

A confiabilidade do PRODES: estimativa da acurácia do mapeamento do desmatamento no estado Mato Grosso

Marcos Adami¹
Alessandra Rodrigues Gomes¹
Amanda Beluzzo²
Andréa dos Santos Coelho¹
Dalton de Morisson Valeriano³
Felipe de Souza Ramos⁴
Igor da Silva Narvaes¹
Irvin Foster Brown⁵
Ivanilson Dias de Oliveira⁵
Lucyana Barros Santos¹
Luis Eduardo P. Maurano³
Orlando dos Santos Watrin²
Paulo Maurício Lima de Alencastro Graça⁴

¹Instituto Nacional de Pesquisas Espaciais - INPE
Av. Perimetral, 2651- 66077-830 - Belém - PA, Brasil
{marcos.adami, alessandra.gomes, igor.narvaes}@inpe.br, {andrea.coelho,
lucyana.santos}@funcate.org.br

²Embrapa Amazônia Oriental - EMBRAPA/ CPATU
Caixa Postal 48 - 66095-903 - Belém - PA, Brasil
orlando.watrin@embrapa.br, amanda.belluzzo@gmail.com

³Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12227-010 - São José dos Campos - SP, Brasil
dalton@dsr.inpe.br, maurano@dpi.inpe.br

⁴Instituto Nacional de Pesquisas da Amazônia – INPA
Av. André Araújo, 2936, Petrópolis- CEP 69067-375. Manaus- AM - Brasil
FS_Ramos@hotmail.com, pmalencastro@gmail.com

⁵Universidade Federal do Acre – UFAC
BR 364, Km 04, s/n - Distrito Industrial - 69920-900 - Rio Branco - AC - Brasil
ivanilson.geo@gmail.com, fbrown@uol.com.br

Abstract. PRODES is completing almost thirty years of uninterrupted monitoring of clear-cut deforestation over the Brazilian Amazon. Until now, no estimate of its mapping accuracy has been made. In this sense, this article brings a first approximation of mapping accuracy estimation of PRODES deforested areas, taking as example the state of Mato Grosso for the year 2014. For this, a random sampling panel was constructed, stratified with two strata, the deforestation of 2014 and the remaining forest. The sample size was calculated using the binomial function. In addition, a web platform was built to evaluate the points drawn by three independent evaluators. The global accuracy of the mapping of deforestation for the state of Mato Grosso, for the year 2014 was 94.5%, and may vary between 92.4% and 96.5%, in the evaluated scenario there was no class discordance to be found. Regarding the Forest class, the user accuracy was 90.5% and the producer's accuracy was 88.4%, this imbalance between user accuracy and producer accuracy indicates that there is a tendency for the forest class area to be underestimated for this mapping, in this year.

Palavras-chave: user and producer's accuracy, global accuracy, Amazon rain forest, acurácia do usuário e do produtor, acurácia global, floresta Amazônica.

1. Introdução

No contexto da geração de mapeamentos sistemáticos acompanhados de estimativas periódicas da taxa de desmatamento na Amazônia brasileira, o trabalho de maior alcance vem sendo desenvolvido pelo INPE, a partir da análise das imagens do satélite norte-americano Landsat ou similares. A realização de um primeiro levantamento experimental data de 1979 (Tardin et al. 1979), mas a partir de 1988 o mapeamento se torna anual com a criação do Projeto de Monitoramento do Desmatamento na Amazônia Legal por Satélite (PRODES) visando à avaliação do processo de desmatamento nos estados que compõem a Amazônia Legal.

