

ALGORITMO PARA COMPOSIÇÃO DE SÉRIES TEMPORAIS LANDSAT PARA DETECÇÃO DE MUDANÇAS EM COBERTURAS AGRÍCOLAS

Eduardo Ribeiro Lacerda¹, Raúl Sánchez Vicens²

¹Doutorando em Geografia da Universidade Federal Fluminense, elacerda@id.uff.br; ²Prof. Dr. do departamento de Geografia da Universidade Federal Fluminense, rsvicens@gmail.com

RESUMO

A análise de séries temporais em trabalhos de sensoriamento remoto vem ganhando cada vez mais força ao longo dos últimos anos devido ao aumento da capacidade de processamento e armazenamento dos dados, e também como consequência do surgimento de novas tecnologias que vem facilitando este tipo de análise. O Google Earth Engine surge neste contexto como uma dessas tecnologias, dando possibilidades a pesquisadores que antes não possuíam a estrutura necessária para realizar análises de grande porte. Sendo assim, utilizando ferramentas como a citada, o presente trabalho busca contribuir com um método de criação de séries temporais para análise de cultivos através de composições de imagens de média resolução utilizando a plataforma do Google com o objetivo de beneficiar trabalhos que busquem maior detalhamento espacial sem a necessidade de perda de informação devido a resolução temporal dos satélites Landsat.

Palavras-chave — landsat, séries temporais, composite, google earth engine.

ABSTRACT

The analysis of time series data in remote sensing papers has been gaining strength in recent years due to the increase in data processing and storage capacity, as well as the emergence of new technologies that have facilitated this type of analysis. Google Earth Engine appears in this context as one of these technologies, giving possibilities to researchers who previously did not have the necessary structure to perform large analyzes. Thus, using tools such as this one, the present work seeks to contribute with a method of creating time series for crop analysis through compositions of medium resolution images using the Google platform in order to benefit works that seek greater spatial detail without the need for loss of information due to the time resolution of the Landsat satellites.

Key words — landsat, time series, composite, google earth engine.

1. INTRODUÇÃO

O estudo de séries temporais voltadas para a análise da produção agrícola tem uma longa história no sensoriamento remoto, e apesar de ter dividido as atenções devido ao surgimento de outros tipos de análise ao longo do tempo, vem retomando a sua importância nos últimos anos devido ao desenvolvimento de novas tecnologias e também por conta do aumento na capacidade de armazenamento e poder computacional disponíveis. Ao mesmo tempo, tecnologias como o Google Earth Engine [1] surgiram dando a possibilidade antes quase que inexistente de processar uma quantidade realmente grande de dados geoespaciais sem possuir o hardware necessário para tal. Isso vem possibilitando estudos em escalas não só geográficas como temporais cada vez maiores por um número cada vez maior de pesquisadores, incluindo pesquisas ligadas a laboratórios e centros de pesquisas em países periféricos onde o investimento em pesquisa normalmente é mais reduzido e consequentemente sofrem com as consequentes limitações estruturais.

A análise de séries temporais na agricultura com o objetivo de entender melhor as mudanças no uso do solo tem ligação direta com questões envolvendo segurança alimentar e planejamento estratégico por parte dos mesmos centros de pesquisa e de organizações governamentais. No entanto, grande parte das análises existentes atualmente utilizam dados derivados de satélites como o MODIS que possui uma resolução temporal diária, o que permite uma melhor compreensão das variações espectrais dessas áreas. Além disso, o menor tempo de revisita do satélite significa uma maior quantidade de imagens disponíveis e uma maior chance de se criar composições sem a presença de nuvens. No entanto, satélites como o MODIS possuem uma baixa resolução espacial, gerando imagens muito agregadas e pouco detalhadas, além de possuir uma série temporal disponível somente a partir do ano 2000, o que limita algumas análises. Já as imagens derivadas da série Landsat possuem melhor resolução espacial e espectral, assim como uma série temporal constante desde a década de 1970. No entanto, possuem um tempo de revisita de 16 dias, o que impossibilita este tipo de estudo.

