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

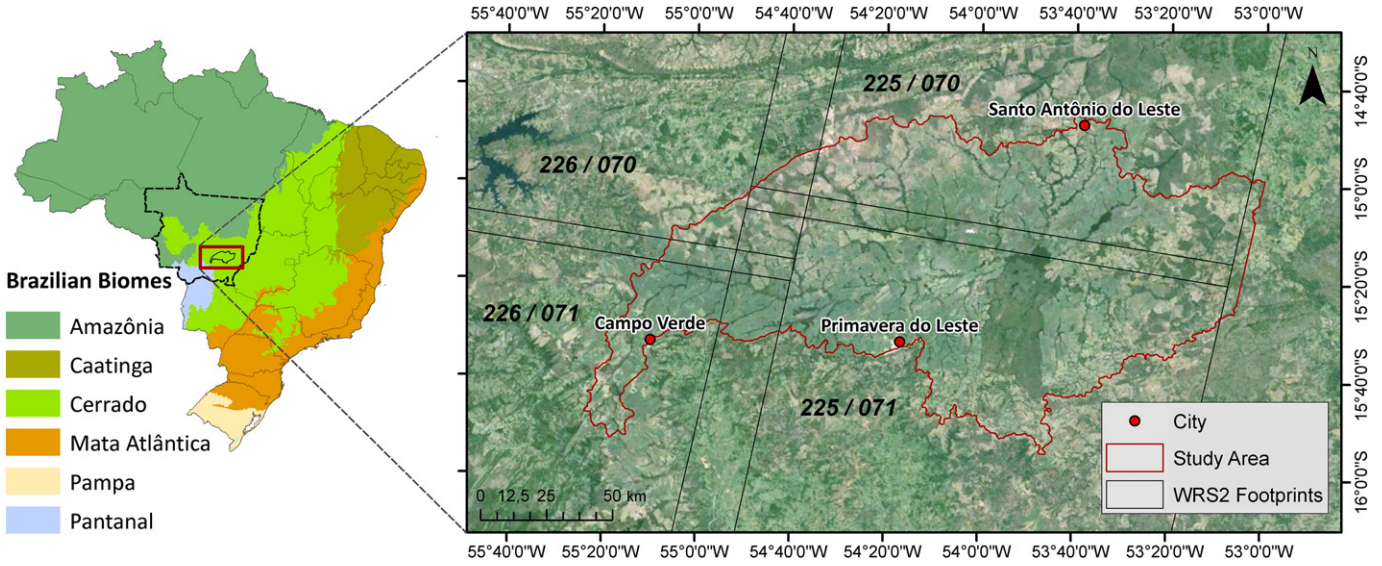

Fig. 1. Left: Distribution of the predominant biomes in Brazil and Mato Grosso state. Right: Close-up of the study area based on Google Earth imagery.

ecological functions, the Brazilian forest code allows conversion of 65% of natural vegetation on private properties in the Cerrado (Federal Law 12.727, 2012 summarized by Soares-Filho et al., 2014), which has permitted major land cover changes over the last 40 years. Land conversion started in the 1970s with the establishment of pastures for cattle ranching. In the 1980s, soybean cultivation was introduced, and corn and cotton have been cultivated in the study area since the 1990s. Production of soy boomed in the early 2000s, and national statistics show that the cropped area reached its maximum in 2005 for the six major municipalities in the study area (approx. 10,600 km^2, IBGE, 2010). After 2005, cropland expansion stopped due to changes in world market prices and yield gaps caused by a new soybean disease, the "Asian rust" (Arvor, Dubreuil, Villar, Ferreira, & Meirelles, 2009). To avoid new outbreaks of the disease, a phytosanitary measure was introduced that prohibits the cultivation of soy from July to September (Seixas & Godoy, 2007). Due to these restrictions, the main sowing season for soy starts with the wet season between September and October. Sowing and harvesting practices differ for individual rotation systems (single- or "double-cropping") and crop type (e.g. cotton, corn), but the production cycle usually ends with the last harvests at

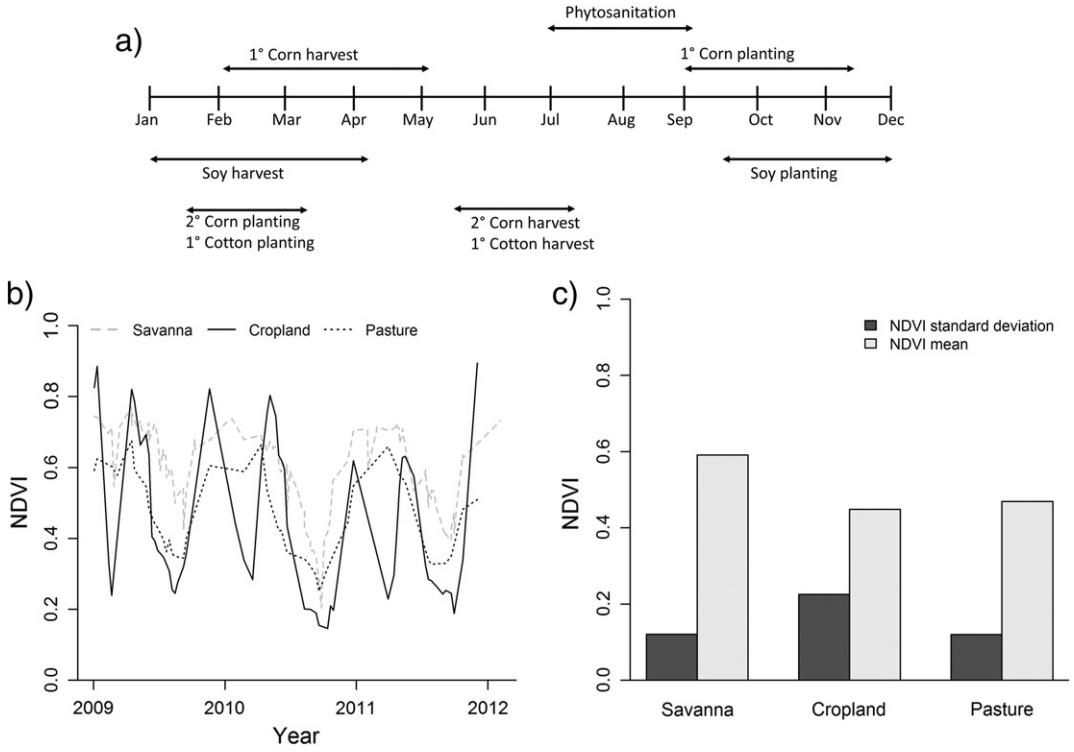

Fig. 2. a) Crop calendar of common cropping practices in Mato Grosso since 2006 (modified after CEGN, 2008). b) Representative NDVI time series for a double cropping system, pasture area, and savanna vegetation. c) Mean and standard deviation of the respective NDVI values over time in b).

the beginning of the dry season in July (Fig. 2a, Arenas-Toledo, Epiphanio, & Galvão, 2009; Arvor, Meirelles, Dubreuil, Bégué, & Shimabukuro, 2011).

