Journal of Arid Environments 166 (2019) 43–50

Contents lists available at ScienceDirect

Journal of Arid Environments

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

Predicting plant species richness with satellite images in the largest dry forest nucleus in South America

Edna Samara e Silva Medeiros[a,*], Célia Cristina Clemente Machado[b], Josiclêda Domiciano Galvíncio[c], Magna Soelma Beserra de Moura[d], Helder Farias Pereira de Araujo[a]

[a] Universidade Federal da Paraíba, Campus II, Centro de Ciências Agrárias, Cep: 58.397.000, Areia, PB, Brazil
[b] Universidade Estadual da Paraíba, Campus V, Centro de Ciências Biológicas e Sociais Aplicadas, Cep: 58071-160, Cristo Redentor, Cep: 58071-160, Joao Pessoa, PB, Brazil
[c] Universidade Federal de Pernambuco, Centro de Filosofia e Ciências Geográficas, Cep: 50.670901, Recife, PE, Brazil
[d] Empresa Brasileira de Pesquisa Agropecuária, Centro de Pesquisa Agropecuária do Trópico Semiárido, Cep: 56302970, Petrolina, PE, Brazil

ARTICLE INFO

Keywords:
Biodiversity
Caatinga
Landsat
Remote sensing

ABSTRACT

Biodiversity assessment is considered an important indicator of ecosystem health by various initiatives worldwide. Satellite remote sensing (SRS) has allowed the development of tools that can assist with the practical search of information related to species richness. The aim of this study was to test whether Landsat satellite spectral variables could be used as indicators of plant species diversity in the Caatinga, the largest nucleus of dry forest in South America. To obtain plant diversity data (richness and Shannon's index), an exhaustive search of plant phytosociological studies carried out in Caatinga was conducted. Pearson's correlation and PCA analysis was used to test the association between spectral variables and plant diversity. Regressions were used to test the models that best explain species richness. The results indicate that a positive correlation exists between richness and the near-infrared (NIR) spectral band ($r = 0.744$; $p < 0.001$). This spectral band was also responsible for explaining better the variation of leaf level reflectance among eight species that occur in the region ($df = 7$; $F = 26317.55$; $p < 0.001$). Therefore, the NIR band variable can be used as an indicator of species richness using power and quadratic regression models, because they were one of the best fit association recorded between spectral variable and plant diversity index, when compared to other studies in natural environments. Thus, we provide important information about biodiversity that can be used in different researches, from ecological modeling for theoretical approaches to practical applications in Caatinga. The potential use of Landsat satellite imagery to estimate species richness makes biodiversity assessments easier and provides a continuous source of data for monitoring in Brazilian semiarid region.

1. Introduction

Biodiversity assessment is considered an important indicator of ecosystem health by various initiatives worldwide (Skidmore et al., 2015), such as the Group on Earth Observations (GEO), the International Geosphere-Biosphere Program (IGBP) under the International Council for Science, and the World Climate Research Programme (WCRP) (Rocchini et al., 2016). Species diversity assessments are usually based on surveys of local diversity (alpha diversity – α) within a given habitat or community; surveys of regional diversity (gamma diversity – γ), which correspond to the diversity of large areas; and surveys of beta diversity (β), which reflect changes in species composition in an environmental gradient (Rocchini et al., 2016; Valentin, 2012). Several indices are used to estimate α-diversity, with species richness and the Shannon-Wiener index being among the most commonly used. Species richness is a basic biodiversity indicator and the center of several ecological models and conservation policies (Giorgini et al., 2015; Rocchini et al., 2007, 2016).

A complete inventory of all species inhabiting large areas is utopian, as it is impractical for field biologists to inspect every organism in a region and monitor changes in species composition over time. Additionally, biodiversity assessments through field surveys face several challenges, for example: (1) defining the number of sampling units; (2) choosing the sampling design, which can impact the results and

* Corresponding author.
E-mail address: sa_medeiros.slv@hotmail.com (E. Samara e Silva Medeiros).

https://doi.org/10.1016/j.jaridenv.2019.03.001
Received 23 July 2018; Received in revised form 25 January 2019; Accepted 11 March 2019
Available online 22 March 2019
0140-1963/

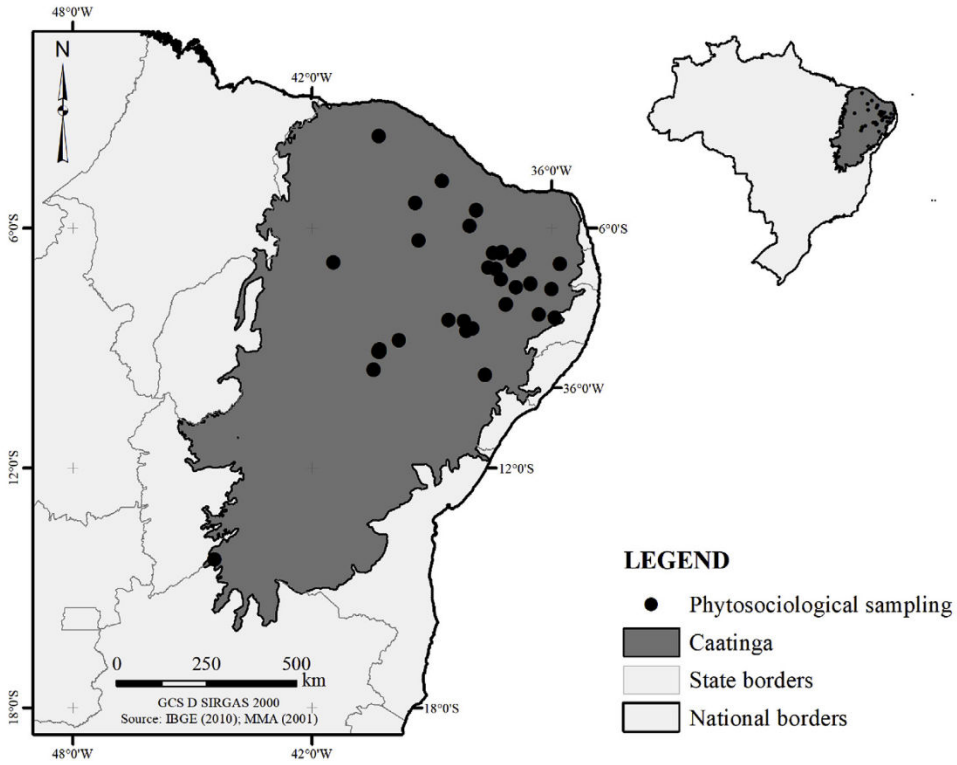

Fig. 1. Location of the study area, highlighting the Caatinga region and state borders (IBGE, 2010).

comparisons of distinct areas; (3) operationally defining the community to be considered; (4) experiencing a slow and costly process; and (5) experiencing difficulties in accessing remote areas (Rocchini et al., 2016; Scott and Hallam, 2003).

