

O desenvolvimento de um sistema de monitoramento da mudança da cobertura terrestre na Amazônia Brasileira, através da integração de dados de satélites de resolução fina e ampla

Marcelo Lopes Latorre^{1,2}
Paulo Eduardo Ubaldino de Souza²
Alexandre da Silva Xavier²
Yosio Edemir Shimabukuro¹

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

²IEAv – Instituto de Estudos Avançados, Rodovia dos Tamoios
Km 5,5, 12.228.890 São José dos Campos, SP, Brasil
(souza, asxavier)@ieav.cta.br

Abstract: The motivation of this work has been the growing need to help the SIVAM in one of your missions: monitor the change of the covering of the terrestrial surface (illegal degradation) in repeated intervals for Brazilian Amazon. Now, countless efforts to quantify the monitoring of the changes of earth covering are based on analyses of Landsat satellite data, consuming for this, an expressive number of hours in laboratories to process those data. This work has as general objective to develop efficient, precise methods and you trusted for the monitoring of the Amazon earth covering dynamics. A larger emphasis will be given in the integration of data of different sensor and resolutions, which will be applied as much automated and practical methods as possible. The preliminary results showed that the system developed is feasible.

Keywords: Sensoriamento Remoto, campos contínuos, Amazônia Brasileira, sistema de monitoramento, IDL.

1. Introdução

A poucos meses atrás, a Aeronáutica iniciou as operações das primeiras aeronaves de Sensoriamento Remoto R-99B, entregues ao 2^o/6^o GAv pelo Projeto Sistema de Vigilância da Amazônia (SIVAM). Essas aeronaves incorporam sistemas sensores imageadores, que têm como um dos objetivos atender às inúmeras atribuições do SIVAM, entre elas o monitoramento da cobertura terrestre (desmatamentos/queimadas ilegais) na Amazônia Brasileira.

No entanto, a utilização única destes sensores torna-se inviável, uma vez que os custos para estes levantamentos seriam elevadíssimos devidos, entre outros aspectos, à grande extensão territorial da Amazônia.

Nos últimos anos, muito progresso tem sido feito no monitoramento da mudança da cobertura terrestre na Amazônia com dados de satélites. Os primeiros esforços foram realizados com os dados do Landsat/TM, embora alguns experimentos tenham sido obtidos com outros tipos de sistemas: satélite SPOT e o Radar de Abertura Sintética.

Embora o imageamento pelos sensores de resolução espacial mais fina tenha sido usado com uma certa frequência e sucesso para mapear e monitorar a Amazônia, sua aplicação tem-se mostrado limitada a estudos em uma escala local, devido aos seguintes fatores: a não obtenção freqüente de imagens livres de nuvens, além do alto custo e o volume de dados gerados.

Deste modo, o mapeamento de florestas em escala regional e global tem-se realizado com o emprego de sistemas de resolução espacial mais ampla, como é o caso do NOAA/AVHRR e mais recentemente, o sensor MODIS.

Esses sistemas são usados para diversas aplicações do meio ambiente devido ao fato de adquirirem imagens de extensas regiões a custos relativamente baixos, além de sua alta resolução temporal. Sua capacidade de revisita estabelece, também, uma elevada

probabilidade de um imageamento livre de cobertura de nuvens, possibilitando assim, um estudo mais apurado da dinâmica de coberturas florestais em escalas regionais e globais.

No entanto, alguns problemas têm sido associados a estes tipos de sistemas, onde, um dos maiores, está no fato da resolução espacial mais ampla dificultar a observação local de um detalhe, além de que, para este tipo de imageamento, cada “pixel” irá representar uma área, onde, em geral, venham a conter duas ou mais classes de cobertura terrestre. Como conseqüência, as estimativas de florestas baseadas nas classificações de extensas coberturas florestais, a partir destes tipos de sistemas, costumam ser errôneas.

Assim, nota-se que, ambos os produtos gerados por estes sensores (resolução ampla e fina) apresentam deficiências, onde se verifica que a aplicação isolada destes na obtenção do monitoramento da cobertura terrestre em uma escala regional e global, não tem sido a solução mais adequada.