To characterize the predominant LULC classes in the study region, we differentiate forest, savanna, cropland, pasture, non-vegetated land and water bodies. The non-vegetated land class includes the built-up environment, open soils and rock outcrops. The savanna class comprises woody grassland (Campo Cerrado), woodland Cerrado (Cerrado sensu stricto) and wooded Cerrado (Cerradão). Croplands include both "single" and "double-cropping" systems for soy, corn and cotton. During our field visits it was apparent that these systems can also include fallow cycles with grassland and temporary cattle ranching. Pasture areas contain planted grasslands used exclusively for cattle ranching. These areas are often characterized by scattered residuals of woody savanna vegetation, indicating the absence of large-scale mechanized agriculture.

2.2. Analysis strategy

Our analysis strategy comprised two stages. First, we evaluated the overall accuracy and spatial consistency of the LULC classification (Section 2.6), which was based on 31 spectral–temporal variability metrics. These metrics were obtained from all available Landsat acquisitions within a three-year period, from January 2nd 2009 to February 28th 2012 (hereafter referred to as "benchmark dataset", Table 1). The analysis method requires that no major land conversions occurred during this period of time. This assumption is supported by field interviews, existing literature and agrarian statistics (Grecchi et al., 2013; IBGE, 2010).

Second, we performed a sensitivity analysis to better understand limitations of spectral–temporal metrics as input for LULC classifications in relation to varying observation densities across the dry and wet seasons. For this sensitivity analysis, we performed a stepwise reduction of the number of observations included in the benchmark dataset, resulting in three additional datasets of different temporal depth (Table 1). For each of the additional datasets, a decision tree based LULC classification was conducted, and classification accuracy, confusion between different classes of interest and spatial patterns in the mapping results were compared.

The first reduction of the benchmark dataset limits the available observations to a 1-year instead of a 3-year period (2011–2012, "single year dataset"). The single year dataset allows investigating the robustness of our classification approach towards a considerably reduced observation density (70% reduction compared to benchmark dataset, Table 1). Also, it resembles the often considered "yearly case" that avoids inter-annual change in a time series. The second reduction limits the available observations to the 2011 dry season. This is a frequent setting, as observations can be poor or entirely missing for the wet season due to dense cloud cover. The third reduction refers to a best observation composite without information from spectral–temporal variability metrics (Section 2.5). This last setup allows us evaluating the added value of phenological information included in the three spectral–temporal datasets (benchmark, single year, dry season) and serves as a comparison to single image classification approaches.

2.3. Satellite imagery, atmospheric correction and cloud masking

We exclusively used TM and ETM+ Landsat data in precision terrain corrected (L1T) format and converted all imagery to surface reflectance using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) atmospheric correction algorithm (Masek et al., 2006) to ensure comparability of input data across different Landsat sensors, footprints, and acquisition dates. Cloud and shadow masks were calculated for all imagery using FMask (Zhu & Woodcock, 2012). The FMask algorithm uses textural and spectral information to detect cloud and cloud shadows on a probabilistic basis. To capture as many clouds and cloud shadows as possible, the cloud probability threshold was set to a highly conservative value (1%). Scenes from different UTM zones were reprojected to WGS 84 UTM 21 South.

2.4. Background and extraction of spectral–temporal variability metrics

Spectral similarities between cropland, pasture, and natural savanna vegetation complicate differentiating LULC in the Brazilian Cerrado (Grecchi et al., 2013; Sano et al., 2010). In contrast, temporal profiles of the normalized difference vegetation index (NDVI) show differences in amplitude and frequency for representative cropland, pasture, and savanna sites (Fig. 2b). Accordingly, the use of spectral–temporal variability metrics should allow for a better separation of relevant LULC classes through the incorporation of seasonal characteristics. In the Brazilian Cerrado, the cropland profile is typically dominated by rapid changes between full vegetation cover and bare soil twice a year, indicating a "double-cropping" system (Fig. 2a & b). Savanna vegetation is characterized by a higher and more stable NDVI signal, which decreases sharply during the dry season. NDVI values of pasture follow a moderate sinusoidal curve, with pronounced phases of transition between dry and wet season. Since these phenological patterns translate into a low standard deviation for pasture and savanna areas and a low mean value for pasture and cropland areas, the combination of both variability metrics should support distinguishing the three major LULC classes in the Brazilian Cerrado (Fig. 2c).

We chose five statistical metrics (mean, median, standard deviation, 75% quartile, and interquartile range), which should capture most phenological characteristics of the investigated LULC classes. We carefully selected these metrics to minimize potential outlier effects across different observation densities of the investigated datasets (Table 1). To include all spectral information, the five statistical metrics were calculated for each spectral band individually. We also calculated an SWIR-Index as the sum of SWIR/NIR reflectance normalized over the number of clear observations. The SWIR-Index serves as an integrated measure which potentially supports the separation of pasture and cropland areas. Metrics were computed from all cloud/cloud shadow-free observations as described by Griffiths, van der Linden, Kuemmerle, and Hostert (2013) and resulted in a total of 31 spectral–temporal metrics per pixel (Table 2).

Table 1
Temporal and spectral properties of the datasets.

	Benchmark dataset	Single year dataset	Dry season dataset	August composite
Period	Jan 2nd 2009–Feb 28th 2012	Jan 1st 2011–Feb 28th 2012	Apr 29th 2011–Oct 22nd 2011	Aug 2011*
No. predictor variables	31 spectral–temporal metrics	31 spectral–temporal metrics	31 spectral–temporal metrics	6 spectral bands
No. of images	344	104	54	6
Cloud free obs. per pixel	54	17	12	1
Wet–Dry season ratio	0.52	0.65	n.a.	n.a.

The number of cloud free observations per pixel is calculated as the median number of clear observations for each spectral–temporal dataset. A wet–dry season ratio of 0.5 indicates twice as much dry season observations as wet season observations.
* For creating the target date image, best pixel observations from mid-July to mid-August were used.

Table 2
Overview of the 31 spectral–temporal variability metrics.

No. of variables	Metric	Spectral bands
6	Mean	1–5, 7
6	Median	1–5, 7
6	Standard deviation	1–5, 7
6	75% Quantile	1–5, 7
6	Interquartile range (25%–75%)	1–5, 7
1	SWIR-Index	4, 5, 7

2.5. Creating a best observation composite

We created a best observation composite over the four footprints using the compositing algorithm developed by Griffiths, van der Linden, Kuemmerle, and Hostert (2013). The algorithm uses scoring functions to identify the best available observation based on each pixel's spatial distance to clouds as well as the temporal distance (i.e. days) to the target day-of-year (DOY). We processed all available imagery for the year 2011 and defined August 8th 2011 as the target date for compositing (DOY = 220). This date was chosen because it falls in the middle of the dry season, which facilitates comparisons across datasets (e.g. dry season dataset) and minimizes temporal variations in the final composite triggered by cloud contamination. The final best observation composite showed high temporal consistency, with 68% of the pixels observed on August 12th 2011 and 99% of the pixels observed between July 19th 2011 and August 12th 2011.

