

COMPARAÇÃO ENTRE RESULTADOS DE CLASSIFICAÇÃO POR MÁXIMA DIFERENÇA ESPECTRAL E POR ANÁLISE MULTITEMPORAL NA DETECÇÃO DE ÁREAS DE EUCALIPTO NO ESTADO DO RIO DE JANEIRO

Debora da Paz Gomes Brandão Ferraz¹, Raúl Sanchez Vicens^{1 2}

¹ Universidade Federal Fluminense – UFF

Posgeo – Programa de Pós-Graduação em Geografia

Campus Praia Vermelha - Instituto de Geociências

Rua Gal. Milton Tavares, s/n, Sala 508, Boa Viagem- Niterói – Rio de Janeiro, Brasil

debora.ferraz93@gmail.com¹

rsvicens@gmail.com²

RESUMO

O crescimento da área plantada de eucaliptos no Brasil vem sendo constatado por dados censitários, e a utilização de dados de sensoriamento remoto contribui para entender onde, quando e como este processo de crescimento acontece. Este trabalho tem como objetivo comparar e analisar os resultados de duas abordagens distintas de classificação das áreas de Eucalipto no Estado do Rio de Janeiro: uma classificação orientada a objetos, com o uso do descritor Máxima Diferença, e outra com base na trajetória da série temporal, com o uso do algoritmo Landtrendr, ambas feitas com imagens do satélite Landsat. A comparação através de teste McNemar mostrou diferenças significativas entre os dois métodos. A classificação orientada a objeto apresentou um melhor desempenho do que a classificação com base na trajetória, com índices kappa de 0,867 e 0,219 respectivamente.

Palavras-chave — silvicultura, séries temporais Landsat, GEOBIA.

ABSTRACT

The growth of the eucalyptus crop area in Brazil has been verified by census data, and the use of remote sensing data helps to understand where, when and how this process happens. The work aims to compare and analyze the results of two distinct approaches to classification of Eucalyptus areas in the State of Rio de Janeiro: An object-based classification with the use of the Maximum Difference feature, and a trajectory-based classification of annual time serie with the use of the Landtrendr algorithm. Both method were made with Landsat data. The McNemar test showed significant differences between the two methods. The object-based classification performed better than the trajectory-based, with kappa index of 0.867 and 0.219 respectively..

Key words — forestry, Landsat time-series, GEOBIA.

1. INTRODUÇÃO

As áreas de Florestas Plantadas têm um grande papel como um dos instrumentos para a obtenção de matéria-prima

destinada ao atendimento de diversas demandas, que vão do pequeno agricultor à grande indústria de base florestal. Ela tem se tornado também importante instrumento a contemplar as práticas de plantações de Eucalipto, por exemplo [1].

Neste sentido, diversos estudos que buscam compreender como, onde e quando esse processo acontece vem crescendo. E para melhor compreender, o Sensoriamento Remoto se apresenta como um grande auxílio principalmente nas análises multitemporais, que buscam entender a dinâmica e evolução da cobertura vegetal e o uso da terra [2].

Para isso, diversas abordagens metodológicas vêm sendo utilizada para atingir uma detecção e classificação cada vez mais precisa dos Eucaliptos a partir das imagens de satélite.

Dentre as diversas técnicas de classificação do Sensoriamento Remoto. A “Análise de Imagens de Objetos Geográficos” (GEOBIA) não considera os elementos como pixels isolados, mas grupos deles (regiões de pixels). Com isso, obtém-se um significativo aumento no número de variáveis consideradas na classificação, uma vez que, além das informações espectrais, torna-se possível descrever cada região usando parâmetros em relação à forma do objeto (área, altura, largura, densidade, altura da borda, entre outros), à textura, às relações entre objetos vizinhos, entre outros [3] [4]

Porém, com o aumento no volume de imagens que pode ser agora automaticamente processado permite realizar análises com base na trajetória das variáveis físicas de cada pixel atravessando uma grande coleção de imagens de diferentes datas [5] [6]. O que permite analisar mais ampla do que o estado atual da imagem, mostrando o comportamento temporal do objeto de estudo. Considerando a resposta temporal gerada pelo manejo da silvicultura, é de se esperar um bom desempenho no uso de series temporais numa classificação com base em trajetória deste tipo de cobertura vegetal

Sendo assim, o presente trabalho tem por objetivo de comparar e analisar os resultados de duas abordagens distintas de classificação das áreas de Eucalipto no Estado

do Rio de Janeiro. Uma com base em GEOBIA com o uso do descritor de máxima diferença espectral (*Max. Diff*), disponível no sistema eCognition, e a outra com base em trajetória utilizando o algoritmo *Landtrendr*, ambas feitas com imagens do satélite *Landsat* (figura 1).