Considerando então essas limitações, este trabalho busca contribuir com o desenvolvimento de uma metodologia que busque reduzir as mesmas através da utilização de novas tecnologias. O trabalho tem como objetivo específico criar composições com imagens Landsat de uma forma que as limitações do próprio satélite possam

ser diminuídas buscando perder o mínimo de informação possível, possibilitando assim que estudos sobre produtividade agrícola possam ser realizados para décadas anteriores ao lançamento do MODIS e/ou para estudos que necessitem de um melhor detalhamento, como para as áreas escolhidas para este estudo em específico no estado do Rio de Janeiro, onde muitas das áreas produtivas estão em propriedades de pequeno porte

2. MATERIAIS E MÉTODOS

2.1. Área de Estudo

A área de estudo se localiza no município de São Francisco de Itabapoana, no estado do Rio de Janeiro. Possui uma área de 122,438 quilômetros quadrados e população de 47,247 habitantes. A principal base da economia municipal se baseia na lavoura canavieira, da pecuária de corte e leiteira, da fruticultura e da pesca. Parte dos pontos coletados em campo que indicavam áreas de interesse com plantações de cana de açúcar, abacaxi e silvicultura foram utilizados neste trabalho. Os pontos escolhidos foram justamente os de áreas onde houveram mudanças no tipo de produção, em especial, da cana de açúcar para o plantio de abacaxi.

2.2. Métodos

O método se baseia na criação de composições de imagens de média resolução espacial derivadas dos satélites da série Landsat utilizando a plataforma Google Earth Engine (GEE). Todo o código gerado foi escrito utilizando a API JavaScript da plataforma devido à facilidade de visualização dos dados principalmente para a verificação de erros.

Após o acesso à plataforma através de um registro gratuito no próprio website, inicia-se o processo de codificação e criação das variáveis e funções. Inicialmente foram criadas variáveis que pudessem ser utilizadas como delimitadoras da área de estudo. Sendo assim, criou-se um ponto em formato vetorial para especificar a área de estudo. Este ponto delimita qual *path/row* Landsat será processado. A coordenada escolhida para o processamento dos dados deste trabalho foi um ponto de longitude -41.7291 e latitude -21.6867. Após isto, uma lista contendo todos os anos a serem processados foi criada. Inicialmente foram escolhidos os anos de 1988 até 2012 devido primeiramente a maior facilidade de se testar o método utilizando somente imagens derivadas do sensor TM do satélite Landsat 5, e também por limitações presentes na quantidade de imagens sem nuvem disponíveis de anos anteriores.

Uma função para o cálculo do índice de vegetação NDVI (*Normalized Difference Vegetation Index*) foi criada para processar de forma mais prática todas as imagens selecionadas. A função utiliza as bandas espectrais da imagem de entrada para criar o índice e posteriormente adiciona o mesmo como uma nova banda na imagem original. Além disso, uma outra função foi implementada para filtrar as imagens e retirar as nuvens presentes.

Utilizou-se a função *simpleCloudScore* já presente na plataforma para criar uma banda com informações sobre a presença de pixels com nuvem.

Foram então criadas máscaras (*raster*) onde apenas pixels com quantidade igual ou inferior a 20% foram selecionadas. Os pixels com valor superior passaram a ter valor “no data”.

Após esta primeira etapa, foram então criadas funções para filtrar a série temporal contendo todas as imagens Landsat 5 registradas para a área de estudo escolhida em períodos de dois em dois meses. A função cria primeiramente composições anuais (1988 - 2012) selecionando todas as imagens disponíveis em cada ano de dois em dois meses (Janeiro/Fevereiro, Março/Abril, Maio/Junho, Julho/Agosto, Setembro/Outubro, Novembro/Dezembro). Após selecionar as imagens disponíveis para cada dupla de meses, foram aplicadas as funções para a criação dos índices e de máscara de nuvens. Por fim, um composite é gerado aplicando-se uma função que calcula o valor mediano entre todos os pixels válidos (diferentes de *no data*) em todas as imagens disponíveis para aquele ano, dentro do período de dois meses. Foram então criadas listas (*arrays*) para armazenar as composições criadas para cada dupla de meses. Sendo assim, todas as composições de Janeiro/Fevereiro para cada ano entre 1988 e 2012 foram armazenadas em uma lista, já as geradas para o período entre Março/Abril em uma outra, e assim por diante.

