Early View (EV): 1-EV

Effects of Quaternary climatic fluctuations on the distribution of Neotropical savanna tree species


In order to develop niche models for tree species characteristic of the cerrado vegetation (woody savannas) of central South America, and to hindcast their distributions during the Last Glacial Maximum and Last Inter-Glacial, we compiled a dataset of tree species checklists for typical cerrado vegetation (n = 282) and other geographically co-occurring vegetation types, e.g. seasonally dry tropical forest (n = 355). We then performed an indicator species analysis to select ten species that best characterize typical cerrado vegetation and developed niche models for them using the Maxent algorithm. We used these models to assess the probability of occurrence of each species across South America at the following time slices: Current (0 ka pre-industrial), Holocene (6 ka BP), Last Glacial Maximum (LGM – 21 ka BP), and Last Interglacial (LIG – 130 ka BP). The niche models were robust for all species and showed the highest probability of occurrence in the core area of the Cerrado Domain. The palaeomodels suggested changes in the distributions of cerrado tree species throughout the Quaternary, with expansion during the LIG into the adjacent Amazonian and Atlantic moist forests, as well as connections with other South American savannas. The LGM models suggested a retraction of cerrado vegetation to inter-tableland depressions and slopes of the Central Brazilian Highlands. Contrary to previous hypotheses, such as the Pleistocene refuge theory, we found that the widest expansion of cerrado tree species seems to have occurred during the LIG, most probably due to its warmer climate. On the other hand, the postulated retractions during the LGM were likely related to both decreased precipitation and temperature. These results are congruent with palynological and phylogeographic studies in the Cerrado Domain.

There is strong evidence that global climate fluctuations, and Pleistocene glacial/interglacial cycles in particular, have played a key role in determining both the origin and distribution of living organisms (Hewitt 2000). While at one time, tropical regions were considered to have been more stable than temperate regions during Pleistocene climatic cycles, a great number of studies have suggested otherwise (Haffer 1969, 1982, France 1982). Savanna is one of the main Neotropical biomes (Bourlière and Hadley 1983), and it is thought to have shifted its distribution significantly during the Pleistocene (Ledru 2002). The Cerrado Phytogeographical Domain contains the largest expanse of the savanna in the Neotropics, and there has been extensive research aimed at understanding its distribution during the Pleistocene (Ledru 1993, 2002, Oliveira-Filho and Ratter 1995, 2002, Ledru et al. 1996, Salgado-Labouriau 1997, Salgado-Labouriau et al. 1998, Werneck et al. 2012). In addition to savannas, which are the main vegetation type, the Cerrado Domain also contains grasslands, semideciduous and seasonally dry tropical forests. The savannas, generally referred to as ‘cerrado’ (note that throughout we use lower case to refer to the vegetation type), are the subject of this study.

Ledru et al. 1996, Salgado-Labouriau et al. 1997, Barberi et al. 2000, Lima-Ribeiro et al. 2004). Understanding the nature of any LGM retraction of cerrado vegetation in the Cerrado Domain has important implications, because refugial areas may contain higher overall species richness and higher genetic diversity within individual species (Collevatti et al. 2012, Lima et al. 2014), and therefore should be priorities for conservation. Understanding whether any refugial areas were numerous and scattered micro-refugia, or fewer, larger areas is therefore of great relevance (Rull 2009, 2011, Vegas-Vilarrubia et al. 2011).

The key question of whether cerrado vegetation may have expanded into Amazonia or contracted during the LGM can be addressed by modelling species distributions. Recent investigations, based on modelling species distributions and patterns of species richness, endemism and genetic variation, have provided increased evidence that climatically stable areas could have played the role of refugia for moist forest species in the Neotropics during Quaternary climatic fluctuations (Graham et al. 2006, Carnaval and Moritz 2008, Carnaval et al. 2009, 2014, Keppel et al. 2012, Werneck et al. 2011, 2012, Montade et al. 2014). Most of these recent studies have focused on moist forests and the existence of such refugia for cerrado vegetation has not been sufficiently tested using newer approaches, such as species distribution modelling. In addition to this, there has been little investigation of the distribution of savannas, and the cerrado in particular, before the LGM (although see Werneck et al. 2012). During the Last Interglacial (LIG), which began ∼ 130 000 to 116 000 BP (130–116 Ka), the climate was significantly warmer than during the Holocene maximum, registering globally higher temperatures (ca. + 2°C) and higher summer insolation (Otto-Bliesner et al. 2006). Thus, expansion of the cerrado vegetation and contraction of moist vegetation may have actually occurred during the LIG.