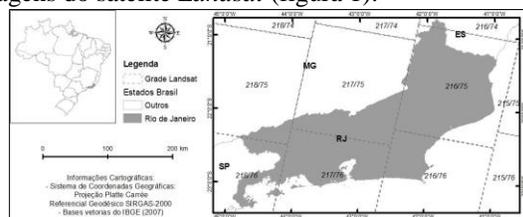


Figura 1 Localização do Estado do Rio de Janeiro e cobertura total de cenas Landsat TM.

2. MATERIAIS E MÉTODOS

Desde seu início, em 1972, com o sensor Multispectral Scanner e posteriormente com o advento dos sensores *Thematic Mapper* em 1982 e *Operational Land Imager* (OLI) em 2013, o imageamento pelos satélites *Landsat* tem sido praticamente contínuo e soluções técnicas têm sido implementadas para maximizar a resolução espacial e espectral assim como aperfeiçoar a consistência radiométrica entre os sensores [7]. O *Landsat 8* entrou em operação em 2013 com dois instrumentos imageadores: O *Operacional Terra Imager* (OLI) gera nove bandas multiespectrais com resolução espacial de 30 metros (bandas de 1 a 7 e de 9). Já a banda 8 do instrumento OLI é a pancromática e possui resolução espacial de 15 metros. O os produtos do *Thermal Infrared Sensor* (TIRS) equivalem às faixas do termal nas bandas 10 e 11 e têm pixels com 100 metros de resolução espacial [8], sendo úteis no fornecimento de temperaturas de superfície mais precisas.

Com base nas imagens *Landsat 8* do ano de 2014, o primeiro mapeamento foi realizado com uma abordagem orientada a objeto. O objetivo deste trabalho era classificar silvicultura no estado do Rio de Janeiro no ano de 2014. Para isso, o mapeamento e os testes de descritores foram realizados no software *eCognition 8.8*. Ao final, foi escolhido para a classificação o índice *Max. Diff*. (Máxima diferença do Brilho) está relacionado à maior diferença encontrada entre as refletâncias médias de cada objeto.

Foram definidos os melhores parâmetros de segmentação para atender à escala, utilizando-se o segmentador *multiresolution segmentation*. Foi considerada a escala como 50, a forma como 0.7 e a compacidade como 0.3. Já para a classificação, o modelo utilizado foi de lógica *fuzzy*, relacionados à probabilidade de pertencimento a uma determinada classe. Os intervalos de *Max. Diff*. variaram entre 0.55 e 0.85 [9].

Já o segundo mapeamento, também com imagens do satélite de *Landsat* foi realizada a partir de uma abordagem em trajetórias. O objetivo deste trabalho era mapear a cobertura florestal no estado do Rio de Janeiro entre os anos de 1984 a 2016. Para níveis de comparação, foi considerado tudo que o algoritmo classificou como recuperação.

Após a normalização radiométrica, foram utilizados dois algoritmos para a segmentação. Primeiramente o algoritmo *multiresolution segmentation*, com parâmetro de escala 30 e os critérios 0.1 de forma (ou seja, 0.9 para o fator cor) e 0.5 de compacidade (resultando em 0.5 para suavidade). Depois foi utilizado o algoritmo *spectral difference segmentation* com o valor espectral máximo de diferença igual a 120 pixels. A identificação das trajetórias evolutivas da paisagem foi realizada no *Envi*, para uma série temporal de NDV com resolução anual, segundo método desenvolvido por Kennedy *et al.* (2007) – *LandTrendr* [10]. A curva da série temporal anual de cada pixel, é particionada numa série de segmentos retos sequenciais (com uma duração arbitrária) que capturam tanto as perturbações abruptas quanto os processos de acumulação ou perda lenta [11].

Para a comparação dos métodos foi utilizado como referencia (verdade terrestre) os resultados do Projeto “*Olho No Verde*” da Secretaria Estadual do Ambiente, que gerou um mapeamento de alta resolução, com imagens do satélite *WorldView2* reamostradas para dois (2) metros do ano de 2014, do estado do Rio de Janeiro, sendo uma das classes utilizadas no projeto a silvicultura [12].

Além disso, também foi realizado o teste estatístico de McNemar que é explicado pela seguinte equação:

$$X^2 = \frac{(|N1 - N2| - 1)^2}{N1 + N2}$$

Onde N1 representa o número de vezes que o primeiro classificador classificou erroneamente e o segundo classificou corretamente, e N2 o número de vezes que o primeiro classificador classificou corretamente e o segundo classificou erroneamente. Assume-se a hipótese nula quando ambos classificadores tem desempenhos semelhantes, ou seja, se o valor de X^2 for maior do que $p = 0.025$, pode-se rejeitar a hipótese nula e concluir que os classificadores tem desempenhos distintos [16].

