



Species distribution modeling in the tropics: problems, potentialities, and the role of biological data for effective species conservation

Authors: Cayuela, L., Golicher, D. J. , Newton, A. C. , Kolb, M., de Albuquerque, F. S. , et al.

Source: Tropical Conservation Science, 2(3) : 319-352

Published By: SAGE Publishing

URL: <https://doi.org/10.1177/194008290900200304>

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Research Article

Species distribution modeling in the tropics: problems, potentialities, and the role of biological data for effective species conservation

L. Cayuela^{1,2,*}, D. J. Golicher^{3,4}, A. C. Newton⁴, M. Kolb⁵, F. S. de Albuquerque⁶, E. J. M. M. Arets^{7,8}, J. R. M. Alkemade⁸, and A. M. Pérez⁹

¹ Fundación Internacional para la Restauración de Ecosistemas (FIRE), Edificio de Ciencias Ambientales, Despacho B, Campus Universitario, Universidad de Alcalá, Carretera de Barcelona km. 33,600, E-28871 Alcalá de Henares, Madrid, Spain.

² Grupo de Ecología Terrestre, Departamento de Ecología, Centro Andaluz de Medio Ambiente, Universidad de Granada – Junta de Andalucía, Avenida del Mediterráneo s/n, E-18006, Granada, Spain.

³ Departamento de Ecología y Sistemática Terrestres, División de Conservación de la Biodiversidad, El Colegio de la Frontera del Sur, Carretera Panamericana y Periférico Sur s/n, C.P. 29290, San Cristóbal de las Casas, Chiapas, Mexico.

⁴ School of Conservation Sciences, Bournemouth University, Talbot Campus, Poole, Dorset BH12 5BB, United Kingdom.

⁵ Dirección Técnica de Análisis y Prioridades, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), Liga Periférico – Insurgentes Sur No. 4903, Col. Parques del Pedregal, Delegación Tlalpan, 14010 México, D.F., Mexico.

⁶ Departamento de Ecología, Universidad de Alcalá, Carretera de Barcelona km 33,600, E-28871 Alcalá de Henares, Madrid, Spain.

⁷ Alterra, Wageningen UR, P.O. Box 47, 6700 AA Wageningen, The Netherlands.

⁸ Netherlands Environmental Assessment Agency, P.O. Box 303, 3720 AH Bilthoven, The Netherlands.

⁹ Asociación GAIA. Biodiversidad, Medio Ambiente y Sociedad. Colonia 10 de Junio. Casa B-501. Managua, Nicaragua.

* Corresponding author: lcayuela@ugr.es; Phone: (+34) 958241000 (ext. 31202); Fax: (+34) 958137246.

Received: Received 10 March 2009; Accepted 18 April 2009, Published: 10 August, 2009

Copyright: © L. Cayuela, D. J. Golicher, A. C. Newton, M. Kolb, F. S. de Albuquerque, E. J. M. M. Arets, J. R. M. Alkemade, and A. M. Pérez. This is an open access paper. We use the Creative Commons Attribution 3.0 license <http://creativecommons.org/licenses/by/3.0/> - The license permits any user to download, print out, extract, archive, and distribute the article, so long as appropriate credit is given to the authors and source of the work. The license ensures that the published article will be as widely available as possible and that the article can be included in any scientific archive. Open Access authors retain the copyrights of their papers. Open access is a property of individual works, not necessarily journals or publishers.

Cite this paper as: Cayuela, L., Golicher, D. J., Newton, A. C., Kolb, M., de Albuquerque, F. S., Arets, E. J. M. M., Alkemade J. R. M. and Pérez, A. M. 2009. Species distribution modeling in the tropics: problems, potentialities, and the role of biological data for effective species conservation. *Tropical Conservation Science* Vol. 2(3):319-352. Available online: www.tropicalconservationscience.org

Abstract

In this paper we aim to investigate the problems and potentialities of species distribution modeling (SDM) as a tool for conservation planning and policy development and implementation in tropical regions. We reviewed 123 studies published between 1995 and 2007 in five of the leading journals in ecology and conservation, and examined two tropical case studies in which distribution modeling is currently being applied to support conservation planning. We also analyzed the characteristics of data typically used for fitting models within the specific context of modeling tree species distribution in Central America. The results showed that methodological papers outnumbered reports of SDMs being used in an applied context for setting conservation priorities, particularly in the tropics. Most applications of SDMs were in temperate regions and biased towards certain organisms such as mammals and birds. Studies from tropical regions were less likely to be validated than those from temperate regions. Unpublished data from two major tropical case studies showed that those species that are most in need of conservation actions, namely those that are the rarest or most threatened, are those for which SDM is least likely to be useful. We found that only 15% of the tree species of conservation concern in Central America could be reliably modelled using data from a substantial source (Missouri Botanical Garden VAST database). Lack of data limits model validation in tropical areas, further restricting the value of SDMs. We concluded that SDMs have a great potential to support biodiversity conservation in the tropics, by supporting the development of conservation strategies and plans, identifying knowledge gaps, and providing a tool to examine the potential impacts of environmental change. However, for this potential to be fully realized, problems of data quality and availability need to be overcome. Weaknesses in current biological datasets need to be systematically addressed, by increasing collection of field survey data, improving data sharing and increasing structural integration of data sources. This should include use of distributed databases with common standards, referential integrity, and rigorous quality control. Integration of data management with SDMs could significantly add value to existing data resources by improving data quality control and enabling knowledge gaps to be identified.

Keywords: Collecting effort, Biodiversity conservation, Data shortage, Geographical bias, Niche-based models, Statistical modeling.

Resumen

En este trabajo se investigan los problemas y potencialidades de los modelos de distribución de especies (SDM de sus siglas en inglés) como una herramienta para el desarrollo e implementación de estrategias y políticas de conservación en regiones tropicales. Se revisaron 123 estudios publicados entre 1995 y 2007 en cinco de las revistas científicas más influyentes en el campo de la ecología y la conservación, y se examinaron dos casos de estudio en regiones tropicales en los cuales los modelos de distribución de especies se han utilizado para diseñar planes de conservación a nivel regional. También se analizaron las características de datos típicamente utilizados para ajustar modelos de distribución de especies, tomando como ejemplo la distribución de árboles en Centroamérica. Los resultados mostraron que los trabajos metodológicos superan a los estudios que utilizan los modelos de distribución de especies en un contexto más aplicado para definir prioridades de conservación, especialmente en los trópicos. La mayoría de las aplicaciones de los modelos de distribución de especies se han desarrollado en regiones templadas y están sesgadas hacia ciertos organismos, como mamíferos y aves. Además, en regiones tropicales, los modelos están, por lo general, menos validados que en zonas templadas. Dos casos de estudio no publicados en revistas científicas mostraron que las especies que más necesitan la implementación de medidas de conservación, concretamente las más raras y amenazadas, son aquellas para las cuales los modelos de distribución de especies tienen menos probabilidades de ser aplicados con éxito. En Centroamérica, tan solo el 15% de las especies de árboles pudieron ser modeladas utilizando una fuente sustancial de registros obtenidos de la base de datos VAST del *Missouri Botanical Garden*. La ausencia de datos limita la validación de los modelos en áreas tropicales, lo que a su vez restringe considerablemente la aplicación de los modelos de distribución de especies. A pesar de dichas limitaciones, los modelos de distribución de especies tienen un gran potencial para la conservación de la biodiversidad de los trópicos, pudiendo contribuir al desarrollo de planes y estrategias de conservación, la identificación de vacíos de información y la generación de herramientas que examinen el impacto potencial de los cambios ambientales sobre la distribución de las especies. Sin embargo, para que este potencial sea aprovechado al máximo, es importante considerar algunos de los problemas relacionados con la falta de datos y su calidad. Para ello, es necesario abordar sistemáticamente las debilidades de algunas de las bases de datos biológicos existentes incrementando los muestreos de campo, mejorando el intercambio de datos y la integración estructural de distintas fuentes de datos. Esto incluye el uso de bases de datos distribuidas bajo estándares comunes, la integridad referencial y el control riguroso de la calidad de los datos. La integración de la gestión de datos a través de modelos de distribución de especies aportaría un valor añadido a las fuentes de datos actuales al mejorar el control de su calidad y permitir la identificación de vacíos de información.

Palabras clave: Ausencia de datos, Conservación de la biodiversidad, Esfuerzo de colecta, Modelación estadística, Modelos de nicho, Sesgos geográficos

Introduction

Predictive species distribution models are empirical models relating field observations to environmental variables, based on statistically or theoretically derived response surfaces [1, 2]. The most common strategy for estimating the potential geographic distribution of a species is to characterize the environmental conditions that are suitable for that species. The spatial distribution of environments that are suitable for a species can then be estimated across a given study region. A wide variety of modeling techniques have been developed for this purpose (see Appendix 1), including generalized linear models, generalized additive models, bioclimatic envelopes, habitat suitability indices, and the genetic algorithm for rule-set prediction (GARP).

Species distribution modeling (SDM) has become increasingly popular in recent years among researchers. It has been used to address a variety of different problems at various scales, with a range of different species occurring in different geographic areas. Applications of SDM methods include quantifying the environmental niche of species [3, 4], testing biogeographical, ecological, and evolutionary hypotheses [5, 6, 7], assessing species' invasions [8, 9], assessing the impact of climate, land use, and other environmental changes on species distribution [10, 11, 12], suggesting unsurveyed sites of high potential of occurrence for rare species [13, 14, 15], and supporting conservation planning and reserve selection [16, 17].

There are several particular advantages to using SDM to support conservation planning: (1) Maps of documented occurrences of species convey no information on the likelihood of occurrence in areas that have not been surveyed. Range maps from field guides and similar data are often too coarse to be useful for on-the-ground conservation action or research. (2) Accurate predictive distribution maps make field inventories more efficient and effective. They show where to commit the limited available resources for inventories by highlighting the areas where a targeted species or habitat type is most likely to be found. (3) Predictive distribution maps for multiple species or habitat types, produced with consistent and reliable methods, are well suited for identifying spatial patterns in biological diversity, which can be of value for assessing conservation priorities. (4) Predictive distribution maps are very useful for conservation planning efforts at a range of different scales. As a result of these advantages, in the last decade, a number of international organizations have employed species modeling in order to address key policy objectives at a global scale (e.g., UNEP, the Convention of Biological Diversity, Organization for Economic Co-operation and Development, European Union, Conservation International, IUCN, WWF, etc.).

Several statistical issues, however, stand as obstacles for species distribution analysis. The first and foremost is data availability [18]. Much biodiversity has yet to be formally described and catalogued. In general, this problem—the so-called “Linnean shortfall” [19]—appears to be of increasing relevance as the organisms decrease in size [20]. In addition, knowledge of the global, regional, and even local distributions of many taxa is currently inadequate, a problem that Lomolino [21] named the “Wallacean shortfall.” Many areas of the world remain seriously under-collected for most taxa, with the result that even for higher plants, reliable systematic species range maps are available only for a fraction of the earth's surface [20]. Many of these problems are particularly intense in tropical areas. Whereas it is widely appreciated that most megadiverse areas occur in the tropics [22], rates of habitat loss and environmental degradation also tend to be higher in tropical regions [23, 24, 25, 26, 27]. Therefore, the need for tools to assist conservation planning, policy development, and implementation is particularly urgent in tropical regions.

In this paper, we first examine the extent to which SDM techniques are being developed and applied in tropical regions, based on the results of a literature review. We then use this review to evaluate the scope

and objectives of SDM initiatives in tropical areas, in comparison to those undertaken in temperate regions. We then explore the potential limitations to the application of SDM in the tropics, by examining two case studies in detail. Finally, we discuss how the potential value of SDM approaches might best be realized in future, given the current limitations that exist.

Literature review

We conducted a literature search on the *Web of Science* using the keywords “species distribution” AND “model.” We selected the five journals containing the largest number of studies using SDM that could be considered to provide information suitable for guiding conservation policies. The selected journals were *Biological Conservation*, *Conservation Biology*, *Diversity and Distributions*, *Global Ecology and Biogeography*, and *Journal of Applied Ecology*. We reviewed all papers published between January 1995 and May 2007, selecting those that had used SDMs to predict species distributions, but excluding review papers and studies using raw geographical data without the inclusion of statistical models. Our search yielded 123 papers (Appendix 2, references are provided in Appendix 3) that satisfied the inclusion criteria. From each study we extracted the following information: (1) study region; (2) aim of the study (classified as “methodological” if it focused on the development of modeling methods or compared the accuracy of different modeling techniques; or “applied” if it involved the application of SDM methods to practical conservation problems, such as biological invasions, climate change, conservation prioritization, and biodiversity mapping); (3) model validation (classified according to the approaches used as “non reported,” “validated with the same data,” “k-fold partitioning,” “prospective sampling,” and “informal validation”); (4) focal taxa; and (5) data type (classified as “presence only,” “presence-absence,” and “abundance” data).

Most studies in our literature review were conducted using data from relatively well sampled countries (Appendix 2), such as the USA and Canada (28), Australia (10), or European countries (48). Relatively few used data from tropical regions such as Central and South America (8), Africa (10), or Asia (6).

We found that 39% of papers (48) focused on the development or evaluation of methods rather than on their application (Appendix 2). This is an unexpectedly high proportion, taking into account exclusion of methodological journals such as *Ecological Modelling*. These papers documented studies presenting new methodologies for predicting species distributions [e.g., 28, 29, 30], and those evaluating the performance of different models [e.g., 31, 32, 33]. Some also explored a variety of issues related to spatial scale and extent [e.g., 34, 35], model accuracy [e.g., 36, 37], or variable selection [e.g., 38]. Overall, the high proportion of methodological papers may be indicating that the use of these techniques are not free of controversy in their application. Among those studies that used SDM in a more applied context, there was a broad mixture of goals, including species conservation (29 papers), biological invasions (14 papers), climate change (10 papers), autoecology (7 papers), and biogeography (6 papers). Conservation prioritization was mentioned in only six papers and biodiversity mapping in three. When stratified by biomes, it can be observed that most methodological studies were carried out in temperate regions, while applied studies tended to be conducted in tropical regions (**Fig. 1a**).

It is generally accepted that a robust test for the prediction success of a model should include independent data, i.e., data not used to develop the predictive model. However, only 13% of the reviewed studies (17) validated SDMs using a new sample of cases obtained from a different region or time after the model had been developed (referred to as prospective sampling in Appendix 2). Just 8% of studies (10) used some sort of informal validation, e.g., by comparing the SDMs with existing distribution atlases [40, 41] or through literature review [42]. A large proportion of studies (48; 38%) partitioned the data into subsets and used one of these sets for training and the remaining sets for testing purposes. Though this is a common practice, data partitioning is not the same as collecting new independent data

for model testing [43], particularly if predictions are to be tested for their general use. Only if predictions are to be restricted to a homogeneous region can data-partitioning be expected to output similar results than prospective sampling [44]. Finally, 17% of studies (22) used the same data for training and testing, and 24% (31) did not report any validation at all. When stratified by biomes, the number of studies using prospective sampling or some sort of informal validation was proportionally higher in temperate (22%) than in tropical regions (15% respectively, **Fig. 1b**). On the contrary, the number of studies not reporting any form of validation or using data partitioning was proportionally higher in tropical (26% and 41% respectively) than in temperate regions (22% and 38%). The number of studies using the same dataset for validation was roughly the same (18%) for both temperate and tropical regions. Overall, this indicates that a large proportion of the reviewed studies—especially those undertaken in tropical regions—reported model results without rigorous testing of model properties. This was particularly the case in applied studies, where lack of reporting of validation and testing SDM with the same data was commonplace (**Fig. 1c**). In contrast, methodological studies accounted for the largest proportion of cases validated through prospective sampling.

