Remote Sensing of Environment 156 (2015) 490–499

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape

Hannes Müller [a,*], Philippe Rufin [a], Patrick Griffiths [a,c], Auberto José Barros Siqueira [b], Patrick Hostert [a,c]

[a] Geography Department, Humboldt Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany
[b] Department of Sanitary and Environmental Engineering, Universidade Federal de Mato Grosso, Av. Fernando Corrêa da Costa, B-78060-900 Cuiabá, Brazil
[c] Integrative Research Institute on Transformations of Human–Environment Systems, Humboldt Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany

ARTICLE INFO

Article history:
Received 29 January 2014
Received in revised form 25 September 2014
Accepted 16 October 2014
Available online 11 November 2014

Keywords:
Compositing
Spectral–Temporal metrics
Landsat time series
Land use and land cover classification
Brazilian Cerrado
Savannas

ABSTRACT

Better spatial information on the global distribution of croplands and pastures is urgently needed. Without reliable cropland–pasture separation it will be impossible to retrieve high-quality information on agricultural expansion or land use intensification, and on related ecosystem service provision. In this context, the savanna biome is critically important, but information on land use and land cover (LULC) is notoriously inaccurate in these areas. This is due to pronounced spatial–temporal dynamics of agricultural land use and spectral similarities between cropland, pasture, and natural savanna vegetation. In this study, we investigated the potential to reliably separate cropland, pasture, natural savanna vegetation, and other relevant land cover classes employing Landsat-derived spectral–temporal variability metrics for a savanna landscape in the Brazilian Cerrado. In order to better understand the surplus value and limitations of spectral–temporal variability metrics for classification purposes, we analyzed four datasets of different temporal depth, using 344 Landsat scenes across four footprints between 2009 and 2012. Our results showed a reliable separation between cropland, pasture, and natural savanna vegetation achieving an adjusted overall accuracy of 93%. A similar accuracy and spatial consistency of LULC classification could not be achieved based on spectral information alone, indicating the high additional value of temporal information for identifying LULC classes in the complex land use systems of savanna landscapes. There is great potential for transferring our approach to other savanna systems which still suffer from inaccurate LULC information.

1. Introduction

Food security is of global concern, and land-based production has to meet the demands of an estimated 9 billion people in 2050 (Godfray et al., 2010). To define an optimum trade-off or even create win–win situations between agricultural expansion and intensification on the one hand, and sustainable ecosystem service provision on the other hand, reliable information about the global distribution of croplands and pasture areas is required (Fritz et al., 2013; Garnett et al., 2013). In this context, the savanna biome is critically important for three reasons: firstly, it suffers from enormous land conversion pressure (MEA, 2005; Ramankutty et al., 2006). Secondly, sustainable land use is a key concern in savanna ecosystems because they are biodiversity hotspots hosting much of the world's last remaining mega-fauna and provide ecosystem services of global importance such as carbon storage and climate regulation (MEA, 2005). Thirdly, spatially explicit mapping of land use and land cover (LULC) is notoriously challenging and inaccurate in these areas, which hampers monitoring of ecosystem changes and therefore does not support the development of appropriate policies to steer land conversion in a sustainable manner (Fritz et al., 2011; Herold, Mayaux, Woodcock, Baccini, & Schmullius, 2008).

Savanna ecosystems cover approximately 20% of the Earth's surface, with important shares in Asia, Australia, Africa, and South America. Over the past 35 years, the Brazilian Cerrado (savanna woodlands) has experienced one of the largest expansions of agro-pastoral lands worldwide (Ramankutty, Foley, & Olejniczak, 2002; Ramankutty et al., 2006). The Brazilian Cerrado is one of the world's hotspots of biodiversity (Myers, Mittermeier, Mittermeier, Da Fonseca, & Kent, 2000), comprising 2 M km^2 of woodlands, savannas, grasslands, gallery and dry forests (Eiten, 1972; Ribeiro, Sano, & Silva, 1981). Despite its outstanding ecological value, only 2.2% of the biome is protected, and 40%–50% of the original vegetation had been cleared for agro-pastoral land uses by 2002 (Klink & Machado, 2005; Machado et al., 2004; Sano, Rosa, Brito, & Ferreira, 2010). These major transformations increased carbon emissions, trigger biodiversity loss and reduce ecosystem services (Batlle-Bayer, Batjes, & Bindraban, 2010; Silva, Fariñas, Felfili, & Klink, 2006). In addition, the Cerrado biome became subject to widespread land use intensification (Barretto, Berndes, Sparovek, & Wirsenius, 2013; Galford et al., 2008), such as the expansion of cash crops on

