

Additions of landscape metrics improve predictions of occurrence of species distribution models

Érica Hasui¹ · Vinícius X. Silva¹ · Rogério G. T. Cunha¹ · Flavio N. Ramos¹ · Milton C. Ribeiro² · Mario Sacramento^{1,5} · Marco T. P. Coelho^{1,3} · Diego G. S. Pereira⁴ · Bruno R. Ribeiro^{1,3}

Received: 25 August 2016 / Accepted: 27 September 2016
© Northeast Forestry University and Springer-Verlag Berlin Heidelberg 2017

Abstract Species distribution models are used to aid our understanding of the processes driving the spatial patterns of species' habitats. This approach has received criticism, however, largely because it neglects landscape metrics. We examined the relative impacts of landscape predictors on the accuracy of habitat models by constructing distribution models at regional scales incorporating environmental variables (climate, topography, vegetation, and soil types) and secondary species occurrence data, and using them to predict the occurrence of 36 species in 15 forest fragments

where we conducted rapid surveys. We then selected six landscape predictors at the landscape scale and ran general linear models of species presence/absence with either a single scale predictor (the probabilities of occurrence of the distribution models or landscape variables) or multiple scale predictors (distribution models + one landscape variable). Our results indicated that distribution models alone had poor predictive abilities but were improved when landscape predictors were added; the species responses were not, however, similar to the multiple scale predictors. Our study thus highlights the importance of considering landscape metrics to generate more accurate habitat suitability models.

Project funding: This work was supported by the Biota Minas Program (Proc. No. APQ 03549- 09) and FAPEMIG (Proc. No. PCE-00106-12).

The online version is available at <http://www.springerlink.com>

Corresponding editor: Hu Yanbo

✉ Rogério G. T. Cunha
rogcunha@hotmail.com

¹ Laboratório de Ecologia de Fragmentos Florestais (ECOFRAG), Instituto de Ciência da Natureza, Universidade Federal de Alfenas, Rua Gabriel Monteiro da Silva, 700, Alfenas, MG 37130-000, Brazil

² Laboratório de Ecologia Espacial e Conservação (LEEC), Departamento de Ecologia, UNESP, Rio Claro. Av. 24A, 1515, Rio Claro, SP 13506-900, Brazil

³ Present Address: Programa de Pós-Graduação em Ecologia e Evolução da Universidade Federal de Goiás, Universidade Federal de Goiás, Goiânia, GO, Brazil

⁴ Departamento de Ciências Florestais, Universidade Federal de Lavras, Câmpus Universitário, Caixa Postal 3037, Lavras, MG 37130-000, Brazil

⁵ Present Address: Estação de Hidrobiologia e Piscicultura de Furnas – EHPF, Rua Lavras, 288 Bairro de Furnas, São José da Barra, MG CEP: 37947-000, Brazil

Keywords Ecological niche model · Generalized linear models · Habitat suitability · Landscape structure · Maxent

Introduction

Understanding the processes driving species distributions is fundamental for both theoretical and practical reasons, including conservation planning (Rosenzweig 1995; Margules and Pressey 2000; Engler et al. 2004; Brotons et al. 2004). Distribution patterns are not always easy to discern or understand, however, because they are influenced by environmental variables operating at multiple spatial and temporal scales (Foltête et al. 2012). Limiting or regulating factors such as temperature, water, and soil composition have emerged as primary predictors of species distributions on broad spatial scales (Pulliam 2000), although the histories and frequencies of disturbances (whether natural or human-induced), types of land use, and the heterogeneity of landscape features (e.g., topography and habitats) may be more important in explaining species distribution patterns at finer scales (Guisan and Thuiller 2000).

There has been increasing use of predictive species distribution models (SDMs) since the early 1990s to describe and quantify species-environment relationships and then predict species occurrences (Guisan et al. 2006). These models relate species occurrences or abundances to a set of environmental variables using statistical methods such as multiple regression, classification techniques, environmental envelopes, or Bayesian approaches (Guisan and Zimmermann 2000). The fitted model is then projected onto a geographic space, providing spatially based probabilities of occurrence (Guisan et al. 2007).

SDM approaches (Guisan and Thuiller 2005) have not been consistently accurate, however, especially at finer scales or in fragmented landscapes (Titeux et al. 2007; Ashcroft et al. 2012; Foltête et al. 2012). A possible explanation for these results is that SDMs are based solely on limiting or regulating factors that operate at broad scales (principally climatic, topographic, and vegetation variables) (Guisan et al. 2006), while other ecological processes that can strongly interfere with species distribution patterns, such as human-induced habitat changes, biotic interactions, or accessibility (considered as the dispersal ability of a given species, sensu Peterson 2006) are considered to lesser extents in this approach (Elith et al. 2006; Ashcroft et al. 2012). If this view is correct, the addition of variables describing landscape contexts that influence species accessibility should improve the predictive performance of SDMs, independent of the taxonomic groups considered (Ashcroft et al. 2012; Foltête et al. 2012).

The absence of landscape variables may thus compromise SDM performance and increase the numbers of false positives and negatives (Hirzel and Le 2008) for several reasons: (1) local extinction: populations are more extinction-prone in remnants of otherwise suitable sites due to environmental or demographic stochasticity (Robinson and Quinn 1988; Henle et al. 2004); (2) fragmentation may disrupt species dispersal or movement, reducing re-colonization of otherwise suitable sites (Fahrig 2003); (3) metapopulation dynamics: local extinctions and re-colonization rates depend on patch sizes and their isolation, implying that a target species may become temporarily extinct at a particular suitable site, but re-colonization may occur if the patches are functionally connected in the landscape (Hanski 1998); (4) source/sink dynamics: depending on a species' prevalence and tolerance, individuals may be located outside the bounds of their potential distribution ranges as based on their ecological niche (Lawton 1996; Pulliam 2000).