Satellite remote sensing (SRS) includes tools that can assist with the practical search of information related to species richness and, consequently, can attenuate the problems presented above, given its low cost, greater operability, greater spatial coverage, and its potential to provide a continuous source of information on biodiversity distribution (Wang, 2012). In a time of intense and rapid environmental changes, SRS presents an opportunity for the acquisition of critical data for spatiotemporal biodiversity monitoring (Pettorelli et al., 2014; Skidmore et al., 2015). However, most studies on environmental monitoring through SRS use vegetation indices, which are indicators of the quantity and conditions of the green vegetation, or classification techniques, which group spectrally similar areas and map vegetation according to its main physiognomies (Beuchle et al., 2015; Hansen et al., 2013; IBGE, 2010). The Brazilian National Institute of Space Research (Instituto Nacional de Pesquisas Espaciais – INPE), for example, is currently mapping the deforestation of the Caatinga ecosystem, the semiarid region of Brazil, and has classified the area into five general classes: preserved vegetation, degraded vegetation, exposed soil, crops, water, and urban areas. Preliminary results based on 2013–2014 Landsat-8 satellite data indicate that approximately 40% of the Caatinga is preserved. A different recent estimate of natural vegetation remnants indicates contradictory results, one suggest that approximately 63.2% of the natural vegetation remains in this region (Beuchle et al., 2015), other suggest that at least 63.3% of the Caatinga is composed of anthropogenic ecosystems (Silva and Barbosa, 2017). These estimates, based on classes generalized by indices or vegetation cover, can mask natural vegetation biodiversity data, since they may identify areas of monospecific arboreal vegetation, such as plantations or areas dominated by invasive species, as natural areas.

One strategy to address these methodological limitations is to use SRS to assess species richness, a challenging and unusual approach (Nagendra et al., 2010). Several studies that tested remote sensing to predict species richness have obtained positive and promising results (Fricker et al., 2015; Nagendra et al., 2010; Palmer et al., 2002; Rocchini et al., 2007). These are based on the spectral variability hypothesis (SVH), which suggests a relation between species diversity and spectral heterogeneity (Palmer et al., 2002; Rocchini et al., 2007).

Different approaches in different environments have shown advances in alpha-diversity assessment using remote sensing, such as the successional fields from USA, tropical dry forests from Florida-USA, Venezuela and India, wetlands from Italy, boreal forests from Finland and Bornean tropical rainforests from Malaysia and Panamá (Aneece et al., 2017; Feeley et al., 2005; Fricker et al., 2015; Gillespie, 2005; Nagendra et al., 2010; Parviainen et al., 2009; Rocchini et al., 2007). However, spectral information may not be crucial in some environments because the moderate species richness may require spectral information for accurate species discrimination, or the temporal variations of spectral properties can be related to differences in phenology and physiology in seasonal environments (Rocchini et al., 2016). Although spectral information can be a good proxy of diversity estimate, caution must be taken considering additional multiscale drivers like climate, soil types, topographic variables, and biotic interactions (Rocchini et al., 2016).

Therefore, the aim of this study was to test whether spectral variables can be used as indicators of plant species diversity in the Caatinga, the largest nucleus of dry forest in South America, where no such evaluation was performed. The Caatinga region is the richest Seasonally Dry Tropical Forest and Woodlands area in the New World (Queiroz et al., 2017). Although Caatinga dry land vegetation is part of a global biome that has been variously treated as dry forests, its vegetation structure is extremely variable, with strong floristic links between the different vegetation types, ranging from open cactus scrub to semi-deciduous forests (Queiroz et al., 2017). For that reason, we questioned: can any of the spectral variables be used to predict diversity in the Caatinga? If so, what is the power of this prediction? This type of information can bring practical applications for actions in a region that needs urgent measures for conservation and sustainable development (Tabarelli et al., 2017).

2. Material and methods

2.1. Study area

This study was conducted in the Caatinga region, an area of approximately 844,453 km^2. The region includes the Brazilian states of Alagoas, northern and central Bahia, Ceará, Pernambuco, Paraíba, Rio Grande do Norte, southeastern Piauí, Sergipe, and northern Minas Gerais (IBGE, 2010) (Fig. 1). According to the Köppen-Geiger climate classification system, updated by Alvares et al. (2013), the Caatinga region contains areas of tropical climate with dry summers (As) and hot semi-arid climate in hot steppes at low latitude and altitude (Bsh). The annual mean temperature is between 24 and 28 °C, with low and irregular rainfall varying between 250 and 1000 mm, as well as water deficit conditions during most of the year. The vegetation displays a variety of physiognomic types, from shrubby areas to seasonally dry forests, consisting mainly of small woody and herbaceous species, usually with spines and reduced and deciduous leaves, as well as many succulents and therophitic herbs that efficiently respond to the precipitation levels (Queiroz et al., 2017).

2.2. Plant diversity data

To obtain plant diversity data (richness and Shannon's index), an exhaustive search was conducted for phytosociological studies carried out in the study area, which included researches published between 1989 and 2015. From 32 studies, we filtered 60 of 91 sites with diversity information (Fig. 1). We excluded sites with outliers and problems during the image processing. All 60 sites have species richness data (n = 60) and 25 have Shannon index data (n = 25). The difference between the number of samples (n) in the two categories was due to the smaller number of publications reporting the use of the Shannon's index. A point vector file with associated species richness and/or Shannon index value was generated.

2.3. Image acquisition and processing

Thematic Mapper (TM) and Operational Land Imager (OLI) images from the Landsat 5 and 8 satellites were selected due to their ubiquity, extensive database (32 years of image acquisition), spectral resolution, adequate spatial resolution (30 m) for environmental analysis (Townshend et al., 2012), their previous use in several species biodiversity modeling studies (Duro et al., 2014; Gillespie, 2005; Nagendra et al., 2010), and free access.

Fifteen images were obtained from the U.S. Geological Survey (USGS) Global Visualization Viewer (GloVis). The orbit-point and date of the images varied according to the sampling date and publication date. Images captured as near as possible to the sampling date were selected, considering the limitations imposed by cloud cover and the satellite's temporal resolution (16 days).

2.3.1. Radiometric calibration

The pixel digital number (DN) of the Landsat 5 TM sensor images was converted to monochromatic spectral radiance (Lλ,i - W m-2 sr-1 μm-1) using the equation proposed by Markham and Barker (1987):

$$L_{\lambda i} = a_i + \left(\frac{b_i - a_i}{255}\right) DN \quad (1)$$

where a and b are the minimum and maximum spectral radiances, respectively (W m^{-2} sr^{-1} μm^{-1}); DN is the pixel intensity (integer between 0 and 255); and i corresponds to Landsat 5 TM bands 1 through 7. The calibration coefficients used were proposed by Chander et al. (2007, 2009) (Table 1).

