How species distribution models can improve cat conservation – Jaguars in Brazil

Katia MPMB Ferraz
Escola Superior de Agricultura "Luiz de Queiroz"

Beatriz Beisiegel
Instituto Chico Mendes de Conservação da Biodiversidade

Rogerio Cunha de Paula
Instituto Chico Mendes de Conservação da Biodiversidade

Tadeu Oliveira
Maranhão State University - Universidade Estadual do Maranhão

ISSN 1027-2992

Special Issue

N° 7 | Spring 2012

CAT news

Jaguar in Brazil

CONSERVATION BREEDING SPECIALIST GROUP

ICMBio CENAP

SPECIES SURVIVAL COMMISSION
CAT SPECIALIST GROUP

CATnews is the newsletter of the Cat Specialist Group, a component of the Species Survival Commission of The World Conservation Union (IUCN). It is published twice a year, and is available to members and the Friends of the Cat Group.

For joining the Friends of the Cat Group please contact Christine Breitenmoser at ch.breitenmoser@kora.ch

Original contributions and short notes about wild cats are welcome **Send contributions and observations to ch.breitenmoser@kora.ch.**

Guidelines for authors are available at www.catsg.org/catnews

This special issue of CATnews has been produced with support from Chester Zoo and San Diego Zoo.

Design: barbara surber, werk'sdesign gmbh
Layout: Christine Breitenmoser
Print: Stämpfli Publikationen AG, Bern, Switzerland

ISSN 1027-2992

Editors: Christine & Urs Breitenmoser
Co-chairs IUCN/SSC
Cat Specialist Group
KORA, Thunstrasse 31, 3074 Muri,
Switzerland
Tel ++41(31) 951 90 20
Fax ++41(31) 951 90 40
<urs.breitenmoser@ivv.unibe.ch>
<ch.breitenmoser@kora.ch>

Contributing Editor: Peter Jackson
29 Lake Close, London SW19 7EG
United Kingdom
Tel/Fax: ++44 (20) 89 47 01 59
<peterfr.jackson@virgin.net>

Associate Editors: Keith Richmond
Brian Bertram
Sultana Bashir

Cover Photo: Jaguar in the Pantanal
Photo P. Meier/mywilderness.net

KATIA M. P. M. B. FERRAZ[1], BEATRIZ M. BEISIEGEL[2], ROGÉRIO C. DE PAULA[2], DÊNIS A. SANA[3], CLÁUDIA B. DE CAMPOS[2], TADEU G. DE OLIVEIRA[4], ARNAUD L. J. DESBIEZ[5]

How species distribution models can improve cat conservation - jaguars in Brazil

Modeling species distribution is a promising field of research for improving conservation efforts and setting priorities. The aim of this study was to produce an environmental suitability map for jaguar distribution in two biomes in Brazil – Caatinga and Atlantic Forest – , where the species is Critically Endangered as part of the Jaguar National Action Plan workshop (Atibaia, São Paulo state). Species occurrence (N = 57 for Caatinga and N = 118 for Atlantic Forest), provided by jaguar specialists, and ten environmental predictors (elevation, land cover, distance from water and bioclimatic variables) were used to generate species distribution models in Maxent. Both models presented high predictive success (AUC = 0.880 ± 0.027 for Caatinga and AUC = 0.944 ± 0.022 for Atlantic Forest) and were highly significant (p < 0.001), predicting only 18.64% of Caatinga and 10.32% of Atlantic Forest as suitable for jaguar occurrence. The species distribution models revealed the low environmental suitability of both biomes for jaguar occurrence, emphasizing the urgency of setting conservation priorities and strategies to improve jaguar conservation such as the implementation of new protected areas and corridors for species dispersal.

Predicting species distribution has made enormous progress during the past decade. A wide variety of modeling techniques (see Guisan & Thuiller 2005) have been intensively explored aiming to improve the comprehension of species-environment relationships (Peterson 2001). The species distribution modeling (SDM) relates species distribution data to information on the environmental and/or spatial characteristics of those locations. Combinations of environmental variables most closely associated to presence points can then be identified and projected onto landscapes to identify areas of predicted presence on the map (Soberón & Peterson 2005, Elith & Leathwick 2009). The geographic projection of these conditions (i.e., where both abiotic and biotic requirements are fulfilled) represents the potential distribution of the species. Finally, those areas where the potential distribution is accessible to the species are likely to approximate the actual distribution of it.

The jaguar, the largest felid in the Americas, has been heavily affected by retaliation killing for livestock predation, fear, skin trade, prey depletion, trophy hunting (e.g. Smith 1976, Conforti & Azevedo 2003) and habitat loss (Sanderson et al. 2002). As a consequence, it is now restricted to ca. 46% of its former range (Sanderson et al. 2002).

Table 1. Environmental predictor variables used in jaguar distribution model.

Variables	Description
Land cover	Land cover map from GlobCover Land Cover version V2.3, 2009
Elevation	Elevation map by NASA Shuttle Radar Topography Mission
Distance from water	Map of gradient distance from water obtained from vector map of rivers from IBGE
Bioclimatic variables	Maps of bioclimatic variables from Worldclim:
	Bio1 = Annual mean temperature
	Bio2 = Mean diurnal range (mean of monthly (max temp - min temp))
	Bio5 = Max temperature of warmest month
	Bio6 = Min temperature of coldest month
	Bio12 = Annual precipitation
	Bio13 = Precipitation of wettest month
	Bio14 = Precipitation of driest month

Environmental suitability models have been produced for jaguar distribution in Brazil during the Jaguar National Action Plan Workshop, facilitated by IUCN/SSC CBSG Brazil and organized and funded by CENAP/ICMBio, Pró Carnívoros and Panthera, in November 2009, Atibaia, São Paulo state, Brazil. During the workshop, jaguar specialists provided occurrence point data for species distribution modeling. A jaguar database was composed only by recent (less than five years) and confirmed records (e.g., signs, telemetry, camera-trapping, chance observations). All models and detailed information about the procedure and the results are included in the Jaguar National Action Plan. Background information on SDM and necessary considerations are summarized in the Supporting Online Material Appendix I (www.catsg.org/catnews). Here, to illustrate the potential of the use of the SDM for cat conservation, we presented the environmental suitability models for jaguar in two biomes (Caatinga and Atlantic Forest, Fig. 1), where the species is considered Critically Endangered in Brazil (de Paula et al. 2012, this issue; Beisiegel et al. 2012, this issue).

