Using Spatial Features to Reduce the Impact of Seasonality for Detecting Tropical Forest Changes from Landsat Time Series

Eduarda M. O. Silveira 1,*, Inácio T. Bueno 1, Fausto W. Acerbi-Junior 1, José M. Mello 1, José Roberto S. Scolforo 1 and Michael A. Wulder 2

1 Forest Science Department, Federal University of Lavras, Lavras 37200-000, Brazil; inaciobueno@gmail.com (I.T.B.); fausto@dcf.ufla.br (F.W.A.-J.); josemarcio@dcf.ufla.br (J.M.M.); jscolforo@dcf.ufla.br (J.R.S.S.)
2 Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, 506 West Burnside Road, Victoria, BC V8Z 1M5, Canada; mike.wulder@canada.ca

* Correspondence: dudalavras@hotmail.com or eduarda.silveira@posgrad.ufla.br; Tel.: +55-35-3822-1700

Received: 19 March 2018; Accepted: 17 May 2018; Published: 23 May 2018

Abstract: In forested areas that experience strong seasonality and are undergoing rapid land cover conversion (e.g., Brazilian savannas), the accuracy of remote sensing change detection is affected by seasonal changes that are erroneously classified as having changed. To improve the quality and consistency of regionally important forest change maps, we aim to separate process related change (for example, spectral variability due to phenology) from changes related to deforestations or fires. Seasonal models are typically used to account for seasonality, but fitting a model is difficult when there are insufficient data points in the time series. In this research, we utilize remotely sensed data and related spectral trends and the spatial context at the object level to evaluate the performance of geostatistical features to reduce the impact of seasonality from the NDVI (Normalized Difference Vegetation Index) of Landsat time series. The study area is the São Romão municipality, totaling 2440 km², and is part of the Brazilian savannas biome. We first create image objects via multiresolution segmentation, basing the objects on the characteristics found in the first image (2003) of the 13-year time series. We intersected the objects with the NDVI images in order to extract semivariogram indices, the RVF (Ratio Variance—First lag) and AFM (Area First lag—First Maximum), and spectral information (average and standard deviation of NDVI values) to generate the time series from these features and to derive Spatio-Temporal Metrics (change and trend) to train a Random Forest (RF) algorithm. The NDVI spatial variability, captured by the AFM semivariogram index time series produced the best result, reaching 96.53% of the overall accuracy (OA) to separate no-change from forest change, while the greatest inter-class confusion occurred using the average of the NDVI values time series (OA = 63.72%). The spatial context approach we presented is a novel approach for the detection of forest change events that are subject to seasonality (and possible miss-classification of change) and mitigating the effects of forest phenology without the need for specific de-seasoning models.

Keywords: remote sensing; geostatistical; semivariogram; change detection; NDVI; forest phenology; savannas

1. Introduction

Savannas are the dominant biome in South Hemisphere, covering approximately 45% of the area of South America [1]. In Brazil, this biome consists of a mosaic of land cover types, undergoing a strong seasonality in climate, accompanied with a widespread occurrence of fires that impose environmental
pressures with the most rapid land conversion in Brazil, exceeding that of the tropical forests [2].
Given the high rates of deforestation and few formally protected areas, the savanna biome is among
the most endangered ecoregions on Earth [3]. Accurate mapping and monitoring of areas undergoing
conversion are needed to indicate the priority areas for conservation and to inform sustainable land
use management, as well as to improve the understanding of the dynamics and related impacts on
carbon balance, nutrient cycling, and water resources [1]. At present, much of the effort estimating
forest changes in Brazil has been focused on the tropical rainforests with less attention dedicated to the
less humid seasonal regions [4]. Many different models can be used to estimate deforestation trends in
tropical regions [5], however, in Brazil, most of these models were developed for the Amazon region
with a few studies representing the Savannas biome [4,6]. Although human-induced disturbances are
threatening these ecosystems, systematic investigations of land cover change are lacking [7].

Remotely sensed data have been widely recognized as essential data sources for comprehensive
mapping and quantification of land cover changes [8–14]. Studies have used the NDVI (Normalized
Difference Vegetation Index) time series from moderate resolution sensors (that is, MODIS) for land
cover change detection benefitting from their frequent revisit time [15,16]. However, images from
medium spatial resolution sources such as Landsat offer more detailed spatial information, providing
insights at smaller spatial scales over a longer period of time. Initiated in 1972, Landsat is the longest
running cross-calibrated globally consistent record of the Earth’s surface at a medium resolution [13].
The ability to utilize the dense time series of imagery from Landsat has been demonstrated for mapping
land cover changes, such as deforestation, forest degradation, and impacts of fire [17–21]. Given the
long time series of Landsat data at the global scale [22], the production of forest change maps is common
as first steps towards estimating biomass and related CO₂ emissions [23] and biodiversity loss. Based on
investigations using Landsat data, the greatest total forest change due to deforestation from 2000 to 2012
of the four climate domains (tropical, subtropical, temperate, and boreal) is the tropical domain [24].