2.3.2. Planetary spectral reflectance

Planetary reflectance is given by the ratio between the hemispheric integration of the monochromatic radiance and the monochromatic solar irradiance incident on a horizontal surface. For the images of the Landsat 5 TM sensor, the equation used by NASA (1998) was applied:

$$\rho_{\lambda i} = \frac{\pi \,.\, L_{\lambda i}}{ESUN_{\lambda}.\cos\theta \,.\, d_r} \quad (2)$$

where $L_{\lambda i}$ is the spectral radiance of each band; $ESUN_{\lambda}$ is the spectral solar irradiance of each band at the top of the atmosphere (W m^{-2} μm^{-1}) (Table 2); θ is the solar zenith angle; and d_r is the inverse of the square of the relative distance between the Earth and the Sun (Table 2).

In the case of the Landsat 8 OLI sensor images, as the Earth-Sun distance and ESUN data had already been integrated into the digital number of each pixel, calculation of the reflectance directly from the raw image was possible using the multiplicative and additive scaling factors from Table 2, which were available in the image metadata file.

$$\rho_{\lambda i} = \frac{M_L .\, DN + A_L}{Cos\ \theta} \quad (3)$$

M_L is the multiplicative conversion factor, and AL is the additive conversion factor specific to each band, available in the image metadata file.

Table 1
Description of Landsat 5 TM bands and spectral bands used, minimum (a) and maximum (b) calibration coefficients, and spectral solar irradiance at top of atmosphere (ESUNλ) (Chander et al., 2007, 2009).

Bands	Spectral bands (μm)	Calibration coefficient (W m^{-2} μm^{-1})				ESUN$_{\lambda}$ (W m^{-2} μm^{-1})
		03/01/1984–12/31/1991		after 01/01/1992		
		A	B	a	B	
BLUE	0.45–0.52	−1.52	169.00	−1.52	193.00	1983
GREEN	0.52–0.60	−2.84	333.00	−2.84	365.00	1796
RED	0.63–0.69	−1.17	264.00	−1.17	264.00	1536
NIR	0.76–0.90	−1.51	221.00	−1.51	221.00	1031
SWIR 1	1.55–1.75	−0.37	30.20	−0.37	30.20	220
SWIR 2	2.08–2.35	−0.15	16.50	−0.15	16.50	83.44

2.4. Vegetation index

We explored eight vegetation indices (VI) cited by literature to evaluate possible association with species diversity variables (Table 3). Some VI were chosen for their ability to measure spectral variability and the relation with plant diversity (eg. Feeley et al., 2005; Gillespie, 2005), while others were tested for being Caatinga specific vegetation indices (eg. Machado, 2014; Ribeiro et al., 2016).

Table 2
Description of Landsat 8 OLI bands and spectral bands, multiplicative factors (M_L), and additive factors (A_L) used to calculate the radiance and spectral reflectance for 08/04/2013 and spectral solar irradiance at the top of the atmosphere (ESUN$_{\lambda}$).

Bands	Spectral bands (μm)	Calibration coefficients reflectance		ESUNλ (W m^{-2} μm^{-1})
		M_L	A_L	
BLUE	0.45–0.51	0.00002	−0.1	2067.00
GREEN	0.53–0.59	0.00002	−0.1	1893.00
RED	0.64–0.67	0.00002	−0.1	1603.00
NIR	0.85–0.88	0.00002	−0.1	972.60
SWIR 1	1.57–1.65	0.00002	−0.1	245.00
SWIR 2	2.11–2.29	0.00002	−0.1	79.72

Table 3
Vegetation indeces used in present manuscript and respective description and model.

Vegetation Index	Description	Model
Simple Ratio Index (SR)	Birth and Mcvey (1968)	$SR = \frac{\rho NIR}{\rho R}$
Normalized Difference Vegetation Index (NDVI)	Rouse et al. (1974)	$NDVI = \frac{\rho NIR - \rho R}{\rho NIR + \rho R}$
Soil-Adjusted Vegetation Index (SAVI)	Huete (1988), following Allen et al. (2002)	$SAVI = \frac{(1 + L)(NIR - R)}{(L + \rho NIR + \rho R)}$
Enhanced Vegetation Index (EVI)	Huete et al. (1997).	$EVI = G\frac{\rho NIR - \rho R}{\rho NIR + C_1\rho R + C_2\rho B + L)}$
Leaf area index (LAI)	We used the LAI as proposed by Galvíncio et al. (2013) (LAIGalv) and Machado (2014) (LAIMac) for values adjusted to Caatinga conditions	$LAI_{Galv} = EXP(1.426 + (-0.542/NDVI)$ $LAI_{Mac} = 0.102e^{5.341NDVI}$
Normalized Difference Moisture Index or Water Index (NDMI or NDWI)	Hardisky et al. (1983), Hardisky et al.(1983), Gao (1996)	$NDMI \text{ or } NDWI = \frac{\rho NIR - \rho SWIR}{\rho NIR + \rho SWIR}$
Difference Vegetation Index (DVI)	Richardson and Wiegand (1977)	$DVI = \rho NIR - \rho R$

2.5. Leaf-level spectral reflectance

Due to the variation of floristic and vegetation structure in Caatinga (Queiroz et al., 2017), we analyzed leaf-level spectral reflectance of eight species in order to verify if these species can be distinguished by their particular reflectance in the BLUE, GREEN, RED and NIR spectral regions. The studied species were *Croton sonderianus*, *Croton conduplicatus* Kunth, *Manihot glaziovii*, *Jatropha mollissima* (Pohl) Bail, *Bauhinia* sp., *Capparis flexuosa* L., *Commiphora leptophloeos*, *Cereus jamacaru*. These species show taxonomic heterogeneity, some belong to the same genus, others to different families; different sizes and shapes of the reflecting units (leaf or stem for cacti), ranging from broad to small, simple or compound leaves, and different life forms (trees, shrubs and cacti).

We used a spectrorradiometer (model FieldSpec® HandHeld Pro) to measure the leaf spectral reflectance between 336 and 1045 nm with a resolution of 1 nm. It was used a 1 and 10° HH FOV lens foreoptic with Radiometric Calibration. We expected that leaf spectral reflectance among species may be associated to the best indicators from spectral variation obtained by SRS data.

2.6. Statistical analysis

The average of spectral variables (individual spectral bands – RED, GREEN, BLUE, NIR and SWIR and VIs – NDVI, EVI, LAI Mac, LAI Galv, NDWI, SAVI, SR, and DVI) was obtained from four adjacent pixels to the each point vector. This average was associated to the 60 sites with respective species richness information and 25 sites with respective Shannon index value. Pearson correlations were used to measure the association between these biodiversity measurements and spectral variables. This method was also applied by Gillespie (2005), Chaves et al. (2013) and Machado (2014).