To test our predictions, we created and compared models constructed with single regional scale predictors to models with predictors operating at landscape scales. We performed these comparisons for 36 species from four phylogenetically distinct taxonomic groups (amphibians, birds, primates, and spermatophytes). At the regional scale,

we built SDMs with environmental variables (considering climate, topography, vegetation, and soil types), using Maxent, and assessed their performance by using them to predict species occurrences in 15 selected fragments in which we conducted rapid surveys. We used the same occurrence data to build models with patch/landscape predictors at the landscape scale.

Materials and methods

Study region

We carried out this study in 15 forest fragments in southern Minas Gerais State (MG), Brazil, within an area of 65,000 km² within the Atlantic Forest domain (Fig. 1, Table 1). The vegetation at these sites was either semi-deciduous or ombrophilous forest. We employed two criteria for fragment selection: (1) the set of fragments was required to show a gradient of estimated species richness based on SDMs (see details in item c—SDM); (2) all fragments were required to be located within priority conservation areas in Minas Gerais State (Drummond et al. 2005). The only exception to criteria number 2 was one of the largest protected semi-deciduous fragments of Atlantic Forest in southern Minas Gerais State (the Pouso Alegre Municipal Natural Park) in the municipality of Pouso Alegre. We also included data from nine additional fragments near Alfenas, MG, for the bird species, to improve model quality. We collected data from these fragments using the methodologies described below.

Study design

Our study design employed seven principal steps: (a) selection of the study species; (b) a search for secondary data on the occurrence sites of those species; (c) building the SDMs using the Maxent algorithm; (d) selecting study sites across a gradient of species richness (low, medium, and high) using SDMs; (e) rapid surveys of species presence/absences within the selected sites; (f) computation of landscape metrics; (g) data analysis.

Selection of the study species

We selected 36 species from four phylogenetically distinct taxonomic groups (spermatophytes, amphibians, birds, and primates), according to the criteria described below.

Spermatophytes

We chose seven tree species a priori (*Campomanesia guazumifolia*, *Chrysophyllum gonocarpum*, *Chrysophyllum*

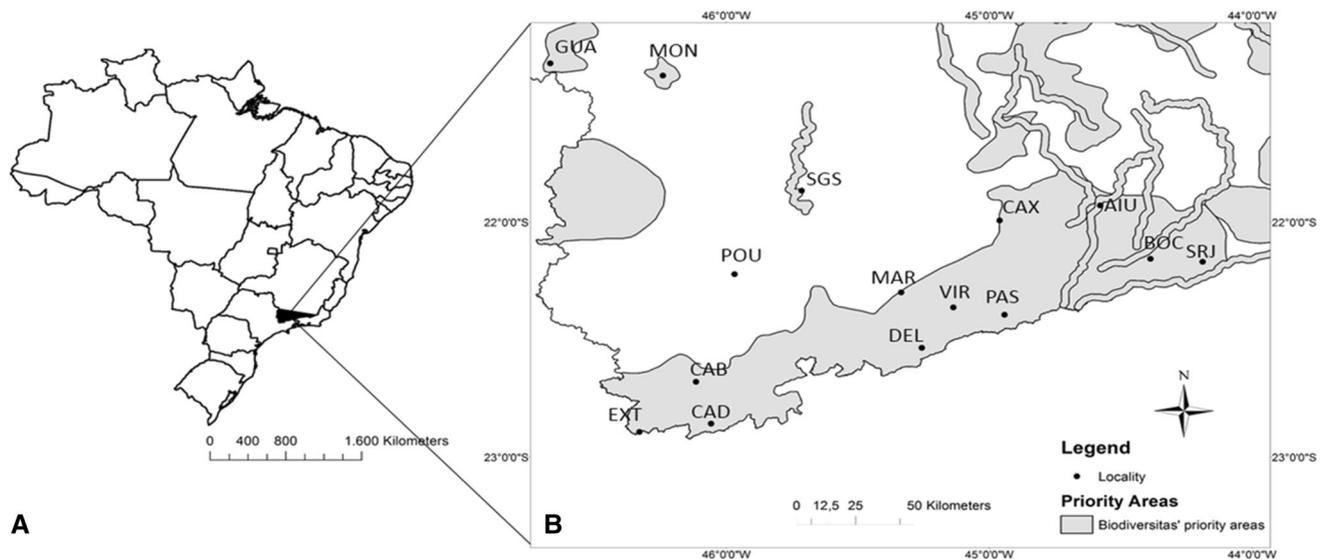


Fig. 1 Study area. **a** Location of Minas Gerais State in Southeastern Brazil, South America. **b** Sites sampled with rapid surveys in priority areas for conservation in Minas Gerais State (Drummond et al. 2005): (GUA) Guaxupé, (MON) Monte Belo, (SGS) São Gonçalo do Sapucaí, (POU) Pouso Alegre, Serra da Mantiqueira priority area

(EXT) Extrema, (CAD) Camanducaia, (CAB) Cambuí, (DEL) Delfim Moreira, (MAR) Maria da Fé, (VIR) Virgínia, (PAS) Passa Quatro, (CAX) Caxambu, (AIU) Aiuruoca, (BOC) Bocaina de Minas, (SRJ) Santa Rita do Jacutinga)

Table 1 Characteristics of the 15 sampled patches

Sampled locality	Forest type	Longitude	Latitude	NNDist (m)	AREA (ha)	NumP	MedPS (ha)	AWMPFD	Pland (%)
Guaxupé	SF	326,806	7,641,042	56.93	647.33	15	30,068.6	1.32	35.96
Monte Belo	SF	367,288	7,636,518	14.68	506.89	12	40,209.8	1.22	41.3
São Gonçalo do Sapucaí	SF	436,726	7,581,693	114.34	11.78	15	53,889.7	1.3	38.08
Pouso Alegre	SF	400,222	7,542,311	188.63	241.11	7	23,647.4	1.28	25.96
Extrema	OM	363,572	7,467,690	29.33	299.91	6	4861.52	1.33	65.73
Camanducaia	OM	391,485	7,471,603	28.85	3832.49	10	21,514.1	1.35	39.57
Cambuí	SF	385,460	7,491,535	72.32	17.3	20	36,970.8	1.31	8.61
Delfim Moreira	OM	473,705	7,507,979	19,196	5502.18	2	4,874,985	1.32	78
Maria da Fé	SF	461,716	7,536,448	45.83	125.57	15	81,888.1	1.32	37.61
Virgínia	OM	486,079	7,526,940	59.43	72.61	15	28,857.9	1.36	26.26
Passa Quatro	OM	506,106	7,523,553	93	81.95	16	69,960.5	1.31	22.12
Caxambu	SF	504,168	7,567,917	26.5	159.43	12	53,731.1	1.26	17.31
Aiuruoca	SF	543,681	7,574,960	33.44	70.97	12	28,751	1.35	19.14
Bocaina de Minas	SF	563,544	7,549,539	80.19	53.93	13	29,125.3	1.36	34.66
Santa Rita de Jacutinga	OM	583,824	7,548,284	26.44	111.27	10	7061.88	1.38	44.68