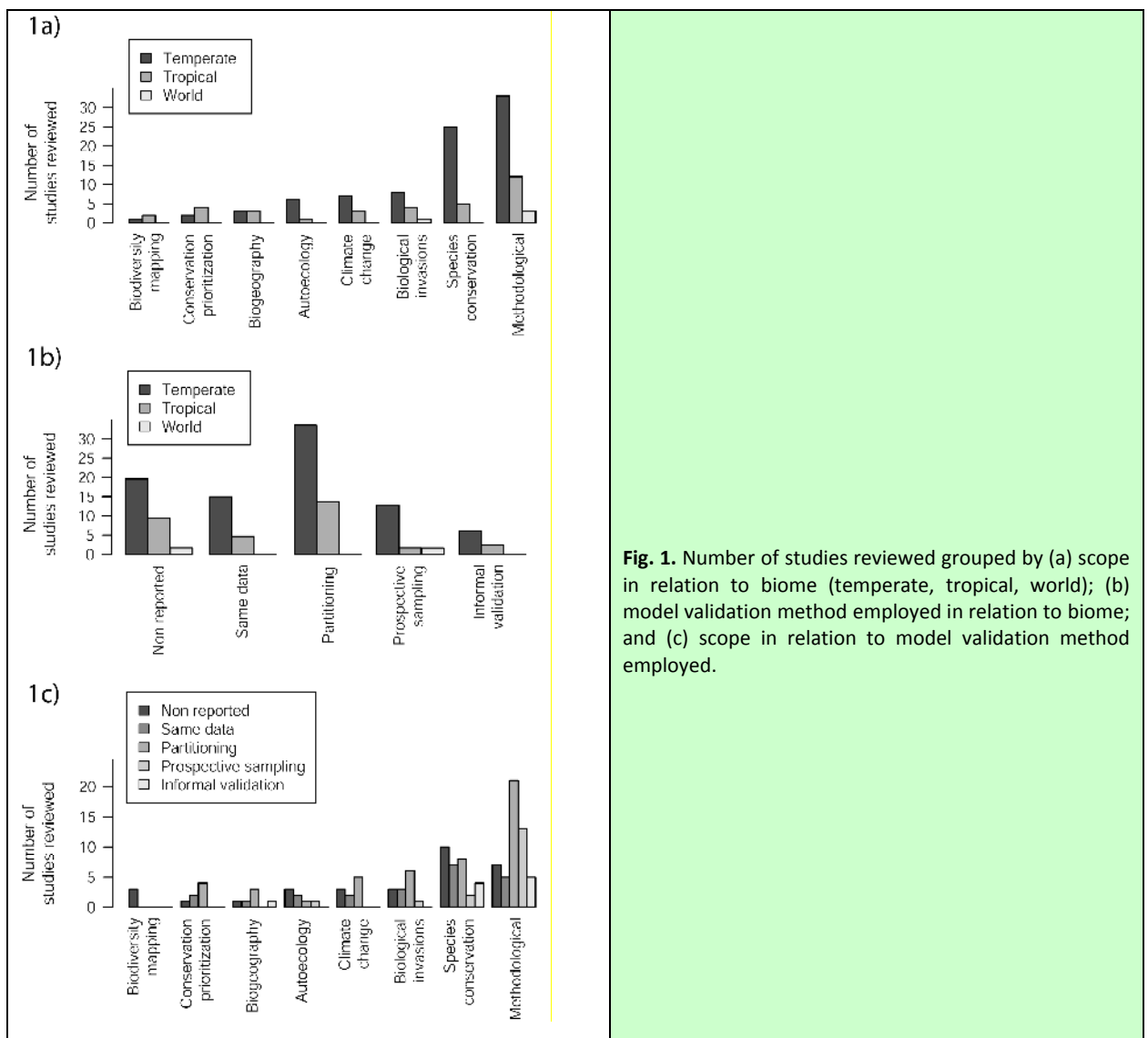


Fig. 1. Number of studies reviewed grouped by (a) scope in relation to biome (temperate, tropical, world); (b) model validation method employed in relation to biome; and (c) scope in relation to model validation method employed.

Most papers were published on birds (46), plants (34), or mammals (29, Appendix 2). Groups receiving less attention were reptiles (11), amphibians (7), fishes (2), lower plants (2), and invertebrates (24). Invertebrates encompassed many different taxa such as insects, arachnids, snails, crustaceans, and rotifers. The shortage of studies focusing on this group contrasts markedly with the high diversity of organisms that are represented within it.

Finally, 60 out of 123 papers used presence-only data for modeling species distributions, 60 used presence-absence data, and only seven used abundance data (from which four studies also used presence-absence data, Appendix 2). A high proportion of the studies conducted in tropical regions (73%) used presence-only data, whereas studies conducted in temperate regions most often used presence-absence or abundance data (59%).

Overall, these results show that studies using SDM in the tropics are more scarce than in temperate regions, which stands in contrast with the high biodiversity held by tropical ecosystems. Fairly complete datasets from well-sampled regions make possible the development or evaluation of methods. However, these methods are not always as effective when applied to conservation case studies in tropical regions. Results from such case studies reveal data-driven constraints that limit the applicability of SDM, such as lack of independent information to validate the predicted distribution of species or lack of reliable absence data.

Application of SDM in the tropics: two case studies

To explore current approaches to applying SDM techniques in the tropics and highlight the problems encountered in greater depth, two case studies are described here in detail. Although none of these case studies have been published previously in scientific journals, they were selected because they were commissioned by conservation organizations, so they can potentially illustrate how well SDMs meet the expectations posed by governmental and non-governmental conservation organizations in tropical countries. The first study reported in this section was published on-line [45] and was part of a larger Andes-Amazon project commissioned by NatureServe. The second study was provided by one of the authors, who works for the Mexican Commission for the Knowledge and Use of Biodiversity (CONABIO).

The first study was conducted on the eastern slope of the Andes in Peru and Bolivia to model the real distribution of endemic species [45]. The study aimed to fill knowledge gaps in support of conservation planning in the Tropical Andes. The list of focal species included 115 birds, 55 mammals, 177 amphibians, and 435 plants. The Maxent algorithm was selected for modeling species distribution because previous comparative studies had shown that it performs well even with small sample sizes [46, 47, 48]. Maxent output predicted four distributions for each species using all the available locality data but varying the input environmental layers (see [45] for further details). Because of the scarcity and low spatial precision of available locality data, it was not possible to partition the data into records used for training the model and those set aside for a statistical model evaluation. Alternatively, specialists in each group reviewed and selected which, if any, of the four models reflected a realistic depiction of the distribution. This decision was based partly on validation with the same dataset used for modeling and partly on expert judgment. In the cases in which a Maxent model was considered to be reasonable, the reviewers then selected a cut-off threshold to convert the continuous Maxent predictions to presence-absence maps. Despite the alleged suitability of this method for this purpose [46, 47, 48], there were many cases where the Maxent models did not produce a realistic distribution map for the species. In such cases, deductive and hybrid models were relied on. Deductive models were created by defining the maximum and minimum elevations at which the species was expected to occur. Hybrid models used part of the Maxent prediction in one portion of the species' range and a deductive model for the remaining area. Table 1 shows that a large proportion of the target species could not be effectively modelled with Maxent. Endemic

amphibians and plants were particularly challenging; 52.0% and 39.3% of the Maxent models for these two groups, respectively, did not produce realistic distribution maps. Similarly, the mean number of records per species and the number of species with one record for endemic amphibians (6.0 and 36.7% respectively) and plants (7.0 and 28.3%, respectively) were comparatively smaller than for birds (21.2 and 2.6%, respectively) and mammals (11.2 and 7.3%, respectively) (Table 1). This highlights the problems of using objective modeling approaches for analyzing range delimitation when quantitative data are lacking.

Table 1. Summary of biological data and modeling methods used in the prediction of endemic species distributions on the east slope of the Andes in Peru and Bolivia [45]. The Maxent method was used for modeling when possible. Where this method did not output a realistic distribution map for the species, deductive and hybrid models were used. Deductive models were created by defining the maximum and minimum elevations at which the species was expected to occur. Hybrid models used part of the Maxent prediction in one portion of the species' range and a deductive model for the remaining part. In brackets, the proportion of the species in each taxon modelled with each of the three methods is given.

Taxon	Number of species	Maximum number of records per species	Number of species with one record	Mean number of records per species	Modeling methods used		
					Maxent	Hybrid	Deductive
Birds	115	94	3 (2.6)	21.2	99 (86.1)	6 (5.2)	10 (8.7)
Mammals	55	70	4 (7.3)	11.2	47 (85.5)	0 (0t.0)	8 (14.5)
Amphibians	177	64	65 (36.7)	6.0	85 (48.0)	8 (4.5)	84 (47.5)
Plants	435	84	123 (28.3)	7.0	264 (60.7)	18 (4.1)	153 (35.2)

Our second example refers to the Gap Analysis for terrestrial biodiversity in Mexico (http://www.conanp.gob.mx/pdf_vacios/terrestre.pdf), an initiative coordinated by CONABIO and CONANP (National Commission for Protected Areas). This initiative aims to assess the efficiency of the current network of protected areas to conserve a representative part of the country's biodiversity and generate a strategy for adapting the protected area system [48]. As part of this task, species distribution maps were generated by experts on several taxonomic groups using the GARP algorithm [50] (Table 2). Species geographical distributions were constructed from raw occurrence data obtained from the National Biodiversity Information System (SNIB, CONABIO) and the World Information Network on Biodiversity (REMIB), in combination with datasets of environmental variables believed to affect species distributions in Mexico. Data layers included climatic variables from Worldclim (<http://www.worldclim.org/>), topographic and hydrologic parameters from Hydro1k (<http://edc.usgs.gov/products/elevation/gtopo30/hydro/index.html>), and thematic national datasets from INEGI and CONABIO. The species selected for this purpose were catalogued under the Mexican Red List of Threatened Species NOM-059-SEMARNAT-2001, were range restricted or rare, or belonged to taxonomic groups of particular conservation concern (e.g., *Agaveacea* spp., *Opuntia* spp.). Because the GARP algorithm had proved to be ineffective in earlier studies when few records were available [e.g., 51], the decision was made to model only species with at least eight records. This excluded 56% of the species catalogued under the Mexican Red List, and 35% of the non-threatened species (Table 2). Amphibians and reptiles distribution maps were validated with the best subset procedure, using half of the data to build the model and the other half to test the predictive ability of the model (i.e., data partitioning). SDMs for

mammals and plants were validated with data from a literature search and expert knowledge (i.e., informal validation). In addition, knowledge on the ecology and biogeography of the species, coarse-scale maps (ecoregions, biogeographic realms), and auxiliary datasets held by experts (e.g., sightings and field specimens not included in the datasets used for modeling) were used to trim SDMs in order to eliminate, at least to some extent, possible model over-predictions and improve estimates of the actual species distributions. This highlights the limitations of SDM for conservation of endangered species, as sufficient data for effective application of SDM are only available for a small minority of species. Consequently, those species that are most likely to require conservation action, namely those that are the rarest or most threatened, are those for which SDM is least likely to be useful.

Table 2. Number of species targeted for distribution modeling by CONABIO, as part of a national conservation assessment initiative between 2004 and 2006 [49]. A total of 1,843 were initially selected for modeling, including 166 amphibians and 435 reptiles [102] (Flores-Villela, unpublished data), 336 trees, 294 agave species and 612 plant species not included in any of the previous categories (shrubs, grasses, etc., CONABIO, unpublished data). The table shows the number of species whose distributions were and were not successfully modelled using GARP [49]. Species distribution models (SDM) were not successful when there were fewer than eight records available. 71.1% of the target species (1311) were catalogued under the Mexican Red List of Threatened Species NOM-059-SEMARNAT-2001 and were, therefore, of particular conservation concern. In brackets, percentage of the species successfully and unsuccessfully modelled under each of the three categories is given.

	Threatened		Non threatened		Total	
	Successful	Unsuccessful	Successful	Unsuccessful	Successful	Unsuccessful
Amphibians	87 (52.4)	79 (47.6)	0 (0.0)	0 (0.0)	87 (52.4)	79 (47.6)
Reptiles	234 (53.8)	201 (46.2)	0 (0.0)	0 (0.0)	234 (53.8)	201 (46.2)
Trees	90 (60.4)	59 (39.6)	180 (96.3)	7 (3.7)	270 (80.4)	66 (19.6)
Agave spp.	18 (37.5)	30 (62.5)	90 (36.6)	156 (63.4)	108 (36.7)	186 (63.2)
Other plants	139 (27.1)	374 (72.9)	76 (76.8)	23 (23.2)	215 (35.1)	397 (64.9)
Total	568 (43.3)	743 (56.7)	346 (65.0)	186 (35.0)	914 (49.6)	929 (51.4)

The main lesson that can be drawn from these case studies is that statistical modeling is not effective when few data points are available. SDMs fail to produce reliable predictions in cases where the distribution data is very limited and, in such cases, predictions must rely mostly on subjective judgment. We believe that the problems encountered in these two case studies are common to many tropical taxa and regions, as explored in the following section.

Problems of applying SDM approaches in tropical regions

Data shortage

The most important problem that species distribution modellers in tropical regions often have to face is the small number of available data points. Lack of information about the distribution of organisms, what has been referred to as the “Wallacean shortfall” [21], is widely recognised to be a major constraint to conservation planning in the tropics [22, 52, 53, 54]. In addition, it represents a problem to SDM approaches. Previous studies have shown that a sample size lower than around 70 observations decreases the performance of SDMs [55, 56, 57]. Drake et al. [29] studied how model performance depends on the sample size of the training dataset, and concluded that at least 40 observations were necessary to obtain consistent models using support vector machine-based methods (see Appendix 1). The GARP algorithm

was reported to consistently under-predict the distribution of mammal species in poorly surveyed regions of west-central Guyana [51].

