

Ferramentas Web-Gis para avaliar exatidão de mapeamento de uso e cobertura da terra no Cerrado brasileiro

Luis Eduardo P. Maurano¹
Marcos Adami¹

¹ Instituto Nacional de Pesquisas Espaciais – INPE
Caixa Postal 515 - 12227-010 - São José dos Campos - SP, Brasil
maurano@dpi.inpe.br, marcos.adami@inpe.br

Abstract: The elaboration of land use and land cover thematic maps, based on satellite remote sensing data, for the monitoring of natural resources, has been quite common in Brazilian governmental institutions related to the environment, agriculture, geography and statistics areas. Most of the time these mappings are performed on a large territorial scale, covering entire biomes or several aggregate states, which makes difficult to carry out field work to assess the accuracy of the mapping done. Specific applications that use the Web-GIS concept have been developed and successfully employed for validation tasks, aiming to associate an accuracy value with the mapping information. This paper demonstrates the development and practical implementation of an application to validate land use mapping data in the Brazilian Cerrado (TerraClass Cerrado), which can be easily adapted to other biomes.

Palavras-chave: mapeamento de uso e cobertura da terra, sensoriamento remoto, Web-GIS, avaliação de exatidão.

1. Introdução

O Cerrado é o segundo maior bioma da América do Sul ocupando cerca de 25% do território brasileiro (2.039.243 km²). Fundamental para a conservação da biodiversidade e o desenvolvimento sustentável do Brasil, o Cerrado tem passado por uma significativa conversão de habitat devido à abertura de novas fronteiras agrosilvopastoris que contribuiu para que o país se tornasse um dos maiores produtores e exportadores de *commodities* agropecuárias do mundo.

Apesar de se tratar de importante região brasileira, com alta biodiversidade o Cerrado não dispunha, até recentemente, de um mapeamento detalhado que fornecesse informações atualizadas sobre o seu uso e cobertura da terra. Em 2013, o Ministério do Meio Ambiente (MMA) fomentou a união de um grupo de instituições públicas brasileiras com ampla experiência em sensoriamento remoto, geoprocessamento e mapeamentos de larga escala, para realizar a primeira versão do projeto intitulado Mapeamento do Uso e Cobertura Vegetal do Cerrado – TerraClass Cerrado. Esse projeto, idealizado para gerar dados compatíveis com a escala cartográfica 1:250.000, utilizou como base de mapeamento 118 cenas do satélite Landsat 8, sensor *Operational Land Imager* (OLI), do ano de 2013. A área mínima mapeável foi de 6,25 hectares.

Entretanto, para que o mapa possa ser utilizado em inferências estatísticas, suas medidas de erro devem ser calculadas (McRoberts, 2011). Jensen (2009) aponta que para minimizar os erros introduzidos no processo de mapeamento temáticos deve-se: a) registrar dos passos operacionais utilizados, b) documentar os erros geométricos e temáticos, c) definir a legenda e d) ter métodos de avaliação de exatidão. A incorporação de medidas de rastreamento de erros acarreta em informações mais exatas e conferem credibilidade ao mapeamento proposto.

Exatamente por isso, medidas de acurácia de mapeamentos temáticos elaborados com base em dados de sensoriamento remoto devem ser asseguradas e encorajadas para, além de prover informação de boa qualidade, também determinar níveis de confiança para estes mapeamentos. Porém, garantir processos de avaliação de erros em levantamentos em larga escala, como no TerraClass Cerrado, não é tarefa simples e nem barata. Aferições em campo de mapeamentos de grande extensão territorial são prejudicadas em parte pela dificuldade de

acesso, custos operacionais e logísticos além de demandarem tempo elevado de planejamento e execução (Adami et al., 2012).

Nos últimos 10 anos, alguns pesquisadores têm estudado métodos alternativos para avaliar a exatidão e a acurácia de mapeamentos de larga escala, quase todos baseados em soluções que utilizam a internet como plataforma operacional, desenvolvendo ferramentas que integram a base de imagens de alta resolução disponíveis no ambiente *Google Maps/Earth* e séries temporais extraídas dos produtos do sensor *Moderate Resolution Imaging Spectroradiometer* (MODIS) (Dorais e Cardille (2011) e Cohen et al. (2010)).