SF semideciduous forest, OM ombrophilous dense forest, NNDist distance from the nearest neighbor fragment, AREA fragment area, NumP number of patches in the landscape, MedPS average size of the fragments, AWMPFD fractal dimension of the fragment average weighted by the area, Pland percentage of forest cover in the landscape. Coordinates datum (SAD69/UTM/23South)

marginatum, *Cryptocarya aschersoniana*, *Eugenia florida*, *Lecythis pisonis*, and *Pouteria torta*) as prospective study species, as they are normally associated with relatively undisturbed forest environments in the focal region. We attempted to include species that encompassed a range of variations in terms of the following biological characteristics: (a) pollination syndromes, (b) sexual systems, (c) reproductive systems, (d) seed dispersal syndromes,

(e) flower sizes, (f) floral resources, (g) fruit sizes, and (h) successional class (Denslow 1980). Most of these species, however, were not encountered or were found in only a few fragments. Only *C. aschersoniana* was encountered with sufficient frequency to perform the analyses. We therefore chose 11 additional species a posteriori (*Amaioua intermedia*, *Dalbergia villosa*, *Eugenia sonderiana*, *Machaerium villosum*, *Miconia chartacea*,

Myrcia splendens, *Nectandra oppositifolia*, *Ocotea corymbosa*, *Prunus myrtifolia*, *Psychotria velloziana*, and *Tapirira obtusa*). In addition to the seven criteria cited above, we required that each species occur in at least five of the study fragments.

Amphibians

We selected six prospective amphibian species associated with natural forest environments in the study region (*Aplastodiscus leucopygius*, *A. perviridis*, *Haddadus binotatus*, *Ischnocnema guentheri*, *Proceratophrys boiei*, and *Vitreorana eurygnatha*). Each species had to comply with the following criteria (Fleishman et al. 2000): (1) it could not be extremely rare in terms of abundance; (2) it could not be a generalist in terms of habitat occupation; (3) it could not have a wide geographic distribution; and, (4) it needed to be moderately sensitive to human disturbance.

Birds

We selected ten bird species a priori: *Automolus leucophthalmus*, *Dysythamnus mentalis*, *Habia rubica*, *Lochmias nematura*, *Mackenziana severa*, *Pyriglena leucoptera*, *Sittasomus griseicapillus*, *Turdus albicollis*, *Xiphorhynchus fuscus*, and *Xenops rutilans*. We required the species to vary widely in terms of the following criteria: geographic distribution, dependence on forest habitats, dispersal ability, home-range size, and degree of micro-habitat specialization.

Primates

Given the low species richness of this group, we surveyed all four primate species known to naturally occur in the study region (*Alouatta guariba clamitans*, *Callicebus nigrifrons*, *Callithrix aurita*, and *Sapajus nigritus*). *Brachyteles arachnoides* may occur in the study area, but its rarity prevents it from being encountered in rapid surveys and it does not naturally occur over the entire study area.

Secondary data

We obtained occurrence data on the species studied here from the Virtual Reference Center on Environmental Information (CRIA), the specialized literature, and biological collections (Herbário do Departamento de Botânica da UFMG (Herbário BHCb), Herbário ESAL (Escola Superior de Agronomia de Lavras), Museu de Zoologia da Universidade de São Paulo; Museu Nacional do Rio de Janeiro/UFRJ; Museu de História Natural Capão da Imbuia de Curitiba/PMC; Museu de História Natural da Unicamp –

MHNU; Coleção Zoológica de Referência da Seção de Vírus Transmitidos por Artrópodos - Instituto Adolfo Lutz - Banco de aves; Terrestrial Vertebrate Specimens – The Museum of Vertebrate Zoology, Berkeley; the American Museum of Natural History, New York; Museu da Usina Hidrelétrica de Segredo de Foz do Jordão, COPEL, the Field Museum of Natural History de Chicago/EUA, Coleção de Aves do Museu do Parque Nacional do Itatiaia).

SDM using Maxent

We modeled species occurrences within the spatial range of six Brazilian states (Espírito Santos, Rio de Janeiro, Minas Gerais, São Paulo, Paraná, and Santa Catarina) between 19° S and 28° S, extending inland from the coast to 51°W.

The modeling process employed 19 bioclimatic variables from the WORLDCLIM Ver. 1.4 database (available at <http://www.worldclim.org/current.htm>) related to climate, topography, vegetation, and soil type, at a 30-arc second resolution (900 m resolution, 1790 columns × 1846 rows), projected using WGS 1984. We also calculated the potential evapotranspiration rates (PET ratio), following Loiselle et al. (2008), as the mean annual temperature (°C) divided by total annual precipitation (mm); this ratio was then multiplied by an empirically derived constant (approximately 60) (Holdridge et al. 1971). At a PET ratio of 1.0, potential evapotranspiration approximately equals total precipitation for an average year; values above 1.0 indicate increasing aridity and values below 1.0 indicate increasing humidity.

We extracted the topographic variables from Shuttle Radar Topography Mission (SRTM) maps (available at <http://www.worldclim.org/current.htm>). We used this layer to derive slope angles and slope aspects. We obtained vegetation and drainage data, as well as soil type, fertility and texture from Embrapa Solos maps (available at <http://mapoteca.cnps.embrapa.br/geoacervo>) at 1:5,000,000 scale, and re-sampled them to a resolution of approximately 1 km, to standardize all variable scales.

