Development of Indicators of Compliance with the *2010 Target* of the Convention on Biological Diversity Using Primary Biodiversity Data Provided by GBIF

May 24, 2007

A. T. Peterson and J. Soberon

Biodiversity Research Center

University of Kansas

_____

I. Introduction

In April 2002, at the VI Conference of the Parties of the Convention on Biological Diversity (CBD), most countries of the world decided to: "*achieve by 2010 a significant reduction of the current rate of biodiversity loss at the global, regional and national level*" (Decision VI/26). In order to verify the fulfillment of the decision, the several bodies and advisors to the CBD began proposing indicators to measure changes in biodiversity. Thus far, the only indicators of species-level biodiversity status that have been implemented, circulated, and quoted widely in CBD documents are aggregated indicators, compiled at the global level by consortia of NGOs and United-Nations bodies. In particular, Red List Indicators (RLI) (Butchart et al. 2005; Butchart et al. 2004) and the Living Planet Index (LPI) (Loh et al. 2005; Loh and Wackemagel 2004) are quoted as such indices. These indices are global in scope (they can sometimes be downscaled to regional levels), are based on secondary data, and are assembled by groups of experts in the developed world. As such, they are difficult to scale and extend to other spatial extents and other situations.

The existence of global-level indicators should be welcomed since they provide a useful and practical view of biodiversity at the planetary level. However, the CBD, in its decision VIII/15.11, urged parties and other governments to "*develop national and/or regional goals and targets and related indicators*" and also emphasized the "*need for capacity building, access to and transfer of technology … in order to enable [countries] to develop knowledge, including taxonomic knowledge, to gain access to their biodiversity, and to better implement activities to achieve and monitor progress towards the goals and targets*" (UNEP/CBD/COP/8/31/VIII/15.12).

In this report, we will argue, with detailed, real-life, worked examples, to demonstrate:

1) That the primary biodiversity data served through the GBIF portal can be used to provide indices of "*trends in the abundance and distribution of selected species*" (UNEP/CBD/COP/8/31/VIII/15.12).
2) Following wording reiterated over many documents and decisions of the CBD, that GBIF-data-based indices can in principle be applied to any kind of "*selected species*" and not only to those that are globally endangered, as the RLI inevitably does.
3) That such new indices, based on publicly available information such as that enabled by GBIF, can be calculated at global, regional or national levels by trained local experts and, by resorting to ancillary data, can be expressed as time series.
4) And finally, that the very existence of GBIF is a major effort of transferring taxonomic knowledge, consistent with CBD exhorts and decisions; training experts in developing countries, as GBIF has been doing in the last two years, represents a concrete example of capacity-building in those skills required for countries to develop their own indicators of achievement of the 2010 target.

In our view, primary data of the kind that GBIF has been enabling, improving, and making freely available represents a fundamental asset for developing countries,

specifically to develop indicators of achievement of the 2010 Target. The Governing Council of GBIF should made be fully aware of this capacity and functionality.

II. Current GBIF Data Holdings as They Relate to the Indicator Process

Biodiversity informatics projects are very numerous at present (Canhos et al. 2004). Different efforts focus on aspects of taxonomy (e.g., Species2000[1], ITIS[2], GTI[3]), conservation status (e.g., WCMC[4], BirdLife International[5]), and natural history (e.g., FishBase[6]). This field is burgeoning with the appearance of funding programs, conferences, symposia, and courses to fill the need for relevant information.

A key role, however, is played by what we will call *primary biodiversity information*. Primary biodiversity information is made up of data records that place a particular taxon in a particular place at a particular point in time. All other sorts of biodiversity data—names, conservation status estimates, range maps, etc.—are secondary in nature. That is, secondary sources represent the product of interpretation by humans, and are invariably based in some way on primary biodiversity information—as such, the primary data are the key infrastructural element for the field of biodiversity informatics.

The traditional repositories of primary biodiversity information have been natural history museums worldwide, which are estimated to hold on the order of 1×10^9 scientific specimens (Krishtalka and Humphrey 2000). These specimens and associated data are particularly crucial, as they hold the key to linking to biological nomenclature, via type specimens and the rich scientific literature over the past three centuries. More recently, however, numerous observation-based efforts have been initiated, which are now providing extensive additional primary biodiversity information, and with magnitudes of numbers of records much larger than the specimen-based data sets. Both types of primary biodiversity information, however, are useful—the specimen information has the advantage of being vouchered, permitting reidentification and verification, and of providing a rich historical context; the observational data, on the other hand, have the force of numbers to document complex phenomena in greater detail.

Although GBIF has several primary biodiversity data sets that can be assembled very naturally in time series (for example bird observations, on the order of 10^7 data records), the statement has been made that the museum holdings cannot be used to establish trends since they cannot be interpreted simply as time series (UNEP-WCMC 2006). This statement is inaccurate, as we will demonstrate later, but we first begin describing some features of the current GBIF database.

GBIF currently (as of September 2006) has just crossed the landmark of 10^8 data records from 179 data providers. Data providers are scattered worldwide, although they are concentrated in Europe and North America. This pattern is perhaps appropriate, given the well-known concentration of biodiversity information in the biodiversity-poor countries of the North. GBIF provided the authors of this report with a broad sample of the data available on the site as of early September 2006, in all 7.9×10^7 data records, of which we will present a brief analysis herein.

---

[1] http://www.sp2000.org/.
[2] http://www.itis.usda.gov/.
[3] http://www.biodiv.org/programmes/cross-cutting/taxonomy/default.asp.
[4] http://www.unep-wcmc.org/.
[5] http://www.birdlife.org/.
[6] http://www.fishbase.org/home.htm.
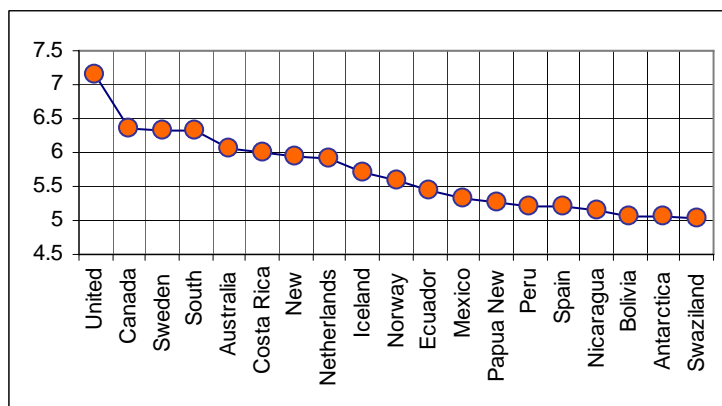
The earliest records in the GBIF network come from around 1800, indicating that 200 yr of scientific information is summarized in the network. Interestingly, the accumulation of information per year is quite-nearly linear on a logarithmic scale (see figure at right, which summarizes number of records in the GBIF-mediated archive per year over the past 200 years). With each 50-55 yr of time, the amount of information available *per year* increases by an order of magnitude. Although it could be argued that this tendency may

not continue far into the future, many large-scale observational data-gathering efforts promise extremely dense biodiversity sampling, and new technologies (e.g., remote sensing at fine scales) may indeed make such data fluxes possible.
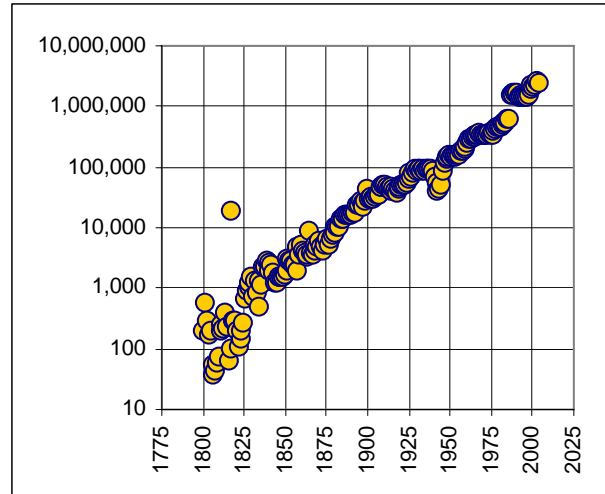
Countries ranged in representation in the GBIF data resources in terms of numbers of primary biodiversity records from 2 (Liechtenstein) to >16 x 10^6 (United States) primary biodiversity records (see graph at left). The highest ranks are populated by countries that might be expected (United States, Canada, Sweden, South Africa, Australia, Costa Rica, New Zealand, etc.). However, it is worthy of note that countries represented by >10^5 records include several for which organized biodiversity informatics efforts are lacking—nonetheless GBIF now makes considerable biodiversity information available to Ecuador, Papua New Guinea, Peru, Nicaragua, Bolivia, and even Swaziland.

One key to 'enabling' biodiversity data fully is that of adding quantitative geographic references, in the form of geographic coordinates. Overall, at present, 79.6%

of data records in the GBIF network are assigned geographic coordinates, which cover much of the surface of the Earth, although at widely varying densities (see map below, which shows the $\log_{10}$ of the number of records per 1° cell globally). The remaining 20.4% records served via GBIF will represent an important future

challenge for the network—enabling these records spatially via adding geographic references will greatly enrich the utility of the GBIF network for spatial questions. Newly available Internet-based software tools[7] will clearly assist in this challenge.