A acurácia de mapeamentos de dados globais é fundamental tanto para estabelecer a exatidão dos produtos para a comunidade científica. E diante disso, diversas técnicas de análises vêm sendo desenvolvidas, principalmente referentes a classificação a partir de imagens digitais.

No presente trabalho foi utilizada a matriz de confusão ou erro. Como uma tabulação cruzada simples, a matriz de confusão constitui a base sobre que descreve a precisão da classificação e caracteriza os erros, podendo ajudar a refinar a classificação ou estimativa derivada dela [13].

Com a finalidade de verificar a acurácia da classificação deste trabalho, foi realizada uma avaliação através do índice Kappa, utilizando como verdade terrestre pontos coletados no campo e mapas de referências com uma maior escala de detalhes. O coeficiente de acurácia Kappa para classes individuais, parece ser o mais indicado na

avaliação da classificação de imagens de sensoriamento remoto, por levar em conta os erros comissão e omissão [14].

Sendo assim, para a comparação entre mapeamentos realizados com imagens de satélite com resoluções espaciais diferentes foi adotado como área mínima para a comparação polígonos maiores que dois hectares.

3. RESULTADOS E DISCUSSÕES

Tanto os dois mapeamentos comparados, como a classificação de referência, foram gerados por métodos diferentes de classificação, gerando diferenças significativas no número de polígonos como é possível observar na figura 2. O mapeamento realizado com o descritor *Max Diff.* gerou 1481 polígonos e 1632 com o algoritmo *LandTredr.* Já a classificação de referência possui 2392 polígonos. Para o cálculo do índice Kappa foi levado em consideração as áreas em comum ou não, no lugar da tradicional contagem do número de acertos e erros.

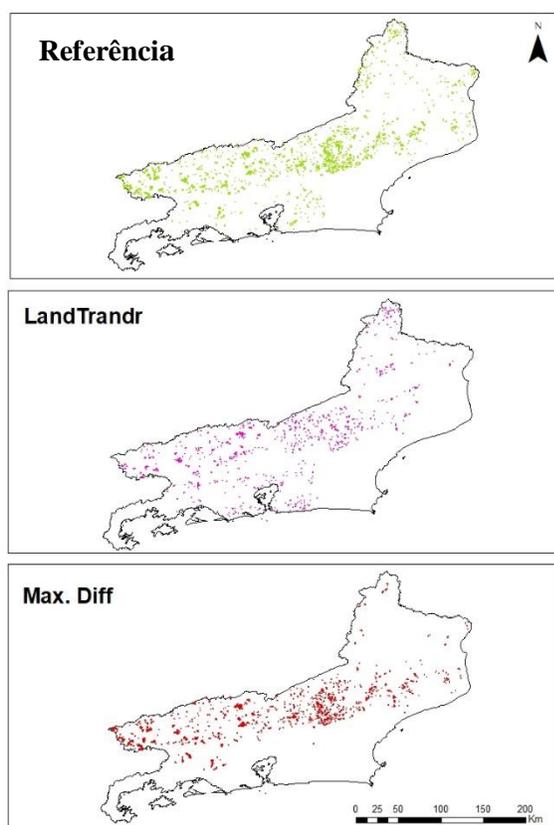


Figura 2: Comparação entre as áreas mapeadas nos mapeamentos utilizados.

Com o objetivo de melhor comparar os dois métodos de classificação, o orientado em objeto e o em trajetória evolutiva, foi realizado um outro teste estatístico. O teste de McNemar auxilia na compreensão do quanto se aproxima ou se afastam os mapeamentos.

A comparação é feita entre o que os dois mapeamentos classificam como certo, ao serem comparados com a referência. O que um acertou, mas o outro errou, e o que

ambos erraram (tabela 1). Diferente do índice kappa, que levou em consideração a área, para este foram utilizados os números dos polígonos.

Tabela 1: Matriz de contingência 2x2 para um total de n=3628 polígonos,

| LandTredr | Max Diff. | |
|-----------|-----------|--------|
| | Certo | Errado |
| | Certo | 1436 |
| Errado | 1144 | 956 |

O cálculo do teste estatístico foi realizado no software R com base na matriz apresentada anteriormente. Sendo assim, o resultado do X^2 encontrado foi de 893,69, com um valor de $p < 2.2e-16$. O que levou a rejeitar a hipótese inicial de não diferença entre os métodos comparados.

