Remote Sens. **2015**, *7*, 12160-12191; doi:10.3390/rs70912160

OPEN ACCESS

remote sensing

ISSN 2072-4292

www.mdpi.com/journal/remotesensing

Article

Comparative Analysis of MODIS Time-Series Classification Using Support Vector Machines and Methods Based upon Distance and Similarity Measures in the Brazilian Cerrado-Caatinga Boundary

Natanael Antunes Abade [1,2], Osmar Abílio de Carvalho Júnior [2,*], Renato Fontes Guimarães [2] and Sandro Nunes de Oliveira [2,3]

[1] Instituto do Meio Ambiente e dos Recursos Hídricos do Distrito Federal, Gerência de Reserva Legal, SEPN 511, Bloco C, Edifício Bittar, Brasília 70750-543, Brazil; E-Mail: natanael.abade@ibram.df.gov.br

[2] Departamento de Geografia, Campus Universitário Darcy Ribeiro, Universidade de Brasília (UnB), Asa Norte, Brasília 70910-900, Brazil; E-Mails: renatofg@unb.br (R.F.G); sandro.oliveira@ifb.edu.br (S.N.O)

[3] Instituto Federal de Brasília (IFB), Campus Gama, DF 480, Setor de Múltiplas Atividades, Lote 01, Gama 72429-005, Brazil

* Author to whom correspondence should be addressed; E-Mail: osmarjr@unb.br; Tel.: +55-61-3367-5001.

Academic Editors: Clement Atzberger and Prasad S. Thenkabail

Received: 13 March 2015 / Accepted: 25 August 2015 / Published: 18 September 2015

Abstract: We have mapped the primary native and exotic vegetation that occurs in the Cerrado-Caatinga transition zone in Central Brazil using MODIS-NDVI time series (product MOD09Q1) data over a two-year period (2011–2013). Our methodology consists of the following steps: (a) the development of a three-dimensional cube composed of the NDVI-MODIS time series; (b) the removal of noise; (c) the selection of reference temporal curves and classification using similarity and distance measures; and (d) classification using support vector machines (SVMs). We evaluated different temporal classifications using similarity and distance measures of land use and land cover considering several combinations of attributes. Among the classification using distance and similarity measures, the best result employed the Euclidean distance with the NDVI-MODIS data by considering more than one reference temporal curve per class and adopting six mapping classes. In the majority of tests, the SVM classifications yielded better results than other methods. The best

result among all the tested methods was obtained using the SVM classifier with a fourth-degree polynomial kernel; an overall accuracy of 80.75% and a Kappa coefficient of 0.76 were obtained. Our results demonstrate the potential of vegetation studies in semiarid ecosystems using time-series data.

Keywords: Caatinga; Cerrado; spectral angle mapper; spectral correlation mapper; euclidian distance measure; MODIS

1. Introduction

Studies of the spatial distribution and the processes governing the Brazilian Cerrado-Caatinga transition can significantly inform the management of these biomes. Ecotones are key indicators of local and global changes that facilitate an understanding of the landscape's responses to changes in climate regimes and human influences. The Cerrado and Caatinga biomes have great biodiversity adapted to the seasonal stress that coexists in a complex mosaic [1–3]. The transition zone includes different vegetation types such as dry forest islands, gallery forests, and savannas that vary in their patch number and size. However, anthropogenic alterations are increasing in this environment, which is causing landscape fragmentation and harm to ecological functioning.

Semiarid ecosystems have been threatened worldwide, with a reduction in area of approximately 50% on a global scale [4]. Portillo-Quintero and Sanchez-Azofeifa [5] estimate that 66% of dry tropical forests in Latin America have been deforested, mainly due to agricultural expansion; only fragmented landscapes remain. In Brazil, human activities have altered about 48% of the Caatinga [6,7] and 53% of the Cerrado [8]. These ecosystems have received less attention and consequent research than tropical rainforests. As a result, questions still remain about their extent, limit, biogeographical relationships, and land use processes, for instance [9,10].

Time-series satellite imagery has great potential for vegetation monitoring in semi-arid regions. In this context, the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor instrument aboard the Terra and Aqua platforms [11] has been widely used for terrestrial monitoring due to its high temporal resolution and sensitivity to radiometric and geometric properties. Moreover, several phenological studies have been conducted using the vegetation indices of MODIS images in both natural ecosystems [12–14] and in anthropic landscapes such as agricultural expansion [15–17], crop phenology [18,19], pastures [20], and eucalyptus plantations [21].

In semi-arid environments, MODIS data are being applied to studies of seasonality, phenology, fire events, and conservation [22]. Hüttich *et al.* [23] distinguish savanna types in South Africa from the *in situ* botanical survey combined with Landsat and MODIS images. Baldi *et al.* [24] quantify the connections among landscape patterns, ecosystem functioning, and agricultural use in the Dry Chaco and Chiquitania ecoregions that cover parts of Argentina, Bolivia, and Paraguay using MODIS-Normalized Difference Vegetation Index (NDVI) time-series data. Portillo-Quintero *et al.* [25] study the correlation between MODIS active fires and forest cover change in the tropical dry forest.

Here, we aim to map the land use and land cover in the Brazilian Cerrado-Caatinga transition using phenological signatures from MODIS vegetation indices over a two-year period (2011–2013). However,

we used different aspects to identify how temporal features affect classification performance. In this context, we evaluated and compared different models for the support vector machines (SVMs) and temporal-signature classifications. We tested different compositions for temporal-signature classification in order to identify the best-fit method: (a) the number of temporal curves by class (single or multiple); (b) the number of classes used in the classification (six or eight); (c) the input data (NDVI-MODIS time series or signal components of the Minimum Noise Fraction (MNF) transform); and (d) three measurement types. The measures that we used to cluster similar phenological patterns included the spectral angle mapper (SAM), the spectral correlation mapper (SCM), and the Euclidian distance (ED). In addition, we compared the four most widely used kernel functions for SVM classification: linear, polynomial, radial basis function (RBF), and sigmoid.

2. Study Area

The study area is located in the southern Brazilian semi-arid region between 14°33'50" and 16°19'55" south latitude and 42°24'32" and 44°08'20" west longitude. The study area is 36,485 km^2, 89.4% of which is located in the Minas Gerais state, and the remainder is contained in the Bahia state (Figure 1).

The climate is tropical semi-arid with two well-defined seasons: a dry season (May through October) and a rainy reason (November through April). The average annual rainfall is approximately 818 ± 242 mm [26], and roughly 50% of this precipitation is concentrated in the months of November, December, and January [27] (Table 1). The average temperature ranges from 21–25 °C, and the coldest months are June and July [26,27]. The geomorphology of the study area includes the São Francisco depression, highlands, and the Espinhaço ranges; the primary soil types are Entisols, Oxisols, Inceptisols, and Gleysols [28].

The native vegetation is composed of six predominant phytophysiognomies, of which two derive from the Caatinga domain (deciduous seasonal forest and semi-deciduous seasonal forest), and four derive from the Cerrado domain (Savanna Woodland and the following grassland formations: Savanna Grassland, Shrub Savanna, and Rupestrian Fields) [29]. The Cerrado predominates in the eastern part of the area in the Espinhaço Range, and the Caatinga occurs in the center and west of the area at elevations below 800 m between the São Francisco River and the Espinhaço Range. We also observed vegetation transition zones in the Parque Estadual Caminho dos Gerais, a protected area where Cerrado physiognomies [30] (mainly grassland) surrounded by Caatinga prevail (Figure 1).

In the Espinhaço Range, the Cerrado occurs in acidic and dystrophic soils from quartzite rocks [31]. However, the vegetation is diverse, with the Savanna Woodland (Cerrado *stricto sensu*) found on more developed and deeper soils (e.g., Oxisol) and grassland formations (Savanna Fields) found on soils associated with rocky outcroppings [31,32].

The Caatinga region near the Espinhaço Range and the São Francisco River encompasses particular combinations of soil type, terrain, and hillsides that act as barriers to regular rainfall [33,34]. The Caatinga physiognomies include species such as *Cavanillesia umbellata* (Barriguda), *Myracrodruon urundeuva*, *Tabebuia impetiginosa*, and *Aspidsperma pyrifolim* occurring in soil erosion resulting from limestone or crystalline rocks of the Bambuí group [28].

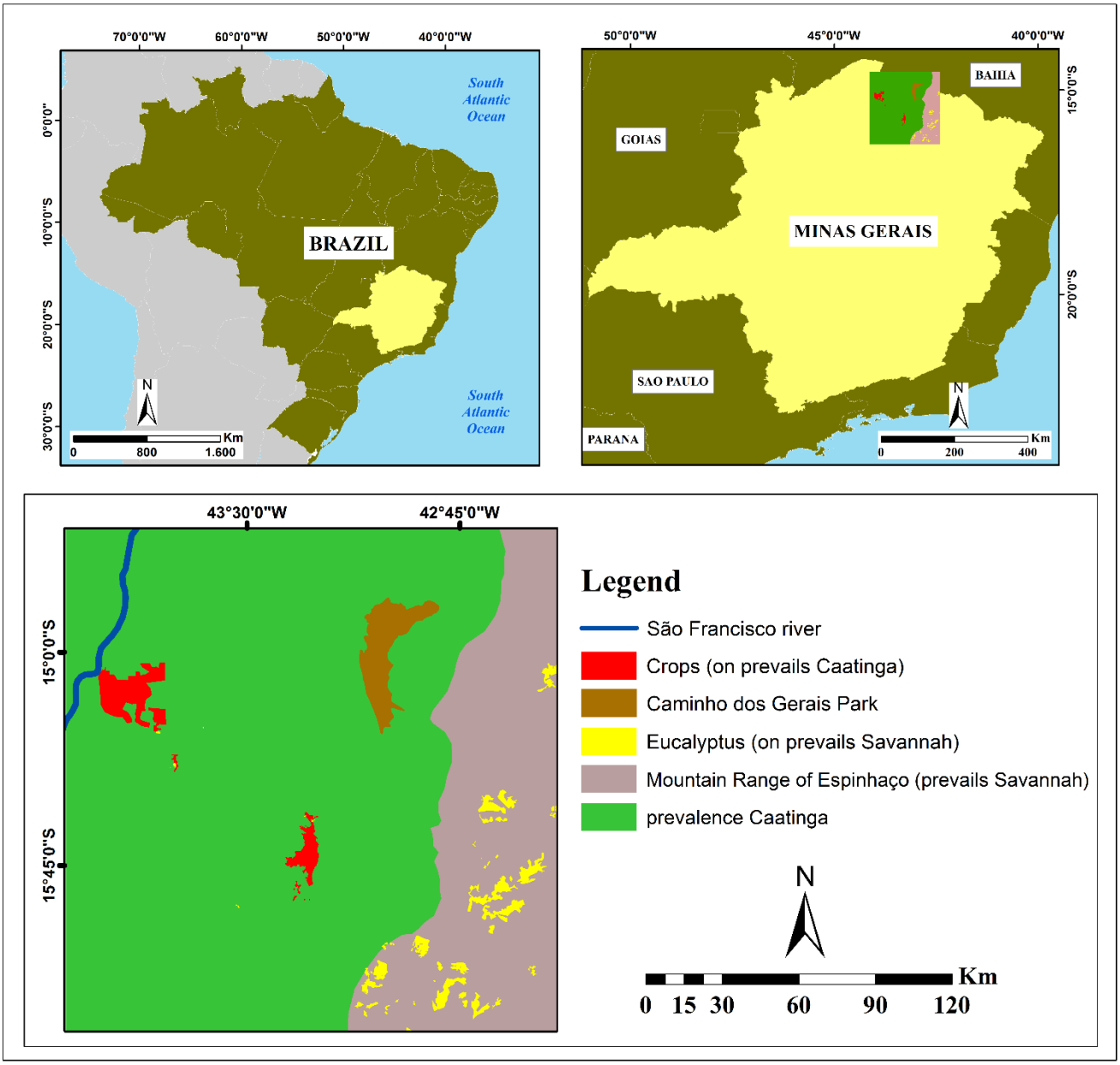

Figure 1. Location map of the study area, with the spatial distribution of Caatinga and Cerrado domains indicated.

Favorable soil conditions, mainly characterized by eutrophic soils, and water from the São Francisco River and its tributaries have enabled technical agriculture in the central and northwest regions of the study area. This agricultural area includes annual crops (corn, cane sugar, sorghum, beans, and cotton) and perennial crops (bananas, lemons, mangoes, and coffee) [35].

Extensive cattle and agricultural plantations, including a large irrigated area (the Jaíba Project), have suppressed the native vegetation [36]. In the eastern part of the study area, the Cerrado predominates due to the presence of low-fertility soil and deficient water. These environmental conditions restrict the development of agriculture on the site that specializes in planting species in the genus Eucalyptus [33]. In addition, scientific studies that could contribute to vegetation monitoring have been scarce, which has hampered conservation [37].

Table 1. Meteorological data of the study area between 1977 and 2007 [27].

