

# Progress in Physical Geography

<http://ppg.sagepub.com/>

---

## **Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models**

Tim Newbold

*Progress in Physical Geography* 2010 34: 3

DOI: 10.1177/0309133309355630

The online version of this article can be found at:

<http://ppg.sagepub.com/content/34/1/3>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Progress in Physical Geography* can be found at:**

**Email Alerts:** <http://ppg.sagepub.com/cgi/alerts>

**Subscriptions:** <http://ppg.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations:** <http://ppg.sagepub.com/content/34/1/3.refs.html>



# Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models

**Tim Newbold**

*University of Nottingham, UK*

Progress in Physical Geography

34(1) 3–22

© The Author(s) 2010

Reprints and permission:

[sagepub.co.uk/journalsPermissions.nav](http://sagepub.co.uk/journalsPermissions.nav)

DOI: 10.1177/0309133309355630

[ppg.sagepub.com](http://ppg.sagepub.com)



## Abstract

To conserve biodiversity, it is necessary to understand how species are distributed and which aspects of the environment determine distributions. In large parts of the world and for the majority of species, data describing distributions are very scarce. Museums, private collections and the historical literature offer a vast source of information on distributions. Records of the occurrence of species from these sources are increasingly being captured in electronic databases and made available over the internet. These records may be very valuable in conservation efforts. However, there are a number of limitations with museum data. These limitations are dealt with in the first part of this review. Even if the limitations of museum data can be overcome, these data present a far-from-complete picture of the distributions of species. Species distribution models offer a means to extrapolate limited information in order to estimate the distributions of species over large areas. The second part of this paper reviews the challenges of developing species distribution models for use with museum data and describes some of the questions that species distribution models have been used to address. Given the rapidly increasing number of museum records of species occurrence available over the internet, a review of their usefulness in conservation and ecology is timely.

## Keywords

ecological niche modelling, georeferencing, GIS, museum data, natural history collections, species distribution modelling

## 1 Introduction

Museums, private natural history collections, herbaria and the historical literature contain a wealth of information on the distributions of species in the form of recorded occurrences of species (hereafter ‘museum data’ or ‘museum records’). In recent years, more and more of these data have been captured in electronic databases and made available through data portals on the internet (for a list of some of the major databases available, see Graham *et al.*, 2004),

although a vast number of records still remain to be computerized (O’Connell *et al.*, 2004). The largest of these portals, and the most wide-ranging in geographical scope, is the Global Biodiversity Information Facility (<http://www.gbif.org>), which at the time of writing contained

---

### Corresponding author:

School of Biology, University of Nottingham, Nottingham NG7 2RD, UK

Email: [plxtn@nottingham.ac.uk](mailto:plxtn@nottingham.ac.uk)



**Figure 1.** Examples of specimen material from museums, private collections and herbaria. From left to right: bottled specimens of Egyptian reptiles in the field museum in Chicago (photo: Sherif Baha El Din); butterfly specimen from the Egyptian Ministry of Agriculture, with label (photo: butterfly – Torben Larsen; label – Mahmoud Saleh Abdel Dayem); herbarium sheet from the Alexandria Natural History Society (photo: Adel Gazzar).

177,448,319 species occurrence records from all over the world. These records show patchy spatial coverage and some countries have much better data than others. This is reflected in the existence of several national databases, such as the National Biodiversity Gateway (NBN) for the United Kingdom (<http://www.nbn.org.uk>), which contains 40,397,129 records, and Egypt's BioMAP project. The BioMAP project (<http://www.biomapegypt.org>), which ran between 2004 and 2007, collected records of Egyptian species from museums, private collections and the literature in an electronic database. By the end of the project, the database contained about 500,000 records for mammals, butterflies, reptiles, amphibians, aquatic invertebrates and plants.

Museum data can be of enormous value to conservation biologists and ecologists for studying the distribution and abundance of species. However, there are a number of limitations of museum data in this context. Of particular concern are errors in the records, and bias in their scope (reviewed previously in Graham *et al.*, 2004). In this paper, I revisit these limitations, with reference to examples from the GBIF database and also from Egypt's BioMAP database. In the second part of the review, I deal with some of the applications of museum data in

conservation ecology. In particular, I focus on one method for filling gaps in our knowledge about the distributions of species: species distribution modelling. Most previous reviews of species distribution modelling have focused on technical aspects of their use (eg, Pearson and Dawson, 2003; Wintle *et al.*, 2005; Araújo and Guisan, 2006; Hirzel and Le Lay, 2008). Here, I focus on the challenges of developing distribution models using museum data and review some of the applications of distribution models, focusing on studies that have used museum data.

## II Limitations of museum data

### I Errors

The accuracy of museum data is often uncertain. There may be errors in the identification of species or in the location of records (reviewed in Graham *et al.*, 2004). Errors in the identification of species can only be detected by very careful checking of all records. One of the main advantages of using records from museums, private collections and herbaria is that the original material (see Figure 1) can be re-examined to check the species identification and to update the identification in the event that accepted taxonomies have changed (Graham *et al.*, 2004; O'Connell

*et al.*, 2004). An additional complication in the identification of species is that accepted taxonomies are constantly being updated, with the result that different names may be given to the same species depending on when the record was made. Therefore, it is necessary to check lists of synonyms to ensure that all records are assigned to the correct species. For some taxonomic groups, such as mammals, this information is readily available through global species lists (Wilson and Reeder, 2005). For other groups, the information is scattered throughout several different works or is not available at all.

Museum records are often accompanied by a textual description of the locality from which they were taken, although coordinates are occasionally given. Assigning geographical coordinates to records, a process referred to as 'georeferencing', is subject to a number of possible errors and uncertainties (Wieczorek *et al.*, 2004). First, textual descriptions may refer to anything from very precise locations to very broad areas. For example, some locations in the BioMAP database have very accurate descriptions (eg, 'St. Katherine's Monastery'), while other descriptions are very vague (eg, 'Egypt'). The coordinates assigned to vague descriptions will obviously be much less accurate. Second, where the description of a location includes an offset (eg, '5 km northwest of St. Katherine's Monastery'), there will be uncertainty over the accuracy of the measurements of distance and direction. Third, where coordinates have been provided with a record, uncertainty will be generated if the corresponding coordinate system has not also been given. There are many different coordinate systems in use, each of which gives slightly different coordinates for a location. Finally, where records are georeferenced using a map, the accuracy of the map will affect the accuracy of the resulting coordinates. Wieczorek *et al.* (2004) present a method for combining these various sources of uncertainty to calculate the overall uncertainty in each georeferenced location. Their method assumes that

uncertainty is equal in all directions around a point and that the true locality has an equal chance of falling anywhere within a sphere of a specified radius around the given coordinates. However, some of the uncertainties operate only in one direction and the true coordinates will be more likely to fall in some locations around the assigned coordinates than in others, meaning that Wieczorek *et al.*'s (2004) method overestimates the actual degree of uncertainty. Guo *et al.* (2008) present a new method that accounts for this directionality and describes the uncertainty around assigned coordinates as a probability density function.

Additional problems are generated when records come from countries whose languages use different alphabets to the language in which the records are made (usually the Latin alphabet). Site names are often transliterated into the Latin alphabet at the time of collection. However, different collectors may produce very different spellings for the same site. This can make it very difficult to identify all records that came from the same site (McGowan and Gillman, 1997). As part of the BioMAP project, a gazetteer was developed for Egypt listing coordinates for all of the collecting sites used with the museum records along with all known transliterations, from Arabic to English, for these sites. This enabled records to be matched to sites with a reasonably high degree of confidence. This was actually the hardest and most time-consuming part of the entire project.

## 2 Bias

Another major limitation of museum data is that they are often biased (Graham *et al.*, 2004). Bias in the records may be of four types: spatial, environmental, temporal and taxonomic (Soberón *et al.*, 2000).

