STILF - A spatiotemporal interval logic formalism for reasoning about events in remote sensing data

Adeline Marinho Maciel[1]
Lubia Vinhas[1]
Gilberto Câmara[1]
Victor Wegner Maus[2]
Luiz Fernando Ferreira Gomes de Assis[1]

[1] National Institute for Space Research – INPE
Caixa Postal 515 – 12227-010 – São José dos Campos - SP, Brasil
[2] International Institute for Applied Systems Analysis – IIASA
{adeline.maciel, lubia.vinhas, gilberto.camara}@inpe.br, {luizffga, vwmaus1}@gmail.com

Abstract. Although several studies perform time series analysis using remote sensing data provided by Earth observation satellites, few have been explored concerning the reasoning about land use change using these data. Besides, exists the challenge of make the best use of big Earth observation data sets to represent change. In this context, this work presents a new formalism - STILF (Spatiotemporal Interval Logic Formalism), and shows how to use it for reasoning about land use change using big Earth observation data. Extending the ideas from Allen's interval temporal logic, we introduce predicates $holds(o, p, t)$ and $occur(o, p, T_e)$ to build a general framework to reason about events. Events can be defined as complete entities on their respective time intervals and their lifetime is limited while objects persist in time, with a defined begin and end. Since events are intrinsically related to the objects they modify, a geospatial event formalism should specify not only what happens, but also which objects are affected by such changes. The formalism proposed and predicates extended from Allen's ideas can model and capture changes using big Earth observation data, and also allows reasoning about land use trajectories in regional or global areas. Examples for tropical forest area application is presented to better understand our proposal using STILF. For the future, the proposed formalism will be include other temporal analysis tools to thinking about events related the land use and cover change.

Keywords: land use and land cover, spatiotemporal representation, Allen's interval, events, logic formalis, remote sensing

1. Introduction

One of the recent trends in applications of remote sensing data is the use of big data sets for obtaining information about land use and land cover. Using long-term time series, scientists can obtain new information to understand how mankind is using natural resources. Satellite image time series data provides a new perspective in remote sensing data analysis (CAMARA et al., 2016a).

An example of big Earth Observation data analysis is the work by Hansen et al. (2013). Using more than 650,000 LANDSAT images and processing more than 140 billion pixels, the authors compared data from 2000 to 2010 to produce maps of global forest loss. The results for 2000 and 2010 were compared to account for forest loss during the 2000-2010 decade. The method classifies each 2D image one by one.

By contrast, methods such as the time-weighted dynamic time warping (TWDTW) (MAUS et al., 2016) and TIMESTAT (JöNSSON; EKLUNDH, 2004) work on remote sensing time series to extract long-term information for each pixel. These algorithms work on individual time series and combine the results for selected periods to generate classified maps.

The benefits of remote sensing time series analysis arise when the temporal resolution of the big data set is able to capture the most important changes. Here, the temporal autocorrelation of the data can be stronger than the spatial autocorrelation. Given data with adequate repeatability, a pixel will be more related to its temporal neighbours than to its spatial ones. In this case, *time-first, space-later* methods lead to better results than the *space-first, time-later* approach (CAMARA et al., 2016a).

Given the possible new results that can be obtained with big remote sensing data, the scientific challenge is how to best represent and detect change. Issues about representation, reasoning, modelling of changes have been researched in GIScience (PEUQUET; DUAN, 1995; GALTON, 2004). In general, these studies show the usefulness of using the concept of "events" to represent changes in spatiotemporal data. The objective of this paper is to apply the concept of "events" for representing change in big remote sensing data sets, following the ideas from Galton (2015). Additionally, we extend the interval temporal logic proposed by Allen (1984) to build a logic formalism which allows reasoning about events of change in land use and land cover data. This paper extends and improves on earlier work by our research group (CAMARA et al., 2016b).

2. A Spatiotemporal Interval Logic Formalism - STILF

To describe land use and land cover changes, we consider an approach based in time intervals. We extend the interval temporal logic from Allen (1984) to build a logic formalism for reasoning about events. Allen (1983) defines a set of thirteen relationships between two time intervals: *before*, *meets*, *during*, *starts*, *finishes*, *overlap*, with inverse relationship, and *equal*.