* Corresponding author at: Geographisches Institut, Humboldt-Universität zu Berlin, Unter den Linden 610099 Berlin, Germany. Tel.: +49 30 2093 6893; fax: +49 30 2093 6848.
E-mail address: Hannes.mueller@geo.hu-berlin.de (H. Müller).

http://dx.doi.org/10.1016/j.rse.2014.10.014
0034-4257/

previously extensively managed pasture areas and shifting from "single-cropping" to "double-cropping" systems. Statistical analyses suggest that this development leads to a displacement of pasture areas to regions of natural vegetation and can be regarded as an indirect driver of deforestation in the Amazon and Cerrado biome (Arima, Richards, Walker, & Caldas, 2011; Barona, Ramankutty, Hyman, & Coomes, 2010). However, consistent spatially and temporally explicit data on the distribution of cropland and pasture is still missing, which causes high uncertainties for estimating environmental and socio-economic impacts of the land changes in the Brazilian Cerrado. Across large areas, remote sensing plays a key role in detecting, characterizing, and monitoring land cover and land use reliably, consistently, and at appropriate spatial scales over time (Kuemmerle et al., 2013; Lambin & Geist, 2006).

So far, the majority of remote sensing studies in Brazil focused on the Amazon biome, including governmental programs like the annual deforestation monitoring (PRODES), the real time system for detection of deforestation (DETER), and the program for land use classification in deforested areas (TerraClass) (Assunção, Gandour, & Rocha, 2013; INPE, 2008). First attempts to build similar knowledge for the Cerrado region are conducted by the remote sensing department of the Federal University of Goiás, who created a Systematic Monitoring of Deforestation in the Cerrado Biome (SIAD) in the year 2006 (Ferreira, Ferreira, Huete, & Ferreira, 2007; LAPIG, 2014; Rocha, Ferreira, Ferreira, & Ferreira, 2011). However, no program has been established for monitoring general land use in the Cerrado biome, and remote sensing faces numerous challenges there.

Besides the strong seasonality of natural vegetation and the high diversity of crop types in space and time (Sano, Rosa, Brito, & Ferreira, 2007), spectral similarities between cropland, pasture, and natural savanna vegetation complicate differentiating LULC in the Brazilian Cerrado (Grecchi, Gwyn, Bénié, & Formaggio, 2013; Sano et al., 2010). This confusion results from spectral similarity between land cover types as well as a high spectral heterogeneity within each land cover type (Brannstrom et al., 2008; Ferreira, Yoshioka, Huete, & Sano, 2003; Hill et al., 2011). Due to these spectral ambiguities, most classification approaches predominantly rely on dense temporal information, such as time series data obtained from the moderate resolution imaging spectrometer (MODIS) for separating natural savanna vegetation, cropland, and pasture areas (Morton, DeFries, & Shimabukuro, 2005; Galford et al., 2008; Adami et al., 2011; Arvor, Milton, Meirelles, Dubreuil, & Durieux, 2011; Trabaquini, Bernardes, Mello, Formaggio, & Rosa, 2011). However, MODIS-based analyses cannot monitor LULC before 2000 (when the satellites became operational) and often do not capture fine-scale patterns in heterogeneous savanna ecosystems such as the Brazilian Cerrado or South-African savannas (Fritz et al., 2011; Huete et al., 2002; Munyati & Mboweni, 2013). Recent approaches therefore include Landsat imagery (30 m resolution) to overcome the limitation of the MODIS spatial scale (250 m resolution) and temporal extent in heterogeneous savanna regions (Grecchi et al., 2013; Schmidt, Udelhoven, Gill, & Röder, 2012).