We performed Pearson correlation matrix analysis to exclude highly correlated predictors from the SDM by extracting the values of the 27 continuum variables from 50,000 randomly selected points in the study region. This we followed with Next, we performed correlation analysis between all possible pairs of variables to detect highly correlated pairs (correlation threshold Pearson's $r = 0.7$). In cases of correlation, we used only one of those variables in the modeling process. The final reduced data set comprised 18 continuum and two categorical variables (Table 2).

We used the Maximum entropy algorithm (Maxent version 3.0.6, available at www.cs.princeton.edu/

~schapire/maxent/) to model the geographic distributions of suitable habitats for the 36 selected species, using occurrence data set as the dependent variable and the 20 environmental variables as the predictors (for a detailed mathematical description of Maxent and its application to SDM, see Phillips et al. 2006). We calibrated the models using 70% of the occurrence data (randomly chosen), as a training sample, and evaluated them with the remaining 30% (test data). We replicated the data by bootstrapping 100 random partitions with replacements.

We used the methods described by Phillips et al. (2006) to evaluate presence/random data using Receiver Operator Characteristic (ROC) curve analysis. The ROC curve represents the relationship between the proportion of correctly predicted presences (sensitivity) and 1 minus the proportion of the correctly predicted absences (specificity). The Area Under the ROC curve (AUC) measures the ability of the model to correctly classify a species as present or absent. The AUC is interpreted as the probability that random false positives and random true positives are correctly predicted by the Maxent model; an AUC value of 0.5 indicates that model predictions were no better than random guessing. While there is no full consensus in the literature, AUC values >0.70 are generally considered a

baseline for model accuracy. According to Pearce and Ferrier (2000), values below 0.70 indicate poor model performance, as they suggest similar rates of correct and erroneous predictions; values between 0.70 and 0.90 indicate moderately useful models; and values exceeding 0.90 indicate excellent accuracy.

Study site selection

We built occurrence maps (binomial) for the 36 study species from the reclassification of the continuous raster maps of occurrence probabilities (ranging from 0 to 100 percent) provided by Maxent. We then established a threshold of 66% occurrence probability, with probabilities below this value indicating species absence, and above it, species presence. We then combined these species occurrence maps into a synthesis-map of species richness. We classified the municipalities represented in this synthesis-map into three categories of predicted richness: low (<12 species), medium (12–24) and high (>24); we then selected four municipalities from each richness category. Finally, we selected one fragment in each municipality to carry out the rapid ecological surveys of the four taxonomic groups.

Table 2 Environmental variables (18 continuum +2 categorical variables) used in the Maxent niche models to predict the distribution of 36 species in four phylogenetically distant taxonomic groups (spermatophytes, amphibians, birds and primates)

Environmental variable	Source	Data type	Unit	Resolution (arc-second)
BIO1 = annual mean temperature	Worldclim	Continuous	(°C * 10)	30
BIO2 = mean diurnal range (mean of monthly (max temp – min temp))	Worldclim	Continuous	(°C * 10)	30
BIO3 = isothermality (P2/P7) (*100)	Worldclim	Continuous	(°C * 10)	30
BIO4 = temperature seasonality (standard deviation *100)	Worldclim	Continuous	(°C * 10)	30
BIO7 = temperature annual range (P5–P6)	Worldclim	Continuous	(°C * 10)	30
BIO10 = mean temperature of warmest quarter	Worldclim	Continuous	(°C * 10)	30
BIO11 = mean temperature of coldest quarter	Worldclim	Continuous	(°C * 10)	30
BIO12 = annual precipitation	Worldclim	Continuous	(mm)	30
BIO13 = precipitation of wettest month	Worldclim	Continuous	(mm)	30
BIO16 = precipitation of wettest quarter	Worldclim	Continuous	(mm)	30
BIO18 = precipitation of warmest quarter	Worldclim	Continuous	(mm)	30
Potential evapotranspiration	Derived from Worldclim	Continuous	(mm)	30
Elevation	SRTM	Continuous	metros	30
Slope angle	Derived from SRTM	Continuous	degree	30
Slope aspect	Derived from SRTM	Continuous	degree	30
Soil	Embrapa Solos	Categorical	248 categories	
Texture	Embrapa Solos	Continuous	11 categories	
Drenagem	Embrapa Solos	Continuous	15 categories	
Fertility	Embrapa Solos	Continuous	9 categories	
Vegetation	Embrapa Solos	Categorical	47 categories	

Rapid survey methods

We applied the rapid ecological assessment method to identify key target species—not to conduct full species inventories (Sayre et al. 2000). This method inevitably results in a tradeoff between the efficiency of encountering target species and the area that can be covered—but it is recommended in studies with objectives such as ours (Sayre et al. 2000).

Spermatophytes

We employed the point-centered quarter method (Cottam and Curtis 1956) to survey the tree community. We distributed twenty points (4 individuals sampled per point) spaced 20 m apart along a 300 m transect at the central region of each forest fragment (total: 1200 individuals, 80 per fragment). We sampled the closest phanerophyte with a diameter at breast height ≥ 3 cm in each quadrant. We deposited voucher specimens in the herbarium of the Universidade Federal de Alfenas (UALF), and made species identifications by comparisons with herbarium specimens and the botanical literature. Nomenclature follows the Angiosperm Phylogeny Group (APG) system (APG III 2009).

Amphibians

We actively searched for amphibians in potential breeding sites (such as streams, wetlands, and temporary ponds) within the forest fragments and along forest edges and trails used to access the study sites (adapted from Crump and Scott 1994) for a specified time period (between 19:00 and midnight). We also used playbacks of frog and toad vocalizations to locate species that could not otherwise be seen or heard.