2.6. Land cover mapping and validation

LULC mapping for all datasets was performed using a random forest (RF) classifier (Breiman, 2001). Being a machine learning classifier, the RF classifier creates multiple decision trees, which are calibrated by a random subset of training data and input variables. Key advantages for remote sensing applications are computational performance, classification accuracy and the ability to handle large and heterogeneous feature spaces based on relatively few training samples (Prasad, Iverson, & Liaw, 2006; Rodriguez-Galiano, Ghimire, Rogan, Chica-Olmo, & Rigol-Sanchez, 2012). Model calibration and prediction was carried out using the statistical software CRAN R (R Development Core Team, 2011) and the "randomForest" package (version 4.6–7 by Liaw & Wiener, 2002).

For model calibration, we randomly selected 500 training points. These were amended with 126 interactively selected points to account for underrepresented classes (Table 3). To classify the training points, we utilized very high-resolution imagery of Google Earth from 2009 to 2012 (Quickbird, GeoEye). If recent Google Earth imagery was not available, Landsat images and calculated variability metrics were used as reference data. In uncertain cases, the labeling was conducted in favor of conservative cropland estimates, as it is the most dominant and heterogeneous land use class in the study area. A minimum distance criterion of 350 m was defined to avoid spatial clustering of training points and related autocorrelation. Due to spatial coherence of the water class and areas of non-vegetated land, no minimum distance criterion was employed for these classes.

Table 3
Number of training and validation samples for each LULC class.

Class name	Training samples	Validation samples
Forest	65	93
Cerrado	108	84
Cropland	232	115
Pasture	143	62
Water	37	74
Non-vegetated land	41	42
Total	626	470

Unequal sample sizes in the training data can lead to an underestimation of classes with low sample sizes during the classification processes. We tested for negative effects of an imbalanced training dataset by running the RF model in a balanced mode and setting the "sample" parameter to a constant number of 20 samples per class. In this case, the balanced RF uses all reference data, but draws each bagged sample in a stratified manner. Hence, the overall training data are imbalanced, but each tree is trained in a balanced way. As the balanced RF prediction resulted in higher class confusions, we employed the unbalanced RF model for all further analyses, using a standard parameterization for the number of trees (500), the number of training data for each tree (63% of sample size per class), and the number of predictor variables tried at each split (5, approximated by the square root of 31 predictor variables).

The predicted LULC maps were validated using a stratified random sampling design. We initiated the validation with 62 random samples per class and a minimum distance of 350 m to avoid spatial autocorrelation. Due to the minimum distance criterion, only 42 points for the non-vegetated land class were selected, resulting in 362 validation points. We used this initial validation dataset to derive a first estimate of user's and producer's accuracies. Based on these estimates we calculated a sufficient number of validation samples following the procedure of Cochran (1977) and targeting a standard error for the overall accuracy of 0.01. We determined the adequate sample size of 470 validation points and added 108 additional, stratified randomly selected samples to the initial validation dataset for the final accuracy assessment (Table 3).

We performed the validation point labeling in the same way as the training data labeling, primarily using Google Earth imagery. To account for the stratified random sampling design, overall accuracies and area estimates were adjusted after Olofsson, Foody, Stehman, and Woodcock (2013). In addition, confidence intervals for the adjusted overall accuracy were calculated to report uncertainty of the accuracy estimates introduced by the number of validation points (Congalton & Green, 2009).

3. Results

3.1. Benchmark classification

Cropland was identified as the dominant land use class, covering approximately 51% (9,096 km^2) of the study area (Fig. 3, Fig. 4). Savanna areas were ranked as the second largest land cover type (24%, 4,298 km^2) and can be found in the indigenous reserve, along gallery forests and within the extensively managed eastern and northern parts of the watershed.

The savanna areas form a mosaic with patches of cattle farming which cover 15% (2,682 km^2) of the study site. Forest areas are dominantly located along river streams and account for 9% (1,737 km^2) of the total land cover. In the southwestern part of the watershed, small amounts of forest show a characteristic rectangular pattern which is typical for eucalyptus plantations. Non-vegetated land was identified in the cities of Campo Verde, Primavera do Leste and Santo Antônio do Leste. In addition, this class was detected in close proximity to cropland and pasture areas representing 0.8% of the study area.

The LULC classification for the benchmark dataset achieved an adjusted overall accuracy of 93% with a 95% confidence interval margin of ±2%. Highest reliability was observed for the forest, savanna and water classes that revealed user's and producer's accuracies greater than 90% (Table 4). No commission error was observed for the cropland class.

Main class confusion arose from the overestimation of pasture at the expense of cropland. This effect resulted in a commission error of 30% for the pasture class and an omission error of 11% for the cropland class after error adjustment. A minor omission error of 1% for pasture originated from the overestimation of non-vegetated land on pasture areas, leading to a commission error of 16% for non-vegetated land.

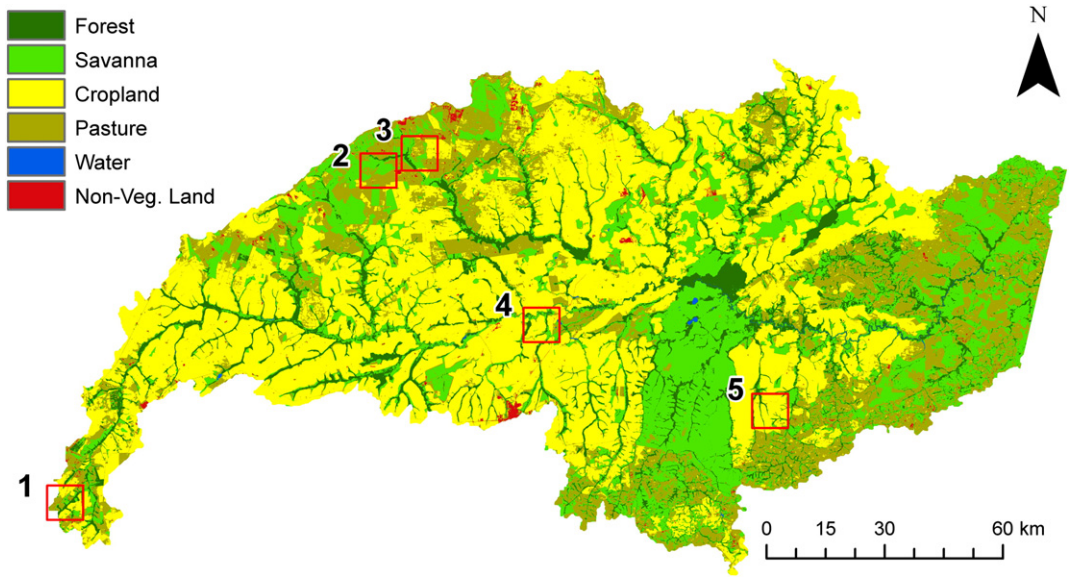

Fig. 3. Classification results of the Rio das Mortes watershed based on the benchmark dataset. Red squares indicate subsets enlarged in Fig. 6.