However, estimating forest change by remote sensing is not a trivial task. Satellite images
can contain a combination of natural forest phenology, cloud cover, atmospheric scattering, and
geometric errors [25], impairing the precision of change estimates, especially in tropical areas that
exhibit strong seasonality in photosynthetic activity [12]. When images from different seasons are
acquired, changes caused by phenological differences are present and can result in the mapping of
a change in condition, rather than a change in land cover or land use [26]. To reduce the seasonal
variations captured by remotely sensed images, techniques for de-seasoning time series have been
used [27] on the assumption that seasonal patterns can be identified and removed [15,16,26,28,29].
However, this assumption may not always hold if the time series is for satellite images which are
not acquired at regular intervals or have gaps due to the presence of clouds [12]. To minimize the
impacts of seasonality, methods combining different satellites [30,31] and others that explore the spatial
context [12] have been developed.

Previous studies have shown that pixel-based change detection can suffer from not including
additional information regarding the spatial context [32]. The inclusion of spatial information can
reduce the small or spurious changes [33] and also allow for the incorporation of supplemental spatial
information, such as geostatistical features [34–37]. Although pixel-based approaches also generate
spatial information, the neighborhood used to retrieve the information is often defined by a square
window that is easy to implement. However, they are computational expensive [27], biased along their
diagonals, and can straddle the boundary between two landscape features, especially when a large
window size is used [38].

Here, we propose a new approach to reduce the impacts of the seasonal changes introduced by
forest phenology on forest change detection using the spatial variability of Landsat NDVI time series at
the object level. We hypothesize that variability measures are not affected by vegetation seasonality as
the NDVI variability remains constant in the presence of seasonal changes and substantially increases
in the presence of land cover changes. Thus, we investigated if the geostatistical features can be used
to eliminate seasonal variations in satellite image time series and improve the change detection results.
We used an object-based approach associated with a set of Spatio-Temporal Metrics by computing geostatistical features, specifically the semivariogram indices developed by Reference [35]. We also compared the change detection results to those provided by traditional spectral features (average and standard deviation of NDVI values). Implementation of this approach enables us to enrich object-based remote sensing methods using the spatial information provided by semivariograms that capture the spatial variability of images which are directly related to land cover changes, improving the change detection classification performance in highly seasonal and ecologically important environments.

2. Study Area

The study area is the São Romão municipality in the state of Minas Gerais (MG), Brazil (Figure 1a). Totaling 2440 km², approximately 84% of this area is occupied by land cover classes comprising Brazilian savannas [39] (Figure 1b). The Brazilian savanna’s natural vegetation consists of mixtures of grasslands, shrublands, and woodlands which are found over a gradient of vegetation density, growth form, vertical structure, and biomass [40] (Figure 1c). Riparian forests are found along the rivers and streams of the study area. As vegetation progresses from grasslands to woodlands and forests, there is an increase in cover, height, density, and the basal area of the trees present [41]. Our analysis focused on the main vegetation types found in the study area (Table 1): open grassland, open grassland with sparse shrubs, mixed grassland, shrublands, trees up to seven meters, and riparian vegetation (palm swamps and semideciduous forests) [42,43]. The climate is typically tropical (Aw) with rains concentrated in the October–May period, which characterizes the high seasonality in the region [44].

![Figure 1](image-url)
Table 1. The examples of the savannas land cover classes found in the study area. Landsat TM images Red-NIR, Green-SWIR 1 and Blue-Red composite.

<table>
<thead>
<tr>
<th>Land Cover Classes</th>
<th>Description</th>
<th>Field View</th>
<th>Landsat TM Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grasslands</td>
<td>Open grassland</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shrublands</td>
<td>Open grassland with sparse shrubs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Savanna Woodlands</td>
<td>Mixed grassland, shrubland, and trees up to seven meters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Riparian forests</td>
<td>Palm swamps and Semideciduous forests</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Methodology

We introduce a method for object-based change detection using geostatistical features to derive Spatio-Temporal Metrics from the time series, eliminating the effect of forest phenology. The method is divided into three main steps (Figure 2).

We chose as the change target the Brazilian savannas land-cover classes and the following change agents (Figure 3):

- No-change: seasonal changes due to phenological effects (natural events)
- Forest change: land cover changes such as deforestation (human-induced) and fires (both natural events and human induces)

3.1. Satellite Data and Processing

We downloaded all available cloud-free Landsat TM (Thematic Mapper) and OLI (Operational Land Imager) images between 2003 to 2016 (total of 66 images) from the free and open access USGS Earth Resources Observation and Science (EROS) Center archive (Table 2). The data as downloaded is already processed to surface reflectance using the Landsat Ecosystem Disturbance Adaptive Processing
System (LEDAPS) and L8SR algorithm, respectively. Clouds and cloud shadows have been masked with the CFmask function [45].