Species distribution modelling can be used to complement palynological studies and enhance our capacity to hindcast and forecast changes in population and vegetation dynamics (Scoble and Lowe 2010, Mellick et al. 2012). This study is the first to hindcast the distributions of tree species characteristic of the cerrado vegetation to the Last Interglacial (LIG) and Last Glacial Maximum (LGM). Werneck et al. (2012) modelled the distribution of cerrado vegetation based both on a map of the Cerrado Domain from Brazilian Inst. of Geography and Statistics (IBGE 1998) and a broader spatial definition, as geographically defined by Olson et al. (2001). However, this approach is less realistic biologically than studying the responses of individual species (Collevatti et al. 2013), which is the approach that we use here.

Our main questions were: a) was there an expansion or contraction of the cerrado vegetation during the LGM and/or LIG; b) if cerrado vegetation contracted during one of these time periods, were there areas of higher environmental suitability that could have operated as refugia; and c) if and when cerrado vegetation expanded, was it extensive enough to fragment Amazonian forest and/or establish connections between the cerrado and the savannas of northern South America?

Methods

Study area

The Cerrado Domain spreads across the Central Brazilian Highlands, which comprise 1/4 of Brazil’s surface, and to smaller areas in northwestern Paraguay and eastern Bolivia (Olson et al. 2001, Oliveira-Filho and Ratter 2002) (Fig. 1). It is the second largest Phytogeographical Domain in South America, surpassed in area only by the Amazon (Ribeiro and Walter 2008). The Cerrado Domain extends over 20 degrees of latitude and from altitudes of 100 m in the Pantanal (western floodplains) to 1500 m in the highest tablelands of the Central Brazilian Highlands (Ribeiro and Walter 2008).

There is remarkable variation in mean annual temperatures across the region, ranging from 18 to 28°C. Rainfall also varies widely, from 800 to 2000 mm yr⁻¹, with a long-lasting dry season during the austral winter (approximately April–September) (Ab’Saber 2003).

The prevalent vegetation type of the Cerrado Domain bears the same name, cerrado. It is a woody savanna that varies from fairly open grasslands to forests with a nearly closed canopy called cerradão (Ribeiro and Walter 2008). The typical cerrado vegetation grows on acidic, dystrophic soils and is one of the richest savanna floras of the world, with over 12 000 species of vascular plants (Mendonça et al. 2008).

Dataset

The floristic dataset was extracted from NeoTropTree (Oliveira-Filho 2014), a database that consists of tree (defined as free-standing woody plants > 3 m in height) species checklists for > 2000 geo-referenced sites compiled from the literature and herbarium specimen records. We extracted all 638 sites and 2155 species from the Cerrado Domain, representing 288 sites of typical cerrado vegetation, 112 sites of semideciduous forest, 116 of seasonally dry tropical forest and 122 of mesotrophic cerradão.

The cerrado vegetation is essentially a vegetation of poor dystrophic soils, and where more fertile soils occur in the Cerrado Domain, they are occupied by seasonally try tropical forests or mesotrophic cerradão, which is transitional between seasonally dry tropical forests and typical cerrado vegetation (Ratter 1973, Ratter et al. 1977, 1978a, b, Oliveira-Filho and Ratter 2002, Ribeiro and Walter 2008, Bueno et al. 2013). Seasonally dry tropical forests are notable for experiencing little fire and are thus occupied by a different set of plant lineages (e.g. Cactaceae) than those in typical cerrado vegetation, which experiences frequent and more intense fires (Pennington et al. 2000, 2009). Meanwhile, mesotrophic cerradão is an almost closed forest with a canopy cover of 50–90%, with trees often growing to 8–12 m (casting a considerable shade so that the ground layer is much reduced), and including a blend of species from both typical cerrado vegetation and seasonally dry tropical forests (Ratter 1973, 1992, Ratter et al. 1977, 1978a, Furley and Ratter 1988, Oliveira-Filho and Fontes 2000, Oliveira-Filho and Ratter 2002, Bueno et al. 2013). Semideciduous forests are found in more humid areas than typical cerrado vegetation, such as along river courses.
(i.e. gallery forest) or in transition zones with moist forests of the Amazon or the Mata Atlantica. Semideciduous forests tend to be richer in species than the other vegetation types (Oliveira-Filho and Ratter 1995, 2000, 2002).

We then performed an indicator species analysis, ISA (Dufrêne and Legendre 1997), of the same matrix from the Cerrado Domain to extract the species that indicate typical cerrado vegetation communities. The ISA produces an IV (indicator value) obtained by a combination of a species’ frequency within a group compared with other groups (specificity) and the species’ presence in most sites of that group (fidelity). We performed the analysis using the labdsv package (Roberts 2013) in the R Statistical Software (R Core Team). We then selected the ten species with the top IVs in typical cerrado and extracted the geographical coordinates of floristic lists in which the species were present from NeoTropTree (Table 1).