To extend the predicates from Allen (1984) to spatiotemporal data, we aggregate the notion of geo-objects, which are related to space. This way, the formalism is composed for a set of elements: (1) discrete geo-objects ($O = o_1, o_2, ..., o_n$); (2) properties of geo-objectos ($P = p_1, p_2, ..., p_n$); and (3) time intervals ($T = t_1, t_2, ..., t_n$). We also use the predicates: (1) $holds(o, p, t) \rightarrow bool$, which asserts that a properties p from geo-object o holds during a interval t; and (2) $occur(o, p, t_e) \rightarrow bool$, given a interval $T_e \subset T$, the properties p from geo-object o will be true during all sub-interval T_e.

The start point of the spatiotemporal interval logic formalism (STILF) is a set of time series data classified from remote sensing images. This images were previously classified by means of data mining algorithm, such as TWDTW (MAUS et al., 2016), and stored in a array database. This is a important stage for the application of the formalism, once allows that Earth observation data were stored in a database which support a large amount of remote sensing data, and subsequently, can will be used for different applications. Next, this data set will be processed, for extraction of the set of elements. Each element is composed of a discrete geo-object, its properties of geo-object and time intervals for which these properties hold.

The set of elements will be used as input data for our formalism. Combining the $holds(o, p, t)$ and $occur(o, p, t_e)$ predicates with Allen's relations, we can ask questions about trajectories of land use and land cover change. The answers will be data sets with the events that have occurred during the whole interval for which we have data. In the last stage, individual events will be combined in terms of their characteristics into recurring events or transition events. Figure 1 shows a overview of our proposed formalism.

3. Application: Examples of Reasoning About Events from Classified Land Use and Land Cover Time Series

In this section we show three examples of application using remote sensing data. The formalism presented was developed in a R programming language and applied on sample

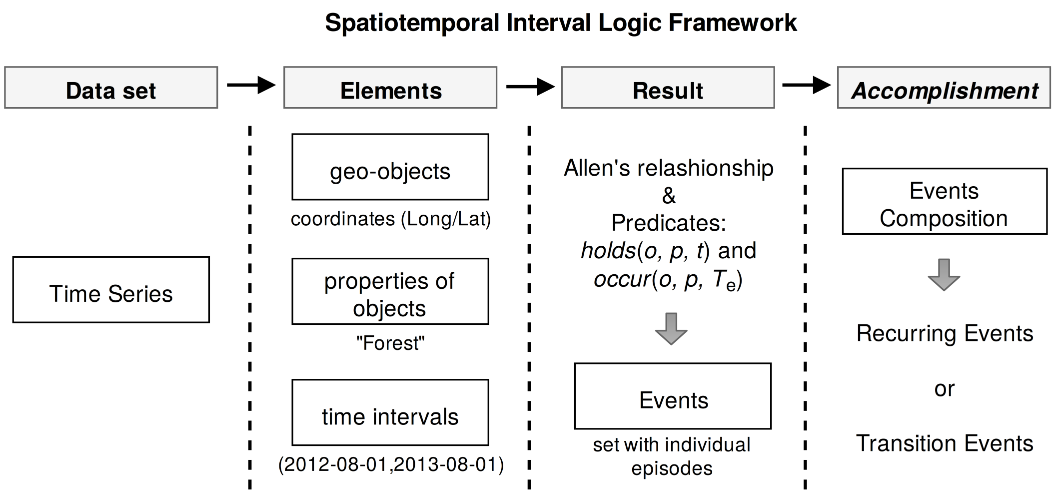


Figure 1: Spatiotemporal interval logic formalism (STILF) design

regions. The input data is composed for a set of time series classified to the municipality of Porto dos Gaúchos, located in northwest Mato Grosso (MT) state, Brazil. With territorial area 6,862.118 km^2, geographical coordinates, latitude 11° 31' 31" South and longitude 57° 24' 50" West and population of 5,400 inhabitant in 2010, according to IBGE statistics (IBGE, 2016). Porto dos Gaúchos is an area into Amazon biome.

Figure 2: Municipality of Porto dos Gaúchos with highlight for three sample regions selected to application of the formalism.

We extracted information from each sample region to discover what events had happened. These events allow us to establish the trajectories of land use and land cover. For example, the results may indicate the increase of deforestation in the municipality after earlier expansion of areas. We can also detect the conversions from pasture to soybean and from soybean to double cropping (soybean-corn or soybean-cotton).