To analyze the characteristics of the data typically available for fitting SDMs in the tropics, we downloaded presence-only data (hereafter referred to as “records”) of all known tropical tree species in Central America from the VAST database of the Missouri Botanical Garden (MOBOT, <http://mobot.mobot.org/W3T/Search/vast.html>). We excluded records without geographical coordinates, or with coordinates derived from political districts. This provided a list of 3,359 species with a total of 135,241 records. We found that 8% of the species consisted of a single record, 21% had fewer than five records, and 50% fewer than 17 records. If a limit of 40 known occurrence points is considered to be the minimum for rigorous modeling (following Drake et al. [29]), the distribution of only 30% of the species could be effectively modelled using this dataset. In addition, 313 of these species have been categorized as at risk of extinction (CR, EN, VU) in the IUCN Red List, and are thus potentially of some conservation concern. Only 15% of these were found to have more than 40 records, while 40% had fewer than five. The most-collected species in the list were those considered to be of least conservation concern. This pattern seems to be the case in many other regions of the world [58, 59].

Data paucity has a range of causes. It can occur even when collection effort is relatively intense. Small-bodied or nocturnal species can be difficult to detect. Many species are genuinely rare, but rarity takes many different forms [60, 61]. A species can have a broad range but have low population abundances, or a narrow range, but be locally abundant. Additionally, data shortage can result from sampling bias towards certain taxa [62]. As noted in the literature review, some groups of organisms (such as birds and mammals) tend to attract greater interest from researchers than others. This bias has been well documented; Keddy [63] refers to it as “moose-goose” syndrome. The distribution of many groups of taxa, such as invertebrates, reptiles, amphibians, bryophytes, and fungi, tends to be relatively poorly documented.

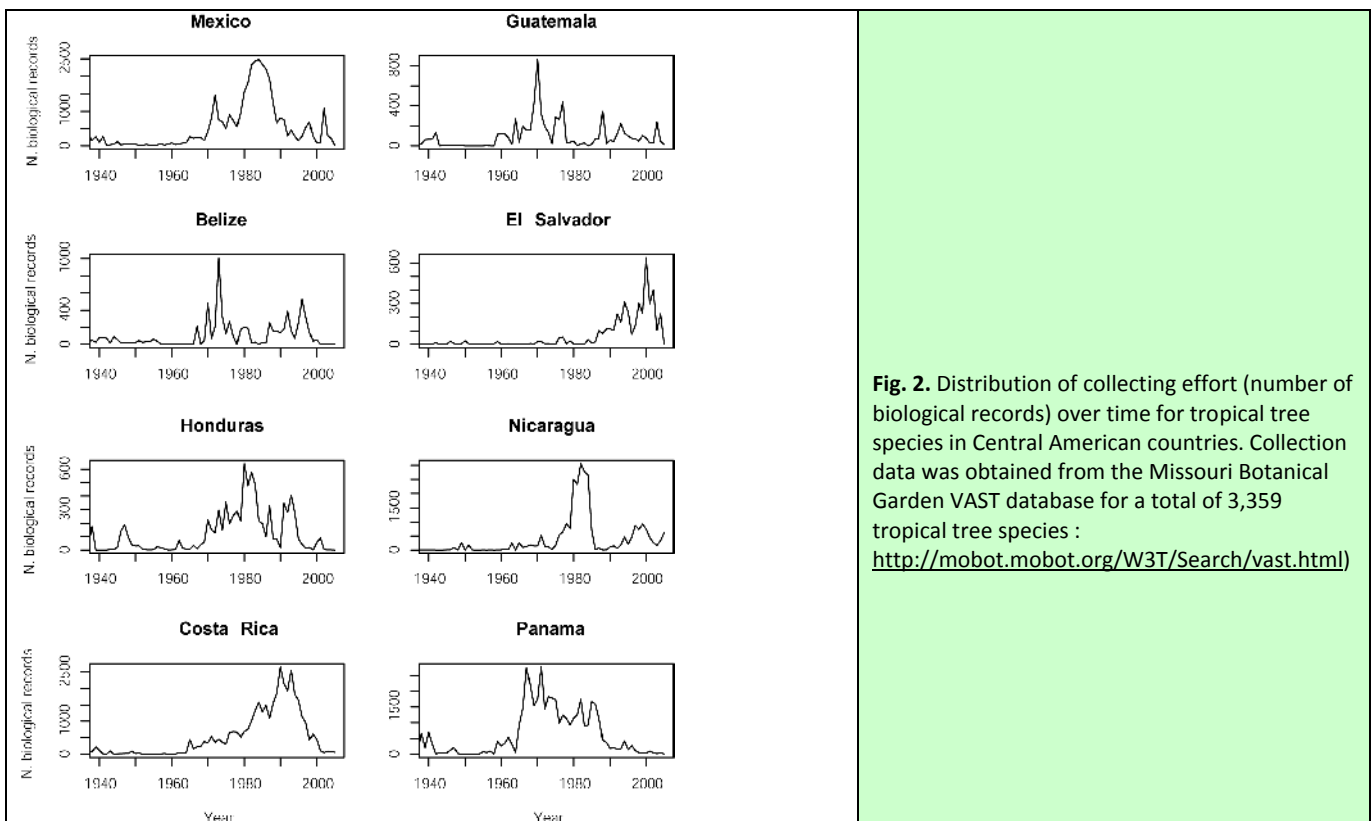
Table 3. Density of tropical tree herbaria specimens for Central America and Great Britain. Collection data for Central America has been obtained from the Missouri Botanical Garden VAST database (<http://mobot.mobot.org/W3T/Search/vast.html>) for a total of 3,359 tropical tree species. Collection data for Great Britain has been obtained from the New Atlas of British and Irish Flora and from the National Biodiversity Network (<http://www.searchnbn.net>) for a total of 137 tree species. Collection data for the Netherlands has been obtained from the Florbase (www.florbase.nl) database and the 'landelijke vegetatie database' (www.synbiosys.alterra.nl/lvd) for a total of 206 tree species.

Country	Number of specimens	Land area (km ²)	Data density (specimens/100 km ²)	Number of tree species in collections
Tropical Mexico ¹	18,756	442,858	4.2	1672
Guatemala	6575	108,889	6.0	1162
Honduras	8470	112,088	7.6	1105
Nicaragua	27,837	130,000	21.4	1301
El Salvador	4820	21,041	22.9	681
Panama	25,262	75,517	33.4	2253
Belize	8508	22,966	37.0	820
Costa Rica	34,372	51,100	67.3	1954
Great Britain and Ireland	590,560	315,100	187.1	137
The Netherlands	1,377,537	41,528	3,317.3	206

¹ Tropical Mexico includes the states of Yucatan, Quintana Roo, Campeche, Tabasco, Chiapas, Oaxaca and Veracruz.

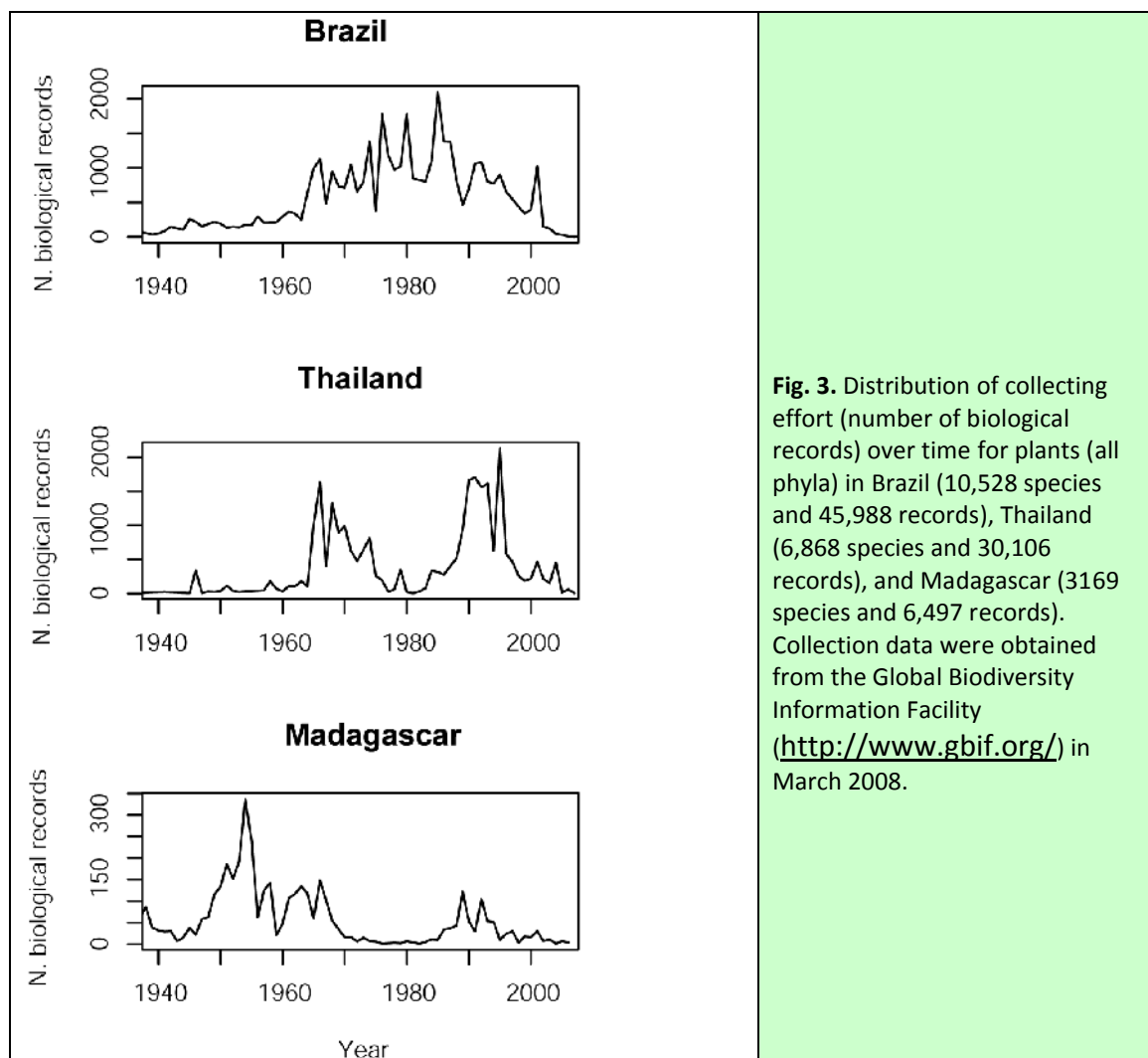
In addition, collecting effort is also unevenly distributed across countries (Table 3). If we compare the spatial density of tree collection specimens for different countries of Central America we find large differences in proportional collecting effort. Based on these data, the biodiversity of large countries appears to receive proportionally less research attention than that of smaller countries. In the Neotropics, large countries also tend to attract proportionally fewer visiting researchers per unit area. Tropical Mexico, Guatemala, and Honduras have low data density (4.2, 6.0 and 7.6 collection data 100 km⁻², respectively), while Costa Rica and Belize both have relatively high data densities of up to 67.3 collection data 100 km⁻² (Table 3). Yet all these data densities are low when compared with those of well-sampled countries such as Great Britain and Ireland or the Netherlands, where even with fewer tree species (137 and 206 species, respectively), much higher data densities have been recorded (187 and 3,317 records 100 km⁻², respectively) (Table 3).

There is also some evidence that the collecting effort for tropical taxa is declining, at least for some groups. The Mexican butterfly database described by Llorente et al. [64] contained 36,685 records collected between 1900 and 1990. When analyzing the utility of this database for conservation, Soberón et al. [65] found an abrupt increase in collecting effort in the 1970s and 1980s. Collecting effort in the 1990s, however, decreased to levels similar to the average between 1910 and 1950. We found similar trends for: (1) tropical tree species in Central American countries (**Fig. 2**), with collecting efforts mostly peaking between 1980 and 1990 and decreasing progressively from 1990 onwards; and (2) plants (all phyla) in Brazil, Thailand, and Madagascar (**Fig. 3**), with collecting effort dropping after peaking in the 1980's, 90's and 50's, respectively.



Data quality issues

Guisan et al. [66] clearly demonstrated the importance of data quality for model performance. Based on a high-quality database they compared several modeling techniques for predictive accuracy and sensitivity to, among others, location error, changes in map resolution, and sample size. They found that sample size and location error affected model performance in particular. Ideally, in order to model species distributions, sampling effort should be uniform across the species' range, so that all recorded variations in distribution patterns are real and not the result of variation in sampling effort [67, 68]. However, systematic surveys of large areas are rare, and therefore models focusing on large-scale patterns of species distribution often rely on incomplete and geographically biased information [68, 69, 70, 71, 72, 73]. This is particularly true for models based on specimen collections in herbaria and natural history museums [74]. Collection data are inherently biased in many respects [73, 75, 76]; therefore, models based on such data may lead to inaccurate predictions.



Geographic bias can come in many different forms, though bias owing to accessibility and a focus on priority areas are probably the most important [68]. The existence of roadside bias (the so-called “highway effect”) in survey and collection data has often been emphasized [3, 62, 77, 78, 79, 80], but less frequently quantified [65, 68, 73, 81, 82]. A similar sampling bias has been observed along rivers [68] and near cities [68, 81]. The effect of nature reserves or priority areas on biological record collection intensities is potentially complex. A common pattern is for such areas to receive attention by collectors prior to their declaration as reserves, followed by a decline in collecting due to restrictions imposed when reserve status is granted. We could not detect, for example, any difference in a historically pooled sample of collections within and outside nature reserves in Mexico for species such as the deer mouse (*Peromyscus* sp.), or various endangered agave species (*Agave* spp., M. Kolb, unpublished data). Likewise, Freitag et al. [70] detected no bias of small mammal survey records towards nature reserves in South Africa. However, they found that large mammal data had been mostly collected within existing conservation areas. Similarly, Reddy and Dávalos [68] reported a bias of passerine bird samples towards areas now designated as conservation priorities in sub-Saharan Africa.

The degree to which geographical bias affects the performance of distribution models has rarely been explored, but may be case specific. According to Kadmon et al. [73], a negative effect of roadside bias on predictive accuracy of bioclimatic models must follow from two necessary conditions: (1) climatic bias should affect the accuracy of model predictions; and (2) the road network should be biased climatically. Both conditions are met in tropical mountain regions where roads can be at low elevations and vegetation patterns are linked to altitudinally determined climatic gradients (D. Golicher and L. Cayuela, unpublished data). But even in apparently well surveyed regions and groups, such biases still produce inaccurate geographical model representations [83, 84], because the process of discovery of species distribution has occurred in a climatically or spatially structured fashion [85]. Stockwell and Peterson [86] suggest methods to correct this bias. However, these are difficult to implement when there are limited available data.

There is a well known general effect of geographical sampling bias in the context of SDM. When presence-only data are used, pseudo-absences or background absences (hereafter both referred to as pseudo-absences) are often used in order to fit models. Procedures for this are frequently integrated within the SDM software programs used by researchers, such is the case of GARP [50] or Maxent [48]. However, model users often do not explicitly investigate the properties of pseudo-absences or the impact of using pseudo-absences on overall model results. This is of paramount importance since, if there are not reliable absence data, the method of pseudo-absence selection strongly conditions the obtained model, generating different model predictions in the gradient between potential and realized distributions [87]. In addition, if a large area is being modelled, pseudo-absences may be taken from well beyond the species’ actual distribution limit. This can provide over-optimistic evaluation of a model’s predictive ability from inspection of ROC curves [88]. This has been referred to as the “naughty noughts” effect [89]. It is not easily avoided if the data available to suggest credible bounds of a species distribution are the same data that are later used in the SDM.

