

Biodiversity and Conservation

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

--Manuscript Draft--

Manuscript Number:	BIOC-D-18-00138R3	
Full Title:	Delimiting floristic biogeographic districts in the Cerrado and assessing their conservation status	
Article Type:	S.I. : CERRADO	
Keywords:	Neotropical savanna; phytogeography; Indicator species; Brazilian savanna; Biogeographic Regionalization.	
Corresponding Author:	Renata Françoso Universidade de Brasília Instituto de Ciencias Biologicas Brasília, DF BRAZIL	
Corresponding Author Secondary Information:		
Corresponding Author's Institution:	Universidade de Brasília Instituto de Ciencias Biologicas	
Corresponding Author's Secondary Institution:		
First Author:	Renata Françoso	
First Author Secondary Information:		
Order of Authors:	Renata Françoso	
	Kyle G. Dexter	
	Ricardo B. Machado	
	R. Toby Pennington	
	José Roberto R. Pinto	
	Reuber A. Brandão	
	James A. Ratter	
Order of Authors Secondary Information:		
Funding Information:	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (4893/13-1)	Mrs. Renata Françoso
	Conselho Nacional de Desenvolvimento Científico e Tecnológico (303838/2016-8)	Mr. Ricardo B. Machado
	Natural Environment Research Council (NE/I028122/1)	Mr. Kyle G. Dexter Mr. R. Toby Pennington
	Conselho Nacional de Desenvolvimento Científico e Tecnológico (307701/2014-0)	Mr. José Roberto R. Pinto
Abstract:	<p>The Cerrado is a biodiversity hotspot in central Brazil that represents the largest expanse of savanna in the Neotropics. Here, we aim to identify and delimit biogeographic districts 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 districts, we clustered sites based on their similarity in tree species composition. To investigate why districts differ in composition, we 1) determined the proportion of tree species in different districts that derive from other biomes, to assess the influence of neighbouring biomes upon geographically marginal districts and 2) assayed key climatic differences between districts, to test the effect of environmental factors upon compositional differences. We found seven biogeographic districts within the Cerrado.</p>	

	<p>Marginal districts have a large proportion of tree species characteristic of Amazonia and Atlantic Forest, but the Cerrado endemic species are also important. Further, districts differed significantly for multiple climatic variables. Finally, to provide a preliminary conservation assessment of the different districts, we assessed their rate of land conversion and current coverage by protected areas. We found that districts 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 districts merit greater conservation and management attention.</p>
<p>Response to Reviewers:</p>	<p>Dr. David L. Hawksworth Editor-in-Chief Biodiversity and Conservation</p> <p>I am pleased to inform you that all the corrections were made. We were requested to exclude the author citations after scientific names in Table 2. We exclude both, in table 2 and in the online resource.</p> <p>Thank you for your attention. Please do not hesitate to contact us if you have any question.</p> <p>kind regards, Renata Françoso</p>

[Click here to view linked References](#)

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

3 Renata D. Françoso, Kyle G. Dexter, Ricardo B. Machado, R. Toby Pennington, José

4 R.R. Pinto, Reuber A. Brandão, James A. Ratter.

5 Françoso, R.D. (Corresponding author, arenatafrancoso@gmail.com; +55 61 3107-
6 5645; ORCID ID 0000-0002-6369-8317)¹

7 Dexter, K.G. (kyle.dexter@ed.ac.uk; ORCID ID 0000-0001-9232-5221)^{2,3}

8 Machado, R.B. (rbmac@unb.br; ORCID ID 0000.0002.6508.9005)⁴

9 Pennington, R.T. (t.pennington@exeter.ac.uk; ORCID ID 0000-0002-8196-288X)^{2,5}

10 Pinto, J.R.R. (jrrpinto@unb.br; ORCID ID 0000-0003-2028-6176)⁶

11 Brandão, R.A. (reuberbrandao@gmail.com; ORCID ID 0000-0003-3940-2544)⁶

12 Ratter, J.A. (j.ratter@rbge.ac.uk)²

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

15 ² Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh EH3 5LR, U.K.

16 ³ School of GeoSciences, University of Edinburgh, Edinburgh EH9 3FF, U.K.

17 ⁴ Zoology Department, Biological Science Institute, University of Brasília. Address:
18 Campus Universitário Darcy Ribeiro, Brasília, DF, Brasil

19 ⁵ Geography, University of Exeter, Exeter EX4 4RJ, U.K.

20 ⁶ Department of Forest Engineer, Faculty of Technology, University of Brasília.

21 Address: Campus Universitário Darcy Ribeiro, Brasília, DF, Brasil.

22 ABSTRACT

23 The Cerrado is a biodiversity hotspot in central Brazil that represents the largest
24 expanse of savanna in the Neotropics. Here, we aim to identify and delimit
25 biogeographic districts within the Cerrado, to provide a geographic framework for
26 conservation planning and scientific research prioritisation. We used data from 588 sites
27 with tree species inventories distributed across the entire Cerrado. To identify districts,
28 we clustered sites based on their similarity in tree species composition. To investigate
29 why districts differ in composition, we 1) determined the proportion of tree species in
30 different districts that derive from other biomes, to assess the influence of neighbouring
31 biomes upon geographically marginal districts and 2) assayed key climatic differences
32 between districts, to test the effect of environmental factors upon compositional
33 differences. We found seven biogeographic districts within the Cerrado. Marginal
34 districts have a large proportion of tree species characteristic of Amazonia and Atlantic
35 Forest, but the Cerrado endemic species are also important. Further, districts differed
36 significantly for multiple climatic variables. Finally, to provide a preliminary
37 conservation assessment of the different districts, we assessed their rate of land
38 conversion and current coverage by protected areas. We found that districts in the south
39 and southwest of the Cerrado have experienced the greatest land conversion and are the
40 least protected, while those in the north and northeast are less impacted and better
41 protected. Overall, our results show how biogeographic analyses can contribute to
42 conservation planning by giving clear guidelines on which districts merit greater
43 conservation and management attention.

44

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

47

48 **ACKNOWLEDGMENTS**

49 R.D.F. thanks the Coordination of Improvement of Higher Level Personnel (CAPES)
50 for the 6-month study period under the Science Without Borders Programme (Process
51 4893/13-1). R.B.M. (Process 303838/2016-8) and J.R.R.P. (Process: 307701/2014-0)
52 received a research fellowship grant from National Council for Scientific and
53 Technological Development (CNPq). K.G.D. and R.T.P. were supported by the Natural
54 Environmental Research Council (grant NE/I028122/1). K.G.D. was supported by a
55 Leverhulme Trust International Academic Fellowship.

56 INTRODUCTION

57 Human activity has affected natural resources to such a high level that it has
58 generated a global biodiversity crisis (Jenkins 2003; Maxwell et al. 2016). Biodiversity
59 threats are distributed unevenly across the globe (Brooks et al. 2006), with developing
60 countries in the tropics currently representing the most vulnerable regions (FAO 2015).
61 Land conversion will persist into the next decades due to agricultural expansion and
62 intensification, especially in South America and sub-Saharan African (Jenkins 2003),
63 affecting mainly tropical savannas (Grace et al. 2006). Brazil is one of the top four
64 countries in South America in terms of predicted habitat loss (FAO 2015), which is
65 concentrated in the Brazilian Cerrado (MMA/IBAMA 2011), a global biodiversity
66 hotspot (Myers et al. 2000). Several thousand hectares of natural vegetation are
67 converted every year in the Cerrado, at rates higher than observed in the Amazon
68 (MMA 2017).

69 Despite the biological importance of the Cerrado, which originally covered more
70 than 2 million km², nearly 50% of its natural vegetation has been cleared, chiefly due to
71 agricultural expansion (MMA 2015). This continuous and intensive conversion is not
72 randomly distributed, but prevalent in some geographic regions and vegetation types
73 (Bianchi and Haig 2012). For example, land conversion has tended to follow the
74 implementation of roads and other infrastructure, which took place first in the south of
75 the Cerrado. Further, additional large declines of the Cerrado vegetation have been
76 predicted over the next 50 years (Ferreira et al. 2012), especially in tableland areas with
77 open vegetation formations, which are more suitable for the establishment of
78 mechanized agriculture. By 2030, we may expect natural vegetation to be found mostly
79 in protected areas (Klink and Machado 2005). Currently, only 3% of the remaining
80 natural vegetation in the Cerrado is maintained in areas of strict protection equivalent to

81 the IUCN categories I to III (Françoso et al. 2015). Regional variation in species
82 composition and the non-uniform human occupation of the Cerrado implies the need for
83 specifically tailored conservation policies, based on regional planning. However,
84 conservation efforts in the Cerrado have not followed any clear plan, with protected
85 areas being established opportunistically on a case-by-case basis (Françoso et al. 2015).
86 Among nine described global approaches to conservation prioritization (Brooks et al.
87 2006), the Cerrado represents a reactive conservation scenario, with decisions based on
88 threat, contrasting with Amazonia where decisions are often based on opportunity.

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

95 Biogeographic regionalization aims to represent distinct biological natural areas
96 on a map (Morrone 2018), which can support conservation policies and scientific
97 investigations. The identification of homogeneous natural areas, based on animal and
98 plant communities, at regional, continental or global scales, is a common approach in
99 ecology and biogeography (e.g. Wallace 1876; Clements and Shelford 1939; Dice 1943;
100 Udvardy 1975). To unify the nomenclature used for floral and faunal biogeographic
101 regions, Udvardy (1975) proposed a hierarchical division with realms, biotic provinces
102 and districts. Realms occur at continental scales and follow the large faunal regions of
103 Wallace (1876). Provinces are subdivisions of realms, comprising large subcontinental
104 regions, characterized by the major biome that occupies the area. The third
105 biogeographical level, the district, encompasses smaller differences within provinces.

106 Districts are essential to drive conservation efforts, since they represent unique features
107 of the provinces (Udvardy 1975). Higher or lower levels, such as regions or dominions,
108 may also be used (Morrone 2014).

109 The identification of biogeographic units in a large and threatened ecosystem,
110 such as the Cerrado, is necessary for recognizing distinct biological communities with
111 different conservation needs, and to subsequently adjust conservation actions for
112 different parts of the biome. Several studies have been conducted to identify
113 conservation priority areas in the Cerrado. These have used different approaches, such
114 as the distribution of endemic species (Simon and Proença 2000; Silva and Bates 2002;
115 Diniz-Filho et al. 2008; Nogueira et al. 2011; Carmignotto et al. 2012; Azevedo et al.
116 2016), the identification of vicariant processes (de Mello et al. 2015), macroecology
117 (Diniz-Filho et al. 2008, 2009) or species community composition (Ratter and Dargie
118 1992; Ratter et al. 1996, 2003; Aguiar et al. 2015; Amaral et al. 2017). The Cerrado
119 biome harbours three to five main areas of endemism, depending on the studied group.
120 These areas (the Central Plateau, Veadeiros Mountain Range, Guimarães Mountain
121 Range, Espinhaço Mountain Range, and Araguaia Valley) have been recorded in studies
122 based on distribution patterns of vertebrates (Diniz-Filho et al. 2008), birds (Silva and
123 Bates 2002), herpetofauna (Nogueira et al. 2011; de Mello et al. 2015; Azevedo et al.
124 2016) and species of *Mimosa* (Simon and Proença 2000).

125 Here we focus on the ecological approach of clustering localities based on
126 similarity in tree species composition because it is relevant for guiding conservation
127 planning and the design of protected areas networks (Whittaker et al. 2005; de Mello et
128 al. 2015; DRYFLOR 2016). Ecological biogeography often relies on cluster methods
129 for identifying patterns in the distribution of organisms across landscapes, and group
130 localities based on their similarities in species composition (Kreft and Jetz 2010). An

131 alternative approach of historical biogeography is to delimit areas of endemism, where
132 the distribution of two or more endemic taxa overlap (Morrone and Url 1994; Szumik
133 and Goloboff 2004). In this case, overlapping species distributions are assumed to result
134 from vicariant processes, such as tectonic-isolating events (Sanmartín 2012). We
135 consider these assumptions to be unreasonable in the Brazilian Cerrado due to its young
136 geological age (<10 MY; Simon et al. 2009) and because of evidence that neotropical
137 tree communities are assembled by dispersal (Pennington and Dick 2004; Dexter et al.
138 2017). In contrast with the historical biogeography approach, ecological biogeography
139 searches for patterns in the current distribution of organisms that are determined by
140 recent dispersal processes and environmental filters (Morrone et al. 1995).

141 Biogeographic studies based on community composition in the Cerrado show
142 large areas that are relatively homogeneous in species composition (Ratter and Dargie
143 1992; Ratter et al. 1996, 2003; Aguiar et al. 2015; Mews et al. 2016; Amaral et al.
144 2017). In a series of studies published from 1996 to 2003, Ratter and colleagues
145 proposed six Floristic Provinces within the core area of Cerrado, and another two
146 disjunct areas in the Amazon (Ratter and Dargie 1992; Ratter et al. 1996, 2003, 2011).
147 These studies were based on an extensive sampling effort for woody plants of the
148 Cerrado, including more than 900 species of trees and large shrubs, and representing the
149 most extensive botanical biogeographic study of the Cerrado to date.