We used Principal Components Analyses (PCA) with variance and covariance matrix to verify possible association between spectral variables and plant diversity indicators. For this, all the variables ranges were proportionally standardized in an interval between 0 and 1 and, consequently, a proportional matrix of variation data was elaborated for PCA.

In order to obtain an equation for predicting the diversity of plant species in the Caatinga, we tested linear, quadratic, exponential, and power regression models between the spectral variable that best correlated with diversity variable. The selected model had (1) the highest probability of association between the variables ($p < 0.05$); (2) the best coefficient of determination (R^2 closest to 1), corresponding to the best fit curve running through the points; and (3) the best Akaike's Information Criterion (AIC) value, as it was the most parsimonious model that could satisfactorily fit the data (Sobral and Barreto, 2011).

A One Way Variance Analysis (ANOVA) was performed to verify if there is variation in leaf-level spectral reflectance in the eight species, using the BLUE, GREEN, RED and NIR regions. Sixty-one reflectance units for each species were used in this comparison. The p-values from significance Dunnett's T3 Mean Test was used as post-hoc test for paired comparison. Next, we used the discriminant function analysis to determine which variables discriminate between the studied species.

3. Results

The species richness varied between 2 and 55 and showed a significant and positive correlation with all spectral variables, except SWIR2. The correlation coefficients of the VIs ranged from 0.343 (Richness – LAI_GALV) to 0.536 (Richness – DVI). NIR had the highest correlation coefficient ($r = 0.744$; $p < 0.001$) (Table 4).

The Shannon diversity index varied between 0.61 and 3.09 and showed weaker correlation with spectral variables compared to richness (Table 4). LAI_MAC, LAI_GALV, SAVI, SR, NDVI, and NDWI presented a positive and significant correlation to Shannon diversity. The correlation coefficients ranged from 0.396 (Shannon diversity – LAI_MAC) to 0.466 (Shannon diversity – NDWI). Among individuals bands BLUE, GREEN, RED, SWIR1, and SWIR2 showed a negative correlation, being SWIR2 band the most sensitive to Shannon diversity ($r = -0.460$, $p < 0.05$).

The two first principal components accumulate about 87% of the variability in spectral and diversity variables (PC1 = 69, 7%, PC2 = 17y, 19%). The spectral variables demonstrated different performances in PCA and NIR showed similar performance to species richness when observed the proportional variation of data (Fig. 2).

Table 4
Mean and standard deviation of spectral variable values and the Pearson correlation coefficient between diversity indicator and spectral variables. The highest value obtained for r is in bold (see abbreviations in the methods section).

Spectral variable	Mean	Standard deviation	Species Richness (N = 60)	Shannon diversity (N = 25)
NDVI	0,279291	0,116004	0.378**	0.407*
EVI	0,19925	0,107655	0.519***	0.392
LAI_MAC	0,596231	0,669343	0.444***	0.396*
LAI_GALV	0,566785	0,413299	0.343**	0.405*
NDWI	−0,0997	0,124302	0.426***	0.466*
SAVI	0,226867	0,099911	0.458***	0.397*
SR	1,869449	0,748523	0.428***	0.402*
DVI	0,0836	0,045136	0.536***	0.380
BLUE	0,10587	0,012777	0.290*	−0.409*
GREEN	0,101848	0,017495	0.404**	−0.316
RED	0,105018	0,028149	0.281*	−0.384
NIR	0,185792	0,046225	**0.744*** **	0.210
SWIR1	0,231352	0,05627	0.344**	−0.367
SWIR2	0,182357	0,10236	−0.045	−0.460*

*P < 0.05, **P < 0.01, ***P < 0.001.

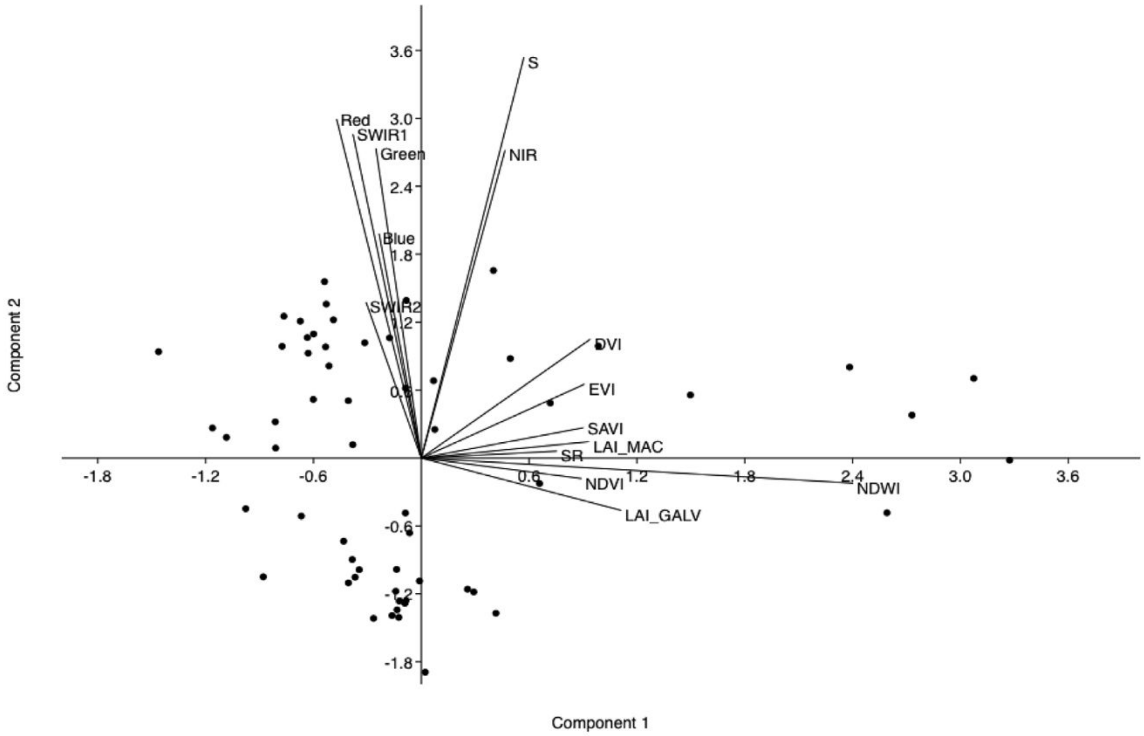

Fig. 2. Principal Components Analyses demonstrating association between spectral variables and species richness (*S*).

