Disciplina: Geoinformática
Profª. Dra. Karine Ferreira

APRESENTAÇÃO DE ARTIGO

remote sensing

Article

Satellite Image Time Series Analysis for Big Earth Observation Data

Rolf Simoes [1,*], Gilberto Camara [1], Gilberto Queiroz [1], Felipe Souza [1], Pedro R. Andrade [1], Lorena Santos [1], Alexandre Carvalho [2] and Karine Ferreira [1]

https://doi.org/10.3390/rs13132428

Aluna: Raquel Zózimo Molinez

São José dos Campos – SP, Abril de 2025

Introdução

- Grandes volumes de dados de satélites disponível;
- Séries temporais revelam a dinâmica que imagens isoladas não mostram;
    - distúrbios florestais; mudanças no uso da terra; dinâmica ecológica; intensificação agrícola; monitoramento do desmatamento.

“espaço-primeiro, tempo-depois” (Camara, et al. 2016)

- Necessidade de novas soluções;

“tempo-primeiro, espaço-depois” (Camara, et al. 2016)

OBJETIVO:

(i) Este artigo descreve o sits, um pacote R de código aberto para análise de séries temporais de imagens de satélite utilizando aprendizado de máquina;

(ii) Estudo de caso no bioma Cerrado, uma das frentes agrícolas em rápida expansão no mundo, análise entre o ano de 2017e 2018.

Índice

- ✓ Cubos de Dados e Observação da Terra
- ✓ Projeto do Software e Método de Análise
- ✓ Exemplo concreto Cerrado Brasileiro
- ✓ Conclusões

remote sensing

MDPI

Article

Satellite Image Time Series Analysis for Big Earth Observation Data

Rolf Simoes [1,*], Gilberto Camara [1], Gilberto Queiroz [1], Felipe Souza [1], Pedro R. Andrade [1], Lorena Santos [1], Alexandre Carvalho [2] and Karine Ferreira [1]

1. National Institute for Space Research (INPE), Avenida dos Astronautas, 1758, Jardim da Granja, Sao Jose dos Campos, SP 12227-010, Brazil; gilberto.camara@inpe.br (G.C.); gilberto.queiroz@inpe.br (G.Q.); felipe.souza@inpe.br (F.S.); pedro.andrade@inpe.br (P.R.A.); lorena.santos@inpe.br (L.S.); karine.ferreira@inpe.br (K.F.)
2. National Institute for Applied Economics Research, SBS, Quadra 1 Bloco J, Brasília, DF 70076-900, Brazil; alexandre.ywata@ipea.gov.br

* Correspondence: rolf.simoes@inpe.br

Abstract: The development of analytical software for big Earth observation data faces several challenges. Designers need to balance between conflicting factors. Solutions that are efficient for specific hardware architectures can not be used in other environments. Packages that work on generic hardware and open standards will not have the same performance as dedicated solutions. Software that assumes that its users are computer programmers are flexible but may be difficult to learn for a wide audience. This paper describes `sits`, an open-source R package for satellite image time series analysis using machine learning. To allow experts to use satellite imagery to the fullest extent, `sits` adopts a time-first, space-later approach. It supports the complete cycle of data analysis for land classification. Its API provides a simple but powerful set of functions. The software works in different cloud computing environments. Satellite image time series are input to machine learning classifiers, and the results are post-processed using spatial smoothing. Since machine learning methods need accurate training data, `sits` includes methods for quality assessment of training samples. The software also provides methods for validation and accuracy measurement. The package thus comprises a production environment for big EO data analysis. We show that this approach produces high accuracy for land use and land cover maps through a case study in the Cerrado biome, one of the world's fast moving agricultural frontiers for the year 2018.