Methods

Jaguar distribution was modeled for each biome separately considering the differences between the environmental spaces (i.e., conceptual space defined by the environmental variables to which the species responds). The biome map used was obtained from a Land Cover Map of Brazil (1:5.000.000), 2004, by the Brazilian Institute of Geography and Statistics, IBGE (available for download at http://www.ibge.gov.br/).

Predictive distribution models were formulated considering the entire available jaguar dataset as the dependent variable (presence points) and the selected environmental variables as the predictors (Table 1). Jaguar data available for modeling (N = 57 for Caatinga; N = 118 for Atlantic Forest; Fig. 2) were plotted as lat/long coordinates on environmental maps with a grid cell size of 0.0083 decimal degree^2 (~1 km^2).

Models were obtained by Maxent 3.3.3e (Phillips & Dudík 2008) using 70% of the data for training (N = 40 for Caatinga and N = 66 for Atlantic Forest) and 30% for testing the models (N = 17 for Caatinga and N = 28 for Atlantic Forest; Pearson 2007). Data were sampled by bootstrapping with 10 random partitions with replacements. All runs were set with a convergence threshold of 1.0E–5 with 500 iterations, with 10,000 background points.

The logistic threshold output format was used resulting in continuous values for each grid cell in the map from 0 (unsuitable) to 1 (most suitable). These values can be interpreted as the probability of presence of suitable environmental condition for the target species (Veloz 2009). The logistic threshold used to "cut-off" the models converting the continuous probability model in a binary model was the one that assumed 10 percentile training presence provided by the Maxent outputs 0.300 for Caatinga; 0.100 for Atlantic forest. These thresholds were selected by the specialists as the best one to represent the suitable areas for recent jaguar distribution in both biomes.

Models were evaluated by the AUC value, the omission error and by the binomial probability (Pearson 2007).

Fig. 1. Biomes of Brazil with Caatinga and Atlantic Forest biomes with protected areas (green).

Results and Discussion

The SDM for Caatinga and Atlantic Forest biomes presented high predictive success and were highly statistically significant (AUC = 0.880 ± 0.027, omission error = 0.206, $p < 0.001$; AUC = 0.944 ± 0.022, omission error = 0.129, $p < 0.001$, respectively; SOM Fig. 1, 2), predicting about 18.64% of the Caatinga (Fig. 3) and 10.32% of the Atlantic Forest (Fig. 4) as suitable for jaguar occurrence.

Much of the Caatinga biome (844,453 km^2) predicted as suitable (54.77%) for jaguar occurrence encompassed the closed to open (>15%) shrubland. Meanwhile, much of the unsuitable area (26.62%) for the species also encompassed this land cover. This discrepancy is due especially to human development or simply occupation that leads to medium to high level of disturbance in the environment. These habitat alterations are especially due to mining activities, agriculture, timber extraction, firewood production, and lowering of prey items due to excessive hunting activities. The closed to open shrubland covers about 40.67% of total biome area. The closed formations have 60% to 80% of plant cover, whereas the open formations have only 40 to 60% (Chaves et al. 2008). The vegetation type is deciduous, generally with thorny woody species > 4.5 m tall, interspersed with succulent plants, especially cacti. The trees are 7-15 m high, with thin trunks. Several have tiny leaves where others have spines or thorns (Andrade-Lima 1981).

The semi-arid Caatinga domain is one of the most threatened biomes in Brazil with less than 50% of its natural cover and greatly impacted and fragmented by human activities (Leal et al. 2005). Most of the protected areas found in this biome (Fig. 3) presented large areas as suitable for jaguar occurrence, such as Serra Branca Ecological Station (ES) and Serra da Capivara National Park (NP) with 100%, Morro do Chapéu State Park (SP) with 91.29% and Serra das Confusões NP with 71.51%. Nevertheless Serra das Confusões and Chapada Diamantina NPs (with 62.63%) are the only two protected areas that are located in transitional areas with the Cerrado biome, hence the lower suitability within the Caatinga. Serra das Confusões NP is indeed a very important area for jaguars as it is large (5,238 km^2), connected to Serra da Capivara NP/Serra Branca ES and also somehow bridges the Caatinga jaguar population with those of the Nascentes do Rio Parnaíba protected areas complex, likely the most important of the Cerrado domain. The bulk of prime areas for jaguars, located within the center of the Caatinga domain are being proposed as a new NP, created to protect one of the most important populations of the Critically Endangered Caatinga jaguar, Boqueirão da Onça NP (Fig. 3). The creation of this new protected area should be of utmost importance for jaguar conservation in the Caatinga. If the NP will be created according to the proposed limits, it will encompass 24.66% of the highly suitable area for jaguars.

Much of the Atlantic Forest biome (1,110,182 km^2) predicted as suitable (27.44%) for jaguar occurrence encompassed the closed to open (>15%) broadleaved evergreen or semi-deciduous forest (55.26%), while unsuitable areas encompassed mainly mosaic cropland (50-70%)/ vegetation (grassland/shrubland/forest) (20-50%).

Most of the continuous forest remains indicated as suitable for the jaguars at the Atlantic Forest biome correspond to the Brazilian protected areas (Fig. 4) such as Morro do Diabo SP, Mico Leão Preto ES, Caiuá ES, Carlos Botelho SP, Intervales SP, Alto Ribeira Touristic SP and Xitué ES, Iguaçu NP, Serra da Bocaina NP, Tinguá Biological Reserve (BR) and Serra dos Órgãos NP, besides surroundings areas and some isolated forest remains (e.g., Rio Doce SP and Itatiaia NP). The marshlands in the Upper Paraná River, in the west portion of the Atlantic Forest biome, are as important as forest areas to jaguar conservation. The most suitable areas in the region includes continuous protected areas such the Ilha Grande NP, Várzeas do Rio Ivinhema SP and Ilhas e Várzeas do Rio Paraná Environmental Protection Area (EPA).