The spectral reflectance in BLUE region showed significant differences between several species (df = 7; F = 436.19; p < 0.001), but did not show differences between *M. glaziovii* and *J. mollissima* (p > 0.05) and between *M. glaziovii* and *Bauhinia sp*. (p > 0.05), when the leaf-level reflectance were analyzed. In GREEN region, the differences were significant among several species (df = 7; F = 1877.36; p < 0.001) except between *C. sonderianus* and *J. mollissima* (p > 0.05). In RED region there were also significant differences in several species (df = 7; F = 1610.85; p < 0.001), except between *M. glaziovii* and *Bauhinia* sp. (p > 0.05) and between *M. glaziovii* and *C. flexuosa* L. (p > 0.05). The NIR region showed significant differences among all species (df = 7; F = 26317.55; p < 0.001). So, NIR is the region of the spectrum that most contributes to distinguish the eight tested Caatinga species.

Four discriminant functions were extracted from the discriminant analysis, and one of then, mostly defined by the NIR variable, was responsible for explaining to 98.1% of the variability among the eight species (Table 5, Fig. 3).

Because the highest correlation coefficients were obtained using richness, its explanatory power was tested using four regression models (Table 6). Among the spectral variables, the power regression (R^2 = 0.61) and the quadratic regression (R^2 = 0.56) of the NIR band were the best fitting equations (Table 6, Figs. 4 and 5).

Table 5
Standardized coefficients of the variables with discriminant power and percentage of variance between the groups explained in the four discriminant functions to identify the variation among leaf-level reflectance of eight plant species studied. The highest value obtained for r is in bold.

Variables	Coefficients of discriminant functions			
	1	2	3	4
BLUE	−0,473	1588	−0,369	−0,481
GREEN	0,133	1360	0,672	−0,458
RED	0,803	0249	−0,166	0950
NIR	**1450**	−0,277	−0,070	0044
Eigen value	2100,329	40,162	0,345	0081
Variance explained	**98,1**	1,9	0	0

4. Discussion

Our results demonstrate that it is possible to estimate plant species richness for the Caatinga Forest using the reflectance of the NIR spectral band. The power regression model approach recommended here demonstrate one of the best coefficient of determination (R^2), that is, the best fit association recorded between spectral variable and plant diversity index, when compared to other studies in natural environments (Aneece et al., 2017; Duro et al., 2014; Feeley et al., 2005; Fricker et al., 2015; Gillespie, 2005; Nagendra et al., 2010; Parviainen et al., 2009; Rocchini et al., 2016). Among spectral variables tested, the PCA confirms that NIR was the variable associated proportional and similarly to species richness in a matrix of variance and co-variance.

Our results also show the potential of Landsat satellite imagery to estimate species richness lies more in its near infrared wavelengths than in the combination of bands used to calculate the vegetation indices. This result is consistent with that observed in an area of riparian forest in the Tuscany region of Italy (Rocchini et al., 2007). NIR was also responsible for explaining better the variation of leaf level reflectance among eight species that represent a sample of variation in taxonomic, leaf size, shape, and type, and life forms (trees, shrubs and cacti) of the Caatinga region.

Many authors have used parameters that are well established in the literature as a measure of spectral variability, which are based on the infrared/visible wavelength relation, using various combinations that form the vegetation indices. Among them, the NDVI has been one of the most evaluated as a potential indicator of species richness (Duro et al., 2014; Fairbanks and McGwire, 2004; Gillespie, 2005; Nagendra et al., 2010). In the present study, NDVI presented one of the lowest correlations with richness values, slightly higher than LAIGalv. This is consistent with the results observed in wetland area in Italy, tropical forest in Borneo and dry forest in India (Nagendra et al., 2010; Rocchini et al., 2004), who obtained a weak explanatory potential between NDVI and species richness (R^2 < 0.30). From a statistical point of view, these results might occur because the NDVI, as well as other vegetation

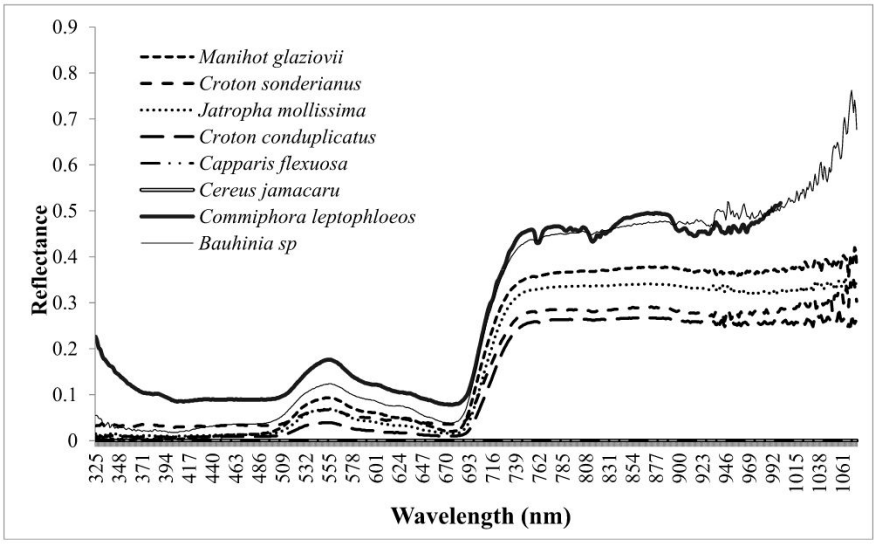

Fig. 3. Distribution of leaf reflectance data among eight species occurring in the Caatinga region.

indices, reduces the range of values by rescaling the original data, considerably affecting the pixel value dispersion (Rocchini et al., 2007). By contrast, the biomass, or amount of GREEN, which was explained by vegetation indices, is not necessarily associated with species richness (Weiher, 2003).

In some studies, individual bands correlated more strongly with diversity measures, a finding similar to our results (Rocchini et al., 2007; Nagendra et al., 2010; Chawla et al., 2010). Among the spectral bands analyzed in this study, the near infrared spectral band stands out due to its high correlation values and explanatory potential for species richness ($r = 0.744$, $R^2 = 0.6108$, respectively), similarly to what was observed by Rocchini et al. (2007). Among the results obtained with individual bands, GREEN presented the second-best result, which was also found by Rocchini et al. (2007). Similar results were also found for a tropical dry forest in India, where the GREEN and BLUE bands seemed to be the most sensitive (Nagendra et al., 2010). A greater sensitivity to richness is observed in the bands where the highest radiation reflectance occurs, in the NIR, where there is a small absorption of the radiation and considerable internal scattering, and in the GREEN, where most of the plants are moderately transparent. This contrasts with what happens with the BLUE, RED, and short-wave infrared regions, where greater radiation absorption takes place.