Birds

We employed a playback approach to survey bird species. This technique consisted of broadcasting bird vocalizations to simulate a territorial invasion by a conspecific (Falls 1981). If present in the vicinity, individuals of the target species will respond by calling and/or approaching the sound source. We established three playback points within each forest fragment, approximately 100 m from the forest edge and 200 m from one another, to ensure sampling independence. We made two visits on consecutive days to each sampling point from 07:00 to 10:00. We broadcast each vocalization for 1 min, followed by three minutes of silence (Boscolo et al. 2006). We carried out this protocol three consecutive times. The species sequence in a given session and the playback volumes were the same throughout the study.

Primates

We employed several complementary approaches for detection of primate species. For *Callithrix aurita*, *Callicebus nigrifrons* and *Sapajus nigrurus* we reproduced long range vocalizations that are known either to play roles in inter-group communication (long calls of another species of *Callithrix* and *C. nigrifrons* duets) or for maintaining contact between widely separated group members (*S. nigrurus* whistles), both within the forest and/or at forest edges. We played back up to four times at each playback point within the fragments (in different directions), to cover a 360° radius; at forest edges, we reproduced the calls in two directions (45° from the edge line). We stopped the playbacks of a given species as soon as a response was obtained. Some occurrences were recorded through direct visual or auditory contacts (Rosales-Meda 2007). We also interviewed farmers and people living or working near the fragments about the occurrence of primates. This was necessary because *Alouatta guariba clamitans* is quite insensitive to playbacks, and *C. aurita* and *S. nigrurus* might not respond to playbacks even if they are present. If the interviewees mentioned the occurrence of callithrichids, we would include these data only if we subsequently located and identified the species (given the occurrence of an exotic species of the genus in the region, *C. penicillata*). We also obtained information regarding the presence of primates in the Pouso Alegre and Passa Quatro localities by consulting management plans of those protected areas. Given the lower richness of this group, we did not restrict ourselves to the two-day sampling regimes, and if there were indications that a species might occur in a locality (through interviews, for example), but we were unsure about it (which *Callithrix* species, for example, or conflicting or apparently inaccurate reports), we returned to the fragments on other occasions to attempt to confirm the findings. We did not use doubtful occurrences.

Landscape metrics

We extracted landscape metrics from an Atlantic Forest vegetation map produced by SOS Mata Atlântica/INPE (2008), using Fragstats 3.0 (McGarigal et al. 2002), with the V-LATE and Patch Analyst 5.0 plug-in for ArcGIS. The mapping scale was 1:50,000 in a vector format, which we subsequently converted to raster in ArcGIS with 30 m spatial resolution to improve metric computations. We calculated landscape metrics within a circular buffer (2000 m radius) around the geometric center of each sample patch. We selected six non-correlated metrics to consider different species perceptions of landscape structures that could affect their occurrences: patch area, median patch size, number of patches, Euclidean nearest-neighbour

distance, area-weighted mean patch fractal dimension, and the percentage of forest cover (Table 3). To avoid spatial autocorrelation, we tested correlations among the metrics using the Pearson correlation test (coefficient >0.7).

Data analysis

We extracted the probability of occurrence of each species in each fragment from the SDMs. The continuous raster maps of occurrence probabilities and the landscape maps had different resolutions (cell sizes of 900 m × 900 m and 30 m × 30 m respectively). We thus re-sampled each of the 36 continuous raster maps of occurrence probabilities to a resolution of 30 m to standardize their resolutions. We then extracted the grid cell values of occurrence probabilities from each study fragment, assuming that the probability of the species occurring in the grid cell was related to its presence/absence data collected during the rapid survey.

We then ran a set of general linear models (GLM) for each studied species, with binomial distributions, which could be either simple or additive. In each model, species presence/absence was related to: (1) the probability of occurrence only (SDMs); (2) landscape metrics only; and, (3) additive models of the probability of occurrence and

landscape metrics, incorporating one landscape metric at a time.

We evaluated the likelihood of each model using the Akaike Information Criterion, corrected for small sample sizes (AICc) (Burnham and Anderson 1998). We ranked the models in ascending AICc order, and Δ AICc was calculated for each model as the difference between its own AICc and the lowest AICc value (best model). We considered all models with Δ AICc values less than 2 as equally plausible. We considered that a species produced a valid model if the null model was not plausible. We conducted the analyses using the R program, version 2.14.2 (R development core team 2012).

Results

The SDMs of all taxonomic groups had excellent accuracies, with mean training AUCs near 0.95 (± 0.05) (Suppl. Mat.). The models could therefore be considered “reliable” for predicting local presence or absence. When we compared the actual observations of occurrence to the predicted occurrences in the SDMs, however, the models were rarely plausible. Actually, just 3% of the species ($n = 1$) showed plausible models using SDM alone (Δ AICc < 2; Table 4).

Table 3 Metrics used to describe the sample patches or the landscape structure around them

Acronym	Metric	Description	Ecological relevance	Reference
Area	Patch area	Area (ha) of each patch	Patch area is related to minimum area requirements of species and therefore associated with the probability of local extinction	Blake (1983), Temple and Cary (1988)
MedPS	Median patch size	Median area of patches (ha) in the landscape	It is the same relevance as Patch area, but refers to the landscape metrics	Blake (1983), Temple and Cary (1988)
NumP	Number of patches	Total number of forest patches in the landscape –an indication of the extent of subdivision or forest fragmentation	Subdivision of patches may lead to population subdivision and has implications for metapopulation dynamics	Hanski (1998)
NNDist	Euclidean nearest-neighbour distance	Shortest straight-line distance (m) between a focal patch and its nearest neighbor patch of the same class—represents a measure of isolation (range ≥ 0 , without limit)	Patch isolation affects inter-patch movement of organisms and is related to colonization dynamics and rescue effects	Brown and Kodric-Brown (1977) and Martensen et al. (2008)
AWMPFD	Area-weighted mean patch fractal dimension	Patch shape complexity measure weighted by patch area (range $1 \leq \text{AWMPFD} \leq 2$). Low values are found when a patch has a compact quadrangular or rectangular form with relatively small perimeter relative to the area. If the patches are more complex and fragmented, the perimeter increases relative to the area, which causes a higher fractal dimension	Shape complexity is related to edge effects, which has implications for interior-sensitive species or affects inter-patch movements for the other species	Ewers and Didham (2006)
Pland	Percentage of forest	Percentage of the landscape comprised of forest	The amount of habitat and its configuration affects population persistence	Andr�n (1994) and Fahrig (2003)

The percentage of species with plausible models increased to 22% ($n = 8$ species) when we combined SDMs with landscapes metrics. There were expressive increments in the wAICc in these models as compared to the wAICc of the model with SDMs alone (Table 4), although we did not find similar responses across all of the taxonomic groups. These models showed good performances with 40% of the birds ($n = 4$ species), 33% of the amphibians ($n = 2$ species), 16% of the plants ($n = 2$ species), and 0% of the primates.

