



OPEN

Protected areas are not effective for the conservation of freshwater insects in Brazil

Karina Dias-Silva¹, Thiago Bernardi Vieira^{1✉}, Felipe Ferraz Figueiredo Moreira², Leandro Juen³ & Neusa Hamada⁴

Biodiversity conservation has faced many challenges, especially the conversion of natural areas that compete with use for agriculture, energy production and mineral extraction. This problem is further aggravated by lack of knowledge of the biodiversity that exists and the geographical distribution of different groups. The objectives of our study were to examine the distributional pattern of Gerromorpha diversity in Brazil, create a map of conservation priority areas, estimate the degree of protection that the current network of protected areas guarantees to this insect group, and identify the size thresholds in geographical distributions that would allow species to be protected. We used species occurrences from the Water Bugs Distributional Database, and we used 19 bioclimatic variables to build models of the potential distributions of species using the MaxEnt program. Using the potential model results, we calculated diversity metrics and overlapped them with the current state and federal “conservation units” (protected areas for biodiversity) in Brazil. Total beta diversity and turnover portions were separated into two faunistic groups, one in northern and the other in southern Brazil. The Amazon has higher beta diversity than what was predicted by the null models. We detected a positive relationship between species distribution area and occurrence in conservation units. Conservation units with less than 250 km² do not protect Gerromorpha species. Our results reinforce the necessity of formulating new conservation strategies for this group, contemplating species with both restricted and ample distributions, because rare and specialist species are the most harmed by habitat reduction, given that they are more sensitive to environmental disturbance.

Biodiversity conservation has faced many challenges in the past few years, especially because of the speed with which natural areas have been converted to areas for agriculture, energy production, mineral extraction and real-estate development¹. This challenge is even bigger when areas of economic interest overlap with areas that are important for biological conservation^{2,3}, creating socio-economic conflicts. In addition, the definition of areas for conservation is usually intended for flag species, synergistic species or species with a social appeal^{4,5}, leaving aquatic and terrestrial insects out of this planning⁶. One way of diminishing this problem is by creating and implementing conservation units to conciliate nature conservation with sustainable resource use, making human presence in protected areas compatible with biodiversity conservation. In Brazil there are 12 types of “conservation units” (CUs) (*Unidades de Conservação* in Portuguese), or protected areas for biodiversity⁷. The different types of CUs differ in the level of access permitted, some are like ecological sanctuaries and have effective protection, being restricted to scientific research; others are less restricted. Some unit types are less restricted and allow use for recreation (e.g., National Parks) and even for resource exploitation (Extractive Reserves). Generally, each protected area has rules regarding its use. These areas shelter a high diversity of species, ecosystem services and traditions. CUs provide biodiversity maintenance, and (in some cases) biodiversity recovery⁷. However, not all protection measures are effective because some CUs are inefficient as a result of the areas being selected arbitrarily based either on the empirical knowledge of researchers or by being placed in areas with difficult access, which limits competing interests for agriculture or for real-estate development. Selection may be

¹Programa de Pós-Graduação em Ecologia, Universidade Federal do Pará, Av. Perimetral, Guamá, Pará CEP 66075-110, Brazil. ²Instituto Oswaldo Cruz, Laboratório de Biodiversidade Entomológica, Av. Brasil, 4365, Pavilhão Mourisco, Sala 214, Manguinhos, Rio de Janeiro, RJ CEP: 21040-360, Brazil. ³Laboratório de Ecologia E Conservação, Instituto de Ciências Biológicas, Universidade Federal Do Pará, Av. Perimetral, Guamá CEP 66.075-110, Belém, Pará, Brazil. ⁴Instituto Nacional de Pesquisas da Amazônia-INPA, Av. André Araújo, 2936, Petrópolis, Manaus, Amazonas CEP 69067-375, Brazil. ✉email: vieirاتب@ufpa.br

Number of occurrences	Number of species	AUC mean	<i>p</i>
4 to 15	76	0.915	
16 to 30	18	0.954	0.63
31 to 45	10	0.951	0.82
46 to 60	2	0.918	0.52
61 to 75	4	0.902	0.96

Table 1. Values of Monte Carlo tests performed to identify the relationship between the numbers of occurrences and the observed AUC values. AUC area under the curve.

done considering only the specific characteristics of a restricted biological group, usually only accounting for terrestrial biodiversity⁸. Brazil's current CU network is insufficient and does not adequately protect biodiversity^{6,9}.

The selection of priority areas for conservation must consider an area's representativeness, irreplaceability, flexibility, complementarity and persistence¹⁰. Area selection should be based on biogeographical and ecologically georeferenced information, so as to permit analyses and interpretation at different scales¹¹. However, it is often impossible to fully protect insect biodiversity due to the lack of biogeographical information on species^{3,12}. In the Neotropical region the lack of knowledge is acute regarding real species distributions (the "Wallacean shortfall")¹². For instance, in Brazil, biogeographical information on species is scarce, even for biodiversity hotspots. Lack of information, together with the lack of financial resources and the overlapping interests between agricultural development and conservation are the main challenges in defining priority areas for conservation^{6,13}. In addition, we do not know whether the current reserve network is sufficient to protect all the taxonomic groups. A study carried out on bats in the Cerrado showed that species with restricted distributions tend not to be covered by conservation units¹⁴. This result leads us to question the effectiveness of conservation units, especially for key species in ecosystems (such as aquatic insects, which are not 'charismatic').

Inefficiency of reserve networks may be related to non-protection of rare species⁶ or to choosing areas that are marginally adequate for species distributions¹⁵. One way of reducing inefficiency is by using other criteria for area selection, e.g., species richness, beta diversity, fauna complementarity and the number of endemic species^{3,16}. These parameters, along with use of species-distribution models (SDMs) could reduce the problem of lack of knowledge about distributions that leads to protection of marginally adequate areas. This knowledge would allow selection of areas having high probability of occurrence of a species and that provide favorable environmental conditions. Although this approach has been shown to be robust for designing prioritization measures e.g.^{17–21} the criteria used are generally based on a few groups of vertebrates and plants, which reduces the effectiveness of protected areas for conserving the biodiversity of all groups. Consequently, several biodiversity components, such as aquatic invertebrates, might not be included in the protected areas^{6,9}.