After we classify the time series, we apply a post-processing rule to distinguish natural, intact forests from areas that had been deforested and then were allowed to regrow. This is required to be able to differentiate primary forest, without degradation, from secondary vegetation, forest areas that happened after other land use or land cover classes. The new classes generated after this stage were called "Secondary vegetation".

In the first sample region, with an area of 50.23 km^2, located in Northeast of the

ISBN: 978-85-17-00088-1
Anais do XVIII Simpósio Brasileiro de Sensoriamento Remoto - SBSR

municipality, we explored the ability of the formalism to detect events composition. Our query searched for events preceding and following the year of 2008, associated to the "Soybean-Millet" crop areas Question 1. The result were three graphics with information for analysing land use trajectories: (1) a custom map that highlights events that show the transition from "Pasture" to"Soybean-Millet", Figure 3; (2) a bar graph which counts the total area (km^2) for each event by year; and (3) a graph which represents the temporal sequence of the events for each pixel in the time. This type of graph show what pasture areas were transformed into crop areas (Figure 4(b)).

Question 1: Example of application of the spatiotemporal interval logic formalism for mapping of changes in land use and cover for the first sample region

a) *Which "Pasture" areas before 2008 have been turned in "Soybean-Millet" cropping areas?*

$$\forall o \in O, occur(o, \text{“}Pasture\text{”}, t_1) \wedge occur(o, \text{“}Soybean - Millet\text{”}, t_2) \wedge$$
$$occur(o, \text{“}Soybean - Millet\text{”}, t_3) \wedge$$
$$preceding(t_2, t_1) \Leftrightarrow (metBy(t_1, t_2) \vee after(t_1, t_2)) \wedge$$
$$next(t_2, t_3) \Leftrightarrow (meets(t_1, t_2) \vee before(t_1, t_2))$$
$$\textit{where } t_1 = \{2000, ..., 2007\}, t_2 = 2008, t_3 = \{2009, ..., 2015\}$$

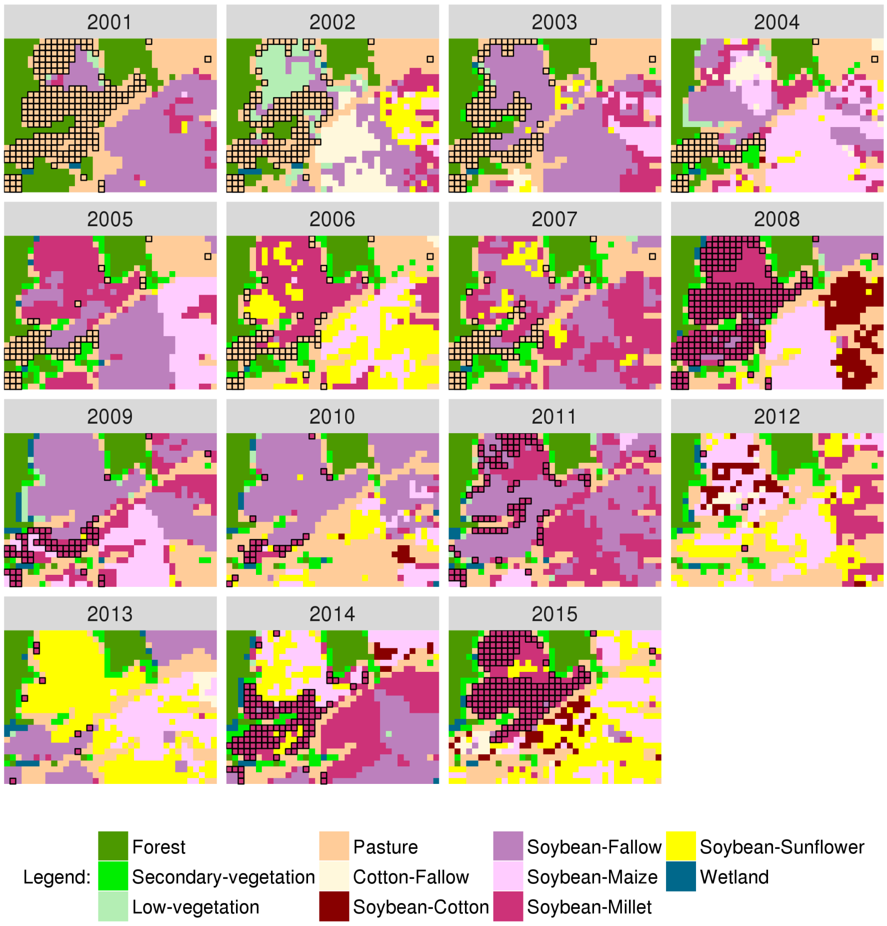
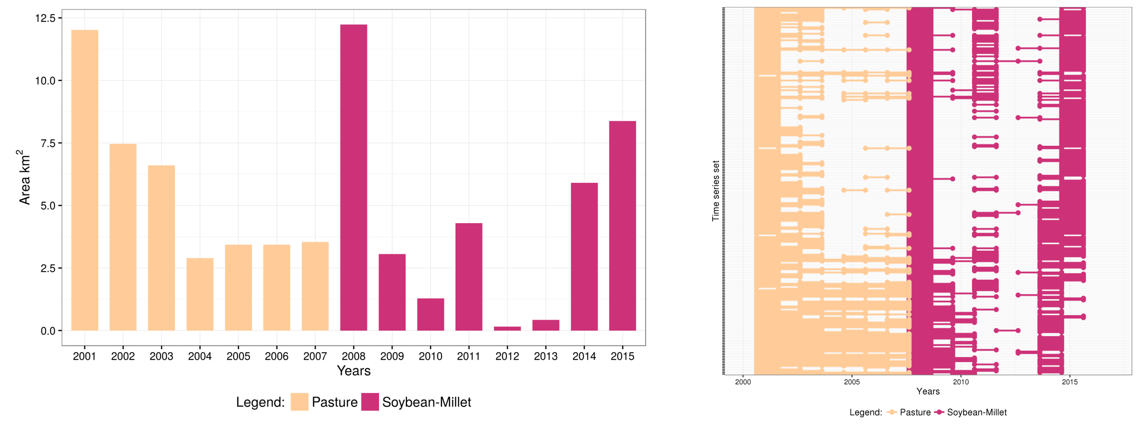