Os resultados do índice Kappa permite observar o quão próximo da verdade os mapeamentos estão e assim melhor analisar os erros. Para a comparação foram consideradas as áreas de interseção dos polígonos, em hectare. No mapeamento de referência foram encontrados 33.079,21 ha de eucalipto, e 27.467,00 ha no mapeamento realizado com o Max. Diff. Já no mapeamento com o algoritmo LandTredr, a área encontrada foi de 24.496,46. Para as áreas de Não-Eucalipto foram considerados o restante de área do estado. As tabelas 2 e 3 apresentam as respectivas matrizes de confusão:

Tabela 2: Matriz de Confusão da classificação pela máxima diferença do brilho (Max.Diff)

| Referência | Classificação | | |
|-------------------|---------------|-----------------|---------|
| | Eucalipto | Não - Eucalipto | Total |
| Eucalipto | 25364,33 | 2102,67 | 27467 |
| Não - Eucalipto | 5612,21 | 4339600 | 4345212 |
| Total: | 30976,54 | 4341702,67 | 4372679 |
| Erro de Comissão: | 12,8% | | |
| Erro de Omissão: | 4,8% | | |
| Exatidão Total: | 99,8% | | |
| Índice Kappa: | 0,867 | | |

Tabela 3: Matriz de Confusão da classificação pelo algoritmo LandTredr

| Referência | Classificação | | |
|-------------------|---------------|-----------------|----------|
| | Eucalipto | Não - Eucalipto | Total |
| Eucalipto | 5215,63 | 27863,58 | 33079,21 |
| Não - Eucalipto | 8582,75 | 4374582,72 | 4383165 |
| total: | 13798,38 | 4402446,3 | 4416245 |
| Erro de Comissão: | 19,43% | | |
| Erro de Omissão: | 63,09% | | |
| Exatidão Total: | 46,1% | | |
| Índice Kappa: | 0,219 | | |

Ao comparar os resultados da classificação baseada em objeto com o uso do descritor da Máxima Diferença (*Max Diff.*) foi encontrado uma exatidão total de 99,8% e um kappa de 0,867. Já no mapeamento com base em trajetórias realizado, com o algoritmo LandTredr, foi encontrado uma exatidão total de 46,1% e um kappa de 0,219.

De acordo com a literatura, o resultado do índice Kappa varia de -1 a 1e quanto mais próximo de 1, maior a acurácia da classificação. A qualidade da classificação pode ser definida em função dos seguintes intervalos de índice Kappa: [$-0,00$]: muito ruim; [$0,00 - < 0,20$]: ruim; [$0,20 - < 0,40$]: aceitável; [$0,40 - < 0,60$]: bom; [$0,60 - < 0,80$]: muito bom; [$0,80 - 1,00$]: excelente [16] [17]. Sendo assim, o Kappa obtido entre a verdade e a *Max. Diff.* é considerado excelente (0,867) e o gerado entre a verdade e a LandTrendr é considerado apenas razoável (0,291)

A classificação pela *Max. Diff.*, apresenta também menores erros de comissão (12,8%) e omissão (4,8%), comparados com o resultado do algoritmo Landtrandr, (19,43% e 63,09%, respectivamente). Em ambos os mapeamentos não há um padrão de erros encontrados. Os erros de comissão correspondem às confusões entre as áreas de eucalipto e as áreas de florestas mais homogêneas. Já os erros de omissão predominaram nas áreas de encostas, dificultado pelo sombreamento e pela diferença de resoluções entre as imagens Landsat e as imagens de referência (WorldView2).

Portanto, o método de classificação pelo descritor de máxima diferença do brilho foi o que apresentou melhor desempenho na classificação. É possível notar que nem tudo que o algoritmo Landtrandr considera como revegetação pode ser classificado como reflorestamento por eucalipto. Os resultados indicam que é necessário redefinir os parâmetros e modelos de trajetórias de revegetação que correspondam ao comportamento temporal da silvicultura sob manejo.

4. CONCLUSÕES

O estudo testou e comparou o desempenho de duas classificações com abordagens diferentes: uma baseada na análise orientada a objeto, e outra baseada na trajetória evolutiva. Tendo como objeto as áreas plantadas de Eucalipto no estado do Rio de Janeiro realizadas com imagens do satélite Landsat.

Ambas técnicas constituem abordagens bem diferentes: em primeiro lugar, uma considera apenas um momento das cenas (o ano de 2014) e a outra trabalha com a série temporal das cenas (de 1984 a 2016), em segundo lugar, uma tem com o objetivo classificar diretamente as áreas de silvicultura e a outra identificar trajetória que incluam revegetação pós-distúrbio, Isto ficou evidente no teste não paramétrico de comparação entre os métodos.

