Journal of Statistical Software

MMMMMM YYYY, Volume VV, Issue II. doi: 10.18637/jss.v000.i00

SITS: An R Package for Data Access, Visualisation, Filtering, Clustering, Event Detection and Classification of Satellite Image Time Series

Rolf Simoes, Gilberto Camara
INPE, Brazil

Victor Maus
IIASA

Alexandre Iwata
IPEA, Brazil

Abstract

Using time series derived from big Earth Observation data sets is one of the leading research trends in Land Use Science and Remote Sensing. One of the more promising uses of satellite time series is its application for classification of land use and land cover, since our growing demand for natural resources has caused major environmental impacts. Given this motivation, this package provides a set of tools for data access, filtering, clustering and classification of satellite image time series.

Keywords: satellite image time series, big Earth Observation data.

1. Introduction

Earth observation satellites provide a continuous and consistent set of information about the Earth's land and oceans. Most space agencies have adopted an open data policy, making unprecedented amounts of satellite data available for research and operational use. This data deluge has brought about a major challenge for Geoinformatics research: *How to design and build technologies that allow the Earth observation community to analyse big data sets?*

Since remote sensing satellites revisit the same place repeatedly, we can calibrate their images so measures of the same place in different times are comparable. These observation can be organised, so that each measure from sensor is mapped into a three dimensional array in space-time. From a data analysis perspective, researchers then have access to satellite image time series (SITS). Using time series derived from big Earth Observation data sets is one of the leading research trends in Land Use Science and Remote Sensing.

A time-series of measurements of the same location in the surface of the Earth can be considered as a historical record. When the images arise for a dense record of frequent revisits, the temporal resolution of the big data set is able to capture the most important land use changes. Such dense time series allow researchers to which changes have taken place in each location.

The benefits of remote sensing time series analysis arise when the temporal resolution of the big data set is sufficient to capture the most important changes. In this case, the temporal autocorrelation of the data can be stronger than the spatial autocorrelation. In other words, given data with adequate repeatability, a pixel will be more related to its temporal neighbours rather its spatial ones. In this case, *time-first, space-later* methods will give better results than the *space-first, time-later* approach.

Time series of remote sensing data show that land cover changes do not always occur in a progressive and gradual way, but they may also show periods of rapid and abrupt change followed either by a quick recovery (Lambin and Linderman 2006). Analyses of multiyear time series of land surface attributes, their fine-scale spatial pattern, and their seasonal evolution leads to a broader view of land-cover change. Satellite image time series have already been applied to applications such as mapping for detecting forest disturbance (Kennedy, Yang, and Cohen 2010), ecology dynamics (Pasquarella, Holden, Kaufman, and Woodcock 2016), agricultural intensification (Galford, Mustard, Melillo, Gendrin, Cerri, and Cerri 2008) and its impacts on deforestation (Arvor, Meirelles, Dubreuil, Begue, and Shimabukuro 2012).

2. Using the WTSS service

Get information about the WTSS (web time series service)

To get a remote sensing time series, one first organises a large set of EO data as a 3D array. From each pixel location in the array, one can extract a time series of one or more variables for a temporal interval. The WTSS service is independent of the actual data architecture used for 3D array store. It can work with solutions such as flat files, MapReduce distributed datasets, array databases or object-relational databases. We have implemented the service using both a set of flat files and the SciDB array database management system (Stonebraker, Brown, Zhang, and Becla 2013), with the same external interface.

```
R> URL <- "http://www.dpi.inpe.br/tws/wtss"
R> wtss_inpe <- sits_infoWTSS(URL)
```

```
--------------------------------------------------------------
The WTSS server URL is http://www.dpi.inpe.br/tws/wtss
Available coverages:
itobi
merge
mixl8mod
mixl8mod_f
mod13q1_512
---------------------------------------------------------------
```