**Bioclimatic variables**

For all sites, we obtained the value, at 2.5 arc-min (approximately 5 km) resolution, of the 19 standard BIOCLIM variables, which reflect various aspects of temperature, precipitation, and seasonality and which are likely to be important in determining species distributions (Hijmans et al. 2005). We cropped the bioclimatic layers to span from 12°47′N to 34°46′S and from 78°31′W to 35°00′W, following Werneck et al. (2012) and which represents a much larger spatial range than that of the Cerrado Domain. After assessing correlations between the bioclimatic variables, we retained 10 of 19 variables, eliminating those with less biological relevance from groups of strongly interrelated variables ($r > 0.9$). This procedure was done to avoid over-parametrization of our modelling with redundant variables. The final selected variables were: annual mean temperature, mean diurnal range, isothermality, temperature annual range, mean temperature of wettest quarter, mean temperature of the driest quarter, mean temperature of warmest quarter, annual precipitation, precipitation of wettest month and precipitation of the driest month.

**Model construction**

We modelled the ecological niche of the ten selected indicator tree species (Table 1) using Maxent ver. 3.3 (Phillips et al. 2006). It has been demonstrated that Maxent often
outperforms other modelling techniques to identify areas critical to the maintenance of species populations (Elith et al. 2006, 2011, Pearson et al. 2007, Phillips and Dudík 2008, Pena et al. 2014). In addition, an important reason for choosing Maxent was that it allowed us to use presence-only species data, which is of great utility because the vast majority of the biotic data available, including those used here, come in this form (Elith et al. 2006, Phillips and Dudík 2008).

To calibrate and evaluate the quality of the models, we divided the data for each species into a training set (75% of occurrences) and a test or validation set (25% of occurrences). We constructed models five times and averaged the output to produce the final results used in downstream analyses. Next, for each species, we defined a threshold value above which grid cells were considered to have environmental characteristics suitable for the maintenance of viable populations of the species (Pearson et al. 2007). We used the ‘minimum training presence’ as the threshold selection method because it assumes that the species presence is restricted to sites at least as suitable as those at which the species has been observed so far (Pearson et al. 2007).

In order to produce models to infer the palaeodistribution of the cerrado indicator tree species, we produced projections of the suitability of occurrence during the Current (0 ka pre-industrial), Mid-Holocene (6 ka BP), Last Glacial Maximum (LGM – 21 ka BP), and Last Interglacial (LIG – 130 ka BP) time periods based on climatic simulations (<www.worldclim.org>; Hijmans et al. 2005). For the Last Glacial Maximum (21 ka LGM), Holocene (6 ka) and Current (0 ka pre-industrial) time periods, we employed the Community Climate System Model – CCSM4 (Gent et al. 2011), which represents downscaled climate data from simulations with Global Climate Models (GCMs) based on the Coupled Model Intercomparison Project Phase 5 (CMIP5; Taylor et al. 2012). The paleo-climatic model for the Last Interglacial (120 ka LIG) used the approach of Otto-Bliesner et al. (2006). We summed the projections of the ten indicator tree species for each time period, which together represent the probability of occurrence of typical cerrado vegetation during that time period. We performed all geographic information system (GIS) analyses in ArcGIS ver. 10 (ESRI 2011).

To indicate potential areas of climatic stability for cerrado tree species during the whole of the Quaternary, we adopted protocols similar to those used in recent studies for other Neotropical Domains (Carnaval and Moritz 2008, Werneck et al. 2011, 2012). Spatial models were converted from continuous outputs into presence/absence maps by applying the lowest occurrence threshold for each model. This approach maximizes agreement between observed and modelled distributions, balancing the cost arising from an incorrect prediction against the benefit gained from a correct prediction (Pearson et al. 2007). By summing up the presence/absence maps obtained under Current, Holocene (6 ka BP), LGM (21 ka BP) and LIG (130 ka BP) projections, we generated a map of areas showing historical stability. This combined map depicted areas that were potentially occupied by typical cerrado species during the climatic oscillations of the Quaternary. These historically stable areas, which we considered to be potential refugia, were defined as those grid cells for which the presence of all indicator species was inferred across all time projections.

### Species distribution modelling validation

We calculated the sensitivity (the proportion of observed presences in relation to those that were predicted, which quantifies omission errors), the specificity (the proportion of observed absences compared to those that were predicted, which quantifies commission errors) and the TSS (true skill statistic), following Allouche et al. (2006). The TSS test corrects the overall accuracy of the model prediction by the accuracy expected by chance. This test provides a score between −1 and +1, with values >0.6 considered to be good, 0.2–0.6 to be fair to moderate and <0.2 to be poor (Jones et al. 2010). The TSS is a threshold-dependent measure that is appropriate for evaluating predictive accuracy in cases where the model prediction is formulated as a presence–absence map (Allouche et al. 2006, Jones et al. 2010). These analyses used the occurrence points of the ten indicator tree species and 202 occurrences of *Eugenia uruguaiana* obtained from NeoTropTree (Oliveira-Filho 2014), which has a restricted niche that differs from the typical cerrado species and is ideal for simulating absence points for typical cerrado vegetation.

