Delimiting floristic biogeographic districts in the Cerrado and assessing their conservation status

Citation for published version:
Francoso, RD & Dexter, K 2019, 'Delimiting floristic biogeographic districts in the Cerrado and assessing their conservation status', *Biodiversity and Conservation*, pp. 1-24. https://doi.org/10.1007/s10531-019-01819-3

Digital Object Identifier (DOI):
10.1007/s10531-019-01819-3

Published In:
Biodiversity and Conservation

Delimiting floristic biogeographic districts in the Cerrado and assessing their conservation status

Renata D. Françoso, Kyle G. Dexter, Ricardo B. Machado, R. Toby Pennington, José R.R. Pinto, Reuber A. Brandão, James A. Ratter.

Françoso, R.D. (Corresponding author, renatafrancoso@yahoo.com.br; +55 61 3107-5645; ORCID ID 0000-0002-6369-8317)[1,2]

Dexter, K.G. (kyle.dexter@ed.ac.uk; ORCID ID 0000-0001-9232-5221)[3,4]

Machado, R.B. (rbmac@unb.br; ORCID ID 0000.0002.6508.9005)[5]

Pennington, R.T. (t.pennington@rbge.ac.uk; ORCID ID 0000-0002-8196-288X)[3,6]

Pinto, J.R.R. (jrrpinto@unb.br; ORCID ID 0000-0003-2028-6176)[7]

Brandão, R.A. (reuberbrandao@gmail.com; ORCID ID 0000-0003-3940-2544)[7]

Ratter, J.A. (j.ratter@rbge.ac.uk)[3]

[1] Graduate Program in Ecology, University of Brasília. Address: Campus Universitário Darcy Ribeiro, Brasília, DF, Brasil.

[2] Instituto Federal de Brasília, Campus Planaltina. Address: Rodovia BF 128 - km 21, S/N - Zona Rural Planaltina, Brasília – DF.

[3] Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh EH3 5LR, U.K.

[4] School of GeoSciences, University of Edinburgh, Edinburgh EH9 3FF, U.K.

[5] Zoology Department, Biological Science Institute, University of Brasília. Address: Campus Universitário Darcy Ribeiro, Brasília, DF, Brasil

[6] Geography, University of Exeter, Exeter EX4 4RJ, U.K.

[7] Department of Forest Engineer, Faculty of Technology, University of Brasília. Address: Campus Universitário Darcy Ribeiro, Brasília, DF, Brasil.

ABSTRACT

The Cerrado is a biodiversity hotspot in central Brazil that represents the largest expanse of savanna in the Neotropics. Here, we aim at identifying and delimiting Biogeographic Districts (BDs) within the Cerrado, to provide a geographic framework for conservation planning and scientific research prioritisation. We used data from 588 sites with tree species inventories distributed across the entire Cerrado. To identify BDs, we clustered sites based on their similarity in tree species composition. To determine why BDs differ in composition, we 1) determined the proportion of tree species in different BDs that derive from other biomes, to test the idea that geographically marginal BDs are influenced by neighbouring biomes and 2) assayed key climatic differences between BDs, to test the idea that environmental factors underlie compositional differences. We found seven BDs within the Cerrado, and found support for both ideas. Marginal BDs have a large proportion of tree species characteristic of Amazon (in CW and NW BDs) and Atlantic Forest (S BD), but the Cerrado endemic species are also important (in CE BD). Meanwhile, BDs differed significantly for multiple climatic variables. Finally, to provide a preliminary conservation assessment of these different BDs, we assessed their rate of land conversion and current coverage by Protected Areas. We found that BDs in the south and southwest of the Cerrado have experienced the greatest land conversion and are the least protected, while those in the north and northeast are less impacted and better protected. Overall, our results show how biogeographic analyses can contribute to conservation planning by giving clear guidelines on which BDs merit greater conservation and management attention.

Key words: Neotropical Savanna; Phytogeography, Indicator Species, Brazilian Savanna, Biogeographic Regionalization.

ACKNOWLEDGMENTS

R.D.F. thanks the Coordination of Improvement of Higher Level Personnel (CAPES) for the 6-month study period under the Science Without Borders Programme (Process 4893/13-1). R.B.M. (Process 303838/2016-8) and J.R.R.P. (Process: 307701/2014-0) received a research fellowship grant from National Council for Scientific and Technological Development (CNPq). K.G.D. and R.T.P. were supported by the Natural Environmental Research Council (grant NE/I028122/1). K.G.D. was supported by a Leverhulme Trust International Academic Fellowship.

INTRODUCTION

Human activity has affected natural resources at such a high level that it has generated a global biodiversity crisis (Jenkins 2003). Biodiversity threats are distributed unevenly across the globe (Brooks et al. 2006), with developing countries in the tropics currently representing the most vulnerable regions (FAO 2015). Land conversion will persist into the next decades due to agricultural expansion and intensification, especially in South America and sub-Saharan African (Jenkins 2003), affecting mainly tropical savannas (Grace et al. 2006). Brazil is one of the top four countries in South America in terms of predicted habitat loss (FAO 2015), which is concentrated in the Brazilian Cerrado (MMA/IBAMA 2011), a global biodiversity hotspot (Myers et al. 2000). Several thousand hectares of natural vegetation are converted every year in the Cerrado, at rates higher than observed in the Amazon (MMA 2017).

Despite the biological importance of the Cerrado, which originally had more than 2 million km^2, near 50% of its natural vegetation has been cleared, most of them caused by agricultural expansion (MMA 2015). This continuous and intensive conversion is not randomly distributed, but prevalent in some geographic regions and vegetation types (Bianchi and Haig 2012). For example, land conversion has tended to follow the implementation of road and other infrastructures, which starts from the south to the north. Thus, the southeast region being inhabited longer compared with the central and northern areas. Further, additional large declines of the Cerrado vegetation over the next 50 years have been predicted (Ferreira et al. 2012), especially for tableland areas with open vegetation formations, which are more suitable for the establishment of mechanized agriculture. By 2030, we may expect natural vegetation to be found mostly in existing Protected Areas (PAs) (Klink and Machado 2005). Currently, only 3% of the remaining natural vegetation in the Cerrado is maintained in areas of strict protection equivalent to the IUCN categories I to III (Françoso et al. 2015). Regional variation in species composition and the non-uniform human occupation of the Cerrado implies the need for specifically tailored conservation policies, based on regional planning. However, conservation efforts in the Cerrado have not followed any clear plan, with PAs being established opportunistically on a case-by-case basis (Françoso et al. 2015). Among nine described global approaches to conservation prioritization (Brooks et al. 2006), the Cerrado represents a reactive conservation scenario, with decisions based on threat, contrasting with Amazonia where decisions are often based on opportunity.

Ideally, conservation efforts and resources should be focused on areas that harbor the greatest proportion of regional biodiversity, including a diversity of ecological communities, the majority of regionally endemic species, and characteristic environmental conditions. By conserving representative examples of different biological communities and ecosystems that occur within a region, the majority of species in that region will also be conserved (Groves et al. 2002).

A biogeographic regionalization aims to represent distinct biological natural areas on a map (Morrone 2018), which can support conservation policies and scientific investigations. The use of different tools for the identification of homogeneous natural areas, based on animal and plant communities, at regional, continental or global scales, is a common approach in ecology and biogeography (e.g. Wallace 1876; Clements and Shelford 1939; Dice 1943; Udvardy 1975). Aiming to unify the nomenclature used for floral and faunal biogeographic regions, Udvardy (1975) proposed a hierarchical division with Realms, Biotic Provinces and Districts. Realms have continental scale and follow the large faunal regions of Wallace (1876). Provinces are subdivisions of Realms, comprising large subcontinental regions, characterized by the major biome that occupies the area. A biome is the combination of the predominant climax vegetation, the local biota (some typical species are distributed throughout the biome), and the prevailing climatic patterns (Clements and Shelford 1939). The third biogeographical level, the District, encompass smaller differences within the Provinces, but are essential to drive conservation efforts, since they represent unique features of the Province (Udvardy 1975). Higher or lower levels, such as Regions or Dominions, may also be used (Morrone 2014).

Areas of endemism, where the distribution of two or more endemic taxa overlap (Morrone and Url 1994), are also focus of biogeographic studies. The overlapping species distributions are assumed to be product of vicariant processes, such as tectonic-isolating events (Sanmartín 2012). Areas of endemism are the main units in the approach of historical biogeography (Szumik and Goloboff 2004). These areas may be large, covering a continental region, like the zoogeographic realms themselves (Morrone and Url 1994), or smaller, such as valleys and mountains (e.g. Silva and Bates 2002).

In contrast with the historical approach, ecological biogeography searches for patterns in the current distribution of organisms, which are determined by recent dispersal processes and environmental filters (Morrone et al. 1995). Ecological biogeography uses cluster methods to identify putatively similar localities in a geographic region, based on communities' similarities in species composition (Kreft and Jetz 2010). Cluster methods are useful for identifying repeated patterns of organisms' distributions across landscapes. All biogeographic approaches are useful for guiding conservation planning and reserve networks design (Whittaker et al. 2005; de Mello et al. 2015).