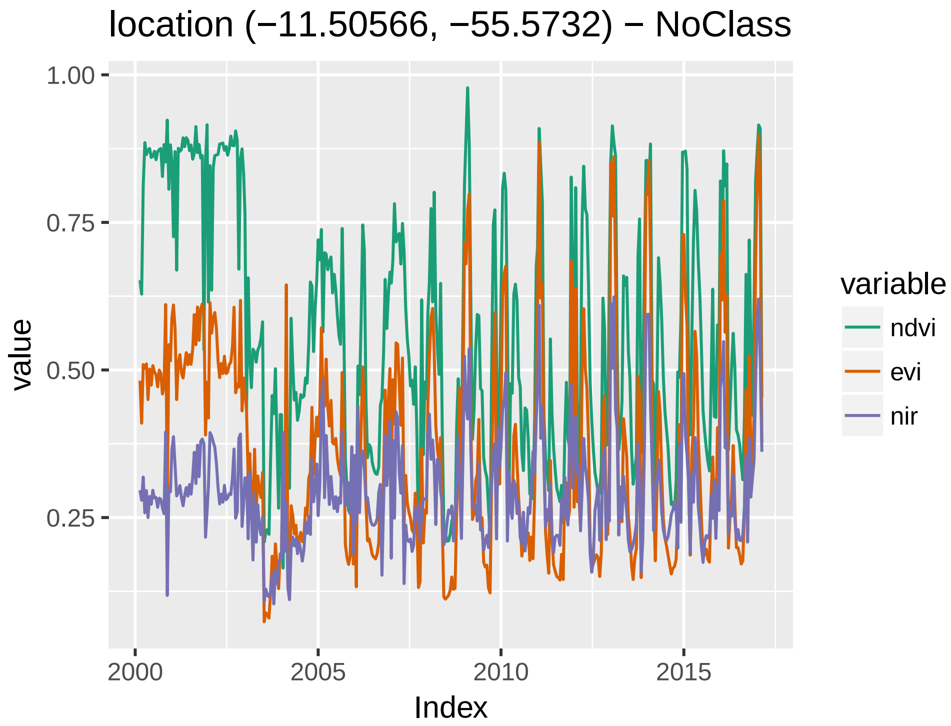
```
R> # get information about a specific coverage
R> sits_coverageWTSS(URL,"mod13q1_512")
```

```
------------------------------------------------------------------------------------
Coverage: mod13q1_512
```

```
Description: Vegetation Indices 16-Day L3 Global 250m
Source: https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13q1
Bands:
  name                            description
1 ndvi                     250m 16 days NDVI
2  evi                      250m 16 days EVI
3  red  250m 16 days red reflectance (Band 1)
4  nir  250m 16 days NIR reflectance (Band 2)
5 blue 250m 16 days blue reflectance (Band 3)
6  mir  250m 16 days MIR reflectance (Band 7)

Spatial extent: (-180, -90) - (180, 90)
Spatial resolution: (0.00208334, 0.00208334)
Projection CRS: +proj=longlat +ellps=WGS84 +datum=WGS84 +no_defs
Time range: 2000-02-18 to 2017-02-18
Temporal resolution: 16 days
----------------------------------------------------------------------------------
```

```
R> # choose a coverage
R> coverage <- "mod13q1_512"
R> # recover all bands
R> bands <- c("ndvi", "evi", "nir")
R> # a point in the transition forest pasture in Northern MT
R> long <- -55.57320
R> lat <- -11.50566
R> # obtain a time series from the WTSS server for this point
R> series.tb <- sits_getdata(longitude = long, latitude = lat, URL = URL, coverage = "mod1
R> # plot the series
R> sits_plot (series.tb)
```



References

Arvor D, Meirelles M, Dubreuil V, Begue A, Shimabukuro YE (2012). "Analyzing the agricultural transition in Mato Grosso, Brazil, using satellite-derived indices." *Applied Geography*, **32**(2), 702–713.

Galford GL, Mustard JF, Melillo J, Gendrin A, Cerri CC, Cerri CE (2008). "Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil." *Remote sensing of environment*, **112**(2), 576–587.

Kennedy RE, Yang Z, Cohen WB (2010). "Detecting trends in forest disturbance and recovery using yearly Landsat time series." *Remote Sensing of Environment*, **114**(12), 2897–2910.

Lambin E, Linderman M (2006). "Time series of remote sensing data for land change science." *IEEE Transactions on Geoscience and Remote Sensing*, **44**(7), 1926–1928.

Pasquarella VJ, Holden CE, Kaufman L, Woodcock CE (2016). "From imagery to ecology: leveraging time series of all available Landsat observations to map and monitor ecosystem state and dynamics." *Remote Sensing in Ecology and Conservation*, **2**(3), 152–170. ISSN 2056-3485.

Stonebraker M, Brown P, Zhang D, Becla J (2013). "SciDB: A database management system for applications with complex analytics." *Computing in Science & Engineering*, **15**(3), 54–62.

Affiliation:

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
MMMMMM YYYY, Volume VV, Issue II *Submitted:* yyyy-mm-dd
doi:10.18637/jss.v000.i00 *Accepted:* yyyy-mm-dd