Seis listas foram criadas representando cada uma dois meses do ano. A partir daí, novas composições de imagens foram geradas criando agrupamentos de 5 em 5 anos. Esta etapa foi feita utilizando novamente a função “*median*” do GEE, que cria uma única imagem escolhendo os valores medianos para cada pixel das imagens de entrada. Isto foi necessário por conta das limitações na resolução temporal dos satélites Landsat. Apesar de terem sido feitas composições anteriores juntando-se o equivalente a 2 meses de imagens ano a ano, muitos “buracos” derivados de imagens com muitas nuvens ainda eram presentes, o que dificultaria a realização de análises temporais. Com isso, optou-se pela criação de novas composições de 5 em 5 anos, ao invés de composições anuais. Esta escolha se deu por conta do trabalho desenvolvido por Patrick Griffiths [2], onde composições pixel à pixel para imagens Landsat foram feitas com este mesmo intervalo e demonstraram bons resultados.

Após a criação das composições finais, dois pontos de interesse em formato vetorial foram importados para dentro da plataforma para a verificação dos resultados. Estes pontos foram coletados originalmente em campo pelos próprios autores. Os pontos de interesse importados contém a localização de propriedades com produção de cana de açúcar/abacaxi.

Com os pontos presentes na plataforma foi possível assim extrair os valores dos locais de interesse para todas as composições e plotar os mesmo em gráficos. Os valores

extraídos se referem aos valores da imagem NDVI criada anteriormente. Os gráficos gerados foram então posteriormente analisados pelos autores e comparados com outros gerados para as mesmas áreas e interesse, mas utilizando a plataforma web disponibilizada pelo Instituto Nacional de Pesquisas Espaciais (INPE) que contém dados de séries temporais MODIS.

3. RESULTADOS

Apesar de conter resoluções espaciais diferentes, as composições Landsat geradas por este estudo e as séries temporais MODIS presentes na ferramenta web do INPE, mostraram semelhanças em suas respostas não só no comportamento fenológico das culturas como também na detecção de mudanças de plantio das mesmas.

Os valores extraídos mostraram padrões interessantes, principalmente no último quinquênio. Foi justamente a partir desse último período da série que houve a mudança no uso do solo. Neste caso em específico, do plantio da cana para o do abacaxi. Este padrão se mostrou presente nos dois pontos de interesse como podemos observar nas figuras 1 e 2 respectivamente:

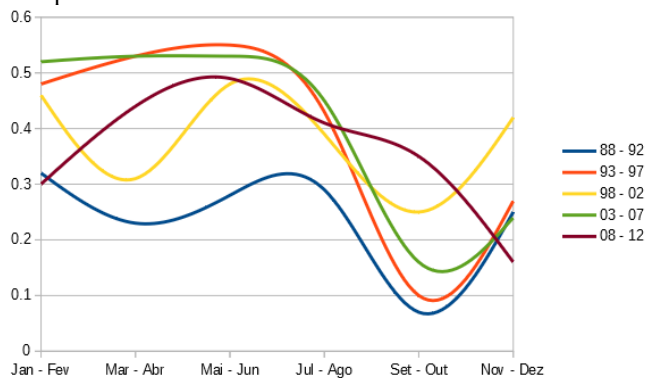


Figura 1. Gráfico com a resposta espectral das composições para cada quinquênio estudado na área 1. Cada quinquênio é representado com uma cor. O eixo x apresenta as duplas de meses e o eixo y os valores do NDVI.

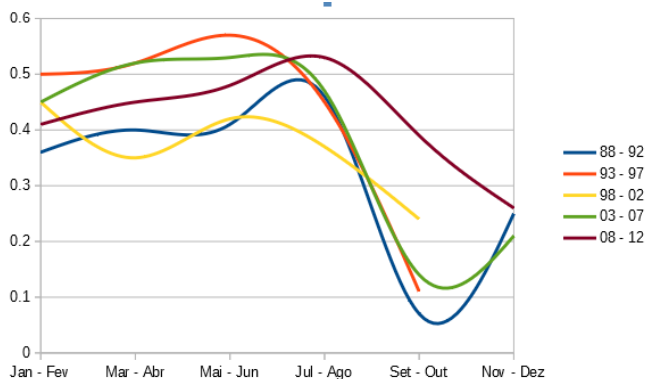


Figura 2. Gráfico com a resposta espectral das composições para cada quinquênio estudado na área 2. Cada quinquênio é representado com uma cor. O eixo x apresenta as duplas de meses e o eixo y os valores do NDVI.