The identification of geographic regions in a large and threatened ecosystem, such as the Cerrado, is necessary for recognizing biological communities with different conservation needs, and to subsequently adjust conservation actions for different parts of the biome. The first step for maximizing the preservation of biodiversity in the Cerrado would be to determine its major biogeographic units that house different species and communities, thus deserving distinct conservation strategies.

Several studies have been conducted to identify conservation priorities areas in the Cerrado. These have used different approaches, such as the distribution of endemic species (Simon and Proença 2000; Silva and Bates 2002; Diniz-Filho et al. 2008; Nogueira et al. 2011; Carmignotto et al. 2012; Azevedo et al. 2016), the identification of vicariant processes (de Mello et al. 2015), macroecology (Diniz-Filho et al. 2008, 2009a) or species community composition (Ratter and Dargie 1992; Castro 1994; Ratter et al. 1996, 2003; Neves et al. 2015; Amaral et al. 2017).

The Cerrado biome harbors three to five main areas of endemism, depending on the studied group. These areas (the Central Plateau, Veadeiros Mountain Range, Guimarães Mountain Range, Espinhaço Mountain Range, and Araguaia Valley) have been recorded in studies conducted with distribution patterns of vertebrates (Diniz-Filho et al. 2008), birds (Silva and Bates 2002), herpetofauna (Nogueira et al. 2011; de Mello et al. 2015; Azevedo et al. 2016), and *Mimosa* species (Simon and Proença 2000).

Biogeographic studies based on community composition in the Cerrado show large areas that are relatively homogeneous in species composition (Ratter and Dargie 1992; Castro 1994; Ratter et al. 1996, 2003; Neves et al. 2015; Mews et al. 2016; Amaral et al. 2017) In a series of studies published from 1996 to 2003, Ratter and colleagues proposed six Floristic Provinces within the core area of Cerrado, and another two disjunct areas in the Amazon (Ratter and Dargie 1992; Ratter et al. 1996, 2003, 2011). These studies were based on an extensive sampling effort for woody plants of the

Cerrado, including more than 900 species of trees and large shrubs, and representing the most extensive botanical biogeographic study of the Cerrado to date.

Here, we aim to identify biogeographic districts within the Cerrado biome, based on a large dataset for woody plants, primarily trees, and propose specific regions as the first level of biodiversity surrogates for conservation planning in the Cerrado. Therefore, we are not interested in areas of endemism, because we do not want to neglect any part of the Cerrado, even if there are no endemic species within a given region. We expanded the woody plant floristic database of Ratter et al. (2003) from 376 to 588 sites, and delimited Biogeographic Districts in this dataset using up-to-date analytical methods, that account for biases that may have been present in previous analyses. We also determine which species are characteristic for each selected Biogeographic District of the Cerrado using indicator species analysis (Dufrêne and Legendre 1997; De Cáceres et al. 2010). We verify climatic differences amongst the Biogeographic Districts, and finally, present a conservation assessment of each region in terms of land conversion and protected area coverage, to guide future conservation efforts in the Cerrado.

METHODS

Study area and database

We used floristic data from 588 inventories and floristic surveys distributed across the Cerrado. The biome is a geographic region delimited by IBGE (2004), which is largely covered by savanna vegetation, but also includes other major vegetation types such as grasslands and deciduous and evergreen riparian forests. We focused on cerrado *sensu lato*, which includes savanna vegetation and woodland or tall-savanna (*cerradão*), since they are floristically similar (Ribeiro and Walter 2008). We did not include deciduous, semi-deciduous, or gallery forests sites, because of sample gaps for these vegetation types, differences in sample methods and effort, and because the savanna cover almost 70% of the biome (Coutinho 2006). We also included some samples of savanna sites in the transition zones with adjacent biomes.

As few studies in our data compilation included all vascular plants, and most focused only on trees and large shrubs, we restricted our analyses to large woody species. We checked the scientific names, the species habits and distribution in the Flora do Brasil website (Flora do Brasil 2020 2016), which follows the APG IV taxonomy updates (APG IV 2016). We used the *flora* package (Carvalho 2017) in R to extract the species information. The final database includes 814 species, belonging to 77 plant families, with 202 species restricted to one site. Most of these unique samples are species more associated with other biomes or vegetation types, occurring only occasionally in savanna habitats. Thus, few unicates actually represent Cerrado-endemic species.

Analyses

Since different tools have been developed for different biogeographic approaches, there is a great variety of methods that can be used to identify biogeographic entities (see Morrone 2018). Considering various cluster methods, there are several options that can give divergent results (Leger et al. 2015). Among the most used methods, the k-means has shown good performance for biogeographic studies (Tichý et al. 2011; Vavrek 2016). For delimiting the Cerrado Biogeographic Districts (BDs), we performed a K-means cluster analysis, using a distance matrix. To compute the distance matrix, we excluded singletons, since they provide no information in similarity analysis (Magurran 1988).

We calculated the fuzzy matrix *a priori* in the fuzzySim package (Barbosa 2016) in R Statistical Software (R Development Core Team 2013). The fuzzy version of species' occurrence is a way to solve gaps and differences in sample methods, since the fuzzy logic searches for a probability of occurrence for each species per site (Barbosa 2015). The fuzzySim package provides three solutions for the fuzzy distribution: the prevalence-independent environmental favorability models produce a generalized linear model for each species using environmental variables. This approach was not used because many species did not have enough occurrences to run the GLM analysis. The second solution is the Spatial Trend Surface (TSA) model, which provides the spatial structure in species distribution by regressing occurrence data on the spatial coordinates. The third option is the Inverse Squared Distance to Presence (ISDP) for each species, which calculates a spatial interpolation model of the species' distribution. We tested the last two methods and compared the results with the original incidence matrix with mantel correlations. We used the ISDP matrix, which has greater correlation with the incidence matrix (ISDP $r=0.67$, $p<0.001$; TSA $r=0.56$, $p<0.001$). We calculated the *jaccard* distance of the ISDP matrix in the *vegan* package (Oksanen et al. 2014) in R.

We used the k-means method to cluster the sites using the *cascadekm* function (in the *vegan* package). In the k-means clustering, the observations are associated with the nearest mean point, according to the number of groups imposed. The cascade k-means creates several data partitions according to the required number of groups, where a range between the smallest and the largest number of groups is stated *a priori*. Considering our proposal to identify Biogeographic Districts (BD) in the Cerrado, the number of groups could neither be so many as to limit utility for conservation policies, nor so few, such that major differences in the spatially extensive and dynamic Cerrado would be not represented. Because of this, we restricted the possible number BDs to between two and 20 groups, inclusive. The number of groups can be chosen according to an SSI (Simple Structure Index) and "*calinski*" criteria. Both are good predictors for groups equal in size, but they may not be taken literally in differently sized groups (Oksanen et al. 2014). Thus, we explored both results considering the best values of each criteria, and the congruence between them, to select the best number of groups for our cluster.

To test the robustness of the groups in capturing vicariant patterns, we tested if the composition of Cerrado endemic species could explain the groups, using the ANOSIM test with 1000 permutation in the *vegan* package (Oksanen et al. 2014). The ANOSIM provides analysis of similarities for matrix data by permutations aiming to identify significant differences between groups. We also selected the endemic species that most explain the differences between the groups, by variable selection with Random Forest (described below), and verified the classification error rate.

To document the association between individual species and the BDs, we conducted an Indicator Species Analysis (ISA) (Dufrêne and Legendre 1997) using the *labdsv* package (Roberts 2013), with 100,000 randomizations. The ISA calculates how a species can be associated with one or more groups, and how statistically significant is the association. The index is based on the relative species' frequency or relative average abundance in clusters using a null model. Our data are presence/absence of species, and only the frequencies were considered. The indicator species value is greatest if all occurrences of the species are restricted to one single group, and if the species occurs in all sites of this group.

Many of the Cerrado tree species are widely distributed, being shared with one or more other biomes (Rizzini 1963; Heringer et al. 1977; Oliveira-Filho and Ratter 1995; Françoso et al. 2016). Those widely distributed species are important to the community composition in the savannas of the Cerrado. In our data, only 10% of the species are endemic to the Cerrado biome. Thus, we cannot ignore the role of widely distributed species in defining biogeographic patterns. We classified the indicator species according to their distribution across all Brazilian biomes, to understand in which BDs the endemic and shared species occur.

We initially examined climatic variation among the BDs. We used 35 bioclimatic variables based on precipitation, temperature, radiation, and moisture (Kriticos et al. 2012). These climatic variables are the mean interpolation of monthly data over a period of 30 to 50 years (reference year 2000) (Hijmans et al. 2004). For data reduction, we excluded some variables that were highly correlated with others (correlation greater than 0.70 or lower than -0.70), focusing on keeping those variables that were correlated with the greatest number of other variables. These surrogate variables are: annual mean temperature (°C), temperature seasonality (unitless coefficient of variation, or CV), temperature annual range (Bio05-Bio06) (°C), annual precipitation (mm), highest weekly radiation (W m^{-2}), lowest weekly radiation (W m^{-2}), radiation of coldest quarter (W m^{-2}), mean moisture index of coldest quarter.