Climatic Data	Month											
	Jan.	**Feb.**	**Mar.**	**Apr.**	**May**	**June**	**July**	**Aug.**	**Sept.**	**Oct.**	**Nov.**	**Dec.**
Rainfall (mm)	175.6	96	93.3	33.1	8,0	5.4	15.1	31.4	31.2	54.2	117.8	150.7
Average Temperature (°C)	25.6	26	26.1	25.4	24.5	23.1	22.6	23.8	25.2	26.3	25.8	25.7
Average Maximum Temperature (°C)	30.8	31.5	31.5	31	30.4	29.1	28.8	30.3	31.6	32.5	31.2	30.8
Average Minimum Temperature (°C)	20.4	20.5	20.6	19.7	18.5	17	16.4	17.3	18.8	20.1	20.4	20.6

3. Methodology

The image processing included the following steps (Figure 2): (a) the acquisition of MODIS images; (b) the development of a three-dimensional cube composed of NDVI-MODIS time series; (c) the removal of noise; (d) the selection of temporal signatures and classification of time series using distance and similarity measures (ED, SAM, and SCM); and (e) classification using SVMs.

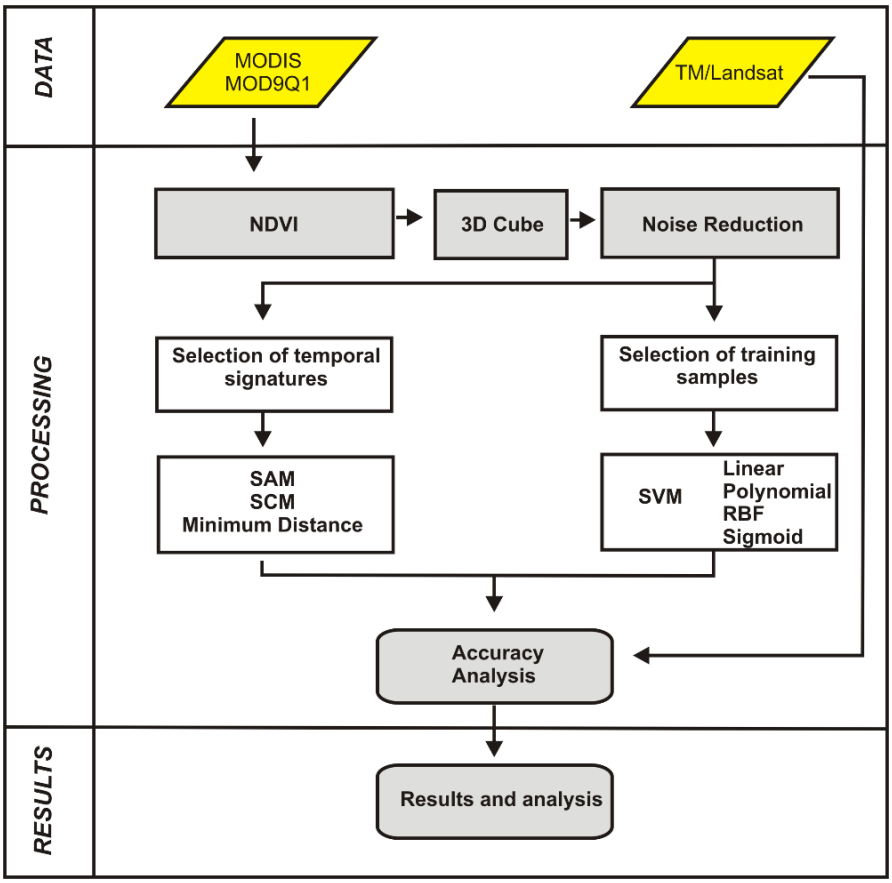

Figure 2. Methodological flowchart of the digital image processing.

3.1. MODIS/Terra Time-Series Dataset

We used the 8-day MODIS/Terra MOD09Q1 product relative to the surface reflectance data. This product includes bands 1 and 2 corresponding to red (620–670 nm) and near-infrared (841–875 nm) wavelengths, respectively. The MOD09Q1 product includes all 8-day inputs at 250-meter resolution [38]. The MOD09Q1 product does not provide direct data vegetation indices like the MOD13Q1 and MOD13A1 products [11,38,39]. However, the MOD13 products are only available as 16-day composite images [40], which is not appropriate for describing the phenologies of the savanna. The images were converted to geographical coordinates, datum World Geodetic System 84 (WGS84), using the MODIS Reprojection Tolls program.

The NDVI-MODIS time series data can be used to identify the savanna types since the evaluation period is not fewer than two years [23]. In addition, the best results consider the dry-season images that are largely free of clouds [23]. Thus, the long-term phenological observations, including the inter-seasonal variability, are useful for a detailed characterization of semi-arid vegetation [23,41]. For this study, the MODIS images were acquired over a period from September 2011 until October 2013, totaling 96 scenes. The initial and final date coincides with the peak of the dry season when the deciduous vegetation loses its leaves [26,41].

3.2. Image Cube of NDVI-MODIS Time Series

NDVI was calculated from the MOD09Q1 bands according to the following equation [42]:

$$\mathrm{NDVI} = \frac{\rho^{*}\mathrm{NIR} - \rho^{*}\mathrm{Red}}{\rho^{*}\,\mathrm{NIR} + \rho^{*}\,\mathrm{Red}} \quad (1)$$

where "ρ*NIR" (band 2) and "ρ*Red" (band 1) are the reflectance values for the near infrared and red ranges, respectively. All MODIS-NDVI images obtained over 2011–2013 were merged in a three-dimensional image cube; "X" and "Y" are related to geographical coordinates (longitude and latitude), and "Z" is the behavior of the target over time [43,44].

3.3. Image Denoising

Noise is very common in time series of NDVI images from the cloud cover, shade, and instrumental defects that affect the identification and quantification of ground targets. Thus, noise elimination is necessary to acquire a high-quality seasonal curve. In this work, we combined two filtering techniques: median filter and Savitzky-Golay (S-G) [45].

The median filter is a nonlinear smoothing technique that preserves signal edges or monotonic changes in trend and particularly removes short-duration impulse noise, which is not possible using linear algorithms [46]. The median filter employs a window moving over the temporal curve and obtains the median value, a particular case of the order statistic (or rank statistic) of a finite set of real numbers, that is taken as the output. However, the temporal median filter is effective only at low noise densities; its effectiveness decreases, and image details diminish given high noise-density interference in successive time series of images [47].

The S-G combines the effective noise removal and the waveform-peak preservation (height, shape, and asymmetry) [48] and is adequate for phenological analyses of remote-sensing data [49–51].

Geng *et al.* [52] compare eight techniques to filter NDVI time-series data and conclude that the S-G filter yields the best results in the majority of situations. However, there is a lack of consensus about the most appropriate technique for smoothing time-series data [53,54]. Each method for smoothing time-series satellite sensor data has its own advantages and disadvantages in accordance with the purpose of the study [55].

3.4. Classification Using Distance and Similarity Measures

The types of measures used to assess spectral distance and similarity are important topics in spectral classification theory. Spectral classifiers compare image spectra to a reference spectrum from spectral libraries or to spectral endmembers in order to capture the dependencies among variables [56]. Spectral classifiers are widely used in hyperspectral image classification in which the difference between the methods is based on the measure adopted. In this study, we apply these classifiers to the temporal curves. Thus, the application of this method depends on two steps: the selection of reference signatures and the calculation of a similarity or distance measure between the reference and image curves. The advantage of this method is that the establishment of a temporal-curve library facilitates the classification of other images or the same image at other times.

3.4.1. Reference Temporal-Signature Selection

Reference-signature selection for the spectral classification employs a different strategy than other classifiers using training samples such as maximum likelihood or SVM. In this approach, only one signature represents the desired class. Therefore, a few key spectra (endmembers) can be used to explain the rest of the dataset [57]. Spectral signature extraction typically uses the sample mean in order to obtain an improved signal-to-noise ratio over that of single spectra. However, this sample has high similarity and low standard deviation unlike other classification procedures, wherein the representation of each class requires an extensive sampling in order to obtain the variations present on the target.

An important feature of the temporal images is that a single class can possess different temporal curves for various reasons such as variations in cultivation and harvesting time, a heterogeneous spatial distribution of rainfall, and fire events, among others. However, the spectral classification considers only a reference signature for each class. An alternative is to use more than one temporal signature in order to independently map the different variations of a land use class and then merge them into a single class. We accordingly propose a refinement to identify subsamples within the initial sample that best represent the class variation; this methodology improves the classification results. In this context, convex geometry-based methods are widely used for isolating reference pixels (training points) from initial data [58–60]. This technique aims to detect endmembers that can be applied to the unmixing processing and spectral classification. In this study, the convex geometry was adapted to refine the initial samples of land-use/land-cover classes. This approach can be organized according to the following steps: (a) initial sample selection containing approximately 10,000 pixels for each class (with the exception of the water class that has a poor representation in the study area); (b) the use of the MNF transform [61] separately for each initial sample; (c) the identification in the initial sample of three subsamples with approximately 200 pixels using the n-dimensional viewer from MNF space; and (d) extracting average signatures from the subsamples considering the temporal data and the MNF signal components (Figure 3).

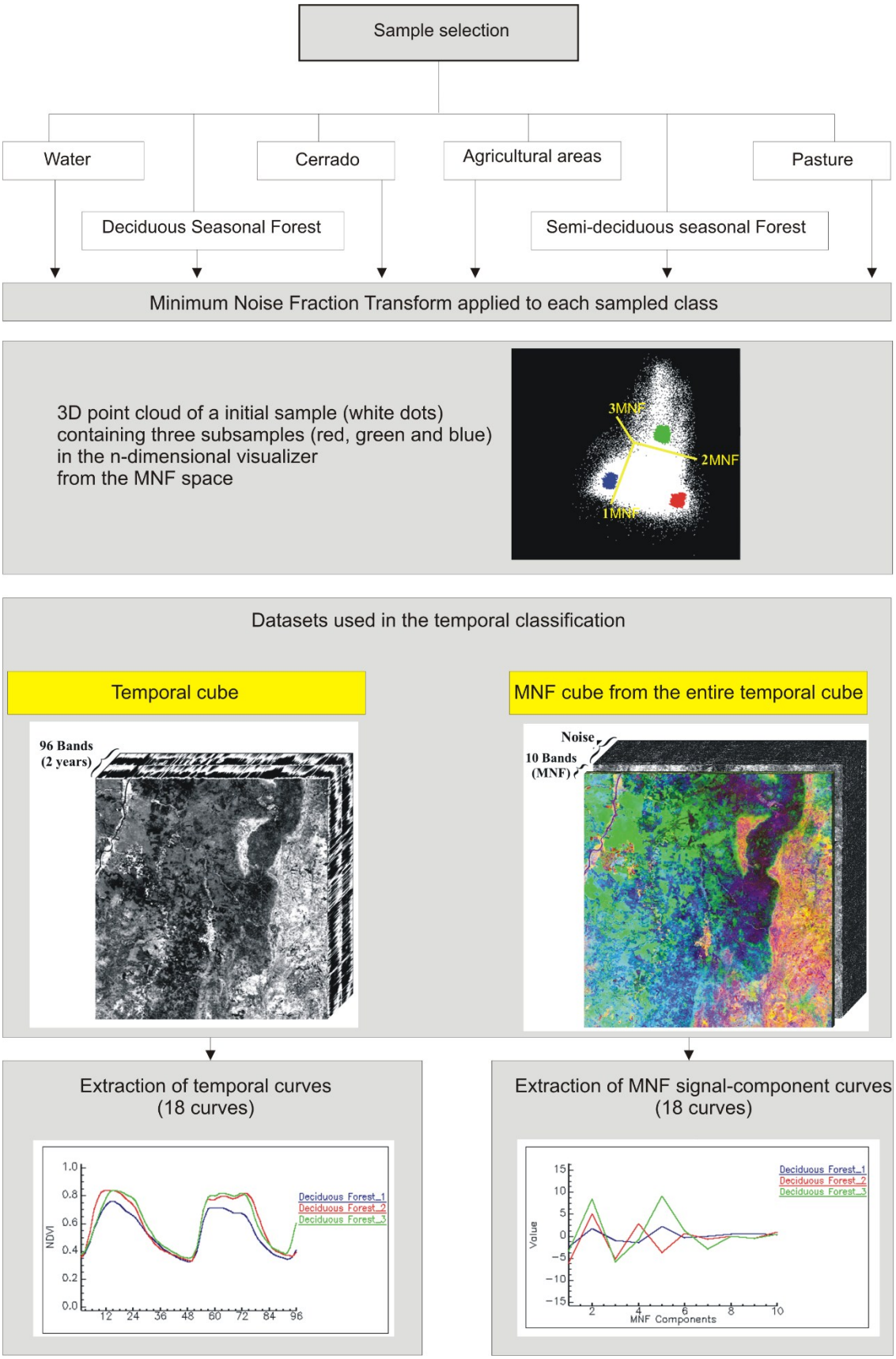

Figure 3. Procedure for the temporal-signature selection of land use and vegetal cover considering both the MODIS-NDVI data dimension and the actual dimensions of images from signal components of the MNF transform.