We also used a threshold-independent method of model validation, the receiver operating characteristic (ROC) curve analysis. The ROC curve is obtained by plotting sensitivity values (the true positive fraction) on the y-axis against their

### Table 1. The ten tree species selected as indicators of typical cerrado vegetation. IV – indicator value; IV-p – the probability of obtaining as high an indicator value as that observed for typical cerrado vegetation.

<table>
<thead>
<tr>
<th>Species</th>
<th>Species records</th>
<th>Typical cerrado</th>
<th>Mesotrophic cerrado</th>
<th>Seasonally dry tropical forests</th>
<th>Semideciduous forests</th>
<th>IV</th>
<th>IV-p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connarus suberosus Planch.</td>
<td>296</td>
<td>0.85</td>
<td>0.59</td>
<td>0.02</td>
<td>0.00</td>
<td>0.50</td>
<td>0.001</td>
</tr>
<tr>
<td>Erythroxylum suberosum St. Hl.</td>
<td>234</td>
<td>0.84</td>
<td>0.58</td>
<td>0.01</td>
<td>0.00</td>
<td>0.50</td>
<td>0.001</td>
</tr>
<tr>
<td>Palicourea rigida Kunth</td>
<td>182</td>
<td>0.65</td>
<td>0.20</td>
<td>0.00</td>
<td>0.01</td>
<td>0.50</td>
<td>0.001</td>
</tr>
<tr>
<td>Kielmeyera coriacea Mart. &amp; Zucc.</td>
<td>244</td>
<td>0.87</td>
<td>0.58</td>
<td>0.03</td>
<td>0.09</td>
<td>0.49</td>
<td>0.001</td>
</tr>
<tr>
<td>Annona crassiflora Mart.</td>
<td>194</td>
<td>0.69</td>
<td>0.26</td>
<td>0.02</td>
<td>0.04</td>
<td>0.48</td>
<td>0.001</td>
</tr>
<tr>
<td>Caryocar brasiliense Cambess.</td>
<td>263</td>
<td>0.75</td>
<td>0.40</td>
<td>0.03</td>
<td>0.06</td>
<td>0.47</td>
<td>0.001</td>
</tr>
<tr>
<td>Couepia grandiflora (Mart. &amp; Zucc.) Benth.</td>
<td>242</td>
<td>0.70</td>
<td>0.33</td>
<td>0.02</td>
<td>0.02</td>
<td>0.46</td>
<td>0.001</td>
</tr>
<tr>
<td>Qualea parviflora Mart.</td>
<td>254</td>
<td>0.91</td>
<td>0.84</td>
<td>0.04</td>
<td>0.03</td>
<td>0.45</td>
<td>0.001</td>
</tr>
<tr>
<td>Byrsonima coccolobiflora Kunth</td>
<td>218</td>
<td>0.78</td>
<td>0.48</td>
<td>0.03</td>
<td>0.06</td>
<td>0.45</td>
<td>0.001</td>
</tr>
<tr>
<td>Qualea grandiflora Mart.</td>
<td>266</td>
<td>0.95</td>
<td>0.92</td>
<td>0.12</td>
<td>0.08</td>
<td>0.43</td>
<td>0.001</td>
</tr>
</tbody>
</table>
equivalent specificity values (1 – specificity, the false positive fraction) on the x-axis for all possible thresholds (Fielding and Bell 1997). The ROC analysis characterizes the predictive performance of a model at all possible thresholds by a single number, the area under the curve (AUC) (Fawcett 2003, Phillips et al. 2006). A single AUC value was calculated for each species, representing the average across the five iterations of model construction. The value of the AUC can fall between 0.5 and 1.0. If the value is 0.5, the model is no better than random, while models with values above 0.75 are generally considered potentially useful and models with a value near one are considered to be strongly supported (Fielding and Bell 1997, Elith 2002, Rushton et al. 2004, Phillips et al. 2006).

Results

The ten tree species identified by our indicator species analyses as the most important indicators of typical cerrado and therefore chosen to generate ecological niche models and predict current and past distributions are given in Table 1. The quality of the models, according to AUC and TSS values computed for the ten indicator tree species, showed that sample and background predictions generated by Maxent were generally in agreement (Table 2). That is, the Maxent model performance in this study is much better than random. This was confirmed by the correct assignment of the test data using the models, indicating that the models showed a good performance in predicting species occurrences with bioclimatic variables.