Como podemos observar, mesmo com a técnica aqui apresentada, ainda sim houveram espaços com nuvem presente mesmo após a composição das imagens. No entanto, de forma bem menos presente que o normal e esperado em séries temporais Landsat.

Além disso, ao analisar as séries temporais ao longo das duplas de meses estudadas e durante os 24 anos, podemos observar melhor (Figura 3 e 4) como os dois pontos seguem o mesmo padrão:

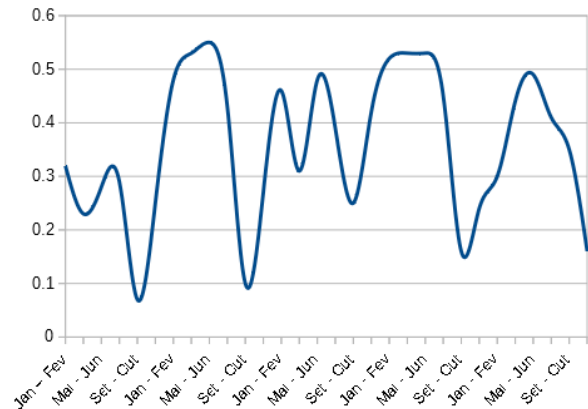


Figura 3. Comportamento das respostas para o ponto 1 ao longo dos 24 anos (1988 - 2012).

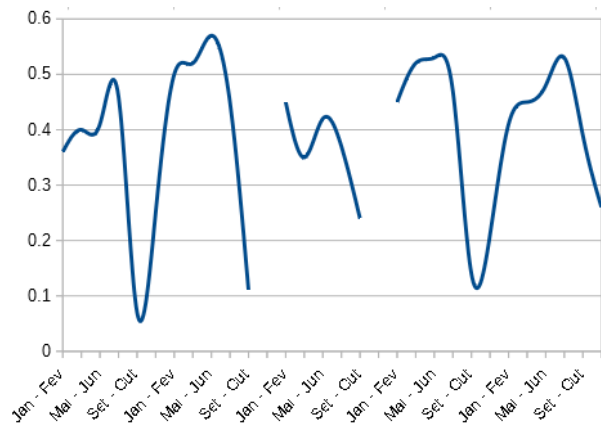


Figura 4. Comportamento das respostas para o ponto 2 ao longo dos 24 anos (1988 - 2012).

Novamente, podemos observar que a falta de dados está presente, mas não o suficiente para comprometer a análise final.

Quando comparamos os resultados obtidos com a curva MODIS do web app do INPE, podemos ver claramente não só uma semelhança na resposta encontrada até o ano de 2008 (Figura 5) como também uma mudança significativa a partir do mesmo ano (Figura 6), onde a curva apresentou uma mudança mais significativa:

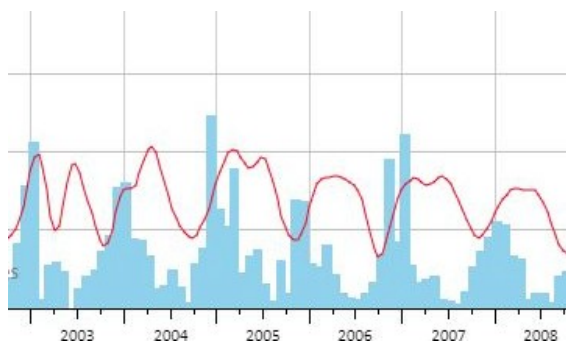


Figura 5. Comportamento espectral do ponto de interesse do ano 2003 até 2008.