150 Here, we aim to identify biogeographic districts within the Cerrado biome, based
151 on a large dataset for woody plants, primarily trees, and propose specific regions as the
152 first level of biodiversity surrogates for conservation planning in the Cerrado.

153 Therefore, we are not interested in areas of endemism *per se*, because we do not want to
154 neglect any part of the Cerrado, even if there are no regionally endemic species present.

155 We expand the woody plant floristic database of Ratter et al. (2003) from 376 to 588

156 sites and delimit biogeographic districts using up-to-date analytical methods, accounting
157 for biases that may have been present in previous analyses. We also determine which
158 species are characteristic for each district using indicator species analysis (Dufrêne and
159 Legendre 1997; De Cáceres et al. 2010). We verify climatic differences amongst the
160 biogeographic districts and, finally, present a conservation assessment of each district in
161 terms of land conversion and protected area coverage, to guide future conservation
162 efforts in the Cerrado.

163

164 **METHODS**

165 *Study area and database*

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

177 We restricted our analyses to trees and large shrubs (plants with a woody stem
178 that reaches 2 m tall or more), because few studies in our data compilation included
179 other vascular plants such as herbs. We checked the scientific names, habits and
180 distributions of the species in the Flora do Brasil website (Flora do Brasil 2016), which

181 follows the APG IV taxonomy updates (APG IV 2016). We used the *flora* package
182 (Carvalho 2017) in R to extract species information. The final database includes 814
183 species, belonging to 77 plant families, with 202 species restricted to one site. Most of
184 these unique samples are species more associated with other biomes or vegetation types,
185 occurring only occasionally in savanna habitats. Thus, few single-site occurrences
186 actually represent Cerrado endemics.

187

188 *Analyses*

189 Since different tools have been developed for different biogeographic
190 approaches, there is a great variety of methods that can be used to identify
191 biogeographic entities (see Morrone 2018). Considering various cluster methods, there
192 are several options that can give divergent results (Leger et al. 2015). Among the most
193 used methods, the k-means clustering has shown good performance in biogeographic
194 studies (Tichý et al. 2011; Vavrek 2016). For delimiting Cerrado biogeographic
195 districts, we performed a k-means cluster analysis excluding singletons, since they
196 provide no information in similarity analysis (Magurran 1988).

197 We calculated a fuzzy version of the Jaccard similarity index using the *fuzzySim*
198 (Barbosa 2016) and the *vegan* (Oksanen et al. 2014) packages in R (R Core Team
199 2016). This involved two steps: (1) the calculation of a fuzzy version of the species
200 occurrence matrix and (2) the calculation of the Jaccard similarity matrix. The fuzzy
201 version of species occurrences is a way to solve gaps and differences in sample
202 methods, since the fuzzy logic searches for a probability of occurrence for each species
203 per site (Barbosa 2015). The *fuzzySim* package provides three solutions for the fuzzy
204 distribution: the prevalence-independent environmental favourability model produces a
205 generalized linear model for each species using environmental variables. We avoided

206 this approach, because many species lacked enough occurrences to run the GLM
207 analysis. The second solution is the Spatial Trend Surface (TSA) model, which provides
208 the spatial structure in species distribution by regressing occurrence data on the spatial
209 coordinates. The third option is the Inverse Squared Distance to Presence (ISDP) for
210 each species, which calculates a spatial interpolation model of the species distributions.
211 We tested the last two methods and compared the results with the original incidence
212 matrix using Mantel correlations. We selected the ISDP matrix, which had greater
213 correlation with the incidence matrix (ISDP $r = 0.67$, $P < 0.001$; TSA $r = 0.56$, $P <$
214 0.001).

215 We implemented the k-means method using the *cascadekm* function in the *vegan*
216 package. In the k-means clustering, the observations are associated with the nearest
217 mean point according to the number of groups imposed. The cascade k-means creates
218 several data partitions according to the required number of groups, where a range
219 between the smallest and the largest number of groups is stated *a priori*. Considering
220 our proposal to identify biogeographic districts in the Cerrado, the number of groups
221 could neither be so large as to limit their utility for conservation policies, nor so few that
222 major differences in the spatially extensive and dynamic Cerrado would not be
223 represented. Therefore, we restricted the possible number districts to a range between
224 two and 20 groups. We used the simple structure (SSI) and the Calinski–Harabasz
225 indices to select the optimal number of groups. Both are good predictors when groups
226 are equal in size, but they may not be interpreted literally for differently sized groups
227 (Oksanen et al. 2014). Thus, we considered the best values of each criteria and their
228 congruence to select the best number of groups for our cluster.

229 To test the robustness of the groups in capturing vicariant patterns, we tested if
230 the composition of Cerrado endemic species, as a subset of the entire data per group,

231 could explain the groups, using the ANOSIM test with 1000 permutation in the *vegan*
232 package (Oksanen et al. 2014). The ANOSIM provides analysis of similarities for
233 matrix data by permutations aiming to identify significant differences between groups.
234 We also selected the endemic species that most explain the differences between the
235 groups, by variable selection with Random Forest (described below).

236 To document the association between individual species and the biogeographic
237 districts, we conducted an Indicator Species Analysis (ISA) (Dufrêne and Legendre
238 1997) using the *labdsv* package (Roberts 2013), with 100,000 randomizations. The ISA
239 calculates how a species can be associated with one or more groups, and how
240 statistically significant the association is. The index is based on species relative
241 frequencies or relative average abundances in clusters using a null model. Since our
242 dataset consists of species occurrences, only their frequencies were considered. The
243 indicator species value is greatest when all occurrences of the species are restricted to a
244 single group, and when the species occurs in all sites of this group.

245 In our dataset, only 10% of the species are endemic to the Cerrado, whereas
246 most tree species are widely distributed, being shared with one or more other biomes
247 (Rizzini 1963; Heringer et al. 1977; Oliveira-Filho and Ratter 1995; Françoso et al.
248 2016). These widely distributed species are important components of Cerrado
249 communities; thus, we cannot ignore their role in defining biogeographic patterns. We
250 classified indicator species according to their distribution across all Brazilian biomes, to
251 understand in which districts the endemic and shared species occur.

252 We initially examined climatic variation among the biogeographic districts,
253 using 35 bioclimatic variables based on precipitation, temperature, radiation, and
254 moisture (Kriticos et al. 2012). These climatic variables are the mean interpolation of
255 monthly data over a period of 30 to 50 years (reference year 2000) (Hijmans et al.

256 2004). For data reduction, we excluded some highly correlated variables (correlation
257 greater than 0.70 or lower than -0.70), keeping those that were correlated with the
258 greatest number of other variables. In the end, we retained: mean annual temperature
259 (°C), temperature seasonality (unitless coefficient of variation, or CV), temperature
260 annual range (BIO 5 – BIO 6) (°C), mean annual precipitation (mm), highest weekly
261 radiation (Wm^{-2}), lowest weekly radiation (Wm^{-2}), radiation of coldest quarter (Wm^{-2})
262 and mean moisture index of coldest quarter.

263 To determine the best climatic predictors of biogeographic districts, we used
264 variable selection with Random Forest in the *varSelRF* package (Diza-Uriarte 2014),
265 with 50,000 trees. We evaluated the error of the variables by quantifying the number of
266 correct predictions in *randomForest* package (i.e. 'out-of-bag' error; Liaw and Wiener
267 2002). Random Forest is a machine learning method that uses several decision trees
268 with different random combinations of the explanatory variables and samples to make a
269 robust variable selection. It is particularly amenable to datasets with many explanatory
270 variables (Liaw and Wiener 2002).

271 We summarized all species occurrences by generating a matrix where each row
272 was one biogeographic district. We determined the relationships among districts using a
273 consensus tree of 100 resamples, each based on Ward's hierarchical cluster method and
274 the Jaccard distance, with the *recluster* package (Dapporto et al. 2013).

275 Biogeographic areas are often limited by natural features (Morrone 2018).
276 Therefore, we used ArcGIS 10.2.1 to produce a map of biogeographic districts, with
277 their boundaries corresponding to known geographic features when this was logical and
278 feasible. To assist in determining district boundaries, we used a digital elevation map
279 (based on images of the Shuttle Radar Topography Mission; NGA and NASA 2000), a

280 map of river catchments, and state boundaries when they coincided with natural
281 features, e.g. the “Serra Geral” mountain chain.

282 We quantified land conversion and the coverage of protected areas for each
283 biogeographic district. We separated protected areas into strict protection and
284 sustainable use, following the Brazilian legal definitions (Brasil 2000). Strict protection
285 areas correspond to IUCN categories I to III, and sustainable use to categories IV to VI.
286 We also quantified the overlap of districts with Priority Conservation Areas (MMA
287 2016), to further understand the conservation status of the Cerrado and discuss threats
288 and conservation opportunities. We created the land conversion map for the Cerrado by
289 quantifying the area that was converted during the period from 2010 to 2015, using
290 natural vegetation distribution during 2010 as a baseline. We obtained all geographic
291 data from <http://mapas.mma.gov.br/i3geo/datadownload.htm>.

292

293 **RESULTS**

294 The optimal solution for the k-means clustering varied with the selection criteria.
295 The Calinski–Harabasz index was highest for two, four, and eight groups, in that order,
296 while the simple structure index favoured nineteen, eighteen, twenty, and eight groups.
297 Despite the difference between the two criteria, both considered eight groups as a good
298 solution (Figure 1). Seeking a balanced solution, we chose eight as the optimal number
299 of groups. The eight groups showed high spatial aggregation, with little overlap, which
300 was crucial to spatial delimitation of the biogeographic districts (Figure 2).

301 Most of the spatial boundaries defining the districts followed landscape
302 geomorphological attributes. We named districts based on their geographic position
303 within the Cerrado: South (S), Southeast (SE), Southwest (SW), Central (Ce), West
304 (We), Northwest (NW), and Northeast (NE). One of the groups is composed of sites in

305 transition zones to other biomes in the south, north, and southwest of the Cerrado.
306 Because these sites occurred outside of the Cerrado's boundary, we refer to this group
307 as the "Extralimital" group. To separate this group from the NE district, we used a
308 shapefile of vegetation classes from IBGE (2004b), excluding the non-savanna classes,
309 such as evergreen and deciduous forest, scrub, and other transitional vegetation. There
310 were two major groups of districts in the hierarchical cluster (Figure 3). The first
311 included the northern and western districts (NW, NE, CW, and SW), and the second
312 included the central and southern districts (CE, SE, and S). The Extralimital group does
313 not have a direct connection with either of these overarching groups. Thus, we excluded
314 it from further analyses, since most of its sites are not in the Cerrado biome, and it does
315 not have spatial cohesion.

316 The ANOSIM results indicated significant differences in endemic species
317 composition among the groups ($r = 0.304$; $P = 0.001$). In the Indicator Species Analysis,
318 394 species were significantly associated with at least one biogeographic district as
319 presented in the Online Resource 1. The highest numbers of indicator species are in the
320 S (109), NW (89), and CE (73) districts (Table 1). We found 14 species with average
321 frequency higher than 60% considering all the biogeographic districts (*Qualea*
322 *parviflora*, *Bowdichia virgilioides*, *Connarus suberosus*, *Hymenaea stigonocarpa*,
323 *Dimorphandra mollis*, *Byrsonima coccolobifolia*, *Handroanthus ochraceus*, *Pouteria*
324 *ramiflora*, *Kielmeyera coriacea*, *Erythroxylum suberosum*, *Roupala montana*, *Tocoyena*
325 *formosa*, *Diospyros hispida*, *Tabebuia aurea*, *Caryocar brasiliense*, and *Davilla*
326 *elliptica*). The districts with the greatest number of endemic indicator species are CE
327 and NW, with 19 and 15 endemic indicator species each. In the Random Forest
328 selection, 39 endemic species were selected as the best predictors of the districts (Table

329 2). The out-of-bag error rate was 22.6% (Online Resource 1). Most of these species are
330 indicators in the CE and NW districts.

331 The best climatic predictors of the districts, based on the Random Forest
332 analysis, were mean annual temperature, temperature seasonality, mean annual
333 precipitation, highest weekly radiation, lowest weekly radiation, and radiation of the
334 coldest quarter. The out-of-bag error rate was 4.8% (see confusion matrix in the Online
335 Resource 1). Mean annual temperature plays an important role splitting the two main
336 groups of districts found in the dendrogram (CW, NE, NW, and SW versus CE, S, and
337 SE) (Figures 3-4). Each districts is different from the others for at least two climatic
338 parameters (Table 3).

339 Conservation status varies substantially across the biogeographic districts (Table
340 4, Figure 5). The conversion rate ranges from 19% in the SW to 90% in the S. The
341 highest protected area coverage is in CE (28.5%), in contrast with 2.7% in SE,
342 highlighting the unbalanced conservation effort across the Cerrado. The protected areas
343 coverage of strict protection and sustainable use varies not only among the districts, but
344 also within districts. The CE district, for example, is covered by 26.6% of sustainable
345 use but only by 1.9% of strict protection areas. Priority Conservation Areas cover more
346 than 23% of all districts, reaching 58% in the CE (Table 4, Figure 6).