Historical sampling of species has clearly been biased spatially and there are still major gaps in our knowledge. Global sampling has been particularly poor in tropical areas and in

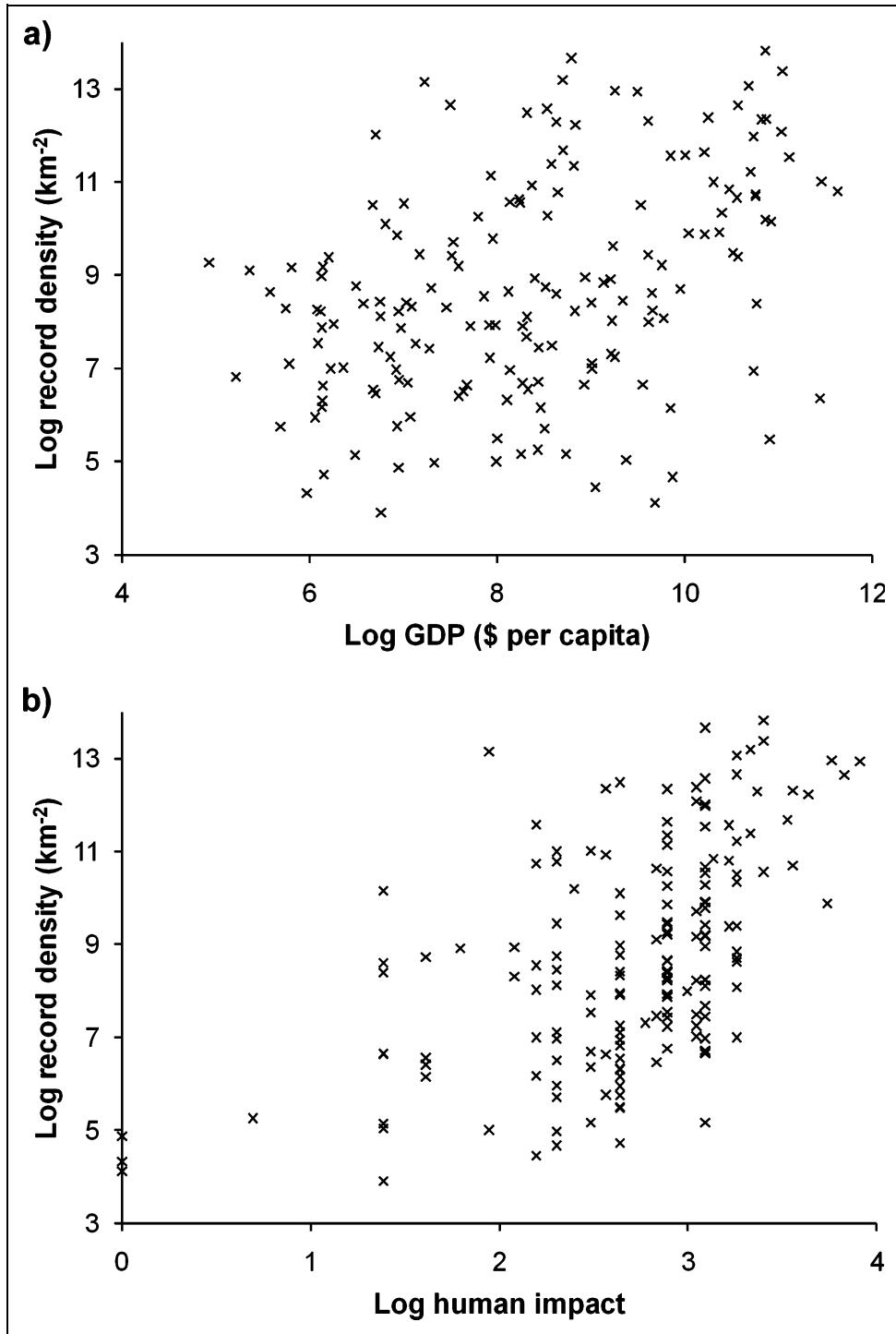
arid environments (Stockwell and Peters, 1999; Anderson *et al.*, 2002a; Soria-Auza and Kessler, 2008). To test for spatial bias in records from GBIF, I calculated the density of records for each country with records in the database. Record density was then analysed with respect to two variables: an index of human impact and per capita gross domestic product (GDP). The index of human impact was derived from the Last of the Wild project's 1 km resolution map of human impact (Sanderson *et al.*, 2002); for each country, the median value across all grid squares was taken. Values of record density, GDP and human impact were all log-transformed to meet assumptions of normality made by the statistical tests used.

If the records were completely unbiased, I should expect the number of records per country to show no relationship with either GDP or human impact. In reality, the number of records showed a strong positive relationship with both GDP (ANOVA:  $F_{1, 167} = 29.44$ ,  $P < 0.001$ ; Figure 2a) and median human impact index values ( $F_{1, 167} = 66.83$ ,  $P < 0.001$ ; Figure 2b), suggesting that collecting has been biased towards more developed countries. GDP and human impact did not correlate strongly with one another (Spearman's rank correlation:  $r_s = 0.2$ ).

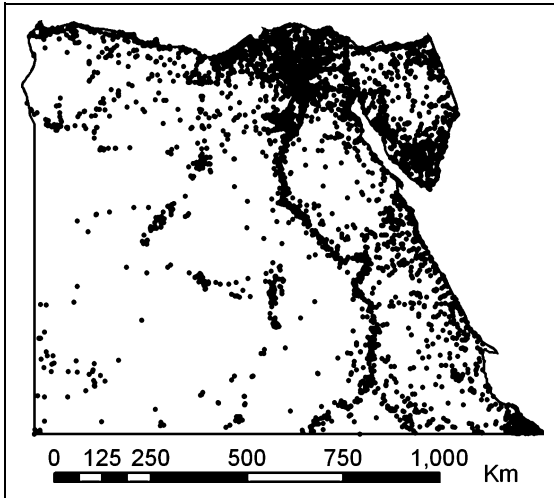
At the level of individual countries, even when sampling has been extensive, gaps and biases are often apparent in the spatial coverage of species records. Several studies on different taxonomic groups and in different countries have noted unevenness in the spatial coverage of sampling (Austin and Meyers, 1996; Peterson *et al.*, 1998; Dennis and Hardy, 1999; Dennis *et al.*, 1999; Hijmans *et al.*, 2000; Soberón *et al.*, 2000; Reddy and Dávalos, 2003; Kadmon *et al.*, 2004; Hortal *et al.*, 2008; Sánchez-Fernández *et al.*, 2008). Records are often closer to roads, rivers, coasts, towns and cities than they would be if sampling were completely random (Hijmans *et al.*, 2000; Soberón *et al.*, 2000; Reddy and Dávalos, 2003; Kadmon *et al.*, 2004). Furthermore, sampling tends to be close to the homes of active collectors (Dennis and Thomas,

2000) or concentrated in areas that are of more interest to collectors, such as protected areas (Reddy and Dávalos, 2003) or hotspots of diversity (Dennis and Thomas, 2000). Localities with records in the BioMAP database show reasonably good coverage of Egypt's area. However, there are important gaps in collecting, especially in the Western Desert, but also in parts of the Eastern Desert and in the northern part of the Sinai Peninsula (Figure 3). Of course, dot maps only show visited locations where organisms were recorded, and not locations visited but which yielded no records. Large areas of the Western Desert of Egypt make up the Great Sand Sea, not an environment from which many records are to be expected, nor an area to which one can anticipate many visitors. Nevertheless, there is clearly a need for more systematic sampling in the future in all countries, focusing on areas that have been overlooked in the past, including the less diverse places outside protected areas.

Many applications of museum data involve analyses of the environment that species inhabit. For these applications, gaps in the spatial coverage of records may not be a problem as long as the data are not environmentally biased. However, spatial bias may result in environmental bias. For example, Hortal *et al.* (2008) showed that museum records for dung beetles in Madrid did not completely capture the environmental conditions inhabited by species. On the other hand, other studies have shown that museum records may be spatially uneven without there being major biases in environmental space (Austin and Meyers, 1996; Kadmon *et al.*, 2004). I conducted a test of environmental bias in the BioMAP data for butterflies with respect to four environmental gradients: elevation, mean annual temperature, total annual precipitation and an index of human impact from the Last of the Wild project (Sanderson *et al.*, 2002). The butterfly data set is the most spatially patchy of those currently completed by the BioMAP project. Although records were clearly biased with respect to the human impact index, there was no



**Figure 2.** (a) The relationship between a country's gross domestic product (GDP) and the density of species occurrence records in the Global Biodiversity Information Facility (GBIF) database. (b) The relationship between an index of human impact, derived from the Last of the Wild project (Sanderson *et al.*, 2002), and the density of occurrence records in the GBIF database.