In aquatic ecosystems, macroinvertebrates stand out from other groups because of their sensitivity to environmental impact^{22,23} and their role in nutrient cycling and energy transfer in food chains. Insects in the suborder Heteroptera are predators, are at the top of food chains, can respond to changes happening in lower trophic levels, and are considered to be good models for evaluating environmental impact²⁴. Heteroptera occupy an ample variety of habitats, including waterbodies that are either lotic or lentic and either perennial or temporary; these insects have an important role in biological control in waterbodies^{25–28}. Heteroptera is composed of three infraorders: Nepomorpha (truly aquatic), Gerromorpha and Lepdomorpha (both semi-aquatic). The infraorder Gerromorpha is divided into 20 families, 325 genera, and approximately 4700 species inhabiting freshwater ecosystems: 28% of these species are in the Neotropics²⁹. Most Gerromorpha species live over the water column, either on floating plants or between plant roots on the edges of freshwater bodies³⁰. Thus, they are affected by forest removal and increased waterflow in streams^{27,31}. Approximately 2100 species have been described for Gerromorpha in the Neotropics, which currently has eight families and about 160 genera²⁹, of which 208 species and subspecies occur in Brazil³². This group was chosen because it occupies various types of aquatic ecosystems, has an ample geographical distribution, and perhaps acts as an umbrella or surrogate group encompassing additional aquatic species that inhabit the same sites^{33,34}.

Due to the speed with which natural areas have been modified by anthropogenic activities and the lack of knowledge about the state of conservation of aquatic insects, our objectives are to: (i) describe distributional patterns of Gerromorpha diversity (Richness and Beta Diversity) among Brazilian biomes; (ii) create a map of conservation priority areas for this group; (iii) assess the importance of the current network of conservation units and (iv) identify the threshold for the extent of geographical distribution needed to protect species under the current Brazilian system of conservation units.

Results

Of 208 Gerromorpha species recorded in Brazil, we found 3541 occurrences and modeled 111 species (only species with three or more occurrences are suitable for this procedure) (Table S1). The Monte Carlo simulation tests to evaluate the dependency of AUC on the number of occurrences were not significant for all classes (Table 1), which means that there is no relationship between the number of occurrences and the AUC values obtained from the model³⁵.

Significant portions of the species-occurrence sites were in the Southeastern Region of Brazil, in the Atlantic Forest biome, and in the Northern Region (Amazon) (Fig. 1). The areas with the highest richness were the northern portion of the Amazon, eastern and western Cerrado (central Brazilian savanna), and a small coastal

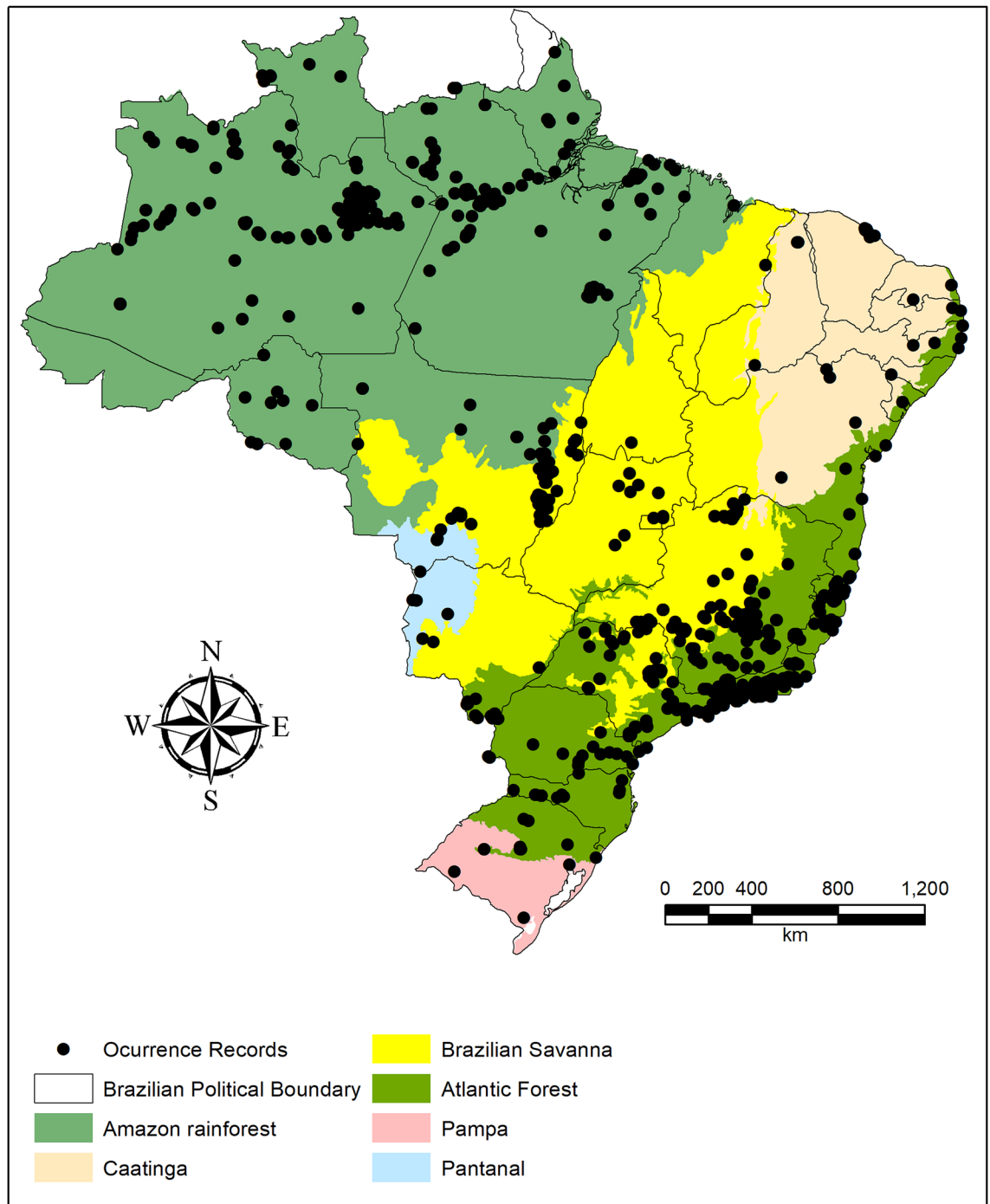


Figure 1. Gerrormorpha species occurrences used in the species distribution models (SDM). The figure was built on the R environment⁶³.

portion in the northern Atlantic Forest (Fig. 2a). The Atlantic Forest and the Amazon had higher richness than predicted based on random sampling of grid cells (Table 2). When evaluating total beta diversity, the highest values were found for the Northern Region of Brazil (Fig. 2b), and the Amazon biome had higher beta diversity than predicted (Table 2). However, when evaluating the components of beta diversity, the turnover portion (Fig. 2d) was greatest in the Southern Region of Brazil (Fig. 2d). On the other hand, in the nested portion we found the areas with the lowest species richness values (the northern Cerrado and Caatinga, as well as the central Amazon); these are the areas with the highest nestedness values for the Gerrormorpha (Fig. 2c).