O resultado pelo descritor *Max. Diff.* Implementado na classificação GEOBIA teve melhor desempenho do que a classificação com base na trajetória dos valores de NDVI, do algoritmo Landtrandr. No entanto, o resultado não permite descartar o potencial de classificação de áreas sob manejo de silvicultura e deixam novos desafios na modelagem da detecção de mudanças na cobertura vegetal através de séries temporais.

6. REFERÊNCIAS

[1] Buainaim, A.M. & Batalha, M.O. (Coord.). “Cadeia produtiva de madeira”. Ministério da Agricultura, Pecuária e Abastecimento,

Secretaria de Política Agrícola, Instituto Interamericano de Cooperação para a Agricultura, p.84.2007.

[2] Jaackson e Jensen,. “An Evolution of Remote Sensing derived Landscape Ecology Metrics for Reservoir Shoreline Environmental Monitoring”, *Photogrammetric Engineering & Remote Sensing*, 71 (12): 1387-1397.2005.

[3] Blaschke, T. “Object based image analysis for remote sensing”. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 65, pp. 2-16. 2010.

[4] Weckmüller, R.; Vicens, R.S. “Método híbrido de detecção de mudanças: uma associação entre classificação baseada em objetos e baseada em pixels”. *Revista Brasileira de Cartografia*, v. 68 n. 05, p. 883-899. 2016.

[5] Kennedy, R.E., Yang, Z., & Cohen, W.B. “Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr --Temporal segmentation algorithms.” *Remote Sensing of Environment*, 114, 2897-2910. 2010.

[6] Fragal, E.H., Silva, T.S.F. & Novo, E.M.L.M. “Reconstructing historical forest cover change in the Lower Amazon oodplains using the LandTrendr algorithm.” *Acta Amazonica*, 46 (1), 13-24. 2016.

[7] NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA). Landsat Science. Disponível: <<https://landsat.gsfc.nasa.gov/landsat-8/>> Acessado em: 08 de outubro de 2018.

[8] UNITED SERVICE GEOLOGICAL SURVEY (USGS). Earth Explorer. Disponível em: <<https://earthexplorer.usgs.gov/>> Acessado em: 16 de outubro de 2018.

[9] Ferraz, D.P.G.B. “Dinâmica espaço-temporal da produção de silvicultura no estado do Rio de Janeiro entre 1985 e 2014.” POSGEO, *Dissertação de Mestrado em Geografia*, UFF. Niterói, RJ. 83 p. 2017.

[10] Kennedy, R.E., Cohen, W.B., & Schroeder, T.A. “Trajectory-based change detection for automated characterization of forest disturbance dynamics”. *Remote Sensing of Environment*, 110, 370-386. 2007.

[11] Weckmüller, R. Trajetórias Evolutivas das Coberturas Florestais do Estado do Rio de Janeiro utilizando os algoritmos LANDTRENDR. POSGEO, *Tese de Doutorado em Geografia*, UFF. Niterói, RJ, 106 p. 2018.

[12] Laboratório ESPAÇO de Sensoriamento Remoto e Estudos Ambientais – UFRJ. Mapa de Cobertura da Terra e Detecção de Mudanças na Cobertura Florestal do estado do Rio de Janeiro. Escala 1:25.000. Ano base 2014+-1. Projeto Olho no Verde. SEA-RJ.

[13] Congalton, R. G.; Green, K. “A practical look at the sources of confusion in error matrix generation.” *Photogrammetric Engineering & Remote Sensing*, v. 59, n. 5, p. 641–644, 1993.

[14] XU, Y.; SHI, J.; DU, J. “An improved endmember selection method based on vector length for MODIS reflectance channels.” *Remote Sensing*, v. 7, n. 5, p. 6280–6295, 2015.

[15] Foody, G. M. “Thematic Map Comparison : Evaluating the Statistical Significance of Differences in Classification Accuracy.”. *Photogrammetric Engineering & Remote Sensing*, v. 70, n. 5, p. 627–633, 2004

[16] Landis, J.; Koch, G. “The measurement of observer agreement for categorical data.”. *Biometrics*, v. 33, p. 159-174, 1977.

[17] Francisco, C. e Almeida, C., “Interpretação de imagens orbitais por meio de sistema especialista para o mapeamento de cobertura da terra em região montanhosa.” *Soc. & Nat.*, Uberlândia, ano 24 n. 2, 283-302, mai/ago. 2012.