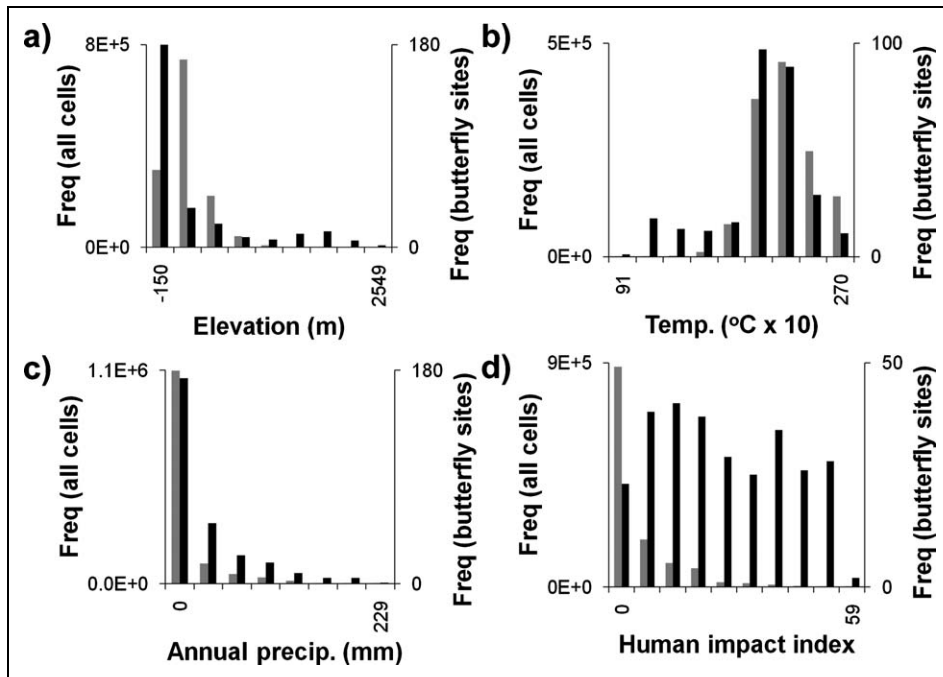
**Figure 3.** Sites with species occurrence records in Egypt's BioMAP (<http://www.biomapegypt.org>) database. Note that collecting effort has been highest in the Nile Valley and Delta, along the Mediterranean Coast and in the Sinai Peninsula. There are gaps in collecting effort in the Western Desert and, to a lesser extent, in the Eastern Desert and the North Sinai Peninsula.

obvious bias with respect to temperature, precipitation or elevation (Figure 4). A quantitative analysis of the same data showed that records were distributed almost as evenly along the main environmental gradients as would be expected if the sites had been located randomly (Newbold *et al.*, 2009). These results suggest that the butterfly collectors have sampled the main environmental gradients relatively evenly, although they have tended to visit sites that are more accessible – ie, those that are closer to roads, towns and cities. For uses of museum data that involve analyses of the environments in which species are found, it will be important to ensure that the species records are not environmentally biased. Even if systematic sampling of the whole area of interest is not possible, future sampling should be designed to cover all of the main environmental gradients present (Hirzel and Guisan, 2002; Wintle *et al.*, 2005).

Species records from museums will almost always show temporal bias, with peaks in collecting when experts on a given taxonomic group are most active (Soberón *et al.*, 2000). For example, records of species in Egypt are unevenly distributed temporally. Sampling for all groups shows clear peaks during periods of time when collectors were most active; these peaks were different for different taxa (Figure 5a). As well as being biased towards certain years, biodiversity sampling may also be seasonally biased (Peterson *et al.*, 1998). Records for mammal, reptile and amphibian species in Egypt were biased seasonally towards the summer months, but not dramatically (Figure 5b). The fact that butterflies were sampled much more often in the summer months is not surprising since the adults of most butterfly species fly only in the summer. Removing all temporal bias in sampling is almost impossible, because there are not enough biologists to sample all taxonomic groups and all geographical areas continuously. However, in rare cases where data sets with records that cover long periods of time do exist, they can be used to make important inferences about temporal changes in the distribution and abundance of species (eg, Peach *et al.*, 1996; Fitter and Fitter, 2002).

The final type of bias often present in museum data is taxonomic bias. Collectors have tended to focus on the larger, charismatic groups of organisms, and on groups that are more easy to detect in the field (Soberón *et al.*, 2000; Williams *et al.*, 2002; Graham *et al.*, 2004). This has led to a bias in sampling towards vertebrates, flowering plants and some insect groups. For example, in the GBIF database, vertebrates and plants are massively overrepresented while most other groups are underrepresented, compared to the estimated total number of species in each group (Figure 6). In the future, there needs to be a greater focus on collecting records for the less-well-studied taxonomic groups. However, we do not have infinite resources and thus it will be impossible to comprehensively sample all





**Figure 4.** Frequency distribution of sites from the BioMAP database with butterfly records, with respect to four environmental gradients – (a) elevation, (b) mean annual temperature, (c) total annual precipitation and (d) an index of human impact (Sanderson *et al.*, 2002) (black bars) – and the frequency distribution of all 30 arc-second grid cells in Egypt with respect to the same gradients (grey bars)

taxonomic groups. This has led to considerable interest in conservation ecology in the possibility that the distributions of well-studied taxa could act as surrogates for the distributions of other taxa (Loyola *et al.*, 2007; Pinto *et al.*, 2008).

It is crucial to know about the errors and biases in museum data so that their limitations are known. Nevertheless, when used with caution, museum data can be invaluable in efforts to understand patterns in the distribution of species, as described in the next section.

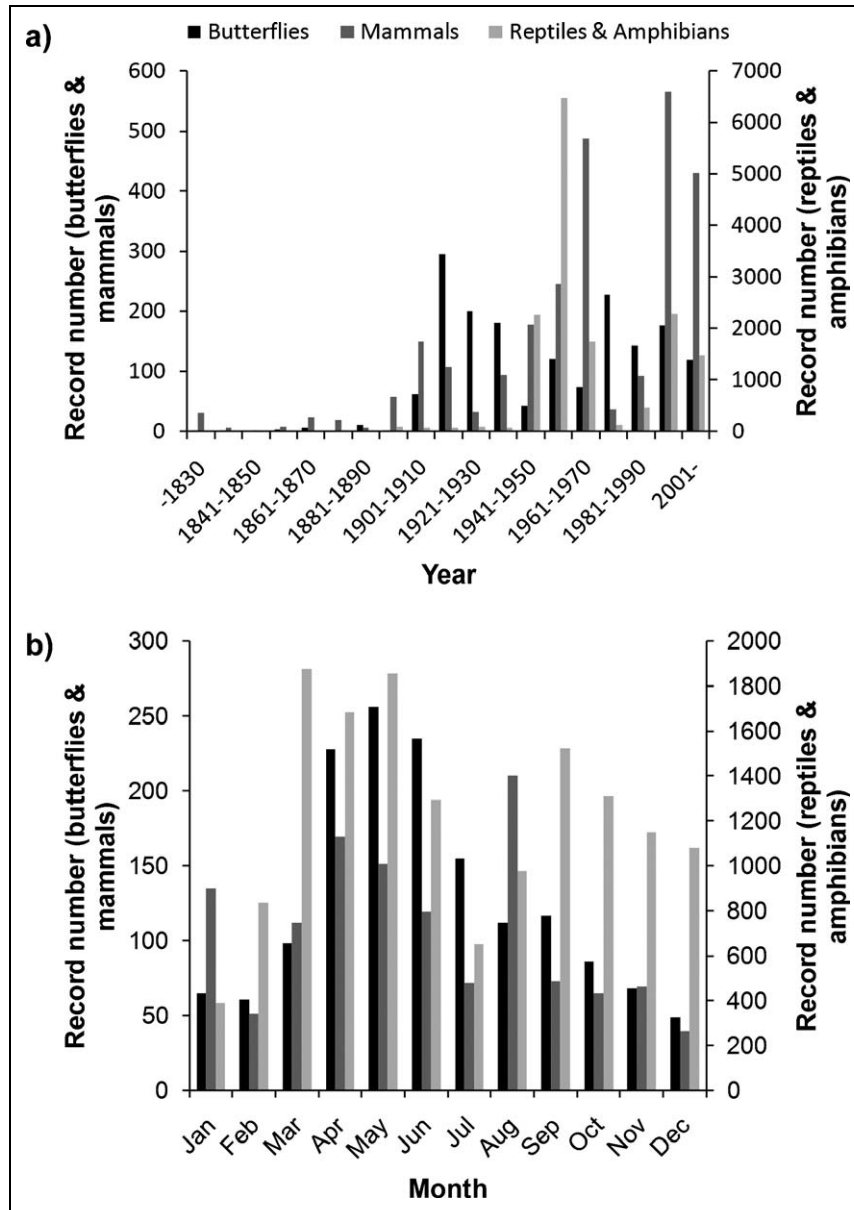
### III Using museum data to model species' distributions

#### I Determining species distributions

Even if efforts are made to reduce bias in data on the occurrence of species, it will almost never be

possible to sample comprehensively. Furthermore, we need to act now to save species from extinction. We cannot wait indefinitely for better information, but must use the knowledge that we already have. In spite of their limitations, museum data may be very useful for conserving biodiversity. For example, several studies have used museum data to estimate the size of species ranges in order to predict their risk of extinction, according to the guidelines set out by the International Union for the Conservation of Nature (IUCN) (eg, Randrianasolo *et al.*, 2002; Greenbaum and Komar, 2005). Gilbert and Zalut (2007) presented a similar assessment of the conservation status of Egypt's butterfly species, using museum records from the BioMAP database. The number of records and the area occupied by each species (measured using extents of occurrence and areas of occupancy;

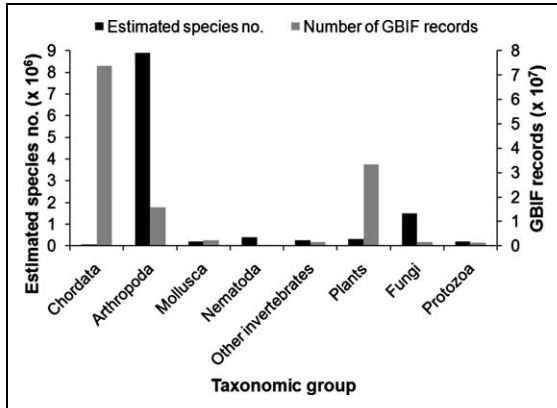




**Figure 5.** (a) Temporal and (b) seasonal distributions of records for butterflies (black bars), mammals (dark grey bars), and reptiles and amphibians (light grey bars) in Egypt's BioMAP database

Gaston, 1991) were used to assign each species to one of the IUCN categories of risk. Of 63 species, one was listed as critically endangered, one as endangered, 14 as vulnerable and five as data-deficient.