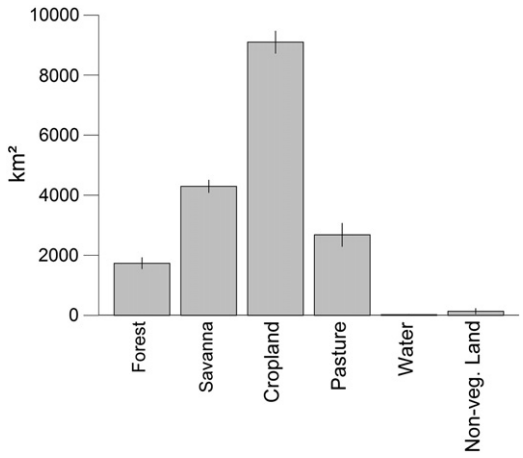

Fig. 4. Error adjusted area estimates and 95% confidence intervals for all six LULC classes. Results are based on the benchmark dataset.

3.2. Sensitivity analysis of temporally reduced datasets

Error adjusted overall accuracies decreased with decreasing temporal depth from 93% to 72% (Table 5).

Errors of omission and commission increased for most LULC classes, including the savanna, cropland, and pasture class (Fig. 5). Here, reducing the temporal observation window led to an overestimation of savanna and cropland areas at the expense of the pasture class. As a result, the omission error of the pasture class increased to 20%–50% when not using the benchmark dataset. In contrast, omission and commission errors for the forest and water classes remained stable for all datasets. This picture changed when adjusting omission errors after Olofsson et al. (2013) to account for the heterogeneity of spatial extents of LULC classes. As a result, land use classes of small spatial extent yielded very high omission errors after adjustment (80–88% for water and non-vegetated land), especially for the dry season dataset and the August composite.

In general, the overestimation of cropland at the expense of pasture was most visible for the dry season dataset and August composite (Fig. 6, subset 3). The August composite also showed SLC-off effects and some unrealistic spatial patterns for all LULC classes, indicating a strong bias due to the cropland and pasture harvest regimes (Fig. 6, subset 1, 2, and 4). Irrigated cropland areas, for example, were classified as savanna or forest areas depending on the growth stage of the respective crops (circular features in Fig. 6, subset 4). In contrast, high agreement was observed between classification results of the benchmark and the single year dataset (90% of pixels are classified identically).

4. Discussion and conclusions

LULC classification of the benchmark dataset achieved an overall accuracy of 93%, and class-wise uncertainty was generally low, including

Table 4
Confusion matrix of validation results for the benchmark dataset.

	Reference data							
Classified data	Forest	Savanna	Cropland	Pasture	Water	Non-veg. land	Total	User's accuracy
Forest	90	3	0	0	0	0	93	0.97
Savanna	2	78	0	0	0	0	80	0.98
Croplands	0	0	92	0	0	0	92	1.00
Pasture	1	2	21	58	0	1	83	0.70
Water	0	0	0	0	73	0	73	1.00
Non-veg. land	0	1	2	4	1	41	49	0.84
Total	93	84	115	62	74	42	470	
Producer's accuracy	0.97	0.93	0.80	0.94	0.99	0.98		
(Error adjusted)	(0.91)	(0.97)	(0.89)	(0.99)	(0.92)	(0.67)		

User's and producer's accuracy is normalized between 1 (100%) and 0 (0%). User's and producer's accuracy translates into omission and commission error as follows: omission error = 1–producers accuracy, commission error = 1–user's accuracy.

Table 5
Adjusted and non-adjusted overall accuracies for all four datasets and the corresponding 95% confidence interval margins (for dataset description see Table 1).

	Benchmark	Single year	Dry season	August composite
Adjusted overall accuracy [%] (Overall accuracy)	92.61 ± 2.15 (92.19)	85.95 ± 3.11 (87.87)	79.43 ± 3.67 (84.89)	71.61 ± 4.05 (79.15)

the target classes of cropland and pasture. We therefore conclude that our spectral–temporal classification approach provided a reliable separation between cropland, pasture, and natural savanna vegetation. Remaining errors in the benchmark dataset mainly relate to the slight overestimation of the pasture class and can be explained by two factors: firstly, we labeled our training data in favor of a conservative estimation of the cropland class, as it is the most dominant and heterogeneous land use class in the study area. Therefore, the commission error of the cropland class is generally low at the expense of a slightly overestimated pasture class. Secondly, there are extensively managed cropland areas that are used as pastures during one or two rotational cycles. These mixed systems appear as croplands in the validation data but are most likely classified as pastures. Major land conversions during our observation period could also hamper the explanatory power of our spectral–temporal variability metrics. However, such land conversions are unlikely, given our analysis period compared to knowledge on recent land use dynamics (Section 2.2).

To better understand the reliability of our approach towards annual or intra-annual classifications we conducted a sensitivity analysis. The sensitivity analysis showed that annual classifications are still rather reliable with an overall classification accuracy of 86% compared to 93% for the benchmark classification. The observed classification error results from the overestimation of croplands at the expense of pastures. As cash crops are mainly grown in the wet season, the amplitude of croplands between the dry and the wet season is unique (Fig. 2a & b). For the single year dataset, less wet season observations are available resulting in a higher phenological similarity between cropland and pasture and higher commission errors for the cropland class (Table 1). However, the influence of missing wet season observations becomes more important for the dry season dataset. Here classification accuracy decreases from 85% (single year dataset) to 79% indicating that seasonal information is crucial for separating the chosen LULC classes. This outcome is in line with results from Prishchepov, Radeloff, Dubinin, and Alcantara (2012), who investigated the importance of image dates for LULC classifications in Eastern Europe using multi-temporal Landsat imagery.

The overall importance of temporal information for our LULC classification was investigated by conducting a classification based on spectral information only, using a composited and cloud-free image of 6 spectral bands (August composite). The classification results revealed considerably lower overall accuracies and inconsistent spatial patterns of LULC classes compared to multitemporal classifications (benchmark, single year and dry season dataset). The spatial patterns relate to different harvesting times and growth stages of croplands in August, when transitioning between the phytosanitary phase (July to September) and the first planting of soy and corn (Fig. 2a). In general, these results confirm that classifications based on monotemporal data only will not allow separating LULC classes with similar spectral properties during different phenological stages. Therefore, specifically the classification of phenologically complex land use systems will profit from our deep time-series approach.