The high explanatory potential of the near infrared spectral band for species richness could be due to the interaction of the electromagnetic radiation in this spectrum range with plant canopies. The occurrence of

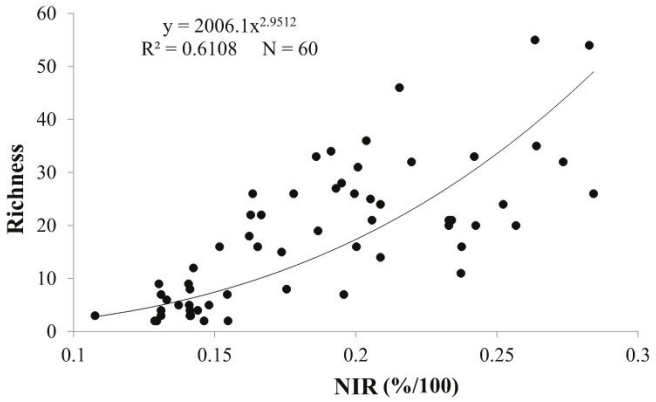

Fig. 4. Power regression between richness and reflectance of band 4 (NIR), based on a Landsat image.

the highest radiation reflectance in the near infrared region (Lillesand et al., 2004; Jensen, 2011) due to internal radiation scattering at leaf level and the multiple reflection among the various leaf layers at the canopy level is well documented in the literature. At the leaf level, different species exhibit differences in mesophyll structure, water content, and air-water ratio and the variability of photosynthetic pigments, consequently, reflect radiation differently. At the canopy level,

Table 6
Results of the regression models for richness using the spectral variables (VIs and spectral bands). The highest R^2 value and the AIC lowest value are in bold.

Spectral variable	Linear		Exponential		Power		Quadratic	
	R^2	AIC	R^2	AIC	R^2	AIC	R^2	AIC
NDVI	0.1425	8555.8	0.053	8018.2	0.018	7930.8	0.236	7625.9
EVI	0.2695	7289.1	0.1567	7291.2	0.1394	7284.7	0.2696	7291.1
LAI_MAC	0.1974	8008.1	0.0938	8010.5	0.0424	7988	0.1975	8010
LAI_GALV	0.1179	8801.6	0.0357	7945.6	0.0007	9848.3	0.2649	7337.7
NDWI	0.1815	8167.1	0.091	7945.3	–	–	0.2254	7731.2
SAVI	0.2094	7889	0.1041	7700.2	0.065	7640.9	0.2403	7582.9
SR	0.1833	8149.6	0.0838	8110.4	0.0621	8073.7	0.1916	8068.5
DVI	0.2878	7106.9	0.1946	7110.2	0.211	7055.9	0.2948	7039.7
BLUE	0.0841	9138.2	0.2096	9139.7	0.2093	9138.9	0.0843	9138.9
GREEN	0.1636	8345.7	0.3069	8352	0.3194	8294.2	0.1779	8205.5
RED	0.0792	9186.9	0.2072	9115.4	0.1831	9082	0.1016	8965.7
NIR	0.5539	4452.7	0.5654	4462.1	**0.6108**	4310.3	0.5688	**4306.5**
SWIR1	0.1184	8796.7	0.2308	8800.7	0.2492	8711.2	0.1458	8525
SWIR2	0.002	9956.8	0.0023	9959	0.0124	9979.4	0.0252	9728.1

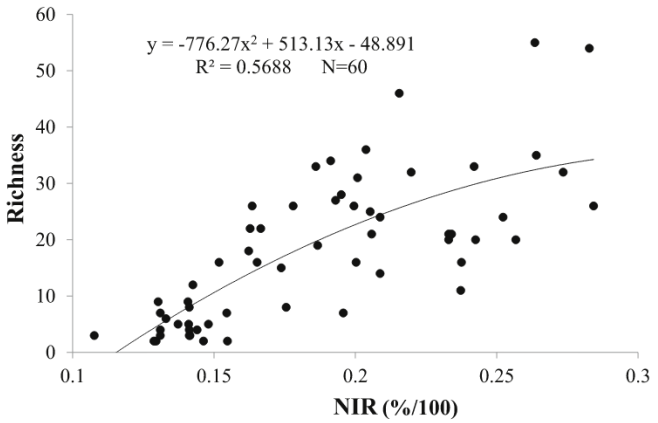

Fig. 5. Quadratic regression between richness and reflectance of band 4 (NIR), based on a Landsat image.

differences in the density and spatial orientation of the plants influence the reflectance dynamics in response to the variation in lighting and target geometries (Rocchini et al., 2007; Ponzoni and Shimabukuro, 2010).

According to Ponzoni and Shimabukuro (2010), the higher the number of layers in a canopy, the greater the reflectance is due to multiple scattering. This implies that an area with more biodiversity will have a greater number of layers in its canopy structure and, consequently, a higher reflectance in NIR, which explains the high and positive correlation coefficient obtained in this study. This multiple scattering also can be associated to different reflectance from different taxa, as observed in the eight species evaluated in this study.

The lower correlation obtained between spectral variables and the Shannon index may seem contradictory to that obtained for richness. However, statistically, similarly to what occurred with the vegetation indices, Shannon's index reduces the range of values by rescaling the original data. From a biological point of view, richness and Shannon's diversity index are conceptually different: the former corresponds to the total number of species in a sample unit, and the latter is a function of the number of species and the equitability of importance values of the species; that is, it represents the relative proportion of each species, which indicates whether the different species have similar or divergent abundances (Valentin, 2012). Still, Aneece et al. (2017) demonstrated that near infrared spectral could be used to estimate species diversity, but the relationships depend on the spectral region examined and the spectral transformation technique used, and we do not evaluate this here.

As observed in other studies, the results obtained confirm the usefulness of the Landsat satellite as a tool for evaluating plant diversity because the heterogeneity of the landscape in its scale (30 m) correlates with the scale in which species richness is observed. In addition to being free, another advantage of using Landsat is its availability of worldwide data, with information being captured at regular intervals (from 1972 to present), facilitating spatiotemporal surveys that assess biodiversity. Nagendra et al. (2010) compared high spatial resolution images from IKONOS with Landsat's medium resolution images and observed correlations between Landsat spectral variability and species diversity. They argue that the IKONOS data scale is too low (1 m), which makes it difficult to evaluate diversity (richness). Similar results were obtained by Rocchini et al. (2007), who attributed Landsat's better performance in evaluating plant biodiversity to its superior spectral resolution when compared to Quickbird.

4.1. Implications and future research

In summary, our study demonstrates that it is possible to quantify species richness in Brazilian semiarid region using satellite images of medium spatial resolution. The power or quadratic regression model approach of NIR is recommended to do this quantification. Therefore, we provide important information about biodiversity that can be used in different studies, from ecological modeling for theoretical approaches to practical applications. For example, the Caatinga region was recently characterized as a region with low rates of sustainable development, principally due to low values of green-structure (Silva et al., 2017). Designing landscapes for biodiversity-based ecosystem services is a key strategy for planning sustainable and resilient systems (Landis, 2017) and, therefore, our results present an essential tools to designing landscapes for biodiversity assessment and conservation planning.