---

[7] http://www.biogeomancer.org.
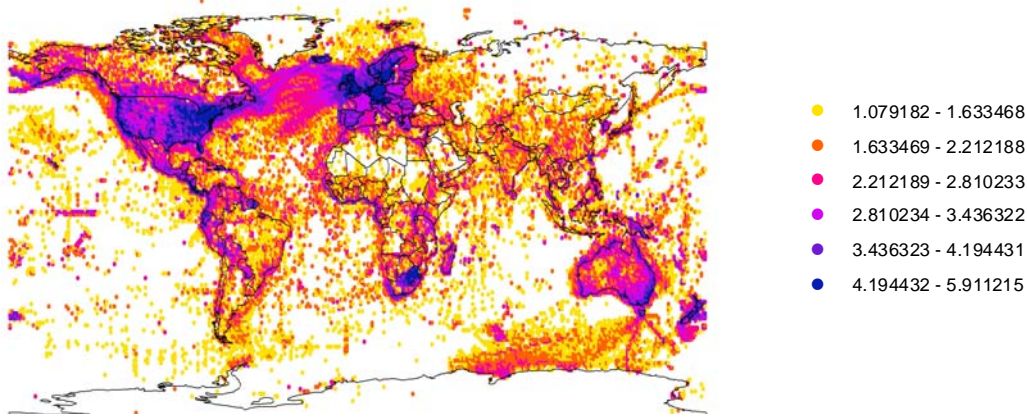
The above map clearly displays the predominant amount of records in the developed world. However, it also shows that a number of developing countries appear to be well covered. This can be shown more precisely in the following figure where
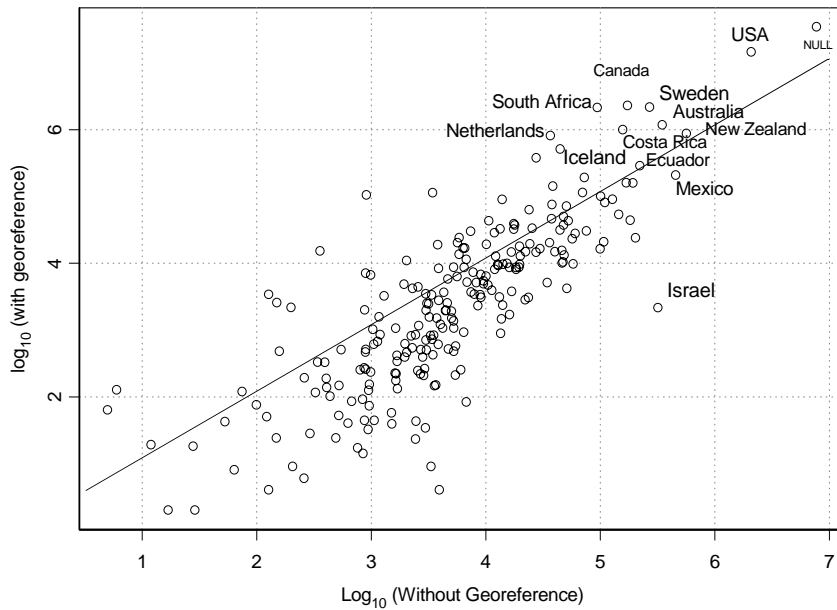
numbers of records georeferenced and not georeferenced for each country are displayed (graph has selected countries labeled). In general, a positive relationship exists between the two, but some countries are weighted on the side of more records georeferenced than not (e.g., South Africa, Canada, Netherlands); other countries are weighted towards many records not yet being georeferenced (e.g., Israel, Mexico)—these countries are certainly those for which a modicum of effort can yield immediate benefits in terms of improved biodiversity knowledge.

In sum, data sets served via the GBIF portal represent a sizeable amount of primary information about biodiversity. Indeed, the GBIF data facility is without doubt the largest single data resource regarding biodiversity in the world. GBIF-mediated data have grown impressively in the short life of the facility, and show every promise of continuing this trajectory.

III. 2010 Target Monitoring – Existing Work and Possibilities for GBIF.

A sizeable literature has emerged regarding measurement of progress towards the 2010 goal of reducing biodiversity loss globally—this work has taken two general tracks:

1. Use of biodiversity and biodiversity loss indices based on aggregations of disparate data assembled from a number of sources. An example is the Living Planet index (LPI) (Loh and Wackemagel 2004).
2. Use of biodiversity indices based on number of species assigned by experts to different endangerment categories, particularly the 'Red Lists' maintained from the UNEP-World Conservation Monitoring Centre. These indices are called Red List Indices (RLI) (Butchart et al. 2005; Butchart et al. 2004; Smith et al. 1993; Stuart et al. 2004).

The result of broad application of these approaches has been the conclusion that biodiversity loss continues, and that conservation activity in recent decades have had little impact on the rate of this loss. Regional differences—of course—are detected, and these differences map onto known shifts in land use and human population distributions.

While these approaches have yielded interesting insights into the global status and broad-scale trends of biodiversity through time, they are nonetheless surprisingly indirect or opaque. That is, each is dependent on numerous assumptions to connect the index with biodiversity loss. For example, the LPI approach depends on pooling (taking the geometric mean of) a hodgepodge of independent species trends (disaggregated down to major biome or biogeographic realm) provided by secondary sources including gray literature. The idea is simple and probably valid, but independent reconstruction of the index would be impossible without access to the original worked-out databases. The reader has no choice but to trust the authors.

The endangerment index-based approach, on the other hand, depends on the accuracy and consistency of the expert opinion upon which those indices are based. What is more, RLIs are a function only of the status of threatened and endangered species (the categories, although operationally defined, are arbitrary), and are limited to those groups for which comprehensive endangerment assessments have been developed. As such, the RLI approach is little scalable and extendable to include larger sectors of biodiversity.

In both cases, the inferences used to date as indicators for the 2010 target tracking are based on secondary information sources, and are little able to be applied at other scales (e.g., single countries, regions). As such, these previous efforts are limited, their applicability is not general and their acceptance by countries in fora like the CBD is at best problematic. GBIF-based information represents an interesting alternative that is scalable and extendable, and that can be implemented by scientists and technicians in countries of high biodiversity concern.

Trend Analysis – A GBIF-based Example

Primary biodiversity data can form the basis for development of detailed summaries of species' geographic and ecological distributions (Elith et al. 2006; Guisan and Zimmermann 2000; Soberón and Peterson 2004). These data can be integrated in diverse manners with information on land use change and other environmental change to derive species-specific assessments of range loss (or gain) (Araújo et al. 2006; Pearson and Dawson 2003; Pearson et al. 2002; Peterson et al. 2006; Peterson et al. 2005;

Sánchez-Cordero et al. 2005; Thomas et al. 2004). The large-scale data resources served by GBIF offer opportunities for development of biota-wide assessments (Peterson et al. 2002b), but based on real species and their individual ecological needs and geographic distributions.

Here, we develop a small-scale illustration of the potential of these approaches—based on a comparison of two regions of Mexico in terms of land-use mediated biodiversity loss. This example is not intended to be comprehensive in any way, nor is it developed in sufficient detail to be satisfactory—rather, we intend to illustrate the feasibility and potential of the approach. We emphasize three points: (1) Such analyses can be developed using GBIF data and can provide detailed trend assessments. (2) They can be performed at any spatial scale, from that of continents down to the resolution of the georeferencing of the occurrence points themselves. (3) These analyses can be performed after relatively little training, and are accessible to technicians in any developing country (note that GBIF has to date carried off 2 global training workshops in these techniques).

Predicting Geographic Distributions of Species

A first challenge in development of this illustration is the question of whether GBIF-served primary biodiversity data are sufficient as the basis for models designed to predict geographic distributions of species. We base our distributional predictions on the technique of ecological niche modeling (ENM). ENM is a suite of quantitative tools aimed at reconstructing the ecological requirements of species—the "ecological niche," which can be defined as the set of environmental conditions within which a species can maintain populations without immigrational subsidy (Grinnell 1917; Grinnell 1924). Once the ecological niche is characterized, it can be used to query a landscape (at any scale) to identify areas of potential distribution for the species; these potential distributional estimates are extremely useful in anticipating the potential for species' invasions (Peterson 2003a; Peterson and Vieglais 2001), but can also be reduced to hypotheses of actual distributions of species under quantitative and repeatable assumption sets. A sizeable literature now documents the predictive power of these approaches (Araújo and Guisan 2006; Araújo et al. 2005b; Berry et al. 2002b; Elith and Burgman 2002; Elith et al. 2006; Guisan et al. 2007; Guisan and Zimmermann 2000; Illoldi et al. 2004; Lehmann et al. 2002; McNyset 2005; Pearson et al. 2002; Pearson et al. 2006a; Peterson 2001; Peterson et al. 2002a; Phillips et al. 2004; Thuiller et al. 2004; Wiley et al. 2003).

As a GBIF-data-based illustration of these approaches, we focused on the Northern Anteater *Tamandua mexicana*. The GBIF network provided 90

occurrences for the species, drawn from 12 institutional databases (Cornell University Museum of Vertebrates, Field Museum of Natural History, Florida Museum of Natural History, Museum of Comparative Zoology, Michigan State University Museum, Museum of Natural Science, Louisiana State University, Museum of Vertebrate Zoology, Paleobiology Database, Royal Ontario Museum, Instituto de Biología UNAM, University of Tennessee – Chattanooga, and the University of Washington Burke Museum). These occurrences were sparsely scattered across the species' known distribution in the northern Neotropics (see map above).

An ENM was developed for *Tamandua mexicana* using a genetic-algorithm approach[8], which provided a prediction of the species' geographic distribution. To assess the robustness of this prediction, we overlaid more detailed data available to us from within Mexico—another 45 occurrence points. Interestingly, these occurrence data corroborated several of the ENM predictions that may, at first glance, have appeared to be modeling errors or artifacts, in particular the occurrence of the species broadly along the west coast of Mexico and across most or all of the Yucatan Peninsula (see map at left,
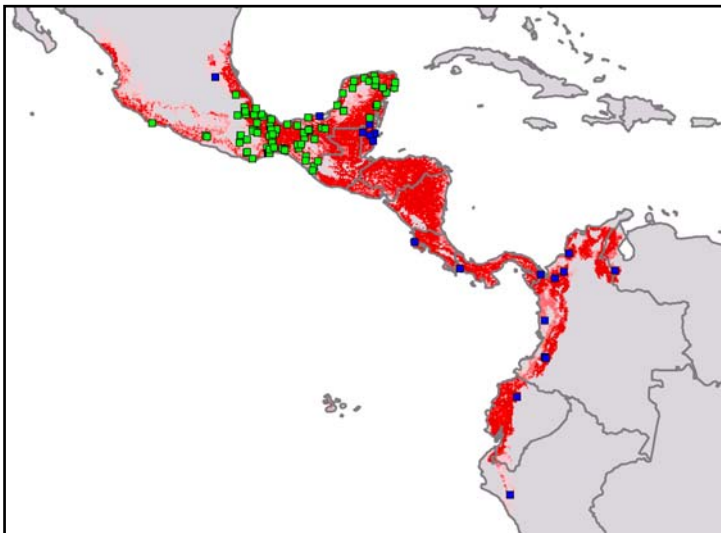
which shows GBIF-mediated data on which the ENM was based as blue squares, and independent national data as green squares). Thus we see that the GBIF-served data were more than sufficient to anticipate the spatial distribution of an additional 50% more information and to characterize the species' geographic distribution.