To our best knowledge, this is the first study employing Landsat-derived, spectral–temporal information for classifying LULC in the heterogeneous Brazilian Cerrado. In a regionally comparable Landsat-based study, Grecchi et al. (2013) reported an overall accuracy of 85% when including MODIS data and pre-existing land cover maps, and not separating forest and savanna areas. This supports our finding that multitemporal information is important in such a setting to separate spectrally similar classes reliably. Other studies that exclusively employed Landsat data reached overall accuracies of above 80% only when fusing cropland and pasture into an agro-pastoral class (Brannstrom et al., 2008; Jepson, 2005). Such aggregation, however, limits linking remote sensing based results with ecological or management processes of great importance, e.g. how pastoralism versus intensive cropping alters savanna ecosystems. Our results therefore emphasize the potential of employing high-resolution spectral–temporal variability metrics for identifying LULC in heterogeneous savanna regions.

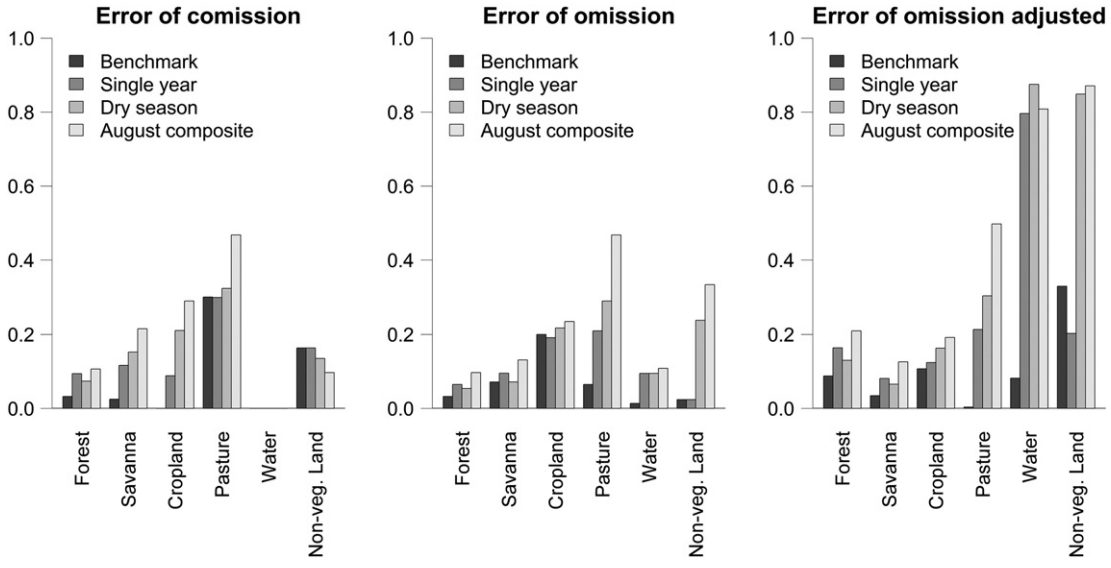

Fig. 5. Errors of omission and commission for the four different datasets (for dataset description see Table 1). Adjustment of the omission error was conducted following Olofsson et al. (2013) and strongly affected the LULC classes with small spatial extent (water and non-veg. land).

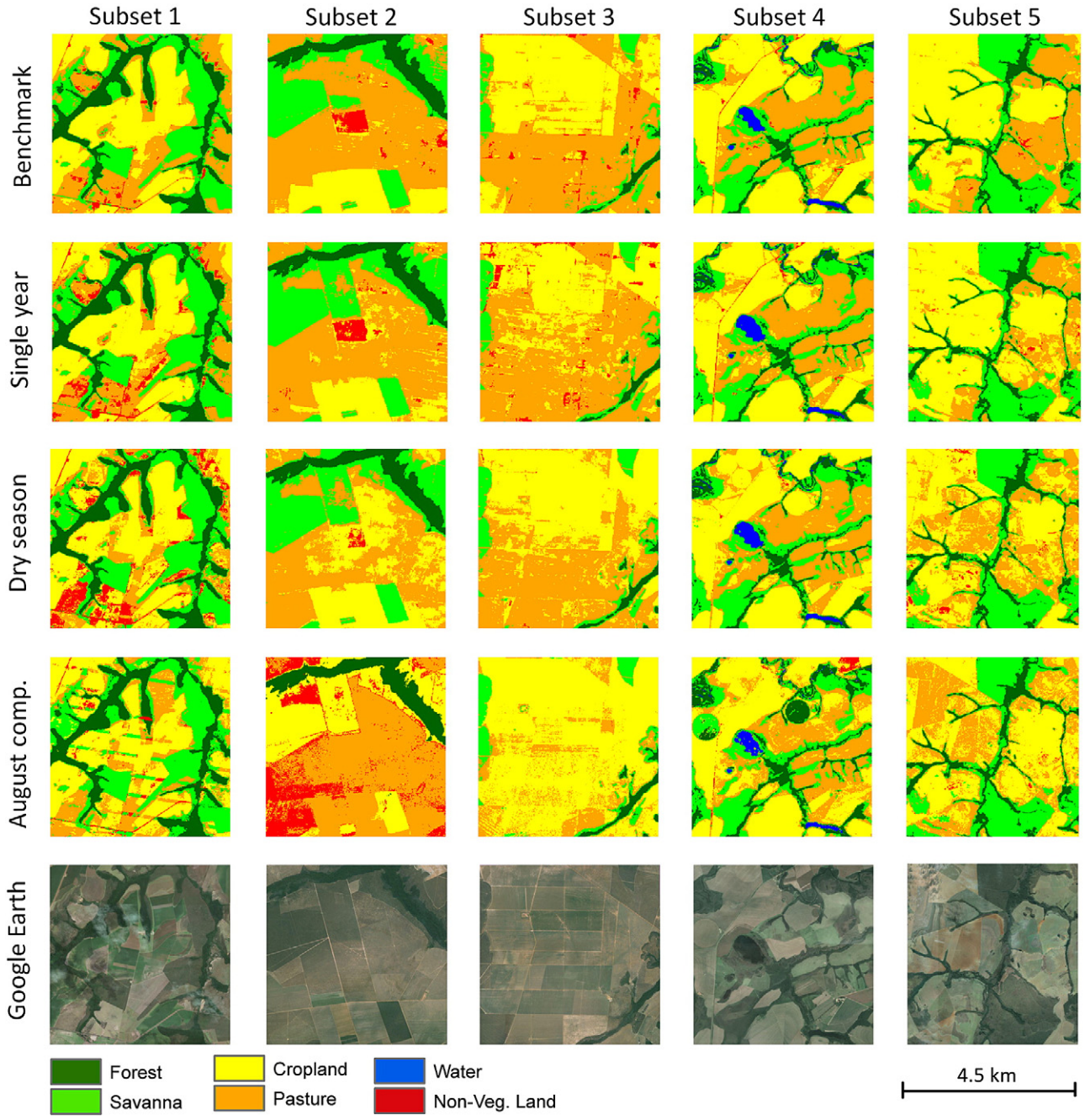

Fig. 6. Spatial patterns of LULC for five subsets in the study area and the four datasets. High resolution imagery from Google Earth is shown for visual comparison (imagery acquisition dates of subsets 1–5: June 2010, unknown, unknown, March 2013, June 2003). For subset locations see Fig. 3.