Some suitable areas indicated by the model such as Cantareira SP and its surrounding did not present any recent record of the species presence. The depauperate quality of forest cover of these areas with high human pressure probably explains the absence of the species there. This clearly illustrates the overprediction (i.e., commission error), frequently observed in SDM. In this particular situation, the degraded vegetation and human pressure are not contemplated in the environmental variables input in the modeling, decreasing

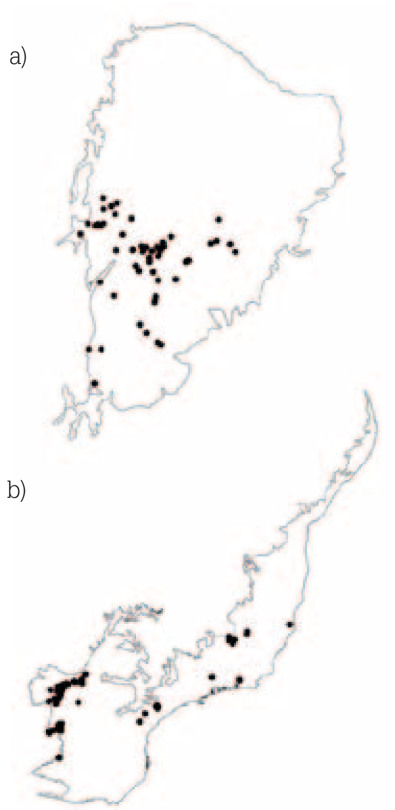


Fig. 2. Jaguar presence points for (a) Caatinga (N = 57) and (b) Atlantic Forest (N = 118) biomes in Brazil.

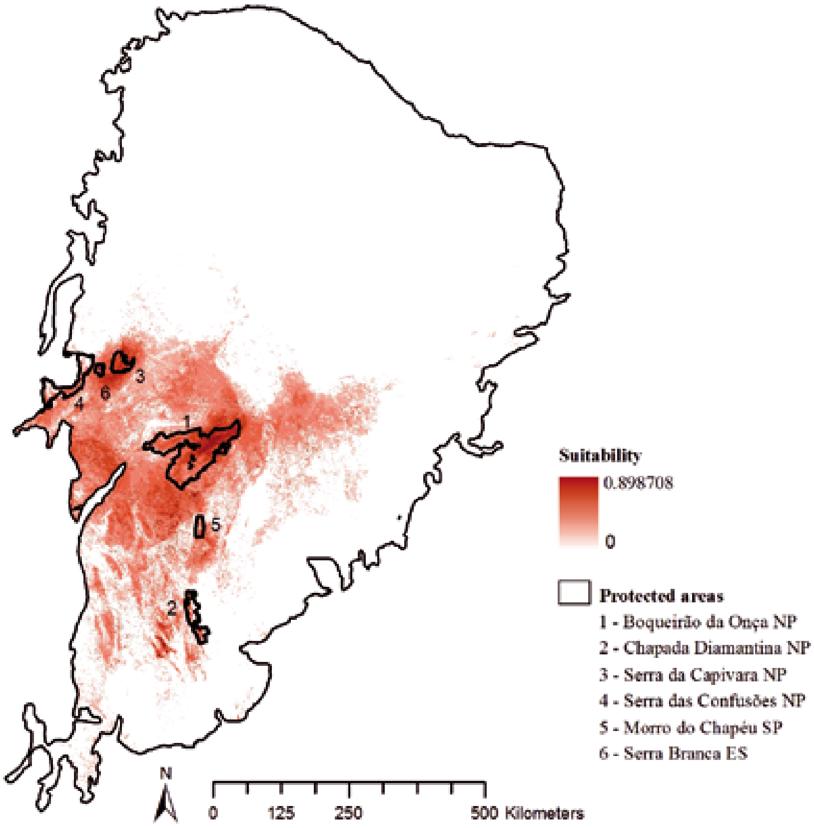


Fig. 3. Potential distribution model for jaguar in Caatinga biome with some protected areas highlighted.

its predictive power. On the other hand, some areas with recent records of the species (not included in the modeling) were not indicated as suitable by the model such as the Juréia-Itatins ES and Caraguatatuba area of Serra do Mar SP. The omission and commission errors are common and frequent in SDM (Fielding & Bell 1997, Pearson 2007), emphasizing the need of cautious interpretation as local characteristics could decrease the model predictive success.

Most of the cropland areas (rainfed croplands, mosaic croplands/vegetation, mosaic croplands/forest; 64.67%) were considered unsuitable for the species occurrence. Jaguars depend on large prey such as peccaries, which are very susceptible to environmental degradation and poaching (e.g. Cullen Jr. et al. 2000), which is intense throughout the Atlantic forest, with the exception of a few well preserved areas. Accordingly, Cullen Jr. et al. (2005) had already verified that jaguars display a strong selection for primary and secondary forests, a strong avoidance of pastures and a weak use of agricultural areas.

The probability of jaguar presence was associated differently to the environmental predictor variables. Elevation (19.03%), the precipitation of driest month (Bio14; 18.08%) and

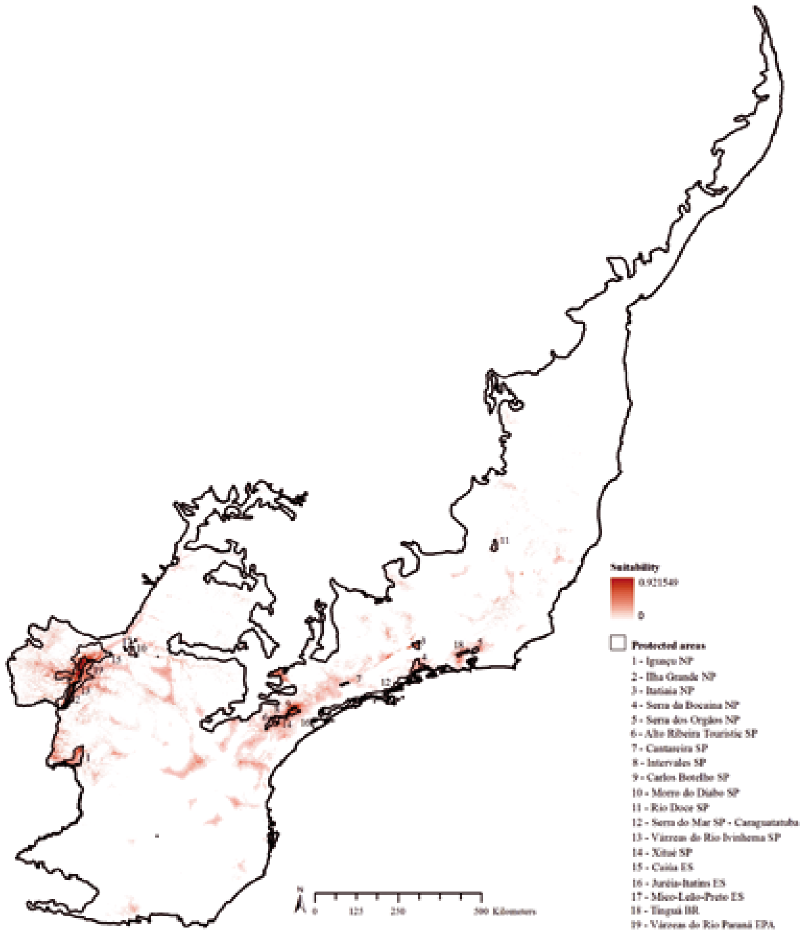


Fig. 4. Potential distribution model for jaguar in Atlantic Forest biome with some protected areas highlighted.