The spermatophytes showed the most diverse responses in relation to the best predictors (Table 4). For 42% of the species, the SDMs showed poor predictive abilities, even when we added landscape metrics. The best models for these species were composed of landscape metrics only. We also observed species (e.g., *Eugenia sonderiana*) that responded only to SDMs, or even species (e.g., *Calyptranthes clusiifolia*) that responded better to the combination of multiple scale predictors. The SDMs showed poor performances overall for primates, and we only found valid models using landscape metrics.

Discussion

SDMs have been intensively used in recent years to generate probability distributions of species (or group of species) at broad scales in ecological evaluations (Engler et al. 2004; Brotons et al. 2004; Elith and Leathwick 2009). Model performance in most of these studies was evaluated using AUC values (Hijmans 2012). In the present study, the AUC values indicated excellent abilities of the models to correctly distinguish between presence/absence in the sites based on secondary occurrence data. For most of our species, however, the AUC values did not reflect actual observations of occurrences at the 15 sites surveyed. Just one species had its occurrence adequately predicted by SDM alone. The poor predictive abilities of those models reflect the failure to take into account factors that are important in determining species distributions (Filz et al. 2013), such as those related to landscape contexts. The percentages of species that were adequately predicted increased when landscape metrics were added to the models. Contrary to our expectations, however, the number of species that benefited from such inclusion was still limited (22% of the species studied) and varied across (and within) taxonomic groups. Previous studies, likewise carried out at multiple spatial scales, have arrived at similar conclusions, namely, the greater efficiency of models incorporating predictors operating at multiple scales, as opposed to those with predictors working at only a single scale (Hopkins 2009; Cabeza et al. 2010; Ashcroft et al. 2012; Foltête et al. 2012). There seem to be two main explanations for this

result in the present work. The first explanation relates to the fact that we examined a highly fragmented landscape. SDM predicts species occurrences based on their fundamental niches (see Warren 2012 for a thorough evaluation of SDMs), but this approach assumes species distributions to be at equilibrium with current environmental conditions and that relevant environmental gradients have been adequately sampled (Thuiller et al. 2004). Most ecosystems, however, have experienced anthropogenic impacts over long periods of time and are also highly fragmented. Under such conditions, the presence of a given species in a fragment may depend less on its fundamental niche and more on the characteristics of the fragments themselves, and the landscapes in which they are located (Elith and Leathwick 2009). The SDMs may over-predict species distributions in these fragments. Adding landscape metrics to the models therefore adds the advantage of allowing human transformations of the landscape to be taken into account and incorporates species accessibility into the modeling process (sensu Peterson 2006).

The second explanation for the greater efficiency of models with multiple scale predictors is related to their spatial resolution (grain cell size). The grain describes properties of the data or the analysis, such as spatial accuracy and the precision of the data records. Recent studies have demonstrated the implications of using different spatial resolution data in SDMs (Ferrier and Watson 1997; Pearman et al. 2008), with predictor effects on species occurrence-accuracy being dependent on the spatial accuracy of the data (associated with the characteristics of the terrain and the species itself). Therefore, the finer the spatial resolution used, the better the predictive ability of SDMs to identify niche shifts. It is therefore likely that the poor predictive ability of SDMs encountered here was due to the 30-arc second resolution (900 m), which was too coarse to represent true high habitat heterogeneity inside each grid cell in the surveyed fragment. This heterogeneity is, however, considered in landscape approaches (30 m of resolution) and could be crucial to predicting current species occurrences.

This result suggests that multiple-scale predictor studies can allow us to consider relevant environmental variables that vary and effect species at different spatial scales (Elith and Leathwick 2009); they can also help identify species that are especially prone to shifts in their environmental niche (Pearman et al. 2008). In agreement with our hypothesis, the addition of landscape predictors improved the predictive performance of SDMs—although the number of species that benefited was limited, and their responses varied across taxonomic groups (amphibians, birds, primates, and spermatophytes, ranging from 0 to 40%); similar results were reported by Thuiller et al. (2004) and Meyer and Thuiller (2006).

Table 4 Logistic regression models of species probability of occurrence ranked according to the Akaike information criterion corrected for small samples (AICc)