To determine the best climatic variables to predict differences among the BDs, we used a variable selection with Random Forest in *varSelRF* package (Diza-Uriarte 2014), with 50,000 trees, and quantified the prediction error of the selected variables in *randomForest* package (Liaw and Wiener 2002). The Random Forest approach is a machine learning method that uses several decision trees with different random combinations of the explanatory variables and samples to make a robust variable selection. It is particularly amenable to datasets with many explanatory variables (Liaw and Wiener 2002).

We summarized all species occurrences by generating a matrix where each row was one BD. We observed the relationship among the BDs with the WARD hierarchical cluster method in the *recluster* package (Dapporto et al. 2013), generating the consensus tree with 100 re-samples, using the *jaccard* distance.

The map of the Biogeographic Districts (BDs) was drawn in a ArcGIS 10.2.1, with divisions among BDs set to correspond to known geographic features, where this was logical and feasible. These natural features usually limit the biogeographic areas (Morrone 2018). To assist in determining the boundaries between BDs, we used a digital elevation map (based on images of the Shuttle Radar Topography Mission; NGA and NASA 2000), a map of river catchments, and boundaries between states when they coincided with natural features, e.g. the “Serra Geral” mountain chain.

We quantified land conversion and the Protected Area (PA) coverage for each BD. We separated the PAs into Strict Protection (SP) and Sustainable Use (SU) groups, following the Brazilian legal definitions (Brasil 2000). The PA of SP correspond to IUCN I to III categories, and the PA of SU to categories IV to VI. We also quantified the Priority Conservation Areas (PCA, MMA 2016) for the BDs to understand further the conservation status of the Cerrado and discuss threats and conservation opportunities. We created the land conversion map for the Cerrado by quantifying the area that was converted during the period from 2010 to 2015, using natural vegetation distribution during 2010 as a baseline. We obtained all geographic data from http://mapas.mma.gov.br/i3geo/datadownload.htm.

RESULTS

The number of groups defined by the k-means varied based on selection criteria. The *calinski* criteria selected two, four, and eight groups, in that order, while the SSI selected nineteen, eighteen, twenty, and eight groups. Despite the difference between the two criteria, both did consider eight groups to be a good solution (Figure 1). Searching for a consensus solution, we selected eight as the best number of groups. The groups showed high spatial aggregation, with little overlap, which was crucial to spatially delimiting the Biogeographic Districts (Figure 2).

Most of the spatial boundaries defining the BDs followed landscape geomorphological attributes. We named the BDs based on their geographic position within the Cerrado biome: South (S), Southeast (SE), Southwest (SW), Central (Ce), West (We), Northwest (NW), and Northeast (NE). Only the External group (Ex) is spatially disaggregated, with samples in transition zones of south, north, and southwest. To separate the NE BD from the external group, we used a shape file of vegetation classes from IBGE (2004b), excluding the non-savannas classes, like evergreen and deciduous forest, scrub, and other transitional vegetation. Most of the external group sites are not within the limits of the Cerrado. In the hierarchical cluster, we found two main composition groups for the BDs (Figure 03). The first includes the northern and western BDs (NW, NE, CW, and SW), and the second includes the central and southern BDs (CE, SE, and S). The external group does not have a direct connection with either of these overarching groups. Thus, we did not consider this group in the further analysis, since most of its sites are not in the Cerrado biome, and it does not have a unique identity. In this way, we compared the seven Biogeographic Districts mentioned above, excluding the external group.

The ANOSIM results indicate significant differences in endemic species composition among the groups (R=0.304; p=0.001). In the Indicator Species Analysis, 394 species are significantly associated with at least one BD as presented in the Online Resource 1. The highest numbers of indicator species are in the S (109), NW (89), and

CE (73) BDs (Table 1). The BDs with the greatest number of endemic indicator species are CE and NW, with 19 and 15 endemic indicator species each. In the Random Forest selection, 39 endemic species were selected as the best for separating the groups (Table 2). The error rate in the confusion matrix was 22.6% (Online Resource 1). Most of these species are indicators in the CE and NW BDs.

The climatic variables selected as the best predictors of the compositional groups or BDs, based on the Random Forest analysis, were mean annual temperature, temperature seasonality, annual precipitation, highest weekly radiation, lowest weekly radiation, and radiation of the coldest quarter (Table 3). The classification rate was 4.8% (see confusion matrix in the Online Resource 1). Mean annual temperature plays an important role splitting the two main groups of BDs (CW, NE, NW, and SW versus CE, S, and SE) (Figure 4), which correspond to the groups found in the dendrogram (Figure 3).

Conservation status varies substantially across the BDs (Table 4; Figure 5). The conversion rate ranges from 19% in the SW to 90% in the S. The highest protected area coverage is in the CE BD (28.5%), in contrast with 2.7% in the SE BD, exemplifying the unbalanced conservation effort across the Cerrado. Not just the PA cover vary among the BDs, but they also vary inside the BDs according to the groups of SP and SU. The CE BD, for example, is covered by 26.6% of PA of SU and only by 1.9% of PA of SP. Priority Conservation Areas are greater than 23% in all the BDs, reaching 58% in the CE (Table 4; Figure 6).

Biogeographic District description

The Central (CE) Biogeographic District, with 24,411 km^2, occupies the central portion of the Cerrado biome, covering the Distrito Federal and neighbouring areas in

Goiás and Minas Gerais states (Figure 2). It occupies mainly the highlands of the Central Plateau, including the heads of the Tocantins, Corumbá and Preto rivers. Most of this area is over 900 m a.s.l. This BD has low annual mean temperature and low temperature seasonality, despite the high radiation rate of the coldest quarter, because of the marked dry season, when clouds are very rare. Seventy-three species are indicators of the CE BD, and it has the greatest number of endemic indicator species (19). Previous studies conducted by the Brazilian Ministry of the Environment suggested that 50.8% of this BD overlaps with extremely high PCA, and it is the BD with highest proportion of this PCA class within its limits. However, this is one of the most populated areas in the entire Cerrado region, and its coverage by Strict Protection UCs is low, with high land conversion rates.

The Central-west (CW) BD covers 417,983 km^2 in the northern portion of the state of Goiás and southern portion of the state of Mato Grosso. This large BD spans the watersheds of the Xingu, Araguaia, and part of Tocantins rivers, occupying a large area in the central and western portion of the Cerrado biome. It includes in its limits highland areas such as Chapada dos Veadeiros (over 1500m a.s.l.) and lowland areas along the Araguaia river and along the border with the Pantanal. This District has high temperatures with low seasonal variation. Radiation is also high during the dry season, which corresponds to the coldest quarter with respect to temperature in the Cerrado biome. It has only 21 indicator species, and most of them are widespread, occurring in more than two biomes (Table 1). Natural vegetation covers 48% of the CW BD, but only 6.2% of it is protected, with only 1.2% in PA of SP (Table 4).

The Northeast (NE) BD occupies the western parts of Bahia and Piauí and southern Maranhão, and northern Minas Gerais with and area of 403,248 km^2. The mean annual temperature is high and the annual precipitation is low. Seventy percent of its land is covered by natural vegetation, which suggests an opportunity to increase coverage by Protected Areas in this region. The current protected area coverage is 13.6%. Some important Protected Areas in the Cerrado are found in the NE BR, including the system of protected areas named *Veredas-Peruaçu*. This systems is composed by close or overlapping areas, which considers a management model in a regional context, named Mosaic of Protected Areas (MMA 2010). However, there is still 23.2% of land in the NE BD under Extremely High or Very High conservation priority. Furthermore, the most degraded Cerrado municipalities over the last years are placed in this BD, mainly along the western borders of the State of Bahia (MMA/IBAMA 2011).

The North West (NW) Biogeographic District covers mainly the state of Tocantins, spreading over 204,646 km^2. The mean annual temperature is extremely high, with very low seasonality *i.e.*, the temperature is high during all the year, as is the radiation (both highest weekly radiation and radiation of the coldest quarter). It has 89 indicator species, with 15 endemic and 14 shared with the Amazon biome. More than 70% of its area has natural vegetation. The percentage of PA coverage is the highest among the BDs (SU = 8.7%, SP = 6.7%), including an important portion of the *Jalapão* Mosaic. The Indigenous Territory coverage is also high (9.4%).

The South (S) Biogeographic District covers nearly all the Cerrados in São Paulo state, with 74,902 km^2. The mean annual temperature is the lowest among all BDs, and the seasonality is high, due to the proximity to the subtropical zone. The highest weekly radiation and the radiation of the coldest quarter are the lowest among the BDs. The number of indicator species is high (109), but most of them also occur in the Atlantic Forest (Table 1). The climatic particularities and the great influence of the Atlantic Forest make it a consistent natural division of Cerrado (Ratter et al. 2003). This unique vegetation is the most threatened among the BDs, with only 10% currently consisting of natural vegetation, and the PA of SP is less than 0.5%. The 23.4% extent of High and Very High conservation priority suggest important opportunities for protected area creation.