Assessing Effects of Change on Species' Distributions

Several approaches for assessing the effects of environmental change on species' geographic distributions have been developed. That which has seen the most exploration is that of development of ENMs based on present-day climates, and identification of areas fitting those conditions under future climate conditions (Anciães and Peterson 2007; Araújo et al. 2005a; Araújo et al. 2006; Bakkenes et al. 2002; Berry et al. 2002a; Erasmus et al. 2002; Huntley et al. 1995; Kadmon and Heller 1998; Midgely et al. 2002; Midgley et al. 2003; Parra-Olea et al. 2005; Pearson and Dawson 2003; Pearson et al. 2002; Peterson 2003b; Peterson et al. 2004; Peterson et al. 2002b; Peterson et al. 2001; Roura-Pascual et al. 2005; Siqueira and Peterson 2003; Thomas et al. 2004; Thuiller et al. 2005a; Thuiller et al. 2006), with a growing number of quantitative tests now confirming the predictive ability of these models (Araújo et al. 2005a; Martínez-Meyer 2002; Martínez-Meyer and Peterson 2006; Martínez-Meyer et al. 2004). However, as the climate change effects modeled in these studies are generally at a 2050 horizon, they are

[8] http://www.lifemapper.org/desktopgarp/.

not so directly relevant to assessing trends relevant to the 2010 target, although more will be said about climate change effects on biodiversity below.

More directly relevant to the 2010 target are approaches oriented at capturing the effects of land use change and conversion of primary habitats into anthropogenic landscapes on species' distributions (Peterson et al. 2006; Sánchez-Cordero et al. 2005; Thuiller et al. 2004), which have indeed been the subject of a GBIF-sponsored demonstration project[9]. Here, ENMs were developed based on data layers summarizing climate variables, topographic information, and original or primary vegetation, and then are projected onto the same conditions, but with present-day land use instead of the original vegetation maps. The result is a before-and-after view of the geographic distribution of individual species, from which range loss can be calculated easily.

As an illustration of the possibility of measuring trends in species' distributional areas and conservation status, focusing on Mexican mammals, we chose six species in the endangered list of Mexico: three that are restricted to lowland rain forest in southeastern Mexico (*Tamandua mexicana*, *Ateles geoffroyi*, *Alouatta palliata*) and three species of highland pine-oak forest in northwestern Mexico (*Sciurus nayaritensis*, *S. aberti*, *Peromyscus spicilegus*). We also included a non-endangered but commercially important tree species, *Guaicum sanctum*. For each species, we obtained all of the georeferenced occurrence data available via the GBIF site, which invariably were drawn from 8-15 institutions. These occurrence data were fed into a genetic algorithm designed to develop ENMs (Stockwell and Peters 1999)[10], along with digital environmental data sets summarizing aspects of climate and topography. Then, as these 7 species are all connected with primary forest habitats, we reduced ENM predictions to those areas retaining primary forest as of 1976, 1994, and 2000 [details on the land-cover data sets provided elsewhere (Peterson et al. 2006)], and measured loss of range area, in km^2, for each species individually.

These projections showed a range of effects of land use conversion on species' distributions, from near nil (e.g., *Sciurus* spp.) to dramatic (e.g., *Tamandua mexicana*). In each case in which range loss was significant, most of the range loss took place prior to 1976, but with losses continuing through 2000 (see graph below, which shows decline in distributional area for 7 species for 1940-2000—species in purple inhabit lowland rain


---

[9] http://www.unibio.unam.mx/splm/index.faces.
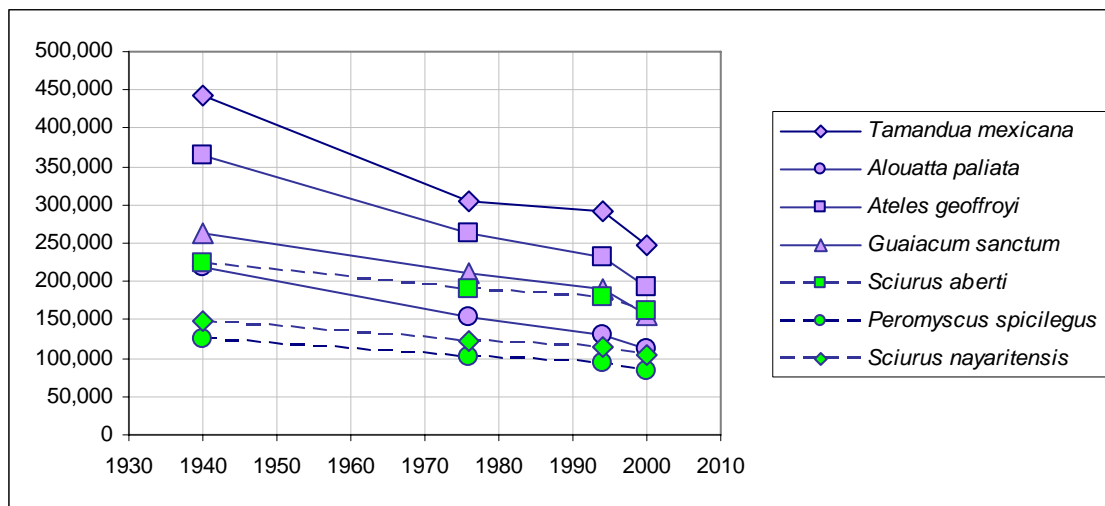
forest species, whereas species in green inhabit pine-oak forest). Both the pattern and the magnitude of range loss varied from species to species, emphasizing the key importance of individual species-based projections and the minimal utility, at the country-level, of aggregated global indices. The examples presented herein are—of course—for the purposes of illustration only, and could be improved greatly in several ways; however, our point is that GBIF-served data can be used quite effectively to develop species-specific indicators of trends in biodiversity loss.

Summarizing Trends and Indicators

The two habitats/regions of Mexico evaluated above contrast sharply in their rates of biodiversity loss, estimated as the slope (in km^2/year) of range loss (DA) among the mammals in the tropical forests (right box) and those in the pine forest (left box) (see graph at right, which shows slopes of linear regressions of distributional areas from 1940-2000, in which boxes show standard errors, and whiskers show ranges; bce = pine-oak forest, bt = lowland rain forest). While range loss among mammal species of pine-oak forest in northwestern Mexico is on the order of 735 km^2/year, that among species of rainforest in southeastern Mexico is considerably greater, on the order of 2500 km^2/year. These contrasting rates of loss illustrate the effectiveness of the species-specific indicators of biodiversity loss explored herein.
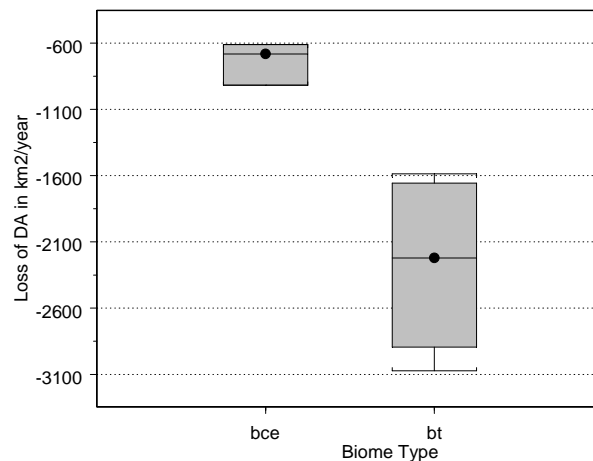
Summary Comments

The species' range loss indicators developed above for 7 exemplar species from 2 habitats/regions of Mexico—if implemented broadly—can then be used to ask more general questions about biodiversity status trends in those regions. Of course, the evaluations presented herein are based on few species, and would have to be extended to many more species before they would be useful for real trend analysis, but the example is provided for the purpose of illustration. Each of these species took <2 hours for preparation of the analyses from start to finish; processing large numbers of species, with data drawn directly from the GBIF site, would be even more efficient, and thousands of species could be processed with comparatively little time and effort. Training and building capacity in application of these methods is completely feasible, as has been demonstrated by GBIF in two training courses to date—more than 40 technicians, researchers, and students from around the world are now reasonably well-versed in the ENM methodology.

Most importantly, the methods explored herein are both scalable and extendable (see summary table below). That is to say, we illustrate the approach with Mexican mammals, but implementation for trees or butterflies or beetles or any other group is

---

[10] http://www.lifemapper.org/desktopgarp/.

completely feasible, providing that occurrence data such as those enabled by GBIF are available. What is more, we have illustrated the methodology at the level of regions within a country—it can easily be implemented at coarser scales (e.g., countries to continents), or at finer scales (e.g., comparisons of areas inside and outside of protected areas on local landscapes). This broad spectrum of applicability is both enormously promising, and contrasts sharply with the previous efforts to develop 2010 target indicators (Butchart et al. 2005; Butchart et al. 2004; Loh and Wackemagel 2004; Smith et al. 1993; Stuart et al. 2004), which are neither easily scalable nor easily extendable to other taxa.

Table 1. Summary comparison between Red-List Indices (RLI) versus ENM-based distributional analysis approaches.