the mean diurnal range (Bio2; 17.25%) were the highest contributor variables for jaguar model at the Caatinga biome. The probability of jaguar presence increased as elevation and the mean diurnal range increased, but decreased as the precipitation of driest month increased (Fig. 5). The presence of jaguar in Caatinga is associated with higher areas probably because of the lower human pressure and more pristine vegetation (e.g., Boqueirão da Onça NP). Although variables Bio14 and Bio2 had important contributions to the model its relationships with jaguar presence were not so clear.

Land cover (41.29%) was the highest contributor variable for the jaguar model in the Atlantic Forest biome. The high probability of jaguar presence was related to the closed to open (>15%) grassland or woody vegetation regularly flooded (Fig. 6). Wetland areas and riparian vegetation (Fig. 7) are core areas and dispersal corridors for jaguars (Cullen Jr. et al. 2005). However, only 30% of the original area of the Paraná River is left because of the construction of hydroelectric power stations (Agostinho & Zalewski 1996).

Future for SDM as a tool for cat conservation

The field of SDM is promising for improving conservation efforts and priorities (e.g. Thorn et al. 2009, Costa et al. 2010, Marini et al. 2010). SDM is a useful tool for resolving practical questions in applied ecology and conservation biology, but also in fundamental sciences (e.g. biogeography and phylogeography) (Guisan & Thuiller 2005). It represents an empirical method to draw statistical inferences about the drivers of species' ranges under different conservation, ecological and evolutionary processes (Zimmermann et al. 2010).

The SDM approach can improve our knowledge about cat species worldwide by 1) highlighting areas where the species might occur but confirmed observation is missing, 2) identifying gaps in data collection and guiding the sampling efforts, 3) identifying key areas for conservation efforts and potential corridors linking protected areas and/or populations, 4) contributing for the assessment of IUCN red list categories, 5) helping to reduce conflicts (e.g., zoning), among others. Moreover, this modeling technique can provide a comprehensive understanding of the historical, current and future ranges of cat species, providing insights to conservation planning (e.g., Marini et al. 2010). Modeling should also be of paramount importance for predicting threatened species range in a world of climatic change. In fact, this kind of prediction could be vital for setting proper and effective action plans for critically endangered populations/species.

In practice, one of the most useful contributions from SDMs could be the prediction of suitable areas for species occurrence as well as helping to delineate potential corridors which link populations on a continental scale. The environmental suitability maps in a modeling framework could be used as a basis to improve the already existing extraordinary initiatives that seek to create such linkages (e.g. jaguar corridor initiative). This, in turn, has been considered one of the most effective conservation strategies to guarantee cat species conservation (Macdonald et al. 2010).

The assessment of conservation priorities for felids should consider the environmental suitability of landscape in a modeling framework. Suitability maps could be considered by stakeholders for defining priority areas for the establishment of new protected areas or corridors. However, conservation inferences should rely on robust models, avoiding omission and overprediction in species distribution range.

The modeling exercise defining priority areas for conservation efforts should be a useful first evaluation. In this workshop one of the most valuable contributions of this exercise was the participatory manner in which this model was constructed. Furthermore the resulting maps provided stakeholders with distribution information and clear results to discuss, and it stimulated debates and discussions which otherwise may not have occurred. However, for reliable conservation decisions suitability models must rely on well-delineated field inventories (Costa et al. 2010) and model results must be validated.

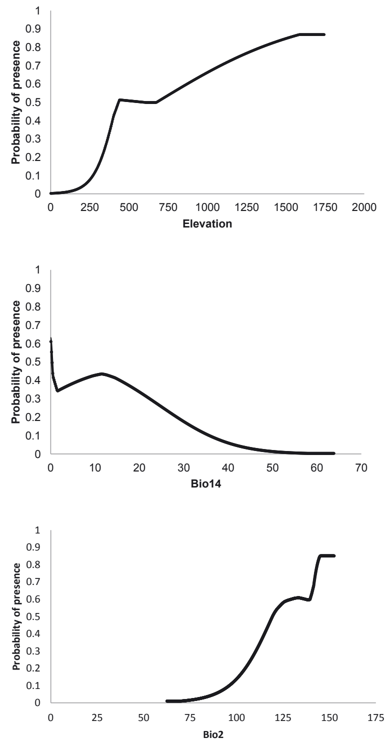


Fig. 5. Marginal response curves of the predicted probability of jaguar occurrence at the Caatinga biome for the environmental predictor variables that contributed substantially to the SDM.

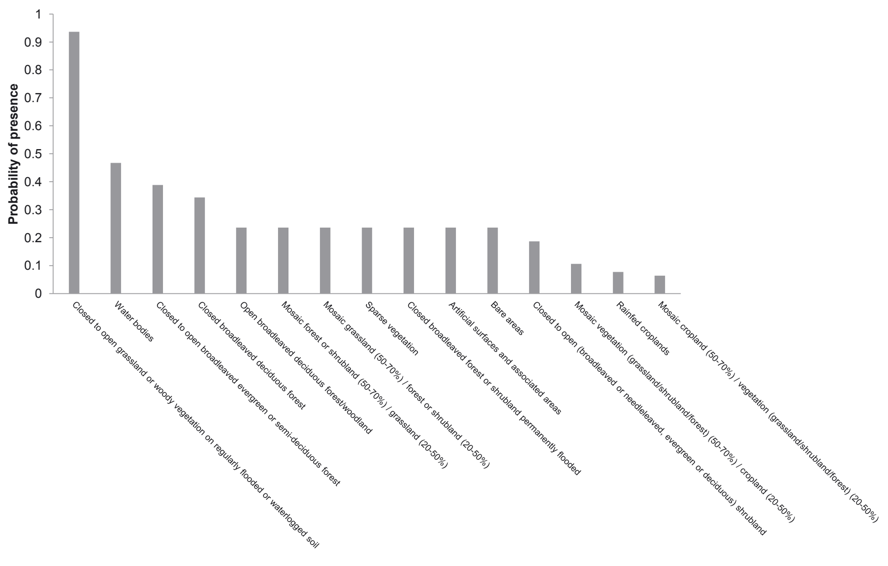


Fig. 6. . Marginal response curve of the predicted probability of jaguar occurrence at the Atlantic Forest biome for the environmental predictor variable that contributed substantially to the species distribution model.