Future researches about biodiversity and spectral information in Caatinga are needed. A sampling design to test the relation between species diversity and the spectral heterogeneity should be evaluated to verify how spectral variability hypothesis (SVH) (Palmer et al., 2002) is performed in Caatinga. Our plant diversity data were obtained from an exhaustive search among phytosociological studies, but it did not allow to evaluate this relation due to need of standardized spatial replicates. We verified the association between alpha diversity and spectral variable, but other studies shown association with beta diversity and this can be evaluated in Brazilian semiarid region. Still, these analyses need to be conducted at multiple spatial scales using different approaches to assess the scale most suitable for biodiversity monitoring (Rocchini et al., 2016).

Acknowledgements

We thank the Coordination for the Improvement of Higher Education Personnel (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES) for the scholarship granted to E.S.S. Medeiros. We also thank Santiago Veron and Damian Ravetta for the contributions during manuscript review.

References

Allen, R.G., Trezza, R., Tasumi, M., 2002. Surface energy balance algorithms for land. Advance training and user's manual 1, 98.

Alvares, C.A., Stape, J.L., Sentelhas, P.C., Gonçalves, J.L.M., Sparovek, G., 2013. Köppen's climate classification map for Brazil. Meteorol. Z. 22 (6), 711–728. https://doi.org/10.1127/0941-2948/2013/0507.

Aneece, I.P., Epstein, H., Lerdau, M., 2017. Correlating species and spectral diversities using hyperspectral remote sensing in early-successional fields. Ecology and Evolution 7, 3475–3488. https://doi.org/10.1002/ece3.2876.

Beuchle, R., Grecchi, R.C., Shimabukuro, Y.E., Seliger, R., Eva, H.D., Sano, E., Achard, F., 2015. Land cover changes in the Brazilian Cerrado and Caatinga biomes from 1990 to 2010 based on a systematic remote sensing sampling approach. Appl. Geogr. 58, 116–127. https://doi.org/10.1016/j.apgeog.2015.01.017.

Birth, G.S., Mcvey, G., 1968. Measuring the color of growing turf with a reflectance spectroradiometer. Agron. J. 60, 640–643. https://doi.org/10.2134/agronj1968.00021962006000060016x.

Chander, G., Markham, B., Barsi, J.A., 2007. Revised landsat-5 thematic mapper radiometric calibration. IEEE Geosci. Remote Sens. Lett. 4 (3), 490–494. https://doi.org/10.1109/LGRS.2007.898285.

Chander, G., Markham, B.L., Helder, D.L., 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Rem. Sens. Environ. 113 (5), 893–903. https://doi.org/10.1016/j.rse.2009.01.007.

Chaves, L.B., Francisco, P.R.M., Lima, E. R. V. De, Silva, B. B. Da, Brandão, Z.N., Chaves, L.H.G., 2013. Índices espectrais, diagnóstico da vegetação e da degradação da Caatinga da Bacia do Rio Taperoá-PB. In: Silva, B.B. da (Ed.), Aplicações ambientais brasileiras de geoprocessamento e sensoriamento remoto. EDUFCG, Campina Grande, pp. 1–31.

Chawla, A., Kumar, A., Rajkumar, S., Sing, R.D., Thukral, A.K., Ahuja, P.S., 2010. Correlation of multispectral satellite data with plant species richness vis-à-vis soil characteristics in a landscape of Western Himalayan Region, India. Applied Remote Sensing 1 (1), 1–13.

Duro, D.C., Girard, J., King, D.J., et al., 2014. Predicting species diversity in agricultural environments using Landsat TM imagery. Rem. Sens. Environ. 144, 214–225. https://doi.org/10.1016/j.rse.2014.01.001.

Fairbanks, H.K., McGwire, K.C., 2004. Patterns of floristic richness in vegetation communities of California: regional scale analysis with multi-temporal NDVI. Glob. Ecol. Biogeogr. 13, 221–235. https://doi.org/10.1111/j.1466-822X.2004.00092.x.

Feeley, K.J., Gillespie, T.W., Terborgh, J.W., 2005. The utility of spectral indices from Landsat ETM+ for measuring the structure and composition of tropical dry forests. Biotropica 37 (4), 508–519. https://doi.org/10.1111/j.1744-7429.2005.00069.x.

Fricker, G.A., Wolf, J.A., Saatchi, S.S., Gillespie, T.W., 2015. Predicting spatial variations of tree species richness in tropical forests from high-resolution remote sensing. Ecol. Appl. 25, 1776–1789. https://doi.org/10.1890/14-1593.1.

Galvíncio, J.D., Moura, M.S.B., De, Silva, T.G.F., Da, Silva, B. B. Da, Naue, C.R., 2013. LAI improved to dry forest in semiarid of the Brazil. International Journal of Remote Sensing Applications 3 (4), 193–202. https://doi.org/10.14355/ijrsa.2013.0304.04.

Gao, B.C., 1996. NDWI-A normalized difference water index for remote sensing of vegetation liquid water from space. Rem. Sens. Environ. 58, 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3.

Gillespie, T.W., 2005. Predicting woody-plant species richness in tropical dry forests: a case study from South Florida, USA. Ecol. Appl. 15, 27–37. https://doi.org/10.1890/03-5304.

Giorgini, D., Giordani, P., Casazza, G., Amici, V., Mariotti, M.G., Chiarucci, A., 2015. Woody species diversity as predictor of vascular plant species diversity in forest ecosystems. For. Ecol. Manag. 345, 50–55. https://doi.org/10.1016/j.foreco.2015.02.016.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., et al., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–853. https://doi.org/10.1126/science.1244693.

Hardisky, M.A., Klemas, V., Smart, R.M., 1983. The influences of soil salinity, growth form, and leaf moisture on the spectral reflectance of spartina alterniflora canopies. Photogramm. Eng. Rem. Sens. 49, 77–83.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Rem. Sens. Environ. 25, 295–309. https://doi.org/10.1016/0034-4257(88)90106-X.

Huete, A.R., et al., 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. Rem. Sens. Environ. 59, 440–451. https://doi.org/10.1016/S0034-4257(96)00112-5.

IBGE - Instituto Brasileiro de Geografia e Estatística, 2010. Mapa de Biomas do Brasil, primeira aproximação. Rio de Janeiro: IBGE. Acessível em. www.ibge.gov.br Acessado em 10 de setembro de 2017.

Jensen, J.R., 2011. Sensoriamento Remoto Do Ambiente: uma Perspectiva Em Recursos Terrestres. São José Dos Campos. Parêntese editora, SP, pp. 672.

Landis, D.A., 2017. Designing agricultural landscapes for biodiversity-based ecosystem services. Basic Appl. Ecol. 18, 1–12. https://doi.org/10.1016/j.baae.2016.07.005.

Lillesand, T.M., Kiefer, R.W., Chipman, J.W., 2004. Remote Sensing and Image Interpretation. vol. 5 John Wiley, , NY.