We used the NDVI vegetation index [46] for further analysis. The advantages of using indices rather than the original spectral band observations include minimizing the soil and other background effects, reducing data dimensionality, providing a degree of standardization for comparison, and enhancing the vegetation signal [47].

**Figure 2.** The methodology flowchart detailing the three main steps to detect land cover changes over the savanna biome: 1. feature extraction from time series using NDVI trends at the object level; 2. Change detection using the random forest to model the change type as a function of Spatio-Temporal Metrics; 3. evaluation of change detection using accuracy measures.

**Figure 3.** The Landsat TM Red-NIR, Green-SWIR 1 and Blue-Red composite. Examples of no-change class: (a) savannas affected by phenological effects; and forest change classes; (b) savannas undergoing a change by fire; (c) savannas undergoing a change by deforestation.
### Table 2. The image acquisition dates used to generate time series.

<table>
<thead>
<tr>
<th>Sensor Type/Date (YYMMDD)</th>
<th>Number of Scenes per Year</th>
<th>Sensor Type/Date (YYMMDD)</th>
<th>Number of Scenes per Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM/20030704</td>
<td></td>
<td>TM/20090704</td>
<td></td>
</tr>
<tr>
<td>TM/20030720</td>
<td></td>
<td>TM/20090720</td>
<td></td>
</tr>
<tr>
<td>TM/20030821</td>
<td>6</td>
<td>TM/20090805</td>
<td></td>
</tr>
<tr>
<td>TM/20030906</td>
<td></td>
<td>TM/20090821</td>
<td></td>
</tr>
<tr>
<td>TM/20031008</td>
<td></td>
<td>TM/20091125</td>
<td></td>
</tr>
<tr>
<td>TM/20031125</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TM/20040401</td>
<td></td>
<td>TM/20100504</td>
<td></td>
</tr>
<tr>
<td>TM/20040706</td>
<td></td>
<td>TM/20100605</td>
<td></td>
</tr>
<tr>
<td>TM/20040823</td>
<td></td>
<td>TM/20100621</td>
<td></td>
</tr>
<tr>
<td>TM/20040908</td>
<td>6</td>
<td>TM/20100707</td>
<td></td>
</tr>
<tr>
<td>TM/20040924</td>
<td></td>
<td>TM/20100824</td>
<td></td>
</tr>
<tr>
<td>TM/20041111</td>
<td></td>
<td>TM/20100909</td>
<td></td>
</tr>
<tr>
<td>TM/20050506</td>
<td></td>
<td>TM/20110421</td>
<td></td>
</tr>
<tr>
<td>TM/20050607</td>
<td></td>
<td>TM/20110710</td>
<td></td>
</tr>
<tr>
<td>TM/20050725</td>
<td>4</td>
<td>TM/20110912</td>
<td></td>
</tr>
<tr>
<td>TM/20050826</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TM/20060626</td>
<td></td>
<td>OLI/20130426</td>
<td></td>
</tr>
<tr>
<td>TM/20060813</td>
<td></td>
<td>OLI/20130512</td>
<td></td>
</tr>
<tr>
<td>TM/20060914</td>
<td>5</td>
<td>OLI/20130613</td>
<td></td>
</tr>
<tr>
<td>TM/20060930</td>
<td></td>
<td>OLI/20130629</td>
<td></td>
</tr>
<tr>
<td>TM/20061117</td>
<td></td>
<td>OLI/20130731</td>
<td></td>
</tr>
<tr>
<td>TM/20070309</td>
<td></td>
<td>OLI/20140208</td>
<td></td>
</tr>
<tr>
<td>TM/20070426</td>
<td></td>
<td>OLI/20140413</td>
<td></td>
</tr>
<tr>
<td>TM/20070629</td>
<td></td>
<td>OLI/20140702</td>
<td></td>
</tr>
<tr>
<td>TM/20070731</td>
<td>7</td>
<td>OLI/20140819</td>
<td></td>
</tr>
<tr>
<td>TM/20070901</td>
<td></td>
<td>OLI/20141006</td>
<td></td>
</tr>
<tr>
<td>TM/20070917</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TM/20080428</td>
<td></td>
<td>OLI/20150619</td>
<td></td>
</tr>
<tr>
<td>TM/20080514</td>
<td></td>
<td>OLI/20150806</td>
<td></td>
</tr>
<tr>
<td>TM/20080530</td>
<td></td>
<td>OLI/20151009</td>
<td></td>
</tr>
<tr>
<td>TM/20080717</td>
<td></td>
<td>OLI/20151110</td>
<td></td>
</tr>
<tr>
<td>TM/20080802</td>
<td>6</td>
<td>OLI/20151212</td>
<td></td>
</tr>
<tr>
<td>TM/20081005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>OLI/20160418</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>OLI/20160707</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>OLI/20161027</td>
<td></td>
</tr>
</tbody>
</table>

### 3.2. Object-Based Image Analysis

Image segmentation is the division of the satellite image into groups of pixels spatially continuous and spectrally homogeneous, minimizing the within-object variability compared to the between-object variability [48]. We used the multiresolution segmentation algorithm [49] implemented in the eCognition software. eCognition is a powerful development environment for object-based image analysis designed to produce homogeneous image objects based on spatial and spectral homogeneity criteria. The eCognition technology examines image pixels not in isolation, but in the spectral and spatial contexts. It builds up a given segmentation iteratively, recognizing groups of pixels as objects, using color, shape, texture, shape, and the size of objects, as well as their context and relationships [50].