Among the IUCN criteria used to categorize extinction risk are those describing trends in populations. Some studies have used museum records spanning several years to investigate population trends in order to make better

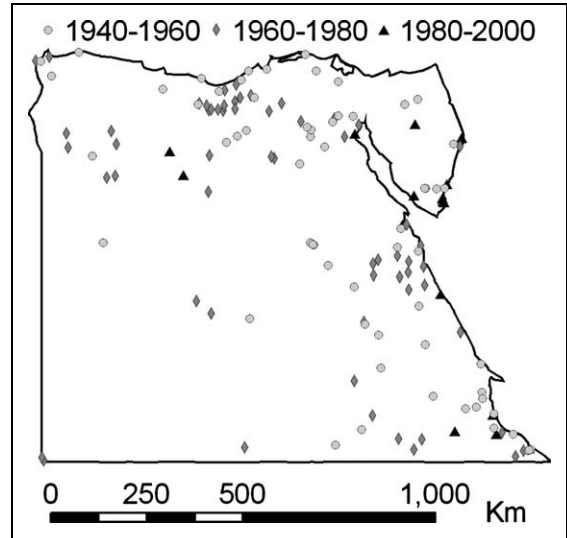


**Figure 6.** Number of records in the Global Biodiversity Information Facility (GBIF) database (grey bars), compared to estimated total numbers of species according to May (1997) (black bars), for a number of major taxonomic groups

assessments of conservation status (McGowan and Gillman, 1997; McCarthy, 1998). Plotting the BioMAP records for the Dorcas gazelle (*Gazella dorcas* Linnaeus) in Egypt reveals a dramatic decline in the range inhabited by this species in the last 20 years (Figure 7). The distribution of the Dorcas gazelle in Egypt has been well studied in recent years (Saleh, 1987; El Alqamy and Baha El Din, 2006). However, for many species occurrence data with sufficient temporal coverage for this kind of analysis are not available.

Another possibility for using museum records in conservation assessments is to combine data for many species to identify areas that are particularly rich in species. Funk *et al.* (1999) used this approach for a selection of genera in 12 taxonomic groups in Guyana, identifying three hotspots of diversity.

If the sampling of biodiversity in an area is incomplete and biased, then the results of conservation assessments using museum data may reflect artifacts of collecting effort rather than the real status of species. We need some way to extrapolate from the incomplete data that exist in order to estimate whether species occur in



**Figure 7.** Temporal trend in the distribution of Dorcas gazelles (*Gazella dorcas*) in Egypt, inferred from museum records taken from Egypt's BioMAP database. Records were divided into three time periods between 1940 and 2000

places that have not been sampled. Species distribution models (or ecological niche models, bioclimate models, or climate envelope models) offer a means to do this. They attempt to capture the ecological niche of species, by relating a set of records of species occurrence to variables describing aspects of the environment thought to be important in determining the distribution of species (for recent reviews, see Wintle *et al.*, 2005; Araújo and Guisan, 2006; Hirzel and Le Lay, 2008). There has been a massive increase in the use of species distribution models in the last two decades, and several dedicated computer programs have been developed for producing them (eg, Stockwell and Noble, 1992; Phillips *et al.*, 2006).

## 2 Species distribution models

A large number of different methods exist for modelling the distributions of species. Among them are traditional statistical techniques, such as generalized linear models. These techniques

require records of both species presence and species absence. However, museum data generally consist only of presence records (Graham *et al.*, 2004). One solution is to draw randomly a set of absence records from cells that do not contain presence records (termed 'pseudo-absence data'; Zaniwski *et al.*, 2002). Alternatively, there are many modelling techniques that do not require absence data. For example, some models simply define an environmental envelope encompassing all of the presence records, within which the species is considered capable of surviving and reproducing. However, in comparisons of the accuracy of models developed with different techniques, envelope methods have been shown to be among the worst-performing (Elith *et al.*, 2006). The most popular modelling techniques, and those that have seen extensive use with museum data, are those that model the presence records with reference to the environmental conditions in the whole study area; examples of these techniques are Maxent (Phillips *et al.*, 2006) and GARP (Stockwell and Noble, 1992). Many of these techniques have been shown to produce very accurate models of species' distributions (Elith *et al.*, 2006).

In order to develop useful distribution models, it is necessary to select environmental variables likely to show a strong association with the distribution of species. Using too many variables will result in overfitting of the model (Chatfield, 1995). It has been shown that the choice of environmental variables can have a significant effect on model accuracy (Parolo *et al.*, 2008; Peterson and Nakazawa, 2008). Models developed with environmental variables that have a direct effect on distributions will be more accurate, more biologically informative and more generalizable than models developed with variables that have only an indirect effect on distributions, through correlations with variables that have a direct effect (Austin and Meyers, 1996; Austin, 2002; 2007; Austin *et al.*, 2006). The determinants of distributions are likely to be different at different spatial

scales and extents of study (Mackey and Lindenmayer, 2001). For example, Anderson *et al.* (2009) showed that distribution models using climate variables explained the distribution of the hen harrier (*Circus cyaneus* Linnaeus) at the European scale very well, but that the British distribution was modelled better by finer-scale habitat data and variables describing habitat management. Climate and habitat variables are the most commonly used variables in distribution model studies and have been shown many times to be very good correlates of species occurrence (Guisan and Hofer, 2003; Araújo *et al.*, 2005a; Wintle *et al.*, 2005; Elith *et al.*, 2006; Guisan *et al.*, 2006b). However, other factors may also be important in determining distributions.

Even the earliest discussions of ecological niches recognized that interactions among species may play an important role in shaping distributions (Grinnell, 1917) and this is a topic that has received much attention since (Pulliam, 2000). Anderson *et al.* (2002a; 2002b) developed distribution models for two species of spiny pocket mouse (*Heteromys australis* Thomas and *H. anomalus* Thompson) in South America, inferring that dominance by one species in areas of predicted sympatry was the result of competitive exclusion by the dominant competitor. Several studies have found that including variables describing the distributions of interacting species can improve the accuracy of distribution models. For example, including the distributions of southern beech (*Nothofagus* sp. Blume) competitors improved the accuracy of distribution models for tree species in New Zealand (Leathwick and Austin, 2001). Similarly, models of Australian marsupial species including variables describing the abundance of competing species were more accurate than models that did not include these variables (Ritchie *et al.*, 2009). The availability of food may also have an important effect on the distribution of species. The accuracy of distribution models for the specialist clouded Apollo

butterfly (*Parnassius mnemosyne* Linnaeus) was improved by including the distributions of its larval host plants as environmental variables (Araújo and Luoto, 2007). Another study (Gutiérrez *et al.*, 2005) found that the distributions of larval host plants were not very important in determining the distribution of a more generalist butterfly species, the silver-studded blue (*Plebejus argus* Linnaeus), but that the accuracy of models of its distribution was improved by including the distribution of its mutualistic ant species, *Lasius niger* Linnaeus.

Distributions will also be affected by the ability of species to disperse to different parts of their potential ranges. An area of suitable climate and habitat will not be occupied by a species if it is unable to disperse there (Pulliam, 2000). Conversely, populations of a species may be able to persist in areas where the climate and habitat would not otherwise be able to support populations (termed 'sink' areas), by continual immigration from suitable areas ('source' areas). Dispersal limitation and source-sink dynamics will result in spatial patterns in species distributions that are at least partly independent of the environment; these spatial patterns are referred to as endogenous spatial autocorrelation (Legendre, 1993). Spatial autocorrelation can be captured using autologistic models (Osborne *et al.*, 2001; Keitt *et al.*, 2002; Lichstein *et al.*, 2002; Segurado *et al.*, 2006; Dormann *et al.*, 2007). However, it is difficult to apply autologistic models to opportunistic records of species occurrence, such as museum records (but see Syartinilia and Tsuyuki, 2008, for one solution to this problem). An alternative method to account for spatial autocorrelation when modelling species' distributions is to include the geographical coordinates (longitude and latitude), and interactions between them, as independent variables in the model; this is referred to as 'trend surface analysis' (Legendre, 1993; Lobo *et al.*, 2002; van Rensburg *et al.*, 2002; Gutiérrez *et al.*, 2005).