So far, it is a great challenge to operationally observe LULC in savanna landscapes on a large spatial scale. Sano et al. (2010) used a monotemporal wall to wall mosaic for manual interpretation of 170 Landsat images for the entire Cerrado Biome and reported high class confusion between croplands and pastures (OA = 71%). These results are in line with our monotemporal classification (August composite, OA = 72%) and suggest great potential for improvement by using multitemporal Landsat data. A preliminary screening of the Landsat archive showed that 80% of the entire Cerrado biome has a similar Landsat coverage compared to our study area (+−20% scene availability in 2009–2012), rendering our multitemporal classification approach highly applicable. However, the northern Cerrado might be a critical region due to higher cloud coverage during the wet season (Sano, Ferreira, Asner, & Steinke, 2007) and frequent land conversions since 2002 (LAPIG 2014; Rocha et al., 2011). These areas need to be analyzed within a more restricted temporal window (annual or seasonal), which limits data availability even more. In case of data scarcity and missing seasonal information, our approach could profit from an external cropland mask. The cropland mask might be useful to minimize the confusion between the pasture and cropland class. Several studies have shown that high temporal coverage of MODIS data allows to reliably identify cropland, even though it fails to separate pastures from natural savanna vegetation because of the limited spatial resolution (Arvor, Milton, Meirelles, Dubreuil, & Durieux, 2011; Fritz et al., 2011; Galford et al., 2008).

The methods presented here use satellite data across multiple seasons to reliably differentiate LULC classes, but our approach could also be used to identify land use change by comparing several multi-season classifications. Since post classification change detection is highly sensitive to error propagation from single classifications (Coppin, Jonckheere, Nackaerts, Muys, & Lambin, 2004), our method will likely improve change detection accuracies if enough imagery is available across phenological cycles—a scenario that will become realistic once Landsat and forthcoming entinel-2S data can be combined (Drusch et al., 2012; Roy et al., 2014).

5. Outlook

Our spectral–temporal classification approach provides a reliable separation between cropland, pasture, and natural savanna vegetation. Similar accuracy and consistency of LULC classification could not be achieved from spectral information alone, proving the additional value of temporal information for LULC classifications in complex savanna landscapes. We are optimistic that our spectral–temporal classification approach will be expandable within the Cerrado biome and transferable to other savanna regions worldwide. This is specifically true once data scarcity in more clouded or dynamic savanna regions is further reduced by combining Landsat and upcoming Sentinel-2 imagery. Multi sensor constellations will therefore be an important step towards annual LULC mapping and operational monitoring of savanna ecosystems, which will be crucial for conservation purposes and sustainable land management strategies.

Acknowledgements

This research is part of the Brazilian–German cooperation project on “Carbon sequestration, biodiversity and social structures in Southern Amazonia (CarBioCial”, financed by the German Federal Ministry of Research and Education (BMBF; project no. 01LL09021). We are further gratefully acknowledge contributions by the Sense-Carbon Project funded by the German Federal Ministry of Economy and Infrastructure (BMWi; project no. 50EE1254). The research presented here contributes to the Global Land Project (http://www.globallandproject.org) and the Landsat Science Team (http://landsat.usgs.gov/Landsat_Science_Team_2012-2017.php).

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.rse.2014.10.014. These data include Google maps of the most important areas described in this article.

References

Adami, M., Rudorff, B.F.T., Freitas, R.M., Aguiar, D.A., Mello, M.P., Adami, M., et al. (2011). Remote sensing time series to evaluate direct land use change of recent expanded sugarcane crop in Brazil. *1st World Sustain. Forum*.

Arenas-Toledo, J.M., Epiphanio, J.C.N., & Galvão, L.S. (2009). Crop patterns extraction derived by classic Fourier analysis of EVI-MODIS time-series data to support crop discrimination. *Simpósio Brasileiro de Sensoriamento Remoto, 14*(SBSR), 83–90.

Arima, E.Y., Richards, P., Walker, R., & Caldas, M.M. (2011). Statistical confirmation of indirect land use change in the Brazilian Amazon. *Environmental Research Letters, 6*, 24010.

Arvor, D., Dubreuil, V., Villar, P.M., Ferreira, C.M., & Meirelles, M.S.P. (2009). Développement, crises et adaptation des territoires du soja au Mato Grosso: l'exemple de Sorriso. *Confins. Revue franco-brésilienne de géographie/Revista franco-brasilera de geografia, 6*, http://dx.doi.org/10.4000/confins.5934.

Arvor, D., Meirelles, M.S.P., Dubreuil, V., Bégué, A., & Shimabukuro, Y.E. (2011b). Analyzing the agricultural transition in Mato Grosso, Brazil, using satellite-derived indices. *Applied Geography, 32*, 702–713.

Arvor, D., Milton, Jonathan, Meirelles, M.S.P., Dubreuil, V., & Durieux, L. (2011a). Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. *International Journal of Remote Sensing, 32*, 7847–7871.

Assunção, J., Gandour, C., & Rocha, R. (2013). DETERring deforestation in the Brazilian Amazon: Environmental monitoring and law enforcement. *A CPI report. Climate Policy Initiative Rio de Janeiro*.

Barona, E., Ramankutty, N., Hyman, G., & Coomes, O.T. (2010). The role of pasture and soybean in deforestation of the Brazilian Amazon. *Environmental Research Letters, 5*, 24002.

Barretto, A.G., Berndes, G., Sparovek, G., & Wirsenius, S. (2013). Agricultural intensification in Brazil and its effects on land-use patterns: An analysis of the 1975-2006 period. *Global Change Biology, 19*(6), 1804–1815, http://dx.doi.org/10.1111/gcb.12174.

Batlle-Bayer, L., Batjes, N.H., & Bindraban, P.S. (2010). Changes in organic carbon stocks upon land use conversion in the Brazilian Cerrado: A review. *Agriculture, Ecosystems & Environment, 137*, 47–58.