Dessa forma, o PRODES está completando quase trinta anos de monitoramento ininterrupto do desmatamento por corte raso na Amazônia brasileira. No início do projeto, a conservação do meio ambiente e os serviços ambientais das florestas tropicais estavam em foco. No cenário internacional, o Brasil encontrava-se sobre forte pressão devido a diversas matérias relatando a destruição da Amazônia por desmatamento e/ou queimadas que colocavam em jogo a governança brasileira sobre a região. Isto motivou o Brasil a retomar as estimativas de desmatamento, utilizando o conhecimento prévio adquirido (Tardin et al. 1979; Tardin et al. 1977), aliado aos avanços tecnológicos da computação e do sensoriamento remoto. No período entre 1988 e 2000 (PRODES Analógico) foram utilizadas aproximadamente 220 imagens Landsat-5/TM por ano, impressas na escala 1:250.000. Sobre estas imagens eram sobrepostos *overlays* para realização da interpretação visual das áreas desmatadas. Posteriormente estes *overlays* eram digitalizados para o cálculo da área e da taxa de desmatamento.

Entre 2000 e 2005, com o avanço das geotecnologias, o PRODES passou a ser realizado utilizando o software SPRING (Câmara et al. 1996). Neste período, todo o processo era realizado utilizando ferramentas computacionais (PRODES Digital) semi-automatizadas (Shimabukuro et al. 1998). Além disto, em 2002, ocorreu a disponibilização de todos os dados produzidos pelo projeto na Internet, estabelecendo um marco, pois tornou o processo de geração e divulgação do dado transparente, permitindo que sejam avaliados tanto pela comunidade científica, quanto por qualquer outro setor da sociedade. Em 2005, o PRODES começou a ser realizado utilizando uma plataforma multiusuário (TERRAAMAZON) com capacidade de armazenar, tratar e distribuir milhões de polígonos, e gerenciar mais de 2.000 cenas Landsat ou similar em uma base de dados única. Com este desenvolvimento, o processo voltou a ser realizado somente por meio de fotointerpretação. A adoção do critério de fotointerpretação foi realizada para garantir maior precisão da classificação e velocidade na classificação das imagens, o que possibilitou que a divulgação dos dados ocorresse ainda durante o ano da detecção.

Dada a importância deste monitoramento, até o presente momento já foram realizadas cerca de 7.000 publicações científicas (Google SCHOLAR) utilizando os dados PRODES. Outro fator importante relacionado com a disponibilização dos dados PRODES é o uso dos mesmos na implementação de acordos voluntários de cadeias produtivas do agronegócio com compromisso de combater o desmatamento na Amazônia, tais como a moratória da soja (Rudorff et al. 2011) e o Termo de Ajustamento de Conduta da pecuária (TAC do Boi) (Nepstad et al. 2014), bem como de acordos intergovernamentais a exemplo da Conferência das Nações Unidas sobre Mudanças Climáticas (COP 21). Desta maneira, este mapeamento contribui para a governança e desenvolvimento de políticas estratégicas que buscam a produção sustentável na Amazônia.

Mesmo dada a sua importância, até o presente momento, ainda não foi realizada nenhuma estimativa da acurácia do mapeamento PRODES. Neste sentido, este artigo traz

uma primeira aproximação da estimativa da acurácia do mapeamento das áreas desmatadas PRODES, tomando como exemplo o estado do Mato Grosso para o ano de 2014.

2. Metodologia de trabalho

O estado do Mato Grosso, com área aproximada de 10% do território nacional tem sua vegetação natural pertencente a três biomas brasileiros, Pantanal, Cerrado e Amazônia. Nestes três biomas, do total de área coberto com formações florestais, até o ano de 2014 já haviam sido desmatados 207.401 km², sendo que o incremento 2013/14 foi de 1.104 km² (Figura 1).