Baseando-se nestes estudos, Adami et al. (2012) conceberam e desenvolveram uma ferramenta que combinava em um mesmo ambiente gráfico as imagens disponíveis no *Google Maps*, imagens Landsat, imagens do sensor MODIS e suas séries temporais para mensurar erros e validar os mapeamentos gerados no Projeto Canasat, determinando assim seus níveis de acurácia e de confiança. Seguindo a ideia dos trabalhos supracitados, o presente artigo tem por objetivo o desenvolvimento e uso de um aplicativo que utiliza conceitos de Web-GIS para auxiliar na validação e cálculo da exatidão do mapeamento realizado pelo TerraClass Cerrado (INPE, 2013).

2. Metodologia de Trabalho

2.1 Suporte Tecnológico

A ferramenta de validação está apoiada em quatro tecnologias básicas:

- a) API (Interface de Programação de Aplicativos) JavaScript do *Google Maps*: Conjunto de funções escritas em linguagem JavaScript que possibilitam customizar aplicações utilizando o ambiente *Google Maps*, incorporando-o a uma aplicação específica, que será responsável pela interface com o usuário (intérprete que fará a validação).
- b) Sistema Gerenciador de Banco de Dados Objeto-Relacional (SGBD-OR): Gerenciador de banco de dados com extensão espacial - PostGIS, para armazenar os dados vetoriais contendo os polígonos do TerraClass Cerrado e os pontos amostrais da validação em padrões estabelecidos pelo OGC (*Open Geospatial Consortium*).
- c) *Web Time Series Service* (WTSS): API JavaScript para extração e plotagem de séries temporais geradas a partir das coleções de imagens do sensor MODIS dos produtos MOD09Q1 e MOD13Q1. Esta API foi desenvolvida pelo grupo da Divisão de Processamento de Imagens (DPI) do INPE, e permite acesso as *coverages* armazenadas em banco de dados espaço-temporal do tipo *multi-dimensional array*, como o SciDB, ou em um sistema de arquivos estruturados para consulta (Queiroz et al., 2015).
- d) GDAL - Biblioteca de Abstrações de Dados Geo-Espaciais: Utilizada para realizar os recortes prévios das imagens Landsat 8/OLI e RapidEye, armazenando-os em formato KML (*Keyhole Markup Language*).
- e) Linguagem PHP (*Hypertext Preprocessor*) é uma linguagem interpretada livre, usada originalmente para o desenvolvimento de aplicações presentes e atuantes no lado do servidor, capaz de gerar conteúdo dinâmico na Internet.

2.2 Arquitetura da ferramenta

A interface WEB, construída sobre a API do *Google Maps* utilizando linguagem HTML, PHP e JavaScript, concentra as principais funções da ferramenta. Através dela o usuário se comunica com o SGBD-OR PostGis, selecionando o ponto a ser classificado e obtendo o respectivo polígono vetorial do TerraClass Cerrado, ela também recupera e plota a série temporal de NDVI (*Normalized Difference Vegetation Index*) do 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 1.

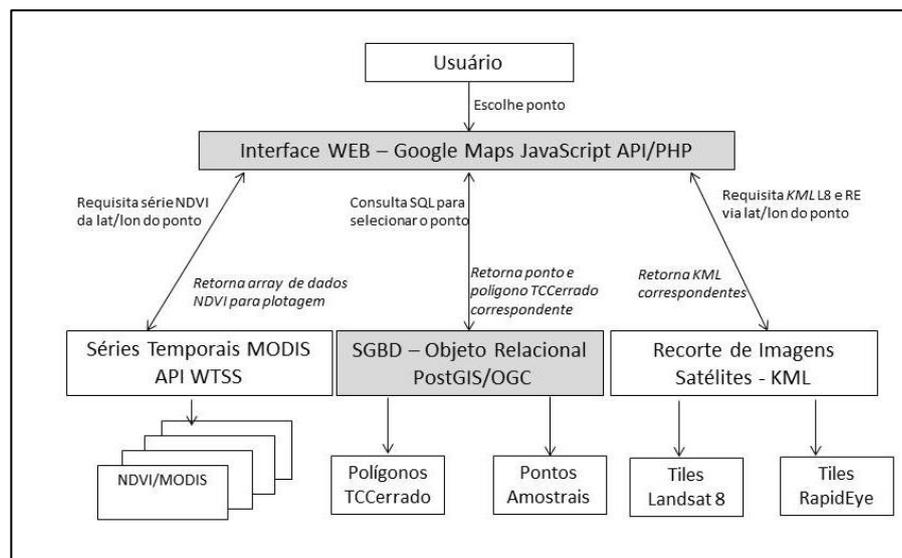


Figura 1 – Arquitetura simplificada do sistema.