In addition to geographic biases, there are frequently errors in the geographical coordinates of specimens and data collections. Before accurate GPS technology became available, specimen collectors used a variety of ad-hoc descriptive protocols to record the localities where collections were made. These textual descriptions were then converted into geographic coordinates using available cartography. Records made before the mid 1990s are therefore inherently imprecise. Where place names are ambiguous, geographical errors may be considerable. Species distribution databases rarely include an explicit measurement of geographic precision. However, the degree of precision can be inferred from examining the last digits of the coordinates. We found that 90% of data points from MOBOT had apparently been

rounded up to the nearest arcminute and 8% to the nearest degree. Small errors have relatively minor consequences when data are used in the context of a traditional distribution atlas. However, the effect of even small positional errors on modern statistical distribution models is potentially more severe. Environmental variables are fed into modeling algorithms as a result of overlaying the points on interpolated raster coverages. In mountainous tropical terrain, temperature and precipitation are strongly correlated with steep elevational gradients. Small positional errors can thus result in markedly different climate parameters being associated with a collection point. This is likely to produce poor or misleading models [66]. The severity of this effect is a function of the size of the error and the specific topography of the region.

A final problem related to data quality is the risk of species misidentification. This is particularly problematic when collating information from different sources, as different datasets may have been generated with different taxonomic concepts [18]. To merge several data sources into one homogenous dataset is an enormous challenge that usually dwarfs the time required to analyze these data [7].

Implications for conservation

Strengthening the applied role of SDM in tropical conservation

Being empirical, SDMs are explicitly data driven. The accuracy of model predictions depends critically on the quality and quantity of data. Biological databases are, by nature, incomplete and have heterogeneous spatial coverage [68, 69, 70, 71, 72, 73, 75, 76]. This has led to the development of techniques aimed at overcoming the analytical challenges posed by incomplete spatial coverage. However, an inevitable tension arises when data-driven models are applied to conservation problems. On the one hand, the fundamental motivation behind the modeling exercise may be to fill data gaps by suggesting present or future species distributions that have not actually been observed. On the other hand, the need for scientifically rigorous tests of model predictions places demands that frequently cannot be met by the limited data available for tropical species of conservation concern.

Logistic difficulties and lack of resources remain a major barrier to data collection in the tropics [54]. At the same time, much invaluable data that have already been collected remain unavailable for modeling owing to unstructured data management. The successful collation of systematic records in relatively well-studied temperate regions can provide a positive role model for strengthening tropical data resources [90]. For example, botanical records in Great Britain and Ireland have been well organized since 1954, when the Botanical Society of the British Isles (BSBI) systematically divided Great Britain and Ireland into 3,500 10-km squares to aid surveying. The BSBI encouraged volunteer recorders to join their network. This was fundamental to the success of the project. By the following year, the BSBI network comprised 1,500 recorders that contributed to the field survey [91]. As a result of such approaches, the UK National Biodiversity Network (NBN) now provides excellent on-line access to detailed wildlife information at the national scale (available at <http://www.searchnbn.net>). Datasets have been contributed by over 50 distinct organizations and hold over 18 million records. A similar monitoring network is present in the Netherlands (Florbase), with more than 20 million records collected from the early 20th century onwards.

The British and the Dutch models rely on strengthening available formal collection data by drawing on qualified volunteer recorders. This bottom-up approach is also being promoted in some parts of the tropics, such as the BERDS database for Belize (<http://www.biodiversity.bz/>). The key feature of the Belizean initiative is the use of a spatially explicit relational database as a key tool for data storage, display, and analysis. Other examples that show potential ways forward for effective collaboration and sharing of plot data are the RAINFOR initiative [92] and the Amazon Plot Network [93].

Top-down initiatives can also strengthen the data available for SDMs, although they provide no new data by themselves. Because volunteer work is likely to be hindered by the high diversity of organisms present in most tropical ecosystems, top-down approaches may work better in the tropics than bottom-up ones. One such top-down initiative is the Global Biodiversity Information Facility (GBIF) [94]. GBIF has collated millions of data entries from natural history collections, library materials, and databases, though this information shows many of the taxonomic and geographic gaps and biases mentioned above [94]. Most of the information currently available still refers to developed countries. For example Spain has 97,295 registered entries for plants, while in countries such as Madagascar, India, or Philippines there are only 6,497, 1,544 and 2,639 available data entries respectively (data referring to March 2008 accessed through www.gbif.org). Another approach is the Conservation Commons (<http://www.biodiversity.org/>), which has identified a set of principles to promote sharing of biodiversity data, information, and knowledge to facilitate the conservation and sustainable use of biodiversity. These principles encourage organizations and individuals alike to ensure open access to data, information, expertise, and knowledge related to the conservation of biodiversity. However, application of these principles has been limited to date.

Such ongoing initiatives to improve data quality and quantity should support the future application of SDMs. Errors in geographic coordinates or taxonomic determination often become evident when SDMs are fitted to data. Although data cleaning is an unrewarding task for research scientists, the development of efficient methods for flagging and correcting dubious records could provide a clear applied role for SDMs [95]. Those involved in the design of data portals and the structure of databases should take into account the need of SDMs to connect directly with database servers as clients in order to automate computationally intensive iterated modeling [73]. This can now be easily achieved through protocols such as Open Data Base Connectivity (ODBC).

The resources being devoted to tropical field research are generally considered to be inadequate [54]. Many important aspects of the distribution and abundance of tropical organisms are likely to remain unknown. However, this also could reveal a positive application for SDMs. If a SDM provides poor predictions, this can be taken as a clear indication that more distribution data are required for the taxon in question. We also suggest that researchers using SDMs should become more open regarding the limitations of SDMs. This is particularly important when poor model results are clearly attributable to weak data rather than poorly constructed algorithms. Reviewers and editors should also be prepared to accept studies that rigorously document model failures as well as successes, in order to prevent the repetition of mistakes through over-optimistic expectations.

Overall, the field of SDM needs a serious reflection about the conceptual basis that underlies species distribution models, as well as about the true meaning of their predictions (potential versus realized distribution) [97]. The design of future works evaluating, comparing, and applying species distribution modeling techniques should be thus rooted in a good understanding of their conceptual background. If species distribution models are to be a common-use tool for biodiversity research and conservation assessment, the foundations of their application must be much more solid than they are now [97].

Finally, since data quality problems and data shortage appear to be very common, a pressing question is: what can be done with such biased and incomplete information? Statistical modeling is not effective when few data points are available. In such cases, input from some sort of expert judgment is inevitably required in order to evaluate which of the outputs from species distribution models are most credible. Expert knowledge is already recognized as an essential source of information for assessing the conservation status of species, given the lack of reliable quantitative information [59]. The development of tools to support the effective integration of expert knowledge with SDM approaches represents a key challenge for the future. Another approach might be to shift the focus of modeling from individual entities

to collective properties of biodiversity [16, 98], such as species assemblages or communities [12, 99, 100]. Although this does not fully circumvent the problem, it might be used to indicate where rare or threatened species are likely to be found in association with which other species [12].

Final remarks

We consider the following steps as vital to a further development of SDM within an applied conservation context in the tropics. First, to reinvigorate SDM applications, more emphasis should be given to mechanisms for improving data sharing and better structural integration of diverse data sources, using distributed databases with common standards, referential integrity, and rigorous quality control. Second, SDMs can be used for strengthening available data, as they provide useful tools for quality control. Finally, SDMs can play a role in prioritizing areas for field survey, by identifying knowledge gaps. While current ongoing initiatives have already implemented mechanisms and stressed the need for open, unrestricted access to data, information and knowledge related to the conservation of biodiversity (e.g., GBIF), none of them have yet considered the applied value of current modeling techniques to improve the applied value of such datasets. We strongly believe that SDMs have great potential to support the conservation of tropical biodiversity in the future, if their value in this context is recognized.

Acknowledgements

We are indebted to Bradford A. Hawkins, Joaquín Hortal, Jorge M. Lobo and three anonymous reviewers for their insightful comments on the manuscript. This study was financed by the Netherlands Environmental Assessment Agency (PBL) in the context of the Biodiversity International project (E/555050). LC was supported by the European project REFORLAN (INCO-CT-2006-032132) and the Andalusian Regional Government project GESBOME (P06-RNM-1890). MK was supported by CONABIO. EJMMA was supported by the Dutch Ministry of Agriculture, Nature and Food Quality (Project BO-10-003-01). This research was initiated in a working group at Managua, Nicaragua.

References

- [1] Guisan, A., and Zimmermann, N. E. 2000. Predictive habitat distribution modeling in ecology. *Ecological Modelling* 135: 147-186.
- [2] Guisan, A., and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8: 993-1009.
- [3] Austin, M. P. 1998. An ecological perspective on biodiversity investigations: examples from Australian eucalypt forest. *Annals of the Missouri Botanical Garden* 85: 2-17.
- [4] Vetaas, O. R. 2002. Realized and potential climate niches: a comparison of four *Rhododendron* tree species. *Journal of Biogeography* 29: 545-554.
- [5] Leathwick, J. R. 1998. Are New-Zealand's *Nothofagus* species in equilibrium with their environment? *Journal of Vegetation Science* 9: 719-732.
- [6] Anderson, R. P., Peterson, A. T., and Gómez-Laverde, M. 2002. Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos* 98: 3-16.
- [7] Graham, C. H., Ron, S. R., Santos, J. C., Schneider, C. J., and Moritz, C. 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in Dendrobatid frogs. *Evolution* 58: 1781-1793.
- [8] Beerling, D. J., Huntley, B., and Bailey, J. P. 1995. Climate and the distribution of *Fallopia japonica*: Use of an introduced species to test the predictive capacity of response surfaces. *Journal of Vegetation Science* 6: 269-282.

- [9] Peterson, A. T. 2003. Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology* 78: 419-433.
- [10] Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., Erasmus, B. F., De Siqueira, M. F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., Van Jaarsveld, A. S., Midgley, G. F., Miles, L., Ortega-Huerta, M. A., Peterson, A. T., Phillips, O. L., and Williams, S. E. 2004. Extinction risk from climate change. *Nature* 427: 145-148.
- [11] Thuiller, W., Brotons, L., Araújo, M. B., and Lavorel, S. 2004. Effects of restricting environmental range of data to project current and future species distributions. *Ecography* 27: 165-172.
- [12] Golicher, J. D., Cayuela, L., Alkemade, R. M., González-Espinosa, M., and Ramírez-Marcial, N. 2008. Applying climatically associated species pools to the modeling of compositional change in tropical montane forests. *Global Ecology and Biogeography* 17: 262-273.
- [13] Elith, J., and Burgman, M. A. 2002. Predictions and their validation: rare plants in the Central Highlands, Victoria, Australia. In: *Predicting Species Occurrences: Issues of Accuracy and Scale*. Scott, J. M., Heglund, P. J., Morrison, M. L., Raphael, M. G., Wall, W. A., and Samson, F. B. (Eds.), pp. 303-314. Island Press, Covelo, CA.
- [14] Raxworthy, C. J., Martinez-Meyer, E., Horning, N., Nussbaum, R. A., Schneider, G. E., Ortega-Huerta, M. A., and Peterson, A. T. 2003. Predicting distributions of known and unknown reptile species in Madagascar. *Nature* 426: 837-841.
- [15] Engler, R., Guisan, A., and Rechsteiner, L. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology* 41: 263-274.
- [16] Ferrier, S. 2002. Mapping spatial pattern in biodiversity for regional conservation planning: Where to from here? *Systematic Biology* 51: 331-363.
- [17] Araújo, M. B., Densham, P. J., and Williams, P. H. 2004. Representing species in reserves from patterns of assemblage diversity. *Journal of Biogeography* 31: 1037-1050.
- [18] Dormann, C. F. 2007. Promising the future? Global change projections of species distributions. *Basic and Applied Ecology* 8: 387-397.
- [19] Brown, J. H., and Lomolino, M. V. 1998. *Biogeography*, 2nd edn. Sinauer Press, Sunderland, Massachusetts.
- [20] Whittaker, R. J., Araújo, M. B., Jepson, P., Ladle, R. J., Watson, J. E. M., and Willis, K. J. 2005. Conservation Biogeography: assessment and prospect. *Diversity and Distributions* 11: 3-23.
- [21] Lomolino, M. V. 2004. Conservation biogeography. In: *Frontiers of Biogeography: new directions in the geography of nature*. Lomolino, M.V., and Heaney, L.R. (Eds.), pp. 293-296. Sinauer Associates, Sunderland, Massachusetts.
- [22] Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. B., and Kent, J. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853-858.
- [23] Laurance, W. F. 1999. Reflections on the tropical deforestation crisis. *Biological Conservation* 91: 109-117.
- [24] Brooks, T. M., Mittermeier, R. A., Mittermeier, C., Da Fonseca, G. A. B., Rylands, A. B., Konstant, W. R., Flick, P., Pilgrim, J., Oldfield, S., Magin, G., and Hilton-Taylor, C. 2002. Habitat loss and extinction in the hotspots of biodiversity. *Conservation Biology* 16: 909-923.
- [25] FAO. 2005. *Global Forest Resources Assessment 2005*. Food and Agriculture Organization of the United Nations, Roma, Italy. URL: <http://www.fao.org/forestry/fra2005>.
- [26] Sodhi, N. S., Koh, L. P., Brook, B. W., and Ng, P. K. L. 2004. Southeast Asia biodiversity: An impending disaster. *Trends in Ecology and Evolution* 19: 654-659.
- [27] Wright, S. J., and Muller-Landau, H. C. 2006. The future of tropical forest species. *Biotropica* 38: 287-301.