Taxonomic group	Species	Model parameters	ΔAICc	wAICc	Taxonomic group	Species	Model parameters	ΔAICc	wAICc
Spermatophytes	<i>Amaiota guianensis</i>	Pland	0	0.3486	Birds	<i>Pyriglena leucoptera</i>	NNDist	0	0.2412
		null	3	0.0783			NNDist + Niche	2	0.0879
		Pland + Niche	3.1	0.0757			null	3.2	0.0495
	<i>Calyptrothanes clusifolia</i>	Niche	5.6	0.0207		<i>Sitassomus griseicapillus</i>	Niche	3.8	0.0356
		Pland + Niche	0	0.9657			log(Area)	0	0.3852
		Pland	8.4	0.0146			log(Area) + Niche	0.2	0.3465
	<i>Cryptocaria aschersoniana</i>	null	11.5	0.003		<i>Turdus albicollis</i>	null	16.4	<0.001
		Niche	13.8	<0.001			Niche	17.3	<0.001
		log(Area)	0	0.245			MedPS + Niche	0	0.1838
	<i>Eugenia sonderiana</i>	Pland	0.9	0.1537			NNDist + Niche	1.2	0.1029
		log(Area) + Niche	2.1	0.0846			Pland + Niche	1.4	0.0928
		null	3.3	0.0463			Niche	3.3	0.0351
	<i>Myrcia splendens</i>	Pland + Niche	3.5	0.0425			null	3.8	0.0273
		Niche	5.1	0.0189			MedPS	4.4	0.02
		Pland + Niche	0	0.1907			Pland	5.7	0.0106
<i>Nectandra oppositifolia</i>	Niche	0.2	0.1737			NNDist	5.9	0.0096	
	null	3.9	0.0261			Pland + Niche	0	0.2144	
	Pland	4.5	0.0128			log(MedPS) + Niche	1.9	0.083	
<i>Tapirira obtusa</i>	log(Area)	0	0.5092			NNDist	2.4	0.0647	
	log(Area) + Niche	3.2	0.1049			null	3.2	0.0453	
	null	5.6	0.0312			Niche	3.5	0.0364	
<i>Aplastodiscus leucopygius</i>	Niche	8.3	0.0081	Primates	<i>Alouatta guariba clamitans</i>	Pland	4.8	0.0197	
	NNDist	0	0.1887			Pland	0	0.618	
	MedPS	1.9	0.0726			Pland + Niche	3	0.136	
<i>Amphibians</i>		null	2	0.0695		<i>Sapajus nigritus</i>	null	9.7	0.0049
		NNDist + Niche	2	0.0681			Niche	12.4	0.0013
		MedPS + Niche	4.5	0.0191			AWMPFD	0	0.4794
<i>Amphibians</i>		Niche	4.5	0.0203			AWMPFD + Niche	3	0.1085
		Pland	0	0.3			null	9.5	0.0041
		Pland + Niche	3.1	0.1019			Niche	11.8	0.0013
<i>Amphibians</i>		null	9	0.0054					
		Niche	11.3	0.0017					
		NumP	0	0.521					
<i>Amphibians</i>		NumP + Niche	1.9	0.197					
		Niche	6.9	0.0165					
		null	6.9	0.0169					

Table 4 continued

Taxonomic group	Species	Model parameters	$\Delta AICc$	wAICc	Taxonomic group	Species	Model parameters	$\Delta AICc$	wAICc
<i>Ischnocnema guentheri</i>		AWMPFD + Niche	0	0.2127					
		AWMPFD	0.9	0.1352					
		null	4.1	0.028					
		Niche	6.5	0.0084					

Only the plausible models ($\Delta AICc < 2$), the null model (null), the model considering only the niche-based probability of occurrence (Niche) and the model with the Niche variable and the best landscape variable are shown. $\Delta AICc$ = difference between the $AICc$ values from a specific model and the model with the smallest $AICc$ value. wAICc = model weight. Codes of variables as in Table 3

Our findings must be interpreted with caution, however, because there were limitations built into our study design. First, the positive factor of encompassing a broad geographic region necessarily decreased the number of fragments that could be efficiently surveyed, and it will be important for future studies to determine the minimum number of fragments necessary to generate consistent habitat suitability models, seeking the best and most cost-effective approaches, whether for conservation or theoretical purposes. Employing abundance data is particularly important for very common species (such as *Callicebus nigrifrons*), as valid models cannot be generated for them using only presence/absence data. Finally, the a posteriori choice of the spermatophyte species in this study increased the chances of finding valid models (and most of the valid models in the present study were of spermatophytes)—but this does not change or invalidate our conclusions, because such an approach did not affect the specific variables that predicted species responses.

In conclusion, we recommend the careful use of SDMs and their AUC values to evaluate model performance, as well as the use of variables from both SDMs and landscape structure, independent of the focal taxonomic group. The combination of both types of variables will generate more accurate species occurrence models. Given the large intra- and inter-group variability of the parameters that predict species occurrences, we further recommend using several species from different groups when employing these models for conservation purposes.

Acknowledgements We thank Mainara Xavier Jordani and Diogo Borges Provete for their assistance in the field; the anonymous reviewers for critically reading the text; the Instituto Chico Mendes (ICMBio) for issuing the capture and transportation licenses (No. 10704-1; 22020-1); the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for its financial support through the Biota Minas Program (Proc. No. APQ 03549-09); and Roy Richard Funch who revised the English translation.

References

- Andrén H (1994) Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat: a review. *Oikos* 71:355–366
- APG. [Angiosperm Phylogeny Group] III (2009) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot J Linnean Soc* 161:105–121
- Ashcroft MB, French KO, Chisholm L (2012) A simple post hoc method to add spatial context to predictive species distribution models. *Ecol Model* 228:17–26
- Blake JG (1983) Trophic structure of bird communities in forest patches in East-Central Illinois. *Wilson Bull* 95(3):416–430
- Boscolo D, Metzger JP, Vielliard JME (2006) Efficiency of playback for assuring the occurrence of five birds species in Brazilian Atlantic Forest fragments. *An Acad Bras Ciênc* 78:629–644