Deste modo, este trabalho tem como objetivo geral desenvolver métodos mais eficientes, precisos e confiáveis para o monitoramento da dinâmica da cobertura da terra na Amazônia, onde uma maior ênfase será dada na integração de dados de diferentes sensores e resoluções (MODIS, LANDSAT/TM e ETM+, CBERS/CCD e HSS (num futuro próximo)).

Com o uso desta técnica, espera-se que o monitoramento repetitivo da dinâmica da cobertura da Terra seja mais tratável, auxiliando, em especial, a missão de monitoramento do SIVAM, possibilitando com isto uma economia considerável para a Força Aérea Brasileira. Espera-se ainda, validar, ou até mesmo desenvolver técnicas mais eficientes através de métodos, os quais serão os mais automatizados e livres possíveis da necessidade da interpretação humana e das inúmeras horas de processamento destes dados.

2. Sensor MODIS

O comitê espacial Norte Americano mantém um programa de longa duração de pesquisas de observação da superfície terrestre, oceanos e atmosfera, e suas interações, incluídas nas medições do “Earth Observing System” (EOS), financiado pelo programa da NASA “Earth Science Enterprises” (ESE).

O objetivo principal da ESE é determinar como a Terra está mudando e quais as conseqüências para a vida neste planeta, desenvolvendo um entendimento de seu funcionamento como um sistema único e interligado (Justice *et al.*, 2002).

Neste projeto (EOS), foram e estão previstos os lançamentos de diversos satélites, sendo o sensor MODIS (“Moderate Resolution Imaging Spectroradiometer”) seu instrumento chave.

O MODIS é o principal sensor desenvolvido para estes satélites, e apresenta, segundo Barker *et al.* (1992) inúmeras características, entre elas:

- Ampla cobertura espacial e espectral (**Tabela 1**);e
- É a primeira ferramenta dos satélites EOS na condução das pesquisas de mudanças globais.

As 36 bandas espectrais deste sensor estão localizadas em função de um comprimento de onda definido para a observação de feições das propriedades das nuvens e sua dinâmica, as propriedades da vegetação na cobertura terrestre e a temperatura da superfície dos oceanos no mundo (Salomonson e Toll, 1990).

O MODIS tende a aperfeiçoar sistemas como o AVHRR, uma vez que, possui a capacidade de rastrear um amplo conjunto de sinais vitais da Terra. Como exemplo, cita-se a capacidade deste sensor em medir o percentual da superfície do planeta coberto por nuvens em quase todos os dias. Esta ampla cobertura espacial irá permitir ao sensor, juntamente com o “Multi-angle Imaging Spectroradiometer” (MISR) e o “Cloud and Earth’s Radiant Energy System” (CERES), determinar o impacto das nuvens e aerossóis no balanço energético da Terra (Strahler *et al.*, 1999).

Tabela 1 – Especificações técnicas do sensor Modis

Órbita	705 Km, síncrona com o Sol, polar; 10:30 a.m descendente; 1:30 p.m ascendente.
Alcance Espectral	0.4-14.4 μm
Cobertura Espectral	$\pm 55^\circ$, 2330 km em fileira (“scans” contínuos em nadir no equador)
Resolução Espacial	250 m (2 bandas), 500 m (5 bandas), 1000 m (29 bandas) em nadir
Precisão Radiométrica	5% Absoluto, $<3 \mu\text{m}$; 1% Absoluto, $>3 \mu\text{m}$; 2% de reflectância
Cobertura de repetição	Diária, a norte da latitude 30° e a cada dois dias, para latitudes inferiores a 30°
Quantização	12 bits
Taxa de Dados	6.2 Mbps (média), 10.8 Mbps (dia), 2.5 Mbps (noite)
Força	162.5 W (média para uma órbita), 168.5 W (pico)

Adaptada de Barker *et al.* (1992, p. 3-4) e Justice *et al.* (2002, p.4).

3. Árvore de decisão

Amplamente utilizadas em algoritmos de classificação/regressão, as árvores de decisão são representações simples do conhecimento e um meio eficiente de construir classificadores que predizem classes baseadas nos valores de atributos de um conjunto de dados.