Feature	Red-list Indices	ENM-based Distributional Analysis
Relevance to CBD Objectives	Conservation of biodiversity	Conservation of biodiversity, Sustainable use, Benefit sharing
Relevance to National Strategies	Indirect: global indices are difficult to interpret at a national level	Direct: species can be selected at national and local levels, and for reasons different from IUCN endangerment criteria
Types of Data	Mostly secondary and gray literature: not widely available	Primary: publicly available (via GBIF, other biodiversity data resources, geospatial data resources)
Spatial Scope and Scaling	Global, sometimes regional	Global, regional, national and even down to local, subject to ENM maximum resolution
Taxonomic Scope	Birds and amphibians. Soon, other terrestrial vertebrates. Subject to completion of status summaries for major taxa	Essentially all species for which data resources are sufficient and that relate to coarse-scale ecological features
Uncertainties	Considerable and poorly characterized, based on changes of endangerment categories and expert opinion	Variable, but quantifiable, based on probabilistic and statistical methods
Ownership	International NGOs and UN agencies	Any researcher with proper training, including national biodiversity agencies
Testability	Dubious, based on arbitrary categories and classifications	Direct: field tests can be carried out readily

IV. Proposed Approach

As we have already said, the fact that GBIF-served data (with the exception of the observational data) are not intrinsically temporal does not prevent their use in developing temporally-organized indicators. In essence, what we propose is to use primary data to estimate areas of distribution and then reduce those areas on the basis of land-use information that can be organized as time series. The approach we propose is based on a published method (Peterson et al. 2006; Sánchez-Cordero et al. 2005) and is summarized in the figure below.
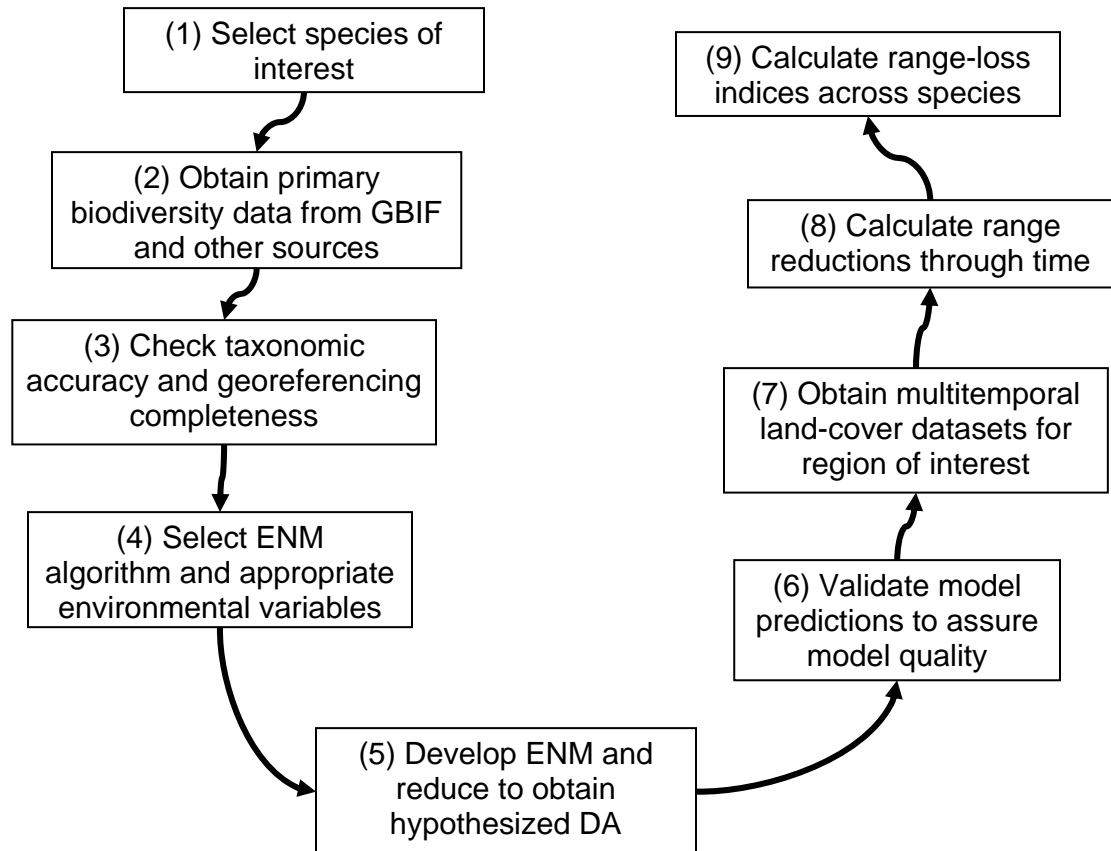
1) **Species selection**. A single species or group of species, or entire faunas, floras, or biotas, can be selected for any reason related to the objectives of the CBD: for example, endangered species subject to national legislation, species of economic importance, or taxonomic groups with potential for bioprospecting. However, since the procedure requires that areas of distribution of these species be reduced using information about land-use change, it is important to select species with clear habitat preferences or associations. For example, only the most drastic changes in land-use will affect species like European Starling (*Sturnus vulgaris*) or Red-winged Blackbird (*Agelaius phoeniceus*), whereas groups like manakins (Pipridae) or forest primates tend to be affected more directly by land-changes. These latter species would be good candidates to select for our method, but habitat generalists would generally be poorer choices.

2) **Data acquisition**. The current GBIF interface allows the user to obtain raw data from hundreds of providers around the world quickly and efficiently. Other data resources, of course, exist and can be consulted, but the GBIF data resources are the largest in the world by orders of magnitude. The entrance to the GBIF query interface is via a catalogue of names that eventually will contain the authoritative catalogues that GBIF is assembling. For the present, however, this interface only enables the user to navigate the most obvious taxonomic hurdles.

3) **Taxonomic and geographic consistency**. The data obtained from the interface must be checked against authoritative taxonomic sources to enable detection and resolution of taxonomic problems such as homonyms and synonyms. This problem becomes more complex when numerous disparate databases are aggregated, but use of mapping tools, good metadata, and taxonomic catalogues can help considerably in detecting problems. Georeferencing should also be checked, at the very least to spot inconsistencies between coordinates and political boundaries, and checking the spatial precision of the georeferencing. Many major collections are now performing such checks, and software capable of dealing with such problems is now available. Unwary users, however, may still believe that every GBIF-provided record has the same quality—rather, GBIF should be up front about the need to check for inconsistencies in taxonomy or georeferencing.

4) **Selection of ENM algorithms and environmental variables**. To estimate the distribution area (DA) of a species, the first step is to model its ecological niche using ENM tools (Soberón and Peterson 2005). At least 15 algorithms have now been used for this purpose—these approaches often differ in their predictive capacities (Elith et al. 2006; Pearson et al. 2006b). Given the fact that GBIF-mediated databases generally do not include information regarding absences of species (although absence information can be extracted from some GBIF databases), the ENM algorithm must be one of the "profile" (presence-only) methods or the algorithm must be provided with pseudoabsence information. Depending on details such as the spatial extent over which predictions will be made, one method might be preferable over another. For example, very dense sampling over relatively small regions can be modeled very well with Maxent (Phillips et al. 2006). Larger regions, with smaller amounts of more clumped occurrence points may be better modeled using GARP (Soberón and Peterson 2005).

ENM requires a suite of informative environmental variables that can be used by the algorithm to characterize the ecological niche. Climatic and topographic variables have shown to be reliable predictors of species' distributions, but the details of these choices depend on spatial scale (Guisan and Zimmermann 2000) and (of course) data availability. Distributions of species normally have geographic extents of about 10^5-10^7 km^2, and can be estimated at resolutions of 10^0-10^3 km^2. At these scales--scales suitable for development of biodiversity loss indicators, experience has shown that digital elevation models and mesoclimate databases like WorldClim can provide rich environmental information; also promising are multitemporal vegetation indices drawn from remotely sensed imagery. These types of variables are available publicly worldwide[11,12]; in addition, increasing numbers of countries, including developing

---

[11] http://www.worldclim.org/.
[12] http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html.

countries, have their own, higher-resolution data layers that also are suitable for ENM.

5) **Modeling**. Application of algorithms like Maxent or GARP generate an ENM, which in turn can be used to identify putatively habitable areas across a geographic region. This geographic expression of the ecological niche constitutes a hypothesis regarding the *potential* (not actual) area of distribution of a species. Depending on technical considerations (Soberón and Peterson 2005), this potential area can be reduced to a hypothesis of an actual distributional area using explicit assumptions about sampling biogeographic regions, or via incorporating information about migration and population expansion (Svenning and Skov 2004). Estimation of DAs has intrinsic uncertainties that are seldom quantified—this area of research is currently quite active, and methods for estimation of probabilistic uncertainties to DA are under development (Argaez et al. 2005).

6) **Validating models**. Hypothesized DAs should be checked empirically. Options range from (best) independent field explorations designed to accumulate independent information with which to test the model predictions (Feria and Peterson 2002; Raxworthy et al. 2003), to (acceptable) checking model predictions against unused subsets of data (Anderson et al. 2003; Elith et al. 2006; Guisan and Zimmermann 2000). Such validations are onerous, but are feasible even for large numbers of species.

7) **Incorporation of land-use change data**. The next step in the process requires reducing estimated DAs using information about land-use change from multiple time periods. This type of information is available world-wide at low resolutions, using data drawn from the AVHRR and MODIS sensors. At higher resolutions, LandSat or even Spot data can be used, but their availability becomes more restricted and costly. For the long-term development of biodiversity indicators, many countries are focusing resources on improving access to higher-resolution land-use data. Development of trend indicators, as exemplified above, requires such information for multiple time periods.

8) **Reduction of DA by land-use change data**. The penultimate step consists of successive maskings of DAs using land-change information. This step is performed via straightforward GIS technology, and yields DA estimates that are functions of time. A distributional estimate can be obtained for each land-use data set available. These reductions in DA can be expressed in km^2, or as a percentage loss of the original distributional area.

9) **Indices of biodiversity loss**. An almost endless number of indices can be calculated from temporal series of reductions in DA[13]. Here, we have used one of the simplest: reduction in DA with time. In a plot of remaining area as a function of time, the slope of a linear regression model (in km^2/year) provides a simple index of biodiversity loss. In this report, we have made no effort to use more sophisticated measures like fragmentation, connectivity, and so on, although border effects are well-known to be important in conservation of viable populations (Stouffer et al. 2006). Rather, the point is that the method we have described yields time-series of DAs that: (i) have direct biological interpretations, (ii) are obtained rigorously and repeatably, using publicly-available primary biodiversity data and well-tested algorithms, and (iii) can

---

[13] http://www.umass.edu/landeco/research/fragstats/fragstats.html

be analyzed statistically and quantitatively in a variety of ways to suit the needs of different users.