Fig. 7. Riparian vegetation is an important part of jaguar core areas and corridors (Photo A. Gambarini),

Acknowledgments

Paper # 01 of the SISBIOTA - Top Predators network. The authors thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for financial support. We thank all participants of the Jaguar National Action Plan Workshop (Atibaia, São Paulo, Brazil) who provided recent jaguar presence points. We also thank the CENAP (The National Center for Research and Conservation of Mammalian Carnivores) for providing useful information for the study. We are grateful to Fridolin Zimmermann and Christine Breitenmoser from KORA for valuable contributions to the manuscript.

References

Agostinho A.A. & Zalewski M. 1996. A planície alagável do alto rio Paraná: importância e preservação. EDUEM. Maringá, Brazil. 100 pp.

Andrade-Lima D. 1981. The Caatinga dominium. Revista Brasileira de Botânica 4, 149-153.

Chaves I. de B., Lopes V. L., Folliott P. F., Paes-Silva A. P. 2008. Uma classificação morfo-estrutural para descrição e avaliação da biomassa da vegetação caatinga. Caatinga (Mossoró, Brasil), 21, 204-213.

Conforti V. A. & Azevedo F. C. C. 2003. Local perceptions of jaguars (*Panthera onca*) and pumas (*Puma concolor*) in the Iguaçu National Park area, south Brazil. Biol. Conserv. 111, 215-221.

Costa G. C., Nogueira C., Machado R. B. & Colli G. R. 2010. Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. Biodiversity and Conservation 19, 883-899.

Cullen Jr. L., Abreu C. K., Sana D. & Nava A. F. D. 2005. As onças-pintadas como detetives da paisagem no corredor do Alto Paraná, Brasil. Natureza e Conservação 3, 43-58.

Cullen Jr. L., Bodmer R. E. & Pádua C. V. 2000. Effects of hunting in habitat fragments of the Atlantic forests, Brazil. Biol. Conserv. 95, 49-56.

Elith J. & Leathwick J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology and Evolution Systematics 40, 677-97.

Fielding A. H. & Bell J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24, 38-49.

Guisan A. & Thuiller W. 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters 8, 993-1009.

Leal I. R., Silva J. M. C. da, Tabarelli M. & Larcher Jr. T. E. 2005. Mudando o curso da conservação da biodiversidade na Caatinga do nordeste do Brasil. Megadiversidade 1, 139-146.

Macdonald D. W. & Loveridge A. J. & Rabinowitz A. 2010. Felid futures: crossing disciplines, borders and generations. *In* Biology and Conservation of Wild Felids. Macdonald D. W. & Loveridge A. J. (Eds). Oxford University Press, Oxford, New York, pp. 599-649.

Marini M. A., Barbet-Massin M., Martinez J., Prestes N. P. & Jiguet F. 2010. Applying ecological niche modelling to plan conservation actions for the Red-spectacled Amazon (*Amazona pretrei).* Biological Conservation 143, 102-112.

Pearson R. G. 2007. Species' Distribution Modeling for Conservation Educators and Practitioners. Synthesis. AMNH (http://ncep.amnh.org).

Peterson A. T. 2001. Predicting species' geographic distributions based on ecological niche modeling. Condor 103, 599-605.

Phillips S. J. & Dudík M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31, 161-175.

Sanderson E., Redford K., Chetkiewicz C., Medellin R., Rabinovitz A. R., Robinson J. G. & Taber A. 2002. Planning to save a species: the jaguar as a model. Conservation Biology 16, 58-72.

Smith N. J. H. 1976. Spotted cats and the Amazon skin trade. Oryx 13, 362-371.

Soberón J. M. & A. T. Peterson. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodiversity Informatics 2, 1-10.

Thorn J. S., Nijman V., Smith D. & Nekaris K. A. I. 2009. Ecological niche modelling as a technique for assessing threats and setting conservation priorities for Asian slow lorises (Primates: Nycticebus). Diversity and Distributions 15, 289-298.

Veloz S. D. 2009. Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. Journal of Biogeography 36, 2290-2299.

Zimmermann N. E., Edwards Jr. T. C., Graham C. H., Pearman P. B. & Svenning J. 2010. New trends in species distribution modelling. Ecography 33, 985-989.

Supporting Online Material SOM available at www.catsg.org/catnews

[1] Departamento de Ciências Florestais, Escola Superior de Agricultura 'Luiz de Queiroz", Universidade de São Paulo, 13418-900, Piracicaba, São Paulo, Brasil
<katia.ferraz@usp.br>

[2] Centro Nacional de Pesquisa e Conservação de Mamíferos Carnívoros – CENAP/ICMBio, Estr. Mun. Hisaichi Takebayashi, 8600, Bairro Usina, 12952-011, Atibaia, SP, Brasil

[3] Instituto para Conservação dos Carnívoros Neotropicais, Cx. P.10, 12940-970 Atibaia, SP, Brasil. / Programa de Pós–Graduação em Ecologia, Instituto de Biociências, Porto Alegre, RS, Brasil

