PRIORITY AREAS FOR ANURAN CONSERVATION USING BIOGEOGRAPHICAL DATA: A COMPARISON OF GREEDY, RARITY, AND SIMULATED ANNEALING ALGORITHMS TO DEFINE RESERVE NETWORKS IN CERRADO


1Laboratório de Ecologia Teórica e Síntese, Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Goiás, CEP 74001-970, Goiânia, GO, Brazil
2DBG, ICB, UFG, CEP 74001-970, Goiânia, GO, Brazil
3Universidade Estadual de Goiás, Departamento de Biologia, Laboratório de Biodiversidade do Cerrado, Av. Juscelino Kubitschek, 146, Bairro Jundiá, CEP 75110-390, Anápolis, GO, Brazil
4Programa de Pós-graduação em Biologia, ICB, UFG, CEP 74001-970, Goiânia, GO, Brazil

Correspondence to: José Alexandre Felizola Diniz-Filho, UFG, ICB, DBG, CP 131, CEP 74001-970, Goiânia, GO, Brazil, e-mail: diniz@icb1.ufg.br

Received July 21, 2003 – Accepted October 27, 2003 – Distributed May 31, 2005

(With 6 figures)

ABSTRACT

Spatial patterns in biodiversity variation at a regional scale are rarely taken into account when a natural reserve is to be established, despite many available methods for determining them. In this paper, we used dimensions of occurrence of 105 species of Anura (Amphibia) in the cerrado region of central Brazil to create a regional system of potential areas that preserves all regional diversity, using three different algorithms to establish reserve networks: “greedy”, rarity, and simulated annealing algorithms. These generated networks based on complementarity with 10, 12, and 8 regions, respectively, widely distributed in the biome, and encompassing various Brazilian states. Although the purpose of these algorithms is to find a small number of regions for which all species are represented at least once, the results showed that 67.6%, 76.2%, and 69.5% of the species were represented in two or more regions in the three networks. Simulated annealing produced the smallest network, but it left out three species (one endemic). On the other hand, while the greedy algorithm produce a smaller solution, the rarity-based algorithm ensured that more species were represented more than once, which can be advantageous because it takes into consideration the high levels of habitat loss in the cerrado. Although usually coarse, these macro-scale approaches can provide overall guidelines for conservation and are useful in determining the focus for more local and effective conservation efforts, which is especially important when dealing with a taxonomic group such as anurans, for which quick and drastic population declines have been reported throughout the world.

Key words: anurans, cerrado, complementarity, reserve network, reserve selection, greedy algorithm, rarity, richness.

RESUMO

Áreas prioritárias para a conservação de anuros usando dados biogeográficos: uma comparação dos algoritmos de “greedy”, raridade e “annealing” para a definição de redes de reservas no Cerrado

Os padrões espaciais da variação da biodiversidade em escalas regionais raramente são considerados na escolha de uma reserva (unidades de conservação), a despeito dos diversos métodos disponíveis para esse fim. Neste trabalho, usamos dados da extensão de ocorrência de 105 espécies de Anura (Amphibia) na região de Cerrado, no Brasil Central, para estabelecer um sistema regional de áreas...
potenciais que preserve toda a diversidade da região. Para tanto, três algoritmos de seleção de redes de reservas foram testados: “greedy”, raridade e “annealing” simulado. Com base na complementação, esses algoritmos geraram redes com 10, 12 e 8 regiões, respectivamente, sendo estas amplamente distribuídas no bioma e contemplando diferentes Estados da União. Esses algoritmos buscam o menor número de regiões que representem todas as espécies pelo menos uma vez. No entanto, 67,6%, 76,2% e 69,5% das espécies foram representadas (“preservadas”) em duas ou mais regiões nas três redes de reservas selecionadas pelos métodos citados. O algoritmo “annealing” simulado resultou na menor rede, no entanto, não incluiu áreas para a preservação de três espécies (sendo uma delas endêmica).

Por outro lado, embora o algoritmo de “greedy” tenha resultado em menor solução, o método baseado na raridade garante que mais espécies sejam representadas mais de uma vez, o que pode ser vantajoso considerando a elevada taxa de perda de habitat no Cerrado. Embora imprecisas, as abordagens em grandes escalas espaciais podem indicar estratégias gerais para a conservação e definir esforços de conservação locais e mais efetivos. Isto é especialmente importante quando se trata de um grupo taxonômico como os anuros, com declínios populacionais acentuados registrados mundialmente.