The map of cell importance for conservation (Fig. 3) shows that cells in the northern Amazon, the eastern and western Cerrado and in the coastal portion of the Atlantic Forest are important for Gerrormorpha conservation, and many locations that are considered priority areas are outside of the conservation units. In general, the biomes did not differ in terms of importance for the conservation of Gerrormorpha (Table 3), and we observed for

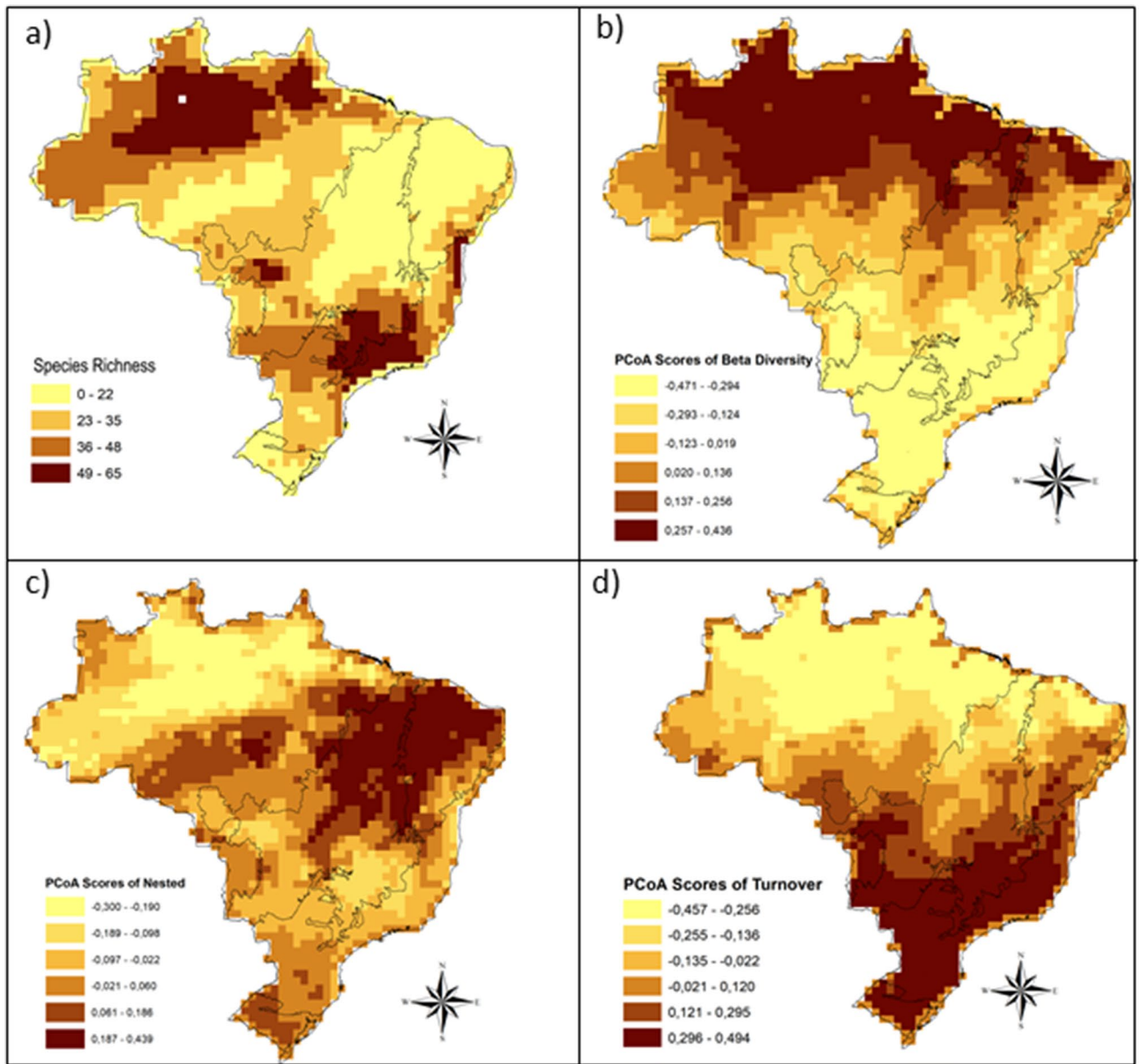


Figure 2. Predicted spatial distributions, based on species predicted presences resulting from the binarization of MaxEnt output of (a) richness, (b) total beta diversity, (c) turnover and (d) nestedness of Gerromorpha in Brazil. The figure was built on the R environment⁶³.

Biome	N	Richness			Beta diversity		
		Observed (M±SD)	Random (M±SD)	p	Observed (M±SD)	Random (M±SD)	p
Amazon	998	36 ± 12	32 ± 1	<0.001	0.159 ± 0.177	<0.001 ± 0.005	<0.001
Caatinga	231	17 ± 8	32 ± 2	0.984	0.030 ± 0.181	<0.001 ± 0.150	0.024
Cerrado	468	29 ± 14	32 ± 1	0.893	(-)0.185 ± 0.211	<0.001 ± 0.010	0.074
Atlantic Forest	194	38 ± 11	32 ± 1	<0.001	(-)0.317 ± 0.143	<0.001 ± 0.017	0.872
Pampas	57	22 ± 5	32 ± 2	0.072	(-)0.220 ± 0.170	<0.001 ± 0.033	0.032
Pantanal	21	30 ± 3	32 ± 4	0.771	(-)0.273 ± 0.137	<0.001 ± 0.056	0.023

Table 2. Random and observed mean (M) and standard deviation (SD) values of Gerromorpha richness and total beta diversity in Brazilian biomes. N—number of observed cells in each biome; p—significance value according to the Monte Carlo randomization test.