The Southeast (SE) Biogeographic District has 462,257 km^2, comprising most of the cerrado of Minas Gerais State and the Paraná River Basin in Goiás. The Espinhaço Mountain-Range is placed in the SE BD, presenting some of the highest elevation areas in the Cerrado. The mean annual temperature and the radiation parameters are average and the seasonality is high. Only 11 species are associated with this BD and most of them are endemic. The SE BD has been greatly transformed, with only 35% under natural cover. The PA coverage is less than 3%, and 20% of its area has Very High conservation priority.

The South-West (SW) Biogeographic District, with 321,068 km^2, comprises sites on the slopes that surround the flooding basin of the Pantanal, and other sites on mountain ranges within it. Interestingly, all localities within the Pantanal flooded basin were classified as SW BD, suggesting a strong resemblance between the Pantanal and the surrounding Cerrado in tree species composition. The mean annual temperature and the temperature seasonality are high, while the highest weekly radiation and the radiation of the coldest quarter are intermediate. The Amazon has an important influence on the SW BD. The floristic composition of this BD indicates great influence of seasonal forest species. Its selected indicator species are commonly found in seasonally dry tropical forests across the Cerrado (Nascimento et al. 2004; Salis et al. 2004; Santos et al. 2007; Kunz et al. 2008; Haidar et al. 2013). Despite the low coverage in PA (1.9%), The Indigenous Territories comprise 12.3% of this region.

DISCUSSION

We have identified seven Biogeographic Districts (BD) in the Cerrado, which are differentiated based on climatic conditions and species composition. These Biogeographic Districts are associated with particular landscapes within the geographic limits of the Cerrado biome, making them of special interest for conservation policies and management purposes. These areas harbor divergent plant communities and have different degrees of habitat loss and coverage by Protected Areas (PA). The use of large and continuous BDs, instead of the discrete endemism centers proposed for the Cerrado in previous studies, allows the formulation and planning of conservation efforts over a much wider region, covering also poorly sampled, but potentially relevant areas.

The patterns recovered in our study were partially observed by Ratter et al. (2003). Nevertheless, we found new Biogeographic Districts and refined delimitations of existing ones, thus representing an increase in the knowledge of distribution patterns of Cerrado woody species. This includes the CE BD, an interesting region placed in the Cerrado core area (Figure 2). Another important finding is the identification of hierarchical patterns in the species composition of woody plant communities in the Cerrado. We detected two main groups, distinguished by mean annual temperature values. We also detected important differences in the communities in transition zones, especially in the northern region of the Cerrado, in Piauí and Maranhão States. On the other hand, the sites inside the Pantanal clustered together with the SW BD, connecting the two portions of this BD. This finding suggests a strong relation between the vegetation of the Cerrado and Pantanal.

We found a high influence of neighboring biomes in all the BDs, particularly the influence of the Atlantic Forest on the S BD, and of the Amazon on the NW BD. Thus, the proximity of neighboring biomes is important to determining the potential of shared species. Nevertheless, other factors, like climate, may explain varying biome influence on the BDs, because their boundaries are dynamic. For example, shifts in vegetation distribution as a consequence of climatic fluctuations in savannas (Cole 1960) may have facilitated the exchange of species among the Brazilian biomes (Salgado-Labouriau 2005; Bueno et al. 2017), especially in ecotonal zones (Castro 1994). This situation may have driven a bidirectional colonization of species between the Cerrado and adjacent biomes (Oliveira-Filho and Ratter 1995; Colli 2005; Salgado-Labouriau 2005; Scariot and Sevilha 2005; Caetano et al. 2008; Ramos et al. 2009; Simon et al. 2009; Novaes et al. 2010), especially from the forest biomes into the Cerrado (Simon et al. 2011). This potential floristic exchange may have driven the influence of species characteristic of other biomes on the Cerrado flora (Rizzini 1963; Heringer et al. 1977; Castro et al. 1998). Nevertheless, and despite the large shared boundary between the Cerrado and Amazon, they share few indicator species, which was also reported in previous studies (Rizzini 1963; Heringer et al. 1977). The Amazon-Cerrado transition represents a complete turnover from savanna to forest communities, even over short distances (Pinto and Oliveira-Filho 1999; Marimon et al. 2006), and this scenario likely affects communities composition and the definition of BDs.

High elevation areas in the Cerrado are known for their high levels of endemism (Silva 1997; Simon and Proença 2000; Alves and Kolbek 2009; Echternacht et al. 2011; Nogueira et al. 2011; Gastauer et al. 2012). These high elevation areas are thought to be refuges for species that were formerly more widespread under past climatic conditions (Antonelli et al. 2010), especially those adapted to lower temperatures. These relictual populations are irreplaceable, bringing great importance to the SE BD. Each BD houses at least one area of endemism (Table 4), placed in highlands or valleys, which deserves special conservation attention.

The following BDs correspond to Ratter's floristic provinces (Ratter et al., 2003): NE (N & NE floristic province), SE (C & SE floristic province), and S (S floristic province). The floristic province Central-west was subdivided in BDs CW, NW, and SW. The CE BD is in the center of BDs and floristic provinces divisions. In Ratter's classification, the CE BD, combined with SE, is part of the C & SE floristic province. The herb–shrub flora grouping (Amaral et al., 2017) provided three main phytogeographic regions within the Cerrado. The phytogeographic region number 3 corresponds to BDs S, SE, and CE, and number 6 corresponds to the NE, NW, and partially CW. The SW BD is the combination of the phytogeographic regions 3 and 7, despite their wide coverage. The small divergences between the regionalization attempts may have arisen from differences in sampling methods and effort, scale, peculiarities of the groups, or methodological approach. Despite the limits of the regions are not identical to the BDs, we have a consistent pattern of plant community that brings confidence to use the BDs as the first layer for conservation policies. Comparisons with other taxonomic groups are also needed for confirm the importance of the BDs as a first layer biodiversity surrogates.

Since several patterns of species distribution, climate characteristics, habitat loss and protected areas coverage arise from BD identification and delimitation, we expect that these BD will be useful in future studies in the Cerrado focusing on biome biogeography or conservation approaches. The two rough groups of BDs, the colder BDs (CE, S and SE) and the hotter BDs (CW, NE, NW and SW), have experienced different patterns of land cover change, related mainly to historical processes in Cerrado colonization.

Colonization of the Cerrado has a main axis from South to North. Consequently, the Cerrado southern regions have experienced extensive land conversion, while the remaining land is poorly protected. New protected areas are urgently needed in these regions to preserve their unique biodiversity, despite the few current opportunities, and include the support for the creation of private reserves. In the northern regions of the Cerrado, given the larger amount of natural vegetation remaining, there is greater conservation opportunity, a plan for which can be defined by subsequent, more-detailed studies. Despite a greater extent of natural vegetation in the Northern region, and more conservation opportunities, the creation of new protected areas is still urgent in the region due to high pressure caused by the expansion of the agribusiness in the biome. The Brazilian Government defined the Northern part of the Cerrado, at the conjunction of the states of Maranhão, Tocantins, Piauí and Bahia (MATOPIBA as it is referred) as a priority region for agricultural occupation (José Roberto Borghetti et al. 2017) and, at present, no conservation strategy has been defined to ensure environmental safeguards for the region.

The remaining natural vegetation and protected area coverage are not evenly distributed across the Cerrado. The S biogeographic district is the least covered by protected areas and is the most impacted by land conversion. The NW biogeographic district is the least impacted, showing larger natural vegetation remnants and protected area coverage. This scenario reflects the south-to-north historical process of human occupation in Central Brazil (Diniz-Filho et al. 2009b). This reality imposes two extreme options for Cerrado conservation, which are different, but complementary, conservation strategies. In Biogeographic Districts of the Cerrado with more cover of natural areas (as NE, NW and SW), the proposition of new protected areas in IUCN groups I – III are urgent to preserve irreplaceable areas from the fast pace of the conversion of natural areas. Conversely, in the CE, S, and CW BDs, the best strategy is promoting the regeneration of natural Cerrado vegetation, including by direct seeding, (Pellizzaro et al. 2017), along with the creation of private reserves. The Brazilian Protected Areas in the category Private Reserves of the Natural Heritage (RPPNs) are an important tool for biodiversity conservation via the engagement of landowners in the challenge of nature conservation, and for ecotourism promotion (Silva et al. 2015). The management and conservation purposes of RPPNs are similar of those for National Parks (Brasil 2000), making this category very attractive for conservation efforts.

