análise comparativa entre o melhor resultado obtido pelo modelo e os gerados pelo Projeto PRODES para a mesma área de estudo no ano de 2002.

Cabe salientar que, utilizou-se como área teste o Estado do Mato Grosso. Esse se caracteriza por apresentar uma grande diversidade de tipos de formações vegetais, além de estar sujeita a uma alta taxa na conversão da sua cobertura vegetal, devido à utilização da agricultura mecanizada e pecuária nas áreas de cerrado, bem como o desmatamento e queimadas nas áreas de floresta. Neste contexto, observou-se que a região mostrava-se interessante para os fins a que se destinava este trabalho.

## **2. MATERIAIS E MÉTODOS**

### **2.1. Dados de Sensoriamento Remoto**

Para a realização deste estudo foram utilizadas imagens dos sensores MODIS (500m), LANDSAT5/TM e ou LANDSAT7/ETM<sup>+</sup> referente aos anos de 2002.

### **2.2. “Softwares”**

Neste trabalho foram empregados os *softwares* ENVI 3.6 (*Environment for Visualisation Images*) para o processamento digital de imagens e a linguagem IDL 5.6 (*Iterative Data Language*) para o funcionamento do sistema de monitoramento, complementado pela árvore de decisão (*Software S-PLUS*). Utilizou-se também, o *Software PCI Geomatic* para realizar o registro das imagens de resolução mais fina (TM/ETM+) com as imagens do MODIS (500m) (Latorre et al, 2005).

### **2.3. Metodologia**

O estudo apresentado neste trabalho foi dividido em três etapas distintas: a) apresentação da obtenção dos dados de treinamento empregando-se a técnica de mistura; b) análise preliminar do mapa final obtido a partir destes dados de treinamento; e c) análise comparativa do mapa final gerado a partir dos dados de treinamento (técnica de mistura) com os dados PRODES obtidos, para a região em estudo, no ano de 2002. Para tanto, empregou-se o melhor resultado (definido através de uma análise visual e comparativa – (acurácia global) dos mapas finais gerados) obtido com os dados de 500m do MODIS, considerando os diversos tipos de dados de treinamentos empregados.

Algumas adaptações foram realizadas, uma vez que, os produtos apresentavam projeções, resoluções e metodologias diferenciadas. Assim:

- O primeiro passo foi colocar os resultados (mapas) na mesma projeção, utilizando como referência a definida para os dados MODIS.
- O segundo passo foi definir as mesmas resoluções, uma vez que os dados PRODES estão na resolução de 120m e estes dados MODIS (empregados como base) em 467.63m.
- O terceiro passo foi estabelecer as mesmas classes, uma vez que o modelo proposto define, floresta, não floresta e água / áreas inundáveis; enquanto o PRODES estabelece floresta, não floresta, desmatamento e hidrografia. Para tanto,

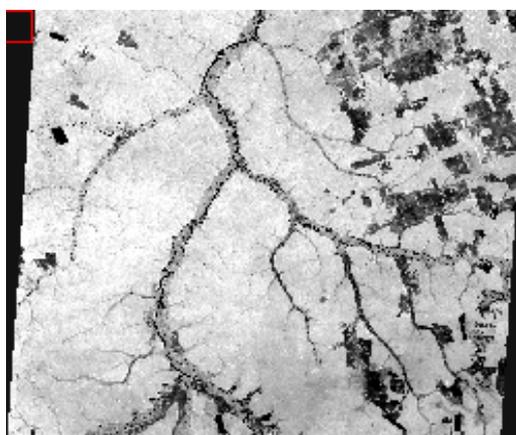
as classes, não floresta e desmatamento foram convertidas para a classe não floresta.

- O quarto passo consistiu em recortar as cenas de maneira correspondente, ou seja, na mesma grade ou área de atuação.
- O último passo foi realizar as comparações em função dos resultados obtidos com o modelo proposto.

### 2.3.1 Obtenção dos dados de treinamento e mapas finais

Os dados de treinamento, de uma maneira geral, são obtidos, inicialmente, por meio da classificação e interpretação de imagens de mais alta resolução considerando áreas mais homogêneas (floresta e não floresta e água / áreas inundáveis ou qualquer outra classe).

Para esta situação, no entanto, a técnica de mistura, substitui o processo anterior utilizando-se das Imagens fração/vegetação (**Figura 1**) dos diferentes sensores, geradas a partir do programa ENVI.