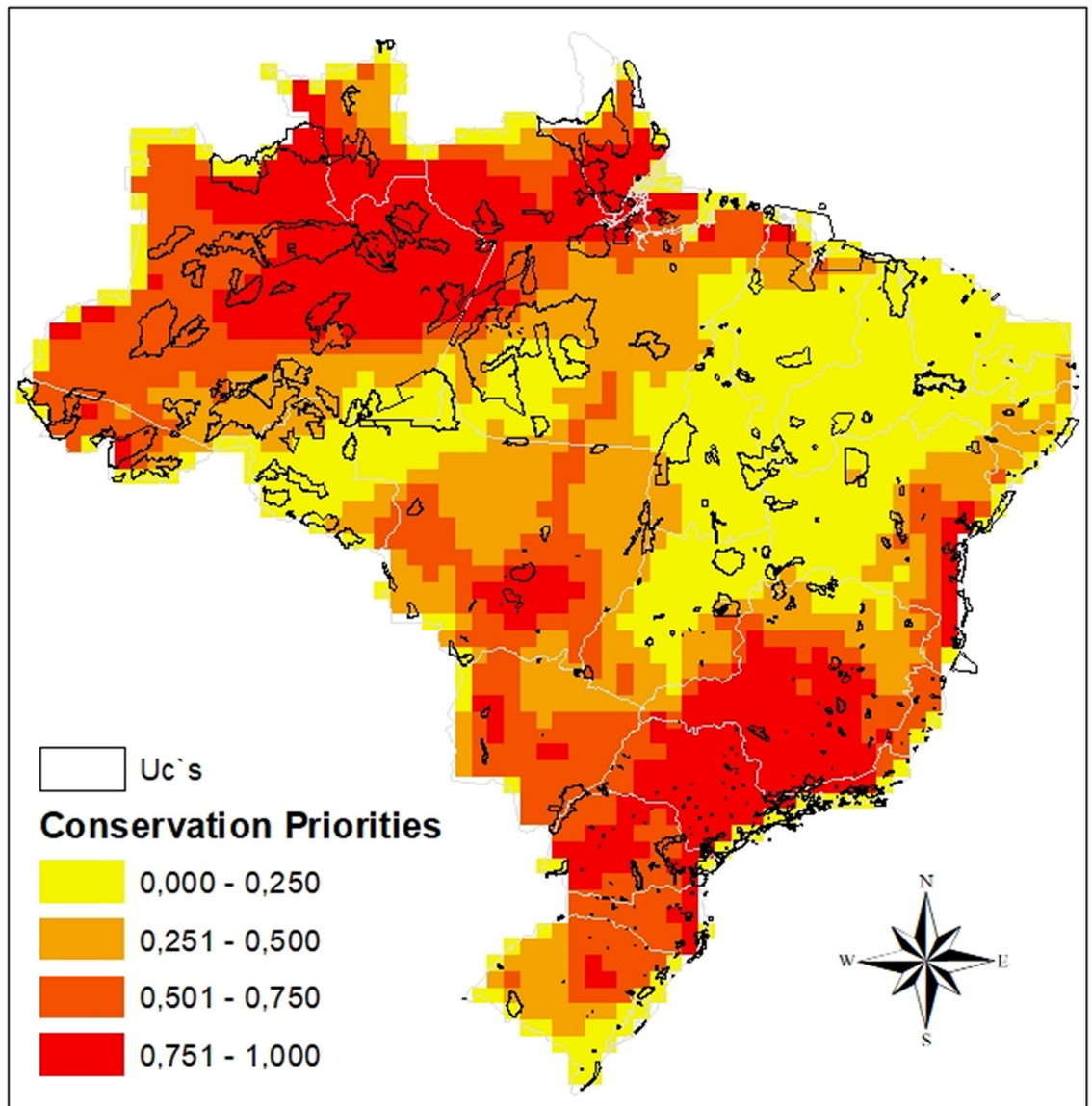


Figure 3. Priority areas for Gerromorpha conservation in Brazil, according to the Zonation algorithm. Values indicate the importance of each cell; the higher the value, the higher the importance. The figure was built on the R environment⁶³.

Biome	N	Conservation unit				Biome		
		NUC	Observed (M ± SD)	Random (M ± SD)	p	Observed (M ± SD)	Random (M ± SD)	p
Total	1969	262	0.462 ± 0.097	0.503 ± 0.066	0.811			
Amazon	998	188	0.509 ± 0.082	0.502 ± 0.019	0.344	0.502 ± 0.081	0.502 ± 0.000	0.965
Caatinga	231	16	0.208 ± 0.053	0.174 ± 0.043	0.210	0.174 ± 0.032	0.174 ± 0.002	0.986
Cerrado	468	34	0.361 ± 0.085	0.455 ± 0.051	0.966	0.455 ± 0.096	0.455 ± 0.003	0.586
Atlantic Forest	194	19	0.492 ± 0.202	0.624 ± 0.062	0.977	0.623 ± 0.082	0.624 ± 0.004	0.987
Pampas	57	4	0.118 ± 0.056	0.294 ± 0.119	0.916	0.294 ± 0.060	0.294 ± 0.014	0.978
Pantanal	21	1	0.549	0.498 ± 0.259	0.985	0.501 ± 0.070	0.498 ± 0.067	0.264

Table 3. Random and observed mean (M) and standard deviation (SD) of importance values for the conservation of Gerromorpha in Brazil (Total) and for each biome. N—number of observed cells in each biome; NUC—number of protected cells in each biome; p—significance value according to the Monte Carlo randomization test.

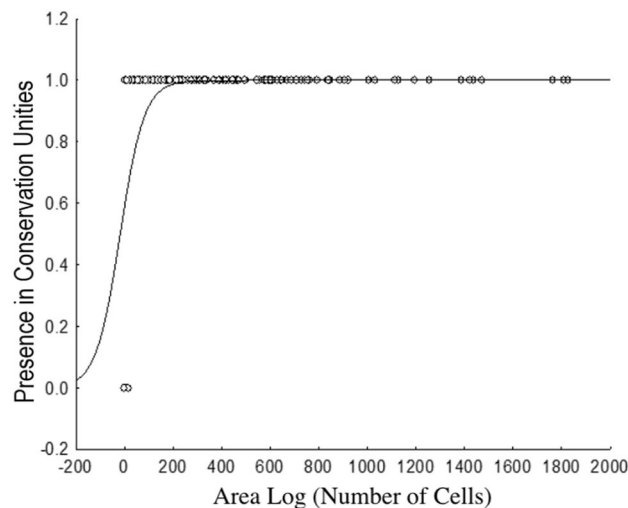


Figure 4. Logistic relationship between Gerrormorpha occurrence and presence of these species in CUs. Four species are located outside the protected area of the biome, where zero (0) means location outside a CU and one (1) means presence in a CU.

both the general model (evaluating conservation units in Brazil as a whole) and the regional models (evaluating the conservation units in each biome) that conservation units do not protect either more or less areas that are important for Gerrormorpha than what was randomly expected (Table 3). Lastly, we observed a positive relationship between species distribution areas and species occurrences in conservation units ($\chi^2_{(1)} = 36.144$ $p < 0.001$). The minimum occurrence for a species to be present within a conservation unit is approximately five cells, i.e., 250 km² (Fig. 4).

Discussion

Cells protected by conservation units (CUs) have a lower degree of importance than expected, based on their values of complementarity and irreplaceability resulting from application of the zonation algorithm. This low importance, determined according to Gerrormorpha biodiversity, may be related to the fact that the CUs have been created arbitrarily, probably based on economic interest or to protect specific taxonomic groups³ that are not aquatic⁹. These results indicate that the current network of CUs does little to protect Gerrormorpha species, especially those with restricted distributions, and especially species with distributions smaller than 250 km². CUs may be protecting sites with marginal environmental suitability for the occurrence of these species e.g.^{6,14,36}, thereby making these species more susceptible to environmental changes, such as those generated by anthropogenic activities and climate change. These results support the conclusion that conservation strategies should be based on conservation policies and not only on the productive value of the areas (which are generally areas with low agricultural interest), or based on political considerations³⁷. When decisions on protected areas are based solely on economic and political interests they reduce the possibility of including species with restricted distributions or that occur in areas with high economic value for agricultural production^{10,14,38–41}.