Between 1990-2010, the Cerrado lost 0.6% of its natural vegetation annually (Beuchle et al. 2015), primarily due to livestock and large-scale intensive agriculture (MMA 2015). This rate of habitat loss represents almost 1,700 ha per day, scattered across the Cerrado biome. At this pace of habitat loss, the creation of protected areas is urgently needed, involving all social actors and spheres of government. It is important to point out that almost the entire Cerrado biome is found within Brazil. Therefore, despite international concern on Cerrado conservation, the maintenance of this unique global biodiversity hotspot is a Brazilian responsibility (e.g. Strassburg et al. 2017).

More broadly, the total PA coverage of the Cerrado (8%) (Françoso et al. 2015) is well below the Aichi targets of the Convention on Biological Diversity, which is 17%. Even the NW, the most preserved BD, is not close to reaching this goal. On the other hand, all BDs except the S BD have more than 17% remaining natural vegetation (Table 4), making it possible to achieve a much larger Protected Area coverage, if conservation efforts increase in the Cerrado. In contrast, at present in Brazil, there seems to be an ongoing process of downsizing protected areas, degazettement, downgrading and reclassification (Bernard et al. 2014).

The Biogeographic Districts can be combined with other approaches for conservation prioritization in the Cerrado to focus on regional conservation needs, providing more realistic and important information for conservation prioritization, and bringing clearer goals for policy makers and for Protected Areas managers. Several approaches can contribute to conservation in the Cerrado and should take into acount the differences in biological communities highlighted herein. Current and future predictions of distribution, based on niche modelling of different taxonomic groups (Siqueira and Peterson 2003; Diniz-Filho 2004; Pinto et al. 2008; Marini et al. 2009; Costa et al. 2010), land conversion prediction modelling (Faleiro et al. 2013), and habitat fragmentation studies (Carvalho et al. 2009; Bianchi and Haig 2012), associated with Systematic Conservation Planning tools (Margules and Pressey 2000), can all contribute to an efficient protected areas system for biodiversity maintenance in the Cerrado. The Biogeographic Districts harbor different plant communities, that reflect differences in Cerrado biophysical and biological characteristics across its wide distribution, and we expect that these same characteristics can also shape ecological communities and biological interactions.

Characterization of Biogeographic Districts in other large tracts of natural habitats can be useful for the conservation of the world's savannas, which are nearly all strongly threatened biomes by human activities (Lima et al. 2018). Since climatic and compositional variation, as we reported here, are also expected to occur in other savannas worldwide (Lehmann et al. 2014), we expected that more detailed sub regions (BD) can be recovered and used as biodiversity surrogates for conservation planning, with the overarching aim to avoid biodiversity loss worldwide.

REFERENCES

Alves RJ V., Kolbek J (2009) Can campo rupestre vegetation be floristically delimited based on vascular plant genera? Plant Ecol 207:67–79. doi: 10.1007/s11258-009-9654-8

Amaral AG, Munhoz CBR, Walter BMT, et al (2017) Richness pattern and phytogeography of the Cerrado herb–shrub flora and implications for conservation. J Veg Sci 28:848–858. doi: 10.1111/jvs.12541

Antonelli A, Verola CF, Parisod C, Gustafsson a. LS (2010) Climate cooling promoted the expansion and radiation of a threatened group of South American orchids (Epidendroideae: Laeliinae). Biol J Linn Soc 100:597–607. doi: 10.1111/j.1095-8312.2010.01438.x

APG IV (2016) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot J Linn Soc 181:1–20

Azevedo JAR, Valdujo PH, de C. Nogueira C (2016) Biogeography of anurans and squamates in the Cerrado hotspot: coincident endemism patterns in the richest and most impacted savanna on the globe. J Biogeogr 43:2454–2464. doi: 10.1111/jbi.12803

Barbosa AM (2015) fuzzySim: Applying fuzzy logic to binary similarity indices in ecology. Methods Ecol Evol 6:853–858. doi: 10.1111/2041-210X.12372

Barbosa AM (2016) fuzzySim: Fuzzy similarity in species distributions. https://cran.r-project.org/web/packages

Bernard E, Penna LAO, Araújo E (2014) Downgrading, downsizing, degazettement, and reclassification of protected areas in Brazil. Conserv Biol 28:939–950. doi: 10.1111/cobi.12298

Beuchle R, Grecchi RC, Shimabukuro YE, et al (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. doi: 10.1016/j.apgeog.2015.01.017

Bianchi CA, Haig SM (2012) Deforestation Trends of Tropical Dry Forests in Central Brazil. Biotropica 45:395–400

Brasil (2000) Lei Nº 9.985. DOU

Brooks TM, Mittermeier R a, da Fonseca G a B, et al (2006) Global biodiversity conservation priorities. Science 313:58–61. doi: 10.1126/science.1127609

Bueno ML, Pennington RT, Dexter KG, et al (2017) Effects of Quaternary climatic fluctuations on the distribution of Neotropical savanna tree species. Ecography (Cop) 40:403–414. doi: 10.1111/ecog.01860

Caetano S, Prado DE, Pennington RT, et al (2008) The history of Seasonally Dry Tropical Forests in eastern South America: inferences from the genetic structure of the tree Astronium urundeuva (Anacardiaceae). Mol Ecol 17:3147–59. doi: 10.1111/j.1365-294X.2008.03817.x

Carmignotto AP, Vivo M De, Langguth A (2012) Mammals of the Cerrado and Caatinga: distribution patterns of the tropical open biomes of central South America. Bones, clones biomes Hist Geogr Recent Neotrop Mamm 307–350. doi: 10.1017/CBO9781107415324.004

Carvalho G (2017) flora: Tools for interacting with the Brazilian Flora 2020. https://cran.r-project.org/web/packages 0.3.0:

Carvalho FM V, Marco P De, Ferreira LG (2009) The Cerrado into-pieces: Habitat fragmentation as a function of landscape use in the savannas of central Brazil. Biol Conserv 142:1392–1403. doi: 10.1016/j.biocon.2009.01.031

Castro AAJF (1994) Comparação florístico-geográfica (Brasil) e fitossociologia (Piauíi - São Paulo) de amostras de cerrado. Instituto de Biologia. Thesis. Universidade Estadual de Campinas

Castro AAJF, Martins FR, Fernandes AG (1998) The woody flora of cerrado vegetation in the state of Piaui, northeastern Brazil. Edinburgh J Bot 55:455–472. doi: 10.1017/S0960428600003292

Clements FE, Shelford VE (1939) Bioecology, First edit. John Wiley & Sons / Chapman & Hall, New York, USA/London,UK

Cole MM (1960) Cerrado, Caatinga and Pantanal: The Distribution and Origin of the Savanna Vegetation of Brazil. Geogr J 126:168–179

Colli GR (2005) As origens e a diversificação da herpetofauna do Cerrado. In: Scariot A, Sousa-silva JC, Felfili JM (eds) Cerrado: Ecologia, Biodiversidade e Conservação, 1st edn. MMA, Brasília, DF

Costa GC, Nogueira CC, Machado RB, Colli GR (2010) Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. Biodivers Conserv 19:883–899. doi: 10.1007/s10531-009-9746-8

Coutinho LM (2006) O conceito de bioma. Acta Bot Brasilica 20:13–23. doi: 10.1590/S0102-33062006000100002

Dapporto L, Ramazzotti M, Fattorini S, et al (2013) Recluster: An unbiased clustering procedure for beta-diversity turnover. Ecography (Cop) 36:1070–1075. doi: 10.1111/j.1600-0587.2013.00444.x

De Cáceres M, Legendre P, Moretti M (2010) Improving indicator species analysis by combining groups of sites. Oikos 119:1674–1684. doi: 10.1111/j.1600-0706.2010.18334.x

de Mello PLH, Machado RB, Nogueira C de C (2015) Conserving Biogeography: Habitat Loss and Vicariant Patterns in Endemic Squamates of the Cerrado Hotspot. PLoS One 10:e0133995. doi: 10.1371/journal.pone.0133995

Dice LR (1943) The Biotic Provinces of North America. University of Michigan Press

Diniz-Filho JAF (2004) Phylogenetic Diversity and Conservation Priorities under Distinct Models of Phenotypic Evolution. Conserv Biol 18:698–704. doi:

10.1111/j.1523-1739.2004.00260.x

Diniz-Filho JAF, Bini LM, Oliveira G, et al (2009a) Macroecologia, biogeografia e áreas prioritárias para conservação no cerrado. Oecologia Bras 13:470–497. doi: 10.4257/oeco.2009.1303.05

Diniz-Filho JAF, Bini LM, Terribile LC, et al (2008) Conservation planning: a macroecological approach using the endemic terrestrial vertebrates of the Brazilian Cerrado. Fauna Flora Int Oryx 42:567–577. doi: 10.1017/S0030605308001129

Diniz-Filho JAF, Oliveira G De, Lobo F, et al (2009b) Agriculture, habitat loss and spatial patterns of human occupation in a biodiversity hotspot. Sci Agric 66:764–771

Diza-Uriarte R (2014) Variable Selection using Random Forest. https://cran.r-project.org/web/packages 23