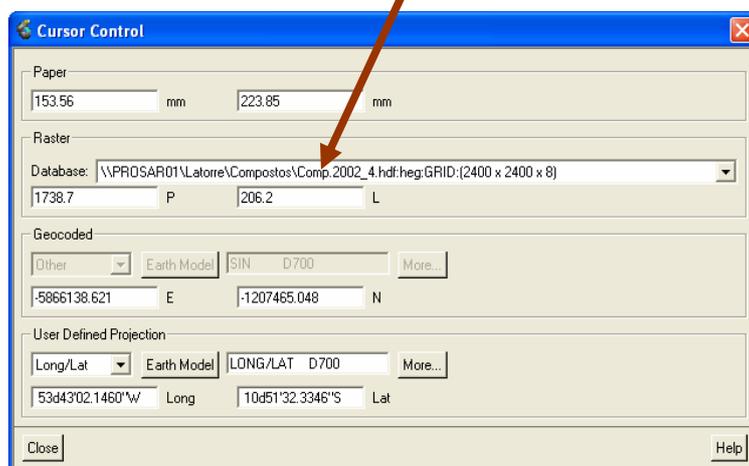
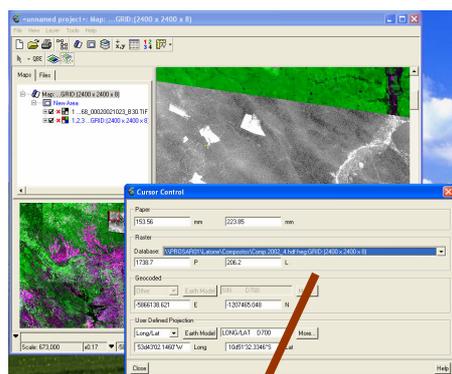


**Figura 1** – Exemplo de cena imagem fração/vegetação da região do Estado do Mato Grosso (Landsat7/ETM+ - ano 2002) obtida no programa ENVI.

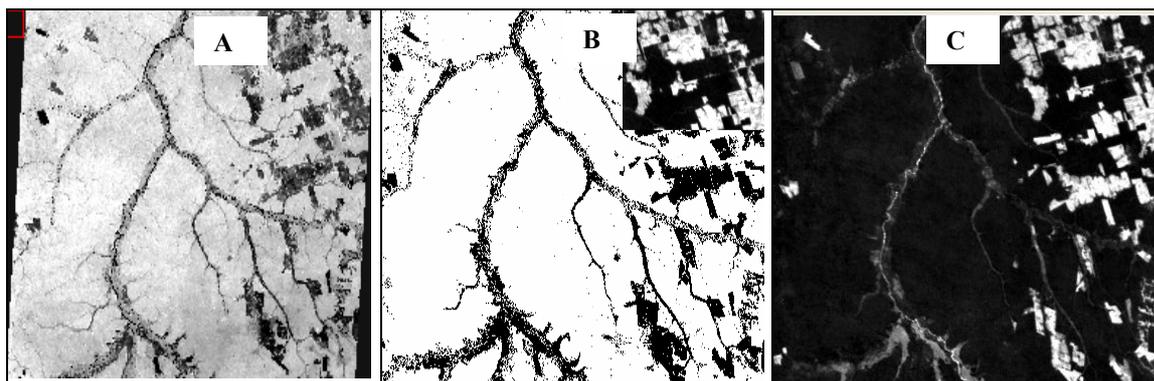
O próximo passo na obtenção dos dados de treinamento é o registro com os dados MODIS. Todo este processo é realizado utilizando-se dos programas ENVI (programa em IDL criado especialmente para esta tarefa) e o PCI 9.1 Geomática.

O programa PCI é utilizado apenas para coletar os pontos de controle necessários ao registro das cenas (**Figura 2**). Em seguida estes pontos são inseridos durante o registro no ENVI.

Nesta fase de registro é, justamente, onde as cenas, obtidas a partir de um modelo de mistura, de melhor resolução são agregadas a escalas mais amplas (dados de 500m), rotulando cada estrato em função dos valores médios definidos no processo (neste caso em valores que variarão de 0 (menor cobertura vegetal) a 1 (maior cobertura vegetal)), sendo em seguida, estabelecido um valor médio sobre as células de resolução mais ampla (**Figura 3**).