Creating new conservation units is a slow process because it depends on financial resources to buy land, provide area maintenance and pay teams working in the unit, besides the various impediments imposed by the conflicts-of-interest mentioned earlier. Thus, an economically viable possibility would be to maintain existing forest fragments and encourage farmers to maintain legal reserves and riparian vegetation. Forest remnants and legal-reserve areas would act as stepping stones for biodiversity, supporting the maintenance of connectivity among CUs. In addition, healthy riparian vegetation keeps catchments fully preserved, further enhancing connectivity among forest remnants, legal reserves and CUs. Finally, we emphasize that, although the maintenance of riparian forests is considered to be sufficient for maintaining the Gerrormorpha community in running waters, other strategic areas (such as those presented in this paper) must be maintained in addition to the riparian strips. Otherwise one risks only species with wide occurrence ranges will be preserved, leaving rare and short-range endemics outside the network of CUs.

In terms of species richness, we found that the Atlantic Forest and the Amazon biomes shelter the highest number of Gerrormorpha species per unit area, and the richness in these biomes is higher than would be expected if random, suggesting that their habitats are more heterogeneous than in the other Brazilian biomes. This idea is supported by the results on total beta diversity, suggesting that the Gerrormorpha composition in these biomes results from steep environmental gradients. On the other hand, we observed the Pampas biome to have fewer species than expected, suggesting that these areas have low heterogeneity for this group.

Regarding priority areas for Gerrormorpha sampling, we observed that all biomes have low sampling rates. Some examples are the “arc of deforestation” in the southern portion of the Amazon biome, large parts of the Pantanal and Caatinga biomes, and part of the Atlantic Forest biome. In general, we observed that most sampling is concentrated near institutions. If we consider the vast river network in Brazil we can conclude that the

number of records we have is small, providing an example of the “Wallacean shortfall”⁴², where information about species distributions is still insufficient. Additional sampling is needed so that information on species distributions can be expanded and more occurrence sites can be identified, enabling the discovery of new species and the generation of information on species distributions to assist in choosing conservation areas³⁹. We also emphasize the importance of evaluating the extreme northern portion of the Amazon and the southeastern Cerrado, which are very relevant for conservation and have high species richness but are not well studied. Furthermore, the southeastern portion of the Cerrado stands out for its large turnover component, suggesting a strong environmental filter, possibly due to the existence of small fragments of vegetation, that can be used as stepping stones by some species (few in this case) but that cannot be used as habitat. The northern portion of the Cerrado and the Caatinga stand out for their nesting components, suggesting that species dispersion is easily observed within and between these biomes, as observed for bats in the Cerrado¹⁴. The nesting component was justified by the existence of large preserved areas in the northern portion of the Cerrado biome, which has favored the mobility of bats¹⁴. The turnover component was justified by the local extinction of some species because the southern and southwestern portions of the Cerrado are highly fragmented, which hampers dispersal and may cause local extinctions of species. The high diversity of Gerromorpha in these regions is also congruent with other biological groups, e.g., birds, reptiles, amphibians and non-flying mammals^{43,44}. Hence, by preserving the diversity of Gerromorpha, we are also preserving the diversity of other groups, and ensuring self-sustaining ecosystems, thus being able to recover from stochastic events or even long-term climate change. We therefore stress that CUs defined with little or no planning have low importance for Gerromorpha conservation and do not sufficiently preserve species with restricted distributions (less than 250 km² for Gerromorpha). New strategies for the conservation of this group are needed, especially addressed to conserve both species with restricted ranges and those with broad distributions⁶ because rare and specialist species are the most affected by habitat reduction and are the most sensitive to environmental disturbance.

Materials and methods

Species occurrence data. Species occurrences (Fig. 1) were obtained from the Water Bugs Distributional Database (available at <https://sites.google.com/site/distributionaldatabase/>). This database represents an extensive research effort to obtain information from the literature. The database contains occurrence information for all Gerromorpha species known in Brazil together with the sources from which the information was obtained. More information about the database is available Dias-Silva et al.⁴⁵.

Initially, coordinates of the occurrence data were transformed from the original geographic reference system into decimal degrees (in datum SAD 69), using the SpeciesLink website (available <http://splink.cria.org.br/convertor?criaLANG=pt>). This transformation is necessary to perform the SDMs because the algorithms only recognize decimal degree as geographic coordinates. In order to avoid any bias caused by geographically imprecise data, we eliminated occurrences where the distance between point coordinates and the seat of the municipality was less than 2 km^{46,47}.

Environmental variables. We used 19 bioclimatic variables to build the models; the values were obtained from monthly data on temperature and precipitation available on the WorldClim 1.4. platform⁴⁸. The climatic variables have a resolution of approximately 9 km (≈ 0.083 decimal degrees) and are highly colinear. In order to solve or reduce the collinearity problem, we performed a principal component analysis (PCA)⁴⁹, from which we extracted seven axes that explained, in total, 95% of the variation in the original dataset. These axes were used as predictive environmental variables in the potential species distribution model (SDM). We tried therefore to reduce the multicollinearity problem of the original variables, making it possible to model species distributions with a small dataset on occurrence⁵⁰.

Distribution models and model evaluation. Potential species distributions were estimated using the MaxEnt (Maximum Entropy Modeling) program version 3.3.3⁵¹. To analyze the performance of the potential species distribution model, we used two techniques: AUC (area under the curve) and TSS (true skilled statistics). The AUC measure, which is the area under the ROC (receiver operating characteristic) curve, has scores varying between 0 and 1, where values ≈ 0.5 (or smaller) correspond to models predicting no better than randomly fitted ones, and values ≥ 0.7 are considered acceptable^{52,53}. True skilled statistics (TSS) scores vary from -1 to 1, where scores close to zero and negative scores are no better than random scores, and scores close to 1 have a perfect fit between the observed distribution and the predicted distribution⁵⁴. Values ≥ 0.5 indicate acceptable models in the field of ecology and for the purpose of protected-area design e.g.:^{14,19,20,55–57}. To convert predicted species distributions into presence/absence maps, we used threshold values^{58,59} that were derived from the ROC curve. Use of the maximum threshold value (the Max TSS threshold) to maximize the specificity and sensitivity of models and tends to produce more-restricted distributions¹⁹, e.g.:^{57,60,61}. Additionally, we performed a Monte Carlo simulation test to evaluate the dependency of AUC on the number of occurrences⁶ (Table 1). Thus, we divided the data into five classes (with an interval of 15 locations between the classes). For each class we calculated the mean of the observed AUC (sum of the AUCs of all models in the class, divided by the number of species present in the class) (Table 1). The first class (4–15 occurrences) was considered as the control, using the observed mean AUC as a critical value for all other classes. For AUC value simulations we maintained the species number inside the classes equal and calculated the mean AUC of a random set of species, called the “random AUC mean.” For this we choose a random set of species (the number of species selected was the same as the number of species present in the class) and calculated the random AUC mean. This procedure (random AUC mean) was repeated 10,000 times, and random AUC mean values greater than or equal to the observed AUC mean of the control class (between 4 and 15 occurrences) were counted. This number was finally divided by 10,000 to assess the

p-value of the class. The null hypotheses in our analysis was that there is no relationship between an increasing number of occurrences and an increasing AUC value. This procedure was repeated for all classes, and the results were not significant.