Dufrêne M, Legendre P (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecol Monogr 67:345–366

Echternacht L, Trovó M, Oliveira CT, et al (2011) Areas of endemism in the Espinhaço Range in Minas Gerais, Brazil. Flora - Morphol Distrib Funct Ecol Plants 206:782–791. doi: 10.1016/j.flora.2011.04.003

Faleiro F V., Machado RB, Loyola RD (2013) Defining spatial conservation priorities in the face of land-use and climate change. Biol Conserv 158:248–257. doi: 10.1016/j.biocon.2012.09.020

FAO (2015) Global Forest Resources Assessment 2015, 1st edn. FAO, Rome, IT

Ferreira ME, Ferreira LG, Miziara F, Soares-Filho BS (2012) Modeling landscape dynamics in the central Brazilian savanna biome: future scenarios and perspectives for conservation. J Land Use Sci 8:403–421. doi: 10.1080/1747423X.2012.675363

Flora do Brasil 2020 (2016) Jardim Botânico do Rio de Janeiro.

http://floradobrasil.jbrj.gov.br. Accessed 5 Jan 2015

Françoso RD, Brandão R, Nogueira CC, et al (2015) Habitat loss and the effectiveness of protected areas in the Cerrado Biodiversity Hotspot. Nat Conserv 13:35–40. doi: 10.1016/j.ncon.2015.04.001

Françoso RD, Haidar RF, Machado RB (2016) Tree species of South America central savanna: endemism, marginal areas and the relationship with other biomes. Acta Bot Brasilica 30:1–9. doi: 10.1590/0102-33062015abb0244

Gastauer M, Teixeira Braga Messias MC, Alves Meira Neto JA (2012) Floristic Composition, Species Richness and Diversity of Campo Rupestre Vegetation from the Itacolomi State Park, Minas Gerais, Brazil. Environ Nat Resour Res 2:115–130. doi: 10.5539/enrr.v2n3p115

Grace J, José JS, Meir P, et al (2006) Productivity and carbon fluxes of tropical savannas. J Biogeogr 33:387–400. doi: 10.1111/j.1365-2699.2005.01448.x

Groves CR, Jensen DB, Valutis LL, et al (2002) Planning for Biodiversity Conservation: Putting Conservation Science into Practice. Bioscience 52:499–512. doi: 10.1641/0006-3568(2002)052[0499:PFBCPC]2.0.CO;2

Haidar RF, Felfili JM, Pinto JRR, et al (2013) Florestas estacionais e áreas de ecótono no estado do Tocantins, Brasil: parâmetros estruturais, classificação das fitofisionomias florestais e subsídios para conservação. Acta Amaz 43:261–290

Heringer EP, Barroso GM, Rizzo JA, Rizzini CT (1977) A flora do Cerrado. In: Ferri MG (ed) IV Simpósio sobre o Cerrado. Editora da Universidade de São Paulo, São Paulo, SP, pp 211–232

Hijmans RJ, Cameron S, Parra J, et al (2004) The WorldClim interpolated global terrestrial climate surfaces. http://www.worldclim.org. Accessed 10 Apr 2014

IBGE (2004a) Mapa de Biomas do Brasil. https://downloads.ibge.gov.br/. doi:

www.ibge.gov.br/home/geociencias/default_prod.shtm#USO

IBGE (2004b) Mapa de vegetação do Brasil. https://downloads.ibge.gov.br/downloads_geociencias.htm

Jenkins M (2003) Prospects for Biodiversity. Science (80-) 302:1175–1177. doi: 10.1126/science.1088666

José Roberto Borghetti, Washington L. C. Silva, Helder Rafael Nocko LN, Loyola GKC (2017) Agricultura Irrigada Sustentável no Brasil : Identificação de Áreas Prioritárias. FAO, Brasília

Klink CA, Machado RB (2005) Conservation of the Brazilian Cerrado. Conserv Biol 19:707–713. doi: 10.1111/j.1523-1739.2005.00702.x

Kreft H, Jetz W (2010) A framework for delineating biogeographical regions based on species distributions. J Biogeogr 37:2029–2053. doi: 10.1111/j.1365-2699.2010.02375.x

Kriticos DJ, Webber BL, Leriche A, et al (2012) CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. Methods Ecol Evol 3:53–64. doi: 10.1111/j.2041-210X.2011.00134.x

Kunz SH, Ivanauskas NM, Martins SV, et al (2008) Aspectos florísticos e fitossociológicos de um trecho de Floresta Estacional Perenifólia na Fazenda Trairão, Bacia do rio das Pacas, Querência-MT. Acta Amaz 38:245–254. doi: 10.1590/S0044-59672008000200007

Leger J-B, Daudin J-J, Vacher C (2015) Clustering methods differ in their ability to detect patterns in ecological networks. Methods Ecol Evol 6:474–481. doi: 10.1111/2041-210X.12334

Lehmann CER, Anderson TM, Sankaran M, et al (2014) Savanna Vegetation-Fire-Climate Relationships Differ Among Continents. Science (80-) 343:548–552. doi:

10.1126/science.1247355

Liaw a, Wiener M (2002) Classification and Regression by randomForest. R news 2:18–22. doi: 10.1177/154405910408300516

Lima DO de, Lorini ML, Vieira MV (2018) Conservation of grasslands and savannas: A meta-analysis on mammalian responses to anthropogenic disturbance. J Nat Conserv 45:72–78. doi: 10.1016/j.jnc.2018.08.008

Magurran AE (1988) Ecological diversity and its measurement. Springer, Dordrecht

Margules CR, Pressey RL (2000) Systematic conservation planning. Nature 405:243–53. doi: 10.1038/35012251

Marimon BS, De S. Lima E, Duarte TG, et al (2006) Observations on the Vegetation of Northeastern Mato Grosso, Brazil. Iv. an Analysis of the Cerrado–Amazonian Forest Ecotone. Edinburgh J Bot 63:323. doi: 10.1017/S0960428606000576

Marini MA, Barbet-Massin M, Lopes LE, Jiguet F (2009) Predicted climate-driven bird distribution changes and forecasted conservation conflicts in a neotropical savanna. Conserv Biol 23:1558–67. doi: 10.1111/j.1523-1739.2009.01258.x

Mews HA, Pinto JRR, Eisenlohr P V, Lenza E (2016) No evidence of intrinsic spatial processes driving Neotropical savanna vegetation on different substrates. Biotropica 48:433–442. doi: 10.1111/btp.12313

MMA/IBAMA (2011) Monitoramento do desmatamento nos biomas brasileiros por satélite: monitoramento do bioma Cerrado 2009-2010. Brasília, DF

MMA (2017) Plano de ação para a prevenção e o controle do desmatamento. Brasília, DF

MMA (2015) Mapeamento do Uso e Cobertura do Cerrado: Projeto TerraClass Cerrado 2013. Proj. TerraClass Cerrado 67

MMA (2016) Áreas Prioritárias para Conservação dos biomas Cerrado, Pantanal e

Caatinga. DOU Portaria:81
MMA (2010) Portaria nº 482, de 14 de Dezembro de 2010. Brasília, DF
Morrone JJ (2018) The spectre of biogeographical regionalization. J Biogeogr 45:282–288. doi: 10.1111/jbi.13135
Morrone JJ (2014) Biogeographical regionalisation of the neotropical region. Zootaxa 3782:1–110. doi: 10.11646/zootaxa.3782.1.1
Morrone JJ, Crisci J V, Plata MD La, et al (1995) Historical Biogeography: Introduction to Methods. Annu. Rev. Ecol. Evol. Syst. 26:373–401
Morrone JJ, Url S (1994) On the Identification of Areas of Endemism. Syst Biol 43:438–441
Myers N, Mittermeier RA, Mittermeier CG, et al (2000) Biodiversity hotspots for conservation priorities. Nature 403:853–858. doi: 10.1038/35002501
Nascimento ART, Felfili JM, Meirelles M (2004) Florística e estrutura da comunidade arbórea de um remanescente de Floresta Estacional Decidual de encosta, Monte Alegre, GO, Brasil. Acta Bot Brasilica 18:659–669
Neves DM, Dexter KG, Pennington RT, et al (2015) Environmental and historical controls of floristic composition across the South American Dry Diagonal. J Biogeogr 42:1566–1576. doi: 10.1111/jbi.12529
NGA, NASA (2000) Shuttle Radar Topography Mission
Nogueira C, Ribeiro S, Costa GC, Colli GR (2011) Vicariance and endemism in a Neotropical savanna hotspot: distribution patterns of Cerrado squamate reptiles. J Biogeogr 38:1907–1922. doi: 10.1111/j.1365-2699.2011.02538.x
Novaes RML, Filho JP de L, Ribeiro RA (2010) Phylogeography of Plathymenia reticulata (Leguminosae) reveals patterns of recent range expansion towards northeastern Brazil and southern Cerrados in Eastern Tropical South America. Mol