Brannstrom, C., Jepson, W., Filippi, A., Redo, D., Xu, Z., & Ganesh, S. (2008). Land change in the Brazilian savanna (Cerrado), 1986-2002: Comparative analysis and implications for land-use policy. *Land Use Policy, 25*, 579–595.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32.

CEGN (2008). *Periodo de safra da soja, milho e cana de acucar no Brasil.* Centro de Estudos em Gestão Naval (http://www.gestaonaval.org.br/arquivos/documentos/Log%C3%ADstica/CEGN%20-%20Per%C3%ADodo%20de%20safra%20da%20soja%20milho%20e%20cana-de-a%C3%A7%C3%BAcar%20no%20Brasil.pdf. Last access: 28.1.2014).

Cochran, W.G. (1977). *Sampling techniques. 98*. (pp. 259–261). New York: Wiley and Sons, 259–261.

Congalton, R.G., & Green, K. (2009). *Assessing the accuracy of remotely sensed data. Principles and practices* (2nd ed.). Boca Raton: CRC Press/Taylor & Francis.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., & Lambin, E. (2004). Digital change detection methods in ecosystem monitoring: A review. *International Journal of Remote Sensing, 25*, 1565-1596–1565-1596.

Development Core Team, R. (2011). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., et al. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment, 120*, 25–36.

Eiten, G. (1972). The cerrado vegetation of Brazil. *The Botanical Review, 38*, 201–341.

Federal Law 12.727 (2012). Brazilian Forest code. www.planalto.gov.br/ccivil_03/_Ato2011-2014/2012/Lei/L12727.htm (Last access 28.08.2014)

Ferreira, N.C., Ferreira, L.G., Huete, A.R., & Ferreira, M.E. (2007). An operational deforestation mapping system using MODIS data and spatial context analysis. *International Journal of Remote Sensing, 28*, 47–62.

Ferreira, L.G., Yoshioka, H., Huete, A., & Sano, E.E. (2003). Seasonal landscape and spectral vegetation index dynamics in the Brazilian Cerrado: An analysis within the Large-Scale Biosphere-Atmosphere Experiment in Amazônia (LBA). *Remote Sensing of Environment, 87*, 534–550.

Fritz, S., See, L., McCallum, I., Schill, C., Obersteiner, M., van der Velde, M., et al. (2011). Highlighting continued uncertainty in global land cover maps for the user community. *Environmental Research Letters, 6*, 44005.

Fritz, S., See, L., You, L., Justice, C., Becker-Reshef, I., Bydekerke, L., et al. (2013). The need for improved maps of global cropland. *Eos, Transactions American Geophysical Union, 94*, 31–32.

Furley, P.A. (1999). The nature and diversity of neotropical savanna vegetation with particular reference to the Brazilian cerrados. *Global Ecology and Biogeography, 8*, 223–241.

Galford, G., Mustard, J., Melillo, J., Gendrin, A., Cerri, C., & Cerri, C. (2008). Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sensing of Environment, 112*, 576–587.

Garnett, T., Appleby, M.C., Balmford, A., Bateman, I.J., Benton, T.G., Bloomer, P., et al. (2013). Sustainable intensification in agriculture: Premises and policies. *Science, 341*, 33–34.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., et al. (2010). Food security: The challenge of feeding 9 billion people. *Science, 327*, 812–818.

Grecchi, R.C., Gwyn, Q.H.J., Bénié, G.B., & Formaggio, A.R. (2013). Assessing the spatio-temporal rates and patterns of land-use and land-cover changes in the Cerrados of southeastern Mato Grosso, Brazil. *International Journal of Remote Sensing, 34*, 5369–5392.

Griffiths, P., Kuemmerle, T., Kennedy, R., Abrudan, I., Knorn, J., & Hostert, P. (2012). Using annual time-series of Landsat images to assess the effects of forest restitution in post-socialist Romania. *Remote Sensing of Environment, 118*, 199–214.

Griffiths, P., Müller, D., Kuemmerle, T., & Hostert, P. (2013). Agricultural land change in the Carpathian ecoregion after the breakdown of socialism and expansion of the European Union. *Environmental Research Letters, 8*, 45024.

Griffiths, P., van der Linden, S., Kuemmerle, T., & Hostert, P. (2013). A pixel-based Landsat compositing algorithm for large area land cover mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 1–14*.

Hansen, M.C., & Loveland, T.R. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment, 122*, 66–74.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., et al. (2013). High-resolution global maps of 21st-century forest cover change. *Science, 342*, 850–853.

Herold, M., Mayaux, P., Woodcock, C.E., Baccini, A., & Schmullius, C. (2008). Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sensing of Environment, 112*, 2538–2556.

Hill, M.J., Hanan, N.P., Hoffmann, W., Scholes, R., Prince, S., Ferwerda, J., et al. (2011). *Remote sensing and modeling of savannas: The state of the dis-union.* International Satellite Remote Sensing of the Environment (ISRSE), Sydney.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., & Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment, 83*, 195–213.

IBGE (2010). *Estatística Agropecuária do Mato Grosso 1988–2010.* Insituto Brasileiro de Geographia e Estatística (http://www.ipeadata.gov.br/. Last access: 28.1.2014).

INPE (2008). *Monitoramento Da Cobertura Florestal Da Amazônia Por Satélites. Sistemas Prodes, Deter, Degrad E Queimadas 2007-2008.* Coordenação Geral de Observação da Terra São José dos Campos (http://www.obt.inpe.br/prodes/Relatorio_Prodes2008.pdf. Last access: 28.1.2014).

Jepson, W. (2005). A disappearing biome? Reconsidering land-cover change in the Brazilian savanna. *The Geographical Journal*, *171*, 99–111.

Kennedy, R., Cohen, W., & Schroeder, T. (2007). Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment*, *110*, 370–386.

Kennedy, R., Yang, Z., Cohen, W., Pfaff, E., Braaten, J., & Nelson, P. (2012). Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan. *Remote Sensing of Environment*, *117–133*.

Klink, C.A., & Machado, R.B. (2005). Conservation of the Brazilian cerrado. *Conservation Biology*, *19*, 707–713.

Kovalskyy, V., & Roy, D.P. (2013). The global availability of Landsat 5 TM and Landsat 7 ETM + land surface observations and implications for global 30 m Landsat data product generation. *Remote Sensing of Environment*, *130*, 280–293.

Kuemmerle, T., Erb, K.H., Meyfroidt, P., Müller, D., Verburg, P.H., Estel, S., et al. (2013). Challenges and opportunities in mapping land use intensity globally. *Current Opinion in Environmental Sustainability*, *5*, 484–493.

Lambin, E.F., & Geist, H.J. (2006). *Land use and land cover change: Local processes and global impacts.* Springer.