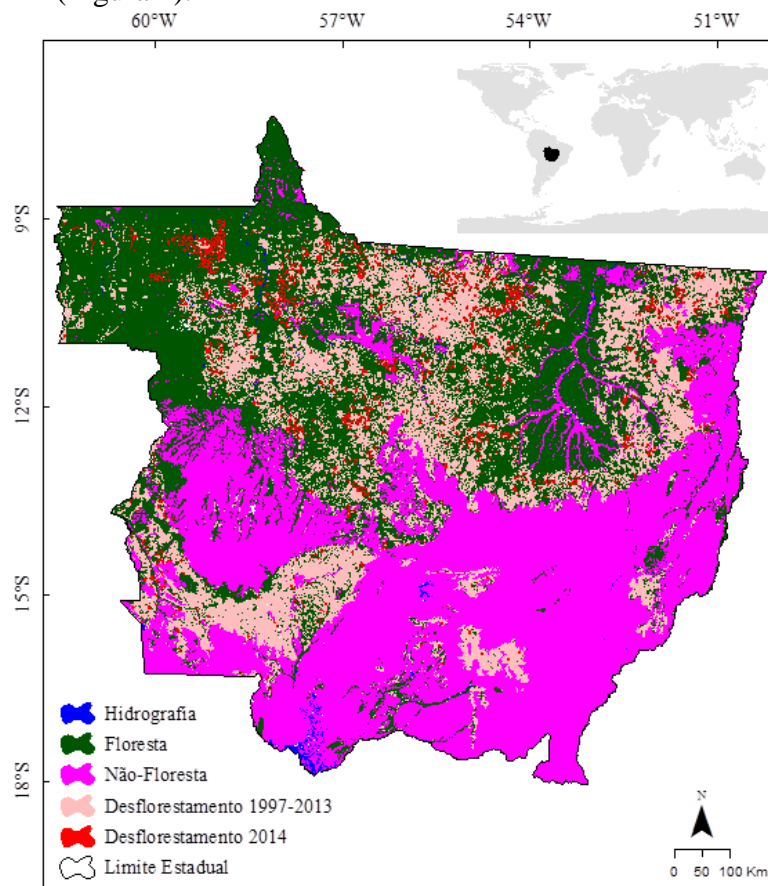


Figura 1. Área de estudo.

2.1. Construção do painel amostral

A construção do painel amostral tem por base amostragem aleatória estratificada, sendo neste caso dois estratos amostrais, desmatamento de 2014 e floresta remanescente no mesmo ano. No caso específico do estrato de floresta, o sorteio das amostras levou em conta a dependência temporal do desmatamento. Para isto, foi calculada a função densidade de probabilidade (FDP) de ocorrência de desmatamento em função da distância do desmatamento do ano anterior. Com base na FDP foi criado um *buffer* em torno dos desflorestamentos acumulados até 2013 e alocado 90% do número de amostras do estrato Floresta os demais 10% de amostras pertencentes a este estrato foram alocadas nas áreas de florestas que não estavam contidas pelo *buffer*.

Para o cálculo do tamanho amostral foi usada a função binomial, que é um caso específico da função multinomial, recomendado quando o mapa temático tem apenas duas classes mutuamente exclusivas segundo a equação 1:

$$n = \left(\frac{z_{\alpha}^2 pq}{E^2} \right) \quad \text{eq. 1}$$

Em que: n = número elementos amostrais (pontos), α = nível de significância, E = erro de amostragem, p = proporção de pontos corretos, $q = 1 - p$ ou proporção de pontos errado e Z = valor crítico tabelado da função normal padrão, para α .

2.2. Ferramenta web

Para realizar o trabalho de validação e estimar as incertezas deste mapeamento, teve por base o trabalho de Adami et al. (2012). Desta forma, foi desenvolvida uma ferramenta web apoiado em: i) Interface de Programação de Aplicativos (API) JavaScript do Google Maps que contém um conjunto de funções que possibilitam customizar aplicações, neste caso a API será responsável pela interface com o usuário (intérprete que fará a validação). ii) Sistema Gerenciador de Banco de Dados Objeto-Relacional (SGBD-OR) com extensão espacial PostGIS. iii) *Web Time Series Service* (WTSS), API Java Script para extração e plotagem de séries temporais geradas a partir das coleções de imagens do sensor MODIS dos produtos MOD09Q1 e MOD13Q1 (QUEIROZ et al. 2015). iv) - Biblioteca GDAL de Abstrações de Dados Geo-Espaciais que foi utilizada para realizar os recortes prévios das imagens Landsat e RapidEye. v) Linguagem *Hypertext Preprocessor* (PHP) usada originalmente para o desenvolvimento de aplicações no lado do servidor, capaz de gerar conteúdo dinâmico na Internet.