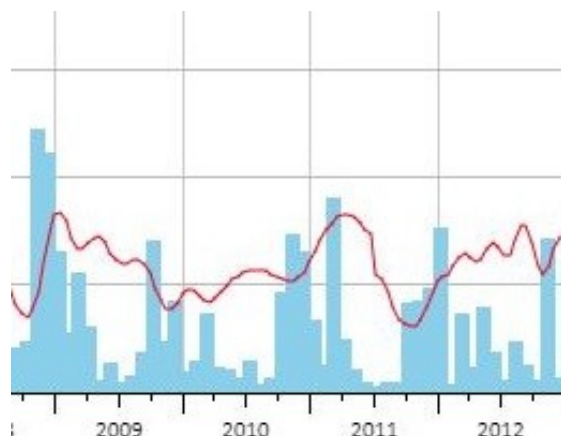


Figura 6. Comportamento espectral do mesmo ponto de interesse do ano 2008 até 2012.

Como podemos observar, a partir do ano de 2008, o tipo de uso do solo foi modificado, deixando de apresentar um comportamento típico de plantio da cana de açúcar e apresentando um comportamento bastante condizente com o do plantio do abacaxi, posteriormente confirmado em campo. Além disso, os resultados nesta área de estudo mostraram coerência mesmo quando comparados a dados monitorados em escalas temporais muito mais detalhadas, o que demonstra que possivelmente é possível encontrar padrões temporais e detectar mudanças específicas do uso do solo em períodos anteriores ao de imageamento do satélite MODIS.

4. DISCUSSÃO

Esperamos que este trabalho possa contribuir metodologicamente para que outros estudos que busquem entender a dinâmica do uso do solo em áreas menores e/ou em períodos anteriores ao ano 2000 possam ser realizados com maior detalhamento.

Este tipo de situação, é bastante comum, por exemplo, no estado do Rio de Janeiro, onde muitas das terras que hoje estão abandonadas e constituem um problema a ser enfrentado pelos próximos governos, começaram neste processo na década de 1980, quando não tínhamos tantas

fontes orbitais de informação disponíveis. Além disso, muita das propriedades em questão são de pequeno porte, o que dificulta a identificação de padrões com pixels de baixa resolução espacial.

Como o todo o processamento foi realizado na plataforma GEE, após o desenvolvimento do código-fonte, os testes e o processamento em si puderam ser realizados em pouquíssimos segundos, o que serve para contribuir com estudos comparativos e em outras áreas de interesse. Sendo assim, a escolha de realizar todo o estudo na plataforma se deu também pela alto nível de replicabilidade que a ferramenta possibilita.

No entanto, algumas limitações foram identificadas durante o processo. Mesmo com a criação dos composites quinquenais e com a escolha de se trabalhar com períodos de dois em dois meses, ainda sim existem “buracos” na imagem, onde o método não conseguiu agregar nenhum tipo de informação devido a alta quantidade de nuvens.

Além disso, ainda é necessário realizar o mesmo teste para outras áreas com outros tipos de cultivo para verificar se o método realmente é eficiente na identificação de outros cultivos. E também testar o resultado obtido utilizando outros índices de vegetação com o EVI (Enhanced Vegetation Index).

Outra limitação, é a dificuldade de se identificar exatamente o ano em que uma mudança realmente ocorreu quando se trabalha com quinquênios ao invés de séries anuais. Pequenas perturbações podem passar sem serem percebidas dependendo do que se está analisando.

Para o futuro planejamos verificar ainda mais o comportamento do plantio do abacaxi na região através da inserção de imagens do sensor OLI (Landsat 8) no estudo para a adição do último quinquênio (2013 – 2017) e posterior comparação com os dados do MODIS, mas ainda é necessário implementar funções específicas para processar dados deste sensor.

5. CONCLUSÕES

Entendemos que o método aqui apresentando representa uma nova possibilidade de análise para muitas áreas que não poderiam ser contempladas por satélites de baixa resolução e para sua aplicação em basicamente qualquer parte do mundo, já que foi totalmente realizado dentro da plataforma do GEE.

Agradecemos ainda o apoio financeiro da CAPES, o que possibilitou a realização deste trabalho.

6. REFERÊNCIAS

- [1] Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, n. 2016, 2017. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0034425717302900>.
- [2] Griffiths, P.; Linden van der S.; Kuemmerle, T.; Hostert, P. A Pixel-Based Landsat Compositing Algorithm for Large Area Land Cover Mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. Volume:6, Issue:5, Oct. 2013.