The initial sample selection considered the following information: (a) a land-cover map of Minas Gerais state produced by the State Forestry Institute and the University of Lavras from 2005–2007 [29] and updated for 2009 [62]; (b) municipal agricultural production information in 2012 from the Brazilian Institute of Geography and Statistics [35]; (c) high-resolution images from Google Earth from 2013; and (d) fieldwork conducted in 2013 to identify vegetation types. The analysis adopted the following classes: water, agricultural areas, pasture, deciduous forest, semi-deciduous forest, and savanna. In addition, a more detailed investigation was conducted in order to separate agricultural areas in annual and perennial crops as well as the savannas in woody and herbaceous formations (Table 2).

Table 2. Classes of land use and land cover used in the classification of MODIS-NDVI time-series data.

Sets	Specifications	Total Classes	Classes
1	Regional. Annual and perennial crops are not separated; vegetation subtypes of the Cerrado biome are not separated.	6	Water; Agricultural Areas [1], Pasture [1], Deciduous Seasonal Forest [2], Semi-deciduous Seasonal Forest [2], and Cerrado[3].
2	Detailed. Annual and perennial crops are separated; woody and herbaceous formations of the Cerrado biome are separated.	8	Water, Annual Crops [1], Perennial Crops [1], Pasture [1], Deciduous Seasonal Forest [2], Semi-deciduous Seasonal Forest [2], Savanna Woodland (Cerrado *stricto sensu*) [3], and Grassland formations [3].

[1] Anthropic use; [2] Forms of occupation related to the Caatinga biome; [3] Forms of occupation related to the Cerrado biome.

The water class necessitates specific sampling. Water is present along the São Francisco River, which has small reservoirs and an average width of approximately 600 meters. The detection of these features is strongly influenced by the spatial resolution of the MODIS images since most of the pixels are spectral mixtures with other targets. Therefore, reference samples at the class edges may include a predominance of other targets and generate an overestimation classification where the water class overlaps with areas of other targets. However, a sampling strategy that considers only the restricted areas in the polygon interior, *i.e.*, the pixels within a homogenous block, can lead to an underestimation of the target area. When the water class is restricted to the study area, we avoid collecting edge sites for the water class.

The MNF transform is a method for both removing noise and compressing the image data into a small group of signal components [61]. The MNF transform is a procedure similar to principal component analysis, and it consists of a linear transformation that maximizes the signal-to-noise ratio to rank order the images, *i.e.*, according to image quality. This procedure is sufficient for reducing data redundancy from hyperspectral images [61], aerial gamma-ray survey data [63], radar datasets [64], and a time series of remote-sensing data [47]. Thus, the MNF is an efficient way to identify a subspace with reduced dimensionality and enable an appropriate selection of reference data. In MNF space, the input data are divided into two parts: one part is associated with the signal components, and the other part is associated

with the noise components. If we consider only the signal components, the data size decreases drastically. Identification of signal components considers the evaluation of image quality and the eigenvalues plot. Low eigenvalues contain little information and concentrate the noise-dominated components. In contrast, large eigenvalues are associated with spatially coherent images and concentrate the signal-dominated components. In this paper, we use the MNF transformation in two different ways: (a) to identify the best temporal signatures in each sample, and (b) over the entire image in order to obtain a different type of input data for classification (*i.e.*, in addition to the temporal data we also used the MNF signal components as input data in the classification).

After we completed the MNF transformation for each initial sample, its points were distributed in the n-dimensional visualizer of the environment for the visualization of images (ENVI) considering the first three MNF components to be the axes of the three-dimensional visualization [58,60]. Normally, the point cloud of the initial sample has a simple geometrical configuration (simplex) such as the triangle or tetrahedron where the vertices are the endmembers [58,60]. In the present case, the point cloud of the sampled data exhibits a triangular shape in which the three vertices are adopted as subsamples for classification. Additional information about endmember detection using the convex geometry is available in the literature [65–69]. The classification procedure used the average temporal curves of the subsamples considering both MODIS-NDVI time-series data and the MNF signal component from the entire image (Figure 3). The image classification that employed the MNF signal components was intended to not only reduce noise but also yield a greater distinction among classes. In this algorithm, the three temporal curves for the same class are classified independently and generate three different subclasses that are then merged into the same class after the map is obtained. This simple methodology allows us to consider different temporal curves for the same class, which yields a significant improvement in the time-series classification.

3.4.2. Distance and Similarity Measures

The similarity measures most frequently used in spectral classification include the cosine correlation adopted in the SAM [70] and the Pearson's correlation coefficient adopted in the SCM [71]; the primary distance measure is the ED used in the minimum distance. The spectral measures yield different information about the target, which justifies conducting testing in order to determine the most appropriate procedure for a particular data type or target [72,73]. In this work, we test the similarity and distance measures for time-series image classification.

The SAM method calculates the angle (expressed in radians: 0–π) formed between the temporal signature present in the pixel "X" and the reference "Y", which can be described by the following equation [70]:

$$\mathrm{SAM} = \alpha = \mathrm{COS}^{-1} \frac{\sum_{i=0}^{N} X_i Y_i}{\sqrt{\sum_{i=0}^{N} (X_i)^2 \sum_{i=0}^{N} (Y)^2}} \quad (2)$$

The variable "I" corresponds to the temporal band, ranging from one to the number of temporal bands "N". The higher the similarity between the temporal curves, the smaller the angle value. The SAM is unable to detect anti-correlated data and is invariant to bias factors.

The SCM is a modification of the SAM method and is invariant to the bias and gain factors; it can be used to detect anti-correlated data. SCM values range from −1 to 1; 1 implies a maximum correlation, 0 implies completely uncorrelated values, and −1 implies that the values are perfect opposites. The major difference between the correlation methods is that SAM uses only values for X and Y; SCM uses data centered by the means "$\overline{X}$" and "$\overline{Y}$" [71]:

$$\mathrm{SCM} = \frac{\sum_{i=1}^{i=N}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{i=N}(X_i - \overline{X})^2 \sum_{i=1}^{i=N}(Y_i - \overline{Y})^2}} \quad (3)$$

The ED is a measure that is sensitive to the bias (additive) factor and gain (multiplicative) factor. The shorter the distance, the greater the likelihood that the temporal signature belongs to the reference group. The ED can be described by the equation:

$$\mathrm{ED} = \sqrt{\sum_{i=1}^{i=N}(X_i - Y_i)^2} \quad (4)$$

3.5. Support Vector Machine

SVMs [74] are a group of methods for supervised non-parametric statistical learning that constitute a powerful technique for general classification (pattern recognition) and regression (function approximation) problems with a perceptive model representation [75,76]. SVMs aim to generate a training data model that yields the target class only based on its attributes. Training vectors are mapped into a higher-dimensional space by a kernel function that describes a hyperplane, which consists of an optimal separation of the dataset into discrete classes according to the training samples. The four most widely used kernel functions are linear, polynomial, RBF, and sigmoid [77,78]. Table 3 lists the mathematical equations of these kernel functions. The choice of kernel function and its parameters are crucial factors for obtaining good results. The SVM learning process establishes what is known from the minimization of structural risk, thereby reducing the classification error on unseen data without prior assumptions of the probability distribution of the data [79]. On the other hand, statistical classification methods assume the data distribution a priori.

Table 3. Mathematical equations of the most widely used kernel functions ("γ" is the gamma term, "d" is the polynomial degree term, and "r" is the bias term).

Linear Kernel	Polynomial Kernel	RBF Kernel	Sigmoid Kernel
$K(x_i, x_j) = x_i^T x_j$	$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$	$K(x_i, x_j) = exp\left(-\gamma \lVert x_i - x_j \rVert^2\right), \gamma > 0$	$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r), \gamma > 0$

In the remote sensing field, Mountrakis *et al.* [79] provide a review of the use of SVM in image processing for remotely sensed data. These authors identify a wide range of application domains and sensors. Carrão *et al.* [80] investigate an arrangement of spectral bands and dates from MODIS data to

land cover classification using the SVM method. Vuolo and Atzberger [81] apply the least squares SVM algorithm to NDVI-MODIS time series for land-cover mapping and achieved good results.

Here, we tested the SVM method considering the four major kernel functions (linear, polynomial, RBF, and sigmoid). We employed the ENVI program based on the LIBSVM library [82]. Our strategy for collecting training samples differed from strategies adopted for the spectral classifiers (SAM, SCM, and minimum distance). The representation of each class requires extensive sampling in order to obtain the variations in the temporal curves. Therefore, a sampling must contain a group of pixels (e.g., constituting a field plot) belonging to a land-cover/land-use class that contains the diversity of temporal curves from several areas of the image. We collected 200 training samples for each class. Our sampling assumed a uniform spatial distribution for each class and considered the same information from the images and maps described for the spectral classification. In the specific case of the water class, the training samples were derived from the samples used in the spectral classifications. These water samples were located preferably in the inner class to avoid pixels with a high percentage of spectral mixture. By adopting these samples, we avoid the reported problems of an overestimated classification for the water class.

3.6. Accuracy Analysis

We used the overall accuracy and the Kappa coefficient for the accuracy assessment of land-use or land-cover classifications [83]. Our calculations of these accuracy indices for the different methodological approaches assumed 1600 ground-truth points (validation sites) collected from a systematic, stratified sampling design. Stratified sampling by mapped land-cover classes guaranteed a specified sample size (200 pixels) for each mapped class. This sampling approach required a previous map from the visual classification of Landsat Thematic Mapper (TM) data on 12 September 2011 (Figure 4a), which was updated for 2013 using high-resolution data from Google Earth and fieldwork observations (Figure 4b). The sampling points were extracted according to land cover classes and their geographical locations, which produced a spatially well-distributed sample. The validation dataset is distinct from the data used in the selection of the reference temporal signatures. The areas of land-use change were not considered in the accuracy analysis. The overall accuracy is the ratio of the number of pixels correctly classified by the total number of pixels, and the Kappa index is a measure described by the following equation:

$$K = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} x_{i+}x_{+i}}{N^2 - \sum_{i=1}^{r} x_{i+}x_{+i}} \quad (5)$$

where "r" is the number of rows in the error matrix, "x_{ii}" is the number of observations in row "i" and column "I", "x_{i+}" and "x_{+I}" are the marginal totals in row "i" and column "I", respectively, and "N" is the total number of observations.

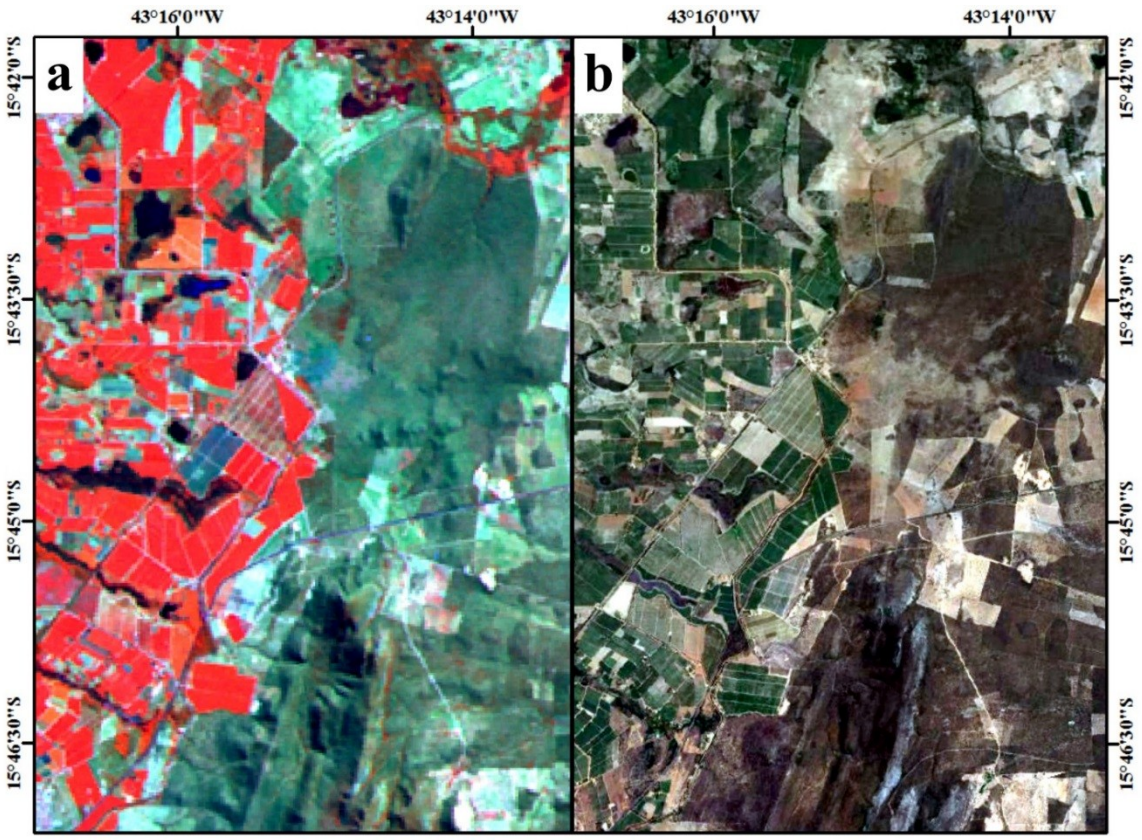

Figure 4. (**a**) Landsat-5 TM image from 12 September 2011, RGB color composite of TM bands 4 (red), 5 (green), and 3 (blue) and (**b**) high-resolution images from Google Earth from 2013.