Three different choices are typically considered when defining the analysis unit (image objects) for multitemporal object-based image analysis [51]: (i) image-object overlay: image-objects are generated by segmenting one of the images in the time series. A comparison against other images is then made by simple overlay; (ii) image-object comparison: image-objects are generated by segmenting each image
in the time series independently; and (iii) multitemporal image-object: image-objects are generated by segmenting the entire time series together. Our study focused on the image-object overlay approach, as expected to capture the intra-object heterogeneity well.

Three key segmentation parameters (shape, compactness, and scale) control the size, shape, and spectral variation of segmented image objects [52]. We set the shape and compactness parameter to 0.1 and 0.5, respectively. The scale parameter controls the size of the image objects and sets a homogeneity threshold that determines the number of neighboring pixels that can be merged together to form an image object [53]. The scale of segmentation directly influences the size of the objects connected to the semivariogram predefined criteria (lag distance) and the minimum number of pixels inside each object necessary to generate the semivariogram makes this segmentation parameter crucial at this point. Based upon starting values found in the literature [48], we then used a “trial-and-error” approach [54] to find the appropriate scale parameter [55] that was set to 150 [37]. The image segmentation outputs were assessed based on a visual assessment of segmentation suitability. The segmentation parameters were also defined to generate sufficiently small objects to not be spatially or spectrally confused, yet large enough to allow for calculation of the semivariogram and related geostatistical features.

3.3. Geostatistical and Spectral Features

We used two semivariogram indices [35] as geostatistical features (Table 3). Inside each object, we extracted the NDVI values to generate the experimental semivariogram and calculate the semivariogram indices. The experimental semivariogram is defined as half of the average squared difference between values separated by a given lag, where this lag is a vector in both, distance, and directions [56]. It was estimated using Equation (1):

\[ \gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x) - Z(x + h)]^2, \]  

where \( \gamma(h) \) is the estimator of the semivariance for each distance \( h \), \( N(h) \) is the number of pairs of points separated by the distance \( h \), \( Z(x) \) is the value of the regionalized variable in the point \( x \) and \( Z(x + h) \) is the value of the point \( (x + h) \).

The graphic for spatial variance versus distance \( h \) represents the semivariogram, which allows obtaining the estimate of the variance value for different combinations of pairs of points. The semivariance functions are characterized by three parameters: sill (\( \sigma^2 \)), range (\( \phi \)), and nugget effect (\( \tau^2 \)). The sill parameter is the plateau reached by semivariance values and shows the quantity of variation explained by the spatial structure of the data. The range parameter is the distance where the semivariogram reaches the sill, showing the distance until where the data are correlated. The nugget effect is the combination of sampling errors and variations that happen in scales smaller than the distance between the sampled points [57].

When the semivariogram is calculated at the object level, an important factor for consideration is the lag distance. It should not be larger than the spatial extent of the object; on the other hand, an exceedingly small distance fails to provide a complete description of textural features. We attempted to find an optimal lag distance to ensure that sill values would provide a concise description of data variability. We fixed the number of lags as 20 pixels and the lag size equivalent to image spatial resolution (30 m), resulting in a lag distance of 600 m [37,58].

Using FETEX 2.0 [59], we calculated the semivariogram indices that describe the shape of the experimental semivariograms and, therefore, the properties that characterize the spatial patterns of each image object. In addition, fitting a model is unnecessary, which reduces both processing time and possible errors from choosing a sub-optimal model. FETEX 2.0 is a software application enabling automatic descriptive feature extraction from image-objects. The input data include a multispectral digital image or single band, such as NDVI, and a vector file in shapefile format containing the boundaries of the objects. The output file is a table containing a vector of features for every object
processed. This table of numeric values describing the objects from different points of view can be then used as the input data to other classification software.

Table 3. The geostatistical features: semivariogram indices [35].

<table>
<thead>
<tr>
<th>Indices</th>
<th>Description</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>RVF</td>
<td>Ratio variance—first lag</td>
<td>[RVF = \frac{\gamma_{\text{max},1}}{\gamma_1}]</td>
</tr>
<tr>
<td>AFM</td>
<td>Area first lag—first maximum</td>
<td>[AFM = \frac{1}{2} \left( \gamma_1 + 2 \left( \sum_{i=2}^{\gamma_{\text{max},1}-1} \gamma_1 + \gamma_{\text{max},1} \right) - (\gamma_1[h_{\text{max},1} - h_1]) ]</td>
</tr>
</tbody>
</table>

RVF is the ratio between the values of the total variance \(\gamma_{\text{max},1}\) and the semivariance at first lag \((\gamma_1)\). Since the semivariogram tends to reach the sill near the total variance, this parameter is an indicator of the relationship between the spatial correlation at long and short distances. Its value increases when the high variability at long distances and low variability at short distances occurs. This feature also presents high values for images with large primitives or periodic patterns [35]. AFM is the area between the semivariogram value in the first lag and the semivariogram. This parameter provides information about the semivariogram curvature and is also related to the variability of the data (Figure 4).