Keywords: big Earth observation data; data cubes; satellite image time series; machine learning and deep learning for remote sensing; R package

check for updates

Citation: Simoes, R.; Camara, G.; Queiroz, G.; Souza, F.; Andrade, P.R.; Santos, L.; Carvalho, A.; Ferreira, K. Satellite Image Time Series Analysis for Big Earth Observation Data. *Remote Sens.* **2021**, *13*, 2428. https://doi.org/10.3390/rs13132428

Academic Editor: Pieter Kempeneers

Received: 29 April 2021
Accepted: 9 June 2021
Published: 22 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

The growing demand for natural resources has caused major environmental impacts and is changing landscapes everywhere. Conversion of land cover due to human use is one of the key factors behind greenhouse gas emissions and biodiversity loss [1]. Spatial quantification of land use and land cover change allows societies to understand the extent of these impacts. Satellites are required to generate land cover products, since they provide a consistent, periodic, and globally reaching coverage of the planet's surface. Thus, satellite-based land cover products are essential to support evidence-based policies that promote sustainability.

There is currently an extensive amount of Earth observation (EO) data collected by an increasing number of satellites. Coupled with the adoption of open data policies by most spatial agencies, an unprecedented amount of satellite data is now publicly available [2]. This has brought a significant challenge for researchers and developers of geospatial technologies: how to design and build technologies that allow the Earth observation community to analyse big data sets?

Cubos de Dados e Observação da Terra

- O que são Cubos de dados (Data Cube EO);
  - abordagem conceitual, seguindo a ideia de cubos de dados EO como **Campos geográficos** (Galton, A., 2004; Camara, G. et al., 2014)

  "E) Não há lacunas ou valores ausentes na extensão tempo-espaço."

- Diferença de coleção ARD (analysis-read data) e Cubos regulares;

ARD – Corrigido, Reprojetado e Recortado

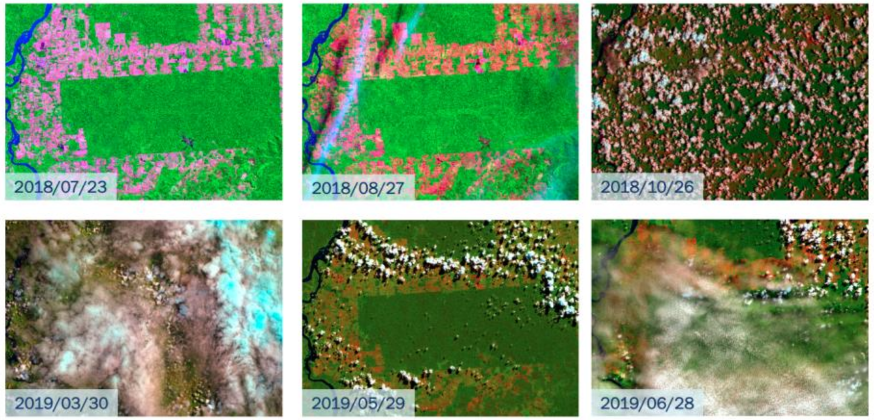

Figure 2. Sentinel-2 image colour composites for tile 20LKP on different dates (source: authors).

- Linha do tempo de diferentes mosaicos;

Sentinel-2, "20LLP"		Sentinel-2, "20LKP"
144 INSTANTES TEMPORAIS	≠	71 INSTANTES TEMPORAIS

O software de análise deve impor uma linha do tempo única para todos os mosaicos.

- Desafios de se trabalhar com séries temporais (ausência de gaps, continuidade).

Projeto do Software e Método de Análise

- **Modular e Intermodular**

- **Fluxo de trabalho: Modular**
  - 1. Criação do cubo de dados. (coleção ARD ou catalogo STAC) ;
  - 2. Extração de series temporais;
    - ✓ Controle de qualidade das amostras (SOM);
  - 3. Treinamento de modelos
    - ✓ (Random Forest, SVM, TempCNN, ResNet);
  - 4. Classificação do cubo.
  - 5. Suavização espacial bayesiana.
  - 6. Validação com amostragem ponderada.