347

348 **Description of Biogeographic districts**

349 The Central (CE) district occupies 24,411 km² of the central portion of the
350 Cerrado biome, covering the Distrito Federal (Federal District) and neighbouring areas
351 in Goiás and Minas Gerais states. It mainly occupies the highlands of the Central
352 Plateau, including the headwaters of the Tocantins, Corumbá and Preto rivers. Most of
353 this area is over 900 m a.s.l. This district has low mean annual temperature and low

354 temperature seasonality, despite the high radiation rate of the coldest quarter, which is
355 because of the marked dry season, when clouds are very scarce. CE has 73 indicator
356 species and the greatest number of endemic indicator species (19). Previous studies
357 conducted by the Brazilian Ministry of the Environment suggested that 50.8% of this
358 district overlaps with extremely high Priority Conservation Areas (MMA 2016), and it
359 is the district with the highest proportion of this class within its limits. However, this is
360 one of the most populated areas in the entire Cerrado region, and its coverage by strict
361 protection areas is low, with high land conversion rates.

362 The Central-west (CW) district covers 417,983 km² in the north of Goiás and
363 southern Mato Grosso states. This large district spans the watersheds of the Xingu,
364 Araguaia, and part of the Tocantins rivers, occupying a large area in the central and
365 western portion of the Cerrado. It includes highland areas such as Chapada dos
366 Veadeiros (over 1,500m a.s.l.) and lowland areas along the Araguaia river and along the
367 border with the Pantanal. This district has high temperatures with low seasonal
368 variation. Radiation is also high during the dry season, which corresponds to the coldest
369 quarter in the Cerrado. It has only 21 indicator species, and most of them are
370 widespread, occurring in more than two biomes. In CW the natural vegetation covers
371 48%, but only 6.2% is protected, of which 1.2% is in the strict category.

372 The Northeast (NE) district occupies 403,248 km², covering western Bahia and
373 Piauí, southern Maranhão, and northern Minas Gerais states. The mean annual
374 temperature is high, and the annual precipitation is low. Natural vegetation covered 70%
375 of this BD, and the current protected area coverage is 13.6%. Some important protected
376 areas in the Cerrado are found in the NE district, including the *Veredas-Peruaçu* system
377 of protected areas. However, 23.2% of the NE district is under Extremely High or Very
378 High conservation priority. Furthermore, the Cerrado municipalities that have suffered

379 most degradation over the last years are placed in this district, mainly along the western
380 borders of the state of Bahia (MMA/IBAMA 2011).

381 The North West (NW) biogeographic district includes mainly the state of
382 Tocantins, covering over 204,646 km². The mean annual temperature is high, with low
383 seasonality *i.e.*, the temperature is high year-round, as is the radiation (both highest
384 weekly radiation and radiation of the coldest quarter). It has 89 indicator species, with
385 15 endemic and 14 shared with the Amazon biome. More than 70% of its area has
386 natural vegetation. The percentage of protected area coverage is the highest among the
387 districts (sustainable use: 8.7%, strict protection: 6.7%), including an important portion
388 of the *Jalapão* mosaic of protected areas.

389 The South (S) biogeographic district covers nearly all the cerrado in São Paulo
390 state, with 74,902 km². The mean annual temperature is the lowest among all districts,
391 and the temperature seasonality is high, due to the proximity to the subtropical zone.
392 The highest weekly radiation and the radiation of the coldest quarter are the lowest
393 among the districts. The number of indicator species is high (109), but most of them
394 also occur in the Atlantic Forest. This unique vegetation is the most threatened among
395 the districts, with only 10% currently consisting of natural vegetation, and the strict
396 protection area represents less than 0.5%. The 23.4% extent of High and Very High
397 conservation priority suggests important opportunities for protected area creation.

398 The Southeast (SE) biogeographic district has 462,257 km², comprising most of
399 the Cerrado in Minas Gerais state and the Paran river basin in Gois. The Espinhao
400 mountain range is placed in this district, presenting some of the highest elevation areas
401 in the Cerrado. The mean annual temperature and radiation parameters are average, and
402 seasonality is high, when compared with the other districts. Only 11 species are
403 associated with this district and most of them are endemic. The SE district has been

404 greatly transformed, with only 35% under natural cover. The protected area coverage is
405 less than 3%, and 20% of its area has Very High conservation priority.

406 The South-West (SW) biogeographic district comprises 321,068 km² on the
407 slopes that surround the flooding basin of the Pantanal, and other sites on mountain
408 ranges within it. All localities within the Pantanal flooded basin were classified as SW,
409 suggesting a strong resemblance between the Pantanal and the surrounding Cerrado in
410 tree species composition. Mean annual temperature and temperature seasonality are
411 high, while the highest weekly radiation and the radiation of the coldest quarter are
412 intermediate in relation to the other districts. Amazonia has an important influence on
413 the SW district. Its floristic composition indicates great influence of seasonal forest
414 species and its selected indicator species are commonly found in seasonally dry tropical
415 forests across the Cerrado (Nascimento et al. 2004; Salis et al. 2004; Santos et al. 2007;
416 Kunz et al. 2008; Haidar et al. 2013). Despite the low coverage in protected areas
417 (1.9%), indigenous lands cover 12.3% of this region.

418

419 **DISCUSSION**

420 We have identified seven biogeographic districts in the Cerrado, which are
421 differentiated based on climatic conditions and species composition. These districts are
422 associated with particular landscapes within the geographic limits of the Cerrado,
423 making them of special interest for conservation policies and management purposes.
424 These areas harbour plant communities divergent in their species composition and have
425 different degrees of habitat loss and coverage by protected areas. The use of large and
426 continuous districts, instead of the discrete endemism centres proposed for the Cerrado
427 in previous studies, allows the formulation and planning of conservation efforts over a
428 much wider region, covering also poorly sampled, but potentially relevant areas.

429 The patterns recovered in our study were partially observed by Ratter et al.
430 (2003). Nevertheless, we found new biogeographic districts and refined delimitations of
431 existing ones, thus representing an increase in the knowledge of distribution patterns of
432 Cerrado woody species. This includes the Central district, which is placed in the
433 Cerrado core area (Figure 2). Another important finding is the identification of
434 hierarchical patterns in the woody plant communities in the Cerrado. We detected two
435 main groups, distinguished by mean annual temperature values. We also detected
436 differences in the communities of transition zones, especially in the northern region of
437 the Cerrado, in Piauí and Maranhão states. The climatic particularities and the great
438 influence of the Atlantic Forest make the S biogeographic district a consistent natural
439 division of Cerrado (Ratter et al. 2003). On the other hand, the sites inside the Pantanal
440 clustered with the SW district, suggesting a strong relation between the vegetation of
441 the Cerrado and Pantanal.

442 We found a high influence of neighbouring biomes in all districts, particularly
443 the influence of the Atlantic Forest on the S and of Amazonia on the NW district. Thus,
444 the proximity of neighbouring biomes is important to determining the potential of
445 shared species. Nevertheless, other factors, like climate, may explain varying biome
446 influence on the districts, because their boundaries are dynamic (Werneck et al. 2012).
447 For example, shifts in vegetation distribution as a consequence of climatic fluctuations
448 in savannas (Cole 1960) may have facilitated the exchange of species among the
449 Brazilian biomes (Salgado-Labouriau 2005; Bueno et al. 2017), especially in ecotonal
450 zones. This situation may have driven a bidirectional colonization of species between
451 the Cerrado and adjacent biomes (Oliveira-Filho and Ratter 1995; Colli 2005; Salgado-
452 Labouriau 2005; Scariot and Sevilha 2005; Caetano et al. 2008; Ramos et al. 2009;
453 Simon et al. 2009; Novaes et al. 2010), especially from the forest biomes into the

454 Cerrado (Simon et al. 2011). This potential floristic exchange may have driven the
455 influence of species characteristic of other biomes on the Cerrado flora (Rizzini 1963;
456 Heringer et al. 1977). Nevertheless, and despite the large shared boundary between the
457 Cerrado and Amazonia, they share few indicator species, which was also reported in
458 previous studies (Rizzini 1963; Heringer et al. 1977). The Amazonia-Cerrado transition
459 represents a complete turnover from savanna to forest communities, even over short
460 distances (Pinto and Oliveira-Filho 1999; Marimon et al. 2006), and this scenario likely
461 affects community composition and the definition of biogeographic districts.

462 High elevation areas in the Cerrado are known for their high levels of endemism
463 (Silva 1997; Simon and Proença 2000; Alves and Kolbek 2009; Echternacht et al. 2011;
464 Nogueira et al. 2011; Gastauer et al. 2012). These high elevation areas are thought to be
465 refuges for species that were formerly more widespread under past climatic conditions
466 (Antonelli et al. 2010), especially those adapted to lower temperatures. These relictual
467 populations are irreplaceable, bringing great importance to the SE district. Each district
468 houses at least one area of endemism (Table 5), placed in highlands or valleys, which
469 deserves special conservation attention.

470 The following districts correspond to Ratter's floristic provinces (Ratter et al.,
471 2003): NE (N & NE floristic province), SE (C & SE floristic province), and S (S
472 floristic province). The Central-west floristic province was subdivided into districts
473 CW, NW, and SW. In Ratter's classification, the CE and SE district are in the C & SE
474 floristic province. An analysis of the herb-shrub flora of the Cerrado (Amaral et al.,
475 2017) suggested three main phytogeographic regions. Their phytogeographic region
476 number 3 corresponds to the S, SE, and CE districts, and number 6 corresponds to the
477 NE, NW, and partially CW. The SW district is the combination of phytogeographic
478 regions 3 and 7, despite their wide coverage. The small divergences between the

479 regionalization attempts may have arisen from differences in sampling methods and
480 effort, scale, and peculiarities of the different life forms studied. Despite the fact that the
481 limits of their regions are not entirely identical to our biogeographic districts, there is a
482 sufficiently consistent geographic pattern of plant community composition to give
483 confidence to using the districts as the first layer for conservation policies. Comparisons
484 with other taxonomic groups are also needed to confirm the importance and generality
485 of the districts we identified here.

486 Since several patterns of species distribution, climate characteristics, habitat loss
487 and protected area coverage arise from the identification and delimitation of
488 biogeographic districts, we expect that they will be useful in future studies in the
489 Cerrado focusing on biogeography or conservation. The two groups of districts, cold
490 (CE, S and SE) and hot (CW, NE, NW and SW) districts, have experienced different
491 patterns of land cover change, mainly related to historical processes in Cerrado
492 colonization.

493 Colonization of the Cerrado has had a main axis from south to north.
494 Consequently, southern regions of the Cerrado have experienced extensive land
495 conversion, and the remaining natural vegetation cover there is poorly protected. New
496 protected areas are urgently needed in these regions to preserve their unique
497 biodiversity, and these should include support for the creation of private reserves. In the
498 northern Cerrado, given the larger amount of natural vegetation remaining, there is
499 greater conservation opportunity, a plan for which can be defined by subsequent, more-
500 detailed studies. Despite this, the creation of new protected areas is still urgent in the
501 region due to high pressure caused by the expansion of agribusiness. The Brazilian
502 Government defined the northern part of the Cerrado, at the conjunction of the states of
503 Maranhão, Tocantins, Piauí and Bahia (“MATOPIBA”) as a priority region for

504 agricultural occupation (Borghetti et al. 2017) and, at present, no conservation strategy
505 has been defined to ensure environmental safeguards there.

506 The remaining natural vegetation and protected areas are not evenly distributed
507 across the Cerrado. The S district is the least covered by protected areas and is the most
508 impacted by land conversion. The NW district is the least impacted, showing larger
509 natural vegetation remnants and better protected area coverage. This reality imposes two
510 extreme options for Cerrado conservation, which are different, but complementary. In
511 districts with more cover of natural areas (as NE, NW and SW), the proposition of new
512 protected areas in IUCN groups I – III are urgent to preserve irreplaceable areas from
513 the fast pace of the conversion of natural areas. Conversely, in the CE, S, and CW
514 districts, the best strategy is promoting the natural regeneration of degraded Cerrado
515 areas, including direct seeding, (Pellizzaro et al. 2017), along with the creation of
516 private reserves. The Brazilian Protected Areas in the category Private Reserves of the
517 Natural Heritage (RPPNs) are an important tool for biodiversity conservation via the
518 engagement of landowners in the challenge of nature conservation, and for ecotourism
519 promotion (Silva et al. 2015). The management and conservation purposes of RPPNs
520 are similar of those for National Parks (Brasil 2000), making this category very
521 attractive for conservation efforts.

522 Between 1990-2010, the Cerrado lost 0.6% of its natural vegetation annually
523 (Beuchle et al. 2015), primarily due to livestock and large-scale intensive agriculture
524 (MMA 2015). This rate of habitat loss represents almost 1,700 ha per day, scattered
525 across the Cerrado. At this pace of habitat loss, the creation of protected areas is urgent,
526 involving all social actors and spheres of government. It is important to point out that
527 almost the entire Cerrado biome is found within Brazil. Therefore, despite international

528 concern for Cerrado conservation, the maintenance of this unique global biodiversity
529 hotspot is a Brazilian responsibility (e.g. Strassburg et al. 2017).