The palaeomodels suggest significant changes in the distributions of typical cerrado tree species during the Quaternary (Fig. 2). The cerrado tree flora experienced its maximum expansion during the Last Inter-Glacial (LIG), when the modelled species spread toward the south and east of the Amazon basin as well as toward the Atlantic coast in both southeastern and northeastern Brazil (Fig. 2A). All of these areas shelter current-day cerrado enclaves within moist forests of both the Amazonian and Atlantic Forest Domains, as well as within the semi-arid thorn-woodlands of the Caatinga Domain.

In contrast with the LIG, the models suggest a maximum retraction of the modelled cerrado species during the Last Glacial Maximum (LGM), with an almost entire withdrawal from both eastern Amazonia and Atlantic coastal areas coupled with a contraction toward central Brazil and eastern Bolivia (Fig. 2B). Further, there was a notable retraction of typical cerrado vegetation to inter-tableland depressions and the slopes of the Central Brazilian Highlands, as well as a low suitability at higher altitudes (Fig. 3).

In the Mid-Holocene (Fig. 2C), the modelled species expanded their distributions to approach those of the current distribution of the typical cerrado vegetation. Lastly, the results obtained for the Current projection (Fig. 2D) showed a distribution similar to that of the Cerrado Domain, as delimited by Olson et al. (2001). Indeed, a map of under vs over-prediction of cerrado vegetation with respect to the map of Olson et al. (2001) shows high congruence, particularly in areas of the central Cerrado Domain. Meanwhile, we overpredicted typical cerrado vegetation in ecotonal areas between the Cerrado Domain and adjacent Domains, i.e. in semideciduous forests that transition to the Amazon and Mata Antártica moist forests (Supplementary material Appendix 1, Fig. A1).

Some areas in the central region of the Cerrado Domain showed a high probability of climatic stability throughout the Quaternary and are shown in yellow in Fig. 4. These postulated refugia for typical cerrado vegetation occur mainly in Minas Gerais and São Paulo states and the Federal District, with smaller areas scattered across the Cerrado Domain in other states, such as Tocantins, Goiás and Mato Grosso. The distribution of Brazilian conservation units shows a low level of coincidence with these postulated refugia (Fig. 4).

Discussion

Modelling cerrado indicator species

Our results demonstrated the greatest extent of typical cerrado tree species in the LIG, the greatest contraction in the LGM, and a subsequent re-expansion in the Holocene. The values obtained by TSS and AUC modelling suggest that the environmental variables used in our models provided important information on the distribution of the tree species selected as indicators of typical cerrado vegetation, and were higher than those obtained by other studies modelling the climatic distribution of neotropical vegetation (Carnaval and Moritz 2008, Werneck et al. 2011, 2012, Pena et al. 2014), but which did not model individual indicator species. The indicator species method has been widely and effectively

Table 2. Evaluation of the model performance for cerrado indicator tree species of cerrado vegetation by Maxent. True skill statistic (TSS) and area under the curve (AUC).

<table>
<thead>
<tr>
<th>Species</th>
<th>Training samples</th>
<th>External test presence points</th>
<th>External test absence points</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>TSS</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connar ceus suberosus Planch.</td>
<td>237</td>
<td>59</td>
<td>202</td>
<td>0.95</td>
<td>1.00</td>
<td>0.95</td>
<td>0.92</td>
</tr>
<tr>
<td>Ery thr oxylum suberosum St.Hil.</td>
<td>176</td>
<td>58</td>
<td>202</td>
<td>0.78</td>
<td>1.00</td>
<td>0.78</td>
<td>0.92</td>
</tr>
<tr>
<td>Palicoria rigida Kunth</td>
<td>137</td>
<td>45</td>
<td>202</td>
<td>0.90</td>
<td>1.00</td>
<td>0.90</td>
<td>0.92</td>
</tr>
<tr>
<td>Kielm yera coriacea Mart. &amp; Zucc.</td>
<td>183</td>
<td>61</td>
<td>202</td>
<td>0.61</td>
<td>0.99</td>
<td>0.60</td>
<td>0.92</td>
</tr>
<tr>
<td>Annona crassilora Mart.</td>
<td>146</td>
<td>48</td>
<td>202</td>
<td>0.95</td>
<td>1.00</td>
<td>0.94</td>
<td>0.93</td>
</tr>
<tr>
<td>Caryocar brasiliensis Cambess.</td>
<td>211</td>
<td>52</td>
<td>202</td>
<td>0.85</td>
<td>1.00</td>
<td>0.85</td>
<td>0.93</td>
</tr>
<tr>
<td>Couepia grandiflora (Mart. &amp; Zucc.) Benth.</td>
<td>194</td>
<td>48</td>
<td>202</td>
<td>0.64</td>
<td>0.99</td>
<td>0.63</td>
<td>0.92</td>
</tr>
<tr>
<td>Qualea parviflora Mart.</td>
<td>191</td>
<td>63</td>
<td>202</td>
<td>0.93</td>
<td>1.00</td>
<td>0.93</td>
<td>0.91</td>
</tr>
<tr>
<td>Byrsonima coccobotilolia Kunth</td>
<td>164</td>
<td>54</td>
<td>202</td>
<td>0.75</td>
<td>1.00</td>
<td>0.75</td>
<td>0.93</td>
</tr>
<tr>
<td>Qualea grandiflora Mart.</td>
<td>200</td>
<td>66</td>
<td>202</td>
<td>0.93</td>
<td>1.00</td>
<td>0.93</td>
<td>0.91</td>
</tr>
</tbody>
</table>