A interface WEB, construída sobre a API do Google Maps utilizando linguagem HTML, PHP e Java Script, concentra as principais funções da ferramenta. Através dela o usuário se comunica com o SGBD-OR PostGis, selecionando o ponto a ser classificado bem como obtendo e apresentando na tela a série temporal *Normalized Difference Vegetation Index* (NDVI) MODIS correspondente ao ponto escolhido, via a API WTSS. Adicionalmente a interface coordena a exibição dos recortes das imagens Landsat 8 e RapidEye, armazenados em formato KML, que servem de auxílio para o processo de validação. De uma maneira resumida a arquitetura básica da ferramenta é mostrada na Figura 2. Este sistema foi implementado para que três avaliadores, orientados por pesquisadores de instituições da Amazônia realizassem, com base nas mesmas imagens utilizadas pelo PRODES, nas imagens de maior resolução espacial (RapidEye) e na série temporal NDVI Modis, a avaliação de cada um dos pontos sorteados. A avaliação da acurácia será executada com base no valor modal destas estimativas.

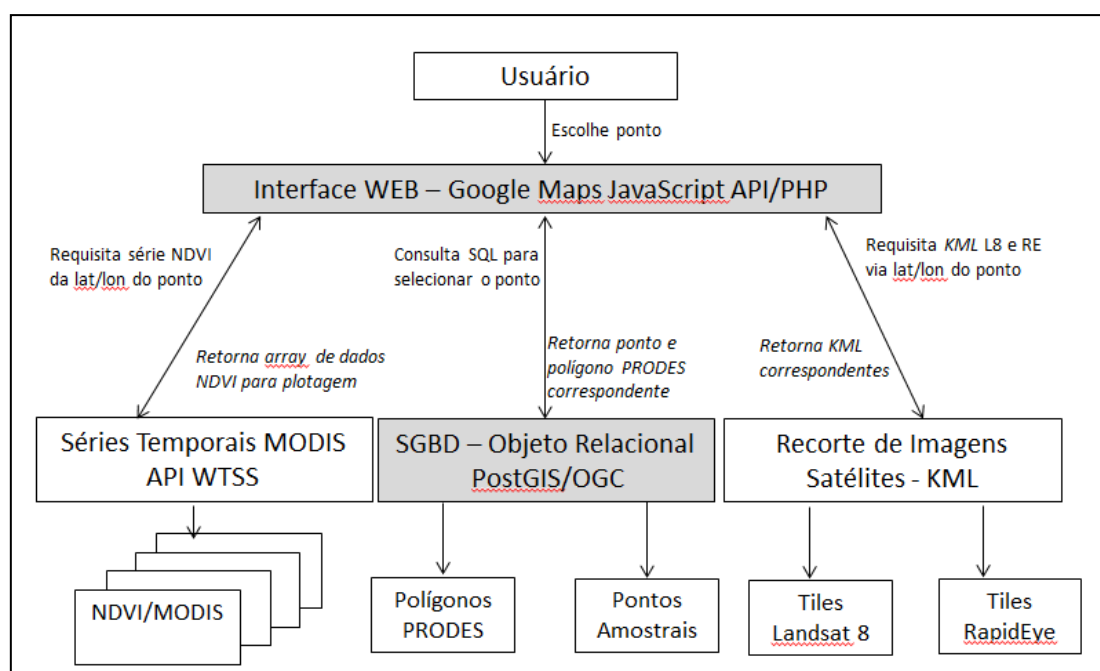


Figura 2 – Arquitetura simplificada do sistema.