Data analysis. To identify the spatial distribution of Gerromorpha diversity in Brazil, we used a grid with a resolution of 0.083° (the same resolution as the SDM) and extracted information on species presence and absence from potential distribution maps. Therefore, each row represents a grid cell and each column represents a species, and the spreadsheet cells represent species presence or absence. This table summarizes the predicted presence or absence for each target species in the Brazilian territory. Using this table, we calculated the total species richness of each cell based on the sum of species presences along the corresponding row and on the beta diversity. For beta diversity we used the procedure described by Baselga et al.⁶², in which diversity is partitioned into portions for nestedness and turnover, and the sum of these portions represents total beta diversity. For the spatialization of beta diversity (the total of the nested and turnover portions), we performed a principal coordinates analysis (PCoA)⁴⁹ on each diversity matrix, and the score of the first PCoA axis was used as a synthesis of the diversity for the cell and is shown on a map (one map for each component). Thus, we obtained values for richness, total beta diversity, nested beta diversity and turnover beta diversity for each grid cell. This allowed us to spatialize diversity values by categorizing the cells into color gradients. This procedure was performed for each of the diversity measures.

We used a Monte Carlo randomization procedure with 10,000 randomizations to evaluate diversity distribution among Brazilian biomes. For this procedure we developed a routine in the R environment that classifies grid cells according to the biome to which cells belong and calculates the mean and standard deviation of the diversity metrics for each biome. We chose cells randomly for the significance test (always controlling the number of observed cells in the biome) and calculated the mean random diversity. For this we assessed the number of cells present in each biome (See Table 2) and calculated the mean and standard deviation of the diversity metrics of these cells. We then selected random cells (the same number present in the focal biome) and calculated the random mean and standard deviation of the diversity metrics of these cells. This procedure was repeated 10,000 times end for each biome. To calculate the significance value, we identified the number of random values higher than the observed values and divided them by 10,001. This procedure was performed for richness and total beta diversity.

To determine priority areas for conservation of aquatic species in Brazil, we used the Zonation algorithm⁶³, along with data on the potential distribution of target group. Zonation is a quantitative method that prioritizes conservation areas, aiming to long-term biodiversity persistence⁶⁴. Evaluation is done by randomly removing cells. This analysis can be performed using three algorithms: Core-area Zonation, Additive benefit function and Target-based planning. For this study, we used the additive-benefit function because it is best suited for a high number of species. This analysis is run with the complete landscape, where sites are classified based on biological values (complementarity and irreplaceability), and subsequently the less-valuable cells are removed one (or more) at a time, producing a sequence of landscape structures with more resources for biodiversity⁶⁵.

To evaluate the difference in importance between protected and unprotected cells, we used a second routine in the R environment based on the Monte Carlo randomization test with 10,000 randomizations. Cells were classified into those that were “protected” (if they were totally or at least 50% within a conservation unit) and those that were “unprotected,” and the mean importance (accessed from the Zonation analysis) of protected cells was estimated. The same number of unprotected cells was selected randomly, and we estimated the mean importance. The estimate of the random importance value was calculated 10,000 times, thus creating the expected distribution of the importance values. This procedure was performed for the reserve network of Brazil as a whole and for the reserve network inside each biome. The significance value was obtained by dividing the number of values that were greater than or equal to the observed values by 10,001⁶⁶. The reserve network considered for the analyses was the official map of conservation units in the “full protection” category, which is available in the Brazilian Ministry of Environment (MMA) website (<http://mapas.mma.gov.br/i3geo/datadownload.htm>). The official biome borders (also available on the MMA website) were considered to represent the historical distribution of the biological biomes.

To evaluate if there is a minimum range size for a species to be protected by a conservation unit, we performed a logistic regression between the potential extent of occurrence as predicted by the binarized SDMs for species distribution areas and their presence/absence inside the units. Species distribution area was estimated using the sum of the numbers of cells in which the species occurs, and the cells were classified as being inside (1) or outside (0) of conservation units. A non-linear logistic regression was performed^{35,67}, where presence/absence was considered to be the dependent variable.

Data availability

All occurrences data are available in Water Bugs Distributional Database (<https://sites.google.com/site/distributionaldatabase/>).

Received: 8 July 2021; Accepted: 15 October 2021

Published online: 28 October 2021

References

1. Brooks, T. M. *et al.* Global biodiversity conservation priorities. *Science* (80-.). **313**, 58–61 (2006).
2. Camacho-Sandoval, J. & Duque, H. Indicators for biodiversity assessment in Costa Rica. *Agric. Ecosyst. Environ.* **87**, 141–150 (2001).