Figure 2. Predicted occupancy across northern South America of ten tree species that are indicators of typical cerrado vegetation during: (A) – the Last Interglacial (LIG 130 ka BP); (B) – the Last Glacial Maximum (LGM 21 ka BP); (C) – the Mid-Holocene (6 ka BP); and (D) – under Current climate (0 ka pre-industrial). Predictions were based on ecological niche models of climatic preference developed separately for each species using the MaxEnt algorithm (Phillips et al. 2006). Predicted occupancy was then summed across all ten species. Warmer colours (red/yellow) of the logistic output correspond to regions with a higher probability of occurrence. Black lines represent the borders of Brazilian states and South America countries.
et al. 2013, Baker and Fritz 2015). One important point that is clear from these studies is that climatic changes were probably not synchronised across lowland tropical South America; rather, different parts of South America may have experienced climatic change in different directions at the same time. Predicting the exact history of dispersal, extinction and recolonization of any typical cerrado tree species across lowland tropical South America is therefore challenging. Thus, here we try and focus on general patterns that can be inferred from our multi-species, palaeodistribution modelling approach.

Cerrado distribution during the LIG, LGM and Holocene

Climate models suggest a warmer and slightly drier climate during the LIG in those areas of the present-day Atlantic and Amazon rain forests into which the cerrado tree species modelled here are suggested to have expanded. Seasonal climates expanded toward the Atlantic coast in southeastern Brazil, and the palaeomodel indicated suitable areas for typical cerrado species near the coast, e.g. in the Paraíba river valley, in Rio de Janeiro and São Paulo. In fact, there were small remnants of cerrado in this region, most of which have disappeared due to habitat alteration in the last century (Matsumoto and Bittencourt 2001).

The modelled expansion of typical cerrado tree species into the Amazon during the LIG is particularly notable within the ‘Amazonian Dry Corridor’, a transverse zone with mean annual precipitation below 1750 mm extending in a northwest-southeast direction near the cities of Óbidos and Santarém (Haffer 1969, Figueroa and Nobre 1990, Van Der Hammen and Aby 1994). Although most of this region is forested, numerous isolated savannas are found there, and it connects the savannas of central Venezuela with the savannas of central and northeastern Brazil (Haffer 1969). Ab’Saber (2003) suggested the existence of savanna corridors in Amazonia during the Quaternary, though he was not certain about the period when such corridors may have existed. He also hypothesised that the corridors probably linked present-day disjunct patches of Amazonian savannas. Our models provide some corroboration for this idea, showing the expansion of cerrado species toward many of these currently disjunct savannas (Sanaiotti et al. 2002), such as Alter do Chão, Amapá, Redenção, Roraima, Humaitá and the Beni in Bolivia.

Our results are congruent with those of Werneck et al. (2012) who also suggested past connections of the cerrado to other areas of savanna in South America during the LIG and a lack of significant savanna areas or corridors across central Amazonia during the LGM. Baker and Fritz (2015) discussed the importance of applying a salinity and temperature correction to δ18O isotopic records. When these corrections are applied to the mean value of δ18O during the LGM in Amazonia (Cheng et al. 2013), this substantially alters previous climatic interpretations that the Amazon was ‘severely dry’ during the LGM. These findings all contrast with one assumption of the Pleistocene Refuge Theory (Haffer 1969, Prance 1982, Whitmore and Prance 1987), which implicated an LGM savanna expansion due to drier climates.

Palaeoecological studies from localities across lowland tropical South America support a decrease in temperature
during the LGM, indicated by the expansion of cold-adapted taxa, which are currently either relictual elements in Amazonia and the Central Brazilian Highlands, such as *Podocarpus*, *Ilex*, *Myrsine* and *Hedyosmum* (Colinvaux et al. 1996, Cardenas et al. 2011), or have vanished completely, like *Anacardia* (Ledru 1993). On the other hand, during the LGM there was a drastic retraction in the occurrence of the tropical palm *Mauritia*, which has been considered as an indicator of higher temperatures (Barberi et al. 2000), as well as the disappearance of tree species characteristic of seasonally dry tropical forest in eastern Bolivia (Whitney et al. 2013).