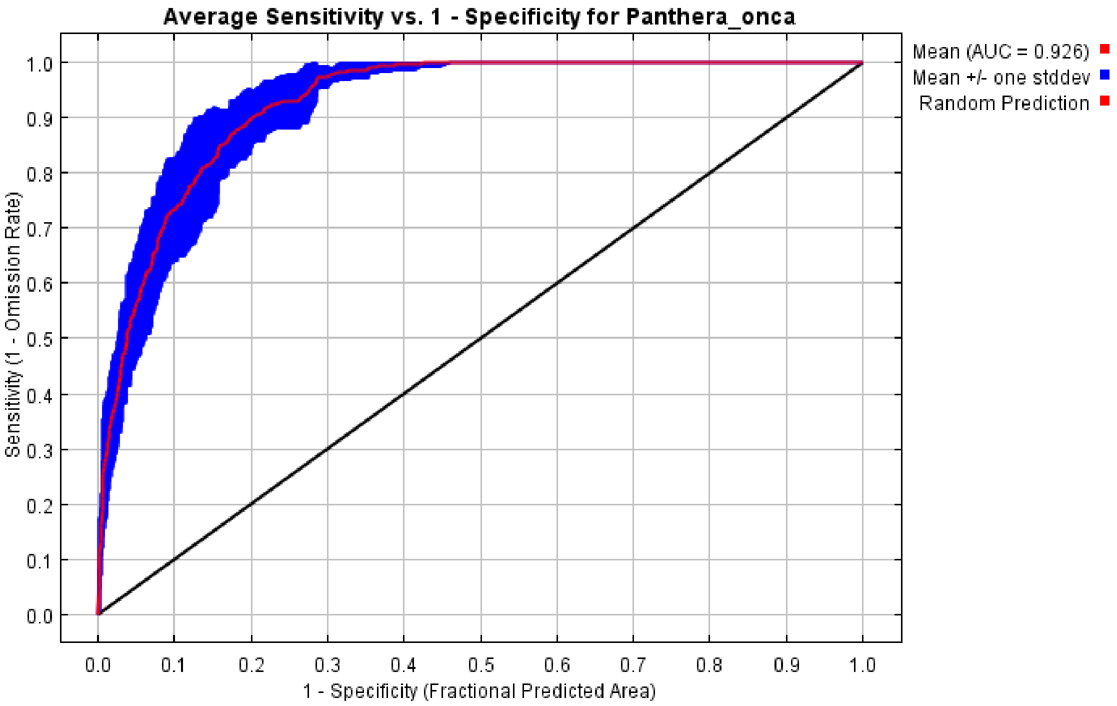
[4] Departamento de Biologia, Universidade Estadual do Maranhão, Cx. P. 09, Cidade Universitária Paulo VI, São Luís-MA, Brasil

[5] Royal Zoological Society of Scotland, Murrayfield, Edinburgh EH12 6TS Scotland, UK

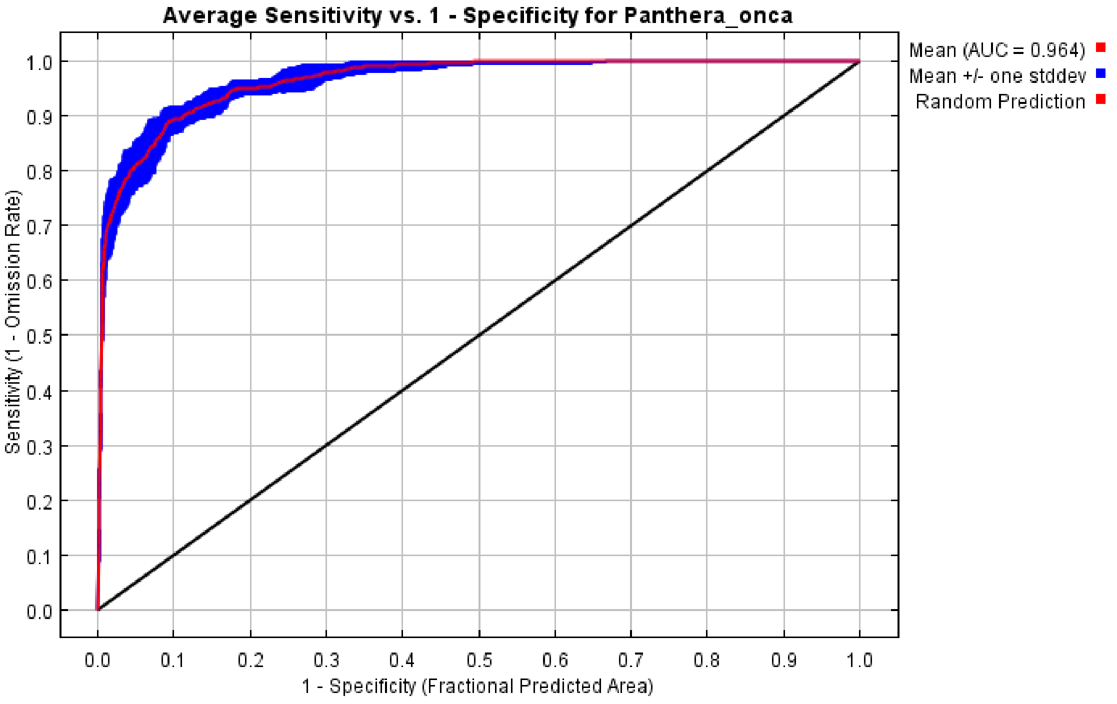
Ferraz et al. 2012. How species distribution models can improve cat conservation - jaguars in Brazil. Cat News Special Issue 7, 38-42.

Supporting Online Material SOM Figures 1 and 2.

a)



b)



SOM Fig. 1. ROC plot curve for (a) Caatinga and (b) Atlantic Forest.

a)

b)

SOM Fig. 2. Jaguar distribution area at (a) Caatinga and (b) Atlantic Forest in Brazil.

Ferraz et al. 2012. How species distribution models can improve cat conservation - jaguars in Brazil. Cat News Special Issue 7, 38-42.

Supporting Online Material SOM Appendix I. Background information on Species Distribution Modeling SDM

Predicting species distribution has made enormous progress in the last decade. A wide variety of modeling techniques (see Guisan & Thuiller 2005) have been intensively explored aiming to improve the comprehension of species-environment relationships (Guisan & Zimmermann 2000, Peterson 2001, Hirzel & Lay 2008, Elith & Leathwick 2009, Franklin 2009). The species distribution modeling (SDM) relate species distribution data to information on the environmental and/or spatial characteristics of those locations. Combinations of environmental variables most closely associated to presence points can then be identified and projected onto landscapes to identify areas of predicted presence on the map (Soberón & Peterson 2005, Peterson 2006). The geographic projection of these conditions (i.e., where both abiotic and biotic requirements are fulfilled) represents the potential distribution of the species. Finally, those areas where the potential distribution is accessible to the species are likely to approximate the actual distribution of the species.

The SDMs have also been termed as ecological niche models (ENMs) or habitat models (sometimes with different emphases and meanings; Elith & Leathwick 2009, Soberón & Nakamura 2009). According to Elith & Leathwick (2009) the use of neutral terminology to describe species distribution models (SDM rather than ENM) seems preferable. Despite its extensive use, there is an enormous debate about terminology and concepts in predictive modeling and a consensus about what we are modeling – habitat, niche, environment, species distribution – does not exists until now (Soberón & Peterson 2005, Kearney 2006, Peterson 2006, Austin 2007, Soberón 2007, Hirzel & Lay 2008, Jiménez-Valverde et al. 2008, Soberón & Nakamura 2009).