530 More broadly, the total protected area coverage of the Cerrado (8%) (Françoso et
531 al. 2015) is well below the Aichi target of the Convention on Biological Diversity,
532 which is 17%. Even the NW, the less impacted biogeographic district, is not close to
533 reaching this goal. On the other hand, all districts except the S have more than 17%
534 remaining natural vegetation (Table 4), making it possible to achieve much larger
535 protected area coverage. Sadly, there currently seems to be an ongoing process of
536 downsizing, degazettement, downgrading and reclassification of protected areas in
537 Brazil (Bernard et al. 2014).

538 The biogeographic districts can be combined with other approaches for
539 conservation prioritization in the Cerrado to focus on regional conservation needs,
540 providing more realistic and important information for conservation prioritization, and
541 bringing clearer goals for policy makers and for protected area managers. Several
542 approaches can contribute to conservation in the Cerrado and should take into account
543 the differences in biological communities highlighted herein. Current and future
544 predictions of distribution, based on niche modelling of different taxonomic groups
545 (Siqueira and Peterson 2003; Diniz-Filho 2004; Pinto et al. 2008; Marini et al. 2009;
546 Costa et al. 2010), land conversion prediction modelling (Faleiro et al. 2013), and
547 habitat fragmentation studies (Carvalho et al. 2009; Bianchi and Haig 2012), associated
548 with systematic conservation planning tools (Margules and Pressey 2000), can all
549 contribute to an efficient protected areas system for biodiversity maintenance in the
550 Cerrado. The biogeographic districts harbour different plant communities, that reflect
551 differences in Cerrado biophysical and biological characteristics across its wide

552 distribution, and we expect that these same characteristics can also shape ecological
553 communities and biological interactions.

554 The characterization of biogeographic districts in other large tracts of natural
555 habitats can be useful for the conservation of the world's savannas, which are nearly all
556 strongly threatened by human activities (Lima et al. 2018). Since climatic and
557 compositional variation, as we reported here, are also expected to occur in other
558 savannas worldwide (Lehmann et al. 2014), we expect that more detailed biogeographic
559 units can be recovered and used as biodiversity surrogates for conservation planning,
560 with the overarching aim to avoid biodiversity loss worldwide.

561

562 **REFERENCES**

- 563 Aguiar LMS, Machado RB, Franoso RD, et al (2015) Cerrado: Terra inc3gnita do
564 s3culo 21. *Ci3ncia Hoje* 330:32–37. doi: 10.1017/CBO9781107415324.004
- 565 Alves RJ V., Kolbek J (2009) Can campo rupestre vegetation be floristically delimited
566 based on vascular plant genera? *Plant Ecol* 207:67–79. doi: 10.1007/s11258-009-
567 9654-8
- 568 Amaral AG, Munhoz CBR, Walter BMT, et al (2017) Richness pattern and
569 phytogeography of the Cerrado herb–shrub flora and implications for conservation.
570 *J Veg Sci* 28:848–858. doi: 10.1111/jvs.12541
- 571 Antonelli A, Verola CF, Parisod C, Gustafsson a. LS (2010) Climate cooling promoted
572 the expansion and radiation of a threatened group of South American orchids
573 (Epidendroideae: Laeliinae). *Biol J Linn Soc* 100:597–607. doi: 10.1111/j.1095-
574 8312.2010.01438.x
- 575 APG IV (2016) An update of the Angiosperm Phylogeny Group classification for the
576 orders and families of flowering plants: APG IV. *Bot J Linn Soc* 181:1–20

577 Azevedo JAR, Valdujo PH, de C. Nogueira C (2016) Biogeography of anurans and
578 squamates in the Cerrado hotspot: coincident endemism patterns in the richest and
579 most impacted savanna on the globe. *J Biogeogr* 43:2454–2464. doi:
580 10.1111/jbi.12803

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

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

585 Bernard E, Penna LAO, Araújo E (2014) Downgrading, downsizing, degazettement,
586 and reclassification of protected areas in Brazil. *Conserv Biol* 28:939–950. doi:
587 10.1111/cobi.12298

588 Beuchle R, Grecchi RC, Shimabukuro YE, et al (2015) Land cover changes in the
589 Brazilian Cerrado and Caatinga biomes from 1990 to 2010 based on a systematic
590 remote sensing sampling approach. *Appl Geogr* 58:116–127. doi:
591 10.1016/j.apgeog.2015.01.017

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

594 Borghetti JR, Silva WLC, Nocko HR, et al (2017) Agricultura Irrigada Sustentável no
595 Brasil: Identificação de Áreas Prioritárias. FAO, Brasília

596 Brasil (2000) Lei N° 9.985 - Institui o Sistema Nacional de Unidades de Conservação da
597 Natureza. DOU

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

600 Bueno ML, Pennington RT, Dexter KG, et al (2017) Effects of Quaternary climatic
601 fluctuations on the distribution of Neotropical savanna tree species. *Ecography*

602 (Cop) 40:403–414. doi: 10.1111/ecog.01860

603 Caetano S, Prado DE, Pennington RT, et al (2008) The history of Seasonally Dry
604 Tropical Forests in eastern South America: inferences from the genetic structure of
605 the tree *Astronium urundeuva* (Anacardiaceae). *Mol Ecol* 17:3147–59. doi:
606 10.1111/j.1365-294X.2008.03817.x

607 Carmignotto AP, Vivo M De, Langguth A (2012) Mammals of the Cerrado and
608 Caatinga: distribution patterns of the tropical open biomes of central South
609 America. *Bones, clones biomes Hist Geogr Recent Neotrop Mamm* 307–350. doi:
610 10.1017/CBO9781107415324.004

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

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

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

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

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

623 Costa GC, Nogueira CC, Machado RB, Colli GR (2010) Sampling bias and the use of
624 ecological niche modeling in conservation planning: a field evaluation in a
625 biodiversity hotspot. *Biodivers Conserv* 19:883–899. doi: 10.1007/s10531-009-
626 9746-8

627 Coutinho LM (2006) O conceito de bioma. *Acta Bot Brasilica* 20:13–23. doi:
628 10.1590/S0102-33062006000100002

629 Dapporto L, Ramazzotti M, Fattorini S, et al (2013) Recluster: An unbiased clustering
630 procedure for beta-diversity turnover. *Ecography (Cop)* 36:1070–1075. doi:
631 10.1111/j.1600-0587.2013.00444.x

632 De Cáceres M, Legendre P, Moretti M (2010) Improving indicator species analysis by
633 combining groups of sites. *Oikos* 119:1674–1684. doi: 10.1111/j.1600-
634 0706.2010.18334.x

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

638 Dexter KG, Lavin M, Torke BM, et al (2017) Dispersal assembly of rain forest tree
639 communities across the Amazon basin. *Proc Natl Acad Sci* 114:2645–2650. doi:
640 10.1073/pnas.1613655114

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

642 Diniz-Filho JAF (2004) Phylogenetic Diversity and Conservation Priorities under
643 Distinct Models of Phenotypic Evolution. *Conserv Biol* 18:698–704. doi:
644 10.1111/j.1523-1739.2004.00260.x

645 Diniz-Filho JAF, Bini LM, Oliveira G, et al (2009) Macroecologia, biogeografia e áreas
646 prioritárias para conservação no cerrado. *Oecologia Bras* 13:470–497. doi:
647 10.4257/oeco.2009.1303.05

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

651 Diza-Uriarte R (2014) Variable Selection using Random Forest. <https://cran.r->

652 project.org/web/packages 23

653 DRYFLOR (2016) Plant diversity patterns in neotropical dry forests and their
654 conservation implications. *Science* 353:1383–1387. doi: 10.1126/science.aaf5080

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

657 Echternacht L, Trovó M, Oliveira CT, et al (2011) Areas of endemism in the Espinhaço
658 Range in Minas Gerais, Brazil. *Flora - Morphol Distrib Funct Ecol Plants*
659 206:782–791. doi: 10.1016/j.flora.2011.04.003

660 Faleiro FV., Machado RB, Loyola RD (2013) Defining spatial conservation priorities in
661 the face of land-use and climate change. *Biol Conserv* 158:248–257. doi:
662 10.1016/j.biocon.2012.09.020

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

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

667 Flora do Brasil (2016) Jardim Botânico do Rio de Janeiro.
668 <http://floradobrasil.jbrj.gov.br>. Accessed 5 Jan 2015

669 Françoso RD, Brandão R, Nogueira CC, et al (2015) Habitat loss and the effectiveness
670 of protected areas in the Cerrado Biodiversity Hotspot. *Nat Conserv* 13:35–40. doi:
671 10.1016/j.ncon.2015.04.001

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

675 Gastauer M, Teixeira Braga Messias MC, Alves Meira Neto JA (2012) Floristic
676 Composition, Species Richness and Diversity of Campo Rupestre Vegetation from

677 the Itacolomi State Park, Minas Gerais, Brazil. *Environ Nat Resour Res* 2:115–
678 130. doi: 10.5539/enrr.v2n3p115

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

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

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

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

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

692 IBGE (2004a) Mapa de Biomas do Brasil. <https://downloads.ibge.gov.br/>. doi:
693 www.ibge.gov.br/home/geociencias/default_prod.shtm#USO

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

696 Jenkins M (2003) Prospects for Biodiversity. *Science* 302:1175–1177. doi:
697 10.1126/science.1088666

698 Klink CA, Machado RB (2005) Conservation of the Brazilian Cerrado. *Conserv Biol*
699 19:707–713. doi: 10.1111/j.1523-1739.2005.00702.x

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

702 2699.2010.02375.x

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

706 Kunz SH, Ivanauskas NM, Martins SV, et al (2008) Aspectos florísticos e
707 fitossociológicos de um trecho de Floresta Estacional Perenifolia na Fazenda
708 Trairão, Bacia do rio das Pacas, Querência-MT. *Acta Amaz* 38:245–254. doi:
709 10.1590/S0044-59672008000200007

710 Leger J-B, Daudin J-J, Vacher C (2015) Clustering methods differ in their ability to
711 detect patterns in ecological networks. *Methods Ecol Evol* 6:474–481. doi:
712 10.1111/2041-210X.12334

713 Lehmann CER, Anderson TM, Sankaran M, et al (2014) Savanna Vegetation-Fire-
714 Climate Relationships Differ Among Continents. *Science* 343:548–552. doi:
715 10.1126/science.1247355

716 Liaw a, Wiener M (2002) Classification and Regression by randomForest. *R news*
717 2:18–22. doi: 10.1177/154405910408300516

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

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

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

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

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

730 Maxwell SL, Fuller RA, Brooks TM, Watson JEM (2016) Biodiversity: The ravages of
731 guns, nets and bulldozers. *Nature* 536:143–145. doi: 10.1038/536143a

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

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

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

739 MMA (2015) Mapeamento do Uso e Cobertura do Cerrado: Projeto TerraClass Cerrado
740 2013. Proj. TerraClass Cerrado 67

741 MMA (2016) Áreas Prioritárias para Conservação dos biomas Cerrado, Pantanal e
742 Caatinga. DOU Portaria:81

743 Morrone JJ (2018) The spectre of biogeographical regionalization. *J Biogeogr* 45:282–
744 288. doi: 10.1111/jbi.13135

745 Morrone JJ (2014) Biogeographical regionalisation of the neotropical region. *Zootaxa*
746 3782:1–110. doi: 10.11646/zootaxa.3782.1.1

747 Morrone JJ, Crisci J V, Plata MD La, et al (1995) Historical Biogeography: Introduction
748 to Methods. *Annu. Rev. Ecol. Evol. Syst.* 26:373–401

749 Morrone JJ, Url S (1994) On the Identification of Areas of Endemism. *Syst Biol*
750 43:438–441

751 Myers N, Mittermeier RA, Mittermeier CG, et al (2000) Biodiversity hotspots for

752 conservation priorities. *Nature* 403:853–858. doi: 10.1038/35002501

753 Nascimento ART, Felfili JM, Meirelles M (2004) Florística e estrutura da comunidade
754 arbórea de um remanescente de Floresta Estacional Decidual de encosta, Monte
755 Alegre, GO, Brasil. *Acta Bot Brasilica* 18:659–669

756 NGA, NASA (2000) Shuttle Radar Topography Mission

757 Nogueira C, Ribeiro S, Costa GC, Colli GR (2011) Vicariance and endemism in a
758 Neotropical savanna hotspot: distribution patterns of Cerrado squamate reptiles. *J*
759 *Biogeogr* 38:1907–1922. doi: 10.1111/j.1365-2699.2011.02538.x

760 Novaes RML, Filho JP de L, Ribeiro RA (2010) Phylogeography of *Plathyenia*
761 *reticulata* (Leguminosae) reveals patterns of recent range expansion towards
762 northeastern Brazil and southern Cerrados in Eastern Tropical South America. *Mol*
763 *Ecol* 19:985–998. doi: 10.1111/j.1365-294X.2010.04530.x

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

766 Oliveira-Filho AT, Ratter JA (1995) A study of the origin of central Brazilian forests by
767 the analysis of plant species distribution patterns. *Edinburgh J Bot* 52:141. doi:
768 10.1017/S0960428600000949

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

772 Pennington RT, Dick CW (2004) The role of immigrants in the assembly of the South
773 American rainforest tree flora. *Philos Trans R Soc Lond B Biol Sci* 359:1611–22.
774 doi: 10.1098/rstb.2004.1532

775 Pinto JRR, Oliveira-Filho AT (1999) Perfil florístico e estrutura da comunidade arbórea
776 de uma floresta de vale no Parque Nacional da Chapada dos Guimarães, Mato