With the opening of the Landsat archive, analysis strategies now shift from a scene-based perspective to more meaningful study area delineations, such as political/administrative units or natural entities, like complete watersheds across multiple Landsat footprints (Hansen & Loveland, 2012; Wulder et al., 2008). In this context, the Brazilian Landsat archive has recently been transferred to the archive of the United States Geological Survey (USGS) and converted to precision terrain corrected (L1T) Landsat format. The high geometric accuracy of this new dataset, coupled with state of the art atmospheric correction (e.g. LEDAPS by Masek et al., 2006) and cloud screening methods (e.g. FMask by Zhu & Woodcock, 2012), now enables the creation of consistent Landsat time series for South America.

So far, most Landsat-based time series have been created on an annual basis to monitor forest cover dynamics (Griffiths et al., 2012; Kennedy, Cohen, & Schroeder, 2007; Kennedy et al., 2012; Powell et al., 2010). Recent research highlights the potential of using intra-annual time series for grassland mapping (Schuster, Schmidt, Conrad, Kleinschmit, & Förster, 2015). We therefore hypothesize that identifying complex LULC classes within agricultural systems will profit from intra-annual information to capture phenological characteristics as well. However, due to Landsat's 16-day repeat cycle, cloud contaminations, and long term acquisition plans, regular spacing of intra-annual acquisitions is difficult, and direct quantification of phenological metrics is challenging (Kovalskyy & Roy, 2013). To overcome these limitations, interpolation and curve fitting methods have been successfully employed for regions with high data availability, notably in North America (Melaas, Friedl, & Zhu, 2013; Zhong, Gong, & Biging, 2014). However, it is questionable whether single dates of phenological characteristics in regions of low observation density should be extracted, given the degree of generalization involved in bridging large temporal data gaps. Less specific but more robust phenological indicators can be derived by simply calculating mean, range, and standard deviation of available observations for seasonal windows, single years or multiple years (Griffiths, Müller, Kuemmerle, & Hostert, 2013; Hansen et al., 2013). These spectral–temporal variability metrics (from now on "spectral-temporal metrics") capture important phenological information and can be analyzed on a seasonal, annual or multi-annual basis, allowing for a wide range of applications for characterizing land use systems in space and time.

In this study, we aimed to investigate the potential of Landsat-derived spectral–temporal metrics to reliably separate cropland, pasture, natural savanna vegetation, and forest in a heterogeneous savanna landscape. To estimate the additional value of phenology in our classification approach and to understand limitations related to varying observation densities, we analyzed five datasets of different temporal depth, based on a total of 344 Landsat scenes across four footprints that were acquired between 2009 and 2012. Our overall research objectives were:

- Use spectral–temporal metrics to separate cropland, pasture, and natural vegetation in a heterogeneous savanna landscape.
- Understand the influence of seasonality and observation density on classification results.
- Compare LULC classification solely based on spectral information with classifications of spectral–temporal information.

2. Data and methods

2.1. Study area and description of LULC classes

To evaluate the potential of the Landsat-derived spectral–temporal information for distinguishing LULC classes in the Brazilian Cerrado, we conducted a case study in the Rio das Mortes watershed, a tributary of the Araguaia River that drains into the Tocantins River, the central fluvial artery of Brazil. This macro-catchment is part of the Chapada dos Guimarães, a plateau region located near the city of Cuiabá in Mato Grosso state. The watershed boundaries were derived from a digital elevation model with 90 m spatial resolution (SRTM, 2008; van Zyl, 2001) and delineate an area of about 18,000 km^2 (Fig. 1). The regional climate exhibits a distinct dry season from May to September and corresponds to a tropical savanna (Aw), according to the Köppen climate classification (Moreno et al., 2005). Annual precipitation of 1,500 mm accumulates mainly between November and March.

The watershed is located in the Cerrado biome, and natural vegetation comprises woody grassland (Campo Cerrado), woodland Cerrado (Cerrado sensu stricto), woody Cerrado (Cerradão), and gallery forest along the rivers (Furley, 1999). Parts of the watershed are also covered by wetlands, which are located close to the northern border of the Sangradouro-Volta Grande Indigenous Reserve, an area of approximately 1,000 km^2.

The Rio das Mortes watershed has undergone significant LULC change since the 1970s and is one of the major agricultural production centers of Mato Grosso, inhabited by 80,000 people. Despite its unique