![Figure 4](https://example.com/figure4.png)

**Figure 4.** The graphic representation of AFM (Area first lag—first maximum) index in the classical semivariogram: \((\sigma^2)\)—sill, \((\varphi)\)—range, \((\tau^2)\)—nugget effect, \(\gamma(h)\)—semivariance, \(\gamma_{\text{max},1}\)—total variance, and \(\gamma_1\)—semivariance at first lag.

To verify the potential of the semivariogram indices, we also used spectral features for comparative purpose. We extracted the average and standard deviation (STDV) of the NDVI values inside the objects.

### 3.4. Time Series

To generate the time series, we included objects intersecting savanna land cover classes in the Landsat TM image acquired in July 2003 (first image of time series) and excluded classes as water, pasture, bare soil, crops, and urban areas. We used a land cover map [39] and subsequently, we performed a post-classification editing of the imagery carried out by a skilled human interpreter using the false color composite of the corresponding image (see Figure 1b). We overlapped the savanna objects with the remaining NDVI images in order to extract the geostatistical (RVF and AFM indices) and spectral features (average and STDV) and then generated the time series (Figure 5).
3.5. Spatio-Temporal Metrics (STM)

We transformed the geostatistical and spectral features time series into two summary variables: Change and Trend, adapted from Reference [16]. As described by Reference [51], change is the maximum interannual absolute difference in the time series data within the period 2003–2016 (Equation (2)). The trend is the slope of linear regression applied to the full time series (Equation (3)).

\[
\text{Change} = \max (\text{Feature}_t - \text{Feature}_{t+1}),
\]

\[
\text{Trend} = \frac{n \left( \sum_{t=1}^{n} t \cdot \text{Feature}_t \right) - \left( \sum_{t=1}^{n} \text{Feature}_t \right) \cdot \left( \sum_{t=1}^{n} t \right)}{n \left( \sum_{t=1}^{n} t^2 \right) - \left( \sum_{t=1}^{n} t \right)^2},
\]

The Change variable contains information on the greatest positive or negative change between consecutive data. A high value indicates that a high magnitude change occurred or there is an influence of seasonal variation of the feature response. The Trend variable indicates the overall (increasing, decreasing, or none) trend in the amount of green vegetation across the entire period [16].
3.6. Random Forest Algorithm

We used the Spatio-Temporal Metrics (STM) derived from geostatistical and spectral features to train the Random Forest (RF) algorithm [60] (Figure 6). These metrics were obtained from the time series of four features (RVF, AFM, Average, and STDV) extracted from NDVI values inside objects, forming four groups of input files as described in Table 4.

Figure 6. The three main steps to obtain the STM (Spatio-Temporal Metrics) from AFM (Area first lag—first maximum) time series to train the RF (Random Forest) algorithm: (a) Example of an object undergoing a forest fire and; (b) Example of an object relating the no-change class.
Table 4. The group of Spatio-Temporal Metrics (STM) used to train the RF (Random Forest) algorithm. The superscripted numbers are related to group names.

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>STM Description</td>
<td>STM Description</td>
<td>STM Description</td>
<td>STM Description</td>
</tr>
<tr>
<td>Change (^1)</td>
<td>Time series of</td>
<td>Change (^2)</td>
<td>Time series of</td>
</tr>
<tr>
<td></td>
<td>NDVI Average</td>
<td></td>
<td>NDVI STDV</td>
</tr>
<tr>
<td>Trend (^1)</td>
<td></td>
<td>Trend (^2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trend (^3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Change (^4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Time series of</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AFM index</td>
</tr>
</tbody>
</table>

RF is an ensemble method which generates a set of individually trained decision trees and combines their results. RF is a robust non-parametric classifier and can accommodate many predictor variables [19]. The advantages of RF include effectiveness in prediction, efficient implementation on large datasets [60], and insensitivity to noise in the training data [61]. We used the Random Forest package [60] available as an R package [62].

Two parameters need to be set in order to produce the random forest trees: the number of decision trees to be generated (Ntree) and the number of variables to be selected and tested for the best split when growing the trees (Mtry) [63]. The Gini criterion is used to select the split with the lowest impurity (highest homogeneity) at each node. For each tree, the predicted class for each observation is obtained and the class with the maximum number of votes among the trees is the predicted class of an observation [64].