Uma árvore de decisão tem a função de dividir um conjunto de treinamento, até que cada subconjunto obtido deste particionamento, contenha casos de uma única classe. Para atingir esta meta, esta técnica examina e compara a distribuição de classes durante a construção da árvore. Os resultados obtidos são dados organizados de maneira compacta, que são utilizados para classificar novos casos (Friedl e Brodley, 1997).

Uma árvore é composta de um nó raiz (formado a partir de todos os dados), um conjunto de nós internos denominados partições (“Split”) e de nós terminais chamados de folhas (“leaves”).

Mais comumente, a divisão definida em cada nó interno de uma árvore é estimada a partir de dados de treinamento, usando procedimentos estatísticos. A técnica específica usada para este trabalho é chamada de “algoritmos de aprendizado” (“Learning Algorithms”). Estes requerem dados de treinamento de alta qualidade, onde as relações entre as feições e as classes existentes dentro dos dados sejam assimiladas.

O passo chave em qualquer problema de estimação de uma árvore está em corrigi-la de uma possível superestimação, através da “poda” da mesma. Convencionalmente, uma árvore é crescente até que todas as observações de treinamento sejam corretamente classificadas.

Inúmeras técnicas de construção de árvores têm sido desenvolvidas nos últimos trinta anos, sendo, o algoritmo ID3 um dos primeiros algoritmos de árvore de decisão. Tem sua elaboração baseada em sistemas de inferência e conceitos de aprendizagem. Logo após, foram elaborados diversos algoritmos, sendo os mais conhecidos: CART (“Classification And Regression Tree”) e C4.5/5.0 (Quilan, 1993; Friedl *et al.*, 1999).

Para este trabalho está se utilizando o algoritmo C4.5/5.0 empregado no “software” estatístico denominado S-Plus, em função de algumas de suas características: flexibilidade, simplicidade intuitiva e eficiência computacional, e também, por possuir uma elevada aceitação nos problemas envolvendo a classificação de cobertura terrestre, em especial para escalas regionais e globais (Strahler *et al.*, 1999).

3. Área de Estudo

A área de estudo será constituída por toda a Amazonia Legal, onde, como área teste está sendo utilizado o Estado do Mato Grosso. A área escolhida foi em função da qualidade das

imagens (desta área) do sensor MODIS obtidas durante este ano de 2001 a 2003 e por já se encontrarem disponíveis para a execução do trabalho.

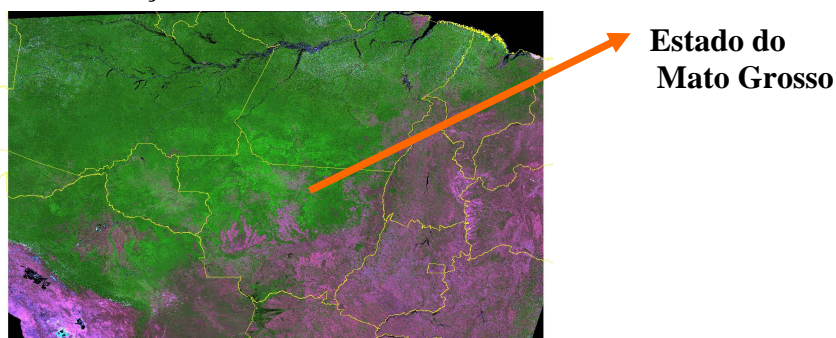


Fig.1- Exemplo de produtos gerados pelo sensor MODIS sobre a Amazonia Brasileira. Mosaico da Amazônia Legal gerado pelo sensor MODIS do mês de agosto de 2002 (Composição: MIR (R), NIR (G) e RED (B)).

3. Materiais e Métodos

3.1. Dados de Sensoriamento Remoto

Para a realização deste estudo estão sendo utilizadas imagens dos sensores MODIS, LANDSAT5/TM e ou LANDSAT7/ETM⁺ correspondente às bandas 1 a 5 e 7 referente aos anos de 2001 a 2003 e CBERS/CCD do ano de 2004.