2.3 Painel amostral

Uma fase importante do trabalho de validação para avaliação da acurácia do mapeamento se refere a confecção do painel amostral. Neste caso o painel foi construído usando amostragem aleatória estratificada. Os estratos tiveram por base as classes mapeadas, que em alguns casos foram agrupadas, sendo proposto os seguintes extratos: 1 - agricultura anual; 2 - agricultura perene; 3 - água e natural não vegetado; 4 - mineração, mosaico de ocupação e área urbana; 5 - outros, solo exposto e não observado; 6 - vegetação natural, classe florestal; 7 - vegetação natural, classes savânico e campestre; 8 - pastagem; e 9 - silvicultura.

Após a definição dos estratos, foi realizado o cálculo do tamanho da amostra utilizando-se a função multinomial. A distribuição da quantidade de amostras por estrato foi realizada utilizando-se do método de alocação ótima de Neyman (COCHRAN, 1977; CONGALTON; GREEN, 2009; STEHMAN, 2012) sendo gerados ao todo 3027 pontos. A Figura 2 mostra a espacialização dos pontos amostrais sorteados.

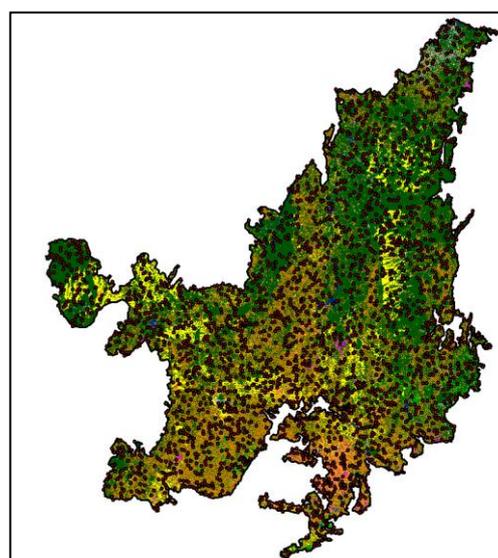


Figura 2 – Espacialização dos 3207 pontos sorteados para a validação do mapeamento.

3. Resultados

A plataforma web foi desenvolvida para operar em ambiente cliente/servidor e permite combinar em uma página de internet escrita em linguagem HTML, PHP e JavaScript: (1) os pontos amostrais sorteados com as classes floresta e desmatamento armazenados no SGBD-OR PostGis, (2) os recortes de imagens do satélite Landsat utilizados no mapeamento, (3) as imagens de alta resolução do satélite RapidEye, e (4) as imagens de altíssima resolução disponíveis no ambiente *Google Maps*. Além deste conjunto de informações, ainda está disponível a série temporal de NDVI do sensor MODIS, para os últimos 10 anos, visando auxiliar o intérprete-auditor no momento da validação de cada ponto.

A operação do aplicativo é bastante simples. Todo processo inicia-se com a escolha do ponto a ser classificado (pelo intérprete-auditor) que é mostrado em uma lista no menu esquerdo da interface. Clicando sobre o ID do ponto, uma consulta em linguagem SQL (Linguagem de Consulta Estruturada) é disparada no SGBD-OR PostGis, recuperando a coordenada de latitude e longitude e demais atributos do ponto. Com base na coordenada, uma segunda consulta SQL é realizada e visa recuperar o polígono do TerraClass Cerrado correspondente ao ponto. Em seguida, polígono e ponto são visualizados de forma centralizada na tela do *Google Maps*. A mesma coordenada é utilizada para consultar, recuperar e plotar o gráfico da série temporal extraída do produto MODIS/MOD13Q1 utilizando-se para isso a API WTSS.

No menu direito encontra-se dois tiles de imagens de satélite da região a ser classificada, sendo que um deles mostra uma área de 10 km ao redor do ponto extraído das imagens Landsat 8/OLI utilizadas no TerraClass Cerrado para o ano de 2013; e o segundo tile mostra uma área de 1,7 km extraído das imagens de alta resolução do satélite RapidEye. É possível sobrepor qualquer um destes tiles a imagem do *Google Maps* e ao polígono do TerraClass Cerrado. Nesta tela também esta disponível um formulário, onde o intérprete-auditor deve escolher a qual classe pertence o ponto a ser classificado. A Figura 3 apresenta a interface principal do sistema.