**Palavras-chave:** anuros, Cerrado, complementação, redes de reservas, seleção de reservas, algoritmos, raridade, riqueza.

**INTRODUCTION**

Delimiting conservation units (reserves) is the main strategy, although sometimes questioned, adopted by most countries in trying to preserve biodiversity (Margules & Pressey, 2000; Aaron et al., 2001). However, for an entire region (country, state, or biome), the group of reserves (i.e., reserve system or network) is usually not defined with respect to specific biodiversity objectives (Possingham et al., 2000). Many reserves encompass areas of unsuitable habitat or are defined based on cultural or scenic reasons. However, at least in theory, the most important criterion in defining a reserve network is its capacity to deliver maximum biodiversity at the smallest possible overall cost.

At present, many algorithms have already been developed to create a reserve network that maximizes regional biodiversity representation (see Cabeza & Moilanen, 2001, for a recent review). These methods can be based simply on trying to preserve biodiversity data for a group of species (Church et al., 1996; Pressey et al., 1997; Araújo & Williams, 2000; Polasky et al., 2000; Briers, 2002). More recently, however, they have been built on optimization of other measures of biodiversity, such as phylogenetic diversity (Polasky et al., 2001; Rodrigues & Gaston, 2002; Sechrest et al., 2002). Also, socio-economic factors including population size, growth rate, and land use associated with the development of human populations at local and regional scales should be taken into account (Abbitt et al., 2000). These human development parameters, which are associated with endemism, have recently been used to define priority areas for conservation (the so-called hotspots; see Myers et al., 2000) at a global scale.

There are basically two groups of algorithms used to define networks of reserves: those based on heuristic criteria and those based on optimization (computational) procedures. Heuristic criteria are based on simple rules, and their use in selecting sites to be included in a reserve network is based on a sequential algorithm determined by the relative importance of each site, as assessed by complementarity (i.e., the relative gain obtained by adding that site to the network). One of the most commonly used heuristic algorithms starts with the richest site in an area and then adds other sites to the network, increasing as far as possible at each step the number of species not previously included in the network (the “greedy” procedure). The other method ranks the sites based on the rarity of the species found there and, appropriately, begins the networks with the sites containing the rare species, regardless of the richness of these areas.

On the other hand, there are now many sophisticated optimization algorithms based on either linear programming or more complex computational procedures (i.e., the Metropolis-Hasting networks) that find optimal, non-sequential solutions for an entire network (Possingham et al., 2000; Cabeza & Moilanen, 2001). Although heuristic algorithms usually produce sub-optimal solutions for the network (i.e., they usually have...
5% to 10% more sites in the network than those obtained by computational procedures; see Pressey et al., 1997), they can easily be applied to very large data matrices and produce sequential solutions (Csuti et al., 1997; but see Rodrigues & Gaston, 2002). At the same time, there are no guarantees that computational algorithms based on iterative linear programming (ILP) find optimal networks when more than 20 or 30 sites are under study, and most landscape-scale conservation problems deal with hundreds, or even thousands, of units. One solution to this problem, however, is to use iterative methods such as the simulated annealing algorithm, despite the sizeable computational effort entailed (Possingham et al., 2000).

Previous attempts to establish conservation priorities in cerrado area have usually been based on subjective criteria, presumed high species richness, and unknown regions (see Colli et al., 2002; Oliveira & Marques, 2002). In this paper, we compared the greedy, rarity-based, and simulated annealing algorithms to establish priority areas for conservation, in the cerrado area of Central Brazil, based on occurrence of Anura (Amphibia) species. We used macro-scale biogeographical data (extension of occurrence, sensu Gaston, 1994) to define the most important regions that should be included in a reserve system designed to protect all species. Although macro-scale approaches are usually coarse, they can provide overall guidelines useful in defining the focus for not only general but also more local and more readily effective conservation efforts in Neotropical regions (Young et al., 2001). This hierarchical approach may be a starting point in prioritizing conservation efforts in a biome where political decisions are usually more common than those determined by scientifically-based criteria. Also, since amphibian diversity is at present severely threatened by habitat destruction (Sala et al., 2000), this approach can be considered a handily available starting point for the more detailed studies that are urgently required (Young et al., 2001, and references therein).