777 Grosso, Brasil. Rev Bras Botânica 22:53–67. doi: 10.1590/S0100-
778 84041999000100008

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

782 R Core Team (2016) R: A language and environment for statistical computing

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

825 Silva JMC, Bates JM (2002) Biogeographic Patterns and Conservation in the South
826 American Cerrado: A Tropical Savanna Hotspot. *Bioscience* 52:225–234. doi:

827 10.1641/0006-3568(2002)052

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

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

834 Simon MF, Proença C (2000) Phylogeographic patterns of Mimosa (Mimosoideae,
835 Leguminosae) in the Cerrado biome of Brazil: an indicator genus of high-altitude
836 centers of endemism? Biol Conserv 96:279–296. doi: 10.1016/S0006-
837 3207(00)00085-9

838 Siqueira MF, Peterson AT (2003) Consequences of global climate change for
839 geographic distributions of Cerrado tree species. Biota Neotrop 3:1–14. doi:
840 10.1590/S1519-566X2006000600003

841 Strassburg BBN, Brooks T, Feltran-Barbieri R, et al (2017) Moment of truth for the
842 Cerrado hotspot. Nat Ecol Evol 1:1–3. doi: 10.1038/s41559-017-0099
843 Szumik CA, Goloboff PA (2004) Areas of endemism: an improved optimality criterion.
844 Syst Biol 53:968–77. doi: 10.1080/10635150490888859

845 Tichý L, Chytrý M, Šmarda P (2011) Evaluating the stability of the classification of
846 community data. Ecography (Cop) 34:807–813. doi: 10.1111/j.1600-
847 0587.2010.06599.x

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

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

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

855 Werneck FP, Nogueira CC, Colli GR, et al (2012) Climatic stability in the Brazilian
856 Cerrado: implications for biogeographical connections of South American
857 savannas, species richness and conservation in a biodiversity hotspot. *J Biogeogr*
858 39:1695–1706. doi: 10.1111/j.1365-2699.2012.02715.x

859 Whittaker RJ, Araujo MB, Paul J, et al (2005) Conservation Biogeography: assessment
860 and prospect. *Divers Distrib* 11:3–23. doi: 10.1111/j.1366-9516.2005.00143.x

861

862

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 - SW). MDA = mean decrease accuracy.

Species	BD	MDA	CE	CW	NE	NW	S	SE	SW
<i>Aspidosperma tomentosum</i>	CE	0.015	0.012	0.019	0.021	0.020	0.005	0.007	0.019
<i>Dalbergia miscolobium</i>	CE	0.013	0.005	0.006	0.003	0.013	0.006	0.024	0.034
<i>Eremanthus glomerulatus</i>	CE	0.019	0.076	0.004	0.015	0.017	0.023	0.014	0.011
<i>Eriotheca pubescens</i>	CE	0.015	0.040	-0.001	0.025	0.008	0.024	0.014	0.012
<i>Erythroxylum tortuosum</i>	CE	0.025	0.011	-0.001	0.071	0.009	0.011	0.037	0.047
<i>Guapira noxia</i>	CE	0.030	0.068	0.004	0.086	0.018	0.017	0.020	0.031
<i>Kielmeyera speciosa</i>	CE	0.008	0.026	0.000	0.013	0.005	0.012	0.006	0.005
<i>Ouratea hexasperma</i>	CE	0.037	0.038	0.010	-0.004	0.027	0.171	0.023	0.029
<i>Salacia crassifolia</i>	CE	0.039	0.116	0.012	0.010	0.053	0.065	0.021	0.049
<i>Styrax ferrugineus</i>	CE	0.034	0.189	0.003	0.025	0.027	0.044	0.017	0.014
<i>Tachigali subvelutina</i>	CE	0.038	0.060	0.011	0.035	0.028	0.099	0.017	0.059
<i>Vochysia thyrsoidea</i>	CE	0.030	0.189	0.009	0.022	0.018	0.026	0.008	0.015
<i>Kielmeyera rubriflora</i>	CW	0.036	0.024	0.083	0.050	0.035	0.006	0.012	0.020
<i>Vochysia rufa</i>	CW	0.019	-0.005	0.015	0.016	0.008	0.071	0.007	0.031
<i>Vochysia gardneri</i>	NE	0.015	0.010	0.004	0.051	0.012	0.009	0.013	0.013
<i>Aspidosperma nobile</i>	NW	0.029	0.026	0.019	0.039	0.027	0.040	0.033	0.019
<i>Callisthene hassleri</i>	NW	0.004	0.001	0.000	0.002	0.020	0.001	0.001	0.000
<i>Caryocar coriaceum</i>	NW	0.026	0.011	0.010	0.017	0.101	0.012	0.016	0.015
<i>Davilla elliptica</i>	NW	0.015	0.002	0.015	-0.002	0.024	0.016	0.022	0.021
<i>Diospyros coccolobifolia</i>	NW	0.011	0.007	0.000	0.000	0.053	0.004	0.004	0.005
<i>Diospyros hispida</i>	NW	0.009	0.004	0.002	-0.004	0.023	0.006	0.021	0.006
<i>Heteropterys byrsonimifolia</i>	NW	0.013	0.009	0.004	-0.001	0.039	0.004	0.011	0.026

<i>Mouriri elliptica</i>	NW	0.039	0.070	0.011	0.008	0.037	0.080	0.064	0.020
<i>Pseudobombax longiflorum</i>	NW	0.022	0.001	0.015	0.059	0.033	0.013	0.024	-0.001
<i>Pseudobombax tomentosum</i>	NW	0.021	0.003	0.015	0.025	0.009	0.039	0.011	0.050
<i>Tachigali aurea</i>	NW	0.012	0.001	0.007	-0.010	0.027	0.019	0.023	0.005
<i>Bauhinia rufa</i>	S	0.011	0.003	-0.001	0.017	0.004	0.038	0.012	0.011
<i>Leptolobium elegans</i>	S	0.055	0.031	0.035	0.039	0.038	0.206	0.020	0.051
<i>Miconia paucidens</i>	S	0.003	0.001	0.001	0.001	0.001	0.019	0.001	0.001
<i>Ouratea spectabilis</i>	S	0.043	0.024	0.005	0.030	0.012	0.216	0.014	0.050
<i>Mimosa laticifera</i>	SE	0.004	0.001	0.005	0.005	0.000	0.003	0.008	0.003
<i>Callisthene mollissima</i>	-	0.002	0.002	0.003	0.001	0.004	0.000	0.001	0.000
<i>Lafoensia pacari</i>	-	0.008	-0.004	0.003	0.023	0.016	0.003	0.005	0.007
<i>Pleroma stenocarpa</i>	-	0.003	0.000	0.001	0.001	0.001	0.014	0.001	0.002

Table 3. Number of climatic parameters of Fig. 3 statistically different between the biogeographic district of the Cerrado are the Central (CE), Central-west (CW), North-east (NE), North-west (NW), South (S), South-east (SE), and South-west (SW).

	CE	CW	NE	NW	S	SE
CW	2					
NE	5	5				
NW	4	5	5			
S	4	6	6	5		
SE	4	5	6	6	5	
SW	5	5	5	6	6	5

Table 4. Biogeographic district 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 km². The proposed biogeographic districts of the Cerrado are the Central (CE), Central-west (CW), North-east (NE), North-west (NW), South (S), South-east (SE), and South-west (SW).

BD	Total area	Conv. rate	Protected Areas						Priority Conservation Areas					
			Sustainable Use		Strict Protection		Indigenous Territory		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 5. Previously identified biogeographic units (areas of endemism or biotic elements) within the biogeographic districts of the Cerrado. The districts are Central (CE), Central-west (CW), North-east (NE), North-west (NW), South (S), South-east (SE), and South-west (SW). The biogeographic units are named according to the original sources.

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

FIGURE LEGENDS

Figure 1. Values of the Calinski-Harabasz and Simple Structure Indices (SSI) for varying number of groups in k-means clustering based on a fuzzy version of the Jaccard distance. The values of each criterion are standardized as *Z*-values. The Calinski-Harabasz is high for low numbers of groups and the simple structure index selected more groups. A balanced solution favours a classification involving eight groups.

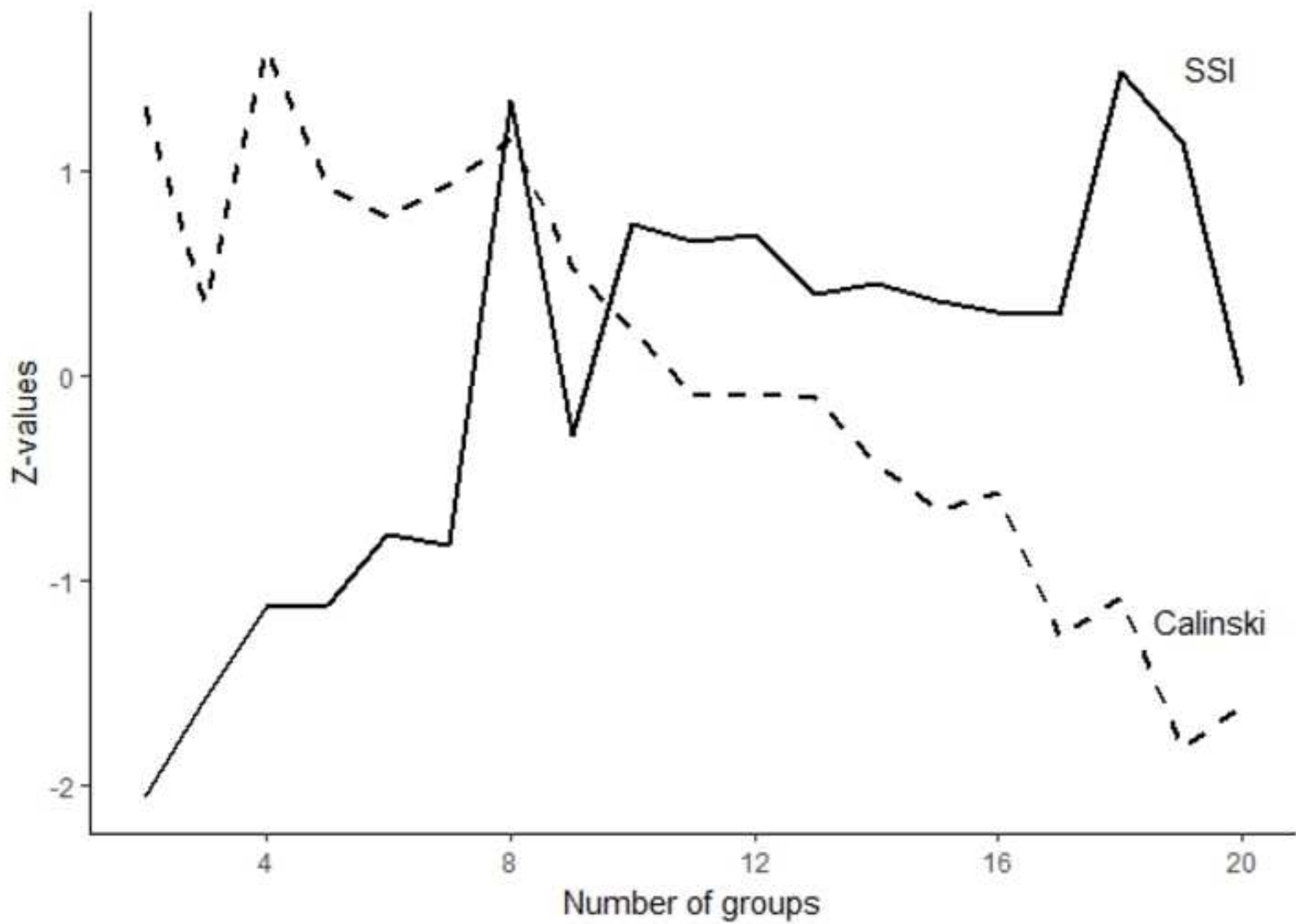
Figure 2. Biogeographic districts of the Cerrado biome in Brazil based on k-means classification and a fuzzy version of the Jaccard distance. The dots represent the surveyed sites used in the cluster analysis and the polygons were based on the distribution of sites in the same group in Fig. 1. The seven districts are Central (CE), Central-west (CW), North-east (NE), North-west (NW), South (S), South-east (SE), and South-west (SW). The group with the marginal cerrado sites in grey was not considered a district due its predominant occurrence outside of the boundaries of the Cerrado and their disjunct nature.