Although distributions may be determined by a great many environmental factors, both

abiotic and biotic, distribution models are often constrained by the availability of maps of these variables in the required format and at an appropriate resolution. Climate and broad-scaled habitat data are now very widely available. WorldClim provides global maps describing elevation and 19 climatic variables, at resolutions as fine as 30 arc-seconds (Hijmans *et al.*, 2005), and there are global classifications of habitat at similar resolutions, derived from satellite imagery (eg, Hansen *et al.*, 2000). Data on microclimate and fine-scale habitat, which may be important in determining distributions over smaller extents, and on the distributions of interacting species are very often not available.

### 3 Challenges of using museum data to develop species distribution models

The problems caused by errors in the location of museum records and bias in sampling effort have received a great deal of theoretical treatment, but the effect of these on the accuracy of distribution models has received very little empirical testing. It seems intuitive that environmental bias in the collecting of species records will result in model predictions that are biased towards environments that have received more intense sampling (Araújo and Guisan, 2006; Wintle *et al.*, 2005). Indeed, Kadmon *et al.* (2003) found that the accuracy of distribution models for woody plants in Israel decreased strongly with an increase in the climatic bias of species records. Stockwell and Peterson (2002) showed that controlling for environmental bias in sampling improved the accuracy of distribution models for the wood thrush (*Hylocichla mustelina* Baird) in the United States. Phillips *et al.* (2009) also showed a decrease in model accuracy with increasing sampling bias, by creating a simulated distribution in Canada and sampling this distribution with increasing degrees of bias. On the other hand, spatial bias may not have a negative impact on distribution model accuracy if it does

not result in environmental bias. Kadmon *et al.* (2004) demonstrated substantial bias towards roads in records of the occurrence of woody plants in Israel. However, the accuracy of models developed using these records was not decreased by this bias because the distribution of roads in Israel was not biased environmentally (Kadmon *et al.*, 2004). Therefore, when using museum data for distribution modelling, it is important to ensure that the records are not biased with respect to the environmental variables used.

The effect of errors in the location of records on the accuracy of species distribution models is less obvious. Maps of environmental variables generally consist of values arranged in a grid of numerous squares. If the locational error is sufficiently large that a record is plotted in the wrong grid square, then the impact on model accuracy may be large. Indeed, one study (Visscher, 2006) showed that introducing errors into records of a simulated species in Canada decreased the accuracy of resource selection function models, which are very similar in principle to species distribution models. In contrast, if the environmental variables are spatially autocorrelated (such that neighbouring cells tend to have similar values of a given variable), as is generally the case, then small errors in the location of records may have a small effect on model accuracy, even if the records are plotted in the wrong grid cell (Graham *et al.*, 2008). Graham *et al.* (2008) found that introducing relatively small errors into the occurrence records for many different species from regions all over the world had, overall, only a minor effect on the accuracy of distribution models; the effect of these errors was greatest in the region that showed the lowest spatial autocorrelation in the environmental variables.

#### 4 Applications of species distribution models

Species distribution models have been applied to answer a wide variety of questions in conservation and ecology. Given the considerable

investment in time and money necessary to conduct surveys of the occurrence of species, many studies using distribution models have used records from museums, collections and the literature. One obvious use for the models is in guiding decisions about the conservation of species. For example, one study (Guisan *et al.*, 2006a) used records of the occurrence of alpine sea holly (*Eryngium alpinum* Linnaeus) in Switzerland, collected by volunteers, to guide a survey to search for more populations; the resulting survey led to the discovery of seven new populations (Guisan *et al.*, 2006a). Likewise, Raxworthy *et al.* (2003) used distribution models based on museum data for chameleons in Madagascar to direct additional surveys. A search of places predicted to have high numbers of chameleon species yielded seven new species (Raxworthy *et al.*, 2003).

Distribution models can also be used to identify areas that are important for species of conservation concern in order that those areas may be protected. Several studies have used distribution models to assess the protection afforded to threatened species by existing protected areas networks and to propose additions to these networks that would benefit the species concerned (Papeş and Gaubert, 2007; Solano and Fera, 2007; Thorn *et al.*, 2009). Another possibility is to use the models to identify sites that are potentially suitable for reintroductions (Klar *et al.*, 2008).

By determining the main correlates of species occurrence, distribution models can also be used to infer the causes for changes in the distribution of species over time. For example, one study used them to show that the decline in the distribution of the bilby (*Macrotis lagotis* Reid) in Australia since European settlement was associated with the introduction of non-native predators and with changes in the frequency and magnitude of fires (Southgate *et al.*, 2007). Another study showed that the extinction of the woolly mammoth (*Mammuthus primigenius* Blumenbach) in the Palearctic was probably the



result of both environmental changes and over-hunting by humans (Nogués-Bravo *et al.*, 2008).

Conservation ecologists often use species richness to assess how important it is to protect an area because species richness is relatively easy to measure, although other criteria may be more informative (Kershaw *et al.*, 1995; Margules and Pressey, 2000; Wilson *et al.*, 2007). However, extensive gaps in our knowledge of the distribution of species mean that even species richness estimates may be unavailable for many areas. One way around this problem is to develop species distribution models to fill the gaps in our knowledge of the distribution of individual species, and then to combine these distribution models to estimate species richness. This approach has been used a number of times to model patterns of species richness in parts of the world where distribution data are patchy (Loiselle *et al.*, 2003; Garcia, 2006; Pineda and Lobo, 2009; Raes *et al.*, 2009; Newbold *et al.*, 2009). Models of species richness can be used to assess whether existing networks of protected areas are effective at conserving the most biologically diverse areas (Garcia, 2006; Newbold *et al.*, 2009).

Another major application of distribution models, and one that is receiving a lot of attention from scientists, policy-makers and the media, is in trying to predict how climate change will affect species. The basic principle is to fit a distribution model using species records and climate data from the current time period, and then to project this model onto variables describing the climate as it is predicted to be in the future (Pearson and Dawson, 2003). Among the many studies that have used this approach, several have used museum records to fit the models (Iverson *et al.*, 1999; Peterson *et al.*, 2001; 2002; Bakkenes *et al.*, 2002; Miles *et al.*, 2004; Hole *et al.*, 2009). There are a number of complications in predicting how climate change will affect the distribution of species (Pearson and Dawson, 2003; 2004; Hampe, 2004; Guisan and Thuiller, 2005): first,

interactions between species may affect how they respond to climate change (Davis *et al.*, 1998; Araújo and Luoto, 2007; Post and Pederesen, 2008; Schweiger *et al.*, 2008; Harmon *et al.*, 2009); second, species may not be able to disperse fast enough to keep pace with the changing climate (Iverson *et al.*, 1999; Engler *et al.*, 2009; Mustin *et al.*, 2009; Willis *et al.*, 2009); third, the climate may not change as predicted (Reilly *et al.*, 2001); and, fourth, species may adapt to the new climate rather than undergoing shifts in their distribution (Davis and Shaw, 2001; Skelly *et al.*, 2007; Charmantier *et al.*, 2008; but see Davis *et al.*, 2005; Visser, 2008). The traditional approach to assessing the accuracy of predictions about future distributions is to evaluate their fit to the data for the current time period. However, given the uncertainties in predicting future distributions, this test of accuracy is likely to be inadequate. A more promising approach is to test the ability of models to predict changes that have already occurred. Several studies have done this, showing variable degrees of success in predicting the changes (Araújo *et al.*, 2005a; 2005b; Green *et al.*, 2008; Mitikka *et al.*, 2008; Gregory *et al.*, 2009; La Sorte *et al.*, 2009).

As well as being projected onto future time periods, distribution models can also be projected onto different geographical areas to predict the potential extent of species invasions. Peterson and Vieglais (2001) used museum data to model the native distributions of several species that are known to be invasive to the United States. These models were then projected onto climate variables in the invaded regions to predict the extent of the invasions. Most previous invasions of these species were successfully predicted by the models (Peterson and Vieglais, 2001). Conversely, another study (Broennimann *et al.*, 2007) found that invasions were predicted poorly by species distribution models, suggesting that species' niches can shift during invasions.