3. Diniz-Filho, J. A. F. *et al.* Ensemble forecasting shifts in climatically suitable areas for *Tropidacris cristata* (Orthoptera: Acridoidea: Romaleidae). *Insect Conserv. Divers.* <https://doi.org/10.1111/j.1752-4598.2010.00090.x> (2010).
4. Morse-Jones, S. *et al.* Stated preferences for tropical wildlife conservation amongst distant beneficiaries: Charisma, endemism, scope and substitution effects. *Ecol. Econ.* **78**, (2012).
5. Verissimo, D., MacMillan, D. C. & Smith, R. J. Toward a systematic approach for identifying conservation flagships. *Conserv. Lett.* vol. 4 (2011).
6. Nóbrega, C. C. & De Marco, P. Unprotecting the rare species: a niche-based gap analysis for odonates in a core Cerrado area. *Divers. Distrib.* **17**, 491–505 (2011).
7. SNUC, (Sistema Nacional de Unidades de Conservação da Natureza). Lei no 9.985, de 18 de julho de 2000. *Mma/Sbf* (2000) doi:<https://doi.org/10.1017/CBO9781107415324.004>.
8. Abell, R., Allan, J. D. & Lehner, B. Unlocking the potential of protected areas for freshwaters. *Biol. Conserv.* **134**, 48–63 (2007).
9. Monteiro, C. da S., Esposito, M. C. & Juen, L. Are the adult odonate species found in a protected area different from those present in the surrounding zone? A case study from eastern Amazonia. *J. Insect Conserv.* **20**, 643–652 (2016).
10. Margules, C. R. & Pressey, R. L. Systematic conservation planning. *Nature* **405**, 243–253 (2000).
11. Whittaker, R. J. *et al.* Conservation biogeography: assessment and prospect. *Divers. Distrib.* **11**, 3–23 (2005).
12. Bini, L. M., Diniz-Filho, J. A. F., Rangel, T. F. L. V. B., Bastos, R. P. & Pinto, M. P. Challenging Wallacean and Linnean shortfalls: knowledge gradients and conservation planning in a biodiversity hotspot. *Divers. Distrib.* <https://doi.org/10.1111/j.1366-9516.2006.00286.x> (2006).
13. Rodrigues, A. S. L. & Gaston, K. J. Maximising phylogenetic diversity in the selection of networks of conservation areas. *Biol. Conserv.* [https://doi.org/10.1016/S0006-3207\(01\)00208-7](https://doi.org/10.1016/S0006-3207(01)00208-7) (2002).
14. Silva, D. C., Vieira, T. B., da Silva, J. M. & de Cassia Faria, K. Biogeography and priority areas for the conservation of bats in the Brazilian Cerrado. *Biodivers. Conserv.* **27**, 815–828 (2018).
15. Salkeld, D. J., Padgett, K. A. & Jones, J. H. A meta-analysis suggesting that the relationship between biodiversity and risk of zoonotic pathogen transmission is idiosyncratic. *Ecol. Lett.* **16**, 679–686 (2013).
16. Juen, L. & de Marco, P. Dragonfly endemism in the Brazilian Amazon: competing hypotheses for biogeographical patterns. *Biodivers. Conserv.* <https://doi.org/10.1007/s10531-012-0377-0> (2012).
17. Mendes, S. L. *et al.* Protected Areas for the Northern Muriqui, *Brachyteles hypoxanthus* (Primates, Atelidae). *Neotrop. Primates* **13**, (2005).
18. Serra, B. D. V., De Marco Júnior, P., Nóbrega, C. C. & Campos, L. A. D. O. Modeling potential geographical distribution of the wild nests of *Melipona capixaba* Moure & Camargo, 1994 (Hymenoptera, apidae): conserving isolated populations in mountain habitats. *Nat. a Conserv.* **10**, 199–206 (2012).
19. Mendes, P. & De Marco, P. Bat species vulnerability in Cerrado: integrating climatic suitability with sensitivity to land-use changes. *Environ. Conserv.* **45**, 67–74 (2018).
20. Brasil, L. S. *et al.* A niche-based gap analysis for the conservation of odonate species in the Brazilian Amazon. *Aquat. Conserv. Mar. Freshw. Ecosyst.* aqc.3599 (2021) doi:<https://doi.org/10.1002/aqc.3599>.
21. da Silva, J. G., Vieira, T. B. & Mews, H. A. Fine-scale effect of environmental variation and distance from watercourses on pteridophyte assemblage structure in the western Amazon. *Folia Geobot.* <https://doi.org/10.1007/s12224-021-09390-y> (2021).
22. Doughty, C. R. Freshwater biomonitoring and benthic macroinvertebrates, edited by D. M. Rosenberg and V. H. Resh, Chapman and Hall, New York, 1993. ix + 488pp. ISBN 0412 02251 6. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **4**, 92–92 (1994).
23. Harper, D. M., Rosenberg, D. A. & Resh, V. H. Freshwater biomonitoring and benthic macroinvertebrates. *J. Appl. Ecol.* **31**, 790 (1994).
24. Cunha, E. J. & Juen, L. Impacts of oil palm plantations on changes in environmental heterogeneity and Heteroptera (Gerromorpha and Nepomorpha) diversity. *J. Insect Conserv.* **21**, 111–119 (2017).
25. Schuh, R. T. & Slater, J. A. *True bugs of the World (Hemiptera: Heteroptera). Classification and Natural History.* (Cornell University Press, 1995).
26. Giehl, N. F. da S., Dias-Silva, K., Juen, L., Batista, J. D. & Cabette, H. S. R. Taxonomic and Numerical Resolutions of Nepomorpha (Insecta: Heteroptera) in Cerrado Streams. *PLoS One* **9**, e103623 (2014).
27. Dias-Silva, K., Cabette, H. S. R., Juen, L. & Jr, P. D. M. The influence of habitat integrity and physical-chemical water variables on the structure of aquatic and semi-aquatic Heteroptera. *Zool.* **27**, 918–930 (2010).
28. Panizzi, A. R. & Grazia, J. *True Bugs (Heteroptera) of the Neotropics.* *True Bugs (Heteroptera) of the Neotropics* vol. 2 (Springer Netherlands, 2015).
29. Polhemus, J. T. & Polhemus, D. A. Global diversity of true bugs (Heteroptera; Insecta) in freshwater. *Hydrobiologia* <https://doi.org/10.1007/s10750-007-9033-1> (2008).
30. Nieser, N. & Melo, A. L. *Os Heterópteros Aquáticos de Minas Gerais.* (UFMG, Belo Horizonte, 1997).
31. Cunha, E. J., de Assis Montag, L. F. & Juen, L. Oil palm crops effects on environmental integrity of Amazonian streams and Heteropteran (Hemiptera) species diversity. *Ecol. Indic.* **52**, 422–429 (2015).
32. Cordeiro, I. & Moreira, F. New distributional data on aquatic and semiaquatic bugs (Hemiptera: Heteroptera: Gerromorpha & Nepomorpha) from South America. *Biodivers. Data J.* **3**, e4913 (2015).
33. Rodrigues, A. S. L. & Brooks, T. M. Shortcuts for biodiversity conservation planning: the effectiveness of surrogates. *Annu. Rev. Ecol. Evol. Syst.* **38**, 713–737 (2007).
34. Andelman, S. J. & Fagan, W. F. Umbrellas and flagships: Efficient conservation surrogates or expensive mistakes?. *Proc. Natl. Acad. Sci.* **97**, 5954–5959 (2000).
35. Fielding, A. H. & Bell, J. F. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* **24**, 38–49 (1997).
36. Abellan, P., Sanchez-Fernandez, D., Velasco, J. & Millan, A. Conservation of freshwater biodiversity: a comparison of different area selection methods. *Biodivers. Conserv.* **14**, 3457–3474 (2005).
37. Fearnside, P. M. Conservation policy in Brazilian Amazonia: understanding the dilemmas. *World Dev.* **31**, 757–779 (2003).
38. dos Santos, A. J., Vieira, T. B. & Faria, K. de C. Effects of vegetation structure on the diversity of bats in remnants of Brazilian Cerrado savanna. *Basic Appl. Ecol.* **17**, 720–730 (2016).
39. Groves, C. R. *et al.* Planning for biodiversity conservation: putting conservation science into practice. *Bioscience* [https://doi.org/10.1641/0006-3568\(2002\)052\[0499:pfbcp\]2.0.co;2](https://doi.org/10.1641/0006-3568(2002)052[0499:pfbcp]2.0.co;2) (2002).
40. Fearnside, P. M. & Ferraz, J. A conservation gap analysis of Brazil's Amazonian vegetation. *Conserv. Biol.* **9**, 1134–1147 (1995).
41. Fearnside, P. M. Introduction: strategies for social and environmental conservation in conservation units. In *The Amazon Várzea* 233–238 (Springer Netherlands, 2011). doi:https://doi.org/10.1007/978-94-007-0146-5_16.
42. Cardoso, P., Erwin, T. L., Borges, P. A. V. & New, T. R. The seven impediments in invertebrate conservation and how to overcome them. *Biol. Conserv.* **144**, 2647–2655 (2011).
43. Marini, M. Â. & Garcia, F. I. Bird conservation in Brazil. *Conserv. Biol.* <https://doi.org/10.1111/j.1523-1739.2005.00706.x> (2005).
44. Young, B. E. *et al.* Population declines and priorities for amphibian conservation in Latin America. *Conserv. Biol.* **15**, 1213–1223 (2001).