**METHODS**

We limited our analyses to 105 species of Anura (Amphibia), all endemic to South America, that can be found in the cerrado area of Central Brazil (Fig. 1). Species of undefined taxonomic status or that probably constitute groups of species (i.e., Scinax gr. rubra or Bufo gr. granulosus) were excluded from our dataset (see Diniz-Filho et al., in press (a), for details). The occurrence of these species in South America was registered on a standardized map thereof (Azimuthal projection, scale 1:40,000,000) covered by a grid with 780 quadrats with sides of approximately 135 km (Bini et al., 2000; Diniz-Filho et al., 2002), based on extensive data in the literature as well as on records of the Museu Nacional (Universidade Federal do Rio de Janeiro). The cerrado biome in this grid was then delineated based on a map of Instituto Brasileiro de Geografia e Estatistica (IBGE) and UNESCO (1981) (Fig. 1). The extents of occurrence were determined based on minimum convex polygons that signaled occurrence of each species in South America (see Gaston, 1994). For most species, more than 10 records from various positions of the continent were available to establish geographic ranges. (Data on extents of occurrence of each species analyzed are available from the main author upon request.)

Grid diversity was estimated by counting the overlaps of extents of occurrence of 105 species, in each of the 82 quadrats in the cerrado biome. To establish a hierarchical scheme, in this paper each quadrat will be considered as a separate region encompassing many local areas, which in turn should support minimum viable populations for the species considered.

Based on the occurrence of 105 species in the 82 quadrats of the cerrado biome, we initially used a greedy algorithm to select the most representative regions for conserving all species with minimum cost (Possingham et al., 2000; Polasky et al., 2001). This algorithm is very simple and starts by identifying the region with highest species richness (grid diversity). It then sequentially searches for the next region that adds most species not found in the previous region. The final purpose is to create a network with a minimum number of regions in which all species are represented at least once (the set covering problem) (Church et al., 1996).

We applied to the same dataset a rarity-based algorithm, which initially ranked the quadrats by the sum of rarity of the species contained by each (see Possingham et al., 2000). The rarity index of a species (here called $R_s$) is given by the inverse of its representation in the entire area (i.e., $1/C$, where $C$ is the number of sites).
where C is the number of quadrats occupied by the species in the cerrado area). Therefore, if a species is found in a single quadrat, its $R_s$ is equal to 1 (rarest species), while a species found in the entire cerrado has a $R_s$ equal to 0.0122 (1/82), the minimum $R_s$ value in this case. Once $R_s$ values are obtained for all species, by summing them for each quadrat, the overall rarity of the quadrat ($R_q$) is obtained. Based on the $R_q$ values, the quadrats in the network were then chosen sequentially and the number of new species added to the network at each step was evaluated.

For the two algorithms used, ties in quadrat selection at each step of the algorithm (especially in the greedy algorithm) were broken by adding the richest region to the network (thus increasing the representativity of the species in the final network) (see also Briers, 2002). Number of new species added to the system can be monitored and the curve can be used to define a minimum cut-off level. Possingham et al. (2000) pointed out that all sequential algorithms are inefficient with respect to guaranteeing optimal solutions.

Thus, we applied the simulated annealing algorithm as implemented in SITES 1.0 software. Simulated annealing (Kirkpatrick et al., 1983) is a computational method that iteratively explores trial solutions for the set covering problem, making sequential random changes in the network. At each iteration, a random quadrat is added (or a previously included one is excluded), the solution is compared with the previous one, and the best is chosen. The SITES software generates many solutions and the best one (minimum number of quadrats with maximum of species preserved) was used. All species and quadrats have the same weight in establishing a network (penalties were set using values of one for all species and quadrats).

*Fig. 1 — Spatial patterns of species richness in 105 species of anurans from the cerrado region, interpolated using a distance weighted least-squares (DWLS) algorithm.*
In the three methods used, the purpose is to keep all species represented at least once in the network (the target). Following Kelley et al. (2002), the three solutions were compared using Jaccard coefficients.

RESULTS

The peak in species diversity for Anuran species was found in the central-southern region of the cerrado biome, as shown by mapping the grid diversity based on extents of occurrence (Fig. 1). The question in using reserve network algorithms is, given these diversity patterns, how to preserve most of biodiversity with the smallest number of regions (quadrats).

The application of the greedy algorithm to the occurrence matrix indicated that 10 regions (quadrats) must be preserved in order to retain all diversity in the biome. These regions were widely distributed in the entire cerrado region and encompass the states of Goiás, Minas Gerais, Tocantins, Bahia, Maranhão, Mato Grosso, and Mato Grosso do Sul (Fig. 2A). The proportion of the species pool added, in relation to sequential introduction of regions in the network system (Fig. 3), indicates that preservation of the first region alone (the region of Alto Paraíso de Goiás, Goiás State, with 62 species), will maintain around 60% of the overall species pool. After increasing the size of the network system to 6 regions, 95% of the pool will be retained.