The Ntree required to maintain a given level of accuracy has been assessed by several authors, and the minimum number of trees for optimal classification appears to be between 100 [65] and 300 [66], with the majority of studies setting the Ntree value to 500 because the errors stabilize before this number of classification trees is achieved [65]. Five hundred trees were grown for each classification and the number of features (Mtry = 1) was left at its default values.

3.7. Change Detection Evaluation

Evaluation of change detection outcomes is communicated using a confusion matrix [67] and its accuracy measures: overall accuracy (OA), producer accuracy (PA), user accuracy (UA), and the kappa coefficient (k). A dataset of 1000 objects well-distributed over the study area, with 500 samples per class (Table 5), was assigned based on visual interpretation (Figure 7a). The objects of no-change were those where no changes were found during the entire analyzed period 2003–2016 (Figure 7b). Otherwise, regardless of the date when changes have occurred, the objects undergoing a deforestation or fires were selected and identified as forest change (Figure 7c). Seventy percent of the samples were randomly chosen as training samples, while the rest were used as validation samples. We used the class-stratified random sampling design that provides the option to increase the sample size in classes that occupy a small proportion of area and is one of the easier designs to implement [68].

Table 5. The classification scheme.

<table>
<thead>
<tr>
<th>Classes</th>
<th>Description</th>
<th>Training Samples</th>
<th>Validation Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>No-change</td>
<td>Areas comprising the same cover in both epochs, albeit with changes caused by vegetation seasonality</td>
<td>350</td>
<td>150</td>
</tr>
<tr>
<td>Forest change</td>
<td>Land cover areas comprising deforestation and fires (savannas that were converted to bare soil and pastures)</td>
<td>350</td>
<td>150</td>
</tr>
</tbody>
</table>

The evaluated performances were compared among the four input features (average, standard deviation, RVF, and AFM) through the confusion matrix measures. We also generated a validation map to identify the savannas land cover changes (deforestation and fires) by visual interpretation to spatialize the areas classified as correct and incorrect. We used as reference data, a land cover map [39]
which described qualitative and quantitative information of savannas vegetation type in the state of Minas Gerais, Brazil.

Figure 7. (a) The location of training and validation samples of No-change and Forest change distributed over the study area; (b) example of an object undergoing a forest fire from August 2015 to October 2015; (c) example of No-change object over the entire time series (July 2003–October 2016).

4. Results

4.1. Semivariogram Analysis

We analyzed the temporal-spatial variability of NDVI data comparing the semivariogram shape and indices. The semivariograms reached the sill ($\sigma^2$) within the calculated distance (600 m), indicating that their spatial extents were sufficiently large to encompass the entire spatial variability. From no-change transitions (landscape vegetation cover that had the same land cover classes in both times, encompassing seasonal variations due to vegetation phenology), NDVI semivariograms had similar behaviors, presenting sill values around 0.0010 (Figure 8a,b). In the presence of changes (that is, fires), NDVI semivariograms considerably increased, reaching sill values around 0.0060 (Figure 8c). Note how the shape and sill parameter of no-change areas remained constant, but changed markedly in areas with human activities (Figure 8d).

These results indicate a very clear trend whereby the overall variability and shape, the semivariogram increased when land cover changes occurred, and were similar when the area had not undergone any land cover change. These results demonstrate that the semivariograms derived from NDVI images, is not affected by phenological changes (remain constant), while also demonstrating the potential of detecting land cover changes (notably increase) due to the mosaic of types of landscape with low NDVI and high NDVI, increases the NDVI spatial variability.
4.2. Change Detection

We used the Change and Trend Spatio-Temporal Metrics derived from geostatistical and spectral features time series as input data to train the random forest (RF) algorithm. The classification results (Figure 9) achieved overall accuracies from 63.7% to 96.5%, with user’s accuracies ranging from 61.5% to 98.2% and producer’s accuracies from 58.9% to 97.7% (Table 6). The AFM index provided the best results, reaching 96.5% of the overall accuracy and a 0.9 kappa index. The standard deviation also produced good results, with a kappa index reaching 0.8. The greatest inter-class confusion occurred using the average of the NDVI values and RVF index, revealing low values of overall, producer and user accuracies and kappa index.

Table 6. The classification results using groups of features tested. PA: Producer’s accuracy in %; UA: User’s accuracy in %; STDV: Standard deviation; RVF: Ratio variance-first lag; AFM: Area first lag-first maximum.