2.3. Estimativa de acurácia

A estimativa da acurácia utilizou uma tabela de contingência (Matriz de Erro), exemplificada na Tabela 1. Os dados tabulados foram utilizados para calcular a Precisão Global (PG), Acurácia do Usuário (AU) que está associada com a omissão e a Acurácia do Produtor (AP) que está associada com a inclusão. Os valores alocados na diagonal principal da Tabela 1 representam os elementos corretamente classificados. Os elementos alocados fora da diagonal principal representam os erros de omissão e inclusão (Foody 2009; Stehman and Foody 2009).

Tabela 1. Exemplo de tabela de contingência (Matriz de Erro ou de Confusão), contendo as funções para cálculo da PG, AU e AP.

		Classes de Referência		Total	Acurácia do Produtor
		Floresta	Desmatamento		
Classes do mapa	Floresta	P_{11}	P_{12}	$Tf+ = P_{11}+P_{12}$	$APf = P_{11} / Tf+$
	Desmatamento	P_{21}	P_{22}	$Td+ = P_{21}+P_{22}$	$APd = P_{11} / Td+$
Total		$T+f = P_{11} + P_{21}$	$T+d = P_{12} + P_{22}$	$T = P_{11} + P_{12} + P_{21} + P_{22}$	
Acurácia do Usuário		$AUf = P_{11} / T+f$	$AUd = P_{22} / T+d$	$PG = (P_{11} + P_{22}) / T$	

Os índices f e d se referem a floresta e desmatamento, respectivamente. Os erros de omissão e inclusão são estimados diminuindo de um os resultados de AU e AP.

O intervalo de confiança fornece uma gama de valores associados a um nível de significância a partir dos quais estima-se a possibilidade de ocorrência de um determinado parâmetro estatístico. Assumindo uma distribuição normal, o intervalo de confiança pode ser obtido por $estimativa \pm Z_{\alpha/2}(SE)$, sendo que $Z_{\alpha/2}$ é o valor tabulado para uma normal padrão associado a um nível de significância (α) e SE é o erro padrão. No caso desta amostragem, foi utilizado o nível de significância de 95% e o SE foi calculado utilizando o teste de McNemar (Foody 2009), em que o erro padrão (SE_M) pode ser calculado segundo a equação 2:

$$SE_M = \frac{\sqrt{P_{12} + P_{21} - (P_{12} - P_{21})^2 / T}}{T} \quad \text{eq. 2}$$

3. Resultados e Discussão

Com base na equação 1 foram sorteados 502 pontos para a área de estudo, 251 pontos para cada estrato. Sendo que dos quais 30 pontos não puderam ser utilizados e estão representados na Figura 3 como sem condições de avaliação.

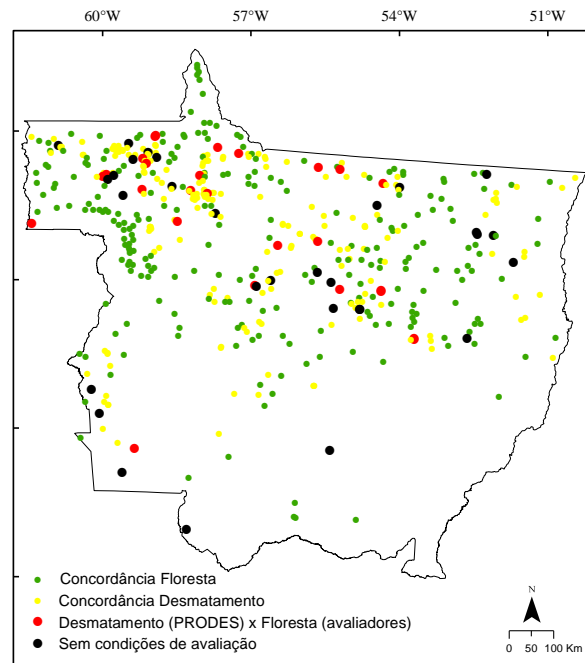


Figura 3. Localização dos pontos sorteados e resultado da validação.