4. Results

4.1. Noise Reduction

The combination of the median filter and S-G exhibited good results for the noise elimination of the MODIS-NDVI time series. The median filter minimized the impulse noise from, for example, cloud cover or shade. However, the median filter can generate some linear levels due to the repetition of values. We found that S-G with a window size of nine [84] refined the result obtained by the median filter and smoothed the temporal profile without interfering with the maximum and minimum values. This methodology accordingly ensured data integrity. Figure 5 shows the results from the application of the filtering techniques in the time series.

4.2. Classification using Distance and Similarity Measures

Reference Temporal Signatures

The temporal signatures of vegetation have the lowest NDVI values during the dry season (May through October) and the highest values in the rainy season (November through April). The formations with more dense canopies in the Caatinga domain (deciduous seasonal forest and semi-deciduous seasonal forest) have NDVI values of 0.80 ± 0.03 during the rainy season. The lower

NDVI values of the deciduous forest (0.33 ± 0.02) are due to loss of leaves. This vegetation has a quick recovery with the rainy season, which is more significant than the shrub-grasslands [85] (Figure 6a). The lower NDVI values for the semi-deciduous forests are above 0.44 (Figure 6b). This vegetation type is photosynthetically active throughout the year, even during periods of lower rainfall. The savanna woodland (Cerrado *sensu stricto*) has a maximum NDVI value of 0.70 (Figure 6c), slightly lower than the arboreal formations of the Caatinga. The savanna fields (grassland formations) exhibit mixed grasses and shrubs, which creates difficulties in the subtype identification [41]. Therefore, the distinction of grassland types (savanna grassland, shrub savanna, and rupestrian fields) was not possible based on temporal curves, and these different grassland types are considered to be a single class. The herbaceous scrub vegetation exhibited NDVI values ranging from 0.27–0.58 (Figure 6d); the woody-herbaceous vegetation exhibited NDVI values ranging from 0.35–0.70 (Figure 6c).

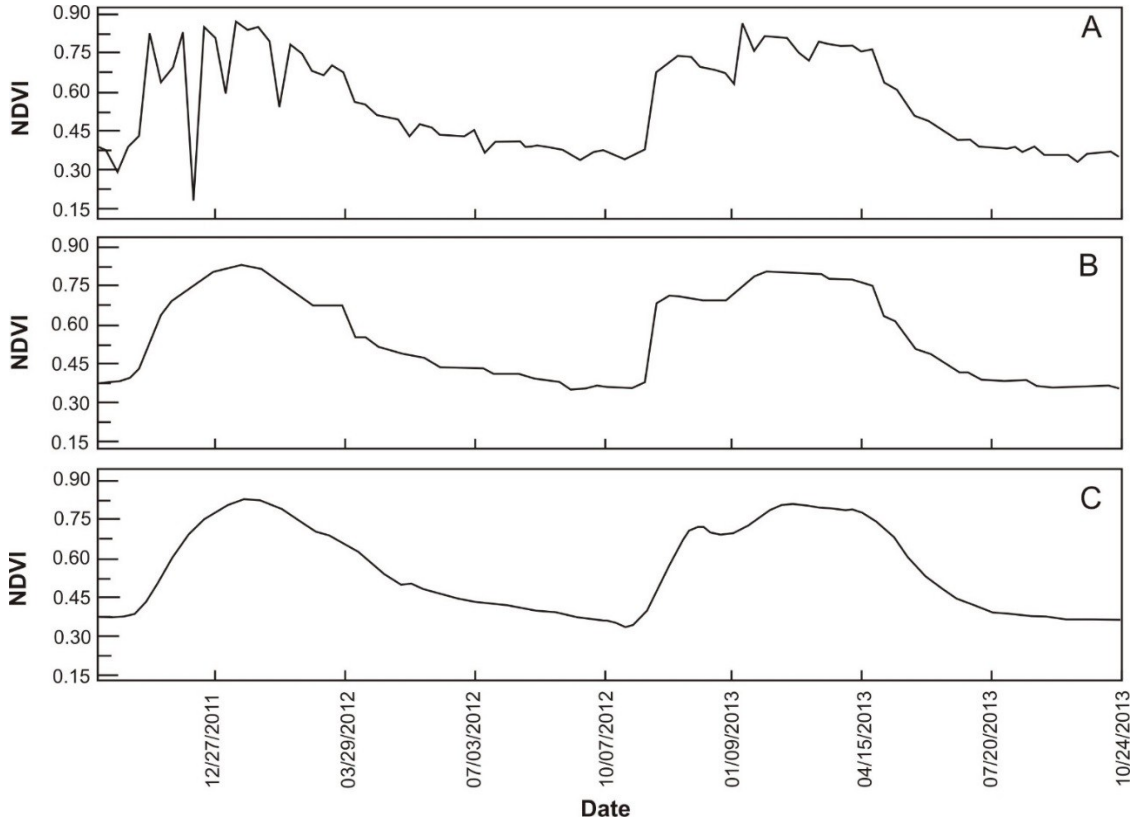

Figure 5. MODIS-NDVI time series, (**A**) original data with noise, (**B**) time profile softened by the median filter, and (**C**) refined by the S-G.

Grassland formations are significantly affected by water deficiency and accordingly exhibited the lowest NDVI values among all of the native vegetation that we studied. However, the phenological behavior between native grasslands and pasture, predominantly of the genus *Brachiaria* [33], can be differentiated. The pasture responds more quickly to the beginning of the rainy season and attains maximum photosynthetic activity (an NDVI value of 0.71) over approximately 8 days (Figure 6e). In the native grasslands, maximum photosynthetic activity occurs after approximately 24 days of the start of the rainy season. Furthermore, pastures exhibit higher NDVI than native grasslands for all of the curves that we analyzed.

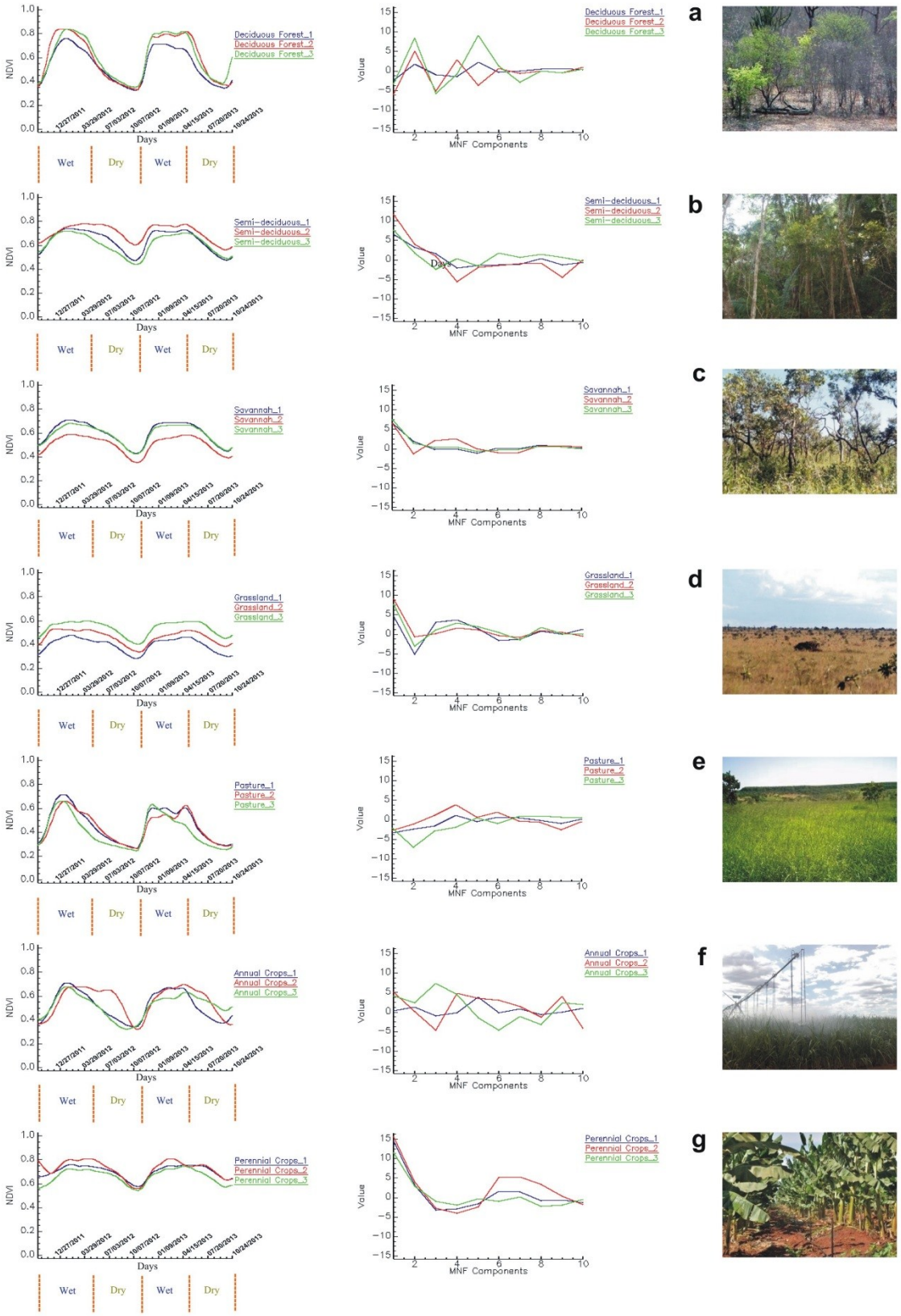

Figure 6. Temporal signatures from the MODIS-NDVI time series and the MNF signal components: (**a**) deciduous seasonal forest, (**b**) semi-deciduous seasonal forest, (**c**) savanna woodland (Cerrado *stricto sensu*), (**d**) grassland formations, (**e**) pasture, (**f**) annual crops, and (**g**) perennial crops.

Differentiating among temporal signatures for all types of short-cycle crops was not possible. Distinct and simultaneous short-cycle crops may be found in the same locations in different configurations. However, annual-cycle crops are individualized even when various species are grown simultaneously. These cultures maintained a maximum NDVI level of 0.70 for a period of 180 days and exhibited similar minimum values (0.31 ± 0.01) (Figure 6f). This finding indicates a cyclical regularity of agricultural practices, nutritional conditions, and water availability, which does not occur as rigorously in nature. Perennial crops are deciduous with maximum NDVI values similar to those of the semi-deciduous forest (0.80 ± 0.03) (Figure 6g). The distinction between these types of vegetation is clearest in their minimum NDVI values: the forest exhibits a range of 0.44–0.60, and the perennial crop has more restricted range of 0.54–0.57. Crop performance compared with native vegetation is motivated by agricultural techniques that standardize the soil fertility, the number of species per unit area, and the availability of water.

4.3. Classification of the MODIS-NDVI Time Series and MNF Signal Components

We evaluated different classifications of land use and land cover considering several combinations of procedures. Table 4 lists the accuracy indices for the different classifications obtained by arrangements of the following attributes: (a) the number of reference curves by class (single or multiple), (b) the number of land use/land cover classes (six or eight), (c) input data (NDVI-MODIS time series or MNF signal components), and (d) classification measures (SAM, SCM, or ED).

The use of multiple temporal curves in the classification yielded better accuracy than the use of a single temporal curve in all cases (Table 4). Figure 7 shows the results from one or three reference curves per class compared with the NDVI-MODIS images, MNF signal components, and Landsat-TM. Furthermore, the increase in the number of land use/land cover classes (from six to eight) results in a decrease in accuracy when we take into account all classification measures and input data (Table 4).

The best result was the combination that employed multiple curves, used ED on the MODIS-NDVI time series, and considered six classes of land use/land cover. We obtained an overall accuracy of 79.06% and a Kappa coefficient of 0.74, which represents a very good classification. The same combination using the MNF signal components also achieved a very good classification with an overall accuracy of 75.18% and a Kappa coefficient of 0.69.

The SCM method from the eight classes and a single signature yielded the poorest result for both the NDVI-MODIS data (overall accuracy of 45.81% and a Kappa coefficient of 0.38) and MNF signal components (overall accuracy of 57.62% and a Kappa coefficient of 0.51). SAM attained slightly better values considering the same configuration for the NDVI-MODIS data (overall accuracy of 50.43% and a Kappa coefficient of 0.43) and the MNF signal components (overall accuracy of 57.87% and a Kappa coefficient of 0.51). This poorer performance for the similarity measures was expected due to the interference of the seasonal behavior of vegetation, which yields similar behavior for the targets of vegetation after canceling the effects of gain and offset [86]. Despite the classes of semi-deciduous forest, savanna woodland, and savanna fields having distinct NDVI values, their temporal curves have similar shapes, which makes it difficult to distinguish these classes using similarity measures (Figures 6b–d).

Table 4. Accuracy assessment for the classifications using two types of input data (NDVI-MODIS time series and MNF signal components), three classification measures (SAM, SCM and ED), number of curves by class (single or multiple), and the two classifications with six or eight classes.