The rarity-based algorithm, on the other hand, suggested a slightly different and geographically more extensive solution, with 12 quadrats also widely distributed in the biome (Fig. 2B). The analysis of $R_s$ values (rarity for the 105 species) revealed that most species are found in a few quadrats of the cerrado, with a few species with a wide distribution with maximum $R_s = 1.0$ (Fig. 4A). Also, there is a nonlinear positive relationship between total geographic range size and the respective geographic range size within the cerrado biome (Fig. 4B). After summing these $R_s$ values to obtain $R_q$ of each quadrat and sequentially selecting quadrats with the highest $R_q$ values, the final network has 12 quadrats, with a slightly more irregular increase of species richness preserved in the entire system (Fig. 3). Also, there is a marked decrease in the $R_q$ values when constructing the network, suggesting that there are a few quadrats concentrating high overall rarity (Fig. 5). Simulated annealing produced a shorter, nonsequential solution, with 8 quadrats (Fig. 2C).

There are 14 common quadrats (regions) in the three networks, and their spatial distributions are similar (Fig. 2). The pairwise Jaccard coefficient ($J$) between them indicates that the greedy is more similar to the simulated annealing ($J = 0.64$) than it is to the rarity-based algorithm ($J = 0.57$). Curiously, although the purpose of the algorithms is to find a small number of regions in which all species are represented at least once, many species appear more than once because of their relatively large extents of occurrence in the area (i.e., low rarity). So, it is possible to check for representation of each species in the established networks (Fig. 6) and, in fact, 67.6%, 76.2%, and 69.5% of the 105 species were represented more than once in the greedy, rarity-based, and simulated annealing networks, respectively.

DISCUSSION

Conservation networks defined by optimal or sub-optimal complementarity solutions, based on regional biodiversity analyses, have been successfully implemented or proposed for different parts of the world (see Csuti et al., 1997; Araújo, 1999). The analyses performed here show that, depending on the strategy adopted, conservation efforts for the anurans in the cerrado biome should be concentrated in at least 10-12 different central Brazilian regions (located in the states of Goiás, Minas Gerais, Tocantins, Bahia, Maranhão, Mato Grosso, and Mato Grosso do Sul). These regions are located far from one another, as expected if beta-diversity increases with increasing geographic distance (Maurer, 1994).
Fig. 2 — Spatial distribution of reserves in the networks obtained by greedy (A), rarity (B), and simulated annealing (C) algorithms, based on the occurrence of 105 species of anurans in the cerrado region. For the two sequential algorithms (A and B), the numbers indicate conservation priority.
All networks suggested by our analysis encompass the environmental heterogeneity in the cerrado biome, including vegetation types, macro-climatic, and relief variation (see Ratter & Dargie, 1992; Ratter et al., 1996). Interestingly, some regions (quadrats) selected by the algorithms used in this paper included various of those previously selected by Cavalcanti & Joly (2002), e.g., Mirador, Chapada dos Veadeiros, and the Cerrado-Pantanal corridor. This suggests that some regions are so important for conservation purposes that they will be prioritized independently of the methods applied. Furthermore, our study partially supports the guidelines suggested by the Ministério do Meio Ambiente regarding prioritized regions and research funding, for example those cited in the Projeto de Conservação e Utilização Sustentável da Diversidade Biológica Brasileira (PROBIO).

The three network systems established, which use greedy, rarity-based, and simulated annealing algorithms are similar, despite differences in their ecological assumptions. It would be difficult to choose among them, although a few guidelines for doing so could be established. The shorter solution, obtained by simulated annealing (8 quadrats), is problematic since it omits three species, one of them endemic to the biome, thus supporting Kelley et al. (2002) who recently found that heuristic, sequential algorithms in some instances may be advantageous in relation to optimized global solutions. In addition, sequential solutions may be preferable for practical applications since they result in a priority rank that is possibly useful in dealing with hotspots (such as cerrado) or with organisms, such as anurans, that are very sensitive to environmental changes. However, simulated annealing may contribute to further improvements in the network design, for example when it becomes necessary to deal with variable costs for implementing a reserve in different regions or when faced with different values for each species (based on phylogenetic importance or rarity, e.g., see Rodrigues & Gaston, 2002).