The conclusion of this section is that procedures have been developed that allow diverse countries to use GBIF-mediated data to create their own indices of biodiversity status for monitoring compliance with the Target 1.2 of the CBD's decision VI/26. Selecting a number of species of global concern and performing the analysis at a global scale is certainly feasible, and this exercise would provide an interesting contrast to existing global indices. Users can be trained in the use of these procedures—in fact, GBIF has already trained >40 students from many parts of the world in its use. International bodies like the CBD should be made aware of the tremendous potential that GBIF data has for the monitoring of the 2010 Targets, and increased training should be encouraged.

V. Improvements to GBIF Data

GBIF data are primary biodiversity data (Soberón and Peterson 2004). Besides being raw data lacking subjective interpretations, the main advantages of primary biodiversity data lies in their sheer magnitude, as well as in scalability and extendability to many uses and applications. No other non-molecular biological data resource can match the magnitude of these data, either in the form of museum specimens or of observations of key taxa. Use of primary data allows maximum flexibility of use, but at the same time requires careful quality control, and intelligent use of statistical and informatics tools (Chapman 1999; Chapman 2005; Chapman et al. 2005; Soberón et al. 1996). Three main problems may hinder use of GBIF data for indicator purposes: (i) improper taxonomic identifications, (ii) adherence to incorrect or outdated taxonomic authorities, and (iii) lack of correct and precise georeferencing. We discuss each of these concerns briefly below.

Taxonomic metadata. Since GBIF does not manage the datasets it serves, nor perform the observations, it cannot control the correctness of taxonomic determinations. What GBIF can and should do is to request data providers to include full documentation about datasets. This documentation ("metadata," in the jargon of the field) should be detailed enough as to enable any user to understand what steps have and have not been taken to assure that correct identifications are assigned to each data records. Data providers can point to doubtful sectors of datasets where misidentifications are most likely to occur. Providing means by which authoritative users can give feedback about misidentifications and other taxonomic problems would constitute a source of value to data providers and assist in improving the overall data resource.

Authoritative taxonomic authority information. When querying a primary biodiversity database, outdated taxonomy can generate type I (rejecting good data) and type II (accepting faulty data) errors. The need for tools and practices that enable users to assess quality of datasets cannot not be overstressed, since outdated nomenclature can populate queries with incorrect presences that will confuse any algorithm, or thin-down an otherwise sizeable dataset. Since databases of heterogeneous provenance are prone to have mixtures of taxonomic treatments, queries should be performed to spot synonyms, invalid names, or other taxonomic problems. A major GBIF program (the Catalogue of Names) aims to organize taxonomic information in a dynamic, distributed system that should improve this situation considerably. GBIF should accelerate the process of providing users with updated, authoritative catalogues that will diminish number of false presences in GBIF-mediated data. Moreover, since numerous algorithms and techniques are being developed to facilitate automated data-cleaning (or at least error-flagging), the catalogues provided by GBIF should not consist of pdfs or other formats that only humans can interpret--rather, taxonomic information should have rigorously-defined ontologies, and should be provided in formats accessible to and readable by other algorithms.

Programs to correct and improve georeferences. ENM algorithms cannot perform without precise georeferences. Enriching primary biodiversity data sets with geographic coordinates of explicitly estimated precision should rank among the major pushes of GBIF. Certain GBIF data providers, like the HerpNet, MANIS, and ORNIS projects in North America, have developed detailed methodologies and protocols for

georeferencing records, thus creating databases rich in geographical coordinates, often at precisions of 1' or better. The North American networks have used a significant degree of community participation in georeferencing efforts, and have achieved georeferencing rates of as fast as 12 records per minute.

A major improvement in GBIF procedures would be to participate in a initiatives for adding and enriching georeferences for the ~25% of its data records lacking them. Such efforts would be particularly beneficial for specimen-based data records, which less frequently include quantitative georeferences. These steps could be achieved by communities, as MANIS did—and GBIF could potentially link community georeferencing with providing detailed national-level biodiversity information for countries wishing to enter the biodiversity informatics arena. Another extremely useful practice that GBIF might start would be providing automatic tools for data cleaning and error flagging, such as the tools already developed by CRIA in Brazil and CONABIO in Mexico.

Finally, many biodiversity data providers may be reticent about advertising the uneven quality of the taxonomic or geographic dimensions of the data they provide. In our experience, no museum in the world is free of this type of inconsistency. Indeed, several major institutions have databases plagued with inconsistent or incorrect records. We need to stress that: (i) this situation is the rule, not the exception—primary biodiversity databases need to be checked for inconsistencies prior to detailed analysis. (ii) These inconsistencies can be detected and either removed from analysis or problems can be corrected. (iii) Although no comprehensive assessments have been published, in our experience (mostly with terrestrial vertebrates), inconsistencies and blatant problems affect 5-20% of a typical database.

In conclusion, problems with misidentifications, problematic taxonomy and lacking, imprecise or faulty georeferencing do exist, and do affect a significant percentage of mixed-provenance databases, but these problems can be spotted, and data sufficient to perform useful analyses can generally be assembled. GBIF can and should perform a comprehensive set of assessment exercises to determine the quality of the data it serves.

VI. Mechanisms for Interactions with Established Data Providers in the Indicators Process

Background

The best means for improving interactions with the leaders of the indicators process is to clarify the sorts of data that each institution/initiative actually houses or serves. As discussed above, GBIF-mediated data are primary biodiversity data, and as such can serve many purposes at many scales of time and space.

The leaders in the indicators process have generally been conservation-oriented organizations. Key and leading examples include the World Conservation Monitoring Centre (WCMC)[14] and the International Union for the Conservation of Nature (IUCN) Species Survival Commission[15], as well as groups focused on specific taxa, such as BirdLife International[16], Butterfly Conservation Initiative[17], and Butterfly Conservation[18]. In each case, these organizations collect, develop, prepare, and serve valuable information regarding threat/endangerment levels, coarse-scale distributional patterns, taxonomy and classification, and (in some cases) compendia of priority areas.

This information is, nonetheless, secondary in nature—invariably, elements of expert opinion (e.g., in setting threat/endangerment levels) or subjectivity (e.g., in drawing range maps) have been involved in the process. This comment is not implying that the importance of this information is somehow lessened. Rather, it simply is an effort to categorize the types of information that are available in a useful manner.

Indicators towards the 2010 target used to date blur these distinctions. Aggregated indicators (Loh and Wackemagel 2004) pool secondary information from the literature. The Red List status-based indicators (Butchart et al. 2005; Butchart et al. 2004) have been based on secondary biodiversity information in the form of changes in status evaluations for species, with only birds and amphibians sufficiently data-rich to date to permit analysis. It is worth noting that these status-based indicators are only applicable to groups with a status summary available, and are difficult to scale down to be able to discern subregional or subnational trends.

Suggested Arrangements

The species-specific trend approach illustrated above could easily be partnered and integrated with status-based approaches to produce a more synthetic product. In this case, the GBIF-mediated data provide two key and novel elements to the indicator evaluation process:

- The potential for development of species-specific, biota-wide, multiscale evaluations of degree and pattern in changing conservation status and biodiversity loss. These computationally intensive, species-based indicators can greatly enrich the approaches already in use by the conservation initiatives by summarizing patterns of biodiversity loss over the past 2-3 decades.
- The deep-time baseline of the museum-based GBIF-enabled data provides a unique perspective on range loss in species. That is to say, the only source of a long-term perspective on biodiversity is that of looking back over the past 200-250 years via the specimen record,

---

[14] http://www.unep-wcmc.org/index.cfm.
[15] http://wwwiucn.org/themes/ssc.
[16] http://www.birdlife.org.
[17] http://www.butterflyrecovery.org.
[18] http://www.butterfly-conservation.org.

and using these sparse and difficult—but invaluable—data to reconstruct distributional patterns prior to extensive disturbance of natural habitats.

The problem in 'selling' this arrangement of mutual respect and dependence is quite simply that no worked examples exist based on the GBIF-mediated data sets. That is to say, most of the indicator analyses published to date are based on difficult-to-replicate aggregates or on status changes, and only small-scale prototypes have been based on primary biodiversity data (Peterson et al. 2006; Williams et al. 2005). Development of several large-scale exemplar applications would go a long way towards remedying this gap, and would certainly provide a platform for more informed and balanced discussions.
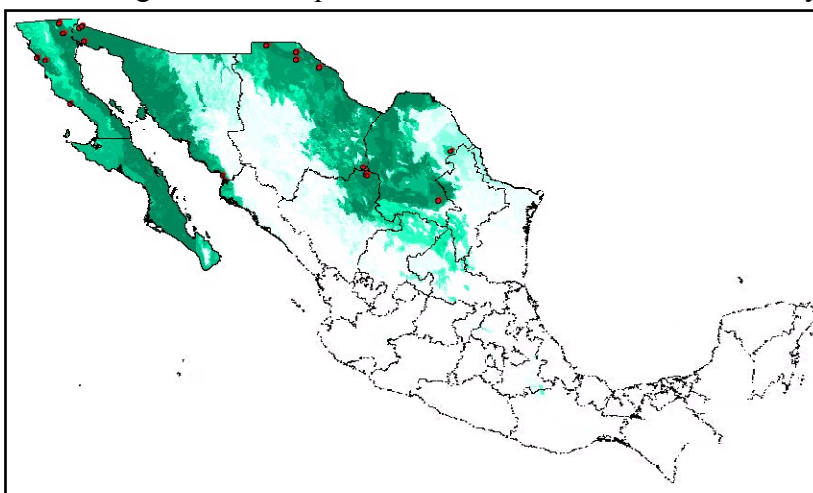
An ideal demonstration of the power of the primary-data-based approach would be that of implementing the trend analysis approach for a country that presently lacks a broad biodiversity information infrastructure (i.e., *not* Mexico or Costa Rica or Australia). Development *de novo* of a broad-scale biodiversity summary for a country for which rich primary data resources exist (e.g., via the GBIF network) but that has not seen intensive development would be an impressive demonstration of the power of these primary-data-based approaches. With such a product or products in hand, tabling a discussion of relative roles and merits will be much more feasible.

VII. Work Programs and Synergies with Other Conventions

The goals and objectives of the Convention on Biological Diversity extend considerably farther than the 2010 targets regarding biodiversity loss. Here, GBIF-mediated data resources become crucial, and secondary data resources are not able to contribute at all. Specific examples relate to invasive alien species, species useful to humans, and effects of global climate change on biodiversity.