As principais causas desta exclusão foram a não existência de um valor modal entre os três avaliadores, a presença de nuvens e outros problemas relacionados com as imagens, tais como ruído. Além disto, observou-se relação entre o formato do polígono e a distância entre o ponto sorteado e a borda dos polígonos desmatados. Polígonos cuja relação área/perímetro era menor e consequentemente, os pontos se localizavam mais próximos da borda (Figura 4) o que aumenta ainda mais a dificuldade para a tomada de decisão pelos avaliadores.

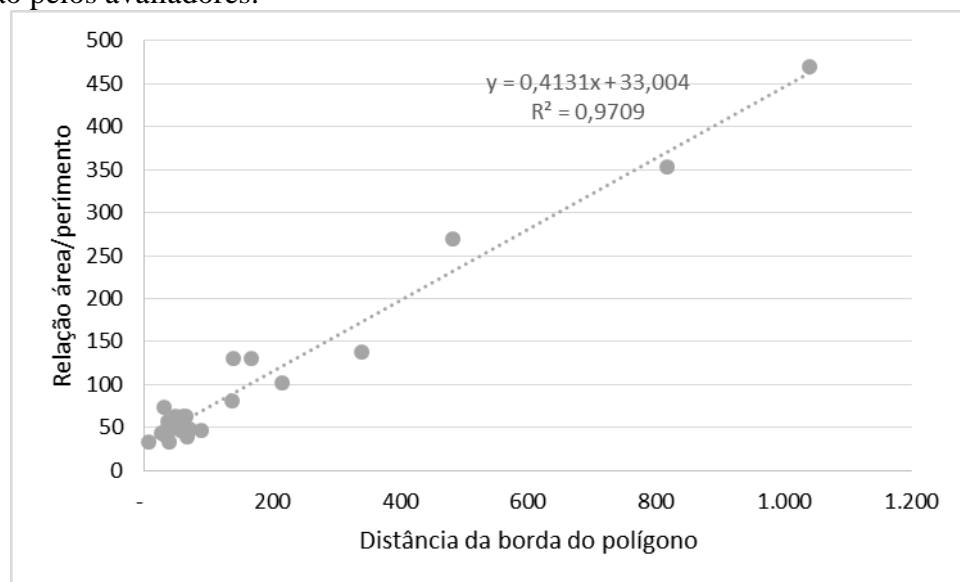


Figura 4. Relação entre o formato dos polígonos e a distância entre o ponto sorteado e a borda dos polígonos desmatados.

Os demais 472 pontos que foram avaliados neste estado estão apresentados na matriz de contingência (Tabela 2). Com base nestes dados, a estimativa da PG do mapeamento do desmatamento para o estado do Mato Grosso, para o ano de 2014 foi de 94,5%, podendo variar entre 92,4% e 96,5%, sendo que a nesta amostragem, não houve nenhuma

discordância da classe mapeada como floresta. Com relação a classe Floresta, a AU foi de 90,5% e a AP foi de 88,4%, este desbalanceamento entre AU e AP indicam que há tendência da área da classe floresta estar sendo subestimada por este mapeamento, neste ano.

Tabela 2. Tabela de contingência contendo a comparação entre os dados de referência e os dados classificados com o PRODES bem como os valores de PG, AP e AU

		Classes de Referência		Total	Acurácia do Produtor (AP)
		Floresta	Desmatamento		
Classes do mapa	Floresta	248		248	100%
	Desmatamento	26	198	224	88,4%
Total		296	198	472	
Acurácia do Usuário (AU)		90,5%	100%	PG=94,5%±2,05	

Ao analisar estes 26 pontos do estado do MT que apresentaram discordância entre o mapeado e a referência foi observado que aproximadamente 70% dos pontos estavam até uma distância de 3 pixels da borda do polígono (Figura 5).

