International Journal of Applied Earth Observation and Geoinformation 47 (2016) 112–124

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

# Reconstructing land use history from Landsat time-series Case study of a swidden agriculture system in Brazil

Loïc P. Dutrieux[a,*], Catarina C. Jakovac[b], Siti H. Latifah[a], Lammert Kooistra[a]

[a] Laboratory of Geo-Information Science and Remote Sensing, Wageningen University, The Netherlands
[b] Forest Ecology and Management Group, Wageningen University, The Netherlands

ARTICLE INFO

*Article history:*
Received 23 July 2015
Received in revised form 2 November 2015
Accepted 30 November 2015
Available online 28 December 2015

*Keywords:*
Landsat
Land trajectories
Swidden agriculture
Segmentation
Time-series
Land use history
Slash and burn agriculture

ABSTRACT

We developed a method to reconstruct land use history from Landsat images time-series. The method uses a breakpoint detection framework derived from the econometrics field and applicable to time-series regression models. The Breaks For Additive Season and Trend (BFAST) framework is used for defining the time-series regression models which may contain trend and phenology, hence appropriately modelling vegetation intra and inter-annual dynamics. All available Landsat data are used for a selected study area, and the time-series are partitioned into segments delimited by breakpoints. Segments can be associated to land use regimes, while the breakpoints then correspond to shifts in land use regimes. In order to further characterize these shifts, we classified the unlabelled breakpoints returned by the algorithm into their corresponding processes. We used a Random Forest classifier, trained from a set of visually interpreted time-series profiles to infer the processes and assign labels to the breakpoints. The whole approach was applied to quantifying the number of cultivation cycles in a swidden agriculture system in Brazil (state of Amazonas). Number and frequency of cultivation cycles is of particular ecological relevance in these systems since they largely affect the capacity of the forest to regenerate after land abandonment. We applied the method to a Landsat time-series of Normalized Difference Moisture Index (NDMI) spanning the 1984–2015 period and derived from it the number of cultivation cycles during that period at the individual field scale level. Agricultural fields boundaries used to apply the method were derived using a multi-temporal segmentation approach. We validated the number of cultivation cycles predicted by the method against in-situ information collected from farmers interviews, resulting in a Normalized Residual Mean Squared Error (NRMSE) of 0.25. Overall the method performed well, producing maps with coherent spatial patterns. We identified various sources of error in the approach, including low data availability in the 90s and sub-object mixture of land uses. We conclude that the method holds great promise for land use history mapping in the tropics and beyond.



## 1. Introduction

Land use and land use dynamics affect elements of the biosphere with local to global impacts (Foley et al., 2005). Forests are of particular importance in this system for the role they play in maintaining biodiversity levels and delivering ecosystem services such as climate regulation and water supplies (Foley et al., 2005; Myers et al., 2000). While old-growth forests are often considered first when accounting for these services, secondary forests, with an estimated 2010 area of 165,230 km$^2$ for the Brazilian Amazon alone (TerraClass, 2011), cannot be ignored since they too play an increasingly important role in the provision of ecosystem services (Bongers et al., 2015). However, multiple factors need to be taken into account when dealing with secondary forests. Secondary forests can succeed to a variety of land use histories such as logging, agriculture, or cattle ranching and these different histories will in turn impact the characteristics of the present forests. It has been shown that the vegetation structure, species composition and resilience of secondary forests are strongly affected by previous land uses (Jakovac et al., 2015; Lawrence et al., 2010; Mesquita et al., 2001). The longer the areas is kept under land use and, in the case of shifting cultivation, the higher the frequency of use, the lower the recovery rate and species diversity of the succeeding secondary forests (Zarin et al., 2005; Jakovac et al., 2015). Intensive land use can also result in a shift in species composition (Mesquita et al., 2001; Jakovac et al., in preparation-a) and can ultimately hinder the secondary succession (Longworth et al., 2014). Examples of previous land uses include swidden agriculture—also known as

* Corresponding author.
*E-mail address:* loic.dutrieux@wur.nl (L.P. Dutrieux).

http://dx.doi.org/10.1016/j.jag.2015.11.018
0303-2434/

slash and burn agriculture, which is a type of shifting cultivation widespread in the tropics and from which a large part of today's secondary forests originate (Van Vliet et al., 2012; Mertz, 2009). Swidden agriculture has a very particular rotation cycle which consists in cutting and burning the forest, cultivating the land for a period of one to three years and leaving the land fallow until the next rotation five to sometimes more than 20 years later (Coomes et al., 2000; Jakovac et al., 2015). For these systems specifically, research carried out at field level has shown that the frequency of land usage as well as the total number of cultivation cycles are important determinants of the current structure and function of the resulting secondary forest (Lawrence et al., 2010; Jakovac et al., 2015).