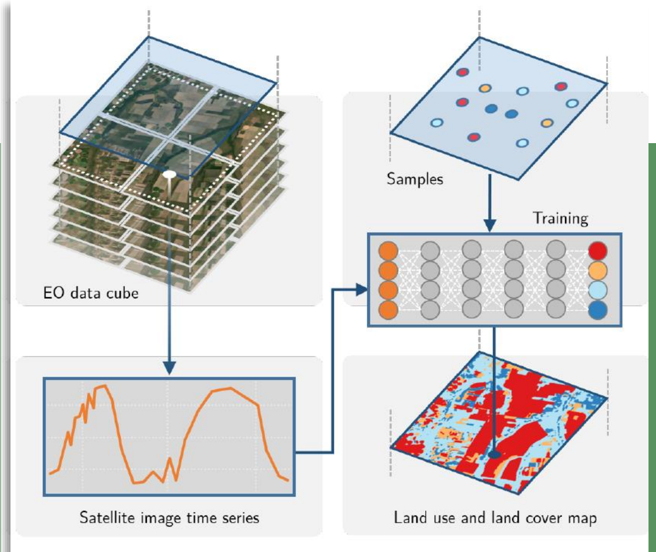

Figure 3. Using time series for land classification (source: authors).

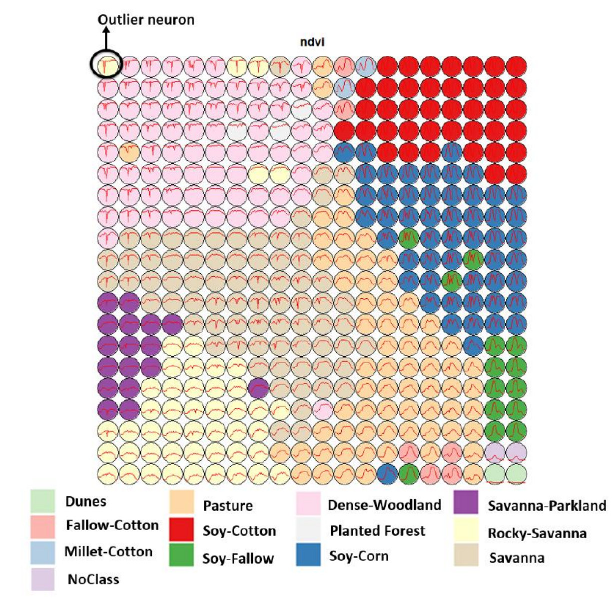

Figure 5. SOM map for Cerrado training samples (source: authors).

Estudo de Caso – Cerrado Brasileiro

- Contexto Cerrado brasileiro:
  - Extensão de **1,9 milhão de km²** (Klink, C. et al, 2005);
  - Mais de **7.000 espécies** de plantas (Klink, C. et al, 2005);
  - Tipos de vegetação: **Cerrado Ralo** (gramíneas, arbustos); **Cerrado Sensu Stricto** (arvores de tronco fino); **Cerradão** (floresta seca) (Goodland, R., 1971; Del-Claro, K. et al, 2019);
  - Fronteiras agrícolas mais devastadas (Walter, B.M.T., 2006).
- Dados:
  - **Landsat-8**, analise 09/**2017–**08/**2018** (calendário agrícola) – **8TB.**
  - Amostras (vegetação; área agrícola) **85.026 -> 48.850 amostras** (classes: MapBiomas, Pastagem.org, IBGE).
- Treinamento / Classificação:
  - TempCNN (conf. Padrão – 3 camadas);
  - **7** Bandas**; NDVI**, **EVI**, **NBR.**
- Pós classificação:
  - - Bayesiano;
  - - Rotulação.

"Lavoura Anual" (6.887)

"Cerradão" (4.211),

"Cerrado" (16.251),

"Área Natural Não Vegetada" (38)

"Cerrado Ralo" (5.658)

"Pastagem" (12.894),

"Lavoura Perenes" (68)

"Silvicultura" (805)

"Cana-de-açúcar"(1.775) "Água" (263).

Estudo de Caso – Cerrado Brasileiro

- Acurácia do mapa
  - Amostragem independente de 5.402 amostras.
  - Técnica ponderada por área (Olofsson, P. et al, 2013/2014);
  - Acurácia geral: 86%.

Table 1. Area-weighted classification accuracy.