Figure 3: Sample region 1, events highlighted.

ISBN: 978-85-17-00088-1
Anais do XVIII Simpósio Brasileiro de Sensoriamento Remoto - SBSR



(a) Area in (km^2) with "Pasture" events turned into "Soybean-Millet" from 2008.

(b) Temporal sequence of the events.

Figure 4: Graphics to analyses of events composition - sample region 1

In the second sample region, located to the south of the municipality and area of 59.568 km^2, we investigated which "Forest" areas that have been turned into "Pasture" or " Low vegetation (a second type of pasture)" after 2001. The formal representation of the question is shown in Question 2. Three output plots were generated with information about events: a map that highlights events that happened yearly, Figure 5; a bar graph with total area for each event by year (Figure 6(a)), and a temporal representation for each pixel over time, which shows the transitions from forest to pasture (Figure 6(b)).

Question 2: Example of application of the STILF for mapping of changes in land use and cover for the second sample region

b) *Which "Forest" areas have been turned into "Pasture" or "Low-vegetation" after the year of 2001?*

$$\forall o \in O, occur(o, \text{“}Forest'', t_1) \wedge$$
$$(occur(o, \text{“}Pasture\text{”}, t_2) \vee occur(o, \text{“}Low - vegetation'', t_2)) \wedge$$
$$next(t_1, t_2) \textit{ where } t_1 = 2001, t_2 = \{2002, ..., 2015\}$$

In a third sample region, located northwest of Porto dos Gaúchos municipality and area of 101.963 km^2, we wanted to know which "Forest" areas have not undergone degradation over years (Question 3) In a similar way to the PRODES system, forest areas that from regrowth after fire or deforestation are called "Secondary vegetation" by our post-processing rule and are not computed. Figure 7 shows a map that highlights the events. Figure 8 displays the amount of forest grouped by year. We can see the expansion of deforestation until 2006, when there was a significant reduction.