LAPIG (2014). *Monitoramento Sistemático dos Desmatamentos no Bioma Cerrado (SIAD-Cerrado)*. Universidade Federal de Goiás (http://www.lapig.iesa.ufg.br/lapig/index.php?option=com_content&view=article&id=38&Itemid=52. Last access: 28.7.2014).

Liaw, A., & Wiener, M. (2002). Classification and regression by random forest. *R News*, *2*, 18–22.

Machado, R.B., Ramos, N., M.B., Pereira, P.G., Caldas, E.F., Gonçalves, D.A., et al. (2004). *Estimativas de perda da área do Cerrado brasileiro.* Conservation International do Brasil, Brasilia http://www.conservation.org.br/arquivos/RelatDesmatamCerrado.pdf. Last access: 28.1.2014.

Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., et al. (2006). A Landsat surface reflectance dataset for North America, 1990-2000. *Geoscience and Remote Sensing Letters, IEEE*, *3*, 68–72.

MEA (2005). *Ecosystems and human well-being. Millennium Ecosystem Assessment.* Washington, DC: Island Press.

Melaas, E.K., Friedl, M.A., & Zhu, Z. (2013). Detecting interannual variation in deciduous broadleaf forest phenology using Landsat TM/ETM + data. *Remote Sensing of Environment*, *132*, 176–185.

Moreno, G., Souza-Higa, T.C.C. de, Maitelli, G.T., & Oliveira, A.U. de (2005). *Geografia de Mato Grosso: Território, sociedade, ambiente: Entrelinhas.*

Morton, D.C., DeFries, R.S., & Shimabukuro, Y.E. (2005). Quantifying cropland expansion in cerrado and transition forest ecosystems with Modis satellite image time series. *Proc. Symp. on Cerrado Land-Use and Conservation: Assessing Trade-Offs Between Human and Ecological Needs.*

Munyati, C., & Mboweni, G. (2013). Variation in NDVI values with change in spatial resolution for semi-arid savanna vegetation: A case study in northwestern South Africa. *International Journal of Remote Sensing*, *34*, 2253–2267.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.A.B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*, 853–858.

Olofsson, P., Foody, G.M., Stehman, S.V., & Woodcock, C.E. (2013). Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment*, *129*, 122–131.

Powell, S.L., Cohen, W.B., Healey, S.P., Kennedy, R.E., Moisen, G.G., Pierce, K.B., et al. (2010). Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sensing of Environment*, *114*, 1053–1068.

Prasad, A.M., Iverson, L.R., & Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems*, *9*, 181–199.

Prishchepov, A.V., Radeloff, V.C., Dubinin, M., & Alcantara, C. (2012). The effect of Landsat ETM/ETM + image acquisition dates on the detection of agricultural land abandonment in Eastern Europe. *Remote Sensing of Environment*, *126*, 195–209.

Ramankutty, N., Foley, J.A., & Olejniczak, N.J. (2002). People on the land: Changes in global population and croplands during the 20th century. *AMBIO: A Journal of the Human, Environment*, *31*, 251–257.

Ramankutty, N., Graumlich, L., Achard, F., Alves, D., Chhabra, A., DeFries, R.S., et al. (2006). Global land-cover change: Recent progress, remaining challenges. *Land-use and land-cover change (pp. 9-39)*. Springer.

Ribeiro, J.F., Sano, S.M., & Silva, J. d. (1981). Chave preliminar de identificação dos tipos fisionômicos da vegetação dos Cerrados. *Congresso Nacional de Botânica* (pp. 124–133).

Rocha, G.F., Ferreira, L.G., Ferreira, N.C., & Ferreira, M.E. (2011). Deforestation detection in The Cerrado Biome between 2002 and 2009: Patterns, trends and impacts. *Revista Brasileira de Cartografia*, *63*.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J.P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *67*, 93–104.

Roy, D.P., Wulder, M.A., Loveland, T.R., CE, W., Allen, R.G., Anderson, M.C., et al. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, *145*, 154–172.

Sano, E., Ferreira, L.G., Asner, G.P., & Steinke, E.T. (2007b). Spatial and temporal probabilities of obtaining cloud-free Landsat images over the Brazilian tropical savanna. *International Journal of Remote Sensing*, *28*, 2739–2752.

Sano, E.E., Rosa, R., Brito, J.L.S., & Ferreira, L.G. (2007a). *Mapeamento de cobertura vegetal do bioma Cerrado.* Planaltina: Embrapa Cerrados.

Sano, E.E., Rosa, R., Brito, J.L.S., & Ferreira, L.G. (2010). Land cover mapping of the tropical savanna region in Brazil. *Environmental Monitoring and Assessment*, *166*, 113–124.

Schmidt, M., Udelhoven, T., Gill, T., & Röder, A. (2012). Long term data fusion for a dense time series analysis with MODIS and Landsat imagery in an Australian Savanna. *Journal of Applied Remote Sensing*, *6*, 63512.

Schuster, C., Schmidt, T., Conrad, C., Kleinschmit, B., & Förster, M. (2015). Grassland habitat mapping by intra-annual time series analysis – Comparison of RapidEye and TerraSAR-X satellite data. *International Journal of Applied Earth Observation and Geoinformation*, *34*.

Seixas, C.D.S., & Godoy, C.V. (2007). *Vazio sanitário: Panorama nacional e medidas de monitoramento.* Londrina: Embrapa (http://www.cnpso.embrapa.br/download/consorcio/simposio2007/PalestrasCAFeSimposio/VazioSanitarioCAF.pdf. Last access: 28.1.2014).

Silva, J.F., Fariñas, Felfili, J.M., & Klink, C.A. (2006). Spatial heterogeneity, land use and conservation in the cerrado region of Brazil. *Journal of Biogeography*, *33*, 536–548.

Soares-Filho, B., Rajão, R., Macedo, M., Carneiro, A., Costa, W., Coe, M., et al. (2014). Cracking Brazil's Forest Code. *Science*, *344*, 363–364.

SRTM (2008). Shuttle radar topographic mission. http://srtm.csi.cgiar.org (Dataset available at)

Trabaquini, K., Bernardes, T., Mello, M.P., Formaggio, A., & Rosa, V.G. (2011). Remote sensing for mapping soybean crop in the Brazilian Cerrado. *AGU Fall Meeting Abstracts* (pp. 566).

van Zyl, J.J. (2001). The shuttle radar topography mission (SRTM): A breakthrough in remote sensing of topography. *Acta Astronautica*, *48*, 559–565.

Wulder, M.A., White, J.C., Goward, S.N., Masek, J.G., Irons, J.R., Herold, M., et al. (2008). Landsat continuity: Issues and opportunities for land cover monitoring. *Remote Sensing of Environment*, *112*, 955–969.

Zhong, L., Gong, P., & Biging, G.S. (2014). Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. *Remote Sensing of Environment*, *140*, 1–13.

Zhu, Z., & Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, *118*, 83–94.