One Temporal Signature per Class					
Method	**Classes**	**Overall Accuracy**		**Kappa Coefficient**	
		NDVI-MODIS	**MNF Signal**	**NDVI-MODIS**	**MNF Signal**
Euclidian Distance Measure	6	73.50	72.43	0.67	0.66
	8	67.25	64.75	0.62	0.59
Spectral Angle Mapper	6	57.68	65.93	0.47	0.58
	8	50.43	57.87	0.43	0.51
Spectral Correlation Mapper	6	52.75	65.31	0.42	0.57
	8	45.81	57.62	0.38	0.51
Three Temporal Signatures per Class					
Method	**Classes**	**Overall Accuracy**		**Kappa Coefficient**	
		NDVI-MODIS	**Signal MNF**	**NDVI-MODIS**	**Signal MNF**
Euclidian Distance Measure	6	79.06	75.18	0.74	0.69
	8	70.56	68.37	0.66	0.63
Spectral Angle Mapper	6	60.81	68.43	0.51	0.61
	8	54.68	61.37	0.48	0.55
Spectral Correlation Mapper	6	55.25	66.18	0.45	0.58
	8	45.18	60.06	0.37	0.54

Figure 8 illustrates the classified images using the distance and similarity measures (SAM, SCM, and ED) and the two data types (NDVI-MODIS and MNF signal components). These maps consider the eight classes described in Table 2. All of the classified images detected the dominance of the Cerrado in the eastern part of the studied area and the Caatinga in the western part. Agricultural activities also had a regional division; annual cycle crops were found in the northwest, and perennial crops were found in the south central and eastern part of the study area largely in the classifications of the ED algorithm.

Table 5 lists the confusion matrices for the ED classifications with eight classes considering the MODIS-NDVI data and the MNF signal components. The Cerrado classes (savanna woodland and grasslands) demonstrated the highest omission and commission error rates in both ED classifications. These vegetation types exhibited a strong overlap with one other. The continuous variations between shrub and herbaceous vegetation generate a mixed behavior that hinders the detection of Cerrado subtypes. Merging the Cerrado classes (savanna woodland and grassland formations) yielded a significant improvement in classification performance.

The deciduous formation exhibited the best accuracy among all of the native vegetation types that we evaluated (Table 4). Unlike the Cerrado, the deciduous forest presents temporal signatures with specific shapes and only few variations (Figure 6a), which enables a relative improvement in accuracy regardless of the algorithm that we adopt. Classification errors occurred near fragment edges where spectral mixing was present. However, the semi-deciduous forest did not achieve the accuracy of the deciduous forest, causing confusion with the savanna types and permanent crops (Table 4); the forest was largely located

in the southeastern part of the study area. The pasture class was characterized by lower omission and commission error rates.

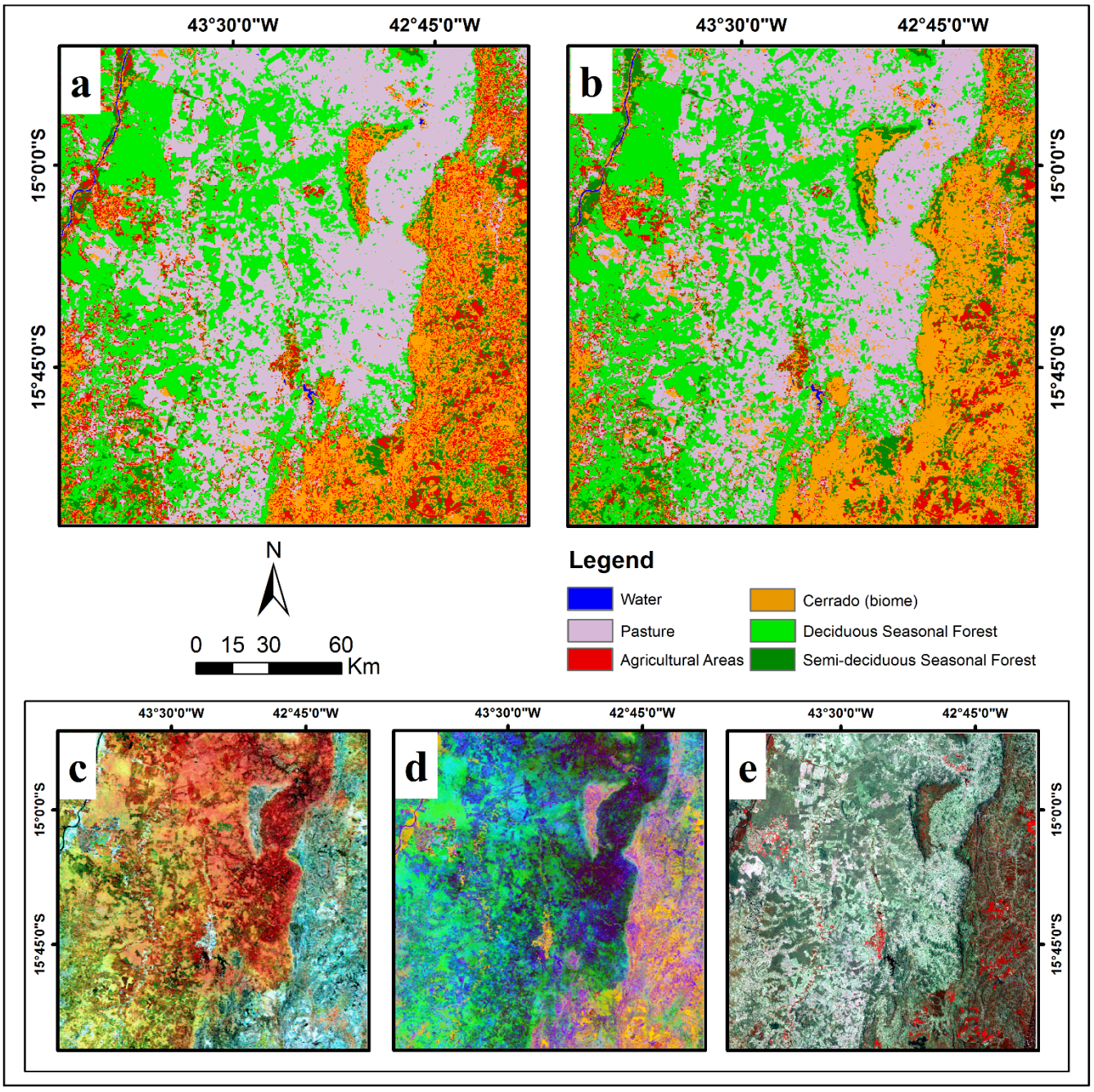

Figure 7. Classified maps built from (**a**) one or (**b**) three temporal curves and holding the remaining factors constant: ED, six classes (land use and land cover), and NDVI-MODIS input data. (**c**) RGB color composite of NDVI-MODIS images (12/27/2011–04/30/2012–09/05/2012). (**d**) MNF (components 1-2-4 as RGB), and (**e**) Landsat-TM (bands 4-5-3 as RGB).

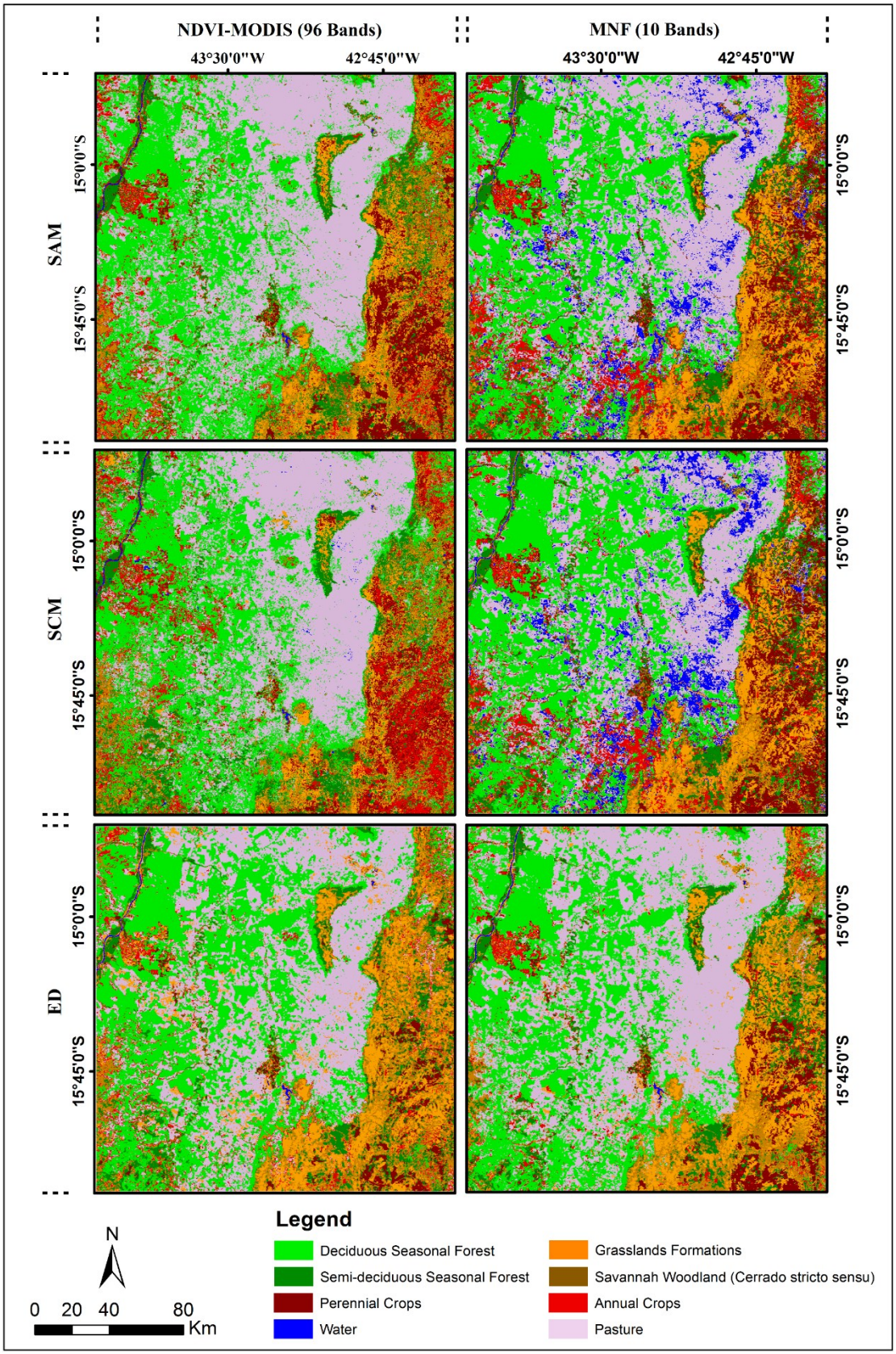

Figure 8. Classified images considering two types of input data (NDVI-MODIS time series and MNF signal components), three types of similarity and distance measures (SAM, SCM, and ED), and the more specific set of use classes (eight classes).

Table 5. Confusion matrices for the classifications based on the ED and eight mapping classes from the MODIS-NDVI data and MNF-signal components.

MODIS-NDVI Data

Class	Water	Annual Crops	Grassland Formations	Savanna Woodland	Deciduous	Pasture	Perennial Crops	Semi-Deciduous	Total
Water	177	0	0	0	0	0	0	0	177
Annual Crops	1	110	11	9	3	7	2	1	144
Grassland Formations	17	23	90	64	1	11	2	2	210
Savanna Woodland	0	23	69	74	1	0	11	14	192
Deciduous	0	26	8	10	181	3	0	1	229
Pasture	5	13	6	1	12	179	0	0	216
Perennial Crops	0	1	5	8	0	0	158	22	194
Semi-deciduous	0	4	11	34	2	0	27	160	238
Total	200	200	200	200	200	200	200	200	1600

MNF-Signal Components

Class	Water	Annual Crops	Grassland Formations	Savanna Woodland	Deciduous	Pasture	Perennial Crops	Semi-Deciduous	Total
Water	171	0	0	0	0	0	0	0	171
Annual Crops	1	112	10	10	2	0	0	4	139
Grassland Formations	17	11	105	60	1	4	0	1	199
Savanna Woodland	2	12	40	60	1	1	6	29	151
Deciduous	0	20	5	11	155	8	0	4	203
Pasture	9	25	8	5	41	187	0	0	275
Perennial Crops	0	9	11	15	0	0	165	23	223
Semi-deciduous	0	11	21	39	0	0	29	139	239
Total	200	200	200	200	200	200	200	200	1600

The distinction between annual and perennial crops occurred only in monoculture areas over large areas in which prevailing crops were irrigated via a center pivot. We found that it was difficult to separate crop types on small farms with intercropping or small-scale monocultures. In addition, the dynamic and multiple uses of land resulted in confusion in the image classification. The confusion matrix regarding the ED reveals some mistakes for the water class (Table 5). Water bodies are contained in rivers and small reservoirs, which are susceptible to interference by spectral mixture; this interference hampers their detection and causes detection errors.

The accuracies of the SAM and SCM methods improve significantly with the use of MNF signal components and begin to approach those of the ED results [44,86]. However, the SAM and SCM classification of the MNF signal fraction overestimated the area of water bodies relative to pastures. Although the MNF signal-component curves for these two classes have different absolute values, the resulting shapes are similar (Figure 9). Therefore, the gain and offset cancellation through the similarity measure results in a normalization, which prevents the two classes from being perfectly disentangled. This example highlights the difficulties of working in MNF space, which can generate similar curves that are not present in the original data. Despite the mistakes in the water bodies, the SAM and SCM

classifications using the MNF signal components as input data yield a higher accuracy for the other classes, which justifies their higher values of a Kappa coefficient and overall accuracy. The ED classification accuracy using the MNF signal fraction produced a worse result.