The Program on Invasive Alien Species within the CBD framework emphasizes the need for information on and risk assessment regarding invasive alien species. A broad body of literature now documents the utility of primary biodiversity data such as that served through GBIF in assessing the geographic potential of invasive species (Arriaga et al. 2004; Beerling et al. 1995; Fonseca et al. 2006; Garnatje et al. 2002; Higgins et al. 1999; Hinojosa-Díaz et al. 2005; Hoffmann 2001; Honig et al. 1992; Iguchi et al. 2004; Martin 1996; Mohamed et al. 2006; Morrison et al. 2004; Nyari et al. 2006; Panetta and Dodd 1987; Papes and Peterson 2003; Peterson 2003a; Peterson et al. 2003a; Peterson and Robins 2003; Peterson et al. 2003b; Peterson and Vieglais 2001; Podger et al. 1990; Richardson and McMahon 1992; Scott and Panetta 1993; Sindel and Michael 1992; Skov 2000; Sutherst et al. 1999; Thuiller et al. 2005b; Thuiller et al. 2005c; Welk et al. 2002; Zalba et al. 2000). These methodologies are now quite well documented as to their ability

to anticipate the geographic potential of invasive species, and so this is clearly an issue to which GBIF and GBIF-mediated information can inform the CBD goals and targets. An exemplar worked case study is shown in the map at the right—occurrence data for *Tamarix ramossisima* were drawn from GBIF-mediated data resources for the species' native range in Eurasia; an ENM was developed, and projected onto Mexican landscapes. The color ramp is from white (predicted absence) to dark green (predicted potential for presence). The red circles that are overlain represent actual occurrences independent of the development of the model, drawn from Mexican national sources; this close correspondence between model prediction and independent test data corroborates the model predictions amply.

Similarly, the CBD's Program on Sustainable Use of Biodiversity places a premium on information relevant to the knowledge about and management of biological species of utility—economic, cultural, or otherwise—to humans. Here again, GBIF-mediated data sets can be enormously relevant. For example, the Guayacan tree (*Guaiacum sanctum*) is of enormous importance to Mexico for both economic and medicinal purposes. Its unknown population status and consequent inclusion on a CITES appendix placed these uses in doubt—primary biodiversity data obtained and analyzed by CONABIO previously clarified its status and distribution, and made clear that sustainable
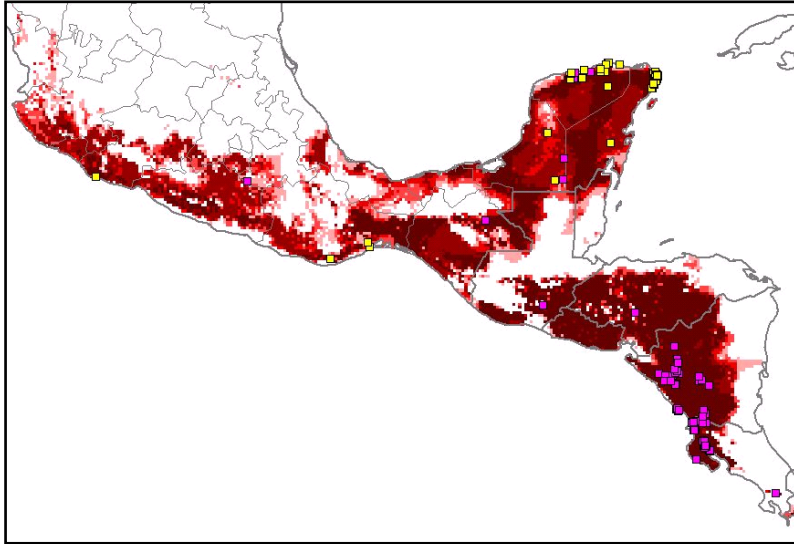
use of this species was feasible and did not endanger populations of the species. To demonstrate the utility of GBIF-mediated data for such questions, we developed a repeat analysis of the distribution of the species based only on GBIF-mediated data (see map at left, which shows GBIF-mediated records as pink squares, Mexican national data sets as yellow squares, and the ENM prediction as a color ramp – white = prediction of absence, dark red = prediction of potential for occurrence)—the result strongly corroborated the ability of ENMs based on data available through GBIF to anticipate the spatial distribution of richer and more detailed data sets.

Finally, the U.N. Framework Convention on Climate Change places emphasis on understanding the effects of ongoing climate change processes on biodiversity. Several summaries have been published laying out likely implications for biodiversity (Chapin et al. 2000; Lovejoy and Hannah 2005), and real-time consequences are increasingly apparent (Allen and Breshears 1998; Brown et al. 1997; Brown et al. 1999; Cresswell and McCleery 2003; Crozier 2003; Inouye et al. 2000; Parmesan 1996; Parmesan et al. 1999; Parmesan and Yohe 2003). Nonetheless, a predictive, proactive, forecasting capability has nonetheless been lacking—particularly one in which individual species' responses can be tracked and anticipated.

The ENM approach, again based on primary biodiversity data such as those served by GBIF, offers an attractive possibility here (Peterson et al. 2005). ENMs are trained based on present-day climates and conditions, and then projected onto modeled future projections of climatic conditions; where possible, these projections have been tested and their forecasts have largely been corroborated (Araújo et al. 2005a). These
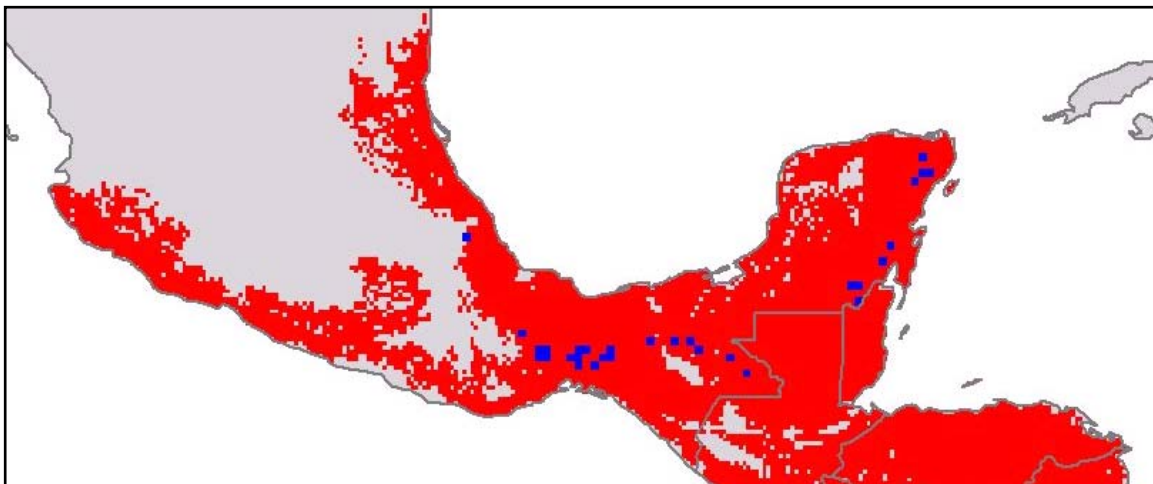
forecasts can be developed for individual species with little difficulty—we have developed the example for *Tamandua mexicana* for the sake of illustration (see map below—red indicates the likely 1940 range of the species; blue indicates areas in 2000 in primary forest habitat that are projected to *remain* habitable for the species through 2050). The comparison between present and future potential distributions of the species can then be evaluated taking into account different scenarios of dispersal ability (Peterson et al. 2002b). These future-climate forecasts can then be examined across many species to see regional trends, or attention can focus on single species of particular interest or concern. As such, the ENM approach based on GBIF-mediated data sets has much to offer to an understanding of climate change consequences for natural systems.

VIII. Conclusions

In this report, we demonstrate the utility of primary biodiversity data sets in developing indices to assess the *"current rate of biodiversity loss at the global, regional and national level,"* one of the primary mandates of the CBD. Several conclusions can be derived from our work:

1) Primary data, as served by GBIF in quantities larger, by orders of magnitude, than any other provider, can be used to obtain effective indices of biodiversity loss.
2) Such indices can be naturally and easily scaled-down to levels of regions, countries, and areas within countries, and can also be based on species selected by countries on the basis of any of the three objectives of the convention.
3) Two types of data are required to calculate the indices; first, the primary biodiversity data served through GBIF by an increasing number of providers, and second, remotely sensed information for summarizing changes in land use. Both kinds of data are publicly available and their quality, resolution and availability are increasing.
4) The method we describe to obtain the basic input for the indices, namely, areas of distribution of different species, is now well established, with numerous research groups in France, Switzerland, Spain, the United Kingdom, the United States, Mexico, Brazil, India, and other countries applying them regularly and with dozens of publications in major journals describing their possibilities and limitations.
5) The basic skills and training required to use the method can be taught in short courses, like the ones GBIF has supported for the last two years.
6) All of the above constitute a major effort in making taxonomic knowledge available, in capacity building and technology transfer to developing countries. We do not think that the different bodies of the CBD are aware of the extent to which GBIF has helped participating countries to advance CBD decisions and programmes of work.
7) We suggest that GBIF begin supporting GBIF-data-based trend evaluations for a country or set of countries, using only GBIF data and a large group of species, in collaboration with local institutions--such demonstrative studies should be communicated to SBSTTA as an example of how countries can use primary biodiversity data for the purposes of the CBD.

IX. Literature Cited

Allen, C. D., and D. D. Breshears. 1998. Drought-induced shift of a forest-woodland ecotone: Rapid landscape response to climate variation. Proceedings of the National Academy of Sciences USA 95:14839-14842.

Anciães, M., and A. T. Peterson. 2007. Climate change effects on Neotropical manakin diversity based on ecological niche modeling. Condor

Anderson, R. P., D. Lew, and A. T. Peterson. 2003. Evaluating predictive models of species' distributions: Criteria for selecting optimal models. Ecological Modelling 162:211-232.

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

Araújo, M. B., R. G. Pearson, W. Thuiller, and M. Erhard. 2005a. Validation of species-climate impact models under climate change. Global Change Biology 11:1504-1513.