- [28] Feria, T. P. A., and Peterson, A. T. 2002. Prediction of bird community composition based on point-occurrence data and inferential algorithms: a valuable tool in biodiversity assessments. *Diversity and Distributions* 8: 49-56.
- [29] Drake, J. M., Randin, C., and Guisan, A. 2006. Modeling ecological niches with support vector machines. *Journal of Applied Ecology* 43: 424-432.
- [30] Elith, J., and Leathwick, J. 2007. Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. *Diversity and Distributions* 13: 265-275.
- [31] Manel, S., Dias, J. M., Buckton, S. T., and Ormerod, S. J. 1999. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *Journal of Applied Ecology* 36: 734-747.
- [32] Guisan A., Zimmermann, N. E., Elith, J., Graham, C. H., Phillips, S., and Peterson, A. T. 2007. What matters for predicting the occurrences of trees: Techniques, data, or species' characteristics? *Ecological Monographs* 77: 615-630.
- [33] Tsoar, A., Allouche, O., Steinitz, O., Rotem, D., and Kadmon, R. 2007. A comparative evaluation of presence-only methods for modeling species distribution. *Diversity and Distributions* 13: 397-405.
- [34] Collingham, Y. C., Wadsworth, R. A., Huntley, B., and Hulme, P. E. 2000. Predicting the spatial distribution of non-indigenous riparian weeds: issues of spatial scale and extent. *Journal of Applied Ecology* 37: 13-27.
- [35] Araújo, M. B., Thuiller, W., Williams, P. H., and Reginster, I. 2005. Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Global Ecology and Biogeography* 14: 17-30.
- [36] Allouche, O., Tsoar, A., and Kadmon, R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* 43: 1223-1232.
- [37] Bulluck, L., Fleishman, E., Betrus, C., and Blair, R. 2006. Spatial and temporal variations in species occurrence rate affect the accuracy of occurrence models. *Global Ecology and Biogeography* 15: 27-38.
- [38] Seoane, J., Bustamante, J., and Díaz-Delgado, R. 2005. Effect of expert opinion on the predictive ability of environmental models of bird distribution. *Conservation Biology* 19: 512-522.
- [40] Brotons, L., Wolff, A., Paulus, G., and Martin, J. -L. 2005. Effects of adjacent agricultural habitat on the distribution of passerines in natural grasslands. *Biological Conservation* 124: 407-414.
- [41] Schulte, L. A., Pidgeon, A. M., and Mladenoff, D. J. 2005. One hundred fifty years of change in forest bird breeding habitat: Estimates of species distributions. *Conservation Biology* 19: 1944-1956.
- [42] Boshoff, A. F., Kerley, G. I. H., and Cowling, R. M. 2001. A pragmatic approach to estimating the distributions and spatial requirements of the medium- to large-sized mammals in the Cape Floristic Region, South Africa. *Diversity and Distributions* 7: 29-43.
- [43] Chatfield, C. 1995. Model uncertainty, data mining and statistical inference. *Journal of the Royal Statistical Society, Series A* 158: 419-466.
- [44] Fielding, A.H., and Bell, J.F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24(1): 38-49.
- [45] Young, B. E. 2007. *Endemic species distribution on the east slope of the Andes in Peru and Bolivia*. NatureServe, Arlington, Virginia, USA. Available on-line in: http://www.natureserve.org/aboutUs/latinamerica/pubs/Endemic_Species_2007.pdf.
- [46] Hernández, P. A., Graham, C. H., Master, L. L., and Albert, D. L. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29: 773-785.
- [47] 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., Lohamnn, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. Overton, J., Peterson, A. T., Phillips, S. J., Richardson, K. S., Scachetti-Pereira,

- R., Schapire, R. E., Soberón, J., Williams, S., Wisz, M. S., and Zimmermann, N. E. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129-151.
- [48] Phillips, S. J., Anderson, R. P., and Schapire, R. E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231-259.
- [49] CONABIO (Comisión Nacional para el Conocimiento y Uso de la Biodiversidad). 2007. *Análisis de vacíos y omisiones en conservación de la biodiversidad terrestre de México: espacios y especies*. CONABIO, Comisión Nacional de Áreas Naturales Protegidas, The Nature-Conservancy-Programa México, Pronatura, A.C., Facultad de Ciencias Forestales, Universidad Autónoma de Nueva León, México.
- [50] Stockwell, D. R. B. 1999. The GARP modeling system: problems and solutions to automated spatial prediction. *International Journal of Geographic Information Systems* 13: 143-158.
- [51] Lim, B. K., Peterson, A. T., and Engstrom, M. D. 2002. Robustness of ecological niche modeling algorithms for mammals in Guyana. *Biodiversity and Conservation* 11: 1237-1246.
- [52] Brooks T., Hannah, L., da Fonseca, G. A. B., and Mittermeier, R. A. 2001. Prioritizing hotspots, representing transitions. *Trends in Ecology and Evolution* 16: 673.
- [53] Olson, D. M., and Dinerstein, E. 1998. The Global 200: a representation approach to conserving the Earth's most biologically valuable ecoregions. *Conservation Biology* 12: 502-515.
- [54] Funk, V. A., Richardson, K. R., and Ferrier, S. 2005. Survey-gap analysis in expeditionary research: where do we go from here? *Biological Journal of the Linnean Society* 85: 549-567.
- [55] Kadmon, R., Farber, O., and Danin, A. 2003. A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications* 13: 853-867.
- [56] Stockwell, D. R. B., and Peterson, A. T. 2002b. Effects of sample size on accuracy of species distribution models. *Ecological Modelling* 148: 1-13.
- [57] Jiménez-Valverde, A., and Lobo, J. M. 2007. Threshold criteria for conversion of probability of species presence to either-or presence-absence. *Acta Oecologica* 31: 361-369.
- [58] Oldfield, S., Lusty, C., and MacKinven, A. 1998. *The World List of Threatened Trees*. World Conservation Press, Cambridge, UK.
- [59] Newton, A. C., and Oldfield, S. 2008. Red Listing the world's tree species: a review of recent progress. *Endangered Species Research* 6: 137-147.
- [60] Rabinowitz, D. 1981. Seven forms of rarity. In: *The biological aspects of rare plant conservation*. Synge, H. (Ed.), pp. 205-217. John Wiley, Chichester, UK.
- [61] Rey Benayas, J. M., Scheiner, S. M., Sánchez-Colomer, M. G. and Levassor, C. 1999. Commonness and rarity: theory and application of a new model to Mediterranean montane grasslands. *Conservation Ecology* 3: art.5. UR: <http://www.consecol.org/vol3/iss1/art5/>.
- [62] Funk, V. A., and Richardson, K. R. 2002. Systematic data in biodiversity studies: use it or lose it. *Systematic Biology* 51: 303-316.
- [63] Keddy, P. A. 2001. *Competition*, 2nd edition. Kluwer Academic Publishers, The Netherlands.
- [64] Llorente, J., Onate, L., Luis, A., and Vargas, I. 1997. *Papilionidae y Pieridae de México: Distribución Geográfica e Ilustración*. UNAM, México.
- [65] Soberón, J. M., Llorente, J. B., and Oñate, L. 2000. The use of specimen-label databases for conservation purposes: an example using Mexican Papilionid and Pierid butterflies. *Biodiversity and Conservation* 9: 1441-1466.
- [66] Guisan, A., Graham, C. H., Elith, J., Huettmann, F. and NCEAS Species Distribution Modeling Group. 2007. Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions* 13: 332-340.
- [67] Williams, P. H., Margules, C. R., and Hilbert, D. W. 2002. Data requirements and data sources for biodiversity priority area selection. *Journal of Bioscience* 27: 327-338.
- [68] Reddy, S., and Dávalos, L. M. 2003. Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography* 30: 1719-1727.

- [69] Nelson, B. W., Ferreira, C. A. C., da Silva, M. F., and Kawasaki, M. L. 1990. Endemism centres, refugia and botanical collection density in Brazilian Amazonia. *Nature* 345: 714-716.
- [70] Freitag, S., Hobson, C., Biggs, H. C., and van Jaarsveld, A. S. 1998. Testing for potential survey bias: the effect of roads, urban areas and nature reserves on a southern African mammal data set. *Animal Conservation* 1: 119-127.
- [71] Peterson, A. T., Navarro-Sigüenza, A. G., and Benítez-Díaz, H. 1998. The need for continued scientific collecting: a geographic analysis of Mexican bird specimens. *Ibis* 140: 288-294.
- [72] Parnell, J. A. N., Simpson, D. A., Moat, J., Kirkup, D. W., Chantaranonthai, P., Boyce, P.C., Bygrave, P., Dransfield, S., Jebb, M. H. P., Macklin, J., Meade, C., Middleton, D. J., Muasya, A. M., Prajaksood, A., Pendry, C. A., Pooma, R., Suddee, S., and Wilkin, P. 2003. Plant collecting spread and densities: their potential impact on biogeographical studies in Thailand. *Journal of Biogeography* 30: 193-209.
- [73] Kadmon, R., Farber, O., and Danin, A. 2004. Effect of roadside on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications* 14: 401-413.
- [74] Ponder, W. F., Carter, G. A., Flemons, P., and Chapman, R. R. 2001. Evaluation of museum collection data for use in biodiversity assessment. *Conservation Biology* 15: 648-657.
- [75] Soberón, J. M., Llorente, J. B., and Benítez, H. 1996. An international view of national biological surveys. *Annals of the Missouri Botanical Gardens* 83: 562-573.
- [76] Funk, V. A., Zermoglio, M. F., and Nasir, N. 1999. Testing the use of specimen collection data and GIS in biodiversity exploration and conservation decision making in Guyana. *Biodiversity and Conservation* 8: 727-751.
- [77] Podger, W. F., Mummery, D. C., Palzer, C. R., and Brown, M. J. 1990. Bioclimatic analysis of the distribution of damage to native plants in Tasmania by *Phytophthora cinnamomi*. *Australian Journal of Ecology* 15: 281-290.
- [78] Hanowski, J. A. M., and Niemi, G. J. 1995. A comparison of on- and off-road bird counts: do you need to go off road to count birds accurately? *Journal of Field Ornithology* 66: 469-483.
- [79] Keller, C. M. E., and Scallan, J. T. 1999. Potential roadside biases due to habitat changes along breeding bird survey routes. *Condor* 101: 50-57.
- [80] Manrique, C. E., Durán, R., and Argáez, J. 2003. Phytogeographic analysis of taxa endemic to the Yucatán Peninsula using geographic information systems, the domain heuristic and parsimony analysis of endemism. *Diversity and Distributions* 9: 313-330.
- [81] Bojórquez-Tapia, L. A., Balvanera, P., and Cuarón, A. D. 1994. Biological inventories and computer data bases: their role in environmental assessments. *Environmental Management* 18: 775-785.
- [82] Hijmans, R. J., Garrett, K. A., Huaman, Z., Zhang, D. P., Schreuder, M., and Bonierbale, M. 2000. Assessing the geographic representativeness of gene bank collections: the case of Bolivian wild potatoes. *Conservation Biology* 14: 1755-1765.
- [83] Hortal, J., Lobo, J. M., and Jiménez-Valverde, A. 2007. Limitations of biodiversity databases: case study on seed-plant diversity in Tenerife (Canary Islands). *Conservation Biology* 21: 853-863.
- [84] Hortal, J., Jiménez-Valverde, A., Gómez, J. F., Lobo, J. M., and Baselga, A. 2008. Historical bias in biodiversity inventories affects the observed realized niche of the species. *Oikos* 117: 847-858.
- [85] Lobo, J. M., Baselga, A., Hortal, J., Jiménez-Valverde, A., and Gómez, J.F. 2007. How does the knowledge about the spatial distribution of Iberian dung beetle species accumulate over time? *Diversity and Distributions* 13: 772-780.
- [86] Stockwell, D. R. B., and Peterson, A. T. 2002. Controlling bias during predictive modeling with museum data. In: *Predicting species occurrences: issues of scale and accuracy*. Scott, J. M., Heglund, P. J., Morrison, M., Raphael, M., Haufler, J., Wall, B., and Samson, F. (Eds.), pp. 537-546. Island Press, Covelo, California, USA.
- [87] Chefaoui, R.M., and Lobo, J.M. 2008. Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling* 210: 478-486.

- [88] Lobo, J. M., Jiménez-Valverde, A., and Real, R. 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* 17:145-151.
- [89] Thuiller, W. 2004. Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology* 10: 2020-2027.
- [90] Pereira, H. M., and Cooper, H. D. 2006. Towards the global monitoring of biodiversity change. *Trends in Ecology and Evolution* 21: 123-129.
- [91] Preston, C. D., Pearman, D. A., and Dines, T. D. 2002. *New Atlas of the British and Irish Flora: An Atlas of the Vascular Plants of Britain, Ireland, the Isle of Man and the Channel Islands*. Oxford University Press, 912 pp.
- [92] Peacock, J., Baker, T. R., Lewis, S. L., López-González, G., and Phillips, O. L. 2007. The RAINFOR database: Monitoring forest biomass and dynamics. *Journal of Vegetation Science* 18: 535-542.
- [93] Ter Steege, H., Pitman, N., Phillips, O. L., Chave, J., Sabatier, D., Duque, A., Molino, J. F., Prevoist, M. F., Spichiger, R., Castellanos, H., van Hildebrand, P., and Vasquez, R. 2006. Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature* 443: 444-447.
- [94] Yesson C., Brewer, P. W., Sutton, T., Caithness, N., Pahwa, J. S., Burgess, M., Gray, W. A., White, R. J., Jones, A. C., Bisby, F. A., and Culham, A. 2007. How Global Is the Global Biodiversity Information Facility? *PLoS ONE* 2: e1124.
- [95] Chapman, A. D. 2005. *Principles and methods of data cleaning – Primary species and species-occurrence data, version 1.0*. Report for the Global Biodiversity Information Facility, Copenhagen, Denmark.
- [96] Guralnick, R. P., Hill, A. W., and Meredith, L. 2007. Towards a collaborative, global infrastructure for biodiversity assessment. *Ecology Letters* 10: 663-672.
- [97] Jiménez-Valverde, A., Lobo, J.M., and Hortal, J. 2008. Not as good as they seem: the importance of concepts in species distribution modeling. *Diversity and Distributions* 14: 885-890.
- [98] Ferrier, S., and Guisan, A. 2006. Spatial modeling of biodiversity at the community level. *Journal of Applied Ecology* 43: 393-404.
- [99] Cawsey, E. M., Austin, M. P., and Baker, B. L. 2002. Regional vegetation mapping in Australia: a case study in the practical use of statistical modeling. *Biodiversity and Conservation* 11: 2239-2274.
- [100] Miller, J., and Franklin, J. 2002. Modeling the distribution of four vegetation alliances using generalized linear models and classification trees with spatial dependence. *Ecological Modelling* 157: 227-247.
- [101] Thuiller, W. 2003. BIOMOD: Optimising predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology* 9: 1353-1362.
- [102] Ochoa-Ochoa, L. M., and Flores-Villela, O. 2006. *Áreas de Diversidad y Endemismo de la Herpetofauna Mexicana*. UNAM-CONABIO, Mexico, D.F.

Appendix 1. Description of modeling methods and examples found in the literature review (references are provided in Appendix 3).

Method

Generalized linear models (GLM)

Statistical models including ordinary regression and analysis of variance. All GLMs have a random component (response variable), a systematic component (explanatory variables), and a link function that describes the relationship between the expected value of the response and the predictors. The most common form of GLM in predictive distribution modeling is logistic regression for presence-absence data. Many examples are found in the literature (e.g. Osborne *et al.* 2001; Bailey *et al.* 2002; Lennon *et al.* 2002; Anadón *et al.* 2007).

Generalized additive models (GAM)

GAMs are a non parametric extension of GLMs, in which the linear or polynomial functions are replaced by smoothed data-dependent functions. Examples included Araújo *et al.* (2005b), Heikkinen *et al.* (2006), García *et al.* (2007) and Luoto *et al.* (2005, 2007).