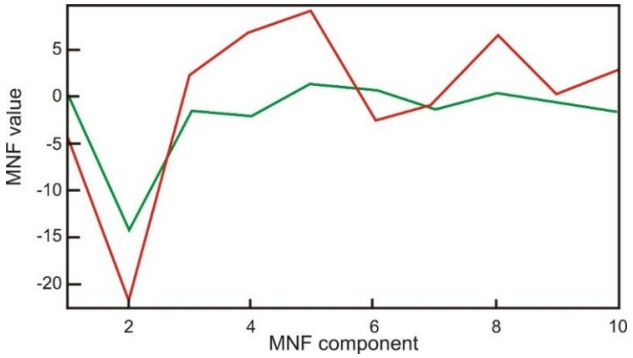

Figure 9. MNF signal-component curve for the water (green line) and pasture (red line) classes, which are characterized by different values and similar shapes.

4.4. Classification Using SVMs

We evaluated different SVM classifications considering main kernel functions (linear, polynomial, RBF, and sigmoid), two input data types (NDVI-MODIS time series and MNF signal components), and the number of classes (six and eight). The parameters that we used included a gamma term of 0.1 for all kernel types except the linear kernel and a bias term of 1 for the polynomial and sigmoid kernels. Table 6 lists the accuracy indices for the different SVM classifications. In the majority of tests, the SVM classifications yielded better results than the classifications using distance and similarity measures. In terms of the spectral classifiers, SVM showed a smaller difference in accuracy values between the classifications with six and eight classes. The results using NDVI-MODIS time series or MNF signal components also exhibited similar values. The only exceptions were the linear and sigmoid kernels, in which the use of the MNF resulted in a significant improvement in the classification.

Table 6. Accuracy assessment for SVM classifications considering two types of input data (NDVI-MODIS time series and MNF signal components), main kernel function (linear, polynomial, RBF, and sigmoid), and six and eight classes.

Method	Classes	Overall Accuracy		Kappa Coefficient	
		NDVI-MODIS	MNF Signal Components	NDVI-MODIS	MNF Signal Components
Linear	6	78.81	79.00	0.739	0.742
	8	72.43	73.00	0.685	0.691
Polynomial 2	6	80.18	78.93	0.756	0.741
	8	73.87	72.93	0.701	0.690
Polynomial 3	6	80.12	79.37	0.755	0.746
	8	73.87	73.18	0.701	0.693
Polynomial 4	6	80.75	79.81	0.763	0.752
	8	74.37	73.75	0.707	0.700

Table 6. *Cont.*

Method	Classes	Overall Accuracy		Kappa Coefficient	
		NDVI-MODIS	MNF Signal Components	NDVI-MODIS	MNF Signal Components
Polynomial 5	6	80.62	80.06	0.762	0.755
	8	74.43	74.12	0.707	0.704
Polynomial 6	6	80.12	80.56	0.755	0.761
	8	74.12	74.56	0.704	0.709
RBF	6	80.43	80.06	0.759	0.755
	8	74.25	73.93	0.705	0.702
Sigmoid	6	65.68	78.18	0.574	0.732
	8	58.12	72.37	0.521	0.684

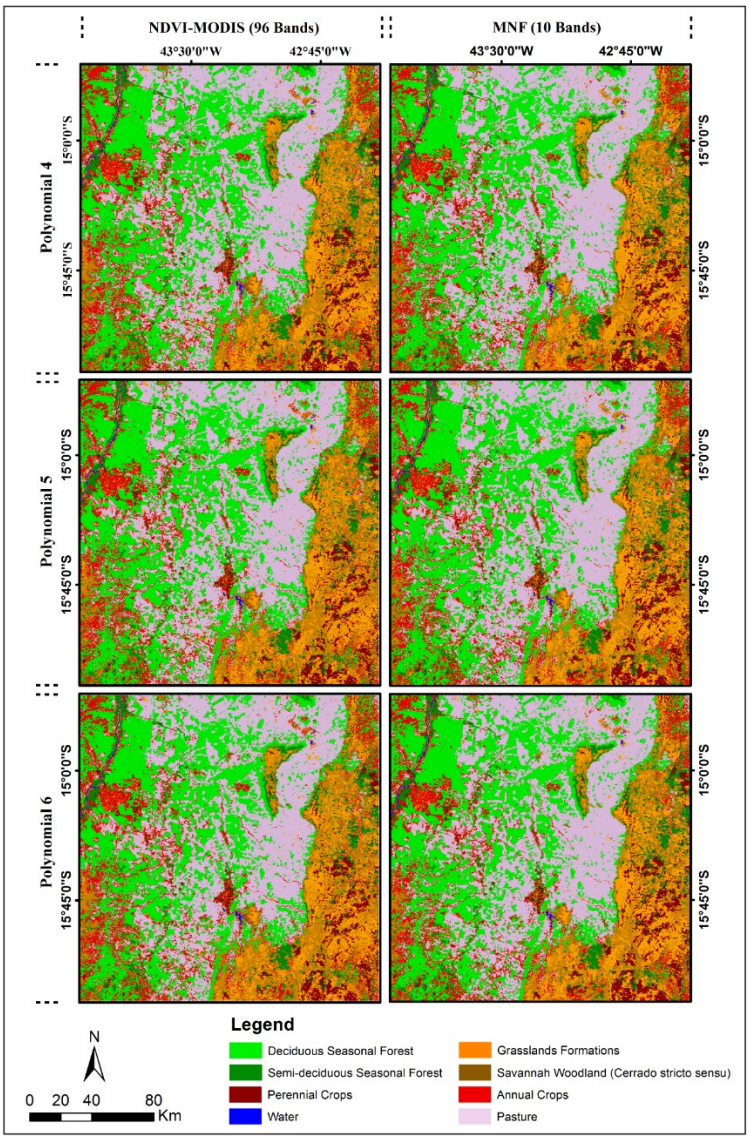

Figure 10. SVM classification maps considering two types of input data (NDVI-MODIS time series and MNF signal components) and three types of polynomial kernel functions (fourth, fifth, and sixth degrees).

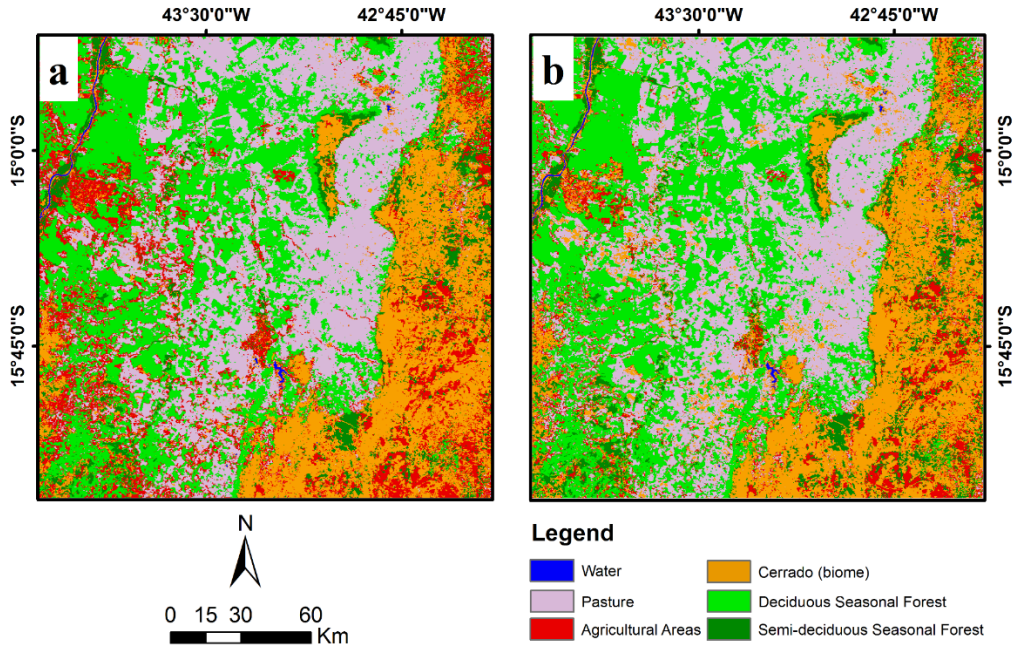

Figure 11. Comparison between the classifications considering six classes: (**a**) SVM (fourth-degree polynomial kernel) and (**b**) ED.

The best result was the SVM classifier using a fourth-degree polynomial kernel in the NDVI-MODIS data with six classes, which achieved an overall accuracy of 80.75% and a Kappa coefficient of 0.765 (Figure 10). The same kernel function using the MNF signal components time series also achieved a very good classification with an overall accuracy of 79.81% and a Kappa coefficient of 0.752. The remaining polynomial and RBF kernels also exhibited very similar accuracy values.

Figure 11 compares the two classification using SVM (a fourth-degree polynomial kernel) and spectral measurements (ED), which exhibit high visual similarity.

5. Discussion

A wide range of densities of woody and herbaceous layers characterizes the vegetation types in the Brazilian Cerrado-Caatinga boundary. The vegetation changes within the Cerrado and Caatinga biomes and along their borders. As a result, an accurate classification using coarse-resolution satellite imagery is challenging. Spectrally separating the vegetation from a single date is difficult due to large seasonal variabilities. A significant improvement is achieved using temporal information that describes different phenological behaviors. We have used temporal trajectories to characterize and classify different vegetation and land-use areas.

We made numerous comparisons in order to identify the optimal image processing for land-cover and land-use mapping in our study region. A characteristic of the phenological classification should consider the different responses to the same type of vegetation caused by environmental variations. In semiarid regions, natural climatic variability results in isolated events such as rain, wind, and fire over distances [87–89]. This variability in physical events is responsible for different behaviors within the same vegetation type, such as regeneration, vegetative growth, flowering, fruiting, and seed dispersal.

In the spectral classifiers, the classification efficiency increases when more than a single temporal signature is applied to describe each land-use/land-cover class. The number of temporal signatures per class should correspond to the number of different temporal patterns described by the analyzed classes. In this approach, the endmember-detection method, which is widely used in hyperspectral imaging, can be adapted for time-series data. Among the classification measures, the ED exhibited the best performance. Different forms of vegetation present seasonal, cyclical behavior that varies largely in magnitude (gain and offset factors). The similarity metrics operate exactly in the curve shapes (nullification of magnitude), resulting in smaller differences in the temporal curves and consequently reducing the effectiveness of the classification. These inherent limitations of SCM and SAM make these techniques less suitable than ED. The MNF transform results in significant data reduction, which yields only approximately 10% of the original data. The use of MNF signal components instead of NDVI-MODIS data did not improve the classification accuracy of the ED. However, these input data caused a significant change when used with the similarity metrics (SCM and SAM) because they removed the redundancies between the curves and the cyclical behavior. In this case, the elimination of the gain and offset factors did not produce similar curves, thereby improving the efficiency of similarity metrics. Classification using distance and similarity measures makes it possible to establish a temporal signature library that can be used in other geographic areas and at other time periods.

The classifications using SVM yielded the most accurate results. A comparison between the different kernel functions demonstrated that the fourth-order polynomial function exhibited the best result. The accuracy values of the different functions and input data (NDVI-MODIS data and MNF signal components) were very similar. The lowest accuracy values were recovered using the sigmoid function. Therefore, this method preserves efficiency while considering different temporal behaviors for the same target.

The separation between annual and perennial crops has limitations, especially in areas with family farming. This agricultural occupation includes multiple activities and high dynamics in smaller territorial extensions, which affects detection quality using MODIS data [24]. Another constraint is to differentiate the savanna types because they consist of a nearly continuous grass layer overlaid by a discontinuous tree layer. The boundaries between the savanna physiognomies are typically gradual and fragmented on the landscape, which impedes proper detection from the 250-meter-resolution images due to their high spectral mixture. Other authors [90,91], even combining MODIS time-series and climate data (precipitation, humidity, soil moisture, light intensity, and day length), have already noted this difficulty. The application of high spatial and temporal resolution images such as the RapidEye can be a possible solution to offer accurate detection [92]. The water class, which possesses a narrow lateral extension, presents spectral mixture interference that affects the detection of bodies in this class.

Future studies should test other classifiers such as the *K*-nearest neighbor [93] and Random Forest [94] algorithm in order to achieve greater accuracy. An additional test that can be performed involves evaluating the time-series interval for phenological classification. Alcântara *et al.* [95] recommend the inclusion of at least two years of MODIS-NDVI data to conduct abandoned agriculture mapping. Huttich *et al.* [23,96] suggest improving the accuracy of the native vegetation classification in savanna environments by using vegetation-index time series spanning more than two years with a temporal resolution of fewer than 16 days. Depending on the application, the analysis of time series can

also require shorter time-series intervals (e.g., seasonal, monthly, or daily) as seen in studies of burned areas [97–99] or change detection [100,101].