Labels	Producer's Accuracy	User's Accuracy
Annual Crop	0.81	0.88
Cerrado	0.89	0.91
Natural Non Vegetated	0.63	0.95
Pasture	0.82	0.76
Perennial Crop	0.51	0.74
Silviculture	0.83	0.91
Sugarcane	0.96	0.81
Water	0.93	0.97

Overall Accuracy: 0.86.

- Discussão:
  - Menores acurácias:
    - ✓ Cultura perene **51% -> 68**;
    - ✓ Área Natural Não Vegetada **63% -> 38**.

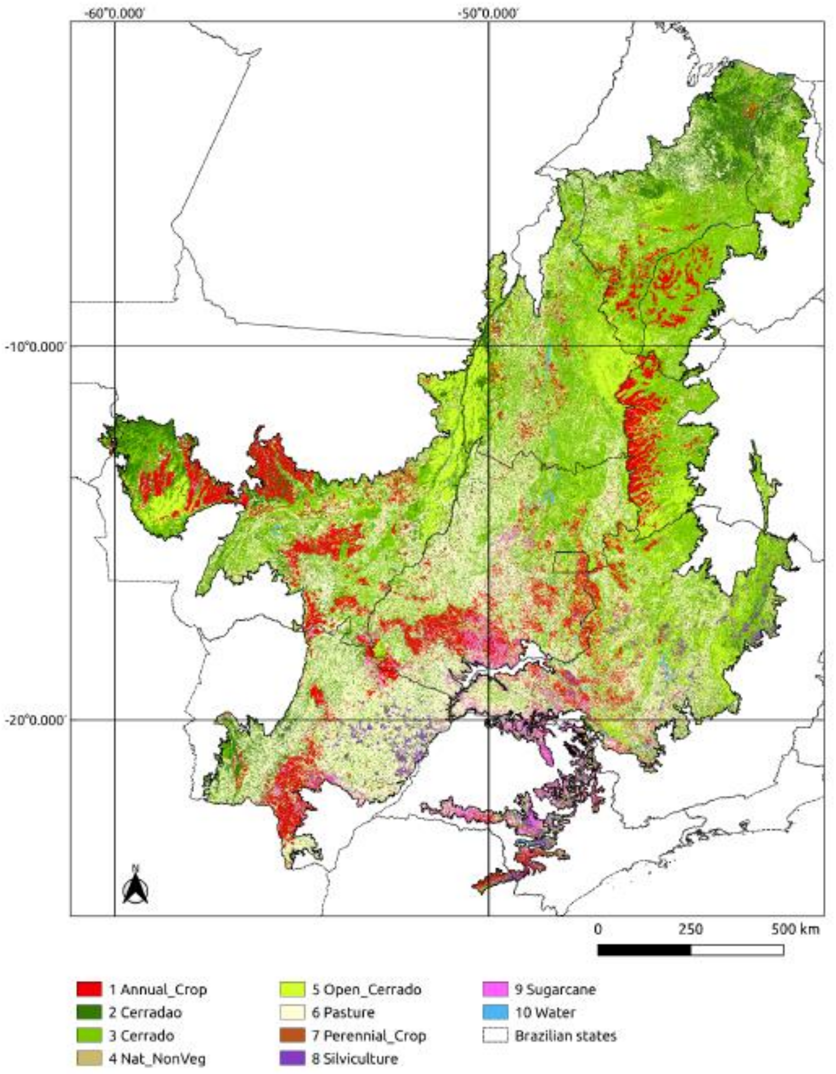

Figure 7. Cerrado land use and land cover map for 2018 (source: authors).

Conclusão

- Séries temporais são essenciais para o monitoramento de larga escala.;
- Apesar das limitações (ex: series temporais baseadas em pixel) é uma inovação;
- Estudo de caso bem sucedido com acurácia considerável em um fluxo de trabalho simples e eficas;
- Futuro: objetos espaciais e aprendizado ativo para melhor acurácia.

Obrigada!

Aluna: Raquel Zózimo Molinez

São José dos Campos – SP, Abril de 2025