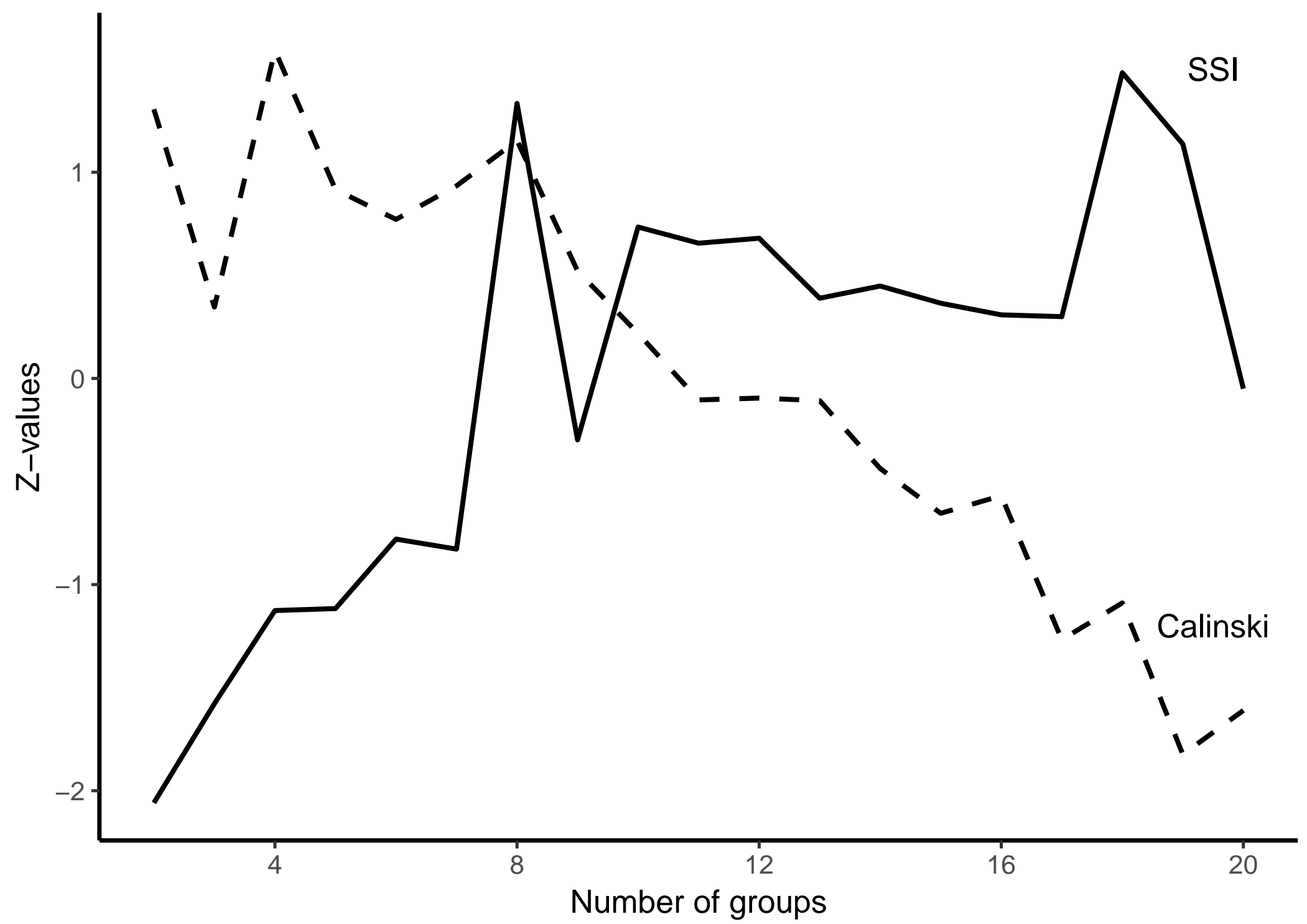
Figure 3. Consensus tree of the Cerrado biogeographic districts relationships: Central (CE), Central-west (CW), North-east (NE), North-west (NW), South (S), South-east (SE), South-west (SW), and the Extralimital (Ex) group with the marginal cerrado sites.

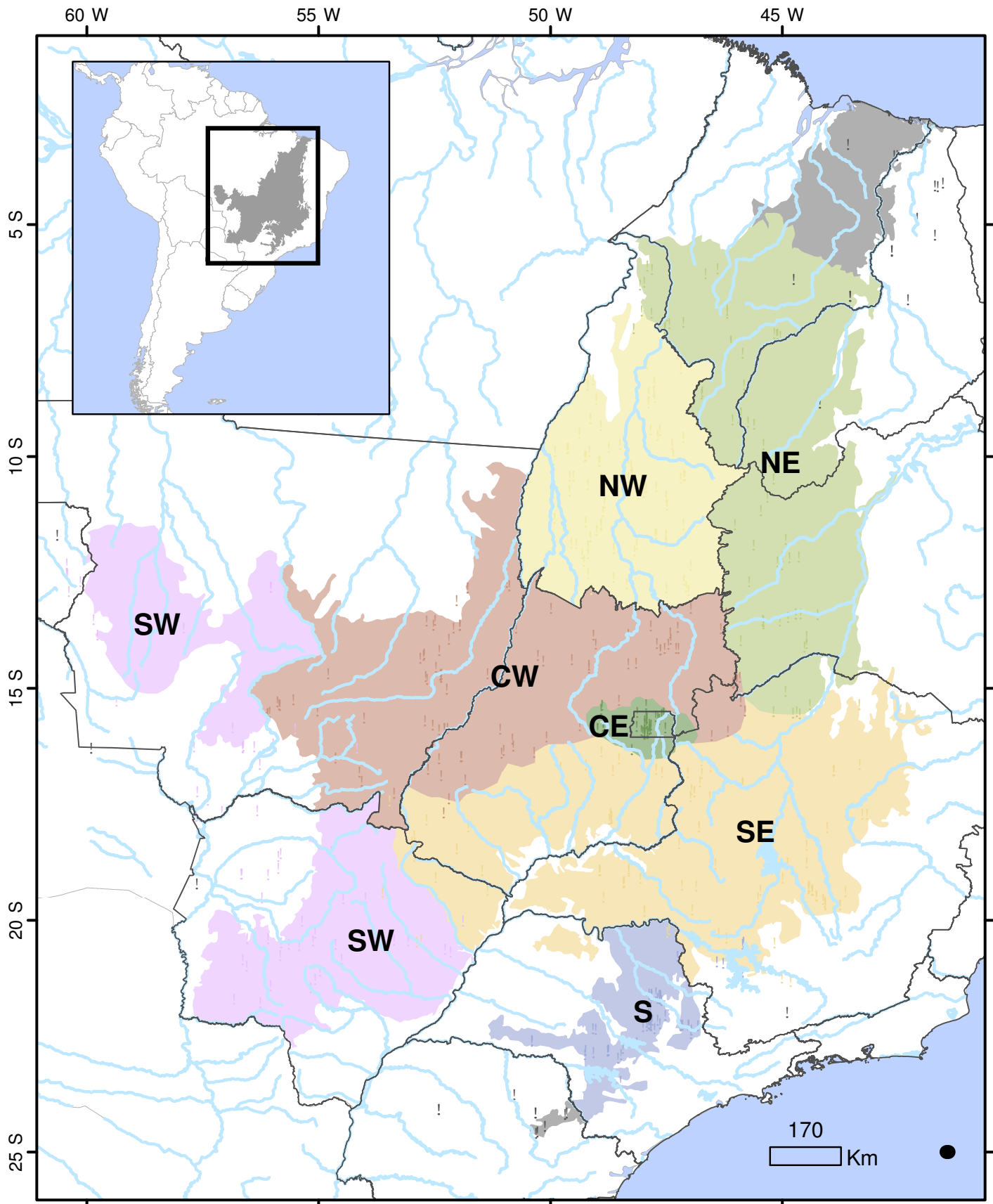
Figure 4. Boxplots showing the bioclimatic predictors selected by Random Forest to classify biogeographic districts of the Cerrado biome. Equal letters indicate no significant differences. Otherwise, all groups are significantly different for a given climatic parameter.

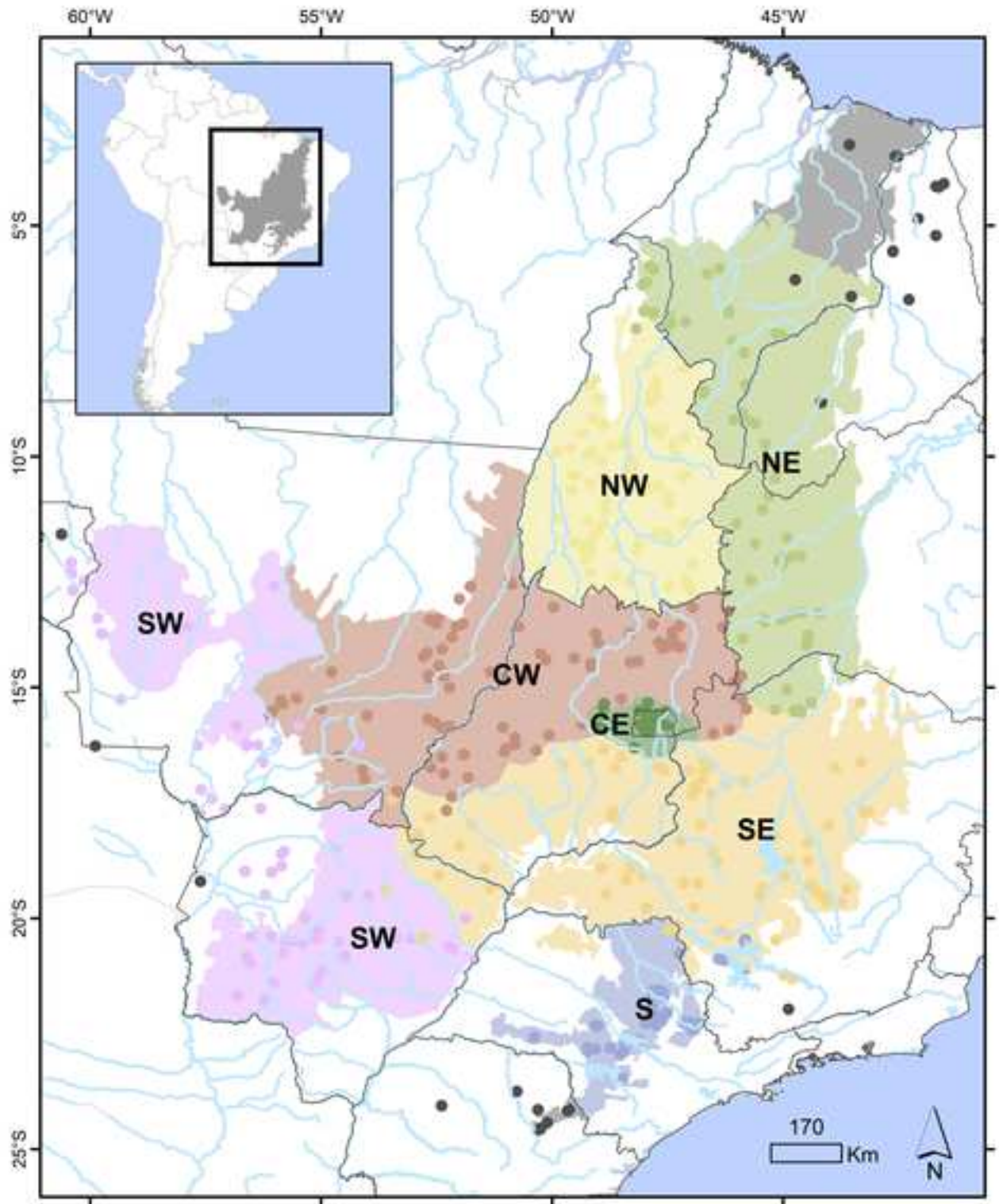
Figure 5. Remaining natural vegetation (light green), strict protection (dark green), and sustainable use areas (brown) in the Central (CE), Central-west (CW), North-east (NE), North-west (NW), South (S), South-east (SE), and South-west (SW) biogeographic districts of the Cerrado. The dashed line delimits the biome and continuous lines mark the districts.