**Figura 2** - Exemplo de obtenção de pontos de controle no programa PCI 9.1 para futuro registro dos dados MODIS com as cenas de melhor resolução (ETM+, TM) classificadas. Coleta realizada empregando cenas do Landsat5/TM da região do Estado do Mato Grosso.



**Figura 3** – Exemplo de: (A) Imagem TM (modelo de mistura); (B) dado de treinamento (obtido através do sensor TM) agregado à resolução dos dados MODIS; e (C) produto MODIS da mesma área de atuação.

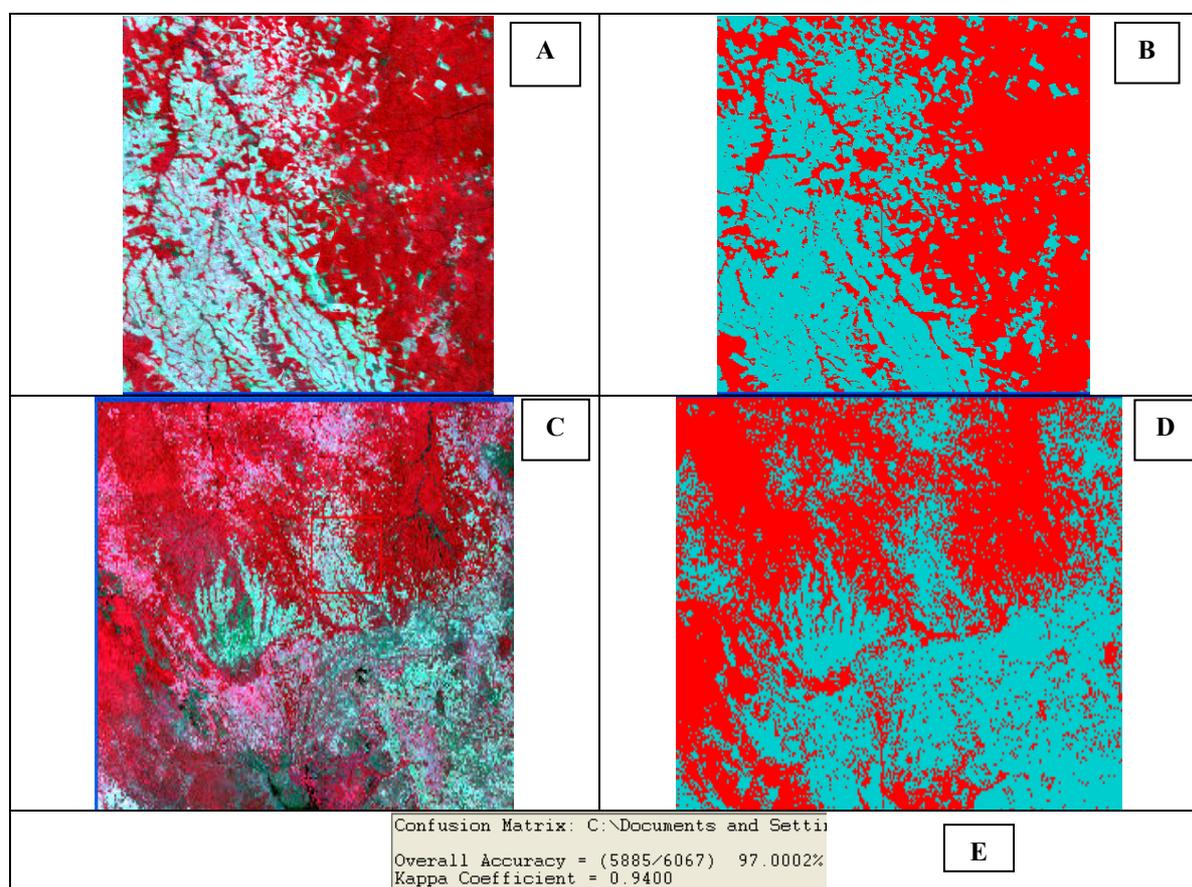
Uma vez definidas todas as variáveis e inseridas no modelo propostos obtêm-se os mapas finais podendo ser obtidos tanto em porcentagem quanto em classes discretas.

### 3. RESULTADOS E DISCUSSÕES

A **Figura 4** apresenta o melhor resultado obtido ao se empregar a técnica de mistura. Observa-se um bom grau de refinamento (separação) entre floresta e não floresta. No entanto, verificaram-se algumas deficiências no quadrante relativo à área do pantanal, causadas, provavelmente, pela alta sazonalidade e complexidade (fisionômica / estrutural) desta área.