6. Conclusions

We have defined a method for native and exotic vegetation mapping in the transition region of semi-arid ecosystems (Cerrado-Caatinga) by comparing different ways of conducting time-series processing. Spatial variations from the climatic factors in semi-arid environments cause a gap between the phenological curves of the same vegetation. Therefore, slightly different temporal signatures can represent the same target. The different factors that can compose the classification highlight the importance of studying dynamics in dry savanna ecosystems. Kappa coefficient statistics for classification sets range between 0.38 and 0.76. Our comparative study of classification methods demonstrated that there is a relationship between the accuracy and complexity of the number of classes. Savannah woodland and grassland formations are difficult to separate, and their unification into a single class significantly improves the classification accuracy.

In the spectral classifier, improved accuracies were achieved by integrating the following parameters: ED, using more than one reference curve per class, and the adoption of six mapping classes (water, farming, pasture, deciduous forest, semi-deciduous seasonal forest, and Cerrado). This configuration yielded an overall accuracy of 79% and a Kappa coefficient of 0.74. The similarity measures yielded poorer performance for the NDVI vegetation time series because of the strong similarity of the curve shapes that followed the climate seasonal variation.

SVMs are suitable for classifying time series in which a target can be characterized by different curves. Our comparison of the different kernel functions yielded very similar results. The best result was obtained using the fourth-order polynomial kernel.

The limitation of the spatial resolution of the MODIS sensor hinders the identification of arboreal and herbaceous physiognomies and activities on small family farms. Future work using other time intervals and images with a periodicity of fewer than 8 days may yield better results. Furthermore, other classifiers such as *K*-nearest neighbor and Random Forest should be tested for time series classification in the Brazilian Cerrado-Caatinga boundary.

Acknowledgments

The authors are grateful for support from the following institutions: the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for sponsoring research grants and the Instituto do Meio Ambiente e dos Recursos Hídricos do Distrito Federal-Brasília Ambiental (IBRAM-DF) for logistical support.

Author Contributions

Natanael Antunes Abade and Osmar Abílio de Carvalho Júnior were responsible for the research design, digital image processing, data preparation and analysis. Renato Fontes Guimarães and Sandro Nunes de Oliveira provided some of the data, conducted the fieldwork and provided relevant technical support. All of the authors contributed to editing and reviewing the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Werneck, F.P. The diversification of Eastern South American open vegetation biomes: Historical biogeography and perspectives. *Quat. Sci. Rev*. **2011**, *30*, 1630–1648.
2. Pinheiro, M.H.O.; MONTEIRO, R. Contribution to the discussions on the origin of the Cerrado biome: Brazilian savanna. *Braz. J. Biol*. **2010**, *70*, 95–102.
3. Pennigton, R.T.; Lavin, M.; Oliveira-Filho, A. Woody plant diversity, evolution, and ecology in the tropics: Perspectives from seasonally dry tropical forests. *Annu. Rev. Evol. Sust*. **2009**, *40*, 437–457.
4. Hoekstra, J.M.; Boucher, T.M.; Ricketts, T.H.; Roberts, C. Confronting a biome crisis: Global disparities of habitat loss and protection. *Ecol. Lett*. **2005**, *8*, 23–29.
5. Portillo-Quintero, C.A.; Sánchez-Azofeifa, G.A. Extent and conservation of tropical dry forests in the Americas. *Biol. Cons*. **2010**, *143*, 144–155.
6. Boori, M.S.; Amaro, V.E. Land use change detection for environmental management: Using multi-temporal, satellite data in the Apodi Vallery of northeastern Brazil. *Appl. GIS* **2010**, *6*, 1–15.
7. Ministério do Meio Ambiente (MMA), Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA). Monitoramento do Bioma Caatinga 2008–2009. In *Monitoramento dos Desmatamentos nos Biomas Brasileiros por Satélite*; MMA: Brasília, Brazil, 2011; pp. 1–46.
8. Sano, E.E.; Rosa, R.; Brito, J.L.S.; Ferreira, L.G. Land cover mapping of the tropical savanna region in Brazil. *Environ. Monit. Assess.* **2010**, *166*, 113–124.
9. Furley, P.A.; Metcalfe, S.E. Dynamic changes in savanna and seasonally dry vegetation through time. *Prog. Phys. Geog*. **2007**, *31*, 633–642.
10. Werneck, F.P.; Costa, G.C.; Colli, G.R.; Prado, D.E.; Sites, J.W., Jr. Revisiting the historical distribution of seasonally dry tropical forests: New insights based on palaeodistribution modelling and palynological evidence. *Global Ecol. Biogeogr*. **2011**, *20*, 272–288.
11. Justice, C.O.; Townshend, J.R.G.; Vermote, E.F.; Masuoka, E.; Wolfe, R.E.; Saleous, N.; Roy, D.P.; Morisette, J.T. An overview of MODIS land data processing and product status. *Remote Sens. Environ*. **2002**, *83*, 3–15.
12. Hammer, D.; Kraft, R.; Wheeler, D. Alerts of forest disturbance from MODIS imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *33*, 1–9.
13. Van Leeuwen, W.J.; Davison, J.E.; Casady, G.M.; Marsh, S.E. Phenological characterization of desert sky Island vegetation communities with remotely sensed and climate time series data. *Remote Sens.* **2010**, *2*, 388–415.
14. Zhao, X.; Xu, P.; Zhou, T.; Li, Q.; Wu, D. Distribution and variation of forests in China from 2001 to 2011: A study based on remotely sensed data. *Remote Sens. Environ*. **2013**, *4*, 632–649.

15. Bernardes, T.; Moreira, M.A.; Adami, M.; Giarolla, A.; Rudorff, B.F.T. Monitoring biennial bearing effect on coffee yield using MODIS remote sensing imagery. *Remote Sens*. **2012**, *4*, 2492–2509.
16. Couto Júnior, A.F.; de Carvalho Júnior, O.A.; Martins, E.S.; Vasconcelos, V. Characterization of the agriculture occupation in the Cerrado biome using MODIS time-series. *Rev. Bras. Geofis*. **2013**, *31*, 393–402.
17. Galford, G.; Mustard, J.F.; Melillo, J.; Gendrin, A.; Cerri, C.C.; Cerri, C.E.P. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sens. Environ*. **2008**, *112*, 576–587.
18. Pan, Y.; Li, L.; Zhang, J.; Liang, S.; Zhu, X.; Sulla-Menashe, D. Winter wheat area estimation from MODIS-EVI time series data using the crop proportion phenology index. *Remote Sens. Environ*. **2012**, *119*, 232–242.
19. Sakamoto, T.; Yokozawa, M.; Toritani, H.; Shibayama, M.; Ishitsuka, N.; Ohno, H. A crop phenology detection method using time-series MODIS data. *Remote Sens. Environ*. **2005**, *96*, 366–374.
20. Ferreira, L.G.; Fernandez, L.E.; Sano, E.E.; Field, C.; Sousa, S.B.; Arantes, A.E.; Araújo, F.M. Biophysical properties of cultivated pastures in the Brazilian savanna biome: An analysis in the spatial-temporal domains based on ground and satellite data. *Remote Sens*. **2013**, *5*, 307–326.
21. Le Maire, G.; Marsden, C.; Nouvellon, Y.; Stape, J.L.; Ponzoni, F.J. Calibration of a species-specific spectral vegetation index for leaf area index (LAI) monitoring: Example with MODIS reflectance time-series on eucalyptus plantations. *Remote Sens*. **2012**, *4*, 3766–3780.
22. Sánchez-Azofeifa, G.A.; Quesada, M.; Rodríguez, J.P.; Nassar, J.M.; Stoner, K.E.; Castillo, A.; Garvin, T.; Zent, E.L.; Calvo-Alvarado, J.C.; Kalacska, M.E.R.; *et al.* Research priorities for neotropical dry forests. *Biotropica* **2005**, *37*, 477–485.
23. Hüttich, C.; Gessner, U.; Herold, M.; Strohbach, B.J.; Schimidt, M.; Keil, M.; Dech, S. On the suitability of MODIS time series metrics to map vegetation types in dry savana ecosystems: A case study in the Kalahari of NE Namibia. *Remote Sens*. **2009**, *1*, 620–643.
24. Baldi, G.; Houspanossian, J.; Murray, F.; Rosales, A.A.; Rueda, C.V.; Jobbágy, E.G. Cultivating the dry forests of South America: Diversity of land users and imprints on ecosystem functioning. *J. Arid. Environ*. **2014**, doi: 10. 1016/j.jaridenv.2014.05.027.
25. Portillo-Quintero, C.A.; Sánchez-Azofeifa, G.A.; Espírito-Santo, M.M. Monitoring deforestation with MODIS active fires in neotropical dry forests: An analysis of local-scale assessments in Mexico, Brazil and Bolivia. *J. Arid Environ*. **2013**, *97*, 150–159.
26. Madeira, B.G.; Espírito-Santo, M.M.; D'Ângelo Neto, S.; Nunes, Y.R.F.; Sánchez-Azofeifa, G.A.; Fernandes, G.W.; Quesada, M. Changes in tree and liana communities along a sucessional gradiente in a tropical dry forest in South-Eastern Brazil. *Plant. Ecol*. **2009**, *201*, 291–304.
27. Instituto Nacional de Meteorologia (INMET). Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP). Available online: http://www.inmet.gov.br (accessed on 21 April 2015).
28. Saadi, A.; Magalhães Júnior, A.P. A geomorfologia do Planalto do Espinhaço setentrional avaliada para a implantação de barragem: A UHE de Irapé-MG. *Geonomos* **1997**, *5*, 9–14.
29. De Carvalho, L.M.T.; Scolforo, J.R. *Inventário Florestal de Minas Gerais: Monitoramento da Flora Nativa 2005–2007*, 1st ed.; Lavras: Universidade Federal de Lavras, Brazil, 2008; p. 357.

30. Dutra, V.F.; Garcia, F.C.P. Three New Species of Mimosa (Leguminosae) from Minas Gerais, Brazil. *Syst. Botany* **2013**, *38*, 398–405.
31. Alves, R.J.V., Kolbek, J. Can campo rupestre vegetation be floristically delimited based on vascular plant genera? *Plant Ecol*. **2010**, *207*, 67–79.
32. Echternacht, L.; Trovó, M.; Oliveira, C.T.; Pirani, J.R. Areas of endemismo in the Espinhaço Range in Minas Gerais, Brazil. *Flora* **2011**, *206*, 782–791.
33. Espírito-Santo, M.M.; Sevilha, A.C.; Anaya, F.C.; Barbosa, R.; Fernandes, G.W.; Sánchez-Azofeifa, G.A.; Scariot, A.; de Noronha, S.E.; Sampaio, C.A. Sustainability of tropical dry forests: Two case studies in southeastern and central Brazil. *Forest Ecol. Manag*. **2009**, *258*, 922–930.
34. Zappi, D. Fitofisionomia da Caatinga associada à Cadeia do Espinhaço. *Megadiversidade* **2008**, *4*, 34–38.
35. Instituto Brasileiro de Geografia e Estatística (IBGE). Sistema IBGE de Recuperação Eletrônica (SIDRA). Available online: http://www.sidra.ibge.gov.br (accessed on 29 August 2014).
36. Domingues, S.A.; Karez, C.S.; Biondini, I.V.F.; Andrade, M.A.; Fernandes, G.W. Economic environmental management tolls in the Serra do Espinhaço biosphere reserve. *J. Sust. Dev*. **2012**, *5*, 180–191.
37. Dos Santos, R.M.; Vieira, F.A.; Fagundes, M.; Nunes, Y.R.F.; Gusmão, F. Floristic richness and similarity of eight forest remnants in the north of Minas Gerais state, Brazil. *Rev. Árvore* **2007**, *31*, 135–144.
38. Vermote, E.F.; Kotchenova, S.Y. *MOD09 (Surface Reflectance) User's Guide*; MODIS Land Surface Reflectance Science Computing Facility: College Park, MD, USA, 2008.
39. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ*. **2002**, *83*, 195–213.
40. Guindin-Garcia, N.; Gitelson, A.A.; Arkebauer, T.J.; Shanahan, J.; Weiss, A. An evaluation of MODIS 8- and 16-day composite products for monitoring maize leaf are index. *Agric. For. Meteorol*. **2012**, *161*, 15–25.
41. Lisenberg, V.; Ponzoni, F.J.; Galvão, L.S. Analysis of the seasonal dynamics and spectral separability of some savanna physiognomies with vegetation indices derived from MODIS/TERRA and AQUA. *Rev. Árvore* **2007**, *31*, 295–305.
42. Rouse, J.W.; Haas, R.H.; Schell J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. In Proceedings of the Third Earth Resources Technology Satellite-1 Symposium, Greenbelt, MD, USA, 10–15 December 1973; pp. 301–317.
43. De Carvalho Júnior, O.A.; Hermuche, P.M.; Guimarães, R.F. Identificação regional da floresta decidual na bacia do rio Paranã a partir da análise multitemporal de imagens MODIS. *Rev. Bras. Geofis*. **2006**, *24*, 319–332.
44. De Carvalho Júnior, O.A.; Sampaio, C.S.; da Silva, N.C.; Couto Júnior, A.F.; Gomes, R.A.T.; de Carvalho, A.P.F.; Shimabukuro, Y.E. Classificação de padrões de savana usando assinaturas temporais NDVI do sensor MODIS no parque nacional Chapada dos Veadeiros. *Rev. Bras. Geofis*. **2008**, *26*, 505–517.
45. Savitzky, A.; Golay, M.J.E. Smoothinbg and differentiation of data by simplified least squares procedures. *Anal. Chem*. **1964**, *36*, 1627–1639.