3.2. “Softwares”

Estão sendo utilizados os “softwares” ENVI 3.6 (“Environment for Visualisation Images”) para o processamento digital de imagens e a linguagem IDL (“Interative Data Language”) 5.6 para a elaboração das funções que geraram o 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⁺ e CBERS/CCD) com as imagens do MODIS, tanto de 250m quanto de 500m.

3.3. Metodologia

O estudo apresentado neste trabalho pode ser dividido em duas etapas consecutivas distintas: a) a análise exploratória da técnica de campos contínuos descrita pelos trabalhos de DeFries *et al.* (1997, 2000) e Hansen *et al.* (2002), a ser utilizada como parâmetro inicial para o desenvolvimento de um sistema semelhante para a Amazônia Brasileira; e b) a elaboração das funções (IDL 5.6) que permitirão realizar o monitoramento da cobertura terrestre da Amazônia, integrando dados de diferentes resoluções.

3.3.1. Análise exploratória do produto CVF (Campos Contínuos)

Nesta etapa foi realizada uma análise exploratória dos métodos desenvolvidos na obtenção deste produto (DeFries *et al.*, 1997, 2000 e Hansen *et al.*, 2002). O estudo visa verificar as evoluções, bem como, as limitações deste produto nos últimos anos, dando ênfase ao trabalho mais recente e obtido por Hansen *et al.* (2002).

Neste trabalho a floresta é descrita como porcentagem de cobertura arbórea, ao invés de floresta ou qualquer outro rótulo existente. São obtidos através de um procedimento automatizado para gerar uma regressão (Venables e Ripley, 1994) com dados de treinamento do Satélite Landsat, entre outros, coincidentes com os dados de resolução ampla.

As entradas para este modelo de regressão são as medidas multitemporais obtidas da reflectância no visível e infravermelho próximo dos alvos, bem como, das bandas que caracterizam a fenologia e as assinaturas espectrais de seus “endmembers”.

Segundo Hansen *et al.* (2002), a principal vantagem deste produto está no fato de que através da descrição de cada “pixel” como uma porcentagem de cobertura arbórea, áreas de heterogeneidade são mais bem representadas, possibilitando assim, uma descrição mais detalhada destas variabilidades espacialmente complexas, as quais, muitas vezes, ocorrem por causa de mudanças antropogênicas da cobertura terrestre.

Desta forma, utilizando-se de estimativas proporcionais, a cobertura ao nível de “subpixel” pode ser mapeada com o objetivo de se medir as mudanças com o tempo. Considerando ainda, que grande parte das mudanças da cobertura terrestre, induzida pelo homem, seja inferior a 250 m, acredita-se que este produto será útil para se criar um sistema de alerta anual, a nível global, das mudanças ocorridas na cobertura terrestre.

3.3.2 Desenvolvimento do sistema de monitoramento

A etapa seguinte consistiu no desenvolvimento de um sistema semelhante para a Amazônia Brasileira, utilizando-se da linguagem IDL. Para um maior entendimento desta fase, serão descritos, resumidamente, os passos criados para se chegar ao resultado proposto.

Basicamente a obtenção do produto pode ser resumida em três fases: 1) criação das variáveis independentes (as métricas) e dependentes (dados de treinamentos gerados a partir de dados de resolução mais fina como: ETM+, TM e CBERS/CCD); b) transformação das variáveis em arquivos .txt para o emprego do “Software” S-PLUS, utilizado para treinar a árvore de regressão; e c) a elaboração de uma função que permitirá a leitura dos dados gerados pelo “Software” S-PLUS, servindo como um classificador automatizado, aplicado a toda a cena MODIS, gerando ao final um mapa de cobertura terrestre da região proposta. Este mapa irá variar em função do número de nós finais (prováveis classes) criados na fase de aprendizado da árvore de regressão, utilizando-se dos dados de treinamentos.

3.3.2.1 Obtenção das métricas/dados de treinamentos

As métricas multitemporais capturam alguns pontos da variação fenológica através dos cálculos de médias (mínimas e máximas) anuais/semestrais/quadrimestrais e as amplitudes das informações espectrais.

Diversos tipos de métricas para a descrição da fenologia da vegetação têm sido sugeridos. Muitos autores têm obtido métricas ligadas a eventos fenológicos, baseados em uma série temporal de dados NDVI, incluindo taxa de senescência, tempo de pico de NDVI, máximo NDVI, entre outros.