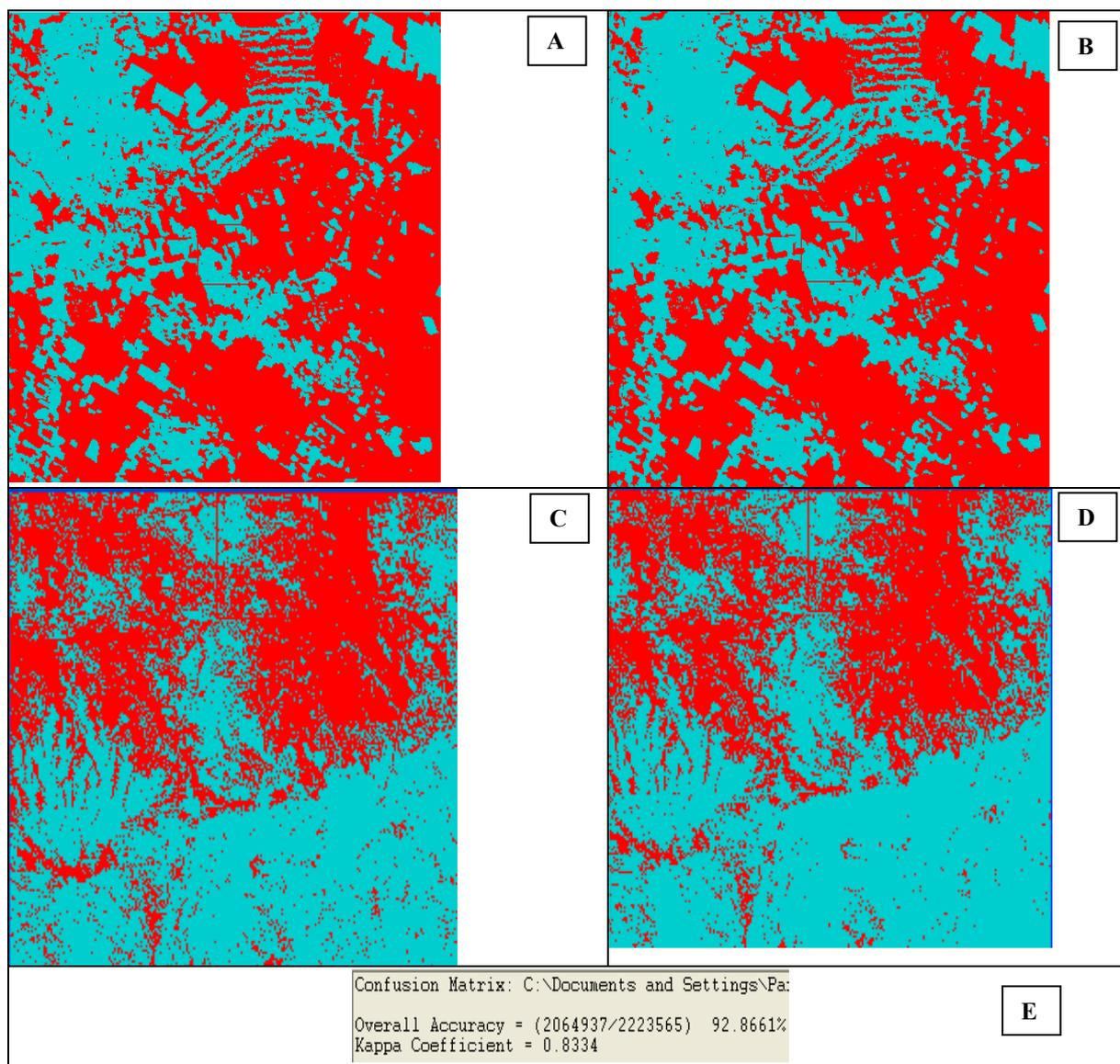
Adicionalmente, observou-se que o emprego do dado de treinamento obtido a partir desta técnica (modelo de mistura - imagem fração / vegetação) não era suficiente para definir com precisão a classe de hidrografia. Desta maneira, optou-se por gerar um produto considerando apenas as classes de floresta e não floresta. Acredita-se que o acréscimo de um dado de treinamento empregando-se a imagem fração / sombra possa minimizar este problema verificado.