Question 3: Example of application of the STILF for mapping of changes in land use and cover for the third sample region

c) *Which "Forest" areas have not undergone degradation during interval of 15 years?*

$$\forall o \in O, occur(o, \text{“}Forest\text{”}, t_1) \textit{ where } t_1 = \{2001, ..., 2015\}$$

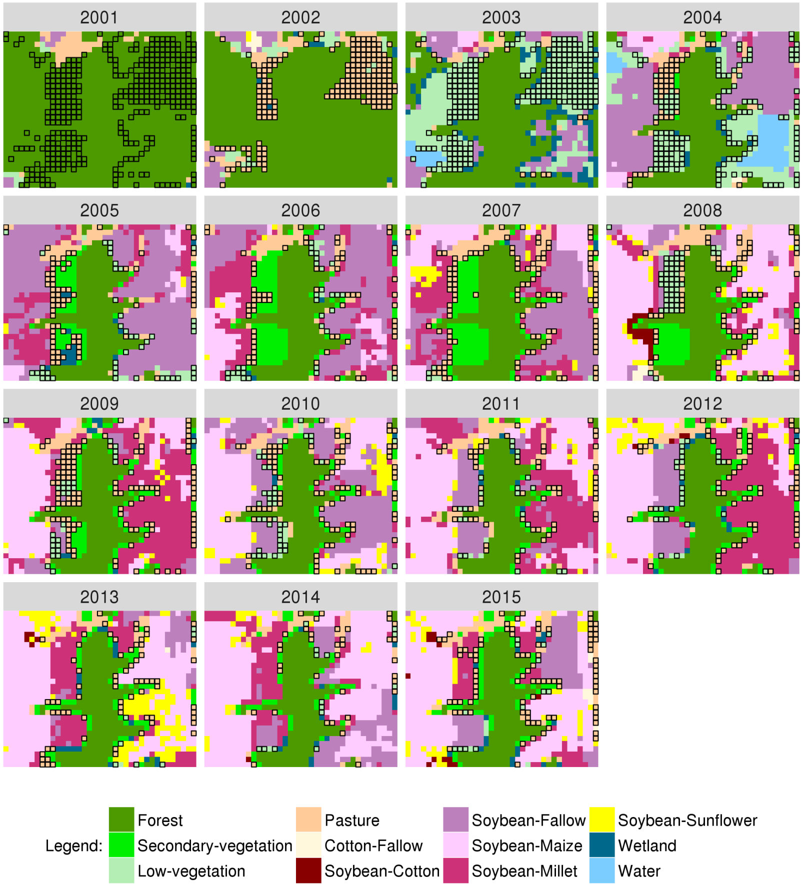


Figure 5: Sample region 2, events highlighted

(a) Area in (km^2) with "Forest" events turned into "Pasture" and/or "Low-vegetation" from 2001.

(b) Temporal sequence of the events to second sample region.

Figure 6: Graphics to analyses of events composition - sample region 2

This spatiotemporal interval logic formalism makes it easy to build questions in a logic representation in order to reason about changes in land use and land cover. We can show the trajectories of change in different perspectives. This makes it easier to understand changes in an environment. The formalism is robust. It allow different logical queries combining Allen's relations and also predicates of geo-objects.