Neste trabalho, para os dados de 500m, por exemplo, as sete bandas referentes à cobertura terrestre foram usadas como entrada: azul (459-479nm), verde (545-565 nm), vermelho (620-670 nm) e infravermelho próximo (841-876 nm) e médio (1230-1250 nm, 1628-1652 nm e 2105-2155 nm). Neste caso, dados compostos do MODIS foram transformados em métricas anuais (68 métricas (médias mínimas e máximas) obtidas a partir destes canais e do NDVI), as quais foram usadas como dados de entrada (variáveis independentes) para se estimar a porcentagem da cobertura arbórea (porcentagem definida em função dos diferentes tipos de classes selecionadas). Segue abaixo um exemplo de métrica gerada nesta fase (**Figura 2**).

A métrica da **Figura 2** foi obtida através da média das três primeiras composições relativas ao ano de 2002 da região do Estado do Mato Grosso, que são cenas compostas de 32 dias, produzidas de modo a minimizar os efeitos atmosféricos, como a eliminação de nuvens.

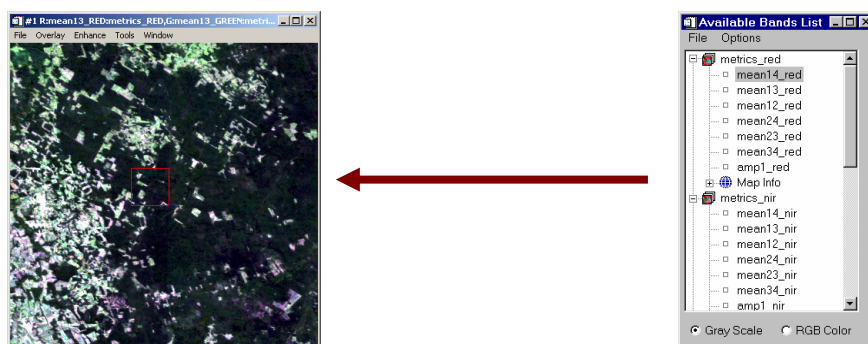


Fig.2 - Métrica gerada da região do Estado do Mato grosso. A composição criada foi a R (média13 (vermelho), G (média 13verde) e B (média 13azul)).

Observou-se que através das métricas obtidas, como a deste exemplo, conseguia-se estabelecer uma melhor distinção entre as áreas de menor (solo exposto) e maior cobertura arbórea (Floresta-parte escura da cena).

Cabe salientar que, o número de métricas é variável, dependendo do tipo de resultado que o usuário deseja e da região que está atuando. Quanto maior o número de métricas, maior a possibilidade de se gerar um produto mais refinado (maior número de variáveis para se treinar a árvore), com um número maior de classes. No entanto, se o usuário quiser apenas duas classes, por exemplo, floresta e não floresta, não há a necessidade de se criar um número tão elevado de métricas. O mesmo procedimento está sendo realizado para os dados de 250m do MODIS.

O próximo passo, ainda dentro desta etapa é a criação dos dados de treinamento. Estes conjuntos de dados são obtidos através da classificação e interpretação de imagens de resolução mais fina, de modo a identificar áreas mais homogêneas (Floresta e não Floresta ou qualquer outro tipo de classe). Nesta técnica, as classificações de resolução mais fina são agregadas a escalas mais amplas (**Figura 3**) (dados de 500 ou 250m), rotulando cada estrato em função dos valores médios de cobertura arbórea (0%, 25%, 50% e 80%); sendo em seguida, estabelecendo um valor médio sobre as células de resolução mais ampla.