Resumindo, apesar do número reduzido de classes e uma provável superestimação do modelo em questão, através desta análise preliminar, verifica-se que o resultado geral do mapa foi bem satisfatório (análise visual e acurácia global / coeficiente kappa) ao se considerar a cena como um todo.



**Figura 4** – Melhor resultado obtido (mapa final) ao se empregar o dado de treinamento – (modelo de mistura). (A) Um exemplo do detalhe empregado para uma análise visual na comparação com os mapas gerados; (B) detalhe do mapa final empregado na análise e com as seguintes classes: vermelho – floresta, ciano - não floresta. (C) Exemplo de métrica (dados 500m - MODIS) composição R (NIR) G (RED) e B (BLUE) da cena H12V10 do Estado do Mato Grosso; e (D) Mapa final gerado da cena toda. (E) Acurácia global e o coeficiente kappa do mapa final.

A **Figura 5** mostra um exemplo de análise comparativa realizada com os dados PRODES. Através desta análise pode-se verificar que os resultados mostraram-se satisfatórios, tendo em vista o grau de precisão apresentado pela estatística kappa. Acredita-se que um estudo mais apurado do emprego destes dados de treinamento permitirá a viabilidade de todo o processo apresentado, podendo, inclusive, auxiliar outros projetos desenvolvidos para a Amazônia e já consagrados, como é o caso do PRODES. Acredita-se ainda que, o seu emprego poderá diminuir o tempo de processamento do modelo em questão, tendo em vista que, para esta situação, não se faz necessário o emprego de classificadores na obtenção dos dados de treinamento do modelo em questão.



**Figura 5** – Exemplo de melhor resultado obtido na comparação com os dados do Projeto PRODES (área do Estado do Mato Grosso – ano de 2002). (A) Detalhes do mapa final obtido com o Projeto PRODES; (B) detalhes do melhor resultado obtido com o modelo proposto (mistura) (classes: floresta e não floresta); e (C)/(D) seus respectivos Mapas finais. (E) Acurácia global e o coeficiente kappa do mapa final obtida empregando como verdade os dados do Projeto PRODES.

#### 4. CONCLUSÕES

A partir da análise e os resultados preliminares apresentados pode-se concluir que apesar do número reduzido de classes, os dados de treinamentos obtidos a partir desta técnica mostraram-se viáveis, necessitando, no entanto, de um estudo mais apurado de sua inserção no modelo proposto, de forma a implementá-lo e permitir gerar resultados similares ao se inserir mais classes, como por exemplo, a de água / áreas inundáveis.

#### 5. REFERÊNCIA BIBLIOGRÁFICA

- Anderson, L.O.; Shimabukuro, Y.E.; Latorre, M.L.; Espírito Santo, F.D.B.; Lima, A.; Duarte, V.; Arai, E.; Defries, R.S.; Morton, D.; Hansen, M.; Jasinsky, E. **Coleta de dados para estudos de cobertura arbórea, mapeamento da cobertura da Terra, desmatamento e avaliação de unidades de conservação/territórios indígenas: metodologias de trabalhos de campo**. São José dos Campos (INPE-12304-NTC/368), 2005, 36p.
- Hansen, M.C.; Defries, R.S.; Townshend, J.R.G.; Sohlberg, R.; Dimiceli, C.; Carroll, M. Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. **Remote Sensing of Environment**, n.83, p.303–319, 2002.
- Hansen M.C.; Defries, R.S.; Townshend, J.R.G.; Carroll, M.; Dimiceli, C. And Sohlberg, A. global percent tree cover at a spatial resolution of 500 meters: first results of the MODIS vegetation continuous fields algorithm, **Earth Interactions**, 7(10):1-15, 2003.
- Hansen, M.C.; De Fries, R.S.; Townshend, J.R.G.; Sohlberg, R. Global land cover classification at 1km spatial resolution using a classification tree approach. **International Journal of Remote Sensing**, 21(6 e7):1331-1364, 2000.
- Latorre, M.L.; Souza, P.E.U; Xavier, A.S.; Shimabukuro, Y.E. O desenvolvimento de sistema de monitoramento da mudança da cobertura terrestre na Amazônia Brasileira através da integração de dados de satélites de resolução fina e ampla. **Anais XII Simpósio Brasileiro de Sensoriamento Remoto**. Goiânia, Brasil, 16-21 Abril 2005, INPE, p. 3085-3092.