In addition to applied problems, distribution models may be used to address more

fundamental ecological questions, such as which factors show the strongest association with species' distributions. For example, Anderson *et al.* (2009) compared four sets of variables for modelling the distribution of the hen harrier in Britain: climate, habitat, gamekeeper activity and habitat management (moorland burning), finding that the European distribution was explained well by climate, whereas the British distribution was better explained by habitat, habitat management and gamekeeper activity. Distribution models have also been used to address questions regarding the evolution of ecological niches. Studies showing that the modelled niches of closely related species are more similar than the modelled niches of more distantly related species have concluded that niches are relatively conserved over evolutionary time periods (Peterson *et al.*, 1999; Eaton *et al.*, 2008). In contrast, other studies have found relatively little evidence for evolutionary conservatism of modelled niches (Peterson and Holt, 2003; Rice *et al.*, 2003; Evans *et al.*, 2009). It would seem that the degree of conservatism is highly dependent on the specific context and the timescale over which the effect is considered (Wiens, 2008).

## IV Conclusions

For many areas of the world and for the majority of species, museum data are the best available data describing distributions. The availability of museum occurrence records is clearly biased towards certain countries and towards species that are charismatic and more easily detectable. Nevertheless, conserving biodiversity requires knowledge of the distribution of species and museum data must play an important role in this process. Museum data can be used to assess the extent of species' distributions and thus the likelihood that species will become extinct. Where records show sufficient temporal coverage, they can be used to infer changes in the distribution of species over time.

Even within countries, sampling of species has been patchy and the results of analyses using museum occurrence records may reflect sampling effort more than they reflect real ecological phenomena. Species distribution models are a powerful tool for filling gaps in distribution data. However, there are a number of limitations of museum data that need to be considered before using them to develop distribution models. Bias in the scope of the records need not be a problem provided that this does not result in bias in environmental space. Errors in the location of records need to be carefully assessed, although small errors may not always decrease the accuracy of models.

Species distribution models have been put to a variety of uses, both applied and conceptual. It is important to remember that the models are correlative and thus care must be taken in inferring causal mechanisms for the distribution of species. This is of particular relevance when projecting models outside the environmental conditions for which the model was developed, such as when predicting the impacts of climate change or predicting the extent of species invasions. Nevertheless, species distribution models have been used to make some important advances in our understanding of the distribution of species and are a valuable tool in efforts to conserve biodiversity.

## Acknowledgements

I thank Francis Gilbert and Tom Reader for valuable comments on an earlier draft, Francis Gilbert, Samy Zalut and all the staff at the BioMAP project in Egypt (Ahmed Yakoub, Alaa Awad, Muhammed Sherif, Shama Omran, Shaimaa Esa, Yasmin Safwat, Nahla Ahmed, Esraa Saber) for collating the distribution data for Egyptian species, the Italian Cooperation (Debt Swap) for funding the BioMAP project, Mustafa Fouda (Director of the Nature Conservation Sector, Egyptian Environmental Affairs Agency), and all of the people and organizations who have contributed data to the GBIF database. This work was



supported by the Natural Environment Research Council (grant number NER/S/A/2006/14170).

## References

- Anderson, B.J., Arroyo, B.E., Collingham, Y.C., Etheridge, B., Fernandez-De-Simon, J., Gillings, S., Gregory, R.D., Leckie, F.M., Sim, I.M.W., Thomas, C.D., Travis, J. and Redpath, S.M. 2009: Using distribution models to test alternative hypotheses about a species' environmental limits and recovery prospects. *Biological Conservation* 142, 488–99.
- Anderson, R.P., Gomez-Laverde, M. and Peterson, A.T. 2002a: Geographical distributions of spiny pocket mice in South America: insights from predictive models. *Global Ecology and Biogeography* 11, 131–41.
- Anderson, R.P., Peterson, A.T. and Gomez-Laverde, M. 2002b: Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos* 98, 3–16.
- Araújo, M.B. and Guisan, A. 2006: Five (or so) challenges for species distribution modelling. *Journal of Biogeography* 33, 1677–88.
- Araújo, M.B. and Luoto, M. 2007: The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography* 16, 743–53.
- Araújo, M.B., Pearson, R.G., Thuiller, W. and Erhard, M. 2005a: Validation of species-climate impact models under climate change. *Global Change Biology* 11, 1504–13.
- Araújo, M.B., Whittaker, R.J., Ladle, R.J. and Erhard, M. 2005b: Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography* 14, 529–38.
- Austin, M.P. 2002: Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling* 157, 101–18.
- Austin, M.P. 2007: Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological Modelling* 200, 1–19.
- Austin, M.P. and Meyers, J.A. 1996: Current approaches to modelling the environmental niche of eucalypts: implication for management of forest biodiversity. *Forest Ecology and Management* 85, 95–106.
- Austin, M.P., Belbin, L., Meyers, J.A., Doherty, M.D. and Luoto, M. 2006: Evaluation of statistical models used for predicting plant species distributions: role of artificial data and theory. *Ecological Modelling* 199, 197–216.
- Bakkenes, M., Alkemade, J.R.M., Ihle, F., Leemans, R. and Latour, J.B. 2002: Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology* 8, 390–407.
- Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T. and Guisan, A. 2007: Evidence of climatic niche shift during biological invasion. *Ecology Letters* 10, 701–709.
- Charmantier, A., McCleery, R.H., Cole, L.R., Perrins, C., Kruuk, L.E.B. and Sheldon, B.C. 2008: Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science* 320, 800–803.
- Chatfield, C. 1995: Model uncertainty, data mining and statistical inference. *Journal of the Royal Statistical Society Series A: Statistics In Society* 158, 419–66.
- Davis, A.J., Jenkinson, L.S., Lawton, J.H., Shorrocks, B. and Wood, S. 1998: Making mistakes when predicting shifts in species range in response to global warming. *Nature* 391, 783–86.
- Davis, M.B. and Shaw, R.G. 2001: Range shifts and adaptive responses to Quaternary climate change. *Science* 292, 673–79.
- Davis, M.B., Shaw, R.G. and Etterson, J.R. 2005: Evolutionary responses to changing climate. *Ecology* 86, 1704–14.
- Dennis, R.L.H. and Hardy, P.B. 1999: Targeting squares far survey: predicting species richness and incidence of species for a butterfly atlas. *Global Ecology and Biogeography* 8, 443–54.
- Dennis, R.L.H. and Thomas, C.D. 2000: Bias in butterfly distribution maps: the influence of hotspots and recorder's home range. *Journal of Insect Conservation* 4, 73–77.
- Dennis, R.L.H., Sparks, T.H. and Hardy, P.B. 1999: Bias in butterfly distribution maps: the effects of sampling effort. *Journal of Insect Conservation* 3, 33–42.
- Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M. and Wilson, R. 2007: Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* 30, 609–28.
- Eaton, M.D., Soberón, J. and Peterson, T. 2008: Phylogenetic perspective on ecological niche evolution in American blackbirds (family Icteridae). *Biological Journal of the Linnean Society* 94, 869–78.