Considering the strong connection that secondary forests have with its previous land uses, land use history is an important aspect to take into account when measuring and modelling the current state of secondary forest ecosystems. However, despite its importance such information is not always straightforward to obtain. Interviewing local stakeholders can provide accurate information, but such effort can only be carried for a few punctual locations, while information on the remaining part of the landscape would still be lacking. There is therefore a need to develop standardized methods capable of deriving quantitative, spatial informations about past land uses in secondary forest systems (Van Vliet et al., 2012). Remote sensing techniques offer great promises to contribute to this effort of studying the dynamics and evolution of these ecosystems in a spatial context. Satellites have been acquiring images of the earth for more than 40 years. This represents large amounts of data which have been collected in an objective, systematic and spatially continuous way. However, exploiting the remote sensing data for ecological purposes contains challenges as well. While variables measured in field inventories have a direct ecological meaning, optical remote sensing measures light reflected by the elements of the earth surface. Reflected light is a biophysical variable which does not have any direct ecological meaning and therefore does not provide useful information to the ecologist or the decision maker. There is therefore a challenge that consists in translating raw satellite measurements into useful and meaningful variables. One way of achieving this is by taking advantage of the temporal dimension of the remote sensing measurements. The earth has been continuously monitored by the various sensors on board the Landsat satellites since roughly the 80s (Goward et al., 2006). By assembling these repeated measurement, land dynamics can be extracted and used to provide useful information about the history and dynamics of the land (Kennedy et al., 2014).

A variety of methods has been developed to investigate change using remote sensing time-series (Lu et al., 2014). The wide range of techniques developed responds to various monitoring needs both in terms of spatial extent and dynamics observed. The land dynamics can be either gradual or abrupt and concern changes about the intra-annual or the inter-annual dynamics (Kennedy et al., 2014). Using simple linear trends, Dutrieux et al. (2012) and Fensholt and Proud (2012) have investigated gradual change regionally to globally using Moderate Resolution Imaging Spectroradiometer (MODIS) and Advanced Very High Resolution Radiometer (AVHRR) coarse resolution time-series. More sophisticated approaches, such as de Jong et al. (2013) and Forkel et al. (2013) used the Breaks For Additive Season and Trend (BFAST) framework (Verbesselt et al., 2010b,a, 2014) to investigate shifts in vegetation trends globally. Studies at higher spatial resolution, nearly always based on Landsat data, include methods for near real time deforestation monitoring (Zhu et al., 2012; Brooks et al., 2014; DeVries et al., 2015b; Dutrieux et al., 2015; Reiche et al., 2015), as well as more general tools for vegetation dynamics and land trajectories monitoring (Kennedy et al., 2010; Huang et al., 2010; Zhu and Woodcock, 2014; DeVries et al., 2015a).

Here we are interested in applying a method to the case of swidden agriculture with the objective to map land use intensity defined as the number of times the land has been cultivated. Swidden agriculture dynamics are fast and complex, which adds additional challenges to the time-series approach we need to develop. This complexity originates firstly from the location of these systems. Because they are in the tropics where vegetation regrowth happens immediately after disturbances, the optical signal related to the swidden agriculture events fades rapidly making the temporal window to detect events very short (Asner et al., 2004a,b). An additional challenge comes from the repeated cycles, all containing gradual changes, such as the periods of forest regrowth, and abrupt changes triggered by the burning events. Finally the high cloud coverage usually present above tropical forest regions further restricts data availability making the observation of dynamics even more challenging (Asner, 2001). These constraints call for a hyper-temporal approach capable of capturing both gradual and abrupt changes.

Here we propose to use a statistically based breakpoint detection method derived from the econometrics literature to retrieve land use history. The method is applicable to Landsat time-series hence producing information on past land uses at medium spatial resolution. We applied the method to the case of swidden agriculture, trying to quantify the land use intensity defined as the number of times an area has been cultivated for two municipalities of the Brazilian Amazon where swidden agriculture is the predominant cultivation practice. Performances of the method are assessed and discussed by comparing method's output with a ground truth dataset about agricultural management collected via farmer's interviews.

## 2. Material and methods

### 2.1. Time-series segmentation

#### 2.1.1. Detecting change in time-series

As described in the introduction of this paper, detecting change may take several forms, depending on whether the change observed is gradual, abrupt, occurs on intra-annual dynamics, or between years (Kennedy et al., 2014). Here we propose an approach to perform full segmentation of the time-series. Abrupt changes, referred to as breakpoints are detected from changes in the coefficients of a time-series regression model (Bai, 1994). Periods between breakpoints, which we call temporal segments can be characterized in different ways, based on their duration, trend, mean values or intra-annual characteristics.

#### 2.1.2. Breakpoint detection—theoretical background

In order to detect abrupt changes in the remote sensing time-series, we use a breakpoint detection method derived from the econometrics literature. The method developed to detect breakpoints in time-series regression models was initially introduced by Bai (1994), and later extended to the detection of multiple breakpoints (Bai, 1997a,b; Bai and Perron, 1998). Examples of applications are given by Zeileis et al. (2003), and Bai and Perron (2003). Given a regression model (Eq. (1)), the method tests the hypothesis that the regression coefficients remain constant over time against the alternative hypothesis that at least one of the coefficients changes. To test this hypothesis, a triangular Residual Sum of Squares (RSS) matrix, which gives the RSS for each possible temporal segment in the time-series, is computed. The optimal number of partitions is then obtained by minimizing the Bayesian Information Criterion (BIC), while the position of the breakpoints is determined by minimizing the RSS among all possible partitioning schemes given by the RSS matrix (Zeileis et al., 2003; Bai and