<table>
<thead>
<tr>
<th>Classes</th>
<th>Average</th>
<th>STDV</th>
<th>RVF</th>
<th>AFM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PA</td>
<td>UA</td>
<td>PA</td>
<td>UA</td>
</tr>
<tr>
<td>No-change</td>
<td>58.9</td>
<td>61.5</td>
<td>88.6</td>
<td>86.9</td>
</tr>
<tr>
<td>Forest change</td>
<td>67.9</td>
<td>65.5</td>
<td>88.1</td>
<td>89.7</td>
</tr>
<tr>
<td>Overall</td>
<td>63.7</td>
<td>88.3</td>
<td>74.4</td>
<td>96.5</td>
</tr>
<tr>
<td>Kappa coefficient</td>
<td>0.3</td>
<td>0.8</td>
<td>0.5</td>
<td>0.9</td>
</tr>
</tbody>
</table>

As a result, the feature average and the semivariogram index RVF were not able to accurately capture the differences between forest change and seasonal changes. These features are sensitive to phenological variations, resulting in the false detection of LULC changes, overestimating the areas undergoing fire or deforestation (Figure 10a,c). The standard deviation spectral feature presented good accuracy measures (overall accuracy reaching 88.3%), however, the map generated was not more accurate (Figure 10b). The AFM semivariogram index provided the best result. This index is directly influenced by the sill ($\sigma^2$) semivariogram parameter (overall variability), which is sensitive to...
land cover changes and is not affected by changes caused by vegetation seasonality, providing good classification results (Figure 10d).

Figure 9. The change maps for the 2003–2016 period generated by the features: (a) average; (b) standard deviation; (c) RVF index; (d) AFM index.

Figure 10. The comparative analysis among the validation map and the estimates using (a) average spectral feature; (b) standard deviation spectral feature; (c) RVF semivariogram index and (d) AFM semivariogram index.
5. Discussion

5.1. NDVI Spatial Variability

To capture the NDVI variability inside the objects we used the semivariogram, that is a graphical representation of the spatial variability in a given dataset [69], provided by the sill ($\sigma^2$), range ($\phi$), and nugget effect ($\tau^2$) parameters. The semivariogram presents the ability to depict landscape spatial heterogeneity [70]. The differences of NDVI spatial variability between landscapes are mainly explained by the type of landscape, where the image spatial variability (sill-$\sigma^2$) increases considerably from homogeneous to heterogeneous areas.

The semivariogram shape and metrics generated from NDVI values, derived from Landsat TM images are useful in detecting deforestation in areas covered by savanna vegetation [62]. In deforested areas, the landscape change caused spatial variations that were quantified by the semivariogram metrics of range, sill, and shape. The range and sill were the two most important and complementary metrics. Both metrics increased their values after deforestation and remain similar if the land cover had not been changed.

Here, instead of using the semivariogram parameters to quantify the NDVI spatial variability we used the RVF and AFM semivariogram indices as they synthesize the most relevant information about the shape of the semivariogram. In addition, fitting a model is unnecessary, which shortens the processing time and reduces the errors generated by choosing a sub-optimal model [35]. From the wet to dry season, the NDVI spatial variability is almost constant and in the presence of land cover changes, such as deforestation and fires, the NDVI variability highly increased (Figure 11). The preserved savannas land cover classes (both in wet and dry seasons) have low NDVI variability compared to disturbed savannas land cover classes. These high values are explained by the presence of high and low NDVI values in the same object, as land cover changes exceed seasonal variability.

![Figure 11. The example of NDVI spatial variability of objects under a seasonal change and deforestation land cover change.](image)

The total variance of NDVI values increases considerably from homogeneous to heterogeneous landscapes. In homogeneous objects as preserved savannas, the NDVI semivariance among the pixels is low, thus, the RVF index is also low because the $y_{\text{max}}$ ($0.0013$) is almost similar with the semivariance in the first lag ($0.0005$). The AFM is also small (0.011) (Figure 12a). In heterogeneous

![Figure 12. Example of NDVI spatial variability of objects (a) under a seasonal change and (b) deforestation land cover change.](image)
objects, such as savannas undergoing a deforestation or fires, the RVF and AFM indices considerably increase, ranging from 3.33 to 8.38 and 0.011 to 0.13, respectively (Figure 12b).

![Figure 12. Example of NDVI spatial variability of objects (a) under a seasonal change and (b) deforestation land cover change.](image_url)

Similar patterns were found by Reference [37] exploring and evaluating the performance of semivariogram indices derived from NDVI images in an object-based approach to detect land cover changes. The image’s spatial variability changed considerably from native vegetation (pre-disaster image) to flooded areas (post-disaster image). The flooded areas had a low overall variability compared to native vegetation, making the detection of change using the semivariogram possible.

5.2. Seasonal Effects on Change Detection

We hypothesized that the semivariogram indices (RVF and AFM) are not affected by vegetation seasonality due to the NDVI variability remaining constant in the presence of seasonal changes and substantially increases in the presence of land cover changes, such as deforestation and fires (both natural events and human-induced).

Remote sensing of vegetation has an implied link to vegetation phenology. Phenology is an environmental process that operates independently of land cover or land use change. The ability to use remote sensing images to detect changes in seasonal areas relies on the capacity to remove the effects of seasonal variations, while still identifying changes caused by deforestation, urbanization, floods, and fires. To avoid the classification of seasonal effects as land cover change, time series have been used to derive the seasonal patterns when mapping deforestation. For instance, Verbesselt et.al developed a powerful change detection approach for time series, using 16-day MODIS NDVI images, involving the detection and characterization of Breaks For Additive Seasonal and Trend (BFAST) that integrates the iterative decomposition of time series into trend, seasonal, and noise components with methods for detecting changes, without the need to select a reference period, set a threshold, or define a change trajectory [15].