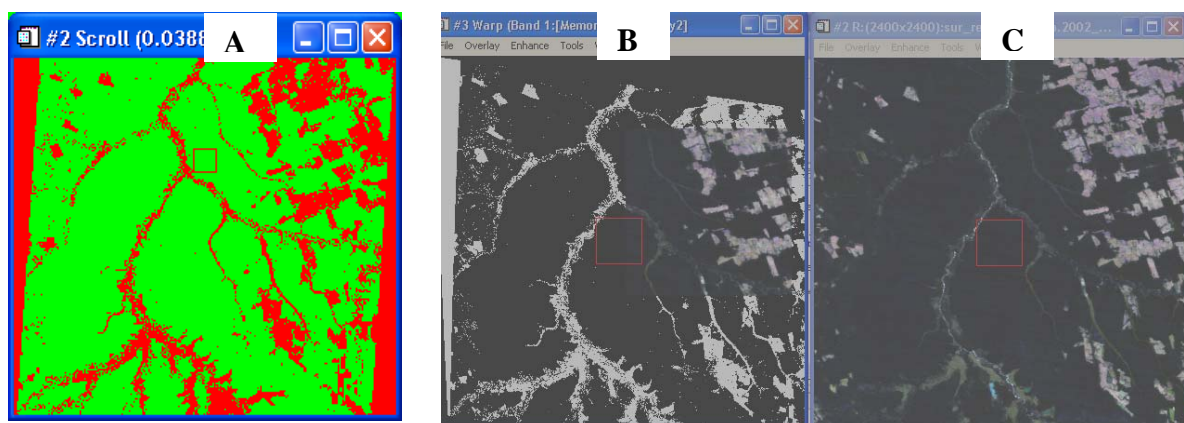


Fig.3 Imagem TM Classificada (dados TM da região – verde (floresta); vermelho (não floresta)) (A); Dado de treinamento (TM) (máscara) registrado e agregado (B) a resolução do produto MODIS (C).

3.3.2.2 “Software” S-PLUS

Uma vez gerados as variáveis independentes (métricas) e dependentes (dados de treinamentos), o próximo passo será o registro, recorte das variáveis para se gerar uma mesma

grade para os dados e conversão dos mesmos em arquivos .TXT. Sequencialmente, estes arquivos serão empregados no treinamento da árvore de regressão existente no “software” S-PLUS.

Este “software” emprega este conjunto de variáveis independentes (métricas) para recursivamente dividir uma variável dependente, ou seja, em subconjuntos, os quais maximizam a redução na soma dos quadrados residuais.

Neste trabalho, duas amostras de dados foram definidas a partir de um conjunto de dados de treinamento. Uma usada para desenvolver a árvore e a outra para “podá-la”. Esta “poda”, se faz necessária, porque este tipo de algoritmo é muito robusto, possuindo a capacidade de delinear um simples “pixel” em um espaço espectral.

Todas as variáveis utilizadas como dados de entrada são analisadas por um valor de número digital, onde divisões à esquerda e à direita são examinadas. A divisão que produz a maior redução na soma dos quadrados residuais, ou desvio, é usada para dividir os dados e o processo começa novamente, para os dois novos subconjuntos criados.

Este algoritmo é descrito, resumidamente, da seguinte forma:

$$\mathbf{D} = \mathbf{D}_s + \mathbf{D}_t + \mathbf{D}_u \quad (1)$$

onde:

s representa o nó “parente”; t e u são as divisões a partir de s .

3.3.2.3 Criação do mapa final

A etapa final está relacionada à conversão dos dados gerados pela árvore para se estabelecer à classificação da região em estudo. Cabe salientar que, neste momento o usuário terá em mãos apenas o resultado gerado, em .TXT, do treinamento realizado com a árvore de regressão. Assim, o resultado obtido pela árvore é expandido, através de uma função definida pelo IDL, para toda a cena proposta, obtendo-se o resultado abaixo (**Figura 4**). Em função do mapa gerado, o usuário poderá trabalhar melhor a árvore de modo a refinar ou não a sua classificação final.

Todas as etapas citadas anteriormente levam, uma vez definido os dados de treinamentos (classificação, registro) a serem utilizados, em média 2:00 hs para se processar o produto final, considerando o usuário possuidor de uma máquina com uma capacidade de processamento razoável.