The use of predictive models of species potential distribution has been increasingly used in many areas related to species ecology and conservation, such as to predict areas that could potentially be re-colonised by an expanding species, to choose the best location for reintroduction/restocking or even to indicate potential areas to be prioritized for conservation purposes, including conservation planning, management and restoration (Guisan & Zimmermann 2000, Ferrier et al. 2002a,b, Soberón & Peterson 2004, Peterson 2006, Franklin 2009, Wilson et al. 2010, Rodríguez-Soto et al. 2011). Published examples indicate that SDMs can perform well in characterizing the natural distributions of species (within their current range), particularly when well-designed survey data and functionally relevant predictors are analyzed with an appropriately specified model (Elith & Leathwick 2009). Despite the widespread use of these models, some authors (Pulliam 2000, Soberón & Peterson 2005, Araujo & Guisan 2006, Peterson 2006, Soberón 2007, Jiménez-Valverde et al. 2008) have pointed out important conceptual ambiguities as well as biotic and algorithm uncertainties that need to be investigated in order to increase confidence in model results, such as 1) clarification of model aims; 2) clarification of niche concept, including the distinction between potential and realized distribution; 3) improved design for sampling data for building model; 4) improved model parameterization; 5) improved model selection and predictor contribution; and 6) improved model evaluation.

Modeling the species distribution

Biological data as good-quality source data

Occurrence data for species distribution models can only include presence or presence-absence data. The type of data available for modeling will determine the algorithm and model procedure selection. Species distribution data can be obtained from museum or scientific collections or by field surveys. Many scientific datasets are available for download such as Global Biodiversity Information Facility (GBIF, http://www.gbif.org/) and SpeciesLink (http://splink.cria.org.br/). There are many problems associated to these data sets mainly related to the species identification, sampling effort bias and precision of records (Soberón & Peterson 2004). Field survey data, generally obtained by species observation, trapping or track surveys, from sampling procedure ensuring a broad environmental coverage of gradients in the species distribution range (Vaughan & Ormerod 2003), avoiding bias and pitfalls, are supposed to be good quality data for species distribution modeling. Occurrence data obtained by interviews are generally not recommended to be used in modeling as they are usually not accurate in regards to the species occurrence site.

Many problems have been faced by modelers due mainly to clustered datasets and biased sampling not covering the full range of environmental conditions (e.g., environmental heterogeneity) within the landscape, especially for wide ranging species. Clustered data, especially when provided by telemetry data, could lead to a potential bias in the final model. An option to solve this apparent problem is to subsample the dataset in order to dilute the oversampling in some parts of the species distribution range (Veloz 2009).

Environmental variables as good predictors

Environmental data sets matter in species distribution modeling (Peterson & Nakazawa 2008). The role of a distribution model may be primarily predictive or, alternatively, may emphasize the relationship between an organism and its habitat (Vaughan & Ormerod 2003). So the environmental predictors should therefore have a biological relationship with the organism. The spatial scale should be carefully defined as it can influence the results and/or not resolve the motivated question of the study (Vaughan & Ormerod 2003). The selection of resolution and extent is a critical step in SDM building, and an inappropriate selection can yield misleading results (Guisan & Thuiller 2005). Ideally, models should examine a series of spatial scales, increasing the understanding of organism-environmental relationship (Vaughan & Ormerod 2003).

Many environmental variables, used as predictors, are available for download by many International Agencies. Some examples of frequently used environmental databases are global climate layers from Worldclim (http://www.worldclim.org/), elevation from the NASA Shuttle Radar Topography Mission (SRTM, http://www2.jpl.nasa.gov/srtm/), climate data from past, present and future from Intergovernmental Panel on Climate Change (IPCC, http://www.ipcc-data.org/), Hidro1K elevation derivative database from Earth Resources Observation and Science (EROS, http://eros.usgs.gov/), global land cover from ESA GlobCover 2009 Project (http://ionia1.esrin.esa.int/), and satellite images from MODIS (https://wist.echo.nasa.gov/api/).

Procedure of species distribution modeling

Some models are presence-only models such as DOMAIN (Carpenter et al. 1993) and BIOCLIM (Busby 1986, Nix 1986), while others demand presence and absence data, such as the GLM (Generalized Linear Model) and GAM (Additive Linear Model; Guisan & Zimmermann 2000). Others demand presence and background points such as Biomapper (Hirzel et al. 2002) and Maxent (Phillips et al. 2004, 2006) or presence and pseudoabsence such as GARP (Stockwell & Peter 1999). The latter was generated by locating sites randomly across the total geographical area, or ‘domain’, of interest (Ferrier et al. 2002a).

Maxent, one of the most recently used algorithm, estimates a target probability distribution by finding the probability distribution of maximum entropy (i.e., that is most spread out, or closest to uniform), subject to a set of constraints that represent our incomplete information about the target distribution (Phillips et al. 2004, 2006). When Maxent is applied to presence-only species distribution modeling, the pixels of the study area make up the space on which the Maxent probability distribution is defined, pixels with known species occurrence records constitute the sample points, and the features are climatic variables, elevation, soil category, vegetation type or other environmental variables, and functions thereof (Phillips et al. 2006). Maxent offers many advantages performing extremely well in predicting occurrences in relation to other approaches (e.g., Elith et al. 2006, Phillips et al. 2006, Elith & Graham 2009) such as the better discrimination of suitable versus unsuitable areas for the species (Phillips et al. 2006), a good performance on small samples (Phillips & Dudik 2008), and theoretical properties that are analogous to the unbiased case when modeling presence-only data (Phillips et al. 2009), this is why it has been frequently used.

Model evaluation can be done by different approaches. One of the most common ones for model evaluation is the calculation of the Receiver Operating Curve (ROC) (DeLong et al. 1988). ROC plot is obtained by plotting all sensitivity values (true positive fraction) on the y axis against their equivalent (1 – sensitivity) values (false positive fraction) for all available thresholds on the *x* axis. The area under the ROC curve (AUC) provides a threshold-independent measure of overall model accuracy. AUC values should be between 0.5 (random) and 1.0 (perfect discrimination). Values lower than 0.5 indicates that prediction is worse than random (Fielding & Bell 1997).