Rather than using the BFAST seasonal model to account for seasonality in the image time series, Hamunyela et.al researches used the spatial context in a pixel-based approach using Landsat data assessing how the size of spatial window influences the deforestation detection and related computational time [12]. They demonstrated that, in dry tropical forests, deforestation events are detected much earlier when the spatial context approach is used than when a seasonal model is used. Smaller spatial windows achieved higher overall spatial accuracy than larger windows and increasing the size of the spatial window resulted in a linear increase in the computational time per pixel.

The window size is an important aspect of the spatial context method since it influences the resultant spatial information due to the amount of variance included [71,72]. The spatial moving window needs to be sufficiently larger than the size of the deforestation event to avoid smoothing out the disturbance impact from the time series, and also for it to capture local variations [12].

Here, we used the spatial context (semivariogram indices) in an object-based approach, without the need of use of the spatial window, basing the objects on the characteristics found in the first image of the time series. The time series of the AFM semivariogram index provided the best separability (global
accuracy = 96.53%) between objects with seasonal changes and objects undergoing a forest change. Similar results were found by Reference [73], assessing the potential of individual geostatistical features derived from bitemporal NDVI images to accurately detect land cover changes. They demonstrated that the accuracies of the sill parameter and AFM semivariogram indices were higher than those of the other indices, which were not affected by vegetation seasonality because they represent the information provided by the entire semivariogram rather than specific parts of it.

5.3. Method Limitations

Despite the good performance of the proposed method, in this section, two limitations to the approach used in this study are elaborated upon, including the quantification of areas changed and the timing of forest changes:

1. Forest change quantification: The method can identify the objects undergoing land cover changes without being impacted by seasonal variations; however, it is not possible to accurately calculate the changed area. Due to use of an object-based approach, some changes do not cover the entire object because the forest change class comprised both small changes reaching up to 20% of the objects and large changes ranging from 50% to 100% of the object area. This limitation is related to the initial image segmentation performed on the first image of the time series (image object overlay) rather than segmenting the entire time series (multitemporal image object). We used this technique because it provides a more meaningful framework for spatial measures [51], capable of capturing the intra-object heterogeneity [74]. Further research into how the proportion of pixels changed inside the objects (real area of forest change) affect the resultant accuracies remain outstanding.

2. Timing of changes: The method allows for the discrimination between forest change and no-change classes, however, it is not possible to indicate when exactly the forest changes have occurred. Although the timing of the change is an important feature of forest monitoring, we did not attempt to model the time of the change, as our main focus is not hinged on the capacity to date forest changes, rather we focus on the advantages of using variability measures as input data to train the RF algorithm.

6. Conclusions

We have developed a method to mitigate the presence of phenological effects from time series using NDVI trends to detect changes over a savannah forest ecosystem. The approach combining spatial and spectral information and allowed for the detection of changes without being impacted by seasonal variations. We demonstrated that NDVI spatial variability, captured by AFM semivariogram index, is not affected by vegetation seasonality and, therefore, can produce time series that accurately differentiate forest changes from seasonal changes, resulting in fewer classification errors. Our proposed method is simple and accurate, efficiently detecting changes while eliminating the effects of phenology.

The spatial context approach we presented in this paper is a novel and useful object-based remote sensing method for the detection of forest change events using Landsat NDVI data in areas where forests exhibit strong seasonality (that is, in Brazilian seasonal biomes). Our results demonstrate that the semivariograms derived from NDVI images are not affected by phenological changes (remain largely constant). The spectral values may differ over time within a given year, but absent actual change (for example, fire or deforestation), the amount and nature of the variability present at the object level remains similar. This method addresses the challenge of accurately detecting deforestation and fires without erroneously labeling seasonal phenological differences as changes without the need to use de-seasoning models.

Author Contributions: These authors equally contributed to this work.

Funding: This research received no external funding.
Acknowledgments: The authors are grateful to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and the Department of Forest Science of the Federal University of Lavras (UFLA) for supporting this work.

Conflicts of Interest: The authors declare no conflicts of interest.

References


44. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* 2007, 11, 1633–1644. [CrossRef]


66. A Language and Environment for Statistical Computing, version 1.0.136; R Core Team: Vienna, Austria, 2009–2016.


66. Akar, Ö.; Güngör, O. Integrating multiple texture methods and NDVI to the Random Forest classification algorithm to detect tea and hazelnut plantation areas in northeast Turkey. *Int. J. Remote Sens.* 2015, 36, 442–464. [CrossRef]


73. Silveira, E.M.O.; de Mello, J.M.; Acerbi Júnior, F.W.; de Carvalho, L.M.T. Object-based land-cover change detection applied to Brazilian seasonal savannahs using geostatistical features. *Int. J. Remote Sens.* 2018, 39, 2597–2619. [CrossRef]


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).