- El Alqamy, H. and Baha El Din, S. 2006: Contemporary status and distribution of gazelle species (*Gazella dorcas* and *Gazella leptoceros*) in Egypt. *Zoology in the Middle East* 39, 5–16.
- Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loizelle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., 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, 129–51.
- Engler, R., Randin, C.F., Vittoz, P., Czaka, T., Beniston, M., Zimmermann, N.E. and Guisan, A. 2009: Predicting future distributions of mountain plants under climate change: does dispersal capacity matter? *Ecography* 32, 34–45.
- Evans, M.E.K., Smith, S.A., Flynn, R.S. and Donoghue, M.J. 2009: Climate, niche evolution, and diversification of the 'bird-cage' evening-primroses (*Oenothera*, sections *Anogra* and *Kleinia*). *The American Naturalist* 173, 225–40.
- Fitter, A.H. and Fitter, R.S.R. 2002: Rapid changes in flowering time in British plants. *Science* 296, 1689–91.
- 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, 727–51.
- García, A. 2006: Using ecological niche modelling to identify diversity hotspots for the herpetofauna of Pacific lowlands and adjacent interior valleys of Mexico. *Biological Conservation* 130, 25–46.
- Gaston, K.J. 1991: How large is a species' geographic range? *Oikos* 61, 434–38.
- Gilbert, F. and Zalut, S. 2007: *The butterflies of Egypt: atlas, red data listing and conservation*. Cairo: BioMAP, EEAA.
- Graham, C.H., Elith, J., Hijmans, R.J., Guisan, A., Peterson, A.T., Loizelle, B.A. and The NCEAS Predicting Species Distributions Working Group 2008: the influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology* 45, 239–47.
- Graham, C.H., Ferrier, S., Huettman, F., Moritz, C. and Peterson, A.T. 2004: New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution* 19, 497–503.
- Green, R.E., Collingham, Y.C., Willis, S.G., Gregory, R.D., Smith, K.W. and Huntley, B. 2008: Performance of climate envelope models in retrodicting recent changes in bird population size from observed climatic change. *Biology Letters* 4, 599–602.
- Greenbaum, E. and Komar, O. 2005: Threat assessment and conservation prioritization of the herpetofauna of El Salvador. *Biodiversity and Conservation* 14, 2377–95.
- Gregory, R.D., Willis, S.G., Jiguet, F., Vorisek, P., Klvanova, A., van Strien, A., Huntley, B., Collingham, Y.C., Couvet, D. and Green, R.E. 2009: An indicator of the impact of climatic change on European bird populations. *PLoS ONE* 4, e4678.
- Grinnell, J. 1917: The niche-relationships of the California Thrasher. *Auk* 34, 427–33.
- Guisan, A. and Hofer, U. 2003: Predicting reptile distributions at the mesoscale: relation to climate and topography. *Journal of Biogeography* 30, 1233–43.
- Guisan, A. and Thuiller, W. 2005: Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8, 993–1009.
- Guisan, A., Broennimann, O., Engler, R., Vust, M., Yoccoz, N.G., Lehmann, A. and Zimmermann, N.E. 2006a: Using niche-based models to improve the sampling of rare species. *Conservation Biology* 20, 501–11.
- Guisan, A., Lehmann, A., Ferrier, S., Austin, M., Overton, J.M.C., Aspinall, R. and Hastie, T. 2006b: Making better biogeographical predictions of species' distributions. *Journal of Applied Ecology* 43, 386–92.
- Guo, Q., Liu, Y. and Wieczorek, J. 2008: Georeferencing locality descriptions and computing associated uncertainty using a probabilistic approach. *International Journal of Geographical Information Science* 22, 1067–90.
- Gutiérrez, D., Fernández, P., Seymour, A.S. and Jordano, D. 2005: Habitat distribution models: are mutualist distributions good predictors of their associates? *Ecological Applications* 15, 3–18.
- Hampe, A. 2004: Bioclimate envelope models: what they detect and what they hide. *Global Ecology and Biogeography* 13, 469–71.
- Hansen, M.C., Defries, R.S., Townshend, J.R.G. and Sohlberg, R. 2000: Global land cover classification at 1km spatial resolution using a classification tree approach. *International Journal of Remote Sensing* 21, 1331–64.
- Harmon, J.P., Moran, N.A. and Ives, A.R. 2009: Species response to environmental change: impacts of food web interactions and evolution. *Science* 323, 1347–50.

- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A. 2005: Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25, 1965–78.
- Hijmans, R.J., Garrett, K.A., Huamán, Z., Zhang, D.P., Schreuder, M. and Bonierbale, M. 2000: Assessing the geographic representativeness of genebank collections: the case of Bolivian wild potatoes. *Conservation Biology* 14, 1755–65.
- Hirzel, A. and Guisan, A. 2002: Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling* 157, 331–41.
- Hirzel, A.H. and Le Lay, G. 2008: Habitat suitability modelling and niche theory. *Journal of Applied Ecology* 45, 1372–81.
- Hole, D.G., Willis, S.G., Pain, D.J., Fishpool, L.D., Butchart, S.H.M., Collingham, Y.C., Rahbek, C. and Huntley, B. 2009: Projected impacts of climate change on a continent-wide protected area network. *Ecology Letters* 12, 420–31.
- Hortal, J., Jiménez-Valverde, A., Gómez, J.F., Lobo, J.M. and Baselga, A. 2008: Historical bias in biodiversity inventories affects the observed environmental niche of the species. *Oikos* 117, 847–58.
- Iverson, L.R., Prasad, A. and Schwartz, M.W. 1999: Modeling potential future individual tree-species distributions in the eastern United States under a climate change scenario: a case study with *Pinus virginiana*. *Ecological Modelling* 115, 77–93.
- Kadmon, R., Farber, O. and Danin, A. 2003: A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications* 13, 853–67.
- Kadmon, R., Farber, O. and Danin, A. 2004: Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications* 14, 401–13.
- Keitt, T.H., Bjørnstad, O.N., Dixon, P.M. and Citron-Pousty, S. 2002: Accounting for spatial pattern when modeling organism-environment interactions. *Ecography* 25, 616–25.
- Kershaw, M., Mace, G.M. and Williams, P.H. 1995: Threatened status, rarity and diversity as alternative selection measures for protected areas – a test using Afrotropical antelopes. *Conservation Biology* 9, 324–34.
- Klar, N., Fernández, N., Kramer-Schadt, S., Herrmann, M., Trinzen, M., Büttner, I. and Niemitz, C. 2008: Habitat selection models for European wildcat conservation. *Biological Conservation* 141, 308–19.
- La Sorte, F.A., Lee, T.M., Wilman, H. and Jetz, W. 2009: Disparities between observed and predicted impacts of climate change on winter bird assemblages. *Proceedings of the Royal Society of London Series B, Biological Sciences* 276, 3167–74.
- Leathwick, J.R. and Austin, M.P. 2001: Competitive interactions between tree species in New Zealand's old-growth indigenous forests. *Ecology* 82, 2560–73.
- Legendre, P. 1993: Spatial autocorrelation: trouble or new paradigm? *Ecology* 74, 1659–73.
- Lichstein, J.W., Simons, T.R., Shriener, S.A. and Franzreb, K.E. 2002: Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs* 72, 445–63.
- Lobo, J.M., Lumaret, J.P. and Jay-Robert, P. 2002: Modelling the species richness distribution of French dung beetles (Coleoptera, Scarabaeidae) and delimiting the predictive capacity of different groups of explanatory variables. *Global Ecology and Biogeography* 11, 265–77.
- Loiselle, B.A., Howell, C.A., Graham, C.H., Goerck, J.M., Brooks, T., Smith, K.G. and Williams, P.H. 2003: Avoiding pitfalls of using species distribution models in conservation planning. *Conservation Biology* 17, 1591–600.
- Loyola, R.D., Kubota, U. and Lewinsohn, T.M. 2007: Endemic vertebrates are the most effective surrogates for identifying conservation priorities among Brazilian ecoregions. *Diversity and Distributions* 13, 389–96.
- Mackey, B.G. and Lindenmayer, D.B. 2001: Towards a hierarchical framework for modelling the spatial distribution of animals. *Journal of Biogeography* 28, 1147–66.
- Margules, C.R. and Pressey, R.L. 2000: Systematic conservation planning. *Nature* 405, 243–53.
- May, R.M. 1997: The dimensions of life on earth. In Raven, P.H., editor, *Nature and human society*, Washington, DC: National Academy Press, 30–45.
- McCarthy, M.A. 1998: Identifying declining and threatened species with museum data. *Biological Conservation* 83, 9–17.
- McGowan, P. and Gillman, M. 1997: Assessment of the conservation status of partridges and pheasants in South East Asia. *Biodiversity and Conservation* 6, 1321–37.
- Miles, L., Grainger, A. and Phillips, O. 2004: The impact of global climate change on tropical forest biodiversity