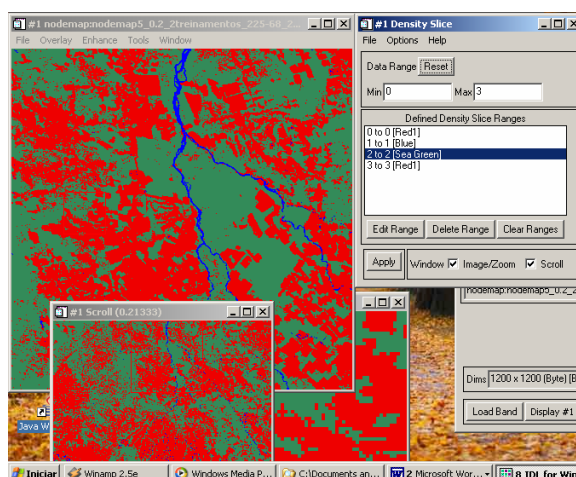


Fig.4 - Mapa final da região do Mato grosso, utilizando-se de dados MODIS 500m para o ano de 2002, contendo três classes: verde - floresta, vermelho - não floresta e azul - água.

Cabe ressaltar que, ainda está em andamento uma fase final no estudo, tanto para os dados de 500m quanto 250m do MODIS, que é a validação dos resultados através de uma comparação com outros projetos de sucesso comprovado, citando como exemplo o PRODES.

4. Considerações finais

Espera-se que com o desenvolvimento deste sistema obtenham-se os seguintes resultados/contribuições:

- ◆ O desenvolvimento de uma metodologia para a detecção dos diversos processos de degradação na Amazônia brasileira, utilizando a integração de dados multisensores;
- ◆ Possibilitar que o monitoramento repetitivo da dinâmica da cobertura da terra torne-se mais freqüente e eficiente, do que os aplicados atualmente;
- ◆ Desenvolver técnicas mais eficientes as quais serão as mais automatizadas e livres possíveis da necessidade da interpretação humana e das inúmeras horas de processamento de seus dados;
- ◆ Possibilitar a utilização de sensores de última geração, como é o caso do sensor HSS do SIVAM, no monitoramento da Amazônia Brasileira;
- ◆ Auxiliar o SIVAM e o INPE em uma de suas inúmeras missões que é o de monitoramento da Amazônia.

5. Referência bibliográfica

- Barker, J.L.; Harden, M.K.; Anuta, E.A.; Smid, J. e Hougt, D. **MODIS Spectral Sensivity Study: Requirements and Characterization**. October, 1992, 84p.
- DeFries, R.S.; Hansen, M.C.; Steininger, M.; Derlayah, R.; Sohlberg R e Townshend, J. **Subpixel forest cover in Central Africa for multisensor, Multitemporal data**. *Remote Sensing of Environment*. n.60, p.228-246, 1997.
- DeFries, R.; Chan, R.S. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. **Remote Sensing of Environment**. n.74, p.503-515, 2000.
- Friedl, M. A.; Brodley, C. E. Decision tree classification of land cover from remotely sensed data. **Remote Sensing of Environment**, n.6, p.399–409, 1997.
- Friedl, M. A.; Brodley, C. E.; Stralher, A.H. Maximizing land cover classification accuracies produced by decision trees at continental to global scales. **IEEE Transactions on Geoscience and Remote Sensing**, 37(2):969-977, march, 1999.
- 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.
- Justice, C. O.;Townshend, J.R.G.; Vermote, E.F., Masuoka, E., Wolfe, R.E., Saleous, N., Roy, D.P., Morisette, J.T. An overview of MODIS Land data processing and product status. **Remote Sensing of Environment** n.83, p.3–15, 2002.
- Salomonson, V.V.; Backer, J.L. **EOS Moderate Resolution Imaging Spectroradiometer: phase C/D status and comments on calibration and georeferencing approaches**, 15th Annual AAS Guidance and Control Conference, Paper AAS 92-004, Keystone, CO, Feb. 8-12, 1992.
- Strahler, A.; Muchoney, D.; Borak, J.; Friedl, M.; Gopal, S.; Lambin, E.; Moody, A. **MODIS Land Cover and Land-Cover Change products algorithm theoretical basis document (ATBD)**. Version 5.0, Center for Remote Sensing, Department of Geography, Boston University, May, 1999, 72p.
- Venables, W.W.; Ripley, B.D. *Modern applied statistics with S-PLUS*. Springer-Verlag, New York, 1994.