Araújo, M. B., W. Thuiller, and R. G. Pearson. 2006. Climate warming and the decline of amphibians and reptiles in Europe. Journal of Biogeography 33:1712-1728.

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.

Argaez, J. A., A. Christen, M. Nakamura, and J. Soberón. 2005. Prediction of potential areas of species distributions based on presence-only data. Environmental and Ecological Statistics 12:27-44.

Arriaga, L., A. E. Castellanos-V., 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.

Bakkenes, M., J. R. M. Alkemade, F. Ihle, R. Leemansand, and J. B. Latour. 2002. Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. Global Change Biology 8:390-407.

Beerling, D. J., B. Huntley, and J. P. Bailey. 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.

Berry, P. M., T. P. Dawson, P. A. Harrison, and R. G. Pearson. 2002a. Modelling potential impacts of climate change on the bioclimatic envelope of species in Britain and Ireland. Global Ecology and Biogeography 11:453-462.

Berry, P. M., T. P. Dawson, P. A. Harrison, and R. G. Pearson. 2002b. Modelling potential impacts of climate change on the bioclimatic envelope of species in Britain and Ireland. Global Ecology and Biogeography 11:453-462.

Brown, J. H., T. J. Valone, and C. G. Curtin. 1997. Reorganization of an arid ecosystem in response to recent climate change. Proceedings of the National Academy of Sciences USA 94:9729-9733.

Brown, J. L., S. H. Li, and N. Bhagabati. 1999. Long-term trend toward earlier breeding in an American bird: A response to global warming? Proceedings of the National Academy of Sciences USA 96:5565-5569.

Butchart, S., A. Stattersfield, J. Baillie, L. Bennun, S. Stuart, H. Akçakaya, C. Hilton-Taylor, and G. Mace. 2005. Using Red List Indices to measure progress towards

the 2010 target and beyond. Philosophical Transactions of the Royal Society B: Biological Sciences 360:255-268.

Butchart, S. H. M., A. J. Stattersfield, L. A. Bennun, S. M. Shutes, H. Akçakaya, J. E. M. Baillie, S. N. Stuart, C. Hilton-Taylor, and G. M. Mace. 2004. Measuring global trends in the status of biodiversity: Red List indices for birds. PLoS Biology 2:e383.

Canhos, V. P., S. de Souza, R. de Giovanni, and D. A. L. Canhos. 2004. Global biodiversity informatics: Setting the scene for a "new world" of ecological forecasting. Biodiversity Informatics 1:1-13.

Chapin, F. S. I., E. S. Zavaleta, V. T. Eviner, R. L. Naylor, P. M. Vitousek, H. L. Reynolds, D. U. Hooper, S. Lavorel, O. E. Sala, S. E. Hobbie, M. C. Mack, and S. Diaz. 2000. Consequences of changing biodiversity. Nature 405:234-242.

Chapman, A. D. 1999. Quality control and validation of point-sourced environmental resource data. Pp. 409-418 *in* K. Lowell, ed. Spatial accuracy assessment: Land information uncertainty in natural resources. Ann Arbor Press, Chelsea, Michigan.

Chapman, A. D. 2005. Principles of Data Quality. Global Biodiversity Information Facility, Copenhagen.

Chapman, A. D., M. E. S. Muñoz, and I. Koch. 2005. Environmental information: Placing biodiversity phenomena in an ecological and environmental context. Biodiversity Informatics 2:24-41.

Cresswell, W., and R. McCleery. 2003. How great tits maintain synchronization of their hatch date with food supply in response to long-term variability in temperature. Journal of Animal Ecology 72:356-366.

Crozier, L. 2003. Winter warming facilitates range expansion: cold tolerance of the butterfly Atalopedes campestris. Oecologia 135:648-656.

Elith, J., and M. Burgman. 2002. Predictions and their validation: Rare plants in the Central Highlands, Victoria. Pp. 303-313 *in* J. M. Scott, P. J. Heglund and M. L. Morrison, eds. Predicting Species Occurrences: Issues of Scale and Accuracy. Island Press, Washington, D.C.

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

Erasmus, B. F. N., A. S. Van Jaarsveld, S. L. Chown, M. Kshatriya, and K. J. Wessels. 2002. Vulnerability of South African animal taxa to climate change. Global Change Biology 8:679-693.

Feria, T., and A. 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.

Fonseca, R. L., P. R. Guimarães, Jr, S. R. Morbiolo, R. Scachetti-Pereira, and A. T. Peterson. 2006. Vulnerability of Brazilian national parks to invasion by the alien weed *Crotalaria pallida* Ait. (Fabaceae). Weed Science

Garnatje, T., R. Vilatersana, C. T. Roché, N. García-Jacas, A. Susanna, and D. C. Thill. 2002. Multiple introductions from the Iberian Peninsula are responsible for invasion of *Crupina vulgaris* in western North America. New Phytologist 154:419-428.

Grinnell, J. 1917. Field tests of theories concerning distributional control. American Naturalist 51:115-128.

Grinnell, J. 1924. Geography and evolution. Ecology 5:225-229.

Guisan, A., C. H. Graham, J. Elith, F. Huettman, and NCEAS Species Distribution Modelling Group. 2007. Sensitivity of predictive species distribution models to change in grain size. Diversity and Distributions

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

Higgins, S. I., D. M. Richardson, R. M. Cowling, and T. H. Trinder-Smith. 1999. Predicting the landscape-scale distribution of alien plants and their threat to plant diversity. Conservation Biology 13:303-313.

Hinojosa-Díaz, I. A., O. Yáñez-Ordóñez, G. Chen, and A. T. Peterson. 2005. The North American invasion of the Giant Resin Bee (Hymenoptera: Megachilidae). Journal of Hymenoptera Research 14:69-77.

Hoffmann, M. H. 2001. The distribution of *Senecio vulgaris*: Capacity of climatic range models for predicting adventitious ranges. Flora 196/5:395-403.

Honig, M. A., R. M. Cowling, and D. M. Richardson. 1992. The invasive potential of Australian banksias in South-African fynbos--A comparison of the reproductive potential of *Banksia ericifolia* and *Leucadendron laureolum*. Australian Journal of Ecology 17:305-314.

Huntley, B., P. M. Berry, W. Cramer, and A. P. McDonald. 1995. Modelling present and potential future ranges of some European higher plants using climate response surfaces. Journal of Biogeography 22:967-1001.

Iguchi, K., K. Matsuura, K. McNyset, A. T. Peterson, R. Scachetti-Pereira, K. A. Powers, D. A. Vieglais, E. O. Wiley, and T. Yodo. 2004. Predicting invasions of North American basses in Japan using native range data and a genetic algorithm. Transactions of the American Fisheries Society 133:845-854.

Illoldi, P., V. Sánchez-Cordero, and A. T. Peterson. 2004. Predicting distributions of Mexican mammals using ecological niche modeling. Journal of Mammalogy 85:658-662.

Inouye, D. W., B. Barr, K. B. Armitage, and B. D. Inouye. 2000. Climate change is affecting altitudinal migrants and hibernating species. Proceedings of the National Academy of Sciences USA 97:1630-1633.

Kadmon, R., and J. Heller. 1998. Modelling faunal responses to climatic gradients with GIS: Land snails as a case study. Journal of Biogeography 25:527-539.

Krishtalka, L., and P. S. Humphrey. 2000. Can natural history museums capture the future? BioScience 50:611-617.

Lehmann, A., J. Leathwick, and J. Overton. 2002. Assessing New Zealand fern diversity from spatial predictions of species assemblages. Biodiversity and Conservation 11:2217-2238.

Loh, J., R. Green, T. Ricketts, J. Lamoreux, M. Jenkins, V. Kapos, and J. Randers. 2005. The Living Planet Index: using species population time series to track trends in

biodiversity. Philosophical Transactions of the Royal Society B: Biological Sciences 360:289-295.

Loh, J., and M. Wackemagel. 2004. Living Planet Report 2004. World Wide Fund for Nature, Gland, Switzerland.

Lovejoy, T. E., and L. Hannah. 2005. Climate Change and Biodiversity. Yale University Press, New Haven, Conn.

Martin, W. K. 1996. The current and potential distribution of the common myna (*Acridotheres tristis*) in Australia. Emu 96:166-173.

Martínez-Meyer, E. 2002. Evolutionary Trends in Ecological Niches of Species. Ph.D. thesis, Department of Geography, University of Kansas, Lawrence, Kansas.

Martínez-Meyer, E., and A. T. Peterson. 2006. Conservatism of ecological niche characteristics in North American plant species over the Pleistocene-to-Recent transition. Journal of Biogeography 33:1779-1789.

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.

McNyset, K. 2005. Use of ecological niche modelling to predict distributions of freshwater fish species in Kansas. Ecology of Freshwater Fish 14:243-255.

Midgely, G. E., L. Hannah, D. Millar, M. C. Rutherford, and L. W. Powrie. 2002. Assessing the vulnerability of species richness to anthropogenic climate change in a biodiversity hotspot. Global Ecology and Biogeography 11:445-451.

Midgley, G. F., L. Hannah, D. Millar, W. Thuiller, and A. Booth. 2003. Developing regional and species-level assessments of climate change impacts on biodiversity in the Cape Floristic Region. Biological Conservation 112:87-97.

Mohamed, K. I., M. Papeş, R. Williams, B. W. Benz, and A. T. Peterson. 2006. Global invasive potential of ten parasitic witchweeds and related Orobanchaceae. Ambio

Morrison, L. W., S. D. Porter, E. Daniels, and M. D. Korzukhin. 2004. Potential global range expansion of the invasive fire ant, *Solenopsis invicta*. Biological Invasions 6:183-191.

Nyari, A., C. Ryall, and A. T. Peterson. 2006. Global invasive potential of the House Crow (*Corvus splendens*) based on ecological niche modelling. Journal of Avian Biology

Panetta, F. D., and J. Dodd. 1987. Bioclimatic prediction of the potential distribution of skeleton weed *Chondrilla juncea* L. in Western Australia. Journal of the Australian Institute of Agricultural Science 53:11-16.

Papes, M., and A. T. Peterson. 2003. Predicting the potential invasive distribution for *Eupatorium adenophorum* Spreng. in China. Journal of Wuhan Botanical Research 21:137-142.