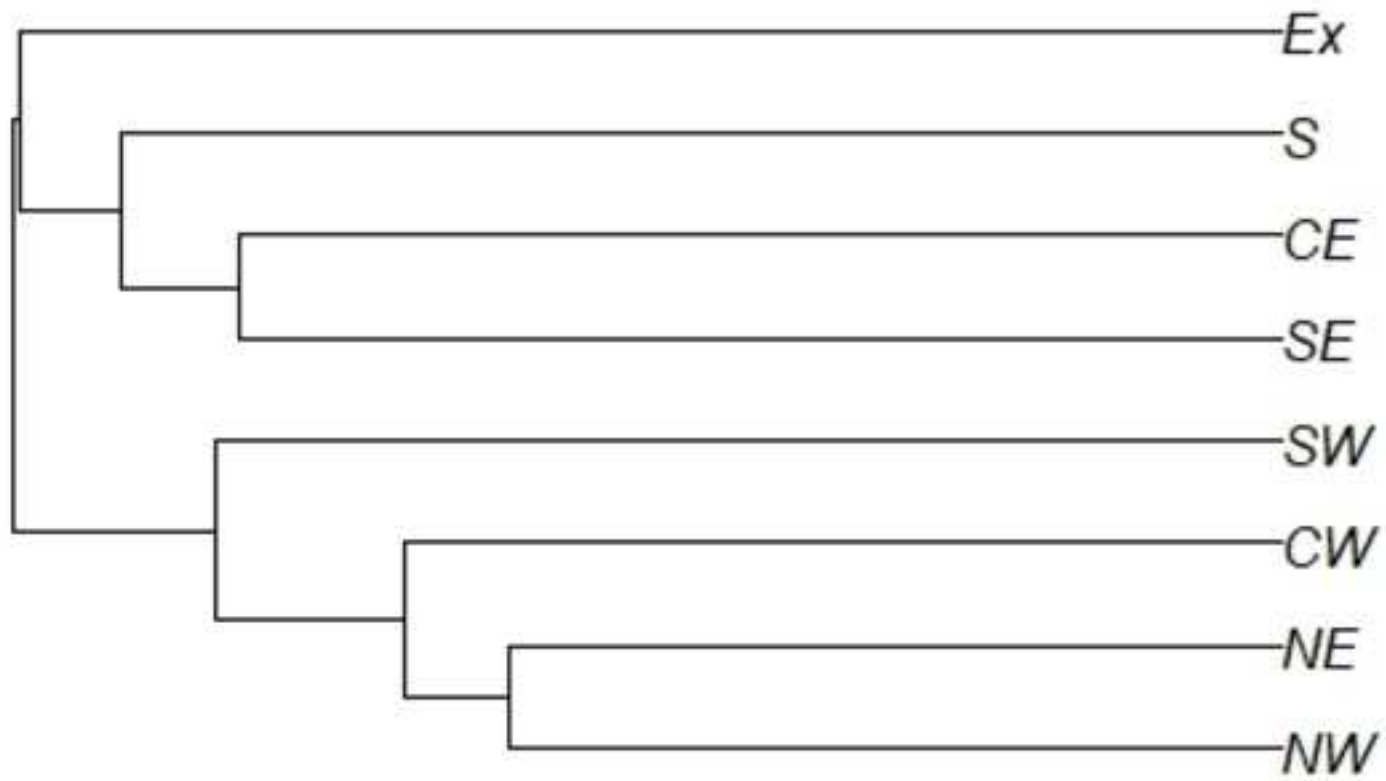
Figure 6. The Brazilian official Priority Conservation Areas (in red) over the remaining natural vegetation (light green), in the Central (CE), Central-west (CW), North-east (NE), North-west (NW), South (S), South-east (SE), and South-west (SW) biogeographic districts of the Cerrado. The shades of red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