Another option for model evaluation is measuring the model predictive success, which is the percentage of occurrence data correctly classified as positive, so measuring the omission error rate. This evaluation requires a specific threshold to convert continuous model predictions to a dichotomous classification of presence/absence (Hernandez et al. 2006). Optimal thresholds are presented and discussed on a comparative study by Liu et al. (2005). Also, Lobo et al. (2008) recommends that sensitivity and specificity should be also reported, so that the relative importance of commission and omission errors can be considered to assess the method performance.

References

Araujo M. B. & Guisan A. 2006. Five (or so) challenges for species distribution modelling. Journal of Biogeography 33(10), 1677-1688.

Austin M. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecological Modelling 200, 1-19.

Busby J. R. 1986. A biogeoclimatic analysis of *Nothofagus cunninghamii* (Hook) Oerst. in southeastern Australia. Australian Journal of Ecology 11, 1-7.

Carpenter G., Gillison A. N. & Winter J. 1993. DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. Biodiversity and Conservation 2, 667-680.

DeLong E. R., DeLong D. M. &, Clarke-Pearson D. L. 1988. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves. Biometrics 44(3), 837-845.

Elith J., Graham C. H., Anderson R. P., Dudík M., Ferrier S., Guisan A., Hijmans R. J., Huettmann F., Leathwick J. R., Lehmann A., Li J., Lohmann L. G., Loiselle B. A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J. M., Peterson A. T., Phillips S. J., Richardson K. S., Scachetti-Pereira R., Schapire R. E., Soberon J., Williams S., Wisz M. S. & Zimmermann N. E. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129-151.

Elith J. & Graham C. H. 2009. Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. Ecography 32, 66-77.

Elith J. & Leathwick J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology and Evolution Systematics 40, 677-97.

Ferrier S., Watson G., Pearce J. & Drielsma M. 2002a. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New SouthWales. I. Species-level modeling. Biodiversity and Conservation 11, 2275-2307.

Ferrier S., Watson G., Pearce J. & Drielsma M. 2002b. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. II. Community-level modeling. Biodiversity and Conservation 11, 2309-2338.

Fielding A. H. & Bell J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24, 38-49.

Franklin J. 2009. Mapping Species Distributions: Spatial Inference and Prediction. Cambridge Univ. Press, Cambridge, United Kingdom. 338 pp.

Guisan A. & Thuiller W. 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters 8, 993-1009.

Guisan A. & Zimmermann N. E. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135, 147-186.

Hernandez P. A., Graham C. H., Master L. L. & Albert D. L. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29, 773-785.

Hirzel A. H. & Lay G. Le. 2008. Habitat suitability modelling and niche theory. Journal of Applied Ecology 45, 1372-1381.

Hirzel A. H., Hausser J., Chessel D. & Perrin N. 2002. Ecological-niche factor analysis: How to compute habitat-suitability map without absence data. Ecology 83, 2027-2036.

Jiménez-Valverde A., Lobo J. M. & Horta J. 2008. Not as good as they seem: The importance of concept in species distribution modeling. Diversity Distributions 14, 885-890.

Kearney M. 2006. Habitat, environment, and niche: What are we modeling? Oikos 115, 186-191.

Liu C., Berry P. M., Dawson T. P. & Pearson R. G. 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28(3), 385-393.

Lobo J. M., Jiménez-Valverde A. & Real R. 2008. AUC: a misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography 17, 145-151.
Nix H. A. 1986. A biogeographic analysis of Australian elapid snakes. *In* Atlas of Australian Elapid Snakes. Longmore R. (Ed). Australian Government Publishing Service, Canberra, Australia, pp. 4-15.
Pearson R.G. 2007. Species' Distribution Modeling for Conservation Educators and Practitioners. Synthesis. American Museum of Natural History. Available at http://ncep.amnh.org.
Peterson A. T. 2001. Predicting species' geographic distributions based on ecological niche modeling. Condor 103, 599-605.
Peterson A. T. 2006. Uses and requirements of ecological niche models and related distributional models. Biodiversity Informatics 3, 59-72.
Peterson A. T. & Nakazawa Y. 2008. Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. Global Ecology and Biogeography 17, 135-144.
Phillips S. J. & Dudik M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31, 161–175.
Phillips S. J., Anderson R. P. & Schapire R. E. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190, 231-259.
Phillips S. J., Dudík M. & Schapire R. E. 2004. A maximum entropy approach to species distribution modeling. *In* Proceedings of the 21st International Conference on Machine Learning. 21st International Conference on Machine Learning. ACM Press, New York, pp. 655-662.
Phillips S. J., Dudík M., Elith J., Graham C. H., Lehmann A., Leathwick J. & Ferrier S. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications 19(1), 181-197.
Pulliam H. R. 2000. On the relationship between niche and distribution. Ecology Letters 3, 349-361.
Rodríguez-Soto C. R., Monroy-Vilchis O., Maiorano L., Boitani L., Faller M. A., Briones M. A., Nunez R., Rosas-Rosas O., Ceballos G. & Falcucci A. 2011. Predicting potential distribution of the jaguar (*Panthera onca*) in Mexico: identification of priority áreas for conservation. Diversity and Distribution 17, 350-361.
Sanderson E., Redford K., Chetkiewicz C., Medellin R., Rabinovitz A. R., Robinson J. G. & Taber A. 2002. Planning to save a species: the jaguar as a model. Conservation Biology 16, 58-72.
Soberón J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. Ecology Letters 10, 1115-1123.
Soberón J. & Peterson A. T. 2004. Biodiversity informatics: managing and applying primary biodiversity data. Philosophical Transactions of the Royal Society B 359, 689-698.
Soberón J. M. & A. T. Peterson. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodiversity Informatics 2, 1-10.
Soberón J. & Nakamura M. 2009. Niches and distributional areas: Concepts, methods, and assumptions. The Proceedings of the National Academy of Sciences (PNAS) 106(2), 19644-19650.
Stockwell D. R. B. & Peters D. P. 1999. The GARP modelling system: Problems and solutions to automated spatial prediction. International Journal of Geographical Information Systems 13, 143-158.
Vaughan I. P. & Ormerod S. J. 2005. The continuing challenges of testing species distribution models. Journal of Applied Ecology 42, 720-730.

Veloz S. D. 2009. Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. Journal of Biogeography 36, 2290-2299.

Wilson C. D., Roberts D. & Reid N. 2010. Applying species distribution modelling to identify areas of high conservation value for endangered species: A case study using *Margaritifera margaritifera* (L.). Biological Conservation 144(2), 821-829.