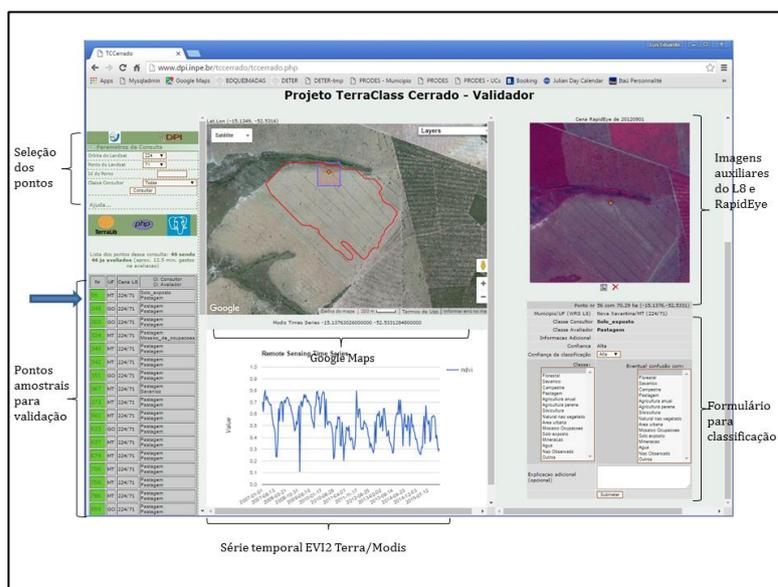


Figura 3 - Interface principal da página desenvolvida para a validação.

Esta abordagem parte do pressuposto que a classificação realizada pelo auditor independente, utilizando a plataforma web, é considerada a “verdade” e assim esta classificação será a base na comparação com a classificação realizada pelos intérpretes do TerraClass Cerrado, para a aferição dos erros e acertos do mapeamento. Ou seja, ela é

altamente dependente da experiência do auditor em conhecer completamente os conceitos e metodologias utilizadas no mapeamento que se pretende avaliar.

Esta condição pode não ser completamente verdadeira e o auditor, mesmo experiente, pode ter dúvidas em relação ao processo que está sendo validado. Para minimizar este fato, a ferramenta também permite que o auditor selecione um fator de confiança (neste caso Alto, Médio ou Baixo) na avaliação que ele está realizando. Com isso a matriz de confusão final expressava também estes fatores de confiança do auditor.

Após a finalização da validação de todos os pontos pelo intérprete-auditor (utilizando a plataforma web), foram calculadas as exatidões globais, os erros de omissão e de inclusão por classe, e os índices de discordância (quantidade, deslocamento e mudança) com base na matriz de contingência (Pontius & Santacruz, 2014).

O índice de concordância geral entre o mapeamento do TerraClass Cerrado e a validação foi de 80,2%, demonstrando uma alta confiabilidade do mapeamento executado com base nos resultados da matriz de contingência conforme dados da Tabela 1. Na Figura 4 são apresentados os valores médios de concordância e discordância, sendo estes últimos divididos em quantidade, deslocamento e mudança (PONTIUS; SANTACRUZ, 2014). Os erros de quantidade referem-se à diferença absoluta entre os erros de omissão e inclusão, fornecendo um indicativo do erro médio da área mapeada. O deslocamento refere-se à componente diferença de alocação não pareada entre as classes analisadas na matriz de contingência; já a mudança refere-se ao componente de erro ocasionado por trocas pareadas entre classes na matriz de contingência.

Desta maneira, podemos afirmar que o erro médio, em área, foi da ordem de 5%, o deslocamento da ordem de 9% e a mudança da ordem de 3%. Ao somarmos estas discordâncias, obtêm-se um erro total de 20%, valor que pode ser considerado baixo, em função da dimensão e da complexidade, tanto dos temas mapeados, quanto da região analisada.

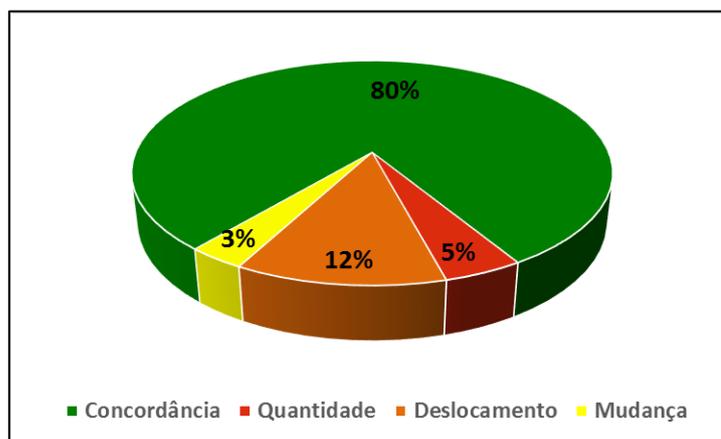


Figura 4: Resultado da exatidão global do mapeamento do TerraClass Cerrado.