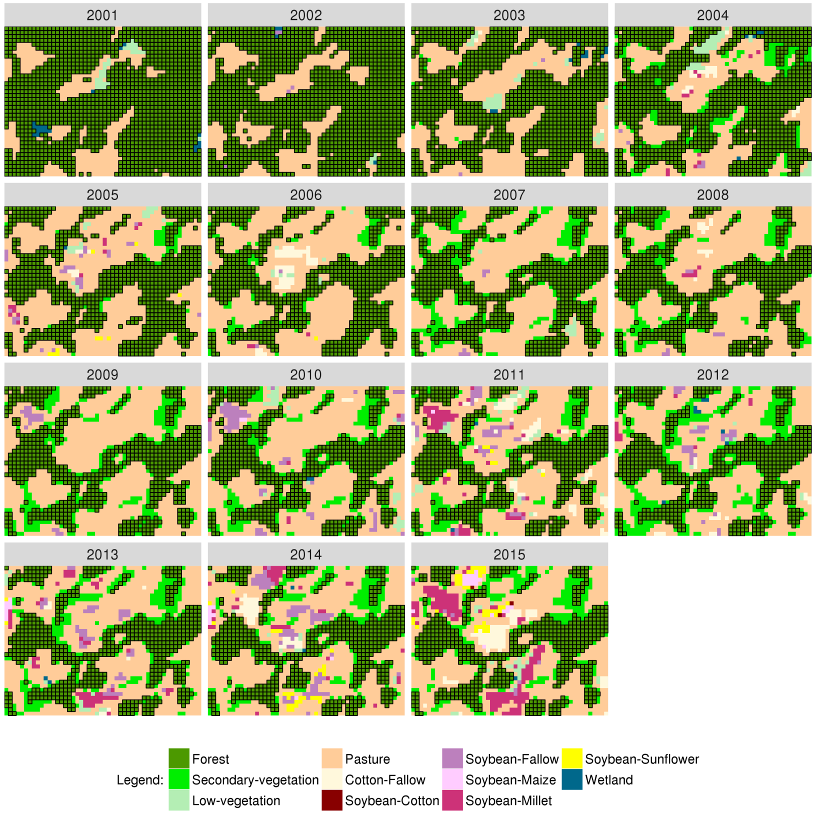


Figure 7: Sample region 3, with events highlighted

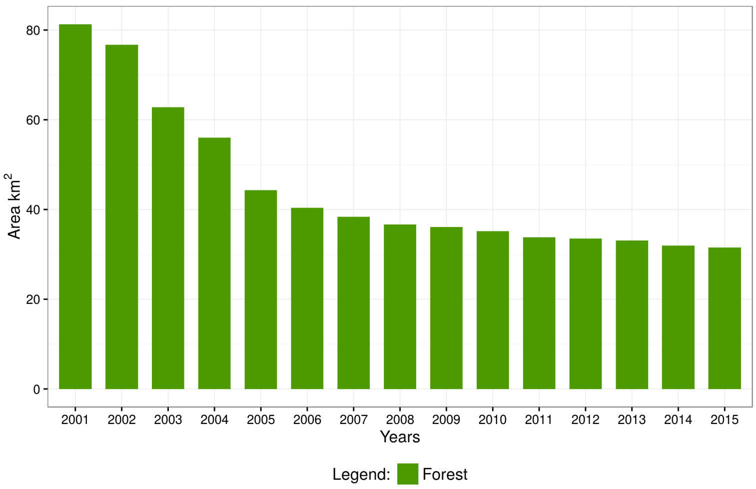


Figure 8: Area in (km^2) with "Forest" events which have not undergone degradation during the period of 15 years.

4. Final Considerations

A spatiotemporal interval logic formalism to reasoning about changes in land use and land cover was presented in this paper. This formalism is an extension from predicates defined by Allen (1984). We introduce geo-objects as new elements for analyses involving spatial data. We show three examples of application where the formalism was implemented in a programming language, take advantaging of the resources for data visualisation and results.

Acknowledgements

The authors thank CAPES and FAPESP e-science program (grants 2014-08398-6 and 2016-03397-7) by financial support.

References

ALLEN, J. F. Maintaining knowledge about temporal intervals. *Communications of the ACM*, ACM, New York, NY, USA, v. 26, n. 11, p. 832–843, 1983.

ALLEN, J. F. Towards a general theory of action and time. *Artificial Intelligence*, Elsevier Science Publishers Ltd., Essex, UK, v. 23, n. 2, p. 123–154, 1984.

CAMARA, G. et al. Big earth observation data analytics: Matching requirements to system architectures. In: *Proceedings of the 5th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data*. New York, NY, USA: ACM, 2016. (BigSpatial '16), p. 1–6.

CAMARA, G. et al. Using dynamic geospatial ontologies to support information extraction from big earth observation data sets. In: *Ninth International Conference on Geographic Information Science (GIScience 2016)*. Montreal, CA: AAG, 2016.

GALTON, A. Fields and Objects in Space, Time, and Space-time. *Spatial Cognition & Computation*, v. 4, n. 1, p. 39–68, 2004.

GALTON, A. Outline of a formal theory of processes and events, and why giscience needs one. In: *Conference on Spatial Information Theory - COSIT 2015*. Santa Fe, NM, USA: Springer International Publishing, 2015. p. 3–22.

HANSEN, M. et al. High-resolution global maps of 21st-century forest cover change. *Science (New York, N.Y.)*, v. 342, n. 2013, p. 850–3, 2013.

IBGE. *The Brazilian Institute of Geography and Statistics*. 2016. Available on `http://www.ibge.gov.br`.

JöNSSON, P.; EKLUNDH, L. Timesat–a program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, v. 30, n. 8, p. 833 – 845, 2004.

MAUS, V. et al. A time-weighted dynamic time warping method for land-use and land-cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, IEEE, PP, n. 99, p. 1–11, 2016.

PEUQUET, D. J.; DUAN, N. An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data. *International journal of geographical information systems*, Taylor & Francis, v. 9, n. 1, p. 7–24, 1995.