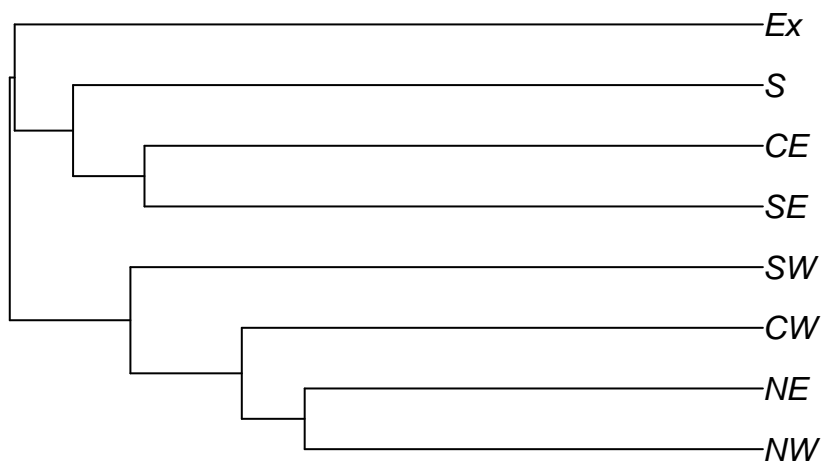


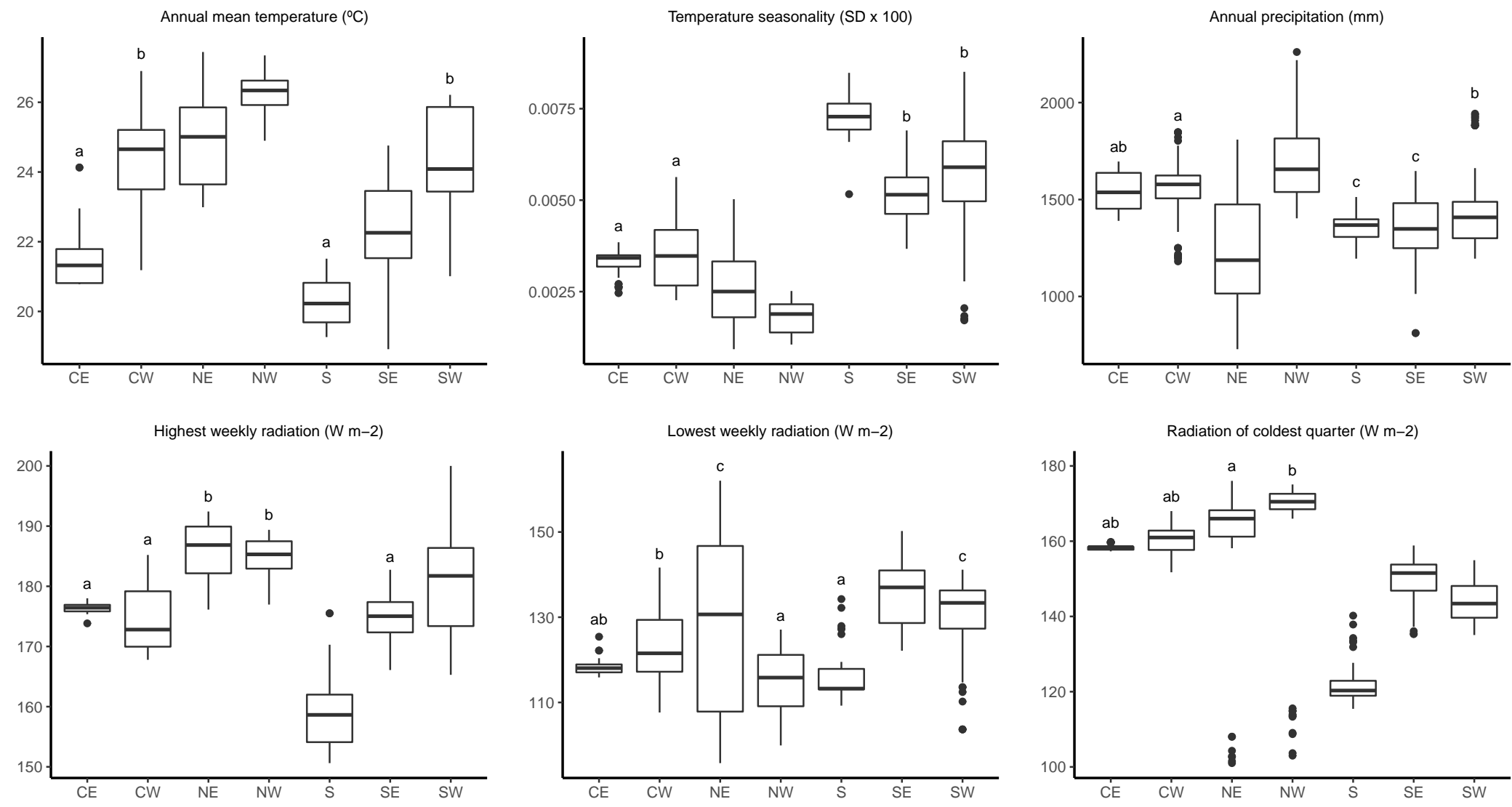


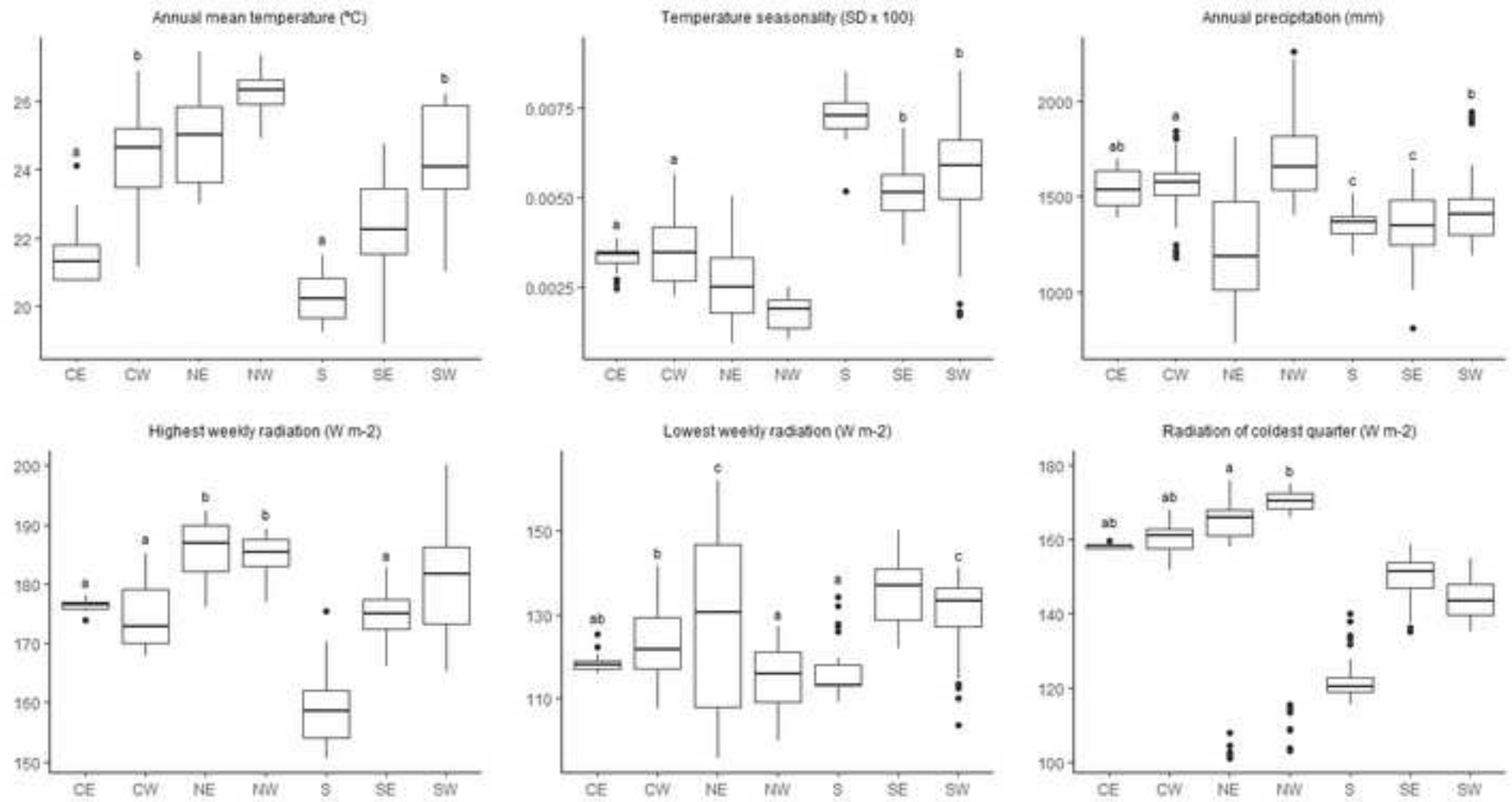


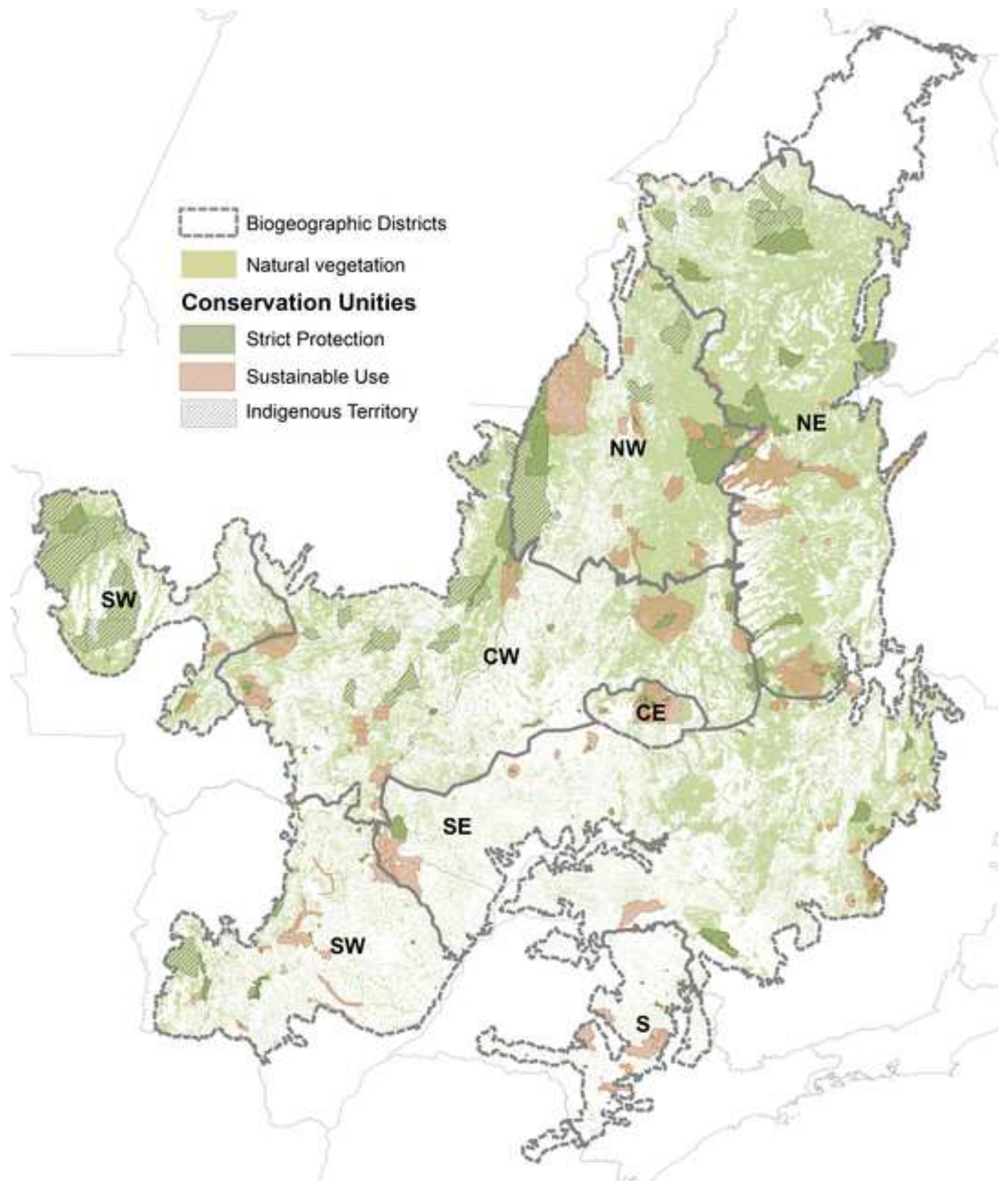


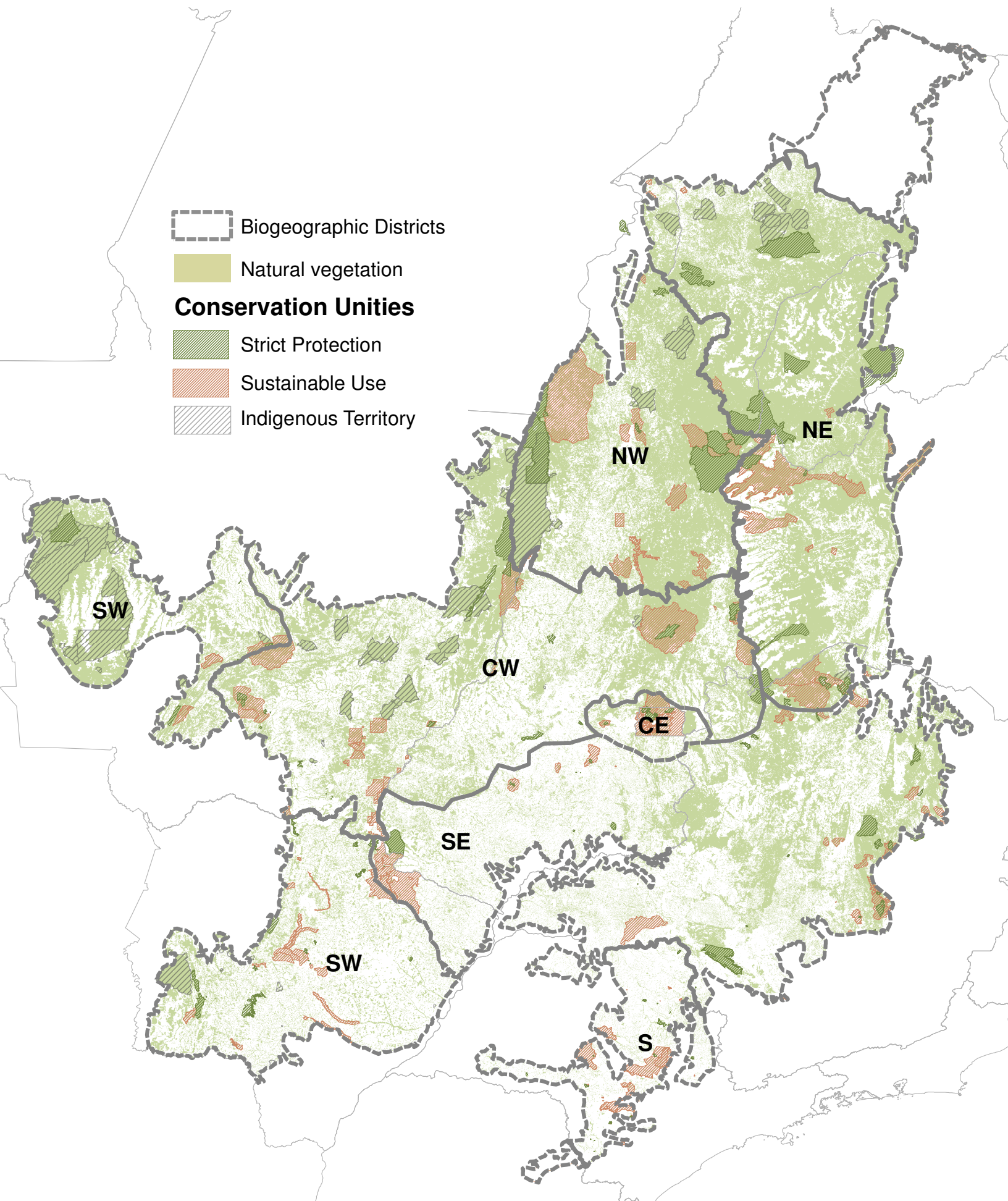


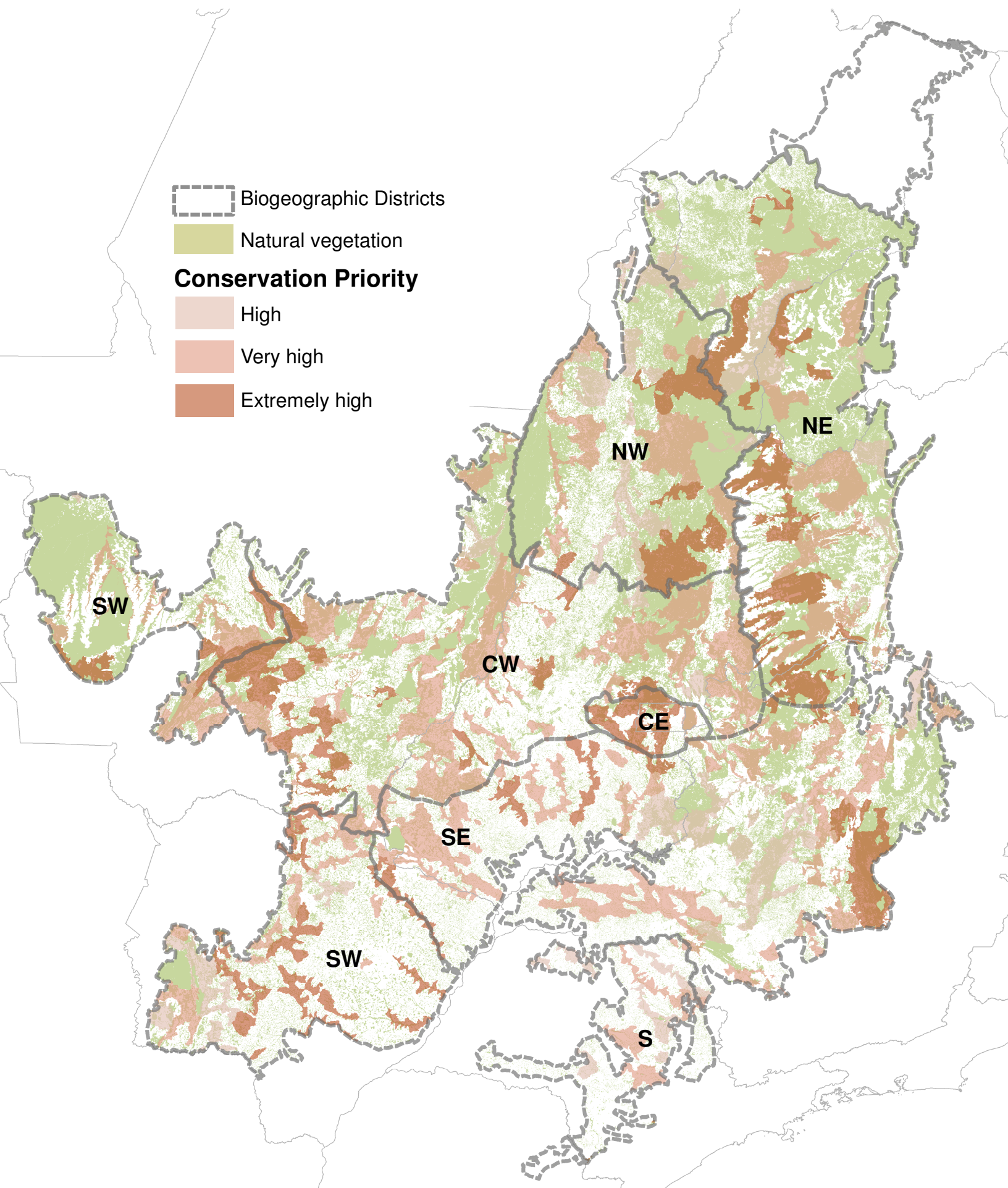


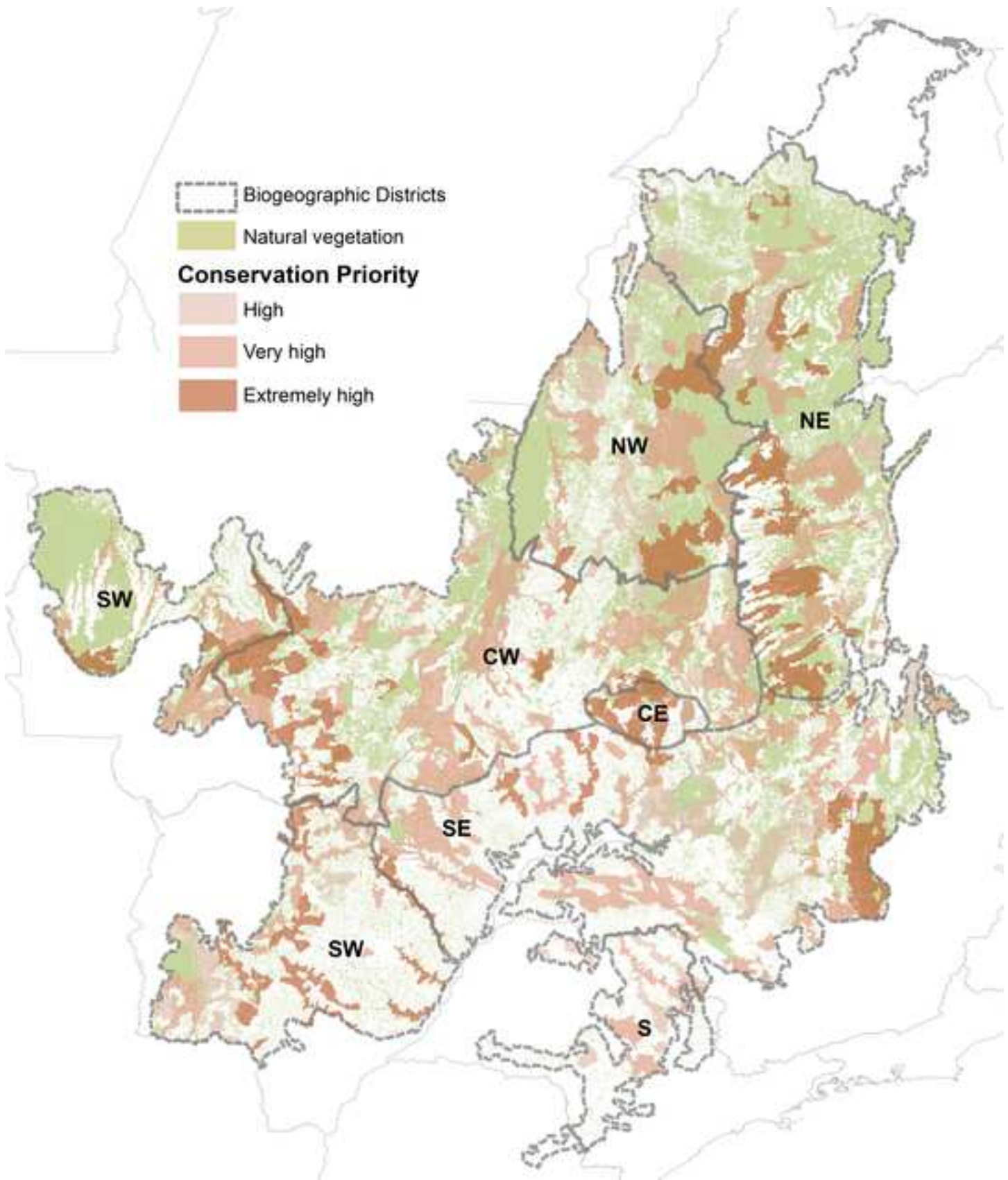




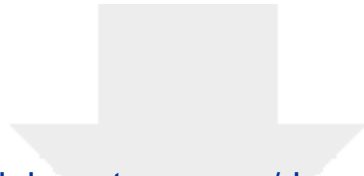








[Click here to view linked References](#)



Click here to access/download
Attachment to Manuscript
ONLINE RESOURCE.pdf