A Figura 5 apresenta o resultado por classe temática, já a Tabela 1 apresenta a matriz de contingência. Quando ponderada com relação à classe, cinco das nove classes apresentaram concordância superior a 70%. Os estratos correspondentes as classes de “Agricultura anual”, “Água e natural não vegetado” e “Silvicultura” tiveram concordância igual ou maior a 80%. Já os estratos “Mineração, mosaico de ocupação e área urbana” e “Outros, solo exposto e não observado” tiveram a maior intensidade de omissão, 61% e 39%, respectivamente.

Com relação a inclusão, os estratos “Outros, solo exposto e não observado”, “Vegetação natural, classe florestal” e “Vegetação natural, classes savânico e campestre” tiveram erros da ordem de 23%, 22% e 24%, respectivamente.

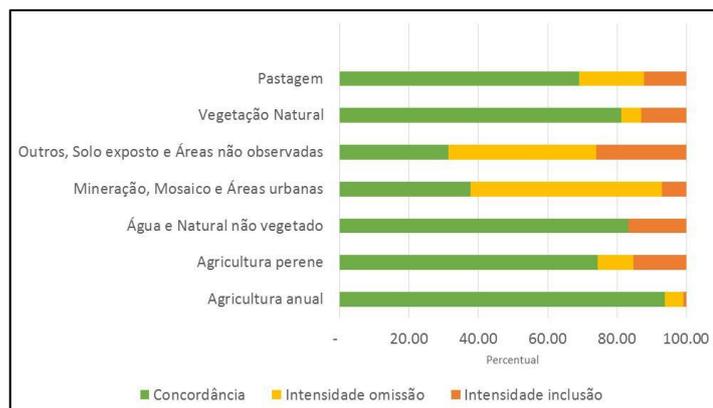


Figura 5: Resultado da exatidão do mapeamento do TerraClass Cerrado por classe.

Tabela 1: Matriz de Contingência

Estratos	Referência								Total
	Agricultura anual	Agricultura perene	Água e Natural não vegetado	Mineração, Mosaico e Áreas urbanas	Outros, Solo exposto e Áreas não observadas	Vegetação Natural	Pastagem	Silvicultura	
Agricultura anual	349	7		1	3	6	20		386
Agricultura perene	7	136		1		3	14	3	164
Água e Natural não vegetado			33	3	2	3	1		42
Mineração, Mosaico e Áreas urbanas	1			38		1	5		45
Outros, Solo exposto e Áreas não observadas				1	39	24	7	2	73
Vegetação Natural	3	2		40	15	1153	154	8	1.375
Pastagem	23	8		49	10	101	851	5	1.047
Silvicultura	1	1		1		1	1	70	75
Total	384	154	33	134	69	1.292	1.053	88	3.207

Este trabalho de validação vem de encontro com as necessidades de avaliação da precisão do mapeamento e com isto possibilitar inferências estatísticas (McRoberts, 2011) em quaisquer trabalhos que utilizem destes dados e possibilita a comparação com outros mapeamentos desta região. Além disto, o ineditismo deste processo aplicado ao mapeamento de larga escala com o uso de ferramentas interativas que possibilitam integrar diversas fontes de dados em um mesmo ambiente, aumentando a capacidade do avaliador na tomada de decisão em região complexa, com grande variação fitofisonômica, além de diversificada forma de ocupação.

O ineditismo deste processo, aplicado ao mapeamento de larga escala com o uso de ferramentas interativas, que possibilitaram a integração de diversas fontes de dados em um

mesmo ambiente, aumentou a capacidade do avaliador na tomada de decisão, especialmente em regiões complexas como o Cerrado que apresenta áreas de grande variação fitofisionômica e diversificadas formas de ocupação.

4. Conclusões

Em mapeamentos do porte do TerraClass Cerrado, mesmo executado por profissionais experientes e respeitando um rigor metodológico na sua confecção, erros são esperados e encontrados no resultado final. Assim, uma tarefa necessária é quantificá-los e informá-los de maneira transparente para garantir parâmetros de confiabilidade mínimos para os mapas gerados. Em virtude da escala territorial envolvida neste projeto, missões de campo para aferição da qualidade da informação gerada são inviáveis por serem dispendiosas financeiramente, pois dependem de logística complicada e também demandam tempo elevado no planejamento e execução.