- Brotans L, Thuiller W, Araujo MB, Hirzel AH (2004) Presence absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27:437–448
- Brown JH, Kodric-brown A (1977) Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology* 58:445–449
- Burnham KP, Anderson DR (1998) *Model selection and inference*. Springer, New York
- Cabeza M, Arponen A, Jaattela L, Kujala H, Van-Teeffelen H, Hanski I (2010) Conservation planning with insects at three different spatial scales. *Ecography* 33:54–63
- Cottam G, Curtis JT (1956) The use of distance measures in phytosociological sampling. *Ecology* 37:451–460
- Crump ML, Scott JRNJ (1994) Standard techniques for inventory and monitoring: visual encounter surveys. In: Heyer WR et al (eds) *Measuring and monitoring biological diversity: standard methods for amphibians*. Smithsonian Institution Press, Washington, pp 84–92
- Denslow JS (1980) Gap partitioning among tropical rainforest trees. *Biotropica* 12:47–55
- Drummond GM, Martins CS, Machado ABM, Sebaio FA, Antonini Y (2005) Biodiversidade em Minas Gerais: um atlas para sua conservação, 2nd edn. Fundação Biodiversitas, Belo Horizonte, p 222
- Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. *Annu Rev Ecol Syst* 40:677–697
- Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakasawa Y, Overton JMCCM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberón J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151
- Engler R, Guisan A, Rechsteiner L (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *J Appl Ecol* 41:263–274
- Ewers RM, Didham RK (2006) Confounding factors in the detection of species responses to habitat fragmentation. *Biol Rev* 81:117–142
- Fahrig L (2003) Effects of habitat fragmentation on biodiversity. *Ann Rev Ecol Syst* 34:487–515
- Falls JB (1981) Mapping territories with playback: an accurate census method for songbirds. *Stud Avian Biol* 6:86–91
- Ferrier S, Watson G (1997) An evaluation of the effectiveness of environmental surrogates and modelling techniques in predicting the distribution of biological diversity. Consult. Rep. NSW Natl. Parks Wildl. Serv. Dep. Environ., Sport Territ., Environ. Aust., Canberra. <http://www.deh.gov.au/biodiversity/publications/technical/surrogates/>
- Filz KJ, Schmitt T, Engler JO (2013) How fine is fine-scale? Questioning the use of fine-scale bioclimatic data in species distribution models used for forecasting abundance patterns in butterflies. *Eur J Entomol* 110:311–317
- Fleishman E, Murphy DD, Brussard PF (2000) A new method for selection of umbrella species for conservation planning. *Ecol Appl* 10:569–579
- Foltête JC, Clauzel C, Vuidel G, Toumant P (2012) Integrating graph-based connectivity metrics into species distribution models. *Landsc Ecol* 27:557–569
- Guisan A, Thuiller W (2000) Predictive habitat distribution models in ecology. *Ecol Model* 135:147–186
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecol Lett* 8:993–1009
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Model* 135:147–186
- Guisan A, Lehmann A, Ferrier S, Austin M, Overton JMCC, Aspinall R, Hastie T (2006) Making better biogeographical predictions of species' distributions. *J Appl Ecol* 43:386–392
- Guisan A, Zimmermann NE, Elith J, Graham CH, Phillips S, Peterson AT (2007) What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? *Ecol Monogr* 77:615–630
- Hanski I (1998) Metapopulation dynamics. *Nature* 396:41–50
- Henle K, Davies KF, Kleyer M, Margules C, Settele J (2004) Predictors of species sensitivity to fragmentation. *Biodivers Conserv* 13:207–251
- Hijmans RJ (2012) Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* 93:679–688
- Hirzel AH, Le LAYG (2008) Habitat suitability modelling and niche theory. *J Appl Ecol* 45:1372–1381
- Holdridge LR, Grenke WC, Hatheway WH, Liang T, Tosi JAJR (1971) *Forest environments in tropical life zones: a pilot study*. Pergamon Press, New York
- Hopkins RL (2009) Use of landscape pattern metrics and multiscale data in aquatic species distribution models: a case study of a freshwater mussel. *Landsc Ecol* 24:943–955
- Lawton JH (1996) Population abundances, geographic ranges and conservation: 1994 Witherby Lecture. *Bird Stud* 43:3–19
- Loiselle BA, Jorgensen PM, Consiglio T, Jimenez I, Blake JG, Lohmann LG, Montiel OM (2008) Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcome? *J Biogeogr* 35:105–116
- Margules CR, Pressey RL (2000) Systematic conservation planning. *Nature* 405:243–253
- Martensen AC, Pimentel RG, Metzger JP (2008) Relative effects of fragment size and connectivity on bird community in the Atlantic Rain Forest: implications for conservation. *Biol Conserv* 141:2184–2192
- Mcgarigal K, Cushman SA, Neel MC, Ene E (2002) FRAGSTATS v3: spatial pattern analysis program for categorical maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. <http://www.umass.edu/landeco/research/fragstats/fragstats.html>. Accessed 30 Mar 2015
- Meyer CB, Thuiller W (2006) Accuracy of resource selection functions across spatial scales. *Divers Distrib* 12:288–297
- Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecol Model* 133:225–245
- Pearman PB, Guisan A, Broennimann O, Randin CF (2008) Niche dynamics in space and time. *Trends Ecol Evol* 23:149–158
- Peterson AT (2006) Uses and requirements of ecological niche models and related distributional models. *Biodivers Inform* 3:59–72
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Model* 190:231–259
- Pulliam HR (2000) On the relationship between niche and distribution. *Ecol Lett* 3:349–361
- R Development Core Team (2012) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. www.R-project.org. Accessed 10 April 2012
- Robinson GR, Quinn JF (1988) Extinction, turnover and species diversity in an experimentally fragmented California annual grassland. *Oecologia* 76:71–82
- Rosales-Meda MM (2007) Caracterización de la población del mono aullador (*Alouatta palliata palliata*) em el Refugio Nacional de Vida Silvestre Isla San Lucas, Costa Rica. *Neotrop Primates* 14(3):122–127

- Rosenzweig M (1995) Species diversity in space and time. Cambridge University Press, Cambridge
- Sayre R, Roca E, Sedagatkish G, Young B, Keel S, Roca R, Sheppard S (2000) Nature in focus: rapid ecological assessment. Island Press, Washington
- SOS Mata Atlântica-INPE (2008) Atlas da evolução dos remanescentes florestais da Mata Atlântica no período de 2000-2005. Fundação SOS Mata Atlântica, São Paulo. www.sosma.org.br e www.inpe.br
- Temple SA, Cary JR (1988) Modeling dynamics of forest interior bird populations in fragmented landscapes. *Conserv Biol* 2:340–347
- Thuiller W, Brotons L, Araújo MB, Lavorel S (2004) Effects of restricting environmental range of data to project current and future species distributions. *Ecography* 27:165–172
- Titeux N, Dufrene M, Radoux J, Hirzel AH, Defourny P (2007) Fitness-related parameters improve presence-only distribution modelling for conservation practice: the case of the red-backed shrike. *Biol Conserv* 138:207–223
- Warren DL (2012) In defense of ‘niche modeling’. *Trends Ecol Evol* 27:497–500