Other regression techniques (RT+)

Alternative regression-based models including multivariate adaptive regression splines (MARS, Elith and Leathwick 2007; Guisan *et al.* 2007), BRUTO (Guisan *et al.* 2007), generalized dissimilarity model of single species (GDMSS, Guisan *et al.* 2007), conditional fixed effects logistic regression (Johnson *et al.* 2004), autologistic regression (McPherson and Jetz 2007), generalized linear mixed models (GLMM, Muñoz and Real 2006; Rhodes *et al.* 2006), geographically weighted regression (GWR, Osborne *et al.* 2007), varying coefficient modeling (VCM, Osborne *et al.* 2007), and Bayesian model averaging with logistic regression (Thomson *et al.* 2007).

Bioclimatic envelopes (BIOCLIM)

Include a variety of methods that use presence-only records to characterize sites that are located within the environmental hyper-space occupied by a species. BIOCLIM is a good example of a climate-mapping approach to modeling (Peterson *et al.* 2000; Loiselle *et al.* 2003; Téllez-Valdés and Dávila-Aranda 2003; Robertson *et al.* 2004; Tsoar *et al.* 2007), but other envelope-style methods are also in use (Yom-Tov and Kadmon 1998; Robertson *et al.* 2004; Wharton and Kriticos 2004).

Habitat suitability indices (HS)

Methods requiring presence-only data based on a comparison between the environmental niche of the species and the environmental characteristics of the entire study area. They include a variety of techniques including ecological niche factor analysis (ENFA, Engler *et al.* 2004; Martínez *et al.* 2006; Stockman *et al.* 2006; Guisan *et al.* 2007; Sérgio *et al.* 2007; Tsoar *et al.* 2007), DOMAIN (Loiselle *et al.* 2003; Manrique *et al.* 2003; Vargas *et al.* 2004; Guisan *et al.* 2007; Tsoar *et al.* 2007), the Mahalanobis distance (Rouget *et al.* 2004; Allouche *et al.* 2006; Tsoar *et al.* 2007), and other related methods (Beard *et al.* 1999; Polasky *et al.* 2000; Robertson *et al.* 2001; Miles *et al.* 2004; Estrada-Peña *et al.* 2005; Gavin and Hu 2005; Beck and Kitching 2007; Tsoar *et al.* 2007).

Multivariate statistics (MS)

Includes a variety of multivariate statistical techniques applied to species distribution modeling, including canonical correspondence analysis (CCA, Gottfried *et al.* 1999; Jaberg and Guisan 2001), correspondence analysis (Wang *et al.* 2003), and discriminant analysis (Manel *et al.* 1999; Curnutt 2000; Tourenq *et al.* 2001; McPherson *et al.* 2004)

Machine learning methods (MLM)

Include classification tree analysis (CART, Thuiller *et al.* 2003; Foxcroft *et al.* 2004; Gavin and Hu 2005; Segurado *et al.* 2006; Bahn and McGill 2007; Usio 2007), artificial neural networks (Manel *et al.* 1999; Tourenq *et al.* 2001; Berry *et al.* 2002; Worner and Gevrey 2006), support vector machines (Drake *et al.* 2006), maximum entropy models (MAXENT, Graham and Hijmans 2006; Guisan *et al.* 2007; Sérgio *et al.* 2007; Pawar *et al.* 2007), and boosted regression trees (BRT, Guisan *et al.* 2007).

Genetic algorithms (GARP)

Genetic algorithms techniques that search iteratively for non-random correlations between species presence and absence and environmental parameter values using several different types of rules. Each rule type implements a different method for building species prediction models. The most widespread method in predictive species distribution modeling is the Genetic Algorithm for Rule-set Prediction (GARP). GARP implements four types of rules: atomic, logistic regression, bioclimatic envelope, and negated bioclimatic envelope rules. Some examples can be found in Anderson *et al.* (2002), Ortega-Huerta and Peterson (2004), Peterson *et al.* (2004), and García (2006). An alternative genetic algorithm technique is the Bayes-based genetic algorithm implemented by McClean *et al.* (2006).

Consensus techniques (BIOMOD)

This refers basically to the BIOMOD application developed by Thuiller [101], which incorporates four familiar modeling techniques to model species distributions: generalized linear models, generalized additive models, classification tree analysis, and artificial neural networks.

Expert knowledge-based methods (EK)

Comprise a variety of procedures based on qualitative information, subjective estimates, and expert judgement (Boshoff *et al.* 2001; Johnson and Gillingham 2004; Schulte *et al.* 2005).

Appendix 2. Methodologies and approaches used in modeling species distribution in 123 articles published from January, 1995 to May, 2007, in *Biological Conservation*, *Conservation Biology*, *Diversity and Distribution*, *Global Ecology and Biogeography* and the *Journal of Applied Ecology*. Full references are provided in Appendix 3.

Reference	Region	Scope	Verification	Taxon	Data ¹
Allouche <i>et al.</i> (2006)	Israel	Methodological	Informal validation	Plants	P
Anadón <i>et al.</i> (2007)	Spain	Species conservation	Same data	Reptiles	P-A
Anderson <i>et al.</i> (2002)	Ecuador, Colombia, Venezuela	Biogeography	Partitioning	Mammals	P
Araújo <i>et al.</i> (2005a)	UK	Climate change	Partitioning	Birds	P
Araújo <i>et al.</i> (2005b)	Europe	Methodological	Partitioning + Informal validation	Amphibians, Birds, Mammals, Plants, Reptiles	P
Arriaga <i>et al.</i> (2004)	Mexico	Biological invasions	Partitioning	Plants	P
Bahn & McGill (2007)	USA, Canada	Methodological	Partitioning	Birds	Ab
Bailey <i>et al.</i> (2002)	UK	Species conservation	Informal validation	Birds, Invertebrates, Mammals, Plants	P-A
Beard <i>et al.</i> (1999)	USA	Methodological	Partitioning	Birds	P-A
Beck & Kitching (2007)	Malesia	Methodological	Non reported	Invertebrates	P
Berry <i>et al.</i> (2002)	UK, Ireland	Climate change	Partitioning	Amphibians, Invertebrates, Mammals, Plants	P-A
Bonaccorso <i>et al.</i> (2006)	Brazil	Biogeography	Partitioning	Birds, Plants	P
Boshoff <i>et al.</i> (2001)	South Africa	Methodological	Informal validation	Mammals	P-A
Brotons <i>et al.</i> (2007)	Spain	Methodological	Informal validation	Birds	P, Ab
Brotons <i>et al.</i> (2005)	France	Species conservation	Non reported	Birds	P-A, Ab
Bulluck <i>et al.</i> (2006)	USA	Methodological	Prospective sampling	Birds	P
Buse <i>et al.</i> (2007)	Germany	Ecology	Prospective sampling	Invertebrates	P-A

Reference	Region	Scope	Verification	Taxon	Data ¹
Carroll <i>et al.</i> (2003)	USA	Species conservation	Non reported	Mammals	P
Cassinello <i>et al.</i> (2006)	Spain	Biological invasions	Prospective sampling	Mammals	P
Castellarini <i>et al.</i> (2007)	France	Conservation prioritization	Partitioning	Invertebrates	P-A
Chefaoui <i>et al.</i> (2005)	Spain	Species conservation	Non reported	Invertebrates	P
Chu <i>et al.</i> (2005)	Canada	Climate change	Partitioning	Fishes	P-A
Collingham <i>et al.</i> (2000)	UK	Methodological	Partitioning	Plants	P
Cooper (2002)	USA	Ecology	Non reported	Birds	P-A
Cowley <i>et al.</i> (2000)	UK	Species conservation	Partitioning	Invertebrates	P-A
Curnutt (2000)	USA, Australia	Biological invasions	Partitioning	Plants	P
Dennis <i>et al.</i> (2002)	France	Methodological	Non reported	Invertebrates	P-A
Domínguez-Domínguez <i>et al.</i> (2006)	Mexico	Species conservation	Partitioning	Fishes	P
Drake <i>et al.</i> (2006)	Switzerland	Methodological	Partitioning	Plants	P
Elith & Leathwick (2007)	Australia, Canada, New Zealand, Switzerland, South America	Methodological	Prospective sampling	Birds, Mammals, Plants, Reptiles	P
Engler <i>et al.</i> (2004)	Switzerland	Methodological	Partitioning + Prospective sampling	Plants	P
Estrada-Peña <i>et al.</i> (2005)	The Americas	Climate change	Partitioning	Invertebrates	P
Feria & Peterson (2002)	Mexico	Methodological	Partitioning	Birds	P
Ficetola & De Bernardi (2004)	Italy	Species conservation	Partitioning	Amphibians	P-A
Fitzpatrick <i>et al.</i> (2007)	South America, USA	Biological invasions	Partitioning	Invertebrates	P

Reference	Region	Scope	Verification	Taxon	Data ¹
Fleishman <i>et al.</i> (2003)	USA	Methodological	Prospective sampling	Invertebrates	P-A
Foxcroft <i>et al.</i> (2004)	South Africa	Biological invasions	Same data	Plants	P-A
García (2006)	Mexico	Biodiversity mapping	Non reported	Amphibians, Reptiles	P
García <i>et al.</i> (2007)	Spain	Species conservation	Partitioning	Birds	P-A, Ab
Gavin & Hu (2005)	USA, Canada	Methodological	Partitioning	Plants	P-A
Germaine & Wakeling (2001)	USA	Ecology	Non reported	Reptiles	P-A
Gibson <i>et al.</i> (2004)	Australia	Species conservation	Same data	Birds	P-A
Gottfried <i>et al.</i> (1999)	Austria	Climate change	Non reported	Plants	P-A
Graham & Hijmans (2006)	USA	Methodological	Partitioning	Amphibians, Reptiles	P
Greaves <i>et al.</i> (2006)	UK	Methodological	Partitioning	Mammals	P-A
Grundel & Pavlovic (2007)	USA	Species conservation	Non reported	Invertebrates	P-A
Guisan <i>et al.</i> (2005)	Switzerland	Methodological	Prospective sampling	Plants	P
Guisan <i>et al.</i> (2007)	Australia, Canada, New Zealand, Switzerland and South America	Methodological	Prospective sampling	Birds, Mammals, Plants, Reptiles	P
Heikkinen <i>et al.</i> (2006)	Finland	Methodological	Same data	Birds	P-A
Jaberg & Guisan (2001)	Switzerland	Species conservation	Partitioning	Mammals	P-A
Jarvis & Robertson (1999)	Namibia	Species conservation	Non reported	Birds	P-A, Ab
Jeganathan <i>et al.</i> (2004)	India	Species conservation	Same data	Birds	P-A
Johnson & Guillingham (2004)	Canada	Methodological	Non reported	Birds, Mammals	?
Johnson <i>et al.</i> (2004)	Canada	Methodological	Partitioning	Mammals	P

Reference	Region	Scope	Verification	Taxon	Data ¹
Lennon <i>et al.</i> (2002)	USA	Biogeography	Non reported	Plants	P-A
Lira-Noriega (2007)	Mexico	Methodological	Partitioning	Birds	P
Lobo <i>et al.</i> (2006)	Iberian Peninsula	Ecology	Same data	Invertebrates	P
Loiselle <i>et al.</i> (2003)	Brazil	Methodological	Informal validation	Birds	P
Luoto <i>et al.</i> (2005)	Finland	Methodological	Partitioning	Invertebrates	P
Luoto <i>et al.</i> (2007)	Finland	Methodological	Partitioning	Birds	P-A
Lütolf <i>et al.</i> (2006)	Switzerland	Methodological	Prospective sampling	Invertebrates	P
Manel <i>et al.</i> (1999)	India	Methodological	Prospective sampling	Birds	P-A
Manel <i>et al.</i> (2001)	India	Methodological	Prospective sampling	Invertebrates	P-A
Manrique <i>et al.</i> (2003)	Mexico, Belice, Guatemala	Biogeography	Informal validation	Plants	P
Martínez <i>et al.</i> (2006)	Spain	Conservation prioritization	Partitioning	Lichens	P
Martínez-Meyer <i>et al.</i> (2004)	USA	Biogeography	Same data	Mammals	P
McClellan <i>et al.</i> (2006)	Africa	Climate change	Same data	Plants	P
McPherson & Jetz (2007)	Southern-Eastern Africa	Methodological	Non reported	Birds	P-A
McPherson <i>et al.</i> (2004)	South Africa, Lesotho and Swaziland	Methodological	Partitioning	Birds	P
Midgley <i>et al.</i> (2006)	South Africa	Methodological	Partitioning	Plants	P-A
Miles <i>et al.</i> (2004)	Brazil, Bolivia, Colombia, Ecuador, Peru	Climate change	Non reported	Plants	P-A, P
Milson <i>et al.</i> (2000)	UK	Species conservation	Non reported	Birds	P-A
Muñoz & Real (2006)	Spain	Biological invasions	Same data	Birds	P-A
Muñoz <i>et al.</i> (2005)	Spain	Species conservation	Non reported	Birds	P-A

Reference	Region	Scope	Verification	Taxon	Data ¹
Ortega-Huerta (2007)	Mexico	Conservation prioritization	Non reported	Birds, Mammals	P
Ortega-Huerta & Peterson (2004)	Mexico	Conservation prioritization	Same data + Partitioning	Birds	P
Osborne <i>et al.</i> (2001)	Spain	Species conservation	Partitioning	Birds	P
Osborne <i>et al.</i> (2007)	Spain, UK	Methodological	Same data	Birds	P-A
Palma <i>et al.</i> (1999)	Portugal	Species conservation	Same data	Mammals	P
Pawar <i>et al.</i> (2007)	India	Conservation prioritization	Same data	Amphibians, Reptiles	P
Pearce <i>et al.</i> (2001)	Australia	Methodological	Same data	Birds, Reptiles, Mammals	P-A
Pearce & Ferrier (2001)	Australia	Methodological	Prospective sampling	Birds, Mammals, Plants, Reptiles	P-A, Ab
Peterson <i>et al.</i> (2000)	Mexico	Conservation prioritization	Partitioning	Birds, Mammals	P
Peterson <i>et al.</i> (2004)	USA, Mexico	Methodological	Non reported	Birds	P
Polasky <i>et al.</i> (2000)	USA	Methodological	Non reported	Amphibians, Birds, Mammals, Reptiles	P-A
Reino <i>et al.</i> (2006)	Portugal	Ecology	Same data	Birds	P-A
Rhodes <i>et al.</i> (2006)	Australia	Species conservation	Same data	Mammals	P-A
Robertson <i>et al.</i> (2001)	South Africa, Lesotho, Swaziland	Methodological	Partitioning	Plants	P
Robertson <i>et al.</i> (2004)	South Africa, Lesotho, Swaziland	Methodological	Partitioning	Plants	P
Robinson <i>et al.</i> (2001)	UK	Species conservation	Non reported	Birds	P-A
Roger <i>et al.</i> (2007)	Australia	Species conservation	Same data	Mammals	P-A