Machado, C.C.C., 2014. Alterações na superfície do Parque Nacional do Catimbau (PE-Brasil): consolidação dos aspectos biofísicos na definição dos indicadores ambientais do bioma Caatinga/Célia Cristina Clemente Machado. Recife: O autor 221.

Markham, B.L., Barker, J.L., 1987. Thematic Mapper band pass solar atmospheric irradiances. Int. J. Remote Sens. 8 (3), 517–523.

Nagendra, H., Rocchini, D., Ghate, R., Sharma, B., Pareeth, S., 2010. Assessing plant diversity in a dry tropical forest: comparing the utility of Landsat and IKONOS satellite images. Rem. Sens. 2, 478–496. https://doi.org/10.3390/rs2020478.

NASA-National Aeronautics and Space Administration, 1998. Landsat 7 Science Data Users Handbook. NASA, USA Available at: 20 de outubro 2019.

Palmer, M.W., Earls, P., Hoagland, B.W., White, P.S., Wohlgemuth, T., 2002. Quantitative tools for perfecting species lists. Environmetrics 13, 121–137.

Parviainen, M., Luoto, M., Heikkinen, R., 2009. The role of local and landscape level measures of greenness in modelling boreal plant species richness. Ecol. Model. 220, 2690–2701. https://doi.org/10.1016/j.ecolmodel.2009.07.017.

Pettorelli, N., Laurance, W.F., O'Brien, T.G., Wegmann, M., Nagendra, H., Turner, W., 2014. Satellite remote sensing for applied ecologists:opportunities and challenges. J. Appl. Ecol. 51, 839–848. https://doi.org/10.1111/1365-2664.12261.

Ponzoni, F.J., Shimabukuro, Y.E., 2010. Sensoriamento Remoto no Estudo da Vegetação. 2010. São José Dos. Parêntese, Campos.

Queiroz, L.P., Cardoso, D., Fernandes, M.F., Moro, M.F., 2017. Diversity and evolution of flowering plants of the Caatinga domain. In: Silva, J.M.C., Leal, I.R., Tabarelli, M. Caatinga (Eds.), The Largest Tropical Dry Forest Region in South America. Springer International Publishing, Cham, Switzerland, pp. 23–64. https://doi.org/10.1007/978-3-319-68339-3_2.

Richardson, A.J., Wiegand, C.L., 1977. Distinguishing vegetation from soil background information (by gray mapping of Landsat MSS data). Photogramm. Eng. Rem. Sens. 43 (12), 1541–1542.

Ribeiro, E.P., Nóbrega, R.S., Mota Filho, F.O., Moreira, E.B.M., 2016. Estimativa dos índices de vegetação na detecção de mudanças ambientais na bacia hidrográfica do rio Pajeú. Geosul 31 (62), 59–92. https://doi.org/10.5007/2177-5230.2016v31n62p59.

Rocchini, D., Chiarucci, A., Loiselle, S., 2004. Testing the spectral variation hypothesis by using satellite multispectral images. Acta Oecol. 26, 117–120. https://doi.org/10.1016/j.actao.2004.03.008.

Rocchini, D., Ricotta, C., Chiarucci, A., 2007. Using satellite imagery to assess plant species richness: the role of multispectral systems. Appl. Veg. Sci. 10, 325–331. https://doi.org/10.1658/1402-2001 (2007)10[325:USITAP]2.0.CO;2.v.

Rocchini, D., Boyd, D.S., Feret, Jean-Baptiste, Foody, G.M., He, K.S., Lausch, A., Nagendra, H., Wegmann, M., Pettorelli, N., 2016. Satellite remote sensing to monitor species diversity: potential and pitfalls. Remote Sensing in Ecology and Conservation. https://doi.org/10.1002/rse2.9.

Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.H., Harlan, J.C., 1974. Monitoring vegetation systems in the great plains with ERTS. In: Proceedings of 2013 IEEE 3rd Earth Resource Technology Satellite (ERTS) Symposium. vol. 1. pp. 48–62.

Scott, W.A., Hallam, C., 2003. Assessing species misidentification rates through quality assurance of vegetation monitoring. Plant Ecol. 165, 101–115. https://doi.org/10.1023/A:1021441331839.

Silva, J.M.C., Barbosa, L.C.F., 2017. Impact of human activities on the Caatinga. In: Silva, J.M.C., Leal, I.R., Tabarelli, M. Caatinga (Eds.), The Largest Tropical Dry Forest Region in South America. *Springer* International Publishing, Cham, Switzerland, pp. 359–368. https://doi.org/10.1007/978-3-319-68339-3_13.

Silva, J.M.C., Barbosa, L.C.F., Pinto, L.P.Z., Chennault, C.M., 2017. Sustainable development in the Caatinga. P. 445-458. In: Silva, J.M.C., Leal, I.R., Tabarelli, M. Caatinga (Eds.), The Largest Tropical Dry Forest Region in South America. *Springer* International Publishing, Cham, Switzerland. https://doi.org/10.1007/978-3-319-68339-3_18.

Skidmore, A.K., Pettorelli, N., C, N., Coops, G.N., Geller, M., Hansen, R., Lucas, et al., 2015. Environmental science: agree on biodiversity metrics to track from space. Nature 523, 403–405. https://doi.org/10.1038/523403a.

Sobral, T.E.L., Barreto, G., 2011. Análise dos critérios de informação para a seleção de ordem em modelos auto-regressivos. https://doi.org/10.5540/DINCON.2011.001.1.0097.

Tabarelli, M., Leal, I.R., Scarano, F.R., Silva, J.M.C., 2017. The future of the Caatinga. P. 461-474. In: Silva, J.M.C., Leal, I.R., Tabarelli, M. Caatinga (Eds.), The Largest Tropical Dry Forest Region in South America. *Springer* International Publishing, Cham, Switzerland. https://doi.org/10.1007/978-3-319-68339-3_19.

Townshend, J.R.G., Masek, J.G., Huang, C.Q., Vermote, E.F., Gao, F., Channan, S., Jo, S., Feng, M., Narasimhan, R., Kim, D., Song, K., Song, D.X., Song, X.P., Noojipady, P., Tan, B., Hansen, M.C., Li, M.X., Wolfe, R.E., 2012. Global characterization and monitoring of forest cover using Landsat data: opportunities and challenges. International Journal of Digital Earth 5, 373–397. https://doi.org/10.1080/17538947.2012.713190.

Valentin, J.L., 2012. Ecologia numérica: uma introdução à análise multivariada de dados ecológicos. 2. Editora *Interciência*, Rio de Janeiro.

Wang, Y., 2012. Remote sensing of protected lands: an overview. In: Wang, Y. (Ed.), Remote Sensing of Protected Lands. CRC Press, Boca Raton.

Weiher, E., 2003. Species richness along multiple gradients: testing a general multivariate model in oak savannas. Oikos 101, 311–316. https://doi.org/10.1034/j.1600-0706.2003.12216.x.