![Graph](image-url)  
**Fig. 3** — Cumulative curve of species preserved when regions (quadrats) are successively included in the greedy and rarity networks shown in Fig. 2. The arrow indicates the relative position of the solution by simulated annealing, which produces a single, non-sequential solution with a fixed number of units (in this case, 8 units preserving 97.14% of the species).
Fig. 4 — (A) Distribution of rarity (1/geographic range in the cerrado) for the 105 species studied and; (B) the relationship between geographic range within the cerrado and overall geographic range (in number of quadrats).

Of the two sequential methods applied, the main advantage of the greedy solution is the smaller number of regions (10 quadrats compared with 12 for the rarity-based algorithm), a key factor if the establishment of a reserve network depends on acting within the limits of existing economic constraints or if most of the area is already the scene of ongoing human activity (both are realistic assumptions in the specific case of cerrado; see Myers et al., 2000). On the other hand, the rarity-based algorithm affords more flexibility because it preserves a large proportion of species more than once in the network, which is a desirable feature if the biome has been widely taken over by human beings and exploited for economic ends, and if maintaining viable populations for all species within each quadrat is in doubt (see below). Conservatively, a combination of the three networks could be developed so that a total of 14 regions would be used.

Flexibility in reserve design is also an important issue, especially in view of the fact that planners usually work in areas in which many conservation units already exist. For example, the first region selected by the greedy algorithm

includes a National Park (Parque Nacional das Chapadas dos Veadeiros). Furthermore, the Mirador National Park is approximately 200 km distant from a region defined by all algorithms in the northern part of the biome. The third region selected by the greedy algorithm (close to the municipality of Pedro Gomes, MS), is situated near the Cerrado-Pantanal Corridor, now under scientific evaluation by Conservation International (CI) (Leandro Silveira, personal communication). Thus, after a more detailed evaluation of local demographic and ecological parameters, these previously established units could be used as starting points for defining priorities in conservation efforts along the entire system (an option allowed by SITES, but based on more detailed planning, including that of landscape). However, a better understanding of how the current system of conservation units preserves anuran diversity requires more detailed study due to differences in scale and focus of each analysis (see below).

Macro-scale approaches can provide overall guidelines for conservation and be helpful in determining the focus of local conservation efforts in Neotropical regions. This is especially important when dealing with a taxonomic group, such as anurans, for which sudden and drastic population declines have been reported globally (Alford & Richards, 1999; Sala et al., 2000; see especially Young et al., 2001, for a recent review of amphibian declines in Latin America). While macro-scale approaches are usually coarse, they are valuable as starting points for strategies elaborated at a local scale (see Diniz-Filho et al., in press(b)).

In the interest of improving research aimed at conservation through reserve networks, we believe that two steps in this research program are necessary: (1) More intensive sampling within each region for the purpose of defining local areas of suitable habitat and estimating population and metapopulation parameters of species found therein; and (2) the characteristics of the entire region under analysis (i.e., cerrado) must be expanded to include landscape patterns according to land use, occupation, existence of previously established conservation units, and also species-level traits, such as rarity, local abundance, and evolutionary history.

Fig. 5 — Decrease in maximum rarity with sequential increase in the network size based on R_q, showing that when the first quadrats with high total rarity are included into the network, maximum rarity across the remaining quadrats quickly decreases.
Our analyses revealed for the cerrado region general patterns of species diversity of Anuran species that are important for conservation purposes. In general, the methods used here produced similar solutions, the best one of which was produced by the simplest algorithm (i.e., greedy). Simulated annealing, however, may be useful in so far as it adds more complex aspects of the landscape and variable weights for the species. The regional systems presented here can furnish guidelines for future anurans-conservation programs based on species diversity patterns, endemism, human development, and land use.

Acknowledgements — We thank Márcia Christianne de Souza for help in data collection, and Thiago Rangel for help with the maps. Financial support for this study is part of a regional program directed to researchers in central Brazil (PCOPG/CNPq, 520804/99-6 and SECTEC/CNPq proc. 23234156). The work of JAFDF, RPB, and LMB was also supported by CNPq productivity grants. Our research program in macroecology and biodiversity was supported throughout by CAPES/MEC and by the Fundação de Amparo à Pesquisa of the Federal University of Goiás (FUNAPE/UFG).

REFERENCES


UNESCO, 1981, Vegetation maps of South America. 1: 5,000,000, Paris, 2 sheets.