Reference	Region	Scope	Verification	Taxon	Data ¹
Rouget <i>et al.</i> (2004)	South Africa, Lesotho, Swaziland	Biological invasions	Same data	Plants	P
Russell & Clout (2004)	New Zealand	Biological invasions	Non reported	Mammals	P-A
Sánchez-Cordero <i>et al.</i> (2005)	Mexico	Species conservation	Partitioning	Mammals	P
Schmidt <i>et al.</i> (2005)	Burkina Faso	Biodiversity mapping	Non reported	Plants	P
Schulte <i>et al.</i> (2005)	USA	Species conservation	Informal validation	Birds	P
Schussman <i>et al.</i> (2006)	USA	Biological invasions	Partitioning	Plants	P
Segurado <i>et al.</i> (2006)	Iberian Peninsula	Methodological	Same data	Reptiles	P-A
Seoane <i>et al.</i> (2005)	Spain	Methodological	Same data + Prospective sampling	Birds	P-A
Sérgio <i>et al.</i> (2007)	Portugal	Species conservation	Informal validation	Bryophytes	P
Silva <i>et al.</i> (2002)	Portugal	Methodological	Partitioning	Birds	P-A
Stockman <i>et al.</i> (2006)	USA	Methodological	Prospective sampling	Invertebrates	P
Suárez-Seoane <i>et al.</i> (2002)	Spain	Methodological	Partitioning	Birds	P-A
Téllez-Valdés & Davila-Aranda (2003)	Mexico	Climate change	Non reported	Plants	P
Thomson <i>et al.</i> (2007)	Australia	Species conservation	Prospective sampling	Birds	P-A
Thuiller <i>et al.</i> (2003)	Spain	Ecology	Partitioning	Plants	P-A
Thuiller <i>et al.</i> (2005)	Europe	Climate change	Same data	Plants	P
Thuiller <i>et al.</i> (2006)	Europe	Climate change	Partitioning	Plants	P
Tourenq <i>et al.</i> (2001)	France	Biological invasions	Partitioning	Birds	P-A
Travaini <i>et al.</i> (2007)	Argentina	Methodological	Partitioning	Mammals	P-A

Reference	Region	Scope	Verification	Taxon	Data ¹
Tsoar <i>et al.</i> (2007)	Israel	Methodological	Prospective sampling	Birds, Invertebrates, Mammals	P
Underwood <i>et al.</i> (2004)	USA	Biological invasions	Partitioning	Plants	P-A
Usio (2007)	Japan	Species conservation	Same data	Invertebrates	P-A
Vanreusel <i>et al.</i> (2007)	Belgium	Species conservation	Partitioning + Prospective sampling	Invertebrates	P-A
Vargas <i>et al.</i> (2004)	Ecuador	Species conservation	Informal validation	Plants	P
Wang <i>et al.</i> (2003)	Mexico	Ecology	Non reported	Mammals	P-A
Weaver <i>et al.</i> (2006)	USA	Biogeography	Partitioning	Invertebrates	P
Wharton & Kriticos (2004)	Australia	Biological invasions	Non reported	Invertebrates	P-A
Wheatley <i>et al.</i> (2005)	Canada	Species conservation	Non reported	Mammals	P-A, Ab
White & Kerr (2007)	Canada	Biodiversity mapping	Non reported	Invertebrates	P
Wilson <i>et al.</i> (2005)	Australia	Methodological	Non reported	Plants	P-A
Worner & Gevrey (2006)	World	Biological invasions	Non reported	Invertebrates	P-A
Yom-Tov & Kadmon (1998)	Israel	Species conservation	Non reported	Mammals	P

¹ P, Presence-only data; P-A, Presence-absence data; Ab, Abundance data; ?, No information given.

Appendix 3. References listed in Appendix 2.

1. Allouche, O., A. Tsoar, and R. Kadmon. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* **43**:1223-1232.
2. Anadón, J.D., A. Giménez, M. Martínez, J. A. Palazón, and M. A. Esteve. 2007. Assessing changes in habitat quality due to land use changes in the spur-thighed tortoise *Testudo graeca* using hierarchical predictive habitat models. *Diversity and Distributions* **13**:324-331.
3. Anderson, R. P., M. Gómez-Laverde, A. T. Peterson. 2002. Geographical distributions of spiny pocket mice in South America: insights from predictive models. *Global Ecology and Biogeography* **11**:131-141.
4. Araújo, M. B., R. J. Whittaker, R.J. Ladle, and M. Erhard. 2005a. Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography* **14**:529-538.
5. Araújo, M. B., W. Thuiller, P. H. Williams, and I. Reginster. 2005b. Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Global Ecology and Biogeography* **14**:17-30.
6. Arriaga, L., A. E. Castellanos, E. Moreno, and J. Alarcón. 2004. Potential ecological distribution of alien invasive species and risk assessment: a case study of buffel grass in arid regions of Mexico. *Conservation Biology* **18**:1504-1514.
7. Bahn, V., and B. J. McGill. 2007. Can niche-based distribution models outperform spatial interpolation? *Global Ecology and Biogeography* **16**:733-742.
8. Bailey, S.-A., R. H. Haines-Young, and C. Watkins. 2002. Species presence in fragmented landscapes: modeling of species requirements at the national level. *Biological Conservation* **108**:307-316.
9. Beard, K. H., N. Hengartner, and D. K. Skelly. 1999. Effectiveness of predicting breeding bird distributions using probabilistic models. *Conservation Biology* **13**:1108-1116.
10. Beck, J., and I. J. Kitching. 2007. Estimating regional species richness of tropical insects from museum data: a comparison of geography-based and sample-based methods. *Journal of Applied Ecology* **44**:672-681.
11. Berry, P. M., T. P. Dawson, P. A. Harrison, and R. G. Pearson. 2002. Modeling potential impacts of climate change on the bioclimatic envelope of species in Britain and Ireland. *Global Ecology and Biogeography* **11**:453-462.
12. Bonaccorso, E., I. Koch, and A. T. Peterson. 2006. Pleistocene fragmentation of Amazon species' ranges. *Diversity and Distributions* **12**:157-164.
13. Boshoff, A. F., G. I. H. Kerley, and R. M. Cowling. 2001. A pragmatic approach to estimating the distributions and spatial requirements of the medium- to large-sized mammals in the Cape Floristic Region, South Africa. *Diversity and Distributions* **7**:29-43.
14. Brotons, L., S. Herrando, and M. Pla. 2007. Updating bird species distribution at large spatial scales: applications of habitat modeling to data from long-term monitoring programs. *Diversity and Distributions* **13**:276-288.
15. Brotons, L., A. Wolff, G. Paulus, and J.-L. Martin. 2005. Effects of adjacent agricultural habitat on the distribution of passerines in natural grasslands. *Biological Conservation* **124**:407-414.
16. Bulluck, L., E. Fleishman, C. Betrus, and R. Blair. 2006. Spatial and temporal variations in species occurrence rate affect the accuracy of occurrence models. *Global Ecology and Biogeography* **15**:27-38.
17. Buse, J., B. Schröder, and T. Assmann. 2007. Modeling habitat and spatial distribution of an endangered longhorn beetle – A case study for saproxylic insect conservation. *Biological Conservation* **137**:372-381.
18. Carroll, C., M. K. Phillips, N. H. Schumaker, and D. W. Smith. 2003. Impacts of landscape change on wolf restoration success: Planning a reintroduction program based on static and dynamic spatial models. *Conservation Biology* **17**:536-548.
19. Cassinello, J., P. Acevedo, and J. Hortal. 2006. Prospects for population expansion of the exotic aoudad (*Ammotragus lervia*; Bovidae) in the Iberian Peninsula: clues from habitat suitability modeling. *Diversity and Distributions* **12**:666-678.
20. Castellarini, F., F. Malard, M.-J. Dole-Olivier, and J. Gibert. 2007. Modeling the distribution of stygobionts in the Jura Mountains (eastern France). Implications for the protection of ground waters. *Diversity and Distributions* **13**:213-224.
21. Chefaoui, R. M., J. Hortal, and J. M. Lobo. 2005. Potential distribution modeling, niche characterization and conservation status assessment using GIS tools: a case study of Iberian *Copris* species. *Biological Conservation* **122**:327-338.
22. Chu, C., N. E. Mandrak, and C. K. Minns. 2005. Potential impacts of climate change on the distributions of several common and rare freshwater fishes in Canada. *Diversity and Distributions* **11**:299-310.
23. Collingham, Y. C., R. A. Wadsworth, B. Huntley, and P. E. Hulme. 2000. Predicting the spatial distribution of non-indigenous riparian weeds: issues of spatial scale and extent. *Journal of Applied Ecology* **37**:13-27.

24. Cooper, D. S. 2002. Geographic associations of breeding bird distribution in an urban open space. *Biological Conservation* **104**:205-210.
25. Cowley, M. J. R., R. J. Wilson, J. L. León-Cortes, D. Gutiérrez, C. R. Bulman, and C. D. Thomas. 2000. Habitat-based statistical models for predicting the spatial distribution of butterflies and day-flying moths in a fragmented landscape. *Journal of Applied Ecology* **37**:60-72.
26. Curnutt, J. L. 2000. Host-area specific climatic-matching: similarity breeds exotics. *Biological Conservation* **94**:341-351.
27. Dennis, R. L. H., T. G. Shreeve, T. H. Sparks, and J. E. Lhonore. 2002. A comparison of geographical and neighbourhood models for improving atlas databases. The case of the French butterfly atlas. *Biological Conservation* **108**:143-159.
28. Domínguez-Domínguez, O., E. Martínez-Meyer, L. Zambrano, and G. Pérez-Ponce de León. 2006. Using ecological-niche modeling as a conservation tool for freshwater species: live-bearing fishes in Central Mexico. *Conservation Biology* **20**:1730-1739.
29. Drake, J. M., C. Randin, and A. Guisan. 2006. Modeling ecological niches with support vector machines. *Journal of Applied Ecology* **43**:424-432.
30. Elith, J., and J. Leathwick. 2007. Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. *Diversity and Distributions* **13**:265-275.
31. Engler, R., A. Guisan, and L. Rechsteiner. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence data and pseudo-absence data. *Journal of Applied Ecology* **41**:263-274.
32. Estrada-Peña, A., C. Sánchez Acedo, J. Quílez, and E. Del Cacho. 2005. A retrospective study of climatic suitability for the tick *Rhipicephalus (Boophilus) microplus* in the Americas. *Global Ecology and Biogeography* **14**:565-573.
33. Fera, T. P. A., and A. T. Peterson. 2002. Prediction of bird community composition based on point-occurrence data and inferential algorithms: a valuable tool in biodiversity assessments. *Diversity and Distributions* **8**:49-56.
34. Ficetola, G. F., and F. De Bernardi. 2004. Amphibians in a human-dominated landscape: the community structure is related to habitat features and isolation. *Biological Conservation* **119**:219-230.
35. Fitzpatrick, M. C., J. F. Weltzin, N. J. Sanders, and R. R. Dunn. 2007. The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography* **16**:24-33.
36. Fleishman, E., R. Mac Nally, and J. P. Fay. 2003. Validation tests of predictive models of butterfly occurrence based on environmental variables. *Conservation Biology* **17**:806-817.
37. Foxcroft, L. C., Rouget, M., Richardson, D. M., and S. MacFadyen. 2004. Reconstructing 50 years of *Opuntia stricta* invasion in the Kruger National Park, South Africa: environmental determinants and propagule pressure. *Diversity and Distributions* **10**:427-437.
38. García, A. 2006. Using ecological niche modeling to identify diversity hotspots for the herpetofauna of Pacific lowlands and adjacent interior valleys of Mexico. *Biological Conservation* **130**:25-46.
39. García, J., S. Suárez-Seoane, D. Mígueles, P. E. Osborne, and C. Zumalacárregui. 2007. Spatial analysis of habitat quality in a fragmented population of little bustard (*Tetrax tetrax*): Implications for conservation. *Biological Conservation* **137**:45-56.
40. Gavin, D. G., and F. S. Hu. 2005. Bioclimatic modeling using Gaussian mixture distributions and multiscale segmentation. *Global Ecology and Biogeography* **14**:491-501.
41. Germaine, S. S., and B. F. Wakeling. 2001. Lizard species distributions and habitat occupation along an urban gradient in Tucson, Arizona, USA. *Biological Conservation* **97**:229-237.
42. Gibson, L. A., B. A. Wilson, D. M. Cahill, and J. Hill. 2004. Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach. *Journal of Applied Ecology* **41**:213-223.
43. Gottfried, M., H. Pauli, K. Reiter, and G. Grabherr. 1999. A fine-scaled predictive model for changes in species distribution patterns of high mountain plants induced by climate warming. *Diversity and Distributions* **5**:241-251.
44. Graham, C. H., and R. J. Hijmans. 2006. A comparison of methods for mapping species ranges and species richness. *Global Ecology and Biogeography* **15**:578-587.
45. Greaves, R. K., R. A. Sanderson, and S. P. Rushton. 2006. Predicting species occurrence using information-theoretic approaches and significance testing: An example of dormouse distribution in Cumbria, UK. *Biological Conservation* **130**:239-250.
46. Grundel, R., and N. B. Pavlovic. 2007. Resource availability, matrix quality, microclimate, and spatial pattern as predictors of patch use by the Karner blue butterfly. *Biological Conservation* **135**:135-144.
47. Guisan, A., O. Broennimann, R. Engler, M. Vust, N. G. Yoccoz, A. Lehmann, and N. E. Zimmermann. 2005. Using niche-based models to improve the sampling of rare species. *Conservation Biology* **20**:501-511.
48. Guisan, A., C. H. Graham, J. Elith, F. Huettmann, and NCEAS Species Distribution Modeling Group. 2007. Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions* **13**:332-340.