Ecol 19:985–998. doi: 10.1111/j.1365-294X.2010.04530.x

Oksanen J, Blanchet FG, Kindt R, et al (2014) The vegan Package Version 1.15-0. https://cran.r-project.org/web/packages

Oliveira-Filho AT, Ratter JA (1995) A study of the origin of central Brazilian forests by the analysis of plant species distribution patterns. Edinburgh J Bot 52:141. doi: 10.1017/S0960428600000949

Pellizzaro KF, Cordeiro AOO, Alves M, et al (2017) "Cerrado" restoration by direct seeding: field establishment and initial growth of 75 trees, shrubs and grass species. Rev Bras Bot 40:681–693. doi: 10.1007/s40415-017-0371-6

Pinto JRR, Oliveira-Filho AT (1999) Perfil florístico e estrutura da comunidade arbórea de uma floresta de vale no Parque Nacional da Chapada dos Guimarães, Mato Grosso, Brasil. Rev Bras Botânica 22:53–67. doi: 10.1590/S0100-84041999000100008

Pinto MP, Diniz-Filho JAF, Bini LM, et al (2008) Biodiversity surrogate groups and conservation priority areas : birds of the Brazilian Cerrado. Divers Distrib 14:78–86. doi: 10.1111/j.1472-4642.2007.00421.x

R Development Core Team (2013) R: A language and environment for statistical computing. http://www.r-project.org. Accessed 2 Jul 2013

Ramos ACS, Lemos-Filho JP, Lovato MB (2009) Phylogeographical Structure of the Neotropical Forest Tree Hymenaea courbaril (Leguminosae : Caesalpinioideae) and Its Relationship with the Vicariant Hymenaea stigonocarpa from Cerrado. J Hered 100:206–216. doi: 10.1093/jhered/esn092

Ratter JA, Bridgewater S, Atkinson R, Ribeiro JF (1996) Analysis of the floristic composition of the Brazilian cerrado vegetation II: Comparison of the woody vegetation of 98 areas. Edinburgh J Bot 53:153. doi: 10.1017/S0960428600002821

Ratter JA, Bridgewater S, Ribeiro JF (2003) Analysis of the Floristic Composition of the Brazilian Cerrado Vegetation III: Comparison of the Woody Vegetation of 376 Areas. Edinburgh J Bot 60:57–109. doi: 10.1017/S0960428603000064

Ratter JA, Bridgewater S, Ribeiro JF, et al (2011) Analysis of the floristic composition of the Brazilian cerrado vegetation IV: Presentation of a Revised Data-Base of 367 Areas. In: Conserv. Manag. Biodivers. Cerrado Biome. http://cerrado.rbge.org.uk/cerrado/download/download.php

Ratter JA, Dargie TCD (1992) An analysis of the floristic composition of 26 cerrado ares in Brazil. Edinburgh J Bot 49:235. doi: 10.1017/S0960428600001608

Ribeiro JF, Walter BMT (2008) As principais fitofisionomias do bioma Cerrado. In: Sano SM, Almeida SP, Ribeiro JF (eds) Cerrado: ecologia e flora. EMBRAPA, Brasília, DF, pp 151–212

Rizzini CT (1963) A flora do Cerrado. In: Simpósio sobre o Cerrado. Editora da Universidade de São Paulo, São Paulo, SP

Roberts DW (2013) Package ‘ labdsv .’ https://cran.r-project.org/web/packages 1–56. doi: 10.1021/es103092a.See

Salgado-Labouriau ML (2005) Alguns aspectos sobre a Paleoecologia dos Cerrados. In: Scariot A, Sousa-silva JC, Felfili JM (eds) Cerrado: Ecologia, Biodiversidade e Conservação. MMA, Brasília, DF

Salis SM, Pereira M, Silva DA, et al (2004) Fitossociologia de remanescentes de floresta estacional decidual em Corumbá , Estado do Mato Grosso do Sul , Brasil. Rev Bras Botânica 27:671–684

Sanmartín I (2012) Historical Biogeography: Evolution in Time and Space. Evol Educ Outreach 5:555–568. doi: 10.1007/s12052-012-0421-2

Santos RM dos, Vieira F de A, Gusmão E, Nunes YRF (2007) Florística e estrutura de uma floresta estacional decidual no Parque Municipal da Sapucaia, Montes Claros (MG). Cerne 13:248–256

Scariot AO, Sevilha AC (2005) Biodiversidade, estrutura e conservação de florestas estacionais deciduais no Cerrado. In: Scariot A, Sousa-Silva JC, Felfili JM (eds) Cerrado: ecologia, biodiversidade e conservação. MMA

Silva DCB, Segalerba MDB, Brandão RA (2015) A representatividade das reservas particulares do patrimônio natural (RPPN) no entorno do Parque Nacional da Chapada dos Veadeiros, estado de Goiás, Brasil. Heringeriana 10:64–78

Silva JMC (1997) Endemic bird species and conservation in the Cerrado Region , South America. Biodivers Conserv 6:435–450

Silva JMC, Bates JM (2002) Biogeographic Patterns and Conservation in the South American Cerrado: A Tropical Savanna Hotspot. Bioscience 52:225–234. doi: 10.1641/0006-3568(2002)052

Simon MF, Grether R, de Queiroz LP, et al (2009) Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. Proc Natl Acad Sci U S A 106:20359–64. doi: 10.1073/pnas.0903410106

Simon MF, Grether R, de Queiroz LP, et al (2011) The evolutionary history of Mimosa (Leguminosae): toward a phylogeny of the sensitive plants. Am J Bot 98:1201–21. doi: 10.3732/ajb.1000520

Simon MF, Proença C (2000) Phytogeographic patterns of Mimosa (Mimosoideae, Leguminosae) in the Cerrado biome of Brazil: an indicator genus of high-altitude centers of endemism? Biol Conserv 96:279–296. doi: 10.1016/S0006-3207(00)00085-9

Siqueira MF, Peterson AT (2003) Consequences of global climate change for geographic distributions of Cerrado tree species. Biota Neotrop 3:1–14. doi:

10.1590/S1519-566X2006000600003

Strassburg BBN, Brooks T, Feltran-Barbieri R, et al (2017) Moment of truth for the Cerrado hotspot. Nat Ecol Evol 1:1–3. doi: 10.1038/s41559-017-0099

Szumik CA, Goloboff PA (2004) Areas of endemism: an improved optimality criterion. Syst Biol 53:968–77. doi: 10.1080/10635150490888859

Tichý L, Chytrý M, Šmarda P (2011) Evaluating the stability of the classification of community data. Ecography (Cop) 34:807–813. doi: 10.1111/j.1600-0587.2010.06599.x

Udvardy MDF (1975) A classification of the biogeographical provinces of the world. IUCN, Morges, Switzerland

Vavrek MJ (2016) A comparison of clustering methods for biogeography with fossil datasets. PeerJ 4:e1720. doi: 10.7717/peerj.1720

Wallace AR (1876) The geographical distribution of animals; with a study of the relations of living and extinct faunas as elucidating the past changes of the Earth's surface. London

Whittaker RJ, Araujo MB, Paul J, et al (2005) Conservation Biogeography: assessment and prospect. Divers <html_ent glyph="@amp;" ascii="&"/> Distrib 11:3–23. doi: 10.1111/j.1366-9516.2005.00143.x

TABLES

Table 1. Number of indicator species significantly associated with the Biogeographic Districts of the Cerrado (Central – CE, Central-west - CW, North-east - NE, North-west - NW, South - S, South-east - SE, and South-west - SE) and their distribution in the Brazilian biomes. The widely distributed species occur in more than two biomes. Only the significant indicator species were counted (See the Online Resource for the indicator species analysis result).

Distribution	CE	CW	NE	NW	S	SE	SW	Total
Cerrado endemic	19	3	3	15	7	9	2	58
Cerrado and Pantanal	1	0	0	0	0	0	2	3
Cerrado and Amazon	9	6	2	14	6	4	8	49
Cerrado and Caatinga	7	1	4	5	0	0	0	17
Cerrado and Atlantic Forest	12	0	0	3	41	4	6	66
Widely	25	11	9	52	55	11	38	201
Total	73	21	18	89	109	28	56	394

Table 2. Importance of endemic species for the delimitation of the Biogeographic Districts of the Cerrado (Central – CE, Central-west - CW, North-east - NE, North-west - NW, South - S, South-east - SE, and South-west - SE). MDA=Mean Decrease Accuracy.