Parmesan, C. 1996. Climate and species' range. Nature 382:765-766.

Parmesan, C., N. Ryrholm, C. Stefanescu, J. K. Hill, C. D. Thomas, H. Descimon, B. Huntley, L. Kaila, J. Kullberg, T. Tammaru, J. Tennent, J. A. Thomas, and M. Warren. 1999. Poleward shift of butterfly species' ranges associated with regional warming. Nature 399:579-583.

Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421:37-42.

Parra-Olea, G., E. Martinez-Meyer, and G. de Leon. 2005. Forecasting climate change effects on salamander distribution in the highlands of central Mexico. BIOTROPICA 37:202-208.

Pearson, R. G., and T. P. Dawson. 2003. Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? Global Ecology and Biogeography 12:361-371.

Pearson, R. G., T. P. Dawson, P. M. Berry, and P. A. Harrison. 2002. SPECIES: A spatial evaluation of climate impact on the envelope of species. Ecological Modelling 154:289-300.

Pearson, R. G., M. Nakamura, A. T. Peterson, and C. Raxworthy. 2006a. Predicting species' distributions from small numbers of occurrence records: A test case using cryptic geckos in Madagascar. Journal of Biogeography

Pearson, R. G., W. Thuiller, M. B. Araujo, E. Martinez-Meyer, L. Brotons, C. McClean, L. Miles, P. Segurado, T. P. Dawson, and D. C. Lees. 2006b. Model-based uncertainty in species range prediction. Journal of Biogeography 33:1704-1711.

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

Peterson, A. T. 2003a. Predicting the geography of species' invasions via ecological niche modeling. Quarterly Review of Biology 78:419-433.

Peterson, A. T. 2003b. Projected climate change effects on Rocky Mountain and Great Plains birds: Generalities of biodiversity consequences. Global Change Biology 9:647-655.

Peterson, A. T., L. G. Ball, and K. C. Cohoon. 2002a. Predicting distributions of Mexican birds using ecological niche modelling methods. Ibis 144:e27-e32.

Peterson, A. T., E. Martínez-Meyer, C. González-Salazar, and P. Hall. 2004. Modeled climate change effects on distributions of Canadian butterfly species. Canadian Journal of Zoology 82:851-858.

Peterson, A. T., M. A. Ortega-Huerta, J. Bartley, V. Sanchez-Cordero, J. Soberon, R. H. Buddemeier, and D. R. B. Stockwell. 2002b. Future projections for Mexican faunas under global climate change scenarios. Nature 416:626-629.

Peterson, A. T., M. Papes, and D. A. Kluza. 2003a. Predicting the potential invasive distributions of four alien plant species in North America. Weed Science 51:863-868.

Peterson, A. T., and C. R. Robins. 2003. Using ecological-niche modeling to predict Barred Owl invasions with implications for Spotted Owl conservation. Conservation Biology 17:1161-1165.

Peterson, A. T., V. Sánchez-Cordero, E. Martínez-Meyer, and A. G. Navarro-Sigüenza. 2006. Tracking population extirpations via melding ecological niche modeling with land-cover information. Ecological Modelling 195:229-236.

Peterson, A. T., V. Sanchez-Cordero, J. Soberon, J. Bartley, R. H. Buddemeier, and A. G. Navarro-Siguenza. 2001. Effects of global climate change on geographic distributions of Mexican Cracidae. Ecological Modelling 144:21-30.

Peterson, A. T., R. Scachetti-Pereira, and D. A. Kluza. 2003b. Assessment of invasive potential of *Homalodisca coagulata* in western North America and South America. Biota Neotropica 3:Online journal: http://www.biotaneotropica.org.br/v3n1/pt/abstract?article+BN00703012003.

Peterson, A. T., H. Tian, E. Martínez-Meyer, J. Soberón, V. Sánchez-Cordero, and B. Huntley. 2005. Modeling distributional shifts of individual species and biomes. Pp. 211-228 *in* T. E. Lovejoy and L. Hannah, eds. Climate Change and Biodiversity. Yale University Press, New Haven, Conn.

Peterson, A. T., and D. A. Vieglais. 2001. Predicting species invasions using ecological niche modeling. BioScience 51:363-371.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231-259.

Phillips, S. J., M. Dudik, and R. E. Schapire. 2004. A maximum entropy approach to species distribution modeling. Proceedings of the 21st International Conference on Machine Learning

Podger, F. D., D. C. Mummery, C. R. Palzer, and M. J. Brown. 1990. Bioclimatic analysis of the distribution of damage to native plants in Tasmania by *Phytophthora cinnamomi*. Australian Journal of Ecology 15:281-290.

Raxworthy, C. J., E. Martínez-Meyer, N. Horning, R. A. Nussbaum, G. E. Schneider, M. A. Ortega-Huerta, and A. T. Peterson. 2003. Predicting distributions of known and unknown reptile species in Madagascar. Nature 426:837-841.

Richardson, D. M., and J. P. McMahon. 1992. A bioclimatic analysis of *Eucalyptus nintens* to identify potential planting regions in Southern Africa. South African Journal of Science 88:380-387.

Roura-Pascual, N., A. Suarez, C. Gómez, P. Pons, Y. Touyama, A. L. Wild, and A. T. Peterson. 2005. Geographic potential of Argentine ants (*Linepithema humile* Mayr) in the face of global climate change. Proceedings of the Royal Society of London B 271:2527-2535.

Sánchez-Cordero, V., P. Illoldi-Rangel, M. Linaje, S. Sarkar, and A. T. Peterson. 2005. Deforestation and extant distributions of Mexican endemic mammals. Biological Conservation 126:465-473.

Scott, J. K., and F. D. Panetta. 1993. Predicting the Australian weed status of southern African plants. Journal of Biogeography 20:87-93.

Sindel, B. M., and P. W. Michael. 1992. Spread and potential distribution of *Senecio madagascariensis pior* (Fireweed) in Australia. Australian Journal of Ecology 17:21-26.

Siqueira, M. F. d., and A. T. Peterson. 2003. Global climate change consequences for cerrado tree species. Biota Neotropica In press

Skov, F. 2000. Potential plant distribution mapping based on climatic similarity. Taxon 49:503-515.

Smith, F. D. M., R. M. May, R. Pellew, T. H. Johnson, and K. S. Walter. 1993. Estimating extinction rates. Nature 364:494-496.

Soberón, J., J. E. Llorente B., and H. Benitez-Díaz. 1996. An international view of national biological surveys. Annals of the Missouri Botanical Garden 83:562-573.

Soberón, J., and A. T. Peterson. 2004. Biodiversity informatics: Managing and applying primary biodiversity data. Philosophical Transactions of the Royal Society of London B 359:689-698.

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

Stockwell, D. R. B., and D. P. Peters. 1999. The GARP modelling system: Problems and solutions to automated spatial prediction. International Journal of Geographic Information Systems 13:143-158.

Stouffer, P. C., R. O. Bierregaard, C. Strong, and T. E. Lovejoy. 2006. Long-Term Landscape Change and Bird Abundance in Amazonian Rainforest Fragments. Conservation Biology 20:1212-1223.

Stuart, S. N., J. S. Chanson, N. A. Cox, B. E. Young, A. S. L. Rodrigues, D. L. Fischman, and R. W. Waller. 2004. Status and trends of amphibian declines and extinctions worldwide. Science 306:1783-1786.

Sutherst, R. W., G. F. Maywald, T. Yonow, and P. M. Stevens. 1999. CLIMEX User Guide--Predicting the Effects of Climate on Plants and Animals. CSIRO Publishing, Collingwood, Victoria, Australia.

Svenning, J.-C., and F. Skov. 2004. Limited filling of the potential range in European tree species. Ecology Letters 7:565-573.

Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. N. Erasmus, M. Ferreira de Siqueira, A. Grainger, L. Hannah, L. Hughes, B. Huntley, A. S. Van Jaarsveld, G. E. Midgely, L. Miles, M. A. Ortega-Huerta, A. T. Peterson, O. L. Phillips, and S. E. Williams. 2004. Extinction risk from climate change. Nature 427:145-148.

Thuiller, W., M. B. Araújo, and S. Lavorel. 2004. Do we need land-cover data to model species distributions in Europe? Journal of Biogeography 31:353-361.

Thuiller, W., S. Lavorel, M. B. Araújo, M. T. Sykes, and I. C. Prentice. 2005a. Climate change threats to plant diversity in Europe. Proceedings of the National Academy of Sciences USA 102:8245-8250.

Thuiller, W., G. F. Midgely, G. O. Hughes, B. Bomhard, G. Drew, M. C. Rutherford, and F. I. Woodward. 2006. Endemic species and ecosystem sensitivity to climate change in Namibia. Global Change Biology 12:759-776.

Thuiller, W., D. M. Richardson, P. Pysek, G. F. Midgely, G. O. Hughes, and M. Rouget. 2005b. Global risk assessment for plant invasions--The role of climatic suitability and propagule pressure. Global Change Biology 11:2234-2259.

Thuiller, W., D. M. Richardson, P. Pysek, G. F. Midgley, G. O. Hughes, and M. Rouget. 2005c. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. Global Change Biology 11:2234-2250.

UNEP-WCMC. 2006. Identifying How the Global Biodiversity Information Facility (GBIF) Can Improve the 2010 Indicators at a Global Level. World Conservation Monitoring Centre, Cambridge.

Welk, E., K. Schubert, and M. H. Hoffmann. 2002. Present and potential distribution of invasive garlic mustard (*Alliaria petiolata*) in North America. Diversity and Distributions 8:219-233.

Wiley, E. O., K. M. McNyset, A. T. Peterson, C. R. Robins, and A. M. Stewart. 2003. Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. Oceanography 16:120-127.

Williams, P., L. E. E. Hannah, S. Andelman, G. U. Y. Midgley, M. Araujo, G. Hughes, L. Manne, E. Martinez-Meyer, and R. Pearson. 2005. Planning for Climate Change: Identifying Minimum-Dispersal Corridors for the Cape Proteaceae. Conservation Biology 19:1063-1074.

Zalba, S. M., M. I. Sonaglioni, and C. J. Belenguer. 2000. Using a habitat model to assess the risk of invasion by an exotic plant. Biological Conservation 93:203-208.