49. Heikkinen, R. K., M. Luoto, and R. Virkkala. 2006. Does seasonal fine-tuning of climatic variables improve the performance of bioclimatic envelope models of migratory birds? *Diversity and Distributions* **12**:502-510.
50. Jaberg, C., and A. Guisan. 2001. Modeling the distribution of bats in relation to landscape structure in a temperate mountain environment. *Journal of Applied Ecology* **38**:1169-1181.
51. Jarvis, A. M., and A. Robertson. 1999. Predicting population sizes and priority conservation area for 10 endemic Namibian bird species. *Biological Conservation* **88**:121-131.
52. Jeganathan, P., R. E. Green, K. Norris, I. N. Vogiatzakis, A. Bartsch, S. R. Wotton, C. G. R. Bowden, G. H. Griffiths, D. Pain, and A. R. Rahmani. 2004. Modeling habitat selection and distribution of the critically endangered Jerdon's courser *Rhinoptilus bitorquatus* in scrub jungle: an application of a new tracking method. *Journal of Applied Ecology* **41**:224-237.
53. Johnson, C. J., and M. P. Gillingham. 2004. Mapping uncertainty: sensitivity of wildlife habitat ratings to expert opinion. *Journal of Applied Ecology* **41**:1032-1041.
54. Johnson, C. J., D. R. Seip, and M. S. Boyce. 2004. A quantitative approach to conservation planning: using resource selection functions to map the distribution of mountain caribou at multiple spatial scales. *Journal of Applied Ecology* **41**:238-251.
55. Lennon, J. J., W. E. Kunin, S. Corne, S. Carver, and W. W. S. Van Hees. 2002. Are Alaskan trees found in locally more favourable sites in marginal areas? *Global Ecology and Biogeography* **11**:103-114.
56. Lira-Noriega, A., J. Soberón, A. G. Navarro-Sigüenza, Y. Nakazawa, and A. T. Peterson. 2007. Scale dependency of diversity components estimated from primary biodiversity data and distribution maps. *Diversity and Distributions* **13**:185-195.
57. Lobo, J. M., J. R. Verdú, and C. Numa. 2006. Environmental and geographical factors affecting the Iberian distribution of flightless *Jekelius* species (Coleoptera: Geotrupidae). *Diversity and Distributions* **12**:179-188.
58. Loiselle, B. A., C. A. Howell, C. H. Graham, J. M. Goerck, T. Brooks, K. G. Smith, and P. H. Williams. 2003. Avoiding pitfalls of using species distribution models in conservation planning. *Conservation Biology* **17**:1591-1600.
59. Luoto, M., J. Pöyry, R. K. Heikkinen, and K. Saarinen. 2005. Uncertainty of bioclimate envelope models based on the geographical distribution of species. *Global Ecology and Biogeography* **14**:575-584.
60. Luoto, M., R. Virkkala, and R. K. Heikkinen. 2007. The role of land cover in bioclimatic models depends on spatial resolution. *Global Ecology and Biogeography* **16**:34-42.
61. Lütolf, M., F. Kienast, and A. Guisan. 2006. The ghost of past species occurrence: improving species distribution models for presence-only data. *Journal of Applied Ecology* **43**:802-815.
62. Manel, S., J. M. Dias, S. T. Buckton, and S. J. Ormerod. 1999. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *Journal of Applied Ecology* **36**:734-747.
63. Manel, S., H. C. Williams, and S. J. Ormerod. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* **38**:921-931.
64. Manrique, C. E., R. Durán, and J. Argáez. 2003. Phytogeographic analysis of taxa endemic to the Yucatán Peninsula using geographic information systems, the domain heuristic and parsimony analysis of endemism. *Diversity and Distributions* **9**:313-330.
65. Martínez, I., F. Carreño, A. Escudero, and A. Rubio. 2006. Are threatened lichen species well-protected in Spain? Effectiveness of a protected areas network. *Biological Conservation* **133**:500-511.
66. Martínez-Meyer, E., A. T. Peterson, and W. W. Hargrove. 2004. Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. *Global Ecology and Biogeography* **13**:305-314.
67. McClean, C. J., N. Doswald, W. Küper, J. H. Sommer, P. Barnard, and J. C. Lovett. 2006. Potential impacts of climate change on Sub-Saharan African plant priority area selection. *Diversity and Distributions* **12**:645-655.
68. McPherson, J. M., and W. Jetz. 2007. Type and spatial structure of distribution data and the perceived determinants of geographical gradients in ecology: the species richness of African birds. *Global Ecology and Biogeography* **16**:657-667.
69. McPherson, J. M., W. Jetz, and D. J. Rogers. 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *Journal of Applied Ecology* **41**:811-823.
70. Midgley, G. F., G. O. Hughes, W. Thuiller, and A. G. Rebelo. 2006. Migration rate limitations on climate-induced range shifts in Cape Proteaceae. *Diversity and Distributions* **12**:555-562.
71. Miles, L., A. Grainger, and O. Phillips. 2004. The impact of global climate change on tropical forest biodiversity in Amazonia. *Global Ecology and Biogeography* **13**:553-565.
72. Milsom, T. P., S. D. Langton, W. K. Parkin, S. Peel, J. D. Bishop, J. D. Hart, and N. P. Moore. 2000. Habitat models of bird species' distribution: an aid to the management of coastal grazing marshes. *Journal of Applied Ecology* **37**:706-727.
73. Muñoz, A. R., and R. Real. 2006. Assessing the potential range expansion of the exotic monk parakeet in Spain. *Diversity and Distributions* **12**:656-665.
74. Muñoz, A. R., R. Real, A. M. Barbosa, and J. M. Vargas. 2005. Modeling the distribution of Bonelli's eagle in Spain: implication for conservation planning. *Diversity and Distributions* **11**:477-486.

75. Ortega-Huerta, M. A. 2007. Fragmentation patterns and implications for biodiversity conservation in three biosphere reserves and surrounding regional environments, northeastern Mexico. *Biological Conservation* **134**:83-95.
76. Ortega-Huerta, M. A., and A. T. Peterson. 2004. Modeling spatial patterns of biodiversity for conservation prioritization in North-eastern Mexico. *Diversity and Distributions* **10**:39-54.
77. Osborne, P. E., J. C. Alonso, and R. G. Bryant. 2001. Modeling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. *Journal of Applied Ecology* **38**:458-471.
78. Osborne, P.E., Foody, G.M. & Suárez-Seoane, S. (2007) Non-stationary and local approaches to modeling the distributions of wildlife. *Diversity and Distributions*, **13**, 313-323.
79. Palma, L., Beja, P. & Rodrigues, M. (1999) The use of sighting data to analyse Iberian lynx habitat and distribution. *Journal of Applied Ecology*, **36**, 812-824.
80. Pawar, S., Koo, M.S., Kelley, C., Ahmed, M.F., Chaudhuri, S. & Sarkar, S. (2007) Conservation assessment and prioritization of areas in Northeast India: Priorities for amphibians and reptiles. *Biological Conservation*, **136**, 346-361.
81. Pearce, J.L., Cherry, K., Drielsma, M., Ferrier, S. & Wish, G. (2001) Incorporating expert opinion and fine-scale vegetation mapping into statistical models of faunal distribution. *Journal of Applied Ecology*, **38**, 412-424.
82. Pearce, J.L. & Ferrier, S. (2001) The practical value of modeling relative abundance of species for regional conservation planning: a case study. *Biological Conservation*, **98**, 33-43.
83. Peterson, A.T., Egbert, S.L., Sánchez-Cordero, V. & Price, K.P. (2000) Geographic analysis of conservation priority: endemic birds and mammals in Veracruz, Mexico. *Biological Conservation*, **93**, 85-94.
84. Peterson, A.T., Martínez-Meyer, E. & González-Salazar, C. (2004) Reconstructing the Pleistocene geography of the *Aphelocoma* jays (Corvidae). *Diversity and Distributions*, **10**, 237-246.
85. Polasky, S., Camm, J.D., Solow, A.R., Csuti, B., White, D. & Ding, R. (2000) Choosing reserve networks with incomplete species information. *Biological Conservation*, **94**, 1-10.
86. Reino, L., Beja, P. & Heitor, A.C. (2006) Modeling spatial and environmental effects at the edge of the distribution: the red-backed shrike *Lanius collurio* in Northern Portugal. *Diversity and Distributions*, **12**, 379-387.
87. Rhodes, J.R., Wiegand, T., McAlpine, C.A., Callaghan, J., Lunney, D., Bowen, M. & Possingham, H.P. (2006) Modeling species' distributions to improve conservation in semiurban landscapes: Koala case study. *Conservation Biology*, **20**, 449-459.
88. Robertson, M.P., Caithness, N. & Villet, M.H. (2001) A PCA-based modeling technique for predicting environmental suitability for organisms from presence records. *Diversity and Distributions*, **7**, 15-27.
89. Robertson, M.P., Villet, M.H. & Palmer, A.R. (2004) A fuzzy classification technique for predicting species' distributions: applications using invasive alien plants and indigenous insects. *Diversity and Distributions*, **10**, 461-474.
90. Robinson, R.A., Wilson, J.D. & Crick, H.Q.P. (2001) The importance of arable habitat for farmland birds in grassland landscapes. *Journal of Applied Ecology*, **38**, 1059-1069.
91. Roger, E., Laffan, S.W. & Ramp, D. (2007) Habitat selection by the common wombat (*Vombatus ursinus*) in disturbed environments: Implications for the conservation of a 'common' species. *Biological Conservation*, **137**, 437-449.
92. Rouget, M., Richardson, D.M., Nel, J.L., Le Maitre, D.C., Eogh, B. & Mgidi, T. (2004) Mapping the potential ranges of major plant invaders in South Africa, Lesotho and Swaziland using climatic suitability. *Diversity and Distributions*, **10**, 475-484.
93. Russell, J.C. & Clout, M.N. (2004) Modeling the distribution and interaction of introduced rodents on New Zealand offshore islands. *Global Ecology and Biogeography*, **13**, 497-507.
94. Sánchez-Cordero, V., Illoldi-Rangel, P., Linaje, M., Sarkar, S. & Peterson, A.T. (2005) Deforestation and extant distributions of Mexican endemic mammals. *Biological Conservation*, **126**, 465-473.
95. Schmidt, M., Kreft, H., Thiombiano, A. & Zizka, G. (2005). Herbarium collections and field data-based plant diversity maps for Burkina Faso. *Diversity and Distributions*, **11**, 509-516.
96. Schulte, L.A., Pidgeon, A.M. & Mladenoff, D.J. (2005) One hundred fifty years of change in forest bird breeding habitat: Estimates of species distributions. *Conservation Biology*, **19**, 1944-1956.
97. Schussman, H., Geiger, E., Mau-Crimmins, T. & Ward, J. (2006) Spread and current potential distribution of an alien grass, *Eragrostis lehmanniana* Nees, in the southwestern USA: comparing historical data and ecological niche models. *Diversity and Distributions*, **12**, 582-592.
98. Segurado, P., Araújo, M.B. & Kunin, W.E. (2006) Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology*, **43**, 433-444.
99. Seoane, J., Bustamante, J. & Díaz-Delgado, R. (2005) Effect of expert opinion on the predictive ability of environmental models of bird distribution. *Conservation Biology*, **19**, 512-522.
100. Sérgio, C., Figueira, R., Draper, D., Menezes, R. & Sousa, A.J. (2007) Modeling bryophyte distribution base don ecological information for extent of occurrence assessment. *Biological Conservation*, **135**, 341-351.
101. Silva, T., Reino, L.M. & Borralho, R. (2002) A model for range expansion of an introduced species: the common waxbill *Estrilda astrild* in Portugal. *Diversity and Distributions*, **8**, 319-326.

102. Stockman, A.K., Beamer, D.A. & Bond, J.E. (2006a) An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. *Diversity and Distributions*, 12, 81-89.
103. Suárez-Seoane, S., Osborne, P.E. & Alonso, J.C. (2002) Large-scale habitat selection by agricultural steppe birds in Spain: identifying species-habitat responses using generalized additive models. *Journal of Applied Ecology*, 39, 755-771.
104. Téllez-Valdés, O. & Dávila-Aranda, P. (2003) Protected areas and climate change: a case study of the cacti in the Tehuacán-Cuicatlán Biosphere Reserve, Mexico. *Conservation Biology*, 17, 846-853.
105. Thomson, J.R., Mac Nally, R., Fleishman, E. & Horrocks, G. (2007) Predicting bird species distributions in reconstructed landscapes. *Conservation Biology*, 21, 752-766.
106. Thuiller, W., Lavorel, S. & Araújo, M.B. (2005) Niche properties and geographical extent as predictors of species sensitivity to climate change. *Global Ecology and Biogeography*, 14, 347-357.
107. Thuiller, W., Lavorel, S., Sykes, M.T. & Araújo, M.B. (2006) Using niche-based modeling to assess the impact of climate change on tree functional diversity in Europe. *Diversity and Distributions*, 12, 49-60.
108. Thuiller, W., Vayreda, J., Pino, J., Sabate, S., Lavorel, S. & Gracia, C. (2003) Large-scale environmental correlates of forest tree distributions in Catalonia (NE Spain). *Global Ecology and Biogeography*, 12, 313-325.
109. Tourenq, C., Aulagnier, S., Durieux, L., Lek, S., Mesléard, F., Johnson, A. & Martin, J.-L. (2001) Identifying rice fields at risk from damage by the greater flamingo. *Journal of Applied Ecology*, 38, 170-179.
110. Travaini, A., Bustamante, J., Rodríguez, A., Zapata, S., Procopio, D., Pedrana, J. & Peck, R.M. (2007) An integrated framework to map animal distributions in large and remote regions. *Diversity and Distributions*, 13, 289-298.
111. Tsoar, A., Allouche, O., Steinitz, O., Rotem, D. & Kadmon, R. (2007) A comparative evaluation of presence-only methods for modeling species distribution. *Diversity and Distributions*, 13, 397-405.
112. Underwood, E.C., Klinger, R. & Moore, P.E. (2004) Predicting patterns of non-native plant invasions in Yosemite National Park, California, USA. *Diversity and Distributions*, 10, 447-459.
113. Usio, N. (2007) Endangered crayfish in northern Japan: Distribution, abundance and microhabitat specificity in relation to stream and riparian environment. *Biological Conservation*, 134, 517-526.
114. Vanreusel, W., Maes, D. & Van Dyck, H. (2007) Transferability of species distribution models: a functional habitat approach for two regionally threatened butterflies. *Conservation Biology*, 21, 201-212.
115. Vargas, J.H., Consiglio, T., Jørgensen, P.M. & Croat, T.B. (2004) Modeling species distribution patterns in a species-rich plant genus, *Anthurium* (Araceae), in Ecuador. *Diversity and Distributions*, 10, 211-216.
116. Wang, H.G., Owen, R.D., Sánchez-Hernández, C. & Romero-Almaraz, M.L. (2003) Ecological characterization of bat species distributions in Michoacán, Mexico, using a geographic information system. *Global Ecology and Biogeography*, 12, 65-85.
117. Weaver, K.F., Anderson, T. & Guralnick, R. (2006) Combining phylogenetic and ecological niche modeling approaches to determine distribution and historical biogeography of Black Hills mountains snails (Oreohelicidae). *Diversity and Distributions*, 12, 756-766.
118. Wharton, T.N. & Kriticos, D.J. (2004) The fundamental and realized niche of the Monterey Pine aphid, *Essigella californica* (Essig) (Hemiptera: Aphididae): implications for managing softwood plantations in Australia. *Diversity and Distributions*, 10, 253-262.
119. Wheatley, M., Fisher, J.T., Larsen, K., Litke, J. & Boutin, S. (2005) Using GIS to relate small mammal abundance and landscape structure at multiple spatial extents: the northern flying squirrel in Alberta, Canada. *Journal of Applied Ecology*, 42, 577-586.
120. White, P.J.T. & Kerr, J.T. (2007) Human impacts on environment-diversity relationships: evidence for biotic homogenization from butterfly species richness patterns. *Global Ecology and Biogeography*, 16, 290-299.
121. Wilson, K.A., Westphal, M.I., Possingham, H.P. & Elith, J. (2005) Sensitivity of conservation planning to different approaches to using predictive species distribution data. *Biological Conservation*, 122, 99-122.
122. Worner, S.P. & Gevrey, M. (2006) Modeling global insect pest species assemblages to determine risk of invasion. *Journal of Applied Ecology*, 43, 858-867.
123. Yom-Tov, Y. & Kadmon, R. (1998) Analysis of the distribution of insectivorous bats in Israel. *Diversity and Distributions*, 4, 63-70.