- in Amazonia. *Global Ecology and Biogeography* 13, 553–65.
- Mitikka, V., Heikkinen, R.K., Luoto, M., Araújo, M.B., Saarinen, K., Pöyry, J. and Fronzek, S. 2008: Predicting range expansion of the map butterfly in Northern Europe using bioclimatic models. *Biodiversity and Conservation* 17, 623–41.
- Mustin, K., Benton, T.G., Dytham, C. and Travis, J.M.J. 2009: The dynamics of climate-induced range shifting; perspectives from simulation modelling. *Oikos* 118, 131–37.
- Newbold, T., Gilbert, F., Zalat, S., El-Gabbas, A. and Reader, T. 2009: Climate-based models of spatial patterns of species richness in Egypt's butterfly and mammal fauna. *Journal of Biogeography* 36, 2085–95.
- Nogués-Bravo, D., Rodríguez, J., Hortal, J., Batra, P. and Araújo, M.B. 2008: Climate change, humans, and the extinction of the woolly mammoth. *PLoS Biology* 6, 685–92.
- O'Connell, A.F., Gilbert, A.T. and Hatfield, J.S. 2004: Contribution of natural history collection data to biodiversity assessment in national parks. *Conservation Biology* 18, 1254–61.
- Osborne, P.E., Alonso, J.C. and Bryant, R.G. 2001: Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. *Journal of Applied Ecology* 38, 458–71.
- Papeş, M. and Gaubert, P. 2007: Modelling ecological niches from low numbers of occurrences: assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents. *Diversity and Distributions* 13, 890–902.
- Parolo, G., Rossi, G. and Ferrarini, A. 2008: Toward improved species niche modelling: *Arnica montana* in the Alps as a case study. *Journal of Applied Ecology* 45, 1410–18.
- Peach, W.J., Buckland, S.T. and Baillie, S.R. 1996: The use of constant effort mist-netting to measure between-year changes in the abundance and productivity of common passerines. *Bird Study* 43, 142–56.
- Pearson, R.G. and Dawson, T.P. 2003: Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography* 12, 361–71.
- Pearson, R.G. and Dawson, T.P. 2004: Bioclimate envelope models: what they detect and what they hide – response to Hampe (2004). *Global Ecology and Biogeography* 13, 471–73.
- Peterson, A.T. and Holt, R.D. 2003: Niche differentiation in Mexican birds: using point occurrences to detect ecological innovation. *Ecology Letters* 6, 774–82.
- Peterson, A.T. and Nakazawa, Y. 2008: Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richeri*. *Global Ecology and Biogeography* 17, 135–44.
- Peterson, A.T. and Vieglais, D.A. 2001: Predicting species invasions using ecological niche modeling: new approaches from bioinformatics attack a pressing problem. *Bioscience* 51, 363–71.
- Peterson, A.T., Navarro-Siguenza, A.G. and Benitez-Diaz, H. 1998: The need for continued scientific collecting: a geographic analysis of Mexican bird specimens. *Ibis* 140, 288–94.
- Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sanchez-Cordero, V., Soberón, J., Buddemeier, R.H. and Stockwell, D.R.B. 2002: Future projections for Mexican faunas under global climate change scenarios. *Nature* 416, 626–29.
- Peterson, A.T., Sanchez-Cordero, V., Soberón, J., Bartley, J., Buddemeier, R.W. and Navarro-Siguenza, A.G. 2001: Effects of global climate change on geographic distributions of Mexican Cracidae. *Ecological Modelling* 144, 21–30.
- Peterson, A.T., Soberón, J. and Sanchez-Cordero, V. 1999: Conservatism of ecological niches in evolutionary time. *Science* 285, 1265–67.
- Phillips, S.J., Anderson, R.P. and Schapire, R.E. 2006: Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190, 231–59.
- Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. and Ferrier, S. 2009: Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications* 19, 181–97.
- Pineda, E. and Lobo, J.M. 2009: Assessing the accuracy of species distribution models to predict amphibian species richness patterns. *Journal of Animal Ecology* 78, 182–90.
- Pinto, M.P., Diniz-Filho, J.A.F., Bini, L.M., Blamires, D. and Rangel, T.F.L.V.B. 2008: Biodiversity surrogate groups and conservation priority areas: birds of the Brazilian Cerrado. *Diversity and Distributions* 14, 78–86.
- Post, E. and Pedersen, C. 2008: Opposing plant community responses to warming with and without herbivores.

- Proceedings of the National Academy of Sciences of the United States of America* 105, 12,353–58.
- Pulliam, H.R. 2000: On the relationship between niche and distribution. *Ecology Letters* 3, 349–61.
- Raes, N., Roos, M.C., Ferry Slik, J.W., van Loon, E.E. and ter Steege, H. 2009: Botanical richness and endemism patterns of Borneo derived from species distribution models. *Ecography* 32, 180–92.
- Randrianasolo, A., Miller, J.S. and Consiglio, T.K. 2002: Application of IUCN criteria and Red List categories to species of five Anacardiaceae genera in Madagascar. *Biodiversity and Conservation* 11, 1289–300.
- Raxworthy, C.J., Martínez-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, 837–41.
- Reddy, S. and Dávalos, L.M. 2003: Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography* 30, 1719–27.
- Reilly, J., Stone, P.H., Forest, C.E., Webster, M.D., Jacoby, H.D. and Prinn, R.G. 2001: Climate change – uncertainty and climate change assessments. *Science* 293, 430–33.
- Rice, N.H., Martínez-Meyer, E. and Peterson, A.T. 2003: Ecological niche differentiation in the *Aphelocoma* jays: a phylogenetic perspective. *Biological Journal of the Linnean Society* 80, 369–83.
- Ritchie, EG, Martin, J.K., Johnson, C.N. and Fox, B.J. 2009: Separating the influences of environment and species interactions on patterns of distribution and abundance: competition between large herbivores. *Journal of Animal Ecology* 78, 724–31.
- Saleh, M.A. 1987: The decline of gazelles in Egypt. *Biological Conservation* 39, 83–95.
- Sánchez-Fernández, D., Lobo, J.M., Abellán, P., Ribera, I. and Millán, A. 2008: Bias in freshwater biodiversity sampling: the case of Iberian water beetles. *Diversity and Distributions* 14, 754–62.
- Sanderson, E.W., Jaiteh, M., Levy, M.A., Redford, K.H., Wannebo, A.V. and Woolmer, G. 2002: The human footprint and the last of the wild. *Bioscience* 52, 891–904.
- Schweiger, O., Settele, J., Kudrna, O., Klotz, S. and Kühn, I. 2008: Climate change can cause spatial mismatch of trophically interacting species. *Ecology* 89, 3472–79.
- Segurado, P., Araújo, M.B. and Kunin, W.E. 2006: Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology* 43, 433–44.
- Skelly, D.K., Joseph, L.N., Possingham, H.P., Freidenburg, L.K., Farrugia, T.J., Kinnison, M.T. and Hendry, A.P. 2007: Evolutionary responses to climate change. *Conservation Biology* 21, 1353–55.
- 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, 1441–66.
- Solano, E. and Ferial, T.P. 2007: Ecological niche modeling and geographic distribution of the genus *Polianthes* L. (Agavaceae) in Mexico: using niche modeling to improve assessments of risk status. *Biodiversity and Conservation* 16, 1885–900.
- Soria-Auza, R.W. and Kessler, M. 2008: The influence of sampling intensity on the perception of the spatial distribution of tropical diversity and endemism: a case study of ferns from Bolivia. *Diversity and Distributions* 14, 123–30.
- Southgate, R., Paltridge, R., Masters, P. and Carthew, S. 2007: Bilby distribution and fire: a test of alternative models of habitat suitability in the Tanami Desert, Australia. *Ecography* 30, 759–76.
- Stockwell, D.R.B. and Noble, I.R. 1992: Induction of sets of rules from animal distribution data – a robust and informative method of data analysis. *Mathematics and Computers in Simulation* 33, 385–90.
- Stockwell, D.R.B. and Peters, D. 1999: The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* 13, 143–58.
- Stockwell, D.R.B. and Peterson, A.T. 2002: Controlling bias in biodiversity data. In Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A. and Samson, F.B., editors, *Predicting species occurrences: issues of accuracy and scale*, Washington, DC: Island Press, 537–46.
- Syartinilia and Tsuyuki, S. 2008: GIS-based modeling of Javan Hawk-Eagle distribution using logistic and autologistic regression models. *Biological Conservation* 141, 756–69.
- Thorn, J.S., Nijman, V., Smith, D. and Nekaris, K.A.I. 2009: Ecological niche modelling as a technique for assessing threats and setting conservation priorities for Asian slow lorises (primates: *Nycticebus*). *Diversity and Distributions* 15, 289–98.
- van Rensburg, B.J., Chown, S.L. and Gaston, K.J. 2002: Species richness, environmental correlates, and spatial



- scale: a test using South African birds. *The American Naturalist* 159, 566–77.
- Visser, D.R. 2006: GPS measurement error and resource selection functions in a fragmented landscape. *Ecography* 29, 458–64.
- Visser, M.E. 2008: Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society of London Series B, Biological Sciences* 275, 649–59.
- Wieczorek, J., Guo, Q.G. and Hijmans, R.J. 2004: The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *International Journal of Geographical Information Science* 18, 745–67.
- Wiens, J.J. 2008: Commentary on Losos (2008): niche conservatism déjà vu. *Ecology Letters* 11, 1004–1005.
- Williams, P.H., Margules, C.R. and Hilbert, D.W. 2002: Data requirements and data sources for biodiversity priority area selection. *Journal of Biosciences* 27, 327–38.
- Willis, S.G., Thomas, C.D., Hill, J.K., Collingham, Y.C., Telfer, M.G., Fox, R. and Huntley, B. 2009: Dynamic distribution modelling: predicting the present from the past. *Ecography* 32, 5–12.
- Wilson, D.E. and Reeder, D.M. 2005: *Mammal species of the world: a taxonomic and geographic reference*. Baltimore, MD: Johns Hopkins University Press.
- Wilson, K.A., Underwood, E.C., Morrison, S.A., Klausmeyer, K.R., Murdoch, W.W., Reyers, B., Wardell-Johnson, G., Marquet, P.A., Rundel, P.W., McBride, M.F., Pressey, R.L., Bode, M., Hoekstra, J.M., Andelman, S., Looker, M., Rondinini, C., Kariva, P., Shaw, M.R. and Possingham, H.P. 2007: Conserving biodiversity efficiently: what to do, where, and when. *PLoS Biology* 5, 1850–61.
- Wintle, B.A., Elith, J. and Potts, J.M. 2005: Fauna habitat modelling and mapping: a review and case study in the Lower Hunter Central Coast region of NSW. *Austral Ecology* 30, 719–38.
- Zaniewski, A.E., Lehmann, A. and Overton, J.M.C. 2002: Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling* 157, 261–80.