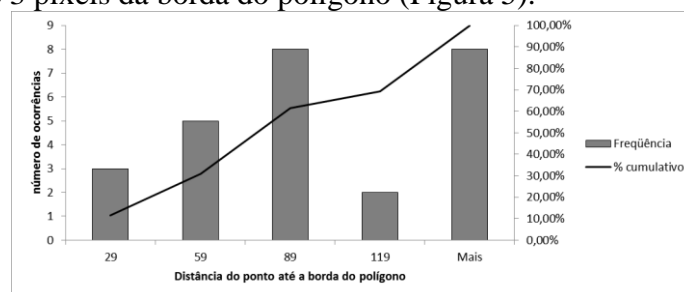


Figura 5. Histograma com a distância entre os pontos sorteados e a borda dos polígonos desmatados.

4. Conclusões

Com base nos resultados apresentados, o mapeamento das áreas desmatadas PRODES, para o estado do Mato Grosso, no ano de 2014 tem uma precisão global (PG) de 94,5% ± 2,05, sendo que somente foram encontrados erros de inclusão de áreas de floresta como sendo desflorestamento.

Dado o desafio de determinar a confiabilidade de estimativas de desmatamento, pretendemos continuar a analisar outros fatores que afetam a acurácia destas estimativas em trabalhos subsequentes.

5. Referências

Adami, M., Mello, M.P., Aguiar, D.A., Rudorff, B.F.T., Souza, A.F.d. A Web Platform Development to Perform Thematic Accuracy Assessment of Sugarcane Mapping in South-Central Brazil. **Remote Sensing**, 4, 3201-3214, 2012.

Câmara, G., Souza, R.C.M., Freitas, U.M., Garrido, J. Spring: Integrating remote sensing and gis by object-oriented data modelling. **Computers & Graphics**, 20, 395-403, 1996.

Foody, G.M. Classification accuracy comparison: Hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. **Remote Sensing of Environment**, 113, 1658-1663, 2009.

Nepstad, D., McGrath, D., Stickler, C., Alencar, A., Azevedo, A., Swette, B., Bezerra, T., DiGiano, M., Shimada, J., Seroa da Motta, R., Armijo, E., Castello, L., Brando, P., Hansen, M.C., McGrath-Horn, M., Carvalho, O., Hess, L. Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. **Science**, 344, 1118-1123. 2014.

Queiroz, G.R.d., Ferreira, K., Vinhas, L., Câmara, G., Costa, R.W., Souza, R.C.M., Maus, V.W.S., A. WTSS: um serviço web para extração de séries temporais de imagens de sensoriamento remoto. In: XVII SBSR João Pessoa, 2015, **Anais...** São José dos Campos:INPE 2015. Artigos, p. 7553-7560

Rudorff, B.F.T., Adami, M., Aguiar, D.A., Moreira, M.A., Mello, M.P., Fabiani, L., Amaral, D.F., Pires, B.M. The Soy Moratorium in the Amazon Biome Monitored by Remote Sensing Images. **Remote Sensing**, 3, 185-202, 2011.

Shimabukuro, Y.E., Batista, G.T., Mello, E.M.K., Moreira, J.C., & Duarte, V. Using shade fraction image segmentation to evaluate deforestation in Landsat Thematic Mapper images of the Amazon Region. **International Journal of Remote Sensing**, 19, 535 – 541, 1998

Stehman, S.V., Foody, G.M. **Accuracy assessment**. In: T.A. Warner, M.D. Nellis, G.M. Foody (Eds.), *The Sage Handbook of Remote Sensing*, p. 297–309. London: SAGE, 2009.

Tardin, A.T., Rodrigues, J.E., Abdon, M.M., Novaes, R.A., Chen, S.C., Duarte, V., Shimabukuro, Y.E., Santos, A.P., Lee, D.C.L., Maia, F.C.S., Mendonca, F.J., Assuncao, G.V. **Levantamento de áreas de desmatamento na Amazônia Legal através de imagens do Satélite Landsat**. São José dos Campos: INPE, 1979, 62p. (INPE-1411-NTE/142).

Tardin, A.T., Santos, A.P., Novo, E.M.L.M., Toledo, F.L. (1977). **Relatório das atividades do projeto SUDAM/INPE durante o ano de 1976/1977**. São José dos Campos: INPE, 144p. (INPE-1034-NTE/088)