Species	**BD**	**MDA**	**CE**	**CW**	**NE**	**NW**	**S**	**SE**	**SW**
Aspidosperma tomentosum Mart.	CE	0.015	0.012	0.019	0.021	0.020	0.005	0.007	0.019
Dalbergia miscolobium Benth.	CE	0.013	0.005	0.006	0.003	0.013	0.006	0.024	0.034
Eremanthus glomerulatus Less.	CE	0.019	0.076	0.004	0.015	0.017	0.023	0.014	0.011
Eriotheca pubescens (Mart. & Zucc.) Schott & Endl.	CE	0.015	0.040	-0.001	0.025	0.008	0.024	0.014	0.012
Erythroxylum tortuosum Mart.	CE	0.025	0.011	-0.001	0.071	0.009	0.011	0.037	0.047
Guapira noxia (Netto) Lundell	CE	0.030	0.068	0.004	0.086	0.018	0.017	0.020	0.031
Kielmeyera speciosa A.St.-Hil.	CE	0.008	0.026	0.000	0.013	0.005	0.012	0.006	0.005
Ouratea hexasperma (A.St.-Hil.) Baill.	CE	0.037	0.038	0.010	-0.004	0.027	0.171	0.023	0.029
Salacia crassifolia (Mart. ex Schult.) G.Don	CE	0.039	0.116	0.012	0.010	0.053	0.065	0.021	0.049
Styrax ferrugineus Nees & Mart.	CE	0.034	0.189	0.003	0.025	0.027	0.044	0.017	0.014
Tachigali subvelutina (Benth.) Oliveira-Filho	CE	0.038	0.060	0.011	0.035	0.028	0.099	0.017	0.059
Vochysia thyrsoidea Pohl	CE	0.030	0.189	0.009	0.022	0.018	0.026	0.008	0.015
Kielmeyera rubriflora Cambess.	CW	0.036	0.024	0.083	0.050	0.035	0.006	0.012	0.020
Vochysia rufa Mart.	CW	0.019	-0.005	0.015	0.016	0.008	0.071	0.007	0.031
Vochysia gardneri Warm.	NE	0.015	0.010	0.004	0.051	0.012	0.009	0.013	0.013
Aspidosperma nobile Müll.Arg.	NW	0.029	0.026	0.019	0.039	0.027	0.040	0.033	0.019
Callisthene hassleri Briq.	NW	0.004	0.001	0.000	0.002	0.020	0.001	0.001	0.000
Caryocar coriaceum Wittm.	NW	0.026	0.011	0.010	0.017	0.101	0.012	0.016	0.015
Davilla elliptica A.St.-Hil.	NW	0.015	0.002	0.015	-0.002	0.024	0.016	0.022	0.021

Diospyros coccolobifolia Mart. ex Miq.	NW	0.011	0.007	0.000	0.000	0.053	0.004	0.004	0.005
Diospyros hispida A.DC.	NW	0.009	0.004	0.002	-0.004	0.023	0.006	0.021	0.006
Heteropterys byrsonimifolia A.Juss.	NW	0.013	0.009	0.004	-0.001	0.039	0.004	0.011	0.026
Mouriri elliptica Mart.	NW	0.039	0.070	0.011	0.008	0.037	0.080	0.064	0.020
Pseudobombax longiflorum (Mart.) A.Robyns	NW	0.022	0.001	0.015	0.059	0.033	0.013	0.024	-0.001
Pseudobombax tomentosum (Mart.) A.Robyns	NW	0.021	0.003	0.015	0.025	0.009	0.039	0.011	0.050
Tachigali aurea Tul.	NW	0.012	0.001	0.007	-0.010	0.027	0.019	0.023	0.005
Bauhinia rufa (Bong.) Steud.	S	0.011	0.003	-0.001	0.017	0.004	0.038	0.012	0.011
Leptolobium elegans Vogel	S	0.055	0.031	0.035	0.039	0.038	0.206	0.020	0.051
Miconia paucidens DC.	S	0.003	0.001	0.001	0.001	0.001	0.019	0.001	0.001
Ouratea spectabilis (Mart.) Engl.	S	0.043	0.024	0.005	0.030	0.012	0.216	0.014	0.050
Mimosa laticifera Rizzini & A.Mattos	SE	0.004	0.001	0.005	0.005	0.000	0.003	0.008	0.003
Callisthene mollissima Warm.	-	0.002	0.002	0.003	0.001	0.004	0.000	0.001	0.000
Lafoensia pacari A.St.-Hil.	-	0.008	-0.004	0.003	0.023	0.016	0.003	0.005	0.007
Pleroma stenocarpa (Schrank et Mart. ex DC.) Triana	-	0.003	0.000	0.001	0.001	0.001	0.014	0.001	0.002

Table 3. Biogeographic Districts' total area, remaining natural vegetation, protected area coverage, and Priority Conservation Areas. Conservation effort was measured for protected areas of sustainable use, strict protection, and indigenous territory. All areas are in km^2. The proposed Biogeographic Districts of the Cerrado biome are the Central (CE), Central-west (CW), North-east (NE), North-west (NW), North-west (NE), South (S), South-east, and South-west (SW).

BD	Total area	Conv. rate	Protected Areas						Priority Conservation Areas					
			SU		SP		IT		High		Very high		Extremely high	
CE	24,411	63%	6491	26.6%	467.6	1.9%	0	0.0%	0	0.0%	1854	7.6%	12408	50.8%
CW	417,983	52%	20941	5.0%	5064.2	1.2%	17739	4.2%	10471	2.5%	113911	27.3%	36533	8.7%
NE	403,248	30%	24500	6.1%	19110.5	4.7%	11175	2.8%	29868	7.4%	43715	10.8%	50182	12.4%
NW	240,646	29%	20904	8.7%	16140.9	6.7%	22621	9.4%	28399	11.8%	38761	16.1%	27786	11.5%
S	74,902	90%	6366	8.5%	232.4	0.3%	16	0.0%	7601	10.1%	9963	13.3%	101	0.1%
SE	469,257	65%	4758	1.0%	7822.2	1.7%	0	0.0%	38281	8.2%	93860	20.0%	31324	6.7%
SW	321,068	19%	2652	0.8%	3656.7	1.1%	39461	12.3%	15260	4.8%	38352	11.9%	37728	11.8%

Table 4. Biogeographic units (areas of endemism or biotic elements) within the Biogeographic Districts (BDs) of the Cerrado found in previous studies. The BDs are Central (CE), Central-west (CW), North-east (NE), North-west (NW), North-west (NE), South (S), South-east, and South-west (SW). The biogeographic units are named according to the original source.

Reference	**Biological group**	**CE**	**CW**	**NE**	**NW**	**S**	**SE**	**SW**
Azevedo et al., 2016	Anurans and squamates	Central plateau	Veadeiros; Guimarães; Caiapônia	Serra Geral; Chapada das Mesas	Tocantins-Araguaia; Jalapão		Espinhaço Canastra	Parecis; Pantanal-Bodoquena ; Paraná plateau
Simon and Proença, 2000	Species in the genus *Mimosa*	Central plateau	Veadeiros; Guimarães				Espinhaço	
Nogueira et al., 2011	Squamate		Guimarães	Serra Geral	Tocantins depression; Upper Tocantins plateaus	Tietê-Rio Grande	Espinhaço	Serra das Araras; Parecis
de Melo et al., 2015	Squamate	Central plateau	Guimarães-Roncador	Serra Geral	Araguaia		Espinhaço	Paraná-Paraguai; Paraguai-Guaporé
Silva and Bates, 2002	Birds		Paranã		Araguaia		Espinhaço	

FIGURE LEGENDS

Figure 1. *Calinski* and SSI (Simple Structure Index) criteria for selection of the optimal number of groups in k-means cluster *jaccard* distance of a fuzzy distribution matrix. The values of each criterion are standardized as z values. The calinski is high for low number of groups and SSI selected more groups, but provided support for a classification involving eight groups.

Figure 2. Biogeographic Districts of the Cerrado biome (Brazil) based on k-means classification of *jaccard* distance. The distance matrix is based on the fuzzy surface of tree communities. The polygons were based on the distribution of sites in the same group in Fig. 1. The seven regions are: Central (CE), Central-west (CW), North-east (NE), North-west (NW), North-west (NE), South (S), South-east, and South-west (SW). The external group in gray was not considered a Biogeographic District due its massive occurrence outside of the Cerrado biome and lack of a coherent geographic identity.

Figure 3. Consensus tree of the Cerrado's Biogeographic Districts of the Cerrado biome. The seven regions are: Central (CE), Central-west (CW), North-east (NE), North-west (NW), North-west (NE), South (S), South-east, South-west (SW), the external group (Ex).

Figure 4. Boxplots showing the bioclimatic variables selected by Random Forest to distinguish each Biogeographic District of the Cerrado biome. Equal letters indicate no significant differences.

Figure 5. Remaining natural vegetation (light green), Protected Areas of Strict Protection (dark green), and Protected Areas of Sustainable Use (brown) in the Biogeographic Districts Central (CE), Central-west (CW), North-east (NE), North-west (NW), North-west (NE), South (S), South-east, and South-west (SW) of the Cerrado biome.

Figure 6. The Brazilian official Priority Conservation Areas (PCA) (in red) over the remaining natural vegetation (light green), in the Biogeographic Districts Central (CE), Central-west (CW), North-east (NE), North-west (NW), North-west (NE), South (S), South-east, and South-west (SW) of the Cerrado biome. The shades of red (light to dark) follow the priority high, very high, and extremely high.