Many authors agree that climate in the central area of the Cerrado Domain during the LGM was characterized by a decrease in both precipitation and temperature (Ferraz-Vicentini and Salgado-Labouriau 1996, Barberi et al. 2000, Lima-Ribeiro et al. 2004). However, according to Salgado-Labouriau et al. (1998), there was no synchronicity of LGM climates inferred from palynological studies in the Cerrado Domain, which they attributed to differences in latitude and regional topography.

Our models emphasize low climatic suitability during the LGM for cerrado tree species at high altitudes (above ~ 800 m) in the Cerrado Domain, particularly in the Central Brazilian Highlands. For example, LGM palaeorecords of Salgado-Labouriau et al. (1997) and Barberi et al. (2000) infer a prevalence of cold and semi-arid climates in those highlands, with strong winds, partial soil exposure and concomitant increased erosion, based on the almost complete absence of arboreal pollen elements. Our LGM models suggest that cerrado tree species persisted at lower altitudes, probably favoured by a warmer climate, deeper soil and higher ground water storage than at higher altitudes. Thus, the inter-tableland depressions and highland slopes of central Brazil may have been refugia for cerrado species, rather than the highlands where climates were too cold and dry, as suggested by Ab’Saber (2003) and Werneck et al. (2012).

Our model, showing a retraction of major areas of typical cerrado vegetation during the LGM, is also corroborated by recent studies of population genetics in cerrado tree species (Ramos 2007, Novaes 2010, Lima et al. 2014). Phylogeographic studies of *Hymenaea stigmocarpa*, *Platymenia reticulata*, *Tabebuia aurea* and *Mauritia flexuosa* found greater genetic diversity in the central region of the Cerrado Domain, which is indicative that this area could have been more stable during the LGM. A study of the  

Figure 4. Predicted regions of historical stability for typical cerrado vegetation across the Quaternary, based on summing the predicted occupancy of ten indicator tree species across Current (0 ka pre-industrial), Mid-Holocene (6 ka BP), Last Glacial Maximum (LGM 21 ka BP) and Last Interglacial (LIG 130 ka BP) climatic scenarios. Areas in yellow are those where all ten species are predicted to occur at all four time periods, and represent postulated refugial areas for typical cerrado vegetation. Areas outlined in red are Brazilian conservation units, while black lines represent the limits of Brazilian states and South America countries. Maps are given for (A) northern South America, (B) the central area of the Cerrado Domain, and (C) the Federal District.
phylogeography of the tree species *Caryocar brasiliense* by Collevatti et al. (2003, 2012) showed that multiple lineages may have contributed to the present-day populations of *Caryocar brasiliense* in the Cerrado Domain, and that populations restricted to refugia in the central region during the LGM may have spread and dispersed to favourable areas in the last 7000 yr. Moreover, in his revision of Neotropical *Andira*, Pennington (2003) highlighted a north to south parapatric distribution of *Andira cuyabensis* and *Andira cordata* across the centre of the Cerrado Domain, perhaps related to a prior separation of the currently continuous typical cerrado vegetation during the LGM, as also suggested by the palynological data of Ledru (1993).

At the end of the LGM, between 17 000 and 11 000 BP, the climate became progressively more humid. However, permanent polar fronts remained at 10°S–20°S latitude (∼ 8500 BP), inferred from the presence of *Araucaria forests* (now confined to southern and southeastern Brazil) and the association of temperate-adapted *Podocarpus* with *Caryocar* in areas presently covered by typical cerrado (Ledru 1993, Ledru et al. 1996). With increasing temperatures, the cold weather elements were probably confined to higher altitudes, principally in gallery forests, a hypothesis supported by the presence of *Podocarpus* in the higher plateaux of the Federal District and Chapada dos Veadeiros (Barberi et al. 2000).