Métodos alternativos têm sido testados com sucesso. Um destes métodos consiste em realizar a validação do mapeamento através de aplicativos específicos, normalmente utilizando-se o conceito de Web-GIS, que reúnem em uma interface gráfica bastante simplificada um conjunto de informações necessárias para que um intérprete execute a tarefa de validação “à distância”.

Este artigo demonstrou o desenvolvimento e utilização de uma ferramenta Web-GIS que foi empregada na validação do mapeamento de uso e cobertura do Cerrado brasileiro realizado pelo projeto TerraClass Cerrado. Ela permitiu ao intérprete-auditor classificar pontos aleatoriamente sorteados sobre o mapa original, apoiando-se em dados auxiliares, como as imagens de alta resolução espacial disponíveis no ambiente *Google Maps*, imagens Landsat/RapidEye e séries temporais NDVI extraídas do conjunto de imagens do MODIS, possibilitando gerar uma medida de incerteza que pode ser agregada aos dados do mapeamento.

O resultado alcançado na validação do TerraClass Cerrado, obtido com uso da ferramenta Web-GIS desenvolvida, se mostrou satisfatório e encoraja sua utilização para aferição da qualidade em projetos de mapeamento similares.

5. Referencias bibliográficas

ADAMI, M.; MELLO, M. P.; AGUIAR, D. A. ; RUDORFF, B. F. T. ; SOUZA, A. F. . A Web Platform Development to Perform Thematic Accuracy Assessment of Sugarcane Mapping in South-Central Brazil. *Remote Sensing*, v. 4, p. 3201-3214, 2012.

JENSEN, J. R. *Sensoriamento Remoto do Ambiente: uma perspectiva em Recursos Terrestres*. Tradução de J. C. N. Epiphânio. São José dos Campos, SP: Parênteses, 2009. 598 p. (Prentice Hall Series in Geographic Information Science) Tradução de: *Remote Sensing of the environment: na earth resource perspective*.

DORAIS, A.; CARDILLE, J. Strategies for incorporating high-resolution google earth databases to guide and validate classifications: Understanding deforestation in Borneo. *Remote Sens.* 2011, 3, 1157–1176.

QUEIROZ, Gilberto Ribeiro de; Ferreira, Karine ; VINHAS, Lúbia ; CÂMARA, Gilberto ; COSTA, R. W. ; SOUZA, Ricardo Cartaxo Modesto ; MAUS, V. W. ; SANCHEZ, A. . WTSS: um serviço web para extração de séries temporais de imagens de sensoriamento remoto. In: XVII Simpósio Brasileiro de Sensoriamento Remoto, 2015, João Pessoa. Anais do 17 Simpósio Brasileiro de Sensoriamento Remoto, 2015. p. 7553-7560.

COHEN, W.B.; YANG, Z.; KENNEDY, R. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync—Tools for calibration and validation. *Remote Sens. Environ.* 2010, 114, 2911–2924.

INPE (2012). Monitoramento do Desmatamento da Floresta Amazônica Brasileira por Satélite, Projeto PRODES: 2012. Disponível em: <<http://www.obt.inpe.br/prodes/>>

INPE (2013). Mapeamento do Uso e Cobertura da Terra no Cerrado, Projeto TerraClass Cerrado: 2013. Disponível em: <http://www.dpi.inpe.br/tccerrado/Metodologia_TCCerrado_2013.pdf>

INPE (2016). Monitoramento da Cana-de-açúcar via imagens de satélite, Projeto Cansat. 2016: Disponível em: <<http://www.dsr.inpe.br/laf/canasat/>>

McRoberts, R.E. Satellite image-based maps: Scientific inference or pretty pictures? *Remote Sensing of Environment*, 115, 715-724, 2011.

Cochran, W.G. *Sampling techniques* New York & Sons, Inc.: John Wiley 1977.

Congalton, R.G., & Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. (2nd ed. ed.). New York: Taylor & Francis Group, 2009.

Stehman, S.V. Impact of sample size allocation when using stratified random sampling to estimate accuracy and area of land-cover change. *Remote Sensing Letters*, 3, 111-120, 2012.

Pontius, R.G., & Santacruz, A. Quantity, exchange, and shift components of difference in a square contingency table. *International Journal of Remote Sensing*, 35, 7543-7554, 2014.