46. Ataman, E.; Aatre, V.K.; Wong, K.M. Some statistical properties of median filters. *IEEE T. Acoust. Speech* **1981**, *29*, 1073–1075.
47. De Carvalho Júnior, O.A.; da Silva, N.C.; de Carvalho, A.P.F.; Couto Júnior, A.F.; Silva, R.S.; Shimabukuro, Y.E.; Guimarães, R.F.; Gomes, R.A.T. Combining noise-adjusted principal components transform and median filter techniques for denoising MODIS temporal signatures. *Rev. Bras. Geofis*. **2012**, *30*, 147–157.
48. Schefer, R.W. What is a Savitzky-Golay filter? *IEEE Signal Process. Mag*. **2011**, *28*, 111–117.
49. Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.; Matsushita, B.; Eklundh, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sens. Environ*. **2004**, *91*, 332–334.
50. Li, L.; Friedl, M.A.; Xin, Q.; Gray, J.; Pan, Y.; Frolking, S. Mapping crop cycles in China using MODIS-EVI time series. *Remote Sens*. **2014**, *6*, 2473–2493.
51. Vrieling, A.; de Leeuw, J.; Said, M.Y. Length of growing period over Africa: Variability and trends from 30 years of NDVI time series. *Remote Sens*. **2013**, *5*, 982–1000.
52. Geng, L.; Ma, M.; Wang, X.; Yu, W.; Jia, S.; Wang, H. Comparison of eight techniques for reconstructing multi-satellite sensor time-series NDVI data sets in the Heihe River Basin, China. *Remote Sens*. **2014**, *6*, 2024–2049.
53. Atkinson, P.M.; Jeganathan, C.; Dash, J.; Atzberger, C. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sens. Environ*. **2012**, *123*, 400–417.
54. De Beurs, K.M.; Henebry, G.M. Spatio-temporal statistical methods for modelling land surface phenology. In *Phenological research, Methods for Environmental and Climate Change Analysis*, 1st ed.; Hudson, I.L., Keatley, M.R., Eds.; Springer: Dordrecht, Netherlands, 2010; pp. 177–208.
55. Hird, J.N.; McDermid G.J. Noise reduction of NDVI time series: An empirical comparison of selected techniques. *Remote Sens. Environ*. **2009**, *113*, 248–258
56. Adams, J.B.; Gillespie, A.R. *Remote Sensing of Landscapes with Spectral Images: A Physical Modeling Approach*; Cambridge University Press: New York, NY, USA, 2006; p. 362.
57. Kruse, F.; Boardman, J.W.; Huntington, J.F. Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping. *IEEE Trans. Geosci. Remote Sens*. **2003**, *41,* 1388–1400.
58. Boardman, J. W. Automating spectral unmixing of AVIRIS data using convex geometry concepts. In *Summaries of the Fourth Annual JPL Airborne Geosciences Workshop*; Robert, O.G., Ed.; Jet Propulsion Laboratory Publication: Pasadena, CA, USA, 25–29 October 1993; pp. 11–14.
59. Craig M. Minimum-volume transforms for remotely sensed data. *IEEE Trans. Geosci. Remote Sens*. **1994**, *32*, 542–552.
60. Boardman, J.W.; Kruse, F.A. Analysis of imaging spectrometer data using n-dimensional geometry and a mixture-tuned matched filtering approach. *IEEE Trans. Geosci. Remote Sens*. **2011**, *49*, 4138–4152.
61. Green, A.A.; Berman, M.; Switzer, P.; Craig, M.D. A transformation for ordering multispectral data in terms of images quality with implications for noise removal. *IEEE Trans. Geosci. Remote Sens*. **1988**, *26*, 65–74.

62. Instituto Estadual de Florestas de Minas Gerais (IEF). Mapeamento da Cobertura Vegetal de Minas Gerais. Available online: http://www.inventarioflorestal.mg.gov.br (accessed on 16 July 2014).
63. Dickson, B.L.; Taylor, G.F. Maximum noise fraction method reveals detail in aerial gamma-ray surveys. *Explor. Geophys.* **2000**, *31*, 73–77.
64. De Carvalho Júnior, O.A.; Maciel, L.M.M.; de Carvalho, A.P.F.; Guimarães, R.F.; Silva, C.R.; Gomes, R.A.T.; Silva, N.C. Probability density components analysis: A new approach to treatment and classification of SAR images. *Remote Sens.* **2014**, *6*, 2989–3019.
65. Bateson, C.A.; Curtiss, B. A method for manual endmember selection and spectral unmixing, *Remote Sens. Environ*. **1996**, *55*, 229–243.
66. Bateson, C.; Asner, G.P.; Wessman, C. Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis. *IEEE Trans. Geosci. Remote Sens*. **2000**, *38*, 1083–1094.
67. Chan, T.H.; Chi, C.Y.; Huang, Y.M.; Ma, W.K. A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing. *IEEE Trans. Signal Process.* **2009**, *57*, 4418–4432.
68. Chang, C.I.; Wu, C.C.; Liu, W.M.; Ouyang, Y.C. A new growing method for simplex-based endmember extraction algorithm. *IEEE Trans. Geosci. Remote Sens*. **2006**, *44*, 2804–2819.
69. Winter, M.E. N-FINDER: An algorithm for fast autonomous spectral endmember determination in hyperspectral data. In Proceedings of the SPIE 3753, Imaging Spectrometry, Denver, CO, USA, October 1999; pp. 266–277. doi:10.1117/12.366289.
70. Kruse, F.A.; Lefkoff, A.B.; Boardman, J.W.; Heidebrecht, K.B.; Shapiro, A.T.; Barloon, P.J.; Goetz, A.F.H. The spectral imageprocessing system (SIPS)—Interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ*. **1993**, *44*, 145–163.
71. De Carvalho Júnior, O.A.; Meneses, P.R. Spectral Correlation Mapper (SCM): An improving on the Spectral Angle Mapper (SAM). In Proceedings of Ninth Annual JPL Airborne Earth Science Workshop, Pasadena, CA, USA, 23–25 February 2000; pp. 65–74.
72. De Carvalho Júnior, O.A.; Guimarães, R.F.; Gillespie, A.R.; Silva, N.C.; Gomes, R.A.T. A new approach to change vector analysis using distance and similarity measures. *Remote Sens*. **2011**, *3*, 2473–2493.
73. De Carvalho Júnior, O.A.; Guimarães, R.F.; Silva, N.C.; Gillespie, A.R.; Gomes, R.A.T.; Silva, C.R.; de Carvalho, A.P.F. Radiometric normalization of temporal images combining automatic detection of pseudo-invariant features from the distance and similarity spectral measures, density scatterplot analysis, and robust regression. *Remote Sens.* **2013**, *5*, 2763–2794.
74. Vapnik, V. *Estimation of Dependences Based on Empirical Data*; Springer Verlag: New York, NY, USA, 1982.
75. Kecman, V. Support vector machines—An introduction. In *Support Vector Machines: Theory and Applications*, 1st ed.; Wang, L. Ed.; Springer Science & Business Media: Heidelberg, Germany, 2005; pp. 1–47.
76. Shmilovici, A. Support vector machines. In *Data Mining and Knowledge Discovery Handbook*, 1st ed.; Maimon, O., Rokach, L., Eds.; Springer: USA, 2005; pp. 257–276.
77. Schölkopf, B.; Smola, A.J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*; The MIT Press: Cambridge, MA, USA, 2002.

78. Lu, J.; Plataniotis, K.N.; Venetsanopoulos, A.N. Kernel Discriminant Learning with Application to Face Recognition. In *Support Vector Machines: Theory and Applications*, 1st ed.; Wang, L., Ed.; Springer Science & Business Media: Heidelberg, Germany, 2005; pp. 275–296.
79. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66,* 247–259.
80. Carrão, H., Gonçalves, P.; Caetano, M. Contribution of multispectral and multitemporal information from MODIS images to land cover classification. *Remote Sens. Environ*. 2008, *112*, 986–997.
81. Vuolo, F.; Atzberger, C. Exploiting the classification performance of support vector machines with multi-temporal Moderate-Resolution Imaging Spectroradiometer (MODIS) data in areas of agreement and disagreement of existing land cover products. *Remote Sens.* **2012**, *4*, 3143–3167.
82. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 1–27.
83. Congalton, R.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; CRC/Lewis Press: Boca Raton, FL, USA, 1999.
84. Miglani, A.; Ray, S.S.; Vashishta, D.P.; Parihar, J.S. Comparasion of two data smoothing techniques for vegetation spectra derived from EO-1 Hyperion. *J. Indian Soc. Remote Sens*. **2011**, *39*, 443–453.
85. Ratana, P.; Huete, A.R.; Ferreira, L.G. Analysis of Cerrado physiognomies and conservation in the MODIS seasonal-temporal domain. *Earth Interact*. **2005**, *9*, 1–22.
86. De Carvalho Júnior, O.A.; Couto Júnior, A.F.; da Silva, N.C.; Martins, E.S.; Carvalho, A.P.F.; Gomes, R.A.T. Distância euclidiana e spectral correlation mapper em séries temporais NDVI-MODIS no campo de instrução militar de Formosa (GO). *Rev. Bras. Cartogr*. **2009**, *61*, 399–412.
87. Miranda, H.S.; Sato, M.N.; Neto, W.N.; Aires, F.S. Fires in the cerrado, the Brazilian savanna. In *Tropical Fire Ecology: Climate Change, Land Use, and Ecosystem Dynamic*, 1st ed.; Cochrane, M.A, Ed.; Springer-Praxis: Berlin, Germany, 2009; pp. 427–450.
88. Araújo, F.M.; Ferreira, L.G.; Arantes, A.R. Distribution patterns of burned areas in the Brazilian biomes: An analysis based on satellite data for the 2002–2010 Period. *Remote Sens*. **2012**, *4*, 1929–1946.
89. Daldegan, G.A.; de Carvalho Júnior, O.A.; Guimarães, R.F.; Gomes, R.A.T.; Ribeiro, F.F.; McManus. C. Spatial patterns of fire recurrence using remote sensing and GIS in the Brazilian savanna: Serra do Tombador Nature Reserve, Brazil. *Remote Sens*. **2014**, *6*, 9873–9894.
90. Archibald, S.; Scholes, R. Leaf green-up in a semi-arid African savanna-separating tree and grass responses to environmental cues. *J. Veg. Sci*. **2007**, *18*, 583–594.
91. Higgins, S.I.; Delgado-Cartay, M.D.; February, E.C.; Combrink, H.J. Is there a temporal niche separation in the leaf phenology of savanna trees and grasses? *J. Biogeogr*. **2011**, *38*, 2165–2175.
92. RapidEye AG. Satellite Imagery Product Specifications. Available online: www.rapideye.net/upload/RE_Product_Specifications_ENG.pdf (acessed on 20 November 2014).
93. Cover, T.M.; Hart, P.E. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27.
94. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.

95. Alcântara, C.; Kuemmerle, T.; Prishchepov, A.; Radeloff, V.C. Mapping abandoned agriculture with multitemporal MODIS satellite data. *Remote Sens. Environ*. **2012**, *124*, 334–347.
96. Hüttich, C.; Herold, M.; Wegmann, M.; Cord, A.; Strohbach, B.; Schmullius, C.; Dech, S. Assessing effects of temporal compositing and varying observation periods for large-area land-cover mapping in semi-arid ecosystems: Implications for global monitoring. *Remote Sens. Environ*. **2011**, *115*, 2445–2459.
97. De Carvalho Júnior, O.A.; Guimarães, R.F.; Silva, C.R.; Gomes, R.A.T. Standardized time-series and interannual phenological deviation: New techniques for burned-area detection using long-term MODIS-NBR dataset. *Remote Sens.* **2015**, *7*, 6950–6985.
98. Lhermitte, S.; Verbesselt, J.; Verstraeten, W.W.; Veraverbeke, S.; Coppin, P. Assessing intra-annual vegetation regrowth after fire using the pixel based regeneration index. *ISPRS J. Photogramm*. *Remote Sens.* **2011**, *66*, 17–27.
99. Veraverbeke, S.; Lhermitte, S.; Verstraeten, W.W.; Goossens, R. The temporal dimension of differenced Normalized Burn Ratio (dNBR) fire/burn severity studies: The case of the large 2007 Peloponnese wildfires in Greece. *Remote Sens. Environ.* **2010**, *114*, 2548–2563.
100. Sakamoto, T.; Van Nguyen, N.; Kotera, A.; Ohno, H.; Ishitsuka, N.; Yokozawa, M. Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. *Remote Sens. Environ*. **2007**, *109*, 295–313.
101. Lunetta, R.S.; Knight, J.F.; Ediriwickrema, J.; Lyon, J.G.; Worthy, L.D. Land-cover change detection using multi-temporal MODIS NDVI data. *Remote Sens. Environ*. **2006**, *105*, 142–154.