The return of warmer, humid conditions in the Mid-Holocene would have favoured the expansion of typical cerrado vegetation in the core area of the Cerrado Domain (Oliveira-Filho and Ratter 2000). The appearance of *Mauritia* (Barberi et al. 2000, Ledru 2002, Lima et al. 2014) and the increasing concentration of charcoal particles, are both associated with increasing temperatures and the re-expansion of cerrado vegetation (Salgado-Labouriau 1997). These changes are corroborated by palaeoecological studies from various localities and supported by our palaeomodels for 6000 BP. During this period, Behling (1995) recorded an increase of species typical of cerrado vegetation, such as *Curatella americana*, in the Lagoa do Pires between the Cerrado and Atlantic Forest Domains in Minas Gerais state. In the state of Rondônia, there was an isotopic enrichment related to the replacement of forest vegetation by typical cerrado vegetation (Pessenda 1998a), as also observed in the region of Humaitá, in the south of Amazonas state (Gouveia et al. 1997, Pessenda 1998b, De Freitas et al. 2001). At the Bolivian border with Brazil, in an area now covered by Amazonian forest, Mayle et al. (2000) and Mayle and Whitney (2012) also recorded the presence of *Curatella* and *Mauritia* during the same period. This find is compatible with a trend of continuously increasing pollen deposition of typical cerrado taxa in the period (Barbieri et al. 2000, Ledru 2002). The patterns may have been accentuated by soil leaching and acidification, which would also have favoured the expansion of typical cerrado vegetation (Oliveira-Filho and Ratter 2000).

**Current distribution of cerrado, stable areas and conservation**

Around 2000 yr BP, palaeoecological studies suggest the onset of present-day climatic conditions. The establishment of a vegetation mosaic with a predominance of typical cerrado species, interspersed with forest and wet grassland formations, is suggested by numerous authors (Oliveira-Filho and Ratter 1995, 2002, Ab’Saber 2003, Ribeiro and Walter 2008). In the modelled current potential distribution, areas of the central Cerrado Domain are maximally suitable for the occurrence of typical cerrado tree species, particularly on the central Brazilian tablelands. This finding is corroborated by Ratter et al. (2003) and Bridgewater et al. (2004), who demonstrated that areas of the central Cerrado Domain show the highest species richness of cerrado tree species.

This high diversity may reflect the stability of the central Cerrado Domain throughout the timespan of our climate models. Other studies have indicated that the stability of climate through time facilitates the accumulation and maintenance of diversity in Neotropical vegetation (Graham et al. 2006, Carnaval et al. 2009, Werneck et al. 2012). The persistence of some species in multiple refugia located throughout their present distribution indicates that these species might have persisted through multiple climatic cycles in heterogeneous environments (Keppel et al. 2012, Turchetto-Zolet 2013). The microrefugia (yellow areas in Fig. 4) are small areas with favourable environmental features within which small populations could have survived when their main distribution area contracted (Rull 2009, 2011, Vegas-Vilarrubia et al. 2011). These areas of historical climate stability likely allowed a number species to persist through time, whereas extinction took place in areas that experienced the most severe climate changes. This then likely resulted in greater diversity in more stable areas (Rull 2008, 2011, Collevatti et al. 2012, Keppel et al. 2012).

There has been a great loss in species diversity and endemism in important areas of the cerrado, as a result of disturbance, and total clearance, by humans, especially due to the expansion of agriculture, cattle ranching, and charcoal production (Ratter et al. 1997, Silva and Bates 2002). There are estimates that less than 20% of the Cerrado Domain vegetation remains undisturbed while only 7.44% is legally protected in conservation units. Meanwhile, many threatened species remain outside any of the region’s parks and reserves (MMA 2011) contributing to the status of the Cerrado Domain as one of the world’s biodiversity hotspots, deserving urgent conservation intervention (Myers et al. 2000).

In our model, climatically stable areas are mostly outside the existing protected areas. The few exceptions are those located in the Federal District and a number of Environmental Protection Areas (APA; a lesser protection level) in Tocantins state (Fig. 4). Larger climatically stable areas in Minas Gerais and São Paulo states have no conservation units (Fig. 4). We suggest that the areas identified as climatically stable in our analyses should be incorporated into systematic conservation planning to preserve the cerrado tree flora, as they represent probable refugial areas with potentially high species and genetic diversity.

**Conclusion**

Palaeodistribution modelling of tree species representative of typical cerrado vegetation showed expansions and contractions related to the climatic fluctuations of the Quaternary,
with the widest expansion during the LIG, related to a warmer, more seasonal climate. The inter-tableland depressions and the highland slopes in the central region of the Cerrado Domain probably operated as refugia for the Cerrado flora during its major retraction in the LGM, a conclusion that is highly congruent with palynological and phylogeographic studies. This central region is indicated as the most species-rich and most stable throughout the climate fluctuations of the Quaternary, and the conservation of such high-diversity and climatically stable areas should be prioritized.

Acknowledgements – This study was in partial fulfilment of the Doctoral requirements of MLB who thanks CNPq for supporting a 12-month study period at the Royal Botanic Garden Edinburgh (grant SWE-202096/2011-4) and Postdoctoral scholarship in UFMG (151002/2014-2). MLB thanks the Royal Botanic Garden Edinburgh for support during the time this research was conducted.

References