45. Dias-Silva, K., Moreira, F. F. F., Giehl, N. F. D. S., Nóbrega, C. C. & Cabette, H. S. R. Gerromorpha (Hemiptera: Heteroptera) of eastern Mato Grosso State, Brazil: checklist, new records, and species distribution modeling. *Zootaxa* <https://doi.org/10.11646/zootaxa.3736.3.1> (2013).
46. Ferraz, K. M. P. M. de B., Ferraz, S. F. de B., Paula, R. C. de, Beisiegel, B. & Breitenmoser, C. Species Distribution Modeling for Conservation Purposes. *Nat. Conserv.* **10**, 214–220 (2012).
47. Marco-Júnior, P. & Siqueira, M. F. Como determinar a distribuição potencial de espécies sob uma abordagem conservacionista? *Megadiversidade* (2009).
48. Hijmans, R. J. *et al.* *DIVA-GIS, version 5.2. A geographic information system for the analysis of biodiversity data. Manual.* . vol. 1 (International Potato Center, 2005).
49. Borcard, D., Gillet, F. & Legendre, P. *Numerical Ecology with R. Numerical Ecology with R* (Springer New York, 2011). doi:<https://doi.org/10.1007/978-1-4419-7976-6>.
50. Serra, B. D. V., De Marco, P. J., Nóbrega, C. C. & Campos, L. A. D. O. Modeling potential geographical distribution of the wild nests of *Melipona capixaba* Moure & Camargo, 1994 (Hymenoptera, Apidae): Conserving Isolated Populations in Mountain Habitats. *Nat. e Conserv.* **10**, 199–206 (2012).
51. Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* <https://doi.org/10.1016/j.ecolmodel.2005.03.026> (2006).
52. Swets, J. Measuring the accuracy of diagnostic systems. *Science* (80-.). **240**, 1285–1293 (1988).
53. Girardello, M., Griggio, M., Whittingham, M. J. & Rushton, S. P. Identifying important areas for butterfly conservation in Italy. *Anim. Conserv.* <https://doi.org/10.1111/j.1469-1795.2008.00216.x> (2009).
54. Allouche, O., Tsoar, A. & Kadmon, R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **43**, 1223–1232 (2006).
55. Vieira, T. B., Mendes, P. & Oprea, M. Priority areas for bat conservation in the state of Espírito Santo, southeastern Brazil. *Neotrop. Biol. Conserv.* **7**, 88–96 (2012).
56. Delgado-Jaramillo, M., Aguiar, L. M. S., Machado, R. B. & Bernard, E. Assessing the distribution of a species-rich group in a continental-sized megadiverse country: Bats in Brazil. *Divers. Distrib.* **26**, 632–643 (2020).
57. Destro, G. F. G., de Fernandes, V., de Andrade, A. F. A., De Marco, P. & Terribile, L. C. Back home? Uncertainties for returning seized animals to the source-areas under climate change. *Glob. Chang. Biol.* **25**, 3242–3253 (2019).
58. Elith, J. *et al.* Novel methods improve prediction of species' distributions from occurrence data. *Ecography* (Cop.). (2006) doi:<https://doi.org/10.1111/j.2006.0906-7590.04596.x>.
59. Pearson, R. G., Raxworthy, C. J., Nakamura, M. & Townsend Peterson, A. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J. Biogeogr.* **34**, 102–117 (2006).
60. de Andrade, A. F. A., Velazco, S. J. E. & De Marco, P. Niche mismatches can impair our ability to predict potential invasions. *Biol. Invasions* **21**, 3135–3150 (2019).
61. Velazco, S. J. E., Villalobos, F., Galvão, F. & De Marco Júnior, P. A dark scenario for Cerrado plant species: Effects of future climate, land use and protected areas ineffectiveness. *Divers. Distrib.* **25**, 660–673 (2019).
62. Baselga, A. Partitioning the turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.* <https://doi.org/10.1111/j.1466-8238.2009.00490.x> (2010).
63. Moilanen, A. *et al.* Prioritizing multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proc. R. Soc.* **272**, 1885–1891 (2005).
64. Moilanen, A. *et al.* *Zonation spatial conservation planning framework and software v. 3.1, User manual.* (2012).
65. Moilanen, A. Landscape zonation, benefit functions and target-based planning: unifying reserve selection strategies. *Biol. Conserv.* **134**, 571–579 (2007).
66. Carvalho, A. R. de. Método de Monte Carlo e Aplicações. *Repositório Inst. da Univ. Fed. Flum.* **84** (2017).
67. Feinleib, M. & Zar, J. H. Biostatistical analysis. *J. Am. Stat. Assoc.* <https://doi.org/10.2307/2285423> (1975).

Acknowledgements

KDS is thankful to the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for granting her a postdoctoral scholarship (Process No 150288/2015-8). LJ and NH received continuous research support from CNPq productivity grants (Processes 304710/2019-9 and 308970/2019-5); NH is supported by INCT ADAPTA II funded by CNPq (465540/2014-7), FAPEAM (0621187/2017). TBV received a scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES). This study was financed in part by CAPES (Finance Code 001). We thank Philip Fearnside for the English correction.

Author contributions

Karina Dias da Silva—Write original draft; Conceptualization. Thiago Bernardi Vieira—Data analyses; Review and editing draft. Felipe Ferraz Figueiredo Moreira—Data Curation; Review and editing draft. Leandro Juen—Review and editing draft. Neusa Hamada—Supervision; Conceptualization; Review and editing draft.

Funding

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)—Process No 150288/2015-8, 304710/2019-9 and 308970/2019-5. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-00700-0>.

Correspondence and requests for materials should be addressed to T.B.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021